

**Three-Dimensional Transmission and Scanning Electron Microscopy of  
Molecular and Cellular Structures**

by

Brian James Caffrey

B.Sc. (Hons), University College Dublin, 2015

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Biochemistry and Molecular Biology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

September 2020

©Brian James Caffrey, 2020

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Three-Dimensional Transmission and Scanning Electron Microscopy of Molecular and Cellular Structures

---

submitted by Brian James Caffrey in partial fulfilment of the requirements for

the degree  
of

Doctor of Philosophy

---

in

Biochemistry and Molecular Biology

---

**Examining Committee:**

Sriram Subramaniam, Professor, Biochemistry and Molecular Biology, UBC

Supervisor

Anthony Phillips, Professor, Psychiatry, UBC

Supervisory Committee Member

Natalie Strynadka, Professor, Biochemistry and Molecular Biology, UBC

Supervisory Committee Member

Calvin Yip, Associate Professor, Biochemistry and Molecular Biology, UBC

University Examiner

Edwin D W Moore, Professor, Cellular & Physiological Science, UBC

University Examiner

Reinhart Reithmeier, Professor, Biochemistry, University of Toronto

External Examiner

## **Abstract**

Unravelling the complex spatial arrangement of networks and interfaces between proteins, cells and tissues is fundamental to our understanding of healthy and pathological processes. Therefore, a three-dimensional ultrastructural understanding of this arrangement is key to developing modern diagnostic and therapeutic applications in disease. Herein, we discuss our application of transmission and scanning electron microscopy techniques such as Cryo-Electron Microscopy (Cryo-EM), Focused Ion Beam-Scanning Electron Microscopy (FIB-SEM) alone and in combination with light microscopy, to develop an understanding of biological processes at the nanoscale.

Cryo-EM is a now established technique for the elucidation of protein structures at atomic resolutions. A key advantage of cryo-EM methods over other structural techniques is its ability to generate 3D maps of proteins in solution, in their native environments. Here, we apply cryo-EM techniques to mutant structures of the protein p97 towards a molecular understanding of the mechanisms of p97-related disease. We solved eight nucleotide-bound cryo-EM structures of four full-length hexameric mutants (R155H, A232E, D592N and E470D) implicated in p97-related neurodegenerative disease.

FIB-SEM is an imaging approach that enables analysis of the 3D architecture of cells and tissues at resolutions that are 1–2 orders of magnitude higher than that possible with light microscopy. The slow speeds of data collection and manual segmentation are two critical problems that limit the more extensive use of FIB-SEM technology. Here, we developed a semi-automated segmentation method that enables rapid, large-scale acquisition of data from tissue specimens. We demonstrate the feasibility of these methods through the 3D analysis of human muscle tissue by showing that our process results in an improvement in speed of up to three orders of magnitude as compared to manual approaches for data segmentation.

Correlated Light Electron Microscopy (CLEM) combines the high-resolution isotropic resolving power of electron microscopies with the discriminatory power of fluorescent light microscopy. Here, we present a CLEM study of molecular interactions between nanoparticles and cells, enabling us to describe the path of a model nanoparticle through the cell and identify key stages of nanoparticle uptake in greater detail than either methodology used independently.

## **Lay Summary**

This thesis is focused on developing tools towards answering key biological questions in human health and disease across several size scales in 3-dimensions, from the cell and its constituent parts, all the way down to individual proteins and the development of drugs to treat the diseases which arise in these systems. Towards this aim we have used electron microscopy, in a variety of different forms, to further understand the biological processes involved in both normal and diseased systems. We have also developed support systems which have increased the throughput and accuracy of both our sample preparation and the yielded data.

## **Preface**

This dissertation is an original intellectual product of the author, B.Caffrey.

Chapter 2. This work was conducted at the Centre for Brain Health at the University of British Columbia, Vancouver Campus in Prof. Sriram Subramaniam's Lab (Department of Biochemistry, University of British Columbia [UBC], British Columbia [BC] Canada). Oda Helene Schiøtz (UBC, Canada) and I performed grid preparation and vitrification. Alison Berezuk (UBC, BC, Canada), the LSI TEM facility (UBC, BC, Canada) and the Pacific Northwest Centre for Cryo-EM (PNCC, Oregon, USA) performed cryo-EM imaging. Xing Zhu (UBC, BC, Canada) performed particle picking, 2D and 3D classification of cryo-EM data. Sriram Subramaniam and I conceived the experiments and interpreted the structural data. I performed all site-directed mutagenesis, protein production/characterisation, biochemical analyses, 3D structural refinement and structural studies on the mutant structures.

Chapter 3. This work was performed in collaboration with Dr. Luigi Ferrucci, (National Institute of Aging [NIA], National Institutes of Health [NIH], Maryland, USA) and Prof. Sriram Subramaniam. A version of this material has been published [Caffrey, B.J., Maltsev, A.V., Gonzalez-Freire, M., Hartnell, L.M., Ferrucci, L. and Subramaniam S. "Automated 3D Segmentation of Human Skeletal Muscle Using Focused Ion Beam-Scanning Electron Microscopic Images." *Journal of Structural Biology* 2019;207(01):1-11]. Alexander Maltsev (NIA, NIH, Maryland, USA) generated machine learning capabilities which allowed the rapid generation of segmented datasets. Marta Gonzalez-Freire (NIA, NIH, Maryland, USA) selected appropriate candidates and provided muscle sample biopsies for the study. Focused Ion Beam – Scanning Electron Microscopic imaging was performed by Lisa Hartnell (National Cancer Institute, NIH, Maryland, USA) and myself. I performed all machine learning training, data processing/interpretation and method generation/validation. Sriram Subramaniam, Luigi Ferrucci and I conceived the experiments and along with Alexander Maltsev wrote the previously mentioned manuscript.

Chapter 4. This work was performed in collaboration with Prof. Kenneth Dawson (Center for BioNano Interactions, UCD, Dublin, Ireland) and Prof. Sriram Subramaniam. Sriram Subramaniam, Kenneth Dawson and I formed the experimental concept and design and I performed the experimental work, data collection and analysis.

# Table of Contents

|   |             |
|---|-------------|
| <b>Abstract</b> .....   | <b>iii</b>  |
| <b>Lay Summary</b> .....  | <b>iv</b>   |
| <b>Preface</b> .....  | <b>v</b>    |
| <b>Table of Contents</b> .....  | <b>vii</b>  |
| <b>List of Tables</b> .....   | <b>x</b>    |
| <b>List of Figures</b> .....  | <b>xi</b>   |
| <b>List of Abbreviations</b> .....  | <b>xiii</b> |
| <b>Acknowledgements</b> .....   | <b>xvi</b>  |
| <b>Dedication</b> .....   | <b>xvii</b> |
| <b>1. Introduction</b> .....  | <b>1</b>    |
| 1.1. General Introduction.....  | 1           |
| 1.2. Structure Elucidation of p97 Mutants Using Cryo-Electron Microscopy.....                                       | 2           |
| 1.2.1. Techniques in Protein Structural Elucidation.....  | 3           |
| 1.2.1.1. Negative Staining.....   | 7           |
| 1.2.1.2. Cryo-Freezing.....   | 7           |
| 1.2.2. The AAA+ ATPase Family and the Role of p97 in the Cell.....  | 8           |
| 1.2.2.1. Structure and Activity of p97.....   | 11          |
| 1.2.3. The Role of p97 in Disease.....  | 17          |
| 1.3. Understanding Tissue Ultrastructure using High-Resolution Focused Ion Beam – Scanning Electron Microscopy..... | 22          |
| 1.3.1. Large Volume Acquisition and Automated Segmentation Techniques.....  | 22          |
| 1.3.2. Human Muscle.....  | 25          |
| 1.3.2.1. Subcellular Morphological Differences Between Skeletal Muscle Types.....                                   | 29          |
| 1.4. Examining Nanotherapeutic Interactions on a Cellular Scale.....  | 31          |
| 1.4.1. Correlative Light Electron Microscopy (CLEM).....  | 31          |
| 1.4.2. Confocal Light Microscopy.....   | 32          |
| 1.4.3. Conventional therapeutics.....   | 32          |
| 1.4.4. Novel Methods for Medicine.....  | 33          |
| 1.4.5. Nanoparticle-Cell interactions.....  | 34          |
| <b>2. Structure Elucidation of p97 using Cryo-Electron Microscopy</b> .....   | <b>36</b>   |
| 2.1. Chapter Introduction.....  | 36          |
| 2.2. Results and Discussion.....  | 38          |
| 2.2.1. Expression and Characterisation of p97 Mutants.....  | 38          |
| 2.2.2. Structural Elucidation of p97 Mutants using Cryo-EM.....   | 40          |
| 2.2.2.1. p97 Mutant Nucleotide Density.....   | 40          |
| 2.2.2.2. p97 <sup>R155H</sup> ADP-Bound.....  | 41          |
| 2.2.2.3. p97 <sup>R155H</sup> ATP <sub>γ</sub> S-Bound.....   | 45          |
| 2.2.2.4. p97 <sup>A232E</sup> ADP-Bound.....  | 47          |
| 2.2.2.5. p97 <sup>A232E</sup> ATP <sub>γ</sub> S-Bound.....   | 51          |
| 2.2.2.6. p97 <sup>D592N</sup> ADP-Bound.....  | 53          |
| 2.2.2.7. p97 <sup>D592N</sup> ATP <sub>γ</sub> S-Bound.....   | 55          |
| 2.2.2.8. p97 <sup>E470D</sup> Cryo-EM Maps.....   | 60          |
| 2.2.2.9. General Observations.....  | 62          |

|  |           |
|--|-----------|
| 2.3. Methodology.....  | 64        |
| 2.3.1. Preparation of Chemically Competent Bacteria.....   | 64        |
| 2.3.2. Site-Directed Mutagenesis (SDM).....  | 64        |
| 2.3.3. Plasmid Isolation.....  | 66        |
| 2.3.4. Sanger Sequencing of Mutant Plasmids.....   | 66        |
| 2.3.5. Protein Expression and Purification.....  | 67        |
| 2.3.6. ATPase Assays.....  | 69        |
| 2.3.7. Cryo-EM Sample Preparation and Imaging.....   | 69        |
| 2.3.8. Single Particle Picking and 3D reconstruction.....  | 70        |
| 2.3.9. Structure Refinement of p97 Mutants.....  | 71        |
| <b>3. Understanding Tissue Ultrastructure using High-Resolution Focused Ion Beam – Scanning Electron Microscopy Volumes.....</b>                     | <b>72</b> |
| 3.1. Chapter Introduction.....   | 72        |
| 3.2. Results and Discussion.....   | 74        |
| 3.2.1. Qualitative Analysis of Mitochondrial Distribution in Human Skeletal Muscle.....  | 74        |
| 3.2.2. Statistical Analysis of Mitochondrial Distribution in Human Skeletal Muscle.....  | 80        |
| 3.2.3. Quantitative Evaluation of Segmentation Pipeline Against Manual Standards.....  | 82        |
| 3.2.4. Evaluation of Semi-Automated Segmentation Performance Using CA1 Hippocampal Test Dataset.....   | 87        |
| 3.3. Methodology.....  | 88        |
| 3.3.1. FIB-SEM Acquisition and Image Processing.....   | 88        |
| 3.3.1.1. Candidate Selection and Muscle Biopsy.....  | 88        |
| 3.3.1.2. Fixation, Contrasting and Embedding.....  | 88        |
| 3.3.1.3. Area Selection for FIB-SEM Analysis.....  | 89        |
| 3.3.1.4. Image Processing.....   | 90        |
| 3.3.2. Automated Segmentation Methodology and Validation.....  | 91        |
| 3.3.2.1. Automated Segmentation Methodology.....   | 91        |
| 3.3.2.2. Methodology Validation.....   | 95        |
| 3.3.3. Statistical Analysis of Mitochondrial Distribution in Human Skeletal Muscle.....  | 96        |
| <b>4. Examining Nanotherapeutic Interactions on a Cellular Scale.....</b>  | <b>97</b> |
| 4.1. Chapter Introduction.....   | 97        |
| 4.2. Results and Discussion.....   | 98        |
| 4.2.1. Nanoparticle Characterisation and Stability.....  | 98        |
| 4.2.2. Characterisation of Nanoparticles and Heavy-Metal Staining in FIB-SEM images.....   | 100       |
| 4.2.3. Determination of Fixation and Permeabilization Parameters for Retention Of Ultrastructural Information from Light to Electron Microscopy..... | 101       |
| 4.2.4. Correlative Light Electron Microscopy Analysis of Nanoparticle-Cell Interactions.....   | 102       |
| 4.3. Methodology.....  | 105       |
| 4.3.1. Human Cell Culture.....   | 105       |
| 4.3.2. Transferrin Absorbed Gold Nanoparticle Preparation.....   | 105       |
| 4.3.3. Tf-Gold Addition to A549 Cells.....   | 106       |
| 4.3.4. Antibody Receptor Staining.....   | 106       |
| 4.3.5. Confocal Imaging of Tf-Gold Incubated A549.....   | 107       |
| 4.3.6. Fixing, Staining and Resin Embedding of Confocal Imaged A549 with.....  |           |

|  |     |
|--|-----|
| Tf-Gold Nanoparticles for FIB-SEM.....   | 108 |
| 4.3.7. FIB-SEM Imaging of Tf-Gold Incubated A549.....  | 109 |
| 4.3.8. Electron Dispersive X-Ray Spectroscopy.....   | 109 |
| 4.3.9. Image Processing and Segmentation.....  | 110 |
| 4.3.9.1. Confocal Image Processing.....  | 110 |
| 4.3.9.2. FIB-SEM Image Processing.....   | 110 |
| 4.3.9.3. Semi-Automated Registration, Alignment and Correlation of<br>Light/Electron Images.....                               | 110 |
| 4.3.9.4. Segmentation and 3D representation of Cellular Structures.....  | 112 |
| <b>5. Conclusion</b> .....   | 112 |
| 5.1. General Conclusion and Future Directions.....   | 112 |
| 5.2. Structure Elucidation of p97 using Cryo-Electron Microscopy.....  | 113 |
| 5.3. Understanding Tissue Ultrastructure using High-Resolution Focused Ion Beam –<br>Scanning Electron Microscopy Volumes..... | 114 |
| 5.4. Examining Nanotherapeutics on a Cellular Scale.....   | 115 |
| <b>Bibliography</b> .....  | 117 |
| <b>Appendix</b> .....  | 128 |

## List of Tables

|   |     |
|---|-----|
| <b>Table 1.1: Examples of Ubiquitin-Associated p97 Cofactors</b> .....                        | 13  |
| <b>Table 2.1: PCR Thermocycling Protocol</b> .....  | 65  |
| <b>Table 3.1: Quantitative Evaluation of Automated vs. Manual Segmentation</b> .....          | 83  |
| <b>Table 4.1: Dynamic Light Scattering Measurements of Nanoparticle Protein Complex</b> ..... | 99  |
| <b>Table A.1: Media Recipes</b> .....   | 131 |
| <b>Table A.2: <i>E. coli</i> Strains</b> .....  | 131 |
| <b>Table A.3: Primer List for Site-Directed Mutagenesis</b> .....                             | 132 |
| <b>Table A.4: EM Data Collection and Reconstruction Statistics</b> .....                      | 133 |
| <b>Table A.5: Model Refinement and Validation Statistics</b> .....                            | 133 |

## List of Figures

|  |    |
|--|----|
| <b>Figure 1.1: Schematic of Image Formation and 3D Particle Reconstruction</b> .....   | 6  |
| <b>Figure 1.2: Schematic Representation of ASCE and AAA+ ATPase Subgroups</b> .....  | 9  |
| <b>Figure 1.3: Homology Structure of the Classic Clade of AAA+ ATPases</b> .....   | 9  |
| <b>Figure 1.4: p97 Structure</b> .....   | 11 |
| <b>Figure 1.5: Analysis of ATP hydrolysis by Wild-Type and Hydrolysis Defective Mutants of p97</b> .....   | 12 |
| <b>Figure 1.6: Molecular Surface Model of Sequential ATP<math>\gamma</math>S Binding States of p97</b> .....   | 14 |
| <b>Figure 1.7: Proposed Models for Substrate Entry and Disassembly by p97</b> .....  | 15 |
| <b>Figure 1.8: Domain Architecture of p97 Co-factors Containing Ubiquitin Regulatory X (UBX) and UBX-Like (UBXL) Domains</b> .....                                 | 16 |
| <b>Figure 1.9: Proposed Co-Factor Binding and Substrate Processing Mechanism for Ubiquitin-Mediated Substrate Threading Model in Mammalian p97</b> .....           | 17 |
| <b>Figure 1.10: Space Filling Model of p97 Mutants</b> .....   | 19 |
| <b>Figure 1.11: FIB-SEM Workflow</b> .....   | 24 |
| <b>Figure 1.12: Major Components of Muscle Fibres</b> .....  | 26 |
| <b>Figure 1.13: Successive Skeletal Muscle Histochemical Staining</b> .....  | 28 |
| <b>Figure 1.14: Skeletal Muscle Ultrastructure and Their Differences Between Subtypes</b> .....  | 29 |
| <b>Figure 1.15: Examples of Different Mitochondrial Morphologies Within Skeletal Muscle Fibre</b> .....  | 30 |
| <b>Figure 1.16: CLEM Workflow</b> .....  | 31 |
| <b>Figure 1.17: Correlation of Light and Electron Images</b> .....   | 32 |
| <b>Figure 2.1: Proposed Mechanistic Model of MSP-1 Mutants</b> .....   | 37 |
| <b>Figure 2.2: Biochemical Analysis of p97</b> .....   | 38 |
| <b>Figure 2.3: Nucleotide Binding Density of p97 Mutants</b> .....   | 40 |
| <b>Figure 2.4: Analysis of Cryo-EM Density Map of ADP-Bound p97<sup>R155H</sup></b> .....  | 41 |
| <b>Figure 2.5: Cryo-EM Structure at 3.0Å Resolution of ADP-Bound p97<sup>R155H</sup> Mutant; Intra-Protomer Effects of R155H Mutant</b> .....                      | 43 |
| <b>Figure 2.6: Comparison Between p97<sup>R155H</sup> Structures in This Thesis and Previously Published Crystallographic Structures</b> .....                     | 44 |
| <b>Figure 2.7: Analysis of Cryo-EM Density Map at 2.5Å Resolution of ATP<math>\gamma</math>S-Bound p97<sup>R155H</sup></b> .....                                   | 45 |
| <b>Figure 2.8: Analysis of Cryo-EM Density Map of ADP-Bound p97<sup>A232E</sup></b> .....  | 47 |
| <b>Figure 2.9: Sequence Alignment of p97 Homologs</b> .....  | 48 |
| <b>Figure 2.10: Cryo-EM Structure at 3.9Å Resolution of ADP-Bound p97<sup>A232E</sup> Mutant; Inter-Protomer Effects of A232E Mutant</b> .....                     | 49 |
| <b>Figure 2.11: Analysis of Cryo-EM Density Map of ATP<math>\gamma</math>S-Bound p97<sup>A232E</sup></b> .....   | 51 |
| <b>Figure 2.12: Cryo-EM Structure at 2.6Å Resolution of ATP<math>\gamma</math>S-Bound p97<sup>A232E</sup> Mutant; Inter-Protomer Effects of A232E Mutant</b> ..... | 52 |
| <b>Figure 2.13: Analysis of Cryo-EM Density Map of ADP-Bound p97<sup>D592N</sup></b> .....   | 53 |
| <b>Figure 2.14: Analysis of Cryo-EM Density Map of ATP<math>\gamma</math>S-Bound p97<sup>D592N</sup></b> .....   | 55 |
| <b>Figure 2.15: Cryo-EM Structure at 2.8Å of ADP-Bound p97<sup>A232E</sup> Mutant; Determination of D1 Nucleotide State</b> .....                                  | 57 |

|   |     |
|---|-----|
| <b>Figure 2.16: Comparison Between Key Residues in p97<sup>D592N</sup> and ATP<math>\gamma</math>S-bound p97<sup>WT</sup> D1 Nucleotide Binding Site</b> .....                        | 59  |
| <b>Figure 2.17: Analysis of Cryo-EM density Map of Nucleotide Bound p97<sup>E470D</sup></b> .....   | 60  |
| <b>Figure 2.18: RMSD Analysis of p97<sup>E470D</sup> Mutants</b> .....  | 61  |
| <b>Figure 3.1: A Typical Example of a Muscle Fibre Acquired With FIB-SEM at a Voxel Size of 15 nm<sup>3</sup></b> .....   | 74  |
| <b>Figure 3.2: Morphological Classification</b> .....   | 76  |
| <b>Figure 3.3: Segmentation Pipeline</b> .....  | 79  |
| <b>Figure 3.4: Boxplot graphical overview of Morphological Analysis of Human Skeletal Muscle Mitochondria Distribution From Two Sub-Populations of Data (Type A vs. Type B)</b> ..... | 81  |
| <b>Figure 3.5: Boxplot of Quantitative Evaluation</b> .....   | 85  |
| <b>Figure 3.6: Model-to-Model Distance Measurement</b> .....  | 86  |
| <b>Figure 3.7: Other Applications of This Software</b> .....  | 87  |
| <b>Figure 3.8: Example of Weka Software Manual Classification in ImageJ</b> .....   | 92  |
| <b>Figure 3.9: Graphical Representation of Key Steps in Segmentation Pipeline</b> .....   | 94  |
| <b>Figure 4.1: DLS Analysis of Transferrin Bound Nanoparticles</b> .....  | 99  |
| <b>Figure 4.2: Energy-Dispersive X-Ray Spectroscopy (EDS) Analysis of Resin Block</b> .....   | 100 |
| <b>Figure 4.3: Comparison of Ultrastructural Features Before (A) and After (B) Optimisation</b> .....   | 101 |
| <b>Figure 4.4: A-G: Segmented Images of Different Stages of Nanoparticle Uptake</b> .....   | 102 |
| <b>Figure A.1: Cryo-EM Structure Validation</b> .....   | 134 |

## List of Abbreviations

µg: microgram  
µl: microlitre  
µm: micrometre/micron  
Å: Angstrom ( $10^{-10}$ m)  
AAA+: ATPases Associated with diverse cellular Activities  
ACC: Accuracy  
ADP: Adenosine DiPhosphate  
ALS: Amyotrophic Lateral Sclerosis  
AR: Aspect Ratio  
ASCE: Additional Strand Conserved E  
ATCC: American Type Culture Collection  
ATP: Adenosine TriPhosphate  
ATP $\gamma$ S: Adenosine 5'-[ $\gamma$ -thio]-TriPhosphate  
Atx3: Ataxin-3  
AVD: Absolute Volume Difference  
Avg.: Average  
BLSA: Baltimore Longitudinal Study and Aging  
BME:  $\beta$ -MercaptoEthanol  
BSA: Bovine Serum Albumin  
CAR-T: Chimeric Antigen-T  
CAV-1: Caveolin-1  
CCD: Charge-Coupled Detector  
CET: Cryo-Electron Tomography  
CFCM: Cross Fibre Connection Mitochondria  
CLEM: Correlative Light Electron Microscopy  
cm: centimetre  
cMEM: Complete Minimum Essential Media  
Co-IP: Co-Immunoprecipitation  
DAPI: 4',6-DiAmidino-2-PhenylIndol  
DDSA: DoDecenylSuccinic Anhydride  
DLS: Dynamic Light Scattering  
DMP-30: Tris-(Dimethylaminomethyl) Phenol  
DPBS: Dulbecco's Phosphate Buffered Saline  
DSC: Dice Similarity Coefficient  
EDS: Energy-Dispersive X-Ray Spectroscopy  
EDTA: EthyleneDiamineTetraacetic Acid  
ER: Endoplasmic Reticulum  
ERAD: Endoplasmic Reticulum-Associated Degradation  
ESB: Energy Selective Back-scattered electron detector  
FBS: Foetal Bovine Serum  
FIB-SEM: Focused Ion Beam-Scanning Electron Microscope  
FLIM: Fluorescence Lifetime Imaging  
FPM: Fibre Parallel Mitochondria

FRET: Förster Resonance Energy Transfer  
FSC: Fourier Shell Correlation  
FTD: FrontoTemporal Dementia  
FWHM: Full Width at Half Maximum  
GA: GlutarAldehyde  
GESTALT: Genetic and Epigenetic Signatures of Translational Aging Laboratory Testing  
GIS: Gas Injection System  
hr: Hour  
IBM: I- Band Mitochondria  
IBM: Inclusion Body Myopathy  
IPTG: Isopropyl  $\beta$ -D-1-ThioGalactopyranoside  
Kan: Kanamycin  
kV: kiloVolts  
L: Litre  
LB: Luria Broth  
LM: Light Microscopy  
M: Moles L<sup>-1</sup>  
MBS: Main Beam Splitter  
min: minute  
ML: Machine Learning  
ml: millilitre  
MSD: Mean Surface Difference  
MSP1: MultiSystem Proteinopathies-1  
MVB: Multi-vesicular body  
MWCO: Molecular Weight Cut-Off  
nm: nanometre  
NMA: Methyl-5-Norbornene-2,3-dicarboxylic Anhydride  
NMR: Nuclear Magnetic Resonance  
NTD: N-Terminal Domain  
OD<sub>600</sub>: Optical Density at 600nm  
ORF: Open Reading Frame  
OsO<sub>4</sub>: Osmium Tetroxide  
PBS: Phosphate Buffered Saline  
PDB: Paget's Disease of the Bone  
PDI: PolyDispersity Index  
Pen-Stryp: Penicillin-Streptomycin  
PFA: ParaFormAldehyde  
PMT: Photomultiplier Tube  
PUB: PNGase/UBA or UBX containing proteins motif  
PVDF: Polyvinylidene fluoride  
PVM: Para Vascular Mitochondria  
RMSD: Root Mean Square Deviation  
ROI: Region Of Interest  
rpm: rotations per minute  
SAV: Surface Area to Volume

SDH: Succinate DeHydrogenase  
SDM: Site-Directed Mutagenesis  
SDS-PAGE: Sodium DodecylSulfate-PolyAcrylamide Gel Electrophoresis  
Sec: Second  
SENS: Sensitivity  
SHP: Suppressor of High-copy PP1-containing motif  
SNR: Signal-to-Noise Ratio  
SOC: Super Optimal broth with Catabolite repression  
SPEC: Specificity  
SPR: Surface Plasmon Resonance  
ssTEM: Serial Section-TEM  
STORM: STochastic Optical Reconstruction Microscopy  
TEM: Transmission Electron Microscopy  
TERA: Transitional Endoplasmic Reticulum ATPase  
Tf-AuNp: Transferrin bound gold nanoparticle  
TfR: Transferrin Receptor  
UA: Uranyl Acetate  
UBL: UBx-Like  
UBX: Ubiquitin Regulatory X  
UN: Ufd1-Npl4  
UPR: Unfolded Protein Response  
UPS: Ubiquitin Proteasome System  
VCP: Valosin-Containing Protein  
VIM: VCP Interacting Motif  
WT: Wild-Type

## **Acknowledgements**

I thank members of Subramaniam, Ferrucci and Dawson laboratories for helpful discussions and scientific resources. I particularly want to thank Alison Berezuk, Oda Helene Schiøtz and Xing Zhu for working with me on the mutant p97 project and Lisa Hartnell for teaching me how to use the FIB-SEM and prepare samples for imaging. I thank my wife, Jessica De Andrade for help validating and proof reading my work. I thank the Wellcome Trust (England), University College Dublin (Ireland), the University of British Columbia (Canada), the Intramural Research Programs of the National Cancer Institute (USA) and the National Institute of Aging (USA), for funding my PhD and for providing resources and equipment to conduct my research. I also owe particular thanks to my PhD supervisor Sriram Subramaniam for his constant guidance, support and focus throughout my PhD.

*To Jessica, my wonderful wife, whose unwavering support and care made it possible to complete this work.*

*I also want to offer thanks to my parents, James and Susan Caffrey, who instilled in me a passion for discovery from an early age.*

*They all have given me crucial support throughout my academic and personal life.*

*Thank you.*

# 1. Introduction

## 1.1. General Introduction

Anomalous individual protein interactions often serve as the nucleation point for most genetic, autoimmune and oncogenic diseases. Over time these aberrant interactions accumulate, leading to the presentation of symptoms in an individual, often only after years of irreparable and silent damage has occurred. Timely, rational and effective therapeutic interventions are necessary to slow this transition from a healthy to diseased state. These treatments require in-depth knowledge of disease mechanisms on several levels, from protein, cells and tissue to patient, in order to rapidly detect and treat these diseases early.

Often disease originates from a defect in the molecular machinery of a protein. p97 is such a protein and is critical to the ubiquitin-dependent degradation pathway in cells, implicated in a late-onset and rather severe combination of neuro-, bone- and muscular degenerative illnesses. Furthermore, exploiting the role of p97 in protein degradation has shown promise as an anti-cancer therapeutic target. Therefore, it is necessary to study the structures of disease-related mutants of p97 at an atomic resolution to reveal the fundamental mechanisms of p97-related diseases.

These errant mutant protein interactions within the cell cause tissues to degenerate and lead to the earliest symptoms associated with disease. A sub-micron ultrastructural understanding of these tissues in the early stages of degradation is key to early detection and treatment. Characterising and defining tissue ultrastructure can be difficult, as the significant variance both within and between individuals, leads to a loosely defined structural map at the nm to  $\mu\text{m}$  level. To account for this, it is necessary to generate a picture of tissue distribution at nanometre resolutions across large volumes, and to quantify this variability and detect the first small but significant shifts towards disease.

Finally, to generate more effective therapies and to overcome issues such as target specificity and biodistribution that have hounded small molecule drugs, it is necessary to produce more complex, multi-faceted scaffolds for accurate delivery of therapeutics to their intended target. As interactions between a therapeutic and target change, so too must our understanding of the underpinning mechanisms of action of these novel therapies. Nanotherapeutics have generated considerable research interest over the past decade as a potential means for the delivery of therapies to the site of action. However, due to the complex three-dimensional nature of nanoparticle preparations, their interactions with cells are still poorly understood and require novel approaches to map these interactions and fates within cells to inform the development of new nanoscale therapies for patients.

During the course of my PhD, I have had the opportunity to work across these 3 structural levels, from protein structure using cryo-Electron Microscopy to elucidate mechanisms of disease related to p97, to mitochondrial architecture in human skeletal muscle tissue using Focused Ion Beam – Scanning Electron Microscopy to derive quantitative methods for the ultrastructural study of ageing and to the use of light and electron microscopies to study the effects of nanoparticle uptake on cells and their fates, towards developing multiple microscopy and biochemical tools to deepen our understanding and treatment of disease.

## 1.2. Structural Elucidation of p97 Mutants Using Cryo-Electron Microscopy

The objective of this section is the structural elucidation of mutants implicated in a variety of diseases that are characterised by heterozygous mutations in the p97 gene.

As presented in chapter 2, we generated atomic resolution maps of the conformational landscape of a series of mutants through the use of cryogenic Electron Microscopy (cryo-EM) techniques. In combination with biochemical assays, we developed a model for the structural mechanism of p97-related disease.

### 1.2.1. Techniques in Protein Structural Elucidation.

The atomic resolution elucidation of protein structure provides invaluable information on the role and function of a protein. It also allows for the determination of possible regions of interest inside the protein for drug targeting and can be used to computationally model protein-drug interactions. Until recently, X-ray crystallography was by far the most widely used technique for atomic resolution determination of protein structure. Once a crystal of the protein is formed, the acquisition and determination of protein structure is relatively straight forward and has produced some of the first atomic scale protein structures ever described.

However, the production of protein crystals is often a tedious and time-consuming exercise, mired in difficulties with often complex unpredictable interactions such as buffer composition, oligomerization and twinning artifacts which can hamper efforts to produce and interpret crystallographic patterns and the resulting product may be marred by crystallographic artifacts. *In vivo*, proteins are often a mixture of several different conformations and it is common to find that the most populated protein structure in X-ray crystallography is not necessarily the most functionally active member of the conformational landscape and due to the nature of crystallization the structure that is less soluble is most likely to be crystallized and selected for structure determination. The nature of crystallisation makes it extremely sensitive to buffer composition that often makes it difficult to crystallize protein-ligand complexes and examine ligand docking outside of *in silico* modelling. Recently, more dynamic and close-to-native strategies such as Nuclear Magnetic Resonance (NMR) and cryo-EM are moving into mainstream scientific application.

NMR can generate 3D structures of macromolecular complexes to sub-nanometer resolutions and has the unique advantage of being able to measure proteins directly in solution providing unique information about the dynamics and intermolecular interactions of proteins in solution. However, NMR spectra of large biomolecules can become very complicated and difficult to interpret, limiting its application in analysing larger structures. Large amounts (>2-10mg) of pure samples are also required to achieve an appropriate signal-to-noise ratio (SNR) for high resolution applications, which can limit interpretation of difficult to express or purify proteins.

In the cryo-EM field, improvements in microscope hardware, sample preparation and imaging software have led to a boom in near-atomic resolution structures over the last couple of years, making cryo-EM methods an attractive alternative to conventional methods for structure determination. This advance was due, in part, to new automated approaches and software for the processing and averaging of EM images as well as the advent of direct electron detectors which provided far superior signal relative to the previously ubiquitous charge-coupled detectors (CCD). CCD's required that incident electrons be converted into photons before detection, this scintillation of electrons led to multiple scattering events creating a cloud of photons per pixel generating often severe attenuation of signal at higher resolutions.

The advantage of cryo-EM methods is the ability to resolve protein and other more dynamic macromolecular structures into multiple discrete conformations in soluble states (Nakane et al., 2018) in physiologically relevant conditions. This method negates the need for potential denaturants such as heavy metal-based contrast agents or crystallizations, required for conventional transmission electron microscopy or X-ray crystallographic analysis.

Cryo-EM includes 3 principal modes for the determination of protein structure, each have their own niche and are often project specific:

### 1. Cryo-Electron Tomography (CET)

CET can provide high resolution structures of large macromolecular assemblies inside and connected to organisms, such as membrane associated ribosomes (Pfeffer et al., 2012), bacterial flagella, (Chen et al., 2011) or viral entry spikes (Harris et al., 2013) in their native state. This is due to the acquisition of multiple images of a single structure taken at several angles relative to the electron beam. This angular distribution allows CET to discriminate between biomolecules above and below the particles of interest, which would appear superimposed on each other in a conventional 2D single particle imaging regime. However, CET often suffers from 3D distortion of the volume along the beam axis, due to the “missing wedge” problem which arises from physical constraints on the angular rotation of the stage inside the electron microscope, e.g. sample thickness greater than  $60^\circ$  relative to the beam become too large for high resolution acquisition. This can be somewhat overcome through the

averaging of multiple tomograms or in combination with single particle EM (Wan & Briggs, 2016).

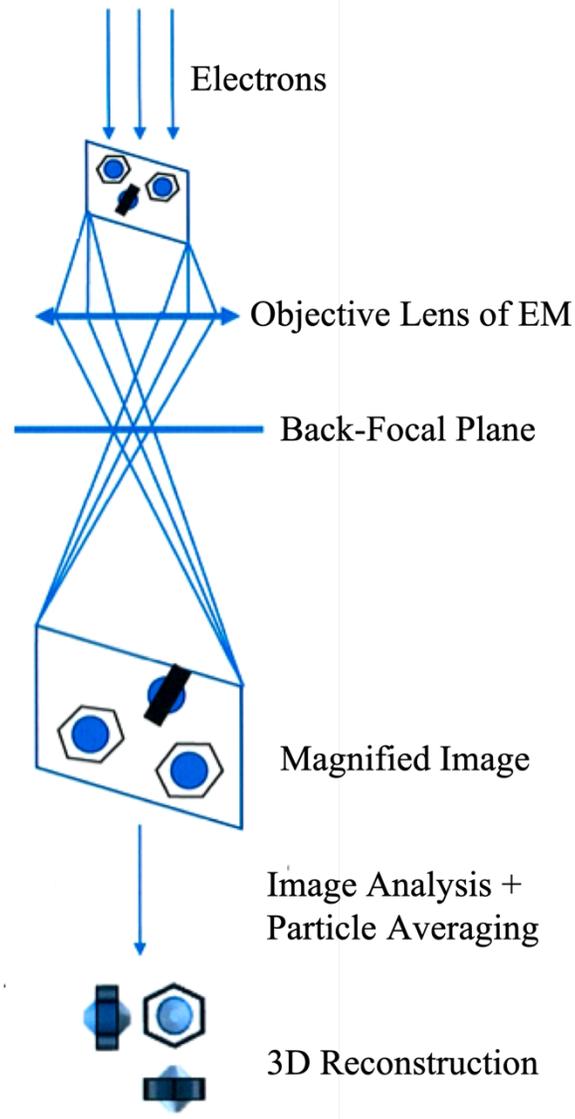
## 2. Micro-Electron Diffraction (Micro-ED)

Micro-ED is a relatively recent technique which bridges the gap between X-ray crystallography and cryo-EM. It uses protein nano-crystals, often several orders of magnitude smaller than those required for X-ray crystallographic analysis and flash freezes them in a similar manner to cryo-electron microscopies. However, instead of being imaged in a conventional manner the crystals are probed in the diffraction mode of the microscope leading to similar diffraction patterns as those found in X-ray crystallography. Using traditional crystallographic methods, analysis of the diffraction pattern produces a structure of the protein which can be resolved well below 3Å (Rodriguez et al., 2015).

## 3. Single Particle Cryo-EM

Single particle cryo-EM is best suited for monodisperse protein solutions where the protein is randomly distributed across the grid in various orientational arrangements. A single image may contain numerous proteins with a variety of orientations, which when isolated and combined is averaged to produce a single high-resolution 3D volume on the order of angstroms (Å) (Figure 1.1). However, problems may arise due to the difficulty in determining the orientation of individual 2D images of particles, compounded by the low SNR due to the use of a limited electron dose, which can complicate the structural elucidation of protein complexes. Despite these complications single particle EM often produces structures below 3.5Å without any axial distortion and the amount and purity of the sample is less of a concern as in NMR.

As p97 and its substrates are relatively small monodisperse soluble proteins, single particle EM was chosen as the best suited approach for its structural determination in this thesis.



**Figure 1.1 Schematic of Image Formation and 3D Particle Reconstruction.** In transmission electron microscopes the sample is irradiated by an electron beam, and a series of lenses enable collection of the scattered electrons to create a magnified image of the sample. This image can be recorded on photographic film or in a detector that counts the incident electrons. When the images are recorded from a specimen that contains multiple copies of the same object present in different orientations such as a protein or ordered virus particle, the set of projection images can be combined and merged to generate a 3D rendering of the imaged object much like that used to create a 3D image from a CAT scan.

### 1.2.1.1. Negative Staining

The first typical step toward the structural determination of a protein with atomic resolution is to examine the state of the protein sample under the electron microscope and use this information to improve the quality of the protein sample used for generating the cryo-EM structure. This is performed by staining the grid with heavy metal salts to produce a highly contrasted TEM image. This stain covers the grid and surrounds the sample but is excluded by the volume occupied by the sample, hence the term “negative stain”. The heavy metal ions scatter the incoming electrons and are filtered by the objective aperture, whereas the relatively unstained protein allows electrons to transmit through the sample and hit the detector. By varying the size of the aperture, the contrast and resolution of the image can be controlled.

This relatively simple screen can deliver valuable information about the sample rapidly, without concern for radiation damage and the formation of freezing artefacts which often accompany cryo-EM analysis. Preliminary 3D models can be reconstructed from this data, information mainly gathered from the sample envelope as much of the structural information in the core of the sample is lost due to a lack of contrast.

### 1.2.1.2. Cryo-Freezing

Once a suitable formulation has been tested on negative stain, a fresh sample is prepared and plunge frozen in cryogenic liquid, such as ethane cooled with liquid nitrogen. This rapid cooling has two main benefits, it prevents the formulation of ice crystals, which will damage the sample making acquisition extremely difficult; and it cryogenically “fixes” (vitrifies) the sample, without the need for chemical fixatives which often denature the protein, allowing for the preservation of high-resolution information. Once the grid is cooled it must remain at cryogenic temperatures for the remainder of the acquisition, even momentary heating can result in the destruction of the protein sample due to freeze-thaw and ice-crystal formation effects.

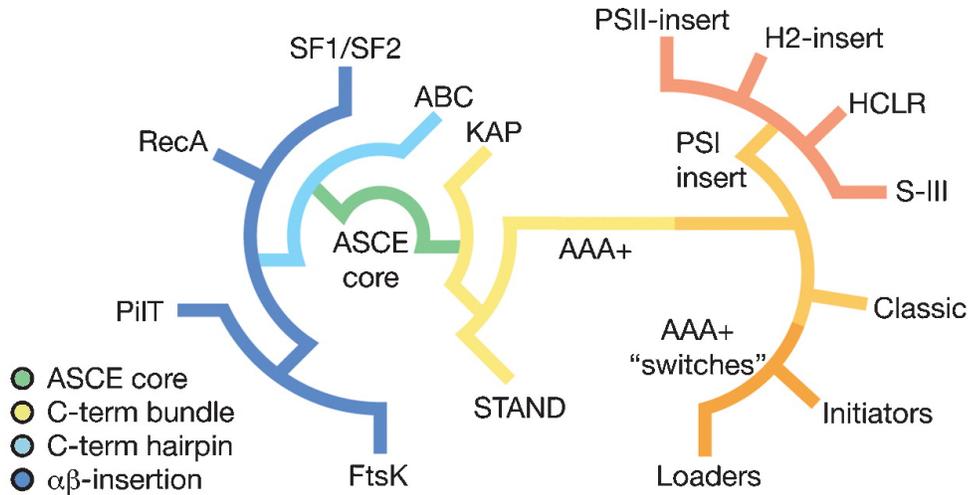
Once inside the microscope, great care must be taken to prevent radiation damage to the sample as the crucial high-resolution protein information rapidly degrades upon imaging (Grant & Grigorieff, 2015), therefore low doses of electrons are used, totalling around 10-25  $e^-/\text{\AA}^2$  per

video, in comparison with the relatively high dose of  $>75 \text{ e}^-/\text{\AA}^2$  for conventional EM. This low dose in combination with the absence of heavy metal stains or other contrasting agents leads to a low SNR image, which is one disadvantage of cryo-EM.

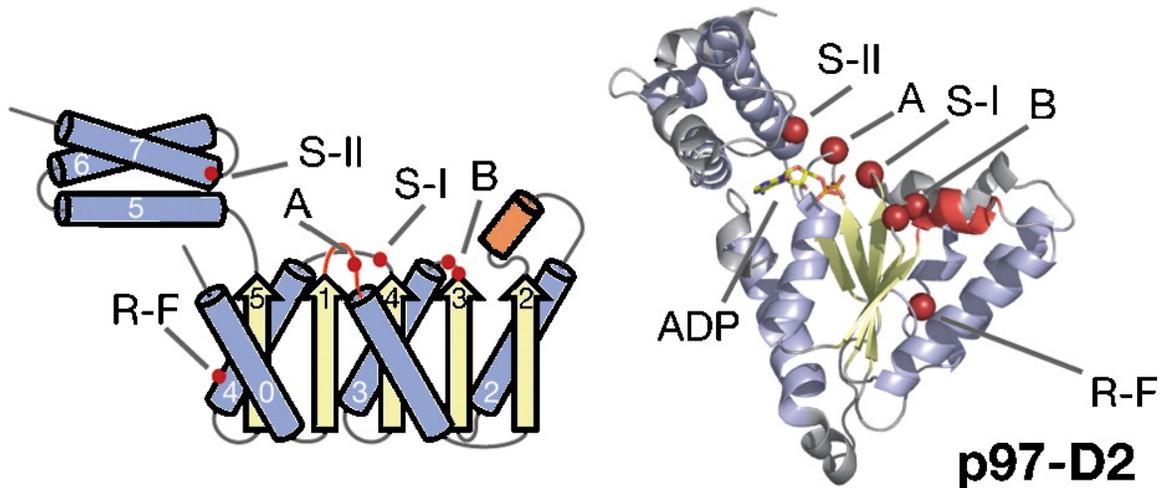
Another potential issue which may affect the contrast of the specimen is ice thickness which should be minimized wherever possible, without effecting the particle distribution, to enable efficient imaging of the sample. However, despite these limitations, judicious application of the electron beam and careful control over the vitrification conditions and grid handling can lead to extracting sufficient information to provide a high-resolution protein structure.

### 1.2.2. The AAA+ ATPase Family and the Role of p97 in the Cell

Often complex biological pathways depend on the activity of proteins that combine both regulatory and enzymatic functions within a multi-protein assembly. An expansive and essential subset of this molecular machinery is comprised of the ATPases Associated with diverse cellular Activities (AAA+) family. As the name suggests, all are related through their characteristic ATP-binding module which oligomerizes into active arrays, despite containing a functionally heterogeneous mixture of proteins. ATP-binding and subsequent hydrolysis leads to conformational changes within the assembly to drive chemo-mechanical motion translated to a specific action on the target substrate. This diverse family of proteins is further classified into clades based on the peripheral structural changes relative to this basic structure.



**Figure 1.2: Schematic Representation of ASCE and AAA+ ATPase Subgroups.** AAA+ ATPases are members of a larger superfamily of proteins which share a common Additional Strand Conserved E (ASCE) structure (Green) and are distinguished by the absence of  $\beta$ -strand additions to the ASCE core (Blue), along with the presence of a small helical bundle fused to the carboxyl-terminus of the central structure (Yellow). ©(Erzberger & Berger, 2006) Page 95. Reprinted with permission from publisher.



**Figure 1.3: Homology Structure of the Classic Clade of AAA+ ATPases.** Basic AAA+ secondary structural features are indicated in blue and yellow. The classic clade specific structural feature is indicated by a short substrate specific helix (orange) preceding the  $\alpha_2$  helix which defines the translocation channel. Conserved catalytic motifs are indicated by red dots, of particular importance are the nucleotide interacting motifs labelled (A, Walker A; B, Walker B) ©(Erzberger & Berger, 2006) Page 95. Reprinted with permission from publisher.

The specific AAA+ ATPase that we have focused on is known as the Transitional Endoplasmic Reticulum ATPase (TERA), also known as Valosin-Containing Protein (VCP) or p97 in mammals (for simplicity only p97 will be used with reference to this protein throughout the rest of this thesis). p97 is a member of the classic clade of AAA+ ATPases which is defined by a short substrate specific helix before the  $\alpha 2$  subunit (Figure 1.3).

