

**CHARACTERIZATION OF THE HUMAN ATYPICAL KINASE,
AARF DOMAIN-CONTAINING KINASE 3, AND ITS SUBSTRATES
REQUIRED FOR COENZYME Q BIOSYNTHESIS**

by

Sohee Yun

A thesis submitted to the Biochemistry and Cell Biology Graduate Program
in the Department of Biomedical and Molecular Sciences
in conformity with the requirements for
the degree of Master of Science

Queen's University

Kingston, Ontario, Canada

(January 2018)

Copyright © Sohee Yun, 2018

Abstract

Aarf domain-containing kinase 3, ADCK3, is a human atypical protein kinase located in the inner membrane of the mitochondria that is required for the biosynthesis of coenzyme Q (CoQ). Stable supply of CoQ is essential to human health as it plays a critical role as an electron transporter in the electron transport chain which is required for proper ATP production. Although the exact role of ADCK3 is unknown, patients possessing mutations within this gene are often diagnosed with autosomal recessive cerebellar ataxia type 2, which results from CoQ deficiency. To better understand the involvement of ADCK3 and other proteins in CoQ biosynthesis, their structure, function, and their interactions with each other were investigated. ADCK3 is hypothesized to regulate the CoQ biosynthesis using its kinase activity to activate several proteins found in the CoQ biosynthetic complex; however, previous work has failed to show direct evidence of its kinase activity and its functional interaction with the other Coq proteins.

In this work, using a soluble construct of ADCK3, we solved its structure by X-ray crystallography. Various constructs of ADCK3 were transfected into Human Embryonic Kidney cells in an attempt to extract and identify its substrates. Further, the genes of the potential substrates, Coq3, Coq5, and Coq7 were cloned into various expression vectors and transformed into *Escherichia coli* expression cell lines. After the extensive expression trials, the combination of N-terminal fusions with maltose-binding proteins and the RIPL *E. coli* cell expression line was found to facilitate the expression of Coq5 and Coq7. However, using radiometric kinase assays, we show that ADCK3 exhibited no kinase activity towards Coq5 and Coq7. Finally, this work also shows the methyltransferase activity of Coq5 using its substrate analog. It is anticipated that our results will serve as a foundation for further characterization of the proteins involved in the human CoQ₁₀ biosynthesis.

Acknowledgements

There are many I wish to thank for all the love and support that I have received throughout this journey. I've had many days where I questioned myself and my decisions, and I wouldn't have fought through this path without them.

First and foremost, I would like to express my highest gratitude to Dr. Zongchao Jia. Joining your research group as a thesis student was one of the most eye-opening experiences I have had in my undergraduate career. Thank you for seeing my potential and taking me on as a graduate student as it has given me the opportunity to further explore the scientific community and the exciting opportunities that await. I appreciate the guidance and input you have given me as they were invaluable for this project. I wish the best of luck to Dr. Jia and his future students.

I'd also like to thank everyone from the Jia lab. The technical assistance as well as the friendship from Natalie Roy was essential in many aspects of this project. Natalie, as well as the rest of the members of the Jia lab, created an overall positive learning experience. Dr. Mona Rahman and Dr. Laura Van Staalduinen were like the mother bears in the lab, always giving advice to optimize and trouble shoot our experiments and even our academic career. It was always nice to discuss the current topics and life in general with Yichen "Jerry" Zhang. Although Greg and Olena were in the other lab, we still had some great hallway chats. I'd also like to thank Brody Wheeler for setting up a great project and Alex Andrew for teaching me how to work with the HEK293 cells. It was also very nice to have met our visiting scholars, Dr. Bin Guo and Dr. Rongfeng Li, and our exchange student Shelley Yang.

It was my pleasure to have met and gotten to know my first and last 4th year thesis student, Gaelen Moore. From the start, he was highly motivated and excited about his opportunity to explore research. He learned extremely quickly, brought many great ideas on the table, and did a tremendous amount of work in the lab. I'm very proud of Gaelen and I would really like to wish him the best of luck with his future endeavors. Thank you so much for making my second year such a rewarding year!

Furthermore, I'd like to thank Michael Lee for having the patience to teach me the basics while I was a 4th year thesis student as well as giving me lots of technical advice throughout my Master's. I've learned so much from Michael and thoroughly enjoyed having him as a mentor and a co-worker. Thank you for enhancing my interest in biochemistry!

I would really like to thank my thesis examination members, Dr. Graham Cote, Dr. John Allingham, Dr. Wayne Snedden, and Dr. Andrew Craig for their thorough suggestions that would help to improve my thesis. It was very gratifying and delightful to have been ensured that I have done an exceptional amount of work and that my extensive trouble-shooting skills will surely help me with my future work.

Also, thank you Kim Munro and David McLeod from Protein Function Discovery lab. Kim ran all of our CD and ITC data while David ran mass spectrometry for our Co-IP experiments. I would also like to thank Dr. Bruce Hill for the countless pieces of advice I've received regarding enzyme kinetics. Diane Sommerfeld and the ladies from the administrative office also deserve many thanks for all of their administrative work.

After a long day at the lab with things going wrong here and there, I would not have come back to the lab as happily as I had if I did not have my friends (and Squirrels) to cheer me on. They were always there to listen to my complaints and make them go away! I think it's safe to say that my friends want only what is the best for me and will always be there to listen when I need them. Thank you for making my time here at Queen's a lot more exciting than it would have been without you.

Finally, I would like to sincerely thank my family back at home in British Columbia. My parents and my siblings have always been supportive of my decisions throughout my entire undergraduate and graduate careers. One day, I hope to be able to pay them back for all the love and support that I've received! I wish Queen's University was a little closer to where you are.

Thank you everyone!

Table of Contents

Abstract	ii
Acknowledgements	iii
List of Figures	vii
List of Tables	x
List of Abbreviations	xi
Chapter 1 Introduction	1
1.1 Literature review	1
1.1.1 Post-translational modification	1
1.1.2 The Features of Human Protein Kinases: typical and atypical	2
1.1.3 Aarf Domain Containing Kinase 3.....	8
1.1.4 ADCK3 mutations lead to CoQ ₁₀ deficiency	11
1.1.5 Protein complex required for CoQ biosynthesis	12
1.1.6 The role of ADCK3 in stabilizing the Coq protein complex	16
1.1.7 Kinase activity of ADCK3 on Coq3, Coq5, and Coq7	17
1.1.8 Characteristics of the potential substrates of ADCK3	19
1.2 Project overview	24
1.2.1 Hypothesis and Objectives.....	24
1.2.2 Approaches	25
Chapter 2 Methods.....	27
2.1 Cloning of ADCK3	27
2.2 Expression and purification of ADCK3	29
2.3 Crystallization of ADCK3	31
2.4 X-ray data collection and structure determination of ADCK3	31
2.5 Co-immunoprecipitation using HEK293 cells.....	32
2.6 Cloning of Coq3, Coq5, and Coq7.....	35
2.7 Expression and purification of Coq3, Coq5, and Coq7	36
2.8 Analysis of protein folding of Q5b and Q7a using Circular Dichroism spectroscopy	36
2.9 Methyltransferase activity of Coq5.....	37
2.10 Crystallization trials of Coq5 and Coq7.....	39
2.11 ATPase activity assay	39
2.12 Radioactive kinase assay.....	40

Chapter 3 Results and Discussion.....	42
3.1 Cloning, mutagenesis, expression, and purification of ADCK3	42
3.2 Crystallization of ADCK3	51
3.3 X-ray data collection and structure determination of ADCK3	55
3.4 Co-immunoprecipitation using HEK293 cell.....	60
3.5 Cloning of Coq3, Coq5, and Coq7.....	66
3.6 Expression and purification of Coq3, Coq5, and Coq7	67
3.7 The protein folding of Q5b and Q7a.....	96
3.8 Methyltransferase activity of Coq5.....	98
3.9 Crystallization trials of Coq5 and Coq7.....	103
3.10 ATPase activity of ADCK3	105
3.11 Kinase activity of ADCK3.....	110
3.12 General Discussion	115
Chapter 4 Conclusion and Future Directions.....	120
4.1 Conclusion	120
4.2 Future directions	121
References.....	124
Appendix A ATP synthase F _o subunit 8 as the potential substrate of ADCK3	133

List of Figures

Figure 1.1 The mechanism of phosphotransferase activity of protein kinase A (PKA).	2
Figure 1.2 The structure of PKA: a representative structure of a typical kinase domain.	3
Figure 1.3 Structures of some atypical kinases.	6
Figure 1.4 The structural comparison between Rio1 and Rio2.	7
Figure 1.5 Schematic representation of ADCK3 protein structure.	8
Figure 1.6 The structure of ADCK3 with and without its N-terminal extension.	10
Figure 1.7 Comparison between MRI image of a healthy individual and a patient with cerebellar ataxia. 12	
Figure 1.8 The proposed arrangement of the yeast Coq proteins in the inner membrane of the mitochondria.	14
Figure 1.9 The structures of human Coq3 and Coq5 proteins predicted by PHYRE2.	19
Figure 1.10 Sequence alignment of Yeast and Human Coq5.	20
Figure 1.11 The structural features of yeast Coq5.	21
Figure 1.12 The structure of Coq7 (45-184) predicted by PHYRE2.	23
Figure 2.1 The HT29 vector engineered by Dr. Chelsey Chesterman.	27
Figure 2.2 Schematic summary of ADCK3 and ADCK4 constructs used for co-immunoprecipitation using HEK293 cells.	32
Figure 2.3 The reaction scheme of the SAM510: SAM Methyltransferase assay.	37
Figure 2.4 The reaction scheme of the SAMFluoro™:SAM Methyltransferase assay.	38
Figure 2.5 The schematic of the P _i ColorLock™ Gold Phosphate Detection System.	40
Figure 3.1 A representative profile of MBP-ADCK3 (245-647) purification using histidine-mediated immobilized metal affinity chromatography.	43
Figure 3.2 Chromatogram of MBP-ADCK3 (245-647) purification via size-exclusion chromatography. 44	
Figure 3.3 Chromatogram of the cleaved ADCK3 (245-647) purification via size-exclusion chromatography.	46
Figure 3.4 A representative profile of ADCK3 (245-647) K276E purification.	48
Figure 3.5 A representative profile of MBP-ADCK3 (329-647) purification using histidine-mediated immobilized metal affinity chromatography.	50
Figure 3.6 Initial crystal hits of ADCK3 (245-647) in medium molecular weight PEG smear crystallization condition.	52

Figure 3.7 Crystals of ADCK3 (245-647).	54
Figure 3.8 The comparison between the apo-structure of ADCK3 and our structure.	58
Figure 3.9 The structural features of AMP-PNP bound ADCK3 (261-644) R611K in comparison to the apo-structure of ADCK3 (254-647).....	59
Figure 3.10 Silver stained SDS-PAGE gel of the co-immunoprecipitation samples for all constructs.	62
Figure 3.11 Silver stained SDS-PAGE gel of the co-immunoprecipitation samples for 3A and 4A constructs.	64
Figure 3.12 Q3a, Q3b, and Q3c purification using histidine-mediated immobilized metal affinity chromatography.	69
Figure 3.13 Q5a purification using histidine-mediated immobilized metal affinity chromatography.....	71
Figure 3.14 DNA sequence map of Coq5 and mutagenesis work done on Q5a plasmid.	73
Figure 3.15 representative profile of Q5b purification using histidine-mediated immobilized metal affinity chromatography.	75
Figure 3.16 Analysis of Q5b following purification using size-exclusion chromatography.	78
Figure 3.17 Analysis of Q5b following purification using ion-exchange chromatography.....	80
Figure 3.18 A representative profile of Q7a purification using histidine-mediated immobilized metal affinity chromatography.....	82
Figure 3.19 Analysis of Q7a following purification using size-exclusion chromatography.....	84
Figure 3.20 A representative profile of Q7b purification using histidine-mediated immobilized metal affinity chromatography.....	86
Figure 3.21 Analysis of Q7b following purification using size-exclusion chromatography.	88
Figure 3.22 Overlay of the size-exclusion chromatograms of Q7b purified with and without reducing reagents.	90
Figure 3.23 Analysis of Coq7 (37-217) following purification using size-exclusion chromatography.....	92
Figure 3.24 Surface charge analysis of Coq7 and BfDSPL.	94
Figure 3.25 Circular dichroism spectra for Q5b and Q7a.	97
Figure 3.26 The molecular structure of Coq5 substrate and its analog.....	98
Figure 3.27 Methyltransferase activity of Q5b.	102
Figure 3.28 Examples of the initial hits for Q7b in screening trays.	104
Figure 3.29 ATPase activity of MBP-ADCK3 (245-647).	106
Figure 3.30 The ATPase activity of ADCK3 and its substrates.	109
Figure 3.31 The kinase activity of ADCK3 in the absence and presence of Coq proteins.	111
Figure 3.32 The nucleotide binding cavity comparison between apo- and holo- ADCK3 structures.	116

Figure 3.33 The comparison of the wild-type residue arginine 611 and mutated residue lysine 611 and their effect on the structure of ADCK3.....	117
Figure A.1 The schematic structure of mitochondrial ATP synthase.	134
Figure A.2 A representative profile of MBP-FoS8 purification using histidine-mediated immobilized metal affinity chromatography.....	139
Figure A.3 Analysis of MBP-FoS8 following purification using size-exclusion chromatography.	140
Figure A.4 Radioactive kinase activity of ADCK3 and ADCK4 with peptide segment and full length of human ATP synthase F ₀ S8 as the substrate.	142
Figure A.5 The Structure of ADCK3.....	144

List of Tables

Table 1.1 Proteins involved in CoQ biosynthesis in yeast and human cells.	13
Table 1.2 Clinical features associated with mutation in the protein involved in CoQ synthesis.	15
Table 2.1 Summary of ADCK3 and ADCK4 constructs used for co-immunoprecipitation using HEK293 cells.	32
Table 2.2 Summary of Coq protein constructs fused to maltose binding protein.	35
Table 3.1 X-ray data collection and refinement statistics.	57
Table 3.2 Identities of the proteins co-immunoprecipitated from HEK293 cells.	65
Table 3.3 Summary of DNA sequence analysis of Coq3, Coq5, and Coq7.	67

List of Abbreviations

2D-IEF	Two-dimensional Isoelectric focussing
ADCK	Aarf domain containing kinase
ADP	Adenosine diphosphate
AMP-PNP	Adenylyl-imidodiphosphate
ARCA-2	Autosomal recessive cerebellar ataxia type 2
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
CD	Circular dichroism
CoQ	Coenzyme Q
CPM	Counts per minute
DDMQ ₆ H ₂	Demethyl-demethoxy-hydroquinone
DMQ ₆	Demthoxy-coenzyme Q ₆
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
eEF2K	Eukaryotic Elongation Factor 2 Kinase
ETC	Electron transport chain
eYFP	Enhanced yellow fluorescent protein
F ₀ S8	ATP synthase F ₀ subunit 8
FPLC	Fast protein liquid chromatography
HEK293	Human embryonic kidney 293
HPLC	High performance liquid chromatography
<i>H. sapiens</i>	<i>Homo sapiens</i>

IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl β -D-1-thiogalactopyranoside
LB	Luria Broth
MBP	Maltose-binding protein
MTS	Mitochondrial targeting sequence
Ni ⁺² -NTA	Nickel-nitrilotriacetic acid
OD	Optical density
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
P _i	Inorganic phosphate
PKA	Protein Kinase A
PTM	Post-translational modification
RMSD	Root-mean-square deviation
SAM	S-adenosyl methionine
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
STPK	Serine/Threonine Protein Kinase
TCEP	Tris (2-carboxyethyl) phosphine
TEV	Tobacco Etch Virus
TPK	Tyrosine Protein Kinase

Chapter 1

Introduction

1.1 Literature review

1.1.1 Post-translational modification

Post-translational modification (PTM) refers to the covalent modifications that can be made to proteins during or after protein biosynthesis. Most PTMs are enzymatic reactions where the synthesized proteins are modified according to the signals received by the cells. Proteins can be modified through several ways including glycosylation, acetylation, alkylation, methylation, biotinylation, glutamylation, glycylation, isoprenylation, lipoylation, phosphopantetheinylation, sulfation, selenation, c-terminal amidation, ubiquitylation, and phosphorylation. In addition to other functional implications, these modifications often assist proteins with their folding and stability which leads to their mature forms. Most importantly, PTMs also play a key role in regulating proteins that needs to be tightly controlled, such as signalling and enzymatic proteins. As a result, instead of synthesizing and degrading specific enzymes as they are needed, PTMs save energy for the cells through simple modifications that activate or deactivate the proteins. Overall, PTM diversifies the functions of proteins under different physiological stimuli.

The most prevalent PTM is phosphorylation and it is known to play critical roles in the regulation of many cellular processes. It is thought to be the most common modification and that nearly 30% of all human proteins are phosphorylated [1]. Phosphorylation of proteins occur when a protein kinase transfers a γ -phosphate from an adenosine triphosphate (ATP) molecule on to a hydroxyl group of an amino acid side chain, specifically the ones of tyrosine, threonine, and serine. Phosphorylation can be reversed when a protein phosphatase removes the phosphate group. The pair of protein kinase and phosphatase works together to appropriately modify specific proteins under different conditions.

1.1.2 The Features of Human Protein Kinases: typical and atypical

Protein kinases are one of the largest gene families as they take up nearly 1.7% of the human genes [2]. As previously mentioned, their kinase activity is important in regulating many cellular processes as they mediate most of the signal transduction in eukaryotic cells. Mutations in kinases have been found to be causative of human diseases such as certain cancers and developmental and metabolic disorders [3]. As a result, it is of great interest to thoroughly investigate the structure and function of kinases to provide the necessary knowledge for the development of therapeutic reagents against such disorders.

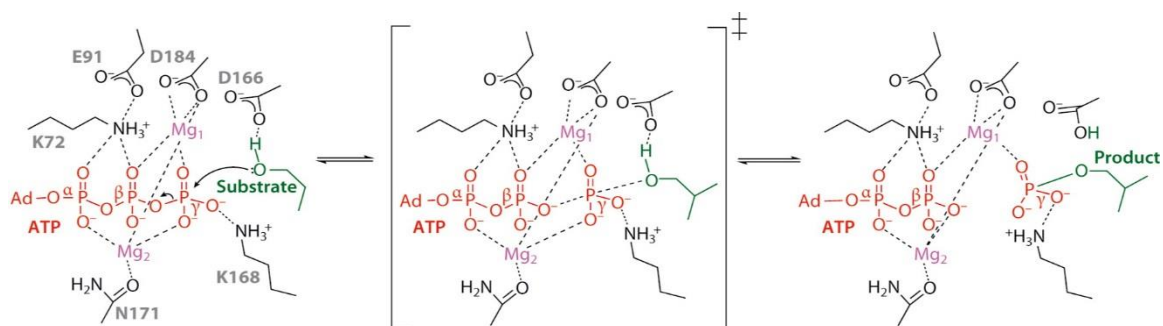


Figure 1.1 The mechanism of phosphotransferase activity of protein kinase A (PKA).

From left to right, the reaction proceeds from enzyme/substrate complex, transition state, and then enzyme/product complex. The catalytic residues of PKA are noted in black and labelled, the hydroxyl group of the substrate is noted in green, the ATP molecule is noted in red, and the magnesium molecules are noted in pink. The initial state shows the lone pair of electrons from the oxygen of the substrate hydroxyl group aligned with the γ -phosphorus atom while the hydrogen of the substrate hydroxyl group is being accommodated by the negative charges of the aspartate residue (D166). In the transition state, it shows that while the two magnesium ions and nearby lysine residue (K72) compensate for the negative charges of the γ -phosphate, the oxygen of the substrate hydroxyl group starts its nucleophilic attack on the γ -phosphorus atom. It is also at this state where the breaking of the bond between the γ -oxygen and γ -phosphorus occurs. As the γ -phosphate is transferred to the oxygen of the substrate hydroxyl group, the sudden decrease in pKa of the hydrogen of the substrate hydroxyl group drives its transfer to the aspartate. In the final state, the γ -phosphate is bonded to the oxygen of the substrate while the hydrogen from the substrate is now stored by the aspartate residue. The by-product ADP is shown to be stabilized by the catalytic residues of PKA. Adapted from [4].

Aforementioned, protein kinases function to phosphorylate substrates by transferring the γ -phosphate of ATP onto the oxygen of the substrate hydroxyl group (Figure 1.1). Depending on the amino acid hydroxyl group a protein kinase phosphorylates, these protein kinases can be categorized as one of three protein kinases: TPK (Tyrosine-specific protein kinase), STPK (Ser-/Thr-specific protein kinase), or dual specificity protein kinase (both tyrosine and serine/threonine). In addition to the ATP molecule, protein kinases typically require two divalent cations, such as magnesium ions or manganese ions. These cations compensate for the negative charges exhibited in the active site during the phosphotransfer and ultimately coordinate the nucleotide.

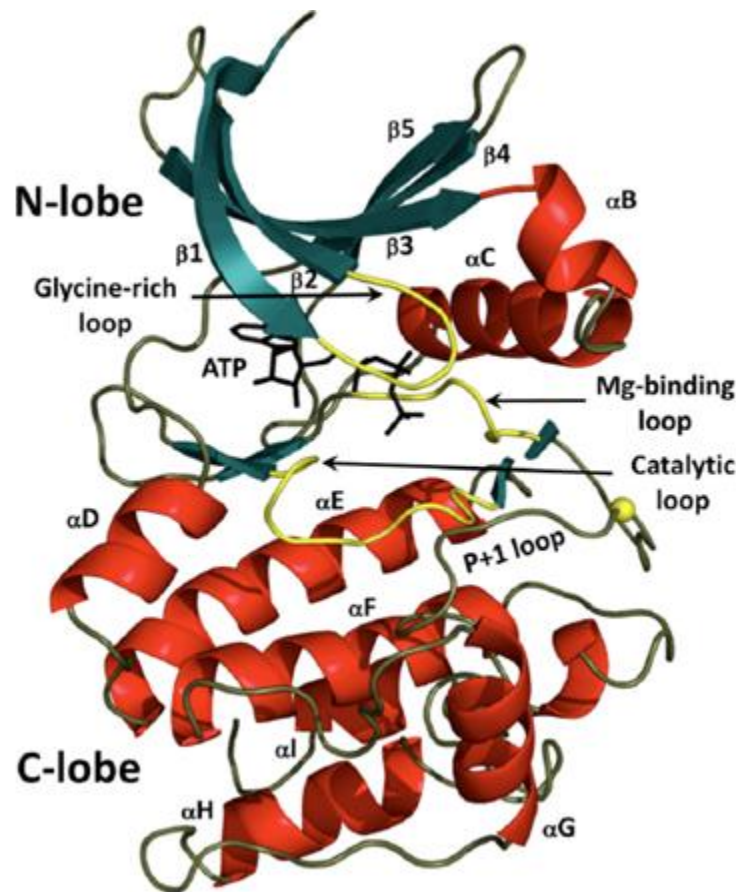


Figure 1.2 The structure of PKA: a representative structure of a typical kinase domain. The N-terminal domain and the C-terminal domain are coloured blue and red, respectively. Some of the essential subdomains, as well as the secondary structures, are labelled. The cleft between N-lobe and C-lobe is the active site in which the ATP is located and where the substrate would enter for its modification. Adapted from [5].

Based on the extensive sequence and structural analysis, the general features of eukaryotic protein kinases (ePK) have been elucidated. As seen in Figure 1.2, the kinase domain of the eukaryotic kinases has an overall bi-lobal architecture with a N-terminal lobe and a C-terminal lobe [3]. Within the kinase domain of the typical kinases, there are 12 conserved subdomains that are required for their phosphotransferase function [6]. First, **subdomain I** is called the Glycine-rich loop (GxGxxGxV) and its residues work together to orient and stabilize the ATP molecule. **Subdomain II** is a highly-conserved lysine residue that helps to anchor and orient the ATP molecule. **Subdomain III** is a centrally located large α -helix C with a well-conserved glutamic acid residue that helps to stabilize the interactions between the subdomain II lysine and the α - and β - phosphates of ATP. Next, **subdomain IV** is merely a hydrophobic β -strand 4 in the N-lobe and its function remains unclear. The **subdomain V** consists of a hydrophobic β -strand 5 and a small α -helix D connected by an extended chain. This subdomain links the N-lobe and the C-lobe together and helps to orient and stabilize the ATP. **Subdomain VIA** is the large hydrophobic α -helix E that acts mainly as a support structure. **Subdomain VIB** folds into β -strands 6 and 7 with a loop connecting them. This domain is also known as the catalytic loop with key residues in its sequence HRDLKxxN. **Subdomain VII** has a highly conserved DFG triplet sequence where these residues chelate the primary magnesium which helps to orient the γ -phosphate of ATP for transfer. **Subdomain VIII** has the APE triplet which seem to play a key role in stabilizing the C-lobe and also in peptide substrate recognition. As well, many kinases have been shown to be activated by phosphorylation of residues in this domain. **Subdomain IX** is a large α -helix F that acts to stabilize the catalytic loop. Finally, **subdomains X and XI** are found in the C-lobe and they are both poorly conserved with an unknown function. Combination of most of these subdomains allow the typical kinases to carry on its function.

