INVESTIGATING THE PROTON-TRANSLOCATING SUBUNIT OF THE ROTARY A-TYPE ATPASE

ELISE M. LAMING

Thesis presented for the
Degree of Doctor of Philosophy

St Vincent's Clinical School
Faculty of Medicine
The University of New South Wales

Month, Year

Supervisor: Dr Daniela Stock
Co-supervisor: Prof Boris Martinac

Molecular, Structural and Computational Biology Division,
The Victor Chang Cardiac Research Institute,
405 Liverpool St,
Darlinghurst,
NSW 2010, Australia.

ORIGINALITY STATEMENT

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed

Date

PLEASE TYPE

THE UNIVERSITY OF NEW SOUTH WALES
Thesis/Dissertation Sheet

Surname or Family name: LAMING

First name: Elise | Other name/s:

Abbreviation for degree as given in the University calendar: PhD

School: St Vincent's Clinical School | Faculty: Faculty of Medicine

Title: Investigating the proton-translocating subunit of the A-type rotary ATPase

Abstract 350 words maximum: (PLEASE TYPE)

All known forms of life are dependent on proton gradients as a source of energy to power essential cellular processes. Rotary ATPases are multi-subunit molecular machines that can harness this proton gradient and couple it to the synthesis of adenosine triphosphate (ATP), the "molecular currency" of the cell, via a rotary catalytic mechanism. All rotary ATPases contain membrane-embedded proton-translocating subunits through which protons are able to traverse the membrane, generating the mechanical energy required to synthesise ATP. The proton translocating subunit is the only subunit common to all rotary ATPases for which there is no high-resolution structural information, and thus the molecular understanding of the mechanism by which proton translocation is coupled to ATP synthesis is yet to be completed. To identify a suitable target for structural characterisation of this subunit by X-ray crystallography, twelve homologues of this subunit from the archaeal/bacterial (A-type) rotary ATPases (subunit I) were cloned into eight different expression vectors, giving 96 expression constructs. Constructs containing subunit I from the bacterium *Meiothermus ruber* routinely gave the highest yield of protein. Purification of one of these constructs was optimised and purified protein was found to be monomeric and monodispersed by multi-angle laser light scattering (MALLS). Furthermore, this protein was able to bind to one of its cognate binding partners from the A-ATPase complex, termed the peripheral stalk. Pulldown analysis, MALLS and isothermal titration calorimetry confirmed this, indicating that subunit I was folded and functional. Subunit I was subsequently subjected to crystallisation trials, however, diffracting crystals of subunit I were not obtained. It was proposed that this was most likely due to subunit I being most stable within the membrane of the native organism, in the context of the intact A-ATPase. Recent advances in the structural characterisation of intact rotary ATPases demonstrate that studying the proton translocating subunit in the context of the intact complex, rather than in isolation, may be a more suitable strategy for obtaining high resolution structural information about this subunit.

Declaration relating to disposition of project thesis/dissertation

I hereby grant to the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or in part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all property rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstracts International (this is applicable to doctoral theses only).

The University recognises that there may be exceptional circumstances requiring restrictions on copying or conditions on use. Requests for restriction for a period of up to 2 years must be made in writing. Requests for a longer period of restriction may be considered in exceptional circumstances and require the approval of the Dean of Graduate Research.

FOR OFFICE USE ONLY | Date of completion of requirements for Award:

THIS SHEET IS TO BE GLUED TO THE INSIDE FRONT COVER OF THE THESIS

Abstract

All forms of life are dependent on proton gradients as a source of energy to power essential cellular processes, and they all store this energy in the form of the nucleotide adenosine triphosphate (ATP). The rotary ATPases are multi-subunit molecular machines that utilise a rotary catalytic mechanism to couple ATP synthesis/hydrolysis to proton translocation across biological membranes. All rotary ATPases contain membrane-embedded proton-translocating subunits, via which protons enter and exit the membrane. This is the only subunit common to all rotary ATPases for which there is no high resolution structural information, and thus the molecular understanding of mechanism by which proton translocation is coupled to ATP synthesis/hydrolysis remains incomplete. To determine this structure, twelve homologues of the proton-translocating subunit from A-type rotary ATPases (subunit I) were cloned into eight different expression vectors, giving 96 expression constructs. Subunit I from the bacterium *Meiothermus ruber* gave the highest expression. Purification of this construct was optimised and purified protein was subjected to multi-angle laser light scattering (MALLS) analysis and crystallisation trials. The interaction between subunit I and the peripheral stalks, one of the binding partners of subunit I in the intact A-ATPase complex, was also investigated by pulldown analysis, MALLS and isothermal titration calorimetry (ITC). The X-ray crystal structure of subunit I was not able to be determined, and it was proposed that this was most likely because subunit I is most stable within the intact A-ATPase complex and in the lipid bilayer of the native organism. Nevertheless, it was found that the recombinantly expressed and purified *M. ruber* subunit I was able to bind to one peripheral stalk heterodimer, indicating subunit I was folded. There are, however, two peripheral stalk binding sites within subunit I, suggesting the two sites have different affinities for the peripheral stalk heterodimers. This has implications for the assembly of the intact A-ATPase complex.

Publication and presentations

Alastair G. Stewart, Elise M. Laming, Meghna Sobti, Daniela Stock. (2014) Rotary ATPases —dynamic molecular machines. *Current Opinion in Structural Biology* 25: 40-48

Elise M. Laming, Meghna Sobti, Daniela Stock (2013). The proton-translocating subunit of the rotary ATP synthase. Poster presentation at the Lorne Conference on Protein Structure and Function

Elise M. Laming, Meghna Sobti, Alastair Stewart, Daniela Stock (2014). The proton-translocating subunit of the rotary ATP synthase. Poster presentation at the Lorne Conference on Protein Structure and Function

Elise M. Laming, Meghna Sobti, Alastair G. Stewart, Daniela Stock (2014). The proton-translocating subunit of the rotary ATP synthase. Poster presentation at The Victor Chang Cardiac Research Institute, 15th International Symposium

Elise M. Laming, Alastair G. Stewart, Meghna Sobti, Daniela Stock (2015). The proton-translocating subunit of the rotary ATPase – the last piece of the puzzle. Poster presentation at the Hybrid Methods in Structural Biology Keystone Symposium

Elise M. Laming, Alastair G. Stewart, Meghna Sobti, Daniela Stock (2015). The proton-translocating subunit of the rotary ATPase – the last piece of the puzzle. Poster presentation at the East Coast Protein Meeting

Elise M. Laming, Alastair G. Stewart, Meghna Sobti, Daniela Stock (2015). The proton-translocating subunit of the rotary ATPase – the last piece of the puzzle. Poster presentation at the 23rd Annual St Vincent's Campus Research Symposium

Acknowledgements

I would first and foremost like to thank my supervisor, Dr Daniela Stock, for accepting me as a PhD student in her laboratory and for allowing me to work on such an exciting project. It was a great honour to work for someone that has made so many important contributions to the rotary ATPase field that are still talked about in papers being published now. Daniela's knowledge of the field is seemingly endless, and I was always in awe of how well she understood everything, and her theories about what we still do not know. I will always feel very privileged to have been mentored by Daniela.

I would also like to thank my co-supervisor Prof Boris Martinac, even though it took a little while for him to realise who I was. I will always have very fond memories of being called "Emily", perhaps the longest running inside joke between a co-supervisor and a student ever, or at least at the Victor Chang Cardiac Research Institute. I always enjoyed our weekly catch-ups at the Wednesday morning seminars.

I am greatly indebted to my colleague and friend Alastair Stewart, without whom I would not have made it through the last three and a half years. He made coming into work every day a lot of fun, and was invaluable in coming up with ideas when I wasn't sure where to take my project next. It means so much to me that he was always so willing to help, however he could, and that he never made it feel like an inconvenience. He has also been incredibly helpful with this thesis, even though I know I didn't always seem grateful for the corrections. I promise I am, and I am also endlessly grateful for everything else he has done for me.

A very big thank you is also due to my other colleagues at the Victor Chang, especially the other members of the Stock lab, Callum Smits, Roberta Davies, Meghna Sobti and Jessica Chaston. I wish to thank them for all of their help with my project and with the equipment, and for helping me to make it through my PhD. I really appreciate the support they all gave me with a project that for the most part was quite difficult. I would also like to thank all of my fellow PhD students, especially the "man lunch" crew. They really helped me to get through the last few months and I will always cherish our friendship and all of the fun we had together. I would like to wish them every success as they finish their PhDs.

I would like to thank the media and wash-up staff at the Garvan, as they were instrumental to me progressing smoothly through my PhD. I would especially like to thank Carolyn for the literally hundreds (possibly thousands?) of litres of media and buffer she made for me, and for always being so pleasant when we bumped into each other. I am so grateful that I had such a capable person to take care of this essential aspect of my work!

I would of course like to thank my friends and family for always supporting and encouraging me through what have been the most difficult three and a half years of my life. They were always there to listen to my woes whenever I was feeling down and it always made me feel like I could keep going. My mother and father have always been very supportive of all of my decisions and have never doubted that I was capable of doing anything I tried. I would especially like to thank Gena (and her lovely dogs), Alex and Daniel for being the very best friends a PhD student could ask for, for all of the coffee and food and late night heart to hearts. I am very lucky to have all of you in my life.

Lastly (and most importantly) I would like to thank Neil for being so loving and caring over the past three and a half years. I absolutely would not have made it through my PhD without him. It was always so nice to have someone that understood exactly what I was going through to talk to at the end of every day. I would like to thank him for always encouraging me and making me feel like I was capable of making it to the end. We've had so many fun times together over the past three and a half years, I'm glad we didn't let our PhDs stop us from doing that, although I feel it is very special that we've shared the experience of undertaking a PhD. I hope that I can be as supportive and caring as he finishes his PhD, and pursues what I know will be a very exciting scientific career.

Table of Contents

List of Figures

List of Tables

Symbols

ºC	Degrees Celsius
Å	Ångstrom
$A_{280\ nm}$	Absorbance at 280 nm
dn/dc	Refractive index increment
× *g*	Times gravity
kDa	kiloDaltons
L	Litre
µg	Microgram
µL	Microlitre
µm	Micrometre
µM	Micromolar
mAu	Milli absorbance units
mg	Milligram
mL	Millilitre
mM	Millimolar
mV	Millivolts
min	Minute
ng	Nanogram
Nm	Nanometre
$OD_{600\ nm}$	Optical density at 600 nm
rpm	Revolutions per minute
rps	Revolutions per second
w/v	Weight per volume

v/v	Volume per volume
U	Unit (enzyme activity)

Abbreviations

A-ATPase	Archaeal-type rotary ATPase
ADP	Adenosine diphosphate
AMP-PNP	Adenylyl-imidodiphsophate
A. fulgidus	*Archaeoglobus fulgidus*
ATP	Adenosine triphosphate
CMC	Critical micelle concentration
CV	Column volume
Cryo-EM	Electron cryo-microscopy
dAb	Domain antibody
DCCD	N,N-Dicyclohexylcarbodiimide
DDM	*n*-Dodecyl β-D-maltoside
DM	*n*-Decyl β-D-maltoside
DMSO	dimethyl sulfoxide
$E_{1\%}$	Absorbance of a 1% (w/v) solution
E. coli	*Escherichia coli*
E. hirae	*Enterococcus hirae*
ECL	Enhanced chemiluminescence
EMD	Electron microscopy data bank
F-ATPase	F_1F_O-type ATP synthase
Fab	Antigen-binding fragment
GFP	Green fluorescent protein
His_6/His_{10}	Hexahistidine/decahistidine
HRP	Horse radish peroxidase
I_{sol}	Soluble domain of subunit I
I_{TM}	Transmembrane domain of subunit I
I. tartaricus	*Ilyobacter tartaricus*
ITC	Isothermal titration calorimetry
LCP	Lipidic cubic phase
MALLS	Multi-angle laser light scattering
MBP	Maltose binding protein
M. ruber	*Meiothermus ruber*
M. jannaschii	*Methanocaldococcus jannaschii*

M. maripaludis	*Methanococcus maripaludis*
M. mazei	*Methanococcus mazei*
Mistic	Membrane-integrating sequence for translation of IM protein constructs
MWCO	Molecular weight cut-off
NCS	Non-crystallographic symmetry
NMR	Nuclear magnetic resonance spectroscopy
OG	*n*-Octyl-β-D-glucopyranoside
OSCP	Oligomycin sensitivity-conferring protein
P. denitrificans	*Paracoccus denitrificans*
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
Pi	Inorganic phosphate
PSI-BLAST	Position-specific iterative basic local alignment search tool
Pmf	Proton motive force
P. denitrificans	*Paracoccus denitrificans*
P. furiosus	*Pyrococcus furiosus*
P. horikoshii	*Pyrococcus horikoshii*
RAVE	Regulator of the ATPase of vacuolar and endosomal membranes
RI	Refractive index
RMSD	Root mean square deviation
S. cerevisiae	*Saccharomyces cerevisiae*
SEC	Size exclusion chromatography
SLIC	Sequence- and ligation-independent cloning
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
T. gammatolerans	*Thermococcus gammatolerans*
T. acidophilum	*Thermoplasma acidophilum*
TEV	Tobacco etch virus
TrxA	Thioredoxin A
V-ATPase	Vacuolar-type ATPase
UV	Ultraviolet

Chapter One: Introduction

1.1 Proton gradients and the emergence of life

For a number of decades chemists and biologists have debated the nature of the conditions in which life emerged. While J.B.S Haldane's theory that life arose in a "primordial soup" of warm water and simple molecules continues to be influential,[1] a number of important studies over the last few decades have exposed the limitations of this theory.[2] Perhaps the most significant of these were the discoveries of hydrothermal vents in 1977,[3] and later alkaline hydrothermal vents in 2000,[4] sites in which ocean water is warmed to 30-70 ºC and where many new life forms were discovered. Unlike a primordial soup, which would only contain ionizing ultraviolet (UV) radiation as a source of energy, alkaline hydrothermal vents are sites of naturally-occurring and continuously replenished proton (H^+) gradients, formed by the flow of water from the ocean that had been acidified by dissolved carbon dioxide (1000 times the concentration of CO_2 dissolved in the ocean today) into the porous alkaline hydrothermal vent. These vents were alkalised by the process of serpentinisation, whereby the mineral olivine ($(Mg^{+2}, Fe^{+2})_2SiO_4$) in these rocks is hydroxylated to serpentine ($Mg_3(OH)_4(Si_3O_5)$).[4,5]. The significance of a H^+ gradient at these sites is that it enabled CO_2 fixation, as the ocean water flowing into the vent contained hydrogen that could reduce the CO_2 to methane or acetate, both of which would have been utilised as the building blocks of early biomolecules.[6] There is evidence that these vents could have contained lipid membranes, produced by inorganic processes, that would have contained metal sulphide clusters,[7] including iron-sulfur clusters identical to those found in many proteins. These minerals could have served as the precursors to enzymes in catalysing essential organic reactions, such as the formation of high-energy thioesters (the likely original energy currency),[8] polymerisation of nucleotides[9] and synthesis of hydrocarbons[10] and amino acids,[11] and the energy to power these reactions would have come from the H^+ gradient. The H^+ gradient across the membranes in these vents has been calculated to be about 200 mV, with an orientation of positive outside and negative inside, which is identical to that of all living cells (Figure 1-1).

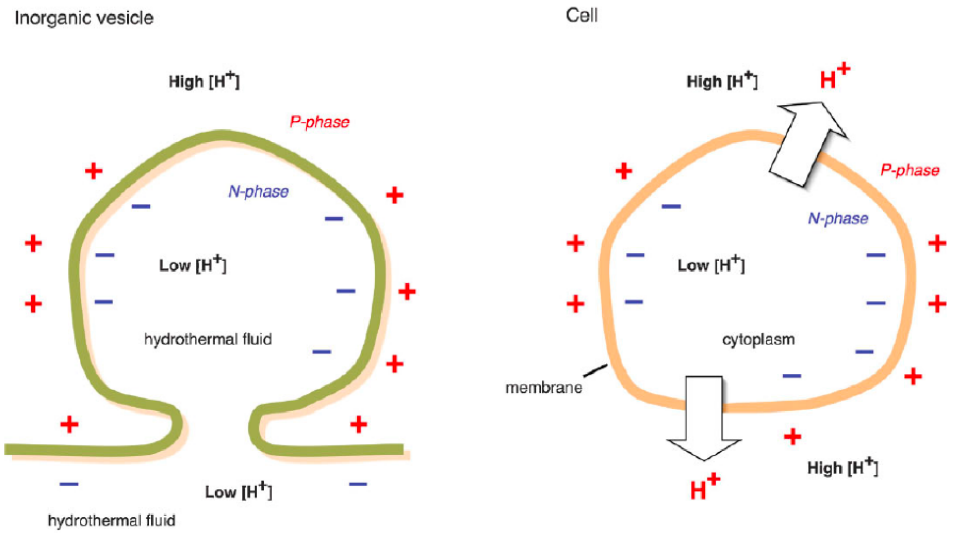


Figure 1-1: Orientation of the proton gradient in hydrothermal vents and the cells of organisms.

The water inside alkaline hydrothermal vents (left) would have been alkalised by serpentinisation, while water outside the vent would have been acidified by dissolved CO_2, and thus across the boundary of the vent there is a H^+ gradient termed the proton motive force. The proton motive force would have been a continuously replenished energy source that would have been used to power chemical reactions occurring in vesicles in the hydrothermal vent. The polarity of all living cells (right) is identical to that of hydrothermal vents, with a membrane separating a low-pH extracellular environment and a high pH intracellular environment. All living cells utitlise the proton motive force to power essential biochemical processes. From Lane et al. (2010), Figure 1.[2]

1.2 Chemiosmosis powers all forms of life

The separation and movement of this hydrothermal H^+ gradient across the membranes of early cells meant that energy could be stored and used to power essential biological reactions, a process that is termed chemiosmosis. Peter Mitchell first proposed this process, which he termed "chemi-osmotic coupling", in 1961.[12] He suggested that the H^+ gradient, or the "proton motive force" (pmf), present in all cells is utilised by an enzyme termed the ATPase (known at the time to be reversible, and can therefore also be called "ATP synthase") to synthesise the high-energy molecule adenosine triphosphate (ATP, Figure 1-2).

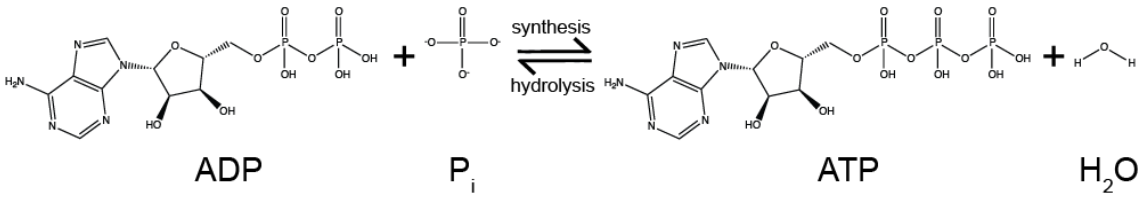


Figure 1-2: Hydrolysis and synthesis of ATP

Hydrolysis and synthesis of ATP, with molecules shown as skeletal formulae.

It is now known that all known forms of life, including archaea, store the majority of their energy in the form of ATP, and they all use the same fundamental cellular machinery, the rotary ATPases, to synthesise ATP (Figure 1-4).[13] The fact that the arachaeal and bacterial rotary ATPases are highly similar indicates they share a common ancestor,[14] which has led to the theory that early cells in the hydrothermal vents, the ancestors of archaea and bacteria, possessed rotary ATP synthases as a means of harnessing the energy of the hydrothermal H^+ gradient.[2] Therefore, when these cells left the hydrothermal vents and were able to synthesise their own H^+ gradients, for example by segregating oxidation and reduction reactions to opposite sides of a membrane.[15] This segregation of redox reactions subsequently evolved into the electron transport chain (Figure 1-3), and as these early cells already possessed the rotary ATPases, the cells were able to utilise the energy of this H^+ gradient from the electron transport chain to power the essential biochemical processes they had developed. The rotary ATPases are therefore ancient and essential molecular machines that utilise chemiosmosis to power all forms of life.

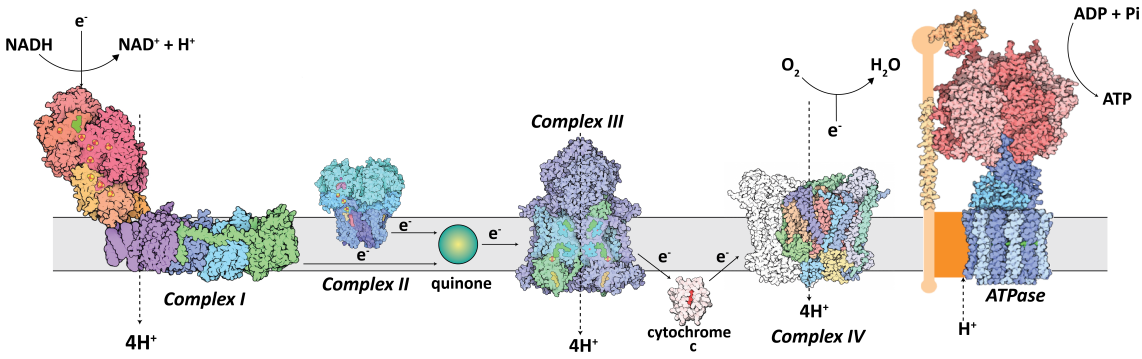


Figure 1-3: The electron transport chain.

Schematic showing the arrangement of the protein complexes of the electron transport chain and the direction of the flow of electrons (e^-) and protons (H^+). H^+ from oxidation of biomolecules flow through Complexes I, III and IV to generate a gradient termed the proton motive force. The proton motive force powers Complex V, the rotary ATPase, to synthesise ATP from ADP and Pi, while e^- reduce the terminal electron acceptor, molecular oxygen (O_2) to water (H_2O). The structures are taken from the PDB's Molecule of the Month and are based on the X-ray crystal structures of these complexes.[16] Complex I: Efremov et al. (2010)[17] and Efremov and Sazanov (2011)[18] DOI: 10.2210/rcsb_pdb/mom_2011_12. Complex II: Yankoskaya et al. (2003)[19] DOI: 10.2210/rcsb_pdb/mom_2012_10. Complex III: Zhang et al. (1998)[20] DOI: 10.2210/rcsb_pdb/mom_2011_5 Cytochrome c: Takano and Dickerson (1980)[21] DOI: 10.2210/rcsb_pdb/mom_2002_12 Complex IV: Yoshikawa et al. (1998)[22] DOI: 10.2210/rcsb_pdb/mom_2000_5. Rotary ATPase: Rastogi and Girvin (1999)[23], Gibbons et al. (2000)[24], Wilkens et al. (2005)[25], Del Rizzo et al. (2002)[26] DOI: 10.2210/rcsb_pdb/mom_2005_12.

1.3 The rotary ATPases

As mentioned in the previous section, the rotary ATPases are ancient molecular machines that couple the translocation of ions across a membrane to the synthesis and/or hydrolysis of ATP via a rotary catalytic mechanism. These enzymes are composed of two connected motors; one is soluble and is powered by ATP, termed the R_1 motor (the "1" comes from coupling factor 1, which will be explained below), and the other is membrane-bound and powered by H^+ or sodium ions (Na^+, Figure 1-4), termed the R_O motor (the "O" stands for oligomycin, an antibiotic that is able to block ion translocation through rotary ATPases). When the enzyme is in synthesis mode, H^+ from the low pH side of the membrane travel down the H^+ gradient to enter the membrane-bound motor and drive the rotation of the R_O rotor ring. This is turn drives the rotation of a central stalk, and the rotating stalk induces conformational changes in

the catalytic, soluble motor of the enzyme that catalyse the reaction of adenosine diphosphate (ADP) with inorganic phosphate (Pi) to produce ATP (Figure 1-4). When the enzyme is in hydrolysis mode, ATP is hydrolysed to ADP + Pi, which causes the central stalk to rotate in the opposite direction, and thus the membrane rotor is powered by the energy of ATP hydrolysis to pump H^+ or Na^+ out of the membrane against the electrochemical gradient (Figure 1-5). The understanding of this mechanism comes from innumerable experiments conducted over many decades. The most important of these experiments will be introduced in the following section.

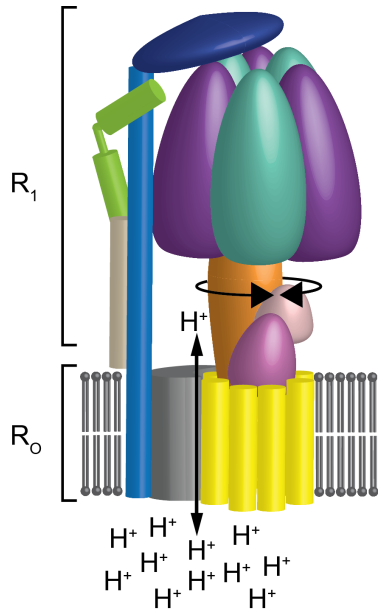


Figure 1-4: The general architecture of the rotary ATPase.

Schematic representation showing the features shared by the rotary ATPase subtypes. All rotary ATPases are composed of a chemically-powered soluble R_1 motor and an electrically-powered transmembrane R_O motor. The parts of the enzyme that rotate during ATP synthesis/hydrolysis are termed the rotor, and are represented by the orange, pink, magenta and yellow subunits. The parts of the enzyme that remain stationary are termed the stator, and are represented by the teal, purple, dark blue, light blue, green, light grey and dark grey subunits. The direction of H^+ translocation and rotation of the rotor are indicated by arrows, and are shown for both ATP synthesis and hydrolysis. Adapted from Lee et al. (2010) Figure 1.[27]

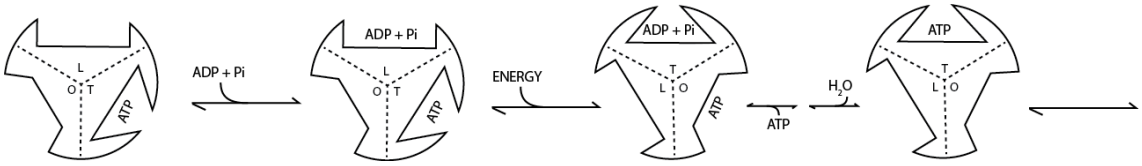


Figure 1-5: The mechanism of rotary catalysis.

Schematic representation of the binding change mechanism proposed by Paul Boyer. In his model, there are three equivalent catalytic sites that undergo sequential conformational changes to synthesise ATP, under conditions of excess ADP and Pi and synthesised ATP is removed. ATP is synthesised in the site labelled T (tight), ADP + Pi bind to the site labelled L (loose) and the site labelled O (open) is empty once ATP has been released. Adapted from Boyer (1989) Figure 7.[28]

1.3.1 The 1997 Nobel Prize

Oxidative phosphorylation, the process of linking oxidation of biomolecules in the mitochondria with ATP synthesis, was discovered by Wladimir Engelhardt in 1932,[29] although the biochemical details of how this occurred were largely unknown for a number of decades. The molecular details of the rotary catalytic mechanism of the rotary ATPase were first hypothesised and later elucidated by Paul Boyer in the 1960s and 1970s. Through a series of experiments with radiolabelled phosphate (containing ^{18}O) and ATP (containing ^{32}P), Boyer was able to show that the release of ATP from mitochondrial extracts could be blocked by chemical uncouplers, which permit H^+ to travel through the mitochondrial membrane without entering the rotary ATPase.[30] This led to the theory that ATP was bound tightly to the catalytic site of ATP synthesis, and energy input would be required to release ATP from the active site, most likely by a conformational change in the enzyme.[31] Further similar experiments coupled with mass spectrometry analysis of the radiolabelled products of ATP hydrolysis led to the conclusion, which Boyer termed the "binding change mechanism", that ATP synthesised by the enzyme cannot be released unless ADP and Pi can bind at a second identical catalytic site in the enzyme.[32,33] It was known at this stage that the catalytic domain of the rotary ATPase contained three β subunits and three almost identical α subunits[34] (Figure 1-4), although there was still significant debate about which and how many of these participated in the catalytic mechanism.[35,36]

In the early 1990s John Walker and colleagues finally resolved the debate by determining the X-ray crystal structure of the $\alpha_3\beta_3\gamma$ sub-complex of the F_1 motor from the bovine mitochondrial ATP synthase, first to 6.5 Å[37] and then to 2.8 Å resolution.[38] In the first structure, the authors observed asymmetry in the arrangement of the $\alpha_3\beta_3$ subunits (they could not be distinguished at this resolution) about a central helix, assumed to be the γ subunit. In the higher resolution structure, the α and β subunits could be distinguished, and the resolution was sufficiently high to visualise bound nucleotides and Mg^{2+}, which the crystallisation solution had been supplemented with (Figure 1-6).[39] It could be seen that each of the α subunits contained a bound molecule of the ATP analogue adenylyl imidodiphosphate (AMP-PNP) along with the β subunit termed β_{TP} (the "tight" site in the binding change mechanism). The second β subunit, termed $\beta_{DP,}$ contained a bound ADP molecule (the "loose" site in the binding change mechanism), and the third, termed β_E contains no bound nucleotide (the "open" site in the binding change mechanism). This structure therefore showed that, as Boyer predicted, the F_1 domain contained multiple identical catalytic sites, and each of the stages of the rotary catalytic mechanism were captured in the high-resolution crystal structure, which also explained the asymmetry seen in the 6.5 Å resolution structure. For their contribution to the "elucidation of the enzymatic mechanism underlying the synthesis of adenosine triphosphate", Boyer and Walker were jointly awarded one half of the 1997 Nobel Prize in Chemistry.[40]

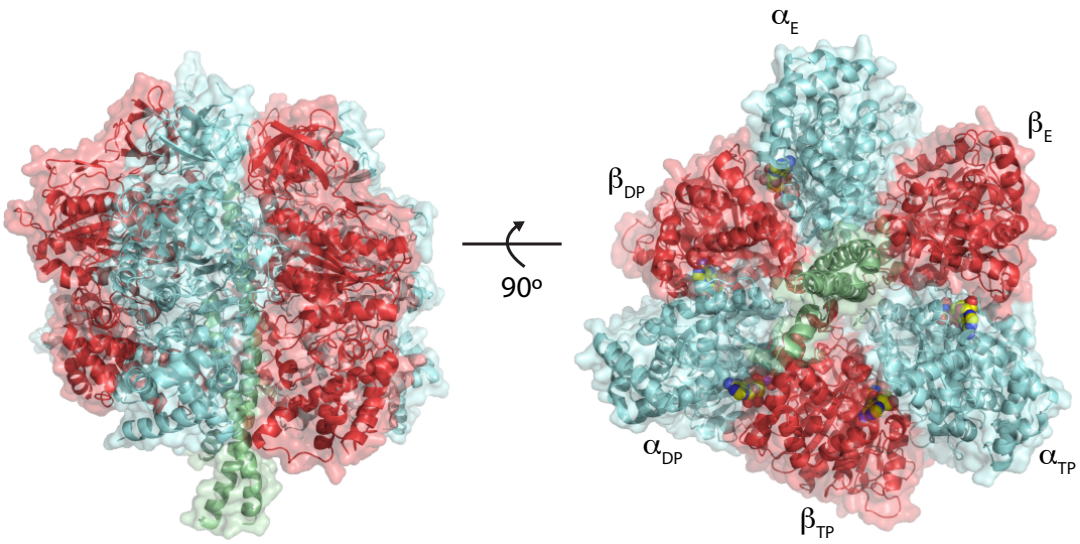


Figure 1-6: F_1-ATPase from bovine mitochondria.

The X-ray crystal structure of the F_1-ATPase from bovine mitochondria at 2.8 Å resolution (PDB 1BMF[37]). The α subunits are coloured red, the β subunits are coloured light blue and the γ subunit is coloured pale green and displayed in cartoon representation. The F_1 complex co-crystallised with nucleotides that are displayed as spheres and coloured by element. An AMP-PNP (non-hydrolysable ATP analogue) molecule was bound to each α subunit and to the β subunit labelled TP, an ADP molecule was bound to the β subunit labelled DP and the β subunit labelled E contained no nucleotide. All figures depicting X-ray crystal structures were made using PyMOL.

1.3.2 Direct observation of F_1 rotation

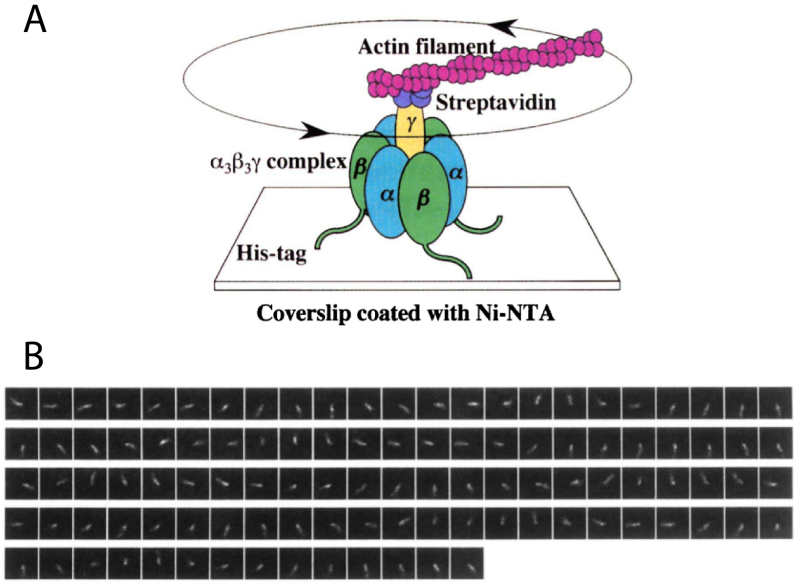
In the paper presenting the 2.8 Å resolution $\alpha_3\beta_3\gamma$ X-ray crystal structure, Walker colleagues proposed that the transition between the different catalytic states the crystal structure may be driven by these subunits rotating relative to the rest of enzyme, although this theory was not immediately accepted. To resolve the about whether the F_1-ATPase did in fact rotate, Hiroyuki Noji and colleagues an experiment to observe the rotation directly.[41] They accomplished this by recombinantly expressing the $\alpha_3\beta_3\gamma$ sub-complex of the *Bacillus* PS3 F-type *Escherichia coli*, with the β subunit expressed as a polyhistidine fusion, and a point mutation was introduced into the γ subunit at position 107 to allow this be biotinylated. The purified $\alpha_3\beta_3\gamma$ complex was then immobilised on a Ni^{2+}-coated glass plate (as Ni^{2+} has high affinity for a polyhistidine tag) and a fluorescently

and biotinylated actin filament was attached to the γ subunit via streptavidin, which has four biotin binding sites (
Figure 1-7A). When ATP was added, the $\alpha_3\beta_3\gamma$ complex could be observed rotating when viewed under an epifluorescence microscope (

Figure 1-7B), although it was shown that it was the γ subunit rotating, which was not what Walker and colleagues had predicted. This provided unambiguous evidence that the ATPase is in fact a rotating molecular machine, and it synthesises or hydrolyses ATP via a rotary catalytic mechanism. Once the rotary catalytic mechanism was established, these enzymes were termed the rotary ATPases to distinguish them from enzymes that hydrolyse ATP by a mechanism distinct from rotary catalysis.

Figure 1-7: Observation of the rotating F_1-ATPase

A: The setup for the single molecule fluorescence experiment that allowed rotation of



the γ subunit of the $\alpha_3\beta_3\gamma$ F_1-ATPase to be observed directly. **B**: Fluorescence images captured of the rotating fluorescently-labelled actin filament attached to the γ subunit. The molecule in each image is viewed from the top of the coverslip (the actin filament is at the top). Length from the rotation axis to the tip of the actin filament, 2.6 μm; rotary rate, 0.5 rps; timer interval between images, 133 ms. Adapted from Noji et al. (1997) Figures 1a and 2a.[41]

1.4 The current picture of rotary ATPases

Structures of the majority of the rotary ATPase subunits and many sub-complexes have now been solved, and medium-resolution electron cryomicroscopy (cryo-EM) structures

of intact complexes are also available.[42-44] While it is now widely accepted that there are three rotary ATPase subtypes, the F-type ATP synthase, the V-type ATPase and the A/V-type ATPase/synthase, these X-ray crystal and cryo-EM structures have revealed they share a similar overall architecture and all couple ion translocation across a membrane to ATP synthesis or hydrolysis in a similar way. The differences are largely confined to function and subunit composition, which will be discussed in the following section, but first the similarities will be discussed to introduce the important structural and chemical features common to all rotary ATPases.

All rotary ATPases contain a soluble catalytic domain composed of a trimer of heterodimers that interacts with a central rotating shaft, termed the central stalk. Depending on whether ATP is being synthesised or hydrolysed, the central stalk rotates clockwise or anticlockwise, respectively, when viewed from above the membrane.[41] The central stalk is attached to a membrane-intrinsic rotor ring (either directly or via another subunit) composed of an oligomer of a varying number of subunits, depending on the enzyme and the organism. The subunit that makes up the rotor ring has been termed a proteolipid as it is soluble in organic solvents.[45] The proteolipid ring is bound to the proton-translocating subunit and the interface between them forms the ion channel through the membrane, either to enable the ion gradient to power ATP synthesis or to pump ions out and create an ion gradient. The soluble and transmembrane domains of the rotary ATPases are also connected by the peripheral stalks, which are composed of elongated coiled-coils. The stalks act to counteract the rotation of the proteolipid ring and central stalk and hold the catalytic domains stationary. The subunits of the enzyme can thus also be grouped according to whether they are rotating or stationary during rotary catalysis (

Table 1-1). The catalytic head domain, the peripheral stalks and the proton-translocating subunit remain stationary and are therefore termed the stator. The central stalk and the membrane ring rotate and are therefore termed the rotor. There are additional stator and rotor subunits in some of the subtypes, and these will be discussed below.

Table 1-1: Subunits and their function in the three rotary ATPase subtypes

Eukaryotic F-ATP synthase	Bacterial F-ATP synthase	V-ATPase	A- ATPase	Motor component	Function
β	β	A	A	Stator	Catalysis
α	α	B	B	Stator	Catalysis
-	-	C	-	Stator	Stator linkage
γ	γ	D	D	Rotor	Axle
OSCP*	δ	E	E	Stator	Membrane-$(AB)_3$ linkage
-	-	F	F	Rotor	Axel/regulatory
δ	ε	-	-	Rotor	Axel/regulatory
ε	-	-	-	Rotor	Axel
b	*b*	G	G	Stator	Membrane-$(AB)_3$ linkage
-	-	H	-	Stator	Regulatory
a	*a*	*a*	I	Stator	Stator element of membrane motor
c	*c*	*c*	L	Rotor	Rotor element of membrane motor
-	-	*d*	C	Rotor	Axel linkage

*OSCP = oligomycin sensitivity conferring protein
Table adapted from Muench et al. (2011) Table 1.[46]

1.4.1 The F-type ATP synthase

The F-type ATPase (Figure 1-8), also distinguished as the F_1F_O-ATP synthase in mitochondria and the F-ATPases in bacteria (F stands for coupling factor, as F_1 was identified as the factor that couples oxidation of biomolecules to phosphorylation of ADP *in vitro*), functions largely as an ATP synthase *in vivo*. The F_1 motor was the first to be studied, and as it was not known originally that F_1 was associated with F_O, only that F_1 hydrolysed ATP, the enzymes were termed ATPases.[47] It was later discovered that these enzymes synthesise ATP when they are powered by the H^+, due to chemisomotic coupling,[12,48,49] and now they are termed F-type ATP synthases. For brevity however, they will be referred to as F-ATPases hereafter.

