Expression and Purification of the Cystic Fibrosis Transmembrane Conductance Regulator from *Saccharomyces cerevisiae* for High-Resolution Structural Studies.

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy (PhD) in the Faculty of Life Sciences

2013

Natasha Cant
List of Contents

List of Contents ............................................................... 2
List of Figures and Tables .................................................. 7
Abstract ............................................................................. 12
Declaration .......................................................................... 13
Copyright Statement ........................................................... 13
Acknowledgements ................................................................ 14
List of Publications ............................................................... 15
Chapter 1 - Introduction ....................................................... 16
  1.1. Structural studies of membrane proteins ......................... 16
  1.1.2. Expression and purification of membrane proteins ............... 16
  1.1.3. Detergents .................................................................. 17
  1.1.4. Lipids ......................................................................... 18
  1.1.5. X-ray Crystallography ................................................ 19
  1.1.6. Electron Microscopy .................................................... 21
  1.1.7. Electron Crystallography .............................................. 22
  1.1.8. Single particle analysis ................................................. 24
  1.2. ATP-binding cassette transporters .................................... 24
  1.2.1. Structure of ABC transporters ..................................... 25
  1.2.2. Mechanism of ABC transporters .................................... 26
  1.3. CFTR .............................................................................. 28
  1.3.1. CFTR Structure ........................................................ 29
  1.3.2. CFTR function and mechanism ..................................... 33
  1.3.3. CFTR biogenesis ........................................................ 36
  1.3.4. Cystic fibrosis ............................................................ 38
1.4.1. CFTR mutations ................................................................. 38
1.4.2. CF therapy ........................................................................... 40
1.5 Strategy of project .................................................................... 42
  1.5.1. CFFT Structure consortium .................................................. 42
  1.5.2. Yeast expression system ....................................................... 45
  1.5.3. CFTR construct .................................................................... 46
  1.5.4. Project aims .......................................................................... 47

Chapter 2 - Materials and Methods .................................................. 48
2.1. Materials .................................................................................. 48
  2.1.1. Chemicals ............................................................................. 48
  2.1.2. Yeast cells ........................................................................... 50
  2.1.3. CFTR gene ........................................................................... 50
  2.1.4. Expression vector ................................................................. 50
  2.1.5. Proteins ............................................................................... 52
  2.1.6. Media .................................................................................. 53
  2.1.7. Inhibitors ............................................................................ 53
  2.1.8. Buffers: purification ............................................................... 54
  2.1.9. Buffers: SDS-PAGE ............................................................... 54
  2.1.10 Buffers: ATPase assay ......................................................... 55

2.2. Methods: expression and purification ........................................ 55
  2.2.1. Transformation of yeast cells ............................................... 55
  2.2.2. Small-scale cell culture ......................................................... 55
  2.2.3. Large-scale fermentation cultures ......................................... 56
  2.2.4. Fluorescence microscopy ...................................................... 56
2.2.5. Small-scale crude membrane preparation ..................................................57
2.2.6. Large-scale microsome preparation ..........................................................57
2.2.7. Bradford assay ..........................................................................................58
2.2.8. GFP-fluorescence spectroscopy .................................................................58
2.2.9. SDS-PAGE ...............................................................................................58
2.2.10. Detergent solubilisation screen ...............................................................59
2.2.11. Detergent solubilisation for CFTR purification ........................................60
2.2.12. Nickel-affinity chromatography purification .............................................60
2.2.13. Size-exclusion chromatography (SEC) purification .................................61
2.2.14. Static Light Scattering ............................................................................61
2.2.15. Mass spectrometry ..................................................................................62
2.2.16. Glycosylation ..........................................................................................62
2.2.17. Phosphorylation ......................................................................................62
2.2.18. TEV cleavage .........................................................................................63

2.3. Methods: stability .........................................................................................63
2.3.1. Static Light Scattering ...............................................................................63
2.3.2. CPM dye binding assay ...........................................................................64
2.3.3. Tryptophan (Trp) fluorescence measurements ..........................................65
2.3.4. EM ............................................................................................................65
2.3.5. Thermal gel analysis (TGA) ....................................................................65
2.3.6. Graphical analysis of the thermal unfolding data ......................................66
2.3.7. Analysis of CFTR structure models .........................................................66

2.4. Methods: homogeneity ................................................................................67
2.4.1. SEC multi-angle light scattering (SEC-MALLS) ........................................67
2.4.4. Cryo- EM ............................................................................................67
2.4.3. NS-EM ................................................................................................68
2.4.4. SPA ......................................................................................................68
2.4.5. 2D crystallographic analysis .................................................................69
2.5. Methods: activity .....................................................................................69
2.5.1. Reconstitution ......................................................................................69
2.5.2. PKA Treatment ....................................................................................70
2.5.3. ATPase activity assay ..........................................................................70
2.5.4. Graphical and statistical analysis of ATPase activity data ..................71

Chapter 3 – Results and Discussion: Expression and Purification .............72
3.1. Establishing CFTR-expressing yeast cell lines. .......................................72
3.2. Optimisation of yeast cell growth conditions for CFTR expression ........73
3.3. Assessing CFTR cellular localisation by GFP-fluorescence microscopy ....76
3.4. Isolating CFTR-containing microsomes from yeast cells .......................77
3.5. Post-translational modification of CFTR in yeast ...................................79
3.6. Detergent solubilisation of CFTR from yeast microsomes .....................82
3.7. Purification of CFTR in LPG ....................................................................84
3.8. Purification of CFTR in DDM .................................................................87
3.9. Identification of contaminating yeast proteins by mass spectrometry ......89
3.10. Cleavage of C-terminal GFP-fusion tag using TEV protease ................91

Chapter 4 – Results and Discussion: Stability .............................................93
4.1. Protein concentration dependence of CFTR stability .............................93
4.2. Thermal stability of CFTR .......................................................................95
4.2.1. CPM dye binding assay.................................................................95
4.2.2. Tryptophan fluorescence ..........................................................100
4.2.3. Static light scattering ..................................................................103
4.2.4. Thermal SDS-PAGE analysis (TGA) ..........................................105

Chapter 5 – Results and Discussion: Homogeneity ...........................................109
5.1. SEC-MALLS analysis of CFTR protein homogeneity .............................110
5.2. EM analysis of CFTR homogeneity .......................................................113
5.2.1. Cryo-EM analysis of LPG-purified CFTR .......................................113
5.2.2. NS-EM of DDM-purified CFTR ......................................................116
5.3. 2D crystallisation ..............................................................................119

Chapter 6 – Results and Discussion: Activity ..................................................124
6.1. Reconstitution of CFTR for activity studies .........................................125
6.2. Effect of lipid and detergent on CFTR ATPase activity .........................129
6.3. Kinetics of CFTR ATPase activity .......................................................131
6.4. Effect of PKA treatment on CFTR activity ...........................................136
6.5. Inhibition of CFTR ATPase activity ......................................................139

Chapter 7 - Conclusions .............................................................................144
7.1. A novel yeast system for expression and purification of CFTR .............144
7.2. Remaining challenges in achieving a high-resolution structure of CFTR ....147
7.3. Future Work .....................................................................................151

References ...............................................................................................154
Appendix – Supplementary Figures ............................................................180

Final word count: 65871
List of Figures and Tables

Figure 1. The ATP switch model for the mechanism of ABC exporters..................27
Figure 2. Schematic of CFTR structure ..............................................................32
Figure 3. Schematic of CFTR molecular mechanism...........................................34
Figure 4. Cloning the CFTR gene into a GFP-fusion yeast expression vector ..........51
Figure 5. CFTR expression was detected in 1/5 transformation colonies ...............73
Figure 6. Optimisation of yeast cell growth conditions for CFTR expression ......75
Figure 7. CFTR has "spotted" localisation at the PM of yeast cells .....................77
Figure 8. Isolation of CFTR-containing microsomes .........................................78
Figure 9. PTM analysis of CFTR expressed in yeast ..........................................81
Figure 10. CFTR can be extracted from yeast microsomes in a range of detergents with varying solubilisation efficiency .........................................................83
Figure 11. Nickel-affinity chromatography purification of CFTR in LPG (A) and DDM (B) ...............................................................85
Figure 12. SEC purification of CFTR in LPG (A) and DDM (B) ..............................86
Figure 13. DDM-purified CFTR appears to aggregate at low concentrations of NaCl ..89
Figure 14. CFTR-GFP tag can be cleaved using TEV protease in DDM, but not LPG .92
Figure 15. CFTR is more soluble at high concentrations in LPG than DDM ...........94
Figure 16. Thermal stability of CFTR as monitored by the CPM dye-binding assay ....99
Figure 17. Thermal stability of CFTR, monitored by intrinsic Trp fluorescence .....102
Figure 18. Thermal denaturation of CFTR does not result in protein aggregation ......104
Figure 19. Thermal stability of CFTR assessed by SDS-PAGE ............................107
Figure 20. SEC-MALLS reveals CFTR is more homogenous in LPG (A) than DDM (B) ........................................................................112
Figure 21. Cryo-EM reveals LPG-purified CFTR is homogenous and monomeric ....115
Figure 22. NS-EM reveals DDM-purified CFTR is a heterogenous population of different sized aggregates .................................................................118
Figure 23. Epitaxial 2D crystals of CFTR are weakly ordered and yield a low-resolution structure map .................................................................121
Figure 24. CFTR was not fully reconstituted into liposomes but rather formed a lipid/detergent/protein mixture after incubation with Biobeads ............128
Figure 25. CFTR has higher ATPase activity when purified in DDM than LPG, and activity is stimulated by the addition of lipids ......................................131
Figure 26. CFTR ATPase activity as a function of ATP concentration (A) and time (B). .................................................................135
Figure 27. PKA treatment does not affect CFTR ATPase activity. .........................138
Figure 28. Inhibition of CFTR ATPase activity.................................................................143

Appendix 1. Quantification of CFTR..............................................................................180
Appendix 2. Mass spectrometry analysis of CFTR purified in LPG and DDM........181
Appendix 3. Location of Cys and Trp residues in chicken-CFTR structure............184
Appendix 4. CPM dye binding assay controls: GFP and ribosomal L3 subunit protein. ......................................................................................................................188
Appendix 5. Thermal stability of CFTR, as monitored by EM. .......................189
Appendix 6. Thermal gel analysis of GFP.................................................................190
Appendix 7. Cryo-EM micrograph of LPG-purified CFTR. .....................................191
Appendix 8. NS-EM micrograph of DDM-purified CFTR HMW aggregates. ........192
Appendix 9. NS-EM micrograph of DDM-purified CFTR. ........................................193
Appendix 10. Vector displacement plot of CFTR 2D crystal...............................194

Table 1. List of CFTR orthologue selected for screening within the CFFT Structure Consortium project.................................................................44
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D</td>
<td>two-dimensional</td>
</tr>
<tr>
<td>3D</td>
<td>three-dimensional</td>
</tr>
<tr>
<td>ABC</td>
<td>ATP-binding cassette</td>
</tr>
<tr>
<td>AEBSF</td>
<td>aminoethylbenzenesulfonyl fluoride</td>
</tr>
<tr>
<td>APS</td>
<td>ammonium persulphate</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>BHK</td>
<td>baby hamster kidney</td>
</tr>
<tr>
<td>bR</td>
<td>bacteriorhodopsin</td>
</tr>
<tr>
<td>BSA</td>
<td>albumin from bovine serum</td>
</tr>
<tr>
<td>C&lt;sub&gt;12&lt;/sub&gt;E&lt;sub&gt;8&lt;/sub&gt;</td>
<td>octaethylene glycol monododecyl ether</td>
</tr>
<tr>
<td>CF</td>
<td>cystic fibrosis</td>
</tr>
<tr>
<td>CFFT</td>
<td>Cystic Fibrosis Foundation Therapeutics</td>
</tr>
<tr>
<td>CFTR</td>
<td>the cystic fibrosis transmembrane conductance regulator</td>
</tr>
<tr>
<td>CIP</td>
<td>alkaline phosphatase from calf intestine</td>
</tr>
<tr>
<td>CMC</td>
<td>critical micelle concentration</td>
</tr>
<tr>
<td>CPM</td>
<td>coumarin maleimide</td>
</tr>
<tr>
<td>CTF</td>
<td>contrast transfer function</td>
</tr>
<tr>
<td>CV</td>
<td>column volumes</td>
</tr>
<tr>
<td>Cys</td>
<td>cysteine</td>
</tr>
<tr>
<td>DDM</td>
<td>dodecyl-maltoside</td>
</tr>
<tr>
<td>DM</td>
<td>decyl-maltoside</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulfoxide</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>E-64</td>
<td>epoxysuccinyl-leucylamido-butane</td>
</tr>
<tr>
<td>ECL</td>
<td>extracellular loops</td>
</tr>
<tr>
<td>EM</td>
<td>electron microscopy</td>
</tr>
<tr>
<td>ENaC</td>
<td>amiloride-sensitive sodium channel</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmic reticulum</td>
</tr>
<tr>
<td>ERAD</td>
<td>endoplasmic reticulum associated degradation</td>
</tr>
<tr>
<td>FC16</td>
<td>fos choline 16</td>
</tr>
<tr>
<td>FDA</td>
<td>food and drug administration</td>
</tr>
<tr>
<td>FEG</td>
<td>field emission gun</td>
</tr>
<tr>
<td>FFT</td>
<td>fast fourier transform</td>
</tr>
<tr>
<td>GD</td>
<td>glow discharged</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>GNS</td>
<td>genestein</td>
</tr>
<tr>
<td>GSH</td>
<td>glutathione</td>
</tr>
<tr>
<td>HCl</td>
<td>hydrochloric acid</td>
</tr>
<tr>
<td>HEK</td>
<td>human embryonic kidney</td>
</tr>
<tr>
<td>His</td>
<td>histidine</td>
</tr>
<tr>
<td>HMW</td>
<td>higher molecular weight</td>
</tr>
<tr>
<td>IC</td>
<td>ATPase inhibitor cocktail</td>
</tr>
<tr>
<td>ICL</td>
<td>intracellular loops</td>
</tr>
<tr>
<td>( K_m )</td>
<td>Michaelis constant</td>
</tr>
<tr>
<td>LDAO</td>
<td>lauryldimethylamine oxide</td>
</tr>
<tr>
<td>LiAc</td>
<td>lithium acetate</td>
</tr>
<tr>
<td>LPG</td>
<td>lyso-phosphatidylglycerol 14</td>
</tr>
<tr>
<td>LPR</td>
<td>lipid-to-protein ratio</td>
</tr>
<tr>
<td>MALLS</td>
<td>multi-angle light scattering</td>
</tr>
<tr>
<td>MgSO(_4)</td>
<td>magnesium sulphate</td>
</tr>
<tr>
<td>mH(_2)O</td>
<td>Millipore Q 0.2 ( \mu )M filtered water</td>
</tr>
<tr>
<td>MWCO</td>
<td>molecular weight cut off</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>NaN(_3)</td>
<td>sodium azide</td>
</tr>
<tr>
<td>NBD</td>
<td>nucleotide-binding domain</td>
</tr>
<tr>
<td>NM</td>
<td>nonyl-maltoside</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>NS</td>
<td>negative-stain</td>
</tr>
<tr>
<td>o/n</td>
<td>overnight</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>OM</td>
<td>octyl-maltoside</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PC</td>
<td>phosphatidylcholine</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Data Bank</td>
</tr>
<tr>
<td>PE</td>
<td>phosphatidylethanolamine</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>PG</td>
<td>phosphatidylglycerol</td>
</tr>
<tr>
<td>Pgp</td>
<td>P-glycoprotein</td>
</tr>
<tr>
<td>( P_i )</td>
<td>inorganic phosphate</td>
</tr>
</tbody>
</table>
PIs  protease inhibitors
PKA  protein kinase A
PKC  protein kinase C
PMSF phenylmethanesulfonylfluoride
PNGaseF peptide-N-glycosidase F
PS  phosphatidylserine
PTM  post-translational modification
R region  regulatory region
RI  refractive index
RPM  revolutions per minute
RTM  room temperature
SAXS  single angle X-ray scattering
SCN  sodium thiocynate
SD  standard deviation
SDS  sodium dodecyl-sulphate
SEC  size-exclusion chromatography
SPA  single particle analysis
SUMO  small ubiquitin-related modifier
TEMED  tetramethylthlenediamine
TEV  tobacco etch virus
TGA  thermal gel analysis
Tm  the mid-point temperature of thermal unfolding
TMD  transmembrane domain
Trp  tryptophan
U  units
UA  uranyl acetate
v/v  volume/volume
Vmax  the maximum rate of velocity at high substrate concentrations
w/v  weight/volume
YNB  yeast nitrogen base
YPD  yeast peptone dextrose
ΔF508  deletion mutation of phenylalanine at position 508
The cystic fibrosis transmembrane conductance regulator (CFTR) is an ABC transporter family protein that acts as an ion channel. Mutations in CFTR cause the most common genetic disease in Caucasian populations, cystic fibrosis (CF). The high-resolution X-ray crystal structure of CFTR is now needed to aid the design of CFTR-targeted drugs for CF treatment and also to elucidate the molecular mechanisms behind its unique function as an ATP-ligand gated ion channel. However, until now, such structural studies have been severely limited by the lack of sufficient quantities of purified full-length CFTR protein. This thesis reports the novel over-expression and purification of milligram quantities of the chicken orthologue of CFTR protein from a *Saccharomyces cerevisiae* (yeast) expression system. A green fluorescent protein (GFP) tag fused to the CFTR C-terminus allowed rapid detection of the protein throughout the purification procedure. CFTR was expressed under an inducible promoter and appeared localised at, or near to, the plasma membrane, where it represented around 1% of total protein after isolation in yeast microsomes. CFTR was solubilised from microsomes and purified using the detergents dodecylmaltoside (DDM) and lyso-phosphatidyl glycerol (LPG), by nickel affinity and size exclusion chromatography (SEC) to yield 1-2 mg of CFTR protein per 18 L fermentation culture. CFTR thermal stability was probed using fluorescent measurements to reveal a two-state cooperative unfolding transition around 40 °C for the DDM-purified protein, but no such transition was observed for the LPG-purified material. Light scattering and electron microscopy techniques revealed that, in LPG, CFTR was a homogenous population of monomeric particles around 60 Å in length that were soluble up to 8 mg/ml protein concentration. In DDM, CFTR was only soluble below 0.4 mg/ml protein concentration where it existed as a very heterogeneous population of different sized amorphous particles, including dimeric particles around 180 Å in length. The DDM-purified CFTR protein could be crystallised as monomers in two-dimensional (2D) crystals with similar lattice parameters to 2D crystals of CFTR purified from mammalian cells. The ATPase activity of DDM-purified and reconstituted CFTR was similar to already published rates, at around 13 nmol P\textsubscript{i}/min/mg integrated over a reaction time of 60 min, with an apparent affinity K\textsubscript{m} for ATP of 0.14 mM. Such a low ATPase rate compared to other ABC transporters may be due to the observed rapid run-down of activity with time and correlation with published CFTR channel gating kinetics. CFTR showed reduced ATPase activity after purification in LPG, suggesting a structural destabilisation in this detergent. The protocols presented here can now be used to provide sufficient quantities of purified CFTR protein for novel biochemical and biophysical studies. The tendency of CFTR to aggregate in a mild detergent remains a major obstacle towards 3D crystallisation trials and a high-resolution structure.
Declaration

I declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by third parties. Such Intellectual Property and Reproductions cannot and must no be made available for use without the prior written permission of the owner (s) of relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialization of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on presentation of Theses.
Acknowledgements

I would like to thank all those who have given their support during this PhD project: Firstly, to my supervisor, Prof. Bob Ford, for the opportunity to work on such an exciting project, the freedom to develop as an independent researcher and for always having an open door. To the BBSRC, the CFFT and Philip Rann for financial support of this work. To all members of the CFFT Structure Consortium, in particular: Drs. John Hunt and Chi Wang (Colombia University) for their warm welcome, technical support and stimulating discussions during my visit to their lab, Dr. Ina Urbatsch (Texas Tech University) for overall help on yeast work and Dr. Christie Brouillette’s lab (University of Alabama at Birmingham) for insights into CFTR thermal stability. To all members of the Ford lab at the University of Manchester: Dr. Liam O’Ryan, Dr. Naomi Pollock, Tracy Rimington and Xin Meng, for providing strong team-work and a positive atmosphere throughout the CFTR project. To Dr. Mark Rosenberg, for his expertise in electron crystallography and tutoring in 2dx. To Dr. Rich Collins, for his assistance with electron microscopy, tutoring in EMAN and lessons on how to make the perfect coffee. Finally, to Jill Turner, for overwhelming understanding and patience during the time spent writing this thesis.
List of Publications


This author contributed to the publications 1 and 3 by writing sections of the Materials and Methods, generating figures and providing results about expression, purification, and activity of mouse- and chicken-CFTR protein from yeast. This author also contributed to publication 2 by providing samples of purified mouse-CFTR from yeast to a collaborating laboratory (University of Padova, Italy) for analysis of CFTR phosphorylation by mass spectrometry.
Chapter 1 - Introduction

1.1. Structural studies of membrane proteins

Membrane proteins are coded for by approximately 30% of all genes, representing a broad range of biological functions, including transporters, channels, receptors and signal transducers (Wallin and von Heijne, 1998). Furthermore, an estimated half of all medical drug targets are membrane proteins (Arinaminpathy et al., 2009, Terstappen and Reggiani, 2001). Obtaining the high-resolution three-dimensional (3D) structure of membrane proteins can therefore provide important information regarding its biological function (Rosenbaum et al., 2007), but can also aid in the rational design of drugs in the treatment of numerous diseases (Murray and Blundell, 2010). However, despite their significance, membrane protein structures represent only a small proportion of all solved 3D protein structures as compared to soluble proteins (White, 2009, Moraes et al., 2013). Less than 1% of protein structures (total = 77,603) deposited in the Protein Data Bank (PDB) (http://www.pdb.org/pdb/statistics/holdings.do; accessed 1/10/2013), are 3D structures of unique membrane proteins (total = 420) (Database of Membrane Proteins of Known Structure, http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html, Stephen White lab at the UC Irvine: accessed 1/10/2013). X-ray crystallography and electron microscopy are the most common biophysical techniques used to determine membrane protein structure, which are discussed in this section. Other techniques include solid state nuclear magnetic resonance (NMR) (Yao et al., 2013) and small angle X-ray scattering (SAXS) (McDevitt et al., 2008).

1.1.2. Expression and purification of membrane proteins

To understand the specific structural and functional properties of a protein, high quantities of purified material are often needed for in vitro studies. Membrane proteins are not normally expressed in natural sources in sufficient quantities so are over-expressed in heterologous systems through genetic engineering (Wagner et al., 2006). Due to simplicity of growing the cells on a large-scale at relatively low cost, bacteria and yeast are commonly chosen to over-express proteins (Schlegel et al., 2012, Drew et al., 2008). Insect and mammalian cell systems are also employed to express eukaryotic membrane proteins because they can support more complex protein synthesis and function, but growing these cells is much more costly, difficult, and often requires specialised equipment (McDevitt et al., 2006, Andréll and Tate, 2013). Through genetic
engineering, sequences can be introduced to the recombinant protein to allow simple purification by affinity chromatography methods, for example the fusion or insertion of a His-tag (SCHMITT et al., 1993), Strep-tag (Schmidt and Skerra, 2007) or FLAG-tag (Einhauer and Jungbauer, 2001). Membrane proteins are notoriously difficult to purify in comparison to soluble proteins, with common problems including very low expression levels, inefficient purification and poor protein stability (Loll, 2003, Carpenter et al., 2008, Lacapere et al., 2007). Purification protocols tend to require several rounds of trial and error in order to establish optimal conditions.

1.1.3. Detergents

Detergents are amphipathic molecules, consisting of a polar head group and one hydrophobic tail (Seddon et al., 2004). Three classes of detergent: ionic, non-ionic and zwitterionic can be defined. Ionic detergents possess a charged head group that is either positive (anionic) or negative (cationic) and are generally known as “harsh” detergents due to their ability to denature protein folds (e.g sodium dodecyl-sulfate, SDS). Non-ionic detergents have uncharged headgroups and are the most commonly chosen detergents for membrane protein study due to their “mild” nature that typically preserves protein-protein interactions during solubilisation from lipids (e.g. dodecyl-maltoside, DDM). Zwitterionic detergents have charged head groups, but overall no net charge and are generally less harsh than ionic detergents (e.g. lauryldimethylamine-N-oxide, LDAO). Recently, a number of new detergent types have been developed that offer potential improvements in protein stability and crystallisation, for example neopentyl glycols and steroid-based facial (Lee et al., 2013, Chae et al., 2010). The choice of detergent is a critical step in membrane protein study as detergents can greatly affect protein structure and function (Seddon et al., 2004, Privé, 2007). The ideal detergent for a given membrane protein is largely protein-specific and a good detergent for purification may not necessarily be the best choice for later crystallisation or functional assays. The type of detergent used is usually established through experimental trials (McDevitt et al., 2006).

Detergents will spontaneously form micelles in solution, where the head groups form hydrogen bonds with the surrounding aqueous solution and the tails aggregate with each other by hydrophobic interaction (Seddon et al., 2004). The critical micelle concentration (CMC) is a key factor in micelle formation and is defined as the minimum concentration of detergent that is needed for detergent monomers to self-
associate and form micelles. Above the CMC, there is constant exchange of detergent between micelles and free detergent monomer populations. The CMC of a detergent is affected by its length, size and charge, and also by its surrounding buffer conditions (e.g. pH, temperature, ionic strength) (Seddon et al., 2004). During membrane protein extraction, detergents are used at concentrations greatly above (~ 10-fold) their CMC. Later, concentrations are limited to just above the CMC to limit any denaturing effects on the protein during purification, and because high levels of detergent can results in phase separation that may inhibit crystal growth (Arachea et al., 2012, Seddon et al., 2004).

1.1.4. Lipids

Lipids are the building blocks of biological membranes. The most common membrane lipids are phospholipids, which are amphipathic molecules with two hydrophobic, fatty acid tails (12-24 carbons long) and one hydrophilic, phosphate-containing head group. There are a variety of different phospholipids types, including phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylserine (PS) and phosphatidylglycerol (PG), with PC constituting around 50 % of all membrane lipids (Dowhan, 1997). The exact lipid composition of a membrane varies between different organisms and cell types, and varies between the inner and outer membrane leaflets (Vance and Vance, 2002, Dowhan, 1997). The fluidity of a membrane can also vary, which is dependent on the length and saturation of the fatty acid tail, and the presence of non-phospholipids, such as cholesterol and sphingolipids (Vance and Vance, 2002, Simons and van Meer, 1988). Such diversity of lipid properties is necessary to support a vast range of membrane protein function (Wallin and von Heijne, 1998, Dowhan, 1997).

In a biological membrane, lipids are thought to directly surround and fit the contours of a membrane protein forming a tight electrochemical barrier that prevents leakage of solutes (Raunser and Walz, 2009). Lipid-protein interactions are mostly non-specific, including van der Waals interactions between protein transmembrane domains (TMDs) and hydrocarbon lipid chains, and ionic interactions between hydrophilic extramembranous protein regions and the polar lipid head groups (Raunser and Walz, 2009). The lateral pressure, elasticity and deformation of the lipid bilayer can non-specifically support membrane protein structural integrity (Phillips et al., 2009). Lipids can also form specific interactions with membrane proteins, which are believed to support
specific protein structure and function (Hunte and Richers, 2008). For example, the high-resolution structure of bacteriorhodopsin (bR) revealed a lipid buried within the cavity of the protein structure, which formed a salt bridge between its phosphate head group and a lysine residue of bR that was crucial for ion channel activity (Mitsuoka et al., 1999).

During purification, a membrane protein is extracted from the lipid membrane in detergent. In some cases, it may be important that a detergent does not remove any protein-associated lipids that are crucial for membrane protein function and structure (Rosenbusch, 2001). The purified protein can be later reintroduced into a synthetic lipid membrane using a method called reconstitution, which involves carefully removing detergent from a purified protein in the presence of lipid molecules (Rigaud and Levy, 2003). Reconstitution is often used to stimulate protein activity that appears lost in the detergent micelle environment (Geertsma et al., 2008). However, the exact lipid composition surrounding a membrane protein in vivo is rarely known, so the choice of lipids for in vitro study is largely empirical (Serebryany et al., 2012, Pollock et al., 2013, Zehnpfennig et al., 2009).

1.1.5. X-ray Crystallography

X-ray crystallography was first used to solve the atomic-resolution structure of a membrane protein in 1985 (Deisenhofer et al., 1985), and has since remained the method of choice for determining the high-resolution (< 4-Å) structures of membrane proteins (White, 2009). This technique involves collecting structural data from the diffraction patterns of an X-ray beam that has been fired through a 3D protein crystal (Kendrew et al., 1958). 3D crystals consist of many protein molecules regularly packed together in the x, y and z directions. Large and highly ordered 3D crystals are desired to yield the highest resolution structural data (Hunte et al., 2003). Methods for crystallising membrane proteins are largely similar to soluble proteins. The most common method is crystallisation by vapour diffusion of a sitting-drop in the presence of a precipitating agent (e.g. polyethylene glycols, ammonium sulphate), which provides the supersaturated state of the protein required to remove the protein from solution, and in the appropriate conditions, formation of stable nuclei that grow into highly-ordered crystals (Hunte et al., 2003).

Crystallisation still remains the major bottleneck in protein structural studies, where large efforts are needed to explore physicochemical properties, such as protein
constructs, buffers, pH, additives, precipitants, and temperatures. For membrane proteins, the preparation of good-quality diffracting 3D crystals is even more challenging than for soluble proteins (Loll, 2003). One major source of problem is the presence of detergent. Detergents increase the solvent content of a crystal, resulting in crystals that are often fragile, difficult to handle, and susceptible to radiation damage (Lacapere et al., 2007, Carpenter et al., 2008, Loll, 2003). Additionally, the volume occupied by the detergent must be small enough to fit between protein-protein crystal contacts, and this is particularly problematic if the protein has small hydrophilic domains (Privé, 2007). The ideal detergent for membrane protein crystallisation is protein-specific, but to date, DDM has been the most successful detergent in yielding structures (Lacapere et al., 2007, Carpenter et al., 2008). Problems with detergents can be overcome with crystallisation in the presence of lipids, either from bicelles (Faham and Bowie, 2002) or from a lipid cubic phase (Landau and Rosenbusch, 1996). These in meso methods have demonstrated improved crystal stability and have resulted in 47 high-resolution structures of unique membrane proteins (http://cherezov.scripps.edu/structures.htm; accessed on 2/10/13).

Other major difficulties in membrane protein crystallography include lack of sufficient quantities of purified protein, poor stability and heterogeneity (Lacapere et al., 2007, Carpenter et al., 2008, Rosenbusch, 2001). There are vast amounts of literature offering advice for overcoming specific challenges facing membrane protein crystallisation (Ostermeier and Michel, 1997, Carpenter et al., 2008, Loll, 2003, Wiener, 2004, Hunte et al., 2003, Lacapere et al., 2007, Sonoda et al., 2011, Chayen, 2004). There are a number of commercially available crystallisation screen kits specifically tailored to conditions that have previously been successful for membrane protein structure determination (e.g. MemGold, Molecular Dimensions). Advances are also being made in high-throughput screening (Carpenter et al., 2008, Stevens et al., 2001, Abola et al., 2000) and in beam-line technologies (Evans et al., 2011) to enable crystallography studies of such problematic proteins. Still, 3D membrane protein crystals tend to have low-diffraction quality and often do not diffract beyond 5-Å (Carpenter et al., 2008). Therefore, determining the 3D structure of membrane proteins by X-ray crystallography still remains a great experimental challenge.
1.1.6. Electron Microscopy

Electron microscopy (EM) is a technique commonly used to study the structure of biological macromolecules, including membrane proteins (Frank, 2006, Rubinstein, 2007). In transmission EM, an electron beam is fired through a thin sample adhered to a flat, specimen grid and projection images are collected along the optical axis. Electrons hit the nuclei of atoms in the sample and undergo elastic interactions, altering the electron phase and creating phase contrast that makes up a large proportion of the contrast in an image (Amos et al., 1982). A smaller fraction of electrons undergo inelastic scattering and lose energy to the sample. Electrons generated in the high voltage fields of electron microscopes have wavelengths of 0.015- to 0.04-Å, so there is theoretically a resolution limit far exceeding that of X-rays for image acquisition of protein molecules (Chiu et al., 1999). However, in practice, resolution is limited due to image and contrast distortions in the microscope, which can be described by the contrast transfer function (CTF) (Frank, 2006). Additionally, factors including astigmatism, beam drift, radiation damage, beam-induced charging effect and background due to inelastic electrons all contribute to a resolution limit of 0.5 to 0.8- Å by EM (Saibil, 2000). EM data is usually collected at low electron dose and under vacuum to limit beam damage by slowing the formation of light-atom free radicals (Saibil, 2000). The use of field emission guns (FEG) in microscopes can also improve resolution by providing good beam coherence and higher beam acceleration voltage (van Heel et al., 2000), which allows greater penetration and less beam damage.

There are two primary methods of sample preparation and analysis for transmission EM: negative-stain (NS) EM (Brenner and Horne, 1959) and cryo-EM (Adrian et al., 1984). NS-EM involves the use of a heavy metal stain, usually uranyl acetate, to coat the surface of the protein on the grid; the heavy metal atoms in the stain scatter electrons more than the light atoms found in proteins (e.g. C, O, N) resulting in good image contrast, as well as keeping the protein sample stable and resistant to electron beam damage in the microscope. Additionally, NS-EM only requires protein at relatively low concentration (10 – 50 µg/ml), which is achievable for even the most difficult to express membrane proteins (Rubinstein, 2007). However, NS can result in image artefacts and sample dehydration and deformation. Resolution is also limited in NS-EM to around 20-Å because of the granularity of the stain which also tends to provide little detail on the structural details of the protein core from which it is excluded (Wang and Sigworth, 2006).
Cryo-EM involves plunging the sample grid in liquid nitrogen-cooled liquid ethane to embed the protein in a thin layer of vitreous ice. The frozen-hydrated protein sample can then be analysed using a cryo-temperature controlled microscope (Saibil, 2000). Typically, liquid nitrogen (77K) is used to cool the specimen stage of the microscope to about 90-100K, but liquid helium (1.5K) has also been used to collect high-resolution cryo-EM data (Fujiyoshi and Unwin, 2008, Gonen et al., 2005, Raunser and Walz, 2009). Proteins are trapped in a native-like hydrated state, resulting in less image artefacts and higher resolution of structural data (>10-Å) than negative-stain EM (Wang and Sigworth, 2006). However, without NS, cryo-EM poses its own experimental challenges, namely images with lower contrast and increased susceptibility of sample to beam damage. Cryo-EM also usually requires a higher concentration of protein (~1 to 3 mg/ml), which may not be possible for some difficult membrane proteins (Rubinstein, 2007).

### 1.1.7. Electron Crystallography

Electron crystallography provided the first structural information concerning a membrane protein, with the 7-Å resolution 3D structure of bR (Henderson and Unwin, 1975). Electron crystallography has since been used to obtain several low- to medium-resolution structures (> 8-Å resolution) of membrane proteins (Lau and Rubinstein, 2010, Korkhov and Tate, 2008, Tsai and Ziegler, 2005, Beuming and Weinstein, 2005, Rosenberg et al., 2005, Purhonen et al., 2005) and also a few near-atomic structures including, the structure of aquaporin-0, solved to 1.9-Å resolution (Gonen et al., 2005). Electron crystallography therefore has the potential to produce protein structures to a resolution comparable with that of X-ray crystallography. The major advantage of electron crystallography is that much lower quantities and concentrations (~0.5 to 1 mg/ml) of purified protein are needed for 2D crystallisation trials, compared to the high concentrations (~10-15 mg/ml) needed for 3D crystallisation trials and X-ray crystallography (Kim et al., 2010, Vink et al., 2007). This is particularly useful for studying membrane proteins that are typically poorly expressed and unstable at high concentrations.

Electron crystallography involves collecting structural data from images and diffraction patterns of electrons that have been fired through a two-dimensional (2D) protein crystal (Amos et al., 1982). 2D crystals consist of well-ordered protein molecules packed together in a single plane, typically one molecule thick, and are ideally large (~1
µm), with a large number (~10^6) of unit cells to give high-resolution structural data (Raunser and Walz, 2009). The single plane of 2D crystals provides the point of reference for symmetry from which phase residuals can be calculated, eliminating the “phase problem” encountered with X-ray crystallography. Using all possible combinations of crystal symmetry operations (rotation, mirror, slide, screw, roto-inversion), there are a seventeen possible protein packing arrangements in a 2D crystal, known as the 2D plane groups (Landsberg and Hankamer, 2007). A 2D crystal will only provide data from a single plane, by convention the x and y plane. In order to get information in the z plane and thus create a 3D image, electron diffraction patterns must also be collected from a range of specimen tilt angles, relative to the direction of the electron beam (Raunser and Walz, 2009). Electron crystallography data can only be collected up to tilt angles of ~70° due to the obstruction of the support grid, and often only data up to 55° to 60° are collectable. As a result, there is missing information in the z plane, compared to the x and y plane, known as the ‘missing cone’ problem, which can result in loss of features, introduction of artificial densities, distortion and smearing in the 3D protein structure images (Ford and Holzenburg, 2008).

There are two main types of 2D crystals: epitaxial and lipid bilayer crystals. Epitaxial crystals consist of detergent-solubilised membrane proteins and are grown on the surface of a grid in the presence of precipitating agent. Subsequent vapour diffusion at the air/water interface leads to an increase in concentration of protein that, under the right conditions, form stable nuclei for crystal growth (Auer et al., 1999). Lipid crystals consist of membrane proteins embedded within a lipid membrane that are packed together at a high enough density to form protein-protein lattice contacts (Lacapère et al., 1998, Aller and Unger, 2006, Larue et al., 2009, Engel et al., 1992). Such lipidic crystals are formed by the reconstitution of a protein into a lipid membrane in carefully controlled conditions at a low lipid-to-protein ratio (LPR) (Rigaud, 2002). Electron crystallography of 2D lipid crystals offers the advantage of studying membrane protein structure in a more native-like lipid environment, compared to the detergent environment of epitaxial crystals. The precise conditions required for successful 2D crystallisation are protein specific and optimal conditions have to be determined through extensive trials. Unfortunately, unlike 3D crystallisation, there are no commercially available robots for reproducible 2D crystallisation screens. Although, high-throughput screens are constantly being developed (Iacovache et al., 2010, Kim et al., 2010, Vink et al., 2007).
1.1.8. Single particle analysis

If a membrane protein is sufficiently large (> 200 kDa), its structure can be studied by EM immediately after purification without the need for 2D crystal growth, using a method known as single particle analysis (SPA) (Rubinstein, 2007). The membrane protein is adhered to a specimen grid as a detergent-solubilised protein solution and is either subject to NS or cryo-EM analysis. Micrographs are acquired containing images of thousands of protein particles, which are classified according to their angular orientation on grid, averaged, and iteratively aligned both translationally and rotationally. Provided the angles relating the different orientations can be estimated, reconstruction of the 3D structure by reprojection is possible (Rubinstein, 2007).

Membrane protein structures obtained by single particle analysis tend to be limited to low resolution (15- to 40-Å) (Nakagawa et al., 2005, Mio et al., 2010), but there are a few examples of medium resolution structures (~10-Å) (Ludtke et al., 2005, Jiang et al., 2004), and the technique has the potential for higher resolution structures (~4-Å) as has been shown with soluble proteins (Cong et al., 2010). Recently, single particle EM analysis has also provided the opportunity to study membrane proteins within a native-like lipid environment, with the low-resolution (1.7 – 2 nm) structure of reconstituted BK potassium channel (Wang and Sigworth, 2009).