Of the 518 human kinases, 40 were categorized as atypical kinases due to their incomplete collection of the essential subdomains. In addition, structure of atypical kinases have shown to vary in which some structures are very similar to those of typical kinases with some differences, while other structures are completely different (Figure 1.3) [7]. Although atypical kinases have little sequence similarity to the typical kinases, some atypical kinases were shown to have kinase activity. For example, Rio1 and Rio2 are two atypical kinases conserved across species from archaea to human [7]. For both, their kinase activity has been demonstrated and in studies done in yeast revealed that both Rio kinases are required for proper cell cycle progression and chromosome maintenance. Also, deletion of either of the two were shown to be lethal, suggesting that these proteins have distinct functions [8,9]. This theory is further supported by the difference in their structures (Figure 1.4). To be specific, both Rio1 and Rio2 kinases have an additional domain on their N-terminal domain that is not conserved in the typical kinase domain. For Rio1, an additional helix was identified that appears to be conserved in the sequence of Rio1 proteins of other species (Figure 1.4, yellow); however, the function of this helix is unknown. For Rio2, an additional winged helix-loop-helix fold was found which is primarily seen in DNA-binding proteins; however, further investigation revealed that this type of fold is also observed in proteins involved in protein-protein and protein-RNA interactions (Figure 1.4, green) [10–12]. Based on the role of Rio2 in rRNA processing, it makes sense for it to possess a RNA binding domain. Furthermore, both Rio1 and Rio2 kinases do not conserve subdomains VIII, X, and XI even though they show phosphorylation activity. Subdomain VIII is the activation loop that is required for peptide substrate recognition. As a result, initially the ability for Rio kinases to bind to their peptide substrates were questioned. Now, it is considered that perhaps the additional domains that are conserved on the N-terminal domain compensates for the lack of subdomains VIII, X, and XI. Nonetheless, in the case of many atypical kinases, such as ADCK3, the structure and the function await to be elucidated.

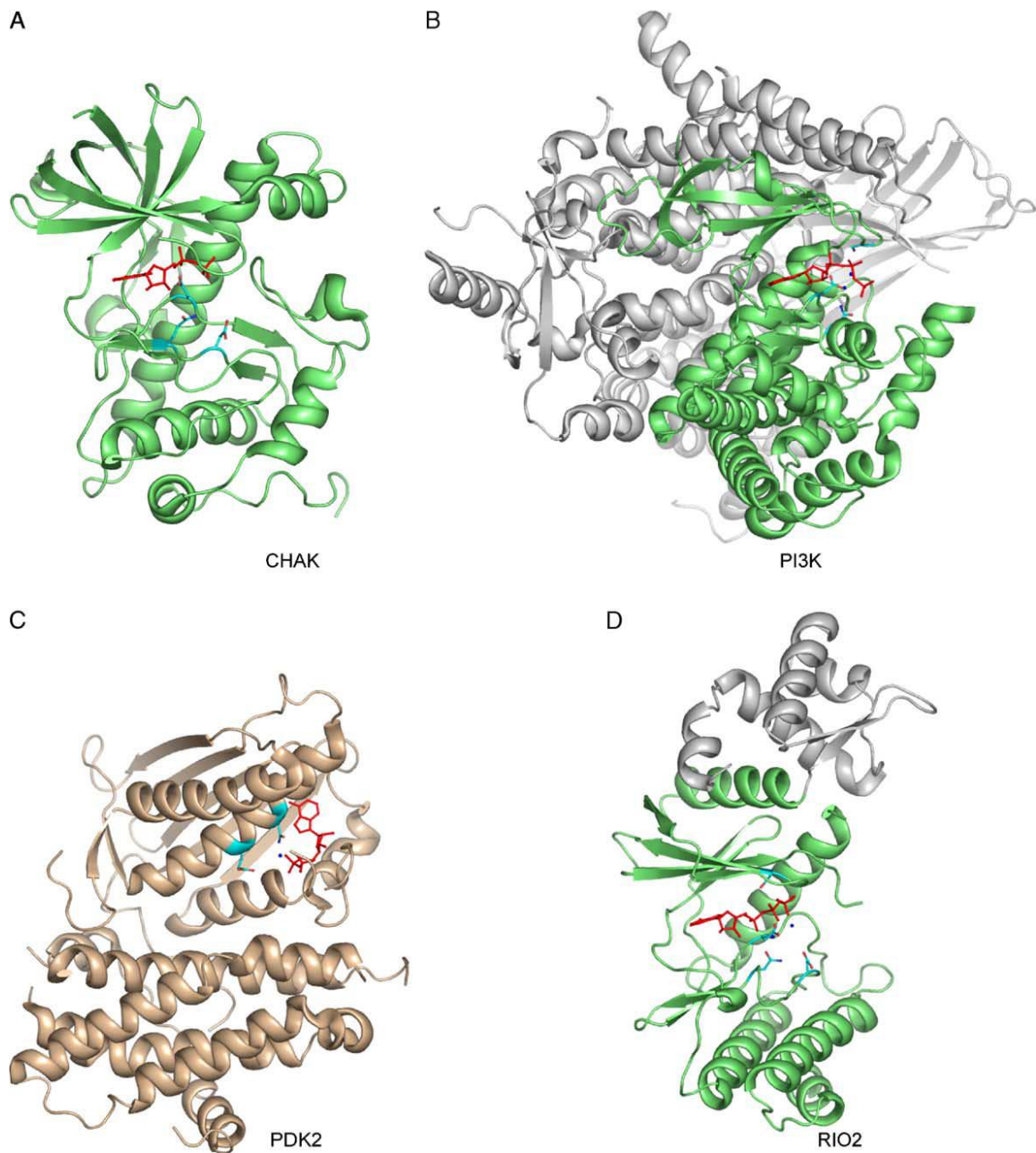


Figure 1.3 Structures of some atypical kinases.

The four atypical kinases shown are: (A) ChaK (PDB: 1IA9); (B) PI3K (PDB: 1E8X); (C) PDK2 (PDB: 1JM6); and (D) Rio2 (PDB: 1ZAO). The domains coloured in green are those homologous to the typical kinase domains. The catalytic residues are highlighted in cyan and a nucleotide bound to the active sites are shown as red sticks [7].

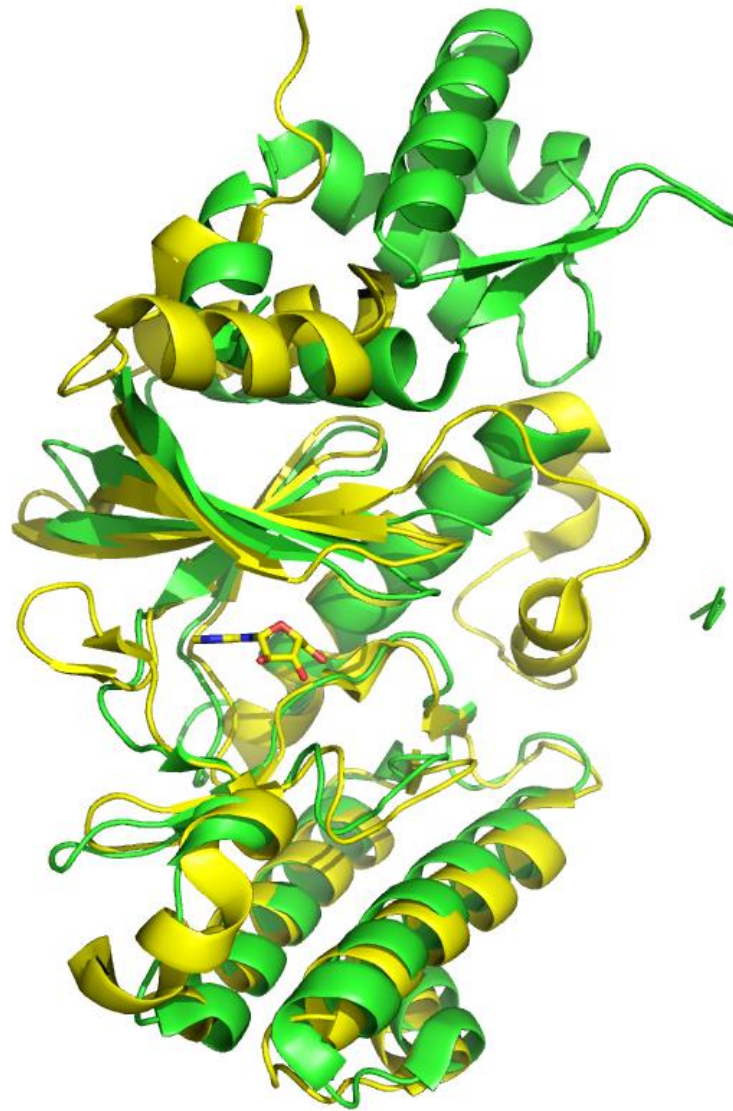


Figure 1.4 The structural comparison between Rio1 and Rio2.

The structure of Rio1 (PDB: 1ZTF) is shown in yellow while the structure of Rio2 (PDB: 1TQI) is shown in green.

1.1.3 Aarf Domain Containing Kinase 3

Human genome encodes 5 ADCK proteins. Of the 5, the third and fourth members, ADCK3 and ADCK4 are co-orthologs of the yeast Coq8 protein. These proteins were identified to be atypical kinases based on bioinformatic analysis of their amino acid sequence. In particular, the amino acid sequence analysis of ADCK3 also revealed a mitochondrial targeting sequence followed by a transmembrane domain. *In vivo* studies have confirmed that mature ADCK3 is found in the mitochondria and localizes to the matrix side of the inner membrane [13]. As shown in Figure 1.5, ADCK3 conserves only 5 of the 12 essential subdomains, which are subdomains I, II, III, VIB, and VII. Interestingly, in addition to the lack of 7 subdomains, for subdomain I, instead of a glycine-rich loop, ADCK3 has an alanine-rich loop. Although the exact implications of conserving an alanine-rich loop instead of a glycine-rich loop are unknown, ADCK3 constructs introduced with a glycine-rich loop instead were shown to have higher auto-phosphorylation activity [14].

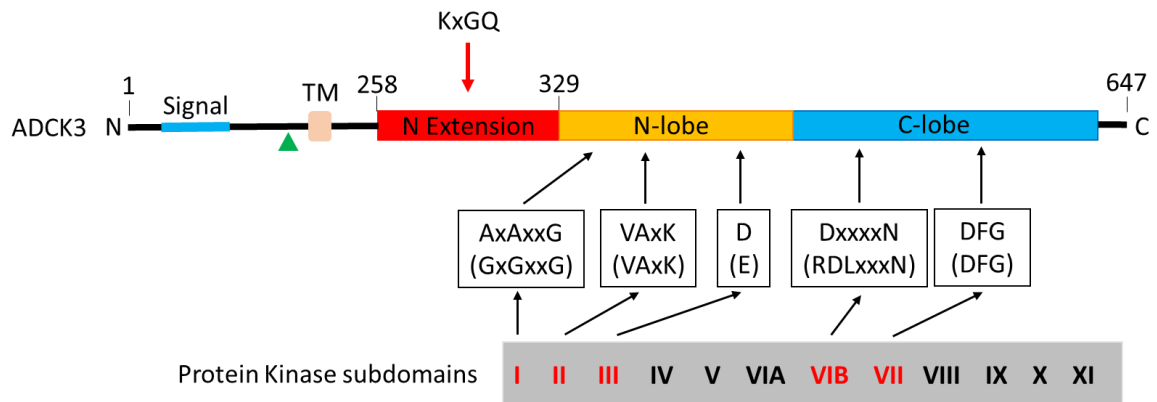


Figure 1.5 Schematic representation of ADCK3 protein structure.

The mitochondrial targeting sequence present on the N-terminal side is indicated by a blue line. The green triangle represents the cleavage site that undergoes cleavage resulting in the mature form of ADCK3 found in the mitochondria. TM represents the transmembrane helix that is embedded in the inner membrane of the mitochondria. As labelled, the N-terminal extension, N-lobe, and C-lobe are represented by rectangles coloured red, yellow, and blue, respectively. The location of the KxGQ motif unique to ADCK proteins is labelled with an arrow. The essential protein kinase subdomains found in ADCK3 are highlighted and compared to the essential motifs found in typical protein kinases. Simplified from [15].

As seen in Figure 1.6, the structure of ADCK3 includes a N-terminal extension with a KxGQ motif that occludes the active site [14]. This N-terminal extension consists of 4 consecutive helices held tightly into the core of the kinase by a salt bridge between the lysine (K276) residue of the KxGQ motif and two glutamic acids (E401 and E405). Currently, this N-terminal extension is labelled as the autoinhibitory domain and its purpose is unknown; however, similar to Rio kinases, this additional domain may also play a role in substrate binding. Similar to Rio kinases, ADCK3 lacks the subdomains VIII, X, and XI required for peptide substrate recognition. Perhaps recognition of the substrate by the N-terminal extension leads to the release of this extension from the active site. At this time, the role of KxGQ motif, conserved in all ADCK proteins, is currently predicted to be an autoinhibitory domain. Although ADCK3 does not conserve all 12 essential motifs and its structure seems to suggest an inhibited or inactive conformation, the possibility of ADCK3 conserving a kinase activity cannot be dismissed. In fact, ADCK3 has been shown to have residual ATPase activity which demonstrates its ability to cleave the γ -phosphate from ATP [16]. As well, studies have shown that ADCK3 has an autophosphorylation activity and that when the KxGQ motif is mutated, the autophosphorylation activity is enhanced [14]. Furthermore, indirect evidence from yeast studies shows phosphorylation of putative protein substrates by ADCK3, which is discussed later in this chapter.

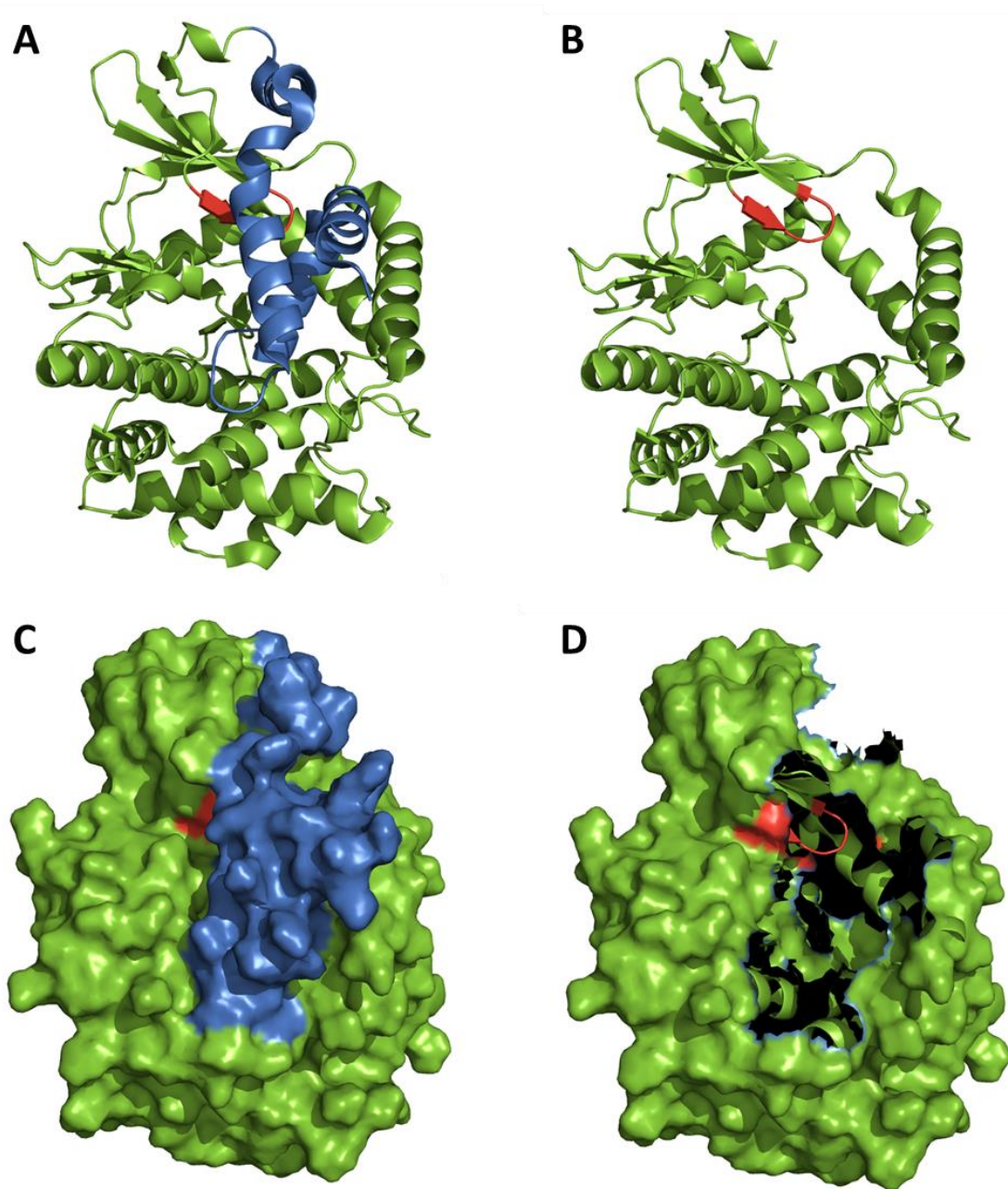


Figure 1.6 The structure of ADCK3 with and without its N-terminal extension.

(A) Crystal structure of ADCK3 including the N-terminal extension (PDB:4PED); (B) ADCK3 structure without the N-terminal extension; (C) Surface representation of ADCK3 structure including the N-terminal extension; and (D) Surface representation of ADCK3 structure without the N-terminal extension. The N-terminal extension structure is shown in blue and the alanine-rich loop is highlighted in red. The structures on the left show the blockage of the active site via the N-terminal extension. The structures on the right show a pocket directly below the alanine-rich loop where a substrate may enter.

1.1.4 ADCK3 mutations lead to CoQ₁₀ deficiency

Patients with mutations in ADCK3 were observed to experience CoQ₁₀ deficiency which leads to a disorder called autosomal recessive cerebellar ataxia 2 (ARCA-2). Cerebellar ataxia is a loss of cerebellar function due to damage or inflammation of the cerebellum of the brain (Figure 1.7). Some symptoms observed from patients with ADCK3 mutations are seizures, impaired movement, instability, muscle weakness, and tremors [17]. In addition, more patients with ADCK3 mutations have been identified with additional symptoms including dysdiadochokinesia, bilateral dysmetria, slurred speech, and truncal ataxia [18]. It makes sense that ADCK3 mutation leads to CoQ₁₀ deficiency because ADCK3 is one of the proteins required for CoQ biosynthesis. Many yeast studies have shown that knocking out ADCK3 hinders the CoQ production as well as the stability of the protein complex which synthesizes CoQ.

Case studies showing the genetic analysis of these patients revealed the specific mutations that led to the onset of CoQ₁₀ deficiency; however, there was no specific trend in the type of genetic mutation nor the region of the resulting protein being altered [15,17–23]. In terms of the type of genetic mutation, patients had either, a single substitution, a double substitution, or a single substitution and deletion. In addition, the mutations identified from these patients were not localized to one specific area of the protein. Instead, the substitution mutations were found in all parts of the protein, whether it be on the N-terminal or C-terminal domains, and whether it be exposed to the surface or hidden in the core of the protein. An important note is that no mutations of the catalytic residues were found. Perhaps it can be concluded that the structure of ADCK3 and the way it is folded by the nature of these specific residues are important to its function.

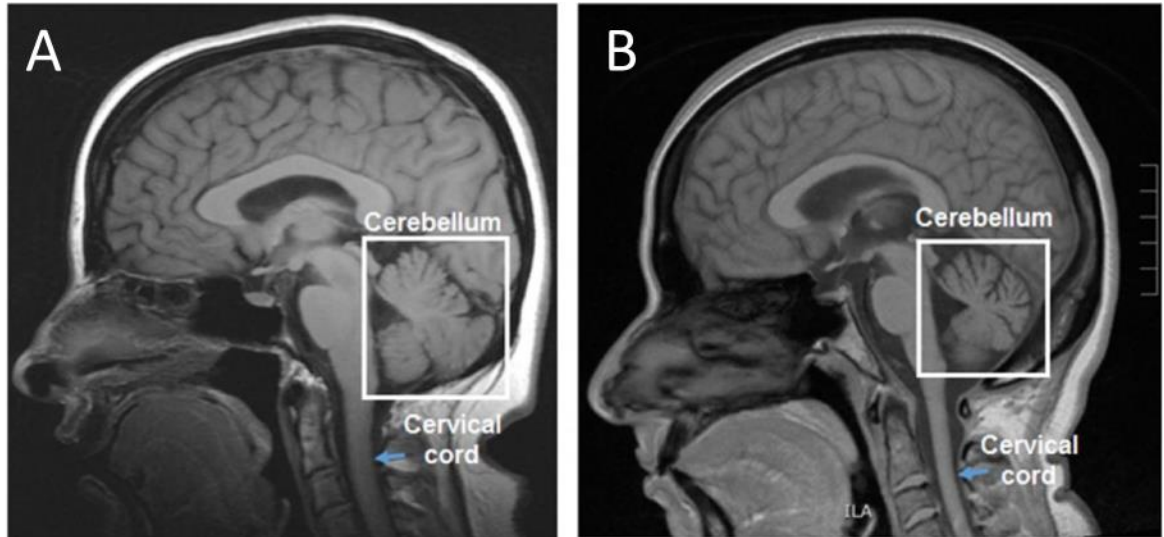


Figure 1.7 Comparison between MRI image of a healthy individual and a patient with cerebellar ataxia.

(A) The MRI of the healthy individual shows the cerebellum taking most of the space in the cavity while (B) the MRI of the patient shows a significantly smaller cerebellum [24].

1.1.5 Protein complex required for CoQ biosynthesis

CoQ is a lipophilic antioxidant comprised of a quinone ring with a polyisoprenoid tail. Maintaining steady and appropriate level of CoQ is essential for the proper human growth and development as CoQ is an electron transporter for the electron transport chain (ETC), transporting electrons from complex I and II to complex III. The proton gradient created by ETC is coupled to the oxidative phosphorylation which produces 90% of the ATP used for maintaining proper bodily functions. Thus, in order to allow for healthy growth and function, proper amount of CoQ must be synthesized and steady level must be maintained. As previously mentioned, eukaryotic CoQ biosynthesis requires a protein complex in the inner membrane of the mitochondria. In yeast, 13 proteins of this complex have been identified to be required for CoQ biosynthesis; their human homologs have been found to include 15 proteins (Table 1.1).

Table 1.1 Proteins involved in CoQ biosynthesis in yeast and human cells.

Adapted from [25].

<i>S. cerevisiae</i>	<i>H. sapiens</i>	Function
COQ1	PDSS1 PDSS2	Prenyl diphosphate synthase
COQ2	COQ2	Prenyltransferase
COQ3	COQ3	O-methyltransferase
COQ5	COQ5	C-methyltransferase
COQ6	COQ6	Mono-oxygenase
COQ7	COQ7	Hydroxylase
YAH1	FDX1L	Electron transfer to COQ6
ARH1	FDXR	Electron transfer to COQ6
COQ4	COQ4	Stabilization of Q complex
COQ8	ADCK3 ADCK4	Putative Kinase
COQ9	COQ9	Lipid-binding protein
COQ11	-	Decarboxylase?
COQ10	COQ10A COQ10B	Necessary for correct localization of CoQ within inner mitochondrial membrane

Proteins required for CoQ biosynthesis can be divided into enzymatic, non-enzymatic, and chaperone roles. First, the enzymatic proteins consist of Coq1, Coq2, Coq3, Coq5, Coq6, and Coq7, along with two electron transporters, YAH1 and ARH1. CoQ production starts with polyisoprenoid chain production by Coq1 (PDSS1 and PDSS2 in human) using its starting material farnesyl-pyrophosphate from the mevalonate pathway [25]. Then, Coq2 condenses the isoprenoid chain onto the benzoquinone ring [26]. This product is passed on to the sequential proteins for the modification of its ring structure. A S-adenosylmethionine-dependent methyltransferase, Coq3, catalyzes two O-methyltransfer steps during the CoQ synthesis [27]. Similarly, Coq5 is also a S-adenosylmethionine-dependent methyltransferase that catalyzes the C-methyltransfer step [28]. Coq6 is a flavin-dependent monooxygenase that adds a hydroxy group to the ring with the help of mitochondrial ferredoxin (YAH1) and its reductase (ARH1) [29]. In addition, Coq7 is also a flavin-dependent monooxygenase that catalyzes the last monooxygenase step in the CoQ synthesis [25].