These enzymes use the energy of the H^+ gradient generated by the electron transport chain to synthesise ATP. This is achieved by the H^+ first entering the enzyme through subunit a, (proton-translocating subunit, for which there is no atomic-resolution structure) which then protonates and neutralises a conserved acidic residue of the membrane rotor ring (c_n-ring, the subunit number varies) to power rotation of the ring.[50] This in turn drives the rotation of the γ subunit (central stalk), which induces sequential conformational changes in the $\alpha_3\beta_3$ subunits (catalytic subunits) that catalyse the synthesis of ATP from ADP and Pi. The F-ATPases contain only one peripheral stalk composed of subunits a, b and OSCP, which spans and connects the soluble and transmembrane regions of the enzyme. While the soluble portion of the peripheral stalk has been structurally characterised to high-resolution,[51] there is no high-resolution structural data for the transmembrane region. Electron cryo-tomography studies of mitochondria have revealed that the F-ATPase dimerises in mitochondria, which induces the inner mitochondrial membrane to form the characteristic curved ridges termed the cristae.[52]

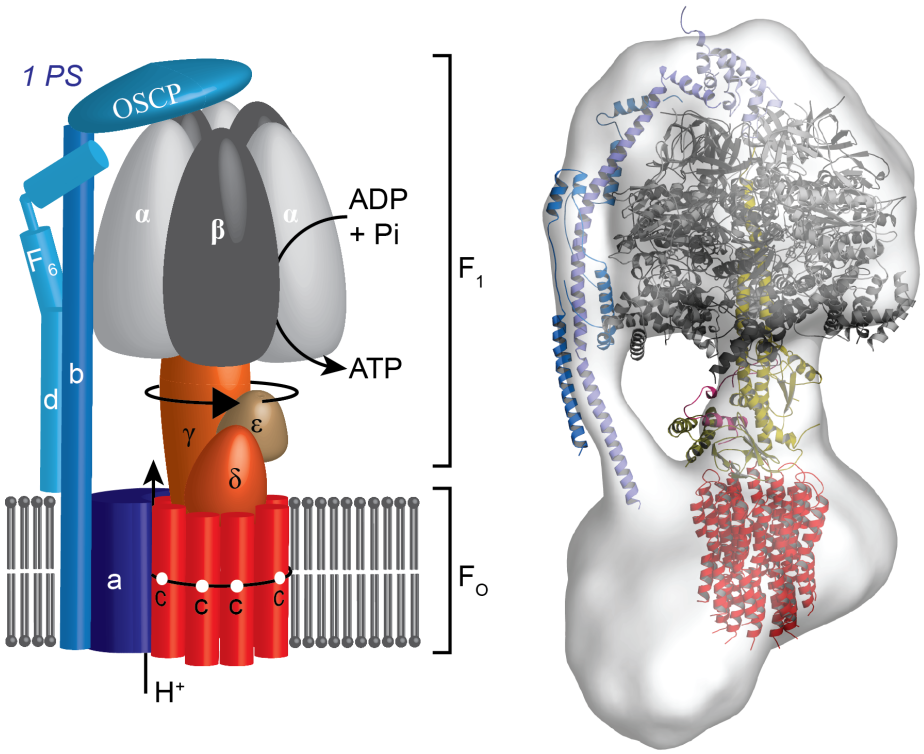


Figure 1-8: Subunit composition and architecture of the rotary F-ATPase.

The image on the left shows a schematic of the eukaryotic F-ATPase. The soluble F_1 motor has the subunit composition $\alpha_3\beta_3\gamma\delta\varepsilon$ along with subunits b (soluble region), F_6 and d of the single peripheral stalk and the regulatory subunit OSCP. The transmembrane F_O motor has the subunit composition a, b (transmembrane region) and the c_n-ring. In the bacterial F-ATPase there is no subunit δ in the central stalk, the peripheral stalk is a homodimer of the b subunit and the regulatory subunit is called δ. Arrows show the direction of H^+ translocation and rotation. The image on the right shows the cryo-EM map of the intact bovine F-ATPase (EMD: 2091)[42] with a fitted model of the yeast F-ATPase generated by subtomogram averaging to show the location of individual subunits (PDB: 4B2Q[52]). Adapted from Stewart et al. (2014), Figures 1 and 2.[53]

1.4.2 The V-type ATPase

The V-type ATPase (V-ATPase) is found in the membranes of a diverse range of eukaryotic organelles, such as vacuoles (hence V-type) and lysosomes, and exclusively hydrolyses ATP to pump H^+ to acidify these organelles (Figure 1-9).[54] These enzymes were first discovered as a series ATP-powered proton pumps in a variety of organisms,[55-58] and when it was discovered they shared identical subunit composition, rotary catalytic mechanism and function, they were identified as a second sub-type of rotary ATPases.[59,60] The V-ATPase A_3B_3 heterotrimer catalyses the hydrolysis of ATP

to ADP and Pi, which induces the DFd central stalk sub-complex to rotate in the opposite direction to the F-ATPase (clockwise).[61] This powers the rotation of the cc'c''$_n$ membrane rotor ring (the number of subunits varies and is composed of three subunit c isoforms) causing H^+ to be pumped into the organelle via the interface of the cc'c''$_n$ ring and subunit a (proton-translocating subunit). There is also no high-resolution structural information for this V-ATPase proton-translocating subunit. The soluble and transmembrane regions of the enzyme are connected by three peripheral stalk heterodimers, formed from subunits E and G. These are anchored to the transmembrane subunits through interactions with the soluble domains of subunits a and subunit C. Subunit H is unique to the V-ATPases and it prevents the V_1 motor from hydrolysing ATP when it has dissociated from the V_O motor, which occurs to prevent unnecessary hydrolysis of ATP.[62] It has been suggested that the peripheral stalks undergo significant bending when they are incorporated into the V-ATPase during enzyme assembly, which would put the enzyme under strain, and thus little energy would be required to trigger disassembly.[63,64] Furthermore, it has been shown that V-ATPase assembly requires the chaperone RAVE (regulator of the ATPase of vacuolar and endosomal membranes), which interacts with the peripheral stalks of dissociated V_1 and may be required to induce these conformational changes in the stalks that allows V_1 to associate with V_O, although this may be isoform-specific.[65]

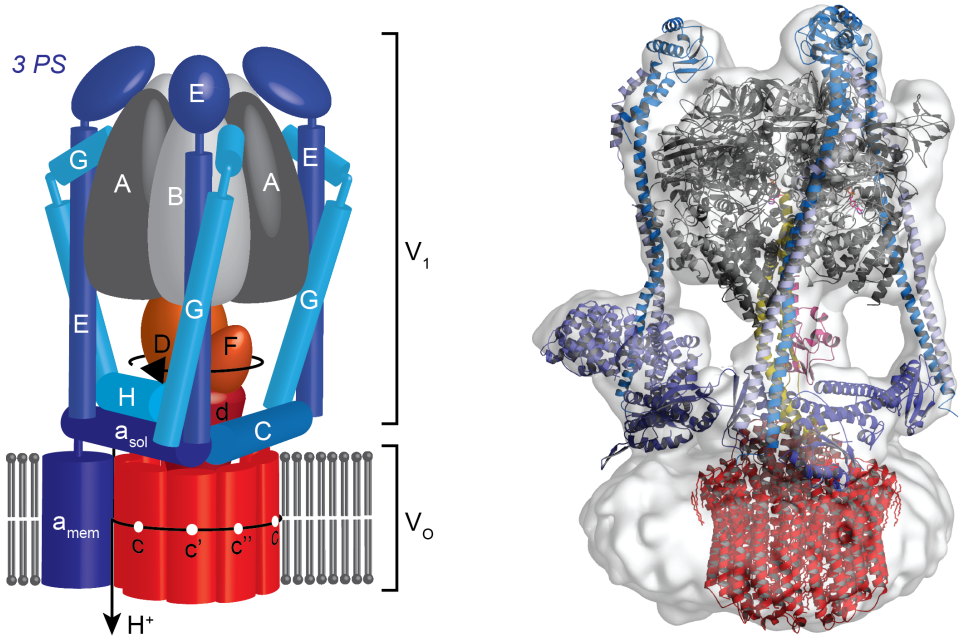


Figure 1-9: Subunit composition and architecture of the rotary V-ATPase

The image on the left shows a schematic of the eukaryotic V-ATPase. The soluble V_1 motor has the subunit composition $A_3B_3CDdE_3FG_3H$ along with soluble domain of subunit a. The transmembrane V_O motor contains subunit the transmembrane domain of subunit a and the cc'c''$_n$-ring. Arrows show the direction of H^+ translocation and rotation. The image on the right shows the cryo-EM map of the intact yeast V-ATPase (EMD: 5476)[43] with fitted X-ray crystal structures to show the positions of individual subunits (PDBs: 3A5C,[66] 3AON,[67] 1R5Z,[68] 2BL2,[69] 4DL0,[63] 3RRK,[70] 1U7I[71] and 1HO8[72]). Adapted from Stewart et al. (2014) Figures 1 and 2.[53]

1.4.3 The A-type ATPase

The A-type ATP synthase (A-ATPase) is found in archaea and in some bacteria, in which it is sometimes called the bacterial V-ATPase. This is because the A-ATPase is more similar to the V-ATPase than the F-ATPase in architecture and subunit composition (Figure 1-10), although the A-ATPase primarily synthesises ATP.[73] When it was established that the A-ATPases are predominantly synthases, and that their subunit composition differs from that of the V-ATPase, it was widely agreed that they represent a distinct subtype of rotary ATPases.[74] The A-ATPases function similarly to the F-ATPases, whereby H^+ first enter the enzyme via the interface between the proton-translocating subunit (subunit I, for which there is also no high-resolution structural information) and the membrane rotor ring (L_n-ring), powering the rotation of the L_n-ring. The ring is connected to the central stalk (subunits D and F) via subunit C,

the equivalent of subunit d in the V-ATPase, and the rotation of the L_n-ring induces these subunits to rotate, which leads to conformational changes in the A_3B_3 heterotrimer and ATP synthesis. Two peripheral stalks connect the A_1 and A_O motors, via two B subunits and the soluble domain of subunit I. Unlike the V-ATPase, the A-ATPase has not been shown to reversibly dissociate, most likely because it lacks most of the subunits that mediate dissociation in the V-ATPase.

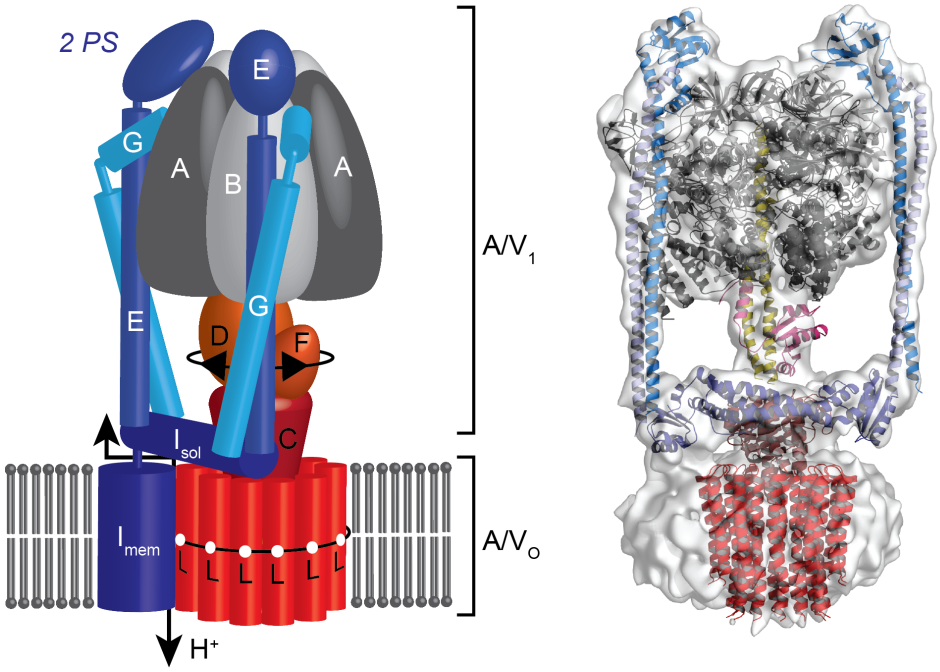


Figure 1-10: Subunit composition and architecture of the rotary A-ATPase.

The image on the left shows a schematic of the archaeal/bacterial A/V-ATPase. The soluble A_1/V_1 motor has the subunit composition $A_3B_3CDE_2FG_2$ along with soluble domain of subunit I. The transmembrane A_O/V_O motor contains subunit the transmembrane domain of subunit I and the L_n-ring. Arrows show the direction of H^+ translocation and rotation. The image on the right shows the cryo-EM map of the intact *T. thermophilus* A-ATPase (EMD: 5335)[44] with fitted atomic models (PDB: 3J0J)[44] and X-ray crystal structures (PDBs: 3RRK[70] and 1C17[23]) to show the positions of individual subunits. Adapted from Stewart et al. (2014) Figures 1 and 2.[53]

1.5 Proton translocation through R_O

As detailed in section 1.3, a number of key experiments were instrumental in understanding the mechanism of the F_1 motor of the rotary ATPase at the molecular level, and subsequent experiments that built on these findings have enabled all

mechanical and chemical features of how the R_1 motor synthesises and hydrolyses ATP to be understood in intricate detail. The same is not true, however, for the R_O motor. As this part of the enzyme is located within the membrane, structural studies of the intact motor have thus far been limited to cryo-EM, due to the challenging nature of membrane protein X-ray crystallography. At the commencement of this project the only high-resolution structures available for R_O subunits were X-ray crystal structures of the membrane rotor ring from a number of different organisms (Figures 1-8 through 1-10).[75-78]

The theory that life evolved around H^+ gradients was introduced in section 1.1, and all forms of life continue to depend on H^+ gradients to power essential biochemical processes. Therefore, along with completing the structural model of rotary ATPases, determining the structure of the proton-translocating subunit would complete the molecular understanding of one of the most fundamental processes of biological energy conversion: how rotary ATPases can harness the power of the H^+ gradient to synthesise ATP, or hydrolyse ATP to generate a H^+ gradient. Nevertheless, the absence of structural information has not entirely thwarted efforts to understand the proton-translocating subunit, and a number of important biochemical analyses, which will be discussed below, have enabled a model of H^+ translocation to be developed (this will be discussed only for ATP synthesis, not hydrolysis).

1.5.1 Biochemical and structural analyses of the R_O proteolipid ring

As there are numerous high-resolution X-ray crystal structures of the rotor ring from different organisms, there is now a detailed understanding of how H^+ powers rotation of the ring through the membrane. A series of mutation studies in *E. coli* identified that there is an acidic residue in the proteolipid c subunit (Asp61) that is required for H^+ translocation through the ring during ATP synthesis.[50,79,80] Sequencing of ATP synthase genes from other sources such as spinach chloroplasts,[81] bovine mitochondria,[82] *Bacillus* sp. PS3[83] and *Saccharomyces cerevisiae*[84] revealed that an acidic residue (in some organisms it is a glutamate rather than an aspartate) is always found at the equivalent position in the proteolipid subunit, suggesting the mechanism of H^+ translocation through the membrane via the rotor ring is conserved. It was also found that the small molecule N,N-Dicyclohexylcarbodiimide (DCCD) covalently binds to

this conserved acidic residue and blocks H^+ translocation.[50] From these findings it was proposed that the requirement for an acidic residue at this position suggests that this residue could be protonated and transport H^+ from one side of the membrane to the other.[85] Rotation of the ring had been directly observed during a similar experiment[86] to that performed with the R_1 motor, and protonation would be necessary to neutralise the acidic residue so it could rotate into the hydrophobic environment of the membrane. For H^+ to access the acidic residue on one side of the membrane and dissociate from the residue on the other side, there would need to be an aqueous channel somewhere in the membrane region of the enzyme. Through cysteine cross-linking experiments, it was shown that there is a direct interaction between subunit a and the proteolipid ring, and it was proposed that the interface between them forms the proton channel.[87] As a result subunit a (and the equivalent subunits in the other rotary ATPase subtypes) was termed the proton-translocating subunit, as it is not strictly a proton channel.

1.5.2 Biochemical characterisation of the proton translocating subunit

In the absence of high-resolution structural information, a wealth of biochemical data has been accumulated to identify functionally important residues of the proton-translocating subunits, propose the position of residues relative to the rotor ring, and develop a model of H^+ translocation. The first essential residue to be identified was Arg210 of the *E. coli* subunit a, as mutation of this residue to any other residue completely abolished ATP synthesis.[88-91] This residue is absolutely conserved in all proton-translocating subunits and does not tolerate mutation in any organism. As arginine is a basic residue, it was suggested that it could form a salt bridge with the conserved acidic residue of one of the proteolipid subunits to stabilise the acidic residue in the ionised form prior to protonation.[92] Through similar mutation studies it was also found that Glu219 and His245 of the *E. coli* subunit a were essential for H^+ translocation, along with Gln252, although this is more tolerant to mutation.[89,93] These residues are now known to be conserved in all proton-translocating subunits. When the double H219E and E245H mutant was made (swapping the positions of Glu219 and His245), the activity of the enzyme was only slightly lower than wildtype, and higher than when each of the mutations were made individually, suggesting this particular pair of residues is required for H^+ translocation.[94]

Secondary structure prediction calculations had indicated that the *E. coli* subunit a has an α-helical fold with 5 transmembrane α helices.[95] Further evidence for this number was provided by small molecule accessibility studies, whereby residues thought to be in aqueous regions of the protein were mutated to cysteine, and accessible cysteine residues were labelled with small molecules such as fluorescein 5-maleimide or N-ethyl-maleimide.[96,97] Modelling the sequence of subunit a as α helices that traverse the membrane would place the essential Arg210, Glu219, His245 and Glu252 in close proximity,[92] and they have been shown to be in a region of the protein that cross-links with the membrane rotor ring, which would place these residues near Asp61 of the rotor ring subunit c.[87] Based on all of these findings, a mechanism of H^+ translocation termed the two half-channel model has been proposed, and it is now widely accepted as the mechanism by which H^+ translocation powers ATP synthesis.

1.6 The two half-channel model of proton translocation

In this model, developed independently by Wolfgang Junge[98] and Steven Vik,[92] H^+ first enter the proton-translocating subunit from the low pH side of the membrane, i.e. from the mitochondrial inner membrane space in eukaryotes or the periplasm in prokaryotes (Figure 1-11). This half channel is predicted to be formed by the interface of the proton-translocating subunit and the membrane rotor ring, and would contain charged or polar residues, as they can be reversibly protonated. This would allow H^+ to access the conserved acidic residue of the membrane rotor ring subunit in the middle of the membrane. Prior to protonation this residue is stabilised by an interaction with the conserved Arg residue of the proton-translocating subunit. Protonation of the acidic residue disrupts this interaction, and as the acidic residue has been neutralised by the H^+, it can rotate into the hydrophobic environment of the lipid bilayer. The protonated acidic residue is then able to travel through the membrane to the other side until it reaches the proposed second half-channel, which would be accessible to the mitochondrial matrix or the cytoplasm. The combination of the high pH of this channel and the presence of the Arg residue of the proton-translocating subunit, which would lower the p*K*a of the acidic residue,[97] would lead to the H^+ dissociating from the acidic residue and exiting to the matrix or cytoplasm via this half channel. The acidic residue would then be free to be protonated again, allowing the rotation of the membrane rotor

ring to continue. This would drive the central stalk of the enzyme to rotate, inducing the conformational changes in the catalytic subunits that result in the synthesis of ATP (Figure 1-4).

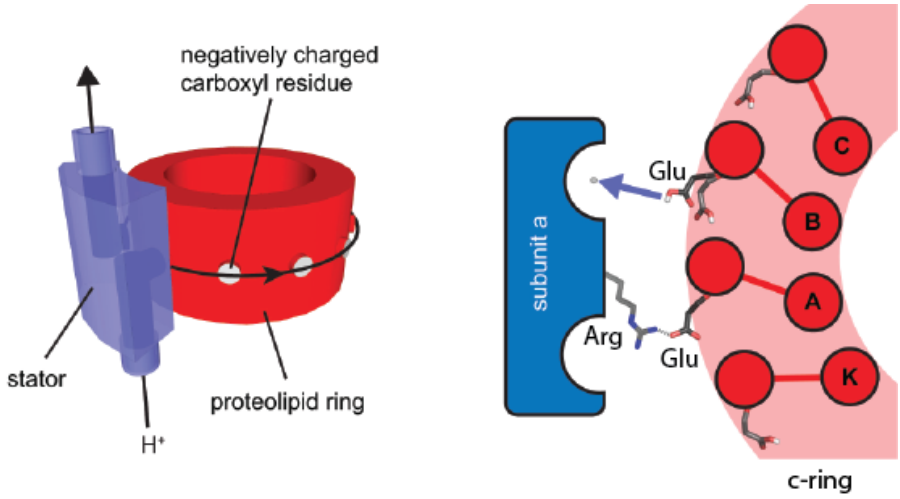


Figure 1-11: Two half-channel model of proton translocation.

The figure on the left shows a schematic of the interaction between the proton-translocating subunit (blue) and the membrane rotor ring (red), the interface of which forms the two half-channels through which H^+ enter and exit the membrane (indicated by the black arrows; the direction corresponds to ATP synthesis). The protonatable acidic residues of the membrane rotor ring are represented as white circles. The right image, viewed from above the membrane, is a schematic showing the interaction of the conserved Arg residue (subunit a, blue) interacting with a conserved acidic residue (Glu for the c-ring, the subunits are labelled A-K). A H^+ dissociating from the Glu residue (causing the Glu residue to adopt a different rotamer) in the matrix/cytoplasmic half-channel is shown as a white sphere, and a deprotonated Glu residue is shown to form an electrostatic interaction with the Arg residue in the inner membrane space/periplasmic half-channel. Adapted from Stewart et al. (2014) Figure 3.[53]

There are a number of important features of this model that elegantly account for the molecular details of H^+ translocation through a membrane. The first of these is that rotation of the rotor ring is driven by Brownian thermal fluctuations, and thus the membrane rotor ring has also been termed a Brownian ratchet. As there is no directionality to thermal fluctuations, the direction of rotation is determined by the probability of the acidic residue being protonated in each of the two half-channels.[98] As the channel accessible to the inner membrane space/periplasm has a lower pH than the channel accessible to the mitochondrial matrix/cytoplasm, the acidic residue is most likely to be protonated in the inner membrane space/periplasmic half-channel, which would impose directionality on the rotation. Another important feature of this model is that, as the name suggests, there are two half-channels rather than a single channel. In

this model, the two half-channels are physically separated by the Arg residue of the proton-translocating subunit.[92] If there were a single continuous channel, H^+ would be able to travel from one side of the membrane to the other without needing to traverse the membrane rotor ring, which would uncouple ATP synthesis from H^+ translocation. The two half-channel model therefore offers an explanation for how H^+ translocation is both physically and chemically coupled to ATP synthesis. Yet despite the plausibility of this model, it is important to note that it is a model, and confirmation requires high-resolution structural information.

1.7 Structural evidence for the two half-channel model

Immediately prior to the beginning of this project two cryo-EM structures of F-ATPases, one of the intact bovine F-ATPase[42] and one of the bacterial F_O motor obtained from 2D crystals,[99] along with a cryo-EM structure of the intact *Thermus thermophilus* A-ATPase,[44] were all determined to resolutions higher than 20 Å. These cryo-EM maps provided the first intermediate-resolution structural evidence for the two half-channel model of H^+ translocation.

1.7.1 Intact bovine F-ATPase cryo-EM map to 18 Å resolution

In the bovine F-ATPase map, which was built to 18 Å resolution (Figure 1-8), a ring of density could be seen in the membrane-embedded region of the enzyme, which was identified as the membrane rotor ring, and adjacent to this was another region of density assumed to be the proton-translocating subunit, as this would form the proton-conducting path through the membrane (Figure 1-12). The authors noted that there was minimal contact between the c-ring and subunit a, with the major region of contact being in the middle of the membrane, where the Arg-Glu (the bovine subunit c contains a Glu rather than an Asp) interaction is predicted to occur, and they proposed that the spaces between subunit a and the c-ring, above and below the middle of the membrane, could be the sites of the two half-channels. The authors also suggested that the minimal contact between the c-ring and subunit a would facilitate rotation of the c-ring, as if there were numerous interactions between the c-ring and subunit a that

would need to be disrupted and reformed, this would most probably slow rotation and ATP synthesis.

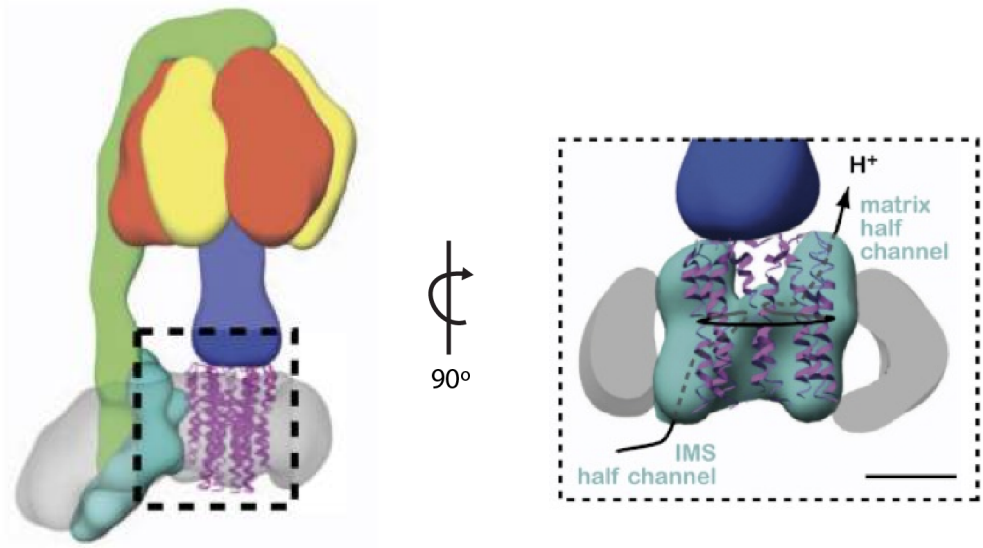


Figure 1-12: Cryo-EM map of the bovine F-ATPase to 18 Å resolution showing possible location of the two half-channels.

The image on the left shows the cryo-EM map of the bovine F-ATPase to 18 Å resolution with subunits coloured as segments and the X-ray crystal structure of the c-ring fitted to the density. The image on the right shows the segment corresponding to subunit a (cyan) interacting with the c-ring. The helices of subunit a appear to be arranged into two bundles and each bundle interacts with a different subunit of the c-ring, which the authors propose represents the two half-channels through which H^+ translocate. Adapted from Baker et al. (2012) Figure 4.[42]

1.7.2 Bacterial FO motor projection map from 2D crystals to 7 Å resolution

The Na^+-translocating F_O rotor from the bacterium *Ilyobacter tartaricus* was isolated and crystallised in two dimensions. The 2D crystals were subjected to cryo-EM analysis and a 7 Å projection electron density map was calculated (Figure 1-13). In this map strong c-ring electron density can be seen, and next to each of the rings are additional electron density peaks. When these additional peaks were aligned and averaged, the density was enhanced relative to the background, suggesting this density was not noise. Following non-crystallographic symmetry (NCS; symmetry that is additional to that imposed by the crystallographic space group, as there were multiple copies of the c-ring in the asymmetric unit) averaging, these additional peaks can be seen to form a

four-helix bundle, which is consistent with the three-dimensional helical arrangement previously predicted for the *E. coli* subunit a.[100] Subunit a from the *I. tartaricus* enzyme is predicted to have six transmembrane helices, and there are additional helices in the electron density that could belong to subunit a, however they could not be unambiguously assigned as they were not seen associated with every copy of the c-ring following averaging. It is possible that these helices were present in multiple different conformations, and therefore would have had weaker electron density. Nevertheless, it was acknowledged that the four-helix bundle could provide at least one of the proposed half channels, as the interface between the c-ring and subunit a in the crystal was large enough to accommodate an aqueous environment for Na^+ entry/exit.

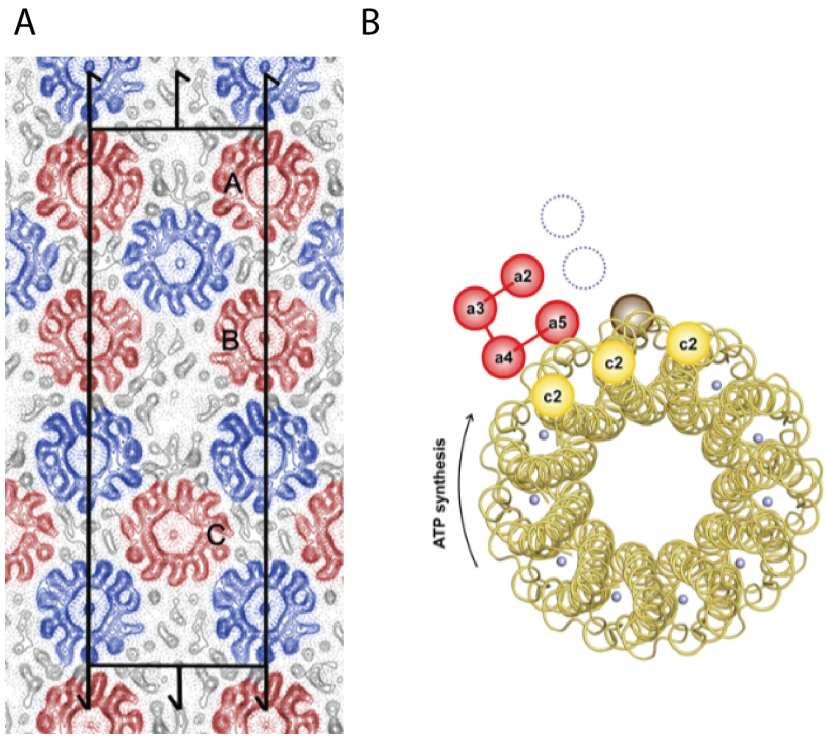


Figure 1-13: Electron density projection map of the F_O motor from *I. tartaricus*.

A: Cryo-EM density projection map of 2D crystals containing the F_O of the *I. tartaricus* F-ATPase at a resolution of 7 Å with symmetry $p12_1$ applied. The rectangle indicates the crystal unit cell and the arrows represent the screw axes parallel to the membrane plane. The C-rings labelled A, B and C correspond to the asymmetric unit, and the rings are coloured red and blue according to their orientation. Additional densities to the c-rings are coloured grey and were assigned to the stator. **B:** Assignment of helices observed in the crystal structure to F_O subunits. After NCS averaging, up to seven additional α-helices were observed in the density. Four of these formed a four-helix bundle and were assigned to subunit a (red). Adapted from Hakulinen et al. (2012) Figures 2 and 4.[99]

1.7.3 Intact *T. thermophilus* A-ATPase cryo-EM map to 9.7 Å resolution

The 9.7 Å resolution cryo-EM map of the intact *T. thermophilus* A-ATPase was the highest resolution structure available for an intact rotary ATPase at the time this project began. Significantly, at this resolution individual helices could be modelled, including the helices of the proton-translocating subunit (subunit I; Figure 1-14). Eight transmembrane α-helices (green) adjacent to the helices of the L-ring (magenta) were observed, and these were assigned to subunit I. As was seen in the *I. tartaricus* motor, the helices assigned to subunit I also cluster into four-helix bundles, although in subunit I there are two bundles rather than one. Cross-sections were taken at different heights through the A_O region, and these show that each of the four-helix bundles contacts a different subunit of the L-ring, and these different contact positions would place each of the L-ring subunits in different chemical environments. This is consistent with the requirements of the two half-channel model. The authors therefore hypothesise that the four-helix bundle closest to the periplasm would create the H^+-entry half channel, while the four-helix bundle closest to the cytoplasm would create the H^+-exit half-channel. Nevertheless, at this resolution individual amino acid residues cannot be visualised, and identification of the residues that line the two theoretical half-channels is required to prove this hypothesis.

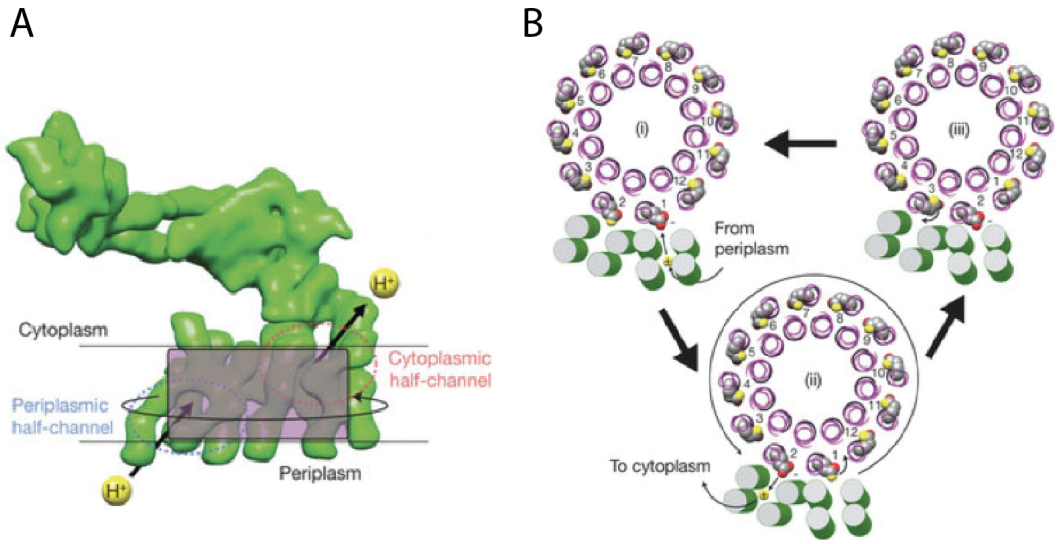


Figure 1-14: Evidence for the two half-channel model in the 9.7 Å resolution *T. thermophilus* A-ATPase cryo-EM map.

A: The cryo-EM desnity corresponding to *T. thermophilus* subunit I (green) is shown parallel to the membrane and the semi-transparent magenta rectangle shows the interface of the interaction with the L_{12}-ring. The sites of H^+ entry and exit are shown as blue and red dashed circles, respectively. The direction of rotation of the L_{12}-ring is shown with a black arrow. **B**: The two four-helix bundles of subunit I (green cylinders) could form the two half-channels for H^+ translocation. Deprotonated Glu residues of the L_{12}-ring are shown with two red oxygen atoms, whereas protonated Glu residues are shown with one red oxygen and one yellow oxygen atom. Protonation of L-subunit "1" at the periplasmic half-channel and deprotonation of subunit "2" allows the ring to rotate. Subunit "1" then enters the membrane, subunit "2" moves into the periplasmic half-channel to be protonated and subunit "3" moves into the cytoplasmic half-channel to be deprotonated. This sequence is repeated with each rotation. From Lau and Rubinstein (2012) Figure 4.[44]

As the experimental work presented in this thesis was being completed, three high-resolution cryo-EM structures and one high-resolution X-ray crystal structure of intact rotary ATPases were published. The findings from these papers will be detailed in Chapter 5, as the majority of the experimental work presented in this thesis was conducted before these structures were available, and these papers did not influence the experimental approach that was pursued.

1.8 Mass spectrometry analysis of intact *T. thermophilus* A-ATPase

The intact *T. thermophilus* A-ATPase was analysed by mass spectrometry, and peaks could be assigned to the intact complex, indicating this is a suitable technique with which to study membrane proteins.[101] At high activation energies, the intact enzyme can be released from the detergent micelle in the gas phase, where it dissociates into the A_1 and A_O sub-complexes. The intact enzyme was also subjected to mass spectrometry with the addition of 50 μM ATP, and it was found that the intact complex and the $A_O(ICL_{12})$ sub-complex were the major species formed. The peaks corresponding to the ICL_{12} complex were found to contain additional masses corresponding to up to two ATP molecules (Figure 1-15). When ATP was depleted, however, subunit I was lost from the ICL_{12} sub-complex, suggesting subunit I could bind to ATP. This was investigated further in a follow up paper, which will be discussed in 5.1.

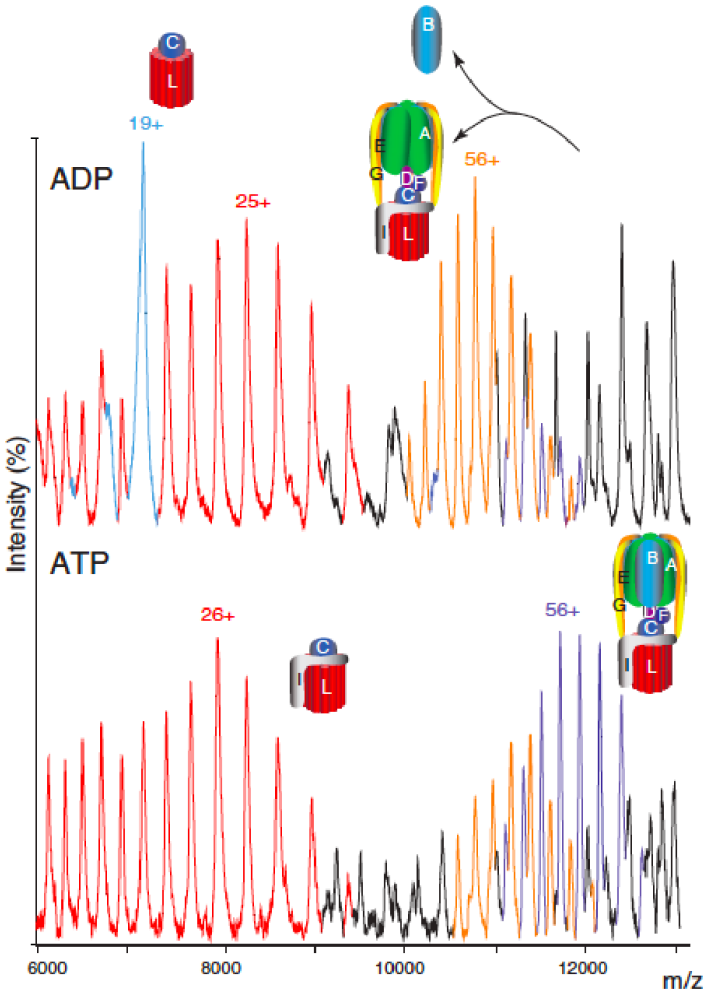


Figure 1-15: Mass spectra of intact *T. thermophilus* A-ATPase with the addition of nucleotides.