Four main functions of p97 in the cell are as follows (Stach & Freemont, 2017):

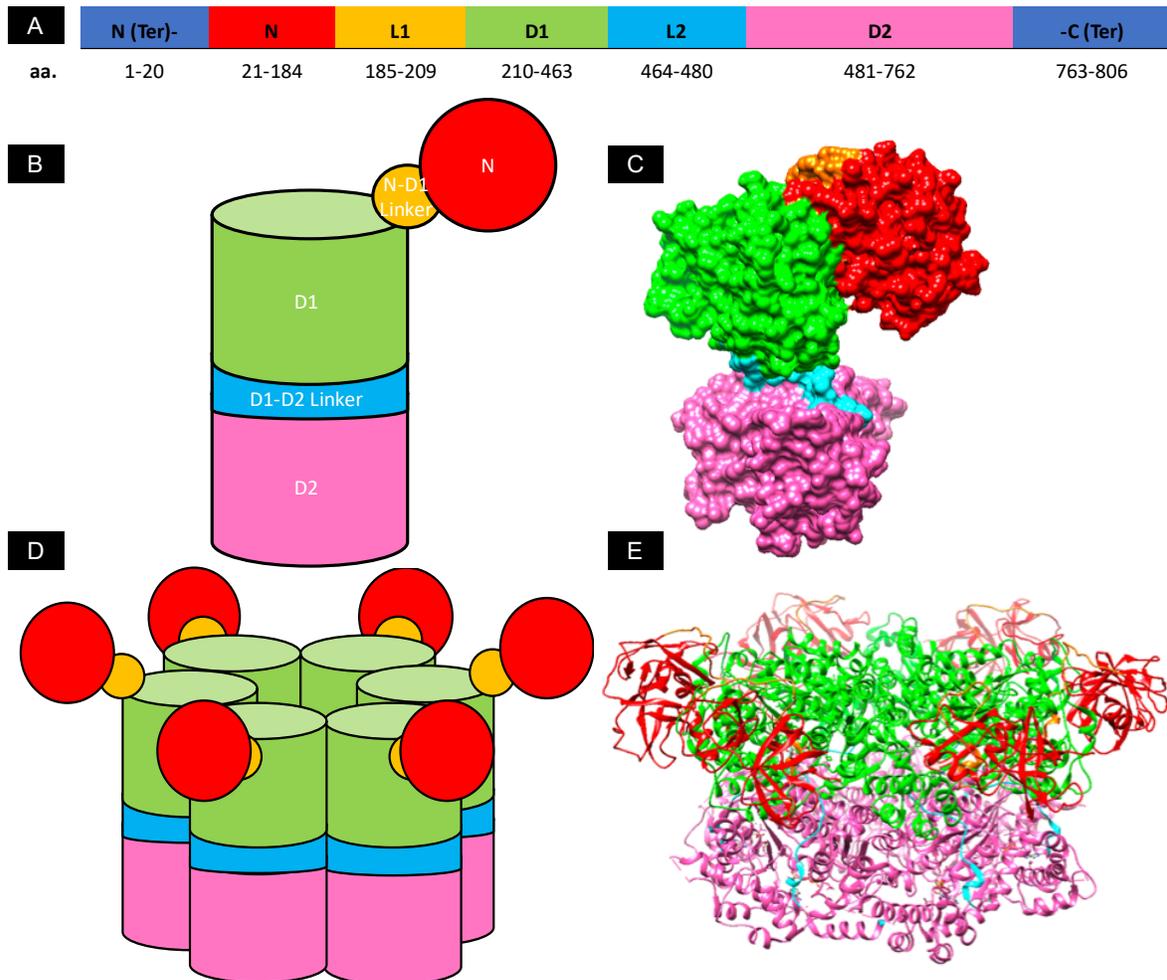
- A. Post-Mitotic Membrane Fusion
- B. Endoplasmic Reticulum-Associated Degradation (ERAD)
- C. Chromatin Related Protein Extraction
- D. NF- $\kappa$ B activation

This highlights the flexibility of this enzyme and its importance in key pathways in the cell. It also provides an insight into the types of issues that may arise when mutations in the structure of p97 lead to aberrant behaviour, which are as multi-faceted as p97.

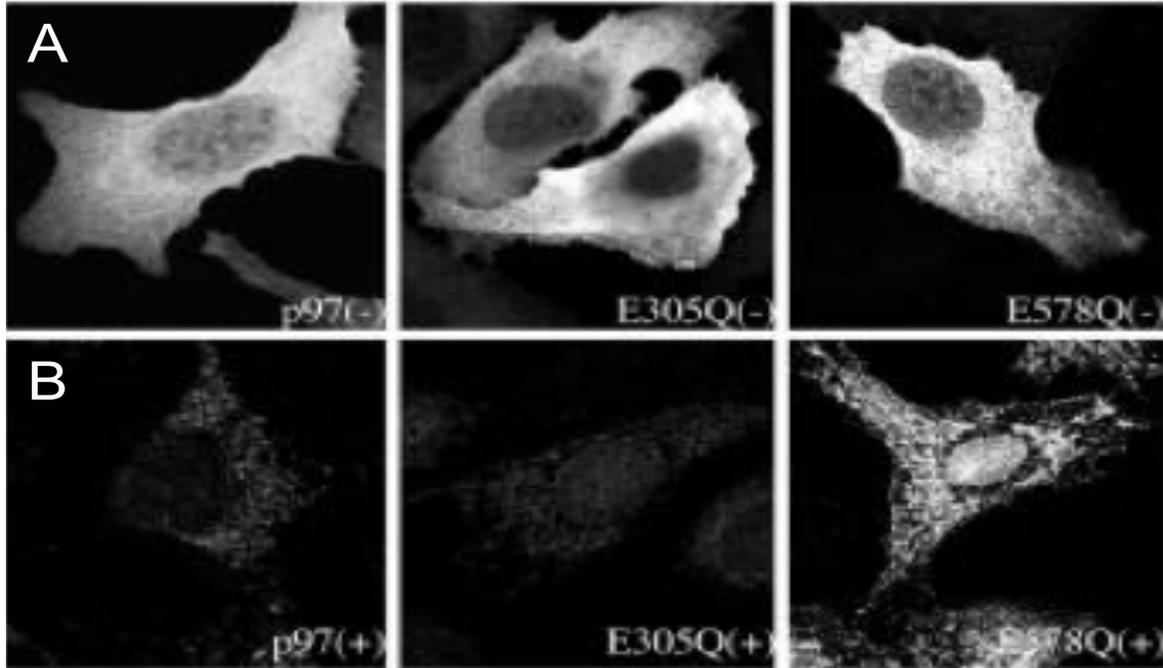
However, the focus of this thesis is primarily on the ERAD pathway which functions as a quality control system for proteins entering the secretory pathway, where both the p97 and the Ubiquitin Proteasome System (UPS) play an essential role. Emerging proteins which become misfolded in the endoplasmic reticulum (ER) must be removed to the cytosol for proteasomal degradation. This relocation of proteins from the ER to the cytosol is energetically unfavourable and requires the mechanical force generated from the ATPase activity of p97 to function.

Accumulation of misfolded proteins in the ER lead to stress and the induction of the Unfolded Protein Response (UPR), moreover, persistent ER stress leads to cell death and apoptosis. As uncontrolled protein synthesis is often a hallmark of cancer cells, they rely on pathways such as ERAD to clear misfolded proteins and prevent the activation of apoptotic pathways. Therapeutic interventions to reduce ERAD and UPR activity in cells have shown promise as a potential target in cancer chemotherapies. The central role of p97 in these processes has led to its investigation as a potential anti-cancer therapeutic (Healy et al., 2009).

### 1.2.2.1. The Structure and Activity of p97



**Figure 1.4: p97 Structure.** p97 is composed of six protomers forming a homo-hexamer with an N-terminal domain (NTD) and two planar ATPase domains stacked on top of each other, with the upper domain labelled D1 and the lower domain D2. Both ATPase domains are thought to hydrolyse ATP and therefore contribute to the enzymatic function of p97. However, studies on the effect of artificial Walker B mutants which disable the catalytic activity of the D1 (E305Q) and D2 (E578Q) domains have indicated that the D2 domain is responsible for the majority of ATPase activity in p97 (Dalal et al., 2004). **A)** Domain organisation of p97, colour coded; **B)** Schematic of p97 monomer structure. **C)** Side view of space filled model of p97 monomer; **D)** Schematic of homo-hexamer architecture; **E)** Cryo-EM structure of ADP-bound p97 (PDB: 5FTK).



**Figure 1.5: Intracellular Distribution of WT-p97 and Walker B Mutants Before (A/-) and After (B/+) Permeabilization with Saponin.** *In-vitro* experiments with the catalytically inactive (Walker B) mutants indicated that mutations in the D2 region result in trapping of the substrate. This is illustrated by the selective retention of p97<sup>E578Q</sup> on reticular membranes after saponin permeabilization indicating that the dominant ATPase activity of the D2 domain plays a central role in p97 substrate processing. ©(Dalal et al., 2004) Page 640. Reprinted with permission from publisher.

More recently, structural studies were performed using cryo-EM, where the mechanical movements within p97 and between the distinct domains were characterised to atomic resolutions (Banerjee et al., 2016). This analysis demonstrated 3 distinct conformational states populated by addition of nucleotide (Figure 1.6 A-C).

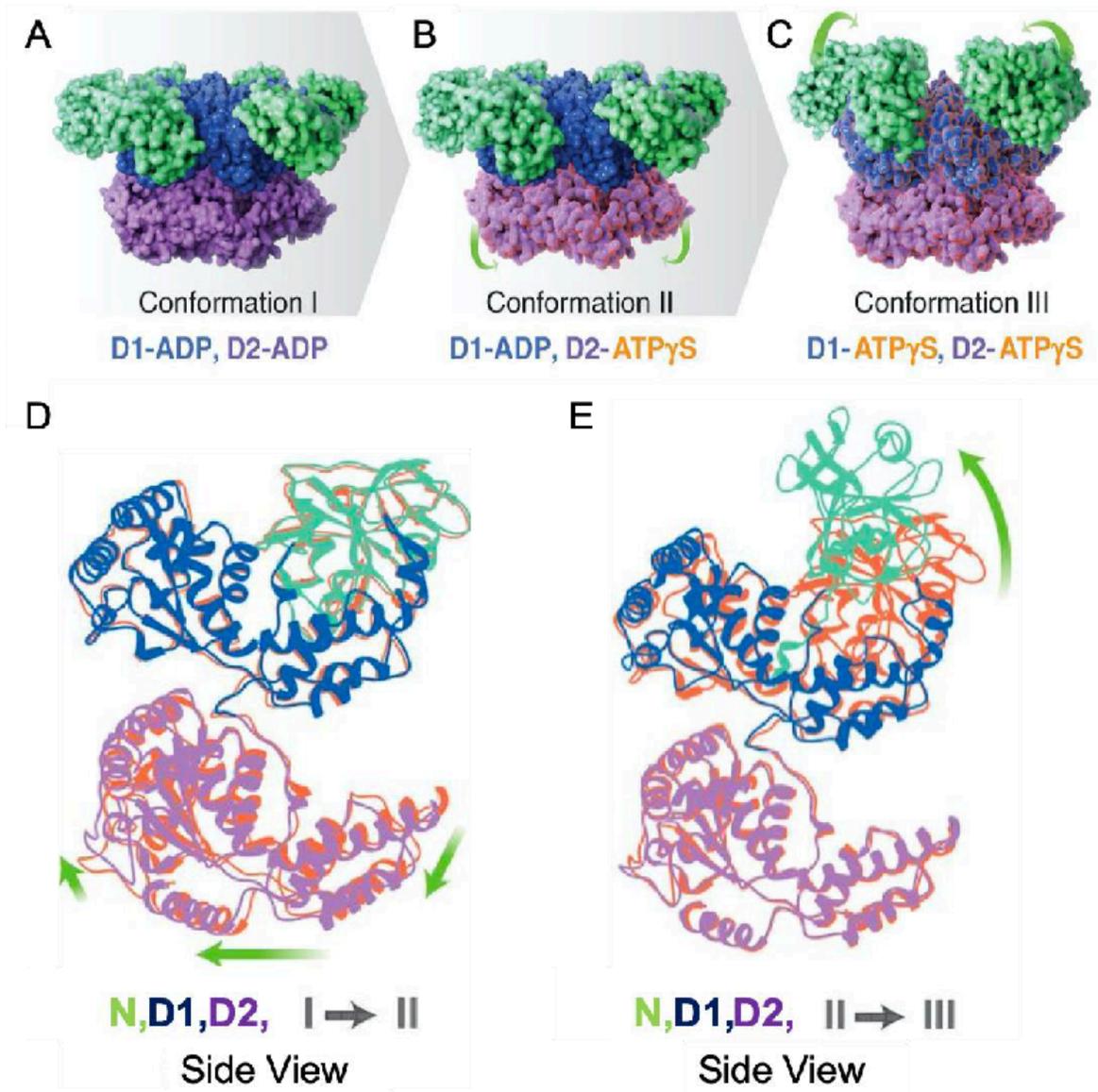
The first “down” conformation is occupied by ADP in both the D1 and D2 domains, this is considered to be the natural state of p97 found immediately after ATP hydrolysis. The second conformation identified was p97 bound to ADP and ATP $\gamma$ S in the D1 and D2, respectively. This caused a rotational twist in the D2 domain (Figure 1.6 D) with relatively little movement in the D1 and N domains. The third conformation identified was p97 bound to ATP $\gamma$ S in both D1 and D2 domains, this led to a very significant movement in the D1 and N domain resulting in an “Up” conformation (Figure 1.6 E). This “Up” conformation is thought to prime p97,

increasing its affinity for endogenous co-factors and for subsequent substrate processing. This observed sequential binding appears to be a consequence of the D1 domains higher affinity for either ADP or ATP ligands relative the D2 domain with a ~40-fold lower  $K_D$  in D1 with respect to D2, as measured by Surface Plasmon Resonance (SPR) experiments (Bulfer et al., 2016).

The main roles of p97 are to translocate and restructure proteins from large cellular structures such as organelle membranes and also to facilitate the degradation of released polypeptides by proteasomes. By recruiting multiple different cofactors, p97 performs a diverse array of applications within the cell despite only exhibiting one enzymatic function, that is, the generation of mechanical energy from ATP hydrolysis.

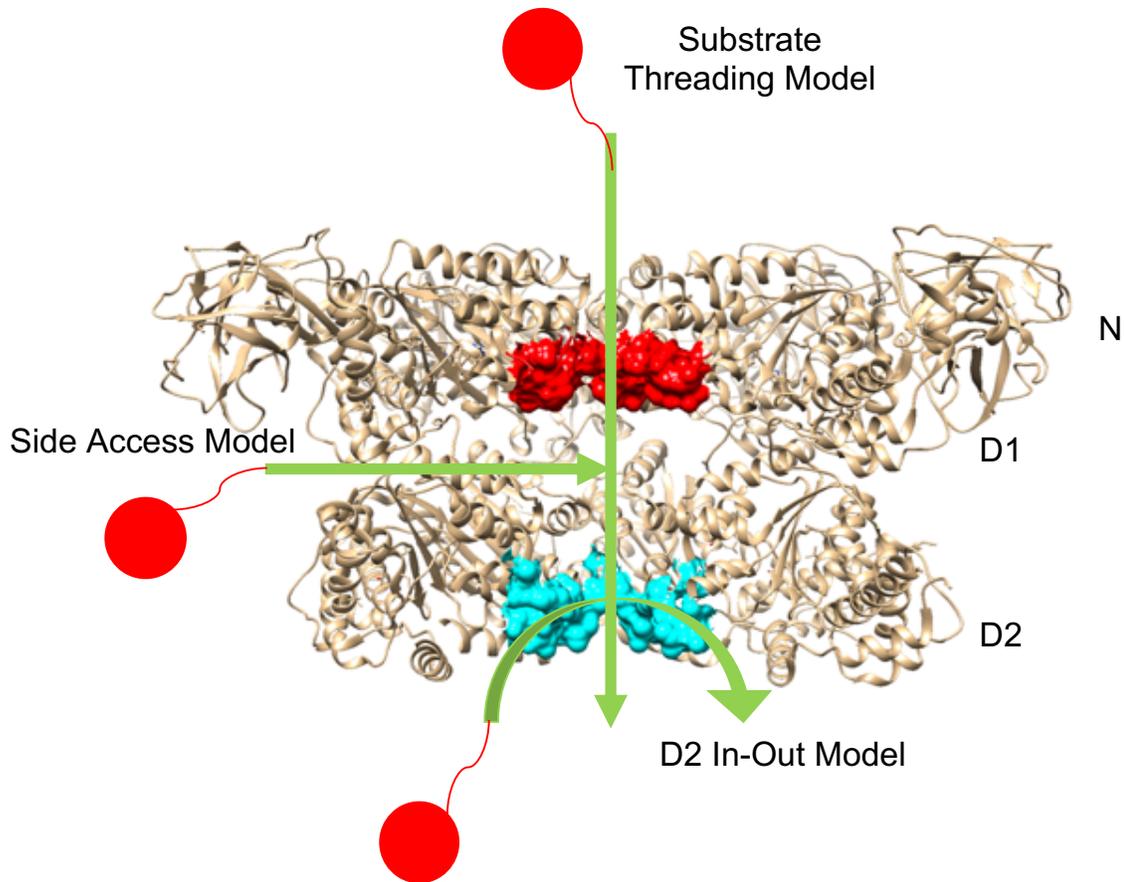
| Co-Factor Class | Example Cofactor | p97-Interacting Domain | Role   |
|-----------------|------------------|------------------------|--|
| UBX             | p47              | N-Domain               | Essential for Membrane Fusion in Golgi Biogenesis                            |
| VIM/VBM         | gp78             | N-Domain               | E3 Ligase Acts on ERAD Substrates and Catalyzes the Formation of K48 Chains. |
| PUB/PUL         | PLAA             | C-Terminus             | Required for the Clearance of K48-Labelled Lysosomes                         |

**Table 1.1: Examples of Ubiquitin-Associated p97 Cofactors.**



**Figure 1.6: Molecular Surface Model of Sequential ATP $\gamma$ S Binding States of p97. (Green: NTD; Blue: D1; Purple: D2.)** A) ADP bound p97 illustrating the “down” conformation of the NTD. B) ATP $\gamma$ S bound D2 domain which leads to a conformational shift in the D2 domain, with the D1 domain conformation remaining largely unchanged. C) ATP $\gamma$ S bound D1 and D2 domains illustrating the “up” co-factor binding conformation of the NTD. D) Superposition of the p97 protomer polypeptide backbones of the I and II (orange) conformations, demonstrating the differences in position of the D2 domain upon ATP $\gamma$ S binding. E) Superposition of the p97 protomer polypeptide backbones of the II and III (orange) conformations, demonstrating the shift in position of the NTD upon ATP $\gamma$ S binding to the D1 domain. ©(Banerjee et al., 2016) Page 874. Adapted by permission from publisher.

Due to the wide array of p97/co-factor relationships it has proved difficult to decipher the exact nature of p97-substrate interactions, with several models proposed (Figure 1.7).

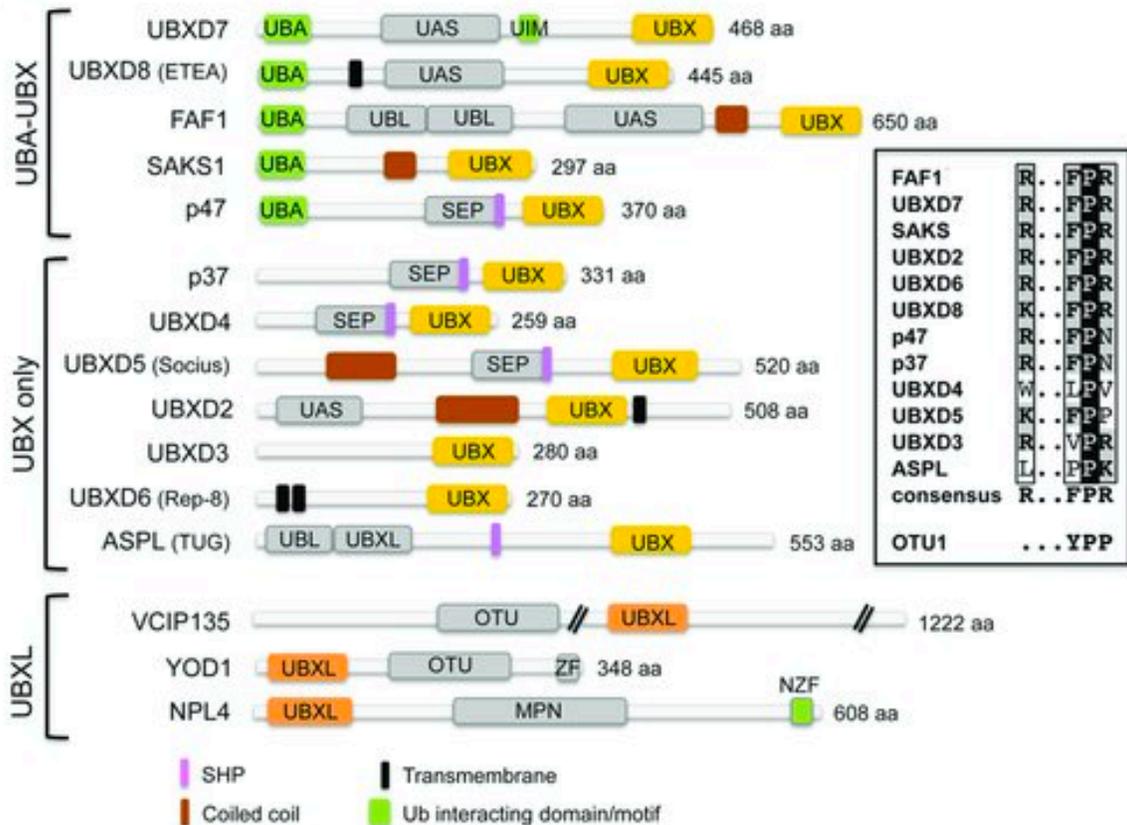


**Figure 1.7: Proposed Models for Substrate Entry and Disassembly by p97.** In recent years, there have been several models proposed for the unfoldase activity of p97. Here we illustrate 3 of the proposed models, all three emphasise the importance of the D1 and D2 pores residues (red and cyan respectively) in the substrate processing pathway of p97 [PDB: 5FTK].

Whilst structural and biochemical evidence has been presented for all 3 proposed processes, only recently have high resolution structural studies of p97 homologs in complex with substrate been published. These studies have indicated that the principle mode of ubiquitin-dependent unfoldase enzymatic activity in Cdc48 (a yeast p97 homolog) is through the substrate threading model (Twomey et al., 2019). This study used a poly-ubiquitinated model substrate, fixed to streptavidin beads to bind the Ufd1-Npl4 (UN) co-factor, (which contains both ubiquitin- and cdc48-binding domains) then cdc48 was added and binds the co-factor/substrate complex,

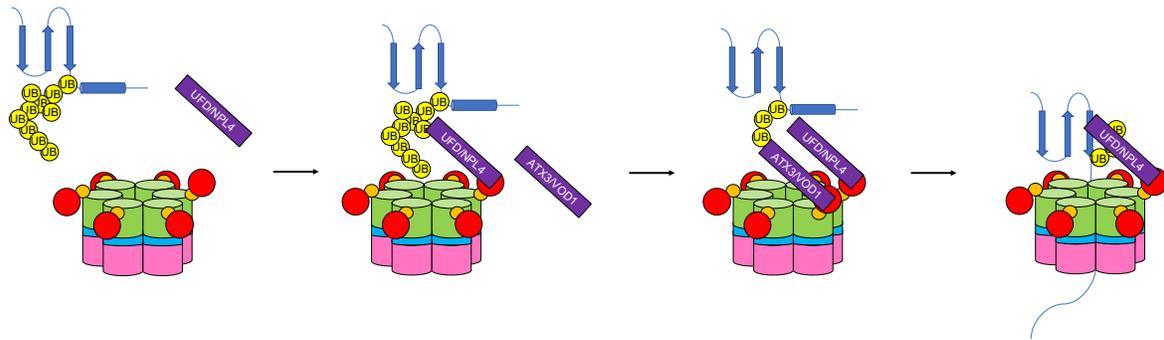
which was then isolated for structural studies. As the amino acid sequence of the substrate was known, it allowed for the modelling of substrate residues within cdc48's pore.

Another study was published concurrently (Cooney et al., 2019), that analysed cell derived substrates of cdc48, by the co-immunoprecipitation (co-IP) and rapid purification of a substrate recruiting cofactor in complex with cdc48 from cell lysates. The cdc48 binding motif used by the substrate recruiting cofactor in both studies was the Suppressor of High-copy PP1-containing (SHP) motif, which is a motif common in p97 co-factors (Figure 1.8). These studies confirmed a generalised model of ubiquitin-dependent SHP-mediated substrate trading through cdc48's pore. While these studies were performed on a p97 homolog, it is thought that mammalian p97 functions in a similar manner.



**Figure 1.8: Domain architecture of p97 cofactors containing Ubiquitin Regulatory X (UBX) and UBXL domains.** SHP binding motifs are labelled in Purple, indicating the relatively high abundance of the SHP binding motif among common ubiquitin binding co-factors of p97. However, as illustrated, even in the ubiquitin-dependent p97 pathway the SHP motif is a common, but by no means universal p97 binding motif. Several other co-factors such as another N-terminus interacting motif, the VCP Interacting Motif

(VIM), binds in a different location to the SHP motif (Hänzelmann & Schindelin, 2011) and yet another class of cofactors, the PNGase/UBA or UBX containing proteins (PUB) motif, binds the c-terminus of p97 (Zhao et al., 2007), suggesting there may still be more than one pathway for protein translocation through p97. Reproduced from (Hänzelmann & Schindelin, 2017).



**Figure 1.9: Proposed Co-Factor Binding and Substrate Processing Mechanism for Ubiquitin-Mediated Substrate Threading Model in Mammalian p97.** In at least one model system of p97 substrate degradation, it is thought that ubiquitinated proteins are bound by the UFD1/NPL4 co-factor, which interacts with p97 protein. Further co-factors are then recruited to control the threading of substrate through the pore of p97.

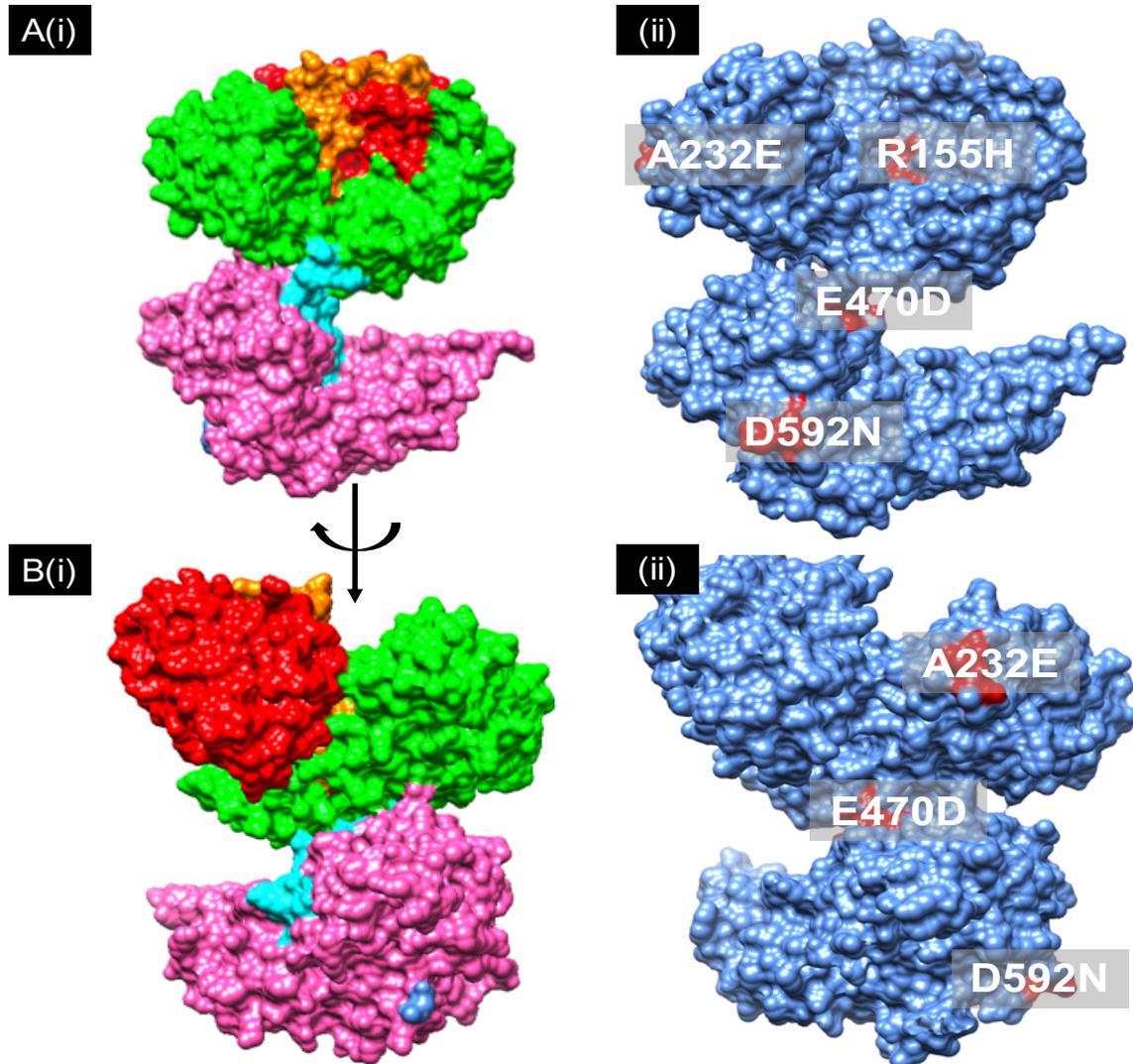
### 1.2.3. The Role p97 in Disease

Whilst the process of protein translocation in p97 has begun to be elucidated through structural biology, yet more work needs to be done to understand its role in disease. The majority of p97 mutations are associated with a rare disease called MultiSystem Proteinopathies-1 (MSP1) which is characterised by a cluster of muscular, bone and neurological diseases known as Inclusion Body Myopathy (IBM); Paget's Disease of the Bone (PDB); and FrontoTemporal Dementia (FTD) (Watts et al., 2004). These mutations tend to cluster in the N- and D1-domains of the p97 protein. As the majority of p97 co-factors bind the N-terminus of the protein this suggests possible co-factor involvement in the mechanism of disease. They are also commonly characterised by an increase in ATPase activity (Niwa et al., 2012), however this does not seem to translate to a higher enzymatic activity (Blythe et al., 2019).

Specific mutations in p97 have also been identified in closely related neurological illnesses such as Amyotrophic Lateral Sclerosis (ALS) (Johnson et al., 2010) which tend to have less pronounced muscular and bone involvement suggesting a subtle but significant difference in

the mechanism of disease. Still other mutations are implicated in cancer and drug resistance (Bastola et al., 2017), reflecting the diverse role of p97 within the cell and its importance in different tissues.

In this thesis we have chosen to study 4 pathogenic mutants of p97. With a view to inform therapeutic interventions, we investigated whether there is a common structural explanation for the mechanism of disease in p97 or if there are indeed significant structural differences in these mutants, which can explain the broad variance and severity in the phenotype of these diseases. The mutants under study in this thesis are the R155H, A232E, E470D and D592N mutants (Figure 1.10), these were chosen due to their relatively high incidence in p97-related disease and drug resistance.



**Figure 1.10: Space Filled Model of p97 Mutants.** A) (i) Side view of space filling model of p97 monomer, domains are colour coded; see figure 1.4. (ii) Mutated regions in red. B) (i-ii) Rotated view of space filled model illustrating the spatial orientation of the different mutants with respect to each other and the specific domains of p97 [PDB: 5FTK].

As these mutants are found in each of the 4 key domains of p97 (N-, D1-, D1-D2 linker and D2-) we expected that despite the relatively large distances between the mutants, a common structural phenotype may arise explaining the similar disease presentation of these mutants.

The p97<sup>R155H</sup> mutant is a missense substitution mutation (c.736G>A)<sup>1</sup>, found on the N-terminal domain of p97 and is the most common mutation location identified in patients with MSP1 (Watts et al., 2004). Studies of p97<sup>R155H</sup> pathology in both mice and human patients have shown a reduction in mutant interaction with caveolin-1 (CAV1) and the cofactor UBXD1. UBXD1 is necessary for the endolysosomal trafficking of ubiquitinated CAV1. A lack of p97 recruitment has been shown to lead to accumulations of CAV1-positive endolysosomes in diseased tissues and consequently a reduction in CAV3, a muscle-specific caveolin (Ritz et al., 2011). Autosomal dominant mutations in CAV3 have been shown to lead to a form of muscular dystrophy with phenotypic similarities to the myopathies associated with MSP1 (Minetti et al., 1998). This may indicate a possible mechanism by which the p97 mutant phenotype is present in MSP1.

Other biochemical studies have also indicated that the population of cofactors bound to p97<sup>R155H</sup> compared to wild-type vary considerably. For example, p97<sup>R155H</sup> has increased associations in co-IP experiments between p47, Npl4, Ufd1 and Ataxin-3 (Atx3) cofactors; and a decrease in association with the ubiquitin ligase E4B cofactor relative to WT (Fernández-Sáiz & Buchberger, 2010). As accumulation of mutant-Atx3 on p97 is associated with neurodegeneration in spinocerebellar ataxia type 3, it is thought that the accumulation of WT-Atx3 on mutant p97 could explain the neurodegeneration observed in patients with MSP-1.

The p97<sup>A232E</sup> mutant is a missense substitution mutation (c.967C>A), found in the D1 domain, positioned at the inter-protomer interface. p97<sup>A232E</sup> has attracted a lot of interest in studies of mutant p97 disruption due to its high relative ATPase activity, it is suggested to be a potential cause of the mutant's increased severity and speed of disease progression (Watts et al., 2004).

Biochemical assays of p97 substrate activity, showed p97<sup>A232E</sup> has an increased ubiquitin dependent unfoldase activity relative to WT (Blythe et al., 2017). However, in another study

---

<sup>1</sup> All nucleotide positions are based on NM\_007126.5 and amino acid numbering is based on NP\_009057.1.

by the same lab, this increase in unfoldase activity was found to have no significant correlation with ATPase activity (Blythe et al., 2019).

Furthermore, some studies (Zhang et al., 2015) have suggested p97 mutants lead to an overall inhibition of ATPase activity upon binding the p37 and p47 cofactors, which suggests that p97 mutant effects may have a more complex mechanism of disease than a simple increase in ATPase activity.

The p97<sup>E470D</sup> mutant is a missense substitution mutation (c.1682G>C), found on the D1-D2 linker in close proximity to the entrance to the D2 nucleotide-binding pocket. p97<sup>E470D</sup> is unique in this study as it has been observed *in-vitro* as an acquired mutation in ovarian cancer cells upon the acquisition of resistance to CB-5083 (Bastola et al., 2017), a first-in-class p97-specific inhibitor to reach clinical trials. It is important to understand the structural components of resistance to p97-specific inhibition in cancer in order to circumvent such resistance when it presents in patients.

The p97<sup>D592N</sup> mutant is a missense substitution mutation (c.2046G>A), found in the D2 domain at the central D2 pore loop (Johnson et al., 2010). This mutant was identified in patients diagnosed with ALS and is one of the few mutations found in the D2 domain that is implicated in neurodegenerative disorders, the majority of mutations identified in this region are linked to cancer. The presence of this mutation in the D2 pore loop suggests substrate processing involvement and due to its large distance relative to the majority of p97 MSP-1 mutations may point to a common source of p97-linked neurodegeneration.

Our aim was to isolate and interpret the structures of each of these mutants and use them to generate a holistic understanding of p97-related disease, which is described in Chapter 2.

### 1.3. Understanding Tissue Ultrastructure using High-Resolution Focused Ion Beam – Scanning Electron Microscopy

The objective for this section is to generate large volumes of high-resolution data of sufficient size for quantitative population-wide analysis.

As described in greater detail in chapter 3, we acquired several hundreds of thousands of  $\mu\text{m}^3$  of human muscle tissue samples using a Focused Ion Beam-Scanning Electron Microscope (FIB-SEM) and segmented their mitochondria. These volumes allowed us to generate a standardised dataset of mitochondrial distribution in healthy individuals and assay the methods sensitivity and power in determining changes in the overall mitochondrial distribution within the sampled population.

#### 1.3.1. Large Volume Acquisition and Automated Segmentation Techniques For Population-Wide Analysis

Quantitation of biological phenomena has largely remained in the realm of bulk assays, by averaging large numbers of constituents it provides a low-resolution picture of the overall system. Examples of such techniques are ubiquitous across both scientific and public life, from handheld blood glucose monitoring to pregnancy tests to more sophisticated population-wide studies in multifactorial disease.

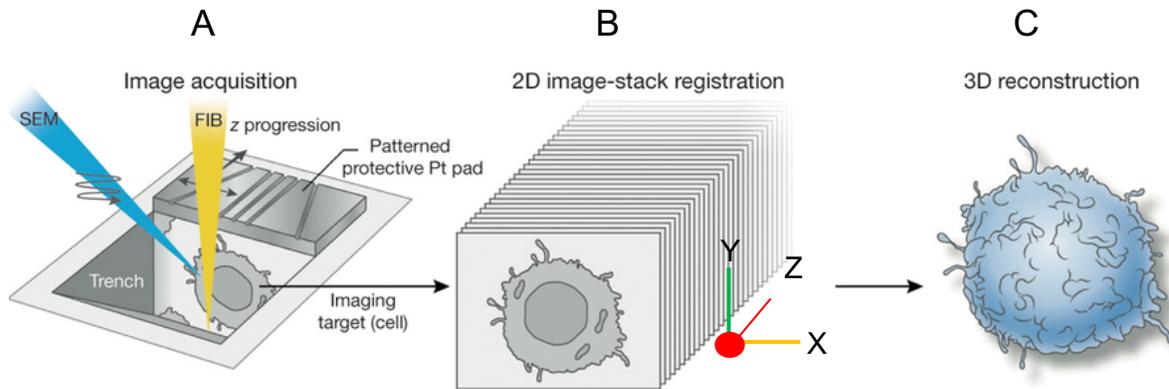
All of these rely on the large-scale amplification and quantification of the overall concentrations of a target in a system of study as is required for any statistical analyses to be robust enough to stand-up to scientific scrutiny.

Most microscopic techniques have been used to perform the qualitative end of scientific analysis, thereby answering the question of “how” rather than the “how much”. Techniques such as Stochastic Optical Reconstruction Microscopy (STORM) and High Resolution – Transmission Electron Microscopy (HR-TEM) have provided mechanistic insights that can inform future biological assays, allowing for a more in-depth understanding of mechanisms underpinning biological functions. However, the immense detail and insight provided by high-

resolution techniques are often difficult to scale and therefore a meaningful statistical description on a population-wide scale at these sub-cellular resolutions remains elusive. This is partly due to instrumental limitations such as sample preparation and acquisition time, but a major limitation is that of the restrictive amount of time required for the manual annotation, measurement and tabulation of the data.

Some recently developed techniques are beginning to address this gap, including methods such as flow cytometry that correlates qualitative signals such as an antibody or gene fluorescent markers to their relative abundance in a cell population, thus providing robust statistical inference at high-resolutions complementary to the light microscopy data. However, there are limitations to this technique, as the analyte must have a high-fidelity label and also requires an understanding of the underlying mechanistic implications of the abundance (or lack thereof) of the marker being tracked. Flow cytometry also cannot supply vital spatial information of the 3D arrangement of biological structures within cells or tissues provided by the specialized techniques of light and electron microscopies.

Transmission Electron Microscopy (TEM) provides an extremely detailed picture of the inner mechanisms of a cell but as the thickness of the sample increases, so too does the scattering of the electrons and subsequent blurring of the image. This limits TEM to giving snapshots of a 2D XY-slice through a cell per image, although some effort has been made to increase the 3D imaging capacity of the TEM using serial section-TEM (ssTEM), which combines images of several successive slices through a sample at high resolution. However, this is extremely time consuming and the Z resolution is distorted due to the variability in the slicing through the resin embedded sample.



**Figure 1.11: FIB-SEM Workflow.** **A)** The imaging face is eroded using the ion beam and imaged using the SEM. **B)** These images are compiled, aligned and binned to create a 3D z-stack. **C)** Regions of interest in the volume are segmented manually or computationally to allow for a 3D rendering of structures of interest. Adapted by permission from **Springer: Nature; Nature Methods; Focused ion beams in biology, Kedar Narayan and Sriram Subramaniam, © 2015.**

FIB-SEM dramatically decreases acquisition time by performing the sectioning inside the instrument (Figure 1.11). The use of an ion beam enables control of Z width and reduces slice to slice variability, generating an isotropic voxel at extremely high spatial resolutions enabling whole cell/tissue imaging on a nanometer scale. The study of the 3D organisation of tissues and their constituent organelles at high resolutions is still underdeveloped but has the potential for detecting and characterising variations in the cellular architecture of tissues early on in disease states, informing prognosis and monitoring progression.

FIB-SEM supplies a unique opportunity to explore tissue organisation. However, the combination of high spatial resolution and large volume acquisition leads to an exponential increase in acquisition time, meaning these two qualities remained mutually exclusive. The application of FIB-SEM in the study of biological tissues is further limited by the prohibitive amount of time required to annotate large datasets manually, until now. By reducing acquisition and segmentation time, the high throughput method developed herein opens the doors to the study of whole tissues and fibres at EM resolutions, which will undoubtedly create interesting avenues of discovery for diseased tissues in general.

### 1.3.2. Human Muscle

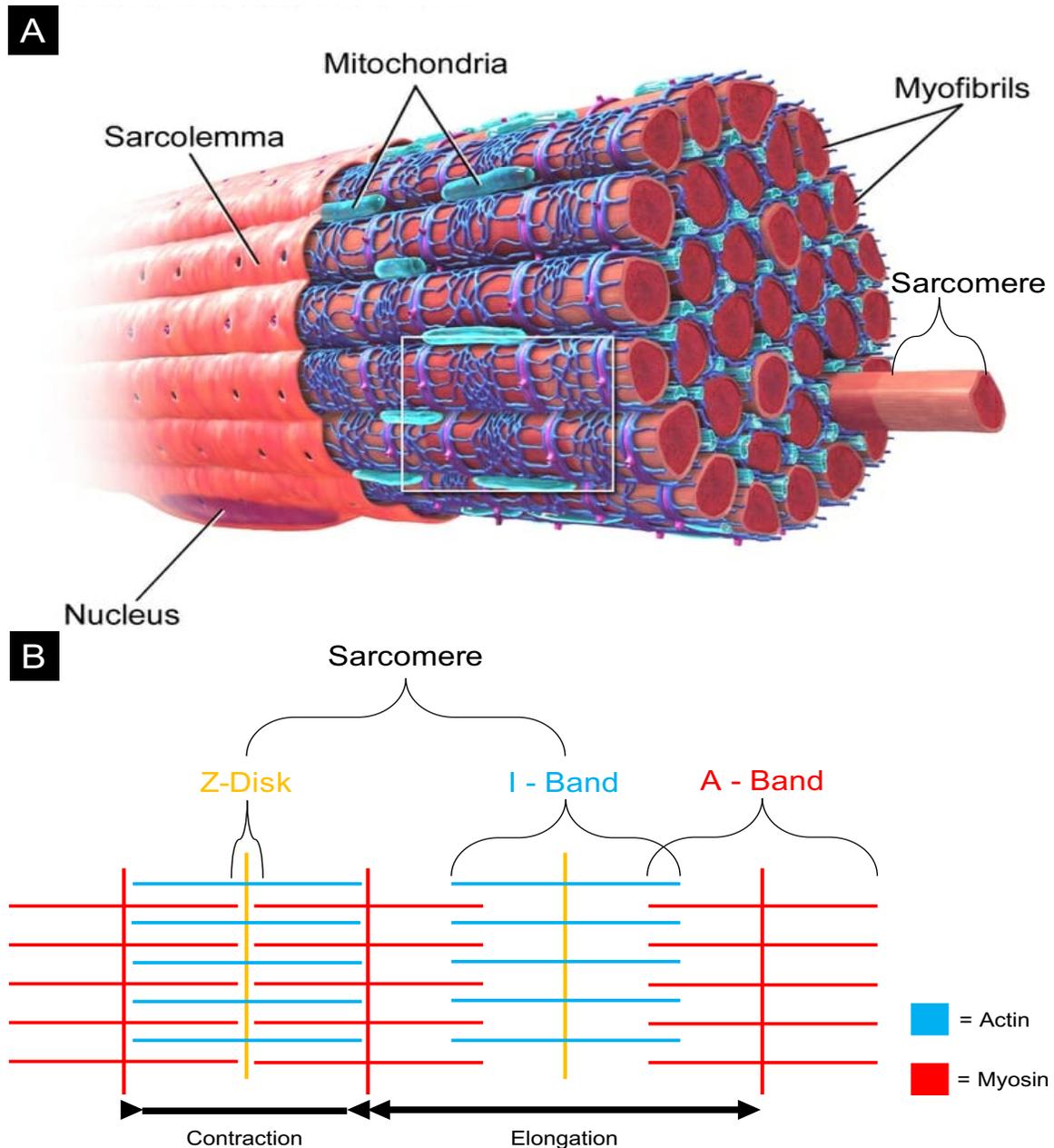
Previously, the Subramaniam lab has developed techniques for the study of mitochondrial distribution in murine muscle tissue (Glancy et al., 2015) . In our study, we chose to apply this knowledge and expertise to work with human muscle due to its relevance in human disease and ageing.

Human muscle is classed into 3 main types: Visceral (smooth), cardiac and skeletal muscle.

These 3 classifications are based principally on their location and function within the body:

1. Visceral muscle is found in the walls of organs and structures such as the intestines and oesophagus, it is an involuntary muscle due to the lack of conscious input required for its function.
2. Cardiac muscle is also an involuntary muscle and is found exclusively in the heart.
3. Skeletal muscle is anchored to the bone and is the only muscle class where most contraction and relaxation occurs due to conscious effort originating in the brain and therefore is considered the only voluntary muscle class. Skeletal muscle provides support of the body, effects skeletal movement and protects internal organs.

We chose to analyse skeletal muscle in this study as it is the safest to extract and most accessible of the 3 muscle tissues. Skeletal muscle fibres are composed of multiple smaller fibres known as myofibrils enveloped by a tubular sheath known as a sarcolemma (Figure 1.12). Myofibrils are surrounded by mitochondria which supplies energy to the fibre and other structural proteins such as the sarcoplasmic reticulum which regulates the calcium ion concentration within the muscle. Each myofibril contains a series of repeating structural units between Z-disks known as sarcomeres which contain both the actin and myosin filaments required for muscle contraction and elongation.



**Figure 1.12: Major Components of Muscle Fibres.** A) Illustration of the ultrastructural components of a single skeletal muscle fibre. Adapted from (*Medical gallery of Blausen Medical 2014, 2014*) B) Illustration of the sarcomere and the myofilament proteins and major structures associated within it.

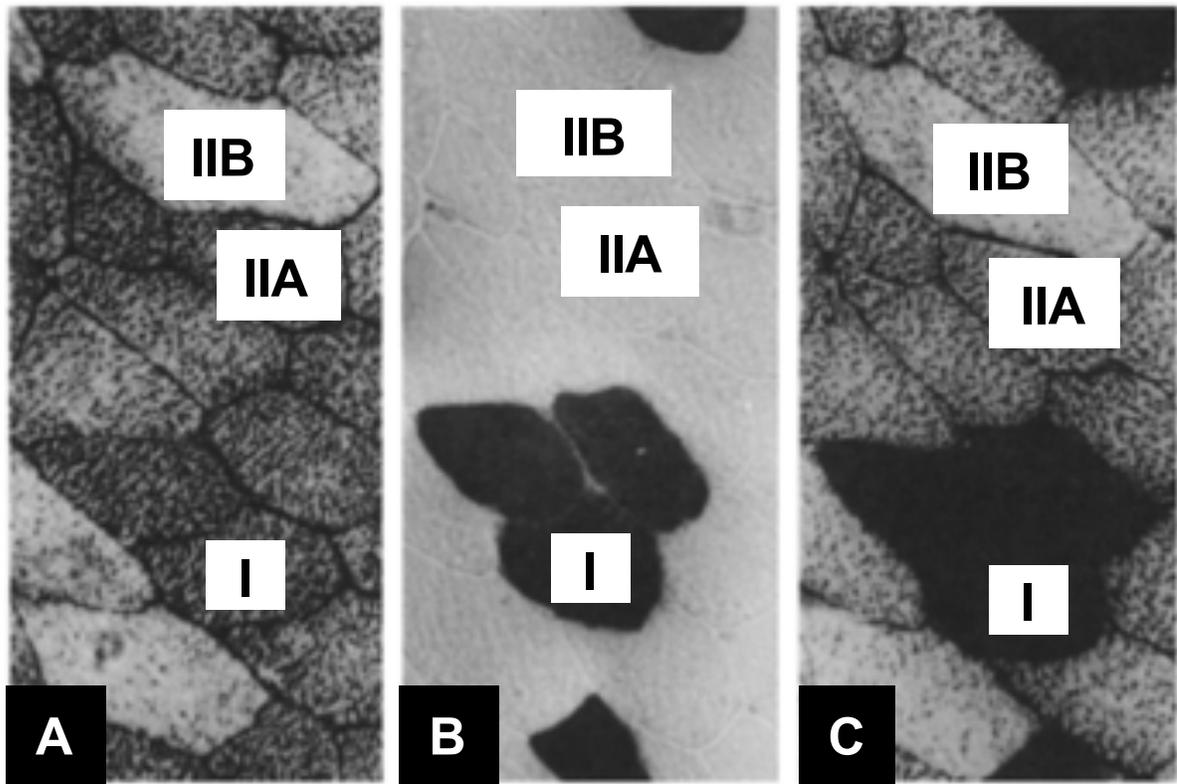
Skeletal muscle can be further divided into 3 subtypes; Type I, Type IIA and Type IIB muscle. These muscle types are primarily defined by their rate of contraction and their main pathway

for ATP production which is: slow-contracting oxidative, fast-contracting oxidative and fast-contracting glycolytic respectively.