1.2. ATP-binding cassette transporters

ATP-binding cassette (ABC) transporters represent the largest family of membrane proteins in the biological world, with examples found in all organisms ranging from archae to higher eukaryotes (Higgins, 1992). There are 48 ABC proteins identified in humans, which can be classified into ABC subfamilies A-to-G based on phylogenetic analysis of their genetic sequences (Dean and Allikmets, 2001). Typically, these proteins use the energy from ATP hydrolysis to pump substrates across cellular membrane against their concentration gradients, acting as either importers (only found in prokaryotes) or exporters. ABC proteins transport a broad range of substrates, including small ions, amino acids, sugars, drugs and proteins, and are involved in a vast number of biological processes, such as nutrient uptake, signal transduction, protein secretion and antigen presentation (Klein et al., 1999, Higgins, 1992). Mutations in ABC transporters are linked to numerous human diseases, for example cystic fibrosis (Riordan et al., 1989), Stargardt macular dystrophy (Allikmets, 1997) and Tangier disease (Rust et al., 1999). Great medical interest also surrounds the role of ABC
proteins in antibiotic resistance (Kerr et al., 2005) and failure of chemotherapy (Tiwari et al., 2011) in cells.

1.2.1. Structure of ABC transporters

ABC transporters share a common architecture of 4 domains, consisting of 2 transmembrane domains (TMDs) and 2 nucleotide-binding domains (NBDs). These 4 domains are the minimal unit required for transport function but can be ordered in any number of ways (Higgins, 1992). In archae and bacteria, ABC proteins can be expressed in up to 4 separate domains that then oligomerise post-translationally to form a functional unit. Higher eukaryote, tend to have all 4 domains fused together in a single polypeptide chain (Schmitt and Tampé, 2002, Hollenstein et al., 2007, Oldham et al., 2008).

To date, there exist high-resolution X-ray crystal structures for 10 different ABC transporters: ABCB10 (Shintre et al., 2013), BtuCD (Locher et al., 2002), an ECF transporter (Karpowich and Wang, 2013), HI1470 (Pinkett et al., 2007), MalFGK (Oldham et al., 2007), MetNI (Kadaba et al., 2008b), ModBC (Gerber et al., 2008), MsbA (Ward et al., 2007), P-glycoprotein (Pgp) (Aller et al., 2009) and Sav1866 (Dawson and Locher, 2006). There are also a number of low-to-medium resolution structures obtained using EM techniques: MRP1 (Rosenberg et al., 2001), CFTR (Rosenberg et al., 2011) and YvcC (Chami et al., 2002). Some of these structures were obtained in the presence of nucleotide (Dawson and Locher, 2006) and inhibitors as analogues of transport ligands (Aller et al., 2009), but no single ABC protein has yet been crystallised in all the different conformational states that represent a full transport cycle. Isolated NBDs have also been studied by X-ray crystallography (Smith et al., 2002, Hung et al., 1998, Diederichs et al., 2000).

Overall, structural studies of ABC transporters have revealed some general features: The TMDs of ABC proteins usually have 6 α-helical membrane-spanning regions with 3 extracellular loops (ECLs) and 2 intracellular loops (ICLs), but display little homology in primary sequence and can vary greatly in terms of length, number (5-11) and packing of α-helices (Kos and Ford, 2009). Such differences in TMDs are probably due to the vast difference in substrates that ABC proteins transport. The NBDs are hydrophilic L-shaped “lobes” that reside in the cytoplasm. Unlike the TMDs, NBDs are highly conserved at both the primary sequence (> 50 % sequence identity) and at the 3D structural level, suggesting a universal mechanism for ATP hydrolysis in all ABC
proteins (Kos and Ford, 2009, Higgins, 1992). Like many other ATPase proteins, the NBDs have conserved Walker A and Walker B motifs that interact and hydrolyse nucleotides (Walker et al., 1982). There is also a highly conserved signature sequence (LSGGQ) that is unique to the NBDs of ABC proteins (Higgins, 1992). Structural data has confirmed that, in the presence of ATP, the NBDs come together in a closed head-to-tail sandwich dimer conformation (Smith et al., 2002, Dawson and Locher, 2006, Ward et al., 2007). There are two ATP molecules bound at the dimer interface, where each ATP interacts with the Walker A and B motif on one NBD monomer and with the signature sequence on the other NBD monomer. The helix that resides at the joint between the NBDs and the TMDs has a highly conserved structure across the ABC family, suggesting a common mechanism for coupling ATP binding and hydrolysis to substrate transport (Locher et al., 2002). Structures of full-length ABC transporters have also revealed that these proteins must undergo large conformational changes during the transport cycle, with TMDs acting as a hinge for flexible NBD movement; recorded distances between NBDs vary between 4- and 41-Å in a range of apo- and nucleotide-bound structures (Shintre et al., 2013). Such large conformational changes agrees with low-resolution EM data of an ABC protein captured during different stages of the transport cycle (Rosenberg et al., 2003).

1.2.2. Mechanism of ABC transporters

ABC transporters function by an alternating access mechanism, where the substrate binding site in the TMDs alternates between open and closed conformations (Rosenberg et al., 2001). The mechanism that couples ATP hydrolysis to substrate transport remains a matter of debate. One of the first mechanisms proposed was the alternating catalytic cycle model, which suggests ATP hydrolysis is the “power stroke” behind transport and that ATP hydrolysis alternates between the two NBDs to control different stages of the catalytic cycle (Senior et al., 1995). However, biochemical and structural evidence have rather indicated that ATP binding alone provides sufficient energy to drive transport and that the NBDs work together in a single step rather than at different stages during the catalytic cycle. Such data has led to the proposal of a new mechanism, known as the ATP switch model (Figure 1) (Higgins and Linton, 2004). For ABC exporters, the ATP switch model can be described in 4 steps: 1. the transport cycle is initiated by the binding of substrate to its site on the TMDs, which then facilitates ATP-binding at the NBDs and the formation of a closed NBD dimer conformation. 2. ATP-dependent
Figure 1. The ATP switch model for the mechanism of ABC exporters. This schematic shows a single transport cycle for a typical ABC exporter, where TMD1 and TMD2 are represented by the blue rectangles (dark and light blue respectively), NBD1 and NBD2 are represented by the green L-shaped boxes (dark and light green respectively), and the transport substrate is an orange hexagon. The transport cycle can be described in 4 steps: 1. The transport substrate binds to an intracellular site on the TMDs that then increases the NBDs affinity for ATP binding and facilitates the formation of an NBD dimer. 2. ATP binding drives the formation of a closed NBD head-to-tail sandwich dimer conformation that acts as the “power stroke” to switch the TMD from a closed-inward to an open-outward conformation, allowing extracellular transport substrate release. 3. Hydrolysis of ATP molecules causes destabilisation of the NBD dimer. 4. Sequential release of P_i and ADP resets the protein into an open-inward facing conformation, ready to accept another transport ligand molecule. Figure adapted from (Higgins and Linton, 2004).
dimerisation of the NBDs provides the “power stroke” behind transport by inducing changes in the TMDs to release the bound substrate. 3. ATP hydrolysis destabilises the NBD closed dimer. 4. sequential release of inorganic phosphate (P_i) then ADP from the NBDs reset the protein to its ground state.

After recent publication of ABCB10 crystal structures trapped in both apo- and nucleotide-bound states, the ATP switch model could be further adapted to suggest that either ATP or ligand could bind first during step 1 of the transport cycle (Shintre et al., 2013). However, ABC proteins have a high affinity for ATP, with apparent sub-millimolar K_m(ATP) concentrations (Lerner-Marmarosh et al., 1999) that are well below average cellular ATP concentrations (Beis and Newsholme, 1975). ABC protein also have high rates of ATP hydrolysis even in a substrate-free basal state (Lerner-Marmarosh et al., 1999). Therefore, it would seem that if ABC proteins could bind ATP in its ground state, then there would be an inappropriately large amount of ATP hydrolysis without necessarily any substrate transport, which would be a great waste of cellular energy. Perhaps, there are more complex mechanisms that keep the NBDs in a hydrolytically inactive state until substrate transport is required. To date, the ATP switch model provides the most consistent explanation to describe the general mechanisms of the ABC family according to current experimental evidence, but further structural and biochemical data are needed to elucidate the mechanistic details, especially when describing the unique transport for a specific ABC protein.

1.3. CFTR

The cystic fibrosis transmembrane conductance regulator (CFTR) is a member of the ABC transporter family of membrane proteins, and is the only known member that acts as an ion channel, rather than a transporter (Gadsby et al., 2006, Dean and Allikmets, 2001). It is found in the apical membranes of epithelial cells, where it acts as an ion channel (Anderson et al., 1992). Mutations in CFTR give rise to the disease cystic fibrosis (CF), which is one of the most common genetic diseases in the world and for which there remains no cure (Riordan et al., 1989, Rommens et al., 1989, Riordan, 2008). Such biological importance has put CFTR as one of the most heavily researched proteins in biological science. The next sections will review current understanding of CFTR structure, function and progress towards CFTR-targeted treatment for CF.
1.3.1. CFTR Structure

CFTR is a 1480-residue long membrane protein, with the typical ABC transporter architecture of 2 TMDs, and 2 NBDs fused together in a single polypeptide chain (Higgins, 1992). Unique to CFTR is an additional regulatory (R) region, as well as long N- and C-terminal extensions about 80 and 30 residues in length, respectively (Hunt et al., 2013). These domains are arranged in order from N- to C-terminus: TMD1-NBD1-R-TMD2-NBD2 (Figure 2A). From the primary sequence, CFTR is classified into the ABCC subfamily, although there is little evidence to suggest similarity in function to these proteins (Dean and Allikmets, 2001). To date, there is no high-resolution X-ray crystal structure of full-length CFTR, which still remains the “holy grail” of CFTR structural biology. Electron crystallography has provided a CFTR structure to 9-Å resolution (Rosenberg et al., 2011), and SPA has been used to solve a number of low resolution CFTR structures (Mio et al., 2008, Zhang et al., 2009, Zhang et al., 2010). A number of computational atomic models of CFTR structure have been generated, based on homology with Sav1866 or Pgp X-ray crystal structures (Serohijos et al., 2008, Mornon et al., 2008, Rahman et al., 2013). Structures of the isolated soluble domains of CFTR have also been solved, with the X-ray crystal structures of NBD1 (Lewis et al., 2004) and NBD2 (PDB: 3GD7, unpublished) and NMR structure of the R region (Baker et al., 2007).

The greatest detail in current CFTR structural knowledge comes from the X-ray structures of its NBDs. Overall, the NBDs of CFTR have similar structural folds to that of other ABC proteins, including the common Walker A, Walker B and signature motifs (Figure 2C) (Lewis et al., 2004). The NBDs of CFTR also have a few unique features compared to other ABC proteins: Firstly, CFTR has additional ~35 residue regulatory insertion and regulatory extension sequences at the N- and C-terminal end of NBD1, respectively. These sequences are highly mobile and believed to play a role in regulating CFTR channel function (Lewis et al., 2004, Aleksandrov et al., 2010). Secondly, the NBDs of CFTR are asymmetrical, with only 29 % sequence identity between NBD1 and NBD2 (Klein et al., 1999), compared to up to 80 % identity in other ABC proteins (Higgins, 1992). CFTR has unique mutations in the signature sequence on NBD2 (LSGGQ to LSHGH) resulting in only one active ATP hydrolytic site, although both sites still have the ability to tightly bind ATP (Figure 2C) (Aleksandrov et al., 2002b, Lewis et al., 2004, Thibodeau et al., 2005). Such NBD asymmetry is also found with a few other ABCC subfamily proteins, e.g. MRP1 (Hou et al., 2000) and has
been speculated to allow a more complex regulation of transport function. Biochemical and biophysical studies have suggested that the NBDs of CFTR form the same ATP-dependent head-to-tail sandwich dimer conformation as other ABC proteins (Figure 2C) (Vergani et al., 2005, Aleksandrov et al., 2009, Rosenberg et al., 2011, Mense et al., 2006).

The TMDs of CFTR each consist of 6 transmembrane spanning α-helices with 3 ECLs and 2 ICLs (Figure 2B). The TMDs are predicted to form a pore, through which ions can passively transport during channel opening (Gadsby et al., 2006, Riordan, 2008). There are no high-resolution structural data available for the TMDs of CFTR. Comparison of CFTR electron crystallography structures to the X-ray crystal structures of homologous Sav1866 has revealed an additional density between TM helices 3, 6, 9 and 12 that may account for the location of the channel pore (Rosenberg et al., 2011), which correlates to locations of biochemically identified pore lining residues (Smith et al., 2001, Lindsell, 2005) and to the predicted pore location in CFTR structure models (Mornon et al., 2008, Serohijos et al., 2008). The pore appears to have a deep wide vestibule on the intracellular side and a shallower vestibule on the extracellular side.

The 4th ECL in CFTR is glycosylated at residues 894 and 900 with N-linked core and complex sugar moieties in the fully mature and folded proteins (Figure 2B) (Chang et al., 2008). The ICLs are predicted to form part of the typical ABC protein coupling helices that interact and transduce information between TMDs and NBDs (Figure 2B) (Serohijos et al., 2008). The ICL from one TMD appears to interact with both adjacent and opposite NBDs, according to comparisons with structures of homologous ABC exporters (Serohijos et al., 2008).

The isolated R region is a highly dynamic and disordered structure that is mostly random coil with around 5% α-helical secondary structure, according to NMR studies (Figure 2B) (Baker et al., 2007, Ostedgaard et al., 2000). NMR data have also confirmed possible intra-molecular interactions between the R region and the NBDs and C-terminus of CFTR, when the isolated proteins are mixed together in vitro (Baker et al., 2007, Kanelis et al., 2010, Bozoky et al., 2013b). Despite the disorder observed with the isolated protein, EM structures revealed the R region resides in a well-defined location in the full-length CFTR protein, spanning both NBDs on the cytoplasmic side of the membrane, with possible interaction to ICLs (Rosenberg et al., 2011, Zhang et al., 2010). The R region is highly charged and contains 9 consensus sequences for
protein kinase A (PKA) phosphorylation (Cheng et al., 1991, Riordan et al., 1989, Csanády et al., 2005), as well as target sites for other kinases (Chappe et al., 2003, French et al., 1995). Phosphorylation of the R region is important for the regulation of CFTR channel gating function (Gregory et al., 1990, Csanády et al., 2005, Dahan et al., 2001, Seibert et al., 1999). The disordered and dynamic structure of the R region is believed to increase accessibility for phosphorylation and maximize its regulatory function in CFTR (Chong et al., 2013).

EM structures of full-length CFTR show the protein is homologous to X-ray crystal structures of other ABC exporters, in particular Sav1866 (Dawson and Locher, 2006) and Pgp (Aller et al., 2009). Such EM data has also shown that CFTR movement is similar to that of other ABC protein exporters during different stages of the transport cycle; CFTR displays overall changes in helical bundle shape that indicate inward-facing conformations in the apo-state and outward-facing conformations in the phosphorylated and nucleotide-bound state, respectively (Rosenberg et al., 2004, Rosenberg et al., 2011, Zhang et al., 2009, Zhang et al., 2010). The inward- and outward-facing conformations are assumed to represent a closed and open CFTR channel, respectively (Wang and Linsdell, 2012). The NBDs cannot be resolved in these low-resolution EM structures, but it appears that they do not move as far apart as distances seen in other ABC structures (Ward et al., 2007, Aller et al., 2009), probably because CFTR need only provide relatively small conformational changes to allow flux of small ions, compared to the transport of larger substrates.

The quaternary structure of CFTR can also be inferred from current structural data. CFTR contains all 4 ABC domains fused in a single polypeptide chain, so it does not theoretically need to oligomerise to form a functional unit (Higgins, 1992). This seems to be the case with the electron crystallography structure, where functional CFTR protein crystallised as a monomer (Rosenberg et al., 2011). Conversely, single particle analysis revealed a dimeric structure for CFTR (Zhang et al., 2009, Awayn et al., 2005, Mio et al., 2008). It is important to note that these EM structures were obtained using CFTR in a detergent micelle and may not reflect the oligomeric state of the protein in a biological membrane. The quaternary structure of CFTR in vivo still remains a matter of debate, especially as both monomers (Ramjeesingh et al., 2001, Haggie and Verkman, 2008, Chen et al., 2002) and dimers (Ramjeesingh et al., 2001, Eskandari et al., 1998, Schillers et al., 2004) have been detected for CFTR that was integrated into a lipid membrane environment.
Figure 2. Schematic of CFTR structure
A. Domain organisation of CFTR. CFTR domains are arranged from N-terminus: TMD1-NBD1-R-TMD2-NBD2. Residue numbers within the amino acid primary sequence are numbered above their predicted domain location. B. CFTR 3D structure. The predicted structure of CFTR is shown integrated into a lipid bilayer (parallel grey lines), with the extracellular side at the top and the cytoplasmic side at the bottom of the diagram. Each TMD consists of 6 transmembrane α-helical regions, shown as blue rectangles, with 3 connecting ECLs, and 2 ICLs. The 4th ECL in TMD2 is glycosylated at two residues N894 and N900, shown as branching lines. CFTR is shown in an outward-facing conformation, similar to that seen with the X-ray crystal structure of Sav1866, and is assumed to be in the open channel state. CFTR has domain-swapping architecture, where the ICLs of one TMD interact with both NBDs. The red asterix indicates the position of amino acid residue F508 in NBD1 and its predicted interaction with the 4th ICL from TMD2. C. Interaction between the NBDs and ATP. Two ATP molecules bind at the interface between NBD1 (dark green) and NBD2 (light green) in what is known as the ATP-dependent closed head-to-tail sandwich dimer conformation. Each ATP binding site involves the Walker A and B motifs from one NBD and the ABC signature sequence (LSGGQ) from the other NBD. The top ATP binding site is catalytically inactive due to non-conserved mutations in the signature sequence, which are highlighted in red. The second ATP-binding site at the bottom of the diagram is capable of hydrolysing ATP to ADP and P₆. Figure adapted from (Hunt et al., 2013) and (Gadsby et al., 2006).
1.3.2. CFTR function and mechanism

Despite its structural similarities, CFTR is functionally unique among ABC proteins as the only known ion channel. CFTR forms a channel in the apical membranes of epithelial cells where it plays a key role in secretion of fluid and electrolytes at epithelial surfaces (Frizzell, 1999, Anderson et al., 1992). CFTR is predominately a chloride channel, although its pore is not very selective and allows a number of other small anions to permeate (Linsdell, 2001). In summary, channels on the basolateral membrane pump chloride ions into epithelial cells to a relatively high cellular concentration so that when CFTR is activated, chloride ions (\(\text{Cl}^-\)) move passively down their electrochemical gradient to the epithelial surface. This in turns causes the epithelial salt secretion that drives water into mucus. CFTR has also been shown to act as a \(\text{HCO}_3^-\) channel (Quinton, 2008) and to regulate other \(\text{Cl}^-/\text{HCO}_3^-\) exchangers (Choi et al., 2001), both of which are important in maintaining healthy mucus. Other roles of CFTR include interactions with the amiloride-sensitive sodium channel (ENaC), a protein involved in sodium secretion and hydration of mucus (Stutts et al., 1995). The CFTR channel is characterized by small single-channel conductance (6-10 ps) and a linear current-voltage (I-V) relationship (Sheppard and Welsh, 1999). Single-channel recordings also show that CFTR channel activity has two phases: a burst interval, during which there is rapid flickering with a closed state, and secondly, a phase when the channel is closed (Winter et al., 1994). Such channel activity can be described by 3 conformational states for CFTR: open, closed and open-ready. The open-ready conformation does not allow chloride flux, but is poised to rapidly transition into the open state. There are likely a number of other discrete conformations during a single gating cycle (Gunderson and Kopito, 1995). CFTR channel activity is also dependent on PKA phosphorylation of its R region and the presence of intracellular ATP (Sheppard and Welsh, 1999, Berger et al., 1991). The mechanisms of CFTR regulation by ATP and phosphorylation are discussed:

Unlike other ABC proteins, CFTR allows passive flux of ions through its pore rather than actively pumping substrates against their electrochemical gradient. CFTR, therefore, does not theoretically require the energy of ATP hydrolysis for its function as an ion channel. This was confirmed by experimental data that showed ATP binding alone provided sufficient energy to open the CFTR channel (Aleksandrov et al., 2000). The current model for the role of ATP in CFTR channel function is that ATP binding opens the channel, and that ATP binding and subsequent release of ADP and P\(_i\) closes
Figure 3. Schematic of CFTR molecular mechanism
A cartoon of CFTR 3D structure is shown integrated into the lipid bilayer (grey parallel lines) with the extracellular side at the top and the cytosolic side underneath. CFTR domains are represented by a dark blue rectangle for TMD1, a light blue rectangle for TMD2, dark green L-shaped box for NBD1, light green L-shaped box for NBD2 and a purple oval for the R region. CFTR channel starts (top left) in the closed state with the NBDs apart in the open dimer conformation. Phosphorylation (red circle) of the R region by PKA disrupts an interaction between the R region and NBD1 that allows for ATP to bind at the NBDs. ATP-binding induces the formation of a closed NBD head-to-tail sandwich dimer (also shown in Figure 2C), leaving the channel in an open-ready transition state conformation, where the channel gate is still closed but poised ready to open (top right). The ATP-induced dimerisation of NBDs causes a conformational change in the TMD region that switches the protein from a closed-inward to an open-outward facing conformation (bottom right). The channel opens and chloride flux occurs across the plasma membrane. Hydrolysis of ATP at the one catalytically active site in the NBDs causes destabilisation of the NBD dimer. Release of P_i and ADP resets the protein into an inward-facing conformation and closes the channel gate. ATP may remain bound to the catalytically inactive ATP-binding site allowing for rapid flickering back to the open-ready state after the addition of another ATP molecule. Alternatively, ATP is released and the CFTR is restored to the closed state. Figure is adapted from (Gadsby et al., 2006).
the channel (Figure 3) (Gadsby et al., 2006, Aleksandrov et al., 2007). In this model, the binding of 2 ATP molecules to the NBDs of phosphorylated CFTR results in the formation of a NBD head-to-tail sandwiched heterodimer (Figure 2C and 3). Such dimerisation induces conformational changes in the TMDs via NBD-TMD coupling helices that switch the protein to an open-outward conformation, and by some unknown mechanism opens the channel gate. The hydrolysis and release of ADP and P$_i$ at the catalytically active site then closes the channel gate and resets the protein into a closed-inward conformation. Evidence for this model include the fact that the duration of CFTR gating cycle (~1 sec) roughly equates to the rate of ATP turnover (Li et al., 1996), and that non-hydrolysable AMP-PNP locks CFTR in open state (Gunderson and Kopito, 1995, Hwang et al., 1994). However, CFTR channel gating can also occur in the absence of ATP and in constructs where NBD2 has been removed, so it appears channel activity can also occur using a non-hydrolytic pathway, although at a much slower rate (Aleksandrov and Riordan, 1998, Ikuma and Welsh, 2000, Aleksandrov et al., 2000, Bompadre et al., 2005). The precise relationship between ATP and CFTR channel gating is still not fully understood, mostly due to the limited number of enzymatic studies on the purified protein, and in these experiments there is still doubt whether the activity of purified protein accurately reflects CFTR behaviour in vivo (Bear et al., 1992, Li et al., 1996, Ramjeesingh et al., 1999a, Aleksandrov et al., 2002a).

CFTR channel opening is dependent on PKA phosphorylation of its R region (Gregory et al., 1990, Csanády et al., 2005, Dahan et al., 2001, Seibert et al., 1999, Dulhanty and Riordan, 1994). The R region is normally kept in its un-phosphorylated state by phosphatases in the cell, when it has an inhibitory role on CFTR activity (Seibert et al., 1999, Csanády et al., 2000); deletion of the R region results in a CFTR channel that is permanently active (Rich et al., 1991). This inhibitory effect is important for CFTR regulation because at cellular concentrations of ATP, which are higher than the apparent sub-millimolar K$_m$ concentration of CFTR for ATP, the channel would otherwise be constantly switched on (Beis and Newsholme, 1975, Li et al., 1996). The R region contains 9 PKA consensus sites, most of which are activating (S660, S670, S700, S712, S753, S795, S813) and two are inhibitory (S737, S768) (Wilkinson et al., 1997). The up-regulation of PKA by cAMP-dependent mechanisms causes phosphorylation of these sites in a specific order, although it is unclear why an order is needed (Wilkinson et al., 1997). No specific phosphostructure is required for channel activation, but activity increases with increasing levels of phosphorylation in a mostly additive manner.
Phosphorylation of the R region reduces its helicity and weakens its interaction with NBD1 (Baker et al., 2007). This is believed to enable ATP binding at NBD1 and allows initiation of the gating cycle (Wilkinson et al., 1997, Winter and Welsh, 1997). CFTR activity can also be regulated by phosphorylation at the R region by other kinases e.g. protein kinase C (PKC) (Chappe et al., 2003). In fact, CFTR activity is likely regulated by a large number other proteins, including PDZ-interacting proteins (Wang et al., 1998b, Wang et al., 2000) and STAS domain interactors (Ko et al., 2004).

### 1.3.3. CFTR biogenesis

During biogenesis, proteins are tightly regulated by cellular quality control systems to prevent accumulation of misfolded and incorrectly processed proteins that are potentially toxic to the cell (Brodsky and Skach, 2011). CFTR is a large, multi-domain and polytopic eukaryotic membrane protein. Such complex proteins are more scrutinised by these checkpoints due to greater chance of misfolding during synthesis. CFTR, however, appears to be especially scrutinised during biogenesis, even compared to other similarly complex membrane proteins; only 30 % of all CFTR nascent polypeptide chains are converted into fully mature protein at the plasma membrane (PM), compared to other ABC proteins (e.g. MRP1 and Pgp) that are processed with nearly 100 % efficiency (Loo et al., 1998). The reasons for such low efficiency of CFTR synthesis still remain unknown, but the mechanisms by which this occurs are becoming more understood, as described in the following:

The synthesis of CFTR begins at the ribosome, where nascent polypeptide chains are co-translationally inserted into the endoplasmic reticulum (ER) membrane by the Sec61 complex translocon (Lu et al., 1998). CFTR folding has relatively slow kinetics, probably to allow time for correct assembly of multiple protein domains (Oberdorf et al., 2005). CFTR folding and domain assembly is cooperative and mostly co-translational, where each domain folds separately and then forms inter-domain interactions with previously translated domain as they are synthesised (Kleizen et al., 2005). Interestingly, CFTR domains have the ability to assemble post-translationally if they are expressed in truncated or separate constructs (Ostedgaard et al., 1997, Sheppard et al., 1994), which is probably an evolutionary relic of other ABC proteins that form multimers from separate polypeptide chains (Locher et al., 2002). Both inter- and intra-domain folding is required for ER exit (Younger et al., 2006), although CFTR can still
exit the ER with only a minimal folded unit of TMD1-NBD1-R-TMD2 (Cui et al., 2007).

Similarly for other membrane and secreted proteins, the ER provides the first steps in CFTR folding quality control. Molecular chaperones, such as calnexin in the ER and Hsp40/70/90 in the cytoplasm, assess the folded state of CFTR during synthesis and target misfolded protein for ER associated degradation (ERAD) by the 26S proteosome (Ward et al., 1995, Jensen et al., 1995, Wang et al., 2006, Yang et al., 1993, Pind et al., 1994). CFTR is highly degraded by the ERAD pathway, with the majority of nascent chains being targeted by ubiquitination early in synthesis during translation (Ward et al., 1995, Jensen et al., 1995). A small percentage of expressed CFTR exits the ER and enters the Golgi via COPII vesicles for final maturation, after which it is enters post-Golgi vesicles and remains stable for 16-24 hours before trafficking to the PM via the endosome network (Gentzsch et al., 2004). CFTR can also enter the post-Golgi network via a non-conventional path that involves tubular structures dependent on syntaxin (Yoo et al., 2002).

The final stages of CFTR processing occur at the PM. CFTR is constantly recycled between the PM and endosomal compartments, where it is often targeted to late endosomal pathways and degradation in lysosomes (Gentzsch et al., 2004). This recycling results in a short half-life for CFTR at the PM, with a 10 % loss every minute (Prince et al., 1994). CFTR folding at the PM is monitored within a large interactome of other proteins, including members of Rab and Rho family of GTPase and cytoskeleton proteins, such as the PDZ-interacting NHERF protein (Okiyoneda and Lukacs, 2007, Farinha et al., 2013). Such a level of recycling is believed to maintain an active pool of CFTR at the PM, which assumes there are some mechanisms that cause a decrease in CFTR activity at the PM in a relatively short time frame (Farinha et al., 2013).

During protein synthesis, CFTR receives post-translational modifications (PTMs) that are important for folding, stability, and biological activity, including glycosylation, palmitoylation, methylation and phosphorylation (Chang et al., 2008, McClure et al., 2012, Seibert et al., 1999, Csanády et al., 2005, Dulhanty and Riordan, 1994). CFTR receives core N-linked glycosylation at residues N894 and N900 in the 4th ECL during translation in the ER and complex glycosylation at the Golgi (Cheng et al., 1990). Core glycosylation is recognised in the ER by the chaperone calnexin, which is part of the checkpoint that allows CFTR to move to the Golgi. Complete glycosylation is therefore
a good indication for fully folded and correctly processed CFTR, although it is not crucial for PM localisation or channel function (Gregory et al., 1990, Morris et al., 1993), but may be required to maintain stability at PM (Lukacs et al., 1993, Wei et al., 1996). CFTR is also kept in a mostly un-phosphorylated state during biogenesis by phosphatases for reasons mentioned earlier (section 1.3.2). Lack of phosphorylation may also prove a check-point for correct CFTR trafficking to the PM (Lukacs et al., 1997).

1.3.4. Cystic fibrosis

Cystic fibrosis (CF) is the most common genetic disease in Caucasian populations, affecting approximately 1 in 2500 live births (Cutting, 2005). The major pathology of CF is the accumulation of dehydrated mucus at the epithelial of organs including the lungs (McCarty, 2000), pancreas (Gray et al., 1995), gut (Vankeerberghen et al., 2002) and testes (Kaplan et al., 1968), which results in blockages, infection, inflammation and ultimately organ failure. CF patients usually die by their 30s, mostly due to chronic lung infection (e.g. *Pseudomonas aeruginosa* and *Burkholderia cepaci*) and inflammation (http://www.cff.org, accessed 3/10/13). The fundamental cause of CF is the mutation of *CFTR* alleles, which cause a decrease in CFTR chloride channel function and a resulting lack of ionic and water homeostasis at epithelial surfaces (Riordan et al., 1989, Frizzell, 1999). The lack of CFTR-dependent water homeostasis can clearly explain mucus dehydration in these effected organs. However, it is difficult to link the large number of CF symptoms directly to the defects in the primary function of CFTR, which has led some to believe that CFTR has other additional functions that are defective in CF, for example, CFTR as a multi-kinase recruiter in cell metabolism pathways (Mehta, 2005), or CFTR as a direct or indirect transporter of HCO$_3^-$ for the correct processing of mucins (Quinton, 2008). CFTR has also gained medical interest in diseases such as polycystic kidney disease (Sullivan et al., 1998) and secretory diarrhea (Gabriel et al., 1994), where CFTR activity levels are abnormally high.

1.4.1. CFTR mutations

To date, there are over 1900 different CFTR mutations that are known to cause CF, but > 90% of cases are caused by a recessive deletion mutation of phenylalanine at position 508 (ΔF508) (Cystic Fibrosis Mutations Database, available at http://www.genet.sickkids.on.ca/cftr/; accessed on 8/10/13). All CF patients are homozygous for CFTR mutations, whereas heterozygotes display normal organ function.
CFTR mutations are functionally classified into 5 groups (Prickett and Jain, 2013): Class 1 mutations cause a defect in CFTR protein synthesis, such as the premature stop codon W1282X, that result in little or no CFTR at the PM. Class 2 mutations, including the common ΔF508, are translated into full-length nascent polypeptide chains but are defective in folding and are thus targeted for degradation rather than trafficked to the PM. Class 3 mutants of CFTR are able to reach the PM but have channel gating defects that decrease channel opening time and decrease chloride flux, e.g. the second most common mutation G551D. Class 4 CFTR mutants reach the PM, but with decreased channel conductance. Class 5 represent a fully functional CFTR at the PM but with reduced abundance due to defective mRNA splicing. Class 1-3 cause severe disease phenotypes, whereas class 4 and 5 are mild-disease causing mutations. It should be noted that some CFTR mutations have more than one phenotype, for example the ΔF508 mutation has reduced channel activity and shorter PM half-life in addition to processing defects (Hwang et al., 1997).

As the vast majority of patients harbour ΔF508-CFTR, it is this mutation that has received the most attention in CF research. The ΔF508 mutation is proposed to destabilise CFTR structure and folding (Lewis et al., 2005, Serohijos et al., 2008, Thibodeau et al., 2005). During protein synthesis, most of the expressed ΔF508-CFTR is targeted by the ERAD pathway for degradation by the proteasome, with only 1 % of translated protein reaching the PM (Ward et al., 1995, Kopito, 1999). The small amount of ΔF508-CFTR that reaches the PM is, in fact, a functional chloride channel, although operating with slightly altered gating kinetic, namely with increased residency in the closed channel state (Cui et al., 2006). Such levels of misfolding would assume gross changes in overall ΔF508-CFTR structure, which was confirmed by its increase susceptibility to limited proteolysis compared to the wild-type protein (Zhang et al., 1998). Surprisingly, however, the X-ray crystal structures of isolated ΔF508-NBD1 protein showed little structural differences compared to its wild-type counterpart, other than small, local perturbations in the vicinity of position 508; there are significant changes only in the backbone at residues 509-511, with the normally buried V510 side chain flipped to a surface exposed position (Lewis et al., 2005, Lewis et al., 2010). This has led to the hypothesis that F508 is important in inter-domain folding and assembly, rather than the folding of NBD1 itself (Lewis et al., 2005, Lewis et al., 2010, Serohijos et al., 2008, Thibodeau et al., 2005, Loo et al., 2010, Rabeh et al., 2012). The F508 residue lies in a shallow groove on the surface of the NBD1 in an area that, in X-ray
crystal structures of other ABC proteins, forms crucial interactions with the coupling helices of the TMD regions (Dawson and Locher, 2006, Aller et al., 2009). CFTR homology models and numerous other biochemical and biophysical evidence have suggested that this F508-mediated interaction in CFTR is between NBD1 and ICL4 of TMD2 (Figure 2B) (Cui et al., 2007, Serohijos et al., 2008, Thibodeau et al., 2005, Lewis et al., 2010, Loo et al., 2010, Rabeh et al., 2012). It is suggested that the peptide backbone at position 508 is important for CFTR folding and the phenyalanine side chain is necessary for inter-domain contacts (Thibodeau et al., 2005).

As well as the destabilisation of NBD1-TMD2 interactions, ΔF508-NBD1 itself appears to be thermally destabilised; isolated ΔF508-NBD1 protein has a thermal unfolding transition about 6 ºC lower than its wild-type counterpart (Protasevich et al., 2010). The ΔF508 mutation is believed to thermodynamically favour the formation of molten-globule state that is prone to aggressive and irreversible aggregation (Wang et al., 2010). The thermal destabilisation effect of the ΔF508 mutation is also apparent in vivo because low temperatures can rescue processing defects and restore ΔF508-CFTR to the PM (Denning et al., 1992). Molecules that can bind and thermostabilise ΔF508-NBD1 in vitro have also been effective in correcting full-length ΔF508-CFTR in cells (Sampson et al., 2011). However, there is conflicting belief as to whether rescuing ΔF508 mutant can be achieved by thermodynamically stabilising NBD1 alone (Aleksandrov et al., 2012), or whether the correction of NBD1-TMD2 interactions is also required (Rabeh et al., 2012). Overall, it appears that the ΔF508-CFTR molecular defect in CF is a combination of decrease in NBD1 thermodynamic stabilisation and destabilisation of NBD1-TMD2 inter-domain interactions.

1.4.2. CF therapy
Currently, the majority of CF therapy targets the symptoms of the disease, with focus on antibiotics, anti-inflammatories, mucus viscosity modulators and nutrient supplements (Hoffman and Ramsey, 2013). Since the cloning of the CFTR gene in 1989, efforts have been ongoing to find a therapeutic that targets the underlying cause of CF, by increasing levels of functional CFTR in patients (Hoffman and Ramsey, 2013). Gene therapy involving the introduction of CFTR cDNA to CF affected cells, has had huge interest, but with little evidence for success, mostly due to problems with cDNA vectors (Prickett and Jain, 2013). There are a number of conditions that can restore CFTR mutants in cell lines, but would prove an impractical solution therapeutically, e.g. low
temperature rescue of ΔF508-CFTR (Denning et al., 1992). Great efforts are now being made towards development of small molecule modulators of CFTR that can increase levels of functional CFTR in patients (Hoffman and Ramsey, 2013, Riordan, 2008). Such drug discovery is currently based on high throughput compound screening. For example, a high through-put screen developed by the Verkman lab (Verkman, 1990) that uses a CFTR-mediated halide flux assays and a halide-sensitive fluorescent dye has identified a number of compounds that activate (Ma et al., 2002b) and inhibit (Ma et al., 2002a) CFTR. Compound screening is also on-going in the pharmaceutical industry (PTC Therapeutics, Vertex Pharmaceuticals).

There are several classes of small-molecules that can rescue CFTR disease phenotypes, for example, potentiators and correctors. Potentiators are compounds that can stimulate channel activity of a CFTR molecule that has already reached the PM, for example the recently FDA-approved drug Kalydeco (or Ivacaftor), which stimulates the activity of the otherwise inactive G551D-CFTR mutant (http://www.cff.org/research/drugdevelopmentpipeline; accessed 8/10/13). Kalydeco has proven very effective in restoring healthy CFTR phenotype in clinical trials and has caused great excitement in the CF research field as it is the first treatment to specifically target the underlying cause in CF. However, only 4 % of all CF patients carry the G551D mutation, so efforts are continuing to find a similarly successful corrector compound for the more common ΔF508-CFTR mutation. Correctors are molecules that target class 2 mutations in folding and processing defects, to restore CFTR levels at the PM (Wellhauser et al., 2009). Class 2 CFTR mutants, including ΔF508-CFTR, are capable of functioning as a chloride ion channel if they are able to reach the PM, although with slightly slower gating kinetics (Yang et al., 2003). Therefore, correctors that can restore CFTR expression at the PM might be sufficient to alleviate CF pathology. A number of correctors are in various stages of clinical trials e.g. VX-809 (Van Goor et al., 2006). Early evidence, however, suggests that VX-809 has low efficacy (<15 % rescue) for rescuing the ΔF508 phenotype (Van Goor et al., 2006). It is now being trialled in combination with the potentiator Kalydeco as it might be that both folding and channel activity defects need to be restored for the ΔF508 mutation. There are also a number of other drugs in clinical trials that target proteins other than CFTR involved in epithelial homeostasis, for example activators of the calcium-activated chloride channel (Deterding et al., 2005) and inhibitors of ENaC (Hirsh et al., 2004).
1.5 Strategy of project

1.5.1. CFFT Structure consortium

A high-resolution 3D structure of CFTR would greatly advance current understanding of CFTR molecular mechanisms. Additionally, an atomic-resolution structure would allow the design of specifically CFTR-targeted drugs to treat CF, rather than the current approach of random compound screening. One of the major obstacles in CFTR structural biology is the lack of pure protein in sufficient quantities for X-ray crystallography studies. Polytopic membrane proteins are notoriously difficult to purify due to poor solubility and stability that comes from having large hydrophobic regions. CFTR, however, appears to be particularly difficult, even for a membrane protein. The large amount of CFTR degradation in the cell results in little protein at the PM from which to extract and purify (Loo et al., 1998). Furthermore, purified CFTR is particularly unstable and prone to aggregation, especially at the high protein concentrations required for 3D crystal trials (Zhang et al., 2010). Currently, functional CFTR can be purified in low amounts from baby hamster kidney (BHK) mammalian cells (Rosenberg et al., 2004), Sf9 insect cells (Eckford et al., 2012) and yeast (Huang et al., 1996), which has been mostly used for functional assays and low-resolution structural studies.