When ATP is depleted from the *T. thermophilus* A-ATPase and replaced with ADP, subunit B and subunit I are lost from A_1 and A_O, respectively (upper panel). When the complex is supplemented with 50 µM ATP, subunit B and subunit I are retained (lower panel). From Zhou et al. (2011), Figure 3.[101]

1.9 The X-ray crystal structure of the soluble domain of subunit I

In 2011 the X-ray crystal structure of the soluble domain (residues 1-344) of subunit I (I_{sol}) from the bacterium *Meiothermus ruber* was determined to 2.6 Å (Figure 1-16).[70] The structure contains a central bundle of four α-helices that are bent in the middle, with two lobes that flank the helices at either end, both of which have a similar α/β architecture. The lobe labelled proximal is very similar to the structure of the "head" domain of the *S. cerevisiae* V-ATPase subunit C, which was co-crystallised with one of the V-ATPase peripheral stalk heterodimers.[63] This indicates that the two lobes of I_{sol}

likely contain the binding sites for the two peripheral stalks of the A-ATPase (Figure 1-10C). The *M. ruber* I_{sol} structure was fitted into the 9.7 Å *T. thermophilus* A-ATPase cryo-EM density with very good agreement, indicating this structure adopted a similar fold to that of the soluble domain of the *T. thermophilus* subunit I in the context of the intact A-ATPase.[44]

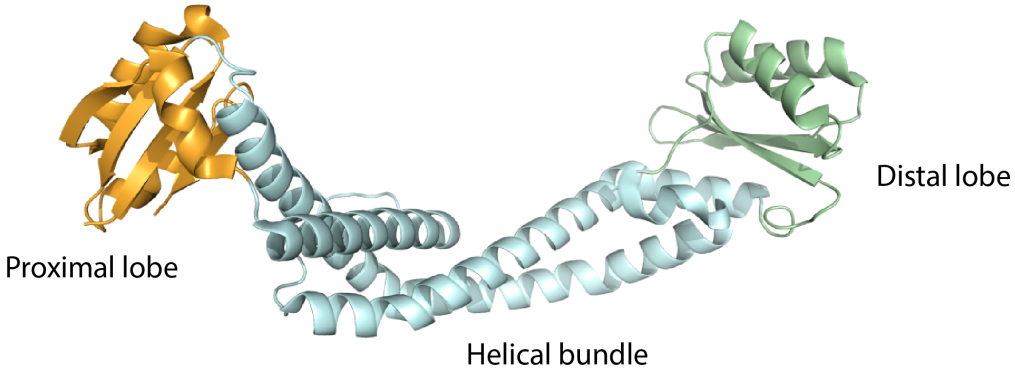


Figure 1-16: X-ray crystal structure of the soluble domain of *M. ruber* subunit I to 2.6 Å resolution.

The X-ray crystal structure of residues 1-344 of subunit I from the *M. ruber* A-ATPase was solved to a resolution of 2.6 Å (PDB: 3RRK[70]). The structure is composed of two lobes (labelled proximal and distal, yellow and blue, respectively), each with an αβ fold, that flank a central helical bundle.

1.10 The A-ATPase as a model for structural characterisation of rotary ATPases

As the A-ATPases are found in thermophilic organisms, they are naturally more tolerant to higher temperatures than rotary ATPases from organisms that grow at lower temperatures, making the A-ATPases attractive structural targets. This is also supported by the fact that the highest resolution structure obtained for an intact rotary ATPase prior to the commencement of this project was the structure of the *T. thermophilus* A-ATPase. The structure of the *T. thermophilus* A-ATPase is therefore now quite well understood, and as the general architecture of all rotary ATPase sub-types is quite similar (Figure 1-4), the *T. thermophilus* enzyme could serve as a model to structurally characterise rotary ATPases, and more specifically, to assign the path of H^+ translocation through the membrane-embedded region of the enzyme.

Previous work in this laboratory had established a reproducible protocol for the purification of the intact *T. thermophilus* A-ATPase from *T. thermophilus* cultured in-house, yet no crystals had been obtained. It was thought that determining the structure of subunit I alone would be a more achievable goal, and constructs for the recombinant expression and purification of subunit I alone had also been tested, although these gave poor yields and the protein was prone to aggregation and degradation. This is largely because membrane proteins have evolved to be most stable within the environment of their native lipid bilayer, yet purification and structural characterisation of membrane proteins requires them to be extracted from this environment into a solubilising agent, usually a detergent, which can be highly destabilising.[102] An added complication of this, which was acknowledged in the *I. tartaricus* F_O 2D crystal structure paper, is that the proton-translocating subunit is quite flexible, which could impede attempts to obtain a high-resolution structure. It was therefore decided to adopt a high-throughput approach to clone and test homologues of subunit I from 11 additional organisms (*T. thermophilus* subunit I would be included as a control), to increase the chance of identifying a homologue of subunit I that is naturally more stable than the *T. thermophilus* protein. These genes would be cloned into eight different expression vectors that had been designed in the laboratory, giving a total of 96 expression constructs to test and potentially optimise for structural characterisation of the proton-translocating subunit of the A-ATPase.

1.11 Aims of this project

1.11.1 Identify and clone eleven homologues of the *T. thermophilus* subunit I

The first aim of this thesis was to identify eleven homologues of the *T. thermophilus* subunit I gene and clone them into the eight different expression vectors to give 96 possible expression constructs to test.

1.11.2 Test the expression of all constructs and select a candidate for optimisation

Once all of these constructs had been generated, the expression of subunit I would be tested in small-scale cultures, and these would be subjected to a small-scale purification to determine whether full-length protein could be purified.

1.11.3 Optimise the expression and purification of the best candidate

The most promising constructs from the small-scale trials would then be expressed and purified on a larger scale, to determine whether sufficient quantities of purified, stable protein could be produced for crystallisation trials. ATP would also be tested as a stabilising agent at this stage.

1.11.4 Obtain protein suitable for crystallisation trials

If pure subunit I could be obtained that was monodispersed and not prone to degradation, the protein would be subjected to crystallisation trials to determine whether diffraction-quality crystals could be grown, which would enable structure determination.

1.11.5 Investigate the interaction between subunit I and the peripheral stalks

As can be seen in Figure 1-10C, the soluble domain of subunit I interacts with the two peripheral stalks of the A-ATPase, yet little is known about this interaction at the molecular level. The final aim of this thesis was to study this interaction by isothermal

titration calorimetry (ITC) and multi-angle laser light scattering (MALLS) to determine the affinity of the interaction and the *in vitro* stoichiometry of the sub-complex.

Chapter Two: Cloning and construct selection

2.1 Selection of 12 subunit I homologues

A position-specific iterative basic local alignment search tool (PSI-BLAST) search was performed using the sequence of the *T. thermophilus* subunit I gene (TTHA1278) as the input. 11 sequences were chosen from the second iteration; nine of the genes were from organisms for which genomic DNA had already been purchased (*Sulfolobus solfataricus, Archaeoglobus fulgidus, Thermoplasma acidophilum, Pyrococcus furiosus, Methanosarcina mazei, Pyrococcus horikoshii, Deinococcus radiodurans, Methanocaldococcus jannaschii, Methanococcus maripaludis*), one was from *M. ruber* for which the crystal structure of the soluble domain had already been solved (section 1.9), one was from *Thermococcus gammatolerans*, an archaeon that possesses the highest known resistance to radiation of any organism, which could be useful if X-ray diffraction data were to be collected[103], and the *T. thermophilus* gene was included as a control.

The genes from *T. thermophilus* and *P. horikoshii* were successfully amplified from the genomic DNA, however the remaining seven organisms for which genomic DNA was available could not be amplified. The operons encoding all of the rotary A-ATPase genes from the nine organisms for which genomic DNA was available had previously been cloned into other expression vectors, so these were used as the templates for PCRs to amplify the subunit I genes, enabling the remaining seven genes from this group to be successfully amplified. Synthetic genes encoding subunit I from *M. ruber* and *T. gammatolerans* were ordered, and these were also successfully amplified.

2.2 Sequence- and Ligation-Independent Cloning (SLIC)

SLIC reactions were performed to clone the 12 subunit I homologues into the eight expression vectors (Table 0-2), and the reactions were transformed into NEB 5-alpha *E. coli* cells. Colonies that grew on the transformation plates were screened for the correct insert by colony PCR. A product was seen at the correct size for 94 of the 96 possible constructs (no colonies grew for the *M. maripaludis* gene in vectors B and C).

Restriction digests with *EcoRI* and Sanger sequencing (AGRF) were performed to confirm that each construct contained the correct insert.

2.3 Small scale expression trials and analysis

Expression of the 94 constructs was tested by transforming them into T7 Express and C41 *E. coli* cells and growing 3 mL cultures in LB medium, followed by a small-scale Ni-NTA affinity purification. All of the constructs for which expression was observed had been expressed in T7 Express cells; no expression was seen in C41 cells for any of the 94 constructs. The presence of intact subunit I was confirmed by Western blots using horseradish peroxidase (HRP)-conjugated antibodies against one of the two tags (Figure 2-1A-G; I) or by measuring in-gel green fluorescent protein (GFP) fluorescence (Figure 2-1H). While most constructs either gave no expression or were substantially degraded, full-length expression was seen for eight of the 94 constructs (*S. solfatricus, T. thermophilus, M.ruber, T. gammatolerans* in vector A, *M. ruber* in vector F, *T. thermophilus, M. ruber, T. gammatolerans* in vector H). The subunit I gene from *M. ruber* routinely gave the highest expression and the highest yield of intact protein.

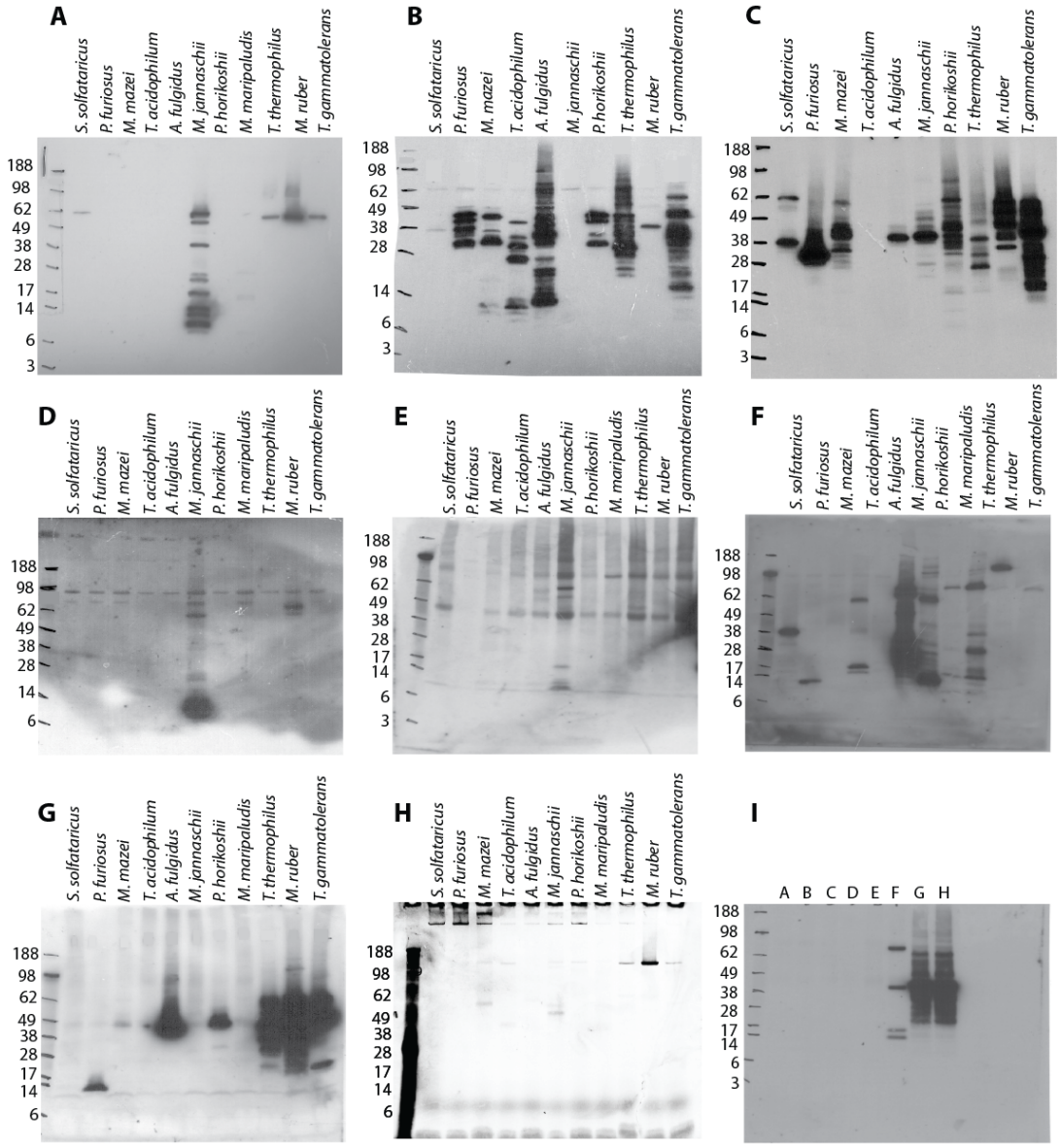


Figure 2-1: Small-scale expression and purification trials for 94 subunit I expression constructs.

Each of the subunit I expression constructs was expressed in 3 mL *E. coli* cultures and captured by cobalt affinity. The cobalt resin was run on SDS-PAGE, and protein was detected by a Western blot against either the N- or C-terminal tag using a HRP-fused antibody. **A**: Anti-FLAG antibody against the C-terminal FLAG tag. **B**: Anti-FLAG antibody against the N-terminal FLAG tag. **C**: Anti-T7 antibody against the N-terminal T7 tag. **D**: Anti-His antibody against the C-terminal His_{10} tag. **E**: Anti-His antibody against the C-terminal His_{10} tag. **F**: Anti-His antibody against the C-terminal His_{10} tag. **G**: Anti-His antibody against the N-terminal His_{10} tag. **H:** In-gel GFP fluorescence. **I**: Anti-His antibody against the N- and C-terminal His_{10} tags. Proteins were visualised by ECL, or in-gel GFP fluorescence was measured using an excitation wavelength of 473 nm and a 510 nm emission filter. The sizes (in kDa) of the SeeBlue Plus2 pre-stained protein markers are labelled on the left of each image.

2.4 Large scale expression and purification trials

2.4.1 *M. ruber* subunit I in vector H

The first construct to be trialled on a larger scale was the *M. ruber* gene in vector H (His_{10}-subunit $I_{M.ruber}$-GFP), as using GFP fluorescence to monitor membrane proteins during purification has been established as a successful screening strategy.[104] A small detergent screen was performed to identify a detergent that gave a high yield of His_{10}-subunit $I_{M.ruber}$-GFP. The protein was solubilised in one of four detergents and immobilised on Ni-NTA resin that was then run on SDS-PAGE, and intact protein was detected by measuring in-gel GFP fluorescence (Figure 2-2). From this screen it was found that Triton X-100 gave the highest yield of protein, although it appeared most of this had precipitated on the resin.

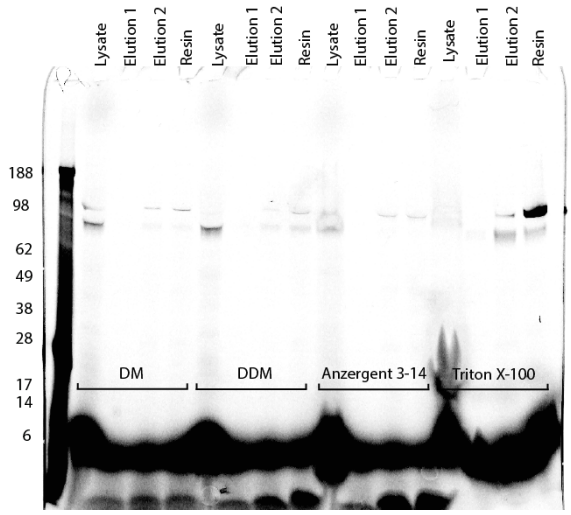


Figure 2-2: Small-scale detergent screen and purification with His_{10}-subunit $I_{M.ruber}$-GFP.

His_{10}-subunit $I_{M.ruber}$-GFP was extracted and purified from 50 mL *E. coli* cultures in one of four detergents (*n*-decyl-β-D-maltoside (DM), *n*-dodecyl-β-D-maltoside (DDM), Anzergent 3-14, Triton X-100) by cobalt affinity. The cell lysate, cobalt elutions and cobalt resin were resolved by SDS-PAGE and intact His_{10}-subunit $I_{M.ruber}$-GFP was detected by measuring the in-gel GFP fluorescence, with an excitation wavelength of 473 nm and a 510 nm emission filter.

The expression of His_{10}-subunit $I_{M.\ ruber}$-GFP was scaled up from 50 mL to 4 L to test whether increasing the culture size could increase the amount of stable protein produced. The protein was solubilised in Triton X-100 and purified by Ni-NTA affinity.

The Ni-NTA elution fractions were pooled and applied to a pre-equilibrated 10/300 superdex 200 (S200) size exclusion column. The peak fractions were analysed for full-length protein by measuring GFP fluorescence with a PHERAstar *FS* microplate reader in conjunction with an anti-His Western blot (Figure 2-3). The peak in the GFP fluorescence profile was quite broad, and each of the fractions from the S200 column gave a number of bands on the anti-His Western blot, suggesting the protein was degraded. This had not been seen in the small-scale expression and purification trial, indicating that either scaling up the expression culture volume led to degradation, or the protein degraded over time due to the addition of a size exclusion step to the purification protocol. It was therefore decided that other constructs that gave full-length bands on the Western blots in Figure 2-1 should be tested to see whether they were less susceptible to degradation.

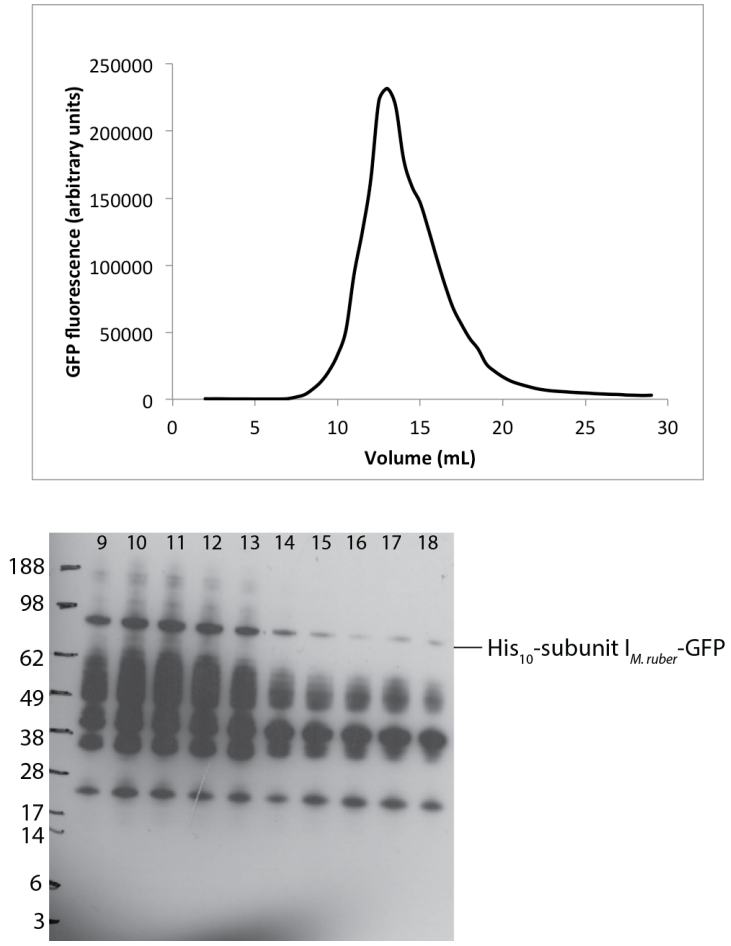


Figure 2-3: Fluorescence size exclusion profile for His_{10}-subunit $I_{M.ruber}$-GFP and anti-His Western blot of the elution fractions.

GFP fluorescence (arbitrary units) was measured for each of the elution fractions and plotted against elution volume (mL) for His_{10}-subunit $I_{M.ruber}$-GFP. The elution fractions were also analysed by a Western blot against the His_{10} tag to determine whether the protein had degraded.

2.4.2 *M. ruber* subunit I in vector F

The construct containing the *M. ruber* gene in vector F (MBP-subunit $I_{M. ruber}$-$TrxAHis_{10}$) gave the most intense band of all of the constructs, with no degradation visible (Figure 2-1F). It was found that the highest expression was achieved when the T7 Express cells transformed with this construct were grown in ZYM-5052 medium (Figure 2-4), possibly because the medium contains glucose and this has been shown to assist the production of proteins fused to MBP by repressing an endogenous *E. coli* amylase.[105] The protein was extracted from a 12 L culture by solubilising the *E. coli* membranes with Triton X-100, and purified by Cobalt affinity followed by amylose

affinity. The Triton series of detergents, however, is generally thought to be unsuitable for structural studies, as each of the detergents is chemically heterogeneous.[106] The most commonly used detergents are DDM and DM, yet as both of these detergents contain maltose derivatives, they are incompatible with amylose resin. Triton X-100 was therefore still used as the detergent during the purification, with the aim of testing other detergents once the purification protocol had been optimised. Elution fractions from the amylose resin were analysed by SDS-PAGE followed by SafeStain staining and a Western blot against the His_{10} tag. The protein running at ~98 kDa in Fig 2-4 was subsequently confirmed to be intact MBP--subunit $I_{M.\ ruber}$-$TrxAHis_{10}$ by mass spectrometry.

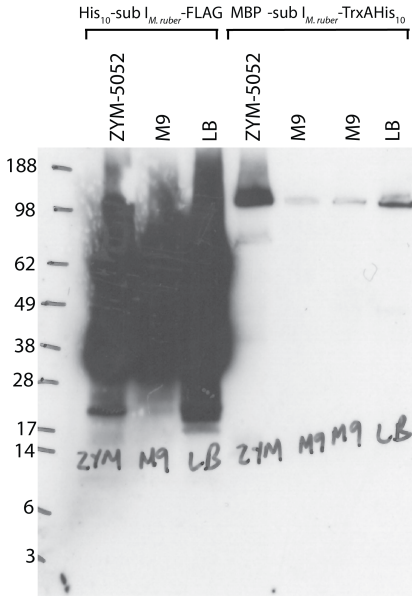


Figure 2-4: Media screen with MBP-subunit $I_{M.\ ruber}$-$TrxAHis_{10}$.

ZYM-5052, M9 and LB media (50 mL) were inoculated with T7 Express *E. coli* cells transformed with the His_{10}-subunit $I_{M.ruber}$-FLAG or MBP-subunit $I_{M.\ ruber}$-$TrxAHis_{10}$ constructs. Expression was induced at an optical density of 0.5-0.7 at 600 nm ($OD_{600\ nm}$), and the cells were incubated at 18 ºC overnight. The protein was extracted from the cells with Triton X-100 (1% (v/v)) and purified by cobalt affinity. The cobalt resin was separated by SDS-PAGE and a Western blot was performed against the His_{10} tag with ECL visualisation.

To test whether the protein was aggregated, the elutions from the amylose resin were applied to a pre-equilibrated 10/300 S200 column (Figure 2-5). The protein eluted at 12.5 mL, indicating it was most likely monomeric, according to the published molecular weight standard elution volumes for that column.[107] The protein aldolase, with a

molecular mass of 158 kDa, elutes at ~12 mL on this column. The molecular mass of MBP-subunit $I_{M.\ ruber}$-TrxAHis$_{10}$ is ~131 kDa, and thus the monomeric protein would be expected to elute at a later volume than 12 mL.

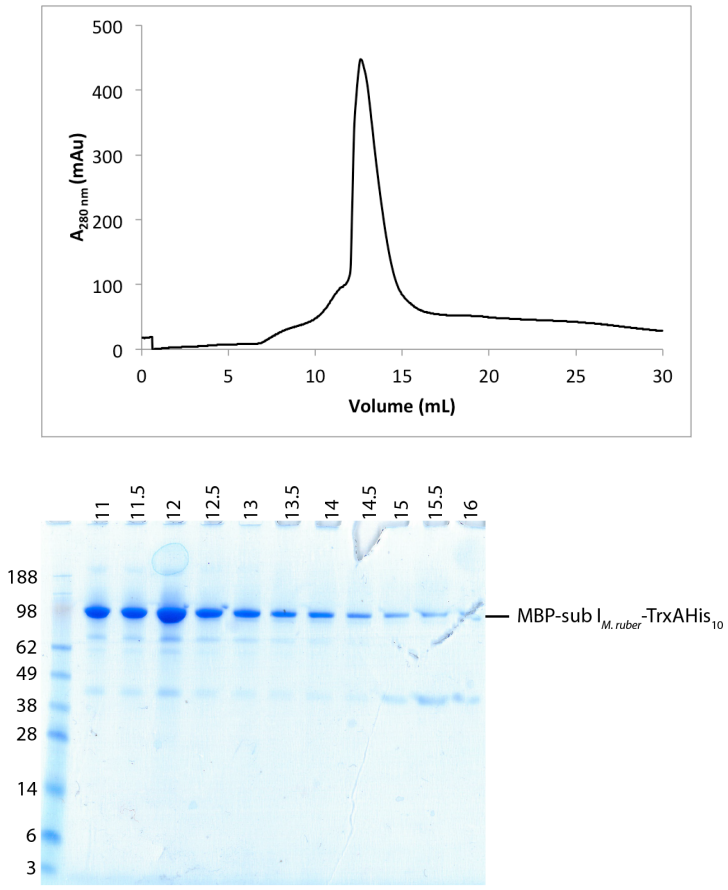


Figure 2-5: S200 size exclusion profile and SDS-PAGE of the elution fractions for MBP-subunit $I_{M.\ ruber}$-TrxAHis$_{10}$.

MBP-subunit $I_{M.\ ruber}$-TrxAHis$_{10}$ that has been purified by cobalt and amylose affinity was applied to a 10/300 S200 size exclusion column and the absorbance at 280 nm was monitored (left). The peak fractions in the size exclusion elution profile were analysed by SDS-PAGE using SafeStain staining to confirm the peak fractions contained MBP-subunit $I_{M.\ ruber}$-TrxAHis$_{10}$. The elution volume of the fraction (mL) is labelled at the top of each lane.

As the presence of solubility tags such as MBP can impede crystallisation, a cleavage trial with HRV 3C protease (3C) was conducted on the purified protein to determine whether the MBP tag could be cleaved. The cleavage trial was analysed by SDS-PAGE using SafeStain staining and it was found there was a faint band present at the size corresponding to the intact protein, and an intense band approximately 20 kDa smaller

than this, which could be the cleaved protein (data not shown). However, this result could not be replicated, and thus requires further investigation. It was decided it would be more efficient to pursue a construct that would not require cleavage of the solubility/affinity tag for crystallisation.

2.4.3 *M. ruber* subunit I in vector A

The construct containing the *M. ruber* gene in vector A (His_{10}-subunit $I_{M.\ ruber}$-FLAG) was chosen for a large scale trial as the FLAG tag is only eight amino acids in length (DYKDDDDK) and thus would be unlikely to interfere with crystallisation. Unlike the MBP-subunit $I_{M.\ ruber}$-TrxA/His_{10} construct, the His_{10}-subunit $I_{M.\ ruber}$-FLAG was found to have the highest expression in M9 minimal medium on a small scale (Figure 2-6). When the culture volume was scaled up to 4 L, however, the T7 Express cells grew too slowly in the M9 medium to reach a sufficient optical density for induction, so LB medium was chosen as the most suitable alternative.

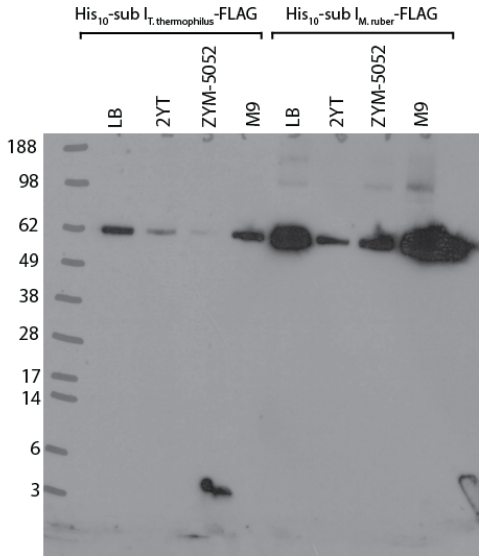


Figure 2-6: Media screen for His_{10}-subunit $I_{M.\ ruber}$-FLAG expression.

LB, 2YT, ZYM-5052 and M9 and media (5 mL) were inoculated with T7 Express *E. coli* cells transformed with the His_{10}-subunit $I_{T.\ thermophilus}$-FLAG or His_{10}-subunit $I_{M.ruber}$-FLAG constructs. Expression was induced at $OD_{600\ nm}$ = 0.5-0.7 and the cells were incubated at 18 ºC overnight. The protein was extracted from the cells with DM (2% (w/v)) and purified by cobalt affinity. The cobalt resin was separated by SDS-PAGE and a Western blot was performed against the FLAG tag with ECL visualisation.

2.4.3.1 Detergent screen with His_{10}-subunit $I_{M.\ ruber}$-FLAG

To improve the likelihood of crystallisation, detergents other than Triton X-100 were screened for their ability to extract His_{10}-subunit $I_{M.\ ruber}$-FLAG from the *E. coli* membrane and minimise degradation during the purification. Five detergents were tested initially (Figure 2-7), including Triton X-100 as a positive control, and each of them were able to solubilise His_{10}-subunit $I_{M.\ ruber}$-FLAG to varying degrees. While Anzergent 3-14 appeared to extract the most protein, there was some degradation evident, so Anzergent 3-12 was chosen to be tested on a larger scale.

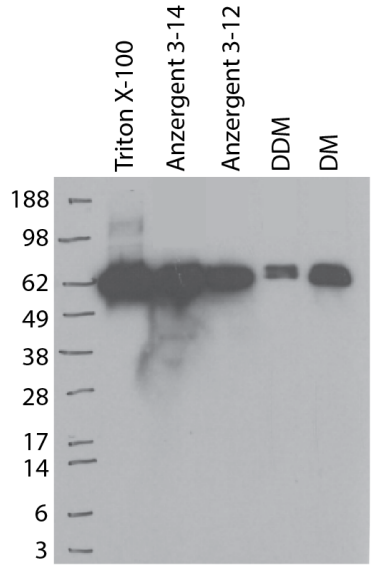


Figure 2-7: Small-scale detergent screen with His_{10}-subunit $I_{M.\ ruber}$-FLAG.
His_{10}-subunit $I_{M.\ ruber}$-FLAG was extracted and purified in one of five detergents: Triton X-100, Anzergent 3-14, Anzergent 3-12, DDM and DM. The protein was immobilised on cobalt resin and run on SDS-PAGE, and a Western blot was performed against the FLAG tag with ECL visualisation.

2.4.3.2 His_{10}-subunit $I_{M.\ ruber}$-FLAG purification trial

To test whether sufficient quantities of His_{10}-subunit $I_{M.\ ruber}$-FLAG could be produced for crystallisation trials, a 12 L T7 Express culture was grown in LB medium. The cultures were incubated for between 14 and 20 hours, although the level of protein expression and the extent of degradation did not change with different growth times. Different growth temperatures were not tested. Anzergent 3-12 was used to extract the protein from the *E. coli* membrane and solubilise the protein throughout the

purification. Following initial purification with Cobalt resin, the protein was applied to a pre-equilibrated 10/300 S200 column to determine whether it was aggregated. While a peak eluted at 11 mL, a volume that would be consistent with monomeric protein in a detergent micelle,[107] this peak had a significant shoulder, and the SDS-PAGE gel on which the elution fractions were run shows the fractions contain multiple bands (Figure 2-8). As with the His_{10}-subunit $I_{M.\ ruber}$-GFP construct, degradation was not evident on the western blot used to analyse the expression and purification trial of the FLAG-tagged construct, suggesting the degradation was occurring during the course of the purification. Nevertheless, the band corresponding to subunit I was the most intense band in the peak fractions, indicating His_{10}-subunit $I_{M.\ ruber}$-FLAG could be produced in sufficient quantities for structural and biophysical studies, provided the purification protocol could be optimised. This construct was therefore selected for further optimisation, characterisation and crystallisation trials.

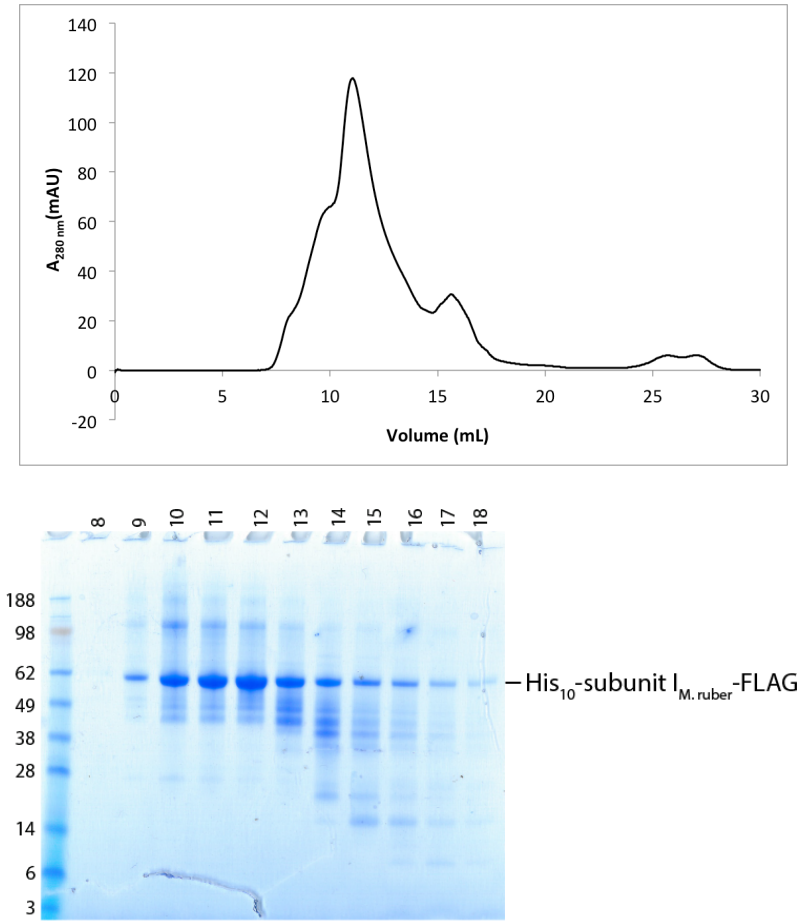


Figure 2-8: S200 size exclusion profile and SDS-PAGE of the elution fractions for His_{10}-subunit $I_{M.\ ruber}$-FLAG.

His_{10}-subunit $I_{M.\ ruber}$-FLAG was purified by cobalt affinity from a 12 L *E. coli* culture in buffer containing Anzergent 3-12 (0.5% (w/v)). Elution fractions containing the protein were applied to a pre-equilibrated S200 size exclusion column, and the absorbance at 280 nm was monitored (upper image). The peak elution fractions were analysed for His_{10}-subunit $I_{M.\ ruber}$-FLAG by SDS-PAGE using SafeStain (lower image).

2.5 Cloning and expressing *M. ruber* subunit I transmembrane domain constructs

As was noted in section 1.10, subunit I is highly flexible, and it has been found that this flexibility is largely localised to the linker between the soluble and transmembrane domains of the protein. This enables the soluble domain to adopt a number of different angles relative to the transmembrane domain,[101] which could reduce the likelihood of subunit I crystallising. To overcome this, constructs containing only the transmembrane domain (I_{TM}) were cloned in tandem with the full-length constructs. As the crystal

structure of the soluble domain of the *M. ruber* protein had already been determined, this was compared with the full-length subunit I sequence, revealing the transmembrane domain expression construct should start at residue T346 to exclude the soluble domain and the flexible linker. The *M. ruber* synthetic gene was used as the template to create the transmembrane domain construct. The desired region of the gene was amplified by using a forward PCR primer designed to clone from the codon for residue T346 onwards, and the reverse primer was the same as that used to clone the full-length gene. The transmembrane domain PCR product was successfully cloned into all of the SLIC vectors with the exception of vector G.

A small-scale expression trial was conducted with each of the vectors as described in section 5.7.1. Intact protein was detected by a Western blot against the His_{10} tag, which enabled all of the constructs to be visualised simultaneously (Figure 2-9). Full-length expression was only seen for vectors C (expected size 39.4 kDa) and D (expected size 50.4 kDa) constructs. There is a band for vector F but this is smaller than the expected size of 80.1 kDa, so it was not tested further.

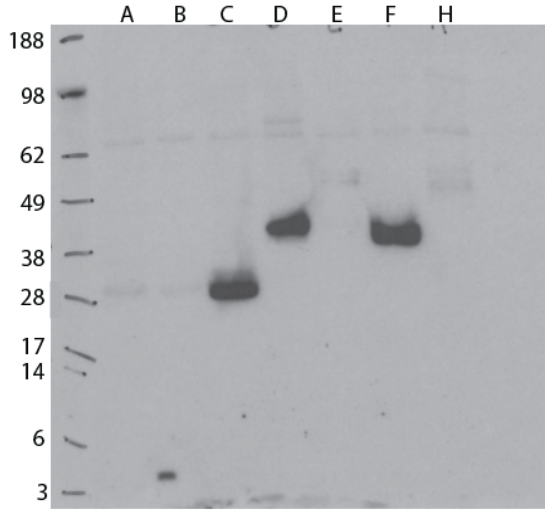


Figure 2-9: Small-scale expression and purification trial for the I_{TM} expression constructs.

Each of the subunit I transmembrane domain constructs was expressed in 3 mL *E. coli* cultures and purified by cobalt affinity. The resin was run on SDS-PAGE and a Western blot was conducted against the His_{10} tag with ECL visualisation.

The constructs in vectors C and D were subsequently expressed in 4 L LB medium, solubilised in DM and purified using Cobalt resin. Elution fractions from the Cobalt resin were visualised by SDS-PAGE using SafeStain (Figure 2-10). Bands were only seen for the vector D construct (mistic-I_{TM}-His_{10}). Mistic (membrane-integrating sequence for translation of IM protein constructs) is a 13 kDa integral membrane protein from *Bacillus subtilis* that does not require the translocon machinery to fold, and has been shown to assist insertion of recombinantly expressed integral membrane proteins into the host cell membrane[108].

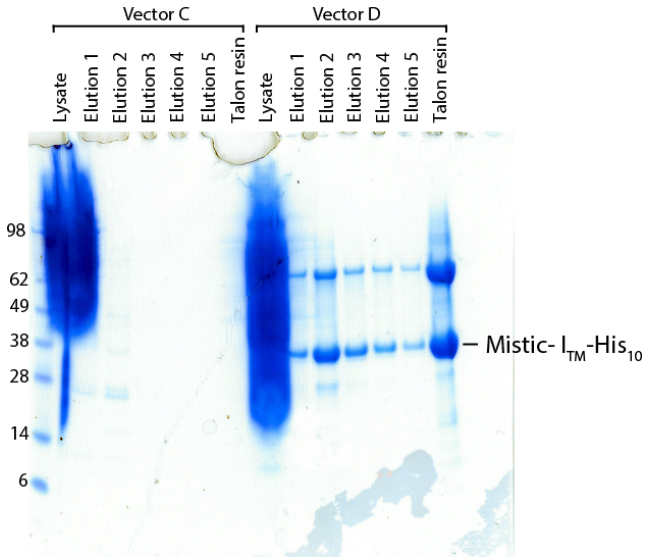


Figure 2-10: Cobalt resin elution fractions for the I_{TM} construct in vectors C and D.