These muscle types can be identified using specific stains in light microscopy, of which the most reliable way to distinguish muscle types is through the use of an ATPase stain which relies on the pH dependent inactivation of myosin ATPase isoforms, which are characteristic of the muscle types (Sciote et al., 1994). Higher concentrations of the Type I myosin ATPase isoform remain active in acidic conditions than in Type II, therefore Type I is heavier stain relative to Type II fibres.

Another distinct feature of the skeletal muscle types is the distribution of mitochondria within the fibre, with Type I and Type IIA having a high, and Type IIB a low, concentration. This can be visualised (Horák, 1983) on a muscle biopsy through the staining for Succinate DeHydrogenase (SDH) activity (Figure 1.13), which is a mitochondrial specific stain and is an indicator of overall oxidative activity. In Type I and Type IIA fibres, the relatively high mitochondrial concentrations result in heavy staining, whereas Type IIB, having relatively few mitochondria, is weakly stained.

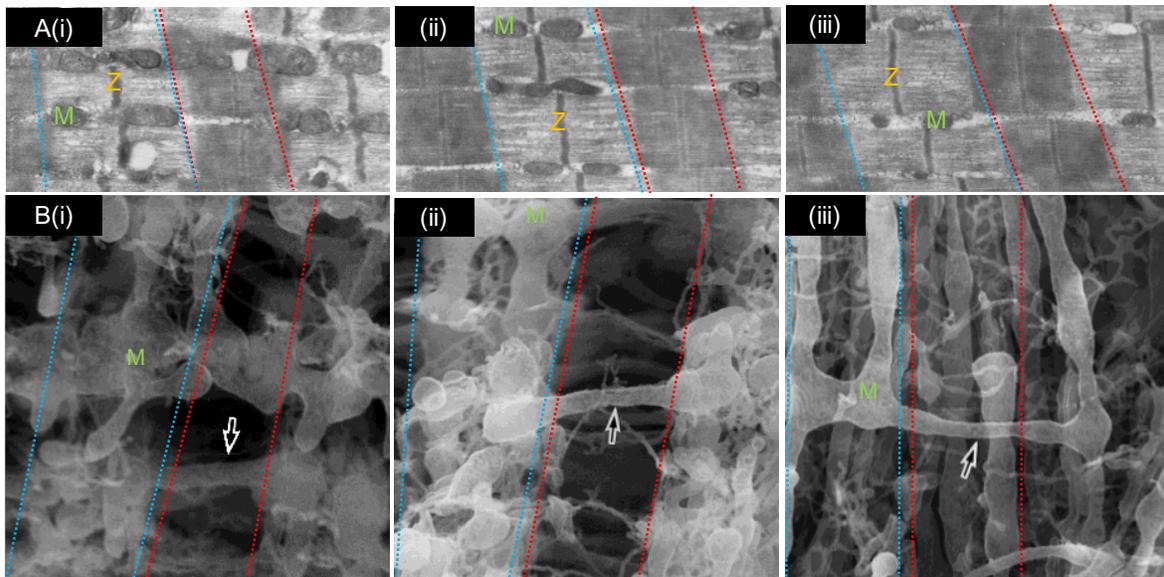
Below is an example of a successively stained skeletal muscle biopsy:



**Figure 1.13: Successive Skeletal Muscle Histochemical Staining.** A) Fibre-dependent Succinate Dehydrogenase (SDH) Stain. B) Fibre-dependent ATPase stain at pH 4.1. C) Successive staining of serial sections for both enzyme activities. Adapted by permission from Springer: Nature; Histochemistry, A successive histochemical staining for succinate dehydrogenase and “reversed”-ATPase in a single section for the skeletal muscle fibre typing, V. Horák © 1983.

### 1.3.2.1. Subcellular Morphological Differences Between Skeletal Muscle Types

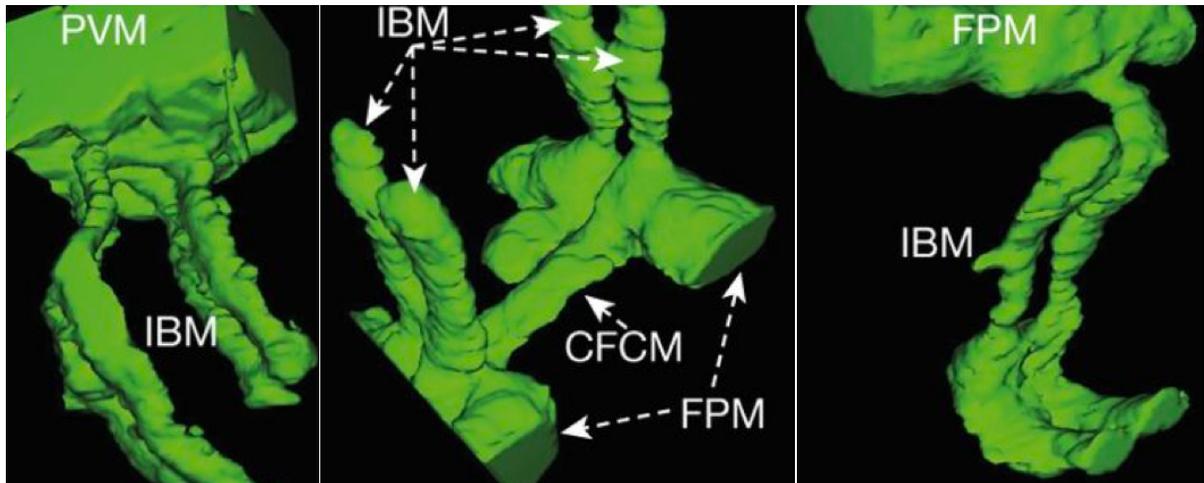
Distinctions between muscle subtypes have been qualified on the subcellular level with electron microscopy and have illustrated some key structural differences in mitochondrial morphology.



**Figure 1.14: Skeletal Muscle Ultrastructure and Their Differences Between Subtypes.** A) Longitudinal TEM sections of Type I (i), Type IIA (ii) and Type IIB (iii) respectively. B) SEM image of macerated tissues samples which results in the removal of the major myofilaments but leaves the mitochondrial structure of Type I (i), Type IIA (ii) and Type IIB (iii) muscle relatively intact. Magnification = x20,000, M = Mitochondria, Z = Z-Disk Blue = I-Band, Red = A-Band, Arrows indicate mitochondria traversing the A-band. (Ogata & Yamasaki, 1997) Pages 216-221. Adapted by permission from publisher. ©1997 WILEY-LISS, INC.

As illustrated in figure 1.14 A, there is a clear decrease in mitochondrial cross-sectional area from type I to type II muscle. In figure 1.14 B, these differences are further emphasised in 3D with marked differences in mitochondrial density and distribution, with type I and IIA both having greater density within the I-band and larger trans-A-band connections than that of type IIB, the morphological difference between the different types of mitochondria is also prevalent with type IIB being flatter and more elongated than type I.

Previously, research from the Subramaniam lab has furthered the understanding of the 3D distribution of mitochondria in muscle, through the use of FIB-SEM methods (Glancy et al., 2015), defining specific structural patterns within skeletal muscle, through the use of murine models, these structural patterns (Figure 1.15) illustrate the interconnectedness of mitochondria within the muscle and their different functions.



**Figure 1.15: Examples of Different Mitochondrial Morphologies Within Skeletal Muscle Fibre.** Manually segmented mitochondrial motifs found in type II murine skeletal muscle, demonstrating the structural heterogeneity and specialisation of individual mitochondria within a fibre. PVM = Para Vascular Mitochondria, IBM = I- Band Mitochondria, FPM = Fibre Parallel Mitochondria and Cross Fibre Connection Mitochondria (CFCM). Reprinted by permission from **Springer: Nature; Nature, Mitochondrial reticulum for cellular energy distribution in muscle**, B.Glancy et al. © 2015.

Here, our aim was to build on this knowledge and generate a quantitative measure of mitochondrial distribution in muscle. This required the development of a rapid acquisition and segmentation workflow for FIB-SEM to accurately define the mitochondrial boundaries within muscle tissue and its evaluation against manually segmented standards to ensure its overall accuracy and fidelity to ground truth measurements.

This workflow enabled us to define the mitochondrial distribution in muscle tissues through a series of statistical and morphological metrics and to generate a quantitative model of different muscle types within human tissues for future studies, as described in detail in Chapter 3.

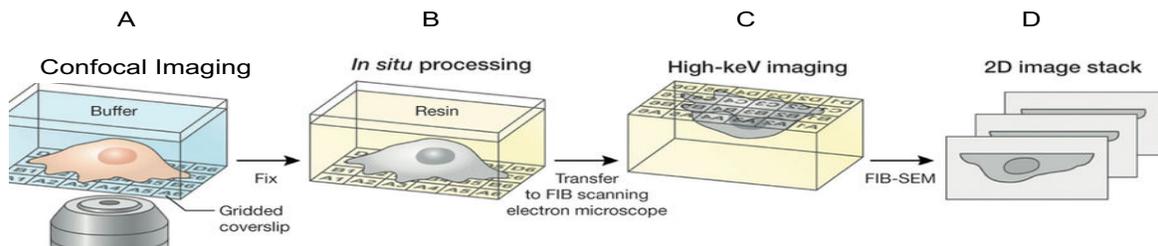
## 1.4. Examining Nanotherapeutic Interactions on a Cellular Scale.

The objective of this section is to develop and use correlative light electron microscopic techniques for the study of nanoparticle uptake in cells.

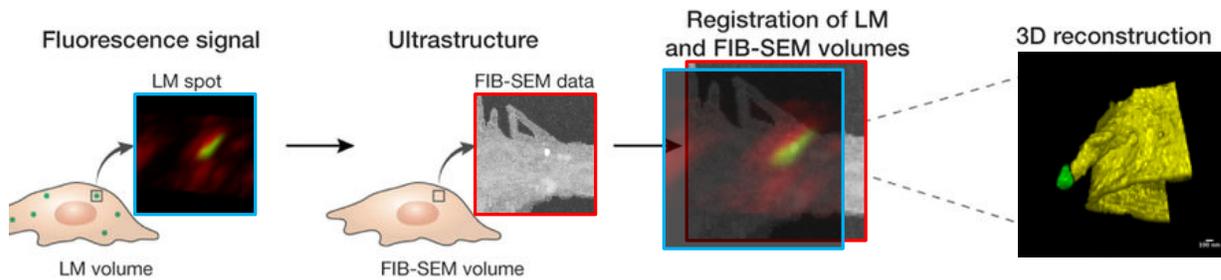
As described in greater detail in chapter 4, we acquired several correlated light and electron microscopy volumes of cells and segmented multiple regions of interest. These provided highly resolved 3D volumes of the different stages of nanoparticle receptor-mediated uptake in cells.

### 1.4.1. Correlative Light Electron Microscopy (CLEM)

CLEM combines both light microscopic techniques such as confocal microscopy and electron microscopic techniques such as FIB-SEM (as described in section 1.3.1), these techniques complement each other, merging the large sampling volume and broad selection of fluorescent labels and applications e.g. Förster Resonance Energy Transfer (FRET), Fluorescence Lifetime Imaging (FLIM) etc...with the high resolving power of the electron microscope. As nanoparticle-cell interactions occur on the scale of the tens of nanometres, combining both techniques provide unique insights greater than the sum of their parts, leveraging the discriminating power of light microscopy with the high precision and high spatial isotropic resolution of electron microscopy.



**Figure 1.16: CLEM Workflow.** **A)** Confocal imaging of sample and logging of region of interest (ROI) positions on gridded coverslip. **B)** Fixation, resin embedding and removal of sample from cell culture dish. **C)** FIB slicing and SEM imaging of ROI. **D)** Computational registration and alignment of 2D image stack. Reprinted by permission from **Springer: Nature; Nature Methods; Focused ion beams in biology**, Kedar Narayan and Sriram Subramaniam. © 2015.



**Figure 1.17: Correlation of Light and Electron Images.** Light and electron stacks are correlated using gold nanoparticles as registration markers and ROIs are segmented manually to visualize nanoparticles in 3D. Adapted by permission from **Springer: Nature; Nature Methods; Focused ion beams in biology, Kedar Narayan and Sriram Subramaniam.** © 2015.

#### 1.4.2 Confocal Light Microscopy

Confocal microscopy has several advantages over conventional widefield microscopy, including its capability to collect z-stacks of optical sections through thick samples and especially the elimination of background information away from the focal plane. Confocal microscopy has become a mainstay of modern molecular biology due to its high resolution in all three dimensions.

#### 1.4.3 Conventional therapeutics

Modern medicine has witnessed a near mastery over previously common infections and illnesses which a century ago proved often fatal such as tuberculosis, syphilis and malaria. Small molecule-based therapies have revolutionised treatment options for humankind paving the way for the immense growth witnessed in the 20<sup>th</sup> century. No such disease highlights the technological and efficacious advancement of small-molecule treatments better than the treatment of HIV, which took less than 20 years from its formal recognition as a new pathogen to the production of a successful combinatorial treatment, which effectively eradicated the symptoms of AIDS and halted the replication and transmission of the virus within the human population.

However, the immense progress of the 20<sup>th</sup> century has brought with it new challenges for the 21<sup>st</sup>. Illnesses thought to be under control have reemerged in more tenacious and aggressive forms than ever; multi-drug resistant cases of tuberculosis are growing at a rate of 20% per annum (Lange et al., 2018), resistant strains of HIV began to appear in populations soon after the introduction of novel retroviral therapies (Zazzi et al., 2018); chloroquine and antifolate resistant malaria is leading to the resurgence in the use of quinine, a centuries old therapy, as a second- and sometimes first-line treatment (Achan et al., 2011).

Indeed, resistance to molecular drugs could begin to outpace their discovery and one of the best demonstrations of this is in oncology where the demand for novel drugs to fight our chameleonic enemy within is ever increasing with every successive round of an individual's chemotherapy. In an effort to sustain efficacy without undermining a patient's health, more and more complex, multi-component regimes tailored to the individual are being prescribed for a given cancer after initial treatments have failed. Most conventional therapeutic agents currently in the clinic act on the molecular level, by passively diffusing across cell membranes with limited selectivity. Many initially promising drug candidates are eliminated in early phase clinical trials because of their adverse side-effects or reduced efficacy *in-vivo*.

#### 1.4.4 Novel Methods for Medicine

Currently, we are witnessing a sea-change in the design approach of novel therapeutics and we now recognize that in order to create more potent and less hazardous pharmaceuticals in the future, a more sophisticated approach is needed (Aggarwal et al., 2007). Medical research is beginning to explore different avenues of enquiry to discover more targeted drug delivery mechanisms. Therapies such as antibody-drug conjugates utilize nature's "silver-bullets" to deliver drugs to disease centers in the body in a targeted fashion. This allows for increased concentrations of a drug that would otherwise be lethal through normal delivery routes and for revival of drug candidates once thought too non-specific on their own for use in disease treatment. Revolutionary approaches such as Chimeric Antigen Receptor-T (CAR-T) cell therapies recruit the patient's own immune system in the fight, paving the way for drug free remission.

These therapies are becoming widely accepted in the medical field, and their high specificity along with low toxicity relative to their efficacy make them attractive platforms for the development of other alternative therapies. However, they are often highly personalized and can be difficult to translate to other pathologies.

Accordingly, we may need to look elsewhere for a modularised platform for delivering therapeutics in a targeted and consistent manner. One such platform attempts to perform such a task through the careful design and presentation of therapeutics in 3-dimensions on the nanoscale similar to viral particles and other cellular machinery. It seems like a logical step to emulate biology's own machinery for the purposes of defeating biological diseases.

Nanoparticles are a heterogeneous class of nanostructures with often very different physical and chemical properties from one another but are related through their characteristic size (<100nm in at least 1 dimension) and the unique biophysical properties these size constraints illicit from these structures. Unable to diffuse passively across cell membranes they lend themselves amenable to exploiting more complex receptor targeting pathways like endogenous processes utilized by our cells daily to grow, communicate and die. This higher level of abstraction allows for a variety of different targets and targeting designs. However, current nanoparticle-based therapeutics have been very slow to make it to market and progress has been mired by the very properties we wish to exploit i.e. nanoparticle's size and it's 3-dimensional nature.

Few therapies have been implemented currently, therapies such as liposomes have had limited success and mainly perform an improved formulation role rather than an active agent for targeting disease.

#### 1.4.5 Nanoparticle-Cell Interactions

Another area of enquiry in this project is the application of the knowledge and experience in the lab in UCD to address questions on how nanoparticles behave on and inside cells and how their behaviour can be modified to increase cell specificity, drug targeting, bio-availability or indeed how modifications to one can affect the others. Research into nanoparticle uptake and bioprocessing tends to focus on two main areas, either through the use of fluorescent labels to

track organelles/receptors of interest relative to nanoparticles of different compositions (Varela et al., 2015) or through the use of electron microscopic methods to study the exact positions and morphology of organelles affected by nanoparticle uptake (Guehrs et al., 2017).

Our aim in this section of the thesis was to use a combination of both techniques, to create a symbiosis where the limitation of one method is mitigated by the other and vice versa. This correlated technique allowed us to pin-point exact positions down to a couple of nanometres while also identifying the substituents with the accuracy of immuno- and transfected labels. This process allowed us to map the course of nanoparticle-cell interactions from receptor binding and endocytosis to excretion and degradation.

To develop a method for correlative imaging of nanoparticle-cell interactions; we chose to focus on the endocytosis of the nanoparticles through the transferrin receptor. Endocytosis is characterized by the internalization of biomolecules into membrane-bound compartments. Vesicular trafficking is divided into two main categories, the clathrin-mediated endocytic pathway and the non-clathrin mediated endocytic pathway. The transferrin receptor (TfR) is a stereotypical example of the family of receptors internalized through the clathrin-mediated pathway. Overexpression of the TfR is characteristic of several different cancer cell populations (Gatter et al., 1983), thus transferrin is a suitable candidate for drug delivery systems (Qian et al., 2002). Therefore, we used transferrin bound gold nanoparticles (Tf-AuNp) as our model particle for this study because of its safety, selectivity, tumour-targeting capabilities and because of its interest in the nanoparticle field (Qian et al., 2002).

The application of CLEM techniques enabled us to define several crucial steps in the uptake, bioprocessing and subsequent excretion of nanoparticles from cells, as described in detail in Chapter 4.

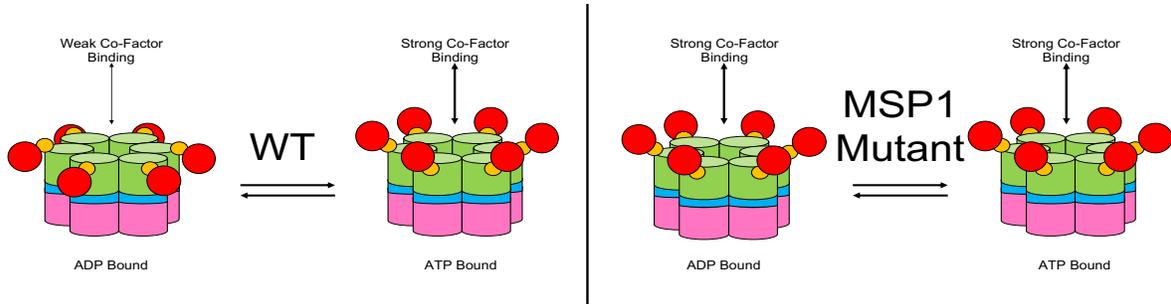
## 2. Structure Elucidation of p97 using Cryo-Electron Microscopy

### 2.1. Chapter Introduction

As discussed in the introductory chapter, p97 has proven to be critical in a number of biological processes and central to several neurodegenerative disorders. Understanding the biomolecular mechanisms of p97-related disease is therefore important for developing pharmacological interventions. While several atomic resolution structures of full-length wild-type (WT) p97 in different nucleotide-bound states have been reported using cryo-EM (Banerjee et al., 2016) there are no published reports of full-length structures of disease related mutants of p97. Previous structural studies of p97 mutants were performed on N-D1 truncated p97 constructs (Tang & Xia, 2013) using x-ray crystallography, protein dynamics analysis of full-length mutants (Schuetz & Kay, 2016) using NMR and on p97 in complex with Ufd1-Npl4 co-factors (Blythe et al., 2019) through cryo-EM. These publications identified potential structural characteristics which may indicate a mechanism for p97-related disease, which we will discuss further in this chapter.

While most mutants identified in patients diagnosed with MSP-1 are almost exclusively found in the N and D1 domains (Watts et al., 2004), the primary source of ATPase activity is derived from the D2 domain and is dependent on nucleotide binding to the D1 domain (Tang & Xia, 2013) suggesting inter-subunit communication between domains. Therefore, it is necessary to study p97 mutants in their entirety to determine their global effects on p97 function.

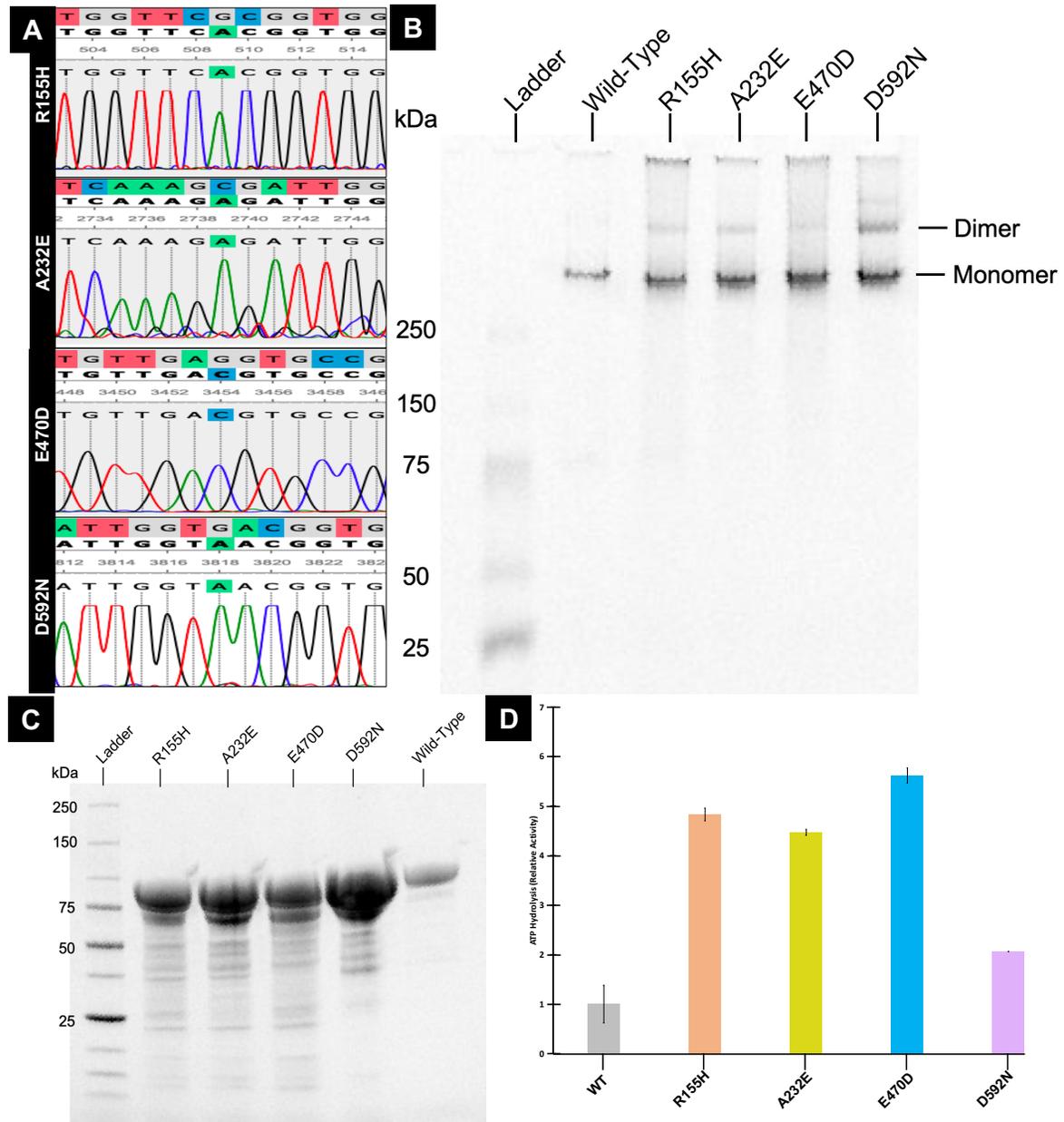
Here, we present our structural studies of both full-length p97<sup>R155H</sup> and p97<sup>A232E</sup> in both the ADP- and ATP $\gamma$ S-bound states and for the first time a full-length D2 domain mutant p97<sup>D592N</sup> implicated in another neurodegenerative disorder, ALS. We also present a cryo-EM structure of p97<sup>E470D</sup>, a mutant implicated in resistance to anti-cancer therapeutics.



**Figure 2.1: Proposed Mechanistic Model of MSP-1 Mutants.** The relative increase in ATPase hydrolysis in MSP1 mutants and structural analysis of MSP1 mutants in the monomeric state suggest p97 mutants are constantly in a pre-activated co-factor binding state, resulting in the sequestration of p97 binding cofactors. However, the increased ATPase activity and cofactor binding does not translate to increased substrate binding (Zhang et al., 2015), which could lead to protein accumulation and explain the etiology of MSP1 and other p97-related disease.

## 2.2. Results and Discussion

### 2.2.1. Expression and Characterisation of p97 Mutants

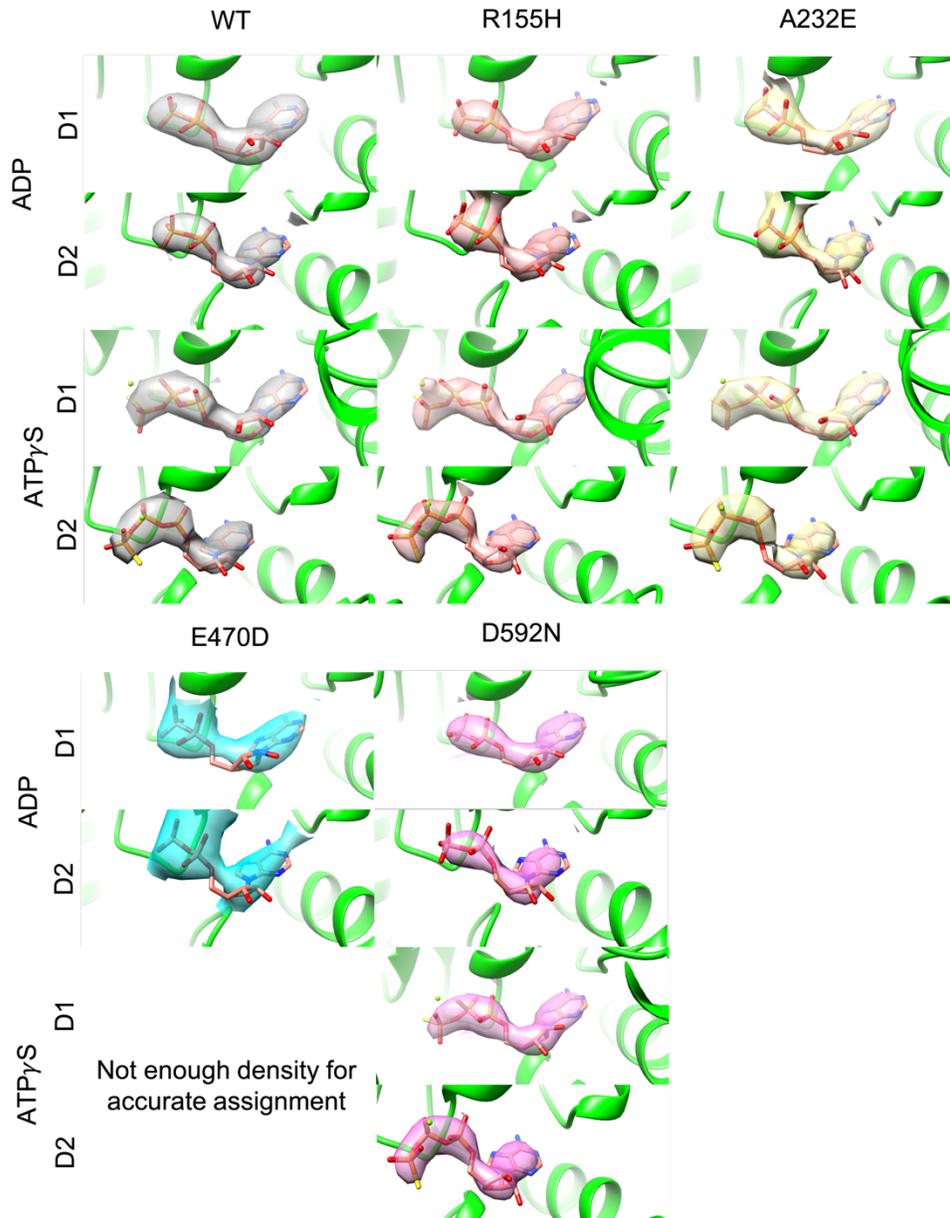


**Figure 2.2: Biochemical Analysis of p97.** **A)** Section of Sanger sequence results containing point mutations for each mutant. **B)** Native PAGE of His-tag purified p97 mutants, showing the two species of p97 found in solution, a single hexamer and a dimer of hexamers. **C)** SDS-PAGE of His-tag purified p97 mutants. **D)** ATPase activity assay of BCA normalised concentrations of p97 and its mutants (n=3).

Figure 2.2 shows the results of the biochemical analysis of isolated p97 mutants. Each mutant plasmid was sequenced to confirm the presence of the correct point mutation (Figure 2.2 A). Mutants were expressed in *E. coli* and isolated after overnight induction, the bacteria were then pelleted and lysed. The lysate was then purified using a gravity flow His-Pur™ column. After His-tag purification, each protein was analysed for yield and purity using nanodrop and SDS-PAGE (Figure 2.2 C), typically ~2mg of protein at a purity of 60-70% with the remaining contaminants being low-molecular weight degradation products of p97. The presence of a native hexameric structure was confirmed using Native-PAGE (Figure 2.2 B). For the ATPase analysis, the protein concentrations were normalised and analysed using a plate-based assay according to section 2.4.1.6. This assay (Figure 2.2 C) indicates each mutant has elevated ATPase activity relative to WT protein which was comparable with previously published data (Niwa et al., 2012).

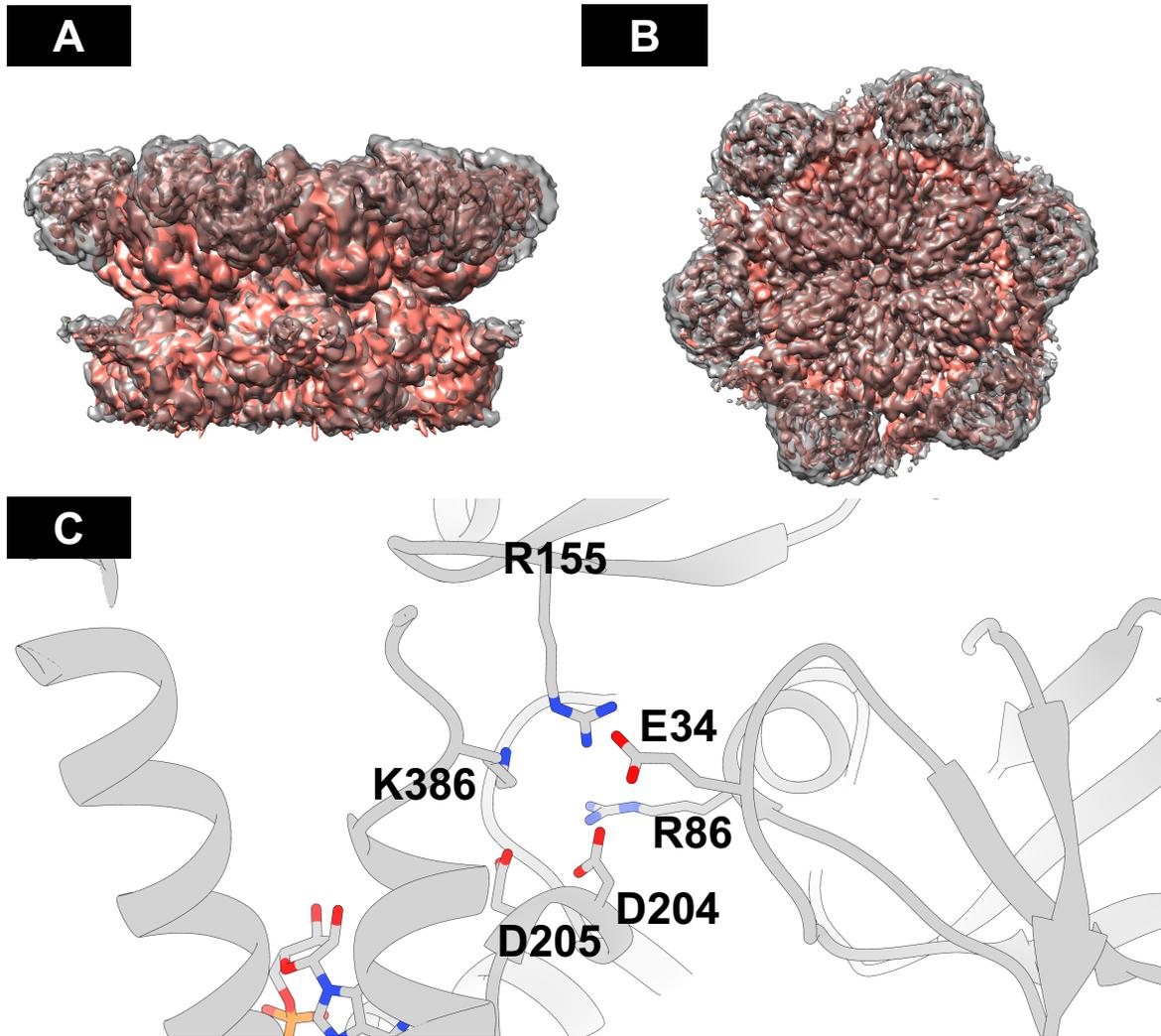
## 2.2.2. Structural Elucidation of p97 Mutants Using Cryo-EM

### 2.2.2.1. p97 Mutant Nucleotide Density



**Figure 2.3 Nucleotide Binding Density of p97 Mutants.** Sharpened cryo-EM density of nucleotide binding density  $<3\text{\AA}$  resolution from nucleotide model. Each mutant nucleotide density agrees with the expected bound nucleotide, showing similar densities to that of the WT maps for ADP-bound [EMDB-3296] and ATP $\gamma$ S-bound [EMDB-3299] nucleotide.

### 2.2.2.2. p97<sup>R155H</sup> ADP-Bound



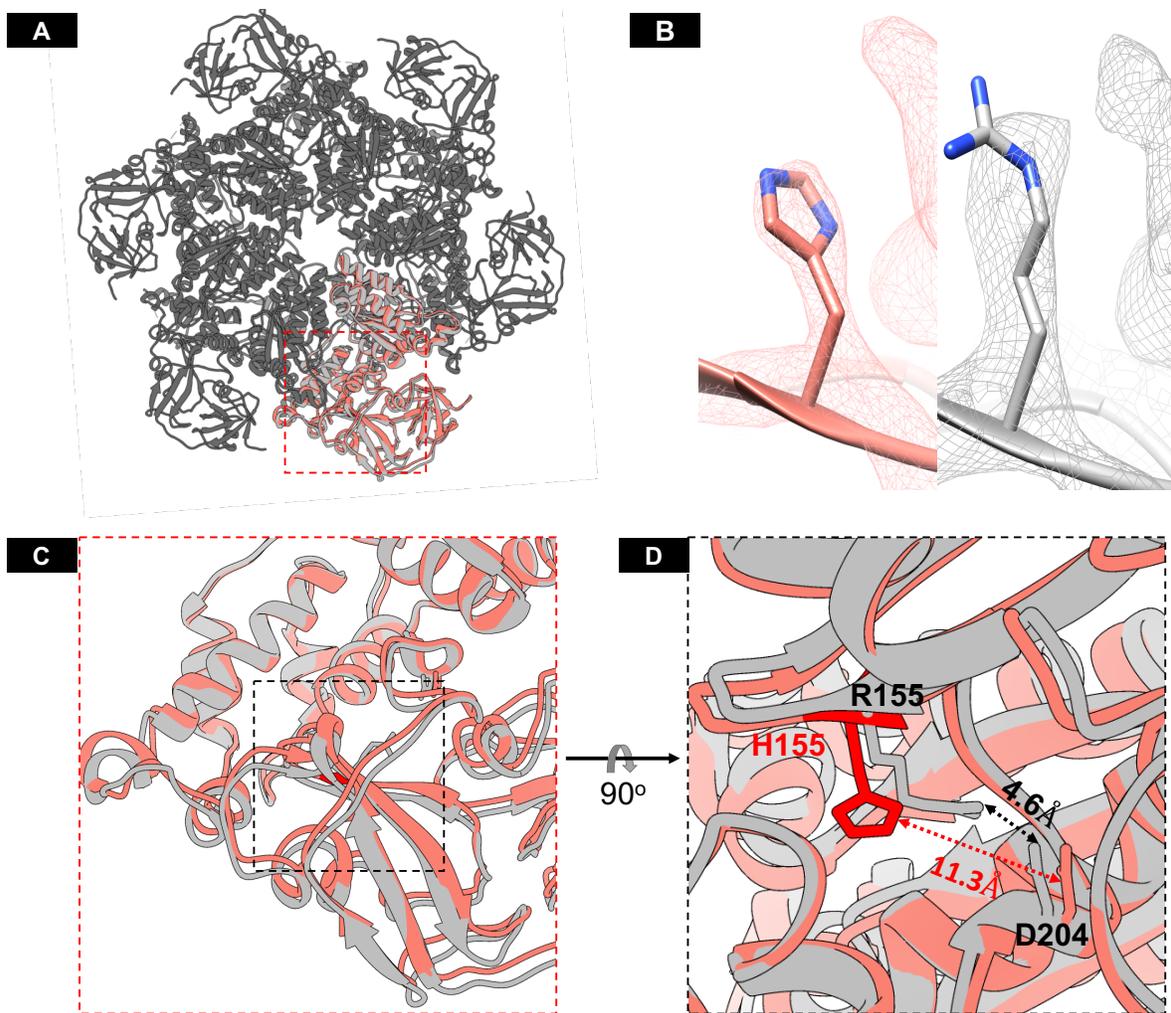
**Figure 2.4: Analysis of Cryo-EM Density Map of ADP-Bound p97<sup>R155H</sup>** **A)** Side view; p97<sup>R155H</sup> density map (brick red) [Current Study] superimposed on p97<sup>WT</sup> density map [EMDB-3296] in ADP-bound “Down” conformation (grey). **B)** Top view of superimposed maps. **C)** Side view of R155 interacting residues between N and D1 domains in WT cryo-EM structure [PDB:5FTK].

Analysis of the cryo-EM map (Figure 2.4 A,B) indicates a decrease in density of the N-domain relative to WT maps, most likely due to an increase in the flexibility of the N-domain motion, which suggests this mutation destabilises the N-domain conformation, as previously observed (Blythe et al., 2019; Schuetz & Kay, 2016; Tang et al., 2010).

Inspection of the structure of WT p97 suggests R155 is important for anchoring the N-domain in the “Down” conformation, due to its proximity to D204 in the N-D1 linker. Similar interactions with D204 were previously predicted as being key for N-domain stability for other MSP-1 disease related mutants in the N domain, R95G and R86A (Tang et al., 2010) (Figure 2.4 C). Superimposition of the N-D1 domains of the p97<sup>WT</sup> and p97<sup>R155H</sup> structures shows that the distance between the carboxyl moiety of D204 and the terminal hydrogen bonding atoms on H155 increases to 11.4 Å compared to the corresponding distance with R155 that is 4.6 Å. (Figure 2.5 D). This indicates a loss in the electrostatic interaction between residues 155 and 204. The loss of this interaction may account for the increased flexibility in the N-domain region of this mutant. This observation appears to agree with previously published NMR studies (Schuetz & Kay, 2016) where the N-domain stability, i.e. exchange between Up and Down states in the ADP-bound conformation, is inversely proportional to the charged/hydrogen bonding nature of the 155 residues.

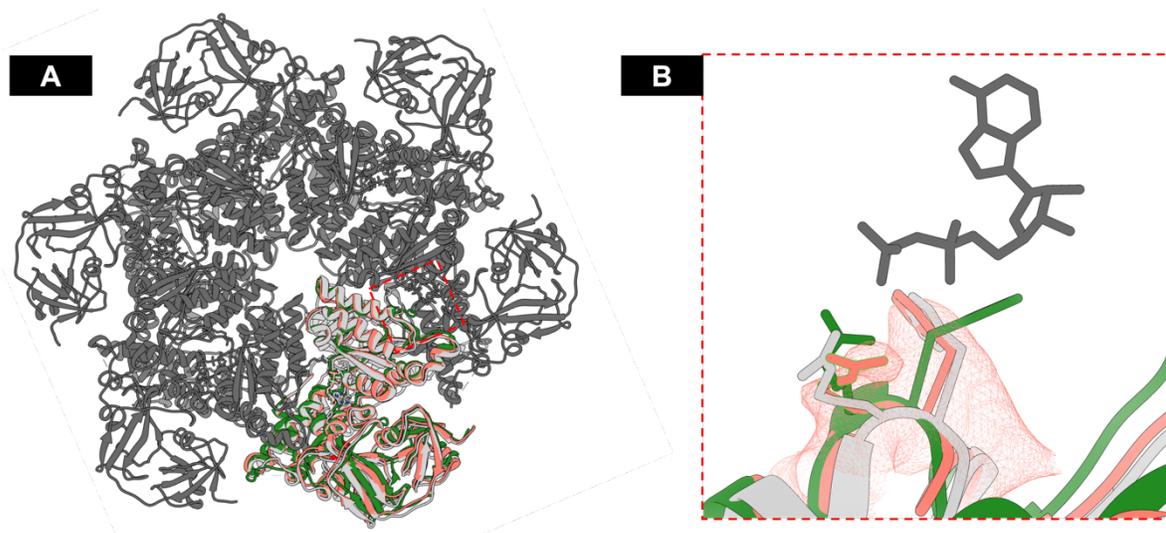
Previous publications (Blythe et al., 2019; Bulfer et al., 2016; Tang et al., 2010; Tang & Xia, 2013) and our current study (Figure 2.2) have shown elevated ATPase activity for the p97<sup>R155H</sup> mutant. However, no significant changes were observed, at the acquired resolution in the immediate vicinity of the nucleotide binding pocket of either the D1, in agreement with previously reported structures (Tang & Xia, 2013) or D2 domains (Figure 2.3), which we present here for the first time. This suggests that the increased flexibility of the N domain conformation alone, may lead to the observed increase in ATPase activity.

The N domain conformation has already been shown to be solely dependent on the nucleotide state of the D1 domain, but not the D2 domain (Banerjee et al., 2016). Therefore, if the N domain is destabilised in the ADP-bound state, it is likely that the mutant will favour either an apo-(nucleotide-free) or ATP-bound state in the D1 domain, both of which have N domain conformations that do not rely on the interaction between R155 and the D1 domain. This is supported by the published observation (Bulfer et al., 2016) that the p97<sup>R155H</sup> mutant has a higher ADP exchange [ $k_{\text{off}}$ ] (~2-3 times) rate than WT in both domains but no significant change is observed in the exchange of ATP between mutant and WT, suggesting that only the ADP-bound conformation is destabilised.



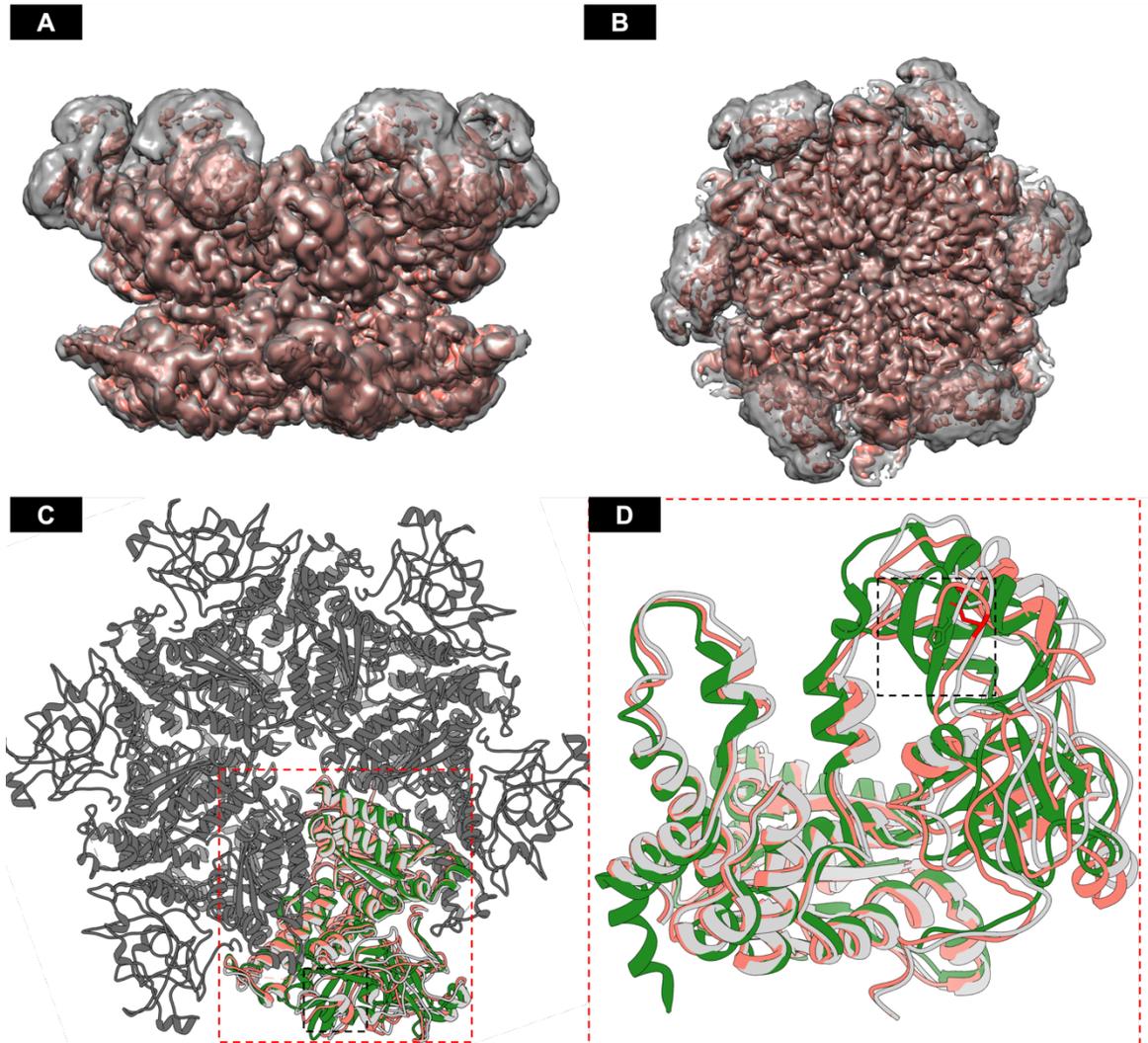
**Figure 2.5 Cryo-EM Structure at 3.0Å resolution of ADP-Bound p97<sup>R155H</sup> Mutant; Intra-Protomer Effects of R155H Mutant.** **A)** Top view of cryo-EM structure of the p97 hexamer presented as a ribbon diagram, showing the full-length WT (grey) [PDB:5FTK] and individual Mutant N-D1 domains (brick red) [Current Study], the mutated residue, His-155 is highlighted in red. **B)** Comparison of map density at the site of mutation using sharpened cryo-EM maps [EMDB:3296]. **C)** Zoom into mutant region. **D)** 90° rotation and further zoom into mutant region, illustrating the loss in a key interaction between residues R155 and D204.