The non-enzymatic proteins required for CoQ biosynthesis consist of Coq4, Coq8 (ADCK3 and ADCK4 in human), and Coq9. First, although the function of Coq4 is unknown, evidence from studies shows the absolute requirement for Coq4 for the assembly and stability of the overall complex [30]. Coq8 is the yeast homolog of the human ADCK3 and ADCK4, which also corresponds to Coq8A and Coq8B, respectively. In terms of Coq9, results from a recent study suggest that Coq9 is involved in stabilizing Coq7 to increase its efficiency [31]. In addition, a small molecule-binding domain was found in the crystal structure of Coq9 consisting of hydrophobic residues. This suggested that Coq9 may also be involved in binding lipids, specifically the lipid tail of CoQ.

Finally, the chaperone protein required for CoQ is Coq10 (Coq10A and Coq10B in human). Although Coq10 has shown to not be required for the synthesis of CoQ, evidence from knockout studies as well as the features of its structure suggest that it is a CoQ trafficking protein [32]. With the evidence gathered from cross-linking studies, He *et al.* proposed an arrangement of these Coq proteins as seen in Figure 1.8 [13].

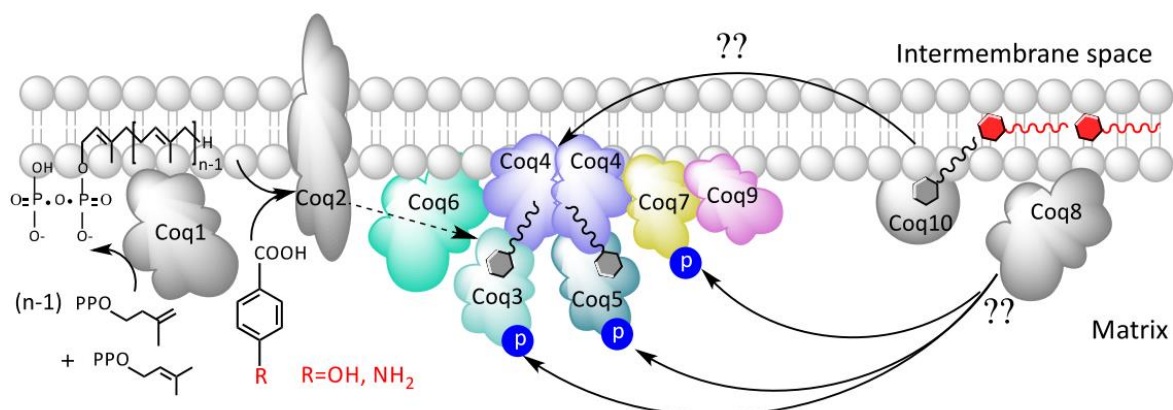


Figure 1.8 The proposed arrangement of the yeast Coq proteins in the inner membrane of the mitochondria.

The CoQ synthesis begins with isoprenoid tail production by Coq1, which is then transferred to Coq2 which connects the tail onto a benzoquinone. The condensed molecule is then passed along to the main complex where a membrane-bound dimer of Coq4 provides the structural basis for the complex. The precursor CoQ molecules goes under series of modifications by Coq3, Coq5, Coq6, and Coq7. The exact role of Coq9 and Coq10 in the modification of CoQ precursor are yet unknown. Coq8, the yeast homolog of ADCK3 is hypothesized to phosphorylate the proteins of the main complex, Coq3, Coq5, and Coq7. Adapted from [13].

In addition to the disorders caused by mutations in ADCK3, mutations in the genes of other proteins found in this protein complex also cause CoQ₁₀ deficiency (Table 1.2). The mild symptoms of CoQ deficiency such as seizures can be alleviated with CoQ supplements; however, due to the uneven distribution of CoQ after ingestion, targeting specific affected areas is difficult. Furthermore, the severe effects of CoQ deficiency such as damages to the brain and physical deforms occurring during growth, cannot be reverse. Thus, it is essential that patients harbouring these mutations are diagnosed and treated before their body goes under serious damage. Consequently, it is important to study the structure and function of these proteins to provide insights for the development of therapeutic reagents against these disorders.

Table 1.2 Clinical features associated with mutation in the protein involved in CoQ synthesis.

Adapted from [33].

Protein	Clinical features	Age at onset	Response to CoQ₁₀ Supplement
PDSS1	Severe infantile multi-systemic disease	1-2 year	Improved and alive
PDSS2	Severe infantile multi-systemic disease Leigh syndrome	3 month	No clinical response
COQ2	Nephropathy Severe infantile multi-systemic disease Multiple system atrophy	Infantile – or early childhood onset Adult-onset	Dramatic improvement of neurological
COQ4	Encephalomyopathy	< 3 year	Signal improvement of neuromuscular symptoms
COQ6	Nephropathy with sensorineural deafness	Infantile – or early childhood-onset	Improvement of nephritic syndrome and hearing loss
COQ8 (ADCK3)	Cerebellar ataxia	Juvenile – or adult-onset	Severe neurological deficit with epilepsy ~mild improvement of ataxia
COQ9	Severe infantile multi-systemic disease	Birth	No clinical response

1.1.6 The role of ADCK3 in stabilizing the Coq protein complex

Several yeast cell biology studies have shown evidence that Coq8, the yeast homologue of ADCK3, has stabilizing effect on the Coq proteins found in the main CoQ biosynthetic complex. First, Xie *et al.* compared the Coq protein expression levels in several Coq protein knock out yeast strains to the levels in wild-type yeast strain [34]. The results showed that in Δ Coq3, Δ Coq4, and Δ Coq7 strains that normally do not express Coq4, Coq7, Coq8, and Coq9, overexpression of Coq8 led to expression of those proteins, except for their respective knocked out proteins. Δ Coq5 strain, which does not express Coq5, Coq7, Coq8, and Coq9, also showed expression of Coq7, Coq8, and Coq9 when induced with overexpression of Coq8. However, overexpression of Coq8 did not affect the expression of Coq4 and Coq7 in Δ Coq9 strain in addition to the lack of expression of Coq7 and Coq9 in Δ Coq1 and Δ Coq2 strains. As a result, it was suggested that Coq8 is required for the stabilization of some Coq proteins.

To further determine whether Coq8 is required for the stabilization of these proteins, Xie *et al.* performed the same experiment using the Δ Coq7 strain; however, this time, the Coq9 level was compared between cells harbouring either no plasmid, multicopy plasmid expressing Coq8, plasmid expressing inactive Coq8 (Coq8-G130D), or a low copy plasmid expressing Coq8 [34]. To clarify, the Coq8-G130D mutation corresponds to the ADCK3-G272D mutation observed in patients with deficiencies in muscle CoQ₁₀ content [22]. Prior to this experiment, yeast mutant harbouring Coq8-G130D showed lack of CoQ₆ and therefore used for further studies. Subsequently, the results showed that the cells with either multicopy plasmid or low copy plasmid expressed Coq9 compared to no Coq9 expression observed in cells with either no plasmid or the plasmid that expresses inactive Coq8. Even though there was expression of Coq8-G130D, the expression of Coq9 was not observed. As a result, it was concluded that active Coq8 is required for the stabilization of these Coq proteins.

1.1.7 Kinase activity of ADCK3 on Coq3, Coq5, and Coq7

Although Coq8 affects many Coq proteins, in this thesis, we focussed on Coq3, Coq5, and Coq7 as there are several studies that showed signs of phosphorylation via Coq8 on Coq3, Coq5, and Coq7. For example, when immunoblot of two-dimensional isoelectric focussing/sodium dodecyl sulfate polyacrylamide gel electrophoresis, 2D-IEF/SDS-PAGE, was performed using wild-type and Coq8 mutant yeast lysates, the results showed disappearance in the phosphorylated version of Coq3 [35]. Initially, when the wild-type lysate sample was analysed, 2 spots were observed when Coq3-cMyc was detected; however, when the same was done with the lysate of the yeast lacking active Coq8, only 1 spot was detected. In order to confirm that the difference between the two spots were due to the phosphorylation, the wild type lysate was incubated with λ -phosphatase before running the sample on 2D-IEF/SDS-PAGE. The results showed a single Coq3-cMyc spot just like the Coq8 mutant yeast strain results confirming that the spot that disappeared was the phosphorylated version of Coq3.

Similarly, Xie *et al.* also performed immunoblots of 2D-IEF/SDS-PAGE using wild-type and Coq8 mutant yeast lysates but instead detected for Coq5 and Coq7 polypeptides [34]. For Coq5, the results showed 4 spots for the wild-type lysate; however, when the wild-type lysate was treated with a phosphatase, only 2 spots remained. For the Coq8 knockout lysate, 3 spots were observed leading to the conclusion that although there are 3 phosphorylation states, Coq8 is responsible for only one state and that there may be another kinase working together with Coq8 on phosphorylating Coq5. For Coq7, the results showed 3 spots for the wild-type lysate; however, when the wild-type lysate was treated with a phosphatase, 1 spot remained. For the Coq8 knockout lysate, Coq7 was not detected. These results suggest that Coq7 has 3 phosphorylation states and that Coq8, the yeast homologue of the human ADCK3, may be an active kinase.

In addition, although no direct phosphorylation of Coq7 by ADCK3 has been shown, 3 phosphorylation sites on Coq7 have been identified [36]. Radioactive kinase assays have been performed with several yeast Coq7 constructs where the predicted phosphorylation residues,

Ser²⁰, Ser²⁸, and Thr³², were mutated to Alanine residues resulting the construct Coq7-AAA. The results showed that the wild-type Coq7 were phosphorylated by the kinase while the constructs with either Alanine or the negatively charged residues did not get phosphorylated. In addition, the CoQ₆ and CoQ₆-precursor molecule demethoxy-coenzyme Q₆ (DMQ₆) levels were monitored between wild-type yeast and yeast strains expressing Coq7-AAA. The results showed that the substitution of phosphorylated residues to alanine increased CoQ₆ levels by 256% and lowered the amount of DMQ₆. This increase in the product formed and decrease in the precursor molecule indicate that Coq7 is more active *in vivo* when it cannot be phosphorylated. The findings of these two experiments support that Coq7 has phosphorylation sites that play a regulatory role in the synthesis of CoQ. Specifically, the results suggest that Coq7 is active when it is not phosphorylated and inactive when it is phosphorylated.

The idea of regulating Coq7 activity via phosphorylation is actually supported by another yeast study done with phosphatase Ptc7 [37]. In this study, the DMQ₆ and CoQ₆ levels in Coq7 knockout cells and Ptc7 knockout cells were compared to the levels in wild-type cells. In Ptc7 knockout cells, when Coq7 was overexpressed, the precursor DMQ₆ accumulated; however, in Ptc7 knockout cells, when non-phosphorylatable Coq7 construct, Coq7-AAA, was overexpressed, there was a dramatic increase in CoQ₆. Thus, these results support that in order for Coq7 to be active, it must be in the dephosphorylated state, possibly from being dephosphorylated by Ptc7. In human, ADCK3 and Ptc3 may be a pair of kinase and phosphatase working together to regulate the hydroxylation activity of Coq7.

1.1.8 Characteristics of the potential substrates of ADCK3

From the evidence described above, it is predicted that the substrates of ADCK3 are Coq3, Coq5, and Coq7. Aforementioned, these proteins are part of protein complex found in the inner membrane of the mitochondria that is required for the biosynthesis of CoQ₁₀. Coq3 and Coq5 proteins are both natural product methyltransferases, which are responsible for adding a methyl group to the CoQ precursor substrate (Figure 1.9). Specifically, Coq3 and Coq5 are O-methyltransferase and C-methyltransferase, respectively, as they methylate their substrates at the oxygen and carbon atoms, respectively. For both, S-adenosyl methionine (SAM) is used as the methyl group donor and S-adenosyl homocysteine (SAH) is released as the by-product.

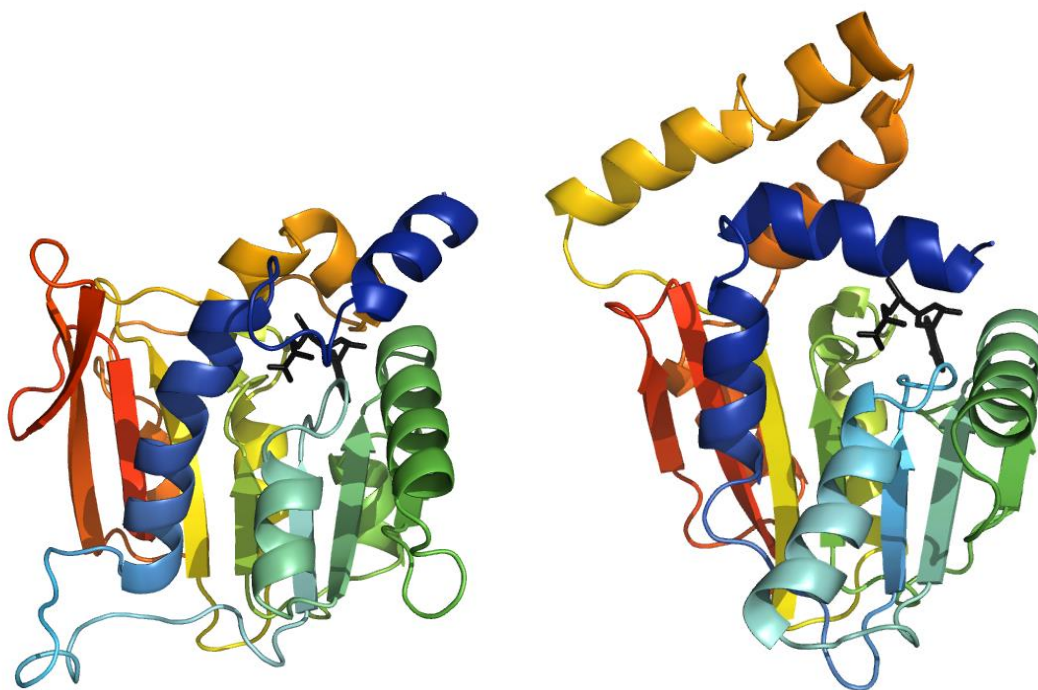


Figure 1.9 The structures of human Coq3 and Coq5 proteins predicted by PHYRE2.

The left structure shows Coq3 (96-330) and the structure on the right shows Coq5 (74-327). Both structures show the presence of the Rossmann-like fold where a β -sheet is sandwiched by 3 α -helices on both sides. The remaining structural differences may account for the placement of these proteins in the CoQ biosynthetic complex in addition to the different region of the CoQ precursor they methylate. By aligning the SAM-bound structure of yeast Coq5 (PDB: 4obw) to these predicted structures, the SAM-binding regions of these structures were identified. SAM molecule is shown in black while both protein structures are shown by rainbow from N (blue) and to C (red) terminus.

From a study done with yeast Coq5, it was established that the protein sequence of Coq5 is conserved across different species (Figure 1.10). The structure of yeast Coq5 has been solved in both apo- and holo- forms to 2.2 and 2.4 Å, respectively [28]. When the two structures were compared, a slight conformation change at the active-site pocket was observed, which is presumed to be due to the binding of SAM (Figure 1.11B and C). Like a kinase, the structure of Coq5 shows a glycine-rich motif, DVAGGSG segment, in the active site required for binding of SAM. Also, the studies revealed the putative substrate binding pocket located directly below the SAM-binding pocket where the methyl group of the SAM is oriented for easy transfer to the substrate (Figure 1.11D and E). Another functional study done with yeast Coq5 showed steady C-methyltransferase activity using a substrate analog 2-farnesyl-6-methoxy-1,4-benzoquinol [38]. In this study, they assessed and compared the C-methyltransferase activity of crude yeast mitochondrial protein from wild-type and Coq5 knockout yeast cells. After incubating the proteins with the substrate analogue and S-adenosyl-L-[methyl-³H] methionine, the lipid products were extracted and assayed for radioactivity. The results showed no radioactivity in the lipid product extracted from the reaction mixture containing Coq5 knockout yeast proteins.

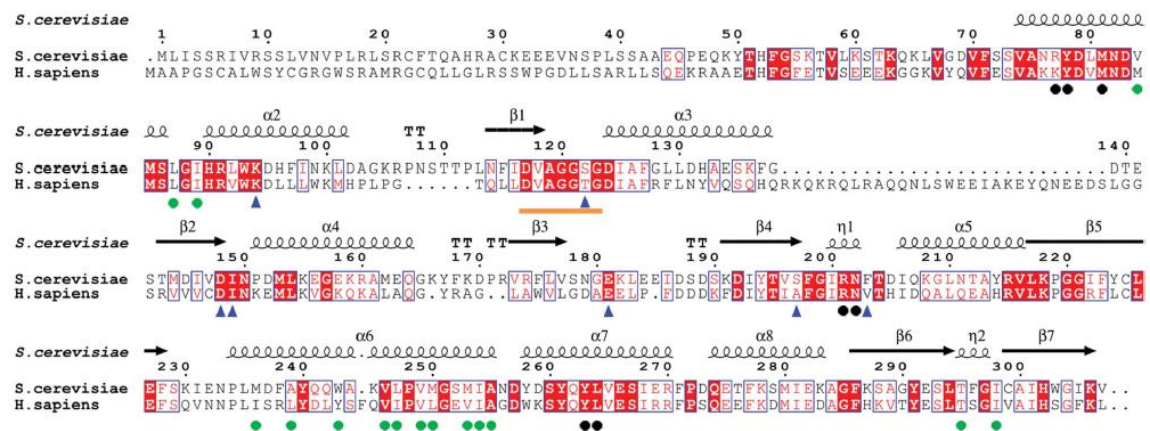


Figure 1.10 Sequence alignment of Yeast and Human Coq5.

The *S. cerevisiae* and *H. sapiens* Coq5 sequences were aligned. The DVAGGSG segment is indicated by an orange line and other residues also involved in SAM-binding are indicated with blue triangles. The residues involved in dimerization are indicated with green dots. The putative residues involved in substrate (DDMQH₂) are indicated with black dots. The conserved residues across species is indicated with red background while similar residues conserved across species is indicated with the red font. Adapted from [28].

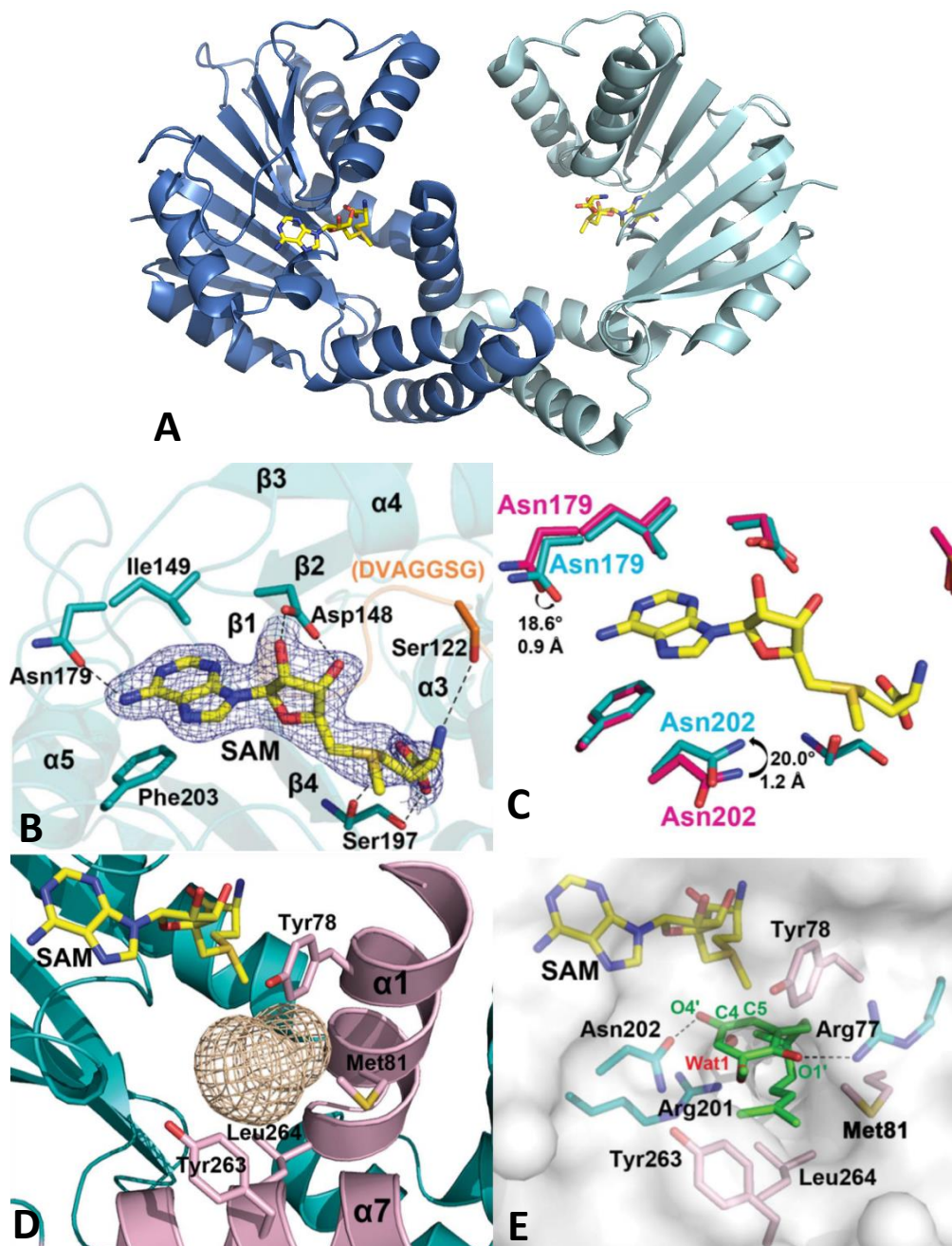


Figure 1.11 The structural features of yeast Coq5.

(A) Dimer of SAM-bound Coq5 (PDB: 4obw), (B) The residues of Coq5 involved in binding of SAM, (C) Comparison of SAM-binding residues between apo- and holo- structures (PDB: 4obx and 4obw), (D) The putative substrate (DDMQH₂) -entrance tunnel calculated by CAVER, and (E) The residues involved in substrate (DDMQH₂) -binding. SAM molecule is shown as yellow sticks, the SAM-bound structure and its residues are shown in blue or cyan, the α -helices 1 and 7 of the holo-structure in part D is shown in light purple, and the residues from the apo- structure is shown in magenta. The putative substrate DDMQH₂ is shown as brown mesh and green sticks. The SAM-bound dimerized Coq5 shows a putative pocket beneath where SAM is bound in which its substrate can enter. Adapted from [28].

On the other hand, Coq7 is a putative hydroxylase, also known as monooxygenase (Figure 1.12). Although the hydroxylation activity of Coq7 has not been directly shown, Coq7 is hypothesized to be a hydroxylase through knockout studies where the accumulated CoQ precursor lipid intermediates from Coq7 knockout strain of yeast was extracted and identified [39]. The results showed that the accumulated intermediate was indeed DMQ₆, the putative substrates of Coq7 which is a precursor of CoQ₆. It has also been shown that knocking out Coq9 reduces the activity of Coq7 [39].

Monooxygenases reduce one oxygen atom into its substrate as a hydroxyl group and the other oxygen into a water molecule. In order for this reaction to take place, the oxygen has to be activated by the donation of electron to disrupt the spin state of oxygen molecule [40]. Monooxygenases can be classified as internal or external depending on whether the source of electron for this activation process is from its substrate or an external source. In addition, these enzymes can be further classified depending on the cofactors they require, such as heme, flavin, or cofactor-independent.