In response to this challenge, the Cystic Fibrosis Foundation Therapeutics (CFFT) has assembled a 3D structure consortium as a collaborative effort to improve CFTR study. The aim of the consortium is to obtain milligrams quantities of pure full-length protein that is suitable for high-resolution crystallography, and for other novel biochemical and biophysical studies. The consortium includes a number of labs that are currently testing different cellular systems to express CFTR, including bacterial cells Escherichia coli and Lactococcus lactis (Slotboom lab, University of Groningen), yeast cells Saccharomyces cerevisiae (Ford lab, University of Manchester) and Pichia pastoris (Urbatsch lab, Texas Tech), mammalian BHK cells (Riordan lab, University of North Carolina) and human embryonic kidney (HEK) cells (Kappes lab, University of Alabama at Birmingham). This work is also supported by labs that specialise in biophysical techniques that characterise the quality of CFTR for X-ray crystallography experiments, such as activity, thermodynamic stability and homogeneity (Riordan lab, University of North Carolina; Hunt lab, Colombia University; Brouillette lab, University of Alabama at Birmingham; Ford lab, University of Manchester).
In addition to expression systems, a number of different CFTR orthologues are also being screened, which are listed in Table 1. Orthologues can be defined as proteins that are found in different organisms, but are coded for by genes derived from common ancestors, often sharing similar structural and functional properties (Fitch, 2000). A total of 13 CFTR orthologues were selected, each one representing a different evolutionary branch of the protein, or examples from well-studied CF animal model e.g. pig, ferret and mouse (Keiser and Engelhardt, 2011). The CFTR amino acid sequence is highly conserved across the 13 orthologues, with the least conserved salmon- and killifish-CFTR still having 58% sequence identity (Table 1). From single-channel recording experiments, all of these CFTR-orthologues have similar ion channel functionality as the human version of the protein, although with slightly altered gating kinetics (Riordan lab, unpublished data). CFTR orthologues also represent a full-length wild-type protein, which should prove more biologically functional than introducing non-native point mutations into the human-CFTR sequence, as is done for many crystallisation projects (Warne et al., 2008). Ultimately, study of the human-CFTR orthologue holds the greatest medical interest, but in its absence, the atomic-resolution structure of other CFTR orthologues should still prove an important breakthrough in CFTR and CF research.

Differences in amino acid sequence are often exploited for 3D crystal trials because they can produce a protein that is more amenable to good-quality, diffracting crystals (Andrykovitch et al., 2003). In fact, several high-resolution structures of eukaryotic membrane proteins have been obtained using orthologues other than human, for example the mouse Pgp (Aller et al. 2009), the rabbit SERCA1a (Jidenko et al., 2005, Toyoshima et al., 2013), the chicken potassium channel Kir2.2 (Tao et al. 2009) and the chicken acid-sensing ion channel ASIC1 (Jasti et al., 2007). Overall, the consortium aims to screen different expression systems and CFTR orthologues to isolate a protein that is more favourable to over-expression, purification and 3D crystallisation, which has not been previously possible.
<table>
<thead>
<tr>
<th>CFTR orthologue</th>
<th>Amino acid sequence identity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>100</td>
</tr>
<tr>
<td>Pig</td>
<td>92</td>
</tr>
<tr>
<td>Ferret</td>
<td>92</td>
</tr>
<tr>
<td>Rabbit</td>
<td>92</td>
</tr>
<tr>
<td>Sheep</td>
<td>91</td>
</tr>
<tr>
<td>Dog</td>
<td>90</td>
</tr>
<tr>
<td>Possum</td>
<td>86</td>
</tr>
<tr>
<td>Platypus</td>
<td>83</td>
</tr>
<tr>
<td>Chicken</td>
<td>80</td>
</tr>
<tr>
<td>Mouse</td>
<td>78</td>
</tr>
<tr>
<td>Rat</td>
<td>78</td>
</tr>
<tr>
<td>Shark</td>
<td>71</td>
</tr>
<tr>
<td>Salmon</td>
<td>58</td>
</tr>
<tr>
<td>Killifish</td>
<td>58</td>
</tr>
</tbody>
</table>

Table 1. List of CFTR orthologue selected for screening within the CFFT Structure Consortium project.
CFTR protein sequences for the 13 listed orthologues were obtained the NCBI protein database (http://www.ncbi.nlm.nih.gov/protein; accessed 3/10/13) and compared in the NCBI BLASTp programme (http://blast.ncbi.nlm.nih.gov/Blast.cgi; accessed 3/10/13). The amino acid sequence identity is shown as a percentage relative to the human-CFTR sequence (100%).
1.5.2. Yeast expression system

Yeast are commonly used as a heterologous over-expression system for recombinant proteins (Cereghino and Cregg, 1999, Cregg et al., 2000, Hays et al., 2010, Newstead et al., 2007a). Yeast provide similar advantages to bacteria over mammalian and insect cell expression systems, including low cost, faster growth conditions, larger biomass production and relatively simple molecular biology step. Such molecular biology steps generally involve the transformation of an expression plasmid that becomes stably integrated into the genome (Cregg et al., 2000, Gietz and Schiestl, 2007). Compared to bacteria, yeast have added machinery to correctly process higher eukaryotic membrane proteins, such as enzymes for glycosylation and disulphide bond formation (Eckart and Bussineau, 1996). It is important to note, however, that yeast lack some of the protein processing machinery and membrane lipid components needed to express a fully functional eukaryotic membrane protein (Eckart and Bussineau, 1996). For example, yeast cannot perform some types of complex glycosylation and are unable to synthesise cholesterol, which can compromise membrane protein function and stability (Gerngross, 2004, Hanson et al., 2008). Although, there are increasing numbers of “humanised” yeast strains that have been genetically engineered to overcome such problems (Kitson et al., 2011, Hamilton et al., 2006). The most commonly chosen yeast strains to over-express membrane proteins are *P. pastoris* and *S. cerevisiae*, which have both provided a number of high-resolution X-ray crystal structures of eukaryotic membrane proteins (Aller et al., 2009, Pedersen et al., 2007). *P. pastoris* are often chosen for their ability to grow to exceptionally high cell density; in minimal media, *P. pastoris* cells can grow to a yield of 200g/L, compared to a 20g/L yield for *S. cerevisiae* cells (Cregg et al., 2000). *S. cerevisiae*, on the other hand, have more widely understood genetics and offer greater availability of different strains and expression vectors (Darby et al., 2012). Previously, CFTR has been expressed in small-quantities by *S. cerevisiae* cells, which was shown to be functional by ion flux and single-channel recording assays of purified and reconstituted protein (Huang et al., 1996, Huang et al., 1998).

Currently, our group is working with the yeast, *S. cerevisiae*, following a protocol developed by David Drew’s lab (Imperial College London) that has been used to express, purify and crystallise a number of eukaryotic membrane proteins (Drew et al., 2008). The strain of yeast used is FGY217, which harbours a *pep4* deletion that prevents the expression of the lysosomal endopeptidase Pep4p (Woolford et al., 1986). Deletion of Pep4p has been shown to limit degradation and improve expression of a
number of membrane proteins in *S. cerevisiae* (Newstead et al., 2007a). This yeast system also uses a C-terminal green fluorescent protein (GFP) fusion tag (see Figure 4 in Materials and Methods) (Drew et al., 2008). GFP is a small (27 kDa), 11-stranded β-barrel soluble protein, which contains a fluorophore within its central cavity that emits green light when excited by blue light (Yang et al., 1996). The GFP used in our yeast cells is the commonly used “enhanced” GFP, which contains mutations to improve folding efficiency (F64L) and increase brightness (S65T) (Tsien, 1998). The GFP-fusion tag provides a fast and easy means to monitor expression and purification trials, using methods including in-gel fluorescence and whole-cell fluorescence microscopy (Drew et al., 2008). The high sensitivity of GFP fluorescence makes it ideal for monitoring such a poorly expressed protein like CFTR and for use in small-scale, high-throughput screens of multiple CFTR orthologues. Additionally, GFP fluorescence is dependent on correct protein folding and is supposed to act as a marker to show that the upstream protein is integrated into the membrane correctly and thus potentially functionally and structurally intact (Drew et al., 2008, Drew et al., 2001).

### 1.5.3. CFTR construct

In addition to the C-terminal GFP-tag, all of our CFTR orthologue constructs have the same N- and C-terminal histidine tags, allowing for purification by nickel affinity chromatography (SCHMITT et al., 1993), as well as an N-terminal SUMO-fusion tag (see Figure 4 in Materials and Methods). SUMO (small ubiquitin-related modifier) is a small, 11 kDa protein that is covalently added to proteins in eukaryotic cells as a means of PTMs that regulates protein stability, activity and cell cycle progression (Hay 2005). SUMO-star is based on the SUMO protein found in *S. cerevisiae*, Smt3, but engineered to be more resistant to proteolysis (Peroutka Iii et al. 2011). The addition of an N-terminal SUMO-tag has been shown to increase solubility, stability and decrease degradation of recombinantly expressed proteins (Marblestone et al. 2006, Zuo et al. 2005). Initial evidence also indicates that the SUMO-tag increases CFTR expression levels and trafficking to the PM in yeast cells (T. Rimington, I. Urbatsch, unpublished data).

This thesis focuses on the expression, purification and characterisation of the chicken-CFTR orthologue. A single orthologue was selected because a full screen of all 13 orthologues was beyond the scope of this PhD project. Recent work has shown the chicken-CFTR orthologue is more thermodynamically stable than the human-CFTR
protein (Aleksandrov et al., 2012), possibly because the biological body temperature of chickens (39-43 °C) is higher than that of humans (37 °C) (Aleksandrov et al., 2012). This work showed that the increased stability of chicken-CFTR was largely due to a number of proline residues in NBD1 that are unique to the chicken orthologue. Molecular dynamics analysis suggests that these prolines reduce flexibility in regions of NBD1 such as the RI insertion that may account for increased stability. Thermodynamic stability correlates well with crystallisation success (Alexandrov et al., 2008). Therefore, the chicken-CFTR appeared to be a good initial candidate for purification and crystallisation trials. This highlights the use of an orthologue screening strategy to select a protein more amenable to crystallisation.

1.5.4. Project aims

In the context of the strategy for the CFFT structure consortium, this PhD project asks the question: can yeast be used to express and purify CFTR protein for high-resolution structural studies? To answer this question, the project aims to:

1. Establish a novel protocol for the purification of milligram quantities of chicken-CFTR protein from a GFP-fusion tag S.cerevisiae expression system.

2. Develop biophysical assays to assess the quality of purified chicken-CFTR for high-resolution structural studies, namely characterising protein thermodynamic stability and homogeneity.

3. Develop biochemical assays to assess the activity of purified chicken-CFTR for functional relevance in future CF research.

From this point, the term “CFTR” will refer to chicken-CFTR, unless otherwise stated.
Chapter 2 - Materials and Methods

2.1. Materials

2.1.1. Chemicals

- Acetic acid (*Fischer Scientific, A/0400/PB17*)
- 30% acrylamide/Bis solution 29:1 (*Bio-Rad, 161-0156*)
- Agar (*Formedium, AGA03*)
- Aminoethylbenzenesulfonyl fluoride (AEBSF) (*Sigma-Aldrich, A8456*)
- Ammonium molybdate (*Sigma-Aldrich, A7302*)
- Ammonium persulphate (APS) (*Fischer Scientific, BPE179-100*)
- Ammonium sulphate (*Fischer Scientific, A/6486/53*)
- Ascorbate (*Sigma-Aldrich, A92902*)
- Bacto-peptone (*Formedium, PEP03*)
- Bacto-yeast extract (*Formedium, YEA04*)
- Benzamidine hydrochloride (*Sigma-Aldrich, 434760*)
- Bestatin (*Sigma-Aldrich, B838*)
- Bradford reagent concentrate (*Bio-Rad, 500-0006*)
- CFTRinh-172 (*Sigma-Aldrich, C2992*)
- Cholesterol (*Sigma-Aldrich, C8667*)
- Chymostatin (*Sigma-Aldrich, C7268*)
- Coumarin maleimide (CPM) (*Life Technologies, D-346*)
- Decyl-maltoside (DM) (*Affymetrix, D322*)
- Dithiothreitol (DTT) (*Sigma-Aldrich, 43817*)
- Dimethylsulfoxide (DMSO) (*Sigma-Aldrich, D8418*)
- Dodecyl-maltoside (DDM) (*Affymetrix, 69227-93-6*)
- *E.coli* total lipid extract (*Avanti lipids, 100500*)
- Epoxysuccinyl-leucylamido-butane (E-64) (*Sigma-Aldrich, E3132*)
- Ethanol (*Fischer Scientific, E/0056DF/P17*)
- Fos choline 16 (FC16) (*Affymetrix, F316*)
- D-galactose (*Fischer Scientific, G/0500/48*)
- Genistein (*Sigma-Aldrich, 92136*)
- Glass beads, acid washed, 425-600µm diameter (*Sigma, G8772*)
- D-glucose (*Fischer Scientific, G/0500/53*)
- L-glutathione reduced (*Sigma-Aldrich, G4251*)
• Glycerol (*Fischer Scientific, G/0650/17*)
• Glycine (*Fischer Scientific, BPE318*)
• Hydrochloric acid (HCl) (*Fischer Scientific, A466-2*)
• Imidazole (*Sigma-Aldrich, I553*)
• Instant Blue Coomassie (*Novexin, 15BIL*)
• Lauryledimethylamine oxide (LDAO) (*Affymetrix, D360*)
• Leupeptin hydrochloride (*Sigma-Aldrich, L9783*)
• Lithium acetate (LiAc) (*Fischer Scientific, L/2050/50*)
• Lyso-phosphatidylglycerol 14 (LPG) (*Avanti, 185092*)
• Magnesium chloride (*Sigma-Aldrich, M0250*)
• Magnesium sulphate (MgSO_4) (*Sigma-Aldrich, M7506*)
• Nonyl-maltoside (NM) (*Affymetrix, N330*)
• Octaethylene glycol monododecyl ether (C12E8) (*Calbiochem, 205528*)
• Octyl-maltoside (OG) (*Affymetrix, O330*)
• Oligomycin (*Sigma-Aldrich, 75351*)
• PageRuler plus pre-stained protein ladder (*Fermentas, SM1811*)
• PEG 4000 (*Sigma-Aldrich, 81240*)
• Pepstatin A (*Sigma-Aldrich, P4265*)
• Phenylmethanesulfonylfluoride (PMSF) (*Sigma-Aldrich, P7626*)
• Polyethylene glycol (PEG) 3350 (*Sigma-Aldrich, 39715*)
• Pre-cast 4-20% Tris/Glycine SDS-polyacrylamide gel electrophoresis (SDS-PAGE) gels (*NuSep, NN10-420*)
• Pre-cast 8% Tris/Glycine SDS-PAGE gels (*NuSep, NN10-008*)
• Pro-Q Diamond phosphoprotein gel stain (*Life Technologies, P33300*)
• SCH-28080 (*Sigma-Aldrich, S4443*)
• Single-stranded salmon sperm carrier DNA (*Sigma-Aldrich, D1626*)
• SM2 BioBeads (*BioRad, 152-3920*)
• *SmaI* restriction enzyme (*New England BioLabs, R0141*)
• Sodium adenosine triphosphate (Na_2ATP) (*Sigma-Aldrich, A2383*)
• Sodium azide (NaN_3) (*Sigma-Aldrich, S2002*)
• Sodium chloride (NaCl) (*Fischer Scientific, S/2920/53*)
• Sodium citrate (*Sigma-Aldrich, W302600*)
• Sodium dodecyl-sulphate (SDS) (*Fischer Scientific, BPE166*)
• Sodium meta-arsenite (*Sigma-Aldrich, S7400*)
• Sodium phosphate (*Sigma-Aldrich, 342483*)
• Sodium thiocyanate (SCN) (*Sigma-Aldrich, 71938*)
• Sucrose (*Fischer Scientific, S/7400/53*)
• Tetramethylethylenediamine (TEMED) (*Bio-Rad, 161-0800*)
• Tris base (*Fischer Scientific, T/P631/48*)
• Uranyl Acetate (UA) (*TAAB, U001*)
• Yeast nitrogen base (YNB) (*Formedium, CYN0410*)
• Yeast synthetic drop out without uracil (*Formedium, DCS0169*)

2.1.2. Yeast cells

*S.cerevisiae* FGY217 yeast cells with a *pep4* deletion were kindly donated by David Drew at Imperial College London. This yeast cell line has been used to express numerous eukaryotic membrane proteins for purification and crystallisation (Drew et al., 2008). The *pep4* deletion prevents expression of the lysosomal endopeptidase Pep4p (Woolford et al., 1986), which has been shown to limit degradation and improve expression of a number of membrane proteins in *S.cerevisiae* (Newstead et al., 2007b).

2.1.3. CFTR gene

The template CFTR DNA was designed and kindly provided by Ina Urbatsch (Texas Tech University) as part of the CFFT Structure Consortium. As shown in Figure 4, the gene contains the full-length, wild-type CFTR sequence from the chicken species *Gallus gallus* (see Appendix 3A or NCBI Reference Sequence: NP_001099136.1 for this chicken-CFTR protein sequence), with an upstream sequence for an N-terminal 10-histidine (His) and SUMO-fusion tag. All DNA sequences were codon-optimised for expression in yeast, as described before (Bai et al., 2011). The template DNA was amplified by Dr Liam O’Ryan to give stocks of PCR product, with 5’ overhangs complementary to the sequences in the expression vector (2.1.4.).

2.1.4. Expression vector

The 2 µ p424GAL1 expression vector was kindly provided by David Drew at Imperial College London (Drew et al., 2008). The key features of the vector are shown in Figure 4. There is a C-terminal tobacco etch virus (TEV)-cleavable yeast enhanced GFP-fusion tag that is codon-optimised for expression in yeast and contains mutations to improve folding efficiency and increase fluorescence brightness (GenBank: AAB18957.1) (Cormack et al., 1997). This GFP has a peak excitation at a wavelength of 488 nm and emission of 512 nm. The GFP is followed by an 8-His tag at the C-terminus. The
expression vector was linearised by Dr Liam O’Ryan using *SmaI* digestion to allow cloning by homologous recombination with the complimentary ends of the linearised CFTR gene (Figure 4).

Figure 4. Cloning the CFTR gene into a GFP-fusion yeast expression vector.
The CFTR gene contains sequences for a 10His -tag, a SUMO-tag and the full-length wild-type chicken-CFTR protein. The gene was introduced into the yeast p424GAL1 expression vector at the *SmaI* restriction site by homologous recombination using the 5’ overhanging complimentary DNA sequences (red lines), downstream to a galactose (GAL1) inducible promoter and upstream to sequences for yeast enhanced GFP and an 8His-tag. The vector also codes for a uracil selection marker (URA3) and ampicillin resistance gene (Amp<sup>r</sup>). This figure was adapted from (Drew et al., 2008).
2.1.5. Proteins

**BSA**
Albumin from bovine serum (*Sigma-Aldrich, A9418*). Dissolved in mH$_2$O at 10 mg/ml immediately prior to use. For BSA standard SDS-PAGE analysis, BSA was heated in SDS load dye for 5 min at 70 °C before loading on the gel.

**CIP**
Alkaline phosphatase from calf intestine (*New England BioLabs, M0290*). 10 000 units (U)/ml stocks, stored at -80 °C in buffer: 10 mM Tris pH 8, 10 mM KCl, 1 mM MgCl$_2$, 50 % glycerol weight/volume (w/v), 0.1 mM ZnCl$_2$.

**hCFTR**
Human-CFTR was expressed in HEK cells by the John Kappes lab (University of Alabama at Birmingham) and purified by Ellen Hildebrandte (Urbatsch lab, Texas Tech University). The human-CFTR construct has the same SUMO- and GFP-fusion tags as the chicken-CFTR (Figure 1), but also an additional FLAG epitope in the 4$^{th}$ extracellular loop. Human-CFTR was purified by nickel-affinity and FLAG-affinity chromatography in the buffer: 50 mM Tris pH 7.5, 150 mM NaCl, 10 % glycerol volume/volume (v/v), 2.5 mM MgCl$_2$, 1 mM ATP, 1 mM DTT, 0.1 % UDM (w/v).

**GFP**
Purified full-length GFP with an N-terminal 6-His tag (*Merck/ Millipore, 14-392*). 1 mg/ml stocks stored at -80 °C in PBS buffer + 20 % glycerol (v/v).

**PKA**
Protein kinase A catalytic subunit (*New England BioLabs, P6000*). 40 µM PKA stocks, stored at -80 °C in buffer: 20 mM Tris pH 7.5, 50 mM NaCl, 2 mM DTT, 1 mM EDTA.

**PNGaseF**
Peptide-N-glycosidase F (*New England BioLabs, P0704*). Also included: 10 x glycoprotein denaturing buffer, 10 x G7 reaction buffer, 10 % NP-40. All reagents stored at -20 °C and kept on ice prior to use.

**TEV**
AcTEV™ protease (*Life technologies, 12575*). 10 U/µl stocks, stored at -80 °C in buffer: 50 mM Tris-HCl, pH 7.5, 1 mM
EDTA, 5 mM DTT, 50 % (v/v) glycerol, 0.1 % (w/v) Triton X-100. Also included: 20 x TEV reaction buffer (1M Tris pH8, 10 mM EDTA) and 100 mM DTT.

2.1.6. Media

All media were made in deionized water that had been sterile filtered through Millipore Q 0.2 µM filter system (mH$_2$O) and sterilised by heating in an autoclave oven.

- **YNB media**: 6.7 g/L YNB without amino acids, 0.77 g/L yeast synthetic drop out without uracil.
- **YNB plate**: YNB media + 8 g/L agar.
- **YPD media**: Yeast peptone dextrose: 10 g/L of bacto-yeast extract, 20 g/L bacto-peptone, 20 g/L glucose.

2.1.7. Inhibitors

- **ATPase (IC)**: For 100 x stocks (v/v): 1 mM sch28080 was dissolved in DMSO, 1M NaSCN in mH$_2$O and 2.5 mM oligomycin in 100 % (v/v) ethanol. Aliquots were stored at -20 °C.
- **Genestein (GNS)**: For 1000 x stocks (v/v): GNS was dissolved in DMSO to 100 mM and stored at -20 °C.
- **Glutathione (GSH)**: GSH stocks were made immediately prior to use by dissolving 0.1 M GSH in the relevant CFTR buffer.
- **CFTR$_{inh-172}$**: For 1000 x stocks (v/v): CFTR$_{inh-172}$ was dissolved in DMSO to 10 mM and stored at -20 °C under protection from light.
- **Protease (PIs)**: For 1000 x stocks (v/v): 200 mM AEBSF, 4 mM chymostatin, 7 mM E-64, 20 mM leupeptin, 15 mM pepstatin, 1M PMSF, 6 mM bestatin were dissolved in DMSO. Benzamidine was dissolved in mH$_2$O at 300 mM for 100 x stocks. All stocks were stored in 200 µl aliquots at -20 °C to limit number of freeze-thaw steps.
2.1.8. Buffers: purification

All buffers were made in mH$_2$O, filtered through a Millipore 0.2 µM filter system and pre-chilled to 4°C.

DDM purification 50 mM Tris pH 8, 1 M NaCl, 20 % glycerol (v/v), 1 mM DTT, PIs, 0.1 % DDM (w/v).

CFTR 50 mM Tris pH 8, 50 mM NaCl, 20 % glycerol (v/v), 1 mM DTT, PIs.

LPG purification 50 mM Tris pH 8, 50 mM NaCl, 20 % glycerol (v/v), 1 mM DTT, PIs, 0.1 % LPG (w/v).

mPiB 0.3 M Tris pH 8, 0.3 M sucrose, 10 % glycerol (v/v), 2 mM DTT, PIs.

SEC-DDM 50 mM Tris pH 8, 1 M NaCl, 20 % glycerol (v/v), 1 mM DTT, PIs, 0.1 % DDM (w/v). Buffer was degassed using a Millipore 0.2 µM filter system under vacuum.

SEC-LPG 50 mM Tris pH 8, 50 mM NaCl, 10 % glycerol (v/v), 1 mM DTT, 0.05 % LPG (w/v). Buffer was degassed using a Millipore 0.2 µM filter system under vacuum.

2.1.9. Buffers: SDS-PAGE

8% gels Lower resolving gel (for 1 x 1.5 mm gel): 4.8 ml mH$_2$O, 2.7 ml of 30 % acrylamide (v/v), 2.5 ml buffer (1.5 M Tris-HCl pH 8.8, 0.4 % SDS (w/v)), 50 µl of 0.1 mg/ml APS, 20 µl TEMED.

Upper stacking gel (for 1 x 1.5 mm gel): 3.1 ml mH$_2$O, 650 µl of 30 % acrylamide, 2.5 ml buffer (0.5 M Tris-HCl pH 6.8, 0.4 % SDS (w/v)), 25 µl of 0.1 mg/ml APS, 10 µl TEMED.

Running buffer For 10 x stocks (v/v): 0.25 M Tris base, 1.92 M glycine, 1 % SDS (w/v).

SDS load dye For 2 x stocks (v/v): 50 mM Tris pH 7.6, 5 % glycerol (v/v), 5 mM EDTA pH 8, 0.02 % bromophenol blue (w/v), 4 % SDS (w/v), 50 mM DTT. Stored in aliquots at -20 °C.
2.1.10 Buffers: ATPase assay

ATPase

50 mM Tris pH 7.4, 150 mM NH₄Cl, 5 mM MgSO₄ and 0.02 % (w/v) NaN₃. Buffer was made in mH₂O and stored at RTM for up to one month.

Buffer A

3 % (w/v) ascorbate, 0.5 % (w/v) ammonium molybdate in 0.5 M HCl. Buffer was made immediately before use and kept under protection from light.

Buffer B

2 % (w/v) sodium citrate, 2 % (w/v) sodium meta-arsenite, 2 % (v/v) acetic acid. Buffer was made in mH₂O and stored at RTM for up to one month.

2.2. Methods: expression and purification

2.2.1. Transformation of yeast cells

Competent yeast cells were prepared by growing 100 µl FGY217 cells in 50 ml of YPD media to an optical density at 600 nm (OD₆₀₀) of 0.5, as measured by a UV/vis scanning spectrophotometer (Jenway 6315), followed by centrifugation at 3000 g for 5 min and resuspended into 1 ml of 100 mM LiAc. The chicken-CFTR gene was cloned into the 2 µM expression vector by homologous recombination and transformed into the competent yeast cells using a rapid, single-step method (Drew et al., 2008, Gietz and Schiestl, 2007). In summary: 50 µl of competent yeast cells, 240 µl of 50 % PEG3350 (w/v), 25 µl of 2 mg/ml single-stranded salmon sperm carrier DNA, 3 µl of 25 ng/µl Smal-digested expression vector, 5 µl of 150 ng/µl CFTR gene PCR product and 42 µl of sterile H₂O were mixed and incubated for 30 min at 30 °C, and then heat shocked for 25 sec at 42 °C. The cells were centrifuged at 8000 g for 15 sec, resuspended in 100 µl of sterile H₂O and grown on a YNB media plate + 2 % glucose (w/v) under minus uracil selection, for 2 days at 30 °C. Plates were stored at 4 °C for up to one month.

2.2.2. Small-scale cell culture

A single yeast transformation colony was taken from plate (section 2.2.1) and used to inoculate 5 ml YNB media + 2 % glucose (w/v) in a sterile 50 ml Falcon tube and grown overnight (o/n) at 30 °C with shaking at 220 revolutions per minute (RPM) until an OD₆₀₀ of 1.0 was reached. Cultures were diluted into 50 ml YNB media + 0.1 % glucose (w/v) in a sterile 250 ml Erlenmeyer baffled flask to an OD₆₀₀ of 0.1 and grown
at 30 °C with shaking at 220 RPM until an OD$_{600}$ of 1.0 was achieved (after ~ 8 h). CFTR expression was induced with 2 % galactose (w/v) and grown for 16 h at 20 °C with shaking at 220 rpm. 8 % glycerol (v/v) was also added at induction to increase CFTR expression levels. Cells were harvested by centrifugation at 4000 g for 5 min at 4 °C, resuspended in mPiB buffer to a final volume of 1 ml and flash-frozen in liquid nitrogen for storage at -80 °C. Glycerol stocks were made by mixing 800 µl cells with 200 µl glycerol in sterile 2 ml screw-top vials and flash-freezing in liquid nitrogen for long-term storage at -80 °C.

2.2.3. Large-scale fermentation cultures

Cells from 2 x transformation plates (section 2.2.1) were used to inoculate 2 x 750 ml YNB media + 2 % glucose (w/v) in sterile 2 L Erlenmeyer baffled flasks and grown o/n at 30 °C with shaking at 220 RPM until an OD$_{600}$ of 1 -1.5 was achieved. Cells were diluted into a final volume of 15 L YNB + 0.1% glucose (w/v) in a sterile 20 L fermenter vessel (Applikon) to an OD$_{600}$ of 0.1 -0.15 and grown at 30 °C with stirrer speed set to 700 RPM and compressed air at 15 dm$^3$ min$^{-1}$. Once the culture reached an OD$_{600}$ of 1 -1.5 (after ~ 8 h), CFTR expression was induced by addition of YNB media + galactose + glycerol, to a final galactose concentration of 2 % (w/v) and glycerol concentration of 8 % (v/v). The temperature was reduced to 20 °C and cells grown for 16 h before harvesting by centrifugation at 4000 g for 5 min at 4 °C. Cells were resuspended in mPiB buffer to a final volume of 200 ml and flash-frozen in liquid nitrogen for storage at -80 °C.

2.2.4. Fluorescence microscopy

Fluorescence microscopy was used to determine the localisation of CFTR in yeast cells by its GFP-fusion tag as described before (Drew et al., 2008). Cells were taken from small-scale cell cultures at 0 and 16 h post-induction and harvested by centrifugation at 4000 g at 4 °C for 5 min. Cells were resuspended in YNB media + 50 % glycerol to limit cell mobility in solution. 50 µl samples were dropped into the reservoir of a glass-based dish microscope slide and analysed on a Delta Vision RT (Applied Precision) restoration microscope with a 100 x objective. Images were acquired using a Coolsnap HQ (Photometrics) camera under a FITC filter (excitation wavelength of 490 nm and emission wavelength of 528 nm) with a Z optical spacing of 0.2 µm, a 2 x 2 binning and a 1 sec exposure time. Raw images were then deconvoluted using the softWoRx
software (*Applied Precision*) and maximum intensity projections of these deconvoluted images are shown in the results. Images were analysed using the Image J software (*National Institutes of Health, USA*).

### 2.2.5. Small-scale crude membrane preparation

Cells from small-scale cultures were thawed on ice and added to 1.5 ml screw-top tubes containing 200 µl glass beads that had been pre-chilled to 4 °C. Cells were broken by 5 x 1 min bursts in a Mini-Beadbeater-24 cell disrupter (*Biospec, 112011*), with 2 min rest on ice between each burst. Beads and unbroken cells were removed by centrifugation at 4000 g for 5 min at 4 °C in a benchtop centrifuge. The cell lysate was collected from the supernatant and centrifuged at 100 000 g for 1 h at 4 °C in a benchtop ultra-centrifuge to pellet crude membranes.

### 2.2.6. Large-scale microsome preparation

Cells from a large-scale fermenter culture were thawed on ice and added to 200 ml glass beads that had been pre-chilled to 4 °C. Cells were broken with 5 x 1 min bursts in a BeadBeater cell disrupter (*Biospec, 1107900*) surrounded by dry ice, with 1 min rests between each burst. A 200 µl sample was taken from the cell lysate and centrifuged at 12 000 g at 4 °C for 5 min to remove unbroken cells and intact organelles. Cell breakage was deemed sufficient if a 1/50 dilution of the supernatant in mPiB had an OD$_{380}$ > 0.1, otherwise breakage in the BeadBeater was repeated until this OD$_{380}$ was achieved. Once the cells were broken, the cell lysate was collected, which involved washing the glass beads in mPiB to collect residual cell lysate. The lysate was centrifuged at 12 000 g at 4 °C for 10 min to remove unbroken cells and intact organelles. The supernatant was collected and centrifuged at 200 000 g for 2 h at 4 °C to pellet microsomes. The supernatant was discarded and the pellet resuspended in 100 ml CFTR buffer using a paintbrush and vortex mixing, before centrifugation at 100 000 g for 1 h at 4 °C. The supernatant was discarded and the pellet initially resuspended in 10 ml of CFTR buffer using a paintbrush and vortex mixing. The total protein concentration was determined using the Bradford assay (section 2.2.7) and the volume of microsomes adjusted to a final concentration of 10 mg/ml in CFTR buffer. Microsomes were flash-frozen in liquid nitrogen and stored at -80 °C for up to 6 months.
2.2.7. Bradford assay

The Bradford assay was used to estimate the total protein concentration of microsomes (Bradford, 1976). BSA protein standards were made up in mH2O (0.1, 0.2, 0.4, 0.6, 0.8, 1 mg/ml). Microsomes were diluted 1/10 with mH2O and 10 µl added to a 1 ml disposable polystyrene cuvette containing 800 µl mH2O and 200 µl Bradford reagent, which was then incubated at room temperature (RTM) for 5 min. Absorbance was measured at OD595. A BSA standard curve was generated by adding 10 µl of BSA standards instead of the 10 µl microsomes and used to determine total microsome protein concentration.

2.2.8. GFP-fluorescence spectroscopy

Fluorescence spectroscopy was used to detect CFTR via its GFP-fusion tag throughout microsome preparations and purification. The sample was added to a quartz cuvette with 1 cm light path-length (Hellma Analytics, 105.250-QS) and the fluorescence emission spectrum measured in a fluorometer (Carys Eclipse, Varian) with excitation wavelength set at 485 nm and emission at 500-600 nm. The maximum peak of GFP fluorescence was detected at 512 nm. The CFTR protein concentration could be estimated using a GFP standard curve with a straight line passing through the origin that is described by the equation of \( y = 1.9524x + 0 \), where \( y \) is the GFP fluorescence intensity at emission wavelength 512nm and \( x \) is the concentration of CFTR in µg/ml, as shown in Appendix figure 1B. All GFP fluorescence values are given as arbitrary units after background subtraction of buffer.

2.2.9. SDS-PAGE

SDS-PAGE analysis was used throughout expression, purification and reconstitution experiments to identify proteins based on their molecular weight (Weber and Osborn, 1969). Protein samples were diluted with an equal volume of SDS load dye and incubated at RTM for 10 min to allow denaturation of CFTR. Samples were loaded onto either 8 % or 4-12 %, home-made or pre-cast SDS-PAGE gels and run at 150 V for 1 h in a Mini-PROTEAN SDS-PAGE system (BioRad, 165-8006). 10 µl sample was loaded per well, unless stated otherwise in the figure legend. 5 µl of PageRuler plus pre-stained molecular weight marker was also loaded per gel (Fermentas, SM1811). It was important that the CFTR-containing sample was not heated and that the gel electrophoresis was performed slowly to prevent denaturation of the GFP tag. Gels were
scanned and digitised in a ChemiDoc MP gel imager (*BioRad, 170-808*), using a blue light and a 530 nm emission filter, to show the GFP-tag fluorescence in the gel. The molecular weight markers could also be viewed under fluorescence conditions using the green light with a 605 nm emission filter and the red light with and a 695 nm emission filter. Gels were scanned within 1 h after electrophoresis to limit protein loss from the gel by diffusion. Gels were next stained with Instant Blue Coomassie stain for 45 min and destained in mH2O o/n to show the total protein content of the fraction. Alternatively, gels were stained with Pro-Q Diamond stain to detect protein phosphorylation, as described in the manufacturer’s guidelines ([http://tools.lifetechnologies.com/content/sfs/manuals/mp33300.pdf; accessed 3/10/13](http://tools.lifetechnologies.com/content/sfs/manuals/mp33300.pdf); Steinberg et al., 2003). Coomassie stained gels were scanned and digitised in the ChemiDoc MP gel using the white light, whereas Pro-Q Diamond stain gels were scanned using the Trans-UV light. Analysis of gels, including quantification of protein band intensity, was performed using the Image J software (*National Institutes of Health, USA*).

### 2.2.10. Detergent solubilisation screen

A range of detergent types were tested for their ability to solubilise CFTR from yeast microsomes, including: fos-choline 16 (FC), lyso-phosphatidylcholine 14 (LPG), lauryldimethylamine oxide (LDAO), decyl-maltoside (DM), dodecyl-maltoside (DDM), octyl-maltoside (OG), nonyl-maltoside (NM), octaethylene glycol monododecyl ether (C12E8), and octyl glucoside (OG). Microsomes were thawed on ice and diluted with an equal volume of CFTR buffer containing 10 % detergent (w/v) to give a total protein concentration of 5 mg/ml and a ratio of detergent to total protein of 10:1 (w/w). This was incubated at 4 °C for 1 h with agitation on a tube rotator. Material was deemed detergent-soluble if it remained in the supernatant after centrifugation at 100 000 g for 1 h at 4 °C (Gutmann et al., 2007). The detergent-insoluble material was collected from the pellet, which was resuspended in CFTR buffer in equal volumes to the supernatant. The soluble and insoluble material for each detergent was analysed by SDS-PAGE under GFP fluorescence conditions (section 2.2.9). The intensity of the CFTR-GFP fluorescence protein bands were quantified in Image J software and the efficiency of detergent solubilisation determined as a percentage of S/(S+I), where S and I represent the intensity of the CFTR-GFP band in the soluble and insoluble fractions, respectively.
2.2.11. Detergent solubilisation for CFTR purification

CFTR was solubilised in either LPG or DDM detergent for later purification steps. Purification in DDM first required an additional step to wash the microsomes in high salt before detergent solubilisation to remove contaminating yeast proteins. In this case, during the last stage of microsome preparation (section 2.2.6), microsomes were resuspended in CFTR buffer supplemented with 1M NaCl, before pelleting by centrifugation at 100 000 g for 1 h at 4 °C. The final microsomal pellet was then resuspended to 10 mg/ml total protein concentration in CFTR buffer + 1M NaCl. This high salt wash was not needed for purification in LPG due to the high purity already achievable in LPG.

Microsomes were thawed on ice and diluted to 5 mg/ml total protein concentration in either DDM or LPG purification buffer, so that the final detergent concentration was 2 % (w/v) and the detergent to protein ratio was 4:1 (w/w). This was incubated at 4 °C for 1 h with agitation on a tube rotator, before centrifugation at 100 000 g to pellet the detergent-insoluble material and leave the detergent-soluble material in the supernatant. The pellet was resuspended in equal volumes of CFTR buffer as the supernatant and kept for later SDS-PAGE analysis to assess the efficiency of detergent solubilisation, as described above (section 2.2.10)

2.2.12. Nickel-affinity chromatography purification

Two HisTrap columns (GE Healthcare, 17-5247-01) were linked in series and attached to an ÄKTA Purifier system (GE Healthcare) maintained at 4 °C. The columns were equilibrated in 10 column volumes (CV) of mH2O and then 10 CV of purification buffer. The detergent-soluble material was supplemented with 5 mM imidazole and passed over the columns at a flow rate of 0.5 ml/min to bind CFTR protein via its histidine tags. The unbound material was collected in equal volumes to the loading volume. The columns were washed with 3 CV buffer + 40 mM Imidazole and 3 CV buffer + 100 mM Imidazole before eluting CFTR with 3 CV buffer + 400 mM Imidazole in 2 ml fractions. All fractions were analysed by fluorescence spectroscopy and SDS-PAGE to track the CFTR protein.
2.2.13. Size-exclusion chromatography (SEC) purification

A Superose 6 10/300 SEC column (GE Healthcare, 7-5172-01) was attached to a ÄKTA Purifier system (GE Healthcare) maintained at 4 °C, and equilibrated in 1.2 CV mH₂O and then 1.2 CV of SEC buffer. During this step, the nickel-purified material was concentrated to a volume less than 500 µl using a centrifugal concentrator with a 100 kDa molecular weight cut off (MWCO) filter (Vivaspin, VS0641). For purification in DDM, the concentration of CFTR was kept below 0.3 mg/ml to prevent protein aggregation and loss on the filter. The retentate from the concentrator was collected and centrifuged at 100 000 g for 30 min at 4 °C to pellet insoluble particles and large aggregates. The supernatant was collected and injected onto the SEC column, which was washed with an isocratic gradient of 1.2 CV of SEC buffer, collecting 0.5 ml elution fractions. All fractions were analysed by fluorescence spectroscopy and SDS-PAGE to track the CFTR protein.