Previously, the ADP-bound state of the N-D1 fragment of p97<sup>R155H</sup> was structurally defined using x-ray crystallography (Tang & Xia, 2013). This paper suggested that the F360 residue was shifted 180° relative to its position in the WT protein as a result of the mutation, in an orientation similar to that seen in the ATP $\gamma$ S-bound WT structure (Figure 2.6). Superimposition of the full-length structure determined in this thesis and that of the published N-D1 fragment seems to indicate that this may not be the case, as F360 in the full-length mutant appears to fit the density map in a similar orientation to that of full-length ADP-bound WT protein.



**Figure 2.6: Comparison Between p97<sup>R155H</sup> Structures in This Thesis and Previously Published Crystallographic Structures** **A)** Top view of cryo-EM structure of the p97 hexamer presented as a ribbon diagram, showing the full-length WT (grey) and adjacent ADP nucleotide (dark grey)[PDB: 5FTK], cryo-EM p97<sup>R155H</sup> mutant N-D1 domain (brick red) structure [Current study] and X-Ray crystallographic mutant N-D1 domain (green) structure [PDB: 4KOD]. **B)** Zoom into region identified by (Tang & Xia, 2013) as a possible explanation for p97 MSP-1 mutant's elevated ATPase activity. Cryo-EM electron density map (red) for F360 and R359 indicates a similar pose to that of ADP-bound p97<sup>WT</sup>.

### 2.2.2.3. p97<sup>R155H</sup> ATP $\gamma$ S-Bound

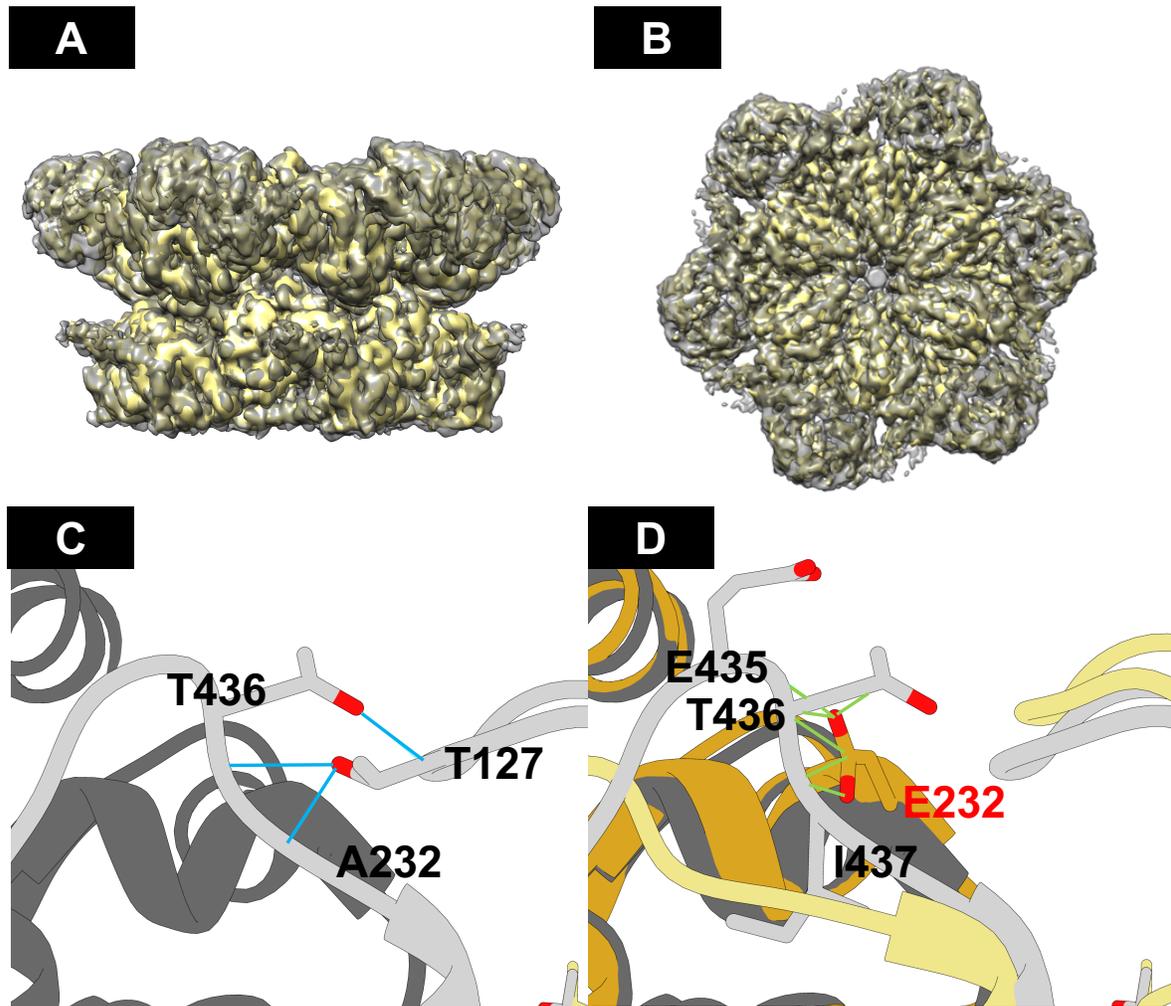


**Figure 2.7: Analysis of Cryo-EM Density Map of ATP $\gamma$ S-Bound p97<sup>R155H</sup> .** **A)** Side view; p97<sup>R155H</sup> density map (brick red) [EMDB-3299] superimposed on p97<sup>WT</sup> density map [EMDB-3299] in ATP $\gamma$ S-bound “Up” conformation (grey). **B)** Top view of superimposed maps. **C)** Top view of cryo-EM structure of the p97 hexamer presented as a ribbon diagram, showing the full-length WT (grey) and individual Mutant N-D1 domains of the cyro-EM structure at 2.5Å resolution of ATP $\gamma$ S -bound p97<sup>R155H</sup> mutant (brick red) and the previously published X-ray crystallography structure [PDB: 3HU3] (green), the mutated residue His-155 is highlighted in red. **D)** Sideview of N-D1 fragment.

Analysis of the cryo-EM map (Figure 2.7 A,B) indicates a decrease in density of the N domain relative to WT maps and similar densities between WT and mutant ATP $\gamma$ S-bound D1 and D2 domains. In the ATP $\gamma$ S-bound state, both R155 and H155 residues are exposed to the solvent and therefore are expected to have little effect on the D1 and D2 domain's nucleotide-binding pocket. This appears to be the case, as there is very little change in the nucleotide-binding region between the mutant and WT (Figure 2.3).

However, the ATP $\gamma$ S-bound p97<sup>R155H</sup> mutant may have an effect on co-factor or substrate binding. It is in close proximity to where several p97 binding co-factors are expected to interact with p97 and could be crucial for the overall stability of p97 substrate complexes (Hänzelmann & Schindelin, 2017). Indeed, the p97<sup>R155H</sup> mutant has been found to have altered responses relative to WT when bound to either p47 or p37 co-factors (Zhang et al., 2015). Comparison between the previously published N-D1 p97<sup>R155H</sup> fragment (Tang et al., 2010) and the full-length mutant in this study indicates good agreement between the structures in the D1 domain (Figure 2.7 C,D). Some differences are observed in the N domain of the structure. However, these are more likely due to the uncertainty in assigning density for the inherently flexible N domain.

#### 2.2.2.4. p97<sup>A232E</sup> ADP-Bound

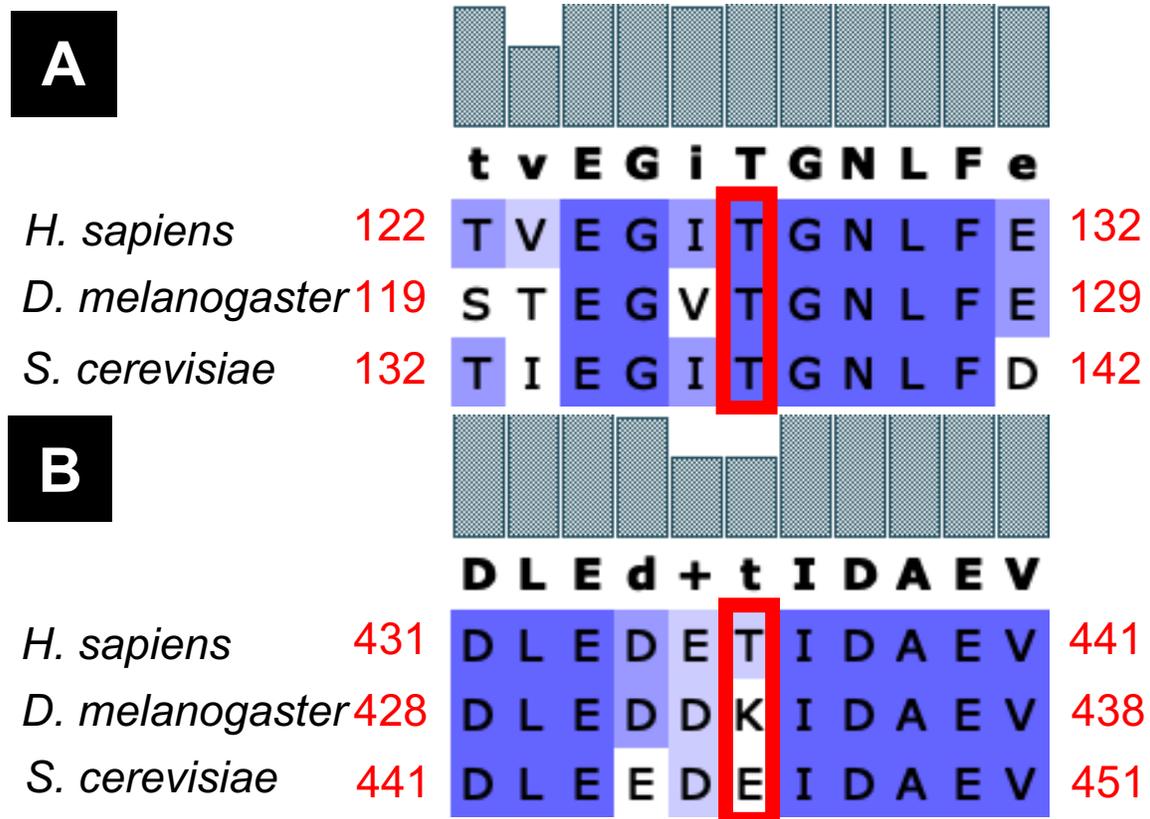


**Figure 2.8: Analysis of Cryo-EM Density Map of ADP-Bound p97<sup>A232E</sup>.** **A)** Side view; p97<sup>A232E</sup> density map (yellow) superimposed on p97<sup>WT</sup> density map in ADP-bound "Down" conformation (grey) [EMDB-3296]. **B)** Top view of superimposed maps. **C)** Side view of A232 interacting residues between N and D1 domains in WT [PDB:5FTK], adjacent protomers labelled grey and dark grey, potential hydrogen bonds modelled in chimera labelled in cyan. **D)** Steric clash between mutant E232 and adjacent WT residues, modelled in chimera (van der Waals overlap of  $\geq 0.6\text{\AA}$ ), labelled in green. Individual Mutant D1-domains labelled in orange/yellow.

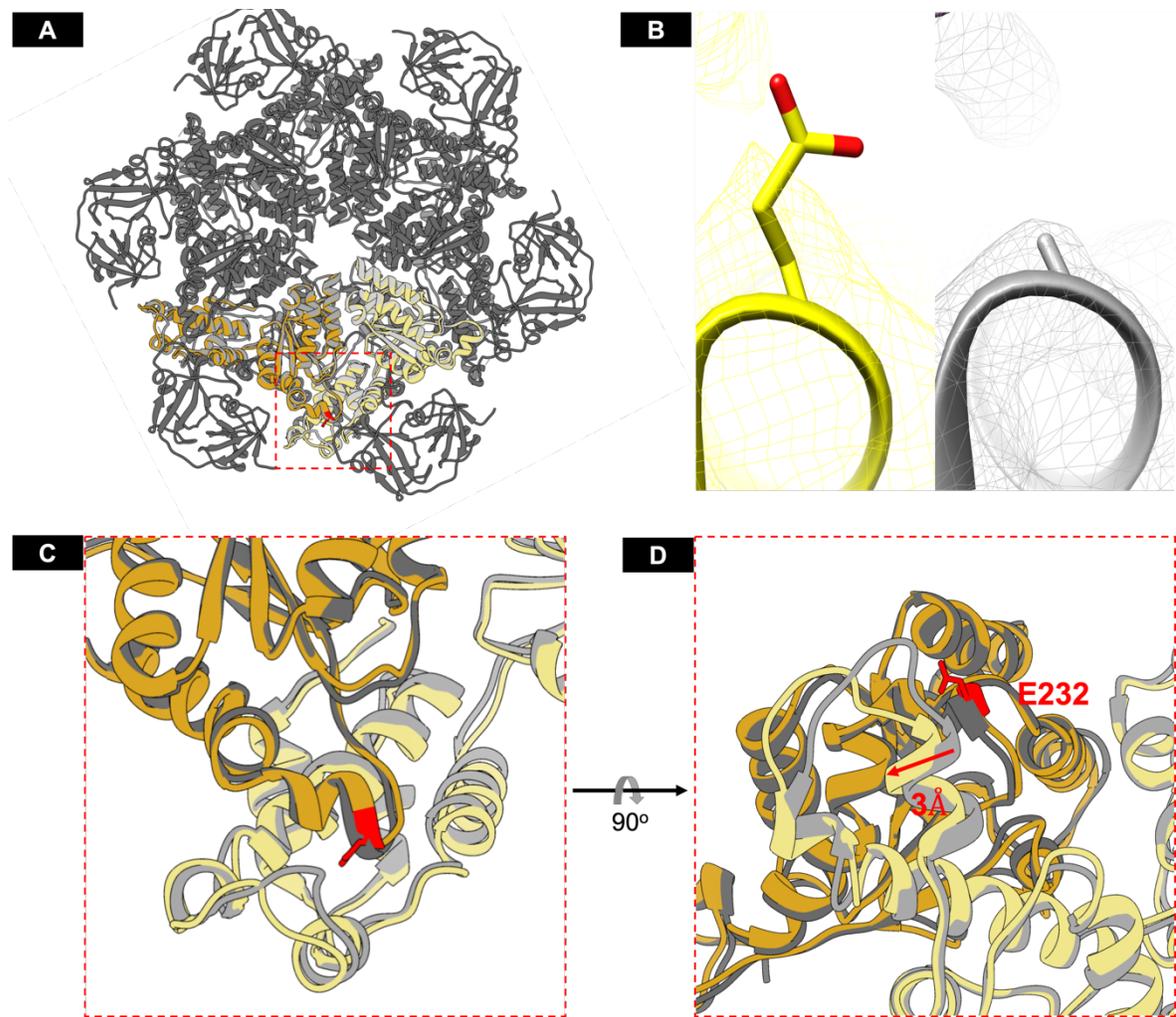
Analysis of the cryo-EM map (Figure 2.8 A,B) indicates a slight decrease in density of the N-domain relative to WT maps. The structure of WT p97 suggests that A232 is located adjacent to two hydrogen bonded threonine residues, located in the N and D1 domains, T127 and T436

respectively (Figure 2.8 C). These residues could be important for stabilising the ADP-bound “Down” conformation of the N domain, and therefore the normal function of p97.

Sequence alignment shows the T127 residue is highly conserved across human, fly and yeast p97 homologs, indicating its importance in the normal function of p97 (Figure 2.9 A). However, T436 does not demonstrate such conservation, as it is substituted for charged lysine and aspartic acid residues in flies and yeast, respectively (Figure 2.9 B). A T127A mutation was identified in an individual with FTD (Shi et al., 2016), further suggesting this hydrogen bonding interaction is indeed important for the normal function of p97.



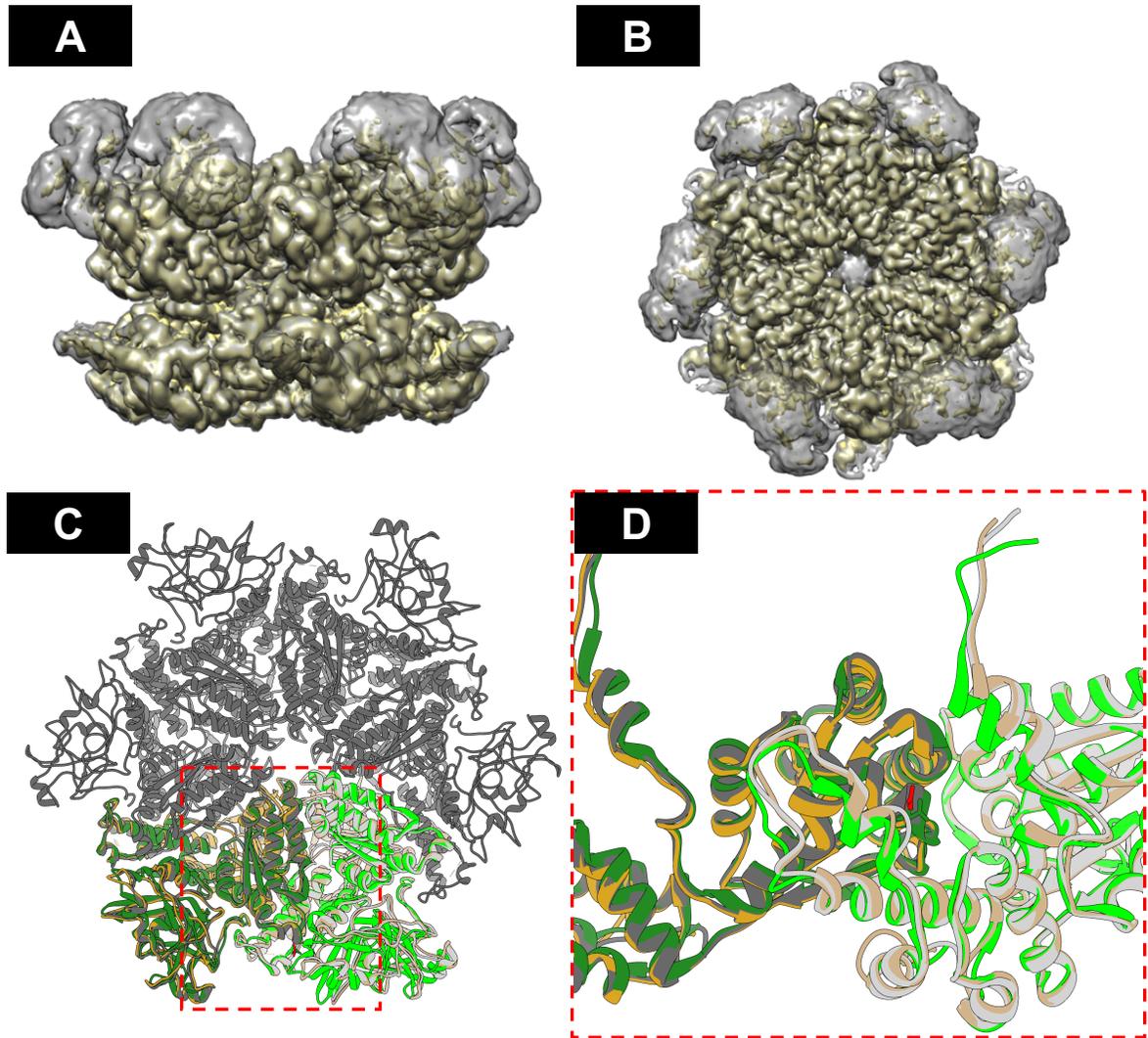
**Figure 2.9: Sequence Alignment of p97 Homologs.** A) T127 and B) T436 are highlighted in red.



**Figure 2.10: Cryo-EM Structure at 3.8Å resolution of ADP-Bound p97<sup>A232E</sup> Mutant; Inter-Protomer Effects of A232E Mutant.** **A)** Top view of cryo-EM structure of the p97 hexamer presented as a ribbon diagram, showing the full-length WT (grey) [PDB:5FTK] and individual Mutant D1-domains (Orange/Yellow) [Current Study], the N domain was removed for clarity, the mutated residue Glu-232 is highlighted in red. **B)** Comparison of map density at the site of mutation using sharpened cryo-EM maps, mutant (yellow) and WT (grey)[EMDB:3296]. **C)** Zoom into mutant region, demonstrating the inter-protomer effects of Glu-232 on adjacent subunit. **D)** 90° rotation of mutant region, illustrating the α-helical shift (p.427-444) in the presence of Glu-232 on the adjacent protomer.

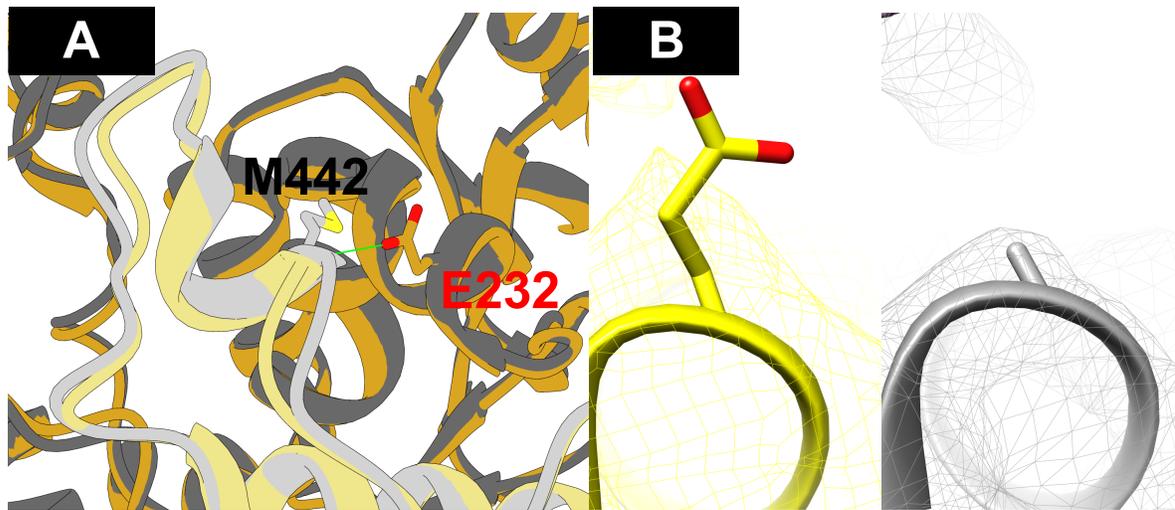
Superimposition of the D1 domains of the p97<sup>WT</sup> and p97<sup>A232E</sup> map density, shows a shift in the loop (p.427-444) density (Figure. 2.10 C) of  $\sim 3\text{\AA}$  relative to WT, consistent with a steric clash between the mutant glutamic acid residue and the adjacent loop (Figure 2.10 D). This shift destroys the hydrogen bonding interactions previously mentioned, decreasing the overall stability of the N domain. Biochemical analyses of full-length p97<sup>A232E</sup> have shown similar behaviour to that of p97<sup>R155H</sup>, i.e. elevated ATPase activity and increased ADP exchange rate. However, as in p97<sup>R155H</sup>, no significant changes were observed at the acquired resolution in the immediate vicinity of the nucleotide binding pocket of either D1 or D2 (Figure 2.3). Therefore, destabilisation of the N domain in ADP-bound complexes, appears to be the principal driver of the elevated ATPase activity in both p97<sup>R155H</sup> and p97<sup>A232E</sup>.

### 2.2.2.5. p97<sup>A232E</sup> ATP $\gamma$ S-Bound



**Figure 2.11: Analysis of Cryo-EM Density Map of ATP $\gamma$ S-Bound p97<sup>A232E</sup>** **A)** Side view; p97<sup>A232E</sup> density map (yellow) [Current Study] superimposed on p97<sup>WT</sup> density map [EMDB-3299] in ATP $\gamma$ S-bound “Up” conformation (grey). **B)** Top view of superimposed maps. **C)** Top view of cryo-EM structure of the p97 hexamer presented as a ribbon diagram, showing the full-length WT (grey) [PDB: 5FTK] and two mutant cyro-EM protomers at 2.6Å resolution of ATP $\gamma$ S -bound p97<sup>A232E</sup> mutant (Orange/Yellow) [Current Study] and the corresponding previously published X-ray crystallography N-D1 fragment [PDB: 4KLN] (dark green/green), the mutated residue Glu-232 is highlighted in red. **D)** Sideview of D1 domain, N domain was removed for clarity.

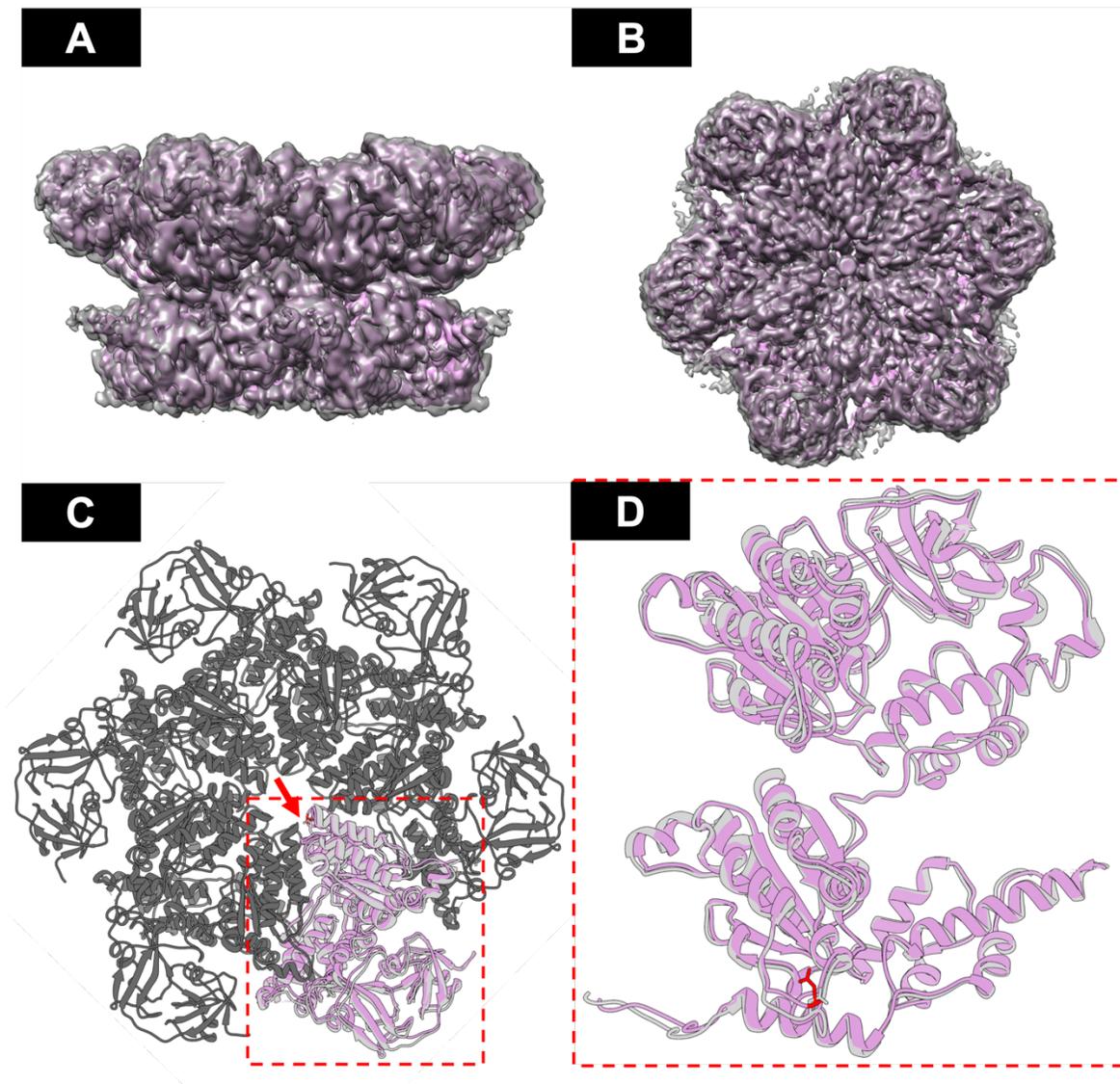
Analysis of the cryo-EM map (Figure 2.11) indicates a decrease in density of the N domain relative to WT maps and similar densities between WT and mutant ATP $\gamma$ S-bound D1 and D2 domains. In the ATP $\gamma$ S-bound state, the E232 residue is exposed to the solvent and therefore is expected to have little effect on the D1 and D2 domain's nucleotide-binding pocket. This appears to be the case, as there is very little change in the nucleotide-binding region between mutant and WT (Figure 2.3).



**Figure 2.12: Cryo-EM Structure at 2.6Å resolution of ATP $\gamma$ S-bound p97<sup>A232E</sup> Mutant; Inter-Protomer Effects of A232E Mutant. A) Zoom into mutant region, steric clash between mutant E232 and WT M442 residues modelled in chimera, labelled in green. B) Comparison of map density at the site of mutation using sharpened cryo-EM maps, mutant (yellow) and WT (grey) [EMDB:3299].**

A slight local deviation near the mutant residue appears to be as a result of a steric clash between E232 and M442 (Figure 2.12 A), however this does not cause significant secondary structural changes as there is good agreement across the rest of the structure relative to WT. This local deviation appears to be more pronounced in the previously published X-ray crystallographic structure (Figure 2.11 D) (Tang & Xia, 2013). As the last assigned amino acid in the x-ray structure of this mutant is less than 20 amino acids from this loop, the lack of D2 protein density may account for the greater flexibility observed in that region.

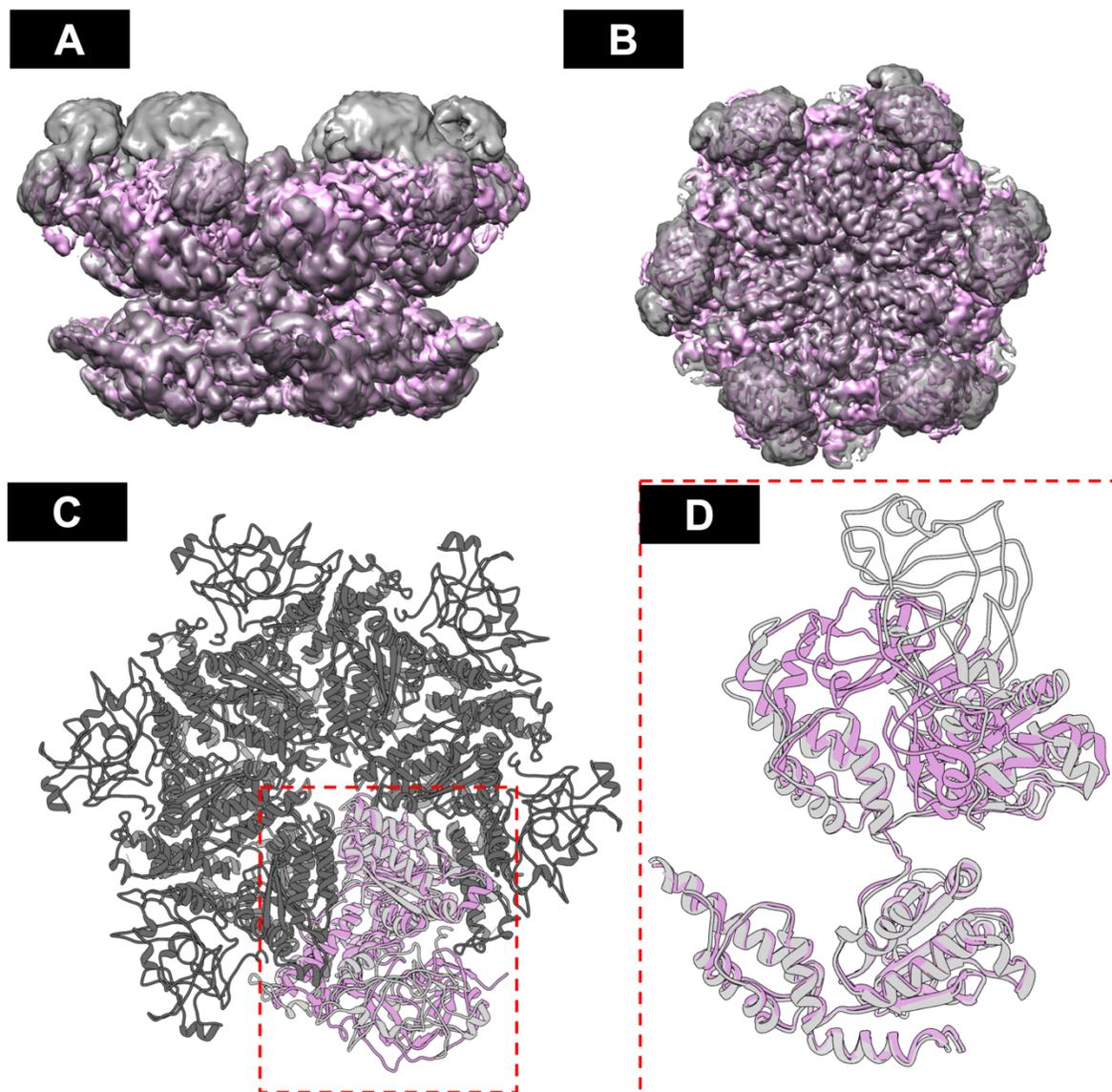
### 2.2.2.6. p97<sup>D592N</sup> ADP-Bound



**Figure 2.13: Analysis of Cryo-EM Density Map of ADP-Bound p97<sup>D592N</sup>.** **A)** Side view; p97<sup>D592N</sup> density map (purple) [Current Study] superimposed on p97<sup>WT</sup> density map [EMDB-3296] ADP-bound “Down” conformation (grey). **B)** Top view of superimposed maps. **C)** Top view of cryo-EM structure of the p97 hexamer presented as a ribbon diagram, showing the full-length WT (grey) [PDB: 5FTN] and full-length mutant structure of the cyro-EM structure at 2.9Å resolution of ADP-bound p97<sup>D592N</sup> mutant (purple) [Current Study], the mutated residue Asn-592 is highlighted in red. **D)** Sideview from inside pore (indicated by red arrow in C) of a full-length protomer.

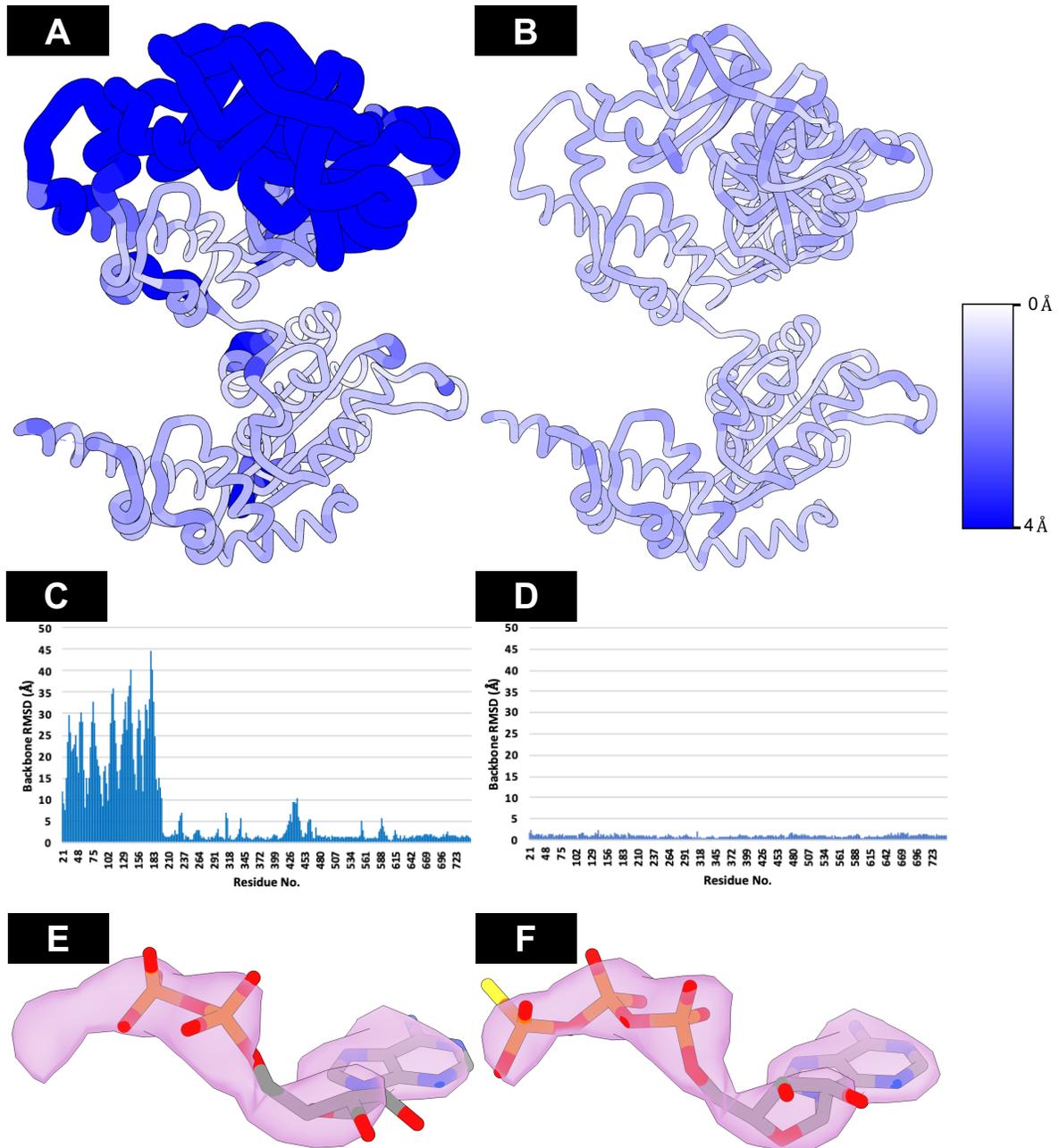
Analysis of the cryo-EM density map of the p97<sup>D592N</sup> mutant (Figure 2.13 A,B) indicates that there are no significant changes as compared to the wild-type protein. Superimposition of the cryo-EM structures (Figure 2.13 C,D) demonstrates excellent correspondence in the atomic models, indicating the D592N mutation appears to have little effect on the structure of p97 in the ADP-bound state at ~ 3 Å resolution.

### 2.2.2.7. p97<sup>D592N</sup> ATP $\gamma$ S-Bound



**Figure 2.14: Analysis of Cryo-EM density map of ATP $\gamma$ S-bound p97<sup>D592N</sup>.** **A)** Side view; p97<sup>D592N</sup> density map (purple) [Current Study] superimposed on p97<sup>WT</sup> density map [EMDB-3299] in ATP $\gamma$ S-bound “Up” conformation (grey). **B)** Top view of superimposed maps. **C)** Top view of cryo-EM structure of the p97 hexamer presented as a ribbon diagram, showing the full-length WT (grey) [PDB: 5FTK] and full-length mutant structure of the cyro-EM structure at 2.5Å resolution of ATP $\gamma$ S-bound p97<sup>D592N</sup> mutant (purple) [Current Study]. **D)** Sideview of full-length p97, shift in N domain conformation indicated by red arrow.

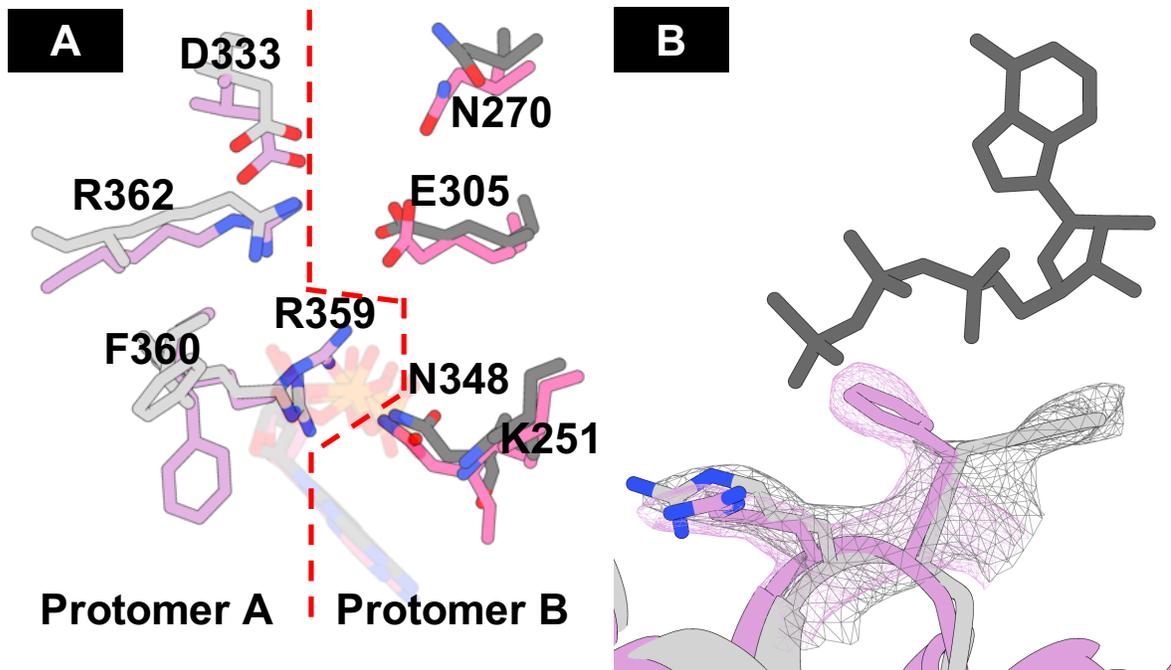
In the ATP $\gamma$ S-bound state, the density map of the p97<sup>D592N</sup> mutant (Figure 2.14 A,B) indicates a similar density in all domains relative to WT maps in the D2 domain and sufficient density in both D1 and N domains. However, the N and D1 domain density are in the “Down” conformation rather than the “Up” conformation typical for ATP $\gamma$ S-bound p97 (Figure 2.14 D). Superimposition of the D2 domains of the cryo-EM structures of ATP $\gamma$ S-bound p97<sup>WT</sup> [PDB: 5FTK] and p97<sup>D592N</sup> both the general conformation and nucleotide binding pocket (Figure 2.3) are similar to that of ATP $\gamma$ S-bound p97<sup>WT</sup>. Superimposition of the cryo-EM structures of ATP $\gamma$ S-bound p97<sup>WT</sup> [PDB: 5FTK] and p97<sup>D592N</sup> indicate both the general conformation and nucleotide-binding pocket of the D2 domain (Figure 2.3) are closely comparable.



**Figure 2.15: Cryo-EM Structure at 2.5Å resolution of ATP $\gamma$ S-Bound p97<sup>D592N</sup> Mutant; Determination of D1 Nucleotide State.** A) p97<sup>D592N</sup> protomer is displayed showing backbone RMSD from D1D2-ATP $\gamma$ S-bound p97<sup>WT</sup> [PDB:5FTK] and B) D1-ADP and D2-ATP $\gamma$ S bound p97<sup>WT</sup> [PDB:5FTM]. C – D) Backbone RMSD plot for A) and B), respectively. D) ADP fit in zoned nucleotide density in D1 domain of p97<sup>D592N</sup> cryo-EM sharpened maps and E) ATP $\gamma$ S fit in the same density. White-Blue bar represents RMSD between WT and mutant backbone models from 0 to 4 Å.

The striking result from these studies is while p97<sup>WT</sup> and p97<sup>D592N</sup> mutants display similar structures in the ADP-bound state, they show a marked difference in the presence of ATP $\gamma$ S. In the mutant, ATP $\gamma$ S binding does not result in transition of the N-domain to the up-conformation despite clear evidence for bound ATP $\gamma$ S in both D1 and D2 nucleotide-binding pockets (Figure 2.15)

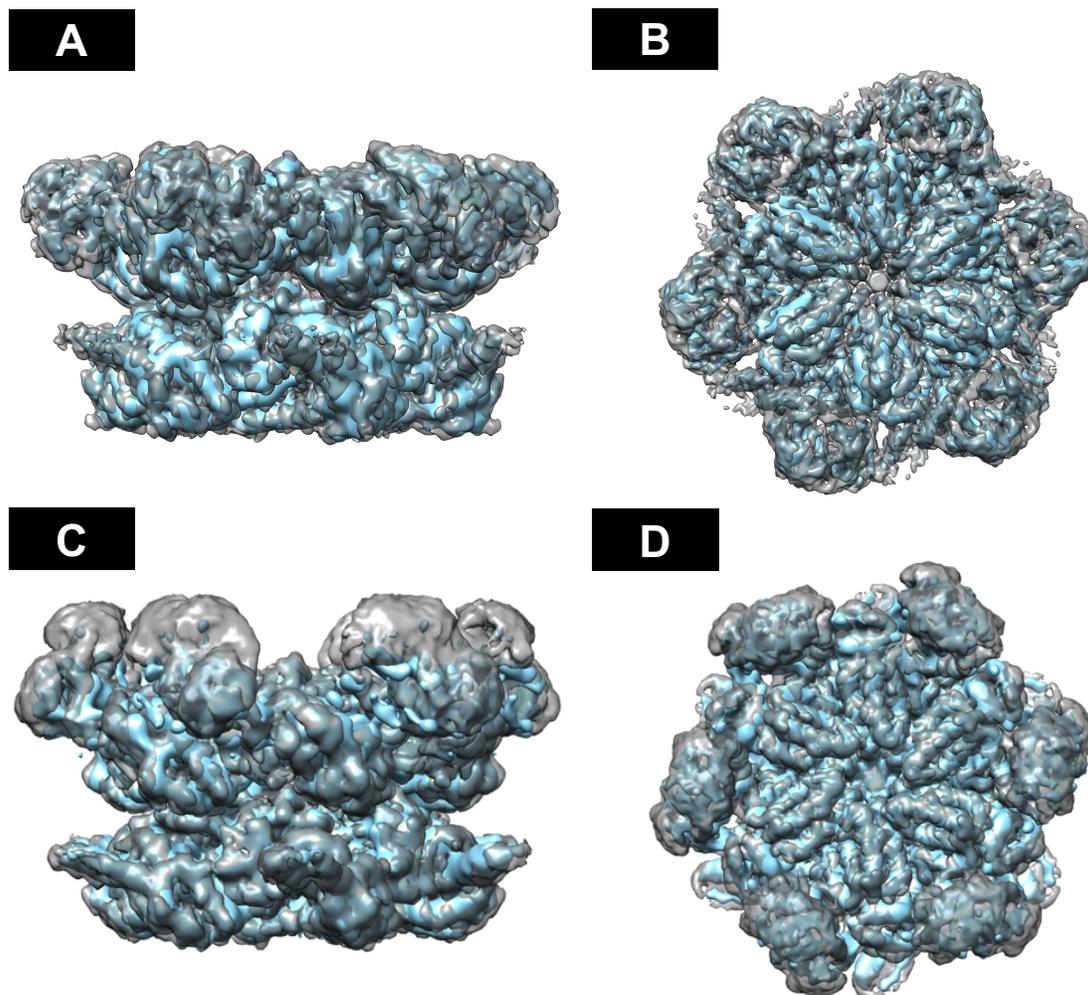
Since the pore loop is very flexible, it is difficult to determine the precise location of the N592 in the mutant density map, but the effects of the mutation can be discerned by the allosteric effects on the quaternary structure of the N and D1 domains. Analysis of the major residues in the nucleotide binding pocket of the D1 domain shows that there is no significant change relative to ATP $\gamma$ S-bound p97<sup>WT</sup>. However, the F360 residue, was identified as a potential sensor for nucleotide status as stated before (Tang & Xia, 2013), is in the ADP-bound conformation (Figure 2.16 A,B), suggesting that the orientation of the F360 residue may be dependent on N domain conformation (Up vs. Down) rather than nucleotide state (ADP vs. ATP bound).