Based on the structure of Coq7, it is unclear as to whether it can be concluded to have its own monooxygenase activity as it does not share any similarities compared to other monooxygenases. It seems that Coq7 is far too small to be able to anchor the lipid tail of DMQ while it adds the hydroxyl group onto the quinone ring. Aforementioned, since it has been shown that knocking out Coq9 reduces the activity of Coq7, it can be suggested that these two proteins work together to carry out the hydroxylation step. According to the cross-linking studies done by He *et al.*, Coq7 and Coq9 physically interact, as shown in Figure 1.8. It would be interesting to see how these two proteins are oriented on the inner membrane of the mitochondria to assist each other.

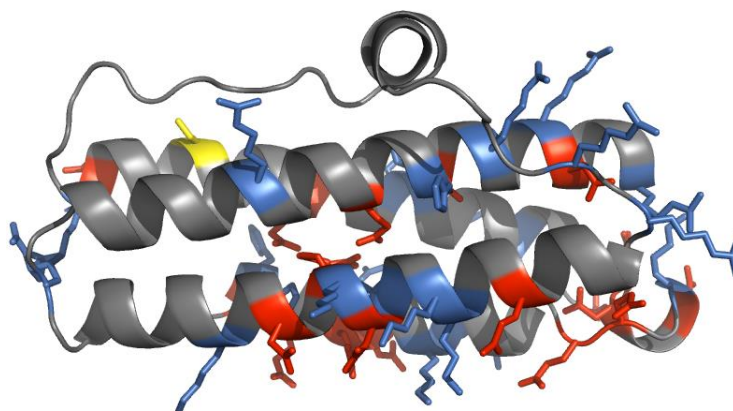
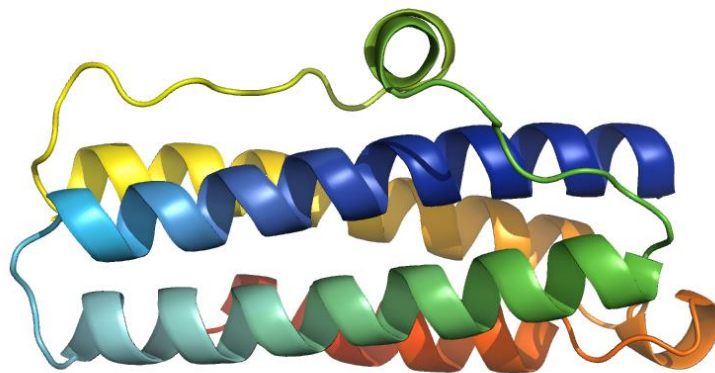


Figure 1.12 The structure of Coq7 (45-184) predicted by PHYRE2.

The left structure shows the general structure of Coq7 while the right structure shows the charged residues as sticks with colour. For the figure on the right, the positively charged residues (lysine and arginine) are shown in blue, the negatively charged residues (glutamate and aspartate) are shown in red, cysteine residue is shown in yellow, and the remain amino acid residues are shown in gray. The structure shows 4 helices forming a tube with a long, disordered region connecting helix 2 and 4. As shown in the right figure, the long, disordered region along with the small helix 3 consist of hydrophobic residues. Thus, we predict that, if Coq7 functions as a monomer, Coq7 is placed on the inner membrane of the mitochondria as such that the small helix on the outer surface of the structure is placed on the membrane side.

1.2 Project overview

1.2.1 Hypothesis and Objectives

The severe symptoms and effects of CoQ deficiency caused by mutations in ADCK3 cannot be reversed with current treatments. In order to design drugs targeting mutant ADCK3, we need to fully characterize the structure of ADCK3 and its function in the biosynthesis of CoQ₁₀.

We predict that ADCK3 is an atypical protein kinase that undergoes a conformational changes to release its N-terminal extension from its active site when activated. This enables it to use its kinase activity to regulate the CoQ₁₀ production. In particular, we predict that ADCK3 phosphorylates Coq3, Coq5, and Coq7, which may play a role in stabilizing them into the multi-subunit complex.

In order to explore our hypothesis, the project has several objectives:

1. To solve the crystal structure of the nucleotide-bound ADCK3 to highlight the key residues involved in the active site.
2. To test that Coq3, Coq5, and Coq7 are the substrate proteins of ADCK3 and further determine any additional binding proteins.
3. To clone and express soluble Coq3, Coq5, and Coq7 proteins for interaction studies with ADCK3.
4. To assess the ATPase and kinase activity of ADCK3 in the presence and absence of the putative substrate proteins.
5. To characterize the structure and to show the activities of the solubilized Coq3, Coq5, and Coq7 proteins.

1.2.2 Approaches

In this thesis, ADCK3 (residues 245-647) was cloned, expressed and solubilized for biochemical and structural characterization. First, this construct was used in co-crystallization with various nucleotides including ADP, ATP, and AMP-PNP. When crystals were obtained, the diffraction data were collected to 2.3 Å and molecular replacement of the apo- structure (PDB 4PED) was used to solve the structure; however, our structure did not contain electron density for a nucleotide. While further optimizing our conditions to trap a nucleotide into the ADCK3 structure, the AMP-PNP bound ADCK3 structure was determined by Stefely *et al.* (PDB 5I35) [41].

To identify the substrates of ADCK3 and ADCK4 through co-immunoprecipitation using HEK293 cell lysates was performed. Various constructs of ADCK3 and ADCK4 were cloned to include a N-terminal FLAG tag and successfully transfected in HEK293 cells. Following the confirmation of protein expression in HEK293 cells by a western blotting, the FLAG tag on the bait proteins were used to pull down their prey proteins using the anti-FLAG M2 resin. Once the purified complex samples were run on a gel and silver stained, several bands not found in the negative control were observed and sent for identification through tandem mass spectrometry. Unfortunately, while the mass spectrometry was being done, the results of the same experiment were published by Stefely *et al.* (PDB 5I35) [41].

Furthermore, with the confirmation that the predicted substrates of ADCK3, Coq3, Coq5, and Coq7, do physically interact, several constructs of these proteins were cloned and attempted to express in *E. coli* expression system. Although all constructs were successfully cloned, Coq3 constructs did not express while Coq5 and Coq7 successfully expressed. Coq5 and Coq7 were successfully solubilized and purified for further biochemical characterization.

In addition, the ATPase activity of ADCK3 was assessed in the absence and the presence of the substrate proteins; however, the substrate proteins, Coq5 and Coq7, were shown to have

higher activity than ADCK3 and masked its activity. Through tandem mass spectrometry, small amounts of contaminant proteins were identified to have contributed to the ATPase activity observed in Coq5 and Coq7 samples.

Kinase activity assays are also included in the thesis showing that ADCK3 (245-647) does not have kinase activity even in the presence of Coq5 and Coq7. Prior to this work, ADCK3 (245-647) has been shown to not have a significant activity. This result confirms the results of Stefely *et al.* where kinase activity done with ADCK3 in the presence of cell-free expressed Coq substrates showed no significant activity [41].

Finally, the purified Coq5 was used to show its SAM-dependent methyltransferase activity. Although the quinone substrate analogue, 2-methoxyhydroquinone molecule proved to be difficult to work with, our preliminary data show the methyltransferase activity in concentration-dependent manner, which provides the first *in vitro* evidence of such an activity for human Coq5.

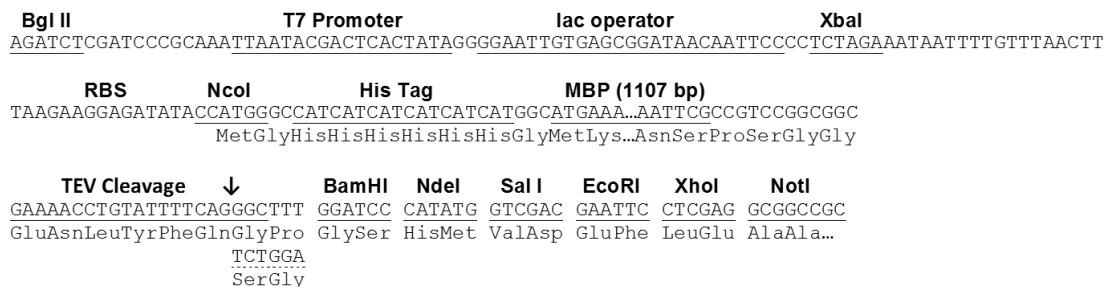
Chapter 2

Methods

2.1 Cloning of ADCK3

Genes of ADCK3 (amino acids 245-647) and ADCK3 (amino acids 329-647) constructs were amplified from the full length human cDNA (Open Biosystems, accession no. BC005171) using primers designed with EcoRI and XhoI sites engineered at the 5' and 3' ends, respectively. The PCR was carried out using the Q5 High-Fidelity Polymerase (New England Biolabs, Ipswich, MA) and the products were checked for the correct amplification using agarose gel electrophoresis. Once the lengths of the PCR products were confirmed to match the theoretical lengths, they were purified using the GeneJET Gel Extraction Kit (Thermo Fisher Scientific). Consequently, the purified products were double digested using EcoRI and XhoI restriction enzymes (New England Biolabs, Ipswich, MA) along with the expression vector HT29 (Figure 2.1).

A



B

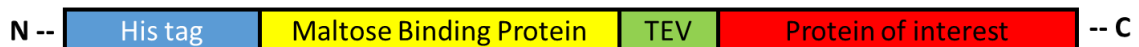


Figure 2.1 The HT29 vector engineered by Dr. Chelsey Chesterman.

(A) The vector map of HT29 vector showing the coding sequence. The major coding regions including key restriction sites are labelled and the exact TEV cleavage location is indicated with an arrow. (B) The schematic map of the resulting protein construct from HT29 vector.

HT29 vector is an expression vector modified from the pET-16b vector (Novagen, Madison, WI) by Dr. Chelsey Chesterman in our lab. The coding sequence starts with the N-terminal 6x histidine tag followed by Maltose Binding Protein (MBP), Tobacco Etch Virus (TEV) cleavage site, and multiple cloning site which includes the EcoRI and XhoI sites.

The digested PCR products and the HT29 vector were then ligated together over night at 16 °C using T4 DNA Ligase (New England Biolabs, Ipswich, MA). The ligated plasmids were then transformed into Top10 chemically competent *Escherichia coli* cells (Thermo Fisher Scientific). For each construct, a single colony was used to inoculate 5 mL of LB containing 100 µg/mL ampicillin and grown overnight at 37 °C. Plasmid DNA was purified from the 5 mL culture using the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific). The successful insertion of the gene was verified by a PCR check followed by DNA sequencing.

Site-directed mutagenesis – Complementary primers were designed so that the DNA base that is being mutated is in the centre of the primers. Using the designed primers, the plasmid containing the ADCK3 gene to be mutated was amplified using Pfu DNA polymerase (Thermo Fisher Scientific). Following the mutagenic PCR, 1 µL of Dpn1 (New England Biolabs) was added to the reaction mixture allowing for the digestion of the original methylated DNA. After incubating the reaction mixture at 37 °C for 1 hour, the digested mixture was transformed into Top10 chemically competent *E. coli* cells (Thermo Fisher Scientific). Colonies from the transformed plates were used to inoculate 5 mL of LB containing 100 µg/mL ampicillin and grown overnight at 37 °C. Plasmid DNA was purified from the 5 mL culture using the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific). The successful mutation of the gene was verified by DNA sequencing.

2.2 Expression and purification of ADCK3

The purified recombinant plasmid was transformed into BL21-CodonPlus (DE3)-RIPL competent *Escherichia coli* cells (Agilent, Santa Clara, CA). A single colony was used to inoculate 25 mL LB containing 100 µg/mL ampicillin and grown overnight at 37 °C. Subsequently, the overnight culture was used to inoculate 1 L LB containing 100 µg/mL ampicillin and 34 µg/mL chloramphenicol. The cells were shaken while being grown at 37 °C until the culture OD₆₀₀ was between 0.6 and 0.8. The culture was cooled to room temperature before protein expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The culture was incubated at 18 °C overnight.

The cells were harvested by centrifugation at 4000 rpm for 30 minutes at 4 °C and then resuspended in lysis buffer A (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol, 10 mM 2-mercaptoethanol, and 1 mg/mL lysozyme). After 20 minutes of incubation on ice, the resuspended cells were lysed by sonication (4 minutes with 5 seconds on, 15 seconds off, 30% amplitude). The debris of the cells were then pelleted by centrifugation at 16000 rpm for 30 minutes at 4 °C and the supernatant was retained for further purification.

Purification – The N-terminal His-tag engineered into the recombinant protein was used to purify ADCK3 using immobilized metal affinity chromatography (IMAC). The supernatant of the cell lysate was passed through 5 mL of nickel-nitrilotriacetic acid (Ni⁺² – NTA) resin (Qiagen, Valencia, CA) that have been pre-equilibrated with 50 mL of buffer A (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol, and 10 mM 2-mercaptoethanol). Once the supernatant has flown through the resin, 50 mL of buffer A was added to the column and allowed to flow through to wash the resin followed by 25 mL of wash buffer A (buffer A with 10 mM imidazole). To elute our target protein from the resin, 25 mL of elution buffer A (buffer A with 400 mM imidazole) was added to the column. The collected fractions were then assessed using SDS-PAGE.

Prior to further purification steps, proteins required for crystallization were subjected to an additional cleaving step, while this step was not carried out for the proteins used for the enzymatic assays. Due to the stability that MBP provides to its fused proteins, ADCK3 constructs used for enzymatic assay were uncleaved. In order to cleave the N-terminal His-tag followed by the MBP-tag, TEV protease (1:50 TEV/fusion protein, mass/mass), which was prepared in our lab, was added to the 3.5 kDa MWCO regenerated cellulose dialysis tubing (Thermo Fisher Scientific) containing the eluted protein. The dialysis tubing containing the protein were then allowed to dialyze overnight at 4 °C in cleavage buffer B (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.3 mM TCEP, 10% glycerol). Next day, the protein sample was subjected to a second IMAC using Ni²⁺ – NTA resin pre-equilibrated with buffer A. The flow through fraction, containing the target protein was kept for further purification.

The protein fractions from the previous purifications were subjected to size-exclusion chromatography over a HiLoad 16/60 Superdex 200 prep-grade column using fast-protein liquid chromatography (FPLC). Following SDS-PAGE analysis of the eluted fractions, the fractions containing purified protein were combined and concentrated using 30K and 50K Amicon Ultra centrifugal filters for cleaved and non-cleaved target proteins, respectively. The optimal centrifugation condition to result in stable protein sample was 2500 rpm at 4 °C in 15 minute increments. In order to determine the concentrations of the protein samples, the absorbance of the diluted protein samples were measured at the wavelength of 280 nm using a quartz cuvette with 1 cm width. Then, using the predicted extinction coefficient value of the protein and Beer's law, the concentration of the protein samples were calculated.

2.3 Crystallization of ADCK3

In order to obtain protein crystals of nucleotide-bound ADCK3, the ADCK3 (245-647) construct lacking the N-terminal disordered signal peptide and the transmembrane helix was used. Before plating, ATP, ADP, or AMP-PNP were individually solubilized in buffer A and was added to the protein sample at a 5:1 ratio and incubated on ice for 15 minutes. Then, using hanging vapour diffusion method, 1 μL of 2.5 mg/mL protein was mixed with 1 μL of the crystallization well solution (50 mM NaHEPES pH 8.3, 22% medium molecular PEG smear, 5 mM MgCl_2).

Once we observed the initial hits in our crystal trays, in an attempt to obtain bigger crystals with better morphology, various protein to buffer condition ratios as well as drop sizes were explored. The crystals grown in 2:2 protein to buffer condition ratio produced the biggest crystals with the best morphology. The crystal trays were incubated at room temperature and the crystals appeared the next day.

2.4 X-ray data collection and structure determination of ADCK3

The crystals were looped and cryopreserved in a cryo-protectant made with 20% glycerol, 50 mM NaHEPES pH 8.3, 22% medium molecular PEG smear [42], 5 mM MgCl_2 . The diffraction data were collected using the 23-ID-B beamline at the Advanced Photon Source (APS) at Argonne National Library (Argonne, IL). The X-ray data were collected for 360° total with 1° oscillation and processed using XDS program package [43]. The Phaser MR program was used for molecular replacement with the apo-protein structure of ADCK3 (PDB: 4PED) to solve the structure [44]. The model was improved using Phenix refinement [45] and manual fitting using Coot [46].

2.5 Co-immunoprecipitation using HEK293 cells

Cloning – Genes of 3 constructs of both ADCK3 and ADCK4 varying in lengths were cloned into pcDNA3.1+ vector as outlined in Table 2.1 (Figure 2.2). Primers were designed to result in an N-terminal FLAG tag with EcoRI and XhoI restriction sites at the 5' and 3' ends, respectively. The preceding steps are the same as the previous cloning work until the insert has been confirmed and the sequence is verified. From then, after the plasmid was transformed into Top10 cells, a colony was used to inoculate 50 mL LB and grown overnight at 37 °C. Then, the plasmid was purified using Qiagen Plasmid Midi Kit (Qiagen) and eluted with water instead of the provided elution buffer, to yield DNA with concentration higher than 1 µg/µL.

Table 2.1 Summary of ADCK3 and ADCK4 constructs used for co-immunoprecipitation using HEK293 cells.

Construct Name	Protein	Length	Molecular weight (kDa)	Construct details
3A	ADCK3	162-647	55.2	No signal peptide
3B		258-647	44.9	No signal peptide and TM domains
3C		329-647	36.6	No signal peptide, TM, and N-term extension
4A	ADCK4	46-544	55.3	No signal peptide
4B		134-544	45.8	No signal peptide and TM domains
4C		209-544	37.4	No signal peptide, TM, and N-term extension

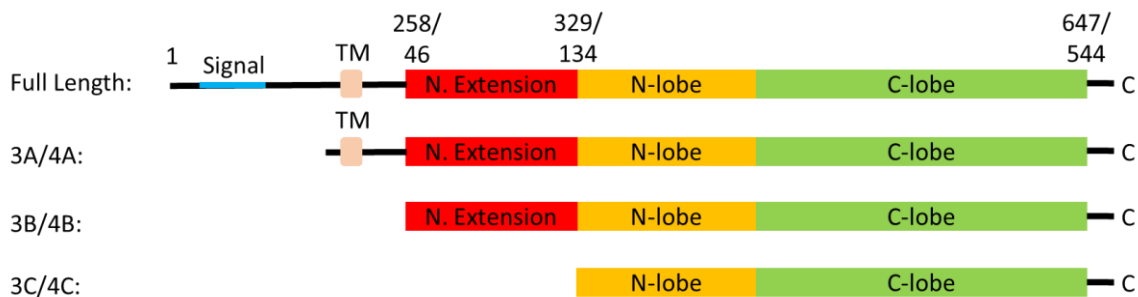


Figure 2.2 Schematic summary of ADCK3 and ADCK4 constructs used for co-immunoprecipitation using HEK293 cells.

The 3 complementary constructs for ADCK3 and ADCK4 are shown in comparison to the full length protein schematic. The residue numbers correspond to ADCK3/ADCK4.

Transfecting HEK293-6E cells with plasmids – HEK293-6E cells were grown and maintained for 1 week before preparing the cells for transfection. Once the cells recovered from the conditioned used for storing the cells, the cells were grown the cells in a T75 flask (TPP tissue culture flask) using Dulbecco's modified Eagle's medium (DMEM, Sigma-Aldrich) supplemented with 10% fetal bovine serum (FBS) at 37 °C. After 2-3 days, when they were 80% confluent, the cells from the flask were detached by adding Trypsin-EDTA and then DMEM + FBS was added to deactivate the trypsin. Using the cell counter, determine the concentration of cells to either concentrate or dilute the cells to 6×10^5 cell/mL. 1 mL of the cells and 1 mL of DMEM are dispensed in a 6 well-plate and incubated at 37 °C to attach. Next day, the old media was replaced with new media and the plasmids encoding our bait constructs were used for transfection using Lipofectamine 3000 (Thermo Fisher Scientific) as described in the manufacturer's protocol. The cells were grown for 2 days and then, cells with the positive control plasmid, eYFP, were viewed under the fluorescence microscope to check for successful transfection. Once successful transfection of the positive control has been confirmed, the sample cells were resuspended using solubilization buffer (50 mM HEPES pH 7.4, 150 mM NaCl, 0.5% (v/v) TritonX-100, 1 mM NaF, 1 mM EDTA, 0.5 mM PMSF, and ½ tablet of EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich)) and collected into a tube. The cells were lysed by passing them through a 20-gauge blunt-end needle six times. The cell debris was pelleted by centrifugation at 15000 rpm for 15 minutes at 4 °C and the supernatant was collected. After mixing the supernatant with 2x SDS-loading dye, the sample was run on a 12.5% polyacrylamide gel and further analysed by a western blot. The blot was probed using an anti-FLAG M2 antibody (Agilent, USA) and rabbit anti-mouse HRP secondary antibody. The rabbit anti-mouse HRP was then activated using Luminata Forte Western HRP Substrate (Millipore, USA) according to the manufacturer's protocols.

Co-immunoprecipitation - Once protein expression was observed, cells were grown in T75 flasks (TPP tissue culture flask) and transfected with higher concentration of lipofectamine 3000 described by the manufacturer's protocol. When the cells have grown for 2 days and the positive control cells have been checked, the media in the sample cell flasks was aspirated off and the cells were resuspended in solubilization buffer (50 mM HEPES pH 7.4, 150 mM NaCl, 0.5% (v/v) TritonX-100, 1 mM NaF, 1 mM EDTA, 0.5 mM PMSF, and ½ tablet of EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich)). The cells were then lysed and subjected to centrifugation for 15 minutes at 15000 rpm at 4 °C. The resulting supernatant was then incubated with pre-equilibrated Anti-Flag M2 resin (Sigma-Aldrich) on ice for 2 hours. The resin was then washed with wash buffer (50 mM Tris-HCl pH 7.45, 150 mM NaCl, 0.5% Triton X-100, 1 mM NaF, and 1mM EDTA) 3 times followed by 3 more washes with PBS buffer. Next, 2x SDS loading-dye was added to the resin and the sample was boiled for 5 minutes at 95 °C before being ran on 12.5% polyacrylamide gel. The resulting gel was initially analyzed after being stained with Coomassie stain and then further stained using a ProteoSilver™ Silver Stain Kit (Sigma-Aldrich). The protein bands of interest were excised and sent for identification using liquid chromatography - mass spectrometry, LC-MS. The samples were digested using Trypsin as a preparation and then the LC-MS was carried out using a Thermo Fisher Orbitrap Velos Pro.

Data analysis – Data was processed using the MASCOT Peptide Mass Fingerprint server [47]. The searches were performed against Human_EST, NCBIprot, and SwissProt databases. For each search, maximum of 1 tryptic cleavage was allowed. The best fitting protein, in terms of the highest peptide matches as well as consideration of the protein molecular weight, from both human and non-human source were chosen.

2.6 Cloning of Coq3, Coq5, and Coq7

The genes of Coq3, Coq5, and Coq7 were amplified from full length human cDNA (Dharmacon, GE Healthcare Cat # MHS6278-202801177, MHS6278-202841399, MHS6278-202828122, respectively). The various constructs designed and amplified are summarized in Table 2.2. The primers were designed with EcoRI and XhoI sites engineered at the 5' and 3' ends, respectively. The PCR was carried out using the Q5 High-Fidelity Polymerase (New England Biolabs, Ipswich, MA) and the products were checked for the correct amplification using agarose gel electrophoresis. Once the lengths of the PCR products were confirmed to match the theoretical length, they were purified using the GeneJET Gel Extraction Kit (Thermo Fisher Scientific). Consequently, the purified products were double digested using EcoRI and XhoI restriction enzymes (New England Biolabs, Ipswich, MA) along with the expression vector HT29.

The digested PCR products and the HT29 vector were then ligated together overnight at 16 °C using T4 DNA ligase (New England Biolabs, Ipswich, MA). The ligated plasmids were then transformed into Top10 chemically competent *E. coli* cells (Thermo Fisher Scientific). For each construct, a single colony was used to inoculate 5 mL of LB containing 100 µg/mL ampicillin and grown overnight at 37 °C. Plasmid DNA was purified from the 5 mL culture using the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific). The successful insertion of the gene was verified by a PCR check followed by DNA sequencing.

Table 2.2 Summary of Coq protein constructs fused to maltose binding protein.