T7-I_{TM}-His_{10} and mistic-I_{TM}-His_{10} were each purified from *E. coli* cultures (4 L) by cobalt affinity. The lysate, cobalt elution fractions, and cobalt resin were analysed for the presence of the intact subunit I transmembrane domain construct by SDS-PAGE using SafeStain.

The mistic-I_{TM}-His_{10} cobalt elutions were pooled and loaded onto a pre-equilibrated 10/300 S200 column. There were two major peaks present in the $A_{280\ nm}$ elution profile. The first eluted at ~8.5 mL, which could be aggregated protein, or protein in a higher oligomeric state, as according to the published molecular weight standard elution volumes for this column, ferritin (440 kDa) elutes at 10 mL.[107] The protein that eluted in this peak is thus even larger than 440 kDa, and many times larger than the monomeric form of mistic-I_{TM}-His_{10} (50.4 kDa).

The second peak eluted at 12.5 mL, which would be consistent with monomeric protein (Figure 2-11). As seen in the SDS-PAGE gel of the S200 elution fractions, however, two major bands are present in each of the peaks, suggesting the protein is present in aggregated or higher oligomeric states as well as the monomeric state.

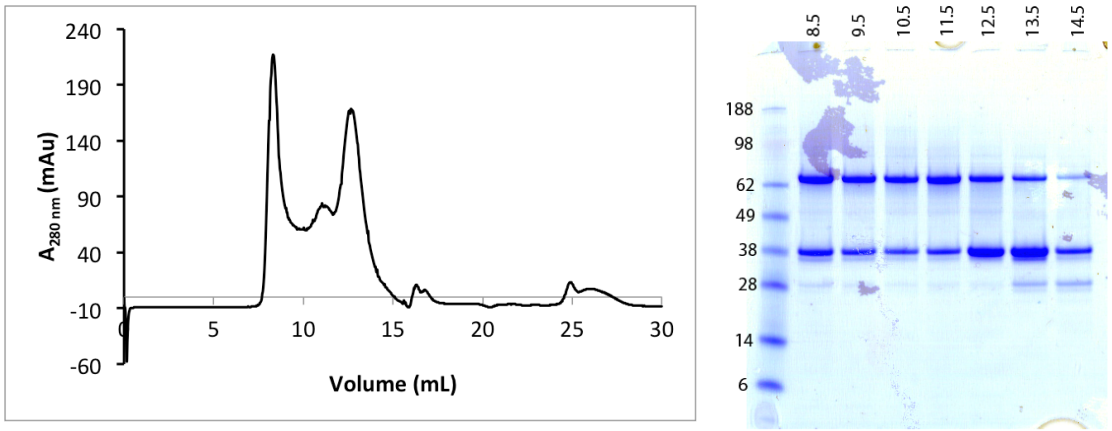


Figure 2-11: S200 size exclusion profile and SDS-PAGE gel of the elution fractions for mistic-I_{TM}-His_{10}.

The cobalt elution fractions containing mistic-I_{TM}-His_{10} were applied to a pre-equilibrated S200 size exclusion column, and absorbance at 280 nm was monitored (left). To determine whether the peak fractions contained mistic-I_{TM}-His_{10}, the fractions were analysed by SDS-PAGE with SafeStain (right).

2.6 Cloning and expressing subunits G and E of the *M. ruber* A-ATPase

2.6.1 Cloning *M. ruber* subunits G and E into the pET-Duet-1 expression vector

In the SDS-PAGE gel shown in Figure 2-8, it can be seen that along with the major His_{10}-subunit $I_{M.\ ruber}$-FLAG band at ~62 kDa, a number of other proteins have co-eluted. As subunit I is bound to both the soluble peripheral stalks and the transmembrane L-ring rotor within the native A-ATPase complex (Figure 1-10), it was thought that the binding sites for these proteins might also be sites prone to forming non-specific interactions with other *E. coli* proteins during purification. Co-purifying His_{10}-subunit $I_{M.\ ruber}$-FLAG with one or both of its partner proteins from the A-ATPase complex could therefore be a strategy to improve the purity and stability of subunit I.

There were thus a number of options for creating A-ATPase sub-complexes with His_{10}-subunit $I_{M.\ ruber}$-FLAG, including the subunit I:peripheral stalks sub-complex, the subunit I:L-ring sub-complex, or a sub-complex containing all three proteins. However, as the L subunit is also a membrane protein, and forms a ring that could be difficult to produce with the correct stoichiometry when expressed recombinantly, it was decided the subunit I:peripheral stalks sub-complex would be the easiest to pursue. Furthermore, cloning, purification and crystallisation protocols had already been established within the laboratory for the *T. thermophilus* peripheral stalks, so these were used as a guide to clone and purify the *M. ruber* peripheral stalks.

Subunit E from *M. ruber* was found through a BLAST search using *T. thermophilus* subunit E as the query sequence, and a codon-optimised synthetic gene encoding *M. ruber* subunit E was ordered. A BLAST search using the protein sequence for *T. thermophilus* subunit G, however, did not return a sequence for *M. ruber* subunit G. The first entry from *M. ruber* was for a hypothetical protein K649_04415 [Meiothermus ruber DSM 1279]. Nevertheless, the gene encoding this protein has been annotated and is located within the *M. ruber* A-ATPase operon. An alignment of the sequences of *T. thermophilus* subunit G and *M. ruber* hypothetical protein K649_04415 using the M-COFFEE program[109] showed that the proteins have a 104 amino acid overlap with 51.9% identity, suggesting they are related proteins. A codon-optimised synthetic gene encoding *M. ruber* hypothetical protein K649_04415 (hereafter subunit G, gene ID K649_04415) was therefore also ordered.

```
M.ru     MGFIFEQVFTCSLFRCAEQSDGDQKGGNVAGSGLIKSLAEREQALAKQLEEAKRAAAAKV  60
T.th     ----------------------------MGGLGLIKSLAEKEKQLLERLEAAKKEAEERV  32
                                     :.* ********:*: * ::** **: *  :*

M.ru     QEAEAKAAQILAEAEQAARDLEAQYRARTTEAVAKVESEAKARAEAEAKAIGEAAAAKVA  120
T.th     KRAEAEAKALLEEAEAKAKALEAQYRERERAETEALLARYRERAEAEAKAVREKAMARLD  92
         :.***:*  :* ***  *:****** *     .  : :. : ********: * * *::

M.ru     EAVQAVLKEVLP  132
T.th     EAVALVLKEVLP  104
         ***  *******
```

Figure 2-12: Sequence alignment for *T. thermophilus* subunit G and *M. ruber* hypothetical protein K649_04415.

The protein sequences of subunit G from *T. thermophilus* and *M. ruber* were aligned using the M-COFFEE Multiple Sequence Alignment tool. Asterisks (*) indicate an amino acid residue that is fully conserved between the two sequences, colons (:) indicate the residues are different but have very similar properties, and periods (.) indicate residues that are different with only slightly similar properties.

The synthetic genes for *M. ruber* subunits G and E were cloned into multiple cloning site (MCS) 1 and MCS2, respectively, of the bicistronic pET-Duet-1 vector, as confirmed by Sanger sequencing. Expression trials with this construct, however, show that only one protein was expressed. An anti-His Western blot confirmed that this was subunit G, as a protein cloned into MCS1 of this vector is encoded with an N-terminal His_6 tag.

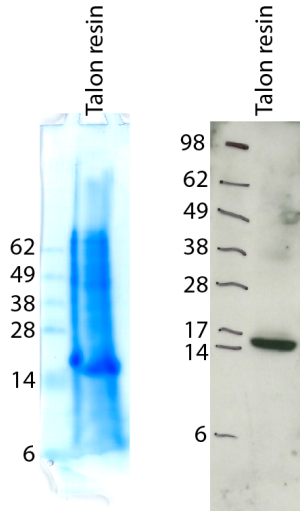


Figure 2-13: SDS-PAGE gel and anti-His Western blot to confirm expression of the *M. ruber* EG complex.

The *M. ruber* EG construct was expressed in T7 Express *E. coli* cells (5 mL) and purified by cobalt affinity. The resin was analysed by SDS-PAGE using SafeStain staining and a Western blot against the His_6 tag at the N-terminus of subunit G.

The sequence alignment of *M. ruber* subunit G with *T. thermophilus* G shows the *M. ruber* protein has an extension of 28 amino acid residues at the N-terminus (Figure 2-12). This had been seen previously with *T. thermophilus* subunit G, for which the most stable expression construct lacked the first 17 codons of the annotated gene.[27] A new construct was therefore cloned using a forward PCR primer that amplified the gene from codon 29 onwards (*M. ruber* G_{29}). The PCR product was then cloned into MCS1 of the pET-Duet-1 expression vector that contained *M. ruber* subunit E in MCS2, as confirmed by Sanger sequencing.

Expression trials with this construct, however, gave a similar result to that seen for the construct with full-length subunit G, with only subunit G_{29} being expressed. To confirm subunit E was not being expressed, a 12 L culture was grown and the protein was purified by Ni-NTA affinity followed by size exclusion chromatography with an S200 column. While there are two bands in the elution fractions run on SDS-PAGE (Figure 2-14), the bands have much lower intensity than that of the *T. thermophilus* EG complex (Figure 3-11). It also appears the two proteins are dissociating from one another during size exclusion, which could hinder crystallisation.

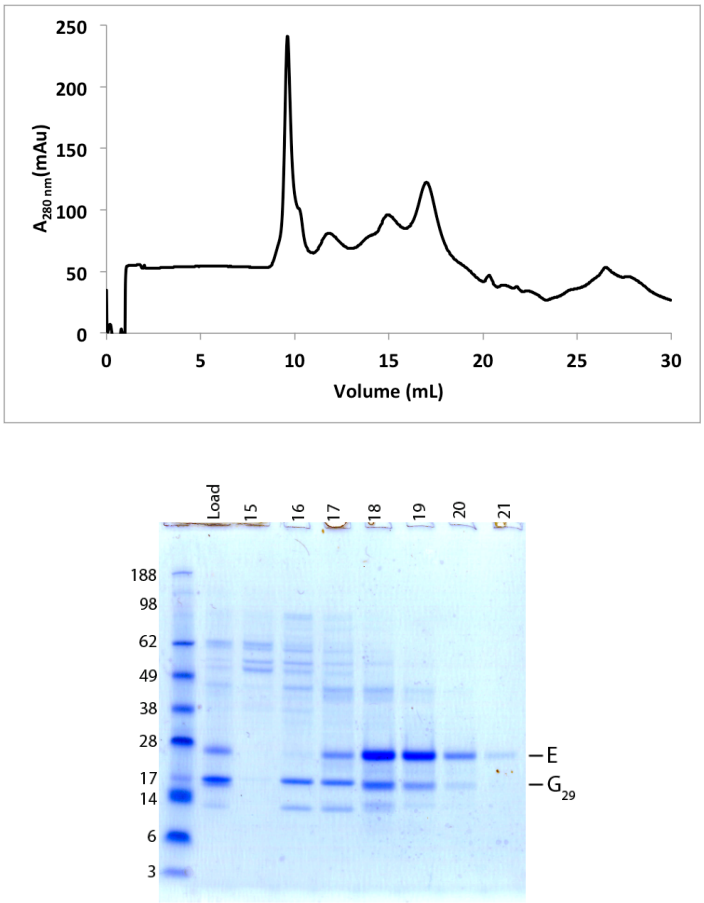


Figure 2-14: Large-scale *M. ruber* EG_{29} expression and purification trial.

The cobalt elution fractions containing EG_{29} were pooled and applied to a 10/300 S200 size exclusion column, and the absorbance at 280 nm was monitored (upper image). The peak fractions in the elution profile were analysed by SDS-PAGE with SafeStain (lower image).

The *M. ruber* EG_{29} complex could not be expressed and purified in quantities sufficient for biophysical and structural analyses, indicating there was a problem with this construct. In the BLAST search using *T. thermophilus* subunit G as the query sequence, there was an entry below the *M. ruber* hypothetical protein K649_04415 termed *M. ruber* subunit C. When the sequences of *T. thermophilus* subunit G, *M. ruber* subunit G_{29} and *M. ruber* subunit C are aligned, it can be seen that they are all very similar (Figure 2-15). However, *M. ruber* subunit C does not contain the 29-residue extension present at the start of the *M. ruber* hypothetical protein K649_04415. There are three additional residue differences; residues 25 and 73 are arginines in *M. ruber* subunit C rather than lysines, and residue 75 in *M. ruber* subunit C is a lysine rather than an arginine (Figure 2-15). In *T. thermophilus* subunit G, the equivalent residues are on the

surface of the protein, so mutating them in the *M. ruber* protein would probably not affect expression, but it was still important to use the wild type sequence.

```
T.thG       MGGLGLIKSLAEKEKQLLERLEAAKKEAEERVKRAEAEAKALLEEAEAKAKALEAQYRER
M.ruG       VAGSGLIKSLAEREQALAKQLEEAKRAAAAKVQEAEAKAAQILAEAEQAARDLEAQYRAR
M.ruC       MAGSGLIKSLAEREQALAKQLEEARRAAAAKVQEAEAKAAQILAEAEQAARDLEAQYRAR
              * ********:*: * ::** *:: *  :*:.***:*  :* ***  *: ****** *

T.thG        ERAETEALLARYRERAEAEAKAVREKAMARLDEAVALVLKEVLP
M.ruG        TTEAVAKVESEAKARAEAEAKAIGEAAAAKVAEAVQAVLKEVLP
M.ruC        TTEAVAKVESEARAKAEAEAKAIGEAAAAKVAEAVQAVLKEVLP
                 .  : :. : :*******: * * *:: ***  *******
```

Figure 2-15: Aligning sequences of *M. ruber* subunits C (M.ruC) and G_{29} (M.ruG) with *T. thermophilus* subunit G (T.thG).

The protein sequences of *T. thermophilus* subunit G, M. *ruber* subunit G_{29} and *M. ruber* subunit C were aligned using the Clustal Omega Multiple Sequence Alignment tool. Symbols underneath indicate the degree of conservation as described in Figure 2-12.

To test whether *M. ruber* subunit C was in fact the true homologue of *T. thermophilus* subunit G, a gBlock containing the genes encoding *M. ruber* subunits C and E was synthesised (section 5.6.4) and was successfully cloned by the Gibson assembly method into the pLICHK expression vector, confirmed by Sanger sequencing. Expression trials with this construct in T7 Express *E. coli* cells grown in LB medium revealed that both protein subunits could be overexpressed, as indicated by two intense bands at the expected sizes (~14 kDa for His_6-G and ~20 kDa for E) in each of the cobalt elution fractions analysed by SDS-PAGE (Figure 2-16). This complex was therefore subsequently used for the expression and purification of the *M. ruber* peripheral stalks, and subunit C will subsequently be referred to as G, as it had clearly been misannotated. Given how difficult it was to express and purify G_{29}, and that it was not possible to form a stable complex between G_{29} and E, it is unlikely that it would be expressed in *M. ruber*.

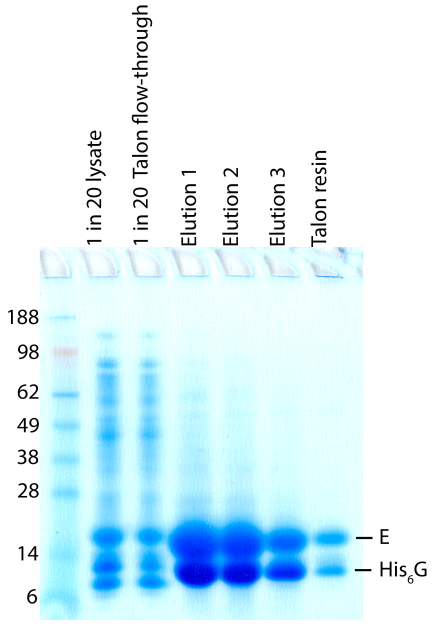


Figure 2-16: *M. ruber* EG small-scale expression and purification trial.

The *M. ruber* EG peripheral stalks construct was expressed in T7 Express *E. coli* cells (100 mL) and purified by cobalt affinity. The lysate, cobalt flow-through, elutions and cobalt resin were analysed for the presence of *M. ruber* EG by SDS-PAGE using SafeStain.

Chapter Three: Protein purification and optimisation

3.1 Optimising His_{10}-subunit $I_{M.\ ruber}$-FLAG in Anzergent 3-12

3.1.1 Purifying His_{10}-subunit $I_{M.\ ruber}$-FLAG in Anzergent 3-12 from isolated *E. coli* membranes

As can be seen in Figure 2-8, Anzergent 3-12 is able to extract His_{10}-subunit $I_{M.\ ruber}$-FLAG (from this point onwards this full-length construct shall be referred to as subunit I) from *E. coli* membranes very efficiently, and is able to prevent aggregation of the protein to some extent. However, the purity and stability of subunit I required further optimisation. To achieve this, the *E. coli* membranes were first isolated as described in section 5.7.5, to reduce contamination from soluble proteins. Sucrose (100 mM) and $MgCl_2$ (5 mM) were added to the purification buffers, as these had been shown to be important for stabilising the intact *T. thermophilus* A-ATPase complex during purification. The NaCl concentration was also increased to 500 mM to minimise the chance of contaminating proteins being co-purified with subunit I. Additionally, a 100 000 MWCO concentrator was used rather than a 30 000 MWCO concentrator to further reduce the chance of co-purification of smaller contaminating proteins, and to avoid concentrating the detergent, as the presence of detergent at concentrations many times above the critical micelle concentration (CMC) can inhibit protein crystallisation and encourage the growth of detergent crystals. Finally, a 16/600 S200 column was used, rather than a 10/300 column, as it was thought the additional volume would allow for greater separation of subunit I from the contaminating proteins.

All of these alterations did lead to improvements in the purity and stability of subunit I (Figure 3-1), as the intensity of the monomer peak (~61 mL, 112 mAu) was now greater than that of the aggregate peak (~54 mL, 28 mAu), and there was less overlap between these two peaks than that seen in Figure 2-8. According to the published molecular weight standard elution volumes for this column, ferritin (440 kDa) elutes at 55 mL, while aldolase (158 kDa) elutes at 65 mL, and these elution volumes were used to assign the peaks in Figure 3-1 as aggregate or monomer, respectively.

There were also fewer contaminants in the peak fractions, with subunit I being the dominant species. Fractions 60-65 were pooled and concentrated to 500 µL at a concentration of 4.5 mg mL^{-1}, giving a total yield of 2.25 mg from the 12 L culture as measured by $A_{280\ nm}$. This yield was quite low, however it is not unexpected for a membrane protein. This sample was subsequently used for crystallisation trials, although as discussed in section 4.1, there were no promising results from these trials. This may have been due in part to the zwitterionic nature of Anzergent 3-12, as zwitterionic detergents are often more disruptive than non-ionic detergents.[102] Other detergents that have been shown to be more favourable for membrane protein crystallisation were therefore subsequently trialled.

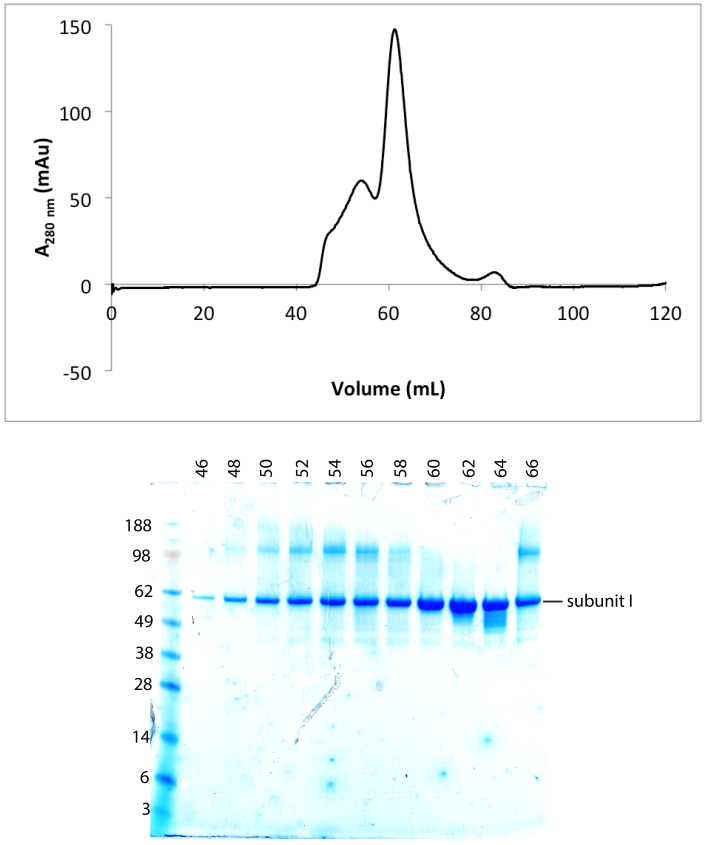


Figure 3-1: Size exclusion profile and SDS-PAGE of elution fractions for His_{10}-subunit $I_{M.\ ruber}$-FLAG purified from isolated membranes.

The absorbance at 280 nm of the S200 size exclusion elution fractions was monitored and plotted against the elution volume (mL) to determine the elution volume of subunit I (upper image). The presence and purity of subunit I in the peak fractions was confirmed by SDS-PAGE using SafeStain (lower image). The mass of the protein standards is labelled (kDa) on the left hand side of the gel, the elution volume of each fraction is labelled (mL) at the top of each lane and the band corresponding to subunit I is labelled.

3.1.2 Further detergent optimisation

A small-scale expression and purification screen was conducted with eight detergents, with Anzergent 3-12 included as a control. Eight × 3 mL expression cultures were grown, and subunit I was solubilised and purified from each one using one of the eight detergents. The lysate was bound to magnetic cobalt resin, and the resin was analysed by SDS-PAGE followed by a Western blot against the FLAG tag to detect the presence of intact subunit I (Figure 3-2).

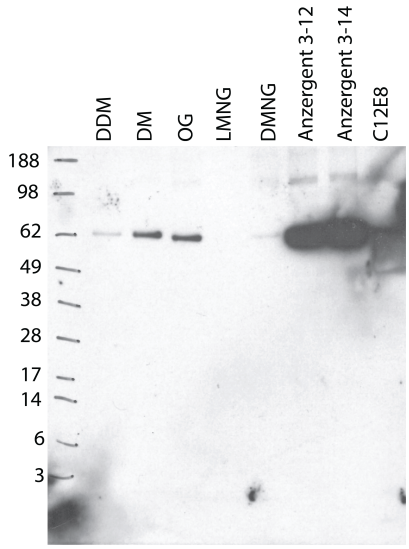


Figure 3-2: Screening further detergents for extracting and purifying subunit I.
A small-scale solubilisation and purification trial was conducted with subunit I in eight different detergents. The samples were analysed for intact subunit I by a Western blot against the FLAG tag. The mass of the protein standards is labelled, in kDa, on the left hand side of the blot, and the detergent used for each samples is labelled at the top of each lane.

Intact subunit I was seen with the detergents DM, *n*-octyl-β-D-glucopyranoside (OG), Anzergent 3-12 and 3-14, C12E8, and DDM to a lesser extent. As DM, OG, Anzergent 3-14 and C12E8 gave the most intense bands on the Western blot, these were subsequently tested on a larger scale to determine whether any of these detergents were able to maintain subunit I in a stable, monodispersed state during the purification. A 12 L expression culture was grown and divided into four batches. Each batch was solubilised in one of the four detergents and the supernatant was applied to cobalt resin. The elutions from the cobalt resin were concentrated and analysed for non-aggregated

protein by size exclusion chromatography using a 16/60 S200 column pre-equilibrated with buffer containing the corresponding detergent at 1 × CMC (Figure 3-3). A non-aggregate peak can be seen in the size exclusion profiles for Anzergent 3-14 and DM, while OG and C12E8 did not yield any monodispersed protein. The difference in the elution volumes of the Anzergent 3-14- and DM-solubilised subunit I is due to the different aggregation numbers of each of these detergents. Aggregation number is the average number of detergent molecules present in the micelle when the detergent concentration is above the CMC. The aggregation number for Anzergent 3-14 is 83-130,[110] while for DM it is 69.[111] As Anzergent 3-14 is a very similar detergent to Anzergent 3-12, and a significant amount of subunit I became aggregated in this detergent, Anzergent 3-14 was not pursued further. DM, the detergent that produced the maximum amount of stable subunit I, was therefore chosen as the detergent in which subunit I was to be solubilised and purified.

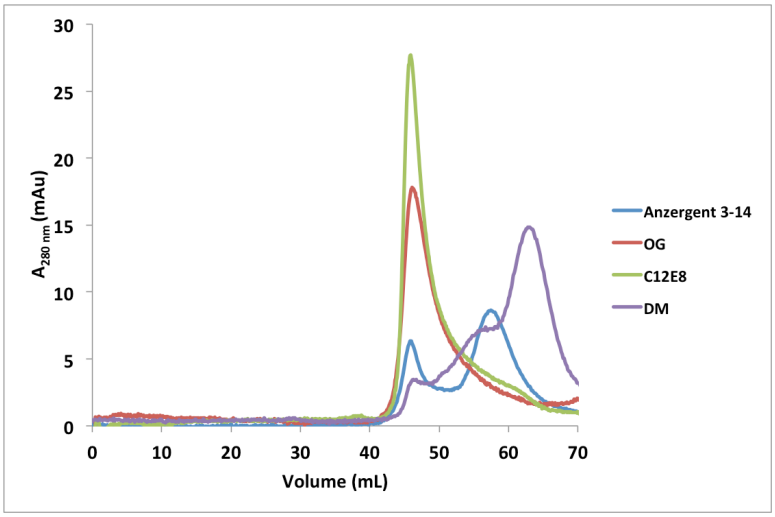


Figure 3-3: S200 size exclusion profiles for subunit I solubilised in four different detergents.

The cell pellet from a 12 L subunit I expression culture was divided into four, and each pellet was purified with one of four detergents: Anzergent 3-14 (blue), OG (red), C12E8 (green) or DM (violet). The clarified lysate was analysed for monodispersed subunit I by size exclusion using a 16/600 S200 column. The absorbance at 280 nm of the elution fractions was plotted against the elution volume (mL) for each detergent.

3.2 Optimising purification of subunit I in DM

Isolated membranes from a 12 L culture were solubilised in lysis buffer containing 2% (w/v) D. Subunit I was purified by cobalt affinity followed by size exclusion chromatography, with DM at a concentration of 0.5% (w/v) in the cobalt purification buffers and 0.09% (w/v) (1 × CMC) in the size exclusion buffer (Figure 3-4a). While the majority of subunit I eluted in the monomeric state (14-16 mL), a significant proportion of the protein was aggregated. According to the published molecular weight standards for this column, ferritin (440 kDa) elutes at 10 mL, and thus this peak (7-8 mL) would contain protein with a molecular weight many times greater than that of the subunit I monomer.[107] Additionally, a number of other contaminating bands of lower intensity can be seen in the peak fractions in the SDS-PAGE, which were not seen when the protein was purified in Anzergent 3-12. To test whether subunit I solubilised in DM continued to aggregate over time or whether it remained stable and the purity could be improved, the fractions that eluted at 14 and 15 mL were pooled, concentrated and loaded onto the same size exclusion column (Figure 3-4b). The aggregate peak is of lower intensity in the second size exclusion profile, suggesting that the majority of the monomeric protein remained stable. There were also fewer contaminating bands in the peak fractions, indicating the purity of the sample was improved, albeit with some sample loss. After the second size exclusion step the final protein yield was 900 μg, as measured by $A_{280\ nm}$.

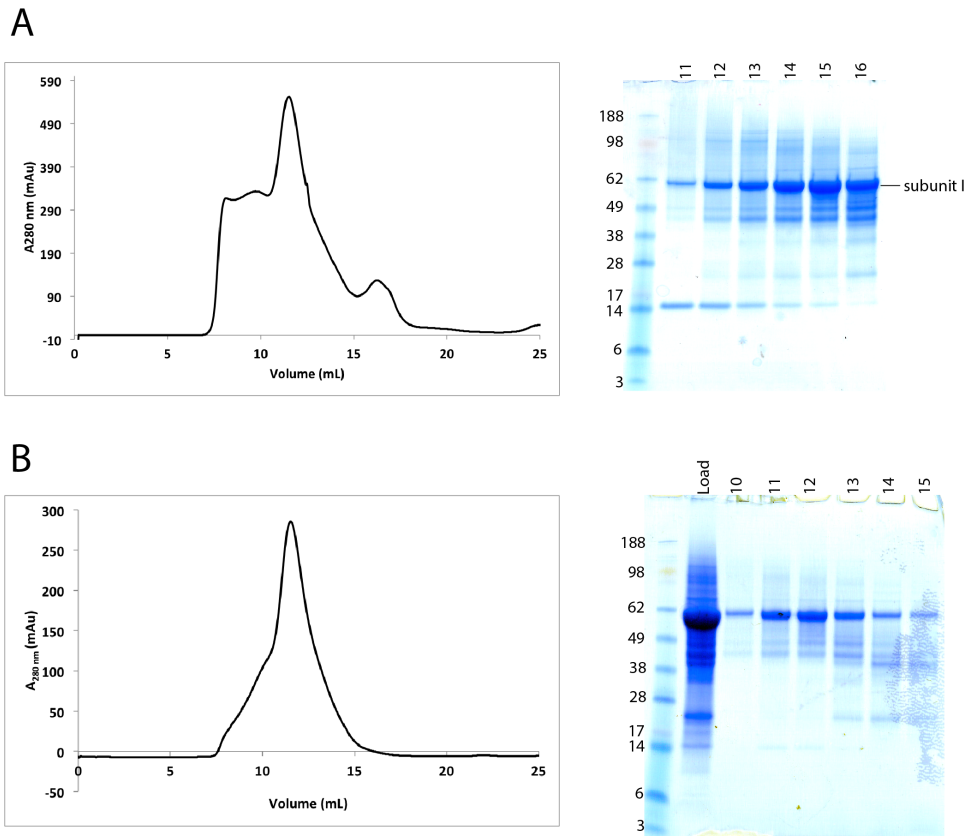


Figure 3-4: S200 size exclusion profiles and SDS-PAGE for subunit I solubilised and purified in DM.

A: Cobalt elutions containing subunit I were pooled and run on a 10/300 S200 size exclusion column. The absorbance at 280 nm of the elution fractions was measured and plotted against the elution volume (mL, left image). The peak fractions were then analysed by SDS-PAGE using SafeStain (right image). The elution volumes are labelled at the top of each lane, and the band corresponding to subunit I is labelled. B: Elutions 14-16 were pooled and concentrated and run on the same S200 column. The S200 $A_{280\ nm}$ elution profile (left) and SDS-PAGE image (right) were generated as described for A.

As the sample purified in Anzergent 3-12 seemed more pure and less prone to aggregation, it was thought that the limitations associated with using DM throughout the purification could be overcome by using Anzergent 3-12 to extract and solubilise subunit I up until size exclusion, and then exchanging subunit I into DM during size exclusion. To test this, subunit I was first extracted from *E. coli* membranes from a 12 L culture using Anzergent 3-12, and the cobalt purification buffers also contained Anzergent 3-12. The size exclusion step, however, was performed using a buffer that contained DM at twice the CMC, to determine whether subunit I could tolerate being exchanged into DM. DM has a lower CMC (1.8 mM) than Anzergent 3-12 (2.8 mM),

and thus DM forms micelles in solution more readily than Anzergent 3-12, theoretically making it able to replace Anzergent 3-12. Additionally, Anzergent 3-12 would be diluted to a concentration many times lower than the CMC in the size exclusion column, and would thus no longer be able to form micelles that could solubilise subunit I, further increasing the likelihood the Anzergent 3-12 micelle would be replaced by a DM micelle. It is evident in Figure 3-5 that a higher yield of subunit I (4 mg) was obtained than in the previous purification in which DM was used as the solubilisation detergent (900 µg). It can also be seen in the size exclusion profile that there is minimal aggregation and the majority of the protein is in the monomeric state. Some degradation can be seen in the SDS-PAGE image, however, ~90% of the protein is intact. This protein was subsequently used for crystallisation trials.

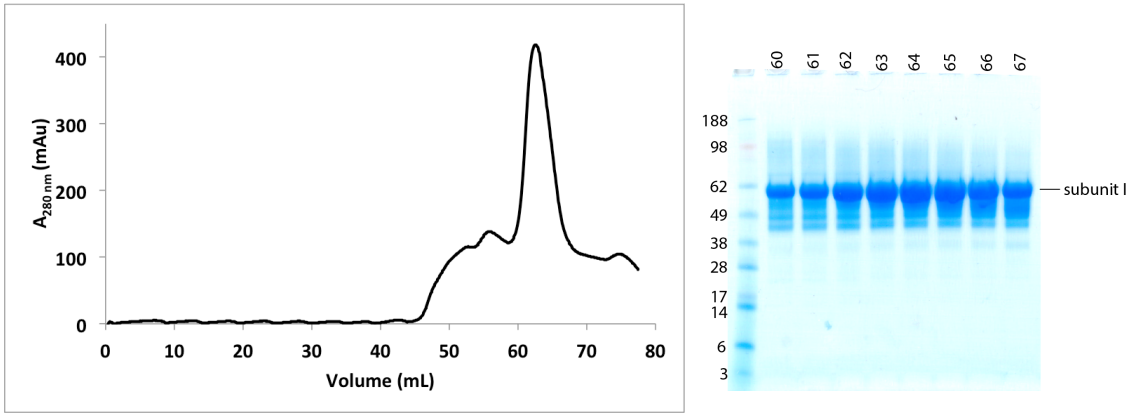


Figure 3-5: Size exclusion profile and SDS-PAGE of the elution fractions of subunit I exchanged into DM.

Subunit I was extracted from *E. coli* cell membranes and cobalt-purified in Anzergent 3-12. The cobalt elutions containing subunit I were pooled and run on a 16/60 S200 column in buffer containing DM (1.7% (w/v); left image). The peak fractions in the $A_{280\ nm}$ elution profile were analysed by SDS-PAGE using SafeStain (right image). The elution volume (mL) of the fractions is labelled at the top of each lane, along with the band corresponding to subunit I.

3.3 Supplementing purification buffers with ATP

As discussed in Section 1-8, there is evidence that *T. thermophilus* subunit I is able to bind to ATP. To determine whether ATP has a stabilising effect on *M. ruber* subunit I, 2 mM ATP was added to each of the purification buffers. The results of these

purifications were variable; for some purifications the addition of ATP was associated with no aggregated subunit I and a well-resolved monomer peak, and for others it did not appear to reduce the amount of aggregated protein compared to purifications performed without added ATP (Figure 3-6). In the upper panel of Figure 3-6, it can be seen that the intensity of the peak corresponding to subunit I is ~4500 mAu, while in the lower panel of Figure 3-6, the intensity of the peak is only ~180 mAu. This is most likely because different Äkta systems (GE Healthcare Life Sciences) were used for the two purifications, with different light sources and detectors. The yields for the samples in the upper and lower panels were 3 mg and 1.3 mg, respectively, which alone would not account for the intensity of the peak in the upper panel being ~25-fold higher than the peak in the lower panel. Yet even though there were differences in both profiles, crystallisation trials were conducted with subunit I purified both with and without ATP (5.8).

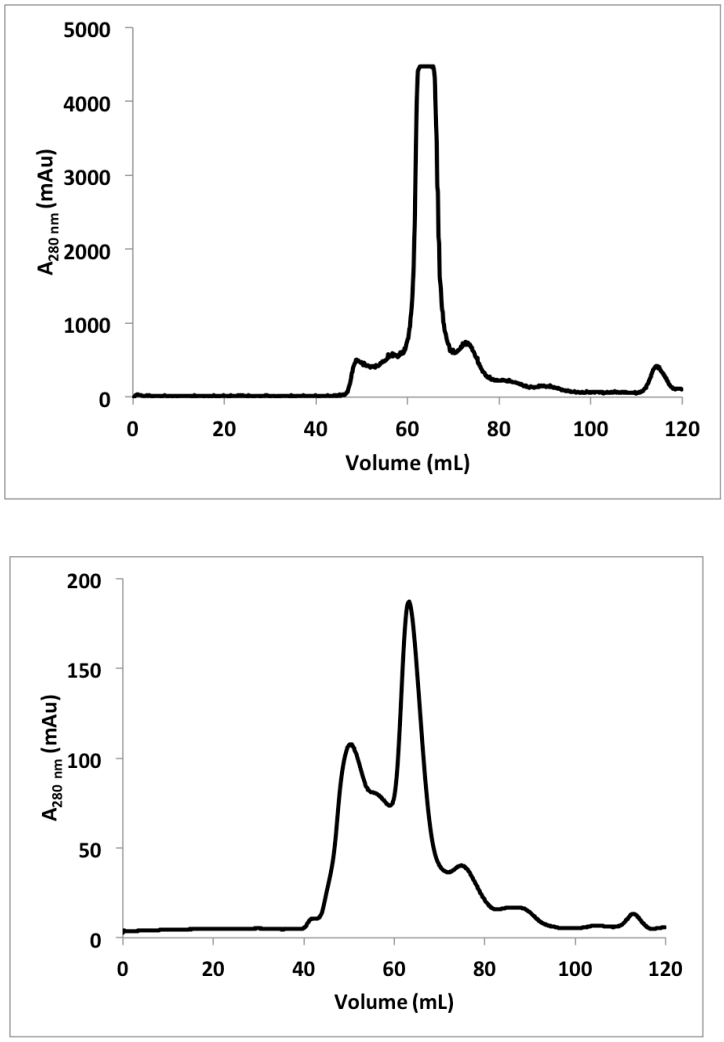


Figure 3-6: Size exclusion profiles of subunit I purified in the presence of 2 mM ATP.

Cobalt elutions containing subunit I purified in the presence of 2 mM ATP were pooled and applied to a 16/60 S200 size exclusion column that had been pre-equilibrated with buffer also containing 2 mM ATP. The upper image shows a size exclusion profile with minimal aggregated protein, while the lower shows a size exclusion profile with an intense aggregate peak.

3.4 Exchanging subunit I into Amphipol A8-35

Identifying a detergent in which a membrane protein is be both stable and amenable to crystallisation is a difficult and time-consuming process. A class of chemicals, termed amphipols, has therefore been developed to replace a detergent once a membrane protein has been extracted from the cell in which it was expressed.[112] Amphipols are amphipathic surfactants, containing a hydrophilic backbone onto which hydrophobic

chains have been grafted, that are able to displace detergent molecules and coat the hydrophobic surfaces of membrane proteins irreversibly without denaturing the protein (Figure 3-7). While amphipols have thus far not been used widely in crystallisation studies, it was thought subunit I could be exchanged into Amphipol A8-35 (the best characterised and most commonly used type of amphipol) for functional studies and crystallisation trials.