2.2.14. Static Light Scattering

Static light scattering measurements were used to assess changes in CFTR aggregation state at different salt concentrations. DDM-purified CFTR was diluted to 20 µg/ml in SEC-DDM buffer containing different concentrations of NaCl (to final concentration of 0.1, 0.2, 0.4, 0.6, 0.8 and 1M NaCl) and incubated in a 1 cm quartz cuvette (Hellma Analytics, 105.250-QS) for 10 min in a fluorometer (Carys Eclipse, Varian) held at 10 °C by a temperature-controlled water bath. The sample was excited by light at 290 nm wavelength through a 10 nm wide slit, and fluorescence detected at 90 ° relative to the direction of the incident light beam. The monochromatic grating of the fluorometer allowed for detection of the second harmonic wave generated from the scattering of 290 nm light, with a peak emission of 565 nm through a 5 nm slit (Franken et al., 1961). This allowed simultaneous collection of light scattering (2.2.1) and Trp fluorescence (2.2.3) measurements in later thermal denaturation experiments. Increases in light scattering fluorescence were used as an indication of protein aggregation (Borgstahl, 2007, Hayashi et al., 1989). Purified GFP was also analysed using the same method to show that changes in light scattering observed for CFTR were not an artefact of NaCl concentrations on the buffer, detergent or GFP-tag. All light scattering values are given as arbitrary units after background subtraction of buffer.
2.2.15. Mass spectrometry

Mass spectrometry was used to identify CFTR and contaminating yeast proteins in the final purification samples. DDM- and LPG-purified CFTR was collected after SEC purification, diluted to 50 µg/ml in relevant SEC buffer and 10 µl analysed by SDS-PAGE gels under Coomassie stain. To analyse individual protein bands, the SDS-PAGE was performed as described in section 2.2.9 and the band of interest excised using a sterile scalpel blade. To analyse the total protein content, the sample was run a few millimetres into the top of the stacking gel before staining, and the entire band cut out for analysis. All mass spectrometry analysis was performed by the Protein Identification Service in the Faculty of Life Sciences at the University of Manchester. Peptide fragments were analysed against the IPI Chicken database (version 3.64) to identify the chicken-CFTR protein (IPI00575922) and against the yeast protein UniProt database (version 2011-05) to identify contaminating yeast proteins.

2.2.16. Glycosylation

The glycosylation state of CFTR was probed by treatment with PNGaseF, an endo-glycosidase that cleaves most N-linked glycans (Mellors and Sutherland, 1994). 2 µl of CFTR-containing microsomes, at a total protein concentration of 10 mg/ml, were mixed with either 2 µl of 10 x Glycoprotein denaturing buffer and 6 µl mH2O for denaturing conditions (D) or with 8 µl of mH2O for non-denaturing conditions (ND) and incubated at RTM for 10 min. The concentration of CFTR in these microsomes was estimated at 1 % of the total protein (=100 µg/ml) as judged by GFP fluorescence spectroscopy; therefore the amount of CFTR in this reaction mixture is about 0.2 µg. Both D and ND samples were mixed with 2 µl of 10 x G7 Reaction Buffer, 2 µl of 10 % NP-40, 4 µl mH2O and either 2 µl PNGaseF or 2 µl of mH2O for the presence and absence of PNGaseF treatment, respectively. Samples were incubated o/n at RTM and analysed by SDS-PAGE scanned for GFP fluorescence. CFTR was deemed to be core glycosylated if there was a shift in protein band mobility in the presence of PNGaseF compared to the absence of PNGaseF treatment.

2.2.17. Phosphorylation

The phosphorylation state of CFTR was assessed by treatment with either PKA or CIP, which are enzymes known to phosphorylate and de-phosphorylate CFTR respectively (Aleksandrov et al., 2002a, Li et al., 1996). CFTR protein was purified in DDM and
stocks made at 20 \( \mu \text{g/ml} \) in SEC-DDM buffer. Immediately prior to use, PKA and CIP stocks were thawed on ice and diluted 1/100 to 400 nM PKA and 100 U/ml CIP in SEC-DDM buffer supplemented with 10 mM MgATP. 10 \( \mu \text{l} \) of purified CFTR was mixed with 10 \( \mu \text{l} \) of diluted enzyme (PKA or CIP) or with 10 \( \mu \text{l} \) of SEC-DDM buffer for the control sample, and incubated o/n at 4 °C. The reaction was stopped by adding SDS-load dye at RTM for 10 min. Samples were analysed by SDS-PAGE gels, which were scanned for GFP fluorescence and then stained with Pro-Q Diamond stain, as described in section 2.2.9. CFTR phosphorylation was shown by a visible band on the gel after Pro-Q Diamond staining (Steinberg et al., 2003).

2.2.18. TEV cleavage

TEV protease was used to remove the C-terminal GFP tag (Figure 1) from CFTR after purification. CFTR was purified in both DDM and LPG and diluted to 35 \( \mu \text{g/ml} \) in the relevant SEC buffer. 15 \( \mu \text{l} \) of purified CFTR (= 0.5 \( \mu \text{g} \)) was incubated with 1.8 \( \mu \text{l} \) TEV protease (= 18 U) o/n at 4 °C. The reaction was stopped by adding SDS-load dye at RTM for 10 min. The efficiency of TEV cleavage was assessed by comparing the intensity of the fluorescence band for full-length CFTR (220 kDa) and cleaved GFP (27 kDa) on a 4-20 % pre-cast gradient SDS-PAGE gel, scanned under GFP fluorescence conditions. For 90 % cleavage efficiency of the GFP-tag a ratio of 36 units of TEV was needed per 1 \( \mu \text{g} \) of CFTR, which is 80 x the amount of TEV recommended in the manufacturer’s guidelines (http://tools.lifetechnologies.com/content/sfs/manuals/12575.pdf; accessed 3/10/13).

2.3. Methods: stability

2.3.1. Static Light Scattering

Static light scattering measurements were used to assess changes in CFTR aggregation state as a function of protein concentrations, temperature and time. Samples were added to a 1 cm quartz cuvette (Hellma Analytics, 105.250-QS) that had been pre-chilled to 10 °C, in a fluorometer (Carys Eclipse, Varian) connected to a temperature-controlled water bath. 90 ° light scattering measurements were collected as described in section 2.2.14, with excitation wavelength of 290 nm and slit width 10 nm, and emission wavelength of 565 nm and slit width 5 nm. For protein concentration experiments, purified CFTR was concentrated using a MWCO 100 kDa spin concentrator (Vivaspin, VS0641) and aliquots were taken every 2 min for analysis at 10 °C. For time-course
experiments, purified CFTR was diluted to 50 µg/ml in SEC buffer and stored at 4 °C, with aliquots taken daily between 0-10 days for analysis at 10 °C. For thermal denaturation experiments, purified CFTR was diluted to 50 µg/ml in SEC buffer and heated from 10 to 80 °C at a scan rate 2.5 °C/min, with readings collected every 2 °C increments. Emission spectra were collected between wavelengths 300-600 nm and slit widths 10 nm to allow simultaneous measurements of Trp fluorescence (section 2.3.3) and 90 ° light scattering during thermal denaturation. All light scattering values are given as arbitrary units after background subtraction of buffer.

2.3.2. CPM dye binding assay

The CPM dye binding assay was used to probe the surface exposure of cysteine residues during thermal unfolding of purified CFTR protein (Alexandrov et al., 2008). Firstly, the DTT had to be removed from the CFTR buffer to prevent its interference with the CPM fluorescence, although this caused some loss in CFTR protein probably due to aggregation. DTT could be removed by diafiltration on a MWCO 100 kDa spin concentrator (Vivaspin, VS0641), on a PD10 desalting column (GE Healthcare, 17-0851-01), or omission from the buffer during the final SEC purification step. Purified chicken-CFTR was diluted in SEC buffer without DTT to a final concentration of 3.33 µg/ml in a volume of 150 µl (= 2 µg CFTR). For the human-CFTR sample, the protein was exchanged and diluted using the same methods into the buffer: 50 mM Tris pH7.5, 150 mM NaCl, 10 % glycerol (v/v), 0.1 % UDM. The sample was added to a 1 cm Quartz cuvette (Hellma Analytics, 105.250-QS) that had been pre-chilled to 10 °C in a fluorometer (Carys Eclipse, Varian) connected to a temperature-controlled water bath. CPM dye stocks were made by dissolving in DMSO at 4 mg/ml, which were stored at -80 °C under protection from light. Immediately prior to use, CPM was diluted to 0.1 mg/ml in the relevant SEC buffer (without DTT) and 1 µl (= 0.1 µg) added to the cuvette with thorough mixing by pipetting. Fluorescence measurements were collected every 10 sec with excitation at 387 nm, emission 463 nm and slit widths of 5 nm. Initially, readings were collected at a constant temperature of 10 °C until the CPM fluorescence had plateaued to signify labelling of surface exposed cysteines in the folded protein. CPM fluorescence was then monitored as the cuvette was heated from 10 to 80 °C at a heat rate of 2.5 °C/min. An increase in fluorescence signified the labelling of cysteine residues as they became surface exposed during thermal unfolding.
2.3.3. Tryptophan (Trp) fluorescence measurements

The intrinsic fluorescence properties of Trp residues in CFTR were used to probe the folded state of CFTR during thermal denaturation (Eftink, 1998). Purified CFTR was diluted into SEC buffer to 50 µg/ml and added to a 1 cm Quartz cuvette (Hellma Analytics, 105.250-QS) that was pre-chilled to 4 °C in a fluorometer (Carys Eclipse, Varian) connected to a temperature-controlled water bath. CFTR Trp fluorescence had peak emission around 330 nm when excited with light at wavelength of 290 nm through 5 nm width slits. Emission spectra were collected between wavelengths 300-600 nm and slit widths 10 nm to allow simultaneous measurements of Trp fluorescence and 90 ° light scattering (section 2.3.1) during thermal denaturation. Emission spectra were collected at every 2 °C increments during heating from 10 to 80 °C at a heat rate of 2.5 °C/min. All light scattering values are given as arbitrary units after background subtraction of buffer.

2.3.4. EM

EM was used to view changes in CFTR particles before and after heating. Purified CFTR was diluted to 50 µg/ml in SEC buffer and incubated at either 10 or 80 °C for 10 min, before being quenched on ice for 10 min. 5 µl of sample was applied to a 400-mesh continuously carbon-coated copper grid (Electron Microscopy Sciences, CF-400-Cu) that had been glow discharged (GD) for 25 sec at 25 mA. The grids were then washed with 2 x 5 µl mH2O for 10 sec and stained with 5 µl 4 % UA (w/v) for 30 sec; grids were blotted with filter paper (Whatman, 1001090) between each step. Grids were then analysed on a Tecnai 12 Biotwin Transmission Electron Microscope (FEI) fitted with an Orius charge-couple device camera (Gatan). Electron Micrographs were analysed using the EMAN 2 software (National Centre for Macromolecular Imaging), where CFTR particles were selected using 48 pixels box size (= 249.6 nm) and 128 pixel box size (= 665.6 nm) for CFTR heated to 10 °C and 80 °C respectively.

2.3.5. Thermal gel analysis (TGA)

Membrane proteins, when heated, form SDS-insoluble aggregates that are unable to enter the SDS-PAGE separating gel or stacking gel (Lysko et al., 1981). SDS-PAGE analysis was therefore used to assess the aggregation state of CFTR after thermal unfolding, in a technique known as thermal gel analysis (TGA). Purified CFTR was diluted using SEC buffer to 50 µg/ml in an eppendorf tube and heated in a water bath from 15 to 80 °C at a heating rate of 2.5 °C/min. Aliquots were taken every 5 °C
increments, quenched on ice for 10 min, and then incubated with equal volume of SDS load-dye for 15 min at RT before loading approximately 10 ng of CFTR onto pre-cast 10% SDS-PAGE gels, which were scanned to detect GFP fluorescence. TGA was also used to assess the stability of CFTR at 4 °C as a function of time, where aliquots were taken daily between 0 and 10 days for similar analysis by SDS-PAGE. CFTR either migrated as a monomer band at approximately 220 kDa, or as a higher molecular weight (HMW) “aggregate” at the top of the stacking gel. The integrated intensities of these fluorescent CFTR-GFP bands were quantified using Image J software (National Institutes of Health, USA).

2.3.6. Graphical analysis of the thermal unfolding data
The mid-point temperature of thermal unfolding (T_m) was calculated for experiments, including light scattering, CPM, Trp and TGA. The CPM data were normalised to represent a fraction of unfolding, using the equation: fraction of unfolding = (Y-Y_D)/(Y_D-Y_N), where Y is fluorescence for a given temperature, Y_D is fluorescence with the denatured protein i.e. at 80 °C, and Y_N is fluorescence with the natively folded protein i.e. at 10 °C. The light scattering fluorescence was normalised as a fraction of change compared to light scattering at 10 °C. The shift in Trp fluorescence was normalised using the formula y = F_D/F_N, where F_D is the fluorescence intensity at the peak emission wavelength (335nm) of the denatured protein, and F_N is the fluorescence intensity at the emission wavelength (331nm) of the native protein, for each temperature point. After normalisation, all data was fitted with a two-state sigmodial curve to show a two-state unfolding model, where the native and denatured protein are represented by the plateaus at 10 °C and 80 °C respectively. The 1st derivative was also calculated from these data and fitted with a van’t Hoff plot, from which the T_m was calculated from the peak, as described previously (John and Weeks, 2000). All graphical analysis was performed in GraphPad Prism 6 software.

2.3.7. Analysis of CFTR structure models
Published human-CFTR structure models were used to predict the 3D location of Cys and Trp residues in the chicken-CFTR protein. This information was used to interpret data from the CPM dye binding and Trp fluorescence assays. The protein sequence of human-CFTR (NP_000483.3) was compared to the chicken-CFTR sequence (NP_001099136.1) using the BLAST protein alignment tool (http://blast.ncbi.nlm.nih.gov/). The PDB coordinates for the inward-facing
conformation of the human-CFTR structure model by Mornon et al. (Mornon et al., 2008) were downloaded from http://www.impmc.upmc.fr/~callebau/CFTR.html and analysed in Chimera. Cys were deemed surface-exposed if the thiol group was clearly accessible to the atomic surface of the structure. The domain localisation of Trp was predicted by comparison to domain allocations within the primary CFTR sequences by Mornon et al. (Mornon et al., 2008).

2.4. Methods: homogeneity

2.4.1. SEC multi-angle light scattering (SEC-MALLS)

SEC-MALLS was used to assess the monodispersity of CFTR particles after purification (Sahin and Roberts, 2012). LPG- and DDM-purified CFTR was diluted in their relevant SEC buffer to 0.4 and 0.2 mg/ml respectively before filtration through a 0.2 µm membrane to remove HMW aggregates. The filtered protein was then injected onto an analytical KW-804 SEC column (Shodex F6989104) attached to a ÄKTA purifier system (GE Healthcare) maintained at 4 °C. As the protein eluted, light scattering and refractive index (RI) measurements were collected by a MALLS detector (Wyatt Technology) that were then transformed into molar mass estimates by the ASTRA programme (Wyatt Technology) using the Zimm expression equation, where the RI was used to estimate protein concentration as described previously (Li et al., 2009). The elution fractions were also analysed by GFP-fluorescence spectroscopy to track presence of CFTR.

2.4.4. Cryo- EM

Cryo-EM was used to assess the homogeneity of LPG-purified CFTR particles. This technique could not be used for DDM-purified CFTR because the protein concentration was too low and the requirement for high salt in the DDM buffer prevented sufficient signal-to-noise to view protein particles. LPG-purified CFTR was diluted to 1 mg/ml into a glycerol-free SEC buffer so that the final concentration of glycerol was 1.25 % (w/w), in order to minimise the scattering of electrons by the buffer alone. 3 µl protein was applied to a holey carbon-coated EM grid (300-mesh/inch, 2 micron diameter holes with 2 micron separation) (Quantifoil, Q325-CR2) at ~ 90 % humidity. The grids were glow discharged for 25 sec at 25 mA to encourage even distribution of particles within the ice. The grid was blotted with filter paper (Whatman, 1001090) with 2 x 1.5 sec blots and frozen in a layer of amorphous ice by vitrification in liquid ethane. Samples
were then analysed at liquid nitrogen temperatures under a high vacuum in a Tecnai G² Polara electron microscope (FEI) fitted with a 300 kV field emission gun and a USC 4000 CCD camera (Gatan). Electron micrographs were initially analysed using the Image J software (National Institutes of Health, USA).

2.4.3. NS-EM

NS-EM was used to assess the homogeneity of DDM-purified CFTR particles. This technique could not be used for LPG-purified CFTR because protein would not adsorb to the surface of the grid in LPG. DDM-purified CFTR was diluted in SEC buffer to 10 µg/ml and 5 µl applied to a 400-mesh continuously carbon-coated copper grid (Electron Microscopy Sciences, CF-400-Cu) that had been glow discharged for 25 sec at 25 mA. The grids were then washed with 2 x 5 µl mH₂O for 10 sec and stained with 5 µl 4% UA (w/v) for 30 sec; grids were blotted with on filter paper (Whatman, 1001090) between each step. Grids were then analysed on a Tecnai 12 Biotwin Transmission Electron Microscope (FEI) and micrographs collected with an Orius charge-couple device camera (Gatan). Electron micrographs were initially analysed in the Image J software (National Institutes of Health, USA).

2.4.4. SPA

SPA was used to analyse CFTR protein particles from electron micrographs and provide information regarding homogeneity and low-resolution structural features of the protein. Image processing was performed using the EMAN 2 software suite (National Centre for Macromolecular Imaging), as described in the following instruction guide: http://blake.bcm.edu/emanwiki/EMAN2 (Ludtke et al., 1999). For the LPG-purified CFTR, 1056 CFTR particles were selected from 5 micrographs using a box size of 48-pixel (= 183.84 Å). Due to the uneven distribution of particles in the ice, well-separated particles were selected manually on the criteria that there was no overlap with other particles within the same box. For DDM-purified protein, 3962 particles were manually selected using a 64-pixel (= 332.8 Å) box size. Particles were selected based on the criteria they were small enough to fit within this box size without overlapping other particles, thereby selecting smaller CFTR particles rather than larger aggregates. All micrographs were CTF corrected using the “interactive tuning” program in EMAN2. After CTF correction, particles were averaged and classified into 2D sets using 8 levels of iteration, in the “e2refine 2d” program in EMAN2. No further SPA was performed.
due to the small size of LPG-CFTR particles and the heterogeneity of DDM-CFTR particle size.

2.4.5. 2D crystallographic analysis

2D crystallisation was performed to assess the quality of CFTR homogeneity. 2D crystals were grown by Dr Mark Rosenberg using the epitaxial crystallisation procedure used previously to crystallise human-CFTR purified from BHK cells (Rosenberg et al., 2004). In summary, DDM-purified CFTR at 100 µg/ml was mixed with an equal volume of crystallisation buffer (50 mM Tris pH8, 0.05 % DDM (w/v), 10 % PEG 4000 (w/v) and 100 mM ammonium sulphate) and applied to a 400-mesh/inch continuously carbon-coated gold grid (TAAB). The protein was concentrated on the grid by vapour diffusion against a droplet of 1 M MgCl$_2$ at 4 °C o/n. The grid was analysed by NS-EM, as described above (2.4.3). Micrographs containing 2D crystals were analysed in the 2dx software, downloaded from: http://www.2dx.unibas.ch/ (Gipson et al., 2007). The crystal lattice parameters were refined from the computed Fast Fourier Transform (FFT) and the CTF corrected. Lattice unbending was also performed to improve the resolution of the crystal. Finally, a first generation projection map was generated using P$_1$ symmetry with a cut-off resolution of 1/20 Å$^{-1}$.

2.5. Methods: activity

2.5.1. Reconstitution

Reconstitution was performed to remove purified CFTR from a detergent micelle into a more native-like lipid environment and stimulate ATPase activity (Rothnie et al., 2001). First, lipid stocks were made by the following method: *E.coli* lipids were mixed with cholesterol at a ratio of 4:1 (w/w) and dissolved in a 1:1 (v/v) mix of methanol and chloroform, which was then dried in a glass vial under a stream of N$_2$ gas for 2 h to give a transparent lipid film. The lipids were dissolved to a final concentration of 40 mg/ml in SEC buffer containing 0.05 % NaN$_3$ and no NaCl by vortex mixing and sonication until the cloudy suspension became transparent. Oxygen was first removed from the buffer by passing nitrogen gas through the buffer for 20 min at RTM. Lipid stocks were made fresh and kept at 4 °C until needed.

Purified CFTR was diluted in SEC buffer containing 0.05 % NaN$_3$ and lipids added so that the final concentration of CFTR was around 0.1 mg/ml and the LPR was 100:1 (w/w). A lipid-only control was also set up that contained equal volumes of SEC buffer.
in place of the purified protein. This mixture was incubated at 4 °C for 1 h. SM2 BioBeads were washed in 5 CV mH₂O, 5 CV 70 % ethanol (v/v), 5 CV mH₂O, 5 CV SEC buffer (without detergent) and added to the reconstitution sample at 200 mg beads per ml sample o/n at 4 °C. The sample was collected from the BioBeads using a thin-ended pipette and analysed by SDS-PAGE to assess efficiency of reconstitution. The reconstitution sample was also analysed by NS-EM using the same method described earlier (section 2.4.3.) except the grid was washed with SEC buffer (no detergent) instead of mH₂O to prevent liposomes bursting by osmotic shock.

2.5.2. PKA Treatment

In some cases, purified CFTR was treated with the catalytic subunit of PKA before reconstitution to establish the effect of phosphorylation on ATPase activity. PKA and ATP were added to purified CFTR to a final concentration of 200 nM PKA and 5 mM ATP, and incubated o/n at 4 °C. A control was similarly set up in the absence of PKA. To remove the PKA, CFTR was bound to a nickel-affinity chromatography column and purified as described before (section 2.2.12). Eluted CFTR was washed into SEC buffer on a PD10 desalting column (GE Healthcare, 17-0851-01) in order to remove the imidazole, which can interfere with CFTR ATPase activity (personal communication with Dr Eckford, University of Toronto).

2.5.3. ATPase activity assay

The CFTR-specific ATPase activity was determined using a modified Chifflet assay in a 96-well plate format (Chifflet et al., 1988, Rothnie et al., 2001). Purified or reconstituted CFTR was incubated with 1:100 (v/v) IC on ice for 10 min. For the inhibitor experiments, samples were next treated with 100 µM GNS, 10 µM CFTR₉₋₁₇₂ or 1-10 mM GSH on ice for 10 min. At least 5 µg of CFTR in 30 µl was added in per well in the 96-well plate. Controls containing buffer or empty liposomes were similarly set up for later background subtraction of phosphate in purified and reconstituted CFTR, respectively. All experiments were set up in triplicate. A stock of 5 mM Na₂ATP was made in ATPase buffer immediately before use and 20 µl added to all wells to a final concentration of 2 mM ATP. In some experiments, CFTR was mixed to different final concentration of ATP (0, 0.2, 0.4, 0.8, 1.25, 1.75 mM). Na₂ATP was used instead of MgATP due to it having less phosphate contamination, which can interfere with the Chifflet assay; Mg²⁺ ions were added separately to the assay as part of the ATPase buffer. Sodium phosphate stock solution (0.65 mM) was used to prepare 0–
20 nmol phosphate in a final volume of 50 µl mH₂O as standards, which were henceforth treated in the same way as the CFTR-containing wells. The plate was incubated at 25 °C for 1 h and the reaction then stopped by adding 40 µl of 10 % SDS (w/v) at RTM. For the time course experiment, CFTR was incubated with a final concentration of 2 mM ATP at 25 °C and 50 µl aliquots taken at different time points (0, 1, 2, 5, 10, 20, 30, 60, 120 and 180 min), which were then added to the 96-well plate and stopped with addition of 40 µl of 10 % SDS (w/v) at RTM. To all wells, 100 µl of Buffer A was added at RTM for 5 min, followed by 100 µl Buffer B at 37°C for 15 min. The plate was measured for absorbance at a wavelength of 800 nm in a 96-well plate-compatible UV/Vis spectrophotometer (Synergy H1 Hybrid Reader, BioTek). The absorbance at 800 nm was converted into amount of liberated phosphate using the phosphate standards and the rate of ATP hydrolysis calculated after subtracting the background signal (liposome-only or buffer-only control wells), in terms of nmol Pᵢ per min per mg of CFTR protein.

2.5.4. Graphical and statistical analysis of ATPase activity data

All analysis was carried out in GraphPad Prism 6 software. The rates of ATPase activity were plotted as the mean of triplicate results ± the standard deviation (SD) after background subtraction of the mean values from the relevant buffer-only or liposome-only controls. Statistical analysis was performed using the Kruskall-Wallis, Dunn’s post-hoc test, to determine if the ATPase rates were significantly different when compared to the control (p<0.05). Data was assumed to be not parametric.
Chapter 3 – Results and Discussion: Expression and Purification

The current strategy of the CFFT Structure Consortium is to screen different expression systems and CFTR orthologues (Table 1), to produce purified CFTR protein suitable for high-resolution structural studies. Our lab uses the yeast *S. cerevisiae* to express CFTR following the GFP-fusion tag method described by Drew *et al.* (Drew *et al.*, 2008). We have recently published a method for expression of mouse-CFTR in *S. cerevisiae* cells (O’Ryan *et al.*, 2012), but this thesis focused on the expression, purification and characterisation of chicken-CFTR. The focus was changed to chicken-CFTR because, after subsequent purification steps, higher yields of purified protein were consistently achieved compared to mouse-CFTR. The use of the chicken orthologue may be a crucial factor in successful expression and purification of CFTR, although a direct comparison between orthologue behaviour is beyond the scope of this thesis. The method for purification of chicken-CFTR is soon to be published (Pollock *et al.*, 2013, in-press). Along with this method, our CFTR-containing yeast plasmids, CFTR-expressing yeast cells and purified CFTR protein are now being released as reagents to the CF research community.

3.1. Establishing CFTR-expressing yeast cell lines.

The gene for chicken-CFTR was cloned into 2 µ GFP-fusion p424GAL1 vector by homologous recombination and transformed into *S. cerevisiae* FGY217 cells, using a rapid, single-step LiAc/SS carrier DNA/PEG method (Figure 4) (Drew *et al.*, 2008, Gietz and Schiestl, 2007). The option of such a rapid cloning step gives *S. cerevisiae* a great advantage, for example, over the yeast *P. pastoris*, where the gene has to be cloned into *E. coli* first (Cregg *et al.*, 2000). Individual colonies were randomly selected from the transformation plate and grown in separate small-scale cultures. Figure 5 shows SDS-PAGE analysis of crude membranes prepared from 5 x transformation colony cultures. The gel was scanned under fluorescence in order to specifically distinguish a CFTR band via its C-terminal GFP-tag (Figure 5, left). In lane 4, a fluorescent band is visible between the 130 and 250 kDa molecular weight band markers, which corresponds to the estimated 220 kDa weight of our CFTR-GFP fusion protein. There are no CFTR bands visible in the other gel lanes, giving a transformation efficiency of 1 in 5 colonies. It was therefore important to screen a number of colonies to detect CFTR expression. Heterogeneity of CFTR expression between transformants could be due to different copy numbers of the 2 µ plasmid (Cereghino and Cregg, 1999). A chicken-
CFTR expressing cell line was established from the highest expressing colony and used for all cell culture work presented in this thesis. The CFTR gene was extracted and analysed by the DNA sequencing facility at the University of Manchester, and confirmed to be the full-length chicken-CFTR plus fusion tags, with no new mutations. After Coomassie-staining the gel, it was not possible to distinguish a CFTR band from the other endogenous yeast membrane proteins (Figure 5, right), showing relatively low levels of CFTR expression, which is typical for over-expression of membrane proteins (Wagner et al., 2006).

![Transformant Colonies](image)

**Figure 5. CFTR expression was detected in 1/5 transformation colonies.** Transformation colonies (1-5) were grown in separate 50 ml cultures and CFTR expression induced by incubation with galactose for 16 h. Crude membranes were prepared and resuspended in 40 µl SDS-load dye before loading on an 8 % SDS-PAGE gel, shown under GFP fluorescence (left) and Coomassie stain (right). Molecular weight markers are indicated on the left in kDa. CFTR protein is detected on the fluorescent gel in lane 4 as a white band between the 250 and 130 kDa markers (left). Data was collected working with Dr Liam O’Ryan (Ford lab, University of Manchester).

### 3.2. Optimisation of yeast cell growth conditions for CFTR expression.

High-resolution structural studies require large amounts of target protein and one strategy to achieve this yield is to produce large biomass of expressing cells. CFTR-expressing *S.cerevisiae* cells were therefore cultured on a large-scale in an 18 L fermenter reactor to maximize cell biomass. Figure 6A shows a yeast cell growth curve during fermenter culture. Pre-induction, cells grew at an expected doubling rate of 2 h (Drew et al., 2008). Post-induction, however, there was little further increase in cell
growth, especially considering the excess of galactose sugar added as a nutrient source. The final OD600 was 2, which is equivalent to $6 \times 10^7$ cells/ml, lower than the possible $2 \times 10^8$ cells/ml density possible for S. cerevisiae growth in minimal media (Bergman, 2001). Induction of CFTR over-expression could be toxic to the yeast cells because of increased metabolic burden (Bonander et al., 2009), up-regulation of cell stress responses (Brodsky and Skach, 2011) or, if CFTR is active, by increasing cell permeability to chloride ions. The use of a tightly regulated, inducible promoter (GAL1) thus appears important to obtain sufficient yield of cells before CFTR expression is induced and such toxicity takes effect. The reduction in cell growth rate could also be accounted for by the drop in temperature to 20 °C at induction. A lower induction temperature was chosen to encourage successful protein folding by slowing the protein expression machinery (Bonander et al., 2005), which may be critical for the slow folding kinetics of CFTR (Oberdorf et al., 2005) and increase trafficking to the plasma membrane (Denning et al., 1992).

CFTR was previously shown to have a short half-life when expressed in yeast cells due to degradation by host yeast proteases (Kiser et al., 2001). To determine if this was similar in our yeast system, CFTR expression levels were monitored at different times post-induction (Figure 6B). CFTR expression was first detected after 8 h, which correlates with the plateau in OD600 in Figure 6A. Again, suggesting CFTR expression is the cause for decrease in yeast cell growth post-induction. The highest level of CFTR expression was seen at 18 h post-induction. After 18 h, CFTR levels decreased with no CFTR detected after 20 h, perhaps due to up-regulation of yeast degradation machinery (Fu and Sztul, 2009). There are a number of GFP-fused C-terminal lower molecular weight fragments visible at 18 h that may represent such degradation of CFTR (Figure 6B, left). Subsequently, all cell cultures were harvested at 16 h post-induction and immediately chilled to 4 °C to prevent CFTR turnover. The wet weight yield of cells is 60 g per 18 L fermenter, or 3.5 g/L culture. Combined, the reduction in cell growth post-induction (Figure 6A) and the short window for cell harvest time (Figure 6B) limits the yield of CFTR-expressing cells achievable from our system.

Chemical chaperones were next tested to try and improve the yield of CFTR per cell (Newstead et al., 2007a, Drew et al., 2006). Figure 6C shows the addition of 8% glycerol in the induction media increased CFTR expression levels by 5.5-fold compared
Figure 6. Optimisation of yeast cell growth conditions for CFTR expression.

A. Yeast cell growth decreases after induction of CFTR expression. The OD600 was monitored throughout yeast cells growth, both before (left of dotted line) and after (right of dotted line) induction by 2 % galactose (right of dotted line). Before induction, the OD600 was taken from one starter culture. At induction, this starter culture was split into 3 cultures and the OD600 is thereafter shown as an average from triplicate results.

B. CFTR is expressed at highest levels 18 h post-induction. Cells were collected at different times post-induction and crude membranes prepared. Membranes were analysed by SDS-PAGE and in-gel detection of GFP fluorescence tag (left), which shows a CFTR band migrating at around 220 kDa. A band at approximately 70 kDa is probably succinate dehydrogenase, a yeast protein with intrinsic fluorescence properties (Robinson and Lemire, 1992). The CFTR protein band intensity was quantified using Image J software and plotted against time (right). CFTR expression was observed between 8 and 20 h after induction, with peak expression at 18 h.

C. Glycerol increases the expression levels of CFTR in yeast. CFTR expression was induced in the presence (+) or absence (-) of 8 % glycerol (v/v). Membranes were prepared and analysed by SDS-PAGE, with 25 µg microsomes loaded per well. The gel is shown under GFP fluorescent conditions and the molecular weight indicated on the left of the gel. CFTR migrates as a fluorescent band between the 250 and 130 kDa marker. There is 5.5-fold more CFTR in membranes prepared from cells induced with glycerol (+) than cells grown without glycerol (-), as judged by integration of the protein band fluorescent intensity in the Image J software.
to cultures grown in its absence, as judged by integration of the CFTR-GFP fluorescent band intensity. The wet cell yield remained the same in the absence or presence of glycerol, indicating glycerol increases CFTR expression levels by mechanisms other than simply providing an additional food source and increasing cell growth. Glycerol has been shown to increase the expression of the other membrane proteins in yeast cells (Figler et al., 2000) and to rescue CFTR folding defects in mammalian tissue culture cells (Sato et al., 1996, Brown et al., 1996).

3.3. Assessing CFTR cellular localisation by GFP-fluorescence microscopy

Yeast have a quality-control system for membrane protein folding, where only correctly folded proteins can exit the ER to the PM and otherwise misfolded proteins get targeted for degradation by ERAD pathways (Brodsky and Skach, 2011). Plasma membrane localisation is therefore a good indication of correct membrane protein folding and this can be easily checked by fluorescence microscopy of a GFP-fusion tag (Newstead et al., 2007a). Additionally, the GFP-fusion tag can itself act as a folding indicator because it should only fluoresce if the upstream protein is correctly folded and inserted into the membrane with correct topology (Drew et al., 2001). To check CFTR cellular localisation, cells were analysed by fluorescence microscopy before and after induction of CFTR expression (Figure 7). Images of pre-induced cells (Figure 7, left) show a uniformly distributed auto-fluorescence, probably due to the presence of naturally occurring fluorescent molecules in yeast (Frey et al., 2011). The increase in yeast fluorescence seen post-induction (Figure 7, right) signifies GFP-fused CFTR expression. It is estimated that 0.1 µM of over-expressed GFP is required to overcome the background fluorescence of yeast cells (Tsien, 1998). This fluorescent signal suggests that CFTR is folded and not in inclusion bodies (Drew et al., 2001). Some GFP fluorescence appears around the cell periphery indicating there may be CFTR PM localisation (Drew et al., 2008) (Figure 7, right). There are also highly fluorescent discrete spots located at, or near to, the PM, which could be compartments of the ER/Golgi/endosome trafficking and recycling pathways (Yoo et al., 2002, Kiser et al., 2001). This “spotted” CFTR-GFP localisation is very similar to the stable PM clusters observed with the expression of ABCA1 in yeast (Bocer et al., 2012). ABCA1 was shown to be fully functional within these membrane sub-domains.
3.4. Isolating CFTR-containing microsomes from yeast cells.

Microsomes is a term used for membrane-vesicles that are re-formed from PM and ER membrane fragments after cell breakage. CFTR is localised to such membranes, as judged by fluorescence microscopy (Figure 7). Our microsome preparation method was adapted from a method developed to make Pgp containing microsomes (Lerner-Marmarosh et al., 1999). Briefly, microsomes were sedimented following yeast cell breakage using a series of centrifugation steps that removed unbroken cells, soluble proteins, and intact nuclei and mitochondrial membranes. The total microsomal protein yield was determined by Bradford assay (Bradford, 1976) as 4.2 mg / g cells, which is comparable to the yield previously described (Lerner-Marmarosh et al., 1999), suggesting efficient cell breakage and recovery of total microsomal proteins. The recovery of CFTR throughout the procedure is also very efficient, as judged by SDS-PAGE analysis (Figure 8A). Analysis of microsomes by fluorescence spectroscopy shows a clear GFP fluorescence emission peak, indicating the microsomes are enriched in CFTR (Figure 8B). The yield of CFTR was estimated from the GFP fluorescence emission peak using a GFP calibration curve (Appendix 1) as 8 µg / mg microsomes or 120 µg / L cell culture. Per cell, the CFTR yield is comparable to other membrane protein yields in yeast, but per litre of culture the yield is relatively low, probably due to
Figure 8. Isolation of CFTR-containing microsomes.

A. Microsomes are prepared by a series of centrifugation steps with efficient recovery of CFTR in the final pellet. Cells were broken and the lysate centrifuged at 14 000 g to pellet unbroken cells, nuclei and mitochondria membranes. Soluble proteins were then discarded from supernatants after centrifugation at 200 000 g and 100 000 g, and microsomes collected in the pellet. The supernatants (S) and pellets (P) were adjusted to comparable volumes and analysed by SDS-PAGE, shown under Coomassie (left) and GFP fluorescence (right), with 50 µg of total microsomal protein loaded in the lane labelled 100K g pellet.

B. CFTR can be detected in microsomes by fluorescence spectroscopy analysis of the GFP-tag. Microsomes at a total protein concentration of 0.5 mg/ml were analysed in a Carys Eclipse Fluorescence Spectrometer, with excitation at 485nm and emission spectra collected between 500 and 600 nm. The GFP fluorescence has a peak emission at 512 nm.

C. Protease inhibitors are required for the integrity of full-length CFTR in microsomes. Microsomes were prepared in the presence (+) or absence (-) of protease inhibitors and 25 µg loaded onto SDS-PAGE gel for analysis by in-gel fluorescence of the GFP-tag. Full-length CFTR is apparent in the presence of inhibitors, whereas only smaller GFP-tagged C-terminal fragments are visible in the absence of inhibitors, including a band for GFP at 25 kDa.
the limitation in cell biomass achievable in our system, as described in section 3.2 (Lerner-Marmarosh et al., 1999, Drew et al., 2008). Proteolysis is a major obstacle when trying to express CFTR in yeast (Zhang et al., 2002, Kiser et al., 2001). The integrity of CFTR in the microsomes can be assessed by fluorescence SDS-PAGE analysis, which shows C-terminal fragments still fused to the GFP-tag. The addition of protease inhibitors proved crucial to prevent CFTR proteolysis. Full-length CFTR is highly degraded in microsomes prepared in the absence of inhibitors (Figure 7C). The cocktail of inhibitors used here can inhibit a wide range of yeast proteases including, serine, cysteine, trypsin, aspartyl- and metallo-proteases. We found that this cocktail of inhibitors was more effective at limiting proteolysis than commercially available inhibitor cocktail tablets (data not shown). Other factors may also be important in overcoming CFTR proteolysis problems, for example a short harvest time and lower temperature of induction, as described in section 3.2.

3.5. Post-translational modification of CFTR in yeast

During protein synthesis, CFTR receives post-translational modifications (PTMs) that are important for folding, stability, and biological activity (Chang et al., 2008, McClure et al., 2012, Seibert et al., 1999). Yeast can perform many eukaryotic PTMs, but do not always correctly process heterologously expressed proteins from higher eukaryotes (Cereghino and Cregg, 1999). CFTR from our yeast system was therefore analysed for PTMs as an indication that CFTR was correctly recognized and processed by the cell, and most likely correctly folded and biologically active. Information regarding CFTR PTMs is also useful for future crystallization trials where homogeneity in modifications is important for successful crystal formation (Reeves et al., 2002). Firstly, the glycosylation state of CFTR was probed using an enzyme, PNGaseF, which is an endo-glycosidase that cleaves most N-linked glycans (Mellors and Sutherland, 1994). If CFTR is extensively N-glycosylated then a shift in gel mobility should be observed after PNGaseF treatment, as shown previously for CFTR expressed in mammalian cells (Zhang et al., 2009). CFTR-containing microsomes were treated with and without PNGaseF at 20 °C o/n under denaturing conditions, and then analysed by fluorescence SDS-PAGE (Figure 9A, left). In the absence of PNGaseF treatment, CFTR was unable to enter the SDS-PAGE gel, and can be seen as a fluorescent band in the top of the gel well. This was expected as CFTR tends to form such SDS-insoluble aggregates during thermal denaturation (see section 4.2.4 for more detail) (Engel et al., 2002). After treatment with PNGaseF, however, CFTR was able to enter the SDS-PAGE gel as a
monomeric protein band. This change in gel mobility after PNGaseF treatment suggests that CFTR is, at least core, glycosylated in yeast. These data also show that the removal of N-glycosylation prevents the formation of SDS-insoluble aggregates that are normally observed after incubating CFTR at increasing temperature (section 4.2.4). The glycosylation state of CFTR may therefore affect its thermal stability, as has been shown with other glycoproteins (Wang et al., 1996). A shift in gel mobility was not observed under non-denaturing conditions, suggesting the N-linked glycosylation sites are not very accessible in the native-folded CFTR structure (Figure 9A, right).