Construct name	Original Protein	Construct length	Molecular weight (Without MBP)	Molecular weight (with MBP)
Q3a	Coq3	1-369 (Full)	41.1 kDa	84.0 kDa
Q3b		85-369	31.5 kDa	74.3 kDa
Q3c		96-333	26.4 kDa	69.2 kDa
Q5a	Coq5	1-327 (Full)	37.1 kDa	80.1 kDa
Q5b		43-327	32.9 kDa	75.5 kDa
Q7a	Coq7	1-217 (Full)	24.3 kDa	67.3 kDa
Q7b		37-217	20.5 kDa	63.2 kDa

2.7 Expression and purification of Coq3, Coq5, and Coq7

The methods for expression and purification are the same as ADCK3 expression with some minor changes. The buffer used to purify Q5b consists of 50 mM NaHEPES pH 7.0 and 300 mM NaCl. Q5b also requires additional ion-exchange purification step using the Resource Q 6 mL column (GE Healthcare). When performing ion-exchange chromatography, the final % Buffer B target was set to 50%, which corresponds to 500 mM NaCl, and the fraction size was set to 1 mL. For Q7a and Q7b, the purification buffer is 50 mM Tris pH 8.0, 300 mM NaCl.

2.8 Analysis of protein folding of Q5b and Q7a using Circular Dichroism spectroscopy

The protein folding of the MBP fusion constructs of Coq5 and Coq7, Q5b and Q7a, respectively, were analyzed using the Chirascan CD spectrometer (Applied Photophysics Inc., Beverly, M.A., U.S.A.) in the Protein Function Discovery Facility at Queen's University (Kingston, ON). Both 20 μ M of Q5b and 17 μ M Q7a samples were analyzed using a 0.1 mm pathlength cuvette. For each sample, 6 scans of absorbance between 195 and 260 nm were measured in 1 nm increments. The average of the 6 scans were used for further analysis. Deconvolution was performed using OLIS SpectralWorksTM and the CONTINLL algorithm (On-Line Instrument Systems, Bogart, GA).

2.9 Methyltransferase activity of Coq5

In order to test the methyltransferase activity of Coq5, Q5b construct was used with 2-methoxyhydroquinone as the substrate. Initially, the SAM510: SAM Methyltransferase Assay kit (GBioscience, Cat # 786-430) was used (Figure 2.3); however, it was later determined that the colorimetric assay is inadequate for use because 2-methoxyhydroquinone changes colour as it auto-oxidizes in the presence of oxygen. The colour change by the substrate masked the absorbance reading at 510 nm. Therefore, an alternative method for assessing the methyltransferase activity was explored.

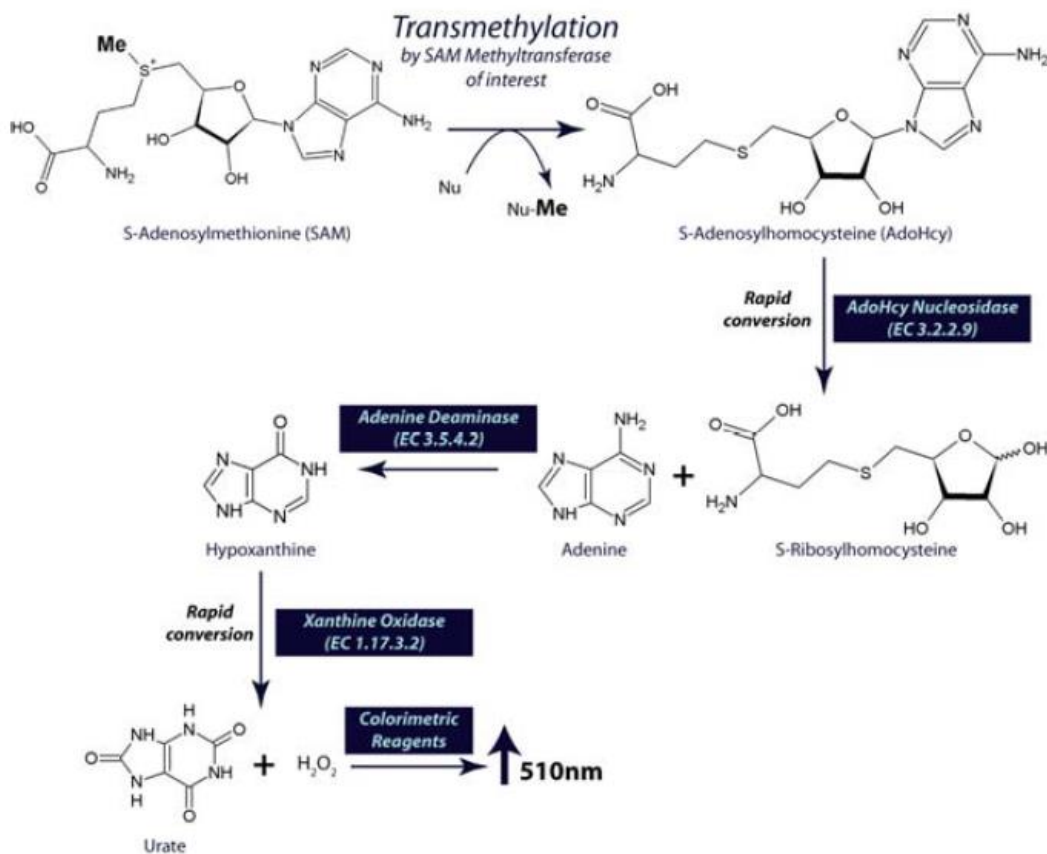


Figure 2.3 The reaction scheme of the SAM510: SAM Methyltransferase assay.

The scheme shows the enzymes involved in this colorimetric continuous enzyme assay. In addition to the methyltransferase enzyme being assessed, the reaction mixture includes AdoHcy Nucleosidase, Adenine Deaminase, and Xanthine Oxidase. Once the methyltransferase carries out its reaction, the cofactor by-product S-adenosylhomocysteine (AdoHcy) is converted into the colourimetric indicator by the 3 enzymes in addition to the colorimetric reagent. The kinetic absorbance measurements of the reaction mixture are taken every 30 seconds at 510 nm.

SAMfluoro: SAM Methyltransferase assay kit (GBioscience, Cat # 786-431) was suggested by the supplier as the alternative assay to overcome the difficulty of working with 2-methoxyhydroquinone (Figure 2.4). Even for the fluorescence assay kit, auto-oxidation of the substrate was a problem. As a result, the substrate was solubilized just before being added to the reaction and the kinetic measurement was started as soon as the master mix was added. The assay was performed according to the protocol provided by the supplier in triplicates in Corning 96 well clear bottom, black walled plates. The final concentration of 2-methoxyhydroquinone was 200 μM while the Q5b concentrations were 0 μM , 1 μM , 2 μM , 3 μM , 4 μM , and 5 μM .

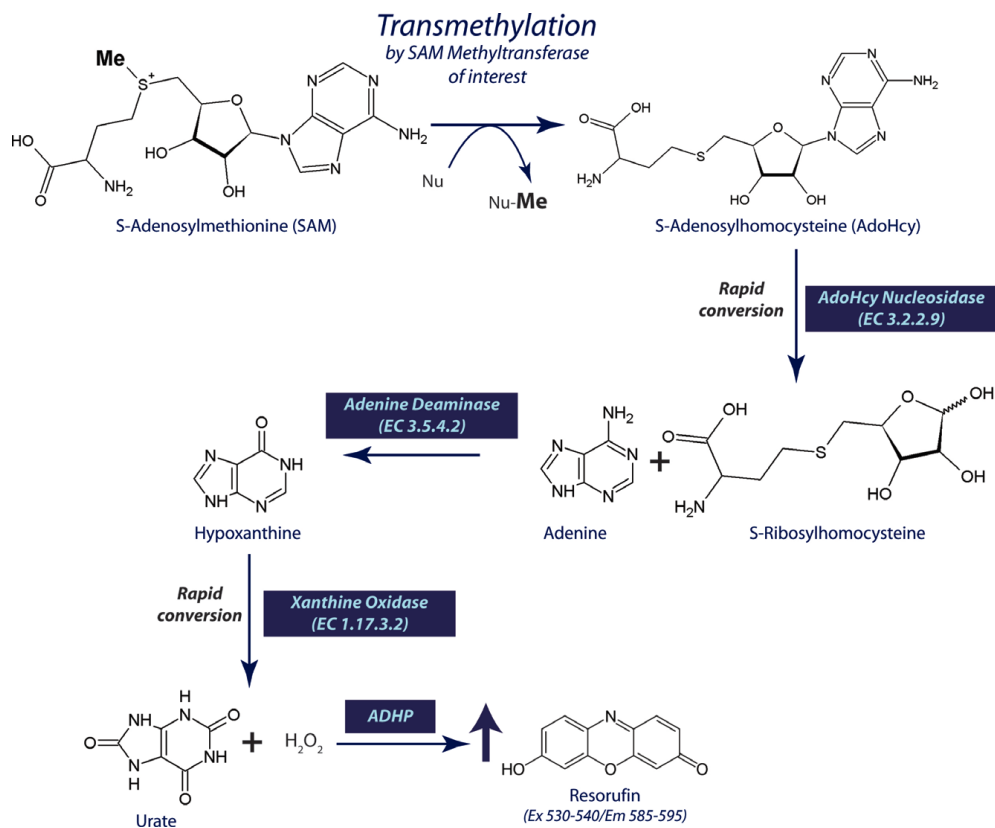


Figure 2.4 The reaction scheme of the SAMfluoro™:SAM Methyltransferase assay.

The scheme shows the enzymes and the intermediates involved in this continuous fluorescent enzyme assay. In addition to the methyltransferase enzyme being assessed, the reaction mixture includes AdoHcy Nucleosidase, Adenine Deaminase, and Xanthine Oxidase. Once the methyltransferase carries out its reaction, the cofactor by-product S-adenosylhomocysteine (AdoHcy) is converted into the colorimetric indicator by the 3 enzymes in addition to the colorimetric reagent, ADHP. The kinetic fluorescence measurements of the reaction mixture are taken every 2 minutes with the excitation and emission set to 535 nm and 590 nm, respectively. Other parameter included using LOW gain and integration time of 1000 ms.

2.10 Crystallization trials of Coq5 and Coq7

Various screening trays were set up with both 10 mg/mL and 5 mg/mL protein samples using the Phoenix Robot (Art Robbins Instruments). The crystallization condition screening kit used were: Classics Suite, Classics II Suite, PEGs Suite, PEGs II Suite, and AmSO₄ Suite (Qiagen). The trays consist of Intelli-Plate 96-2 Shallow Well Low-Profile plates (Art Robbin Instruments) with sitting drops comprised of 0.4 μ L protein and 0.4 μ L of screening condition. The plates were kept in the cold room at 4 °C.

2.11 ATPase activity assay

The ATPase activity of ADCK3 was assessed using the P_iColorLock™ assay kit (Innova Biosciences, Cambridge, UK). The reactions were carried out at 25 °C in the assay buffer containing 20 mM Tris-HCl pH 8.0, 20 mM MgCl₂, and 2 mM ATP. Negative control reaction containing no kinase protein was used to measure the background noise of the reaction. The final protein concentration used in the experimental reactions for ADCK3 (245-647) was 10 μ M, while the concentrations of Q5b, Q7a, and Q7b were 1 μ M. The experiment was performed according to the manufacturer's protocol (Figure 2.5). The absorbance of the sample reactions was measured at 635 nm using a microplate spectrophotometer (Bio-Tek, Winooski, VT). Then, from the standard curve created using known concentrations of free phosphates, the linear equation was used to determine the concentration of free phosphates released during the sample ATPase reactions.

paper squares were dried for 15 minutes before being washed twice in 1.6 % o-phosphoric acid for 15 minutes. Then, the paper squares were placed on a bed of paper towel in a plexiglass chamber to dry over night. Next day, the paper squares were immersed in 5 mL of ScintiVerse™ II cocktail buffer (Thermo Fisher Scientific). The radioactivity of the samples was measured using Beckman LS 9000 scintillation counter (Beckman Coulter, Brea, CA).

Chapter 3

Results and Discussion

3.1 Cloning, mutagenesis, expression, and purification of ADCK3

The successful insertions of the ADCK3 genes in HT29 vector was confirmed via sequencing. In addition, the successful mutation of K276 residue to glutamate residue was also confirmed via sequencing. The *E. coli* cells transformed with the verified plasmids were grown in LB media and induced with IPTG to express ADCK3. The results for MBP-ADCK3 (245-647) expression and purification show successful expression and purification via Nickel chromatography (Figure 3.1). The SDS-PAGE gel of the samples collected during Nickel chromatography shows a purified band at around 90 kDa which corresponds to the expected molecular weight of 89.1 kDa for MBP-ADCK3 (245-647). When the elution fractions were collected for further purification using size-exclusion chromatography, the results showed elution of MBP-ADCK3 (245-647) at the elution volume of 75 mL (Figure 3.2). This elution volume corresponds to the elution volume of proteins with the molecular weight of 90 kDa as indicated by the manufacturer's guidelines. Typical purification of this construct from 1 L of cells yielded 28 mg of protein.

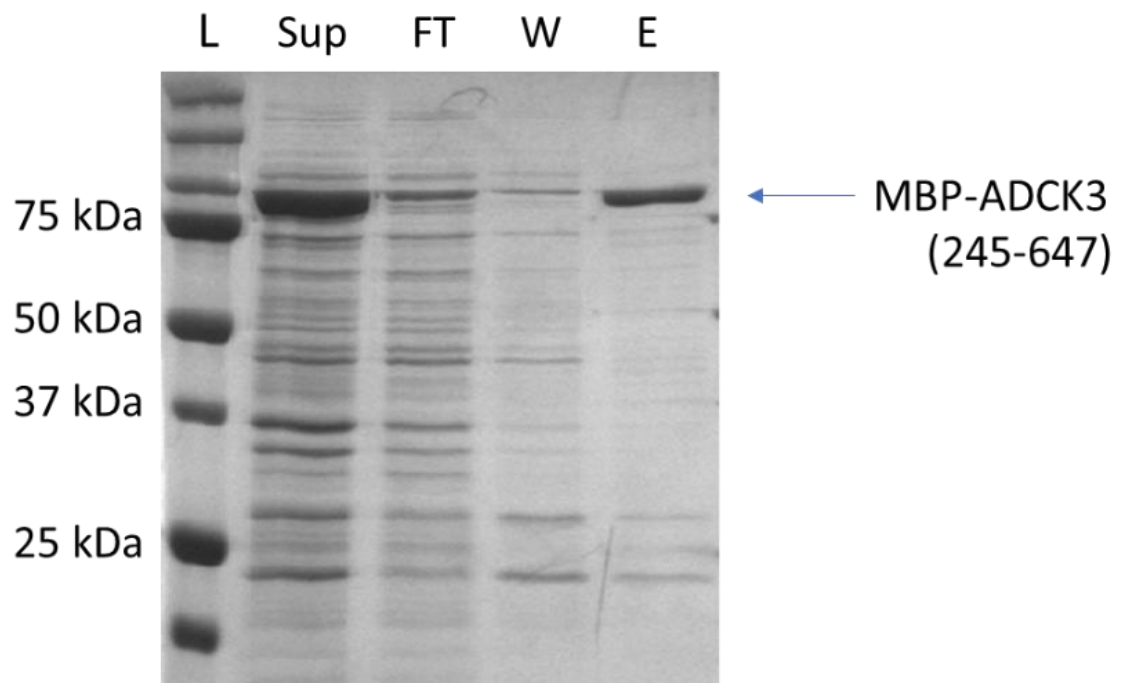


Figure 3.1 A representative profile of MBP-ADCK3 (245-647) purification using histidine-mediated immobilized metal affinity chromatography.

Using the histidine-tag engineered into the construct, MBP-ADCK3 was extracted from the supernatant of the cell lysate using Ni²⁺-NTA column. Throughout the purification process, 10 μ L of sample were collected and analysed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining. From left to right, the lanes represent: (L) ladder; (Sup) supernatant; (FT) flow-through; (W) wash; and (E) Elution. The elution fractions were collected for further purification.

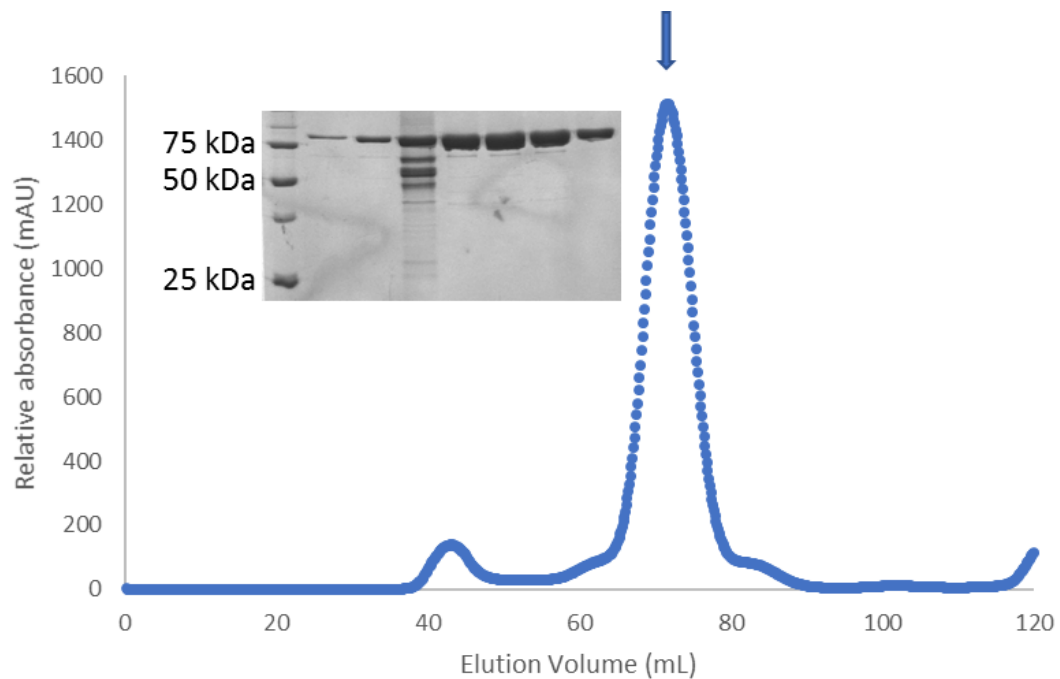


Figure 3.2 Chromatogram of MBP-ADCK3 (245-647) purification via size-exclusion chromatography.

Elution fractions from Ni⁺²-NTA purification were collected and further purified via size-exclusion chromatography over a HiLoad 16/60 Superdex 200 prep-grade column using FPLC. Peak containing ADCK3 protein is indicated by a blue arrow. The fractions of that peak were analyzed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining.

ADCK3 protein samples required for crystallization were subjected to TEV cleavage after the initial purification using Ni⁺²-NTA chromatography. After the His-tagged MBP was separated from ADCK3 via second round of Ni⁺²-NTA chromatography, the flow-through fraction containing ADCK3 was further purified using size-exclusion chromatography. The chromatogram of size-exclusion chromatography show a shift in the protein peak from the elution volume of 75 mL to 85 mL (Figure 3.3). This correctly corresponds to a peak for proteins of the molecular weight of 45 kDa as indicated by the manufacturer's guidelines. At the end of the purification, the yield of ADCK (245-647) was 10 mg of protein. In addition to the elimination of a large protein tag MBP, the decrease in yield may be due to loss of protein during the additional purification steps.

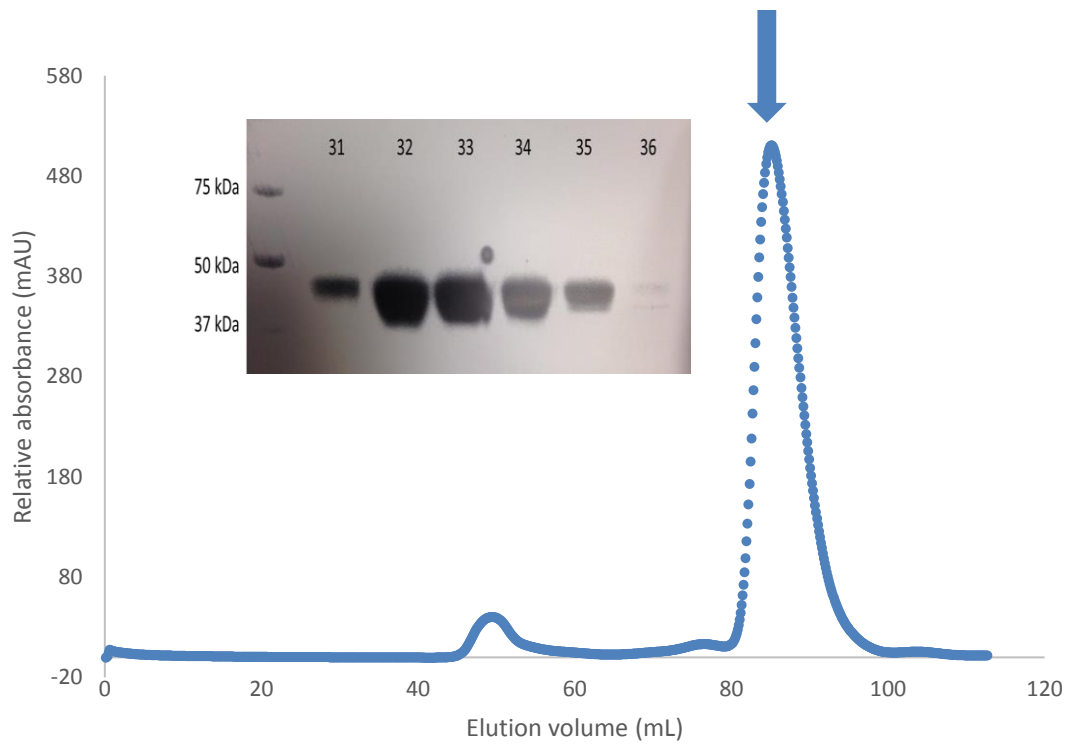


Figure 3.3 Chromatogram of the cleaved ADCK3 (245-647) purification via size-exclusion chromatography.

After the MBP was cleaved off the MBP-ADCK3 (245-647) construct with TEV protease, ADCK3 (245-647) was purified using Ni^{+2} -NTA chromatography. Then the flow-through fraction containing the protein of interest was further purified using size-exclusion chromatography over a HiLoad 16/60 Superdex 200 prep-grade column using FPLC. Peak containing ADCK3 protein is indicated by a blue arrow. The fractions of that peak were analyzed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining.

Site-directed mutagenesis was used to make a K276E mutation in ADCK3 (245-647) construct in an attempt to disrupt one of the salt-bridges between the N-terminal extension and the kinase domain. It was predicted that by changing the charge of this residue, the repulsion between the two residues would release the N-terminal extension away to expose the active site of the kinase domain. The results of the expression and purification of this construct with K276E mutation also show successful expression and purification; however, this construct's yield was significantly less and was prone to degradation (Figure 3.4). The decrease in the yield may be due to the degradation of the protein, as the mutation made for this construct aimed to release the N-terminal extension. If this mutation did lead to the release of the N-terminal extension, then it may lead to a flexible N-terminal extension that may not be as stable compared to the original construct. As well, perhaps active site of ADCK3 by releasing the N-terminal extension adds to the instability of this construct. Although the purified protein was initially stable, heavy precipitation formed once a flash frozen aliquot of the sample was thawed in preparation of an experiment. In order to compensate for this, perhaps different buffers can be explored to accommodate the residues being exposed from the active site. At the same time, perhaps various protease inhibitors can be added in the buffer to reduce cleavage of this construct by unspecific proteases.