**Figure 2.16: Comparison Between Key Residues in p97<sup>D592N</sup> and ATP $\gamma$ S-Bound p97<sup>WT</sup> D1 Nucleotide Binding Site.** **A)** Stick representation of a superimposition of the D1 ATP $\gamma$ S binding sites of p97<sup>WT</sup> (grey) and p97<sup>D592N</sup> (purple). **B)** Sharpened cryo-EM electron density map for ATP $\gamma$ S-bound p97<sup>WT</sup> (grey) [EMDB:3299] and p97<sup>D592N</sup> (purple); F360 indicates a similar pose to that of ADP-bound p97<sup>WT</sup> and R359 is in a similar pose to that of the ATP $\gamma$ S-bound p97<sup>WT</sup>.

While p97<sup>D592N</sup> has an elevated ATPase activity relative to WT (Figure 2.2), it has the lowest activity relative to the other mutants in this study. This relatively low activity may be due to a lack of flexibility in the N-domain region to react to ATP binding in the D1 domain. However, it has been shown before that nucleotide binding to D1 hydrolysis-deficient mutants actually led to an increase in overall ATPase activity (Tang & Xia, 2013). This suggests that while p97<sup>D592N</sup> may abolish N-D1 interdomain communication it may not have any effect on the D1-D2 interdomain communication and may act in a similar manner to the D1 hydrolysis-deficient mutants. Further biochemical analysis is required to determine the mechanistic basis for the reduced ATP activity of the D1 domain in p97<sup>D592N</sup> mutants.

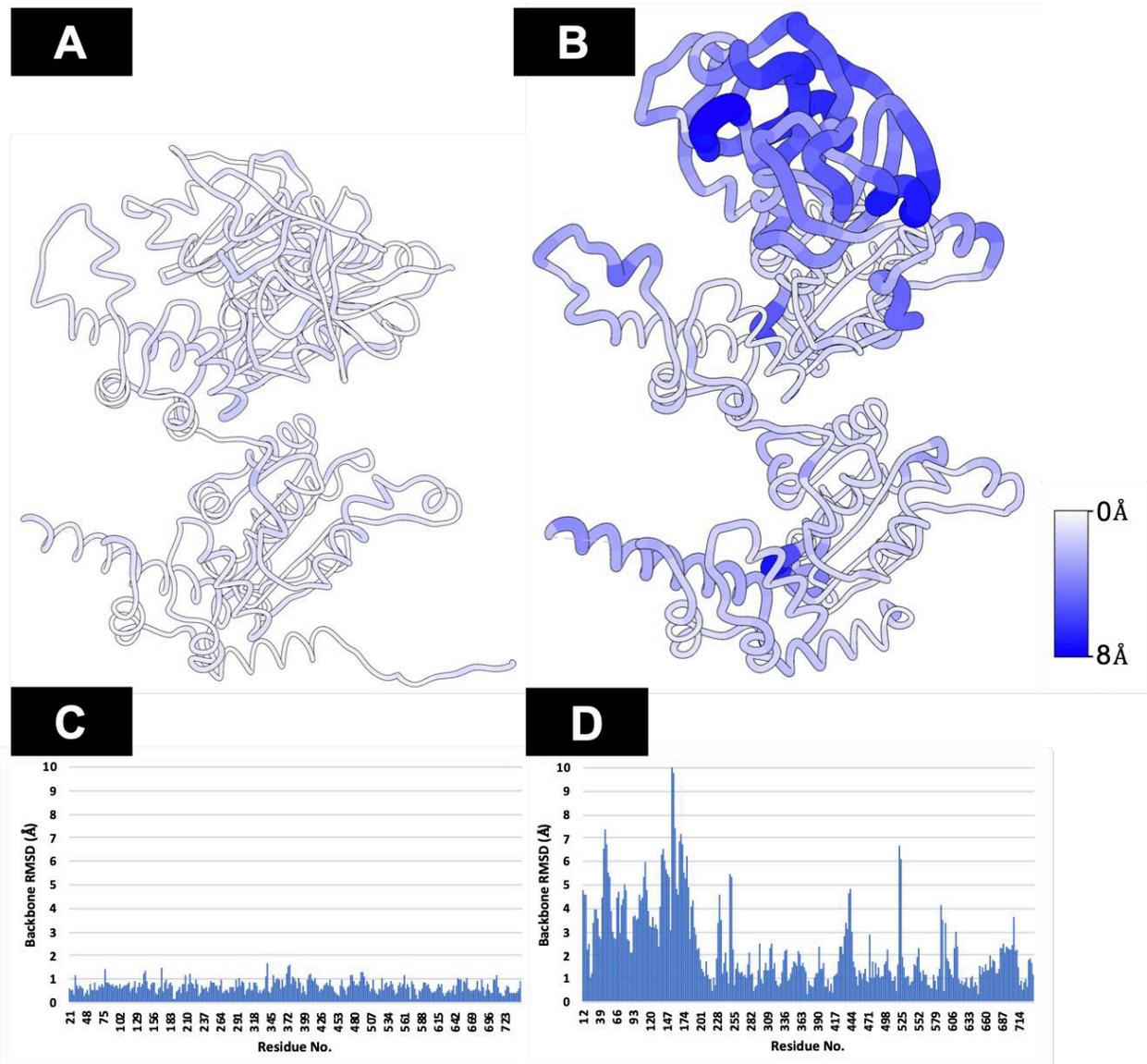
### 2.2.2.8. p97<sup>E470D</sup>Cryo-EM Maps



**Figure 2.17: Analysis of Cryo-EM Density Map of Nucleotide Bound p97<sup>E470D</sup>** **A)** Side view; p97<sup>E470D</sup> density map (cyan) [Current Study] superimposed on p97<sup>WT</sup> density map [EMDB-3296] in ADP-bound “Down” conformation (grey). **B)** Top view of superimposed maps. **C)** Side view; p97<sup>E470D</sup> density map (cyan) [Current Study] superimposed on p97<sup>WT</sup> density map [EMDB-3299] in ATP $\gamma$ S-bound “Up” conformation (grey). **D)** Top view of superimposed maps.

Analysis of the cryo-EM map of ADP-bound p97<sup>E470D</sup> (Figure 2.17 A,B) indicates a similar density in all domains relative to WT maps. At the current resolution of 4.2Å, it is difficult to precisely position all the side chain residues. However, fitting of the secondary structure indicates a near native structure, suggesting p97<sup>E470D</sup> has minimal, if any effect on the structure

of p97 in the ADP-bound state. RMSD backbone analysis demonstrates the difference between WT and mutant is  $<2 \text{ \AA}$ , well within the noise levels of the cryo-EM density (Figure 2.18 A).



**Figure 2.18: RMSD Analysis of p97<sup>E470D</sup> Mutants.** **A)** ADP-bound p97<sup>E470D</sup> protomer is displayed showing backbone RMSD from ADP-bound p97<sup>WT</sup> [PDB:5FTN]. **B)** ATP $\gamma$ S-bound p97<sup>E470D</sup> protomer is displayed showing backbone RMSD from ATP $\gamma$ S-bound p97<sup>WT</sup> [PDB:5FTK]. **C – D)** Backbone RMSD plot for **A)** and **B)**, respectively. White-Blue bar represents RMSD between WT and mutant backbone models from 0 to 8  $\text{\AA}$ .

Analysis of the cryo-EM map of ATP $\gamma$ S-bound p97<sup>E470D</sup> (Figure 2.17 C,D) indicates a decrease in density of the N domain relative to WT maps and similar amount of density between WT and mutant ATP $\gamma$ S-bound D1 and D2 domains. Again, due to the relatively low resolution of this structure (4.6Å), only the secondary structure was fitted to the map. RMSD analysis demonstrates the difference between WT and mutant backbones (Figure 2.18 B).

Despite the relatively low resolution of the ATP $\gamma$ S-bound p97<sup>E470D</sup> cryo-EM map, RMSD analysis reveals multiple areas of the fitted model that appear to deviate from WT. Due to the reduced resolution, it is difficult to say with certainty whether these differences arose from the mutant or indeed if they are significant enough to cause the effect observed in biochemical assays but it appears that there is significant flexibility in the N domain and in the residues closest to adjacent protomers, suggesting that the effect of p97<sup>E470D</sup> is on the ATP $\gamma$ S-bound structure primarily.

#### 2.2.2.9. General Observations

It has been shown that introduction of Walker A (nucleotide-binding defective) mutants into either D1 or D2 domains (K251T and K524T, respectively) in p97<sup>WT</sup> leads to an almost complete loss of ATPase activity in both domains (Tang & Xia, 2013), indicating the importance of nucleotide binding in both domains for normal p97 ATPase activity. Also, introduction of Walker B (ATP-hydrolysis deficient) mutants into the D1 domain (E305Q) in p97<sup>WT</sup> leads to a slight but significant increase in overall ATPase activity and a similar mutant in the D2 domains (E578Q) almost completely abolishes the activity of p97. Similar observations were made for MSP-1 mutants (Niwa et al., 2012; Tang & Xia, 2013), suggesting that while the ATPase activity of p97 is dependent on nucleotide binding to both D1 and D2 domains, ATP hydrolysis in the D1 domain is not necessary for p97 ATPase activity in either mutant or WT. This indicates the primary function of the D1 domain is to facilitate D2 ATPase activity through binding ATP, which leads to a conformational change in the N domain and therefore increasing the overall rate of D2 ATPase activity.

As we have observed very little change in the nucleotide-binding sites in either D1 and D2 domains of the MSP-1 mutants, the previously observed increase in ADP off rate and ATPase activity of the mutants studied may not be due to a direct conformational change in the nucleotide-binding domain. Instead, these changes in activity are likely because of the destabilisation of the N domain in the ADP-bound conformation. This destabilisation reduces the population of ADP-bound p97 complexes in mutant cells relative to WT. Therefore, a higher rate of ADP release will lead to the quicker turnover of ATP hydrolysis observed, due to the increased availability of nucleotide-free p97 relative to WT in solution. This observation is in agreement with previously published NMR studies of MSP-1 mutants, whereby MSP-1 “mutations shift the ADP-bound form of the enzyme towards an ATP-like state in a manner that correlates with disease severity” (Schuetz & Kay, 2016).

As for the p97<sup>D592N</sup> and p97<sup>E470D</sup> mutants, their effect appears to occur in the ATP $\gamma$ S-bound conformations. p97<sup>D592N</sup> appears to prevent communication of D1 nucleotide state between the N and D1 domains, possibly through the loss of electrostatic interactions within the central pore translated up towards the D1 domain from ATP binding to D2. As a large shift of  $\sim 10\text{\AA}$  in the D2 pore occurs upon ATP binding to the D2 domain (Banerjee et al., 2016), it is possible that this shift is not communicated sufficiently to the D1 pore loop through the loss of the aspartic acid preventing the necessary translation in D1 to sense the binding of ATP necessary for the subsequent N domain shift. However, the communication between D1 and D2 through the linker may still remain intact explaining the increased ATPase activity, further analysis is need of the relative activities and nucleotide affinities of the different domains in this mutant to confirm its effect on interdomain communication. p97<sup>E470D</sup> appears to disrupt the communication between D1 and D2 in the ATP-bound state primarily, however it remains unclear as to the exact nature of these interactions which causes the observed increases in ATPase activity.

## 2.3. Methodology

### 2.3.1. Preparation of Chemically Competent Bacteria

A stab from bacterial frozen stocks was added to 5 ml of Luria Broth (LB) (See Table A.1 and A.2 for Media Recipes and Bacterial Strains, respectively.) and incubated overnight at 37°C with shaking. 2 ml of the overnight culture was mixed with 48 ml of LB broth and incubated at 37°C with shaking to an OD<sub>600</sub> of ~0.6-0.7 (~2 hrs). The culture was centrifuged at 1000g for 10 min at 4°C. The cell pellet was resuspended in 5 ml of ice-cold 50 mM CaCl<sub>2</sub> and incubated in ice-water for 20 mins. The suspension was centrifuged at 1000g for 10 mins at 4°C. The cell pellet was resuspended in 2ml of ice-cold 50 mM CaCl<sub>2</sub>. 10ul of 50% (v/v) sterile-filtered (0.22µm) Glycerol was added to 40ul of cell suspension in -80°C tubes and the suspension was stored at -80°C.

### 2.3.2. Site-Directed Mutagenesis (SDM)

The p97 gene was codon optimised for bacterial expression and inserted into an expression vector with the antibiotic kanamycin as the selection agent (see Appendix Seq.1+2, for full nucleotide and amino acid sequence, respectively). Multiple nucleotides were substituted at areas identified in key diseases: c.464G>A (p.Arg155His); c.695C>A (p.Ala232Glu); c.1410G>C (p.Glu470Asp); c.1774G>A (p.Asp592Asn) (See Table A.3, for full list of primer sequences). Each mutation was introduced using the Q5<sup>®</sup> Site-Directed Mutagenesis Kit from New England Biolabs (NEB). A reaction mixture of 12.5 µl of Q5 Hot Start High-Fidelity 2X Master Mix, 1.25 µl of Forward Primer [10 µM], 1.25 µl of Reverse Primer [10 µM], 1 µl of Plasmid DNA [~100 ng] and 9.0 µl of Nuclease-Free water.

The mixture was thermocycled according to Table 2.1:

| Step                 | Temperature (°C) | Time (Sec) |
|----------------------|------------------|------------|
| Initial Denaturation | 98               | 30         |
| 25 Cycles            | 98               | 10         |
|                      | 65               | 30         |
|                      | 72               | 180        |
| Final Extension      | 72               | 120        |
| Hold                 | 4                | ∞          |

**Table 2.1: PCR Thermocycling Protocol.**

After thermocycling the removal of the template plasmid and circularization of the mutant was performed by mixing 1 µl of thermocycled product, 5 µl of 2X KLD Reaction Buffer, 1 µl of 10X KLD Enzyme Mix and 3 µl of Nuclease-Free water at room temperature for 5 minutes. The resulting plasmid was transformed into DH5α chemically competent *E.coli* (NEB #C2987) by adding 5µl of isolated product to a thawed 50 µl aliquot of *E.coli* prepared previously, as described in section 2.4.1.1 on ice for 30 minutes. The sample was heat shocked at 42°C for 30 seconds and placed on ice for a further 5 minutes. 950 µl of room temperature SOC media was added to the sample and incubated at 37°C for 60 minutes with shaking at 250 rpm. 100µl of the sample were spread onto a LB selection plate (LB-Kan) with Kanamycin (Kan) [50 µg/ml] and incubated overnight at 37°C.

### 2.3.3. Plasmid Isolation and Sequencing

A colony was picked from each selection plate and grown in 5 ml LB -Kan broth at 37°C overnight with shaking at 250 rpm. The culture was spun down at 3000rpm for 10 minutes and the pellet was isolated. The DNA was isolated using the PureLink® Quick Plasmid Miniprep Kit from Invitrogen. The pellet was resuspended in 250µl of R3 resuspension buffer, followed by 250 µl of L7 lysis buffer, the mixture was incubated at RT for 5 mins. The cell contents were precipitated by adding 350µl of N4 precipitation buffer and homogenised.

The lysate was spun down at 12,000g for 10 minutes. The supernatant was loaded onto a spin column and centrifuged at 12,000g for 1 minute. The flow-through was discarded and the column was washed with 500 µl W10 wash buffer and incubated for 1 minute at RT, spun down at 12,000g for 1 minute. The flow-through was discarded and the column was washed with 700 µl w9 wash buffer and incubated for 1 minute at RT, spun down at 12,000g for 1 minute. The column was spun down at 12,000g a further minute to ensure the column was dried of ethanol and transferred to a collection tube. 75 µl of nuclease free water, preincubated at 70°C, was added to the column and incubated for 1 minute at RT. The column was centrifuged at 12,000g for 2 minutes to collect the isolated plasmid, the concentration was measured using a nanodrop. 4 aliquots with 9µl of plasmid each were prepared and mixed with 1 µl of 10x TE buffer. All aliquots were frozen at -80°C.

### 2.3.4. Sanger Sequencing of Mutant Plasmids

Plasmids extracted from both DH5α and BL21(DE3) cells were amplified with PCR using the same conditions previously described in section 2.4.1.2. Primer set 5 or 6 (see Table A.3 for sequence) were used for the amplification, depending on the respective positions of the mutant for sequencing in the open reading frame (ORF). The reaction mixture was run on a 2% (w/v) agarose gel in TAE buffer; 40 mM Tris base, 20mM Acetic Acid and 1mM EDTA (pH 8.0). The amplified bands were cut out of the agarose gel and the DNA was isolated using Monarch® DNA Gel Extraction Kit (NEB #T1020L). The section of gel containing the band was placed in a 1.5 ml Eppendorf tube and weighed. 4 volumes of dissolving buffer was added to the gel

and the sample was incubated at 55°C, vortexing periodically until the gel had completely dissolved. The collection column was inserted into the collection tube and the sample was loaded onto the column.

The Eppendorf was spun down at max g for 1 min, the flow-through was discarded and 200µl of DNA wash buffer was added and the Eppendorf was spun down again for 1 min and the flow-through discarded. This step was repeated and the column was transferred to a clean 1.5ml Eppendorf tube, 10µl of pre-warmed (55°C) ddH<sub>2</sub>O was added to the column and incubated for 2 mins. The column was spun down at max g for 2 mins. The eluent was mixed with 5 µl of 5 µM solution of the appropriate reverse primer from primer set 5 or 6. The 15µl solution was submitted for sanger sequencing and the results analysed using Unipro UGENE software (Rose et al., 2019).

#### 2.3.5. Protein Expression and Purification

50 µl of chemically competent BL21(DE3) (Thermo Fisher #C6000-03), prepared previously as described in section 2.4.1.1., was thawed on ice and 10 ng of mutant plasmid DNA was added to the cells. The cells were incubated for 30 minutes on ice. The sample was heat shocked at 42°C for 30 seconds. The sample was placed on ice and 250 µl of preincubated [37°C] Super Optimal broth with Catabolite repression (SOC) media was added to the sample and shaken at 37°C for 1 hour at 250 rpm. 100 µl of sample was plated onto LB-Kan selection plates and incubated at 37°C overnight. A colony was picked and added to 5 ml of LB-Kan selection broth and incubated at 37°C with shaking at 250 rpm overnight.

After overnight incubation the saturated solution was diluted 1:50 in fresh selection broth and incubated to an OD<sub>600</sub> of ~0.7. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM to induce protein expression, the culture was grown at 27°C overnight. After overnight expression, the bacterial sample was spun down at 3000 rpm for 5mins, the supernatant was removed and the pellet is resuspended in 20 ml of lysis buffer; 50 mM Tris-HCL pH:8, 300 mM NaCl, 0.5 mM β-Mercaptoethanol (BME) and 1x Protease Inhibitor Cocktail tablet (Thermo Fisher #A32963). The resuspension was placed on ice and inserted into a tip sonicator. The suspension was sonicated with 20sec bursts with 10sec spaces

between bursts for a total of 5 min sonication time, this was repeated until the solution appeared slightly darker and less viscous. The suspension was spun down at 3000 rpm to remove the cellular debris for 5 mins. The supernatant extracted and spun down at 12,000g to remove any remaining debris and aggregates, the resulting supernatant was combined and imidazole was added to a concentration of 20 mM.

The sample was added in 5ml amounts to a 5 ml His-Pur™ Ni-NTA column (Thermo Fisher #88221) and incubated at RT for 5-10 mins each until all the protein solution had passed through the resin. The resin was washed continuously for 5mins with wash buffer; 50 mM Tris-HCL pH:8, 150 mM NaCl, 20 mM Imidazole, 5 mM MgCl<sub>2</sub> and 0.5 mM BME. 5ml of elution buffer (Wash buffer with 300 mM Imidazole) was added to the column in 1 ml amounts, each 1ml aliquot was collected in a separate tube and placed on ice. An SDS-PAGE was run on the eluent to determine the amount and efficiency of purification, only aliquots with a single visible p97 band were combined. The combined samples were spun down in a concentrator (50kDa MWCO) at 3000 rpm for 15mins, the remaining concentrate was diluted 10-fold in protein storage buffer; 50mM Tris-HCL pH:8, 150 mM NaCl, 5 mM MgCl<sub>2</sub> and spun down again at 3000 rpm for 15mins. This process was repeated twice more. The sample was run on Native-PAGE against a standard and on a Nanodrop to estimate the amount and quality of the protein, only protein samples with a characteristic p97 band were accepted for storage. Several aliquots were prepared in 10ul volumes of 1 and 10mg/ml of protein, they were frozen and stored at -80°C for collection.

### 2.3.6. ATPase Assays

Protein concentrations were first normalised using Pierce BCA Protein Assay kit (Thermo Fischer #23225), a standard series was prepared with the BSA provided (2 mg/ml - 25 µg/µl). 25 µl of each of the standard and unknown sample were added to a 96-well plate in triplicate, 200 µl of working reagent was added to each well; prepared according to manufacturer's guidelines. The plate was covered and incubated at 37°C for 30 mins, the plate was cooled to RT and the absorbance at 562nm was read for each well.

The ATPase assays were performed using the ADP-Glo™ Kinase Assay (Promega #V9101), as previously described in (Anderson et al., 2015). The protein solution was diluted to 20 nM p97 with 20 µM ATP in triplicate in 5 µls of 1X Kinase Reaction Buffer A; 40mM Tris (pH 7.5), MgCl<sub>2</sub> 20 mM, 0.1 mg/ml BSA and incubated at 37°C for 30 mins. After incubation, the solutions were equilibrated to RT for 5 mins and 5 µl of ADP-Glo™ Reagent was added to stop the reaction and consume the remaining ATP. The solution was incubated for 40 mins at RT. 10 µl of Kinase Detection Reagent was added and incubated for 30 mins. The sample's luminescence was recorded using Varioskan LUX Multimode Microplate reader. Each sample was averaged across the 3 measurements and plotted as mean ± standard deviation.

### 2.3.7. Cryo-EM Sample Preparation and Imaging.

10 mg/ml frozen aliquots of the mutant solution were spun down at 12,000g for 10 mins and the supernatant was diluted to a concentration of 2mg/ml in protein storage buffer with 0.5 µM Octyl glucoside, 1 mM TCEP and 1 mM ADP or ATPγS depending on the experiment. Protein vitrification was achieved using the Leica GP2 plunge freezer on Quantifoil (R1.2/1.3 2.0µm hole, 200 or 300 mesh-Cu) holey carbon grids. Firstly, coated grids were negatively glow discharged at 25 mA for 15 sec and placed on a pair of tweezers in the blotting chamber. Chamber temperature was set to 20°C with 100% humidity to avoid evaporation after blotting. 3 µl of prepared protein solution was loaded on the grid. The grid was automatically blotted for 6 sec, followed by plunge-freezing in liquid ethane.

Frozen grids were stored in cryo grid boxes in liquid nitrogen until imaging. Digital micrographs of frozen hydrated protein particles were recorded at 190000 x magnification with defocus range between -1 to -2  $\mu\text{m}$  with a total dose of  $50 \text{ e}^-/\text{\AA}^2$  on either a Glacios (200kV) or Krios (300kV) transmission electron microscope fitted with a K2 (Gatan) or Falcon3 (Thermo Fischer) direct electron detector camera. Exposure was performed for 24 secs per movie. The beam was aligned and eucentric height and astigmatism were adjusted before imaging (See Table A.4, for a full description of the cryo-EM data collection parameters).

### 2.3.8. Single Particle Picking and 3D Reconstruction.

Automated single particle picking and reconstruction was performed using CryoSPARC Live™ described by the developers (Punjani et al., 2017). A spherical blob of diameter  $160\text{\AA}$  was used as an initial template for particle picking. The particles were sorted into 2D classes, classes with dimers of the p97 hexamer were excluded and those which showed high resolution density of monomers of the p97 hexamer were combined to generate an initial 3D map. The initial map was used to run template-based particle picking on the acquired dataset, the particles were run through 2D and 3D classification to generate the final EM map. Given that p97<sup>WT</sup> is a hexameric complex, C6 symmetry was applied. Resolution of the 3D model was determined according to the resulting Fourier Shell Correlation (FSC) curve at a cutoff criterion of  $\text{FSC} = 0.143$  (Rosenthal & Henderson, 2003) (See Table A.5, for a full description of the cryo-EM structure validation and Figure A.1 for graphical representation of structural validation analysis.)

### 2.3.9. Structure Refinement of p97 Mutants

The WT-p97 structure reported from earlier cryo-EM studies (Banerjee et al., 2016) containing ADP (PDB ID:5FTK) and ATP $\gamma$ S (PDB ID: 5FTN) in both nucleotide-binding regions was used as a rough starting model to initiate refinement for the ADP- and ATP $\gamma$ S-bound mutants respectively. Using the program Chimera (Pettersen et al., 2004), the hexamer was fit into the map to get an initial structure for refinement of the atomic model in PHENIX (Liebschner et al., 2019) using the real-space refinement program. The resulting atomic model was examined using the comprehensive validation tool in PHENIX.

The broader implications of these studies will be briefly discussed in Chapter 5.

### **3. Understanding Tissue Ultrastructure Using High-Resolution Focused Ion Beam – Scanning Electron Microscopy Volumes**

This chapter has been published in The Journal of Structural Biology with the title “Semi-automated 3D segmentation of human skeletal muscle using Focused Ion Beam-Scanning Electron Microscopic images” (Caffrey et al., 2019).

#### 3.1. Chapter Introduction

FIB-SEM is an approach for 3D imaging of specimens with thicknesses greater than  $\sim 1$  micron that cannot be imaged using transmission electron microscopy due to their thickness. In biological FIB-SEM imaging, a focused gallium ion beam is used to progressively remove material from the surface of a macroscopic specimen such as a cell pellet or tissue specimen, with the recording of a backscattered electron microscopic image using a scanning electron beam. The resulting volumes contain useful information on subcellular architecture at spatial resolutions as high as  $\sim 10$  nm. FIB-SEM holds great promise for high resolution study and visualization of tissues through segmentation and can provide new and unexpected insights into the organization of organelles and membranes in the cell (Glancy et al., 2015, 2017; Narayan & Subramaniam, 2015)

However, 3D structural analysis of sub-cellular assemblies requires accurate segmentation and delineation of organelle boundaries, as currently used, the speed of interpreting the image stack using manual segmentation approaches is the principal bottleneck in the application of FIB-SEM. To realistically address biologically and medically interesting problems, increases in the speed of segmentation of at least two orders of magnitude are required. A recent estimate put the amount of time required, using present approaches for manual segmentation, to segment a  $1 \times 10^5 \mu\text{m}^3$  volume to take between  $2 \times 10^4$  -  $1 \times 10^5$  work hours to complete (Berning et al., 2015), this is not including the time taken to acquire such large volumes at high resolutions in the first place. We estimate that to obtain a statistically meaningful analysis of the effect of aging on human tissues we would require several volumes from an individual, each a minimum of  $1 \times 10^4 \mu\text{m}^3$  in size across a population dispersed over a wide age range. Using this approach, to discern

potential age-related variations in mitochondrial architecture with manual segmentation could take years. Machine learning, and other advanced computational techniques have begun to dramatically reduce the time taken to convert an imaged volume to discrete and quantifiable structures within the volume (Berning et al., 2015; Camacho et al., 2018; Januszewski et al., 2018; Kasaragod et al., 2018; Lucchi et al., 2012; Meijs et al., 2017).

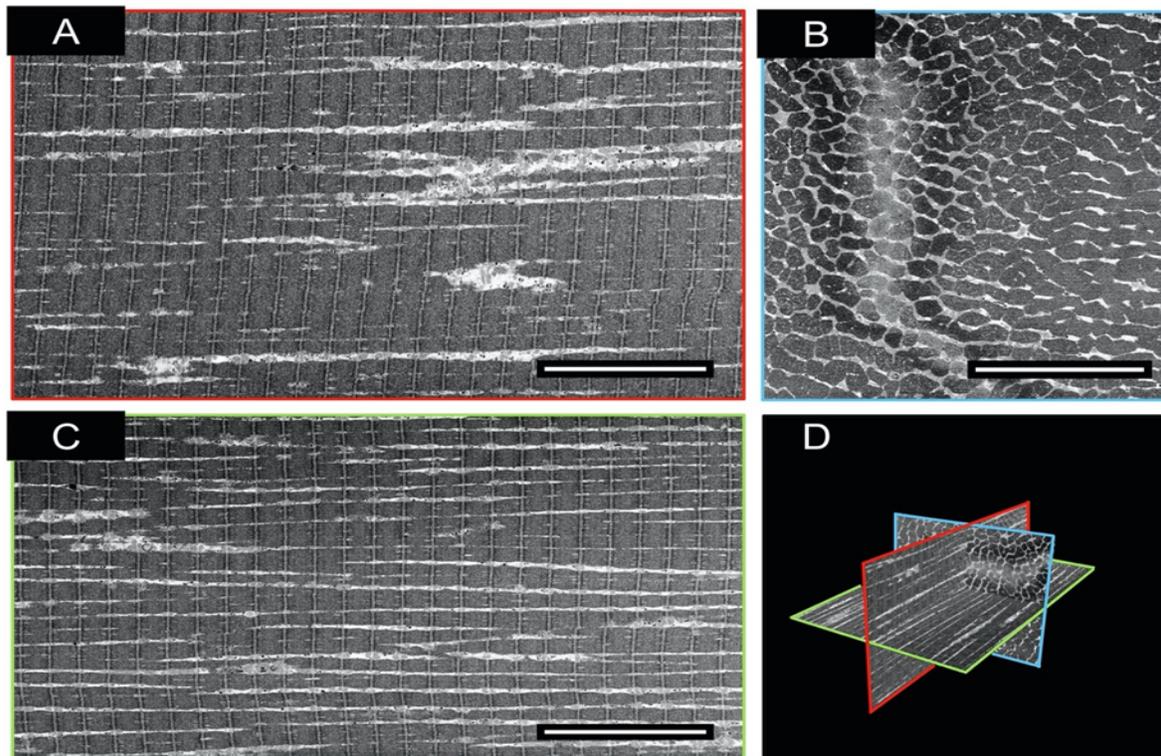
The approach we present here, is a robust, easily accessible, step-by-step workflow from sample preparation to 3D visualization using the readily available Weka segmentation platform on ImageJ (Arganda-Carreras et al., 2017) combined with parallelization scripts for high-throughput segmentation of large volumes. This approach dramatically reduces the time required to describe such structural characteristics on a population wide scale and takes an integrative approach towards sample collection, segmentation and analysis with a view to creating a versatile but accurate methodology for tackling a multitude of biologically relevant problems. Our methodology enabled the collection of  $3.44 \times 10^5 \mu\text{m}^3$  of human skeletal muscle samples from four healthy male individuals in  $\sim 12$  instrument days. Segmentation of the data sets was achieved at the rate of  $\sim 2800 \mu\text{m}^3$  per hour based on actual time estimates of time taken from initial training to the generation of complete 3D segmentations. This corresponds to a  $\sim 500$ - $3000$ -fold increase in speed relative to manual segmentation without loss of the detail required to interpret the image data. Our goal is to combine this high-resolution structural information with biochemical and proteomic analyses of the samples to define the biology of aging. Using our technique described herein, we have taken a major step towards this goal by quantifying multiple 3D characteristics of human skeletal muscle mitochondrial networks in a population of healthy individuals aged 65 and older.

## 3.2. Results and Discussion

### 3.2.1. Qualitative Analysis of Mitochondrial Distribution in Human Skeletal Muscle

In Figure 3.1, we present orthogonal views of the SEM images from a representative FIB-SEM data collection run with a muscle tissue specimen. Early in the design of our experiments, we found that increasing the voxel size from  $5 \times 5 \times 15 \text{ nm}^3$  to  $15 \times 15 \times 15 \text{ nm}^3$  increased the rate of data acquisition  $\sim 3$ -fold from  $500 \mu\text{m}^3$  per hour to  $1500 \mu\text{m}^3$  per hour.

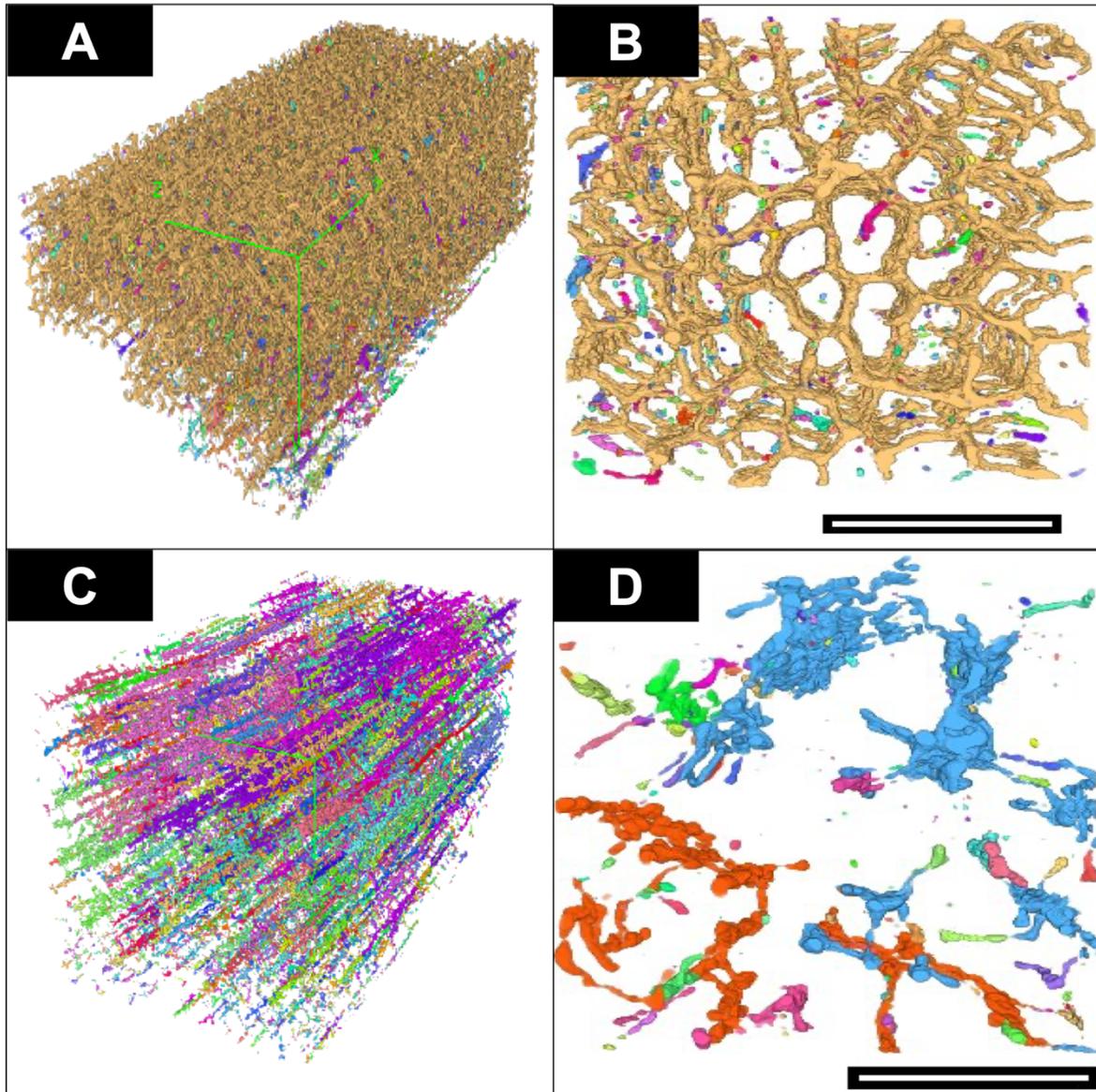
We established that the information required for segmentation was not compromised by the use of larger pixel sizes in the x and y dimensions for the purpose of recognizing mitochondria.



**Figure 3.1: A Typical Example of a Muscle Fibre Acquired With FIB-SEM at a Voxel Size of  $15 \text{ nm}^3$ .** A) Z-Axis (Imaging) face of FIB-SEM volume. B) X-Axis face of FIB-SEM volume. C) Y-Axis face of FIB-SEM volume. D) 3D Orthoslice representation of slices A-C. Scale Bar =  $10 \mu\text{m}$ .

Manual analysis of the 3D image stack shows that the mitochondria display two distinct architectural arrangements, with one class displaying thick, densely packed networks (type A) and those with thin, sparse networks (type B).

The overall spatial arrangements of these mitochondrial types (Figure 3.2) are distinct, with the type A fibers (Figure 3.2 A,B) forming a highly connected assembly, while the type B fibers (Figure 3.2 C,D) are arranged in smaller clusters in addition to being loosely packed.



**Figure 3.2: Morphological Classification.** Each continuous network of connected mitochondria, as determined by ImageJ's "MorphoLibJ" plugin, in the above images were labelled a single color. **A)** Typical "Type A" fiber segmentation volume. **B)** Transverse (X-Axis) "Type-A" image of a mitochondrial sub-volume. The majority of mitochondria in this volume are from a single network, indicated by a uniform label across the whole volume. **C)** Typical "Type B" fiber segmentation volume **D)** Transverse (X-Axis) "Type B" image of mitochondrial sub-volume. The majority of mitochondria in this volume are from multiple discontinuous networks indicated by the multi-colored labelling evident in the volume. Scale bar: 2  $\mu\text{m}$ .

We combined volume acquisition, alignment and normalization, machine learning (ML) training, semi-automated segmentation and statistical analysis into a pipeline and used it to segment multiple tissue volumes (Figure 3.3)

Below, we summarize the main steps of our approach:

1. *Sample Acquisition (Input)*: Once a suitable area was found on a given tissue block, it was imaged with a  $15 \times 15 \times 15$  nm voxel size (xyz) resulting in final volume dimensions of  $60 \times 30 \times 30 \mu\text{m}^3$  ( $54,000 \mu\text{m}^3$ ).

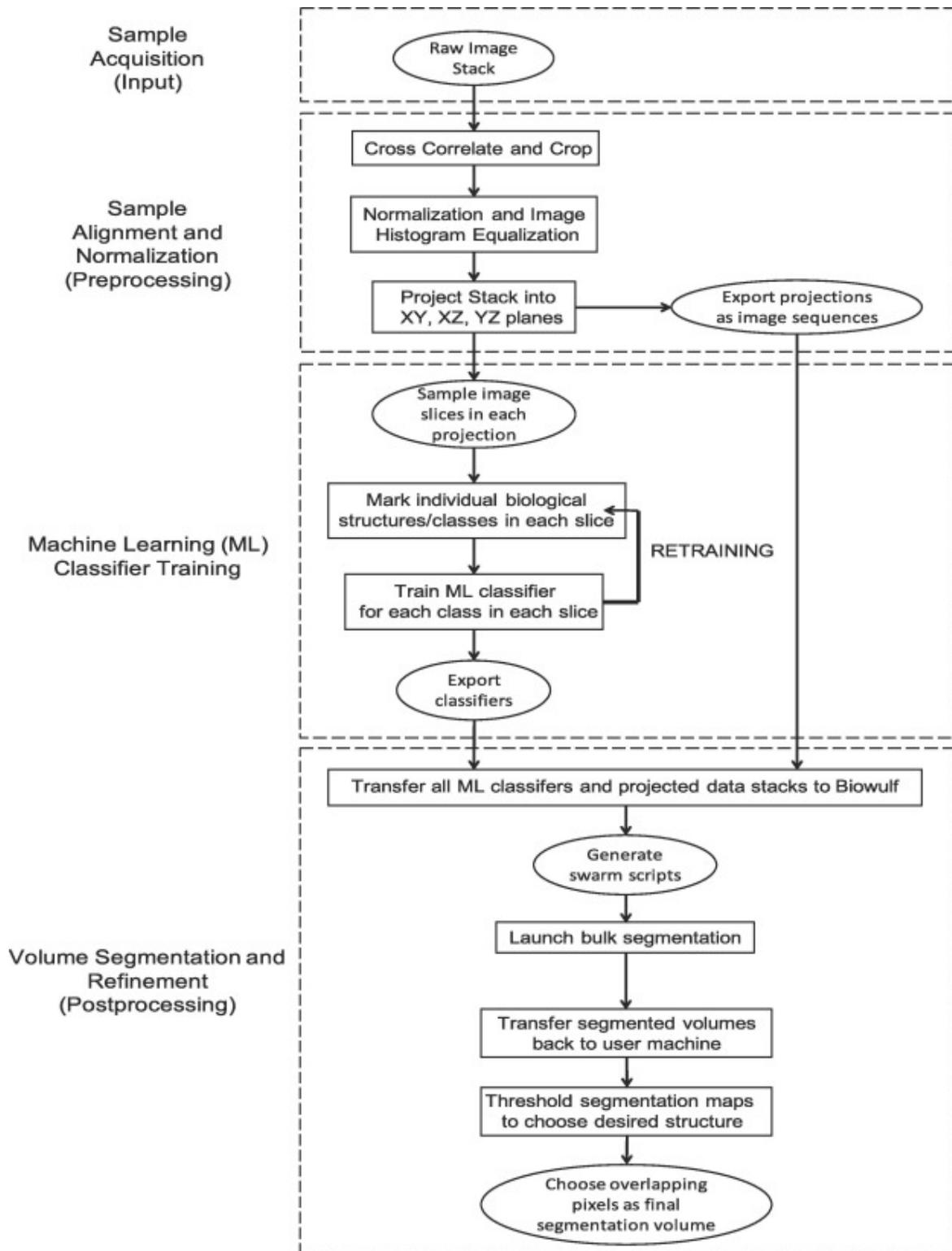
2. *Sample Alignment and Normalization (Preprocessing)*: The individual images (in tiff format) were aligned to a complete 3D stack using a cross-correlation algorithm as described previously (Murphy et al., 2011). The resulting (.mrc) file was opened in ImageJ cropped, median filtered, binned to a voxel size of  $30 \text{ nm}^3$  and the stack histogram was normalized and equalized for reproducibility between volumes. 3 slices were selected from each of the principal axes, evenly spaced across the volume, resulting in 9 images for manual classification.

3. *Machine Learning Training (Manual Classification)*: Each of the major biological structures in the 9 images were classified based on their standard histological features (z-disk, mitochondria, A-band, I-band, sarcoplasmic reticulum and lipids). After sufficient annotation, the Weka segmentation platform was used to train the machine learning software on the images. The output was inspected, and if the software failed to classify the image adequately, the above classification process was repeated iteratively until the software produced an accurate classification of the slice. A (.model) file was exported to the biowulf computing resource at the NIH.

4. *Volume Segmentation and Refinement (Postprocessing)*: The volume prepared in step 2 was exported to biowulf as a series of individual image slices, and each of the 9 classifiers were applied to the image stack, producing  $9 \times 32$ -bit tiff format outputs of classified images, which were imported from biowulf and processed on local computers using the ImageJ image processing package. Images classified as mitochondria were isolated as binary 8-bit tiff format

files. The 3 volumes from each axis were first added together using ImageJ's "Image calculator" function; densities that did not overlap with at least one of the other 2 volumes were removed through simple thresholding. Each axis volume was added together using the previously mentioned function, and density which did not overlap with at least one of the other 2 axes was removed through simple thresholding. The resulting 3D volume, after low-pass filtering was used for statistical analysis of mitochondrial densities.

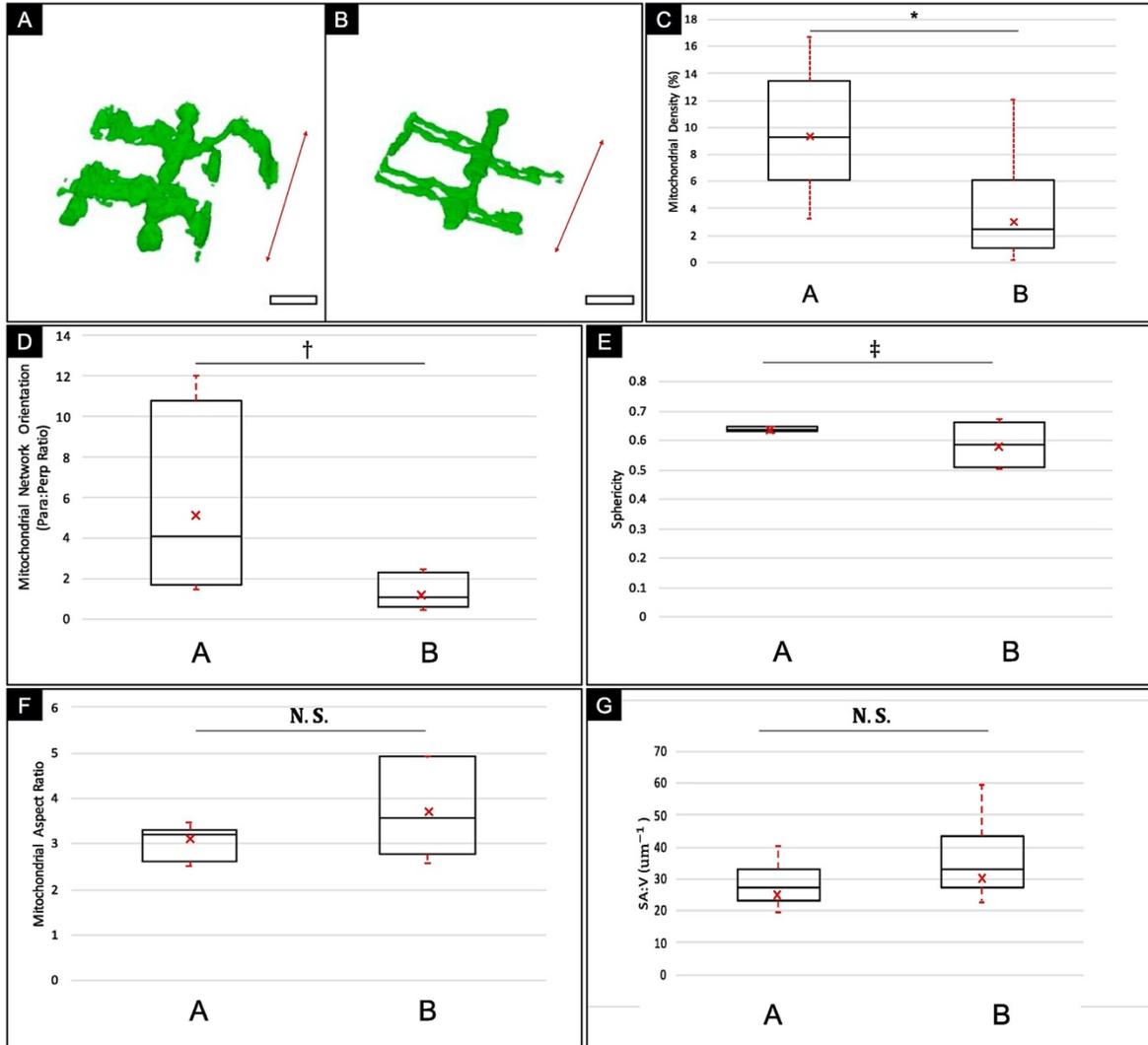
5. *Statistical Analysis:* Each volume was classified according to the mitochondrial density pattern into either a "type A" or "type B" fiber (Figure 3.2). The resulting average densities were analyzed to quantitatively assess the reliability of the segmentation (Figure 3.5). The segmented mitochondrial volume was measured and tabulated (Figure 3.4).