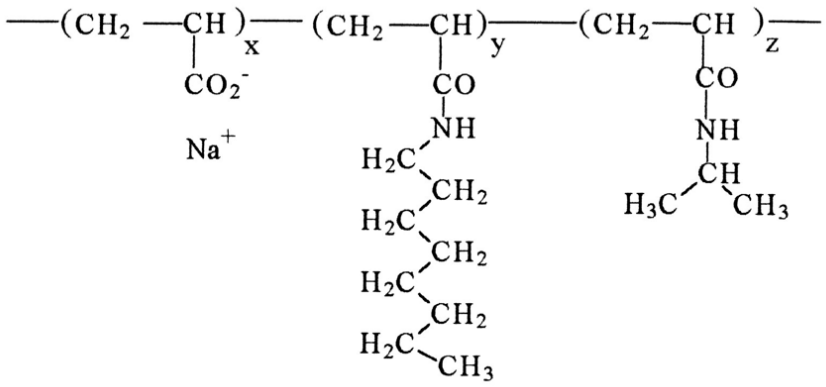


Figure 3-7: Amphipol A8-35 molecular structure.[112]

The molecular structure of Amphipol A8-35 is shown as a condensed structural formula. It contains a polyacrylate backbone with ~35% of the carboxylates free (x), ~25% containing octyl side chains (y) and ~40% containing isopropyl side chains (z).[113]

Subunit I was expressed and purified using the optimised protocol as described (section 3.2), and exchanged into Amphipol A8-35 (section 5.7.7). To test whether the exchange and detergent removal were successful, the subunit I sample was loaded onto a 10/300 S200 column pre-equilibrated with buffer that did not contain detergent, to minimise sample loss (Figure 3-8). The size exclusion elution volume was similar to that observed for subunit I solubilised in detergent (~12.5 mL), and there was minimal aggregation and degradation, although approximately only half the amount of protein was recovered, with a total yield of 1 mg as measured by $A_{280\ nm}$. Attempts were made to improve the yield of subunit I in amphipols, including performing the exchange immediately following cobalt purification and reducing the incubation time with Amphipol A8-35 and the Bio-Beads to two hours each, however, these did not make a significant difference. To overcome this, two amphipol-exchanged subunit I samples were combined to be used for subsequent experiments and crystallisation trials.

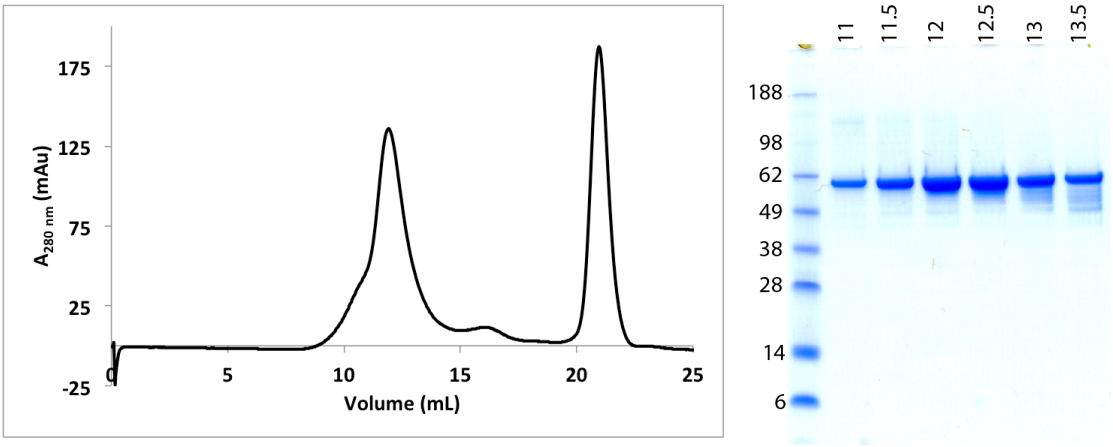


Figure 3-8: S200 profile and SDS-PAGE gel of subunit I exchanged into Amphipol A8-35.

Subunit I that had been exchanged from Anzergent 3-12 into Amphipol A8-35 was purified by running the sample on a 10/300 size exclusion column equilibrated in buffer containing no detergent (left image). The fractions from the first peak in the $A_{280\ nm}$ elution profile were visualised by SDS-PAGE using SafeStain staining (right image). The elution volumes (mL) of the fractions are labelled at the top of each lane on the gel image.

3.5 Optimising mistic-I_{TM}-His_{10} purification

As was noted in section 2.5, the purification of the mistic-I_{TM}-His_{10} construct looked promising, however it required further optimisation to remove the contaminating higher molecular weight protein (Figure 2-11). To address this, a 12 L expression culture was grown in LB medium, and mistic-I_{TM}-His_{10} was extracted with 2% (w/v) DM. The protein was purified as described in section 5.7.6, but instead a 16/60 S200 size exclusion column was used, rather than the 10/300 column, and as mistic contains one cysteine residue, 5 mM β-mercaptoethanol was added to all buffers to prevent disulphide-mediated aggregation. There was a large peak that eluted in the void volume (~46 mL, 71 mAu), with almost equal intensity to that of the monomeric peak (~67 mL, 72 mAu). Nevertheless, the monomeric peak appeared to be reasonably symmetric, and the elution fractions were largely pure when visualised by SDS-PAGE, with a yield of 1 mg. Despite the low yield, this sample was concentrated further and was used to set up crystallisation trials. As detailed in section 4.2, however, there were no promising results from these trials. It was thought that this may be due to the presence of the mistic tag, so based on the results from the expression trial (Figure 2-9), expression and

purification of the T7-I_{TM}-His_{10} construct was trialled once again. Although no protein was seen in the cobalt elutions for this trial, it was possible this was due to a problem that could be overcome with further optimisation.

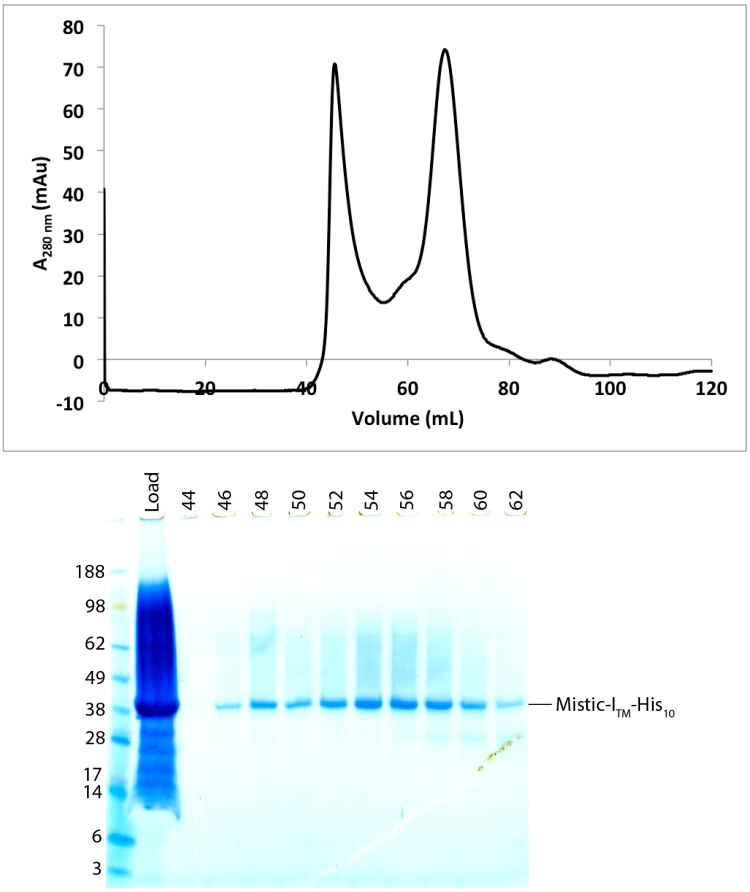


Figure 3-9: S200 size exclusion profile and SDS-PAGE of elution fractions for mistic-I_{TM}-His_{10}.

Cobalt elutions containing mistic-I_{TM}-His_{10} were pooled and run on a 16/60 S200 size exclusion column. Elution fractions from the two peaks in the $A_{280\ nm}$ elution profile (upper image) were analysed for the presence of mistic-I_{TM}-His_{10} by SDS-PAGE using SafeStain (lower image). The elution volume (mL) of the peak fractions is labelled at the top of each lane in the gel image, and the band corresponding the mistic-I_{TM}-His_{10} is labelled.

The T7-I_{TM}-His_{10} construct was expressed and purified using the same protocol as was used for the mistic-I_{TM}-His_{10} construct. As with mistic-I_{TM}-His_{10}, there was an aggregate peak with high intensity, however, the monomeric peak was well resolved and there were no contaminating proteins or degradation products in the peak fractions (Figure 3-10), with a total yield of 750 µg, as measured by $A_{280\ nm}$. Protein from the

peak fractions (75-77 mL) were pooled and concentrated to 3 mg mL^{-1}, and used to set up crystallisation trials.

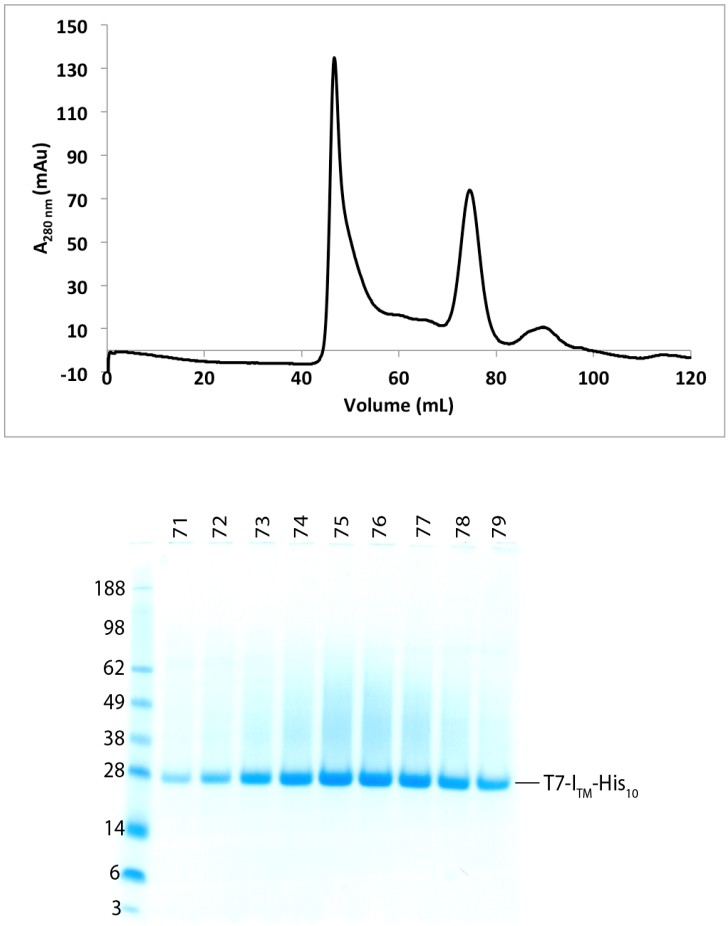


Figure 3-10: S200 size exclusion profile and SDS-PAGE gel of peak elution fractions for T7-I_{TM}-His_{10}.

Cobalt elutions containing T7-I_{TM}-His_{10} were pooled and run on a 16/60 S200 size exclusion column (upper image). Fractions from the two peaks in the $A_{280\ nm}$ elution profile were visualised by SDS-PAGE using SafeStain staining (lower image). The elution volume (mL) of the fractions is labelled at the top of each lane and the band corresponding to T7-I_{TM}-His_{10} is labelled.

3.6 Optimising purification of the *M. ruber* peripheral stalks

The *M. ruber* peripheral stalks (pLICHK-EG construct) were expressed and purified as described in section 5.7.8. In the size exclusion profile, however, two peaks can be seen, which was not expected, as the T. thermophilus peripheral stalks elute as a single peak

(Figure 3-11). Furthermore, when overlaying the elution profiles of the *M. ruber* and *T. thermophilus* peripheral stalks, the *T. thermophilus* peak lies almost exactly in the centre of the two *M. ruber* peaks, so the elution volume of the *T. thermophilus* stalks could not be used as a guide. To resolve this issue, the fractions corresponding to the two peaks were pooled separately and each was analysed by MALLS (Figure 3-12).

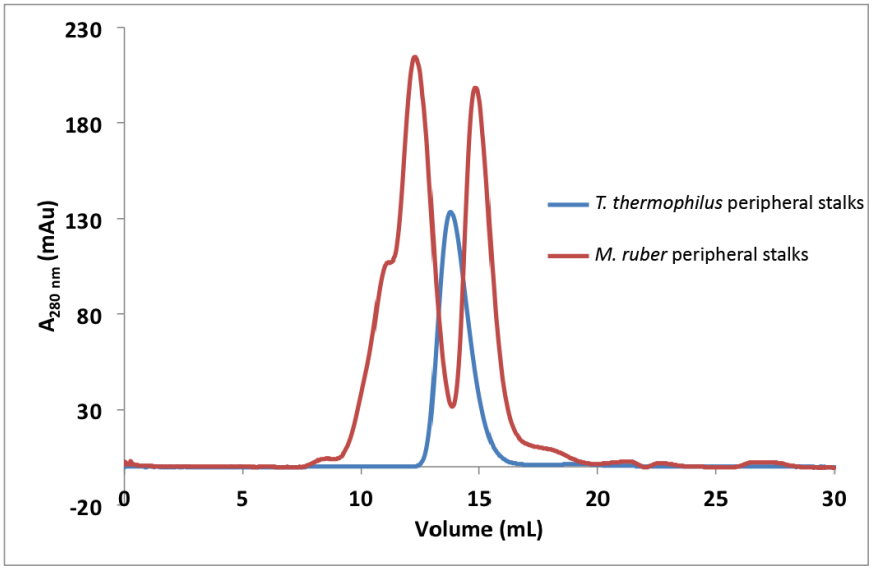


Figure 3-11: S200 size exclusion elution profiles of *M. ruber* and *T. thermophilus* peripheral stalk complexes.

The S200 size exclusion elution profiles of the *M. ruber* (red) and *T. thermophilus* (blue) peripheral stalks were plotted on the same graph to determine which of the peaks in the *M. ruber* profile corresponded to the single heterodimer.

According to the MALLS analysis the first peak had a mass of 58.8 ± 0.2 kDa and the second peak had a mass of 28.9 ± 0.3 kDa. The theoretical mass of a single *M. ruber* peripheral stalks heterodimer is 30.6 kDa, thus the second peak corresponds to the single heterodimer and the first peak contains a dimer of dimers. In the mass fraction analysis for the second peak, a small amount of the dimer was still present, possibly due to incomplete separation, however this percentage was not considered to be high enough to interfere with subsequent experiments. Furthermore, there was no evidence of time-dependent dimerisation of the peripheral stalks heterodimer, suggesting the dimerisation was not in dynamic equilibrium (see section 4.4.2). The protein in the second peak was used for interaction studies with subunit I, with a total yield of 4 mg as measured by the DirectDetect® infrared spectrophotometer.

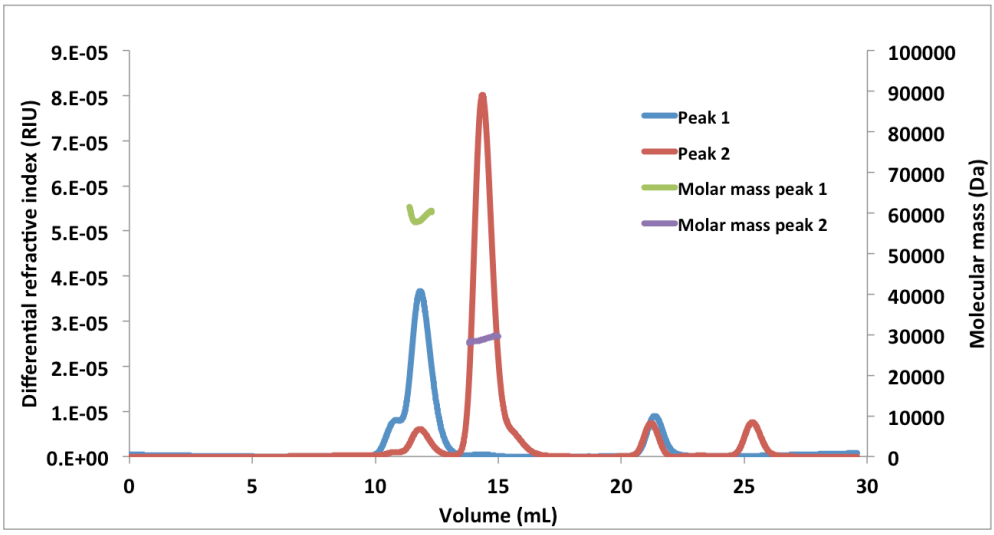


Figure 3-12: MALLS analysis of *M. ruber* EG S200 peak elution fractions.

The molecular mass and polydispersity of each of the peaks in the *M. ruber* EG size exclusion elution profile was analysed by MALLS. Each of the peaks in the size exclusion profile were pooled separately and analysed in separate MALLS experiments. The differential refractive index profiles of peak 1 (blue) and peak 2 (red) are superimposed. The calculated molecular masses of peak 1 (green) and peak 2 (violet) have also been plotted.

Chapter Four: Crystallisation trials and biophysical characterisation

4.1 Crystallisation trials with subunit I

4.1.1 Vapour diffusion crystallisation trials in Anzergent 3-12

All of the crystallisation experiments will be described in detail below and they have also been summarised in Appendix One. Various combinations of the 17 different commercial crystallisation screens (Table 6-3) were used to set up vapour diffusion trials with full-length subunit I purified in Anzergent 3-12 (5 mg mL^{-1}). Trials were initially set up with the PACT, Wizard Cryo 1&2 and Index HT screens and these were incubated at 20 ºC. Light to heavy precipitate was only ever observed in these trials, in approximately 70% of the drops.

Further crystallisation trials were set up with another purified sample of subunit I (5 mg mL^{-1}) using the MembFac, MemPlus and Cryo 1&2 screens. The Cryo 1&2 trial was repeated to determine whether different samples of subunit I gave similar results with the same crystallisation screen. The crystallisation trials were also incubated at 20 ºC and inspected periodically, however, only clear drops or heavy precipitate were observed in again approximately 70% of the drops.

To investigate the effect of temperature, vapour diffusion trials that were set up with 16 of the screens (Index HT was not available) were incubated at 4 ºC, with subunit I at 6 mg mL^{-1}. What appeared to be spherulites (radial aggregates of crystals) formed in condition A3 of the Cryo 1&2 screen (50% (v/v) PEG 200, 100 mM sodium citrate/ citric acid pH 5.5; Figure 4-1). At this time, the purification protocol for subunit I had been modified to exchange the protein into DM during size exclusion, so the only subunit I sample available for crystallisation experiments was solubilised in DM. Therefore, an optimisation tray was set up based on this condition with subunit I in DM at 6 mg mL^{-1}, in which this condition was replicated in all wells of a 96 well plate, and to each well was added one of the detergents of the Detergent Screen HT (Hampton Research, the screen contains 96 unique detergents) to see if this could promote crystallisation. Each detergent is at 10 times the CMC in the screen, and they are all

added at one-tenth the volume of the crystallisation condition so the detergent is present in the reservoir at the CMC. The drops were set up in ratios of 1:1 protein to reservoir (200 nL:200 nL) and also 1:2 protein to reservoir (200 nL:400 nL) to determine whether a higher concentration of a new detergent would increase the likelihood of crystallisation. Spherulites with a similar appearance to those in the original condition also formed in the presence of some of the additive detergents, however, no crystals grew.

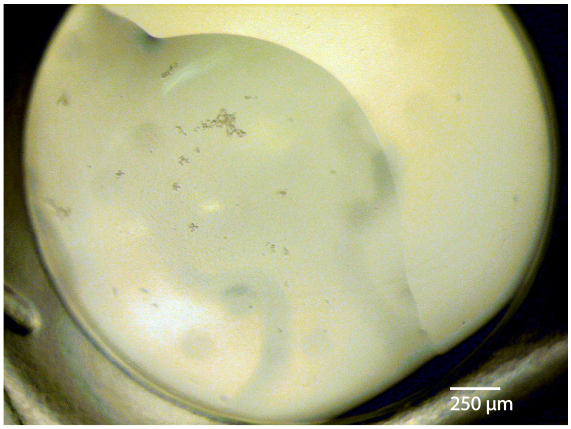


Figure 4-1: Brightfield image of condition A3 of the Cryo 1&2 crystallisation screen with subunit I in Anzergent 3-12.

Spherulites were observed in condition A3 of the Cryo 1&2 crystallisation screen with subunit I purified in Anzergent 3-12 (6 mg mL^{-1}) when the experiment was incubated at 4 ºC.

As no promising results were seen in any of these trials, and Anzergent 3-12 has not been shown to be an effective detergent for crystallising membrane proteins, subsequent crystallisation trials were conducted with subunit I that had been exchanged into DM (3.2).

4.1.2 Vapour diffusion crystallisation trials in DM

Additonal crystallisation trials with subunit I exchanged into DM were set up with MemStart™ + MemSys™ HT-96, MembFac, MemGold, Wizard Cryo 1&2 and MemPlus at 5 mg mL^{-1}, with and without 2 mM ATP. Heavy precipitate was observed in approximately 50% of the conditions, so further trials were set up with another

sample of subunit I purified in DM with the Hampton CS HT, Index and Wizard Classic 1&2 screens at 4 mg mL^{-1}.

Discrete objects that looked like spherulites formed in condition F11 of the Wizard Classic 1&2 screen (15% (v/v) ethanol, 0.1 M HEPES, pH 7.5, 0.2 M $MgCl_2$; Figure 4-2). An additive screen was set up based on this condition using the Hampton Additive Screen HT (Hampton Research; contains 96 unique reagents). No promising results, however, were observed in this screen.

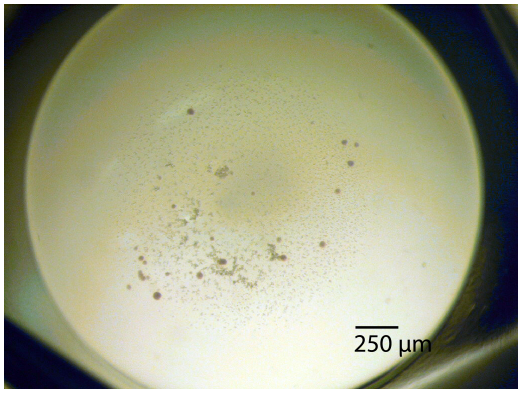


Figure 4-2: Bright-field image of condition F11 of the Wizard Classic 1&2 screen with subunit I in DM.

Spherulites were observed in condition F11 of the Wizard Classic 1&2 screen with subunit I purified in DM (4 mg mL^{-1}) when the experiment was incubated at 20 ºC.

To sample more of the crystallisation space, additional trials were set up with the JCSG-plus, MIDAS™, Classic HTS 1, Classic HTS 2, Hampton CS, Clear Strategy 1, Clear Strategy 2, Morpheus Green and Salt RX screens. These trays were set up with subunit I at 5 mg mL^{-1} (Morpheus Green and Salt RX) or 8 mg mL^{-1} (the remaining screens), to investigate the effect of altering the protein concentration. All that was observed in these trays, however, was light to heavy precipitate, or in some cases phase separation. No difference was observed with the addition of ATP.

4.1.3 Vapour diffusion crystallisation trials in Amphipol A8-35

As it appeared the subunit I sample exchanged into Amphipol A8-35 was less prone to degradation and was of higher purity (Figure 3-8), crystallisation trials were set up with

this sample using the Hampton CS HT, MemStart™ + MemSys HT-96, MemPlus, Wizard Cryo 1&2, MemGold, MembFac, MIDAS, SaltRX, Clear Strategy 1, Clear Strategy 2, Morpheus Green and JCSG-plus at 5.6 mg mL^{-1}. Phase separation was observed in some conditions, although it had a different appearance to the phase separation observed with subunit I in detergent, so this may have resulted from the crystallisation conditions interacting with Amphipol A8-35. An example of the type of phase separation seen is shown in Figure 4-3. Another type of crystallisation screen was also set up with subunit I in amphipols, the PGA Screen™ (Molecular Dimensions).[114] The screen contains poly-γ-glutamic acid polymers (PGA) of different lengths and with different branching chemical moieties, as PGA is an alternative precipitant to those most commonly used in commercial crystallisation screens, such as polyethylene glycol.

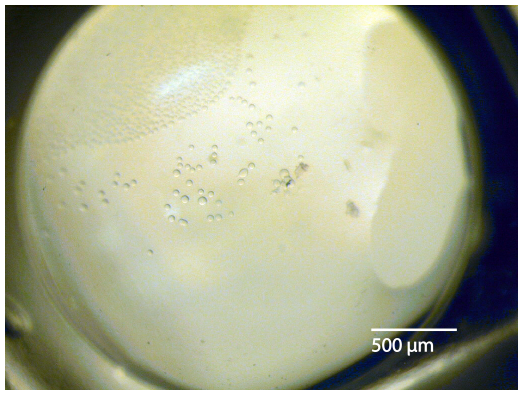


Figure 4-3: Bright field image of condition A9 of the Wizard Classic 1&2 Screen with subunit I exchanged into Amphipol A8-35

Large phase separation droplets were seen in condition A9 (1 M ammonium phosphate dibasic, 100 mM sodium acetate/acetic acid pH 4.5) of the Wizard Classic 1&2 crystallisation screen with subunit I exchanged into Amphipol A8-35 (8.6 mg mL^{-1}) when the experiment was incubated at 20 ºC.

Another sample of subunit I exchanged into Amphipol A8-35 was used to set up trials with the PACT, Classic HTS 1, Classic HTS 2, and Wizard Classic 1&2 screens at 4 mg mL^{-1}. A crystallisation trial was not set up with Index HT as this screen was unavailable at the time. No crystals, however, grew in these trials. Previously, crystallisation of membrane proteins exchanged into amphipols has only been demonstrated in lipidic cubic phase (LCP), whereby the membrane protein is exchanged from the amphipols into the LCP.[115] As most commonly used amphipols carry a net

charge and are weak acids, it may be the case that amphipols are not compatible with aqueous crystallisation to the same extent as detergents.[116]

4.1.4 LCP trials in DM

Given the number of membrane protein structures that have been solved from crystals grown in LCP, this technique was also attempted as a means of crystallising subunit I in DM, with and without ATP (5.8.2). Four different crystallisation screens were set up with the LCP containing subunit I at 10 mg mL^{-1}, with incubation at 20 ºC. The plates were inspected daily for crystal growth using a light microscope with a polarising filter. What appeared to be phase separation was observed in some conditions (Figure 4-4A and B) however, the experiments did not progress from this state. It was thought there might be microcrystals in some conditions, but these subsequently proved to be the conversion of the LCP to the lamellar mesophase, in which the planar lipid bilayers become stacked on top of one another, and are therefore highly birefringent (Figure 4-4C).[117] It was suggested that the concentration of protein was not high enough to form crystals, however there was insufficient time to set up further LCP experiments.

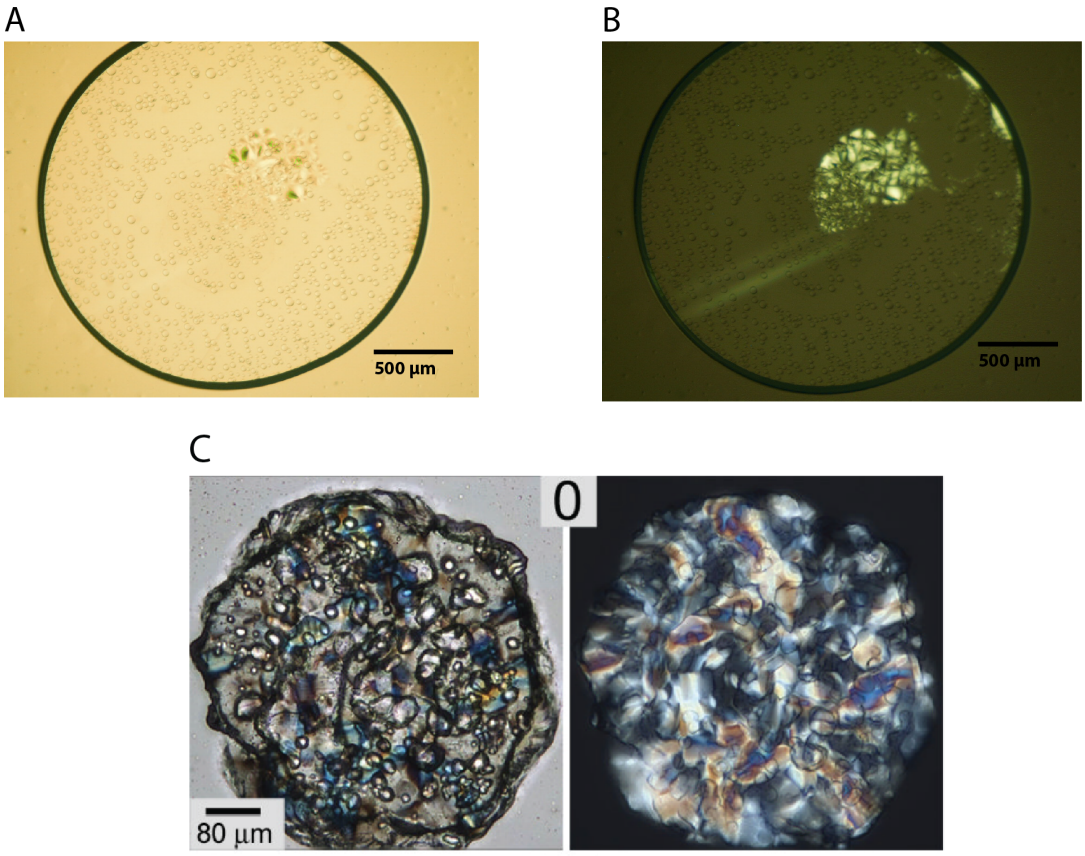


Figure 4-4: Outcomes of LCP crystallisation trials with subunit I

A: Phase separation was seen in condition A10 of the MemGold screen in the LCP trial with subunit I (10 mg mL^{-1}). B: Condition A10 of the MemGold screen seen under plane polarised light. The lamellar mesophase, which is birefringent, had begun to form in this condition. C: Example of an LCP condition that has undergone a complete conversion to the lamellar mesophase under visible light (left) and between crossed polarisers (right). This image was taken from Figure 13, Caffrey and Cherezov (2009).[117]

4.2 Crystallisation trials with I_{TM} constructs

4.2.1 Vapour diffusion crystallisation trials with mistic-I_{TM}-His_{10}

Mistic-I_{TM}-His_{10} that had been purified in DM was used to set up vapour diffusion crystallisation trials with the MemGold, MemStart + MemSys, PACT and Index crystallisation screens at 3.4 mg mL^{-1}, which were incubated at 20 ºC. The next day, two crystalline objects, which exhibited UV fluorescence, appeared in well F10 of the MemStart + MemSys trial (0.1 M NaCl, 0.1 M Tris-HCl, pH 8.5, 0.1 M $MgCl_2$, 30% (w/v) PEG 400) (Figure 4-5). An optimisation screen was set up based on this

condition, in which the concentrations of PEG 400, NaCl and $MgCl_2$ were varied and the concentration and pH of Tris-HCl, pH 8.5 were kept constant. The optimisation tray was incubated at 20 ºC and inspected periodically, however, no crystals were observed, and over time heavy precipitate formed in most of the wells.

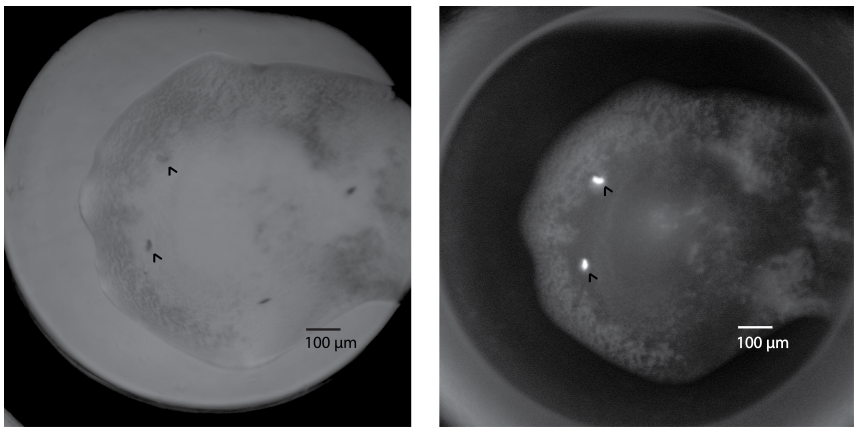


Figure 4-5: Bright-field and UV fluorescence images of condition F10 of the MemStart + MemSys crystallisation trial with mistic-I_{TM}-His_{10}.

Crystalline objects (~50 µm × 25 µm) formed in condition F10 of the MemStart + MemSys screen with purified mistic-I_{TM}-His_{10} (3.4 mg mL^{-1}). An image of the well was taken under visible light (left) and UV fluorescence (right).

As no crystals grew in the optimisation screen, a mistic-I_{TM}-His_{10} sample from a subsequent purification, also concentrated to 3.4 mg mL^{-1}, was used to set up crystallisation trials with the Wizard Cryo 1&2, MemPlus and MemStart + MemSys HT-96 screens. The MemStart + MemSys HT-96 trial was repeated to see if different samples of mistic-I_{TM}-His_{10} produced the same results in different crystallisation trials. Phase separation that had not previously been seen in any condition with this construct was observed in Wizard Cryo 1&2 condition G6 (40% (w/v) PEG 600, citrate pH 5.5; Figure 4-6). An additive screen was conducted based on this condition using the Detergent Screen HT. Similar phase separation to what was seen in the original condition also developed in a number of wells in this screen, however, no protein crystals grew.

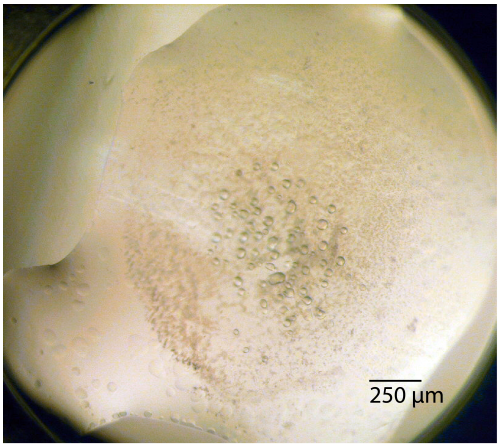


Figure 4-6: Bright-field image of phase separation in condition G6 of the Wizard Cryo 1&2 screen with mistic-I_{TM}-His_{10}.

Phase separation was observed in condition G6 of the Wizard Cryo 1&2 screen with mistic-I_{TM}-His_{10} (3.4 mg mL^{-1}) when the experiment was incubated at 20 ºC.

4.2.2 Vapour diffusion crystallisation trials with T7-I_{TM}-His_{10}

As the mistic tag could not be cleaved, mistic-I_{TM}-His_{10} was replaced with the T7-I_{TM}-His_{10} construct, as it was thought the small T7 tag would be less likely to interfere with crystallisation (3.5). Purified T7-I_{TM}-His_{10} (2.2 mg mL^{-1}) was used to set up crystallisation trials with the MemStart + Memsys HT-96, Wizard Cryo 1&2, MembFac, PACT, MemPlus, MemGold and Wizard Classic 1&2 screens. No promising results, however, were observed in these trials. This may have been because the concentration was too low, although precipitate was observed in approximately 80% of the conditions, so it may be because this construct is not very stable, as there was an intense aggregate peak in the size exclusion profile (Figure 3-10). Due to the limited amount of sample and time constraints, however, no further crystallisation trials were attempted with this construct.

4.3 Subunit I:peripheral stalks pulldown assay

As described in section 5.7.8, following cleavage of the His_6 tag from the peripheral stalks, the cleavage reaction was incubated with Ni-NTA resin to remove the TEV protease and any protein that still contained the His_6 tag. As can be seen in Figure 4-7, the cleaved peripheral stalks do not bind to the resin, and therefore the stalks would not

bind non-specifically to the Ni-NTA resin, which was important to confirm before the pulldown assay was conducted.

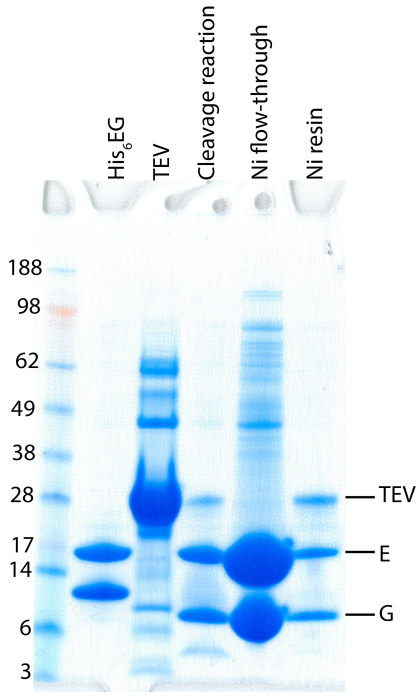


Figure 4-7: SDS-PAGE analysis of peripheral stalks pre- and post-TEV cleavage
Ni-NTA elutions containing the peripheral stalks were pooled and TEV protease was added to cleave the His_6 tag. To confirm the tag had been cleaved, a pre-cleavage sample of the stalks, the TEV sample, the cleavage reaction, the Ni-NTA flow-through post-cleavage and the Ni-NTA resin were analysed by SDS-PAGE using SafeStain staining. The sample that was analysed is labelled at the top of each lane and the different proteins are labelled on the right side of the image.

To determine whether the recombinantly expressed subunit I and peripheral stalks were able to form a complex, the interaction was tested using a nickel-NTA pulldown assay, using purified subunit I and purified peripheral stalks from only the heterodimer peak in the size exclusion profile (Figure 3-5 and Figure 3-11). It is evident in Figure 4-8 that subunit I and both of the peripheral stalk subunits are present in all of the Ni-NTA elutions (labelled Elution 1-3). As the peripheral stalks do not bind to Ni-NTA following removal of the His_6 tag, the assay demonstrates that subunit I is able to bind to the peripheral stalks heterodimer (Figure 4-8). This demonstrates that the interaction that occurs between subunit I and the peripheral stalks in the native A-ATPase enzyme is able to be replicated *in vitro*, suggesting all three recombinantly expressed proteins are correctly folded and functional. Furthermore, if subunit I had not been correctly

folded, the whole sample would most likely have precipitated on the Ni-NTA resin, rather than only a small proportion of the total sample remaining bound to the resin, as can be seen in Figure 4-8 (lane labelled Ni resin). Nevertheless, as a pulldown does not truly test the interaction in solution, it was necessary to confirm this result with MALLS and ITC.