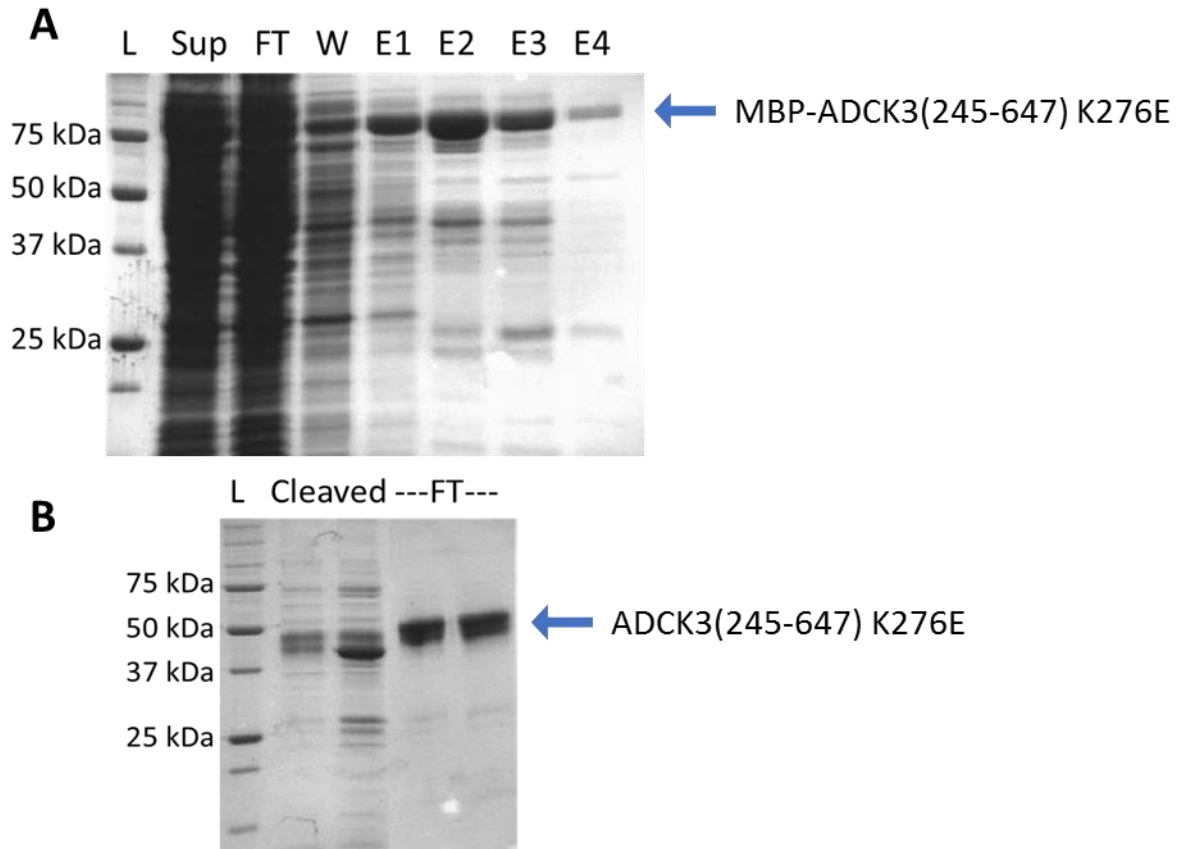


Figure 3.4 A representative profile of ADCK3 (245-647) K276E purification.

(Left) shows the purification of MBP-ADCK3 (245-647) K276E and (Right) shows the purification of the cleaved ADCK3 (245-647) K276E. Using the histidine-tag engineered into the construct, MBP-ADCK3 was extracted from the supernatant of the cell lysate using Ni⁺²-NTA chromatography. Then, the elution fractions were collected and was subjected to TEV protease cleavage overnight. The His-tagged MBP was captured by the Ni⁺²-NTA column while the ADCK3 protein flowed through. This fraction was collected for further purification. Throughout the purification process, 10 μ L of sample were collected and analysed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining. The protein being purified are indicated with the blue arrow.

Furthermore, the results shows unsuccessful expression for the MBP-ADCK3 (329-647) construct (Figure 3.5). SDS-PAGE analysis shows a purified protein band with the molecular weight of 45 kDa. The expected molecular weight of MBM-ADCK3 (329-647) is 79.8 kDa. We predict that this is the result of the expression of His-tagged N-terminal MBP as it was able to bind to the nickel resin. It is intriguing that the cells would be able to express the longer construct of ADCK3 but not the shorter construct. As a result, we predict that the expression problem is not due to the codons of ADCK3 that are rare in *E. coli* cells. In addition, we eliminated the possibility of ADCK3 being cleaved off the tag even before the purification as both the pellet and the supernatant lanes do not show a band with molecular weight of 36.6 kDa (Figure 3.5). It may, however, in fact be that ADCK3 (329-647) is expressed but actively degraded by the proteases in *E.coli*. A possible explanation as to why ADCK3 (329-647) construct is degraded while the ADCK3 (245-647) construct is not degraded might be because the shorter construct that does not have the N-terminal extension is active. Perhaps without the N-terminal extension, ADCK3 is more active and is able to phosphorylate various substrates found in the cell; however, as previously mentioned, the cell regulates and maintains the phosphorylation activity of kinases very tightly. As a result, in order to inhibit this active human kinase that may be damaging the cell, the *E. coli* cell might be actively degrading ADCK3 (329-647) or inhibiting its expression altogether. In order to overcome this difficulty, this construct has been cloned into various vectors with Thioredoxin tag or a Nus tag, in addition to the vectors used for insect cell expression, that may help to express and stabilize this construct.

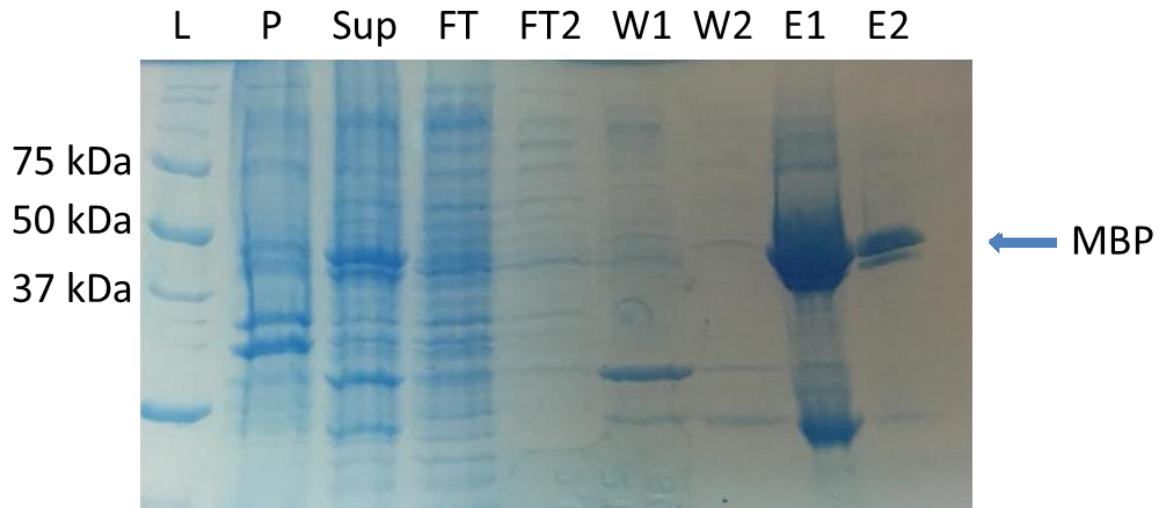


Figure 3.5 A representative profile of MBP-ADCK3 (329-647) purification using histidine-mediated immobilized metal affinity chromatography.

Using the histidine-tag engineered into the construct, MBP-ADCK3 was extracted from the supernatant of the cell lysate using Ni^{+2} -NTA chromatography. Throughout the purification process, 10 μL of sample were collected and analysed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining. From left to right, the lanes represent: (L) ladder; (P) pellet; (Sup) supernatant; (FT) flow-through 1; (FT2) flow-through 2; (W1) wash 1; (W2) wash 2; (E1) Elution 1, and (E2) Elution 2.

3.2 Crystallization of ADCK3

When ADCK3 was initially crystallized, it contained ammonium sulfate in the crystallization condition. In many cases of co-crystallization of protein kinases with a ATP, the sulfate ion of ammonium sulfate is a problem as the sulfate ion is similar to the phosphate of the ATP. As a result, the sulfate ion is often observed in the active site of these kinases, inhibiting the ATP from being able to stay in the active site. As a result, a new crystallization condition for ADCK3 was sought for its co-crystallization with a nucleotide.

Initial protein crystal hits obtained using medium molecular weight PEG smear were either single plates or plate clusters (Figure 3.6). The crystals were verified to be protein crystals by imaging with UV light. Initially, the crystal trays were incubated at 4 °C to grow and store the crystals; however, due to unforeseen circumstances, the crystallization condition had to be re-optimized to be able to grow the crystals at room temperature.

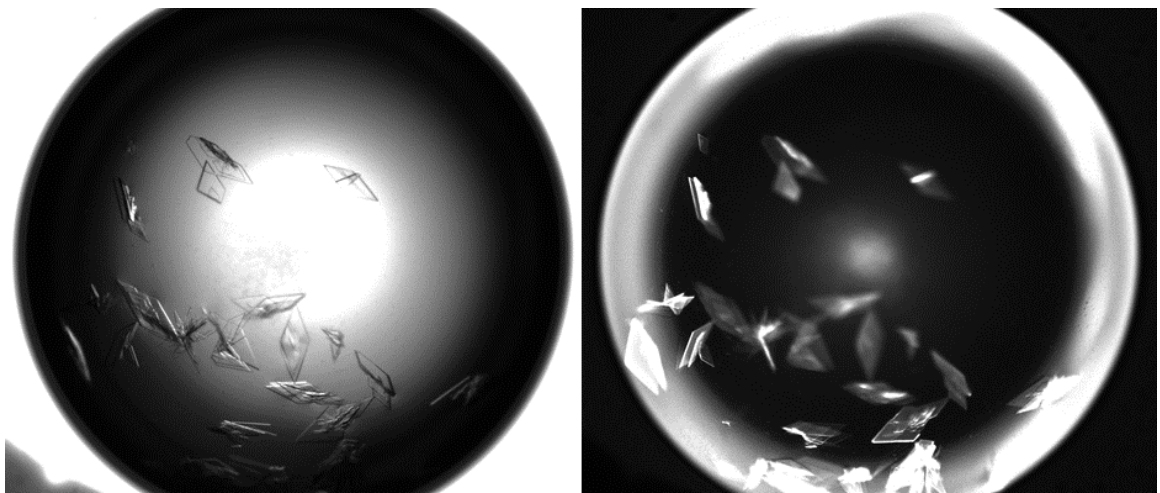


Figure 3.6 Initial crystal hits of ADCK3 (245-647) in medium molecular weight PEG smear crystallization condition.

The crystals were imaged both under visible light (left) and UV light (right).

After finding a crystallization condition that allows ADCK3 to crystallize at room temperature, efforts were made to improve the crystal quality and increase its size. In order to improve the quality of the crystals, the concentration of the MMW PEG smear and $MgCl_2$ were adjusted. Also, additive screens (Hampton research, Aliso Viejo, CA, U.S.A.) were tested to see if additional reagents will help to improve the crystal quality. In order to increase the size of the crystals, various drop sizes of the protein-well solution were explored. At room temperature, the best crystallization condition (50 mM NaHEPES pH 8.3, 22% medium molecular PEG smear, 5 mM $MgCl_2$) in addition to 4 μ l sized drops with 1:1 protein to well-solution ratio, yielded protein crystals after 1 day. The crystals grew as colourless, oblique prisms but quickly degraded within a week (Figure 3.7). As a result, all crystallization trays were set up 4 days before they were looped to be sent for X-ray diffraction at APS in Chicago.

Prior to looping the crystals, 1 μ L of either ADP, ATP, or AMP-PNP was slowly pipetted onto some of the protein crystal drop in an attempt to saturate the kinase with nucleotides; however, the crystals were observed to crack or quickly degrade. This may be due to the change in the protein conformation induced by the addition of the nucleotides. As mentioned, the N-terminal extension may have been “loosened” or released from the core of the kinase domain to expose the active site to accommodate for its cofactors.

In addition, crystallization attempts made with the ADCK3 (245-647) K276E construct resulted in heavy precipitation. Optimization of the crystallization conditions have been made with no progress. This precipitation may be due to degradation as observed during its purification. Furthermore, the inability of this construct to crystallize may also be due to the flexibility of the N-terminal extension caused by the mutation to release the extension from the active site of the kinase.

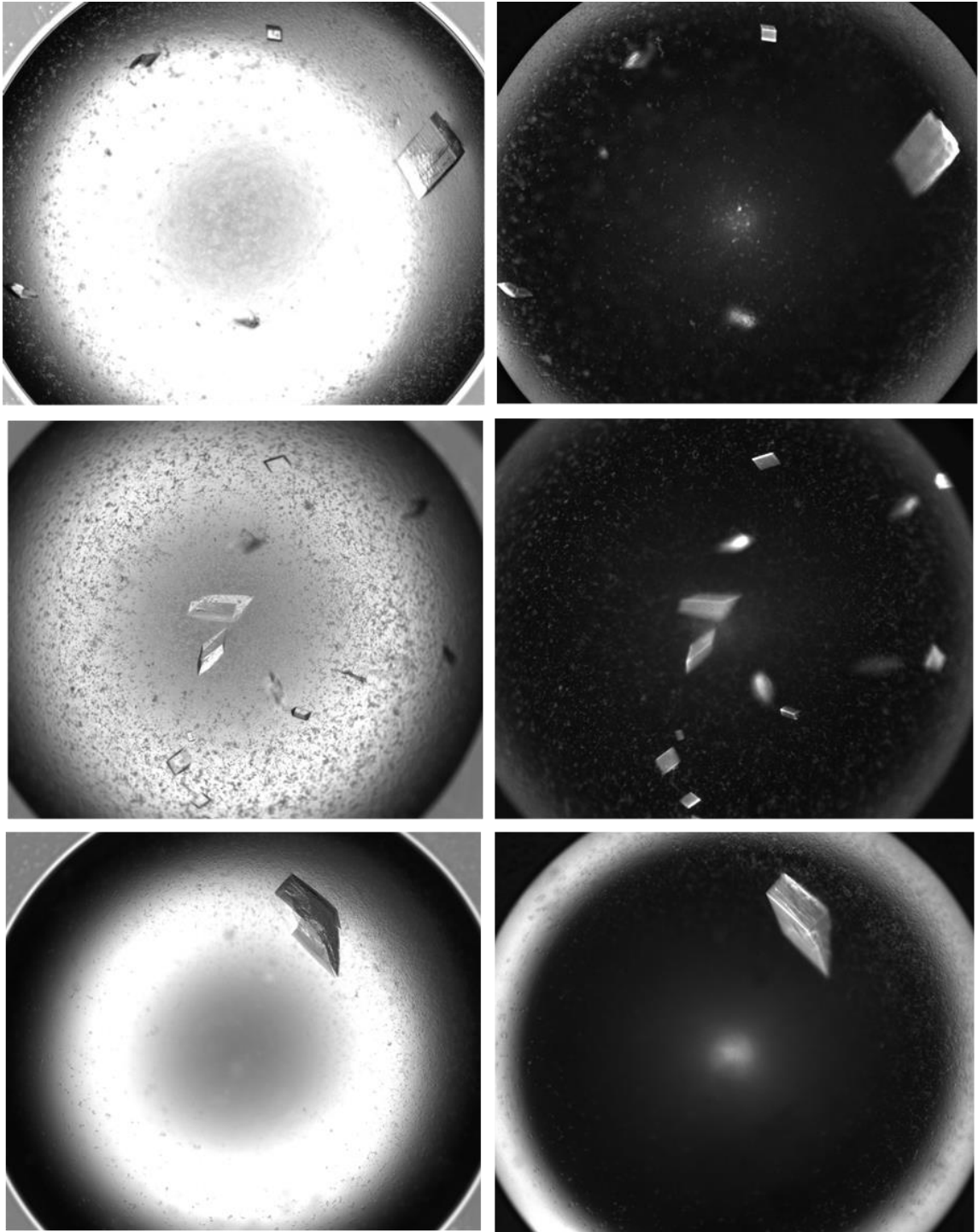


Figure 3.7 Crystals of ADCK3 (245-647).

Prior to crystallization, the protein sample was incubated with either ADP, ATP, or AMP-PNP. The crystals were grown in 50 mM NaHEPES pH 8.3, 22% medium molecular PEG smear, 5 mM MgCl₂. The crystals were imaged both under visible light (left) and UV light (right).

3.3 X-ray data collection and structure determination of ADCK3

The best crystals of ADCK3 (245-647) construct diffracted to 2.4 Å with 99.6% completeness (Table 3.1). When the electron density map was analyzed, the results displayed C2 space group with unit cell dimensions: $a = 152.27$ Å, $b = 54.65$ Å, $c = 49.14$ Å, $\alpha = 90^\circ$, $\beta = 96.62^\circ$, and $\gamma = 90^\circ$. The majority of the protein from residues 258 to 644 were clearly visible in the electron density. Unfortunately, upon analysis of our electron density maps of our crystals, there were no signs of a nucleotide bound to the structure; however, we went on to refine our models to look for any specific data sets that show structures that show a different conformation. Upon several rounds of refinement, when the apo-structure of ADCK3 and our refined structures were compared, our structures were found to overlay the apo-structure and does not show a conformation with significant difference (Figure 3.8).

During the optimization of the crystallization of both ADCK3 (245-647) and ADCK3 (245-647) K276E, a study by Stefely *et al.* published an AMP-PNP-bound ADCK3 structure [41]. In order to do so, Stefely *et al.* also used AMP-PNP in place of ATP and a R611K mutant construct. Stefely *et al.* explained that R611K mutant was used as they were unable to co-crystallize the wild-type protein with a nucleotide. In their previous paper that analyzed the apo-structure, it was noted that when the structure of PKA and ADCK3 were compared, the catalytic residue D488^{ADCK3} was found to be oriented in the opposite direction to the complementary one of the PKA residue D166^{PKA} [14]. In PKA, this residue is oriented toward the terminal phosphate of the ATP to act as a catalytic base. As a result, it was concluded that along with the K276E salt bridge, D488-R611 salt bridge may play a role in inhibiting phosphorylation activity of ADCK3.

Compared to the apo-structure, the nucleotide-bound structure shows a more open conformation (Figure 3.9A). This slight change in the conformation may be due to the mutation made to the ADCK3 construct used for crystallization or AMP-PNP binding. The ADCK3 construct used for their co-crystallization has a mutation in arginine 611 residue that resulted in

lysine 611 residue. Prior to the mutation, the hydrogens of both nitrogen atoms from the arginine side chain are interacting with the two oxygen atoms of the glutamic acid 488 side chain (Figure 3.9B). Although both arginine and lysine side chains harbour positive charges, the shorter length of lysine may have caused the loss of interaction between its positive side chain to the negative charges of the glutamic acid 488 side chain (Figure 3.9C). The loss of 2 hydrogen bonds between residues in the core of the structure may have loosened the overall structure, allowing for enough room for the nucleotide to enter; however, it is uncertain whether the binding of the nucleotide has further changed the conformation in addition to the change induced from the mutation.

Table 3.1 X-ray data collection and refinement statistics.

	ADCK3 (245-647)
<i>Data collection</i>	
Space group	C2
Cell dimensions:	
a (Å)	152.27
b (Å)	54.65
c (Å)	49.14
α (°)	90
β (°)	96.62
γ (°)	90
Resolution range (Å)	2.4 – 43.4
Total reflections	188582
Unique reflections	26595
Completeness (%)	99.6%
R_{sym} (%)	5.4
$I/\sigma I$	18.15
No. of molecules in the asymmetric unit	1
<i>Refinement</i>	
R_{work} (%)	25.57
R_{free} (%)	29.77
RMSD for bond length (Å)/bond angle (°)	0.021/2.08
Ramachandran plot preferred (%)	92.31
Ramachandran plot allowed (%)	5.35
Ramachandran plot outliers (%)	2.34

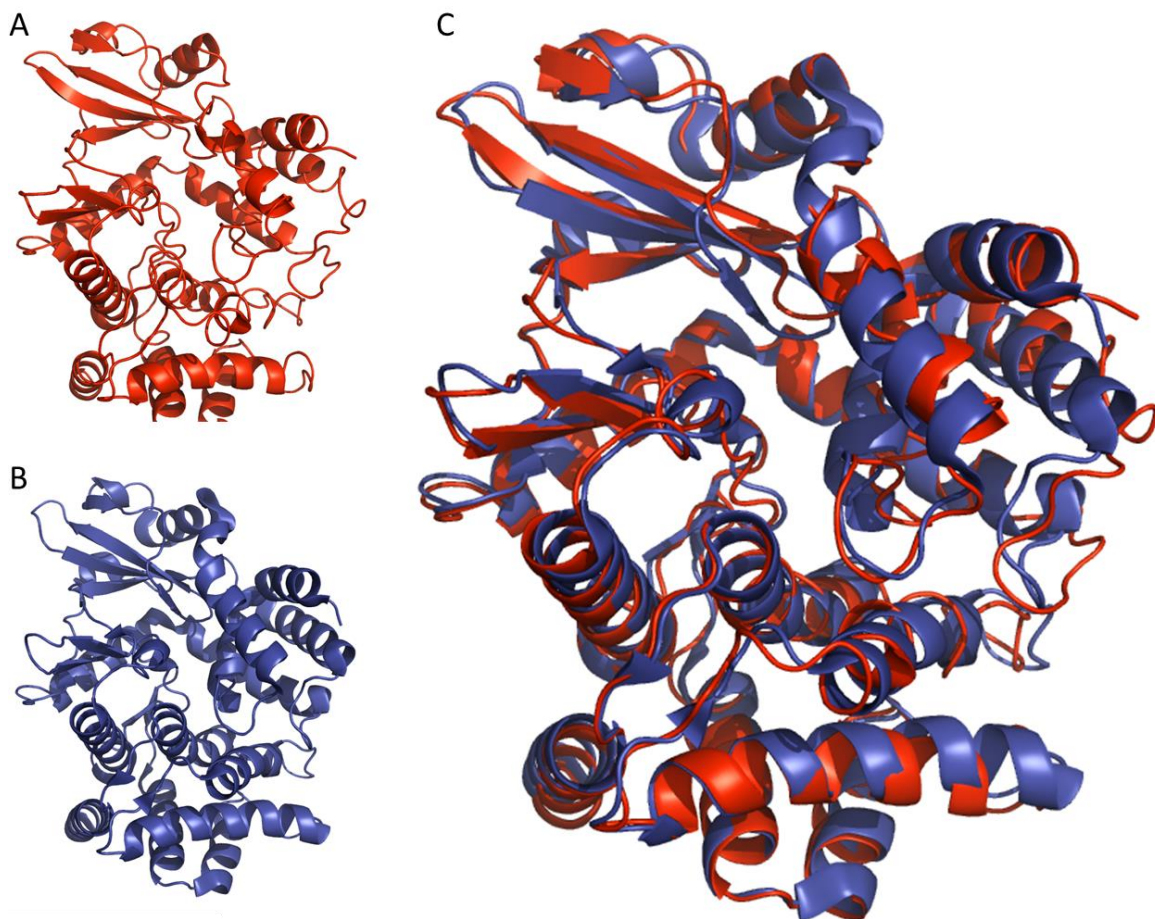


Figure 3.8 The comparison between the apo-structure of ADCK3 and our structure. (A) The structure we solved and refined is shown in red, (B) the structure from the protein data bank is shown in blue (PDB: 4PED), and (C) overlay of the two structures.