**Figure 3.3: Segmentation Pipeline.** Flowchart indicating the significant steps in the acquisition, segmentation and analysis of 3D volumes.

### 3.2.2. Statistical Analysis of Mitochondrial Distribution in Human Skeletal Muscle

An essential step in the evaluation of this method was in determining its sensitivity to subtle differences in 3D volumes. Figure 3.4 demonstrates this by differentiating between two muscle types across 4 healthy individuals. This type of analysis has the potential to generate statistically relevant data for the study of age and disease-related differences in sub-cellular architecture across a population of individuals, where detection of subtle differences between populations may provide a wealth of insight into the mechanism and progression of disease states.



**Figure 3.4: Boxplot Graphical Overview of Morphological Analysis of Human Skeletal Muscle Mitochondria Distribution from Two Sub-Populations of Data (Type A vs. Type B).** **A)** Typical section of a Type A Mitochondrial Network. **B)** Typical Section of a Type B Mitochondrial Network; Red arrow = Orientation of fiber contraction; Scale bar: 1.5 μm. **C)** Mitochondrial Density Measurement. **D)** Mitochondrial Network Orientation. **E)** Mitochondrial Sphericity. **F)** Mitochondrial Aspect Ratio. **G)** Surface Area-to-Volume Ratio (SA:V). The center line indicates the median values; a × indicates the mean, the box edges depict the 5th and 95th percentiles. The error bars show the maxima and minima of each population. \* indicates a statistically significant difference was found between the muscle types with high power (p-value < 0.01, α = 0.05; Power (1 - β) > 0.95). † Indicates that a statistically significant difference was found between the muscle types with moderate power (p-value < 0.01, α = 0.05; Power > 0.7). ‡ Indicates a marginal difference was found between the muscle types (p-value = 0.03, α = 0.05; Power < 0.7). N.S. indicates no statistically significant difference found between muscle types (p-value > 0.05, α = 0.05). Total Sampled volume = 343,600 μm<sup>3</sup> across 4 healthy individuals.

In Figure 3.4 C, the mean density values measured for type A and B muscle are  $8.9 \pm 1.8\%$  and  $3.2 \pm 1.5\%$ , respectively. This mitochondrial density was found to be a defining characteristic of the skeletal muscle types, consistent with similar density values obtained in murine skeletal muscle measurements (Bleck et al., 2018). Analysis of mitochondrial network orientation in Figure 3.4D, showed a higher degree of variation in the type A muscle relative to type B muscle indicating a higher heterogeneity in the structural composition of type A muscle with a trend towards high mitochondrial density parallel to the mitochondrial contraction axis, whereas type B muscle types had similar densities in both the parallel and perpendicular orientations. The mean orientation values were  $5.2 \pm 4.2$  and  $1.3 \pm 0.6$  for type A and type B muscle types respectively. Mitochondrial Sphericity (Figure 3.4 E) values for the two muscle types are similar but there is a higher variance among the type B muscle fibers. The mean sphericity values were  $0.64 \pm 0.01$  and  $0.58 \pm 0.06$  for type A and type B respectively. Both Mitochondrial Aspect Ratio (AR) and Surface area to Volume (SAV) ratio show no significant difference between muscle types in this study, with mean AR values of  $3.7 \pm 0.73$  and  $3.1 \pm 0.37$  and mean SAV values of  $25.6 \pm 3.3 \mu\text{m}^{-1}$  and  $30.8 \pm 5.0 \mu\text{m}^{-1}$  for type A and B muscle respectively.

Normalization of the average densities across different data sets minimized variability between data sets and allowed us to develop a generalized model of mitochondrial distribution across the muscle samples from different individuals. Combining multiple segmented volumes along each of the principal axes further increased the reproducibility of the results of semi-automated segmentation.

### 3.2.3. Quantitative Evaluation of Segmentation Pipeline Against Manual Standards

Table 3.1 shows the quantitative evaluation of the performance of the method vs two independently segmented versions of the same data set by two individuals along with inter-individual variability, calculated according to the equations in section 3.3.2.2.

|                                       | Semi-Automated<br>vs. Observer 1 | Semi-Automated<br>vs. Observer 2 | Inter-Observer  |
|---------------------------------------|----------------------------------|----------------------------------|-----------------|
| Sensitivity (SENS)                    | $0.9 \pm 0.07$                   | $0.92 \pm 0.06$                  | $0.83 \pm 0.01$ |
| Specificity (SPEC)                    | $0.99 \pm 0.02$                  | $0.98 \pm 0.01$                  | $0.99 \pm 0.01$ |
| Accuracy (ACC)                        | $0.98 \pm 0.02$                  | $0.98 \pm 0.02$                  | $0.99 \pm 0.01$ |
| Dice Similarity<br>Coefficient (DSC)  | $0.79 \pm 0.07$                  | $0.79 \pm 0.08$                  | $0.83 \pm 0.05$ |
| Absolute Volume<br>Difference (AVD) % | $26.7 \pm 23.2$                  | $32.7 \pm 27$                    | $21.1 \pm 16$   |
| Cohen's Kappa ( $\kappa$ )            | $0.78 \pm 0.08$                  | $0.76 \pm 0.1$                   | $0.83 \pm 0.05$ |

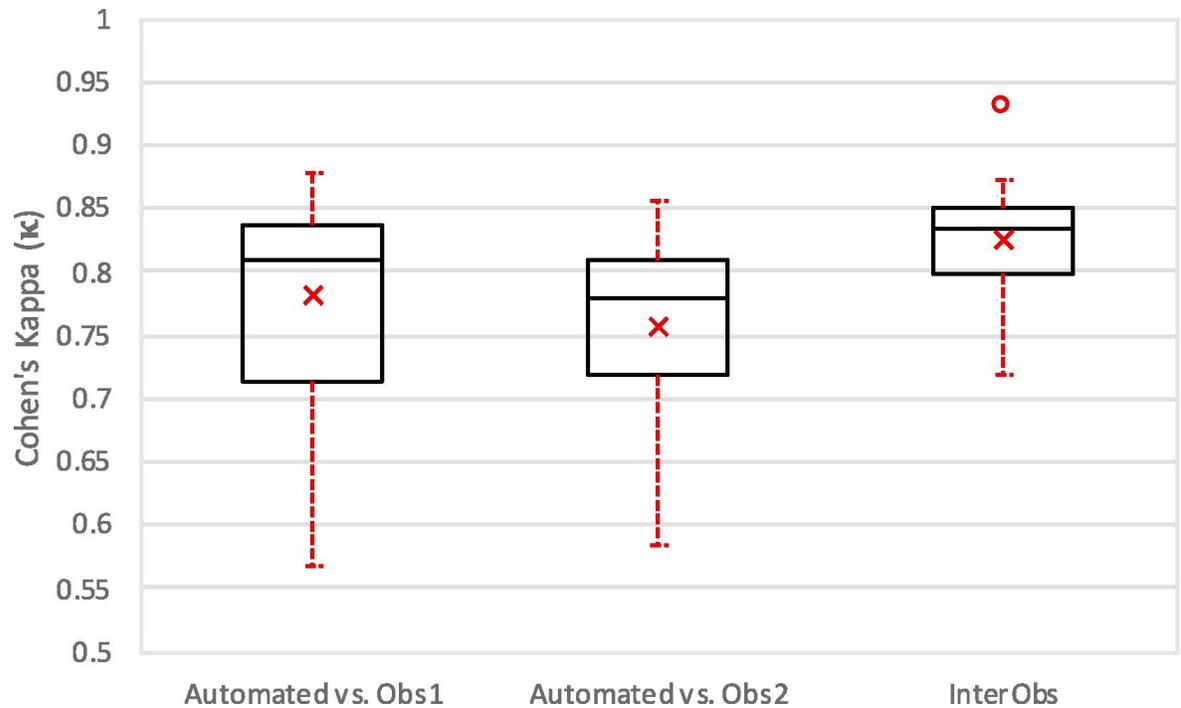
**Table 3.1: Quantitative Evaluation of Automated vs. Manual Segmentation.** Reported as a mean  $\pm$  standard deviation (n=12).

Sensitivity, specificity and accuracy are  $0.83 \pm 0.12$ ,  $0.99 \pm 0.01$  and  $0.99 \pm 0.01$  respectively between the independent manually segmented data sets. A similar relative distribution of the mean sensitivity, specificity and accuracy were found between each manually segmented data set and the semi-automated segmentation with average values of  $0.91 \pm 0.07$ ,  $0.98 \pm 0.01$  and  $0.98 \pm 0.02$  respectively. The relatively low sensitivity and large absolute volume difference (avg.  $26.8 \pm 22.1\%$ ) in all comparisons is indicative of the difficulty in defining the

mitochondrial boundary, even between manual observers, a 1-pixel difference in mitochondrial thickness across a volume can lead to dramatic decreases in sensitivity and an overall large difference between segmented volumes. To illustrate the effect of these deviations on the overall volume and measure the Mean Surface Difference (MSD) between segmented volumes, a 3D volume-to-volume comparison was performed on a typical “type A” and “type B” muscle volume in figure 3.6, which demonstrates that the majority of the difference is within 2–3 voxels across the whole volume.

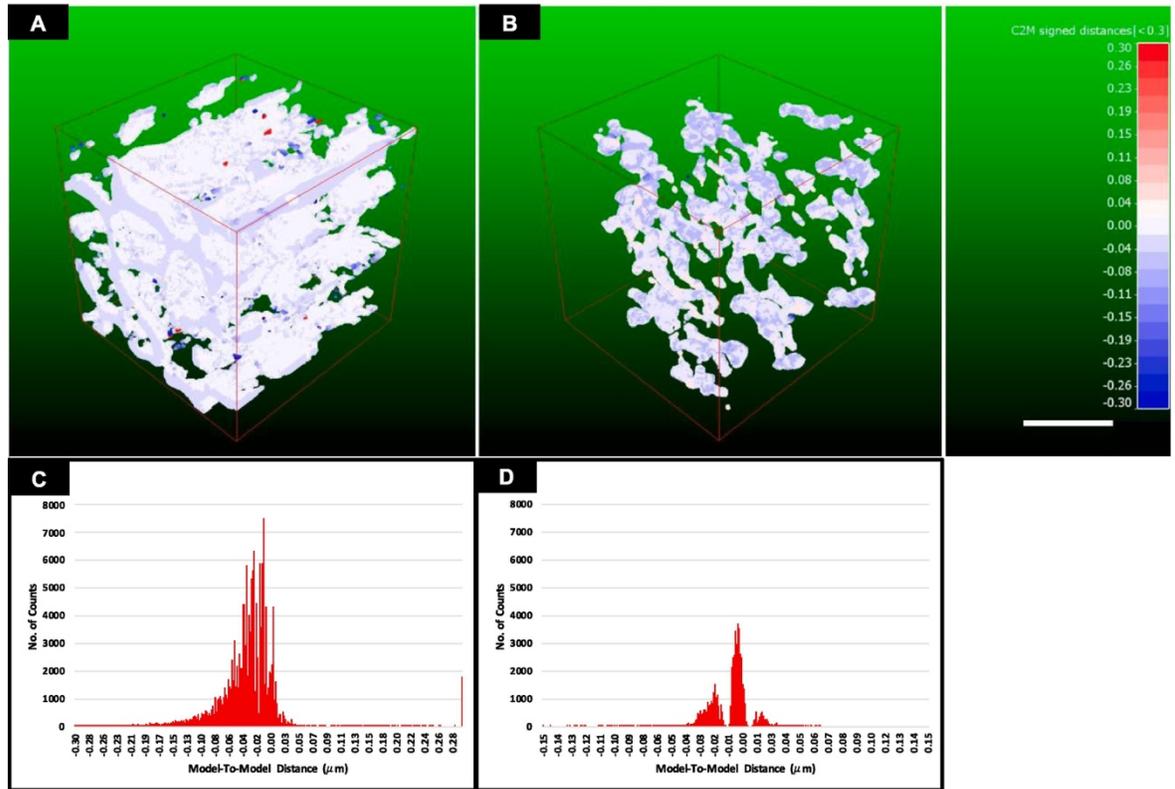
There was excellent agreement between the two manual datasets and between the manual and semi-automated datasets in the overall specificity and accuracy of identification of mitochondria. The high level of agreement between segmentations is illustrated using DSC and Cohen’s Kappa similarity coefficients, between the manually segmented data sets ( $0.83 \pm 0.05$  and  $0.83 \pm 0.05$  respectively) and between each manual and automated segmented data set (with average values of  $0.79 \pm 0.08$  and  $0.77 \pm 0.1$  respectively). Analysis of the similarity coefficients across the sampled population indicated no significant difference in agreement between both the manual and automated segmentations, demonstrating the reliability of our methodology to reproduce structures which approach human level performance ( $p > 0.05$ ,  $\alpha = 0.05$ ).

Figure 3.5 provides a graphical representation of Cohen’s Kappa values showing how the majority of the manual segmentations (75%) are above the widely accepted threshold of 0.7 for semi-automated segmentations (McHugh, 2012). We anticipate that this could be further improved with refinement of the classifiers or increasing the number of classifiers per volume.



**Figure 3.5: Boxplot of Quantitative Evaluation.** Study of inter-observer variability and method versus each observer independently (n = 12), reported as a mean  $\pm$  standard deviation. The center line indicates the median value; a  $\times$  indicates the mean, the box edges depict the 25th and 75th percentiles. The error bars show the extremes at 1.5 inter-quartile range, calculated inclusive of the median, excluding outliers, indicated by  $\circ$ .

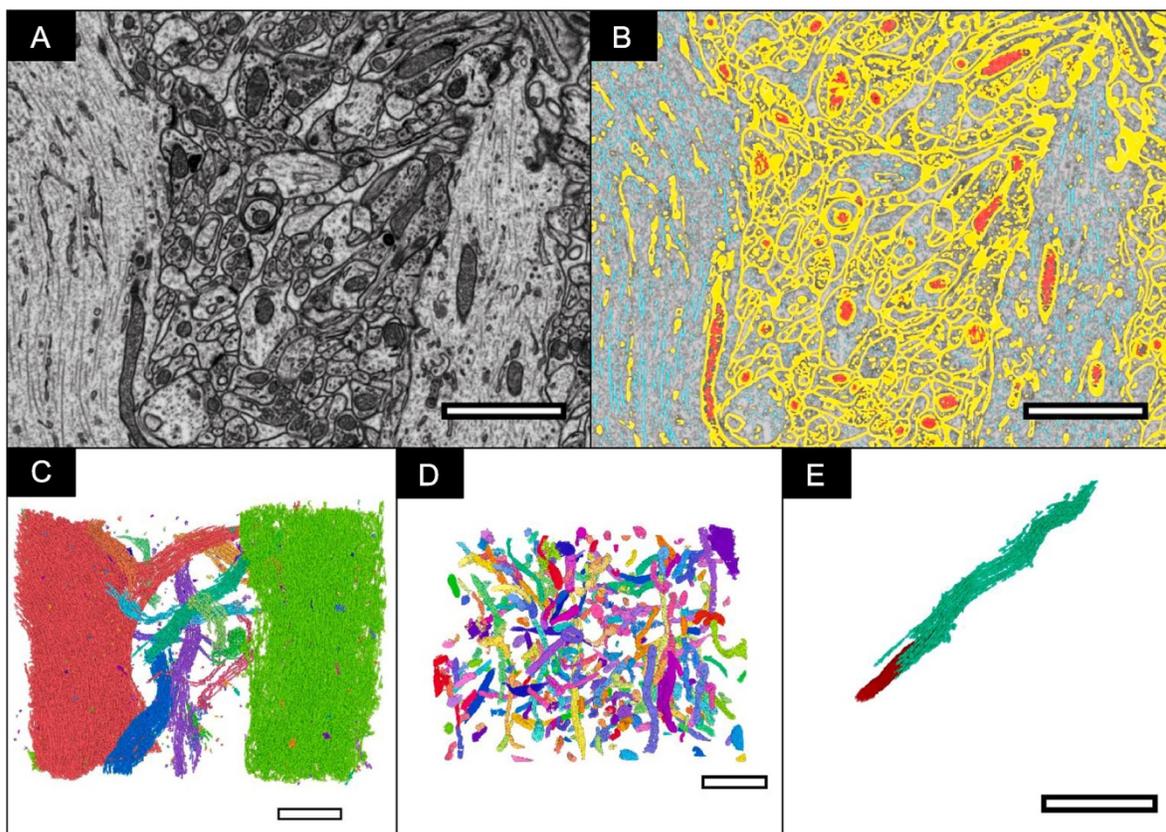
Figure 3.6 provides a graphical representation of the model-to-model distance map between semi-automated and manual segmentations of two muscle types. MSD was calculated by fitting the above distributions (Figure 3.6 C,D) to a Gaussian distribution, and the mean  $\pm$  standard deviation was determined. The MSD showed the semi-automated segmentation was accurate to  $0.03 \pm 0.06 \mu\text{m}$ , indicating a segmentation accurate to 2 or 3 voxels, with a slight bias to overestimate the size of the mitochondria relative to the manual segmentation. Of note, differences of the same magnitude were detected between observers, as mentioned previously and is indicative of the difficulty in defining precisely mitochondrial boundaries.



**Figure 3.6: Model-to-Model Distance Measurement.** **A)** Isometric projection of  $100 \mu\text{m}^3$  3D model-to-model distance map for a typical “Type A” sub-volume. **B)** Isometric projection of  $100 \mu\text{m}^3$  3D model-to-model distance map for a typical “Type B” sub-volume. **C)** A graphical representation of the distribution of the mean surface distances between manual (reference) and semi-automated (comparison) volumes across the 3D mesh map for a typical type A sub-volume. **D)** A graphical representation of the distribution of the mean surface distances between manual (reference) and automated (comparison) volumes across the 3D mesh map for a typical type B sub-volume. Red-White-Blue distance map represents distances in microns, Red: Manual model > Automated model; White: Manual  $\approx$  Automated model ( $\pm 15$  nm); Blue: Manual < Automated model. Scale bar =  $2 \mu\text{m}$ .

### 3.2.4. Evaluation of Semi-Automated Segmentation Performance using CA1 Hippocampal Test Dataset

Figure 3.7 is a demonstration of the performance of the segmentation approach against a hippocampal dataset. The time for obtaining this segmentation of  $400 \mu\text{m}^3$  volume took  $<24$  h. We estimate that a 100-fold increase in the volume of the data to be segmented would not increase the segmentation time considerably, once the classes are produced they can be applied across an extremely large volume with little-added input.



**Figure 3.7: Other Applications of This Software.** **A)** FIB-SEM slice from the CA1 hippocampal region of the brain with a voxel size of  $5 \times 5 \times 5 \text{ nm}^3$  **B)** FIB-SEM slice with automated segmentation overlaid (Yellow = Cell membrane; Red = Mitochondria; Cyan = Microtubules) **C)** 3D volume of segmented microtubules labelled separately, allowing for the straightforward isolation of individual cells for focused study. **D)** 3D volume of segmented mitochondria labelled separately. **E)** Individual microtubule (cyan) and mitochondria (red). Scale bar =  $2.5 \mu\text{m}$ .

### 3.3. Methodology

#### 3.3.1. FIB-SEM Acquisition and Image Processing

##### 3.3.1.1. Candidate Selection and Muscle Biopsy

This study was conducted in healthy men participating in the Baltimore Longitudinal Study and Aging (BLSA) and the Genetic and Epigenetic Signatures of Translational Aging Laboratory Testing (GESTALT) studies. The design and description of the BLSA and GESTALT studies have been previously reported (Kirkland, 1985; Stone & Norris, 1966; Tanaka et al., 2018). Skeletal muscle biopsies were performed in fasting conditions as described elsewhere (Gonzalez-Freire et al., 2018). Briefly, a ~ 250mg muscle biopsy was obtained from the middle portion of the vastus lateralis muscle using a 6-mm Bergstrom biopsy needle inserted through the skin in the muscle. A small portion of muscle tissue (~5mg) was immediately placed in 2% Glutaraldehyde (GA) and 2% Paraformaldehyde (PFA) in 100mM sodium cacodylate buffer, pH 7.3-7.4 at 4°C until required for sample preparation. The rest of the biopsy specimen was snap frozen in liquid nitrogen and subsequently stored at -80°C until used for further analyses.

##### 3.3.1.2. Fixation, Contrasting and Embedding

Muscle biopsy samples from human donors were fixed with 5% glutaraldehyde in 100mM sodium cacodylate buffer at pH 7.4 as in a murine skeletal muscle study (Glancy et al., 2015). In order to achieve the contrast required to be able to consistently identify mitochondria with similar signal to noise ratio the standard post-fixation protocol used for the murine muscle skeletal muscle samples was changed. Here we post-fixed with 2% Osmium Tetroxide (OsO<sub>4</sub>) in sodium cacodylate buffer for 1hr at RT, washed with ddH<sub>2</sub>O and treated with 4% tannic acid in sodium cacodylate buffer.

A second treatment of 2% OsO<sub>4</sub> in cacodylate buffer either reduced or not reduced with 0.6% Potassium Ferrocyanide was performed for 1hr at RT. Samples were washed in ddH<sub>2</sub>O and treated with 2% Uranyl Acetate (UA) in ddH<sub>2</sub>O at 4°C overnight (Gunji et al., 1980; Lewinson,

1989). The samples were washed in ddH<sub>2</sub>O, 5 x 10 min, and dehydrated using a graded ethanol series ending in 100% propylene oxide. Infiltration of embedding media was performed using a ratio of 2:1, 1:1, 1:2 propylene oxide to Eponate12 resin formula (EMS). Samples were embedded in resin molds and placed in an oven set at 60°C overnight for polymerization.

### 3.3.1.3. Area Selection for FIB-SEM Analysis

Areas of muscle were chosen for FIB-SEM data collection following a survey of 0.5-1 µm thick sections of resin-embedded muscle tissue; sections were created using an Ultracut S microtome from Leica Microsystems. The sections were stained with Toluidine blue which stains nucleic acids blue and polysaccharides purple. Once stained, the orientation and morphology of the fiber was assessed using a light microscope. Suitable areas with intact muscle fibers were chosen for FIB-SEM data collection using the last section taken from the top of the block-face, and digital images were taken for reference. These images were used as maps to pinpoint the previously selected areas for data collection in the FIB-SEM (Glancy et al., 2015). The resin was cut to create a suitable sample for SEM. The samples were sonicated in ethanol: water (70:30) for 15 mins to remove dust and particulates which would hinder imaging. The sample was mounted on an aluminium stub using a double-sided adhesive conductive carbon tab, and the sides painted with silver paint to prevent charge build-up.

The sample was allowed to dry, placed in a sputter coater (Cressington model 108), and coated with gold for 40 seconds at 30 mA.

After gold coating, the sample was placed into the sample chamber of the FIB-SEM. FIB-SEM imaging was performed using a Zeiss NVision 40 microscope, with the SEM operated at 1.5 keV landing energy, a 60 µm aperture and backscattered electrons were recorded at an energy selective back-scattered electron (ESB) detector. The user interface employed ATLAS 3D from Carl Zeiss, consisting of a dual 16-bit scan generator assembly to simultaneously control both the FIB and SEM beams and dual signal acquisition inputs, as well as the necessary software and firmware to control the system.

The fiber of interest was located using the SEM, and the instrument was brought to eucentric and coincidence point at a specimen tilt of 54°, i.e. the specimen height where the specimen does not move laterally with a change in tilt and where the focal point of both FIB and SEM coincide. Once the exact milling area was determined with reference to the microscope images, a protective platinum pad was laid down on top of the area using a Gas Injection System (GIS) of size 60 µm x 30 µm and 5 µm in thickness. Alignment marks were etched into the platinum pad using an 80 pA FIB aperture to allow for automated tracking of milling progress, SEM focus and stigmation during acquisition. After alignment etching, the platinum pad was covered with a carbon pad using the GIS to protect the etched marks from the milling process. After deposition of the carbon pad, a trench was dug using a 27nA FIB aperture to allow for line-of-sight for the SEM ESB detector. After the trench was dug, the imaging face was polished using a 13 nA FIB aperture. The FIB aperture was changed to 700 pA and SEM imaging area selected (Typical Image size: 4000 px x 2000 px /Pixel size: 15 x 15 x 15 nm [xyz]) the automated acquisition software was set up and run until all the sample area was acquired.

#### 3.3.1.4. Image Processing

After SEM acquisition the individual image files (.tif) were aligned using a cross-correlation algorithm (Murphy et al., 2011). The images were opened in ImageJ, and the volume was cropped to ensure a minimum distance of at least 1µm away from the cell boundary in any direction, this was performed to reduce measurement variability of mitochondrial density due to the non-uniform distribution of mitochondria near capillaries and cell boundaries (Sjöström et al., 1982).

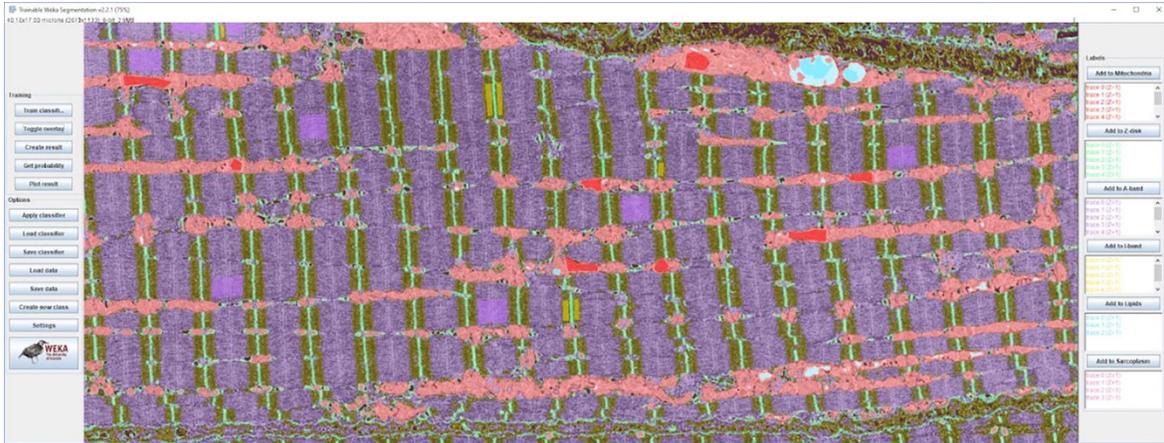
To reduce noise volumes were median filtered by 1 pixel in the x, y and z directions and binned by 2 in all three axes to produce a final voxel size of 30 x 30 x 30 nm. The volume's contrast was normalized and equalized using ImageJ's "Enhance Contrast" function.

### 3.3.2. Automated Segmentation Methodology and Validation

#### 3.3.2.1. Automated Segmentation Methodology

Sample images were required for preliminary training to construct the necessary machine learning classifiers for automatic segmentation. Referring to the schematic in Figure 3.9, three representative slices (one from each 3<sup>rd</sup> of the volume), were taken at random from each of the principal axes: x, y and z (9 slices in total) (Figure 3.9 A). A classifier was trained for each slice by sampling several main structures found in each sample image (Figure 3.9B). The primary structures, based on standard histological examples, were as follows: z-disk, mitochondria, A-band, I-band, sarcoplasmic reticulum and lipids. The classifier was trained using all training features available in the “Trainable Weka Segmentation” plugin for ImageJ Fiji, a robust machine learning plugin that is professionally maintained by The University of Waikato. The Weka algorithm, in brief, extracts image features using common filters that can be categorized as edge detectors (e.g. Laplacian and Sobel filters), texture filters, (such as minimum, maximum, and median filters), noise reduction filters (such as Gaussian blur and bilateral filter), and membrane detectors, which detect membrane-like structures of a specified thickness and size. Furthermore, the feature set also included additional features from the ImageScience suite (<https://imagescience.org/meijering/software/imagescience>).

Since only 2D image features were calculated, classifiers were trained and applied on all three image axes to compensate for the loss of a third dimension. In our machine learning approach, we applied the multi-threaded version of the random forest classifier with 200 trees and 2 random features per node. Probability measurements of each class were generated, allowing for a class-by-class assessment of the performance of each classifier during training. Segmentation masks of the key skeletal structures were outputted based on these probability measurements (Figure 3.8).



**Figure 3.8: Example of Weka Software Manual Classification in ImageJ.**

Once all 9 classifiers were trained (3 for each axis), they were exported as separate ".model" files and applied to each slice in the volume according to the respective axes which they were trained.

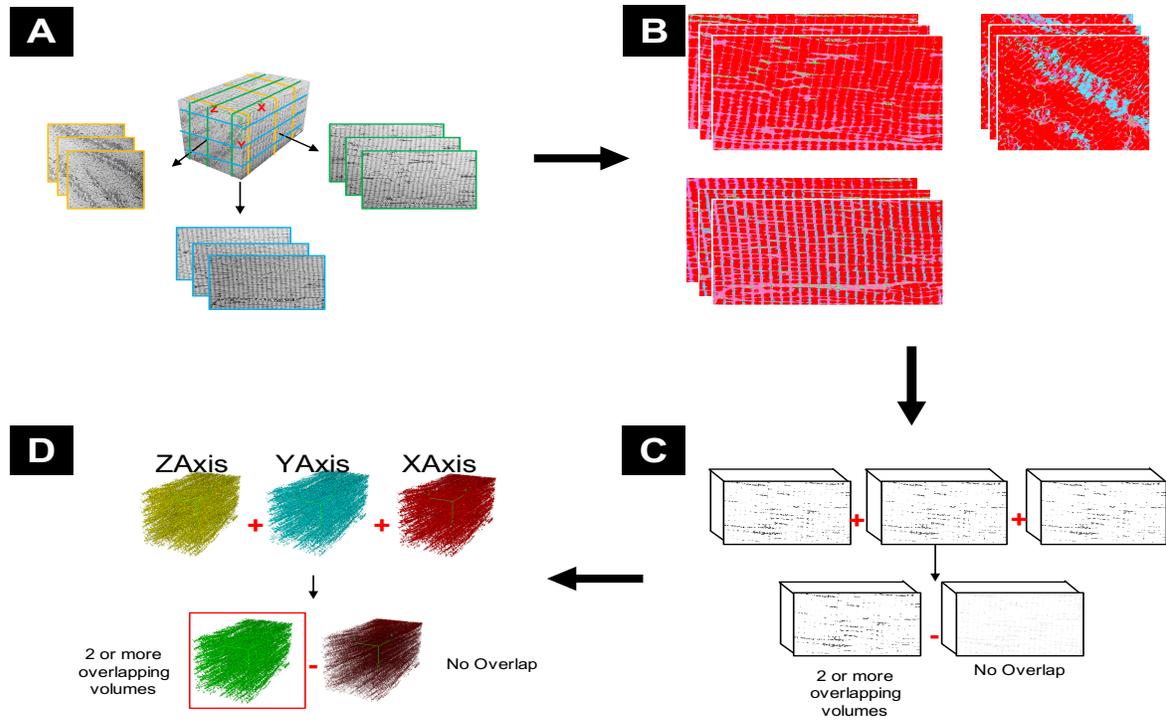
The segmentation of the full volume was performed on the Biowulf supercomputer cluster, and its implementation is as follows:

To simultaneously process many image slices at once, each slice was opened in a separate instance of ImageJ Fiji and Weka machine learning was executed in each instance. Since only 2D image features were calculated, each instance of ImageJ Fiji could effectively classify an image without needing access to any other image data.

Thus, images were simultaneously classified by parallel processors running multiple instances of ImageJ Fiji. Each instance executed a Beanshell script (source.bsh) that automatically performed Weka machine learning on a specified image using a specific classifier file. The process of opening instances of ImageJ Fiji was automated through the command line interface by using the existing "--headless" option that came with the ImageJ Fiji package. Biowulf effectively allocated and launched hundreds of processors at once with the use of the "swarm" command that already existed on the supercluster.

The command required a formatted file containing independent commands to distribute to each processor and to generate such a file quickly we wrote an automated Bash script "generate\_swarm\_script.sh". If a different system other than Biowulf is being used, it is advised to create a script that launches parallel instances of ImageJ Fiji that execute the "source.bsh" Beanshell script. It should be noted that Weka machine learning is optimized to run faster by utilizing a substantial amount of RAM. For the classification of our large FIB-SEM images, we allocated 25 GB RAM per processor per image.

A total of 9 automated segmentation volumes were created. The 3 volumes from each axis were first added together using ImageJ's "Image calculator" function, and density which did not overlap with at least one of the other 2 volumes was removed through simple thresholding (Figure 3.9 C). Each axis volume was added together using the previously mentioned function, and density which did not overlap with at least one of the other 2 axes was removed through simple thresholding (Figure 3.9 D). The volume was filtered by 2 pixels in the x, y and z directions using ImageJ's "Median 3D Filter" function and was used for statistical analysis of mitochondrial densities.



**Figure 3.9: Graphical Representation of Key Steps in Segmentation Pipeline.** **A) Classifier Selection:** Three representative slices are taken from each of the principal axes. **B) Classifier Generation:** Each slice is manually classified based on the organelles within the volume. **C) Volume Classification and Refinement:** The classifiers are applied to the entire volume and produce segmented volumes of each class. The mitochondrial class is isolated, and each of the 3 volumes from the same axes are combined and non-overlapping data removed to produce an axial volume. **D) Axial Volume Combination:** Each of the refined volumes from the principal axes are combined, and non-overlapping data is removed.

### 3.3.2.2. Methodology Validation

Automated segmentation methods were compared to manually segmented volumes using the following metrics:

1. Absolute Volume Difference (AVD) (%): Absolute volume difference measurements were performed to measure the total volume difference between manually and automatically segmented volumes, allowing for a global metric of volume-to-volume difference. For labels with identical volume, %Difference (A, M) = 0, with increasing values indicating a greater volume difference between the two labels.
2. 3D Model-to-Model distance (Mean Surface Distance / MSD): Both the manual and automated volumes were converted to ASCII mesh surfaces using ImageJ's "3D viewer" (Schmid et al., 2010). These meshes were transferred to the Cloud Compare platform (Girardeau-Montaut, 2015), where the manually segmented (reference) volume was compared to the automatically generated (comparison) volume, using the "Compute cloud/mesh distance" tool a map of the model-to-model distance was created. The max distance between the reference and automated datasets was set to 0.3  $\mu\text{m}$  (any greater distance was set to the maximum threshold), and the model-to-model distance distribution was fitted to a Gaussian distribution, and the mean  $\pm$  standard deviation calculated. A tricolor histogram was applied to the map with red representing automated density areas greater than manually segmented density, blue representing automated density areas less than manually segmented density and white represents  $< \sim 15\text{nm}$  difference between structures.
3. Sensitivity, Specificity, Accuracy, Dice Similarity Coefficient (DSC) and Cohen's Kappa ( $\kappa$ ) calculations. Sensitivity, specificity and accuracy were calculated according to the conventional equations.

$$\text{DSC} = [2 * V(A \cap M)] / [V(A) + V(M)]$$

**Equation 1: Dice Similarity Coefficient (DSC) calculation** (Dice, 1945).

Cohen's Kappa ( $\kappa$ ) =  $(P_a - P_e)/(1 - P_e)$

Where  $P_a$  = Actual Observed Agreement = **Accuracy**;  $P_e$  = Expected Agreement =  $[\{(TP+FN)*(TP+FP)\}/n + \{(FP+TN)*(FN+TN)\}/n] / n$

Where  $n$  = total number of observations =  $TP+FP+FN+TN$  and  $TP$  = True positive (**Automated [A]  $\cap$  Manual [M]**);  $FP$  = False Positive ( $A \setminus M$ );  $FN$  = False Negative ( $M \setminus A$ ) and  $TN$  = True Negative ( $U \setminus [A \cup B]$ ).

**Equation 2: Equation for Cohen's Kappa calculation** (McHugh, 2012).

### 3.3.3. Statistical Analysis of Mitochondrial Distribution in Human Skeletal Muscle

Mitochondrial density measurements were made by splitting each volume into 100  $\mu\text{m}^3$  segments and the volume of white pixels in the binary segmented image were measured using the "3D Objects Counter" function in ImageJ.

Mitochondrial surface area to volume measurements were made by dividing the surface area and volume results determined using the "3D Objects Counter" from the binary segmented image. Mitochondrial orientation, aspect ratio and sphericity were calculated according to the methods described in (Bleck et al., 2018). Mitochondrial orientation was determined relative to the fiber contractile axis which is defined by vertically aligning the Z-disks in the raw image data in both XY and XZ planes, using the "Rotate" and "Reslice" tools in ImageJ. The average projection of both XY and XZ planes of aligned mitochondrial networks were produced using the Z Project tool. The OrientationJ Distributions plugin was used to determine the angles of the mitochondrial network. Parallel mitochondria were determined to be values with a  $\pm 0-10^\circ$  angle and perpendicular mitochondria were determined to be values with a  $\pm 80-90^\circ$  angle.

Mitochondrial aspect ratio was calculated by multiplying mitochondrial elongation and flatness values together, determined using the "3D Shape Measure" plugin in ImageJ. Mitochondrial sphericity was also calculated using the "3D Shape Measure" plugin.

The broader implications of these studies will be briefly discussed in Chapter 5.

## **4. Examining Nanotherapeutic Interactions on a Cellular Scale**

### **4.1. Chapter Introduction**

One key area of nanomedicine research is the mechanism of action of nanoparticles within the body, its biodistribution and pharmacokinetic properties (Ernsting et al., 2013). However, insight into the nanoparticle-cell interface is limited, due in part to the heterogeneous nature of nanoparticles as a drug class and the numerous possible interactions between cell and nanoparticle which stem from this diversity such as surface modifications, particle composition and final target tissue.

Knowledge of the specific events which occur once a nanotherapeutic has reached its target are still very early in development, this is in part due to the difficulty of resolving interactions in sufficient spatiotemporal resolutions for a concrete understanding of the mechanism for nanoparticle-cellular uptake. However, this is crucially important for our understanding of the therapeutic capabilities of nanoconstructs as a whole. Traditionally the study of nanoparticle-cell interactions has usually remained exclusively in either the light or electron microscopies, sacrificing either high resolution cellular information or accurate identification of specific biomolecules through fluorescent markers respectively. We have successfully produced a method for the study and categorization of nanoparticle-cell interactions using CLEM, towards monitoring different drug delivery methods and their effect on target tissues at high resolutions.

## 4.2. Results and Discussion

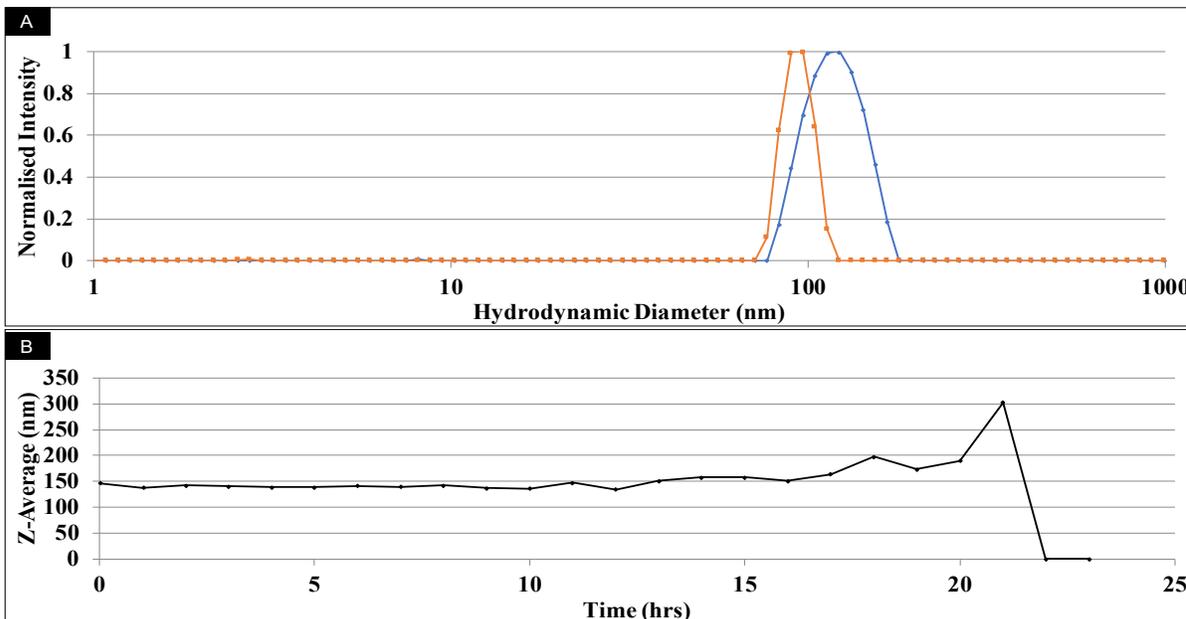
### 4.2.1. Nanoparticle Characterisation and Stability

In order to ensure a uniform particle preparation and the stability of prepared nanoparticles in solution during nanoparticle uptake experiments, it was necessary to analyse the particle size and distribution over time in solution.

| Sample  | Size – Z-Average (nm) (n=4) | PDI (n=4) |
|---|-----------------------------|-----------|
| Gold Bare (100nm) in 1% (v/v) PBS<br>in Water | 92.6                        | 0.05      |
| Transferrin-Gold (100nm) in PBS               | 117.2                       | 0.15      |
| Transferrin Protein in PBS                    | 3.5                         | 0.85      |

**Table 4.1: Dynamic Light Scattering Measurements of Nanoparticle Protein Complex.** Demonstrates a shift in hydrodynamic size (Z-Average), measured by Dynamic Light Scattering (DLS) of 100nm gold particle upon Transferrin absorption.

The PolyDispersity Index (PDI) is used as an indication of the relative distribution of particle sizes within the measured population. PDI values  $< 0.1$  are considered monodisperse i.e. of similar hydrodynamic diameter and those  $>0.4$  are considered to be polydisperse i.e. a mixture of different particle sizes. A PDI value of 0.15 for the Tf-Gold nanoparticles prepared here is considered standard for nanoparticle preparations of protein absorbed nanoparticles.



**Figure 4.1: DLS Analysis of Transferrin Bound Nanoparticles.** A) DLS intensity measurement of 100nm Gold particles before (orange) and after (blue) transferrin absorption. B) DLS Z-Average stability measurement of Tf-absorbed gold nanoparticles at 37°C over a 23hr period.

Figure 4.1 A, demonstrates the increase in hydrodynamic diameter upon addition of transferrin to the solution and the broadening of the peak reflects the increase in the polydispersity of the solution. Figure 4.1B, indicates that the prepared transferrin particle solution remained stable up to 15hrs. The gradual increase in Z-average and subsequent drop over the next 8hrs is most likely due to the gold nanoparticle agglomeration typical of nanoparticles of this size and not aggregation, nevertheless this graph shows that the particles are stable well beyond the timescale of nanoparticle passaging in this project (<4 hrs) and no passage of gold nanoparticles was made 12 hrs after production.

## 4.2.2. Characterisation of Nanoparticles and Heavy-Metal Staining in FIB-SEM Images

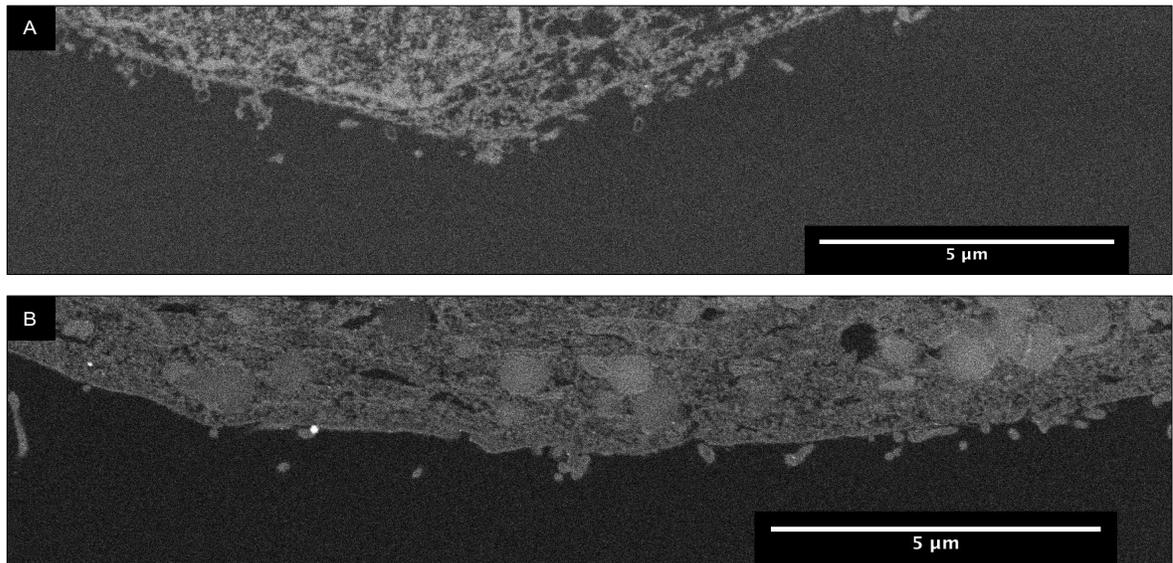


**Figure 4.2: Energy-Dispersive X-Ray Spectroscopy (EDS) Analysis of Resin Block.** **A)** SEM image of typical central slice through nanoparticle treated cell, with high-scattering ROI's selected for EDS analysis circled; **B)** EDS spectra of Osmium rich high-scattering vesicle, indicating a high concentration of lipids present.; **C)** EDS spectra of Uranium rich high-scattering extracellular object, these Uranyl aggregates appeared from time-to-time in different preparations and was important to differentiate from gold nanoparticle accumulations. **D)** EDS spectra of suspected gold nanoparticle, presence of gold signal validate the identification of the gold nanoparticle. Scale bar: 1  $\mu\text{m}$

EDS analysis enabled the discrimination of three main components of the sample with high scattering intensities and allowed for the verification of gold nanoparticle size and shape in the cell sample, for future acquisitions.