Secondly, the phosphorylation of CFTR was probed using Pro-Q Diamond SDS-PAGE gel stain (Life Technologies), a stain that binds to phospho-proteins and can be detected in the gel under UV fluorescence (Steinberg et al., 2003). For this characterisation, CFTR was first purified in DDM, as described in section 3.8, to eliminate co-staining of contaminating yeast phospho-proteins. CFTR was treated with no enzyme, PKA or alkaline phosphatase from calf intestine (CIP) o/n at 4 °C, and subject to SDS-PAGE analysis and Pro-Q diamond stain (Figure 9B). With no enzyme treatment, CFTR can be detected by Pro-Q Diamond stain, suggesting CFTR is phosphorylated in our yeast system. Such phosphorylation can be removed by enzymatic cleavage, as shown by a loss of the CFTR fluorescent band after treatment with CIP (Figure 9B). Interestingly, treatment with PKA gave no change in intensity or gel mobility of Pro-Q Diamond stained CFTR band, suggesting that CFTR cannot be further phosphorylated by PKA under these conditions (Figure 9B). This differs from CFTR purified from mammalian and insect cells, which was mostly de-phosphorylated and required PKA treatment in vitro to stimulate activity (Zhang et al., 2009, Li et al., 1996). Recent mass spectrometry analysis also revealed that CFTR from our yeast system displays additional phospho-residues to those observed in CFTR from other cell systems (Venerando et al., 2013). It seems that the CFTR produced here has higher levels of phosphorylation than previously reported, perhaps because methods used to limit CFTR degradation during expression in yeast (section 3.2) allowed these extra phospho-residues to survive before being processed by the cell.
**Figure 9. PTM analysis of CFTR expressed in yeast.**

**A. CFTR is core N-glycosylated in yeast.** CFTR-containing microsomes were treated with (+) or without (-) PNGaseF in either denaturing (D) or non-denaturing (ND) conditions at 20 °C o/n and analysed by SDS-PAGE, which is shown under GFP-fluorescence conditions. During this experiment, CFTR formed SDS-insoluble aggregates that could not enter the top of the gel (lanes: D -, ND +/-). De-glycosylation by PNGaseF under denaturing conditions prevented aggregate formation and allowed CFTR to enter the gel as a monomeric protein band (D +).

**B. CFTR is phosphorylated in yeast.** Purified CFTR was treated with no enzyme (-), PKA or CIP at 4 °C overnight and analysed by SDS-PAGE gel. Total CFTR is shown by the GFP-fluorescence scan of the gel (left) and phosphorylated CFTR is shown by Pro-Q Diamond staining of the gel (right). CFTR appears to be similarly phosphorylated with and without PKA treatment and this phosphorylation can be removed by CIP enzyme.
3.6. Detergent solubilisation of CFTR from yeast microsomes

Detergent solubilisation is the first stage in the purification of a membrane protein that then allows individual proteins to be separated by purification techniques commonly used for soluble proteins (Carpenter et al., 2008). High solubilisation efficiency is desired to achieve the yields of protein required for high-resolution structural studies. The efficiency of detergent solubilisation is often protein specific and varies with types of detergent used. To assess the efficiency of CFTR solubilisation from our yeast system, microsomes were incubated with detergent at a concentration greatly above the CMC and the detergent-soluble material collected after ultracentrifugation (Gutmann et al., 2007). Both the detergent-soluble and -insoluble material were analysed by SDS-PAGE and the solubilisation efficiency calculated by comparing the GFP-tag fluorescence band intensity in each fraction. Detergents were chosen based on those previously used to successfully purify and crystallise CFTR and other ABC transporter proteins (Oldham et al., 2007, Pinkett et al., 2007, Locher et al., 2002, Aller et al., 2009, Rosenberg et al., 2004, Huang et al., 1998, Shintre et al., 2013). Figure 10 shows the solubilisation efficiency of CFTR for different detergents. CFTR can be extracted in all detergents to some extent under these conditions. The highest solubilisation efficiency is seen with the charged or zwitterionic detergents (e.g. FC16 and LPG14 at around 90 %) compared to the efficiency with non-ionic detergents (e.g. maltosides at around 40 %). Harsh detergents, however, run the risk of destabilising the protein structure and function (Privé, 2007). Choosing a detergent for purification is therefore an important balance between getting high enough solubilisation efficiency for sufficient protein yield whilst maintaining the protein in a folded and active state. For this reason, both the highly solubilising LPG and the mild detergent DDM were selected for subsequent CFTR purification steps, which had solubilisation efficiencies of 88 % and 46 %, respectively. It is important to note that the best detergent for extraction and purification may not correlate to the best detergent for crystallisation and that the best choice of detergent is often protein specific (Rosenow et al., 2002, Privé, 2007, Sonoda et al., 2011). The data presented here suggest that CFTR could be exchanged into a range of detergent types for later crystal screens if required for later experiments.
Figure 10. CFTR can be extracted from yeast microsomes in a range of detergents with varying solubilisation efficiency.

CFTR-containing microsomes at total protein concentration of 5 mg/ml were incubated with 5 % (w/w) detergent for 1 h at 4 °C, to a detergent to total protein ratio of 10:1 (w/w). Detergents included: fos-choline 16 (FC), lyso-phosphatidylcholine 14 (LPG), lauryldimethylamine oxide (LDAO), decyl-maltoside (DM), dodecyl-maltoside (DDM), octyl-maltoside (OG), nonyl-maltoside (NM), octaethylene glycol monododecyl ether (C12E8), and octyl glucoside (OG). Detergent-soluble proteins (S) were then separated from detergent-insoluble (I) material by ultracentrifugation at 100 000 g and analysed by GFP-fluorescence SDS-PAGE (A). The amount of CFTR in each fraction was quantified by Image J analysis of the GFP-fluorescent protein band intensity. The efficiency of solubilisation was calculated as a percentage of the integrated band intensity of S/(S+I), and these values are shown above the relevant bars (B).
3.7. Purification of CFTR in LPG

LPG is a zwitterionic, phospholipid derivative with only one fatty acid acyl chain, which is found naturally occurring in biological membranes (Stafford et al. 1989). LPG is not well characterised in the field of membrane protein crystallisation, but has been used previously to purify and stabilise CFTR in an apparently folded and active state (Zhang et al., 2009, Huang et al., 1998). CFTR was purified from microsomes isolated from an 18 L fermenter culture in two steps: nickel-affinity chromatography and size exclusion chromatography (SEC). Figure 11A shows an SDS-PAGE gel of different fractions throughout the nickel-affinity purification step. CFTR was first solubilised in LPG and then bound, washed and eluted from a nickel HisTrap column via its His tags, which include a C-terminal 8His and an N-terminal 10His sequence. The GFP-tag is shown by in-gel fluorescence (Figure 11A, bottom panel) and is used to specifically track CFTR in the purification. Using Image J to quantify the fluorescence intensity of the GFP-tag, it appears that CFTR was solubilised in LPG with 79 % efficiency, and the binding and recovery of CFTR from the HisTrap column was nearly 100 % efficient. The gel was also stained with Coomassie blue stain to show the total protein content of each fraction (Figure 11A, top panel). The purity of CFTR in the elution was judged at 80 % by Image J quantification of CFTR band intensity, compared to the total intensity of all the stained bands in that fraction.

The nickel-affinity purified material was next injected onto a Superose 6 10/300 column for SEC purification. Figure 12A shows the UV trace elution profile and a Coomassie-stained SDS-PAGE gel of the peak CFTR elution fraction. The purity of CFTR in this peak fraction is judged to be 92 %. The yield was calculated by comparing CFTR band intensity to BSA standards on SDS-PAGE (Appendix 1), and resulted in 2 mg purified CFTR from an 18 L fermenter culture, or 111 µg per L culture. The recovery of CFTR from the membranes is approximately 80 %, largely due to the high efficiency of detergent solubilisation and binding to the HisTrap column. This yield is relatively low compared to other eukaryotic membrane proteins expressed in the same GFP-tagged, yeast expression system (Newstead et al., 2007a), but close to previous yields of CFTR purified from yeast (Huang et al., 1998). Overall, LPG appears to be an efficient detergent for purifying milligram quantities of CFTR to near homogeneity; qualities a protein requires for high-resolution structural studies.
Figure 11. Nickel-affinity chromatography purification of CFTR in LPG (A) and DDM (B). Microsomes were incubated with 2 % detergent (w/v) and the soluble material separated from the insoluble by ultracentrifugation at 100 000 g for 1 h. The soluble material was passed over a nickel HisTrap column to bind CFTR via its His-tags, and the unbound material collected. The column was washed with 3 CV buffer + 40 mM Imidazole and 3 CV buffer + 100 mM Imidazole before eluting with 3 CV buffer + 400 mM Imidazole. The purification was monitored by SDS-PAGE with 10 µl sample loaded per well. The gel is shown under Coomassie stain (top panel) and under GFP-fluorescence conditions (bottom panel). The molecular weight marker is indicated on the left in kDa. A black arrow indicates a band corresponding to CFTR in the elution fraction.
Figure 12. SEC purification of CFTR in LPG (A) and DDM (B).
CFTR from the nickel-affinity purification was injected onto a Superose 6 10/300 column. CFTR eluted between 8 and 11 ml retention volume, as indicated on the UV trace elution profile. Other peaks on the UV trace corresponded to contaminating proteins or buffer components. Protein from the peak elution fraction at 9.5 ml is shown on a Coomassie stained SDS-PAGE gel (10 µl load) where the CFTR protein band is indicated by a black arrow.
3.8. Purification of CFTR in DDM

The second detergent chosen to purify CFTR was DDM. DDM is the most common detergent used in the purification and crystallisation of membrane proteins (Newstead et al., 2008). It has also been used to successfully purify CFTR in an active and folded state compatible with 2D crystallisation (Rosenberg et al., 2004). Figure 11B and 12B shows the analysis of CFTR purified by nickel affinity chromatography and SEC, using a similar method to the purification of CFTR in LPG, described in section 3.7. Purification of CFTR in DDM yielded 60 µg of purified CFTR per L culture with approximately 40 % recovery of CFTR from the microsomes. The lower yield in DDM compared to LPG seems to be caused by the lower detergent solubilisation efficiency and losses during chromatography steps. The purity of CFTR in DDM is also lower than in LPG at 60 % homogeneity, largely due to a single contaminating band at around 45 kDa. This protein was a persistent contaminant that could not be removed by alternative chromatography methods or by filtration through 100 kDa MWCO membrane. This contaminating protein is discussed in more detail in the next section (section 3.9). CFTR can be purified in DDM to a sufficient yield for high-resolution structural studies, but the low purity still remains a large problem for 3D crystallisation trials.

CFTR is purified in DDM using mostly the same purification methods as for LPG (section 3.7). One notable difference is that the DDM buffer is supplemented with 1 M NaCl because it was found to greatly improve the efficiency of purification in this detergent. High NaCl improved the detergent solubilisation efficiency from 40 % (Figure 10) to 62 % (Figure 11B), and allowed 90 % binding efficiency to the nickel-affinity column (Figure 11B). Additionally, high salt seemed to improve CFTR solubility by decreasing levels of CFTR aggregation, as monitored by 90° light scattering measurements (Figure 13). The light scattering increased with decreasing concentrations below 0.6 M NaCl, indicating CFTR aggregation in low salt conditions (Hayashi et al., 1989). Above 0.6 M NaCl, the light scattering remained relatively low, suggesting high salt prevented the aggregation of CFTR. High salt is believed to improve the solubility of membrane proteins by reducing the energetic cost of keeping hydrophobic protein surfaces in solution (Vogel et al., 2001). High salt also decreases the CMC of many detergents, which could further stabilise a membrane protein in solution (Jumpertz et al., 2011). A jump in NaCl concentration from 0 to 1 M results in a 4-fold decrease in the CMC of DDM (Jumpertz et al., 2011). Therefore, at 1M NaCl,
the use of 0.1 % DDM in the purification buffer is likely to be 40-times more than the CMC. The purification buffer also required 20 % glycerol to maintain CFTR stability. Glycerol and other osmolytes are well known to improve protein stability (Gekko and Timasheff, 1981, Street et al., 2006) and their use was important for the purification of isolated NBD domains in a non-aggregated state (Lewis et al., 2010). The high ionic strength, high detergent micelle concentration and concentration of glycerol required to keep CFTR in solution implies the protein is quite unstable in DDM. The LPG method used just 50 mM NaCl in the buffer because LPG already gave a highly efficient purification (Figure 11 & 12), perhaps because the charged head-groups of LPG negates the need for such ionic strength in the buffer. The DDM-buffer thus included 1M NaCl throughout the purification procedure to keep CFTR stable. A number of other ABC transporter proteins were similarly purified in high NaCl concentrations (0.5-1 M NaCl) before 3D crystallisation and structure determination, showing their structures were compatible with high salt (Pinkett et al., 2007, Gerber et al., 2008, Kadaba et al., 2008a). Additionally, NBD1 and NBD2 of CFTR were crystallised using 3.5 M sodium acetate (Lewis et al., 2010) and 1.2 M sodium citrate (PDB = 3DG7) respectively, which implies at least these soluble domains of CFTR are folded in high salt conditions. Still, it is important to note that the 1M NaCl used in the purification buffer is at a higher concentration than the naturally occurring environment of CFTR (Song et al., 2003) and it may later prove difficult to interpret the biological relevance of CFTR characterisation data. High salt is also problematic for certain techniques, e.g. cryo-EM (Rubinstein, 2007).
Figure 13. DDM-purified CFTR appears to aggregate at low concentrations of NaCl.
DDM-purified CFTR was diluted to 20 µg/ml in buffer to different concentrations of NaCl and incubated on ice for 10 min. The 90° light scattering was measured in a fluorometer (excitation = 280 nm, emission = 565 nm). Increased light scattering was used as an indication of protein aggregation. A GFP control was set up by the same method to show that changes in light scattering were not an artefact of NaCl concentrations on the buffer or GFP-tag.

3.9. Identification of contaminating yeast proteins by mass spectrometry

Protein homogeneity is an important quality after purification because any remaining contaminating proteins could interfere with further biochemical and biophysical characterisation. From SDS-PAGE analysis of SEC-purified fractions, CFTR purity appears to be 92 % and 60 % in LPG and DDM, respectively (Figure 12). To identify CFTR and any contaminating yeast proteins, SEC-purified CFTR was analysed by mass spectrometry and the full list of identified proteins are listed in Appendix 2. Firstly, the protein at around 200 kDa visible on both LPG- and DDM-purification gels (Figure 12) was identified from the IPI Chicken (version 3.64) database as chicken-CFTR (IPI00575922), confirming that the protein purified in this method is, in fact, the target CFTR protein. The SEC-purified protein samples were next analysed by mass spectrometry against the yeast protein UniProt (version 2011-05) database. As expected, the majority of contaminants are highly abundant yeast proteins, such as components of the ribosome and proton ATPase. Some identified proteins are known to directly interact with CFTR in mammalian cells during protein folding and degradation pathways, e.g. HSP70 (Yang et al., 1993) and PI3K (Luciani et al., 2010), suggesting CFTR is similarly recognised as a substrate by these proteins in yeast. Other
contaminating proteins possess intrinsic ability to bind metal ions that would enable co-purification on a nickel-affinity column. Interestingly, in LPG there are a number of proteins identified that have a linear His sequences that would otherwise be buried in their folded 3D structure e.g. GTPase-activating protein and adenylate cyclase. To allow co-purification on the nickel-affinity chromatography column, these internal His sequences would have to be exposed to the surface of the protein. Such proteins were not identified in the DDM-purified sample, where only proteins with structurally conformational and accessible metal binding sites were identified, e.g. alcohol dehydrogenase and large ribosomal protein subunit L3. It could therefore be proposed that LPG, at least partially, destabilises protein structure during purification, whereas DDM more likely preserves a protein’s fully folded state. LPG could better expose the CFTR His-tags to nickel-affinity chromatography than DDM, which could partly explain why the purification recovery is far more efficient in LPG than DDM.

In DDM, the most abundant contaminant seen at approximately 45 kDa by SDS-PAGE analysis (Figure 12B) was specifically cut from the gel and identified as the yeast ribosomal subunit L3. Examination of the structure of L3 (PDB=1FFK) shows a potential poly-His cluster in the folded protein that would explain co-purification by nickel-affinity chromatography. Unfortunately, L3 could not be separated from CFTR by SEC purification or by filtration through a 100 kDa sized membrane, despite their greatly differing apparent molecular weights. Another explanation for co-purification could be that L3 directly interacts with CFTR. It has a highly basic, structural “W-finger” peptide extension that penetrates and stabilises rRNA structure by forming interactions with the negatively charged rRNA phosphate groups (Meskauskas and Dinman, 2008). This extension could therefore interact with CFTR, for example via interactions with phospho-residues or acidic regions in the protein. Other chromatography steps that could be used to further purify CFTR from L3 are limited due to the incompatibility of such resins in the high salt required to keep CFTR stable in DDM (Figure 13), for example ion exchange, heparin or hydrophobic interaction chromatography. The contamination of CFTR with L3 remains a limitation in the DDM-purification method and future improvements are needed to produce protein more suitable for high-resolution structural studies.
3.10. Cleavage of C-terminal GFP-fusion tag using TEV protease

The C-terminal GFP-fusion tag has proven a useful tool to easily monitor CFTR throughout the expression and purification method, but its removal might be required for later studies. There is a peptide linker between CFTR and the GFP-fusion tag that contains a TEV protease cleavage site; TEV is a highly specific cysteine protease that cleaves between the Gln and Gly/Ser in the consensus ENLYFQ(G/S) sequence (Kapust et al., 2002). Figure 14 shows the cleavage of GFP-tag from CFTR by TEV, as monitored by in-gel fluorescence. In DDM, there is clear evidence of GFP-tag cleavage shown by a decrease in the fluorescence intensity of the full-length CFTR protein band and an increase in the intensity of cleaved GFP protein band (Figure 14, right panel).

The amount of cleavage was estimated as 90 %, as judged by comparing CFTR band intensity +/- TEV addition. The cleaved GFP-tag appears to migrate as a dimer in the gel, which may be caused by the presence of high salt in the buffer (Yang et al., 1996, Tsien, 1998). After cleavage, the GFP-tag and TEV protease can be separated from CFTR using SEC purification methods but not by filtration through a 100 kDa membrane (data not shown).

According to the manufacturer’s guidelines, the recommended amount of TEV required for 90 % cleavage of a fusion-tag at 4 °C overnight is 0.4 units per µg of target protein. A larger amount of TEV was required (= 36 units per µg of target protein) to observe such a level of cleavage with CFTR-GFP target under the same conditions, perhaps due to steric hindrance of the cleavage site in the CFTR structure. Increasing the concentration of CFTR protein or the incubation time with TEV may increase the efficiency of cleavage, but these conditions do not favour CFTR stability. It is therefore possible to cleave CFTR’s GFP-tag in the presence of DDM but at great cost in TEV protease enzyme.

In LPG, the GFP-tag was not cleaved by TEV under the same conditions used for the DDM-purified CFTR (Figure 14, left panel). TEV cleavage can be affected in the presence of some detergents, for example TEV protease was unable to cleave affinity tags in the harsh detergent, FC (Lundbäck et al., 2008). LPG may similarly destabilise and inactivate TEV protease activity.

Due to these experimental difficulties in TEV protease cleavage, the GFP-tag was left attached to CFTR for subsequent studies. The GFP-tag may not hinder later characterisation of CFTR, for example the addition of a GFP-tag does not affect the
channel gating activity of CFTR expressed in mammalian cells (personal communication with Riordan lab). Also, the presence of a soluble protein tag can improve protein stability and provide more soluble protein interfaces for 3D crystal contacts (Rosenbaum et al., 2007). Nonetheless, it is still important to note that the GFP-tag may interfere with certain characterisation techniques.

Figure 14. CFTR-GFP tag can be cleaved using TEV protease in DDM, but not LPG. CFTR was purified in either DDM or LPG and 0.5 µg incubated with 18 units of TEV protease enzyme overnight at 4 °C. 10 µl of samples were analysed on a 4-12% gradient SDS-PAGE. The gel is shown under GFP fluorescence conditions, with the molecular marker indicated on the left. Full-length CFTR and cleaved GFP protein bands are labeled.
Chapter 4 – Results and Discussion: Stability

Protein stability is an important biophysical characteristic for successful protein crystallisation (Warne et al., 2008, Malawski et al., 2006). All integral membrane proteins of known atomic-resolution structure display high levels of stability (Sonoda et al., 2011, Rosenbusch, 2001). For NBD1 of CFTR, improving protein thermodynamic stability by introducing a series of “solubilising” point mutations was shown to be crucial for crystallisation and structure determination (Protasevich et al., 2010, Wang et al., 2010, Lewis et al., 2005, Lewis et al., 2010). Furthermore, the predominant CF disease-causing mutation (ΔF508) is known to destabilise CFTR protein \textit{in vivo} in a temperature-sensitive manner (Denning et al., 1992, Sharma et al., 2001). Until now, all \textit{in vitro} investigation into CFTR stability has been limited to isolated CFTR domains (Protasevich et al., 2010, Wang et al., 2010, Thibodeau et al., 2005, Lewis et al., 2005, Marasini et al., 2012). This chapter, for the first time, presents data on the stability of purified CFTR protein with the full-length, wild-type protein sequence. Stability measurements are therefore not only useful to assess the quality of our CFTR for crystallisation studies, but could also provide new insights in CFTR study and potential development of CF treatment.

4.1. Protein concentration dependence of CFTR stability.

Three-dimensional crystallisation for high-resolution structural studies requires that a protein is stable at relatively high concentrations. Protein concentrations of 10-15 mg/ml are typically recommended to achieve the supersaturated solutions that induce crystal growth (Chayen and Saridakis, 2008, Newby et al., 2009). Purified CFTR was concentrated using a MWCO 100 kDa spin concentrator, a pore size small enough to retain full-length CFTR particles (220 kDa), but large enough to allow DDM (70 kDa) and LPG (90 kDa) detergent micelles to pass through (Jumpertz et al., 2011, Oliver et al., 2013); it is important that the detergent is not concentrated as this may cause protein denaturation and inhibit any later crystal growth by causing phase separation (Garavito and Ferguson-Miller, 2001). At several points during the concentration process, aliquots were taken and the concentration of CFTR estimated from the GFP fluorescence, which was later confirmed by SDS-PAGE using BSA as a calibration (Appendix 1). The 90° light scattering fluorescence was also monitored throughout the CFTR concentration procedure as a measure of CFTR aggregation (Hayashi et al., 1989), and these data are shown in Figure 15. In DDM, it was not possible to reach a CFTR concentration above
0.4 mg/ml. Above 0.4 mg/ml protein concentration, the retentate volume only reduced by a small degree, despite long periods of centrifugation, and also a white precipitate began to appear. In Figure 15 (left panel), there is a dramatic increase in 90° light scattering between concentrations of 0.3-0.4 mg/ml CFTR, implying a protein aggregation event (Hitscherich et al., 2001, Hayashi et al., 1989). Such aggregation could explain the halt in CFTR concentration due to blockage of the centrifugal filter. This limit on CFTR concentration in DDM remains a major obstacle for biophysical studies and crystallisation trials.

In LPG, CFTR was concentrated to approximately 8 mg/ml without any experimental difficulty. A steady increase in light scattering was observed that appears proportional to the increase in protein concentration, implying no aggregation event as was seen with the DDM-purified protein (Figure 15, right). It therefore appears that CFTR is more soluble at high protein concentrations in LPG than in DDM because LPG prevents the formation of protein aggregates. This is consistent with previous work, where purified CFTR was limited to 0.3 mg/ml in DDM, but was still soluble at 4 mg/ml in LPG (Zhang et al., 2009). At 8 mg/ml, this is the highest reported concentration for purified CFTR from yeast or any other expression system and is, for the first time, in a suitable range for crystallisation and high-resolution structural studies.

**Figure 15. CFTR is more soluble at high concentrations in LPG than DDM.**

Purified CFTR in either DDM or LPG were concentrated using a 100 kDa MWCO centrifugal filter. Aliquots were taken at different points during the concentration process and the GFP fluorescence (excitation 485nm, emission = 512nm) and 90° light scattering (excitation = 290nm, emission = 565nm) measured in a fluorometer. CFTR concentration was estimated from the GFP fluorescence, using the GFP standard curve, and the final protein concentration was confirmed by BSA standard SDS-PAGE analysis (Appendix 1).
4.2. Thermal stability of CFTR

There is a good correlation between protein thermal stability and production of good diffraction-quality crystals, and thermo-stability screens have proved successful for identifying conditions for the crystallisation of a number of membrane proteins (Serrano-Vega et al., 2008, Warne et al., 2008). Probing the thermal stability of an integral membrane protein, however, can be experimentally challenging. Techniques that are commonly used for soluble proteins, such as differential scanning calorimetry (DSC), require relatively large amounts of protein at concentrations that are difficult to achieve for poorly expressed and unstable proteins like CFTR (Weber and Salemme, 2003). The presence of detergent can often mask thermal unfolding transitions that rely on the detection of the exposure of hydrophobic protein residues upon unfolding (Kean J, 2008, Yeh et al., 2006). Detergents can also preserve levels of membrane protein α-helical structure during thermal unfolding, making it difficult to detect unfolding transitions by techniques that monitor secondary structure folds, e.g. circular dichroism (Brouillette et al., 1987). Therefore, a number of techniques that are deemed more compatible with membrane proteins were selected to elucidate the thermal stability of CFTR.

4.2.1. CPM dye binding assay

Recently, an assay has been established for assessing the thermal stability of membrane proteins using a fluorescent dye called N-[4-(7-diethylamino- 4-methyl-3-coumarinyl)phenyl]maleimide (CPM) (Alexandrov et al., 2008). CPM is a thiol-specific fluorochrome that only fluoresces in its thiol-bound state (Ayers et al., 1986). As cysteine (Cys) residues become surface exposed during thermal denaturation, there will be an increase in CPM fluorescence. The advantages of the CPM assay are that it is requires relatively small amounts of protein, is compatible with a range of detergent types and performs at high signal-to-noise ratio (Lee et al., 2013, Alexandrov et al., 2008, Sonoda et al., 2011).

Chicken-CFTR has a total of 15 Cys residues, which, using a Blast protein sequence alignment, were all found to be conserved with Cys locations in the human-CFTR primary amino acid sequence (Appendix 3). Currently, there are no structural data for the chicken-CFTR orthologue, so to predict the location of its Cys residues, the chicken-CFTR sequence was compared to the available structural model of human-CFTR (Mornon et al., 2008) (Appendix 3). The inward-facing model was used because
CFTR was purified in the absence of ATP and is assumed to be in this conformation (Zhang et al., 2009, Gadsby et al., 2006, Vergani et al., 2005). Although in the absence of ATP, the CFTR channel can also be found in the open state (Rosenberg et al., 2011, Kirk and Wang, 2011). In either case, the accessibility of cysteines appears to be the same for both the outward- and inward-facing model (Appendix 3). It is important to note that this model is only a static snapshot of a CFTR structure and, in reality, the degree of Cys exposure may fluctuate due to structural flexibility. CFTR flexibility is particularly expected here due to the absence of lateral pressure from the plasma membrane when in the detergent-soluble state (Marsh, 1996). Still, a prediction can be made that chicken-CFTR has 4 surface-exposed Cys and 11 Cys that are at least partly buried (Appendix 3).

The CPM dye binding assay proved suitable to monitor CFTR thermal unfolding. Figure 16A shows an example of CPM fluorescence during a thermal melt experiment of DDM-purified CFTR (blue line). A pre-incubation with CFTR and CPM at 10 °C for 30 min shows an increase in fluorescence that is probably due to the kinetics of labelling of surface-exposed Cys. The plateau of fluorescence after this initial incubation indicates CFTR remains structurally intact during a short incubation at low temperatures. A subsequent increase in temperature to 80 °C sees a second increase in fluorescence that is likely to be due to the labelling of natively buried Cys that become exposed during thermal unfolding. These data show a two-state cooperative unfolding transition profile, where the natively folded and denatured proteins are represented by the bottom and top plateau, respectively. These data provide evidence that purified CFTR has a folded, cooperative tertiary structure at 10 °C that allows the seclusion of a significant proportion of its Cys residues. CFTR displays a broad unfolding transition between 25 and 60 °C, which is probably due to the large number of buried Cys in CFTR, each exposed at different temperature points during the sequential unfolding of multiple domains combined with some additional kinetic element for the labelling reaction which may cause some hysteresis in the CPM fluorescence versus temperature curve.

A second heating scan of the denatured protein gives a trace devoid of the transition seen in the first scan (Figure 16A). Instead, the CPM fluorescence shows a negative slope, which is probably due to thermal quenching of the CPM fluorophore (Ayers et al., 1986, Alexandrov et al., 2008). This suggests CFTR thermal unfolding is an irreversible process that can be defined as $N \rightarrow D$, where N is natively folded and D is the
denatured protein. This unfolding transition profile is specific to CFTR and not a contribution of its fusion-tags or the contaminating protein, ribosomal L3, as shown by the unfolding profiles of these isolated proteins (Appendix 4); L3 only has 2 surface-exposed Cys that are labelled during the initial incubation, GFP does not thermally unfold until above 80 °C, and SUMO has no Cys in its primary sequence.

The increase in CPM fluorescence is supposedly proportional to the number of thiol-CPM adducts that are formed (Ayers et al., 1986), so could theoretically be used to predict the number of Cys that are exposed and buried in a native-folded protein structure. From the Mornon model of CFTR structure, one would expect a 3.75-fold increase in the CPM fluorescence during thermal unfolding, assuming all of the predicted 4 surface-exposed Cys are labelled during the initial incubation at 10 °C and the remaining 11 buried Cys are labelled after heating to 80 °C (Appendix 3) (Mornon et al., 2008). The results in Figure 16A shows the CPM fluorescence increases by 1.85-fold during the thermal unfolding transition. Assuming all 15 Cys in CFTR become exposed after heating, this data would predict that 7 Cys are surface-exposed and that 8 Cys are buried in the native-folded structure (after background subtraction of the 2 Cys in ribosomal L3 shown in Appendix 4). Therefore, the CPM assay presented here predicts a different model for Cys accessibility in CFTR structure than that proposed by Mornon et al. (Mornon et al., 2008). There are a number of possible reasons why the purified CFTR has apparently different numbers of exposed cysteines in vitro compared to the model in silico, e.g. CFTR flexibility in solution, heterogeneity in oligomeric state, or even some cysteines may remain buried after heating. Further studies would be required to validate the CPM assay as an effective tool to predict the location of Cys in CFTR’s 3D structure.

This thesis presents a novel method to express and purify CFTR from a yeast expression system. To validate that CFTR expressed in our yeast system was correctly folded, the CPM thermal unfolding profile was compared to data collected in an equivalent experiment with human-CFTR expressed in HEK mammalian cells (protein donated from the Urbatsch lab), which has been shown to be fully functional as an ion channel using single channel conductance experiments (personal communication with Riordan lab). Figure 16B (left panel) shows that yeast-expressed CFTR has a very similar thermal unfolding profile to HEK-expressed CFTR, suggesting CFTR is recognised and folded by yeast cells in a similar manner to the more biologically relevant mammalian cells. The mid-point temperature of thermal unfolding (T_m) is slightly shifted to higher
temperatures in the yeast-expressed chicken CFTR ($T_m = 41.5 \, ^\circ C$) than in the HEK-expressed human CFTR ($T_m = 40.8 \, ^\circ C$), as shown in the right panel of Figure 16B. A higher $T_m$ would agree with published experiments that show the chicken-CFTR orthologue is more thermally stable than the human-orthologue in vivo (Aleksandrov et al., 2010). However, multiple repeats of this CPM experiment would be needed to confirm the small $T_m$ difference was statistically significant. These data highlights the use of a multiple orthologue strategy to screen for a CFTR protein that is more thermally stable and favourable for crystallisation. However, a direct in vitro comparison between the stability of human- and chicken-CFTR orthologues cannot be accurately made here as the two proteins have been produced using two different expression and purification protocols.

For both the human- and chicken-orthologues, CFTR begins to unfold above 25 °C (Figure 16B), which is much lower than physiological temperatures of both these organisms (37 and 39-43°C, respectively) and lower than temperatures known to support CFTR channel activity in mammalian cell lines (Sharma et al., 2001). This suggests that the purification of CFTR in a detergent environment has reduced the thermal stability of CFTR compared to its stability in a cellular lipid membrane, as is a common problem for other membrane proteins (Rosenbusch, 2001). The $T_m$ of CFTR reported here is close to the $T_m$ (~40 °C) of the detergent-purified P-glycoprotein that yielded crystals for high-resolution structure determination (Aller et al., 2009, Bai et al., 2011), implying CFTR stability is in a suitable range for further structural study. Although, the exact $T_m$ required for a good-diffraction quality crystal is largely protein specific (Sonoda et al., 2011).

Unlike the DDM-purified protein, CFTR in LPG has no observable cooperative unfolding transition. As shown in Figure 16B, the CPM fluorescence simply increases linearly with increasing temperature (left panel) and thus no $T_m$ can be calculated (right panel). The LPG detergent, itself, may interfere with the CPM dye and obscure any observable unfolding transition, however, this assay has already been robustly validated for a range of detergent types (Alexandrov et al., 2008, Lee et al., 2013, Sonoda et al., 2011). An alternative conclusion is that the tertiary structure of CFTR is destabilised in LPG, where Cys residues are already surface-exposed in a more random, dynamic manner at low temperatures. The latter theory agrees with evidence already reported earlier in this thesis that LPG is potentially destabilising for protein structure (section 3.9 and 3.10). Also, a lack of cooperative unfolding transition was similarly observed
Figure 16. Thermal stability of CFTR as monitored by the CPM dye-binding assay.

A. The CPM dye binding assay can be used to probe thermal unfolding of CFTR. DDM-purified chicken-CFTR was initially incubated with CPM dye for 30 min at 10 °C (left of dotted line) and the observed increase in CPM fluorescence represents labelling of surface-exposed Cys. The temperature was then increased to 80 °C at a heat rate of 2.5 °C/min (right of dotted line) and an increase in CPM fluorescence shows the labelling of natively buried Cys as they become surface-exposed during thermal unfolding. The heated protein was cooled back to 10 °C for 30 min and then a 2nd heat scan was recorded (red line), which was devoid of spectral characteristics observed with the 1st heat scan (blue line). B. Yeast-expressed chicken-CFTR (cCFTR) has a similar thermal unfolding profile to that of mammalian-expressed human-CFTR (hCFTR) when purified in DDM, but not LPG. The CPM dye-binding assay was performed for yeast expressed CFTR in DDM (blue line) and LPG (green line), and for mammalian cell expressed human-CFTR (purple line). The CPM fluorescence was normalised to show fraction of unfolding using the following equation: Fraction of unfolding = (Y - Y_D) / (Y_D - Y_N), where Y is fluorescence for a given temperature, Y_D is fluorescence with the denatured protein i.e. at 80 °C, and Y_N is fluorescence with the natively folded protein i.e. at 10 °C (left panel). A van’t Hoff line was fitted to the first derivative plots of these data using GraphPad software and the T_m determined from the peak as 40.8 °C for human-CFTR and 41.5 °C for the DDM-purified chicken-CFTR. In LPG, there was no apparent cooperative unfolding transition so no T_m could be calculated.
by this CPM assay under denaturing conditions for other membrane proteins (Alexandrov et al., 2008). If this is the case, LPG would only seem to have caused partial destabilisation of CFTR structure, because a fully denatured protein would give a CPM unfolding profile similar to the 2nd scan in Figure 16A. Further investigation is still required to fully understand the effect of LPG on CFTR structure.

Overall, the CPM-dye binding assay appears to be useful for probing the thermal stability of CFTR in DDM, with the main advantages of requiring little protein (minimum of 1 µg at 5 µg/ml). This data was collected in a single 150 µl quartz cuvette, but the assay could be adapted to a multi-well plate format for high-throughput screening of conditions that may thermally stabilise CFTR protein and favour crystallisation, e.g. different orthologues, mutations, constructs, buffer conditions and compounds. This assay could also be adapted for a CF drug screen to find compounds that could rescue the thermodynamic instability of ΔF508-CFTR. One caveat to this assay is that it requires an absence of reducing agents due to their reaction with CPM, but the removal of DTT from the purification buffer resulted in losses of CFTR probably due to a reduction in protein stability and a resulting precipitation (data not shown). Additionally, introducing an extrinsic probe such as the CPM dye always runs the risk of causing non-native perturbations of structure and function that could also interfere with potential CFTR-drug interactions.

4.2.2. Tryptophan fluorescence

Most proteins have intrinsic fluorescence due to the presence of aromatic residues tryptophan (Trp), tyrosine and phenylalanine (Lakowicz, 2006). Trp fluorescence has long been used to probe protein structure due to high signal-to-noise ratio, high sensitivity to changes in local environments such as polarity, and preservation of protein function during analysis (Eftink, 1998). In the hydrophobic core of a folded protein, Trp residues tend to have high quantum yield and high intensity fluorescence, but when the protein becomes unfolded, for example by heating, the residues are exposed to the surrounding solvent, which decreases the quantum yield and fluorescence intensity and often causes a red-shift in the emission spectra peak wavelength (Eftink, 1998). Chicken-CFTR has 19 Trp residues, including 8 in the NBD domains and 11 in the TM regions, as predicted by comparison with the human-CFTR structure model (Appendix 3) (Mornon et al., 2008). Therefore, Trp fluorescence measurements should be suitable to extract information regarding the integrity of CFTR’s tertiary structure.
Figure 17A shows the Trp fluorescence emission spectra for CFTR after heating to 10 or 80 °C. In DDM, heat treatment of CFTR gives a sizeable reduction in Trp fluorescence intensity from 170 to 42 AU, and also a red-shift in emission peak from 331 to 335 nm (left panel). This suggests a change in the local environment of Trp residues corresponding to thermal unfolding of CFTR structure. The magnitude of reduction in Trp fluorescence seems appropriate for a protein with as many Trp as CFTR, but the shift in emission peak is very modest when compared to the shift seen with soluble proteins containing even less Trp residues (Vivian and Callis, 2001). As with most trans-membrane proteins, a large number of CFTR’s Trp residues are in the TM region that, in vivo, provide the hydrophobic interface required for integration in a lipid membrane (Schiffer et al., 1992). During this experiment, these TM regions will be coated in detergent micelles, which could mask solvatochromism as detected by Trp fluorescence. Therefore, it could be hypothesised that the detected shift in Trp fluorescence is primarily due to structural changes within the soluble NBD regions, where the unfolding would cause more significant changes in local Trp environment than unfolding of the TM regions.

For the DDM-purified CFTR, plotting the ratio of the shift in Trp emission peak gave a broad cooperative unfolding transition (Figure 17B) that allowed the calculation of a $T_m = 41.8$ °C. The similarity of this $T_m$ to that determined from the CPM dye-binding assay ($T_m = 41.5$ °C) (section 4.2.1.) implies both assays measure the same structural changes in CFTR during thermal destabilisation. The $T_m$ recorded here is lower than that previously reported for the thermal unfolding of the isolated NBD1 protein ($T_m = 57.8$ °C) (Protasevich et al., 2010) and higher than that of the isolated R region ($T_m = 37$ °C) (Marasini et al., 2012), suggesting the $T_m$ of full-length CFTR is a an average of the unfolding of its multi-domains.