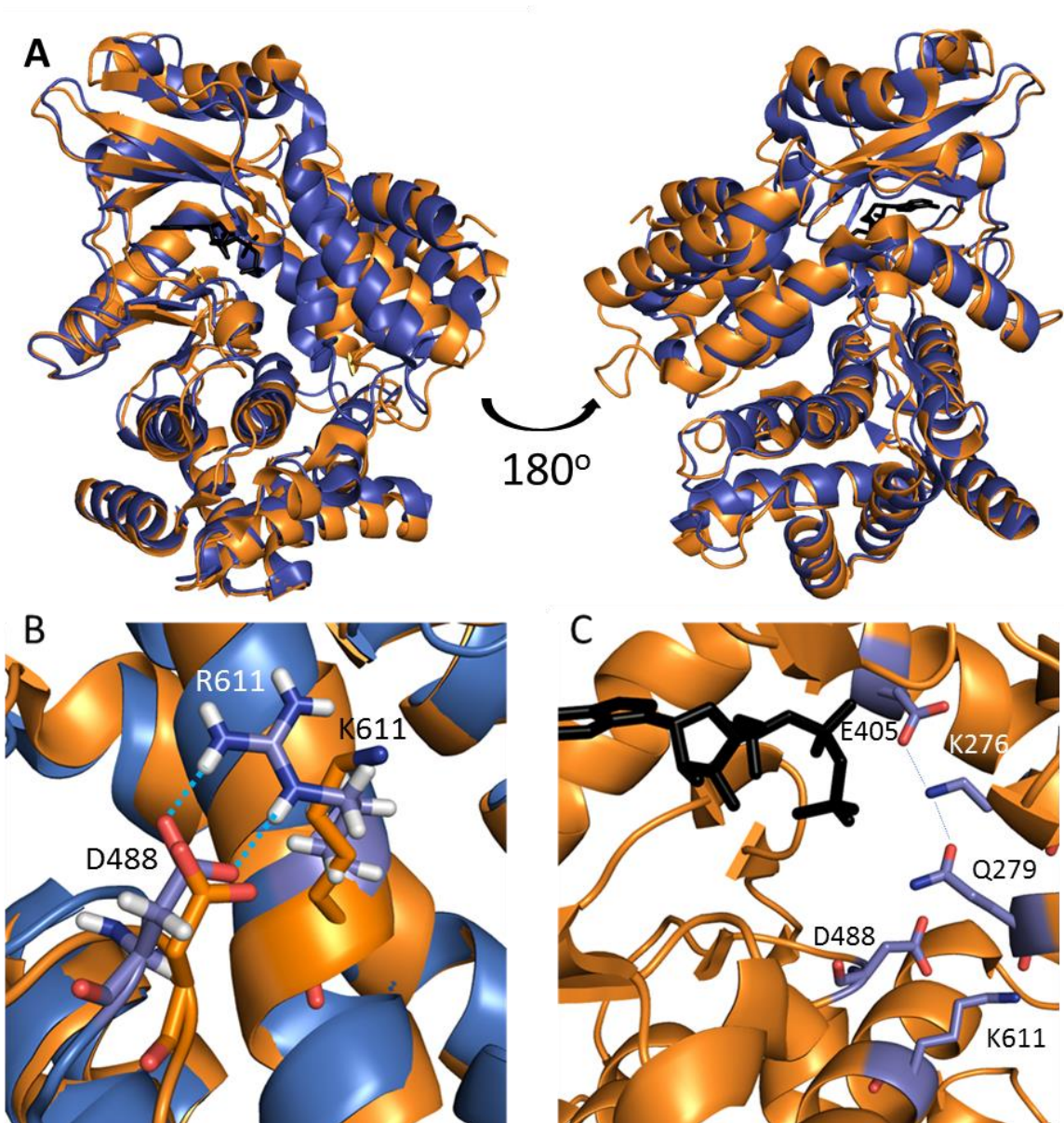


Figure 3.9 The structural features of AMP-PNP bound ADCK3 (261-644) R611K in comparison to the apo-structure of ADCK3 (254-647)

(A) Overlay of the holo- and apo- ADCK3 structures, (PDB: 5I35 and 4PED, respectively), (B) The structural difference between the wildtype residue K611 and the mutated residue R611, and (C) the residues of ADCK3 that interact with the mutate R611 residue. The holo- structure is shown in orange, the apo-structure is shown in blue, and AMP-PNP molecule is shown in black. Key residues are shown as sticks and labelled. The key residues are labelled and their interactions between each other are shown as blue dotted lines.

3.4 Co-immunoprecipitation using HEK293 cell

In order to pull-down potential substrates and binding partners of ADCK3 and ADCK4, various constructs of FLAG-tagged ADCK3 and ADCK4 were over-expressed in HEK293 and purified. The cloning component of the experiment was successful as the sequencing results reveal FLAG-tag followed by our ADCK genes inserted into the correct vector. The constructs A, B, and C were designed for both ADCK3 and ADCK4 with the hypothesis that ADCK3 and ADCK4 proteins in different conformation would pull-down different substrate or binding partners (Table 2.1, Figure 2.2). First, the longest constructs 3A and 4A were designed to exclude the signal peptide to mimic the mature ADCK3 proteins found in the mitochondria. This construct was used in hope to pull-down proteins that may bind to ADCK3 and ADCK4 to activate them. Constructs 3B and 4B were designed to exclude the signal peptide as well as the transmembrane domain. The purpose of using these constructs was similar to the purpose of using 3A and 4A constructs except that 3B and 4B would be easier to purify as it lacks the transmembrane domain. Finally, constructs 3C and 4C were designed to exclude the signal peptide, transmembrane domain, and the N-terminal extension. The purpose of using these constructs was to pull-down the substrates that dock right into the active site of ADCK3 and ADCK4.

Subsequently, the plasmid was purified from 50 mL cells harbouring the plasmid using the midi prep kit, yielding 59.1, 29.3, 31.4, 51.5, 56.3, and 70.9 μg of 3A, 3B, 3C, 4A, 4B, and 4C plasmids, respectively. Then, these plasmids were transfected into the HEK293 cells successfully since the positive control eYFP was expressed by the HEK293 cells based on fluorescence microscopy. Next, to confirm the expression of our proteins and check their over-expression levels, the transfected cells were lysed and the supernatant was collected for SDS-PAGE and western blot analysis. As shown in Figure 3.10, all constructs for both ADCK3 and ADCK4 were over-expressed and were bound to the anti-FLAG M2 resin. Although the western

blot results are not shown, it was clear that the protein bands highlighted with red boxes in Figure 3.10 were indeed the FLAG-tagged ADCK3 and ADCK4 proteins. Since the pulled-down protein bands that appear in all sample lanes were very similar, we chose the full-length constructs, 3A and 4A, to be used for the Co-IP experiment.

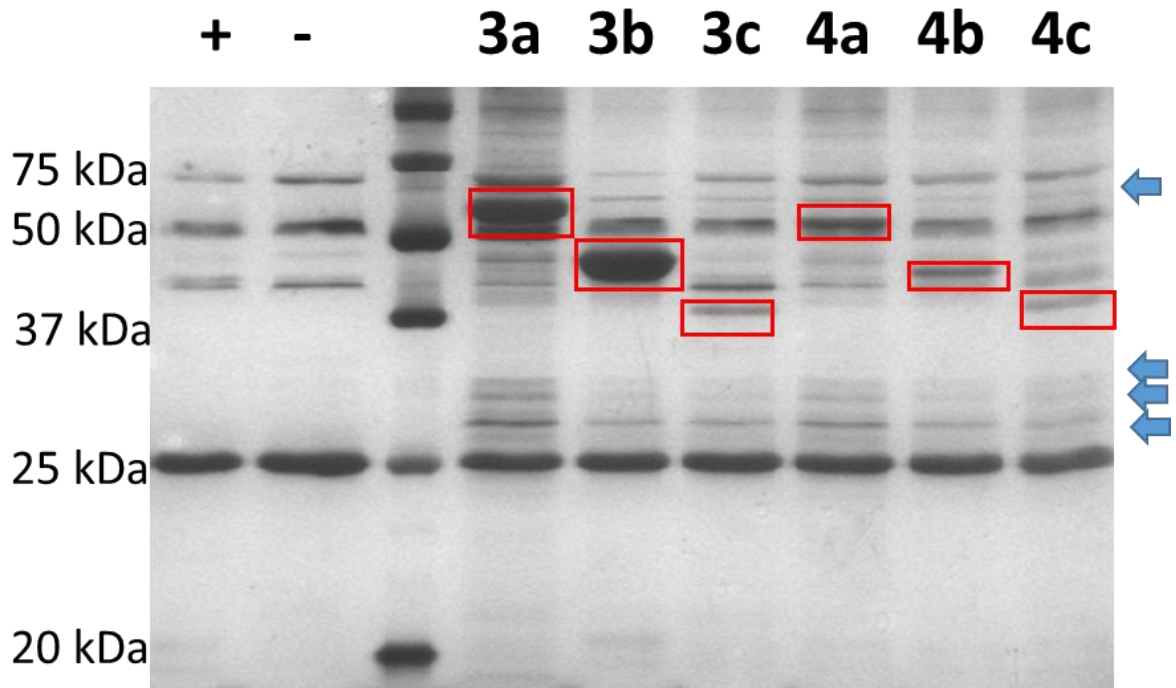


Figure 3.10 Silver stained SDS-PAGE gel of the co-immunoprecipitation samples for all constructs.

The overexpressed FLAG-tagged constructs were purified from HEK293 lysate using anti-FLAG M2 resin. The co-immunoprecipitated protein samples for all constructs were eluted from the anti-FLAG M2 resin by the addition of protein loading buffer and heat denaturation. Then the samples were run on a 12.5% SDS polyacrylamide gel. The gel was then silver stained using the ProteoSilver™ Silver Stain Kit for visualization. From left to right the lanes are: positive control eYFP; negative control – no plasmid; Ladder; 3A construct; 3B construct; 3C construct; 4A construct; 4B construct; and 4C construct. The red boxes indicate the overexpression of the corresponding construct. The blue arrows indicate proteins in the sample lane that do not appear in the negative control lane.

In order to verify the reproducibility of the pull-down of the 4 protein bands using 3A and 4A constructs, the transfection and Co-IP experiments were repeated a couple of times. Finally, last round of transfection with 3A and 4A plasmids were done and the proteins pulled down using the anti-FLAG M2 resin were ran on a gel. 5 protein bands from each of 3A and 4A sample lanes were identified using LC-MS (Figure 3.11). The LC-MS results were matched against pool of *H. sapiens* proteins and the results confirmed the expression of ADCK3 3A construct (Table 3.2); however, for both 3A and 4A results, no significant matches were observed.

One interesting protein identified may be the F₀F₁ ATP synthase subunit a for band 2 from the ADCK4A lane. Although this protein may be a contaminant as its sequence was matched from a bacterial species, it is not the first time that ATP synthase F₀ subunit has been mentioned in the ADCK3 project. In a radiometric kinase assay study done by Brody Wheeler in our lab, ADCK3 was shown to phosphorylate a peptide corresponding to site in ATP synthase F₀ subunit 8 [48]. Prior to identifying Coq proteins as the substrates for ADCK3, this F₀ subunit 8 seemed likely to be the substrate as this protein is also found embedded in the inner membrane of the mitochondria. As a result, the human ATP synthase F₀ subunit 8 protein was cloned, expressed, and used to test the kinase activity of ADCK3 and ADCK4 (Appendix A).

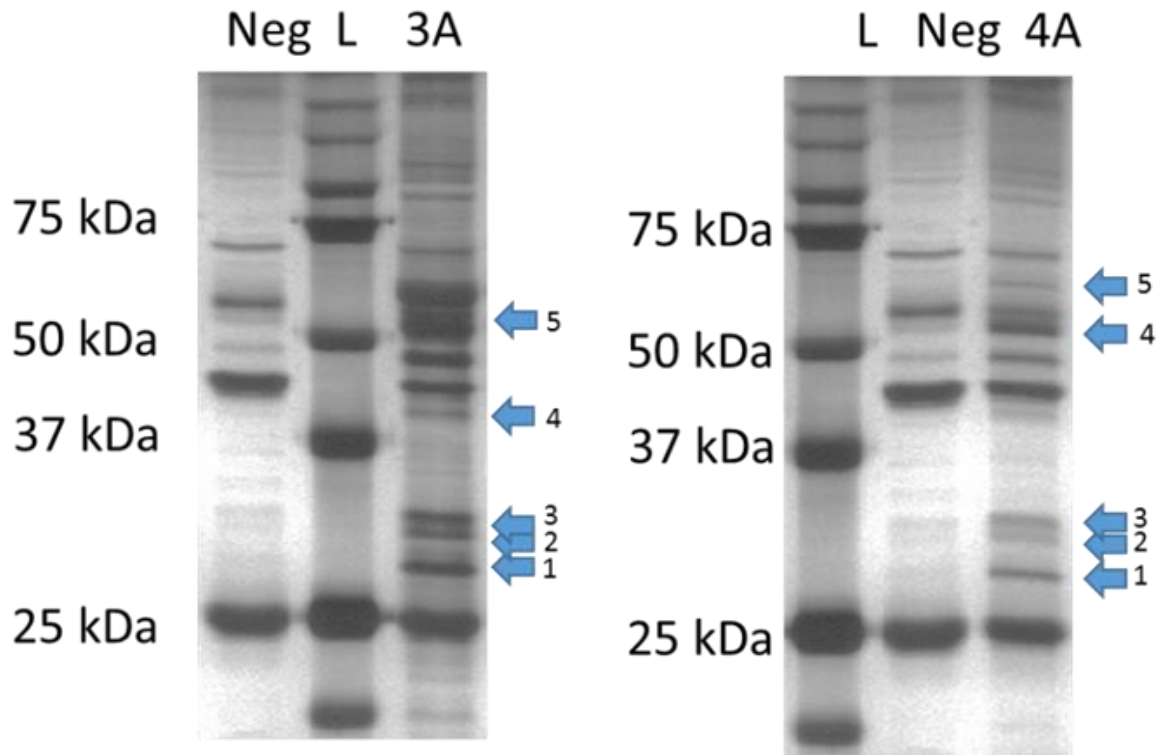


Figure 3.11 Silver stained SDS-PAGE gel of the co-immunoprecipitation samples for 3A and 4A constructs.

Left: ADCK3 sample with ladder and negative control. Right: ADCK4 sample with ladder and negative control. The blue arrows with numbers correspond to the bands chosen for mass spectrometry analysis.

Table 3.2 Identities of the proteins co-immunoprecipitated from HEK293 cells.

	ADCK3 3A		ADCK4 4A	
Band	No confident match	[human]	SLC25A5	[human]
1	ADP/ATP translocase	[Other]	ADP/ATP translocase	[Other]
Band	40S ribosomal protein	[human]	G617756.1 602645274F1	[human]
2	40S ribosomal protein	[Other]	F0F1 ATP synthase subunit a	[Other]
Band	BM006903	[human]	Soares infant brain 1NIB	[human]
3	Ribosome recycling factor	[Other]	Hypothetical protein EGK	[Other]
Band	HY150107 (male brain RIKEN)	[human]	B-Actin	[human]
4	lysophospholipase	[Other]	B-Actin	[Other]
Band	ADCK3	[human]	No confident match	[human]
5	COQ8	[Other]	Hypothetical protein	[Other]

Failure to obtain significant and clear-cut results may have been due to couple of factors. First, the protein bands of interest may not have been separated far from each other, leading to mixture of proteins. For example, for many bands, the results indicated that there are peptides from multiple proteins. To fix this, perhaps 2-dimensional gel electrophoresis can be done to separate even the proteins that are the same sizes that may overlap in SDS-PAGE gels. Also, for the negative control, empty pcDNA3.1+ vector should have been transfected into the HEK293 cells instead of transfecting water. This may have allowed for elimination of unspecific proteins that bind to the anti-FLAG M2 resin.

In the midst of optimizing our protocol for this experiment, a similar study was published showing the binding of Coq 3, Coq4, Coq5, Coq6, and Coq7 to ADCK3 proteins [41]. Similar to our study, HEK293 cells were transfected with pcDNA3.1 plasmid harbouring FLAG-tagged ADCK3 gene; however, the FLAG tag was placed on the C-termini and the ADCK3 construct

length has not been identified, meaning that full length was most likely used. In addition, unlike the lipofectamine 3000 we used, Stefely *et al.* used PEI (linear polyethylenimine MW 25,000). In addition, their solubilization buffer contained an extensive number of protease inhibitors which may have contributed to the identification of more proteins from their samples. Finally, for their experiment, FLAG-peptides were used to elute the protein off the anti-FLAG M2 beads, which is a more specific way to elute the protein off the resin rather than adding the protein loading dye right into the beads. This may be the reason for the lack of contaminants found in their results; however, Stefely *et al.* did not run the eluted samples onto a gel prior to preparing individual bands for mass spectrometry analysis. Instead, the entire elution mixture was digested with trypsin and analyzed by mass spectrometry. This may contribute to the lengthy list of proteins identified by their mass spectrometry analysis. All in all, the results of the Co-IP experiment done by Stefely *et al.* further suggest that Coq3, Coq5, and Coq7 have the potential to be the substrates of ADCK3.

3.5 Cloning of Coq3, Coq5, and Coq7

In order to obtain soluble Coq3, Coq5, and Coq7 proteins for interaction studies with ADCK3, the genes of the various truncations of these proteins were cloned into bacterial expression vectors. The cloning of various truncations of Coq3, Coq5, and Coq7 into expression vectors proved to be difficult for unknown reasons. Specifically, the amplification of the genes from the cDNA required additional dimethyl sulfoxide (DMSO) in the PCR solution and for Coq3 truncation cloning, the resulting plasmids were unable to be sequenced despite showing the successful insertion from a PCR check and a double restriction digest check; however, eventually, all truncations were verified to have successfully cloned into HT29 vector through sequencing.

3.6 Expression and purification of Coq3, Coq5, and Coq7

For the expression of Coq proteins, various *E. coli* cell lines including BL21, Rosetta Gami, Rosetta 2, Artic Express, and pLySs were explored; however, the only cell line that successfully expressed Coq5 and Coq7 was the BL21-CodonPlus (DE3)-RIPL *E. coli* cell line. A possible explanation for this may be the fact that some codons used in human protein expression are rarely found in *E. coli*. The BL21-CodonPlus (DE3)-RIPL *E. coli* cell line was engineered to include tRNA genes to increase the availability of Arginine (encoded by AGA and AGG), Isoleucine (encoded by AUA), Proline (encoded by CCC), and Leucine (encoded by CUA). Upon DNA sequence analysis, it was determined that human Coq3, Coq5, and Coq7 genes contain 29, 29, and 23 rare codons, respectively. As summarized in Table 3.3, the number of rare codons was significantly reduced by using the RIPL *E. coli* cell line.

Table 3.3 Summary of DNA sequence analysis of Coq3, Coq5, and Coq7.

Construct name	Original protein	Construct length (amino acid)	Number of Rare codons in Gene	
			w/ out RIPL cell	w/ RIPL cells
Q3a	Coq3	1-369 (Full)	23	2
Q3b		85-369	9	1
Q3c		96-333	9	1
Q5a	Coq5	1-327 (Full)	29	10
Q5b		43-327	21	6
Q7a	Coq7	1-217 (Full)	23	9
Q7b		37-217	16	6

Expression trials of Coq3 constructs Q3a, Q3b, and Q3c – Although the genes of each construct were successfully inserted into the expression vector with no mutations, upon extensive expression trials, no sign of protein expression was observed (Figure 3.12). At this time, it was uncertain as to why there was no expression of any of these constructs. When transformations of these plasmids were performed, the same aliquot of cells used to transform other vectors successfully expressed their encoded proteins. One hypothesis is that the mRNA transcripts from the plasmid harbouring these genes are forming secondary structures that inhibit the ribosomes from binding to it to begin the translation [49]. It is predicted to be so because during the cloning process, it was difficult to amplify the genes of all three constructs from the Coq3 cDNA. This suggests that perhaps the cDNA plasmid has a secondary structure that inhibited the amplification of the genes. In addition, when the genes were found to have successfully ligated into the vector via PCR check, sequencing of the plasmids often resulted in incomprehensible sequence. In addition, Coq3 gene only has 2 rare codons when using the RIPL *E. coli* cells. If this was the problem, then the cells would have at least expressed the His-tagged MBP encoded prior to the Coq3 insert. In order to successfully express Coq3 protein for its characterization, the cDNA of Coq3 may need further DNA sequence analysis and optimization to inhibit any secondary structure formation.

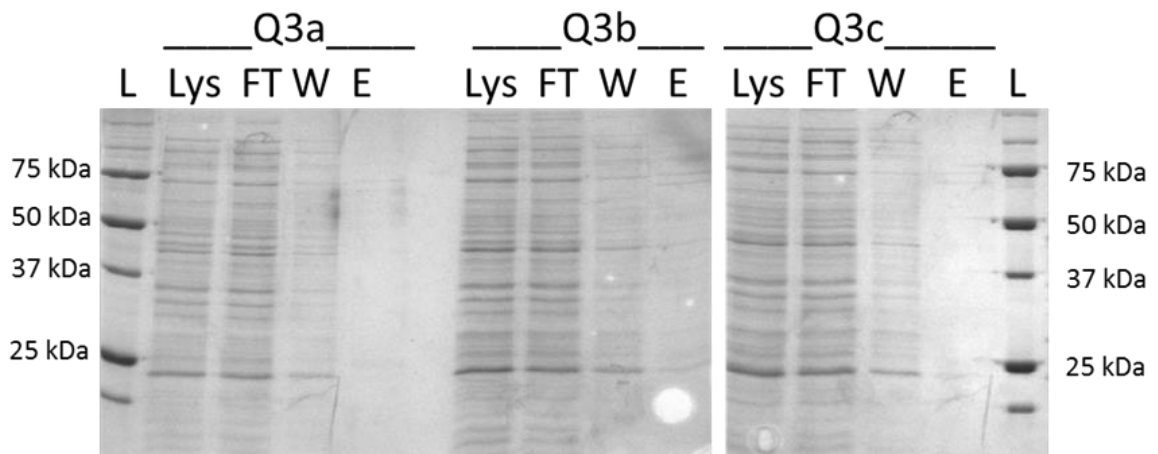


Figure 3.12 Q3a, Q3b, and Q3c purification using histidine-mediated immobilized metal affinity chromatography.

Using the histidine-tag engineered into the construct, Q3a, Q3b, and Q3c were extracted from the supernatant of the cell lysate using Ni^{+2} -NTA chromatography. Throughout the purification process, 10 μL of sample were collected and analysed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining. From left to right, the lanes represent: (L) ladder; (Lys) lysate; (FT) flow-through; (W) wash; and (E) elution. The samples of each lane were taken from the purification of the protein construct listed above the gel.

Expression trials of Q5a and Mutagenesis of Q5a – Although cloning of both Q5a and Q5b constructs into the HT29 vector was successful, the expression trials failed to show any expression of Q5a. Instead, from SDS-PAGE analysis of the purification using Ni⁺²-NTA showed expression and purification of a protein with the molecular weight of approximately 45 kDa (Figure 3.13). This protein is predicted to be His-tagged MBP, with an approximate molecular weight of 43 kDa, as it was able to bind to Ni⁺²-NTA. As a result, sequence analysis was done to investigate the cause of lack of expression of Q5a which was not the case for Q5b.

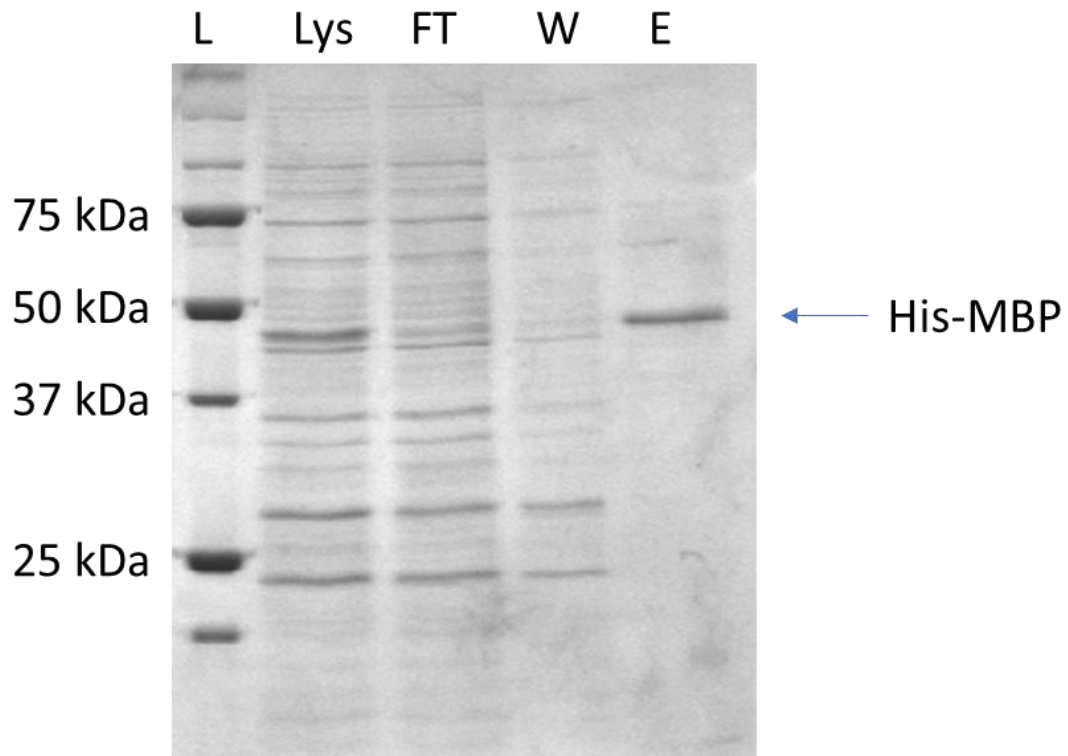


Figure 3.13 Q5a purification using histidine-mediated immobilized metal affinity chromatography.