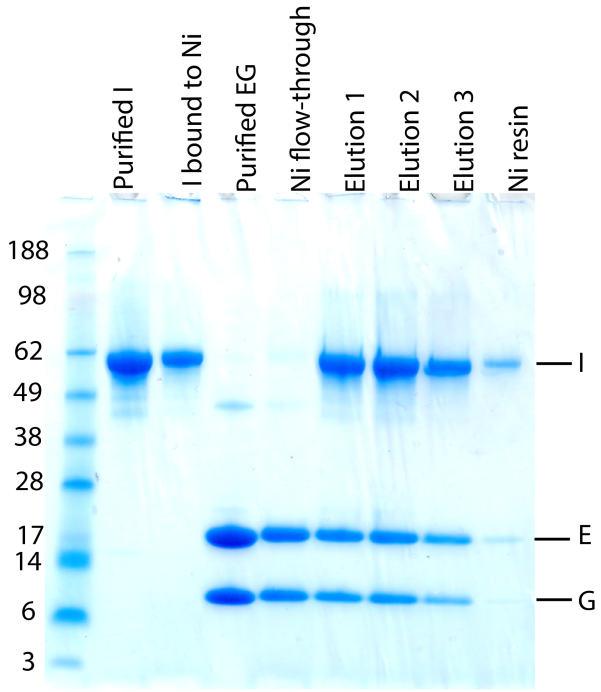


Figure 4-8: SDS-PAGE analysis of subunit I:peripheral stalk Ni-NTA pull-down assay

SEC-purified subunit I was immobilised on Ni-NTA resin and incubated with SEC-purified peripheral stalks single heterodimer. The flow-through was discarded, the resin was washed and bound proteins were eluted with 300 mM imidazole. Aliquots from the purified subunit I sample, the Ni-NTA resin on which subunit I had been immobilised, the purified peripheral stalks, the flow-through post-incubation, the elutions and the Ni-NTA resin post-elution were all analysed by SDS-PAGE using SafeStain staining. The samples are labelled at the top of each lane in the gel image, and the proteins are labelled to the right of the image.

4.4 MALLS analyses

4.4.1 Subunit I in DM

The MALLS analysis of the peripheral stalks was described in section 3.6, and it showed that the single heterodimer could be purified as a monodispersed species. It was

also necessary to determine the oligomeric state and polydispersity of subunit I. Fractions from the middle of the S200 size exclusion peak were pooled (Figure 3-4, lanes labelled 11 and 12), and these were applied to the same column and analysed by an inline MALLS machine. From the light scattering measurements, it can be seen that the majority of the subunit I sample elutes in the monomeric state, with a mass of 65.2 ± 0.5 kDa, and the sample is monodispered (polydispersity = 1 ± 1%) (Figure 4-9). The theoretical molecular mass of subunit I, however, is 77.3 kDa, which is significantly larger than the result that was obtained. This is most likely because the extinction coefficient calculated for subunit I using ProtParam[118] is incorrect. Obtaining the correct extinction coefficient would be quite time consuming, and the main aim of this experiment was simply to determine the oligomeric state and polydispersity of purified subunit I. As these results are consistent with subunit I being monomeric, as a mass of 65.2 kDa could not be anything larger than a monomer, and monodispersed, they were considered sufficient for this aim. As with the ITC data, the MALLS data could perhaps be improved by purifying subunit I in the absence of glycerol and using the DirectDetect® to measure the concentration of the protein.

As can be seen in Figure 4-9, there is a small amount of aggregate, but the intensity is sufficiently low that this was not considered to be a significant problem. Furthermore, this indicated that once the monomeric protein was separated from the aggregated protein, it was not prone to further aggregation. This suggests that the aggregation occurred in the early stages of the purification, although aggregation was not monitored over time. The peak that eluted just after 20 mL contained no protein (protein fraction moment 0.003 ± 21%), and was thus most likely ATP, as the SEC buffer contained 2 mM ATP. A conjugate analysis was performed to calculate the individual masses of the detergent and subunit I within the micelle. This is possible because subunit I and DM have different refractive index increment (dn/dc) values (see section 5.10.1), and only subunit I is able to absorb UV light. The ASTRA conjugate analysis algorithm utilises scattering, refractive index and UV absorbance data to separately calculate the concentrations of the protein and detergent within the micelle. The total mass of the protein + DM micelle is 131.74 ± 1 kDa, thus the composition of the protein + DM micelle is 50% protein, 50% detergent.

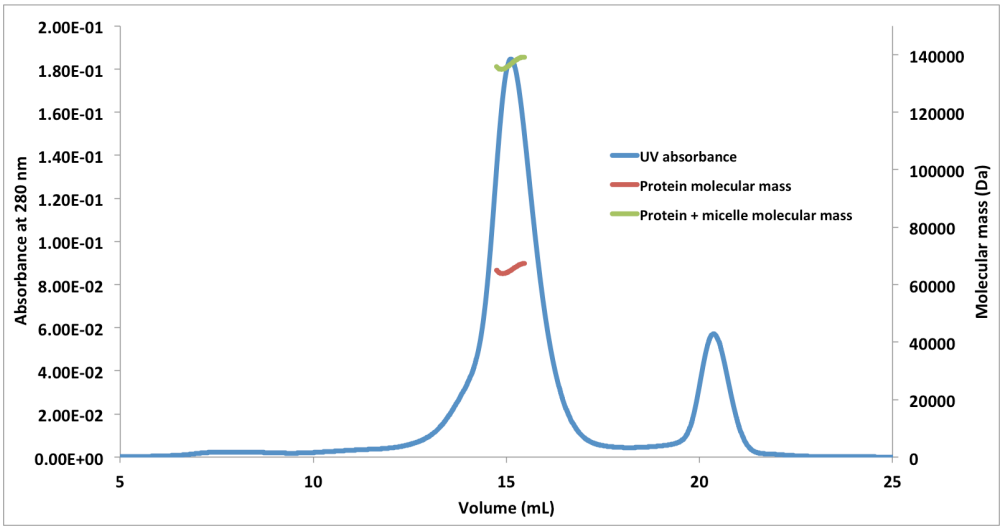


Figure 4-9: MALLS analysis of subunit I purified in DM.

The molecular mass and polydispersity of subunit I purified in DM was calculated using MALLS. The protein was run on a 10/300 superose 6 size exclusion column and protein elution was monitored by measuring absorbance at 280 nm (blue). The molecular masses of subunit I (red) and subunit I + the DM micelle (green) were determined by a conjugate analysis of the UV absorbance and the light scattering data (the dn/dc and the $E_{1\%}$ for protein are 0.80 and 1.14, respectively, and the dn/dc and $E_{1\%}$ for DM are 0.14 and 0.00, respectively).

4.4.2 Subunit I:peripheral stalk complex

To confirm the subunit I:peripheral stalk complex could form in solution, and to determine the stoichiometry of this interaction, the two proteins were mixed together and analysed by MALLS. As Figure 4-10 shows, a peak eluted at 13.5 mL with a mass of 91.6 ± 0.6 kDa and a polydispersity of 1 ±1%. This value would most closely fit the mass of a 1:1 complex of subunit I monomer and one peripheral stalk heterodimer, which has a theoretical molecular mass of 107.9 kDa as calculated by ProtParam. This of course assumes that the complex does not undergo dissociation on the column; although this is possible given the low micromolar affinity of the interaction calculated by ITC (see section 4.5). As was seen with the MALLS analysis of subunit I, the calculated mass was 65.2 kDa, approximately 12.2 kDa smaller than the theoretical mass. Additionally, in this buffer the calculated mass of the peripheral stalks is 26.2 ± 0.6 kDa (data not shown), approximately 3 kDa smaller than the theoretical mass, and together these account for the difference between the calculated mass of the first peak in Figure 4-10 and the theoretical mass of a 1:1 subunit I:peripheral stalk

complex. The calculated molecular mass for the peripheral stalks is different to that calculated in Figure 3-12, and this is most likely because they were purified in different buffers. The molecular masses uncertainties were calculated using the following equation in the ASTRA® analysis software (Wyatt Technologies):

$$M_{uncertainty} = M\sqrt{M^2 * fit_{covariance} * fit_{\chi^2} + c_{uncertainty}}$$

Where M is the protein molecular mass and c is protein concentration. During data processing, data from each detector is weighted according to the level of fluctuation in the baseline (the first or last 10% of the data points); this corresponds to the $fit_{covariance}$ term of the equation. The data are then fit to a theoretical scattering model and a χ^2 test is performed to determine how well the data can be fitted to the model (the fit_{χ^2} term of the equation), and thus whether the weighted baseline can be used to normalise the data points used to calculate the protein molecular mass. The equation also includes a term to account for the uncertainty in protein concentration. Ultimately, the molecular mass uncertainty is a measure of the level of noise in the data.

The fractions from the peak at 13.5 mL were analysed by SDS-PAGE using SafeStain staining, and subunits I, E and G can all be seen in these fractions (Figure 4-11, fractions eluting at 13.5-14.5 mL). The intensity of these bands was not measured, as SafeStain staining is not sufficiently quantitative, however the intensity of the peripheral stalks bands appears to be proportional to the subunit I bands, which is in agreement with the mass calculated for this peak in the MALLS analysis for a 1:1 complex.

A second peak eluted at ~17 mL, however this peak could not be analysed as the intensity of the light scattering was above the range of the detector. Nevertheless, when this peak was analysed by SDS-PAGE, it was found it contained only the peripheral stalks (Figure 4-11), fractions eluting at 16-17 mL), and as the stalks had already been analysed by MALLS it was not necessary to repeat this.

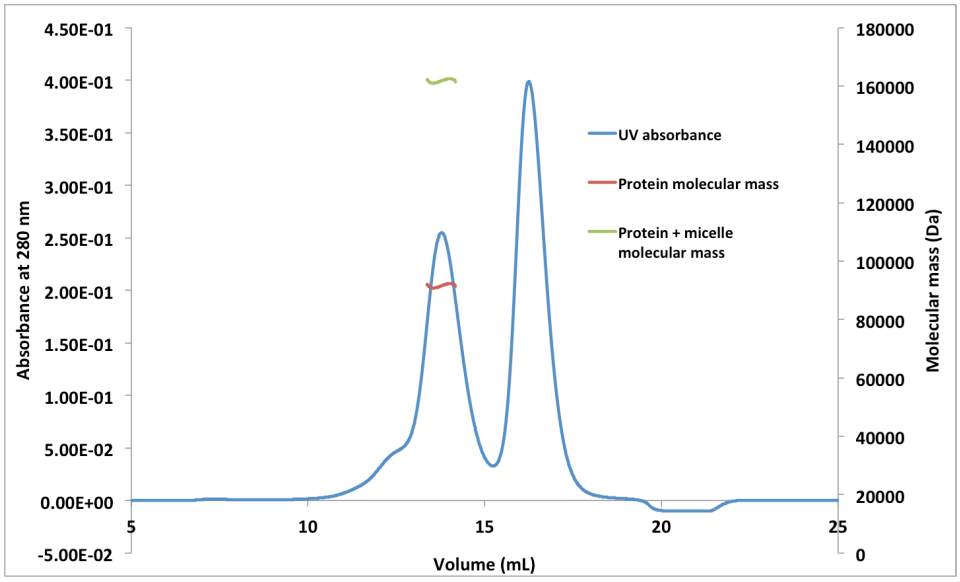


Figure 4-10: MALLS analysis of the subunit I:peripheral stalk complex

Purified subunit I and an ~four-fold excess of the peripheral stalks samples were mixed and applied to a superpose 6 column with an inline MALLS instrument. Protein elution was monitored by absorbance at 280 nm (blue). The molecular masses of the protein complex (red) and the protein complex + DM micelle (green) were calculated by a conjugate analysis of the UV absorbance and the light scattering data (the dn/dc and the $E_{1\%}$ for protein are 0.80 and 1.00, respectively, and the dn/dc and $E_{1\%}$ for DM are 0.14 and 0.00, respectively).

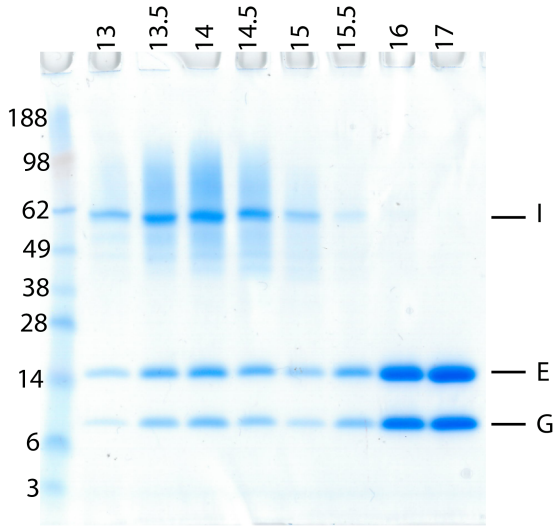


Figure 4-11: SDS-PAGE of the subunit I:peripheral stalk MALLS elution fractions

The peak fractions in the MALLS $A_{280\ nm}$ elution profile were visualised by SDS-PAGE using SafeStain staining. The elution volumes (mL) of the fractions are labelled at the top of each lane and the bands corresponding to the different proteins are labelled

4.5 Isothermal titration calorimetry with subunit I and the peripheral stalks

All ITC data were collected and analysed with the assistance of Prof Jacqueline Matthews. The experiment was first performed with subunit I that had been purified in DM, however, double peaks were observed with each injection of the peripheral stalks, which is a common complication of performing ITC with detergent-solubilised samples.[119] These data were therefore not analysed, and the experiment was subsequently performed with subunit I that had been exchanged into Amphipol A8-35 (5.7.7).

The amphipol-exchanged sample did not give double peaks during the titration, however a complete thermogram (Figure 4-12, upper panel) was not observed, indicating the peripheral stalks were present at a saturating concentration and the actual concentration of subunit I was lower than that measured. A complete thermogram would have a sigmoidal shape, rather than the hyperbolic shape seen for this titration. This is because the first few points of the curve should show minimal binding, as the concentration of subunit I should have been below the dissociation constant for the interaction.[120] The concentration of this subunit I sample had been measured using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific), which most probably gave an inaccurate reading for the concentration of subunit I, as the extinction coefficient calculated for subunit I is unlikely to be correct (4.4.1). These protein samples were purified in the presence of glycerol, and thus the DirectDetect® could not be used to measure protein concentration. The ITC data could perhaps be optimised by obtaining more accurate concentration measurements for subunit I and the peripheral stalks using the DirectDetect®, however, time constraints did not permit this.

Nevertheless, binding of the peripheral stalks to subunit I was observed, although due to the incompleteness of the thermogram an accurate stoichiometry could not be determined (Figure 4-12, lower panel). It is clear, however, that there was only one binding event over the course of the titration, thus the data were fit to a one-site model. Furthermore, while an accurate K_d could not be obtained for the same reason, from this experiment it was determined that one peripheral stalk heterodimer binds to subunit I

with low μM affinity. This is approximately an order of magnitude lower than what was expected, as a previous study with the *Enterococcus hirae* bacterial A/V-ATPase found using surface plasmon resonance that the peripheral stalks bound to the soluble domain of subunit I with an equilibrium dissociation constant of 190 nM,[121] and thus further optimisation is required to obtain a more accurate dissociation constant for the interaction between the *M. ruber* subunits.

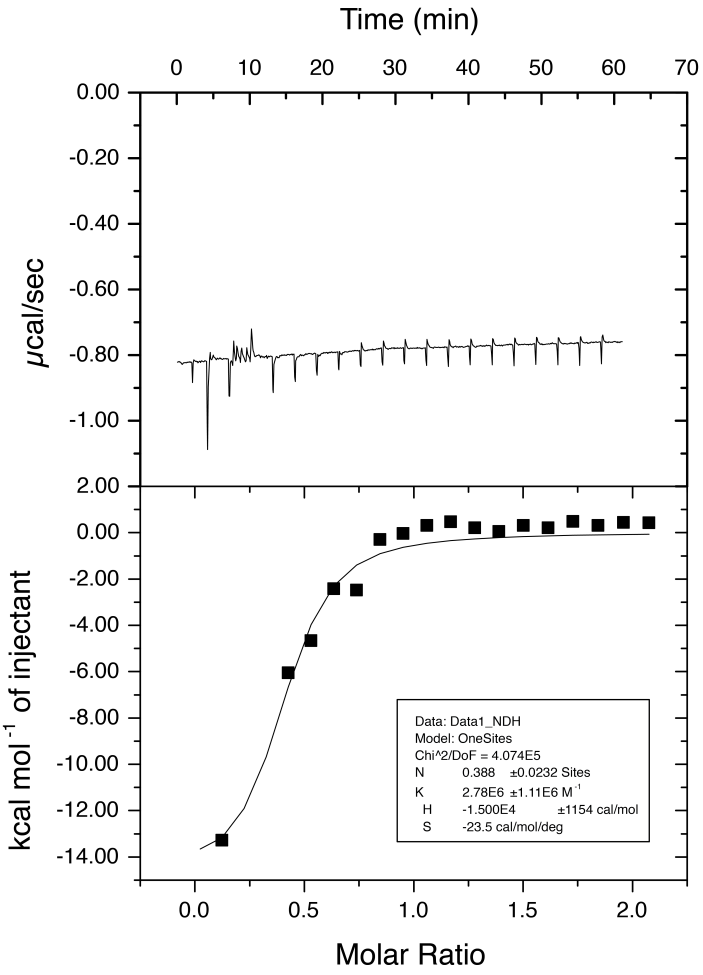


Figure 4-12: ITC titration of peripheral stalks into subunit I.

Purified peripheral stalks (100 μM) were titrated into amphipol-exchanged subunit I (10 μM). The upper panel shows the ITC thermogram. Individual titrations were integrated using Origin, and the resulting data points (coloured squares, lower panel) were fitted to generate the binding isotherm, assuming a 1:1 stoichiometry. The binding isotherm was plotted as the enthalpy per mole of peripheral stalks versus the peripheral stalks to subunit I molar ratio.

This result confirms that the recombinantly expressed subunit I and peripheral stalks constructs are able to form a 1:1 complex. It is unclear why they do not form a 1:2 complex, as is seen in the native A-ATPase enzyme, yet this is consistent with previous findings with the equivalent subunits from the V-type ATPase that had been recombinantly expressed. In these experiments, which included ITC and X-ray crystallography, subunit C (the equivalent of the soluble domain of A-type subunit I; Figure 1-10) and the peripheral stalks from *S. cerevisiae*[122,123] or *E. hirae*[121] only ever formed a 1:1 complex, and in each of the experiments, only the C_{head} binding site was ever occupied. It was suggested in these studies that binding of the second peripheral stalks heterodimer to subunit C most likely only occurs when the rest of the V-ATPase complex has been assembled, as perhaps binding of the second peripheral stalks heterodimer requires the peripheral stalks and/or subunit C (and subunit I) to adopt particular conformations that are not present when the proteins are free in solution (section 5.4).

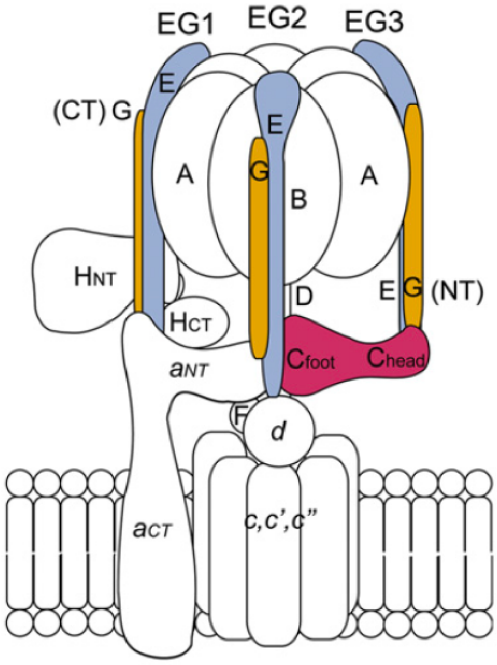


Figure 4-13: Schematic of the *S. cerevisiae* V-type ATPase showing subunit interactions.

This cartoon, adapted from Fig. 1 in Oot et al. (2012),[63] illustrates the arrangement of the subunits within the *S. cerevisiae* V-type ATPase. The third peripheral stalks heterodimer labelled EG3 (blue and orange) interacts with the domain of the C subunit labelled C_{head} (magenta). This is the only EG:C interaction to be observed in *in vitro*

binding experiments; binding of the EG2 heterodimer to C_{foot} has not been observed experimentally.

Chapter Five: Discussion

The ultimate aim of this project was to determine the X-ray crystal structure of the proton-translocating subunit of the A-type rotary ATPase. At the beginning of this project no high-resolution structural information was available for this subunit, nor for the proton-translocating subunit of any of the rotary ATPase subtypes. This aim was unfortunately not achieved, as a stable subunit I construct was not identified, and membrane protein crystallography is still very challenging for many reasons. Stable protein is defined in this context as protein that adopts one major conformation and is therefore less prone to aggregation, and as was seen in the results chapters, all of the subunit I constructs that were tested were prone to aggregation. This section will aim to address these challenges and how they might be overcome to obtain high-resolution structural information for subunit I, specifically in light of recent advances in the field of rotary ATPase/synthase structural biology and membrane protein structural biology more broadly. In addition, the interaction between subunit I and the peripheral stalks was also investigated, and the implications of these findings will also be discussed in this section.

5.1 Subunit I remains a challenging structural target

As discussed previously, the proton-translocating subunit is the only rotary ATPase/synthase subunit for which there is no atomic-resolution structural information. This is primarily due to it being a membrane protein, thereby making it considerably difficult to solubilise and purify, which will be addressed further in the next section. During the course of this work, a number of subunit I homologues were tested for bacterial recombinant expression and purification, along with a number of different construct lengths and affinity/solubility tag combinations. Nevertheless, all of these constructs were prone to aggregation or degradation, or both. This was evident from the fact that there was often an aggregate peak present in the size exclusion profile, and moreover, when the elution fractions from these experiments were visualised by SDS-PAGE, a number of other proteins were present in these fractions, confirmed by Western blot to be degradation products. Furthermore, independent of the protein concentration, medium to heavy precipitation was always seen in at least 70% of the conditions in crystallisation trials with all of the different screens, suggesting the protein

was inherently unstable. It was thought that creating a sub-complex with the peripheral stalks might induce subunit I to adopt a single conformation, however as can be seen from the MALLS data (Figure 4-10), the peak containing the complex had a shoulder and was quite broad, indicating the sample was conformationally heterogeneous, which would make it unsuitable for crystallisation trials. When these fractions were visualised by SDS-PAGE (Figure 4-11) it was evident that the subunit I degradation products were still present, so the peripheral stalks had not displaced them.

It was suggested in section 2.6.1 that a sub-complex could also have been formed with subunit I and the membrane rotor ring. Recent studies have shown however, that even when the proton-translocating subunit is in complex with the membrane rotor ring, it is still inherently flexible. When the A_O sub-complex (subunits ICL_{12}) of the *T. thermophilus* enzyme was analysed by ion mobility mass spectrometry, it was found that the arrival time of the sub-complex at the detector formed a broad peak, indicating the sub-complex was present in a number of different conformations.[124] When subunit I was removed from this sub-complex, the arrival time distribution was significantly reduced, indicating subunit I was the source of the conformational heterogeneity. Furthermore, the addition of ATP to ICL_{12} reduced the arrival time distribution, but this effect was only seen in the presence of subunit I; the addition of ATP had no effect on the arrival time distribution of the CL_{12} subcomplex (Figure 5-1). ATP was tested as a stabilising agent for subunit I by adding it to the buffers during purification, however it did not seem to make a reproducible difference to the yield obtained for subunit I. ATP may therefore only bind to subunit I when it is the A-ATPase complex, and not when it is isolated.

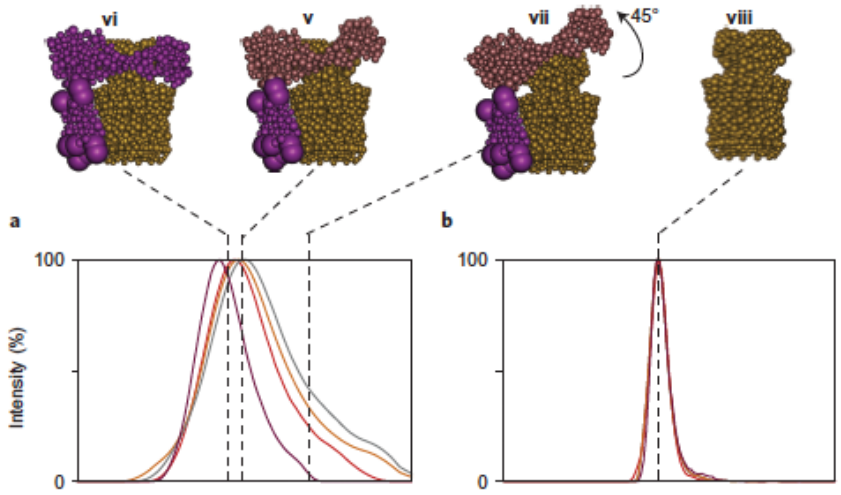


Figure 5-1: Ion mobility mass spectrometry arrival time distributions for ICL_{12} and CL_{12}.

a: Arrival time distributions for ICL_{12} purified with ATP (purple), AMPPNP (orange), EDTA (grey) or no addition (red). The dashed lines indicate calculated arrival times for the models labelled v, vi and vii, in which I_{sol} (purple in vi, pink in v and vii) is at varying angles to I_{TM}. **b**: Arrival time distributions for CL_{12} purified under the same conditions described in **a**. The dashed line indicates the calculated arrival time for the model labelled viii. Adapted from Zhou et al. (2014) Figure 2.[124]

The inherent flexibility of the proton-translocating subunit has also been demonstrated by the fact it is poorly resolved in the recent X-ray crystal structure of the *Paracoccus denitrificans* F-ATPase (discussed in detail below), in which only the α-helices of the a subunit (equivalent of subunit I) in contact with the membrane rotor ring have strong electron density (section 5.3).[125] The authors acknowledge that the electron density for this subunit is relatively featureless with little sidechain density, and thus the protein sequence could not be modelled into the density unambiguously. Both the *T. thermophilus* and the *P. denitrificans* samples were purified from source, which suggests that even when the proton-translocating subunit is in complex with the rest of the native enzyme and has not been expressed recombinantly, it is still highly flexible and challenging to structurally characterise. For future structural studies to be conducted with subunit I, or any proton-translocating subunit, efforts will need to be made to stabilise the protein. Some strategies that have been successfully employed for other membrane proteins are discussed below.

5.2 Strategies for determining membrane protein structures

Due to their hydrophobic nature, integral membrane proteins such as subunit I require detergents to extract them from the membrane and solubilise them in aqueous solutions. The first main problem with the need for detergents is that integral membrane proteins have evolved to be most stable within the cell membrane lipid bilayer of the native organism, and are therefore not stable when they are extracted from this environment. The second complication is that most membrane proteins are more stable in mild detergents (those that form large micelles, e.g. DDM) whereas harsher detergents (those that form smaller micelles, e.g. OG) are more compatible with X-ray crystallography, as due to their smaller micelle size these detergents are less likely to obstruct the formation of crystal contacts between the hydrophilic surfaces of the protein molecules. Nuclear magnetic resonance spectroscopy (NMR) and cryo-EM are more tolerant to harsher detergents, however NMR is often not compatible with proteins larger than 30 kDa, whereas cryo-EM is limited to proteins larger than 100 kDa. A number of membrane protein protomers (individual proteins that have not formed and oligomer or are not part of a protein complex) often fall within the 30-100 kDa range.

To overcome the first problem, a number of different strategies have been developed over the last decade or so to specifically address the challenges of obtaining a stable membrane protein for structural studies. Perhaps the most simple of these is selecting and testing multiple different homologues of the protein of interest, and then creating multiple different truncation constructs of the chosen protein to identify the most stable core construct. This approach was partially adopted for this project, as 12 different homologues of subunit I were selected and expression and purification of each of these was tested on a small scale. Constructs were also created that contained only the transmembrane domain, to determine whether the isolated transmembrane domain was more stable and amenable to crystallisation. As a stable and crystallisable subunit I construct could not be created, it is evident now that further truncations could also have been tested, such as serially truncating the N- and C-termini of the protein to find the most stable construct. Another common approach involves pairing these N- and C-terminal truncations with the incorporation of a stabilising motif into the most flexible region of the protein, such as the insertion of T4 lysozyme into intracellular loop 3 (IL3) of the inactive β-2 adrenergic receptor,[126] the insertion of the

thermostabilised apocytochrome b_{562}RIL triple mutant (one residue was mutated to arginine, another to isoleucine and the third to leucine) into IL3 of the A_{2A} adenosine receptor[127], or the N-terminal fusion of apocytochrome b_{562}RIL with the angiotensin receptor in complex with its selective antagonist.[128] Nevertheless, this would have been a very time-consuming and expensive process, and there is no guarantee a stable construct would have been identified in the timeframe of this project. It is not currently known which regions of the transmembrane domain of subunit I are the most flexible.

Co-crystallisation of membrane proteins with fragment antibodies (usually the antigen-binding fragment, or Fab) or single-domain antibodies (dAb) raised against the protein has also been used quite successfully, such as in the determination of the active, agonist-bound β-2 adrenergic receptor,[129] the human histamine H_1 receptor in complex with doxepin,[130] and the KcsA K^+ channel.[131] The advantage of this approach is two-fold: the Fab or nanobody has the ability to lock the membrane protein in one conformation, and this sample homogeneity improves the likelihood of crystallisation, and secondly the Fab/dAb increases the soluble surface that is available to form crystal contacts, as the detergent micelle often obstructs the limited hydrophilic regions of membrane proteins from forming crystal contacts. The Fab/dAb can be raised by either immunising mice with the protein of interest and purifying the Fab from the blood, or phage display can be used to select high-affinity dAbs *in vitro*. Phage display was performed by Meghna Sobti against subunit I during the course of this project, however no high-affinity dAbs were identified, so this strategy was not pursued further.

Thermostabilisation, whereby point mutations are introduced that increase the stability of the protein of interest, is a strategy that has been utilised very successfully for the G protein-coupled receptors (GPCRs).[132] This process however, as well as being costly and time-consuming, also requires an assay to select a single conformation of the protein and ensure the thermostabilised mutants are functional. This works well for GPCRs for which the endogenous ligand is known or another compound that modulates its activity in a predictable way is available. Yet as subunit I is not functional when isolated from the intact rotary ATPase complex, it was not possible to develop a functional assay, and therefore thermostabilisation through systematic mutagenesis could not be attempted with subunit I during the course of this work.

In the last few months however, a bioinformatics-based approach has been developed that enables a thermostabilised construct to be identified without the requirement for a functional assay.[133] Using this method, the sequences of homologues of a protein of interest, in this case a much larger number than the 12 studied in this project, from organisms that have a range of optimal growth temperatures are collected and multiple sequence alignments are assembled and sorted based on the growth temperature. From this, sequence differences that correlate with organisms living at higher temperatures to the optimal growth temperature of the organism of interest can be identified, and these mutations can be made and tested. The reason this reduces the need for an activity assay is because these thermostabilising mutations arose in proteins that are known to be functional. The obvious issue here for subunit I from *M. ruber* is that a number of the other 11 organisms that were sources of subunit I genes have higher optimal growth temperatures than *M. ruber* (Table 5-1), yet the full-length subunit I genes from these organisms were either not expressed or were extensively degraded (Figure 2-1). Nevertheless, there must be some feature of the *M. ruber* subunit I sequence that enables the full-length protein to be expressed recombinantly in *E. coli*, and so perhaps introducing thermostabilising mutations from subunit I genes from organisms with higher optimum growth temperatures may further stabilise *M. ruber* subunit I and minimise or prevent the degradation and aggregation that was routinely seen during purification. If this project were to be pursued further, identifying thermostabilising mutations using this *in silico* approach would ideally be the next strategy to be pursued, and if a thermostabilised construct could be identified, detergent screens would be conducted with this protein followed by further LCP crystallisation trials.

Table 5-1: Optimum growth temperatures of source organisms for subunit I genes

Organism	Optimal growth temperature (ºC)
Sulfolobus solfataricus	75[134]
Pyrococcus furiosus	100[135]
Methanosarcina mazei	40[136]
Thermoplasma acidophilum	60[137]
Archaeoglobus fulgidus	83[138]
Methanocaldococcus jannaschii	85[139]
Pyrococcus horikoshii	95[140]
Methanocaldococcus maripaludis	38[141]
Thermus thermophilus	65[142]
Meiothermus ruber	60[143]
Thermococcus gammatolerans	88[103]
Deinococcus radiodurans	30[144]

5.3 Recent medium-resolution cryo-EM and X-ray crystal structures of intact rotary ATPases

At the commencement of this project, the highest resolution for an intact rotary ATPase structure was the 9.7 Å cryo-EM structure of the *T. thermophilus* A-ATPase (section 1.7.3), and this therefore represented the highest resolution structure of the proton-translocating subunit for any rotary ATPase subtype. In this structure, the individual helices of subunit I were visualised for the first time, forming two offset bundles which provided structural evidence for the two half-channel model, although the atomic details were yet to be filled in. During the latter stages of this project, a number of medium-resolution cryo-EM and crystal structures of intact rotary ATPases were published. These include the 6.2 Å cryo-EM structure of the *Polytomella* sp. F-ATPase,[145] the 6.9

– 8.3 Å resolution cryo-EM structures of the *S. cerevisiae* V-ATPase in different rotational states,[146] the 6.4 – 7.4 Å resolution cryo-EM structures of the bovine mitochondrial F-ATPase in different rotational states,[147] the 4 Å resolution X-ray crystal structure of the *P. denitrificans* bacterial F-ATPase,[125] and the 6.4 - 9.5 Å resolution cryo-EM structures of the *T. thermophilus* A-ATPase in two different conformations (Figure 5-2).[148] In all of these structures, the proton-translocating subunits and their interactions with the neighbouring subunits within the enzymes were able to be visualised in near-atomic detail. Perhaps the most remarkable feature to be observed in all of these structures was that the proton-translocating subunits contained helical hairpins that were tilted almost completely horizontally with respect to the membrane. As a result of the availability of this new structural information, an updated model of proton translocation is now emerging. While the proton-translocating subunits from each rotary ATPase type have different numbers of transmembrane helices (F-type subunit a has 5-6, V-type subunit a has 8 and A-type subunit I has 8), their overall architecture is very similar, particularly the long tilted helices that are in contact with the membrane rotor, suggesting these features that are conserved among the different proton-translocating subunits must somehow facilitate proton translocation.

Analysis of these structures has led to the proposal that the interactions between the longest of the tilted proton-translocating subunit helices and the other transmembrane subunits form both of the two half-channels: in the *Polytomella* enzyme (Figure 5-2A), the larger helical hairpin (dark blue) can be seen to span the width of the membrane rotor (yellow). It has been suggested that the cavity between the turn in the helix and the rotor forms the luminal half-channel (it is lined with polar and charged residues), where H^+ first enter the enzyme. The second half-channel would then be formed by the ends of the same helical hairpin, which line an aqueous cavity that leads to the mitochondrial matrix. Based on the likely chemical environments of the helical hairpins, the authors were able to dock the subunit a protein sequence into the helices, and it was found that the conserved Arg239 residue would separate the two half-channels, as previously predicted.[149]

In the *S. cerevisiae* structures, the two long α-helices (Figure 5-2B), red arrows pointing to the green subunit a are in contact with three α-helices of the rotor ring (purple

helices). It has been proposed that two of the three rotor ring helices contain the glutamate residues that are protonated and de-protonated during ATP hydrolysis, with one of the glutamates being exposed to the cytoplasm and one being exposed to the organelle matrix (the authors do not distinguish which glutamate is which), and as the two glutamates are separated by the conserved arginine residue, this would create the two half-channels. The observation of the horizontal hairpins in these structures also enabled the 9.7 Å *T. thermophilus* cryo-EM map, which had been published by the same group, to be re-interpreted. They found that what had previously been interpreted as four vertical helices, due to the low resolution, could in fact be interpreted as two long horizontal helices, as have been seen in all of these structures. So it is likely that this feature is also found in subunit I.

In the bovine mitochondrial structures, the two long α-helices (Figure 5-2C), rainbow helices coloured from red to green) form a hairpin that spans both the rotor-ring (not shown) and the peripheral stalk b subunit (red helical segments). It is likely that the matrix half channel is formed by the cavity between the rotor ring and the matrix ends of the helical hairpin (in the region labelled "C term") and the inter-membrane space half-channel is probably formed by the cavity between the loop of the helical hairpin and the b subunit helices (in the region labelled "IM space"). While the authors concede that a higher resolution structure is required to confirm this, as there is no direct evidence that these proposed channels enclose aqueous environments, the authors were able to perform an evolutionary covariance analysis on the sequence of subunit a to trace the sequence through the cryo-EM density map. Covariance analysis involves identifying residues in evolutionarily related sequences that mutate as pairs, either to maintain a structurally important interaction or to physically accommodate one another. This strongly suggests that these residues are in physical proximity to one another within the tertiary structure of the protein, which enables a set of spatial restraints for the protein to be generated. The model of subunit a that was generated from these covariance restraints was fitted into the bovine F-type ATP synthase cryo-EM density, and following molecular dynamics flexible fitting calculations to improve the overall fit of the helices in the map, the final subunit a model within the complex satisfied 94% of the covariance restraints.

In the *P. denitrificans* crystal structure, subunit a has two tilted α-helical hairpins that form a bundle (Figure 5-2C) similar to the arrangement seen in the *Polytomella* structure. As the resolution was 4 Å, the side chain electron density was not sufficient to model the subunit a sequence, but the authors were able to use crosslinking and biochemical data to assign the helices and the position of functionally important residues, such as polar and charged residues known to be accessible to the cytoplasm or periplasm, and the conserved Arg182. It was inferred that, as in the other structures, this residue was in the subunit a helix (labelled aH5) closest to the membrane rotor ring, as this places it close to the essential Glu60 in the closest c subunit. The authors propose that the positioning of the polar and charged residues in both aH5 and aH6 on either side of Arg182 defines what would be the two half-channels. Nevertheless, a higher resolution structure is required to unequivocally assign the sequence of the a subunit, and therefore the position of the half-channels cannot be accurately determined until the position of each residue in subunit a is known.

In these most recent *T. thermophilus* cryo-EM structures, the transmembrane domain of subunit I contains eight elongated α-helices, two of which form a tilted α-helical hairpin that contacts the c-ring (Figure 5-2E), confirming the re-interpretation of the 9.7 Å resolution cryo-EM density. The remaining six helices, all of which are slightly tilted except for helix 3, are distal to the c-ring. Using restraints generated by a covariance analysis of the subunit I sequence, the authors were able to fit the subunit I sequence into the cryo-EM density. The resulting model places the conserved Arg563 residue in the middle of helical hairpin, and notably, is in good agreement with the biochemical data discussed in section 1.5.2. The *S. cerevisiae* subunit a is predicted to have a similar overall fold to the *T. thermophilus* subunit I, and the authors generated a model of the *S. cerevisiae* subunit a by combining data from the cryo-EM density determined previously for the intact *S. cerevisiae* V-ATPase[146] and a covariance analysis of the subunit a sequence, with an assumed identical topology to that of subunit I. From this the authors identify specific residues, all of which are found in the helices of the hairpin, which would most likely form the two half-channels for H^+ translocation, and which has not been done previously. As these residues are conserved in all proton-translocating subunits, they will be discussed here in terms of the A-ATPase subunit I. The cytoplasmic half-channel is proposed to be formed by residues E550 and T553, and the periplasmic half-channel is formed by residues D365, E426 and H616 (H616

correspond to the essential H245 in *E. coli* subunit a, although there is no equivalent residue to Glu219). It was suggested that H616 would facilitate stabilisation of the deprotonated conserved acidic residue of the rotor ring L-subunit as it enters the periplasmic half-channel, and then E550 and T553 would facilitate deprotonation of the acidic residue in the cytoplasmic half-channel. As these residues are conserved between all of the proton-translocating subunits of all rotary ATPase subtypes, it is likely that the proposed mechanism of H^+ translocation is shared by all of the subtypes.