#### 4.2.3. Determination of Fixation and Permeabilization Parameters for Retention of Ultrastructural Information from Light to Electron Microscopy

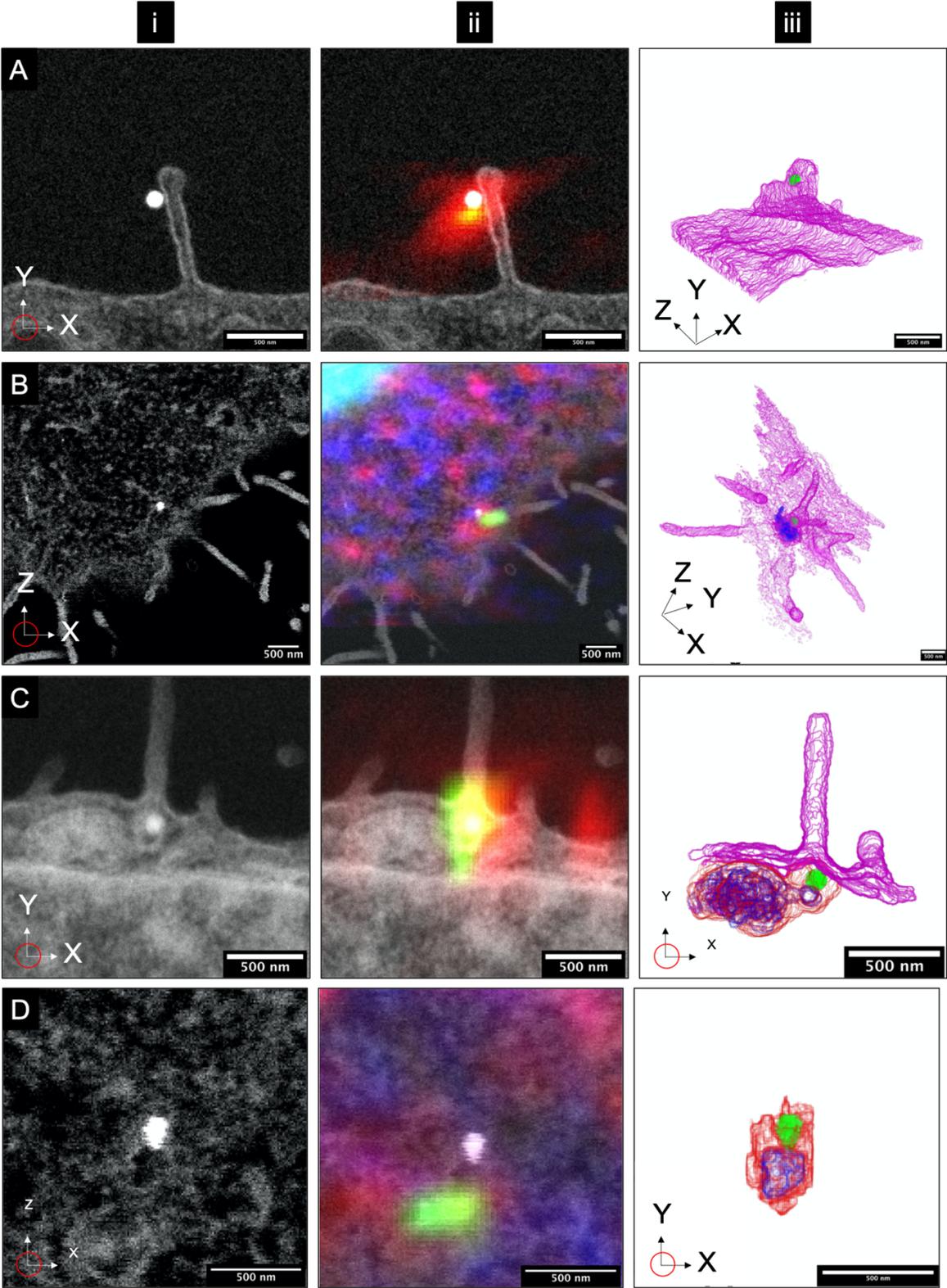
During CLEM analysis it became apparent there was a need for optimisation of cell preparations for electron microscopy. Paraformaldehyde fixation preceding permeabilization was not sufficient to preserve cellular ultrastructure during the harsh dehydration and resin embedding steps necessary for EM acquisition (Figure 4.3 A).

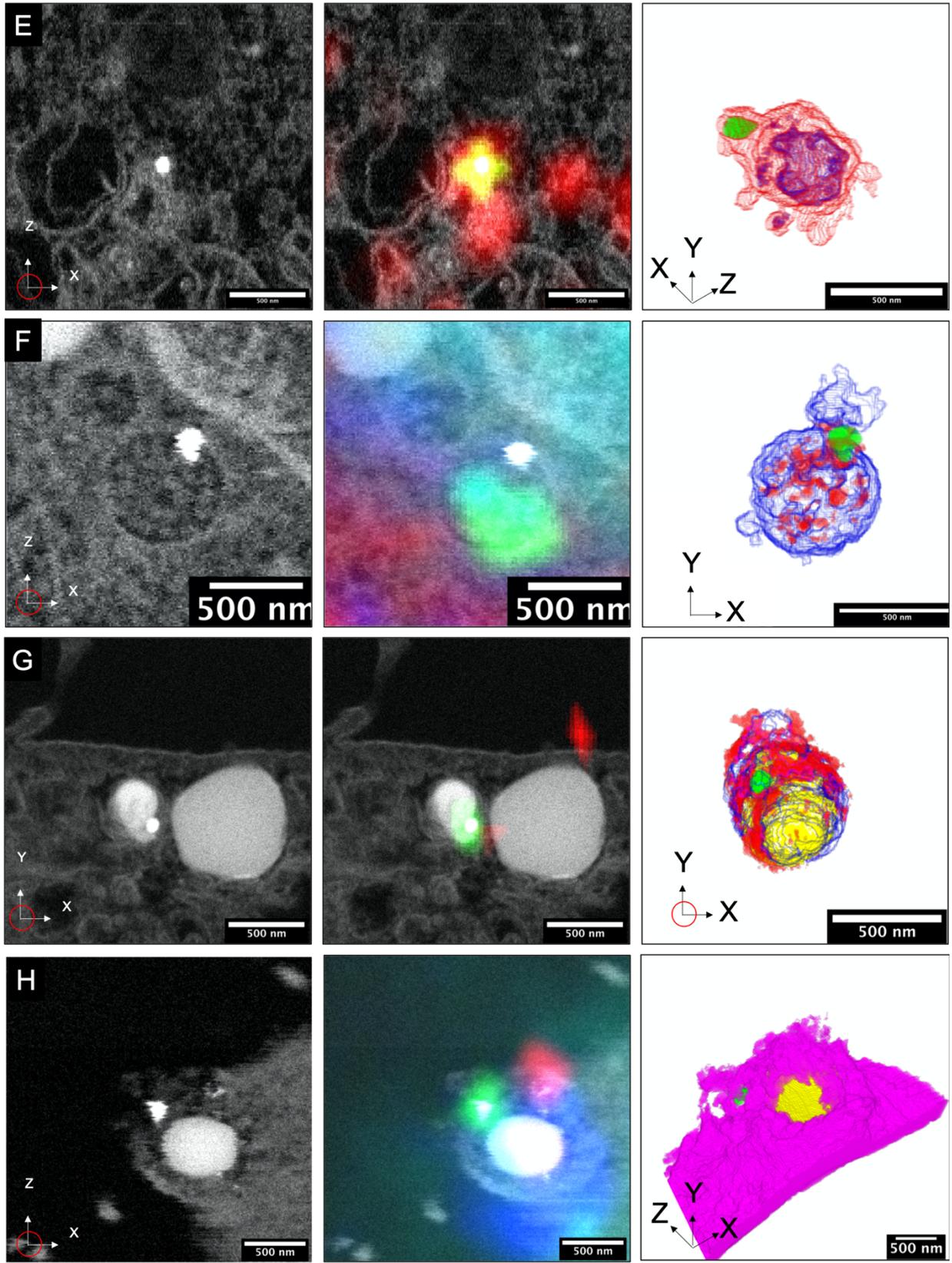


**Figure 4.3: Comparison of Ultrastructural Features of A549 Human Lung Cell Carcinoma Cells Before (A) and After (B) Optimisation.** Fixation and permeabilization conditions for A): 4% v/v PFA fixation for 30mins followed by 1hr incubation in 0.5% w/v Saponin with 1% BSA at RT and for B): 0.2% v/v GA and 4% v/v PFA fixation followed by 1hr incubation in 0.1 w/v Saponin with 3% BSA at 4°C.

GA fixation was avoided in early experiments due to its associated reduction of antigenicity and autofluorescence in LM, however after testing several conditions it was determined that GA was necessary to stabilise cellular structure enough to survive permeabilization and subsequent dehydration and embedding. We determined that a concentration 0.2% v/v of GA, in combination with a reduced saponin concentration (0.1% w/v) allowed for sufficient retention of high-resolution ultrastructural information in the EM (Figure 4.3 B) while also minimising undesirable effects in the LM.

4.2.4. Correlative Light Electron Microscopy Analysis of Nanoparticle-Cell Interactions





**Figure 4.4: A-H Segmented Images of Different Stages of Nanoparticle Uptake.** (i) High resolution EM slices with manually segmented [Pixel size: 5x5 nm (xy); 5x15(xz) / Scale bar: 500 nm]; (ii) High resolution EM with correlated LM slice [Red: TfR / Green: AuNp / Cyan: Nucleus / Blue: Clathrin] [Pixel size: 5x5 nm (xy) / Scale bar: 500 nm]; (iii) Segmented 3D volume of EM data [Pink: Cell Membrane/ Green: AuNp/ Blue: Low UA/OsO<sub>4</sub> labelling density (e.g. Vacuoles)/ Red: High UA/OsO<sub>4</sub> labelling density (e.g. high lipid/protein content)/ Yellow: Lysosome/] [Voxel size: 5x5x15 nm (xyz) / Scale bar: 500 nm]. **A) Np-Tf Receptor (TfR) co-localised on cellular extension:** Nanoparticles bind to transferrin receptors on cellular protrusions.; **B) Np-TfR co-localised at cell surface:** Nanoparticles can also bind receptors on the surface of cells, but it is difficult to determine whether this is a result of the retraction of cellular protrusions or a separate cellular event. Some examples have indicated the presence of microtubules consistent with the retraction of cilia from the surface of the cell, whilst others are absent.; **C) Np transport in endocytic vesicle:** This shows the a nanoparticle completely engulfed by a cell in an endocytic vesicle; **D) Np-TfR co-localised engulfing of Np:** The nanoparticle is engulfed in a pocket of cellular membrane at the base of a cilium before uptake is completed and is connected to an early endosome. **E) Np-TfR co-localised in an early endosome:** As the nanoparticle is transported deeper into the cytosol, it fuses with an early endosome for sorting and further trafficking.; **F) Np in a Late Endosome/Multi-vesicular body (MVB):** In later stages of nanoparticle trafficking the nanoparticle is transported to a late endosome in preparation for degradation and exocytosis.; **G) Lysosomal fusion and formation of Autolysosome:** The nanoparticle is degraded further in a lysosomal vesicle within the autolysosome. ; **H) Particle Excretion:** After trafficking through the cell the nanoparticle is excreted from the cell.

Over the course of several experiments a number of common structural components were observed between cells. Above are exemplar images of the types of structures associated with nanoparticle receptor mediated uptake.

Vesicles in EM volumes were classified as follows:

**Endocytic vesicles:** Small spheroid vesicles (50-200nm diameter) with densely labelled membranes and an inner vacuole.

**Early endosomes:** Large irregular vesicles (200-400nm diameter) with densely labelled reticulated membranes and compartments with an inner vacuole.

**Late endosome/Multi-vesicular bodies:** Large spherical vesicles (400-700nm diameter) containing several smaller (20-100nm) vesicles.

**Lysosome:** Large (400-700nm) very densely labelled spherical vesicles with no visible inner compartments.

**Autolysosome:** Extremely large (0.7-1.5µm) vesicle containing multiple compartments of varying labelling density and containing at least one lysosome.

### 4.3. Methodology

#### 4.3.1. Human Cell Culture

A549 Human lung carcinoma epithelial cells obtained from American Type Culture Collection (ATCC) were grown in cell culture flasks obtained from Stratagene with 10ml Complete Minimum Essential Media (cMEM) which contained Minimal Essential Medium obtained from Corning with 10% Fetal Bovine Serum (FBS) obtained from Gibco and 1% Penicillin-Streptomycin (Pen-Stryp) obtained from Invitrogen in a humidified 37°C incubator with 5% CO<sub>2</sub>. Cells were passaged 2-3 times a week depending on the confluency of the cultures when passaging i.e. between 60-90% coverage of flask surface, the medium was removed from the flasks and washed with Dulbecco's Phosphate Buffered Saline (DPBS) obtained from Gibco, the cells were incubated in 5ml of 0.05% trypsin-EDTA obtained from Gibco at 37°C for 3 min, the flasks were tapped gently to ensure complete detachment of cells from the flask surface. 7ml of cMEM was added to the cell suspension and the solution was transferred to a sterile 50ml Falcon tube. The falcon tube was spun down at 200g for 3 min and the supernatant removed. The cell pellet was resuspended in 10ml cMEM and mixed thoroughly. 10ul of the suspension was placed on a Neubauer haemocytometer and their concentration determined. A volume containing  $5 \times 10^5$  cells was mixed with 10ml cMEM in a fresh flask and incubated in an incubator at 37°C and 5% CO<sub>2</sub>. Cells were tested monthly for mycoplasma using the MycoAlert test kit (Lonza Inc. Allendale, NJ).

#### 4.3.2. Transferrin Absorbed Gold Nanoparticle Preparation

This method was performed with some adaptations as previously described in (Kelly et al., 2015). Gold Nanoparticles (100nm) were obtained from Cytodiagnosics. The nanoparticle solutions were mixed with holo-Transferrin protein powder obtained from Sigma-Aldrich to a final protein concentration of 1mg/ml and incubated for 1hr at 37°C. The solutions were removed from the water bath and spun down three times at 400g for 30 min, after each spin the supernatant was removed and replaced with an adequate volume of DPBS. The final nanoparticle solution concentration was measured using the Nanodrop Spectrometer 2000 and

normalized to a concentration of  $1 \times 10^7$  nanoparticles per ml at peak wavelength 572nm and using the molar extinction coefficients supplied by the Cytodiagnostics ( $1.57 \times 10^{11} \text{ M}^{-1} \text{ cm}^{-1}$ ).

#### 4.3.3. Tf-Gold Addition to A549 Cells

A549 cells were split according to section 4.4.1 and  $5 \times 10^4$  cells were mixed with 3ml of CMEM and added to each 35mm culture dish with Gridded Coverslips (No. 1.5) obtained from MatTek corporation. The dishes were incubated at 37°C with 5% CO<sub>2</sub> for 48 hours. The media was removed from each dish and the cells were washed with 3ml of warm DPBS and 1ml of freshly prepared and normalized Tf-Gold nanoparticle solution, as per section 3.1.2, was added to each dish. The cells were incubated at 37°C with 5% CO<sub>2</sub> for a period of time. The nanoparticle solution was removed and the cells were washed with 3ml of cold DPBS (4°C) and a fixative solution of 1ml of 0.2% (v/v) of glutaraldehyde and 4% paraformaldehyde in DPBS was added, the dishes were placed in the fridge and incubated at 4°C for 1hr. The fixative solution was mixed with 2ml of cold DPBS and removed.

The cells were washed once more with 3ml cold DPBS. For permeabilization experiments, a permeabilization solution of 3% (w/v) BSA and 0.1% (w/v) saponin in DPBS was added to the fixed cell and incubated at room temperature for 1hr. The cells were treated under the following according to section 4.4.4 depending on the experiment.

#### 4.3.4. Antibody Receptor Staining

Anti-Transferrin receptor rabbit IgG polyclonal antibody (1:100 in DPBS with 1% (w/v) Bovine Serum Albumin [BSA]) obtained from Abcam was added to each gridded coverslip and incubated at 4°C for 1hr or o/n. After incubation in primary antibody, the dish was washed with cold DPBS w/1%BSA and Anti-Rabbit IgG Alexa Fluor 647 antibody (1:200 in DPBS w/1%BSA) is added to the coverslip and incubated at 4°C for 1hr or overnight.

Anti-Clathrin heavy chain mouse IgG polyclonal antibody (1:200 in DPBS with BSA) obtained from Thermo Fisher was added to each gridded coverslip and incubated at 4°C for 1hr or o/n. After incubation in primary antibody, the dish was washed with cold DPBS w/1%BSA and

Anti-Mouse IgG Alexa Fluor 568 FluoroNanogold™ antibody (1:200 in DPBS w/1%BSA) obtained from Nanoprobes is added to the coverslip and incubated at 4°C for 1hr or overnight.

For double-stain experiments, the primary and secondary staining steps were performed sequentially with transferrin and anti-rabbit staining being performed first, as described above, followed by several washes of cold DPBS with 1% BSA and subsequent addition of the anti-clathrin and anti-mouse antibodies as described above. Depending on the experiment, 4',6-diamidino-2-phenylindole (DAPI) counterstaining was performed by adding 100µl of 10µg/ml solution to the cell and incubated in the dark for 1 min at room temperature and Phalloidin Alexa Fluor 546 staining was performed by adding 5µl of methanolic stock solution to 200µl DPBS. This solution was added to the cells and incubated in the dark at room temperature for 20 mins. After incubation the dish was washed with cold DPBS with 1% BSA, twice with cold (4°C) DPBS and finally suspended in cold DPBS. The dish was brought to the confocal microscopy core for imaging.

#### 4.3.5. Confocal Imaging of Tf-Gold Incubated A549

Imaging was carried out with an inverted confocal laser scanning microscope in reflection mode, LSM780 (Carl Zeiss, Jena, Germany) using both 63× oil immersion and 10× air objectives.

The argon laser line of wavelength 488nm (Power: 25mW /Pinhole diameter: 49.6µm/ {1.1AU} [0.34µm Z-slice] ) was used in conjunction with a dichroic mirror (MBS T80/R20, Carl Zeiss) and a Main Beam splitter (MBS) MBS-405 to visualize the gold nanoparticles with a photomultiplier tube (PMT) detector (detection window: 463-507nm). Other laser lines of wavelength 561nm (DPSSL561/ Power: 20mW/ Pinhole Diameter: 47.2µm [0.34µm Z-slice]/ Detection Window: 579 - 614nm) and 633nm (HeNe633) (Power: 5mW / Pinhole diameter: 46.9µm [0.34µm Z-slice]/ Detection Window: 650 - 695nm) was used in conjunction with a Main Beam Splitter (MBS 488/561/633, Carl Zeiss) to excite and detect fluorescence from Alexa Fluor 568 and Alexa Fluor 647 secondary antibodies respectively. Widefield mode was

used to locate ROIs and a series of three stacks at magnifications x10, x63 and a digital zoom x126 (2x63) were typically made for correlative work.

The first image at x10 (Image size: 512px x 512px / Pixel size: 1.66 $\mu$ m x 1.66 $\mu$ m [xy]) was to locate a ROI on the grid in widefield mode, the second was a x63 stack (Image size: 512px x 512px / Pixel size: 0.26 $\mu$ m x 0.26 $\mu$ m x 0.34 $\mu$ m [xyz]) of the ROI with widefield, 488nm and 561nm or 633nm laser lines to image the nanoparticles, transferrin receptors and clathrin molecules respectively. The third and final image was a x2 digital zoom of a x63 magnification (Image size: 512px x 512px / Pixel size: 0.130 $\mu$ m x 0.130 $\mu$ m x 0.34 $\mu$ m [xyz]) on a cell of interest. Several locations on the grid were taken in this manner.

#### 4.3.6. Fixing, Staining and Resin Embedding of Confocal Imaged A549 with Tf-Gold Nanoparticles for FIB-SEM

This method was performed with some adaptations as previously described in (Narayan et al., 2014). Immediately after imaging the DPBS was removed and replaced with 2% Glutaraldehyde in 0.1M Cocodylate Buffer at 4°C, the cells were incubated at 4°C for 1hr.

For SEM visualisation of the FluoroNanogold™ antibodies, the cells were incubated 3 times (x3) with 50mM glycine in PBS for 5 mins each to remove trace aldehydes which inhibit gold enhancement. The cells were washed with 1% BSA in PBS for 5 mins (x3) and finally for 5 mins (x3) in ddH<sub>2</sub>O. Using the Gold Enhance™ EM kit provided by Nanoprobes, 10 $\mu$ l each of Solution A and B were mixed and incubated at room temperature for 5 mins, 10 $\mu$ l each of Solution C and D were added. The 40 $\mu$ l were added to the cells and incubated for 5 mins. The cells were washed with ddH<sub>2</sub>O (x3).

After fixation and/or FluoroNanogold™ enhancement, the cells were washed with 0.1M of cold cocodylate buffer (x3) and incubated in 2% v/v OsO<sub>4</sub> in 0.1M cocodylate buffer with 0.5% (w/v) Ferrocyanide for 1hr at RT. The sample was washed with ddH<sub>2</sub>O (x3) and incubated in cold-filtered 2% v/v Uranyl Acetate (UA) in ddH<sub>2</sub>O at 4°C overnight. After o/n incubation the UA solution was removed and the sample washed with ddH<sub>2</sub>O (x3) and replaced with a 50:50

ethanol:water solution and incubated at RT for 5 mins (x3). The ethanol:water solution was replaced with 70:30 ethanol:water solution and incubated at RT for 5mins (x3), this process was repeated for 90:10 ethanol:water (x3) and 100% ethanol solutions for 10mins (x2). During the dehydration procedure with ethanol the epoxy resin was prepared by creating “Solution A” by mixing DoDecenylSuccinic Anhydride (DDSA) with Embed-812 resin both obtained from Electron Microscopy Sciences (EMS) to a ratio of 1g : 0.76g. Another solution, “Solution B” was also created by mixing resin hardener Methyl-5-Norbornene-2,3-dicarboxylic Anhydride (NMA), also obtained from EMS, with Embed-812 resin to a ratio of 0.87g : 1g.

10 min before addition to the cells, Solution A and B were mixed to a ratio of 2ml : 8ml and mixed thoroughly, immediately before addition to the cells the solution was mixed with 0.2ml of accelerator Tris-(Dimethylaminomethyl) Phenol (DMP-30) and heated to 60°C for 5 mins to remove bubbles. After incubation in 100% Ethanol, a 50:50 Resin:Ethanol mixture was added to the cells and incubated for 1hr at RT and placed on an orbital shaker at <5Hz. The resin:ethanol mixture was replaced with 70:30 resin:ethanol mixture and incubated for 1hr at RT on an orbital shaker. This was repeated for 90:10 resin:ethanol and 100% resin solutions. After incubation in 100% resin for 1hr, a fresh batch of 100% resin solution was prepared as previously described and added to the cells. The cells were incubated at RT for 1hr and moved to an oven to cure overnight at 60°C.

#### 4.3.7. FIB-SEM Imaging of Tf-Gold Incubated A549

This method was performed as described in Chapter 3.4.1.3.

#### 4.3.8. Electron Dispersive X-Ray Spectroscopy

EDS detection was performed on a FEI Helios 660 Nanolab FIB-SEM with a 5kV landing energy, a 200pA aperture and a windowless Oxford XMax Xtreme detector. Points of interest were selected for analysis and spectra were acquired from 30 sec exposures at a time. The data was graphed using NIST DTSA-II Jupiter software.

#### 4.3.9. Image Processing and Segmentation

##### 4.3.9.1. Confocal Image Processing

Confocal image file (.lsm) was opened in ImageJ and the channels were split (Red:568/647/Green:488/Widefield). The coverslip bottom was found using the green laser line reflectance and everything below this Z-height was removed. Each image was set to a 10-bit grayvalue (images were acquired at 12-bit) for both the Red and the Green channels. Both channels were merged together, saved as .tif and used for correlative alignment.

##### 4.3.9.2. FIB-SEM Image Processing

After SEM acquisition the individual image files (.tif) were aligned using IMOD's tiltcorr program and binned by 3 in the x and y planes to produce isotropic pixels of size 15nm x 15nm x 15nm (xyz). The images were opened in ImageJ and cropped to isolate ROIs, a representative slice was picked and an area of size 10 x 15 $\mu$ m was selected and using the "Enhance Contrast" function was used to normalize and equalize the whole stack relative to the chosen area. The stack was binned by 2 in the x,y and z axes to produce a final voxel size of 30 $\mu$ m<sup>3</sup>, the final stack was resliced down the Y axis to aid in manual alignment and registration of the light and electron images.

##### 4.3.9.3. Semi-Automated Registration, Alignment and Correlation of Light/Electron Images

Both Confocal and FIB-SEM stacks were moved to Icy software (De Chaumont et al., 2012) where both stacks were opened in the eC-CLEM plugin (Paul-Gilloteaux et al., 2017) and gold nanoparticles in both images were correlated. The eC-CLEM software aligned the stacks based on the coordinates of the shared registration marks. A non-rigid transformation was also performed on alignments that didn't correlate efficiently, this inefficient correlation is mainly due to the contraction of cells after dehydration and resin embedding.

#### 4.3.9.4. Segmentation and 3D representation of Cellular Structures

The correlated data set allowed for identification of nanoparticles of interest, i.e. nanoparticle/Transferrin receptor co-localisation. These nanoparticles of interest were selected in a high-resolution image stack (Voxel size: 5 nm x 5 nm x 15 nm) and cropped in volumes of 2  $\mu\text{m}^3$ . These volumes were moved to Slicer-3D software (Fedorov et al., 2012) for segmentation. Cellular structures of interest i.e. cell membrane, vacuoles and nanoparticles were manually segmented by using threshold values appropriate for the structures.

The broader implications of this study will be briefly discussed in Chapter 5.

## 5. Conclusion

### 5.1. General Conclusion and Future Directions

Our aim in this thesis was to develop and apply both conventional and novel approaches to solving biological problems in disease and treatment with transmission and scanning electron microscopy.

To this end, we produced several high-resolution maps of a key enzyme, p97, to determine the structural implications of disease-related mutants. How these structural changes lead to the variety of diseases presented by p97 mutants, beyond an increased ATPase rate, remains to be seen. However, the findings of this thesis provide critical evidence on key structural interactions which necessarily drive underlying p97-related disease processes. Future steps in answering this question could be to determine changes in co-factor interactions of p97 mutants and its consequent aberrant function on substrate processing using cryo-EM methods.

Furthermore, we developed a segmentation methodology which allows for the accurate and rapid segmentation and quantitation of human tissues on the scale of 10's of nanometres. This methodology has the potential to detect morphological changes in an individual's tissues well before disease may present, providing for early treatment and possible mitigation of some of the more deleterious effects of degenerative diseases. Future directions may include examining affected human tissues for ultrastructural hallmarks of degenerative disease, generating statistical descriptors to define preliminary changes in tissue morphology before major symptoms manifest.

Finally, we used CLEM methods to study the ultrastructural effects of nanoparticles on cells, with the aim of developing a platform for the analysis and study of nanotherapeutic agents' action on target tissues. Using confocal light microscopy, we were able to localise nanoparticles to receptors of interest and then applied electron microscopy to map, in high detail, the cellular environment and interactions around the nanoparticle as it bound the receptor and moved into the cell. Future experiments might include using knowledge gained from our studies to tailor nanoparticle constructs to precisely control their entry and

bioprocessing. This analysis could enable the development of novel nanotherapeutic systems for tackling more aggressive and troublesome diseases which are proving a challenge for modern pharmaceuticals.

## 5.2. Structure Elucidation of p97 Mutants Using Cryo-Electron Microscopy

In chapter 2, we presented biochemical and structural research data of p97 mutants. Prior to the start of our study, limited knowledge with respect to both the structure of p97 mutants and their precise effect on ATPase activity on a biomolecular scale was available. Previously, 3 X-ray crystallography structures of p97 mutants existed (ADP-p97<sup>R155H</sup>, ATP $\gamma$ S-p97<sup>R155H</sup> and ATP $\gamma$ S-p97<sup>A232E</sup>) (Tang & Xia, 2013), however these were of N-D1 truncated mutants, omitting the D2 domain, the principle source of ATPase activity in p97. Through our investigation of p97 mutant structures, we have presented a total of 8 cryo-EM structures below 5Å resolution of 4 different full-length p97 mutants, 5 of which are being presented here for the first time.

These structures allowed us to generate a common structural model for the observed increase in ATPase activity of MSP-1 related mutants, chiefly through the destabilisation of the N domain in the ADP-bound state, either through a loss of inter-domain interactions in p97<sup>R155H</sup> or through inter-protomer steric effects in p97<sup>A232E</sup>. As our structures were imaged in near-native environments we were able to develop a clearer picture of the destabilising effect of these p97 mutants by capturing the increased flexibility in the cryo-EM maps, previously observed in NMR studies (Schuetz & Kay, 2016). We also showed that two mutants, p97<sup>D592N</sup> and p97<sup>E470D</sup> act primarily through their ATP-bound states suggesting a second possible mechanism for the increased ATPase activity observed, through the disruption of interdomain communication. Although it is difficult to determine the exact interactions which lead to the observed structural changes, these mutants give us an insight into the critical importance of inter-protomer and inter-domain interactions which govern p97 activity.

In conclusion, these new unexpected structural insights aid in dissecting the fundamental protein interactions that regulate quaternary protein conformation and may help inform future

therapeutic interventions in p97-related disease. Our work further highlights the central role of the N domain in dictating the activity of p97, potentially generating new avenues of enquiry for the allosteric inhibition of WT p97 through N domain stabilisation in the ADP-bound conformation.

### 5.3. Understanding Tissue Ultrastructure Using High-Resolution Focused Ion Beam – Scanning Electron Microscopy

In chapter 3, we have presented a method for 3D segmentation and statistical analysis of human skeletal muscle volumes using a semi-automated segmentation framework. The results demonstrate that rapid analysis of 3D muscular mitochondrial characteristics in relatively large volumes ( $> 10,000 \mu\text{m}^3$ ) can be achieved consistently with high accuracy across multiple data sets. Our data collection approach enables rapid acquisition of large volumes at a rate of  $>1500 \mu\text{m}^3/\text{hr}$ . The acquisition rate is dependent on, among other variables, the pixel size which in turn determines the scanning area and resolution of subsequent volumes. Therefore, there is an inherent trade-off between resolution and volume acquisition rate. In this study, we determined that a  $15 \text{ nm}^2$  pixel area returned sufficient resolution and volume acquisition rate for statistical analysis of the mitochondria in the 3D image data from muscle tissue. The Weka machine learning (Arganda-Carreras et al., 2017) software was chosen specifically for its segmentation capabilities. The Weka software is a robust software that is professionally maintained by The University of Waikato in New Zealand. Weka's software is powerful and versatile, allowing it to be ported to various operating systems and be used as a component of larger software.

Our approach to full volume segmentation is to manually classify a small set of images and export the manually trained classifier to use on the entire data set. These methods are generalizable to a variety of other data sets. Large-scale, high-resolution volume segmentation and validation of multiple cellular components can be achieved by a single individual in an extremely brief timespan using our approach. We demonstrate this by generating multiple structural benchmark values for human skeletal muscle which will expand with time to cover all healthy age groups, opening the door for its use in the study of the structural implications of disease. We have also been able to delineate between two different morphological types found in muscular tissue demonstrating the sensitivity of this approach to detect and describe

structural differences on an individual scale. We illustrate the versatility of our methodology through a publicly available dataset of neural tissue, used as a standard to test automated segmentation approaches, to produce several 3D segmentations of major cellular organelles in less than 24 hr. This dataset was acquired and segmented at a spatial resolution of 5 nm<sup>3</sup>.

Currently, the majority of neuronal tissue segmentations (Zheng et al., 2018) are performed using manual tracing methods, however, due to its time-consuming nature, much of the intracellular detail is lost. Through the use of our approach, this information can be rescued and used in conjunction with manually traced data to build a complete picture of the sub-cellular environment in neuronal tissues.

In conclusion, we note that our approach, which is available online to any interested user, can be readily applied to a wide variety of biological problems, with minimal human input, from tackling large-scale population-wide studies to the sensitive high-resolution analysis of cellular components.

#### 5.4. Examining Nanotherapeutic Interactions on a Cellular Scale

In chapter 4, we presented methods for assessing the cell's behaviour and interactions which occur in nanoparticle-cell interactions. The study of nanotherapeutics is uniquely suited to the use of correlative light electron microscopic methods, we demonstrated the ability to track multiple biochemical tags of interest to high resolutions and determine the structural environment nanoparticles encounter along their path through the cell.

During our investigation, we were able to generate a correlated 3D map of a model nanoparticle system's interaction throughout the cell, using both light and electron information to identify the key checkpoints from initial binding to final excretion of the nanoparticle. Early in our research, we discovered that sample preparation methods are often designed explicitly for one mode of microscopy, in ways that are regularly to the exclusion of another, e.g. fixation methods that work best for EM samples often degrade protein antigenicity, making fluorescent labelling difficult. Therefore, our method, which retains the ultrastructural and antigenic character of the specimens throughout the light and electron microscopy processes will be

extremely useful in studying nanoparticle action in cells in the future. Although our method relies on gold nanoparticles as both markers for correlation (due to their reflective properties in both light and electron microscopies) and as subjects themselves for study, this method can be applied to a wide variety of nanoparticles with the use of standard fiducial markers to correlate the two microscopies.

In conclusion, our technique can be applied to a wide variety of nanotherapeutic studies and expanded to even higher-resolution techniques such as Stochastic Optical Reconstruction Microscopy (STORM) to identify and differentiate smaller cellular structures in electron microscopy, often key in the accurate targeting and distribution of therapeutic agents. Here we demonstrate a proof-of-concept, that defines the biomolecular makeup of nanoparticle-cell interactions on a nanoparticle by nanoparticle scale.

## Bibliography

- Achan, J., Talisuna, A. O., Erhart, A., Yeka, A., Tibenderana, J. K., Baliraine, F. N., Rosenthal, P. J., & D'Alessandro, U. (2011). Quinine, an old anti-malarial drug in a modern world: Role in the treatment of malaria. In *Malaria Journal*. <https://doi.org/10.1186/1475-2875-10-144>
- Aggarwal, B. B., Sethi, G., Baladandayuthapani, V., Krishnan, S., & Shishodia, S. (2007). Targeting cell signaling pathways for drug discovery: An old lock needs a new key. In *Journal of Cellular Biochemistry*. <https://doi.org/10.1002/jcb.21500>
- Anderson, D. J., Le Moigne, R., Djakovic, S., Kumar, B., Rice, J., Wong, S., Wang, J., Yao, B., Valle, E., Kiss von Soly, S., Madriaga, A., Soriano, F., Menon, M. K., Wu, Z. Y., Kampmann, M., Chen, Y., Weissman, J. S., Aftab, B. T., Yakes, F. M., ... Rolfe, M. (2015). Targeting the AAA ATPase p97 as an Approach to Treat Cancer through Disruption of Protein Homeostasis. *Cancer Cell*. <https://doi.org/10.1016/j.ccell.2015.10.002>
- Arganda-Carreras, I., Kaynig, V., Rueden, C., Eliceiri, K. W., Schindelin, J., Cardona, A., & Seung, H. S. (2017). Trainable Weka Segmentation: A machine learning tool for microscopy pixel classification. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btx180>
- Banerjee, S., Bartesaghi, A., Merk, A., Rao, P., Bulfer, S. L., Yan, Y., Green, N., Mroczkowski, B., Neitz, R. J., Wipf, P., Falconieri, V., Deshaies, R. J., Milne, J. L. S., Huryn, D., Arkin, M., & Subramaniam, S. (2016). 2.3 Å resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. *Science*. <https://doi.org/10.1126/science.aad7974>
- Bastola, P., Wang, F., Schaich, M. A., Gan, T., Freudenthal, B. D., Chou, T. F., & Chien, J. (2017). Specific mutations in the D1–D2 linker region of VCP/p97 enhance ATPase activity and confer resistance to VCP inhibitors. *Cell Death Discovery*. <https://doi.org/10.1038/cddiscovery.2017.65>
- Berning, M., Boergens, K. M., & Helmstaedter, M. (2015). SegEM: Efficient Image Analysis for High-Resolution Connectomics. *Neuron*. <https://doi.org/10.1016/j.neuron.2015.09.003>

- Bleck, C. K. E., Kim, Y., Willingham, T. B., & Glancy, B. (2018). Subcellular connectomic analyses of energy networks in striated muscle. *Nature Communications*. <https://doi.org/10.1038/s41467-018-07676-y>
- Blythe, E. E., Gates, S. N., Deshaies, R. J., & Martin, A. (2019). Multisystem Proteinopathy Mutations in VCP/p97 Increase NPLOC4·UFD1L Binding and Substrate Processing. *Structure*. <https://doi.org/10.1016/j.str.2019.09.011>
- Blythe, E. E., Olson, K. C., Chau, V., & Deshaies, R. J. (2017). Ubiquitin- And ATP-dependent unfoldase activity of P97/VCP•NPLOC4•UFD1L is enhanced by a mutation that causes multisystem proteinopathy. *Proceedings of the National Academy of Sciences of the United States of America*. <https://doi.org/10.1073/pnas.1706205114>
- Bulfer, S. L., Chou, T. F., & Arkin, M. R. (2016). P97 Disease Mutations Modulate Nucleotide-Induced Conformation to Alter Protein-Protein Interactions. *ACS Chemical Biology*. <https://doi.org/10.1021/acscchembio.6b00350>
- Caffrey, B. J., Maltsev, A. V., Gonzalez-Freire, M., Hartnell, L. M., Ferrucci, L., & Subramaniam, S. (2019). Semi-automated 3D segmentation of human skeletal muscle using Focused Ion Beam-Scanning Electron Microscopic images. *Journal of Structural Biology*, 207(1), 1–11. <https://doi.org/10.1016/j.jsb.2019.03.008>
- Camacho, D. M., Collins, K. M., Powers, R. K., Costello, J. C., & Collins, J. J. (2018). Next-Generation Machine Learning for Biological Networks. In *Cell*. <https://doi.org/10.1016/j.cell.2018.05.015>
- Chen, S., Beeby, M., Murphy, G. E., Leadbetter, J. R., Hendrixson, D. R., Briegel, A., Li, Z., Shi, J., Tocheva, E. I., Müller, A., Dobro, M. J., & Jensen, G. J. (2011). Structural diversity of bacterial flagellar motors. *EMBO Journal*. <https://doi.org/10.1038/emboj.2011.186>
- Cooney, I., Han, H., Stewart, M. G., Carson, R. H., Hansen, D. T., Iwasa, J. H., Price, J. C., Hill, C. P., & Shen, P. S. (2019). Structure of the Cdc48 segregase in the act of unfolding an

authentic substrate. *Science*. <https://doi.org/10.1126/science.aax0486>

Dalal, S., Rosser, M. F. N., Cyr, D. M., & Hanson, P. I. (2004). Distinct Roles for the AAA ATPases NSF and p97 in the Secretory Pathway. *Molecular Biology of the Cell*. <https://doi.org/10.1091/mbc.E03-02-0097>

De Chaumont, F., Dallongeville, S., Chenouard, N., Hervé, N., Pop, S., Provoost, T., Meas-Yedid, V., Pankajakshan, P., Lecomte, T., Le Montagner, Y., Lagache, T., Dufour, A., & Olivo-Marin, J. C. (2012). Icy: An open bioimage informatics platform for extended reproducible research. In *Nature Methods*. <https://doi.org/10.1038/nmeth.2075>

Dice, L. R. (1945). Measures of the Amount of Ecologic Association Between Species. *Ecology*. <https://doi.org/10.2307/1932409>

Ernsting, M. J., Murakami, M., Roy, A., & Li, S. D. (2013). Factors controlling the pharmacokinetics, biodistribution and intratumoral penetration of nanoparticles. In *Journal of Controlled Release*. <https://doi.org/10.1016/j.jconrel.2013.09.013>

Erzberger, J. P., & Berger, J. M. (2006). Evolutionary relationships and structural mechanisms of AAA+ proteins. In *Annual Review of Biophysics and Biomolecular Structure*. <https://doi.org/10.1146/annurev.biophys.35.040405.101933>

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J. C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., Buatti, J., Aylward, S., Miller, J. V., Pieper, S., & Kikinis, R. (2012). 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging*. <https://doi.org/10.1016/j.mri.2012.05.001>

Fernández-Sáiz, V., & Buchberger, A. (2010). Imbalances in p97 co-factor interactions in human proteinopathy. *EMBO Reports*. <https://doi.org/10.1038/embor.2010.49>

Gatter, K. C., Brown, G., Strowbridge, I., Woolston, R. E., & Mason, D. Y. (1983). Transferrin receptors in human tissues: Their distribution and possible clinical relevance. *Journal of*

*Clinical Pathology*. <https://doi.org/10.1136/jcp.36.5.539>

Girardeau-Montaut, D. (2015). CloudCompare: 3D point cloud and mesh processing software.

Webpage: <https://www.Cloudcompare.Org>.

Glancy, B., Hartnell, L. M., Combs, C. A., Fenmou, A., Sun, J., Murphy, E., Subramaniam, S., & Balaban, R. S. (2017). Power Grid Protection of the Muscle Mitochondrial Reticulum. *Cell Reports*. <https://doi.org/10.1016/j.celrep.2017.03.063>

Glancy, B., Hartnell, L. M., Malide, D., Yu, Z. X., Combs, C. A., Connelly, P. S., Subramaniam, S., & Balaban, R. S. (2015). Mitochondrial reticulum for cellular energy distribution in muscle. *Nature*. <https://doi.org/10.1038/nature14614>

Gonzalez-Freire, M., Scalzo, P., D'Agostino, J., Moore, Z. A., Diaz-Ruiz, A., Fabbri, E., Zane, A., Chen, B., Becker, K. G., Lehrmann, E., Zukley, L., Chia, C. W., Tanaka, T., Coen, P. M., Bernier, M., de Cabo, R., & Ferrucci, L. (2018). Skeletal muscle ex vivo mitochondrial respiration parallels decline in vivo oxidative capacity, cardiorespiratory fitness, and muscle strength: The Baltimore Longitudinal Study of Aging. *Aging Cell*. <https://doi.org/10.1111/accel.12725>

Grant, T., & Grigorieff, N. (2015). Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *ELife*. <https://doi.org/10.7554/eLife.06980>

Guehrs, E., Schneider, M., Günther, C. M., Hessing, P., Heitz, K., Wittke, D., López-Serrano Oliver, A., Jakubowski, N., Plendl, J., Eisebitt, S., & Haase, A. (2017). Quantification of silver nanoparticle uptake and distribution within individual human macrophages by FIB/SEM slice and view. *Journal of Nanobiotechnology*. <https://doi.org/10.1186/s12951-017-0255-8>

Gunji, T., Wakita, M., & Kobayashi, S. (1980). Conductive staining in SEM with especial reference to tissue transparency. *Scanning*. <https://doi.org/10.1002/sca.4950030317>

- Hänzelmann, P., & Schindelin, H. (2011). The structural and functional basis of the p97/valosin-containing protein (VCP)-interacting motif (VIM): Mutually exclusive binding of cofactors to the n-terminal domain of p97. *Journal of Biological Chemistry*. <https://doi.org/10.1074/jbc.M111.274506>
- Hänzelmann, P., & Schindelin, H. (2017). The interplay of cofactor interactions and post-translational modifications in the regulation of the AAA+ ATPase p97. In *Frontiers in Molecular Biosciences*. <https://doi.org/10.3389/fmolb.2017.00021>
- Harris, A. K., Meyerson, J. R., Matsuoka, Y., Kuybeda, O., Moran, A., Bliss, D., Das, S. R., Yewdell, J. W., Sapiro, G., Subbarao, K., & Subramaniam, S. (2013). Structure and accessibility of HA trimers on intact 2009 H1N1 pandemic influenza virus to stem region-specific neutralizing antibodies. *Proceedings of the National Academy of Sciences of the United States of America*. <https://doi.org/10.1073/pnas.1214913110>
- Healy, S. J. M., Gorman, A. M., Mousavi-Shafaei, P., Gupta, S., & Samali, A. (2009). Targeting the endoplasmic reticulum-stress response as an anticancer strategy. In *European Journal of Pharmacology*. <https://doi.org/10.1016/j.ejphar.2009.06.064>
- Horák, V. (1983). A successive histochemical staining for succinate dehydrogenase and “reversed”-ATPase in a single section for the skeletal muscle fibre typing. *Histochemistry*. <https://doi.org/10.1007/BF00496207>
- Januszewski, M., Kornfeld, J., Li, P. H., Pope, A., Blakely, T., Lindsey, L., Maitin-Shepard, J., Tyka, M., Denk, W., & Jain, V. (2018). High-precision automated reconstruction of neurons with flood-filling networks. *Nature Methods*. <https://doi.org/10.1038/s41592-018-0049-4>
- Johnson, J. O., Mandrioli, J., Benatar, M., Abramzon, Y., Van Deerlin, V. M., Trojanowski, J. Q., Gibbs, J. R., Brunetti, M., Gronka, S., Wu, J., Ding, J., McCluskey, L., Martinez-Lage, M., Falcone, D., Hernandez, D. G., Arepalli, S., Chong, S., Schymick, J. C., Rothstein, J., ... Traynor, B. J. (2010). Exome Sequencing Reveals VCP Mutations as a Cause of Familial ALS. *Neuron*. <https://doi.org/10.1016/j.neuron.2010.11.036>

- Kasaragod, D., Makita, S., Hong, Y.-J., & Yasuno, Y. (2018). Machine-learning based segmentation of the optic nerve head using multi-contrast Jones matrix optical coherence tomography with semi-automatic training dataset generation. *Biomedical Optics Express*. <https://doi.org/10.1364/boe.9.003220>
- Kelly, P. M., Åberg, C., Polo, E., O'Connell, A., Cookman, J., Fallon, J., Krpetić, Ž., & Dawson, K. A. (2015). Mapping protein binding sites on the biomolecular corona of nanoparticles. *Nature Nanotechnology*. <https://doi.org/10.1038/nnano.2015.47>
- Kirkland, J. L. (1985). Normal Human Aging: The Baltimore Longitudinal Study of Aging by Nathan W. Shock, et al. U. S. Department of Health and Human Services. Publication No. 84-2450 (U.S. Supt of Docs) November, 1984. pp. 663. *Canadian Journal on Aging / La Revue Canadienne Du Vieillissement*. <https://doi.org/10.1017/s0714980800015920>
- Lange, C., Chesov, D., Heyckendorf, J., Leung, C. C., Udwardia, Z., & Dheda, K. (2018). Drug-resistant tuberculosis: An update on disease burden, diagnosis and treatment. *Respirology*. <https://doi.org/10.1111/resp.13304>
- Lewinson, D. (1989). Application of the ferrocyanide-reduced osmium method for mineralizing cartilage: Further evidence for the enhancement of intracellular glycogen and visualization of matrix components. *The Histochemical Journal*. <https://doi.org/10.1007/BF01757178>
- Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., ... Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallographica Section D: Structural Biology*. <https://doi.org/10.1107/S2059798319011471>
- Lucchi, A., Smith, K., Achanta, R., Knott, G., & Fua, P. (2012). Supervoxel-based segmentation of mitochondria in em image stacks with learned shape features. *IEEE Transactions on Medical Imaging*. <https://doi.org/10.1109/TMI.2011.2171705>

- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*. <https://doi.org/10.11613/bm.2012.031>
- Medical gallery of Blausen Medical 2014*. (2014). WikiJournal of Medicine. <https://doi.org/10.15347/wjm/2014.010>
- Meijs, M., Patel, A., Van De Leemput, S. C., Prokop, M., Van Dijk, E. J., De Leeuw, F. E., Meijer, F. J. A., Van Ginneken, B., & Manniesing, R. (2017). Robust Segmentation of the Full Cerebral Vasculature in 4D CT of Suspected Stroke Patients. *Scientific Reports*. <https://doi.org/10.1038/s41598-017-15617-w>
- Minetti, C., Sotgia, F., Bruno, C., Scartezzini, P., Broda, P., Bado, M., Masetti, E., Mazzocco, M., Egeo, A., Donati, M. A., Volonte, D., Galbiati, F., Cordone, G., Bricarelli, F. D., Lisanti, M. P., & Zara, F. (1998). Mutations in the caveolin-3 gene cause autosomal dominant limb-girdle muscular dystrophy. *Nature Genetics*. <https://doi.org/10.1038/ng0498-365>
- Murphy, G. E., Narayan, K., Lowekamp, B. C., Hartnell, L. M., Heymann, J. A. W., Fu, J., & Subramaniam, S. (2011). Correlative 3D imaging of whole mammalian cells with light and electron microscopy. *Journal of Structural Biology*. <https://doi.org/10.1016/j.jsb.2011.08.013>
- Nakane, T., Kimanius, D., Lindahl, E., & Scheres, S. H. W. (2018). Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *ELife*. <https://doi.org/10.7554/eLife.36861>
- Narayan, K., Danielson, C. M., Lagarec, K., Lowekamp, B. C., Coffman, P., Laquerre, A., Phaneuf, M. W., Hope, T. J., & Subramaniam, S. (2014). Multi-resolution correlative focused ion beam scanning electron microscopy: Applications to cell biology. *Journal of Structural Biology*. <https://doi.org/10.1016/j.jsb.2013.11.008>
- Narayan, K., & Subramaniam, S. (2015). Focused ion beams in biology. In *Nature Methods*. <https://doi.org/10.1038/nmeth.3623>

- Niwa, H., Ewens, C. A., Tsang, C., Yeung, H. O., Zhang, X., & Freemont, P. S. (2012). The role of the N-domain in the atpase activity of the mammalian AAA ATPase p97/VCP. *Journal of Biological Chemistry*. <https://doi.org/10.1074/jbc.M111.302778>
- Ogata, T., & Yamasaki, Y. (1997). Ultra-high-resolution scanning electron microscopy of mitochondria and sarcoplasmic reticulum arrangement in human red, white, and intermediate muscle fibers. *Anatomical Record*. [https://doi.org/10.1002/\(SICI\)1097-0185\(199706\)248:2<214::AID-AR8>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1097-0185(199706)248:2<214::AID-AR8>3.0.CO;2-S)
- Paul-Gilloteaux, P., Heiligenstein, X., Belle, M., Domart, M. C., Larijani, B., Collinson, L., Raposo, G., & Salamero, J. (2017). EC-CLEM: Flexible multidimensional registration software for correlative microscopies. In *Nature Methods*. <https://doi.org/10.1038/nmeth.4170>
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*. <https://doi.org/10.1002/jcc.20084>
- Pfeffer, S., Brandt, F., Hrabe, T., Lang, S., Eibauer, M., Zimmermann, R., & Förster, F. (2012). Structure and 3D arrangement of endoplasmic reticulum membrane-associated ribosomes. *Structure*. <https://doi.org/10.1016/j.str.2012.06.010>
- Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*. <https://doi.org/10.1038/nmeth.4169>
- Qian, Z. M., Li, H., Sun, H., & Ho, K. (2002). Targeted drug delivery via the transferrin receptor-mediated endocytosis pathway. In *Pharmacological Reviews*. <https://doi.org/10.1124/pr.54.4.561>
- Ritz, D., Vuk, M., Kirchner, P., Bug, M., Schütz, S., Hayer, A., Bremer, S., Lusk, C., Baloh, R. H., Lee, H., Glatter, T., Gstaiger, M., Aebersold, R., Weihl, C. C., & Meyer, H. (2011).