In LPG, there is a much smaller decrease in CFTR Trp fluorescence after thermal denaturation (37 to 23 AU) than for CFTR purified in DDM (170 to 42 AU), and the fluorescence intensity of the native protein in LPG (37 AU) is a similar value to the fluorescence of the denatured protein in DDM (42 AU), considering the same protein concentration was used in each case (Figure 17A). Additionally, there was no shift in the fluorescence emission peak, which remained at 331 nm throughout the heat scan, and no $T_m$ could be calculated (Figure 17B). This indicates little change in the polarity of the environment of the Trp residues in CFTR during heating in LPG. LPG, itself, may mask any observable changes in Trp fluorescence environment. Although, in
Figure 17. Thermal stability of CFTR, monitored by intrinsic Trp fluorescence. 
A. Heat denaturation of CFTR in DDM caused a greater change in Trp fluorescence intensity and a greater Stokes shift than denaturation in LPG. Trp fluorescence emission spectra (excitation at 290 nm) were collected for CFTR (50 µg/ml) purified in either DDM or LPG after incubation at 10 °C and 80 °C, shown as native (solid line) and denatured (dotted line) respectively. In DDM, the emission peak shifted from 331 to 335 nm. There was no shift in emission peak wavelength for CFTR in LPG. B. The shift in Trp fluorescence emission peak revealed a cooperative unfolding transition for CFTR in DDM, but not LPG. The shift in Trp fluorescence was normalised using the formula $y = F_D/F_N$, where $F_D$ is the fluorescence intensity at the peak emission wavelength (335 nm) of the denatured protein, and $F_N$ is the fluorescence intensity at the emission wavelength (331 nm) of the native protein, for each temperature point. For CFTR-DDM, a sigmoidal curve was fitted to the normalised data in GraphPad Prism software showing a two-state unfolding model, where the native and denatured protein are represented by the bottom and top plateau respectively, and a $T_m$ was calculated from a van’t Hoff plot as 41.8 °C. No $T_m$ could be calculated for the CFTR-LPG data due to a lack of cooperative unfolding transition.
combination with the CPM thermal stability assay data, it does seem that CFTR is already structurally destabilised in LPG at low temperatures, hence a minimum change in tertiary structure observed with heating. The extent of CFTR structural destabilisation in LPG is still not clear, and may just be limited to structural folds specifically detected in by these assays.

4.2.3. Static light scattering

Thermal denaturation of proteins can also sometimes induce aggregation, usually by unfolding hydrophobic regions that prefer to self-associate than be exposed to the solvent environment. Such protein aggregation can be described by the well-known Lumry-Eyring model: \( N \rightleftharpoons TS \rightarrow A \), where \( N \) is the native protein and TS is an unfolded transition state (sometimes referred to as molten-globule state) preceeding formation of non-native aggregates, A (Lumry, 1954). Protein aggregation can be monitored by changes in static light scattering measurement (Hayashi et al., 1989, Kopec and Schneider, 2011). Previous works have shown that unfolding of the isolated NBD1 region of CFTR induces significant aggregation that is detectable by an increase in static light scattering (Wang et al., 2010). To see if the same aggregation occurred with the thermal unfolding of full-length CFTR, 90° light scattering measurements were collected simultaneously to the Trp fluorescence measurements shown in the previous results section. Surprisingly, the light scattering of purified CFTR decreased with increasing temperature (Figure 18A). This suggests that the thermal unfolding of CFTR reported here is not coupled with an increase in protein aggregation, but rather a decrease in protein hierarchical structure. The decrease in light scattering is also evident from comparison of CFTR protein particles incubated at 10 versus 80 °C by negative-stain EM, where at low temperatures, CFTR is a tight compact globular protein about 150 Å in length, and at higher temperatures, CFTR is a more random, linearised structure about 300 Å in length (Figure 18B). For DDM-CFTR, the decrease in light scattering could be fitted with a two-state unfolding curve and \( T_m \) calculated as 32.0 °C, which is lower than the \( T_m \) calculated from the Trp fluorescence (41.8 °C) and CPM dye (41.5 °C) assays, suggesting that destabilisation of CFTR starts to occur before the breakage of its tertiary structure bonds. In LPG, there is a decrease in light scattering with increasing temperature, but with no clear cooperative transition that could be used to calculate a mid-point melting temperature, as was the case for the CPM and Trp fluorescence data. Again, suggesting the CFTR in LPG is lacking some levels of cooperative tertiary structure.
Figure 18. Thermal denaturation of CFTR does not result in protein aggregation.
A. **Heat treatment of CFTR caused a decrease in light scattering.** DDM- and LPG-purified CFTR at 50 μg/ml were heated from 10 to 80 °C at scan rate 2.5 °C/min, and 90° light scattering fluorescence data collected (excitation = 290 nm, emission = 565 nm) and background-subtracted for protein-buffer at each temperature point. The fluorescence was normalised as a fraction of change in light scattering for each sample. A two-state sigmodial curve was fitted to CFTR-DDM and the $T_m = 32.0$ °C was calculated from a van’t Hoff plot of this data. No $T_m$ could be calculated for CFTR-LPG due to a lack of a cooperative unfolding transition.

B. **Electron microscopy reveals CFTR unfolds from a compact to a linearised structure after heating.** DDM-purified CFTR was analysed by negative-stain EM after incubation for 10 min at either 10 or 80 °C. Protein particles were selected from the electron micrographs (Appendix 5) using EMAN2 software, which are viewed as white particles against a grey background. The box sizes for CFTR heated to 10 °C (left panel) and 80 °C (right panel) are 249.6 nm and 665.6 nm, respectively.
To date, there has been no reported data for the unfolding of isolated full-length CFTR. Here, it appears that the unfolding of CFTR does not induce the same protein aggregation event that is observed with isolated NBD1 (Wang et al., 2010, Protasevich et al., 2010). The lack of observed aggregation could be due to the presence of detergent that may protect unfolded hydrophobic regions from self-association. Also, the concentration of CFTR in this experiment remained relatively low (50 µg/ml) compared to previous work on NBD1 (~1 mg/ml), due to the restrictions on protein concentration discussed in results section 2.1. Therefore, it is difficult to extrapolate the data from isolated NBDs to the behaviour of the full-length protein, leaving the exact thermodynamic mechanism of CFTR aggregation unclear. Regardless, the decrease in light scattering shown in Figure 18 could prove a useful tool to probe the thermal stability of CFTR for later crystallisation in a detergent-solubilised state, i.e. by screening compounds that can shift the T_m to higher temperatures.

4.2.4. Thermal SDS-PAGE analysis (TGA)

SDS-PAGE is a well-known technique to study the thermal stability of membrane proteins (Lysko et al., 1981) and has been successfully used to screen for different conditions that promote protein crystallisation (Engel et al., 2002). The technique takes advantage of the fact that membrane proteins readily form irreversible, SDS-resistant aggregates after heat denaturation. In summary, this process is visualised by the disappearance of specific protein bands and the appearance of HMW aggregates on the gel. The melting temperatures calculated from this TGA assay correlate well to the transitions detected by other biophysical methods, e.g. DSC (Brouillette et al., 1982, Soler et al., 1984). In combination with the possible detection of CFTR’s GFP tag by in-gel fluorescence, this TGA method appeared to be a quick and easy method to study CFTR thermal stability that only used nanogram quantities of protein.

Figure 19A shows that heating CFTR does result in SDS-PAGE resistant aggregates, as seen by the appearance of a GFP-fluorescent band near the top of the gel at higher temperatures that coincides with a loss in a fluorescent band where CFTR monomers migrate. This change in protein band migration is specific to CFTR and not simply an artefact of temperature effects on in-gel fluorescence, as is shown by the TGA assay of purified GFP in Appendix 6. Quantification of the aggregate protein band shows that conversion of monomers to aggregates is a two-state process that occurs over a narrow
temperature-range for the DDM-purified CFTR and a more broad temperature transition for the LPG-purified protein, with unfolding transitions calculated at 55 and 46 °C, respectively (Figure 19A, right). Again, implying that CFTR structure is more stable in DDM compared to LPG. For DDM, the $T_m$ is the same as that reported previously (55 °C) for the formation of similar SDS-insoluble aggregates of CFTR in membranes from mammalian cells (Sharma et al., 2001). Interestingly, when the TGA method was used to monitor CFTR stability at low temperatures, there was little difference in the time taken (5-8 days) for formation of SDS-insoluble aggregates in either detergent (Figure 19B, left), suggesting the solubility of CFTR during long periods of storage, e.g. in a crystallisation trial, would not be very affected by detergent type.

At first glance, this TGA assay appears to be a very simple method to probe CFTR thermal stability in different buffer conditions (e.g. detergents) for crystallisation, which also reflects the thermal stability of CFTR in a more native-like environment of a cell membrane (Sharma et al., 2001). However, this thesis raises a few points of concern regarding the interpretation of the TGA assay. Firstly, the appearance of HMW gel bands has been previously explained by severe, irreversible protein aggregation (Lysko et al., 1981). However, no such thermally-induced aggregation was detected for CFTR by light scattering or by EM (section 4.2.3) and the TGA aggregation detected after 5-8 days incubation at 4 °C was in fact coupled with a decrease in light scattering that contradicts the formation of any aggregates (Figure 19B). Secondly, whereas published results have shown good correlation between thermal unfolding transitions detected by TGA and other biophysical methods e.g. DSC (Brouillette et al., 1982, Soler et al., 1984), the $T_m$ calculated from the TGA of CFTR here is shifted to higher temperatures (55 °C) than melting temperatures (~41 °C) obtained from the CPM dye and Trp fluorescence assays, even when employing the same protein concentration (50 µg/ml) and heating rate (2.5 °C/min). Finally, the similarity in the time taken for CFTR aggregation in both LPG and DDM at low temperatures (Figure 19B) contradicts the large difference in CFTR solubility observed in Figure 15 between the two detergents. Therefore, it appears that the change in CFTR band migration with increasing temperature is a phenomenon that occurs specifically at the stage of SDS-PAGE analysis and not before, when the protein is in solution.
Figure 19. Thermal stability of CFTR assessed by SDS-PAGE.

A. CFTR forms SDS-PAGE resistant “aggregates” after heat treatment. DDM- and LPG-purified CFTR at 50 µg/ml were heated in a water bath from 15 to 80 °C at a rate of 2.5 °C/min. Aliquots were taken at 5 °C increments, quenched on ice for 10 mins, and then incubated with SDS load-dye for 15 min at RTM before loading approximately 10 ng of CFTR onto a pre-cast 10% SDS-PAGE gel with no stacker. An example gel is shown on the left showing GFP fluorescence where CFTR either migrates as a monomer band at approximately 220 kDa (M), or as a HMW aggregate at the top of the gel (A). Plotting the intensity of the A band (right) at each temperature relative to the intensity at 15 °C, shows that the formation of these SDS-PAGE resistant aggregates occurs with a T_m of 46 and 55 °C in LPG and DDM, respectively.

B. After 5-8 days storage at 4 °C, CFTR formed aggregates that were resistant to SDS-PAGE but not detected by an increase in light scattering. DDM- and LPG-purified CFTR at 50 µg/ml were incubated at 4 °C for 0-10 days. Aliquots were taken at different time intervals and 90° light scattering measurements collected (excitation = 290 nm, emission = 565 nm), which are shown on the right-hand side graph after background-subtraction of protein-free buffer and normalisation to the scattering at 0 days. These aliquots were also stored at -20 °C in SDS-load dye before analysis on a 10% SDS-PAGE gel that was scanned for GFP fluorescence, as shown in Figure 15A. The relative intensity for protein band A is shown on the left-hand fluorescence and the times for the cooperative formation of these aggregates was calculated as 6.1 and 6.5 days for LPG and DDM, respectively.
Possible mechanisms for the appearance of these SDS-PAGE resistant aggregates can be discussed: Because these TGA aggregates appear at higher temperatures compared to the $T_m$ calculated from the CPM dye binding and Trp fluorescence assays, it can be assumed that this phenomenon first requires the unfolding of CFTR tertiary structure. Also, from the early results section (Figure 15), it seems that CFTR aggregation is dependent on protein concentration. Hence, CFTR aggregation could be specifically induced by the locally high concentration of protein as it enters the gel, only if thermal unfolding has already occurred. Previous works have assigned TGA aggregation to irreversible covalent cross-links of disulphide bonds formed by Cys residues as they become exposed during thermal unfolding (Engel et al., 2002, Brouillette et al., 1982, Lysko et al., 1981). However, the high concentration of DTT (50 mM) in the SDS-load dye used here, suggests that disulphide bonds do not play a role in the result of this assay, which leaves the possibility of non-native, non-covalent bonds as the cause of TGA aggregation, as proposed by Galani et al. (Galani and Apenten, 1999). Whatever the reason for the formation of CFTR TGA aggregates, it seems care must be taken when interpreting this TGA data, as it is not clear if this assay is specifically probing the thermal stability of CFTR in solution.
**Chapter 5 – Results and Discussion: Homogeneity**

Protein homogeneity is an important biophysical characteristic for the production of well-diffracting 3D crystals (McPherson, 1999, Bergfors, 1999). The presence of heterogeneous protein populations in the solution will inhibit nucleation and crystal growth (Borgstahl, 2007). Using our novel yeast expression system, it is now possible to produce CFTR protein in sufficient quantities for 3D crystallisation trials (Chapter 3), but producing a homogenous population of purified protein now poses yet another great challenge in the effort towards a high-resolution 3D structure of CFTR. SDS-PAGE analysis has shown that our yeast-expressed CFTR can be purified to 90 and 60 % homogeneity in detergents LPG and DDM respectively (section 3.7 & 3.8). However, due to the denaturing conditions of SDS-PAGE, this analysis is limited to only providing information about the denatured polypeptide chain. In native conditions, protein homogeneity can also be affected by a number of factors that can later inhibit crystal growth including, protein size (Borgstahl, 2007), conformational flexibility (DePristo et al., 2004), PTMs (Reeves et al., 2002), and for membrane proteins, variations in the protein-detergent micelle size (Hitscherich et al., 2001). A number of biophysical techniques can be used to pre-screen the quality of protein homogeneity before crystallisation trials, for example, sedimentation equilibrium (Yoshizaki et al., 2005), light scattering (Borgstahl, 2007), SEC (Newby et al., 2009), gel electrophoresis (Cleverley et al., 2008), and EM (Frank, 2006). In the previous results chapters, 90° static light scattering was used to measure CFTR aggregation state under different conditions e.g. at different NaCl concentration (section 3.8), protein concentration (section 4.1) and temperature (section 4.2.3), but this technique only provided a crude, bulk measure of CFTR, which would disguise details about size heterogeneity. SEC was also used to purify CFTR (sections 3.7 & 3.8), and this data could be used to estimate protein size (Barth et al., 1994, Potschka, 1987). However, estimations by SEC relies on protein calibration standards that may not reflect the same geometric compactness as CFTR, or take into account any unexplained interactions that have been reported to occur between CFTR and the SEC matrix (Zhang, 2008, Zhang et al., 2010), or association with detergent micelles (Hitscherich et al., 2001). Therefore, a technique was required to provide further detail about CFTR homogeneity. This chapter describes two techniques, SEC-MALLS and EM, the focus of which was characterisation of CFTR homogeneity by protein size, also known as protein monodispersity.
5.1. SEC-MALLS analysis of CFTR protein homogeneity

SEC-MALLS is a non-invasive technique that measures dynamic light scattering of particles as they elute from a SEC column (Sahin and Roberts, 2012). Dynamic light scattering is often used to assess protein monodispersity in pre-crystallisation screens (Ferré-D’Amaré, 1997). MALLS allows for the direct calculation of protein size without the need for calibration standards, and SEC can fractionate different sized sub-populations of proteins that would otherwise be hidden in a bulk measurement of the entire solution (Sahin and Roberts, 2012, Li et al., 2010). Full-length CFTR has not been previously analysed by SEC-MALLS, largely due to lack of sufficient quantities of purified protein. Here, using our yeast expression system, we were able to produce enough CFTR for such novel biophysical characterisation.

Figure 20 shows SEC-MALLS data of CFTR purified in LPG (A) and DDM (B). In both detergents, CFTR was first filtered through a 0.2 µm membrane to remove large insoluble aggregates and then injected onto an analytical KW-804 SEC column. As the protein eluted, light scattering and RI measurements were collected by a Wyatt technology detector that were then calculated into molar mass by the ASTRA programme using the Zimm expression equation, where the RI was used to estimate protein concentration (Li et al., 2009). In LPG, CFTR eluted from the SEC column between 7 and 10 ml, as judged by the fluorescence of its GFP-fusion tag (green line). This elution is shortly after the 1 MDa (or 6ml elution volume) size exclusion limit of the KW-804 column, according to the manufacturers guidelines (Shodex). At the peak of CFTR elution (8ml), the RI and light scattering trace overlay well with a ratio of 1:1 relative arbitrary units, which indicates CFTR is mostly monomeric (Li et al., 2009). Ideally, CFTR would elute in a sharp Gaussian peak because this would indicate a good level of homogeneity (Barth et al., 1994, Potschka, 1987). The peak between 7.5 and 8.5ml appears quite symmetrical, suggesting that the CFTR found in these fractions was suitably homogenous for 3D crystallisation trials. The rest of the elution peak is not symmetrical, with a shoulder to the left seen by light scattering and a shoulder to the right seen with RI, suggesting there are also some HMW aggregates and smaller contaminants or CFTR fragments present in the sample. The size of the CFTR monomeric particles can be calculated from the plateau in molar mass (dotted line) at the 8 ml elution volume as approximately 500 kDa. This is higher than the predicted 220 kDa molecular weight for a CFTR monomer, probably due to the added mass of an LPG detergent micelle belt around the protein; an average LPG micelle is predicted to
be around 90-100 kDa (Oliver et al., 2013). Additional modifications to protein amino acids may also add to apparent molecular weight, such as the core glycosylation observed in results section 3.5 (Larkin and Imperiali, 2011).

Whilst the homogeneity of CFTR in LPG appears quite promising for 3D crystallisation trials, in DDM it appears quite the opposite (Figure 20B). CFTR purified in DDM eluted from the SEC column in a very broad, asymmetrical peak between 6.5 and 11.5ml, with a peak elution at 8ml, according to the GFP fluorescence trace. Such an elution profile is consistent with a polydisperse population of protein particles. Due to the limitations in protein concentration of CFTR in DDM (section 4.1), this SEC-MALLS experiment was conducted at the very lower limit of RI detection as seen by the relatively small RI peak at 8 ml. Despite this low concentration of protein, the light scattering at 8 ml was above the scale detectable by the MALLS Wyatt instrument. The very high ratio of light scattering to RI is indicative of non-specific, HMW aggregates (Li et al., 2009). The predicted size of CFTR particles is taken from the plateau of the molar mass line at 8ml elution volume as 10 MDa. A proportion of this mass is from DDM detergent micelles, which potentially could be very large in the high salt buffer used here (Jumpertz et al., 2011). Even taking this into account, it appears that CFTR particles are very large, around 100 monomers of CFTR in size. This result is very interesting because the size exclusion limit of this SEC column is only 1MDa (or 6 ml elution volume) and any insoluble aggregates should have been removed during the initial filtration through a 0.2 µm membrane. It therefore seems that these large aggregates are forming during or shortly after the SEC fractionation. After CFTR peak elution, the calculated molar mass declines steadily between 9 and 11.5 ml elution from around 10 MDa to around 200 kDa with a straight line devoid of individual peaks. This profile points to CFTR existing as a mixed population of particles anywhere between 1 and 100 monomers in size, rather than a few populations of discrete oligomeric states. Also, that aggregate formation is highly dynamic, even at 4 °C. One could hypothesise that, under these conditions, CFTR is continuously aggregating from monomers to non-specific HMW aggregates, and that this SEC-MALLS experiment is simply capturing a snapshot of the different stages in nucleation of this aggregation pathway. Such heterogeneity in particle size would probably make this sample of DDM-purified CFTR unsuitable for 3D crystallisation trials. However, if the hypothesised dynamic equilibrium (between monomers and HMW aggregates) is much faster than the rate of crystal growth, then in principle crystal formation from monomers should be feasible.
Figure 20. SEC-MALLS reveals CFTR is more homogenous in LPG (A) than DDM (B). CFTR purified in LPG (A) and DDM (B) were injected onto an analytical KW-804 SEC column held at 4 °C, at protein concentrations of 0.4 and 0.2 mg/ml respectively. The column was connected to a HPLC with Wyatt technology detectors that allowed simultaneous detection of MALLS (red line), RI (blue line) during protein elution from the SEC column. GFP fluorescence data was also collected (green line) to specifically track CFTR elution via its GFP-fusion tag. The MALLS and RI data was imported into the ASTRA software for calculation of protein molar mass (dotted red line). The MALLS, RI and GFP values are shown as relative arbitrary units. The arrow marked $V_o$ at 6 ml indicates the size exclusion limit of the column (approx. 1MDa). In LPG (A), CFTR mostly eluted from the SEC column in a single, monomeric protein peak with molar mass of approximately $5 \times 10^5$ g/mol (500 kDa). In DDM (B), CFTR eluted as a heterogeneous populations of aggregates. Due to a technical fault, no data was collected for the DDM sample before 5.5 ml elution volume. Data collected in collaboration with Dr Chi Wang (Hunt Lab, Colombia University).
5.2. EM analysis of CFTR homogeneity

EM is a technique used to study the structure of biological macromolecules, including membrane proteins (Frank, 2006, Rubinstein, 2007). Collection of electron micrographs can quickly provide qualitative information regarding the homogeneity and dispersity of a protein (Mio et al., 2009). EM has been used to assess the aggregation state of CFTR after purification from mammalian cells prior to further structural characterisation (Zhang, 2008, Zhang et al., 2009, Zhang et al., 2010). There are two main types of transmission EM techniques that have been applied to the study of protein structures: cryo-EM (Adrian et al., 1984) and NS-EM (Brenner and Horne, 1959). Cryo-EM would be the preferred method to study CFTR because the protein remains in a native-like hydrated state whilst trapped in a thin layer of vitreous ice; NS-EM traps the protein in a layer of heavy metal stain that can limit structural resolution and sometimes induce structural artifacts (Rubinstein, 2007). Also, in certain detergents, including LPG, the protein will not adsorb to the surface of the carbon-coated grids used in NS-EM and can hence only be studied by cryo-EM where the protein is frozen in holes within a perforated grid. The recommended protein concentration for cryo-EM is around 1 mg/ml, which should provide sufficient numbers of non-overlapping particles in a field-of-view for further structural processing. A minimal buffer is also recommended to reduce the density of the buffer relative to the protein molecules and to provide sufficient signal-to-noise to view particles. These sample requirements were easily achieved for LPG-purified CFTR, but not feasible for the DDM-purified protein due to the limitation on protein concentration (section 4.1.) and requirement for high salt in the purification buffer (section 3.8). NS-EM was therefore used instead for DDM-CFTR, which only needs concentrations of protein of about 10 µg/ml because the charge applied to the grid surface in this method will induce a physisorption of protein on the surface of the carbon-coated grid (Rubinstein, 2007).

5.2.1. Cryo-EM analysis of LPG-purified CFTR

CFTR in LPG was collected after the SEC purification step and concentrated to 8 mg/ml, as described earlier (sections 3.7 & 4.1). The protein was then diluted to 1 mg/ml in a glycerol-free buffer, to a final glycerol concentration of 1.25 % (w/w), in order to minimise the scattering of electrons by the buffer alone, before applying to a holey carbon-coated EM grid (Quantifoil – 300mesh/in, 2 micron diameter holes with 2 micron separation). The grid was flash frozen in liquid ethane to trap the protein in thin
layers of vitreous ice within the grid holes. The grid was glow discharged as this is reported to induce a good, even distribution of protein particles within the ice (Rubinstein, 2007). Despite the GD, CFTR particles were unevenly distributed across the ice, with high density of proteins at the edge of the hole and none in the centre (Appendix 7). Increased concentration of protein particles at the edge of the hole can often occur with cryo-EM, particularly with hydrophobic membrane proteins (Rubinstein, 2007, Quispe et al., 2007). For further analysis, areas within the ice were selected where CFTR particles were well separated (Appendix 7). A section of this area is shown in the top panel of Figure 21. From this view, CFTR appears to be a very homogenous population of small particles with no sign of aggregation. As CFTR is analysed here under cryo-EM in a fully hydrated state, the homogeneity reported is likely to reflect the behaviour of CFTR in solution. Such monodispersity is consistent with the SEC-MALLS data (section 5.1), and indicates that this protein sample would be suitable for 3D crystallisation trials.

SPA is a computational technique that uses raw EM data to create a 3D structural model of macromolecules (Henderson and Unwin, 1975, Frank, 2006). The major advantage with SPA is that a structure can be produced of a protein in solution, without the need for crystallisation. In summary, many thousands of randomly orientated particles are selected from EM micrographs, averaged and iteratively aligned both translationally and rotationally to reconstruct a 3D model. Here, around 1000 CFTR particles were selected from across 5 cryo-EM micrographs using the EMAN software. The middle panel in Figure 21 shows a few examples of selected particles. Due to the uneven distribution of particles in the ice (Appendix 7), well-separated particles were selected on the criteria that there was no overlap within a 48-pixel (= 183.84-Å) box. The particles were CTF-corrected and then averaged into 2D classifications, which are shown in the bottom panel of Figure 21. CFTR appears as small, spherical particles that appear similar to the top view of CFTR monomers found in previously reported 2D crystals (Rosenberg et al., 2011). In fact, all of the 2D class views are around 60-Å in length, suggesting CFTR is even smaller than a monomer when viewed from its side. SDS-PAGE (section 3.7) and mass spectrometry (section 3.9) analysis have confirmed that CFTR is purified in its full–length. Therefore, it would seem that during the 2D class averaging there is a loss in density for some part of the CFTR structure, which is indicative of disordered regions in the protein structure.
Figure 21. Cryo-EM reveals LPG-purified CFTR is homogenous and monomeric.
The top panel is a portion of an EM micrograph showing CFTR protein as black particles against a white background; a wider field-of-view can be seen in Appendix 7. Black arrows label a few example CFTR particles and the scale bar shown is 50 nm. The EMAN software was used to select 1056 CFTR particles and perform 2D class averaging. The middle panel shows a few selected particles that are in boxed in areas of 183.84 Å width. The bottom panel shows the 16 x class averages representing different 2D side views of the CFTR protein, which are inverted as white particles on a black background, with scale bar = 100 Å. CFTR particles are around 60-Å in length.
Again, this provides evidence that suggests CFTR structure is partly destabilised in LPG. Interestingly, previous cryo-EM analysis of CFTR purified from mammalian cells revealed CFTR to be folded as a dimer when in LPG (Zhang et al., 2009). In these experiments, CFTR was first solubilised and purified in DDM and only later exchanged into LPG during the final chromatography step. Therefore, it could be proposed that LPG is not denaturing to CFTR structure if only used for short periods of time at relatively low concentrations, and that the long incubation times and high concentrations of LPG used here to purify CFTR from yeast was too harsh to maintain protein structural integrity. The oligomeric state of CFTR is again addressed in the next results section (5.2.2), but in combination with the SEC-MALLS data (5.1), it is likely that CFTR is monomeric after purification in LPG using the methods described in this thesis.

Unfortunately, the CFTR particles in Figure 21 appear featureless and simply spherical because their small size is at the resolution limit for cryo-EM analysis (Frank, 2006). Therefore, no further SPA analysis was performed for this CFTR sample. Overall, it seems that CFTR is a very homogenous populations of monomers and such monodispersity would qualify this CFTR preparation for 3D crystallisation trials. However, it is proposed that the LPG-purification method used here causes partial destabilisation of CFTR structure and any unfolded regions in the protein would prevent the formation of the rigid contacts needed for crystallisation.

5.2.2. NS-EM of DDM-purified CFTR

CFTR in DDM eluted from the SEC column over a broad volume (section 3.8). Protein was therefore taken from different fractions across the SEC elution profile to gain a full understanding of homogeneity by NS-EM analysis. CFTR at around 20 µg/ml protein concentration was adsorbed to a GD continuous carbon-coated grid and washed with water to remove excess detergent molecules, before incubation with UA stain. The grids were viewed by transmission EM and entire micrographs are shown in Appendix 8 and 9 to show a wide field-of-view. Protein taken from elution fractions close to the void volume (7 ml) of the SEC column comprised mostly of large amorphous aggregates (Appendix 8). Individual CFTR particles were better viewed in fractions taken further from the void volume (9.5 ml) (Appendix 9). The SEC purification step thus appeared to be a useful tool for removing aggregates and improving the homogeneity of purified CFTR protein, as has been shown previously (Zhang et al., 2010). Figure 22 (top panel)
shows a section of the NS-EM micrograph of CFTR protein taken from the later elution fractions (9.5 ml). Even after fractionation on the SEC column, CFTR still appears to be heterogeneous in size, consisting of small particles about 10-20 nm and larger, amorphous aggregates ranging from 20-40 nm in diameter. These larger aggregates were still observed by NS-EM, even after filtration through a 1 MDa pore-sized membrane and after ultracentrifugation at 100 Kg (data not shown). Therefore, it seems that these CFTR aggregates are soluble and that self-association is a highly dynamic process. This data is consistent with the SEC-MALLS analysis (section 5.1), indicating the EM results likely represent the actual behaviour of CFTR in solution, and not just some artifact of NS on the grid. The number of aggregates appear greater than that seen in NS-EM micrographs of DDM-purified CFTR from mammalian cells (Zhang et al., 2010), suggesting a lower quality of protein homogeneity for the yeast-expressed CFTR. Such dynamic polydispersity means that the CFTR purified from our yeast expression system in DDM is not suitable for 3D crystallisation trials.

Unlike 3D crystallisation, proteins with any degree of heterogeneity can be applied to an EM grid. If the polydisperse protein can be separated into discrete monodisperse sub-populations, then the particles can be later refined into a number of different 3D structure models using a multivariate statistical analysis methods (Heymann et al., 2003, White et al., 2004). Unfortunately, the larger aggregates shown in Figure 22 (top panel) appear amorphous and not likely to fall within discrete conformations required for such multivariate structural refinement (White et al., 2004). Still, it was possible to use part of this approach and specifically separate the more homogenous, smaller CFTR particles from the larger aggregates using the EMAN software. CFTR particles were selected based on the criteria that they fit in a 64-pixel (332.8 Å) box, without overlapping other particles (Figure 22, middle panel). Around 4000 particles were selected, CTF corrected and processed for 2D class averaging. The different 2D side views of the CFTR particles are shown in Figure 22 (bottom panel). CFTR particles are around 180-Å in length and 100-Å in width; such particle dimensions are comparable to the dimensions previously reported for CFTR dimers viewed by EM and SPA (Zhang et al., 2009, Zhang et al., 2010, Awayn et al., 2005). Unlike the LPG-purified protein (Figure 21), DDM-purified CFTR is large enough for some structural features to be resolved by EM. In some of the 2D views, CFTR appears to have a pseudo 2-fold symmetry that is consistent with a dimeric organisation (Zhang et al., 2009, Zhang et al., 2010, Awayn et al., 2005). It therefore seems that the yeast-expressed CFTR is
Figure 22. NS-EM reveals DDM-purified CFTR is a heterogeneous population of different sized aggregates.

The top panel is a portion of an EM micrograph showing CFTR protein as white particles against a grey background, with scale bar equal to 50 nm; a wider field-of-view can be seen in Appendix 9. CFTR exists as small particles about 15 nm in length (black arrows) as well as larger aggregates ranging 20-40 nm (red arrows). The EMAN2 software was used to select 3962 of the smaller CFTR particles and perform 2D class averaging. The middle panel shows a few selected particles that are boxed in areas of 332.8 Å width. The bottom panel shows 16 x class averages representing different 2D side views of the CFTR protein, with scale bar = 200 Å. CFTR particles are around 180 Å in length and 100 Å in width, which may represent a dimeric quaternary structure.
similarly folded as the CFTR produced in mammalian cells, but this level of quaternary structure is only preserved during purification in DDM, and not in LPG. The exact oligomeric structure of CFTR in a native cell membrane still remains uncertain, with reports in the literature of both monomers and dimers acting as a functional unit (Zhang et al., 2009, Rosenberg et al., 2004, Ramjeesingh et al., 2001, Haggie and Verkman, 2008, Mio et al., 2008, Chen et al., 2002, Eskandari et al., 1998, Schillers et al., 2004). The EM data reported here would suggest a dimeric conformation for CFTR, but it may be that solubilisation in detergent induces a non-native conformation, as is believed to sometimes occur with purification of membrane proteins (Larue et al., 2009, Fujiyoshi and Unwin, 2008, Tate, 2006).

5.3. 2D crystallisation

Electron crystallography is a technique often used to solve the structures of membrane proteins, at medium-to-high resolution (Lau and Rubinstein, 2010, Korkhov and Tate, 2008, Tsai and Ziegler, 2005, Mitsuoka et al., 1999, Gonen et al., 2005). Electron crystallography has been used to investigate the structures of several ABC transporters (Rosenberg et al., 2001, Rosenberg et al., 2005, Ward et al., 2009, Chami et al., 2002, Ferreira-Pereira et al., 2003), and has, to date, provided the highest resolution (9-Å) structural data of CFTR (Rosenberg et al., 2011). Previous 2D crystallisation of CFTR assumes that the purified protein was sufficiently homogenous (at least in the x and y plane) to form ordered crystal-crystal contacts (Rosenberg et al., 2011). Using this protein as the current “gold standard” in CFTR homogeneity, the same 2D crystallisation technique was hence chosen to screen the quality of our yeast-expressed CFTR protein; the rationale being that crystals would only grow if the protein is homogenous.

Currently, 2D crystals of CFTR have been grown using the epitaxial method (Rosenberg et al., 2004, Rosenberg et al., 2011). This method involves applying a droplet of purified protein directly to a carbon-coated EM grid in the presence of a precipitating agent, then allowing vapour diffusion at the air/water interface to promote the increase in protein concentration required for crystal growth (Auer et al., 1999). Epitaxial 2D crystallisation of CFTR is routinely done in the presence of DDM. This method is not however compatible with LPG because this detergent prevents the adsorption of protein to the surface of the grid (personal communication with Dr.
Rosenberg). Therefore, only the DDM-purified CFTR was used for this 2D crystallisation screening experiment.

Epitaxial 2D crystals of our yeast-expressed and DDM-purified CFTR were successfully grown using the same methods as described previously (Rosenberg et al., 2011). Figure 23A shows a 2D crystal of CFTR as viewed by NS-EM, where CFTR particles appear packed together in repeating units across the surface of the grid. The crystal is approximately 500 nm in length, which is smaller than the possible 1 µm size previously achieved using this crystallisation method (Rosenberg et al., 2011), but still should contain a sufficiently large number of unit cells for further processing (Raunser and Walz, 2009). At first inspection, this crystal does not seem very ordered and is relatively mosaic. The computed FFT further agrees with this assessment, showing strong peaks only extending to the second order (Figure 23B); previous NS-EM 2D crystals of CFTR were better ordered, showing sharp FFT spots extending to third order (Rosenberg et al., 2004).

Further assessment of this crystal image was carried out using the 2dx software package (Gipson et al., 2007). In summary, the crystal lattice parameters were refined and corrected for the CTF, and lattice unbending was performed to improve the resolution of the crystal (Appendix 10). Figure 23B (right) shows the resulting resolution plot of the digitised FFT. There are FFT spots of IQ values 1 and 2 extending in all directions in the x and y plane, to approximately 30-Å resolution (Figure 23B). This resolution is lower than the supposed 15-Å limit of negative-stain crystallography (Schmidt-Krey et al., 2007) and lower than the 20-Å resolution obtained from previously published negatively-stained crystals of CFTR (Rosenberg et al., 2004), suggesting this data is itself at low resolution and not simply a limitation of NS-EM.

A first generation structure map was next generated in 2dx (Figure 23C). The map was truncated at 20-Å resolution to accurately represent the low resolution of this data. At this resolution, little detail can be seen about CFTR structure from the map. Still, general parameters of the crystal can be calculated to provide insight into CFTR crystallisation and structure. Firstly, CFTR appears to crystallise with an overall three-
Figure 23. Epitaxial 2D crystals of CFTR are weakly ordered and yield a low-resolution structure map.
CFTR protein appears as white particles repeatedly packed together against a grey background. A particle around 1 nm in length is indicated with a black arrow. The crystal appears weakly ordered and mosaic. **B. Computed FFT of the crystal image (left) and the digitised FFT resolution plot (right).** The crystal image was analysed and FFT calculated using the 2dx software. The resolution rings starting from the centre of the plot represent 50, 30 and 20-Å resolution. The crystal appears ordered with bright spots in the FFT to the second order and to around ~ 30-Å resolution. **C. First projection contour map of CFTR crystal.** The crystal was processed in 2dx for lattice refinement, unbending and CTF correction. The generated projection map shows P1 symmetry, with one CFTR molecule per unit cell. One unit cell is boxed showing the principal reciprocal lattice vectors, ‘a’ and ‘b’. The unit cell dimensions are: a = 77.6 Å, b = 75.8 Å, γ = 119.2°. Data collected in collaboration with Dr Mark Rosenberg (Ford Lab, University of Manchester).
lobed triangular configuration. This shape is very similar to CFTR in crystals denoted “crystal form 2” by Rosenberg et al. 2004 (Rosenberg et al., 2004). This shape was interpreted as an open conformation of CFTR because it was similar to the 2D crystal structure of P-glycoprotein in its open, nucleotide-bound state (Rosenberg et al., 2004, Rosenberg et al., 2003). It could therefore be extrapolated that CFTR is crystallised here in an open-channel, active state, which would be consistent with data in previous results section that suggests CFTR from yeast is already activated by phosphorylation (section 3.5), although further investigation would be required to confirm this hypothesis.

Secondly, the map shows CFTR is crystallised here with one molecule per unit cell and $P_1$ symmetry. There are a total of 17 plane group symmetries that a protein can possibly crystallise in 2D and $P_1$ symmetry is of the lowest order, with no point symmetry in the unit cell (Landsberg and Hankamer, 2007). The unit cell is a rhombus with dimensions: $a = 77.6$ Å, $b = 75.8$ Å, $\gamma = 119.2^\circ$. From a top view of the crystal plane, the length of the CFTR molecule in this crystal is around 70-Å. These unit cell dimensions and $P_1$ symmetry are very similar to the parameters of previously crystallised CFTR, which appeared to be monomeric in the final 3D structure map (Rosenberg et al., 2004, Rosenberg et al., 2011). Also, these dimensions are shorter than those reported for dimeric CFTR by SPA (Zhang et al., 2009). Therefore, the data reported here implies that CFTR was crystallised as a monomer, which is particularly interesting as the very same protein sample appears to form larger complexes in solution, as judged by NS-EM and SPA (Figure 22). It is thus difficult to predict the relevant biological oligomeric state from this data, especially because any CFTR organisation could be artificially induced by detergent or during crystallisation (Larue et al., 2009, Fujiyoshi and Unwin, 2008, Tate, 2006). There is ongoing debate whether CFTR functions as a monomer or dimer in vivo (Larue et al., 2009) (Zhang et al., 2009, Rosenberg et al., 2004, Ramjeesingh et al., 2001, Haggie and Verkman, 2008, Mio et al., 2008, Chen et al., 2002, Eskandari et al., 1998, Schillers et al., 2004).

The crystal reported here is not sufficiently ordered for further study by cryo electron crystallography and generation of a 3D structure model, as was previously possible (Rosenberg et al., 2011). This suggests that the CFTR purified from yeast is less homogenous than the CFTR purified from mammalian cells. The possible reasons for this are next discussed. First, CFTR was only purified in DDM to 60% purity due to contamination by the yeast ribosomal L3 subunit protein (section 3.8 & 3.9). Such level of contamination would be expected to prevent protein crystallisation altogether, but
perhaps the conditions used for 2D crystal growth are able to specifically target and effectively purify CFTR by crystallisation. Second, heterogeneity may be caused by differences in PTMs (Reeves et al., 2002). CFTR purified from yeast is believed to be phosphorylated (section 3.5) compared to the un-phosphorylated CFTR isolated from mammalian cells (Zhang et al., 2009) and could thus exist as a more heterogenous populations of different conformational states that could affect crystal order (DePristo et al., 2004). Glycosylation, on the other hand, does not seem a likely candidate to have caused lower crystal order because it is already known that these 2D crystals can tolerate a fully mature CFTR protein with complex sugar modifications, probably because the glycosylated extra-cellular loops are not involved in the protein-protein contacts in the $x$ and $y$ plane (Rosenberg et al., 2004, Zhang et al., 2009). Next, the CFTR purified here appears more heterogenous in size with a greater amount of HMW aggregates, as judged by comparing single particle EM data (Zhang et al., 2010), which will inhibit crystal growth (Borgstahl, 2007). Finally, the yeast-expressed CFTR has additional GFP- and SUMO-fusion tags that can increase conformational heterogeneity and may not be easily accommodated in the crystal lattice (Smyth et al., 2003).
Chapter 6 – Results and Discussion: Activity

This thesis presents a new method for the purification of CFTR protein from a yeast expression system. Demonstrating CFTR activity is important to validate this yeast system as a suitable host to express a functional protein from a foreign higher eukaryote, and to validate the purification procedure as able to preserve CFTR in a correctly folded state. As a membrane protein, correct CFTR folding is of particular concern because detergents often induce non-native conformations (Larue et al., 2009) and can sometimes cause protein denaturation, which will inhibit later crystallisation efforts (Privé, 2007). It is also important for future structural studies that this purified CFTR is active so that any structures that are obtained are biologically relevant. Furthermore, this thesis presents the purification of chicken-CFTR, which is the first time an orthologue other than human-CFTR has been isolated and characterised in vitro, although there has been some study of the chicken-CFTR orthologue in vivo (Aleksandrov et al., 2012). Characterisation of chicken-CFTR activity is therefore also required to validate this orthologue as a suitable model for human-CFTR in later CF drug design.