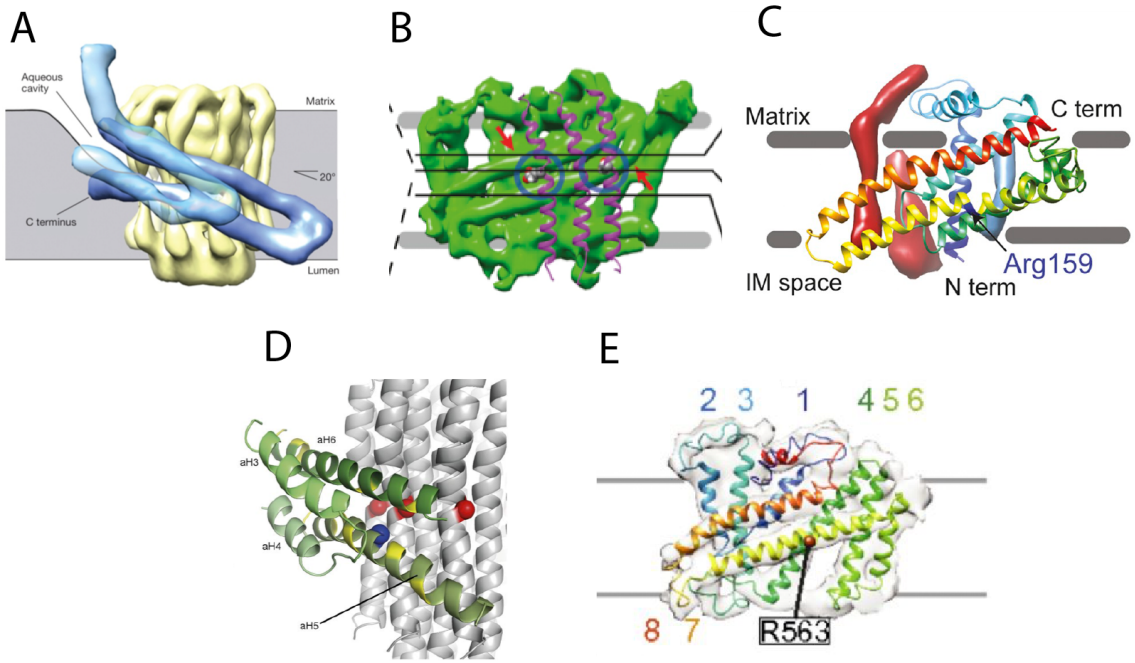


Figure 5-2: Horizontal membrane helices of the proton-translocating subunits in F-, V- and A-type rotary ATPases

A: Horizontal transmembrane helices of subunit a (blue) and their interaction with the membrane rotor ring (yellow) from the 6.2 Å cryo-EM structure of the *Polytomella sp.* F-type ATP synthase. Adapted from Allegretti et al. (2015) Figure 2.[145] **B**: Interaction of subunit a (green) with the membrane rotor ring (purple) from the 6.9-8.3 Å cryo-EM structures of the *S. cerevisiae* V-type ATPase. The horizontal helices of subunit a are indicated by red arrows. Adapted from Zhao et al. (2015) Figure 2.[146] **C**: Interaction of subunit a (red to blue, N- to C-terminus) with subunit b (red helical segments) and A6L (blue helical segments) from the 6.4-7.4 Å cryo-EM structures of the bovine F-type ATP synthase. Adapted from Zhou et al. (2015) Figure 2.[147] **D**: Interaction of subunit a (green) with the membrane rotor ring (grey) from the 4 Å X-ray crystal structure of the *P. denitrificans* F-type ATP synthase. Adapted from Morales-Rios et al. (2015) Figure 4.[125] **E**: Atomic model of transmembrane subunit I derived from the 6.4 Å cryo-EM structure of the intact *T. thermophilus* A-ATPase and the sequence covariance analysis (red to blue, N- to C-terminus) fitted into the cryo-EM density (grey envelope). Adapted from Schep et al. (2016) Figure 2.[148]

From all of these structures, it is highly likely that the two half channels necessary for H^+ translocation are formed by the interface between the long tilted α-helices of the proton-translocating subunit and the two or three closest α-helices of the rotor ring, and perhaps with the other transmembrane subunits. Yet while combining the covariance analysis data with medium-resolution density from cryo-EM or X-ray crystallography provides a very plausible and experimentally-supported model of the proton-trasnlocating subunit structure, individual residues of this subunit are yet to be experimentally visualised at the molecular level. While the *P. denitrificans* structure is the highest resolution structure available for any rotary ATPase/synthase, the authors acknowledge that crystallisation of the intact enzyme was not trivial, and many more medium-resolution rotary ATPase/synthase structures have been obtained using cryo-EM, as this overcomes the requirement to grow crystals. There have been a number of recent advances in cryo-EM technology that have enabled high-resolution structures to be determined. Some of these include the use of direct electron detection cameras, which overcome the need to convert electrons to photons, and the collection of data as movies due to the fast read out of the direct electron detectors, reducing blurring caused by electron beam-induced sample movement and instability of the sample mounting stage.[150,151] There have also been important advances in data collection and processing methods, allowing very large data sets to be collected and to be processed as multiple 3D classes that represent different conformations of the protein that have been captured during the freezing process,[152,153] as was seen with the *S. cerevisiae* V-ATPase, bovine mitochondrial F-ATPase and *T. thermophilus* A-ATPase.

The combination of all of these advances means it is now possible to obtain high-resolution structures using cryo-EM, such as the 3.2 Å structure of the yeast mitochondrial large ribosomal subunit[154], the 2.65-2.9 Å structures of the *E. coli* 70S ribosome-EF-Tu complex[155] and the 2.2 Å structure of β-galactosidase.[156] While a resolution greater than 6.2 Å is yet to be achieved for a rotary ATPase structure, with improved sample preparation and the possible discovery of organisms naturally more amenable to yielding high-resolution structures, it is very likely that a rotary ATPase/synthase structure of sufficient resolution to visualise sidechains will be obtained within the next few years. While *in silico* identification of thermostabilising mutants would likely be the most successful approach if crystallising subunit I in

isolation were still the major aim, it is now apparent that perhaps pursuing a higher resolution structure of an intact A-ATPase, perhaps from *M. ruber* or from another organism, would be more likely to yield a high-resolution structure of subunit I. This would also have the additional benefit of visualising subunit I in the context of the intact enzyme, which would provide details of how subunit I interacts with the neighbouring subunits within the complex.

5.4 The subunit I: peripheral stalks interaction and implications for the assembly of the A-ATPase complex

In the native A-ATPase complex there are two peripheral stalks heterodimers bound to the soluble domain of subunit I (Figure 5-3B). Yet in both of the experiments performed to test the subunit I:peripheral stalk interaction quantitatively, only one peripheral stalk heterodimer was bound to subunit I, even though the concentration of the peripheral stalks exceeded the concentration of subunit I by four- (MALLS) or tenfold (ITC). This indicates that under the conditions tested, there must be some difference between the two peripheral stalks binding sites, which could account for only one of the sites being occupied. As was briefly discussed in Section 4.5, this phenomenon has been observed previously. The first study characterising the interaction between the peripheral stalks (subunits E and G) and subunit C in *S. cerevisiae* found through a combination of sedimentation velocity ultracentrifugation and diffusion coefficient-calibrated size exclusion chromatography that these subunits only ever formed a 1:1:1 complex.[123] In these experiments, the protein concentrations were 2-40 μM, lower than the concentration used in ITC experiment with the *M. ruber* peripheral stalks and subunit I. In another study that was discussed in 4.5, ITC experiments with the *S. cerevisiae* peripheral stalks and subunit C demonstrated that only one EG heterodimer bound to subunit C, although the authors did not give the protein concentrations used in the experiments.[122] The interaction has also been tested with the peripheral stalks and soluble domain of subunit I from the bacterial *E. hirae* V-ATPase.[121] As with the yeast subunits, the authors found through a combination of surface plasmon resonance, analytical ultracentrifugation and small angle X-ray scattering that only one peripheral stalk heterodimer bound to the soluble domain of subunit a. Therefore, the findings from the MALLS and ITC experiments in this study are consistent with previous

results. Furthermore, these experiments also show that using full-length subunit I rather than only the soluble domain does not lead to a second peripheral stalk heterodimer binding event. In all of these studies, the authors acknowledge that binding of a second EG heterodimer to the other binding site present on subunit C/subunit I may require the rest of the ATPase complex to be assembled, as this may stabilise the interaction of a second EG heterodimer with the second binding site.

From the X-ray crystal structure of the *M. ruber* subunit I soluble domain (I_{sol}),[70] the two lobes of this domain that contain the two peripheral stalk binding sites are structurally distinct (aligning the two domains in Coot using least squares fitting gives an RMSD of 4.0 Å), and therefore most likely interact with the peripheral stalks heterodimer in different ways. It is also likely that the two binding sites on subunit I have different affinities for the peripheral stalks, and that *in vitro*, binding only ever occurs at the high affinity site. In the yeast V-ATPase, this high affinity site has been identified as the "head" domain (Figure 4-13)[122]. The I_{sol} crystal structure was aligned with the 2.8 Å X-ray crystal structure of the yeast V-ATPase C_{head} domain bound to an EG heterodimer[63] in Coot[157] using the secondary structure matching calculation (Figure 5-3).[158] The calculation automatically aligned C_{head} with the proximal domain of I_{sol} with a root mean square deviation (RMSD) of 2.05 Å, indicating these are the most similar parts of the two structures. When the distal lobe of I_{sol} was aligned with C_{head} using least squares fitting the RMSD was 4.5 Å, thus these two domains are not structurally similar. The proximal lobe of I_{sol} therefore probably contains the high affinity peripheral stalks binding site. It should also be noted that when the I_{sol} X-ray crystal structure is fitted into the 9.7 Å *T. thermophilus* A-ATPase cryo-EM density, the conformation is very different to when the I_{sol} structure is aligned to the C_{head} X-ray crystal structure, and this is most likely because the helical bundle domain between the two lobes is very different in the two proteins. This alignment was therefore only used to predict which lobe of I_{sol} represents the high affinity binding site, not the conformation when it has bound to the EG heterodimer. The difference in affinity between these two sites, and the likely requirement for the intact A-ATPase to be assembled for the second peripheral stalk to bind, may provide insights into the assembly of this complex.

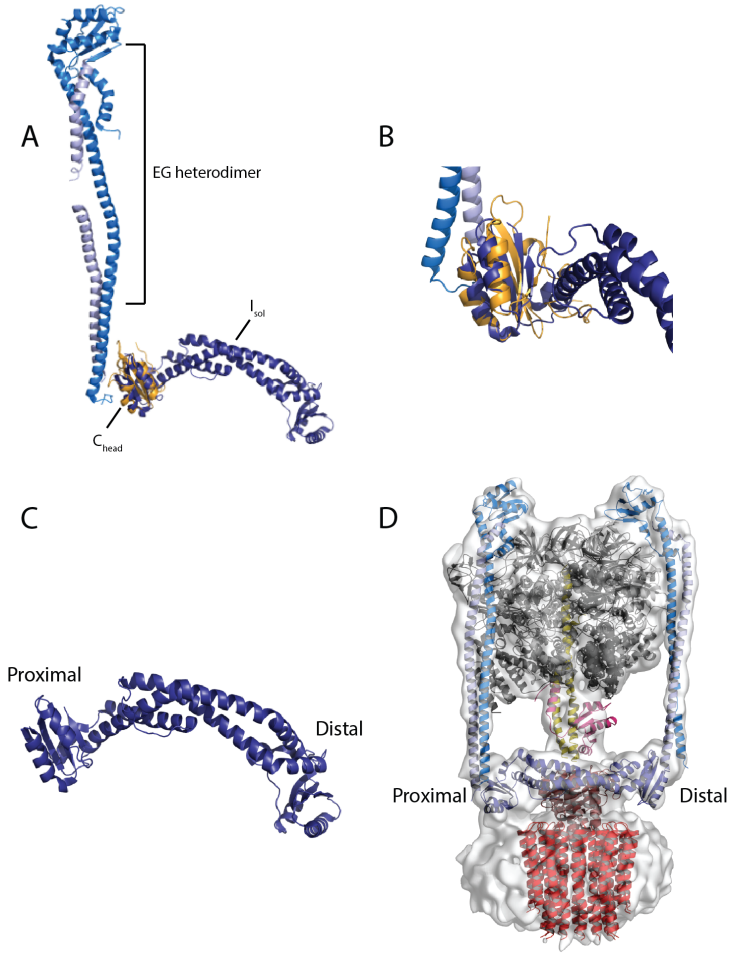


Figure 5-3: The proximal lobes of I_{sol} corresponds to the C_{head} EG heterodimer binding site of the V-ATPase.

A: Alignment of the 2.6 Å X-ray crystal structure of the soluble domain of subunit I from *M. ruber* (dark blue; PDB: 3rrk)[70] with the 2.8 Å X-ray crystal structure of the C_{head} domain (orange) and the EG heterodimer (marine and pale blue, respectively) of the *S. cerevisiae* V-type ATPase (PDD: 4efa).[63] The two structures were aligned in Coot using the secondary structure matching calculation with an RMSD of 2.05 Å. **B**: Zoomed-in view of the alignment of I_{sol} with the C_{head} domain. Colours are the same as for A. **C**: I_{sol} with proximal and distal sites labelled. These are the binding sites for each of the EG heterodimers present in the A-ATPase complex. **D**: The proximal and distal binding sites of I_{sol} shown in the context of the intact *T. thermophilus* A-ATPase. Both sites are bound to a peripheral stalks heterodimer. Adapted from Stewart et al. (2014) Figure 2.[53]

During assembly of the A-ATPase, it is possible that the head (subunits A and B) and central rotor (subunits D and F) of the soluble A_1 sub-complex assemble in the cytoplasm where they were expressed, and the stator (subunit I) and rotor (subunit C and the L-ring) come together at the periplasmic membrane where they have been co-translationally folded (Figure 5-4A). A similar process has been seen for the *S. cerevisiae* F-ATPase and it has been termed modular assembly.[159] The dissociation

constant of the interaction between the A-ATPase peripheral stalks and the catalytic subunit B has been calculated to be ~2 mM for the *T. acidophilum* enzyme using NMR.[160] This is a much lower affinity than that of the EG heterodimer for the subunit I proximal binding site (calculated by ITC to be in the low μM range in 4.5), so it is likely an EG heterodimer binds to the high affinity binding site on subunit I, and this would recruit the head and rotor A_1 sub-complex (Figure 5-4B and C). The binding sites for the second EG heterodimer on subunit B and the subunit I would then be physically connected and stationary, which would greatly increase the likelihood that the second EG heterodimer would bind to these sites (Figure 5-4D). This would complete the assembly of the enzyme, yielding a functional rotary A-ATPase.

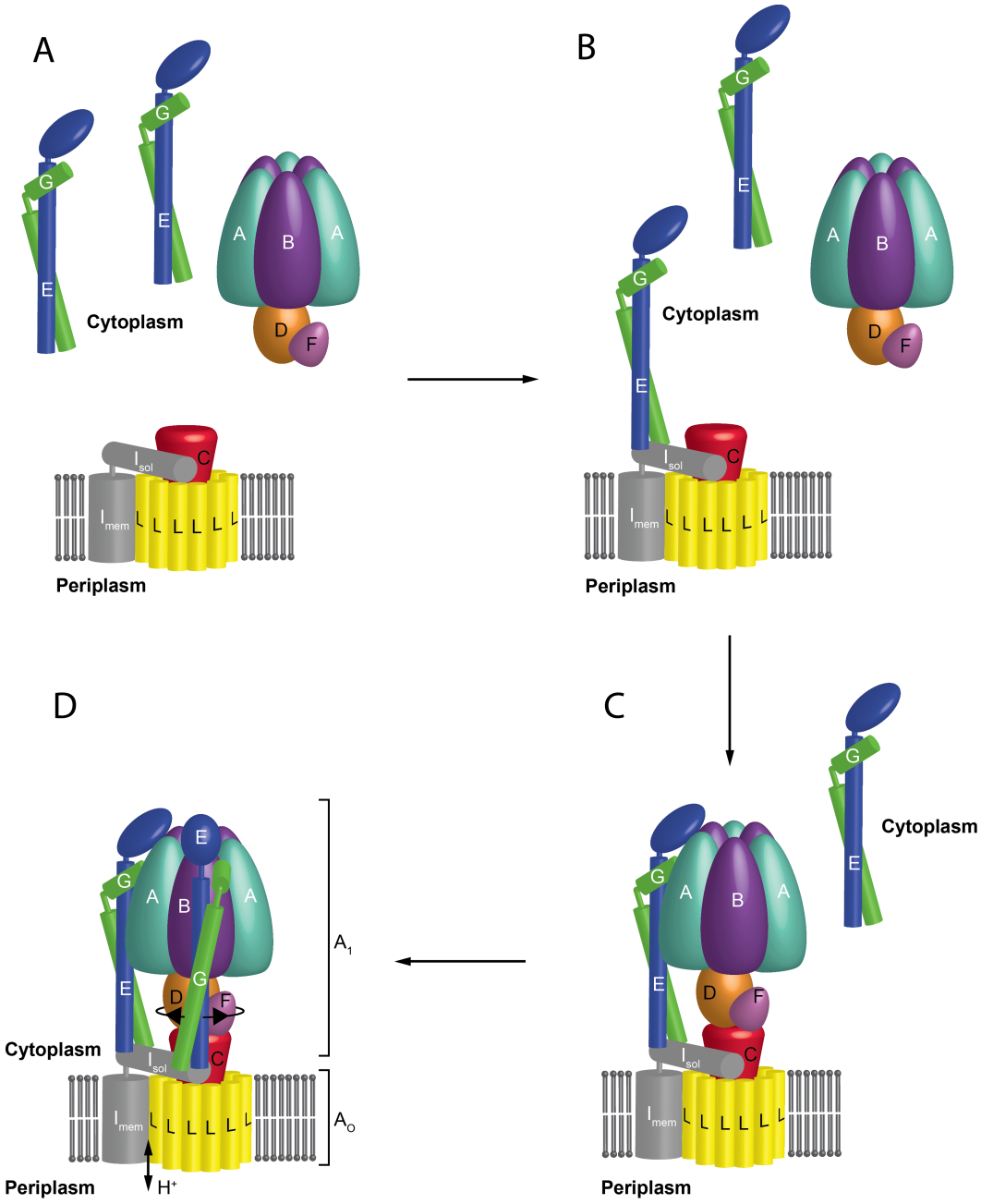


Figure 5-4: Possible assembly pathway of the rotary A-ATPase.

A: During assembly of the A-type ATPase complex, the catalytic head heterotrimer (subunits A and B) is likely associated with the central rotor (subunits D and F) in the cytoplasm, where they have been expressed. The affinity of the EG heterodimer for subunit B (~2 mM) is probably too low for this interaction to occur outside the complex, so the EG heterodimer would be separate from the other A_1 subunits in the cytoplasm. The A_O sub-complex would likely be assembled in the periplasmic membrane to stabilise subunit I. **B**: A peripheral stalks heterodimer binds to the high affinity binding site on the soluble domain of subunit I. **C**: An EG heterodimer engages the A_3B_3DF sub-complex, bringing the two binding sites for the second EG heterodimer together. **D**: The second EG heterodimer is able to bind, completing the assembly of the complex. The complex is now able to synthesis ATP or pump H^+ according to the organism's metabolic requirements. Adapted from Stewart et al. (2013) Figure 1.[161]

5.5 Concluding remarks

While the X-ray crystal structure of the rotary A-type ATPase was not determined during the course of this work, full-length subunit I from *M. ruber* was expressed and the purification was optimised. Furthermore, the purified subunit I was shown to be functional through binding studies with the peripheral stalks heterodimer from the same organism. The findings from studying this interaction were consistent with previous findings, namely that only one peripheral stalk heterodimer is able to bind the soluble domain of subunit I, or its equivalent, *in vitro*. The implications for this are that during assembly, one peripheral stalk heterodimer engages the high affinity binding site on subunit I and once the rest of the complex has been assembled, the second heterodimer can bind. It was noted in the other studies that investigated the interaction between subunit I/subunit C and the peripheral stalks that the individually expressed proteins were less stable in isolation than they were once they had formed the sub-complex.[121,123] This is especially the case for subunit I, which would be far more stable in the context of the native A-ATPase complex in the native lipid bilayer of *M. ruber* than when it is isolated from the complex and solubilised in detergent or amphipols. Therefore, the approach that will most likely yield a high-resolution model of the proton-translocating subunit would be to pursue a higher resolution structure of the intact rotary ATPase, either using cryo EM or X-ray crystallography.

5.6 Construction of expression plasmids

5.6.1 Selecting subunit I homologues

12 homologues of subunit I (Table 0-1) were selected from a position-specific iterated nucleotide basic alignment search (PSI-BLAST)[162] using the sequence of subunit I from *T. thermophilus* as the input. All sequences that were subsequently cloned were identified within two iterations of the search procedure. The genes for each subunit I homologue were obtained from genomic DNA from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ), vectors containing the ATPase operon from the corresponding organism that had been constructed previously or synthetic genes were ordered (GenScript).

Table 0-1: Sources of subunit I genes

Organism	Subunit I gene ID	Source
Sulfolobus solfataricus	SSO0559	ATPase operon vector
Pyrococcus furiosus	PF0177	ATPase operon vector
Methanosarcina mazei	MM_0785	ATPase operon vector
Thermoplasma acidophilum	Ta0008	ATPase operon vector
Archaeoglobus fulgidus	AF_1159	ATPase operon vector
Methanocaldococcus jannaschii	MJ_0222	ATPase operon vector
Pyrococcus horikoshii	PH1981	DSMZ
Methanocaldococcus maripaludis	MmarC6_1620	ATPase operon vector
Thermus thermophilus	TTHA1278	DSMZ
Meiothermus ruber	Mrub_0956	GenScript
Thermococcus gammatolerans	TGAM_RS00720	GenScript
Deinococcus radiodurans	DR_0695	ATPase operon vector

5.6.2 Polymerase chain reaction (PCR)

If the vectors were to be constructed by a sequence and ligation independent cloning (SLIC) reaction, the primers were designed to add 15 bp extensions at the 5' and 3' ends of the PCR product that were complementary to the target vector. The 5X LongAmp™ *Taq* Reaction Buffer (NEB), 300 µM dNTPs (NEB), 0.4 µM forward and reverse primers (Sigma Aldrich), 300-500 ng template DNA, 5 U LongAmp *Taq* DNA polymerase (NEB) and 3% (v/v) dimethyl sulfoxide (DMSO) were combined and made up to 50 µL with MilliQ water. The desired products were amplified by 30 cycles of PCR. The reactions were run on a 1.5% (w/v) gel with 0.005% (v/v) ethidium bromide to confirm the correct product had been amplified. Correctly amplified PCR products were purified using the MinElute PCR Purification Kit (Qiagen).

5.6.3 SLIC

The purified PCR products were cloned into each of the eight *EcoRI*-linearised SLIC vectors (2:1 molar ratio; Table 0-2) in the presence of the InFusion HD EcoDry^{TM} reagent (Clontech). The reaction was incubated for 15 min at 37 ºC, followed by 15 min at 50 ºC. The reaction mixture was diluted four-fold with TE buffer and 2.5 µL was transformed into NEB 5-alpha Competent *E. coli* cells (NEB) by heat shock at 42 ºC for 30 seconds. The cells were recovered for one hour at 37 ºC in Super Optimal Broth, then plated onto LB agar plates with 34 µg mL^{-1} chloramphenicol. The plates were incubated at 37 ºC overnight.

Table 0-2: SLIC vectors

Vector	N-terminal tag	C-terminal tag
A	His_{10}*	FLAG
B	FLAG	His_{10}
C	T7	His_{10}
D	mistic[108]	His_{10}
E	GST	His_{10}
F	MBP	ThioredoxinA (TrxA)/His_{10}
G	His_{10}	MBP
H	His_{10}	GFP

* His_{10} = decahistidine tag

5.6.4 Gibson Assembly®

A gBlock gene fragment (IDT) containing the genes encoding *M. ruber* subunits C/G (gene ID SY28_13195) and E (gene ID Mrub_0958) separated by a non-coding region, designed to transcribe both genes independently, was cloned into the vector pLICHK by the Gibson Assembly method[163]. This vector encodes a hexahistidine tag (His_6) and a Tobacco etch virus protease (TEV) recognition sequence at the N-terminus of the insert, thus subunit C was encoded as a His_6-fusion protein. The lyophilised gBlock was resuspended in Milli Q water to give a final concentration of 10 ng/µL. The Gibson Assembly reaction was then set up by mixing the gBlock (10 ng) with the *SspI*–digested pLICHK vector (50 ng) and the Gibson cloning mix (isothermal buffer (100 mM Tris-Cl; pH 7.5, 10 µM $MgCl_2$, 200 µM dNTPs, 10 µM dithiothreitol, 5% (w/v) PEG 8000, 1 mM NAD+), 0.04 U T5 exonuclease, 0.25 U Phusion polymerase, 37.5 U *Taq* DNA ligase). The reaction was incubated at 50 ºC for 30 min, and was subsequently transformed into DH5α cells as described in 5.6.3. The transformed cells were plated onto LB agar plates containing 50 µg mL^{-1} kanamycin and were incubated at 37 ºC overnight.

5.6.5 Colony PCR

Single colonies that grew on transformation plates were tested for correct gene insert by PCR. A colony was resuspended in 2 µL MilliQ water and stored on ice. 10X ThermoPol Reaction Buffer (NEB), 200 µM dNTPs, 0.2 µM forward (cloning primer) and reverse (DuetDOWN1) primers, 2 µL resuspended colony solution, 0.625 U Taq DNA Polymerase (NEB) and 3% (v/v) DMSO were combined and made up to 25 µL with MilliQ water. The desired product was amplified by 25 cycles of PCR. The reactions were visualised by running them on a 1.5% (w/v) agarose gel with 0.005% (v/v) ethidium bromide, and the colonies that contained vectors with the correct insert were used to inoculate 5 mL LB medium containing the corresponding antibiotic. The culture was incubated overnight at 37 ºC and the plasmid was purified using the QIAprep Spin Miniprep Kit (Qiagen).

5.7 Expression and purification

5.7.1 Small-scale expression trials

The purified plasmid for each construct was transformed into T7 Express Competent *E. coli* cells (High Efficiency) (NEB) using the same method as described in 5.6.3, and the cells were grown on LB agar plates with 34 µg mL^{-1} chloramphenicol overnight at 37 ºC. A scraping of cells was used to inoculate 3 mL LB medium with 34 µg mL^{-1} chloramphenicol and cultures were grown at 37 ºC with shaking at 180 rpm until the $OD_{600\ nm}$ reached ~0.7. Expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and the cultures were incubated overnight at 18 ºC with shaking at 180 rpm. The cells were harvested by centrifugation at 1681 × *g* then the pellets were resuspended in 180 µL of lysis buffer (20 mM Tris pH 8.0, 300 mM NaCl, 2% (v/v) Triton X-100 (Anatrace), cOmplete EDTA-free protease inhibitor cocktail tablet (Roche), lysozyme (Sigma-Aldrich) and DNaseI (Roche)). The resuspended cells were incubated at 4 ºC for 30 min, then insoluble material was cleared by centrifugation at 21 000 × *g* for 20 min. The supernatant was applied to 50 µL EcoMag™His-Co+ Magnetic Particles (BioClone Inc), followed by incubation for 1 hour at 4 ºC. The unbound fraction was discarded and the resin was washed with 2 × 1 mL of wash buffer (20 mM Tris pH 8.0, 300 mM NaCl, 1% (v/v) Triton X-100). The

resin was then analysed by SDS-PAGE on a NuPAGE Novex® 4-12% Bis-Tris protein gel (Thermo Fisher Scientific), along with SeeBlue® Plus2 pre-stained protein standards (Thermo Fisher Scientific). Proteins were visualised by in-gel GFP fluorescence or Western blot against the affinity/solubility tag for that construct (see section 5.7.3).

5.7.2 Large scale expression

Subunit I constructs were transformed into *E. coli* T7 Express cells that were plated onto LB agar. Transformed colonies were used to inoculate 12 L LB medium containing 34 μg mL^{-1} chloramphenicol and the cultures were grown at 37 ºC until $OD_{600\ nm}$ was at ~0.7. Expression was then induced with 1 mM IPTG and the cultures were incubated for 14-20 hours at 18 ºC. Other incubation temperatures were not trialled. The cells were harvested by centrifugation at 5044 × *g*, washed with 1 × PBS, snap frozen in liquid N_2 and stored at -80 ºC.

5.7.3 Measuring GFP fluorescence

5.7.3.1 In-gel fluorescence

The presence of the intact subunit I-GFP fusion protein was detected by measuring in-gel GFP fluorescence of the SDS-PAGE using an FLA-5100 Fluorescent Image Analyzer (Fujifilm Life Science) with an excitation wavelength of 473 nm and a 510 nm emission filter.

5.7.3.2 Fluorescence-detection size exclusion chromatography

Aliquots of size exclusion fractions were transferred to a 96-well black plate and GFP fluorescence was measured by a PHERAstar *FS* microplate reader (BMG Labtech), using an excitation wavelength of 485 nm and a 520 nm emission filter. The absorbance was plotted against fraction number to generate a fluorescence-detection size exclusion chromatography profile (FSEC).

5.7.3.3 Western blot analysis

The presence of the remaining subunit I fusion proteins was analysed by Western blot using an anti-His (Sigma-Aldrich), anti-MBP (NEB), anti-FLAG (Sigma-Aldrich), anti-T7 (Novagen) or an anti-GST (Novagen) antibody, all of which were monoclonal antibodies conjugated to horse radish peroxidase (HRP). After SDS-PAGE the proteins were transferred to a nitrocellulose membrane using the iBlot™ Gel Transfer Device (ThermoFisher Scientific), which was blocked with blocking solution (5% (w/v) skim milk powder in TBS-Tween) for 1 hour at room temperature. The membrane was then incubated with the corresponding antibody (1:6000 dilution in TBS-Tween) at room temperature for one hour. The membrane was washed with TBS-Tween for 3 × 5 min, then the membrane was developed by incubating with Western Lightning Plus enhanced chemiluminescence (ECL) HRP substrate (PerkinElmer) and exposed to Amersham Hyperfilm ECL X-ray film (GE Healthcare Life Sciences).

5.7.4 Detergent screen

3 mL cultures of the His_{10}-*M. ruber* subunit I-FLAG (His_{10}-I-FLAG) construct transformed into T7 Express cells were grown as described in 5.7.1. The cells pellets from each culture were resuspended in 180 µL lysis buffer (5.7.1) with 2% (w/v) of either DDM, DM, OG, lauryl maltose neopentyl glycol, decyl maltose neopentyl glycol, Anzergent 3-12, Anzergent 3-14 or octaethylene glycol monododecyl ether (C12E8). The resuspended cells were incubated at 4 ºC for ~30 min, then centrifuged at 21000 × *g* for 20 min to remove debris. The supernatant was applied to 50 µL EcoMag™His-Co+ Magnetic Particles (BioClone Inc) and incubated for 1 hour at 4 ºC. The unbound fraction was discarded, then the resin was washed with 2 × 1 mL of wash buffer (20 mM Tris pH 8.0, 300 mM NaCl, 1% (w/v) of the corresponding detergent). The resin was then analysed for bound proteins by SDS-PAGE followed by anti-His Western blot as described in 5.7.3.

5.7.5 *E. coli* membrane isolation

Cell pellets from a 12 L *E. coli* overexpression culture were resuspended in 100 mL lysis buffer (20 mM Tris, pH 8.0, 500 mM NaCl, 10% (v/v) glycerol, 5 mM $MgCl_2$, 2 ×

protease inhibitor cocktail tablets, spatula tip of lysozyme and DNase I) and incubated at 4 ºC for 15 min. Cell membranes were disrupted with a Digital Sonifier® Cell Disruptor (Branson) at 70% amplitude for 2 min. Debris was removed by centrifugation at 8030 × *g* for 10 min. The supernatant was centrifuged at 100 000 × *g* for 1 hour to pellet the membranes. The membrane pellets were resuspended in 50 mL wash buffer (20 mM Tris, pH 8.0, 300 mM NaCl, 10% (v/v) glycerol, 5 mM $MgCl_2$) and centrifuged at 100 000 × *g* for 30 min. This wash step was repeated once more, then the membranes were snap frozen in liquid N_2 and stored at -80 ºC until further use.

5.7.6 Full-length subunit I and transmembrane domain purification

Frozen cell pellets or membranes were thawed, resuspended in 100 mL lysis buffer (20 mM Tris pH 8.0, 300 mM NaCl, 10% (v/v) glycerol, 100 mM sucrose, 1.3% (w/v) detergent, ± 2 mM ATP, 20 mM imidazole, two cOmplete protease inhibitor cocktail tablets, spatula tips of lysozyme and DNaseI) and incubated for 30 min at 4 ºC to allow incorporation of the protein into the detergent micelles. Lysis was completed by sonication at 60% amplitude using a pulse program of 1 sec on/1 sec off for 1 min 30 sec in 2 × 70 mL batches. The lysate was cleared by centrifugation at 30 000 × g for 40 min. The supernatant was applied to 3 mL cobalt® Metal Affinity Resin (Clonetech Laboratories Inc) pre-equilibrated with 8 column volumes (CV) of wash buffer (20 mM Tris pH 8.0, 300 mM NaCl, 10% (v/v) glycerol, 100 mM sucrose, 0.5% (w/v) detergent, 20 mM imidazole, ± 2 mM ATP) and was incubated with the resin for 1 hour at 4 ºC. The flow-through was discarded, the resin was washed with the remaining wash buffer and bound proteins were eluted with 4 × 3 mL elution buffer (wash buffer with 300 mM imidazole). The elutions were analysed by SDS-PAGE with SimplyBlue SafeStain (Thermo Fisher Scientific), and those containing subunit I were pooled and concentrated to 1 mL using a Vivaspin 20 100 000 molecular weight cut-off (MWCO) centrifugal concentrator. The 1 mL sample was centrifuged at 21 000 × *g* for 5 min to remove large aggregates, and was loaded onto a HiLoad 16/600 Superdex 200 pg (S200; GE Healthcare Life Sciences) size exclusion column (SEC) pre-equilibrated with SEC buffer (20 mM Tris pH 8.0, 300 mM NaCl, 10% (v/v) glycerol, 100 mM sucrose, 0.17% (w/v) detergent, ± 2 mM ATP). The fractions corresponding to the peak in the $A_{280\ nm}$ elution profile were analysed by SDS-PAGE with SafeStain staining, pooled and

concentrated to 500 µL using a Vivaspin 2 100 000 MWCO concentrator. The concentration was determined by measuring absorbance at 280 nm using the NanoDrop spectrophotometer (Thermo Fisher Scientific) then the protein was snap frozen in 50 µL aliquots in liquid N_2 and stored at -80 ºC.

5.7.7 Amphipol A8-35 exchange

Peak fractions containing subunit I from the S200 column were pooled and the concentration was measured using absorbance at 280 nm. 5 mg Amphipol A8-35 per mg of protein was added to the pooled elutions, and protein:amphipol mixture was incubated for 2 hours at 4 ºC. Excess detergent was removed by adding 20 µL Bio-Beads™ SM-2 resin per µg of detergent in the sample and incubating for 2 hours at 4 ºC. The resin was collected in a gravity column, and the flow-through was concentrated to 500 µL using a Vivaspin 20 50 000 MWCO concentrator. The sample was applied to an S200 column pre-equilibrated with detergent-free size exclusion buffer.

5.7.8 Expression and purification of *M. ruber* peripheral stalks

The pLICHK-EG vector was transformed into *E. coli* T7 Express cells that were plated onto LB agar containing 50 µg mL^{-1} kanamycin. The plate was incubated overnight at 37 ºC. Transformed colonies were used to inoculate 12 L LB medium with 50 µg mL^{-1} and the cultures were incubated at 37 ºC with shaking at 150 rpm until $OD_{600\ nm}$ reached ~0.7. Expression was induced with the addition of 1 mM IPTG. Cultures were incubated for 14-16 hours at 18 ºC. Cells were harvested by centrifugation at 5044 × *g* and the pellet was resuspended in 70 mL lysis buffer (20 mM Tris, pH 8.0, 300 mM NaCl, 20 mM imidazole, 2 × protease inhibitor cocktail tablets, spatula tips of lysozyme and DNase I). The resuspended cells were incubated at 4 ºC for 15 min, then were disrupted with two freeze-thaw cycles. Lysis was completed by sonication at 70% amplitude for 3 min. Insoluble material was cleared by centrifugation at 30 000 × *g* and the supernatant was bound to 3 mL Ni-NTA resin pre-equilibrated with Ni binding buffer (20 mM Tris, pH 8.0, 300 mM NaCl, 20 mM imidazole) for 1 hour at 4 ºC. The flow-through was discarded and the resin was washed with the remaining wash buffer.

Bound proteins were eluted with 3 × 3 mL Ni elution buffer (wash buffer with 300 mM imidazole) and analysed by SDS-PAGE with SafeStain staining. Fractions containing the peripheral stalks were pooled and 800 µg TEV was added. The cleavage reaction was incubated overnight at 4 ºC.

The cleavage reaction was centrifuged at 5000 × *g* for 20 min to remove any precipitated protein. Ni-NTA resin (1 mL) was added to the supernatant to remove the TEV, with incubation at 4 ºC for 30 min. The flow-through was collected and concentrated to 500 µL using a Vivaspin 20 10 000 MWCO concentrator. The sample was applied to an S200 size exclusion column pre-equilibrated with size exclusion buffer (20 mM Tris, pH 8.0, 150 mM NaCl). Protein elution was monitored by measuring the absorbance at 280 nm, and the peak fractions were analysed by SDS-PAGE with SafeStain staining. Fractions containing both peripheral stalks were pooled, concentrated and snap frozen in aliquots in liquid N_2. The aliquots were stored at -80 ºC until further use.

5.8 Crystallisation

5.8.1 Vapour diffusion crystallisation trials

Vapour diffusion crystallisation trials were conducted in 96-well sitting drop plates with 17 commercially available crystallisation screens, each of which contained 96 different combinations of crystallisation reagents (Table 6-3). Purified protein at the desired concentration was mixed with the crystallisation condition (200 nL + 200 nL) using a mosquito® Crystal liquid-handling robot (TTP Labtech) and the mixtures were dispensed onto MRC 2 well crystallisation plate (96-well sitting drop plate; Swissci). The plates were sealed with UV-transparent ClearSeal™ film and incubated at the desired temperature (20 ºC or 4 ºC). Drops were imaged using either a UVEX-p UV fluorescence and absorbance microscope (JANSi) or an Olympus SZ61 light microscope (Olympus).