Endolysosomal sorting of ubiquitylated caveolin-1 is regulated by VCP and UBXD1 and impaired by VCP disease mutations. *Nature Cell Biology*. <https://doi.org/10.1038/ncb2301>

Rodriguez, J. A., Ivanova, M. I., Sawaya, M. R., Cascio, D., Reyes, F. E., Shi, D., Sangwan, S., Guenther, E. L., Johnson, L. M., Zhang, M., Jiang, L., Arbing, M. A., Nannenga, B. L., Hattne, J., Whitelegge, J., Brewster, A. S., Messerschmidt, M., Boutet, S., Sauter, N. K., ... Eisenberg, D. S. (2015). Structure of the toxic core of  $\alpha$ -synuclein from invisible crystals. *Nature*. <https://doi.org/10.1038/nature15368>

Rose, R., Golosova, O., Sukhomlinov, D., Tiunov, A., & Prospero, M. (2019). Flexible design of multiple metagenomics classification pipelines with ugene. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/bty901>

Rosenthal, P. B., & Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *Journal of Molecular Biology*. <https://doi.org/10.1016/j.jmb.2003.07.013>

Schmid, B., Schindelin, J., Cardona, A., Longair, M., & Heisenberg, M. (2010). A high-level 3D visualization API for Java and ImageJ. *BMC Bioinformatics*. <https://doi.org/10.1186/1471-2105-11-274>

Schuetz, A. K., & Kay, L. E. (2016). A dynamic molecular basis for malfunction in disease mutants of p97/VCP. *ELife*. <https://doi.org/10.7554/eLife.20143>

Sciote, J. J., Rowleson, A. M., Hopper, C., & Hunt, N. P. (1994). Fibre type classification and myosin isoforms in the human masseter muscle. *Journal of the Neurological Sciences*. [https://doi.org/10.1016/0022-510X\(94\)90089-2](https://doi.org/10.1016/0022-510X(94)90089-2)

Shi, Z., Liu, S., Xiang, L., Wang, Y., Liu, M., Liu, S., Han, T., Zhou, Y., Wang, J., Cai, L., Gao, S., & Ji, Y. (2016). Frontotemporal dementia-related gene mutations in clinical dementia patients from a Chinese population. *Journal of Human Genetics*. <https://doi.org/10.1038/jhg.2016.92>

- Sjöström, M., Ängquist, K. -A, Bylund, A. -C, Fridén, J., Gustavsson, L., & Scherstén, T. (1982). Morphometric analyses of human muscle fiber types. *Muscle & Nerve*. <https://doi.org/10.1002/mus.880050708>
- Stach, L., & Freemont, P. S. (2017). The AAA+ ATPase p97, a cellular multitool. In *Biochemical Journal*. <https://doi.org/10.1042/BCJ20160783>
- Stone, J. L., & Norris, A. H. (1966). Activities and attitudes of participants in the Baltimore longitudinal study. *Journal of Gerontology*. <https://doi.org/10.1093/geronj/21.4.575>
- Tanaka, R., Takimoto, H., Yamasaki, T., & Higashi, A. (2018). Validity of time series kinematical data as measured by a markerless motion capture system on a flatland for gait assessment. *Journal of Biomechanics*. <https://doi.org/10.1016/j.jbiomech.2018.01.035>
- Tang, W. K., Li, D., Li, C. C., Esser, L., Dai, R., Guo, L., & Xia, D. (2010). A novel ATP-dependent conformation in p97 N-D1 fragment revealed by crystal structures of disease-related mutants. *EMBO Journal*. <https://doi.org/10.1038/emboj.2010.104>
- Tang, W. K., & Xia, D. (2013). Altered intersubunit communication is the molecular basis for functional defects of pathogenic p97 mutants. *Journal of Biological Chemistry*. <https://doi.org/10.1074/jbc.M113.488924>
- Twomey, E. C., Ji, Z., Wales, T. E., Bodnar, N. O., Ficarro, S. B., Marto, J. A., Engen, J. R., & Rapoport, T. A. (2019). Substrate processing by the Cdc48 ATPase complex is initiated by ubiquitin unfolding. *Science*. <https://doi.org/10.1126/science.aax1033>
- Varela, J. A., Åberg, C., Simpson, J. C., & Dawson, K. A. (2015). Trajectory-Based Co-Localization Measures for Nanoparticle-Cell Interaction Studies. *Small*. <https://doi.org/10.1002/sml.201401849>
- Wan, W., & Briggs, J. A. G. (2016). Cryo-Electron Tomography and Subtomogram Averaging. In *Methods in Enzymology*. <https://doi.org/10.1016/bs.mie.2016.04.014>

- Watts, G. D. J., Wymer, J., Kovach, M. J., Mehta, S. G., Mumm, S., Darvish, D., Pestronk, A., Whyte, M. P., & Kimonis, V. E. (2004). Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nature Genetics*. <https://doi.org/10.1038/ng1332>
- Zazzi, M., Hu, H., & Prospero, M. (2018). The global burden of HIV-1 drug resistance in the past 20 years. *PeerJ*. <https://doi.org/10.7717/peerj.4848>
- Zhang, X., Gui, L., Zhang, X., Bulfer, S. L., Sanghez, V., Wong, D. E., Lee, Y. J., Lehmann, L., Lee, J. S., Shih, P. Y., Lin, H. J., Iacovino, M., Weihl, C. C., Arkin, M. R., Wang, Y., & Chou, T. F. (2015). Altered cofactor regulation with disease-associated p97/VCP mutations. *Proceedings of the National Academy of Sciences of the United States of America*. <https://doi.org/10.1073/pnas.1418820112>
- Zhao, G., Zhou, X., Wang, L., Li, G., Schindelin, H., & Lennarz, W. J. (2007). Studies on peptide:N-glycanase-p97 interaction suggest that p97 phosphorylation modulates endoplasmic reticulum-associated degradation. *Proceedings of the National Academy of Sciences of the United States of America*. <https://doi.org/10.1073/pnas.0702966104>
- Zheng, Z., Lauritzen, J. S., Perlman, E., Robinson, C. G., Nichols, M., Milkie, D., Torrens, O., Price, J., Fisher, C. B., Sharifi, N., Calle-Schuler, S. A., Kmecova, L., Ali, I. J., Karsh, B., Trautman, E. T., Bogovic, J. A., Hanslovsky, P., Jefferis, G. S. X. E., Kazhdan, M., ... Bock, D. D. (2018). A Complete Electron Microscopy Volume of the Brain of Adult *Drosophila melanogaster*. *Cell*. <https://doi.org/10.1016/j.cell.2018.06.019>

## Appendix

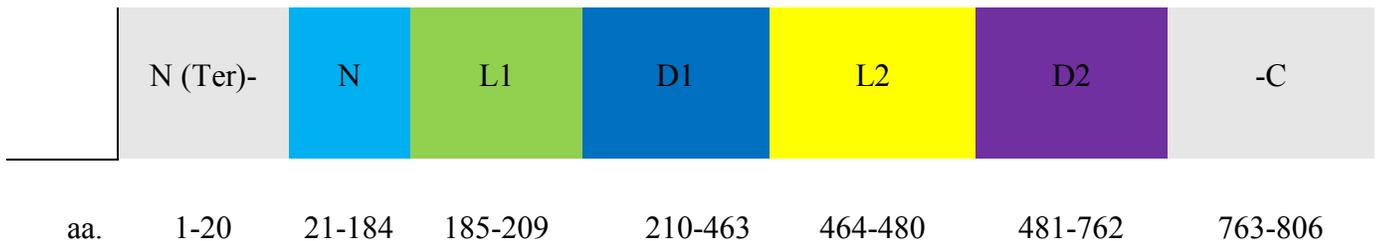
### Sequence 1: Wild-Type p97 Gene Sequence Codon Optimised

ATGGCCAGCGGTGCAGATTCGAAAGGCGACGACTTGAGCACTGCGATCCTGAAG  
CAAAAAAATCGTCCGAACCGCCTGATCGTTGATGAAGCGATCAACGAGGACAAT  
TCTGTTGTGAGCCTGAGCCAGCCGAAAATGGACGAACTGCAATTGTTCCGCGGTG  
ATACGGTCCTGCTGAAGGGTAAGAAACGTCGTGAAGCTGTGTGCATTGTCCTGA  
GCGATGACACCTGTAGCGACGAGAAAATTCGCATGAACCGTGTTGTCCGTAACA  
ATCTGCGTGTTTCGCCTGGGTGACGTTATTAGCATCCAGCCGTGTCCTGACGTCAA  
ATACGGTAAGCGTATTCACGTGTTGCCGATCGACGACACTGTCGAGGGTATCACT  
GGCAACCTGTTTGAAGTCTATTTGAAGCCGTACTTTCTGGAAGCATAACCGTCCGA  
TCCGTAAGGGTGACATCTTTCTGGTTCGCGGTGGCATGCGTGCGGTTGAGTTTAA  
AGTCGTGCAAACGGATCCGTCCCCGTATTGTATCGTTGCGCCGGATAACCGTGATT  
CACTGCGAGGGTGAGCCGATTAACGTGAAGATGAAGAAGAGTCCCTGAATGAG  
GTTGGCTACGATGACATCGGCGGCTGCCGCAAACAGCTGGCACAGATTAAGAA  
ATGGTTGAGCTGCCACTGCGCCACCCGGCTCTGTTCAAAGCGATTGGCGTGAAGC  
CGCCGCGTGGTATTCTGCTGTATGGCCC GCCAGGTACGGGCAAGACGCTGATCGC  
CCGTGCGGTGCGGAACGAAACCGGCGCTTTCTTTTTCTGATTAATGGTCCGGAA  
ATCATGAGCAAACCTGGCGGGTGAGAGCGAGTCTAACTTGCGTAAGGCATTCGAG  
GAAGCCGAGAAAAACGCGCCGGCAATTATTTTTATTGATGAATTAGATGCGATC  
GCTCCGAAACGTGAAAAACGCACGGTGAAGTGGAACGCCGCATCGTCAGCCAG  
CTGCTGACCTTGATGGACGGTCTGAAGCAACGTGCGCACGTCATCGTTATGGCAG  
CAACCAACCGCCCGAATAGCATCGACCCAGCGTTGCGTCGTTTTGGCCGTTTTGA  
TCGCGAAGTTGACATTGGTATTCCTGATGCGACGGGTCGTCTGGAAATTCTGCAG  
ATTCATACCAAGAACATGAAACTGGCAGACGATGTCGACTTGGAACAAGTGGCC  
AATGAAACCCATGGCCATGTGGGTGCGGATCTGGCCGCTCTGTGTAGCGAAGCG  
GCCCTGCAAGCGATTCGCAAGAAAATGGATCTGATTGACCTGGAAGATGAAACG  
ATCGACGCAGAAGTGATGAATTCGCTGGCAGTTACGATGGATGACTTTCGCTGG  
GCACTGAGCCAGTCCAATCCGAGCGCCCTGCGCGAAACTGTTGTTGAGGTGCCG  
CAAGTGACCTGGGAAGATATTGGTGGTCTGGAAGATGTTAAACGTGAGCTGCAA

GAACTGGTGCAGTACCCGGTGGAGCATCCGGATAAGTTCTTGAAGTTCGGTATG  
ACGCCGTCTAAAGGTGTTCTGTTTTATGGTCCGCCAGGTTGCGGCAAGACCCTGC  
TGGCAAAGGCTATCGCGAATGAATGCCAAGCTAACTTCATTAGCATTAAAGGGCC  
CGGAACTGCTGACCATGTGGTTCGGCGAGTCCGAGGCCAACGTCCGTGAGATTTT  
TGACAAGGCAAGACAGGCAGCGCCATGTGTCTTGTTTTTTCGATGAACTGGATAGC  
ATCGCCAAGGCACGTGGCGGCAACATTGGTGACGGTGGCGGGCGCCGCTGACCGT  
GTCATTAATCAAATTCTTACCGAAATGGACGGCATGAGCACCAAGAAAAACGTT  
TTCATCATCGGCGCAACGAATCGTCCGGATATCATCGATCCGGCAATTCTGCGCC  
CGGGTCGCCTGGATCAGCTGATTTACATTCCGTTGCCTGATGAGAAGTCCCGTGT  
CGCAATTCTGAAAGCGAATCTGCGCAAGTCCCCGGTCGCGAAAGACGTTGATCT  
GGAATTCTTAGCTAAGATGACCAACGGCTTCAGCGGCGCGGATTTGACCGAAAT  
CTGCCAGCGTGCCTGCAAACCTGGCGATCCGTGAGAGCATTGAGAGCGAGATCCG  
TCGTGAGCGTGAGCGTCAAACGAACCCGAGCGCCATGGAAGTTGAGGAAGATGA  
CCCAGTTCCGGAAATCCGTGCGGACCACTTTGAAGAAGCGATGCGTTTCGCGCGT  
CGTAGCGTTAGCGACAATGACATTCGTAAATACGAAATGTTTGCGCAGACCCTGC  
AGCAGAGCCGCGGTTTCGGCAGCTTCCGTTTTCTAGCGGCAACCAGGGTGGCGC  
GGGTCCTTCGCAAGGTTCTGGCGGGCGGCACCGGTGGCTCCGTGTACACCGAAGA  
TAATGACGACGACCTGTATGGTTAA

**Sequence 2: Wild-Type p97 protein amino acid sequence.** Mutated amino acids highlighted in red.

MASGADSKGDDLSTAILKQK **N**RPNRLIVDEAINEDNSVVSLSQPKMDELQLFRGDTV  
 LLKGKKRREAVCIVLSDDTCSDEKIRMNRVVRNNLRVRLGDVISIQPCPDVKYGKRI  
 HVLPIDDTVEGITGNLFEVYLKPYFLEAYRPIRKGDIFLV **R**GGMRAVEFKVVETDPSP  
 YCIVAPDVTVIHC **E**GEPIKREDEEESLNEVGYDDIGGC **R**KQLAQIKEMVELPLRHPALF  
**K** **A**IGVKPPRGILLYGPPGTGKTLIARAVANETGAFFFLINGPEIMSKLAGESESNLKKA  
 FEEAEKNAPAIIFIDELDAIAPKREKTHGEVERRIVSQLLTLMDGLKQRAHVIVMAAT  
 NRPNSIDPALRRFGRFDREVDIGIPDATGRLEILQIHTKNMKLADDVDLEQVANETHG  
 HVGADLAALCSEAALQAIRKKMDLIDLEDETIDAEVMNSLAVTMDDFRWALSQSNP  
**S** **A**LRETVV **E**VPQVTWEDIG **G**LEDVKRELQELVQYPVEHPDKFLKFGMTPSKGVLFY  
 GPPGCGKTLAKAIANECQANFISIKGPELLTMWFGSEANVREIFDKARQAAPCVLF  
 FDELDSIAKARGGNIG **D**GGGAADRVINQILTEMDGMSTKKNVFIIGATNRPDIIDPAIL  
 RRGRLDQLIYIPLPDEKSRVAILKANLRKSPVAKDVDLEFLAKMTNGFSGADLTEICQ  
 RACKLAIRESIESEIRRERERQTNPSAMEVEEDDPVPEIRRDHFEEAMRFARRSVSDN  
 DIRKYEMFAQTL **Q**QSRGFGSFRFSPGNQGGAGPSQGSGGGTGGSVYTEDNDDDLYG



**Table A.1: Media recipes.**

For agar preparations 15g L<sup>-1</sup> of agarose was added to the media.

| Medium                  | Ingredients and methods  |
|-------------------------|--|
| Luria Broth (LB)        | 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, 1.5% (w/v) agar (for plate) in 1L of dH <sub>2</sub> O, then adjusted to pH7 with 5N NaOH. Autoclaved and stored at 4°C. |
| LB + Kanamycin (LB-KAN) | After autoclaving, Kanamycin was added to a final concentration of 50 μg ml <sup>-1</sup> .  |

**Table A.2: *E. coli* Strains**

| Strain    | Genotype  | Source                    |
|-----------|---|---------------------------|
| BL21(DE3) | B F <sup>-</sup> ompT gal dcm lon hsdS <sub>B</sub> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB <sup>+</sup> ] <sub>K-12</sub> (λ <sup>S</sup> ) | Thermo Scientific         |
| DH5α      | F <sup>-</sup> endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(r <sub>K</sub> <sup>-</sup> m <sub>K</sub> <sup>+</sup> ), λ <sup>-</sup>                         | New England Biolabs (NEB) |

**Table A.3: Primer List for Site-Directed Mutagenesis**

| Pri<br>mer<br>Set. | Label | Sequence (5'-3') Forward ( <u>Red</u> = <u>Insertion mutation</u> ) | Sequence (5'-3') Reverse    |
|--------------------|-------|---|-----------------------------|
| 1                  | R155H | TTTCTGGTTC <u>A</u> CGGTGGCATG                                      | GATGTCACCCTTAC<br>GGATC     |
| 2                  | A232E | CTGTTCAAAG <u>A</u> GATTGGCGTGAAG<br>CCGCC                          | AGCCGGGTGGCGC<br>AGTGG      |
| 3                  | E470D | CTGTTGTTGA <u>C</u> GTGCCGCAAGTGA<br>CCTGG                          | TTTCGCGCAGGGC               |
| 4                  | D592N | CAACATTGGT <u>A</u> ACGGTGGCGGCG                                    | CCGCCACGTGCCTT<br>GGCG      |
| 5                  | p97-1 | GCGACGAGAAAATTCGCATGA   | GGCAGCTCAACCA<br>TTTCTTTAAT |
| 6                  | p97-2 | GCTGGCAGTTACGATGGATGA   | GCTCACGACGGAT<br>CTCGCT     |

**Table A.4: EM Data Collection and Reconstruction Statistics**

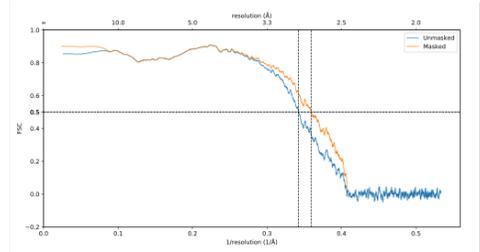
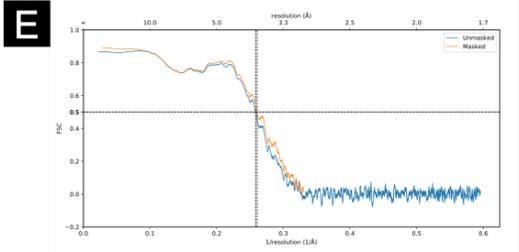
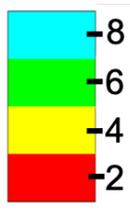
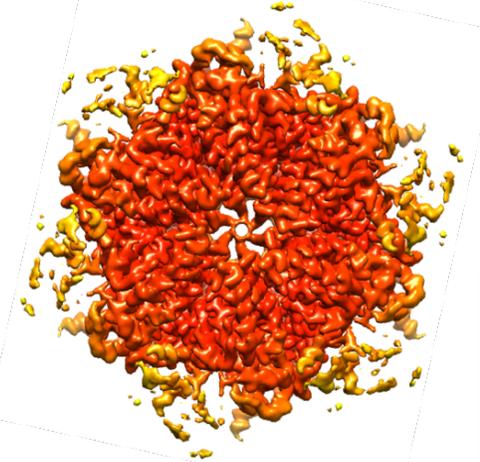
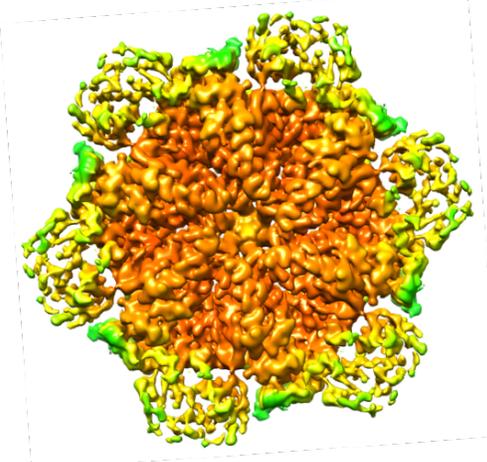
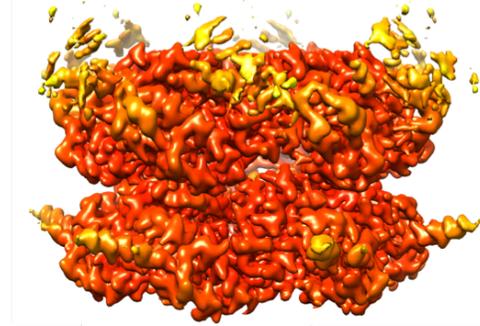
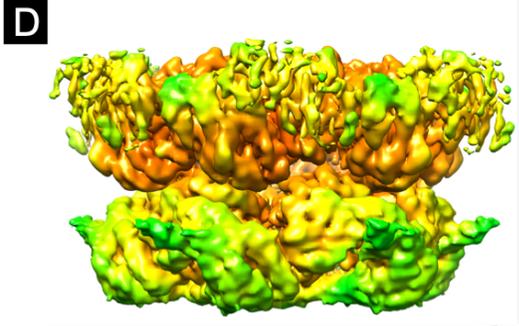
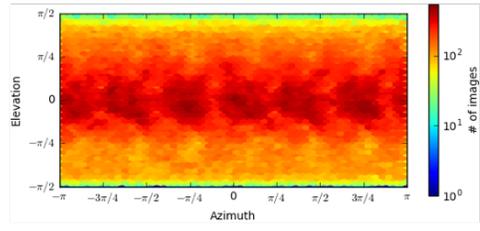
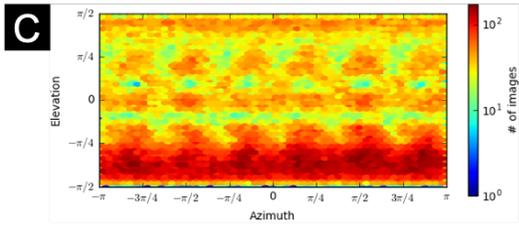
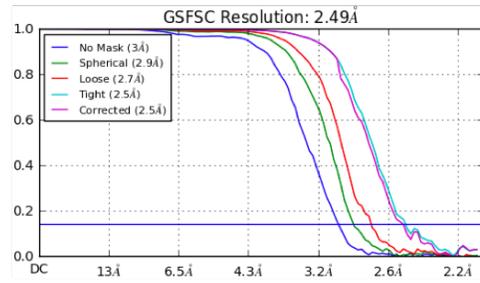
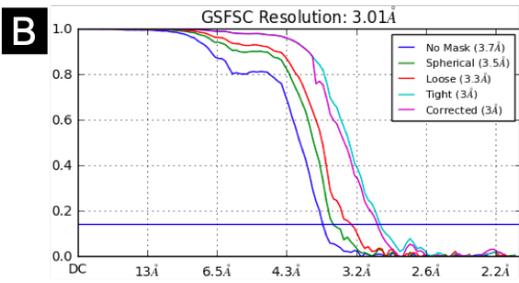
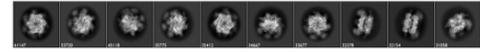
| Protein   | R155H-ADP       | R155H-ATP $\gamma$ S | A232E-ADP   | A232E-ATP $\gamma$ S | E470D-ADP    | E470D-ATP $\gamma$ S | D592N-ADP       | D592N-ATP $\gamma$ S |
|---|-----------------|----------------------|-------------|----------------------|--------------|----------------------|-----------------|----------------------|
| <b>Microscope</b>                               | FEI Titan Krios | FEI Titan Krios      | FEI Glacios | FEI Glacios          | FEI Glacios  | FEI Titan Krios      | FEI Titan Krios | FEI Titan Krios      |
| <b>Voltage (kV)</b>                             | 300             | 300                  | 200         | 200                  | 200          | 300                  | 300             | 300                  |
| <b>Detector</b>                                 | Gatan K3        | Gatan K3             | Falcon 3    | Falcon 3             | Falcon 3     | Falcon 3             | Gatan K3        | Gatan K3             |
| <b>Magnification (Nominal)</b>                  | 105000          | 105000               | 150000      | 150000               | 150000       | 72000                | 105000          | 105000               |
| <b>Pixel Size (Å)</b>                           | 0.324           | 0.324                | 0.76        | 0.98                 | 0.98         | 1.06                 | 0.324           | 0.415                |
| <b>Dose Rate (e<sup>-</sup>/pixel/sec)</b>      | 20              | 20                   |             |                      |              |                      | 20              | 20                   |
| <b>Frames per Exposure</b>                      | 40              | 40                   | 48          | 48                   | 48           | 48                   | 40              | 40                   |
| <b>Exposure (e<sup>-</sup>/Å<sup>2</sup>)</b>   | 50              | 50                   | 25          | 50                   | 50           | 50                   | 50              | 50                   |
| <b>Defocus Range (μm)</b>                       | -1 to -2.5      | -1 to -2.5           | -1 to -2    | -1 to -2             | -1 to -2     | -1 to -2             | -1 to -2.5      | -1 to -2.5           |
| <b>Micrographs Collected</b>                    | 4410            | 5265                 | 2569        | 640                  | 500          | 367                  | 6940            | 5926                 |
| <b>Particles Extracted/Final</b>                | 629745/146973   | 1293922/284092       | 74417/14388 | 310167/163041        | 151431/32694 | 34282/5698           | 697655/118863   | 3138385/1092658      |
| <b>Symmetry Imposed</b>                         | C6              | C6                   | C6          | C6                   | C6           | C6                   | C6              | C6                   |
| <b>Map Sharpening B-Factor</b>                  | 126             | 109.5                | 116         | 126.2                | 177.6        | 108.4                | 104.3           | 115.4                |
| <b>Unmasked resolution at 0.5/0.143 FSC (Å)</b> | 3.9/3.4         | 2.9/2.6              | 4.5/3.9     | 3.1/2.8              | 4.9/4.3      | 7.03/4.91            | 3.49/3.13       | 2.9/2.6              |
| <b>Masked resolution at 0.5/0.143 FSC (Å)</b>   | 3.8/3.3         | 2.8/2.5              | 4.4/3.8     | 3.0/2.8              | 4.8/4.3      | 6.78/4.70            | 3.32/3.09       | -/2.8                |

**Table A.5: Model Refinement and Validation Statistics**

| Protein                     | R155H-ADP | R155H-ATP $\gamma$ S | A232E-ADP | A232E-ATP $\gamma$ S | E470D-ADP | E470D-ATP $\gamma$ S | D592N-ADP | D592N-ATP $\gamma$ S |
|-----------------------------|-----------|----------------------|-----------|----------------------|-----------|----------------------|-----------|----------------------|
| <b>Composition</b>          |           |                      |           |                      |           |                      |           |                      |
| <b>Amino Acids</b>          | 4410      | 4422                 | 4410      | 4422                 | 4410      | 4422                 | 4410      | 4422                 |
| <b>AGS</b>                  | 0         | 12                   | 0         | 12                   | 0         | 12                   | 0         | 12                   |
| <b>ADP</b>                  | 12        | 0                    | 12        | 0                    | 12        | 0                    | 12        | 0                    |
| <b>MG</b>                   | 0         | 12                   | 0         | 12                   | 0         | 12                   | 0         | 12                   |
| <b>RMSD</b>                 |           |                      |           |                      |           |                      |           |                      |
| <b>Bonds (Å)</b>            | 0.024     | 0.019                | 0.004     | 0.009                | 0.005     | 0.008                | 0.004     | 0.005                |
| <b>Angles (°)</b>           | 1.736     | 1.871                | 0.799     | 0.89                 | 1.002     | 1.348                | 0.764     | 0.887                |
| <b>Mean B-Factors</b>       |           |                      |           |                      |           |                      |           |                      |
| <b>Amino Acids</b>          | 101.03    | 101.82               | 256.47    | 134.89               | 260.53    | 66.25                | 311       | 411.75               |
| <b>Ligand</b>               | 53.65     | 25.1                 | 180.69    | 70.25                | 203.57    | 22.51                | 251.08    | 374.15               |
| <b>Ramachandran</b>         |           |                      |           |                      |           |                      |           |                      |
| <b>Favoured (%)</b>         | 92.89     | 91.38                | 93.07     | 92.5                 | 93.55     | 90.38                | 93.25     | 90                   |
| <b>Allowed (%)</b>          | 6.25      | 7.5                  | 6.79      | 7.5                  | 6.32      | 9.21                 | 6.61      | 9.87                 |
| <b>Outliers (%)</b>         | 0.87      | 1.11                 | 0.14      | 0                    | 0.14      | 0.41                 | 0.14      | 0.14                 |
| <b>Rotamer Outliers (%)</b> | 0.96      | 0.96                 | 11.31     | 7.37                 | 1.12      | 16.75                | 0.16      | 1.44                 |
| <b>Clash Score</b>          | 25.08     | 24.49                | 16.46     | 19.37                | 23.59     | 14.27                | 23.42     | 34.68                |
| <b>C-beta outliers (%)</b>  | 0         | 0                    | 0         | 0                    | 0         | 0.15                 | 0         | 0                    |
| <b>CaBLAM outliers (%)</b>  | 3.99      | 6.31                 | 3.71      | 4.25                 | 4.4       | 6.42                 | 3.85      | 5.37                 |
| <b>CC (mask)</b>            | 0.8       | 0.82                 | 0.79      | 0.88                 | 0.82      | 0.77                 | 0.86      | 0.8                  |
| <b>MolProbity Score</b>     | 2.34      | 2.39                 | 2.96      | 2.91                 | 2.33      | 3.13                 | 2.3       | 2.69                 |

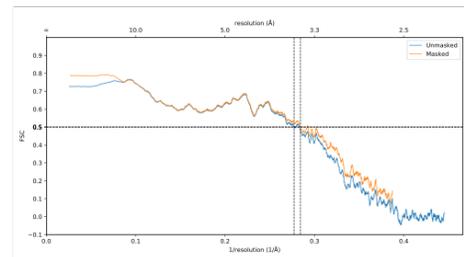
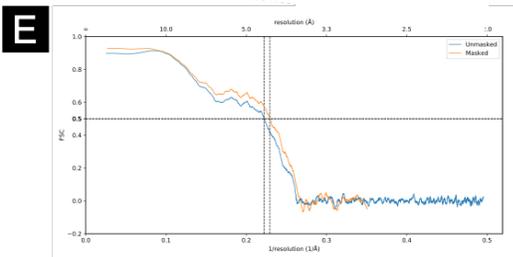
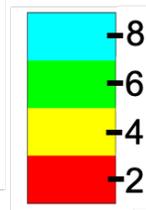
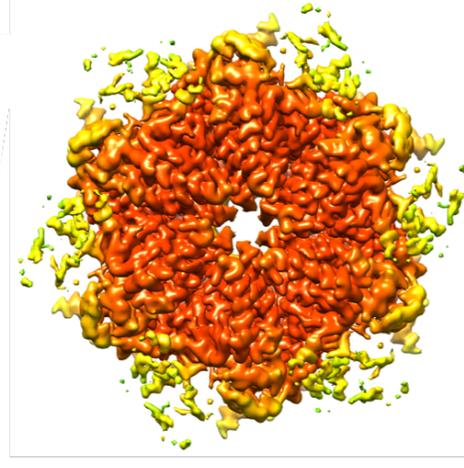
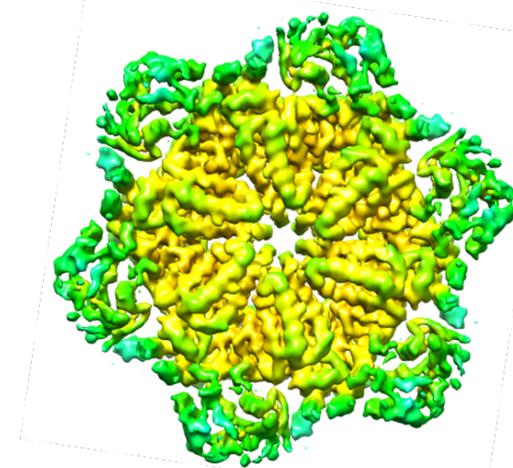
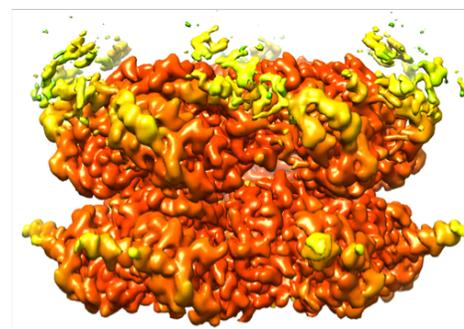
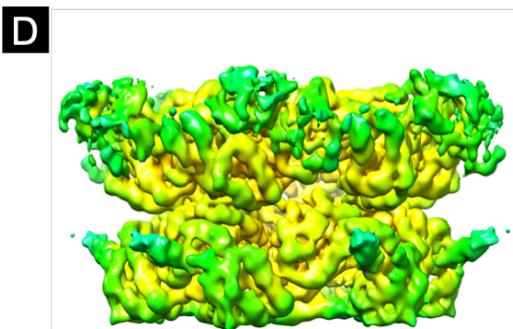
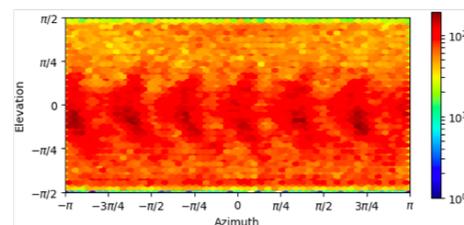
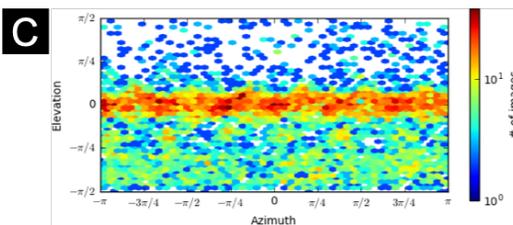
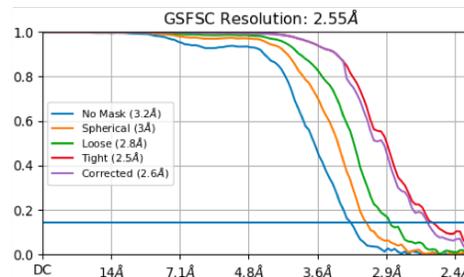
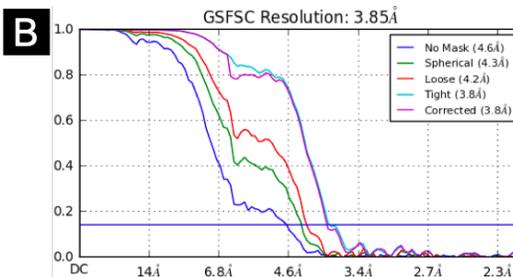
### R155H-ADP

### R155H-ATP $\gamma$ S



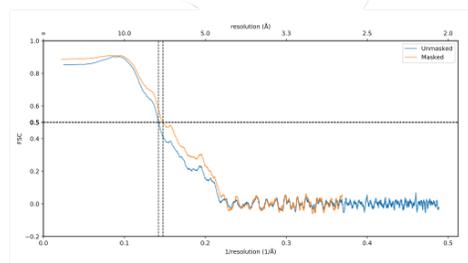
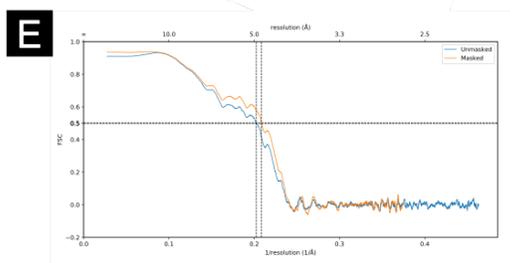
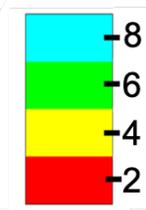
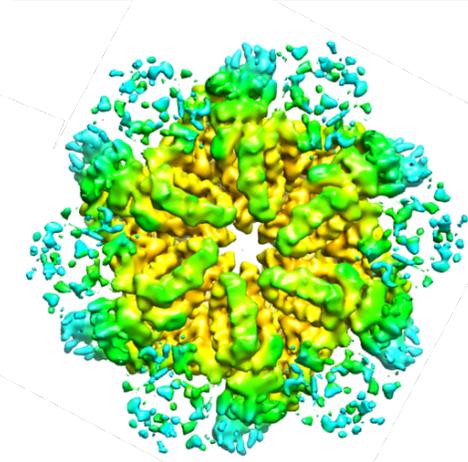
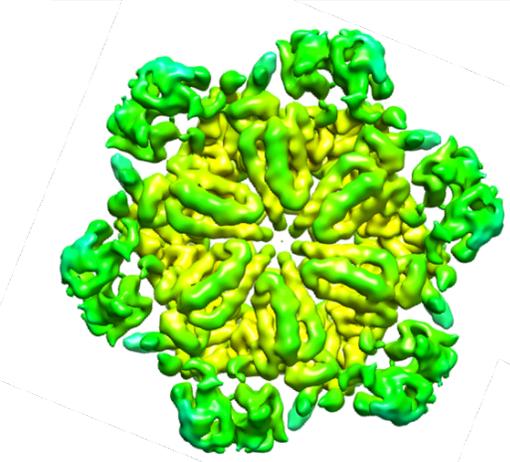
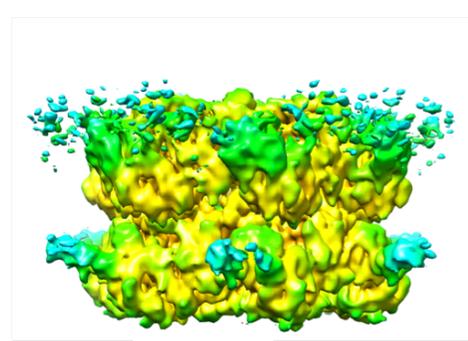
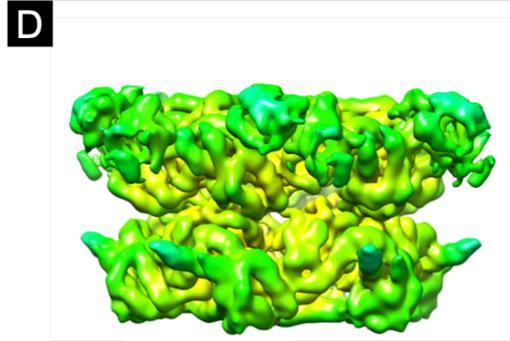
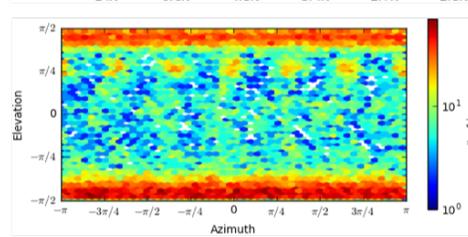
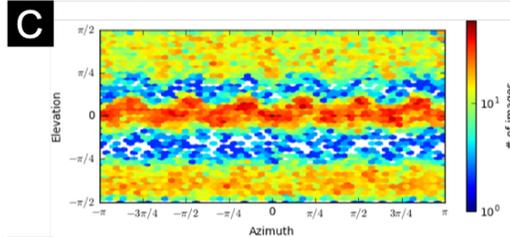
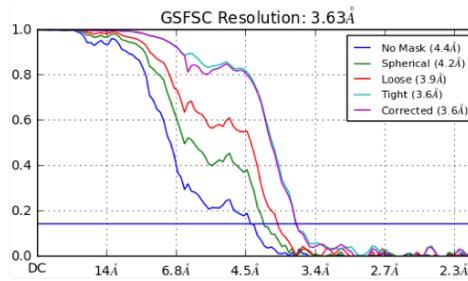
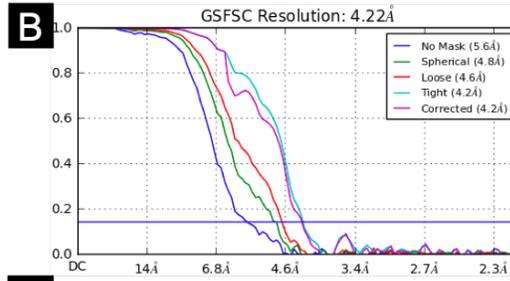
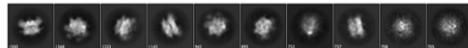
### A232E-ADP

### A232E-ATP $\gamma$ S



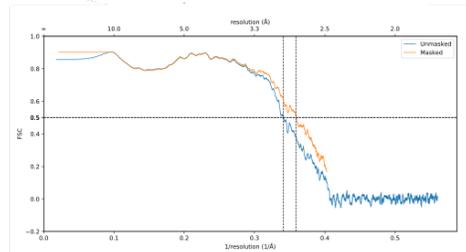
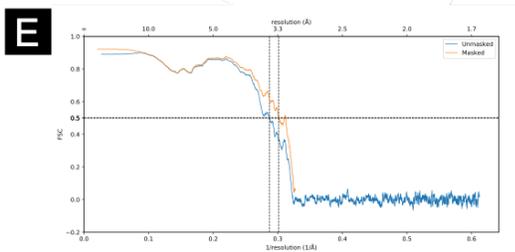
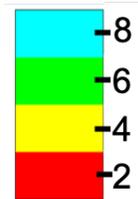
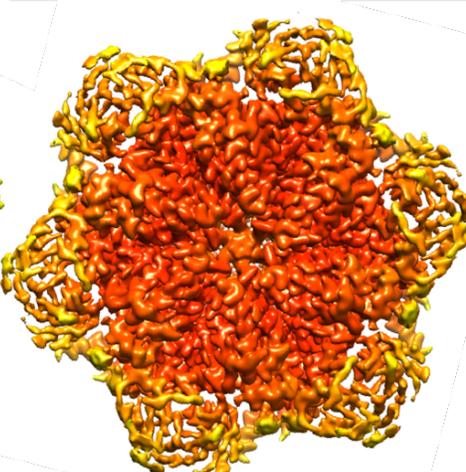
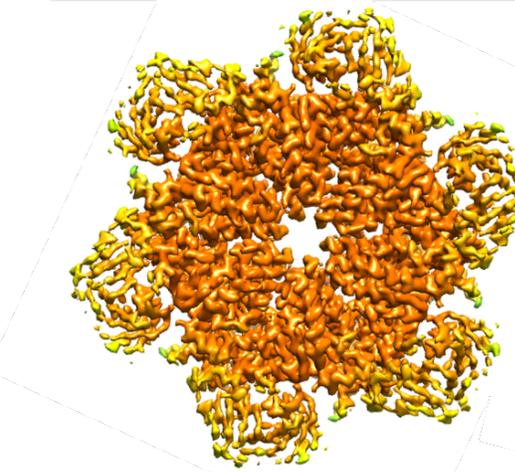
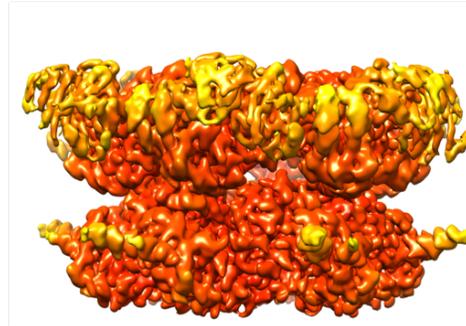
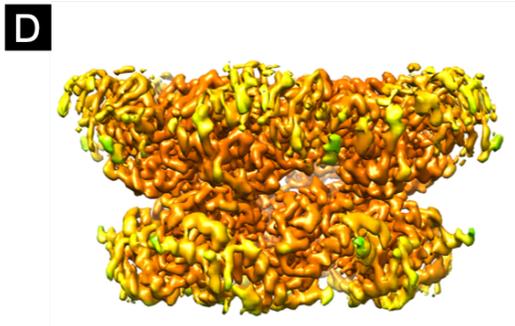
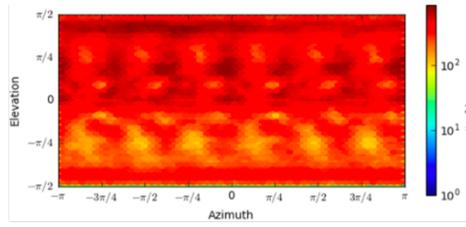
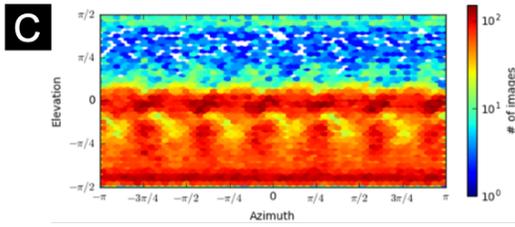
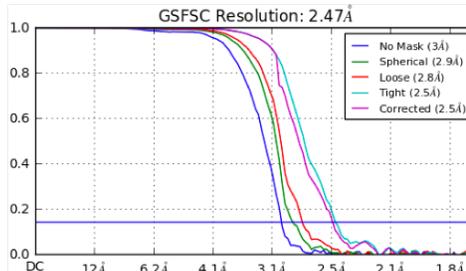
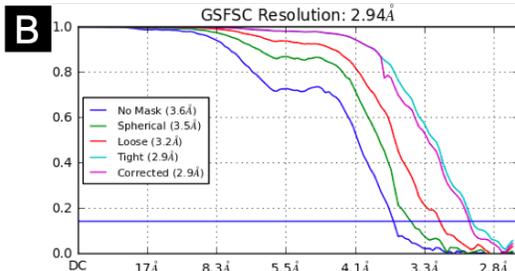
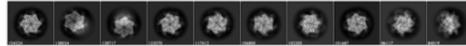
### E470D-ADP

### E470D-ATP $\gamma$ S



### D592N-ADP

### D592N-ATP $\gamma$ S



**Figure A.1: Cryo-EM Structure Validation.** **A)** Top ten 2D classes sorted by number of particles **B)** FSC curves and **C)** The viewing direction distribution plot for mutant density. **D)** Cryo-EM density of side-view of p97 mutants (top) and top-view (bottom) is shown, colored according to local resolution. Image on the right is a central slice through the density. **E)** Model-to-map FSC.