The biological function of CFTR is to act as a chloride ion channel at the surface of epithelial cell membranes (Riordan, 2005). CFTR channel function has been extensively studied by measuring the flux of ions across a membrane either biochemically, e.g. using chloride sensitive fluorescent probes (Verkman, 1990) or biophysically, e.g. single channel recordings (Aleksandrov and Riordan, 1998) and patch clamping (Sheppard et al., 1993). Assaying for chloride ion flux is, however, experimentally very difficult using purified proteins in vitro. Such assays require overcoming the challenge of reproducibly inserting the purified protein from a detergent micelle into a tightly-sealed, synthetic membrane that is able to mimic the ion impermeability of a cell plasma membrane with distinct internal versus external environments (Demarche et al., 2011). Although one group has succeeded in this type of assay for CFTR (Eckford et al., 2012), we were unable to reproduce the required reconstitution of CFTR into tightly sealed lipid vesicles (section 6.1). Therefore, an alternative activity assay was sought that did not rely on such CFTR reconstitution.

As a member of the ABC transporter family, CFTR possesses well-conserved nucleotide binding domains (NBD1 and NBD2) and NBD2 is predicted to have ATP hydrolysis capability whilst NBD1 retains ATP binding (Lewis et al., 2004,
Aleksandrov et al., 2002b); ATP hydrolysis is believed to act as the driving force behind CFTR conformational changes required for channel gating (Aleksandrov et al., 2007, Gadsby et al., 2006). ATP hydrolysis reactions are commonly used to probe the activity of purified ABC transporter proteins (Swartz et al., 2013, Infed et al., 2011, Rosenberg et al., 2010, Chavan et al., 2013, Shintre et al., 2013), including CFTR (Li et al., 1996, Aleksandrov et al., 2002a). There are a number of assays that can be used to test ATPase activity, including the malachite green assay (Baykov et al., 1988), the NADH-linked assay (Trentham et al., 1972), radio-labelled ADP release (Li et al., 1996) and the Chifflet assay (Chifflet et al., 1988). This thesis uses a modified version of the Chifflet assay (Rothnie et al., 2001), which is a colorimetric assay that measures the amount of P as they are hydrolysed from ATP by the target enzyme.

6.1. Reconstitution of CFTR for activity studies

Studying the activity of purified membrane proteins is a greater challenge than for soluble proteins, as extraction from a lipid membrane into a detergent micelle can alter native protein function (Privé, 2007, Seddon et al., 2004, Garavito and Ferguson-Miller, 2001, Infed et al., 2011). Detergents can remove associated lipids that are required for protein structure and function, or a detergent can itself cause non-native protein conformations. The effect of detergents on protein activity can vary with detergent type; mild detergents are believed to preserve the native fold and activity of a protein during purification, whereas more harsh detergents risk denaturation and loss of protein function. Reconstitution is a method that involves inserting a purified protein back into a more native-like lipid membrane environment by careful removal of detergent in the presence of lipids (Rigaud and Levy, 2003). This technique is commonly employed to recover membrane protein function after purification, and has been shown to stimulate activity of a number of purified ABC transporter proteins (Rothnie et al., 2001, Chavan et al., 2013, Telbisz et al., 2013, Zehnpfennig et al., 2009), including CFTR (Aleksandrov et al., 2002a, Eckford et al., 2012). Reconstitution is, however, a difficult method because detergent removal can easily result in protein aggregation. At present, there exists no single reproducible method of reconstitution; variable factors include lipid composition, lipid-to-protein ratio, type of detergent, and type of method to remove detergent (e.g. dialysis, hydrophobic adsorption, dilution) (Rigaud, 2002, Rigaud and Levy, 2003). The search for optimal conditions for successful reconstitution of a given membrane protein is largely empirical.
CFTR was purified from our yeast system in the detergents LPG and DDM, both of which have low CMCs. A popular method for the removal such low CMC detergents during protein reconstitution is the use of BioBeads (Rigaud et al., 1998). BioBeads are polystyrene beads, which are covered in small, hydrophobic pores around 90-Å diameter that adsorb detergent molecules (Rigaud et al., 1998). Purified CFTR was reconstituted using BioBeads, following the same method described previously for the reconstitution of Pgp (Rothnie et al., 2001). Firstly, detergent-purified CFTR at 0.1 mg/ml was mixed with an excess of lipids (E.coli total lipid extract, with estimated PE/PG/CA at 6/1.5/1 w/w) and cholesterol to a lipid/cholesterol/protein/detergent ratio of 100/25/1/1 (w/w). A high LPR was recommended to increase the efficiency of reconstitution (Geertsma et al., 2008). Cholesterol was included because it is often important for the function of reconstituted membrane proteins (Telbisz et al., 2013, Rothnie et al., 2001, Perez-Castiñeira and Apps, 1990, Ohhashi et al., 1992) and sterols have been used previously in lipid mixtures for CFTR reconstitution and activity experiments (Aleksandrov et al., 2002a, Eckford et al., 2012). Secondly, the lipid and protein mixture was incubated with SM2 BioBeads (BioRad), which was done overnight at 4 °C due to the poor stability of CFTR (chapter 4). The result was a cloudy solution, no longer with the surfactant properties of dissolved detergent, indicative of liposome formation (Geertsma et al., 2008, Rigaud and Levy, 2003). Figure 24A shows that there was nearly 100 % recovery of CFTR after this first BioBead incubation, by comparison of the CFTR band intensity across lanes 1 and 2 on the SDS-PAGE gel. This sample was also analysed by NS-EM. The micrograph in Figure 24B shows lipid vesicles ranging in size up to around 150 nm in diameter (red arrows), as well as a number of smaller, spherical lipid “patches” only 20-30 nm in diameter (black arrows). Due to the high LPR and low protein concentration on this grid (1 µg/ml), it is difficult to observe any individual CFTR protein molecules in the lipid structures on this micrograph. Still, the number of CFTR molecules per vesicle can be estimated using the following calculation: the 150 nm vesicles would have a surface area of $14.1 \times 10^5 \text{Å}^2$, and assuming the average area occupied by a phospholipid molecule is 50 Å² (Levitzki, 1985), there should be $2.8 \times 10^3$ lipid molecules per vesicles. For this experiment, 2 mg of lipids were used, which equates to $1.6 \times 10^{18}$ molecules, given 746.35 as the average molecular weight of the E.coli lipids and cholesterol used (http://avantilipids.com). The average number of vesicles should therefore be around $5.7 \times 10^{14}$ vesicles. Assuming 100 % recovery of CFTR, there should be 20 µg of CFTR, which equates to $5.5 \times 10^{13}$ molecules, given 220 kDa molecular weight for CFTR, as calculated using the online
ExPASy software (http://web.expasy.org/compute_pi/). Therefore, there should be around one CFTR molecules for every ten 150 nm lipid vesicles.

The presence of liposomes indicates that detergent was sufficiently removed by the BioBeads to a concentration below its CMC, but the lipid “patches” are typical of a reconstitution intermediate, suggesting there was still some residual detergent in the sample preventing the vesicles from fully sealing (Vinson et al., 1989). Subsequently, CFTR was incubated with a second BioBead addition to try and fully remove the detergent. Unfortunately, this second incubation step resulted in 100 % loss of CFTR, which appeared to aggregate on the surface of the BioBeads (Figure 24A, lane 3 & 4). This problem was common in reconstitution efforts with our purified CFTR protein in both DDM and LPG, indicating that CFTR is an unstable protein. It was therefore concluded that CFTR could be reconstituted into a lipid/detergent/protein mix but, under these conditions, could not be fully inserted into tightly-sealed lipid vesicles.

To date, only one research group has demonstrated the ability to reconstitute purified CFTR into tightly-sealed lipid vesicles, which has been validated by the development of a robust ion flux assay (Eckford et al., 2012). This group similarly used CFTR purified in low CMC detergents, but there are a number of differences that could have accounted for a more successful generation of sealed vesicles. Firstly, instead of BioBeads, they used an Extracti-Gel D detergent-binding column (Pierce), although this column method is still based on the same surface adsorption of detergent as BioBeads. The lipid composition is different (PE/PS/PC/ergosterol at 5:2:1:1 w/w), although they estimate a similarly low number of CFTR molecules per vesicle, as reported here (around 1 CFTR per 5 vesicles). Additionally, their CFTR was purified from insect cell lines, which may be more stable and able to survive detergent removal during reconstitution, although there is no published data on the stability of CFTR purified from insects cells to confirm this theory. It would be also be interesting to collect NS-EM data of their reconstituted CFTR to confirm that this group is, in fact, producing uni-lamellar, sealed vesicles. Extensive trials would be needed to find conditions that allow reconstitution of our purified CFTR into tightly sealed vesicles, such as alternative lipid compositions and detergent removal methods. Due to time restrictions on this project, CFTR was simply collected after the 1st BioBead addition for further activity studies described in this chapter, and this sample is henceforth referred to as reconstituted CFTR.
Figure 24. CFTR was not fully reconstituted into liposomes but rather formed a lipid/detergent/protein mixture after incubation with Biobeads.

A. SDS-PAGE analysis revealed significant losses of CFTR protein during detergent removal by Biobeads. Samples throughout the reconstitution process were taken and analysed by SDS-PAGE, shown under GFP fluorescence. DDM-purified CFTR was mixed with lipids (1) and then incubated with Biobeads (2) to remove detergent. The Biobead incubation was repeated twice to fully remove detergent (3), after which the Biobeads were collected and denatured by SDS to release any bound CFTR (4). CFTR protein migrated as either an aggregate (A) or a monomer (M).

B. NS-EM analysis after the 1st Biobead addition. Sample was taken after the 1st Biobead incubation (corresponding to lane 2 in Figure A) and analysed by NS-EM. A mixture of larger liposomes around 150 nm in diameter (red arrows) and smaller lipid patches around 20-30 nm (black arrows) were observed, suggesting CFTR is part of a lipid/detergent/protein mixture.
6.2. Effect of lipid and detergent on CFTR ATPase activity

As mentioned in the previous section, the activity of purified membrane proteins can be affected by detergents (Infed et al., 2011) and reconstitution of the protein back into a lipid environment can stimulate biological function (Zehnpfennig et al., 2009, Telbisz et al., 2013). To assess the effect of detergent and lipid on the activity of our purified CFTR, both LPG- and DDM-purified CFTR was tested for ATPase activity before and after reconstitution, using an adapted version of the Chifflet assay (Rothnie et al., 2001, Chifflet et al., 1988). The assay measured the amount of P\(_i\) released from the hydrolysis of 2 mM MgATP by CFTR in 1 hour at 25 °C. Mass spectrometry analysis revealed a number of contaminating yeast proteins with ATPase capability (section 3.9 and Appendix 2). To reduce background ATPase activity of contaminating yeast proteins, these experiment were carried out in the presence of an inhibitor cocktail (IC), consisting of oligomycin, sodium thiocyanate and Sch28080, which inhibit F-, V- and P-type ATPases respectively; these ATPase inhibitors do not affect the gating activity of CFTR (Schultz et al., 1996) and hence were predicted to have no major effect on CFTR ATPase activity. Furthermore, a fraction rich in ribosomal protein L3 (see SDS-PAGE gel in Appendix 4A) - the major contaminating yeast protein in the DDM purification procedure (results 3.8 & 3.9) - was shown to have no ATPase activity as judged by the Chifflet assay (data not shown). Therefore, all detectable ATPase activity of purified material after addition of the inhibitor cocktail was predicted to be largely the result of CFTR.

The ATPase rates for CFTR purified in DDM (D) and LPG (L), either in the detergent-solubilised (-lipid) or reconstituted (+lipid) state are shown in Figure 25. The first thing to note is that activity in both detergent types is stimulated by the addition of lipid. In DDM, the activity is increased by 4.4-fold, from 2.99 ± 0.256 to 13.21 ± 0.256 nmol P\(_i\)/min/mg (n=3). In LPG, the activity was very low in the detergent-solubilised state (0.356 ± 0 nmol P\(_i\)/min/mg, n=3) unless lipids were added, causing an increase in activity by 3.5-fold (1.25 ± 0.33 nmol P\(_i\)/min/mg, n=3). It therefore seems that the lipid/detergent mixture used here is sufficient to stimulate CFTR activity, as has been observed previously for other membrane proteins, for example, the addition of *E.coli* lipids to P-glycoprotein purified from yeast saw an increase in ATPase activity by 5-fold (Lerner-Marmarosh et al., 1999). If complete detergent removal could be achieved, then this ATPase activity rate may be even further stimulated, but unfortunately this was not possible to CFTR aggregation (Figure 24A).
Activation of CFTR could simply be caused by non-specific lipid-protein interactions that, for example, increase lateral pressure and improve protein thermal stability (Morin et al., 1990, Phillips et al., 2009). Stimulation of activity may also be due to highly specific protein-lipid interactions (Raunser and Walz, 2009, Opekarová and Tanner, 2003). For instance, CFTR was reconstituted with mostly PE lipids, which are known to induce specific structural changes in the phosphorylation domain of Ca^{2+}ATPase that increase its catalytic activity (Hunter et al., 1999). It would be interesting to further study alternative lipid compositions for CFTR reconstitution, as this may identify lipids that specifically stimulate CFTR activity.

A second interesting observation in these data is a difference in CFTR activity when the protein is purified in DDM or LPG. After reconstitution with lipids, the activity of CFTR purified in DDM was 10.5-fold higher than the activity detected for CFTR purified in LPG (13.21 ± 0.256 compared to 1.25 ± 0.33 nmol P_i/min/mg respectively, n=3). Another experiment was conducted where the LPG-purified CFTR was exchanged into DDM before reconstitution and ATPase measurements were performed (bars L→D in Figure 25). The ATPase activity of this exchanged sample was 2.35 ± 0.681 nmol P_i/min/mg (n=3), suggesting the low activity is specifically caused by the effect of LPG on CFTR during purification, and not simply interference between residual LPG detergent and components of the Chifflet assay. Again, the low ATPase activity indicates that LPG is a relatively harsh detergent that destabilises CFTR structure. Previous data has suggested that CFTR structure is only partially destabilised by LPG and not fully denatured (sections 4.2 & 5.2.1). Due to the loss in ATPase activity in LPG seen with the data in Figure 25, it can be proposed that LPG at least partly unfolds the NBDs of CFTR, although other domains may be similarly denatured that cannot be detected by this assay. Addition of lipid to the LPG purified material showed an increase in ATPase activity, which may indicate some level of refolding or retention of structure as has been shown for the reconstitution of other membrane proteins after denaturing conditions (London and Khorana, 1982). Even so, this recovery in function after reconstitution is still small compared to the activity of CFTR when purified in DDM, suggesting LPG denaturation is mostly irreversible or that its low CMC reduces the efficiency of detergent removal after one Biobead addition. Alternatively, it could be that CFTR-LPG associates into a more active oligomeric state during reconstitution, e.g. from its monomeric state to the dimers seen for CFTR-DDM (section 5.2), as was similarly achieved with the reassembly of membrane protein complexes during
reconstitution (Hagan et al., 2010). Although previous works have shown that CFTR has similar ATPase activity when reconstituted as either a dimer or monomer (Ramjeesingh et al., 2001). Due to concerns about the folded state of CFTR in LPG, all further ATPase activity assays were carried out using reconstituted, DDM-purified CFTR.

![Figure 25](image)

**Figure 25. CFTR has higher ATPase activity when purified in DDM than LPG, and activity is stimulated by the addition of lipids.**

CFTR was purified in either DDM (D), LPG (L) or exchanged into DDM after purification in LPG (L→D), and then tested for ATPase activity using the Chifflet assay in the absence (-lipid) or presence of lipid (+lipid) at a fixed concentration of 2 mM MgATP at 25 °C. For each sample, the ATPase rate is shown with error bars showing the mean ± S.D. of triplicate experiments.

### 6.3. Kinetics of CFTR ATPase activity

DDM-purified and reconstituted CFTR was incubated with different concentrations of its substrate, ATP. The rate of ATPase activity is shown as a function of ATP concentration in Figure 26A, which can be described by Michaelis-Menten kinetics. The experimental data points were fitted using the Michaelis-Menten equation \( y = \frac{V_{\text{max}} \cdot x}{K_m + x} \), using the GraphPad Software, where \( V_{\text{max}} \) is the maximum rate of velocity at high substrate concentrations and \( K_m \) is the apparent affinity for substrate, also referred to as the concentration of substrate required to achieve half the \( V_{\text{max}} \). The \( V_{\text{max}} \) of CFTR ATPase activity = 13.43 ± 0.53 nmol P\(_i\)/min/mg of CFTR, and the \( K_m = 0.14 ± 0.03 \) mM. The rate of ATP turnover per CFTR molecule is therefore 0.049 s\(^{-1}\), or
about 20 seconds for each CFTR molecule to hydrolyse a molecule of ATP. The CFTR ATPase rate is approaching its maximum at around 2 mM ATP, therefore this concentrations of ATP was used for all other ATPase measurements in this chapter to saturate the binding site, maximise detectable signal and minimise the amount of CFTR required for this assay.

The ATPase activity of CFTR was also measured as a function of hydrolysed P\textsubscript{i} produced over time (Figure 26B). In the first 5 min of the reaction, the amount of P\textsubscript{i} increased linearly with time, giving an initial rate of 89 nmol P\textsubscript{i}/min/mg of CFTR and an ATP turnover rate of 0.33 s\textsuperscript{-1}, or 3 seconds per ATP. After 5 min, however, the amount of P\textsubscript{i} produced begins to plateau relative to time, and between 60 and 180 min there is no further increase in P\textsubscript{i}. The concentration of ATP (2 mM) in this reaction mixture far exceeds the concentration of CFTR (0.76 µM), and even at the maximum rate of 0.33 ATP s\textsuperscript{-1} the ATP concentration would only drop by 0.3 mM over a period of 20 min. This suggests that CFTR is losing its ATPase activity over time for a reason other than limitations in substrate concentration. Because the protein is purified, it seems unlikely that the run down in activity is due to inactivation by other proteins e.g. phosphatases. Such time-dependent run down of CFTR activity was also seen for CFTR channel current readings in whole pancreatic duct cells, indicating the short lifetime is an inherent property of the protein and not just an artefact of purification (Winpenny et al., 1995, Berger et al., 1991, Becq et al., 1994).

A few other research groups have similarly measured the ATPase activity of purified and reconstituted CFTR, namely Christine Bear’s group studying CFTR purified from insect cells (Li et al., 1996) and the Riordan lab studying CFTR purified from mammalian cell lines (Aleksandrov et al., 2002a). The ATPase activity of our yeast-expressed CFTR integrated over 60 min is slightly lower (13.4 nmol P\textsubscript{i}/min/mg) than the reported V\textsubscript{max} of CFTR from insect cells, which ranged from 15 – 73 nmol P\textsubscript{i}/min/mg (Bear et al., 1992, Li et al., 1996, Ramjeesingh et al., 1997, Ramjeesingh et al., 1999a, Ramjeesingh et al., 1999b, Ramjeesingh et al., 2001, Ketchum et al., 2004) and from mammalian cells, which was calculated as 47 nmol P\textsubscript{i}/min/mg (Aleksandrov et al., 2002a). However, when the initial rate at short timescales is considered, then the chicken-CFTR has a higher V\textsubscript{max} than reported elsewhere. Studies of the CFTR derived from insect cells showed a similar affinity for ATP (K\textsubscript{m} = 0.13 - 0.18 mM) (Ketchum et al., 2004, Ramjeesingh et al., 1997) as for our yeast-expressed CFTR (K\textsubscript{m} = 0.14 mM) calculated from the data in Figure 22A. It therefore appears that the CFTR purified from
our yeast expression system has similar affinity for ATP as CFTR purified from other systems. This seems reasonable given the above rationale, as the $K_m$ determination should be relatively unaffected by the integration time used for determining the rate. This emphasises the difficulty of obtaining satisfactory kinetic data when sampling at a single time interval. Clearly, the differences between different reports could also be due to different purification and reconstitution procedures, for example the ATPase activity can be affected by protein-to-detergent ratio, detergent incubation time and purity of the detergent (Infed et al., 2011). Also, the experiments conducted here are the first reported ATPase activity measurements for the chicken-CFTR orthologue; all currently published ATPase measurements of purified CFTR used the human orthologue. Chicken-CFTR has slower channel gating activity than human-CFTR at the temperature of this experiment (25 °C), perhaps because this temperature is further from the biological temperature of chickens (39-43°C) than humans (37°C) and channel opening rates are very temperature sensitive (Aleksandrov and Riordan, 1998, Mathews et al., 1998). This slower gating for chicken-CFTR may account for a slower ATPase rate described in this thesis if the hypothesis is accepted that ATP hydrolysis is associated with channel closing whilst ATP binding is associated with channel opening (Gadsby et al., 2006). Purification of human-CFTR protein from our yeast system is now needed to make a direct comparison of ATPase activity between orthologues and confirm this theory.

Comparison of these activity results to similar experiments with other ABC proteins reveals that CFTR ATPase activity is somewhat lower than other members of this protein superfamily, and the run-down is more striking. For example, the ATPase activity of yeast-expressed and -purified Pgp is reported to reach 1-2 µmol P/min/mg, in the absence of transporter substrate, which is 60-120 times the ATP turnover rate than observed for CFTR (Lerner-Marmarosh et al., 1999, Bai et al., 2011). As already discussed, the relatively low ATPase activity of CFTR is consistently observed in the literature, so it does not appear to be a specific to our particular expression system and purification procedure. The $K_m$ of CFTR, on the other hand, is in a more similar sub-millimolar range as other purified and reconstituted ABC proteins, e.g. the $K_m$ of Pgp = 0.14 mM (Ketchum et al., 2004), and MRP1 = 0.1 mM (Mao et al., 1999). It therefore seems that CFTR binds to ATP with similar affinity as other ABC proteins, but hydrolyses the ATP at much slower rates.
So, why might the ATPase activity of CFTR be so much lower than other ABC proteins? The most obvious explanation for low CFTR ATPase activity is that CFTR, unlike all other ABC proteins, does not technically require the energy from ATP hydrolysis to function; CFTR allows for the passive diffusion of anions through its gated channel pore, rather than other ABC proteins that actively pump substrates against an electrochemical gradient (Aleksandrov et al., 2007, Higgins, 1992). It seems appropriate than the slower rates of CFTR ATP hydrolysis are a reflection of the slow gating kinetics observed for ion channels (Foskett, 1998). ATP binding and hydrolysis is commonly shown to be necessary for the opening and closing of the CFTR channel pore, although the exact mechanisms that link ATPase activity to channel gating remain unclear (Aleksandrov et al., 2002a, Anderson et al., 1991, Vergani et al., 2005, Hwang et al., 1994). The general idea is that binding of ATP to the NBDs causes a conformational change that in turn opens the CFTR channel pore, and subsequent ATP hydrolysis closes the channel (Gadsby et al., 2006). If this model were true, then we would expect a direct coupling between rates of CFTR ATPase and rates of CFTR channel gating. The ATP turnover rate reported here for the V_{max} (20 sec) is 4-fold slower than the rate of chicken-CFTR channel gating cycle, at 25 °C (around 4.5 sec) (Aleksandrov et al., 2012). This discrepancy has led to the proposal of a more complex model to describe the relationship between ATP hydrolysis and gating, including multiple gating events per ATP hydrolysis and rapid transient states between channel opening and closing (Ramjeesingh et al., 1999b, Weinreich et al., 1999, Gunderson and Kopito, 1995). However, if the initial rate of CFTR ATPase activity is used the ATP turnover rate (3 sec) now appears closer to the time taken for an average CFTR gating cycle for chicken CFTR at 25°C (4.5 sec), estimated from single channel recordings (Aleksandrov et al., 2012). In this case, the ATPase activity does seem to be linked to the gating activity of CFTR, at least in the first few minutes after the reaction is initiated, although comparison of the two types of data is problematic due to uncertainty of the exact proportion of purified CFTR that is functional in the ATPase assay. Still, linkage with channel activity would explain CFTR’s uniquely low ATPase activity compared to other ABC proteins.

Another hypothesis for CFTR’s relatively low ATPase is that only one of its NBDs (NBD2) has the complete catalytic “ABC signature” sequence (conserved LSGGQ) and ability to hydrolyse ATP (Aleksandrov et al., 2002b). Isolated NBD1 can bind ATP with the same affinity (K_{m} = 0.14 mM) as shown for the full-length protein in Figure
Figure 26. CFTR ATPase activity as a function of ATP concentration (A) and time (B).
DDM-purified CFTR was reconstituted with lipids and its ATPase activity measured using the Chifflet assay. The rate of ATPase activity was measured as a function of ATP concentration at 25 °C, and this relationship was fitted with a Michaelis-Menten dose response curve using non-linear regression in GraphPad Prism software ($V_{\text{max}} = 13.4 \pm 0.54 \text{ nmol/min/mg}$, $K_m = 0.14 \pm 0.03 \text{ mM}$). Error bars show the mean ± S.D. of triplicate experiments (A). The amount of hydrolysed $P_i$ ions were measured as a function of time to show that CFTR activity stopped after around 2 hours after incubation with 2 mM MgATP at 25 °C (B).
22A \((K_m = 0.14 \text{ mM})\), but its rate of ATP turnover is almost negligible, at around 1 ATP per 8 minutes (Lewis et al., 2004). There are no published data for the ATPase activity of the isolated NBD2, but it can be assumed that the activity detected for full-length CFTR in figure 26A is due to activity solely at the NBD2. Other, more active ABC proteins have a complete catalytic consensus sequence that allows for ATP hydrolysis at both NBDs (Higgins, 1992). The asymmetry between CFTR’s NBDs thus may be the cause of low ATPase activity. In fact, a close relative of CFTR, MRP1, also has relatively low symmetry between the NBD protein sequence, compared to other ABC proteins. Purified MRP1 also shows low rates of ATP hydrolysis to CFTR, at around 5-10 nmol P/min/mg (Manciu et al., 2000). Perhaps these proteins have evolved to have functions that expend less cellular ATP. For example, controlling levels of ATP have been implicated in cross-talk between CFTR and other proteins in the basolateral membrane of epithelial cells (Foskett, 1998).

Finally, low ATPase activity could be a result of damage to CFTR during purification. Rates of ATPase may only therefore appear low because only a small proportion of CFTR protein is actually folded and active during the assay. As CFTR has similarly low ATPase rates throughout the literature, this would suggest a common loss in CFTR activity during all published CFTR purification procedures. Earlier EM (section 5.2.1) and SEC-MALLS (section 5.1) data suggest that CFTR protein readily forms aggregates during purification in DDM. Further evidence for CFTR instability is again highlighted in Figure 26B, where the ATPase activity is lost after incubation at 25 °C for greater than 30 min, unlike other ABC proteins that see a stable ATPase activity over several hours (Mao et al., 1999). This instability is also seen for CFTR activity in whole-cells, where current activity runs down within 20 min after initiating the reaction at 21-23 °C (Winpenny et al., 1995). Such instability again raises concerns about the suitability of this protein for crystallisation trials. However, ATPase activity assays, such that shown in Figure 26B, could prove a useful tool in future screens to find conditions that may stabilise CFTR towards a more crystallisable form.

### 6.4. Effect of PKA treatment on CFTR activity

CFTR channel opening is strictly dependent on phosphorylation of its R region by PKA (Cheng et al., 1991, Seibert et al., 1999). It is believed that phosphorylation of the R region disrupts its interaction with NBD1 that then allows, in the presence of ATP, the formation of an NBD1-NBD2 sandwich heterodimer and subsequent channel opening.
channel gating is therefore modulated by both phosphorylation of the R region and ATPase activity at the NBDs. However, whether the R region phosphorylation directly affects the hydrolysis of ATP itself still remains unclear. Some groups propose that PKA is required to stimulate ATPase activity (Cheung et al., 2008), though others have shown that changes in R region phosphorylation does not alter the binding of ATP to the NBDs (Basso et al., 2003).

The effect of PKA on the ATPase activity was tested using purified CFTR (Figure 27). DDM-purified CFTR was incubated either in the absence or presence of PKA overnight at 4°C, before PKA was removed, CFTR reconstituted and ATPase activity measured using the standard Chifflet assay. The ATPase rate in the absence of PKA treatment was comparable to the rate reported from similar experiments in this chapter (Figure 25 & 26) at 13.05 ± 2.9 nmol P_i/min/mg, implying the additional overnight incubation at 4°C for this PKA experiment did not reduce CFTR activity. After PKA treatment, the rate of ATPase was not significantly changed to 14.00 ± 2.8 nmol P_i/min/mg, as judged by the Kruskall-Wallis statistical test (n=3, p>0.05). Therefore, it can be concluded that PKA treatment did not affect the ATPase activity of our purified CFTR.

In contrast to our results, others have regularly shown that PKA treatment significantly stimulates ATPase activity, for CFTR purified from their insect cell system (Cheung et al., 2008). They observed a 3-fold increase in ATPase activity after PKA treatment, which was attributed to an increase in apparent affinity for ATP (Li et al., 1996). The level of CFTR phosphorylation during expression may depend on cell type and expression protocol, which could in turn account for differences in PKA affect on CFTR activity. For example, data reported in section 3.5 indicates that the yeast-expressed CFTR is more phosphorylated than CFTR purified from mammalian cells (Zhang et al., 2009). Recent mass spectrometry analysis also shows that CFTR purified from yeast cells is phosphorylated at more residues than for CFTR purified from other cell systems (Venerando et al., 2013). Therefore, it could be postulated that the lack of an effect of PKA treatment (Figure 23) is because CFTR from the yeast expression system is already activated and in a highly phosphorylated state. The affinity for ATP reported in Figure 26A (K_m = 0.14 mM) is closer to the ATP affinity reported for phosphorylated (K_m = 0.13 mM) versus de-phosphorylated (K_m = 2.9 mM) CFTR (Ramjeesingh et al., 1997). If the data in Figure 26A was instead fitted with a Hill equation using the GraphPad Prism software (\( v = \frac{S^m V_{max}}{S^m + K} \)), where v and V_{max} are
the initial and maximal rate, $S$ is the concentration of substrate, $m$ is the Hill coefficient and the constant $K$ is $(K_m)^m$, then the Hill coefficient ($m=1.8$) would be closer to that calculated previously for phosphorylated ($m=1.9$) versus non-phosphorylated ($m=1$) CFTR (Ramjeesingh et al., 1997).

It is important to note that the lack of a PKA affect on CFTR ATPase activity shown in Figure 27 could simply be due to damage of CFTR during the purification procedure. This reason was given by Ketchum and co-workers, who also showed no affect of PKA on CFTR affinity for ATP or ATPase rate, despite purifying CFTR from the same insect cell line as others (Ketchum et al., 2004). Overall, it is still difficult to conclude if there is direct coupling between PKA phosphorylation and CFTR ATPase activity. It would be interesting for future experiments to alter the phosphorylation of our purified CFTR e.g. using alternative kinases, phosphatases, serine/threonine mutants and, in combination with mass spectrometry studies, identify if and which phospho-residues are responsible for modulating ATPase activity.

Figure 27. PKA treatment does not affect CFTR ATPase activity.
DDM-purified CFTR was treated with (+) or without (-) PKA overnight at 4 °C. PKA was removed by binding and stringent washing of CFTR to a nickel affinity column. CFTR was eluted, reconstituted with lipids and tested for ATPase activity using the Chifflet assay at a single concentration of 2 mM ATP at 25 °C. For each sample, the ATPase rate is shown with error bars showing mean ± S.D. of triplicate experiments. There is no significant difference between CFTR ATPase activity with or without PKA treatment, as judged by the Kruskall-Wallis test (n=3, p>0.05).
6.5. Inhibition of CFTR ATPase activity

The aim of this project was to express and purify CFTR from a novel yeast expression system, for high-resolution structural studies. Such high-resolution structures could then be used for rational drug design for CF treatment. Also, high yields of purified protein would enable development of new high-throughput drug screens. So far, only the chicken-CFTR orthologue has been successfully purified from our system. Therefore, it is important to validate the pharmacological relevance of the chicken-CFTR orthologue, as well as the use of a foreign yeast expression system, in the treatment of a human disease. Inhibitors are useful to probe specific protein structure and function. There are a number of compounds known to inhibit CFTR activity, which have had medical interest for diseases where CFTR is abnormally over-active e.g. secretary diarrhoea (Verkman et al., 2006). For this experiment, three inhibitors were selected based on the evidence that they inhibit CFTR ATPase activity, as well as channel gating function.

CFTR\textsubscript{inh}-172 is a very potent and specific inhibitor of human CFTR channel activity (Ma et al., 2002a, Verkman, 1990). It is believed to stabilise the closed channel state by binding to a site on CFTR other than the channel pore, possibly on the NBDs in a non-competitive manner with ATP (Taddei et al., 2004). To test the effect of CFTR\textsubscript{inh}-172 on our yeast-expressed chicken CFTR, DDM-purified protein was reconstituted and ATPase measured after treatment with or without 10 µM CFTR\textsubscript{inh}-172, at a single concentration of 2 mM ATP for 1 hour at 25 °C. This data is shown in the first and third bars of Figure 28A, where the ATPase rate is 13.6 ± 0.44 and 13.44 ± 0.45 nmol Pi/min/mg for the untreated and treated, respectively. There was no significant difference in ATPase activity in the presence of CFTR\textsubscript{inh}-172 compared to the untreated sample, as judged by the Kruskall-Wallis test (n=3, p < 0.05). This experiment was repeated a number of times for CFTR purified from different fermenter cultures, but still no inhibition was seen after treatment with CFTR\textsubscript{inh}-172 (data not shown). Therefore, CFTR\textsubscript{inh}-172 seemed to have no affect on the function of our purified CFTR, at least in terms of ATPase activity.

This result on first sight, contradicts previous work that showed CFTR\textsubscript{inh}-172 caused a significant inhibition of ATPase activity of purified and reconstituted human-CFTR from insect cells (Wellhauser et al., 2009, Eckford et al., 2012). To date, this is the only evidence that suggests CFTR\textsubscript{inh}-172 affects human CFTR ATPase activity, although it is very clear from the literature that CFTR\textsubscript{inh}-172 specifically inhibits human CFTR
channel activity (Ma et al., 2002a, Verkman, 1990). The differences in inhibitory effect of CFTR$_{inh}$-172 with our results may simply be due to difference in purification procedure. However, from the data in Figure 28A, it could be argued that CFTR$_{inh}$-172 does not actually alter CFTR ATPase activity. In fact, the inhibition seen by C.Bear’s group would have to be inhibition of ATP binding rather hydrolysis because inhibition of ATP hydrolysis is often associated with prolonged channel opening (Hwang et al., 1994, Gunderson and Kopito, 1994), rather than the prolonged closed state observed during treatment with CFTR$_{inh}$-172 (Ma et al., 2002a, Verkman, 1990). Overall, it is still unclear if inhibitor CFTR$_{inh}$-172 alters CFTR ATPase activity.

The difference in inhibition profiles could also be due to the different CFTR-orthologues used. To date, there has been no study of the effects of CFTR compounds on the chicken-CFTR orthologue. Studies of wild-type mouse-, pig- and killifish-CFTR have shown that these orthologues display similar sensitivity to CFTR$_{inh}$-172 inhibition as the human-CFTR orthologue (Stahl et al., 2012). Shark-CFTR, however, is completely insensitive to this drug (Stahl et al., 2012). The lack of CFTR$_{inh}$-172 inhibition seen for chicken-CFTR does not appear to be simply due to divergence in amino acid sequence because the chicken-CFTR orthologue has higher protein sequence identity (80 %) to human-CFTR than mouse-CFTR (78 %) and killifish-CFTR (58 %) which are all sensitive to CFTR$_{inh}$-172 (Table 1). Also, Caci et al. identified R347 as critical for CFTR$_{inh}$-172 interaction with CFTR, a residue that is conserved in all these orthologues (Caci et al., 2008). Perhaps the interactions with CFTR$_{inh}$-172 and CFTR are dependent on more higher order structural differences between orthologues (Stahl et al., 2012). Overall, these data raise caution for the use of alternative CFTR-orthologues to extrapolate the effect of drugs on the human-CFTR protein.

The second inhibitor used to probe CFTR activity was genistein (GNS). GNS is a flavanoid that, at concentrations above 50 µM, inhibits CFTR channel activity (Cai et al., 2003, Wang et al., 1998a, Cai et al., 2004). At high concentrations, GNS is believed to act by competing for the same binding site in the NBDs as ATP. If this were the case, GNS should inhibit the ATPase activity of CFTR as well as the ion channel function. The effect of 100 µM GNS was tested on purified and reconstituted CFTR, as was described earlier for treatment with CFTR$_{inh}$-172. The ATPase rate after GNS treatment was $13.48 \pm 0.2699$ nmol P$_i$/min/mg, which has no significant difference to the untreated sample (Figure 28A). The effect of GNS on the ATPase activity of CFTR has never before been investigated for the purified protein. It could be that GNS inhibits
CFTR ion flux activity independently of CFTR’s ATPase activity. Alternatively, GNS may not directly interact with CFTR, but regulates its activity by other interacting proteins in the cell. Again, it remains unclear whether this inhibitor does not specifically reduce CFTR ATPase activity or if the protein purified here is not responding to the drug in a biologically relevant manner.

The final compound to be selected as a potential CFTR inhibitor was reduced glutathione (GSH). GSH is a large organic anion and is the most common water-soluble reducing agent in the body (Hudson, 2001). Previously, GSH has been shown to directly interact with purified CFTR and inhibit its ATPase activity (Ketchum et al., 2004, Kogan et al., 2001). No such inhibition was seen with DTT or the oxidised version of GSH, GSSG, suggesting an alternative mechanism for CFTR inhibition other than protein reduction (Ketchum et al., 2004, Kogan et al., 2001). GSH was proposed to stabilise the open channel conformation of CFTR by inhibiting its ATP hydrolysis ability, a mechanism that would require some cross talk between the NBDs and channel pore regions (Kogan et al., 2001). The effect of GSH on the ATPase activity of our purified CFTR was tested and the results are shown in Figure 24A. The ATPase rate of CFTR after treatment with 10 mM GSH was 6.56 ± 2.50 nmol P$_i$/min/mg, which was a significant difference (*) in activity when compared to the untreated control. These experiments were carried out in the presence of 2 mM DTT, therefore it does not seem that inhibition in ATPase activity was simply caused by GSH affecting the redox potential of the CFTR environment, which is a mechanism believed to modulate CFTR channel gating (Harrington et al., 1999). To investigate if the inhibition was biologically important, CFTR ATPase activity was also tested with a range of physiological GSH concentrations, from 1-10 mM (Cantin et al., 1987, Kogan et al., 2001). As shown in Figure 28B, there is a decrease in CFTR activity with increasing GSH concentration, with 50 % inhibition in activity at 10 mM GSH. This inhibition profile is very similar to that seen for ATPase activity of CFTR purified from insect cells, where 10 mM GSH gave around 40 % inhibition (Ketchum et al., 2004, Kogan et al., 2001). Interestingly, these previous reports showed that GSH only inhibited CFTR in its fully phosphorylated state (Kogan et al., 2001). This is consistent with data implying that our yeast-expressed CFTR is purified in an activated phosphorylated state (section 6.4).