Table 0-3: Commercial crystallisation screens

Screen name	Manufacturer	Screen type
Crystal Screen HT	Hampton Research	Sparse matrix
MembFac HT	Hampton Research	Sparse matrix
Index HT	Hampton Research	Sparse matrix and grid
SaltRX HT	Hampton Research	Sparse matrix
Wizard Classic 1&2	Rigaku Reagents	Sparse matrix
Wizard Cryo 1&2	Rigaku Reagents	Sparse matrix
JBScreen Classic HTS I	Jena Bioscience	Sparse matrix and grid
JBScreen Classic HTS II	Jena Bioscience	Sparse matrix and grid
Clear Strategy Screen I HT	Molecular Dimensions	Sparse matrix and grid
Clear Strategy Screen II HT	Molecular Dimensions	Sparse matrix and grid
PACT premier™ HT-96	Molecular Dimensions	Grid
MIDAS™ HT-96	Molecular Dimensions	Sparse matrix and grid
Morpheus® HT-96 Green Screen	Molecular Dimensions	Sparse matrix
JCSG-plus HT-96	Molecular Dimensions	Sparse matrix
MemGold™ HT-96	Molecular Dimensions	Sparse matrix
MemPlus™ HT-96	Molecular Dimensions	Sparse matrix
MemStart™ + MemSys™ HT-96	Molecular Dimensions	Sparse matrix

5.8.2 Lipidic cubic phase (LCP) crystallisation trials

The LCP trials were set up at the CSIRO Collaborative Crystallisation Centre (www.csiro.au/C3), Melbourne, Australia, with the assistance of David Aragão. Molten 1-oleoyl-rac-glycerol (monoolein, ~35 mg, Sigma) was added to the barrel of a pre-weighed 100 µL gastight syringe (Hamilton). The syringe was weighed to determine the exact mass of the lipid added to the syringe, then the subunit I sample was added to a second 100 µL gastight syringe to a volume that would be 40% of the mass

of the total protein:monoolein mixture. The two syringes were connected with a coupler (Formulatrix) and the protein:monoolein mixture was transferred between the two syringes until the mixture was homogenous and transparent, which indicated the LCP had formed. The LCP was then dispensed onto a 96-well glass sandwich plate (Hampton Research) in 100 nL drops using a mosquito® LCP (TTP Labtech). 1 µL of the crystallisation condition was then dispensed onto each LCP drop. The drops were sealed by mounting a second glass sandwich plate on top of the first, and the LCP trials were incubated at 20 ºC.

Four screens were used in the LCP trials: two commercial screens, the Wizard Cubic LCP Screen (Rigaku Reagents) and MemGold, and two custom screens for LCP designed in Vadim Cherezov's laboratory, both of which contain PEG 400 in all conditions along with different buffers and salts.[164] Each of the four screens was set up with subunit I at a concentration of 10 mg mL^{-1}, purified in the presence or absence of 2 mM ATP. Drops were imaged using an Olympus SZ61 light microscope with an SZ-PO polarising filter.

5.9 Nickel-NTA pulldown assay

Purified subunit I (400 µg) was thawed and centrifuged at 21 000 × g to remove large aggregates. The sample was diluted to 1 mL with binding buffer (20 mM Tris pH 8.0, 300 mM NaCl, 10% (v/v) glycerol, 100 mM sucrose, 20 mM imidazole, 8.4 mM Anzergent 3-12 (this subunit I sample had been purified in Anzergent 3-12)) and immobilised on 250 µL Ni-NTA resin for 1 hour at 4 ºC. The unbound fraction was discarded, and the resin was washed with 3 × 1 mL binding buffer. The resin with immobilised subunit I was then incubated with the untagged and purified peripheral stalks single heterodimer (700 µg) for 1 hour at 4 ºC. The unbound fraction was discarded and the resin was washed with 50 mL binding buffer. Bound proteins were eluted from the resin with 3 × 500 µL elution buffer (binding buffer with 300 mM imidazole), and the elutions were analysed along with the input samples, flow-through and resin by SDS-PAGE using SafeStain staining.

5.10 Biophysical techniques

5.10.1 Multi-angle laser light scattering (MALLS) analysis

5.10.1.1 Individual proteins

Subunit I and peripheral stalks samples were applied to an S200 or superose 6 column pre-equilibrated with SEC buffer (2.2.5) and eluting fractions were analysed by inline MALLS and refractive index (RI) measurements. MALLS intensity was measured using a miniDawn™ TREOS triple-angle light scattering detector (Wyatt Technology), with detectors at 43.6º, 90.0º and 136.4º. Protein elution was monitored by absorbance at 280 nm using an SPD-20A UV/VIS detector (Shimadzu Scientific Instruments), and protein concentrations were obtained from inline RI measurements using an RID-10A differential refractive index detector (Shimadzu Scientific Instruments). A conjugate analysis was performed to calculate the proportions of subunit I and DM in the protein:detergent micelle using a refractive index increment (dn/dc) of 0.18 mL g^{-1} for the protein component and a dn/dc of 0.14 mL g^{-1} for the DM component.[165] The weight-average molecular masses of the protein and detergent components were calculated using the ASTRA software (Wyatt Technology).

5.10.1.2 Subunit I:peripheral stalk complex

Samples of subunit I and peripheral stalks prepared for this experiment were purified in the absence of glycerol so the concentration of each sample could be measured by a DirectDetect® Infrared Spectrophotometer (Merck Millipore), as this is a more accurate method than using the absorbance at 280 nm. The subunit I:peripheral stalk complex was prepared by mixing subunit I (500 µg) and peripheral stalks (1.8 mg) samples that had been purified independently, and incubating this mixture for four hours at 4 ºC. The sample was centrifuged at 21 000 × g for 5 min to remove large aggregates, then the sample was injected onto a superpose 6 column pre-equilibrated with MALLS buffer (20 mM Tris pH 8.0, 300 mM NaCl, 2 mM ATP, 5 mM $MgCl_2$, 3.2 mM DM) and analysed by MALLS as described above. The $E_{1\%}$ used for the MALLS analysis of the subunit I:peripheral stalk complex was 1.00 $mg^{-1}mL\ cm^{-1}$ (1:1 complex). Both proteins were also analysed individually as controls.

5.10.2 Isothermal titration calorimetry (ITC)

ITC data was recorded using a MicroCal iTC200 (Malvern) at 16 ºC using a reference power of 2 µcal/s and a stirring speed of 1000 rpm. Purified peripheral stalks (100 µM) that had been buffer-exchanged into the subunit I amphipol SEC buffer (3.4) were titrated into purified subunit I (10 µM) in 20 steps of 1 µL every 400 sec. As a control, the experiment was repeated with the peripheral stalks injected into buffer rather than subunit I. The data were processed using the Origin 7.0 software package (OriginLab) using the "one set of sites" model.

Chapter Six: References

1. Haldane, J.B.S. The origin of life. *The Rationalist Annual* **3**, 3-10 (1929).
2. Lane, N., Allen, J.F. & Martin, W. How did LUCA make a living? Chemiosmosis in the origin of life. *Bioessays* **32**, 271-80 (2010).
3. Corliss, J.B. et al. Submarine thermal sprirngs on the galapagos rift. *Science* **203**, 1073-83 (1979).
4. Kelley, D.S. et al. An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30 degrees N. *Nature* **412**, 145-9 (2001).
5. Sleep, N.H., Meibom, A., Fridriksson, T., Coleman, R.G. & Bird, D.K. H2-rich fluids from serpentinization: geochemical and biotic implications. *Proc Natl Acad Sci U S A* **101**, 12818-23 (2004).
6. Proskurowski, G. et al. Abiogenic hydrocarbon production at lost city hydrothermal field. *Science* **319**, 604-7 (2008).
7. Russell, M.J., Daniel, R.M., Hall, A.J., Sherringham, J.A. A hydrothermally precipitated catalytic iron sulphide membrane as a first step toward life. *Journal of Molecular Evolution* **39**, 231-243 (1994).
8. Wachtershauser, G. On the chemistry and evolution of the pioneer organism. *Chem Biodivers* **4**, 584-602 (2007).
9. Baaske, P. et al. Extreme accumulation of nucleotides in simulated hydrothermal pore systems. *Proc Natl Acad Sci U S A* **104**, 9346-51 (2007).
10. McCollum, T.M., Ritter, G., Simoneit, B. R. T. Lipid synthesis under hydrothermal conditions by Fischer-Tropsch-type reactions. *Origins of Life and Evolution of Biospheres* **29**, 153-166 (1999).
11. Hennet, R.J.C., Holm N.G., Engel, M.H. Abiotic synthesis of amino acids under hydrothermal conditions and the origin of life: a perpetual phenomenon? *Naturwissenschaften* **79**, 361-365 (1992).
12. Mitchell, P. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* **191**, 144-8 (1961).
13. Schafer, G., Engelhard, M. & Muller, V. Bioenergetics of the Archaea. *Microbiol Mol Biol Rev* **63**, 570-620 (1999).
14. Muller, V. & Gruber, G. ATP synthases: structure, function and evolution of unique energy converters. *Cell Mol Life Sci* **60**, 474-94 (2003).
15. Koch, A.L. & Schmidt, T.M. The first cellular bioenergetic process: primitive generation of a proton-motive force. *J Mol Evol* **33**, 297-304 (1991).

16. Goodsell, D.S. et al. The RCSB PDB "Molecule of the Month": Inspiring a Molecular View of Biology. *PLoS Biol* **13**, e1002140 (2015).
17. Efremov, R.G., Baradaran, R. & Sazanov, L.A. The architecture of respiratory complex I. *Nature* **465**, 441-5 (2010).
18. Efremov, R.G. & Sazanov, L.A. Structure of the membrane domain of respiratory complex I. *Nature* **476**, 414-20 (2011).
19. Yankovskaya, V. et al. Architecture of succinate dehydrogenase and reactive oxygen species generation. *Science* **299**, 700-4 (2003).
20. Zhang, Z. et al. Electron transfer by domain movement in cytochrome bc1. *Nature* **392**, 677-84 (1998).
21. Takano, T. & Dickerson, R.E. Redox conformation changes in refined tuna cytochrome c. *Proc Natl Acad Sci U S A* **77**, 6371-5 (1980).
22. Yoshikawa, S. et al. Redox-coupled crystal structural changes in bovine heart cytochrome c oxidase. *Science* **280**, 1723-9 (1998).
23. Rastogi, V.K. & Girvin, M.E. Structural changes linked to proton translocation by subunit c of the ATP synthase. *Nature* **402**, 263-8 (1999).
24. Gibbons, C., Montgomery, M.G., Leslie, A.G. & Walker, J.E. The structure of the central stalk in bovine F(1)-ATPase at 2.4 A resolution. *Nat Struct Biol* **7**, 1055-61 (2000).
25. Wilkens, S., Borchardt, D., Weber, J. & Senior, A.E. Structural characterization of the interaction of the delta and alpha subunits of the Escherichia coli F1F0-ATP synthase by NMR spectroscopy. *Biochemistry* **44**, 11786-94 (2005).
26. Del Rizzo, P.A., Bi, Y., Dunn, S.D. & Shilton, B.H. The "second stalk" of Escherichia coli ATP synthase: structure of the isolated dimerization domain. *Biochemistry* **41**, 6875-84 (2002).
27. Lee, L.K., Stewart, A.G., Donohoe, M., Bernal, R.A. & Stock, D. The structure of the peripheral stalk of Thermus thermophilus H+-ATPase/synthase. *Nat Struct Mol Biol* **17**, 373-8 (2010).
28. Boyer, P.D. A perspective of the binding change mechanism for ATP synthesis. *FASEB J* **3**, 2164-78 (1989).
29. Engelhardt, W.A. Correlation between respiration and conversion of pyrophosphate in avian erythrocytes. *Biochemische Zeitschrift* **251**, 343-368 (1932).
30. Boyer, P.D., Cross, R.L. & Momsen, W. A new concept for energy coupling in oxidative phosphorylation based on a molecular explanation of the oxygen exchange reactions. *Proc Natl Acad Sci U S A* **70**, 2837-9 (1973).
31. Rosing, J., Kayalar, C. & Boyer, P.D. Evidence for energy-dependent change in phosphate binding for mitochondrial oxidative

phosphorylation based on measurements of medium and intermediate phosphate-water exchanges. *J Biol Chem* **252**, 2478-85 (1977).
32. Kayalar, C., Rosing, J. & Boyer, P.D. An alternating site sequence for oxidative phosphorylation suggested by measurement of substrate binding patterns and exchange reaction inhibitions. *J Biol Chem* **252**, 2486-91 (1977).
33. Hutton, R.L. & Boyer, P.D. Subunit interaction during catalysis. Alternating site cooperativity of mitochondrial adenosine triphosphatase. *J Biol Chem* **254**, 9990-3 (1979).
34. Bruist, M.F. & Hammes, G.G. Further characterization of nucleotide binding sites on chloroplast coupling factor one. *Biochemistry* **20**, 6298-305 (1981).
35. Boyer, P.D. The binding change mechanism for ATP synthase--some probabilities and possibilities. *Biochim Biophys Acta* **1140**, 215-50 (1993).
36. Cross, R.L. The mechanism and regulation of ATP synthesis by F1-ATPases. *Annu Rev Biochem* **50**, 681-714 (1981).
37. Abrahams, J.P. et al. Inherent asymmetry of the structure of F1-ATPase from bovine heart mitochondria at 6.5 A resolution. *EMBO J* **12**, 1775-80 (1993).
38. Abrahams, J.P., Leslie, A.G., Lutter, R. & Walker, J.E. Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria. *Nature* **370**, 621-8 (1994).
39. Lutter, R. et al. Crystallization of F1-ATPase from bovine heart mitochondria. *J Mol Biol* **229**, 787-90 (1993).
40. AB, N.M. Press Release: The Nobel Prize in Chemistry 1997. Vol. 2016 (1997).
41. Noji, H., Yasuda, R., Yoshida, M. & Kinosita, K., Jr. Direct observation of the rotation of F1-ATPase. *Nature* **386**, 299-302 (1997).
42. Baker, L.A., Watt, I.N., Runswick, M.J., Walker, J.E. & Rubinstein, J.L. Arrangement of subunits in intact mammalian mitochondrial ATP synthase determined by cryo-EM. *Proc Natl Acad Sci U S A* **109**, 11675-80 (2012).
43. Benlekbir, S., Bueler, S.A. & Rubinstein, J.L. Structure of the vacuolar-type ATPase from Saccharomyces cerevisiae at 11-A resolution. *Nat Struct Mol Biol* **19**, 1356-62 (2012).
44. Lau, W.C. & Rubinstein, J.L. Subnanometre-resolution structure of the intact Thermus thermophilus H+-driven ATP synthase. *Nature* **481**, 214-8 (2012).

45. Yokoyama, K., Nakano, M., Imamura, H., Yoshida, M. & Tamakoshi, M. Rotation of the proteolipid ring in the V-ATPase. *J Biol Chem* **278**, 24255-8 (2003).
46. Muench, S.P., Trinick, J. & Harrison, M.A. Structural divergence of the rotary ATPases. *Q Rev Biophys* **44**, 311-56 (2011).
47. Pullman, M.E., Penefsky, H.S., Datta, A. & Racker, E. Partial resolution of the enzymes catalyzing oxidative phosphorylation. I. Purification and properties of soluble dinitrophenol-stimulated adenosine triphosphatase. *J Biol Chem* **235**, 3322-9 (1960).
48. Kagawa, Y. & Racker, E. Partial resolution of the enzymes catalyzing oxidative phosphorylation. 8. Properties of a factor conferring oligomycin sensitivity on mitochondrial adenosine triphosphatase. *J Biol Chem* **241**, 2461-6 (1966).
49. Green, D. The electromechanochemical model for energy coupling in mitochondria. *Biochim Biophys Acta* **346**, 27-78 (1974).
50. Hoppe, J., Schairer, H.U. & Sebald, W. Identification of amino-acid substitutions in the proteolipid subunit of the ATP synthase from dicyclohexylcarbodiimide-resistant mutants of Escherichia coli. *Eur J Biochem* **112**, 17-24 (1980).
51. Dickson, V.K., Silvester, J.A., Fearnley, I.M., Leslie, A.G. & Walker, J.E. On the structure of the stator of the mitochondrial ATP synthase. *EMBO J* **25**, 2911-8 (2006).
52. Davies, K.M., Anselmi, C., Wittig, I., Faraldo-Gomez, J.D. & Kuhlbrandt, W. Structure of the yeast F1Fo-ATP synthase dimer and its role in shaping the mitochondrial cristae. *Proc Natl Acad Sci U S A* **109**, 13602-7 (2012).
53. Stewart, A.G., Laming, E.M., Sobti, M. & Stock, D. Rotary ATPases--dynamic molecular machines. *Curr Opin Struct Biol* **25**, 40-8 (2014).
54. Forgac, M. Vacuolar ATPases: rotary proton pumps in physiology and pathophysiology. *Nat Rev Mol Cell Biol* **8**, 917-29 (2007).
55. Arai, H. et al. Subunit composition and ATP site labeling of the coated vesicle proton-translocating adenosinetriphosphatase. *Biochemistry* **26**, 6632-8 (1987).
56. Kakinuma, Y., Ohsumi, Y. & Anraku, Y. Properties of H+-translocating adenosine triphosphatase in vacuolar membranes of SAccharomyces cerevisiae. *J Biol Chem* **256**, 10859-63 (1981).
57. Uchida, E., Ohsumi, Y. & Anraku, Y. Purification and properties of H+-translocating, Mg2+-adenosine triphosphatase from vacuolar membranes of Saccharomyces cerevisiae. *J Biol Chem* **260**, 1090-5 (1985).

58. Xie, X.S. & Stone, D.K. Isolation and reconstitution of the clathrin-coated vesicle proton translocating complex. *J Biol Chem* **261**, 2492-5 (1986).
59. Pedersen, P.A.a.C., E. Ion motive ATPases. 1. Ubiquity, properties and significance to cell function. *Trends Biochem Sci* **12**, 146-150 (1987).
60. Nelson, N. Structure, molecular genetics, and evolution of vacuolar H+-ATPases. *J Bioenerg Biomembr* **21**, 553-71 (1989).
61. Imamura, H. et al. Evidence for rotation of V1-ATPase. *Proc Natl Acad Sci U S A* **100**, 2312-5 (2003).
62. Jefferies, K.C. & Forgac, M. Subunit H of the vacuolar (H+) ATPase inhibits ATP hydrolysis by the free V1 domain by interaction with the rotary subunit F. *J Biol Chem* **283**, 4512-9 (2008).
63. Oot, R.A., Huang, L.S., Berry, E.A. & Wilkens, S. Crystal structure of the yeast vacuolar ATPase heterotrimeric EGC(head) peripheral stalk complex. *Structure* **20**, 1881-92 (2012).
64. Stewart, A.G. & Stock, D. Priming a molecular motor for disassembly. *Structure* **20**, 1799-800 (2012).
65. Smardon, A.M. et al. The RAVE complex is an isoform-specific V-ATPase assembly factor in yeast. *Mol Biol Cell* **25**, 356-67 (2014).
66. Numoto, N., Hasegawa, Y., Takeda, K. & Miki, K. Inter-subunit interaction and quaternary rearrangement defined by the central stalk of prokaryotic V1-ATPase. *EMBO Rep* **10**, 1228-34 (2009).
67. Saijo, S. et al. Crystal structure of the central axis DF complex of the prokaryotic V-ATPase. *Proc Natl Acad Sci U S A* **108**, 19955-60 (2011).
68. Iwata, M. et al. Crystal structure of a central stalk subunit C and reversible association/dissociation of vacuole-type ATPase. *Proc Natl Acad Sci U S A* **101**, 59-64 (2004).
69. Murata, T., Yamato, I., Kakinuma, Y., Leslie, A.G. & Walker, J.E. Structure of the rotor of the V-Type Na+-ATPase from Enterococcus hirae. *Science* **308**, 654-9 (2005).
70. Srinivasan, S., Vyas, N.K., Baker, M.L. & Quiocho, F.A. Crystal structure of the cytoplasmic N-terminal domain of subunit I, a homolog of subunit a, of V-ATPase. *J Mol Biol* **412**, 14-21 (2011).
71. Drory, O., Frolow, F. & Nelson, N. Crystal structure of yeast V-ATPase subunit C reveals its stator function. *EMBO Rep* **5**, 1148-52 (2004).
72. Sagermann, M., Stevens, T.H. & Matthews, B.W. Crystal structure of the regulatory subunit H of the V-type ATPase of Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **98**, 7134-9 (2001).

73. Nakano, M. et al. ATP hydrolysis and synthesis of a rotary motor V-ATPase from Thermus thermophilus. *J Biol Chem* **283**, 20789-96 (2008).
74. Muller, V., Ruppert, C. & Lemker, T. Structure and function of the A1A0-ATPases from methanogenic Archaea. *J Bioenerg Biomembr* **31**, 15-27 (1999).
75. Pogoryelov, D., Yildiz, O., Faraldo-Gomez, J.D. & Meier, T. High-resolution structure of the rotor ring of a proton-dependent ATP synthase. *Nat Struct Mol Biol* **16**, 1068-73 (2009).
76. Pogoryelov, D. et al. Microscopic rotary mechanism of ion translocation in the F(o) complex of ATP synthases. *Nat Chem Biol* **6**, 891-9 (2010).
77. Preiss, L., Yildiz, O., Hicks, D.B., Krulwich, T.A. & Meier, T. A new type of proton coordination in an F(1)F(o)-ATP synthase rotor ring. *PLoS Biol* **8**, e1000443 (2010).
78. Matthies, D. et al. High-resolution structure and mechanism of an F/V-hybrid rotor ring in a Na(+)-coupled ATP synthase. *Nat Commun* **5**, 5286 (2014).
79. Hoppe, J., Schairer, H.U. & Sebald, W. The proteolipid of a mutant ATPase from Escherichia coli defective in H+-conduction contains a glycine instead of the carbodiimide-reactive aspartyl residue. *FEBS Lett* **109**, 107-11 (1980).
80. Hoppe, J., Schairer, H.U., Friedl, P. & Sebald, W. An Asp-Asn substitution in the proteolipid subunit of the ATP-synthase from Escherichia coli leads to a non-functional proton channel. *FEBS Lett* **145**, 21-9 (1982).
81. Sebald, W.a.W., E. Amino acid sequence of the proteolipid subunit of the ATP synthase from spinach chloroplasts. *FEBS Lett* **122**, 307-311 (1980).
82. Walker, J.E., Lutter, R., Dupuis, A. & Runswick, M.J. Identification of the subunits of F1F0-ATPase from bovine heart mitochondria. *Biochemistry* **30**, 5369-78 (1991).
83. Hoppe, J. & Sebald, W. Amino acid sequence of the proteolipid subunit of the proton-translocating ATPase complex from the thermophilic bacterium PS-3. *Eur J Biochem* **107**, 57-65 (1980).
84. Sebald, W., Wachter, E. & Tzagoloff, A. Identification of amino acid substitutions in the dicyclohexylcarbodiimide-binding subunit of the mitochondrial ATPase complex from oligomycin-resistant mutants of Saccharomyces cerevisiae. *Eur J Biochem* **100**, 599-607 (1979).
85. Fillingame, R.H. H+ transport and coupling by the F0 sector of the ATP synthase: insights into the molecular mechanism of function. *J Bioenerg Biomembr* **24**, 485-91 (1992).

86. Sambongi, Y. et al. Mechanical rotation of the c subunit oligomer in ATP synthase (F0F1): direct observation. *Science* **286**, 1722-4 (1999).
87. Jiang, W. & Fillingame, R.H. Interacting helical faces of subunits a and c in the F1Fo ATP synthase of Escherichia coli defined by disulfide cross-linking. *Proc Natl Acad Sci U S A* **95**, 6607-12 (1998).
88. Cain, B.D. & Simoni, R.D. Proton translocation by the F1F0ATPase of Escherichia coli. Mutagenic analysis of the a subunit. *J Biol Chem* **264**, 3292-300 (1989).
89. Eya, S., Maeda, M. & Futai, M. Role of the carboxyl terminal region of H(+)-ATPase (F0F1) a subunit from Escherichia coli. *Arch Biochem Biophys* **284**, 71-7 (1991).
90. Howitt, S.M. & Cox, G.B. Second-site revertants of an arginine-210 to lysine mutation in the a subunit of the F0F1-ATPase from Escherichia coli: implications for structure. *Proc Natl Acad Sci U S A* **89**, 9799-803 (1992).
91. Lightowlers, R.N., Howitt, S.M., Hatch, L., Gibson, F. & Cox, G.B. The proton pore in the Escherichia coli F0F1-ATPase: a requirement for arginine at position 210 of the a-subunit. *Biochim Biophys Acta* **894**, 399-406 (1987).
92. Vik, S.B. & Antonio, B.J. A mechanism of proton translocation by F1F0 ATP synthases suggested by double mutants of the a subunit. *J Biol Chem* **269**, 30364-9 (1994).
93. Hartzog, P.E. & Cain, B.D. Mutagenic analysis of the a subunit of the F1F0 ATP synthase in Escherichia coli: Gln-252 through Tyr-263. *J Bacteriol* **175**, 1337-43 (1993).
94. Cain, B.D. & Simoni, R.D. Interaction between Glu-219 and His-245 within the a subunit of F1F0-ATPase in Escherichia coli. *J Biol Chem* **263**, 6606-12 (1988).
95. Senior, A.E. Secondary and tertiary structure of membrane proteins involved in proton translocation. *Biochim Biophys Acta* **726**, 81-95 (1983).
96. Valiyaveetil, F.I. & Fillingame, R.H. Transmembrane topography of subunit a in the Escherichia coli F1F0 ATP synthase. *J Biol Chem* **273**, 16241-7 (1998).
97. Fillingame, R.H., Angevine, C.M. & Dmitriev, O.Y. Coupling proton movements to c-ring rotation in F(1)F(o) ATP synthase: aqueous access channels and helix rotations at the a-c interface. *Biochim Biophys Acta* **1555**, 29-36 (2002).
98. Junge, W., Lill, H. & Engelbrecht, S. ATP synthase: an electrochemical transducer with rotatory mechanics. *Trends Biochem Sci* **22**, 420-3 (1997).

99. Hakulinen, J.K. et al. Structural study on the architecture of the bacterial ATP synthase Fo motor. *Proc Natl Acad Sci U S A* **109**, E2050-6 (2012).
100. Schwem, B.E. & Fillingame, R.H. Cross-linking between helices within subunit a of Escherichia coli ATP synthase defines the transmembrane packing of a four-helix bundle. *J Biol Chem* **281**, 37861-7 (2006).
101. Zhou, M. et al. Mass spectrometry of intact V-type ATPases reveals bound lipids and the effects of nucleotide binding. *Science* **334**, 380-5 (2011).
102. Seddon, A.M., Curnow, P. & Booth, P.J. Membrane proteins, lipids and detergents: not just a soap opera. *Biochim Biophys Acta* **1666**, 105-17 (2004).
103. Jolivet, E., L'Haridon, S., Corre, E., Forterre, P. & Prieur, D. Thermococcus gammatolerans sp. nov., a hyperthermophilic archaeon from a deep-sea hydrothermal vent that resists ionizing radiation. *Int J Syst Evol Microbiol* **53**, 847-51 (2003).
104. Drew, D. et al. A scalable, GFP-based pipeline for membrane protein overexpression screening and purification. *Protein Sci* **14**, 2011-7 (2005).
105. Magasanik, B. Catabolite repression. *Cold Spring Harb Symp Quant Biol* **26**, 249-56 (1961).
106. Prive, G.G. Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41**, 388-97 (2007).
107. Limited, G.H.U. Gel Filtration Calibration Kits: Product booklet. (Buckinghamshire, UK, 2011).
108. Roosild, T.P. et al. NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science* **307**, 1317-21 (2005).
109. Notredame, C., Higgins, D. G., Heringa, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* **302**, 205-217 (2000).
110. LeMaire, M., Champeil, P. and Moller, J. V. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim Biophys Acta* **1508**, 86-111 (2000).
111. Garavito., A.i.c.w.P.R.M. D322LA - n-Decyl-β-D-Maltopyranoside, Anagrade.
112. Tribet, C., Audebert, R. & Popot, J.L. Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc Natl Acad Sci U S A* **93**, 15047-50 (1996).
113. Tehei, M. et al. Thermal fluctuations in amphipol A8-35 particles: a neutron scattering and molecular dynamics study. *J Membr Biol* **247**, 897-908 (2014).

114. Hu, T.C., Korczynska, J., Smith, D.K. & Brzozowski, A.M. High-molecular-weight polymers for protein crystallization: poly-gamma-glutamic acid-based precipitants. *Acta Crystallogr D Biol Crystallogr* **64**, 957-63 (2008).
115. Polovinkin, V. et al. High-resolution structure of a membrane protein transferred from amphipol to a lipidic mesophase. *J Membr Biol* **247**, 997-1004 (2014).
116. Popot, J.L. et al. Amphipols: polymeric surfactants for membrane biology research. *Cell Mol Life Sci* **60**, 1559-74 (2003).
117. Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nat Protoc* **4**, 706-31 (2009).
118. Gasteiger E., H.C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook* (ed. Walker, J.M.) 571-607 (Humana Press, 2005).
119. Rajarathnam, K. & Rosgen, J. Isothermal titration calorimetry of membrane proteins - progress and challenges. *Biochim Biophys Acta* **1838**, 69-77 (2014).
120. Freyer, M.W., Lewis, E. A. Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand ibinding and kinetic interactions. *Methods in Cell Biology* **84**, 79-113 (2008).
121. Yamamoto, M. et al. Interaction and stoichiometry of the peripheral stalk subunits NtpE and NtpF and the N-terminal hydrophilic domain of NtpI of Enterococcus hirae V-ATPase. *J Biol Chem* **283**, 19422-31 (2008).
122. Oot, R.A. & Wilkens, S. Domain characterization and interaction of the yeast vacuolar ATPase subunit C with the peripheral stator stalk subunits E and G. *J Biol Chem* **285**, 24654-64 (2010).
123. Fethiere, J., Venzke, D., Madden, D.R. & Bottcher, B. Peripheral stator of the yeast V-ATPase: stoichiometry and specificity of interaction between the EG complex and subunits C and H. *Biochemistry* **44**, 15906-14 (2005).
124. Zhou, M. et al. Ion mobility-mass spectrometry of a rotary ATPase reveals ATP-induced reduction in conformational flexibility. *Nat Chem* **6**, 208-15 (2014).
125. Morales-Rios, E., Montgomery, M.G., Leslie, A.G. & Walker, J.E. Structure of ATP synthase from Paracoccus denitrificans determined by X-ray crystallography at 4.0 A resolution. *Proc Natl Acad Sci U S A* **112**, 13231-6 (2015).
126. Rosenbaum, D.M. et al. GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science* **318**, 1266-73 (2007).

127. Liu, W. et al. Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **337**, 232-6 (2012).
128. Zhang, H. et al. Structure of the Angiotensin receptor revealed by serial femtosecond crystallography. *Cell* **161**, 833-44 (2015).
129. Rasmussen, S.G. et al. Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. *Nature* **469**, 175-80 (2011).
130. Shimamura, T. et al. Structure of the human histamine H1 receptor complex with doxepin. *Nature* **475**, 65-70 (2011).
131. Uysal, S. et al. Crystal structure of full-length KcsA in its closed conformation. *Proc Natl Acad Sci U S A* **106**, 6644-9 (2009).
132. Vaidehi, N., Grisshammer, R. & Tate, C.G. How Can Mutations Thermostabilize G-Protein-Coupled Receptors? *Trends Pharmacol Sci* **37**, 37-46 (2016).
133. Sauer, D.B., Karpowich, N.K., Song, J.M. & Wang, D.N. Rapid Bioinformatic Identification of Thermostabilizing Mutations. *Biophys J* **109**, 1420-8 (2015).
134. Grogan, D.W. Phenotypic characterization of the archaebacterial genus Sulfolobus: comparison of five wild-type strains. *J Bacteriol* **171**, 6710-9 (1989).
135. Fiala, G., Stetter, K. O. . *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100 ºC. *Archives of Microbiology* **145**, 55-61 (1986).
136. Liu, Y., Boone, D.R., Sleat, R. & Mah, R.A. Methanosarcina mazei LYC, a New Methanogenic Isolate Which Produces a Disaggregating Enzyme. *Appl Environ Microbiol* **49**, 608-13 (1985).
137. Darland, G., Brock, T.D., Samsonoff, W. & Conti, S.F. A thermophilic, acidophilic mycoplasma isolated from a coal refuse pile. *Science* **170**, 1416-8 (1970).
138. Knaevelsrud, I., Kazazic, S., Birkeland, N.K. & Bjelland, S. The pH optimum of native uracil-DNA glycosylase of Archaeoglobus fulgidus compared to recombinant enzyme indicates adaption to cytosolic pH. *Acta Biochim Pol* **61**, 393-5 (2014).
139. Jones, W.J., Leigh, J. A., Mayer, F., Woese, C. R., Wolfe, R. S. *Methanococcus jannaschii* sp. nov., an extremely thermophilic methanogen from a submarine hydrothermal vent. *Archives of Microbiology* **136**, 254-261 (1983).
140. Gonzalez, J.M. et al. Pyrococcus horikoshii sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough. *Extremophiles* **2**, 123-30 (1998).
141. Jack Jones, W., Paynter, M. J. B., Gupta, R. Characterization of *Methanococcus maripaludis* sp. nov., a new methanogen isolated from salt marsh sediment. *Archives of Micorbiology* **135**, 91-97 (1983).

142. Oshima, T., Imahori, K. Description of *Thermus thermophilus* (Yoshida and Oshima) comb. nov., a onsporulating thermophilic bacterium from a Japanese thermal spa. *International Journal of Systematic and Evolutionary Bacteriology* **24**, 102-112 (1974).
143. Loginova, L.G., Egorova, L. A., Golovacheva, R. S., Seregina, L. M. *Thermus ruber* sp. nov., nom. rev. *International Journal of Systematic and Evolutionary Bacteriology* **34**, 498-499 (1984).
144. Murray, R.G.E. The family Deinococcacaea. in *The prokaryotes.* (ed. Balows, A., Trüper, H. G., Dworkin, M., Harder, W., Schleifer, K.-H.) (Springer-Verlag, New York, 1992).
145. Allegretti, M. et al. Horizontal membrane-intrinsic alpha-helices in the stator a-subunit of an F-type ATP synthase. *Nature* **521**, 237-40 (2015).
146. Zhao, J., Benlekbir, S. & Rubinstein, J.L. Electron cryomicroscopy observation of rotational states in a eukaryotic V-ATPase. *Nature* **521**, 241-5 (2015).
147. Zhou, A. et al. Structure and conformational states of the bovine mitochondrial ATP synthase by cryo-EM. *Elife* **4**(2015).
148. Schep, D.G., Zhao, J. and Rubinstein, J.L. . Models for the a subunits of the *Thermus thermophilus* V/A-ATPase and *Saccharomyces cerevisiae* V-ATPase enzymes by cryo-EM and evolutionary covariance. *Proc Natl Acad Sci U S A* (2016).
149. Mitome, N. et al. Essential arginine residue of the F(o)-a subunit in F(o)F(1)-ATP synthase has a role to prevent the proton shortcut without c-ring rotation in the F(o) proton channel. *Biochem J* **430**, 171-7 (2010).
150. Cheng, Y. Single-Particle Cryo-EM at Crystallographic Resolution. *Cell* **161**, 450-7 (2015).
151. Campbell, M.G. et al. Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. *Structure* **20**, 1823-8 (2012).
152. Sigworth, F.J., Doerschuk, P.C., Carazo, J.M. & Scheres, S.H. An introduction to maximum-likelihood methods in cryo-EM. *Methods Enzymol* **482**, 263-94 (2010).
153. Scheres, S.H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-30 (2012).
154. Amunts, A. et al. Structure of the yeast mitochondrial large ribosomal subunit. *Science* **343**, 1485-9 (2014).
155. Fischer, N. et al. Structure of the E. coli ribosome-EF-Tu complex at <3 A resolution by Cs-corrected cryo-EM. *Nature* **520**, 567-70 (2015).
156. Bartesaghi, A. et al. 2.2 A resolution cryo-EM structure of beta-galactosidase in complex with a cell-permeant inhibitor. *Science* **348**, 1147-51 (2015).

157. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
158. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60**, 2256-68 (2004).
159. Rak, M., Gokova, S. & Tzagoloff, A. Modular assembly of yeast mitochondrial ATP synthase. *EMBO J* **30**, 920-30 (2011).
160. Kish-Trier, E. & Wilkens, S. Interaction of the Thermoplasma acidophilum A1A0-ATP synthase peripheral stalk with the catalytic domain. *FEBS Lett* **583**, 3121-6 (2009).
161. Stewart, A.G., Sobti, M., Harvey, R.P. & Stock, D. Rotary ATPases: models, machine elements and technical specifications. *Bioarchitecture* **3**, 2-12 (2013).
162. Altschul, S.F. & Koonin, E.V. Iterated profile searches with PSI-BLAST--a tool for discovery in protein databases. *Trends Biochem Sci* **23**, 444-7 (1998).
163. Gibson, D.G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-5 (2009).
164. Xu, F., Liu, W., Hanson, M.A., Stevens, R.C. & Cherezov, V. Development of an Automated High Throughput LCP-FRAP Assay to Guide Membrane Protein Crystallization in Lipid Mesophases. *Cryst Growth Des* **11**, 1193-1201 (2011).
165. Anatrace Maltosides. (2016).

Appendix One

Table 6-1: Summary of all crystallisation trials

Protein sample	Detergent or amphipol	Protein conc. (mg mL-1)	Crystallisation screens	Incubation temp. (°C)
Subunit I	Anz. 3-12	5	PACT Cryo 1&2 x 2 Index MembFac MemPlus	20
Subunit I	Anz. 3-12	6	PACT Hampton CS Cryo 1&2 Index MembFac MemPlus Crystal Screen HT SaltRX Classic 1&2 Classic HTS I Classic HTS II Clear Strategy I Clear Strategy II MIDAS Morpheus Green JCSG-plus MemGold MemStart + MemSys Detergent Screen	4
Subunit I	DM	5	MemStart + MemSys MembFac MemGold Cryo 1&2 MemPlus	20
Subunit I	DM	4	Hampton CS Index Classic 1&2 Additive Screen	20
Subunit I	DM	5	Morpheus Green	20
Subunit I	DM	5	SaltRX	20

Subunit I	DM	8	JCSG-plus MIDAS Classic HTS I Classic HTS II Hampton CS Clear Strategy I Clear Strategy II	20
Subunit I	Amphipol A8-35	5.6	Hampton CS MemStart + MemSys MemPlus Cryo 1&2 MemGold MembFac MIDAS SaltRX Clear Strategy I Clear Strategy II Morpheus Green JCSG-plus PGA Screen	20
Subunit I	Amphipol A8-35	4	PACT Classic HTS I Classic HTS II Classic 1&2	20
Mistic-I_{TM}-His_{10}	DM	3.4	MemGold MemStart + MemSys x 2 PACT Index Optimisation screen Cryo 1&2 MemPlus Detergent Screen	20
T7-I_{TM}-His_{10}	DM	2.2	MemStart + MemSys Cryo 1&2 MembFac	20
T7-I_{TM}-His_{10}	DM	2.2	PACT	20
T7-I_{TM}-His_{10}	DM	2.2	MemPlus	20

			MemGold Wizard Classic 1&2	
Subunit I*	DM	10	Wizard Cubic LCP Screen x 2 MemGold x 2 LCP Custom Screen 1 x 2 LCP Custom Screen 2 x 2	20

*LCP trials are shaded grey, vapour diffusion trials are unshaded