As GSH is a naturally occurring molecule, these data raise the question: what would be the biological role of GSH inhibition on CFTR ATPase activity? Linsdell and Hanrahan were the first to show that CFTR-expressing cells can directly transport GSH across the
plasma membrane (Linsdell and Hanrahan, 1998). These findings led to the hypothesis that CFTR, in addition to being a chloride ion channel, can also transport GSH. This hypothesis has had great interest within the CF research field because CF patients have a GSH deficiency in their airway surface liquid (Gao et al., 1999). This GSH deficiency is suggested as a potential cause for the increased mucus viscosity and elevated lung inflammation found in CF patients (Hudson, 2001). It has been speculated that CFTR is modulated by cellular events, for example oxidative stress, to alter its preference for GSH versus chloride in a mechanism similar to that proposed by Quinton and colleagues for the carbonate transport function of CFTR (Reddy and Quinton, 2001). For example, GSH inhibition of ATP hydrolysis could provide the switch that turns off chloride flux and allows the GSH flux function to be switched on. The idea that CFTR is a GSH transporter is further fuelled by comparison with its close relatives within the ABCC sub-family, MRP1 and MRP2, which are known to transport GSH, mostly as a conjugate for larger substrates (Deeley and Cole, 2006). This has led some to believe that CFTR is, in fact, an ATP hydrolysis driven transporter like the rest of the ABC protein family. However, this remains a controversial idea and there is still much debate as to whether the GSH deficiency is simply a secondary effect of CF symptoms rather than a primary cause of the disease (Hudson, 2001).
Figure 28. Inhibition of CFTR ATPase activity.

A. CFTR activity was inhibited by GSH. DDM-purified and reconstituted CFTR was tested for ATPase activity using the Chifflet assay after either no treatment (-) or treatment with inhibitors: 100 µM genistein (GNS), 10 µM inhibitor CFTRinh-172 (172), 10 mM reduced glutathione (GSH). For each sample, the ATPase rate is shown with error bars showing mean ± S.D. of triplicate experiments. Only treatment with GSH showed a significant (*) inhibition of CFTR ATPase activity when compared to untreated sample as judged by the Kruskall-Wallis test (n=3, p < 0.05).

B. CFTR activity was inhibited by increasing concentrations of GSH. DDM-purified and reconstituted CFTR was incubated with increasing concentrations of GSH and then tested for ATPase activity using the Chifflet assay. CFTR ATPase activity was normalized as a percentage of the activity with no GSH. Error bars show mean ± S.D. of triplicate experiments. CFTR ATPase activity shows 50% inhibition with 10 mM GSH.
Chapter 7 - Conclusions

7.1. A novel yeast system for expression and purification of CFTR

The high-resolution structure of CFTR would provide novel insights into its molecular mechanism and also enable the design of new drugs to treat CF. Until now, structural studies of CFTR have been severely limited by insufficient quantities of purified full-length protein. This thesis describes new methods for the expression and purification of CFTR protein from a novel S. cerevisiae expression system. The protocol for CFTR expression has been recently published (O'Ryan et al., 2012) and the protocol for purification is soon to be released (Pollock et al., 2013, in-press). With financial support from the CFFT, these publications are also associated with the release of reagents, including CFTR-vector DNA, CFTR expressing yeast cell lines and purified CFTR protein, which are now being used by the CF research community. For the first time, CFTR protein can be easily produced on a large-scale to milligram quantities, paving the way for high-resolution structural studies, as well as novel biochemical and biophysical characterisation.

Yeast offer great advantages for heterologous expression of recombinant proteins, namely ease and low cost of culturing cells on a large-scale, and simple molecular biology steps that would enable screening of many constructs and mutations (Hays et al., 2010). Despite their allure, yeast have been rarely used before to express CFTR (Huang et al., 1996, Kiser et al., 2001, Fu and Sztul, 2009). In mammalian cells, the majority of CFTR is targeted in the ER for degradation by the proteasome (Loo et al., 1998). In yeast, virtually all expressed CFTR appeared to meet this fate, with little protein ever reaching the PM (Kiser et al., 2001), probably because yeast have particularly stringent ER quality control mechanisms compared to other cell types (Hays et al., 2010). CFTR purified from yeast ER compartments displayed typical CFTR channel gating activity, showing that these cells are capable of producing a functional CFTR protein (Huang et al., 1998). However, purification of CFTR from yeast has not been since reported in the literature, probably due to such problems with degradation. The development of our yeast system for CFTR expression met similar challenges and likely required a combination of factors to limit CFTR degradation. Using a combination of an inducible promoter and a short induction time, expressed CFTR can be collected before it is turned over by the yeast cell degradation pathways (section 3.2). The use of protease inhibitors after cell harvest also greatly improved the resistance of
CFTR to yeast proteases and degradation (section 3.4). Unfortunately, the short harvest
time and the apparent cell toxicity of CFTR induction gave a relatively low yield of
CFTR-expressing cells per litre of cell culture (section 3.2). However, the use of a
large-scale fermenter culture and the addition of glycerol to the induction medium
(section 3.4) can allow for the production of a suitable yield for 3D crystallisation
studies. Another factor may be the use of an N-terminal SUMO-fusion tag, which is
supposed to decrease degradation of recombinantly-expressed proteins (Marblestone et
al., 2006, Zuo et al., 2005, Denuc and Marfany, 2010). Initial evidence suggests that the
SUMO-tag increases CFTR expression and PM trafficking in our *S. cerevisiae* cells (T.
Rimington, I. Urbatsch, unpublished data).

The CFTR purification method described here uses two steps: nickel-affinity and size
exclusion chromatography in the detergents, LPG and DDM (sections 3.7 and 3.8). LPG
gives a better purification, with higher purity (90 %) and yield (111 µg /L) than DDM
(60 % and 60 µg/ml, respectively). Also, CFTR is much more soluble at high protein
concentrations in LPG than DDM, up to 8 mg/ml and 0.4 mg/ml, respectively (Figure
15). In LPG, this is the first time CFTR protein has been purified and concentrated
within a range suitable for 3D crystallisation trials. However, evidence throughout this
thesis raises concerns of the possible destabilising affects of LPG on CFTR structure
and function, including: mass spectrometry (section 3.9), TEV protease cleavage
experiments (section 3.10), thermal stability (section 4.2), cryo-EM (section 5.2.1) and
ATPase activity (6.2) measurements.

Our yeast expression and purification protocol offers the first opportunity to easily
produce large quantities of purified CFTR. However, the use of yeast often raises
concern surrounding the ability of these simple organisms to produce a correctly folded
and active protein that is as complex as CFTR. So, how does the quality of our CFTR
compare to protein purified using existing mammalian (Rosenberg et al., 2004) and
insect cell expression systems (Eckford et al., 2012)? This thesis describes a number of
methods to assess the functional quality of our purified CFTR protein. From
fluorescence microscopy and analysis of PTMs, it appears that yeast, at least partially,
recognises CFTR for correct processing and localisation during protein synthesis
(sections 3.3 & 3.5). Fluorescent spectroscopy techniques, including the CPM dye
binding and Trp fluorescence assays, were used to probe CFTR thermal stability and
showed that the purified protein had a clear two-state thermal unfolding transition after
purification in DDM (N→D), occurring over a broad temperature range between 25 and
60 °C and with a T\textsubscript{m} of approximately 41 °C (section 4.2). Until now, there has been no characterisation of full-length CFTR thermal stability, probably due to lack of sufficient quantities of purified protein. However, the thermal unfolding data reported here are consistent with similar experiments on other membrane proteins known to be folded and active, (Alexandrov et al., 2008) and are very similar to the unfolding data collected for CFTR purified from mammalian HEK cells (donated from Kappes and Urbatsch labs) (section 4.2.1). Therefore, it appears that CFTR expressed in our yeast system is correctly folded into a cooperative tertiary structure. In LPG, no such cooperative thermal unfolding transitions were observed (section 4.2), which agrees with the idea that LPG is a relatively harsh detergent and, at least partially, destabilises CFTR structure during purification. Therefore, it seems that CFTR is expressed in our yeast using DDM with an intact cooperative tertiary structure that is preserved during purification in DDM but not LPG.

Another read-out for CFTR functional quality described in this thesis is its ATPase activity (chapter 6). DDM-purified and reconstituted chicken-CFTR from our yeast system showed a similar ATPase activity as that reported for human-CFTR purified and reconstituted from mammalian (Aleksandrov et al., 2002a) and insect cells (Li et al., 1996), suggesting a common function for CFTR, regardless of expression system or orthologue. These ATPase activity rates are relatively low compared to the rates observed for other ABC proteins (Lerner-Marmorosh et al., 1999). In combination with data for CFTR channel activity (Aleksandrov et al., 2012), there is mounting evidence that suggests CFTR’s low ATP turnover rate is probably due to coupling with its relatively slow channel gating kinetics, as described in the model by Gadsby (Gadsby et al., 2006). The run-down in CFTR ATPase activity shown in Figure 26 also suggests that the run-down in channel activity seen for CFTR \textit{in vivo} (Winpenny et al., 1995) is an intrinsic property to the protein and is not caused by interactions with other proteins because the ATPase assay was conducted with purified protein. Therefore, CFTR function appears to be self-regulated, in addition to regulation from other proteins (Dulhanty and Riordan, 1994). This run-down in ATPase activity is unique to CFTR among ABC proteins and likely correlates to its similarly unique short half-life at the PM. Such a short half-life for CFTR function at the PM is probably required for fine-tuning of such a complex ion channel, although this seems to be a great waste of protein and reasons for such high levels of CFTR recycling still remain a puzzle. LPG-purified and reconstituted CFTR showed a reduced level of ATPase activity again indicating a
partial denaturation of CFTR structure and function by this detergent (section 6.2) or alternatively a faster run-down of ATPase activity in LPG. Interestingly, the protocol for purifying CFTR from insect cells first solubilises the protein using the harsh detergents PFO or FC16 before exchanging into the milder DDM (Eckford et al., 2012), and LPG was previously used to solubilise CFTR purified from yeast (Huang et al., 1998). In both cases, CFTR retained channel activity after purification, suggesting CFTR folding must be, to some level, resistant to denaturation by these charged detergents. This highlights the importance for selecting the correct detergent when purifying a membrane protein for functional studies (Privé, 2007).

The similar observations of ATPase inhibition by GSH (section 6.5) to that previously reported for purified human-CFTR helps validate our yeast system as a tool for development of drugs in the treatment of CF (Ketchum et al., 2004). The effect of GSH is also interesting because of various hypotheses that CFTR has an alternative function as a GSH transporter, and deficiency in GSH transport can lead to CF (Linsdell and Hanrahan, 1998, Gao et al., 1999). The lack in response to CFTR<sub>inh</sub>-172 and GNS, however, raises concerns about the pharmacological relevance of the chicken-CFTR orthologue for screening compounds to target human-CFTR (section 6.5). There is still little known about the function of purified CFTR, so it remains unclear if these compounds directly affect CFTR ATPase activity or if they inhibit channel activity through other mechanisms. Overall, CFTR purified in DDM from our novel yeast expression system appears to be of similar functional quality to protein purified from already established insect and mammalian cell systems, suggesting yeast are a suitable host from which to purify CFTR for novel biochemical and biophysical characterisation.

7.2. Remaining challenges in achieving a high-resolution structure of CFTR

We can now obtain sufficient yield of purified CFTR for 3D crystal trials, but how close are we to achieving an X-ray crystal structure of CFTR? This thesis presents the purification of CFTR in two detergents: LPG and DDM. In LPG, CFTR appears to have two of the main qualities required for successful crystallisation that have not been previously achievable for purified CFTR (Zhang et al., 2009, Zhang et al., 2010): solubility at high protein concentrations (<8 mg/ml) (section 4.1) and homogeneity (Figure 21). However, as discussed above (section 7.2), it appears that the benefits of LPG come at the cost of partial unfolding or destabilization of CFTR structure, which
could inhibit rigid protein-protein contacts needed for crystal growth. The choice of detergent for CFTR purification is thus crucial for its later crystallisation success. Conversely, in DDM, CFTR appears folded and active, but lacks the qualities in solubility and homogeneity required for crystallisation; DDM-purified CFTR is only soluble at concentrations below 0.4 mg/ml (section 4.1) and appears to be a very heterogenous population of different sized particles (section 5.1 & 5.2.2). From the SEC-MALLS data, it seems that CFTR aggregation is very dynamic in a DDM buffer solution (section 5.1), probably fluctuating between a population of monomers and soluble HMW aggregates. The dimeric particles selected from EM micrographs in Figure 22 likely represent a discrete population of these HMW particles. From light scattering measurements, CFTR forms irreversibly insoluble aggregates with increasing concentrations of protein (section 4.1). The aggregation behaviour of CFTR in DDM buffer solution can be hence be described: M ⇌ HMW → A, where M is monomers, HMW are soluble higher molecular weight aggregates and A are insoluble, irreversible aggregates. LPG seems to inhibit the formation of M→HMW (sections 5.1. & 5.2.1) and high salt appears to slow the process down (section 3.8). Perhaps LPG coats the protein in a negative charge that causes electrostatic repulsion between CFTR molecules and high salt provides the buffer with sufficient ionic strength to stabilise CFTR monomers. At low protein concentrations, it seems that 2D crystals can grow from the population of CFTR monomers (section 5.3), but CFTR would enter an irreversibly aggregated state before ever reaching protein concentrations required for 3D crystallisation.

From this thesis, it can be said the biggest obstacle now facing X-ray crystallography studies of CFTR is the aggregation behaviour of the purified protein in solution. This aggregation problem seems to be common with CFTR purified from other cell expression systems. CFTR purified in DDM from BHK cells similarly has a maximum protein concentration limit of 0.4 mg/ml before aggregation, which was also inhibited by LPG (Zhang et al., 2009, Zhang et al., 2010). NS-EM analysis of BHK-expressed CFTR also revealed a heterogeneous population of HMW aggregates, even after fractionation on a SEC column, indicating similar dynamics to CFTR aggregation as our yeast-expressed protein (Zhang et al., 2010). Additionally, this CFTR aggregation problem seems to be common for different orthologues. This thesis presents data for chicken-CFTR protein, which is the first time another orthologue to human-CFTR has been studied in its purified form. Interestingly, the aggregation problem seems be just as
apparent for the purified chicken-CFTR protein (this thesis) as human-CFTR (Zhang et al., 2010), despite a 20 % difference in amino acid sequence identity (Table 1). In fact, it was surprising that the chicken-CFTR offered no clear improvements in CFTR purification and stability despite the reported increased thermostability of this orthologue in live cells (Aleksandrov et al., 2012). Perhaps the higher thermostability of chicken-CFTR protein in vivo is not preserved in vitro after purification in a detergent micelle, as indicated by the similar unfolding profile compared to the human-CFTR protein in Figure 16. It is important to note that a direct comparison between the chicken- and human-CFTR orthologues cannot be accurately made until these two proteins have been purified from the same expression system. Still, it can be said that the propensity of purified CFTR protein to aggregate in vitro is an intrinsic property of the protein that is conserved across orthologues and is not related to its thermodynamic stability in vivo, and will likely prove a similar problem regardless of the expression system used.

CFTR is an integral membrane protein and purification requires its solubilisation from a cellular membrane into a non-native detergent micelle environment (section 3.6). The lipid bilayer of cell membranes provides a lateral pressure that is important for membrane protein structural and functional integrity (Phillips et al., 2009) and is lacking in a detergent micelle (Privé, 2007). Therefore, it would seem easy to account for the CFTR instability and its tendency to aggregate by the use of detergents during purification. However, many other ABC proteins can be stably isolated in detergent micelles at relatively high protein concentrations, without the same problems of aggregation observed for CFTR (Oldham et al., 2007, Pinkett et al., 2007, Locher et al., 2002). Also, an increasing number of X-ray crystal structures of eukaryotic ABC proteins show that it is possible to isolate these complex membrane proteins as a sufficiently homogenous population of protein particles for successful 3D crystallisation (Shintre et al., 2013, Aller et al., 2009, Jin et al., 2012). This poses the question: what is different about CFTR compared to other ABC proteins that could cause such a unique problem with aggregation at relatively low protein concentrations? Structurally, one potential answer is the presence of the R region. The R region is highly disordered protein region that is considered a dynamic “hub” to which numerous binding partners can interact during regulation of CFTR function (Chong et al., 2013). The disordered nature of the R region allows for transient and highly dynamic interactions that allow quick on-off regulation and complex fine-tuning of CFTR
channel gating (Baker et al., 2007). Interestingly, recent NMR studies have shown that the R region not only interacts with other proteins (e.g. PKA), but can also interacts with other domains on the same CFTR molecule as part of its regulatory function, namely NBD1, NBD2 and the C-terminus (Baker et al., 2007, Bozoky et al., 2013a, Bozoky et al., 2013b). In live cells, these R region interactions with CFTR sub-domains would primarily occur across one CFTR molecule because evidence indicates that CFTR monomers are kept at well-separated and restricted regions of the plasma membrane by interactions with the cytoskeleton (Haggie and Verkman, 2008). During microsome preparation and detergent solubilisation, the restrictions on CFTR movement by the cytoskeleton will be lifted, hence allowing individual CFTR molecules to freely self-interact. It can therefore be speculated that the tendency of purified CFTR to aggregate in vitro is a kinetic problem due to interactions between the R region on one molecule and other sub-domains on adjacent molecules. This model would be consistent with the observation that the formation of CFTR HMW aggregates is transient and dynamic (section 5.1 & 5.2.2) and that this aggregation is protein concentration dependent (section 4.1). It would also explain the heterogeneity in CFTR HMW aggregates morphology seen by NS-EM (section 5.2.2) because R region intra-molecular interactions are very heterogenous with multiple and overlapping binding sites. Finally, the R region and its intra-molecular interacting sites are highly conserved in sequence and structure across the animal kingdom, which would explain unifying mechanism of aggregation across CFTR orthologues (Kanelis et al., 2011). According to this model, R region intra-molecular interactions are responsible for driving M→HMW, which would increase the chance of HMW→A formation, possibly due to interactions between disordered R regions on adjacent molecules. Therefore, purification of CFTR is a race against time following detergent solubilisation, in order to isolate the protein before it irreversibly aggregates, where low temperatures proved vital throughout the procedure to slow kinetics of aggregation. The ability of LPG to prevent M→HMW aggregation could be due to denaturation of protein structures that are required for such R region intra-molecular interactions, for example folds within the NBD1 that are known to interact with the R region α-helices (Baker et al., 2007). Denaturation of the CFTR NBDs by LPG is already implicated by a reduced ATPase activity compared to the DDM-purified protein (section 6.2).

Overall, it appears that the highly dynamic and disordered nature of CFTR, which is inherent to its function as a complex ion channel, is the reason this protein has had little
success in purification and crystallisation. Using current CFTR constructs, it seems unlikely that we will achieve a monodisperse population of folded and active CFTR at the high protein concentrations required for successful 3D crystallisation experiments, regardless of orthologue or expression system used. Even so, encouragement can be found in the fact that CFTR can form well-ordered 2D crystals (Rosenberg et al., 2011). In the full-length protein, the R region seems to have a relatively well-defined location that is accommodated within the crystal lattice in CFTR 2D crystals (Rosenberg et al., 2011, Zhang et al., 2010), even if it is probably too disordered to crystallise as an isolated protein (Baker et al., 2007).

**7.3. Future Work**

Our protocol for CFTR expression and purification from *S. cerevisiae* still has a few limitations: Firstly, yeast-expressed CFTR purified in DDM is contaminated by the ribosomal L3 subunit protein (section 3.8). Higher levels of purity are now needed for further characterisation and structural studies. Unfortunately, L3 could not be removed by SEC purification (section 3.8) or filtration (data not shown), and a variety of other techniques were not compatible with the high salt buffer, e.g. ion exchange or heparin chromatography. Engineering additional affinity tags into the CFTR construct could prove useful for further purification, for example FLAG-tag (Einhauer and Jungbauer, 2001) and STREP-tag (Schmidt and Skerra, 2007). Secondly, our purified CFTR has not yet been tested for channel activity, which is needed to both validate our yeast system as a suitable host for CFTR production and for further characterisation of CFTR molecular mechanisms. Channel activity can be measured by ion flux assays or single channel recording measurements. Within this project, CFTR could not be reconstituted into vesicles that were sufficiently sealed for such experiments, as discussed in section 6.1. Extensive reconstitution trials are now needed to achieve this, which could explore variables such as lipid composition, detergent type, LPR, temperature and method of detergent removal. Thirdly, the purification of the human-CFTR orthologue from our yeast expression system is now needed to make accurate comparisons to chicken-CFTR protein. So far, the chicken-CFTR orthologue seems to offer no clear advantage to purification compared to previous work on the human-CFTR protein (Zhang et al., 2009, Zhang et al., 2010), despite amino acid residues that confer greater thermodynamic stability *in vivo*. If this is the case, then work should continue primarily with the more medically relevant human-CFTR protein, especially as data in chapter 4
implies that the chicken-CFTR protein may not respond similarly to CF drugs as the human-CFTR counterpart.

The final challenge that remains in this project is producing a purified CFTR protein that is sufficiently homogenous for 3D crystallisation studies. As described in section 7.1, purified CFTR has a tendency to aggregate into a heterogeneous population of differently sized particles, which may be caused by dynamic interactions between the R domain and other CFTR sub-domains. If so, future experiments could trial methods to limit such R region interactions, for example: newly available detergents (e.g. neopentyl glycols and steroid-based facials) (Lee et al., 2013, Chae et al., 2010), truncation mutations in the C-terminus, mutations that could potentially lock CFTR in specific conformations (e.g. G551D), mutations in R region phosphoresidues, incubation with CFTR-binding proteins (e.g. NHERF) or Fab fragments to provide a scaffold around flexible regions.

Improving CFTR thermostability may also be important for successful 3D crystallisation, as is the case for other membrane proteins (Serrano-Vega et al., 2008). Analysis of the chicken-CFTR protein indicates that difference in amino acids residues that confer improved thermostability over human-CFTR in vivo do not necessarily translate to improved thermostability for the purified protein in vitro (Figure 16). Still, the remaining orthologues (Table 1) should be screened for improvements in purification and crystallisation. Additionally, there are a number of thermostabilising mutations that could be introduced into our CFTR gene that are known to improve CFTR biogenesis in vivo (Teem et al., 1993) and were used to obtain X-ray crystal structures of the isolated NBD regions (Lewis et al., 2004). Alterations in the glycosylation of CFTR may improve thermostability, as indicated by de-glycosylation experiments in section 3.5. Also, different buffer conditions and detergents can often improve stability of membrane proteins (Alexandrov et al., 2008). It is important, however, that any improvements in stability or homogeneity do not risk denaturing CFTR structure, as appears to be the case with LPG detergent. Also, caution should be taken when introducing mutations into CFTR because this may result in a non-native protein behaviour and false interpretation of biological function in later biophysical and biochemical studies. For example, the thermostabilising mutations that were required to crystallise the isolated NBDs may explain why there has been limited evidence for their dimerisation in vitro, despite clear evidence that NBD1 and NBD2 form a functional heterodimer in vivo (Hunt et al., 2013). In fact, the detection of a tight NBD1/NBD2
complex by SAXS analysis on constructs containing no such thermostabilising mutations would suggest that this is the case (Galeno et al., 2011).

It is likely that a large trial of constructs and conditions will be needed before an X-ray crystal structure of CFTR is achieved. The novel yeast expression system presented here offers the first opportunity to easily carry out such a screen. Biochemical and biophysical assays described in this thesis are now in place to assess the quality of CFTR purified for 3D crystallisation trials and should next be adapted for high-throughput screening assays. The GFP-fusion tag could also prove a useful probe for CFTR protein quality (Kawate and Gouaux, 2006).
References


154


ROSENBERG, M., KAMIS, A., CALLAGHAN, R., HIGGINS, C. & FORD, R. 2003. Three-dimensional structures of the mammalian multidrug resistance P-


TELBISZ, Á., ÖZVEGY-LACZKA, C., HEGEDŰS, T., VÁRADI, A. & SARKADI, B. 2013. Effects of the lipid environment, cholesterol and bile acids on the


Appendix – Supplementary Figures

Appendix 1. Quantification of CFTR.
A. BSA standard SDS-PAGE gel. BSA of known concentration was loaded onto an 8 % SDS-PAGE gel next to different volumes of purified CFTR, and stained with Coomassie blue (left). The BSA band intensity was quantified using Image J and plotted against BSA concentration (right). The concentration of CFTR was determined as 448 µg/ml, by comparing the protein band intensity with the standard curve line equation B. GFP fluorescence standard curve. CFTR was diluted to different concentrations and the GFP fluorescence measured in a fluorometer (excitation = 485nm, emission = 512nm).
Appendix 2. Mass spectrometry analysis of CFTR purified in LPG and DDM.
SEC-purified CFTR in both LPG and DDM were digested by trypsin and peptides separated by reverse phase chromatography. Peptides were matches against the IPI Chicken (version 3.64) and yeast Uniprot (2011-05) database and the proteins identified are listed in the table, with the most certain identifications at the top of the list. Proteins that are commonly identified in both LPG- and DDM-purification are indicated *.

### LPG-purification

<table>
<thead>
<tr>
<th>Identified Protein</th>
<th>Database accession number</th>
<th>Molecular weight (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cystic fibrosis transmembrane conductance regulator *</td>
<td>IPI00575922</td>
<td>168</td>
</tr>
<tr>
<td>GTPase-activating protein for Rho1 and Rho2</td>
<td>BEM2_YEAST (+1)</td>
<td>245</td>
</tr>
<tr>
<td>Adenylate cyclase</td>
<td>CYAA_YEAST</td>
<td>228</td>
</tr>
<tr>
<td>Heat shock protein 70*</td>
<td>HSP71_YEAST</td>
<td>70</td>
</tr>
<tr>
<td>Acetyl CoA carboxylase *</td>
<td>A6ZS90_YEAS7 (+5)</td>
<td>250</td>
</tr>
<tr>
<td>Fas2p</td>
<td>C7GU70_YEAS2 (+3)</td>
<td>207</td>
</tr>
<tr>
<td>Microtubule-associated protein</td>
<td>A6ZPT5_YEAS7 (+1)</td>
<td>155</td>
</tr>
<tr>
<td>Ubiquitin carboxyl-terminal hydrolase</td>
<td>A6ZRA8_YEAS7 (+3)</td>
<td>102</td>
</tr>
<tr>
<td>Fas1p</td>
<td>E7NJT1 YEASO (+1)</td>
<td>229</td>
</tr>
<tr>
<td>1-phosphatidylinositol-3-phosphate 5-kinase</td>
<td>A7A257_YEAS7 (+4)</td>
<td>257</td>
</tr>
<tr>
<td>H(+)-ATPase *</td>
<td>A6ZUY5_YEAS7 (+7)</td>
<td>100</td>
</tr>
<tr>
<td>Translational activator *</td>
<td>A6ZTZ3_YEAS7 (+5)</td>
<td>297</td>
</tr>
<tr>
<td>Stress-seventy subfamily A protein</td>
<td>A6ZR59_YEAS7 (+6)</td>
<td>70</td>
</tr>
<tr>
<td>Alpha mannosidase</td>
<td>A6ZU29_YEAS7 (+1)</td>
<td>125</td>
</tr>
<tr>
<td>Ubi4p</td>
<td>E7Q6N3_YEASB</td>
<td>28</td>
</tr>
<tr>
<td>Stress-seventy subfamily B protein *</td>
<td>A6ZRN1_YEAS7 (+15)</td>
<td>67</td>
</tr>
<tr>
<td>Identified Protein</td>
<td>Accession number</td>
<td>Molecular weight (kDa)</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Elongation factor 1-alpha</td>
<td>A6ZL63_YEAS7 (+11)</td>
<td>50</td>
</tr>
<tr>
<td>Adh1p</td>
<td>E7KTZ4_YEASL (+10)</td>
<td>37</td>
</tr>
<tr>
<td>Glyceraldehyde-3-phosphate dehydrogenase *</td>
<td>A6ZPX6_YEAS7 (+11)</td>
<td>36</td>
</tr>
<tr>
<td>Chaperonin</td>
<td>A6ZMP7_YEAS7 (+14)</td>
<td>81</td>
</tr>
<tr>
<td>Meiotic sister-chromatid recombination-related protein</td>
<td>A6ZYG2_YEAS7 (+2)</td>
<td>81</td>
</tr>
<tr>
<td>Nucleolar protein</td>
<td>B3LJX4_YEAS1 (+11)</td>
<td>57</td>
</tr>
<tr>
<td>Proteinase B</td>
<td>A6ZQP1_YEAS7 (+10)</td>
<td>70</td>
</tr>
<tr>
<td>Vacuolar protein sorting-associated protein 13 *</td>
<td>VPS13_YEAST</td>
<td>358</td>
</tr>
<tr>
<td>Ubiquitin protease cofactor</td>
<td>A6ZSC5_YEAS7 (+5)</td>
<td>58</td>
</tr>
</tbody>
</table>

**DDM-purification**

<table>
<thead>
<tr>
<th>Identified Protein</th>
<th>Accession number</th>
<th>Molecular weight (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cystic fibrosis transmembrane conductance regulator*</td>
<td>IPI00575922</td>
<td>168</td>
</tr>
<tr>
<td>Ribosomal protein L3</td>
<td>A6ZNR5_YEAS7 (+10)</td>
<td>44</td>
</tr>
<tr>
<td>Acetyl CoA carboxylase*</td>
<td>A6ZS90_YEAS7 (+5)</td>
<td>250</td>
</tr>
<tr>
<td>Heat shock protein 70 *</td>
<td>HSP71_YEAST</td>
<td>70</td>
</tr>
<tr>
<td>Translation elongation factor 2 (EF-2)</td>
<td>A7A1D2_YEAS7 (+11)</td>
<td>116</td>
</tr>
<tr>
<td>Alcohol dehydrogenase</td>
<td>A6ZNC5_YEAS7 (+10)</td>
<td>37</td>
</tr>
<tr>
<td>H(+)-ATPase *</td>
<td>A6ZUY5_YEAS7 (+7)</td>
<td>100</td>
</tr>
<tr>
<td>Glyceraldehyde-3-phosphate dehydrogenase *</td>
<td>A6ZPX6_YEAS7 (+11)</td>
<td>36</td>
</tr>
<tr>
<td>ADP/ATP carrier</td>
<td>A6ZKR7_YEAS7 (+3)</td>
<td>34</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession Number</td>
<td>Position</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>-----------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Eukaryotic initiation factor 4A</td>
<td>B3LRD4_YEAS1 (+9)</td>
<td>45</td>
</tr>
<tr>
<td>Translational activator *</td>
<td>A6ZTZ3_YEAS7 (+5)</td>
<td>297</td>
</tr>
<tr>
<td>Pyruvate kinase</td>
<td>A7A0C8_YEAS7 (+10)</td>
<td>55</td>
</tr>
<tr>
<td>Ribosomal protein L10</td>
<td>A7A0Y0_YEAS7 (+10)</td>
<td>25</td>
</tr>
<tr>
<td>ATP synthase subunit alpha</td>
<td>A6ZKK5_YEAS7 (+8)</td>
<td>59</td>
</tr>
<tr>
<td>Porin</td>
<td>A6ZS23_YEAS7 (+11)</td>
<td>30</td>
</tr>
<tr>
<td>Vacuolar protein sorting-associated protein 13 *</td>
<td>VPS13_YEAST</td>
<td>358</td>
</tr>
<tr>
<td>Rps1bp</td>
<td>C8ZEE2_YEAS8 (+10)</td>
<td>29</td>
</tr>
<tr>
<td>Obg-like ATPase</td>
<td>A6ZKX1_YEAS7 (+11)</td>
<td>44</td>
</tr>
<tr>
<td>Stress-seventy subfamily B protein *</td>
<td>A6ZRN1_YEAS7 (+15)</td>
<td>67</td>
</tr>
<tr>
<td>Phosphatidylethanolamine N-methyltransferase G</td>
<td>CHO2_YEAST</td>
<td>101</td>
</tr>
<tr>
<td>Ribosomal protein S3</td>
<td>A6ZRQ7_YEAS7 (+9)</td>
<td>27</td>
</tr>
</tbody>
</table>
Appendix 3. Location of Cys and Trp residues in chicken-CFTR structure.

A. Alignment of chicken- and human-CFTR protein sequence. The protein sequence of human-CFTR (NP_000483.3) was compared to the chicken-CFTR sequence (NP_001099136.1) using the BLAST protein alignment tool (http://blast.ncbi.nlm.nih.gov/), shown here as Query and Sbjct, respectively. Cys residues are highlighted in yellow and Trp in green. The overall protein sequence identity is 80%.

B. Cys residues in the human-CFTR structure model. The inward-facing conformation of the human-CFTR structure model by Mormon et al. (Mornon et al., 2008) was analysed in Chimera (PD coordinates were downloaded from http://www.impmc.upmc.fr/~callebau/CFTR.html) and two views are shown rotated 180° relative to one another. Red arrows indicate surface-exposed Cys residues. Cys were deemed surface-exposed if the thiol group (shown in yellow) was clearly accessible to the surface of the structure.

C. Prediction of Cys locations in the chicken-CFTR structure. The 15 Cys in chicken-CFTR are listed by residue number. Their location and surface-exposure was predicted by comparison with equivalent Cys in both the inward- and outward-facing conformation human-CFTR structure model (Mornon et al., 2008).

D. Prediction of Trp locations in the chicken-CFTR structure. The 19 Trp in chicken-CFTR are listed by residue number. Their locations were predicted by comparison to the human-CFTR structure model, shown in part B.

A.
C.

<table>
<thead>
<tr>
<th>Cys Residue</th>
<th>Location</th>
<th>Surface exposed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Inward</td>
</tr>
<tr>
<td>77</td>
<td>N-terminus</td>
<td></td>
</tr>
<tr>
<td>129</td>
<td>TM2</td>
<td>✔</td>
</tr>
<tr>
<td>277</td>
<td>Linker between ICL2 and TM5</td>
<td></td>
</tr>
<tr>
<td>344</td>
<td>TM6</td>
<td></td>
</tr>
<tr>
<td>525</td>
<td>NBD1</td>
<td></td>
</tr>
<tr>
<td>591</td>
<td>NBD1</td>
<td>✔</td>
</tr>
<tr>
<td>593</td>
<td>NBD1</td>
<td>✔</td>
</tr>
<tr>
<td>625</td>
<td>NBD1</td>
<td>✔</td>
</tr>
<tr>
<td>836</td>
<td>R region</td>
<td>✔</td>
</tr>
<tr>
<td>869</td>
<td>Linker between R region and TM7</td>
<td>✔</td>
</tr>
<tr>
<td>1346</td>
<td>NBD2</td>
<td>✔</td>
</tr>
<tr>
<td>1357</td>
<td>NBD2</td>
<td>✔</td>
</tr>
<tr>
<td>1397</td>
<td>NBD2</td>
<td>✔</td>
</tr>
<tr>
<td>1412</td>
<td>NBD2</td>
<td>✔</td>
</tr>
<tr>
<td>1469</td>
<td>C-terminus</td>
<td>✔</td>
</tr>
<tr>
<td>Trp Residue</td>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>N terminus</td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>N terminus</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>TM1</td>
<td></td>
</tr>
<tr>
<td>203</td>
<td>TM3</td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>TM3</td>
<td></td>
</tr>
<tr>
<td>240</td>
<td>TM4</td>
<td></td>
</tr>
<tr>
<td>278</td>
<td>TM5</td>
<td></td>
</tr>
<tr>
<td>362</td>
<td>TM6</td>
<td></td>
</tr>
<tr>
<td>402</td>
<td>NBD1</td>
<td></td>
</tr>
<tr>
<td>497</td>
<td>NBD1</td>
<td></td>
</tr>
<tr>
<td>850</td>
<td>TM7</td>
<td></td>
</tr>
<tr>
<td>1077</td>
<td>TM11</td>
<td></td>
</tr>
<tr>
<td>1093</td>
<td>TM11</td>
<td></td>
</tr>
<tr>
<td>1102</td>
<td>TM12</td>
<td></td>
</tr>
<tr>
<td>1149</td>
<td>NBD2</td>
<td></td>
</tr>
<tr>
<td>1206</td>
<td>NBD2</td>
<td></td>
</tr>
<tr>
<td>1276</td>
<td>NBD2</td>
<td></td>
</tr>
<tr>
<td>1284</td>
<td>NBD2</td>
<td></td>
</tr>
<tr>
<td>1312</td>
<td>NBD2</td>
<td></td>
</tr>
<tr>
<td>1318</td>
<td>NBD2</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 4. CPM dye binding assay controls: GFP and ribosomal L3 subunit protein.
A. SDS-PAGE analysis of L3 and GFP. DDM-solubilised CFTR was bound to a nickel-affinity chromatography column and washed with an increasing gradient of 200 to 400 mM Imidazole. A fraction enriched in ribosomal large subunit L3 protein is shown by Coomassie-stained SDS-PAGE analysis as a protein band migrating at approximately 45 kDa (left gel). Purified GFP was purchased from Millipore (cat. No. 14-392) and analysed by SDS-PAGE, which appears as a protein band migrating at approx 27 kDa (right gel). B. CPM dye binding assay of ribosomal large subunit L3 protein and GFP show thermal unfolding profiles that do not overlap with that of CFTR (see Results figure 16). L3 and GFP were incubated with CPM dye from 10 to 80 °C at heat rate of 2.5 °C/min and CPM fluorescence monitored. Ribosomal L3 protein shows no increase in CPM fluorescence with increasing temperature probably because its 2 x surface-exposed cysteine residues (PDB=1FFK) are already labeled during the initial incubation with CPM at 10 °C. GFP shows an increase in CPM fluorescence above 70 °C which is likely due to the labeling of its 2 x buried cysteines (PDB=1GFL) as they become exposed during thermal denaturation.
Appendix 5. Thermal stability of CFTR, as monitored by EM.
DDM-purified CFTR was incubated for 10 min at 10 and 80 °C and analysed by NS-EM. CFTR protein is viewed as white particles on a grey background. At 10 °C (top panel), CFTR appears as compact globular shaped particles. At 80 °C (bottom panel), CFTR particles appear to be more random and linearised structures.
Appendix 6. Thermal gel analysis of GFP.

Purified GFP at 50 µg/ml was heated from 20 to 85 °C in a water bath at a heating rate of 2.5 °C/min. Aliquots were taken at 5 °C increments and quenched on ice for 10 min before incubating with SDS-load dye for 15 min at RTM. Approximately 10 ng of GFP were loaded per well on a 10 % pre-cast SDS-PAGE with no stacking gel, which is shown under GFP-fluorescence (top panel). GFP fluorescent protein band intensities were quantified from the digitalized gel image using Image J software and are plotted on the graph below relative to the band intensity at 20 °C. GFP in-gel fluorescence appears to be unaffected at temperatures below 75 °C.
Appendix 7. Cryo-EM micrograph of LPG-purified CFTR.
LPG-purified CFTR at 1 mg/ml was applied to a holey Quantifoil carbon grid and flash frozen in liquid ethane. CFTR particles were then trapped in a thin layer of vitreous ice within the holes of the grid and analysed by cryo-EM. An example micrograph is shown below. The hole containing the ice layer is to the left of the curved black line and to the right of the black line is the continuous carbon surface. CFTR protein appears as black particles against a white background. The scale bar is 100nm. CFTR particles appear to concentrate toward the edge of the hole (from right-to-left on this image), where they become too densely packed for further single particle processing. The area between the two dashed red lines represent CFTR particles that are well separated and suitable for selection by EMAN2 for single particle processing.
Appendix 8. NS-EM micrograph of DDM-purified CFTR HMW aggregates.
DDM-purified CFTR was collected from the SEC column elution fraction that represented the void volume (7 ml) and analysed by NS-EM at protein concentration of 50 µg/ml. An example micrograph is shown below, where the CFTR protein appears as white particles against a grey background that are heavily stained in large aggregates.
Appendix 9. NS-EM micrograph of DDM-purified CFTR.
DDM-purified CFTR was collected from the SEC column elution fraction shortly after the void volume (9.5 ml) and analysed by NS-EM at protein concentration of 50 µg/ml. An example micrograph is shown below, where the CFTR protein appears as white particles against a grey background. CFTR appears as small particles about 10-20 nm in length, as well as larger aggregates about 20-30 nm in size.
Appendix 10. Vector displacement plot of CFTR 2D crystal.
2D crystal of CFTR was processed for lattice unbending using the 2dx software. The plot below shows the vector displacement after the second round of unbending. The length of the lines represent the degree of unbending required at that point to flatten the crystal relative to the reference area at the centre of the crystal.