Using the histidine-tag engineered into the construct, Q5a was extracted from the supernatant of the cell lysate using Ni^{+2} -NTA chromatography. Throughout the purification process, 10 μL of sample were collected and analysed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining. From left to right, the lanes represent: (L) ladder; (Lys) lysate; (FT) flow-through; (W) wash; and (E) elution.

DNA sequence analysis revealed that, although the number of rare codons was significantly reduced by using the RIPL *E. coli* cell line, there were 4 more rare codons that were unaccounted for in the Q5a construct that are not found in the Q5b construct (Figure 3.14). These 4 codons are found in the DNA segment that codes for the first 43 residues of Coq5. In addition, two of the 4 codons are found to be placed adjacent to each other. During mRNA translation, elongation pauses can be caused by mRNA secondary structures or runs of rare codons, which can eventually lead to ribosomal abandonment of translation [50]. As a result, using multiple rounds of site-directed mutagenesis, both of these codons were mutated into more common codons readily found in *E. coli*.

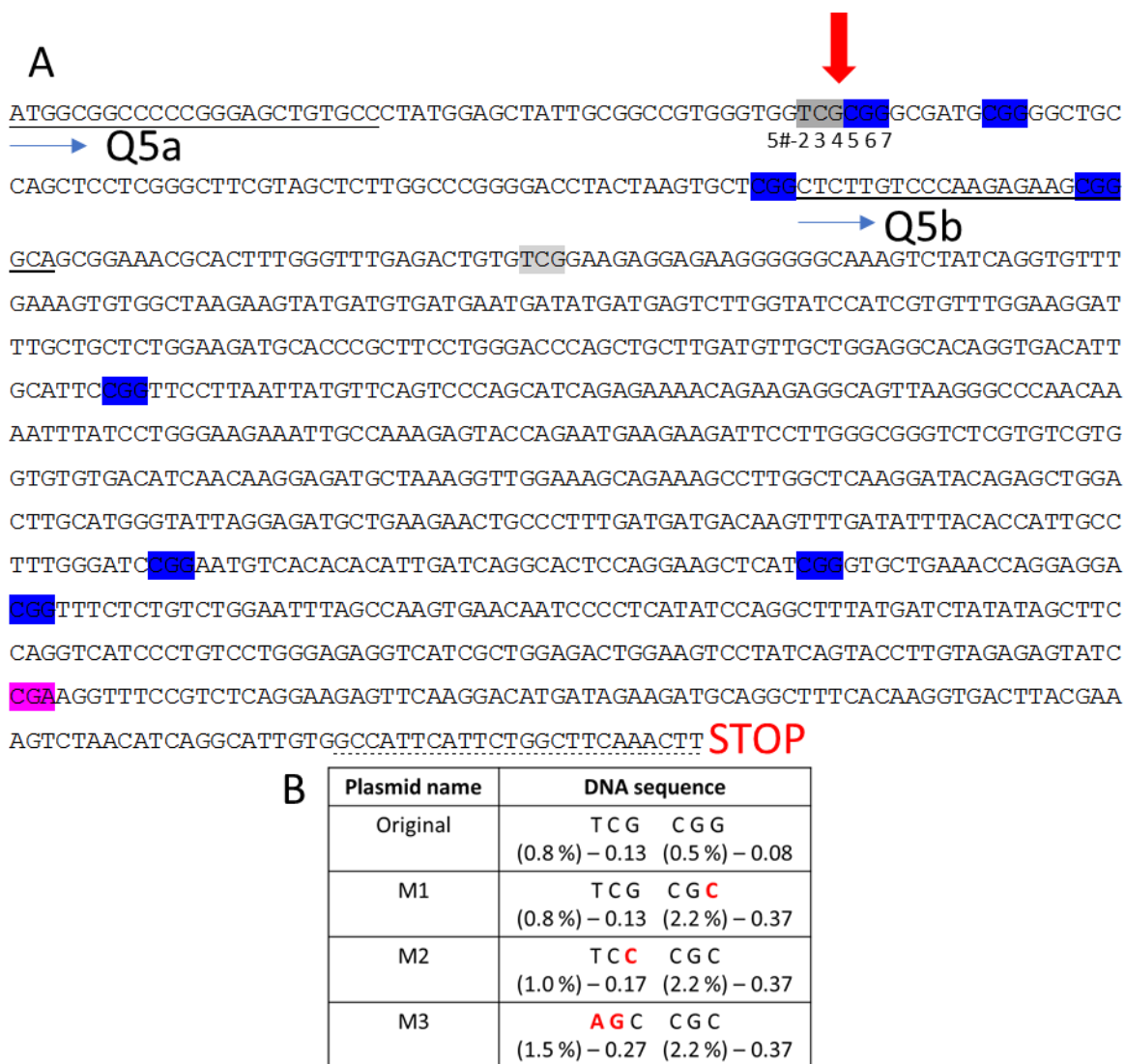


Figure 3.14 DNA sequence map of Coq5 and mutagenesis work done on Q5a plasmid.

(A) The single underline and bold underline show the segment where the forward primer for Q5a and Q5b construct would bind. The dotted underline shows the segment where the reverse primer for both Q5a and Q5b would bind. The rare codons for *E. coli* cells that are not accounted for by the BL21-CodonPlus (DE3)-RIPL *E. coli* cell line are highlighted with gray, pink, and blue colours. The red arrow indicates the predicted area in which the ribosome cannot continue mRNA translation. The base pair number of the 2 adjacent rare codons are indicated by number underneath (bases 52-57). (B) Multiple rounds of site-directed mutagenesis were performed to mutate the bases (52-57) highlighted in red. The % shown under each codon represent the average frequency this codon is used per 100 codons in *E. coli* and the following number represents the abundance of that codon relative to all of the codons for that particular amino acid [51].

Thus far, the sequencing results verified that the first two rounds of mutagenesis were successful. As a result, both M1 and M2 plasmids were transformed into BL21-CodonPlus (DE3)-RIPL *E. coli* cells to be used for expression trials; however, SDS-PAGE analysis of the samples collected from expression and purification process showed no improvement in the expression of the Q5a construct. Although the codons were mutated to those with higher % use and abundance, the expression may have not occurred due to the relatively low % use and abundance of the first codon. In order to overcome this obstacle, the third mutation should be pursued. Expression of the full length of Coq5 is important as it was determined that for Coq7, the possible phosphorylation sites are located in the N-terminal disordered region. As a result, in order to fully characterize the kinase activity of ADCK3, the full length of Coq5 should be expressed and purified in order to be used as the substrate.

Expression and purification of Q5b – The results showed overexpression and successful purification of Q5b via IMAC (Figure 3.15). The predominant band in the elution lane of SDS-PAGE shows an approximate molecular weight of 75 kDa which matches the expected weight of 75.5 kDa. The initial IMAC purification profile SDS-PAGE gel showed significant amount of Q5b left in the pellet. As a result, the sonication time was increased from 2 minutes to 4 minutes to further lyse the cells to release more Q5b into the lysate. In addition, although the His-tagged Q5b was captured by the nickel, the SDS-PAGE gel showed contaminating proteins being co-purified along with Q5b. These contaminants may possess histidine-rich areas that helped them to bind to the nickel resin or they may have simply bound to Q5b. In order to further purify Q5b, size-exclusion chromatography step was added to the purification procedure.

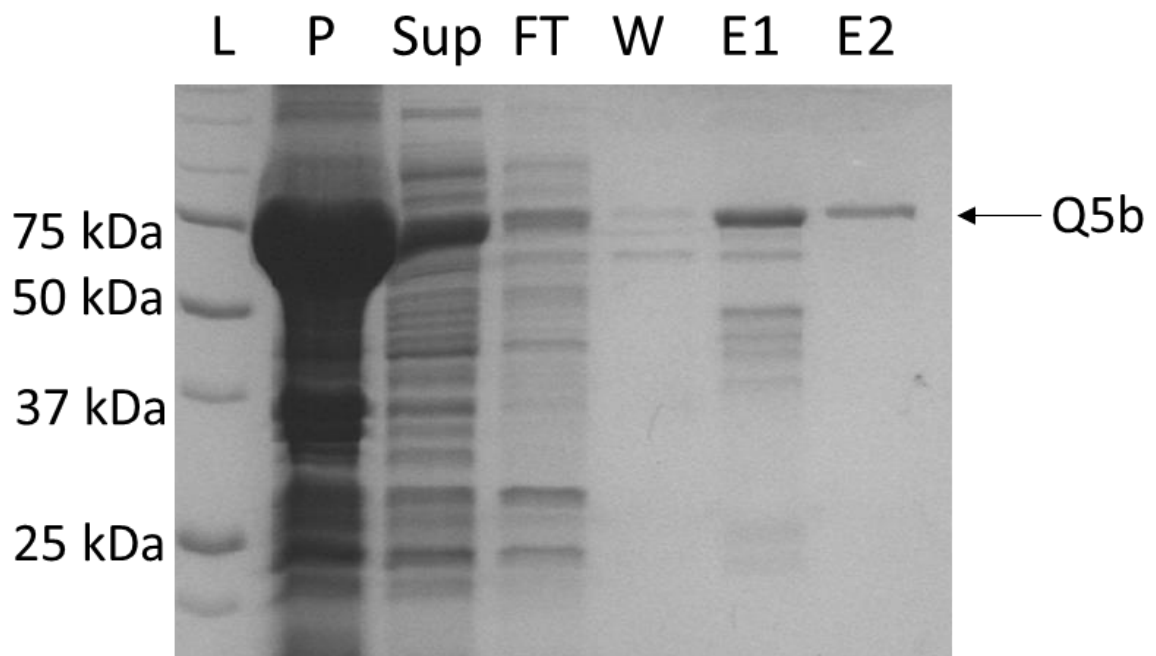


Figure 3.15 representative profile of Q5b purification using histidine-mediated immobilized metal affinity chromatography.

Using the histidine-tag engineered into the construct, Q5b was extracted from the supernatant of the cell lysate using Ni^{+2} -NTA chromatography. Throughout the purification process, 10 μL of sample were collected and analysed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining. From left to right, the lanes represent: (L) ladder; (P) pellet; (Sup) supernatant; (FT) flow-through; (W) wash; (E1) Elution 1; and (E2) Elution 2. The elution fractions were collected for further purification.

The results from the size-exclusion chromatography step showed the elution of Q5b shortly after the void volume around 40 mL (Figure 3.16A). This elution volume corresponds to the elution volume of proteins with molecular weight higher than 440 kDa, as indicated by the manufacturer's guidelines. An explanation for this might be that the buffer condition for Q5b is not optimal which leads to protein oligomerization. It is also important to note that the yeast Coq5 has been shown to dimerize in solution [28]. As previously mentioned, the yeast Coq5 and human Coq5 have high sequence homology. As a result, human Coq5 may follow the trends of the yeast Coq5 and dimerize in solution. However, even though MBP of the Q5b construct may increase the overall molecular weight, dimer of Q5b would result in the molecular weight of 151 kDa. As a result, in addition to the dimerization, there may be other factors causing Q5b to oligomerize in solution. However, the possibility of the MBP tags interacting with each other to cause the oligomerization can be eliminated. This is because the purification of MBP-tagged ADCK3 elutes at its appropriate elution volume when subjected to size-exclusion chromatography. Alternatively, it is still possible that Coq5 of one Q5b interacts with the MBP of another Q5b to cause oligomerization.

Furthermore, the SDS-PAGE gel showing the protein content of the proteins eluted at 40 mL from size-exclusion chromatography showed incomplete purification of Q5b (Figure 3.16A). The contaminating proteins shown in the gel may be the factor which caused Q5b to further oligomerize in solution as they elute in the same peak; however, at the same time, that may not be the case as the earlier and later fractions of the peak only showed Q5b and no contaminants. Thus, it may be that the contaminants are oligomerizing among themselves that caused their elution volume to be the same as the oligomerized Q5b.

Prior to optimization of the purification procedure of Q5b, the resulting protein was used for further experiment including ATPase assays (Chapter 4). However, the contaminating proteins observed in the SDS-PAGE proved to be significant as they hindered protein

quantification as well as the ATPase activity of the Q5b sample (Figure 3.16B). Using mass spectrometry, contaminant 1 and contaminant 2 were identified as *E. coli* proteins, chaperonin GroL and outer membrane protein F, respectively. Upon further investigation, it was found that GroL has ATPase activity which was the protein hindering our ATPase assays. As a result, ion-exchange chromatography step was added to further purify Q5b from these contaminants.

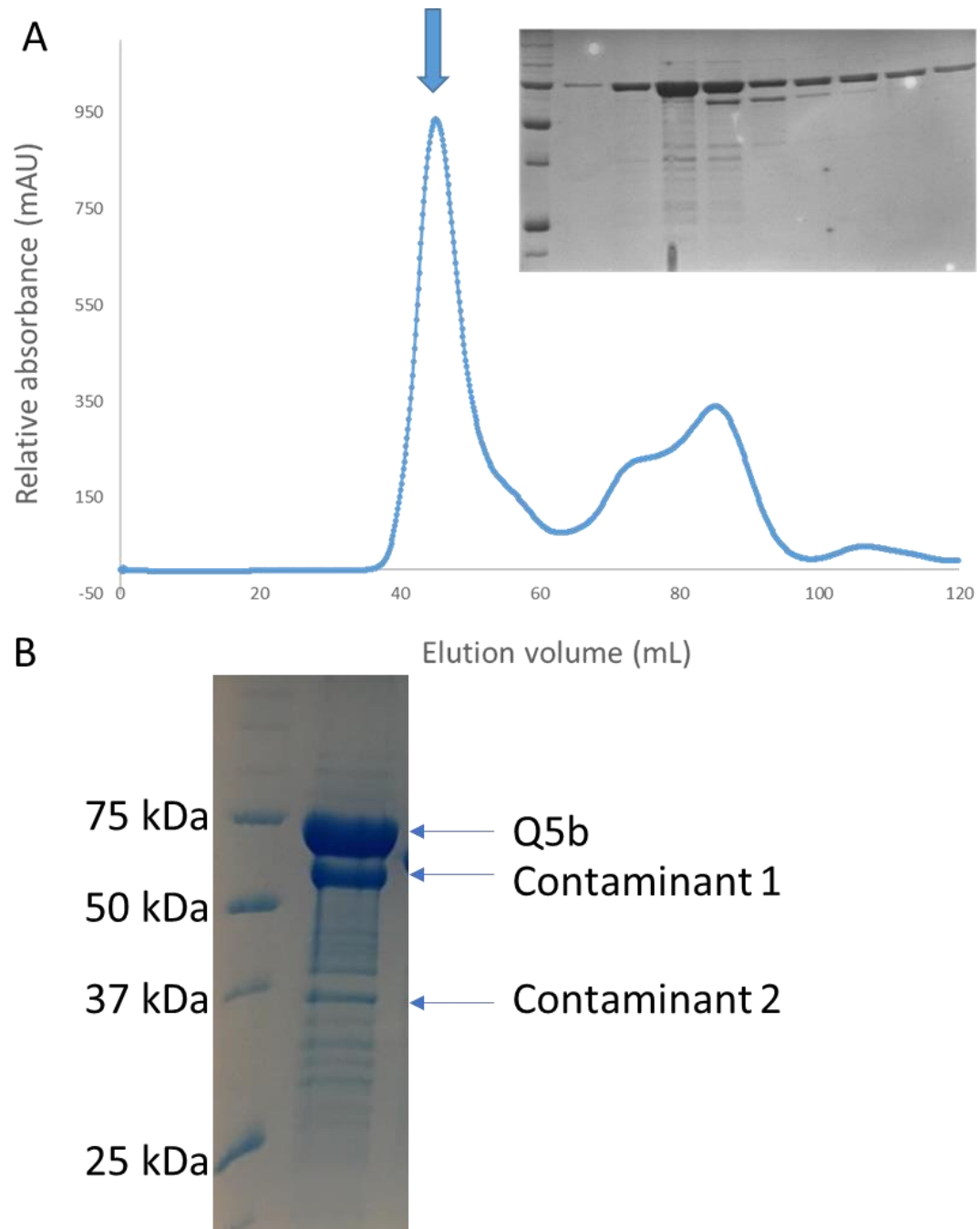


Figure 3.16 Analysis of Q5b following purification using size-exclusion chromatography. Elution fractions from Ni²⁺-NTA purification were collected and further purified via size-exclusion chromatography over a HiLoad 16/60 Superdex 200 prep-grade column using FPLC. (A) The peak containing Q5b is indicated by a blue arrow. The fractions of that peak were analyzed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining. (B) A sample of the pooled fractions were then analyzed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining.

The results from ion-exchange chromatography step in the purification of Q5b showed successful separation of Q5b and contaminant 1, GroL (Figure 3.17). The ion-exchange chromatogram shows elution of GroL at the elution volume of 27 mL which corresponds to 270 mM NaCl. Also, it shows the elution of Q5b at the elution volumes 30 – 36 mL which corresponds to 300 – 360 mM NaCl content. However, the fractions containing Q5b remains to have traces of contaminants and will need further purification for future experiments.

The best purification of Q5b yielded 8.4 mg of protein from 1 L of cells.

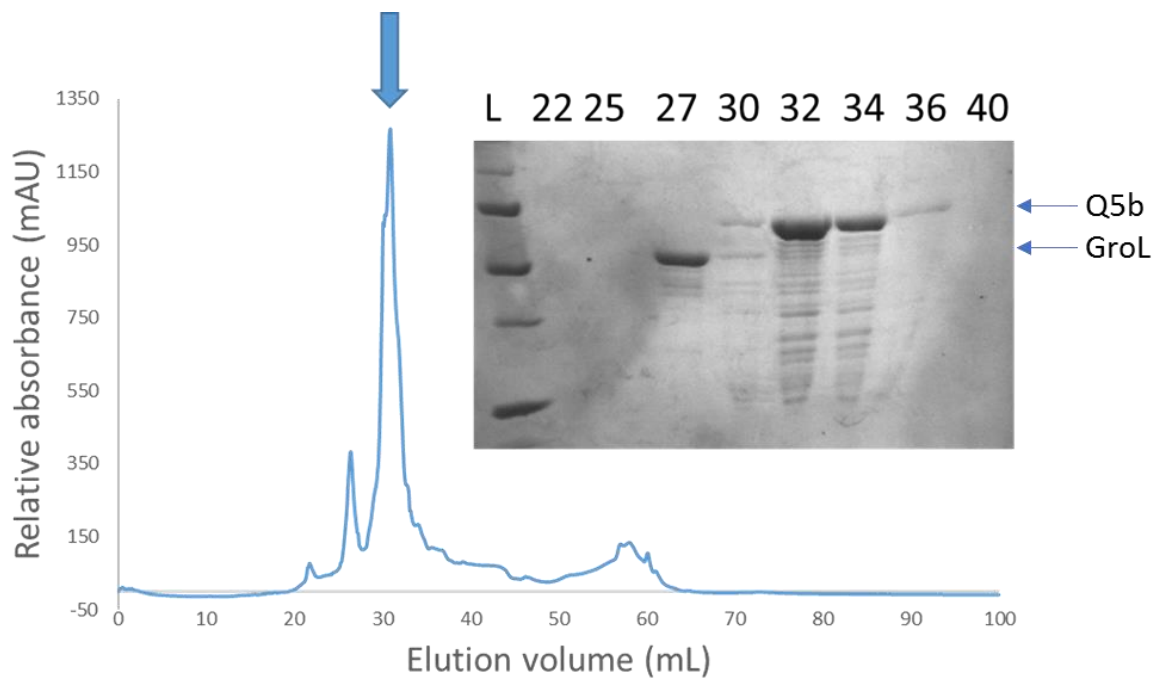


Figure 3.17 Analysis of Q5b following purification using ion-exchange chromatography.

Fractions containing Q5b from size-exclusion chromatography were collected and further purified via ion-exchange chromatography through a Resource™ Q 6 mL column using FPLC. The peak containing Q5b is indicated by a blue arrow. The fractions of two peaks shown in the chromatogram were analyzed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining. The lanes of the gel are labelled with the fraction number which also corresponds to the elution volume. Fractions 32-36 were pooled for further experiments.

Expression and purification of Q7a – The results showed the overexpression of Q7a and the successful purification of Q7a using IMAC (Figure 3.18). The predominant bands in the elution lanes of SDS-PAGE were proteins with approximate molecular weights of 70 kDa and 45 kDa. The molecular weights match the expected molecular weights of Q7a and MBP which are 67.3 kDa and 42.5 kDa. The presence of MBP in the elution may be due to the unspecific cleavage of Coq7 (1-217) off from the Q7a construct. This cleavage may be due to the flexible disordered region present in the N-terminus of Coq7. The abundance of protein around 25 kDa in the pellet, supernatant, and wash lanes match the expected molecular weight of Coq7 (1-217) which is 24.3 kDa.

In addition, the initial IMAC purification profile SDS-PAGE gel showed significant amount of Q7a left in the pellet. As a result, the sonication time was increased from 2 minutes to 4 minutes to further lyse the cells to release more Q7a into the lysate. In addition, in order to separate Q7a from the residual MBP, size-exclusion chromatography step was added to the purification procedure of Q7a.

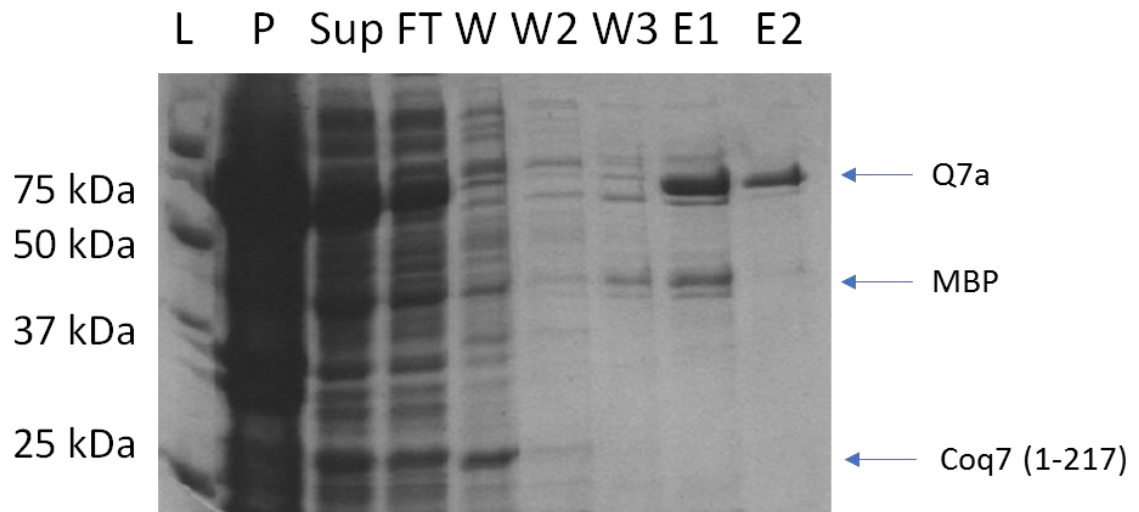


Figure 3.18 A representative profile of Q7a purification using histidine-mediated immobilized metal affinity chromatography.

Using the histidine-tag engineered into the construct, Q7a was extracted from the supernatant of the cell lysate using Ni^{+2} -NTA chromatography. Throughout the purification process, 10 μL of sample were collected and analysed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining. From left to right, the lanes represent: (L) ladder; (P) pellet; (Sup) supernatant; (FT) flow-through; (W) wash; (W2) wash2, (E1) Elution 1; and (E2) Elution 2. The elution fractions were collected for further purification.

The results from size-exclusion chromatography step showed elution of Q7a shortly after the void volume, around 40 mL (Figure 3.19). Similar to the size-exclusion chromatography results of Q5b, elution of Q7a at this volume suggests that Q7a is oligomerized in solution. Similar to Q5b, dimerized Q7a results in molecular weight of 134.6 kDa and does not match the molecular weight of proteins that elute shortly after the void volume. Unlike Q5b, SDS-PAGE analysis of the eluted protein peak containing Q7a does not show a significant number of contaminants. This may suggest that in the current purification buffers, Q7a oligomerizes and may require further buffer optimization. On the other hand, SDS-PAGE gel showed successful separation of Q7a from MBP. Thus, fractions containing Q7a were pooled for further experiments.

The best purification of Q7a yielded 3.2 mg of protein from 1 L of cells.

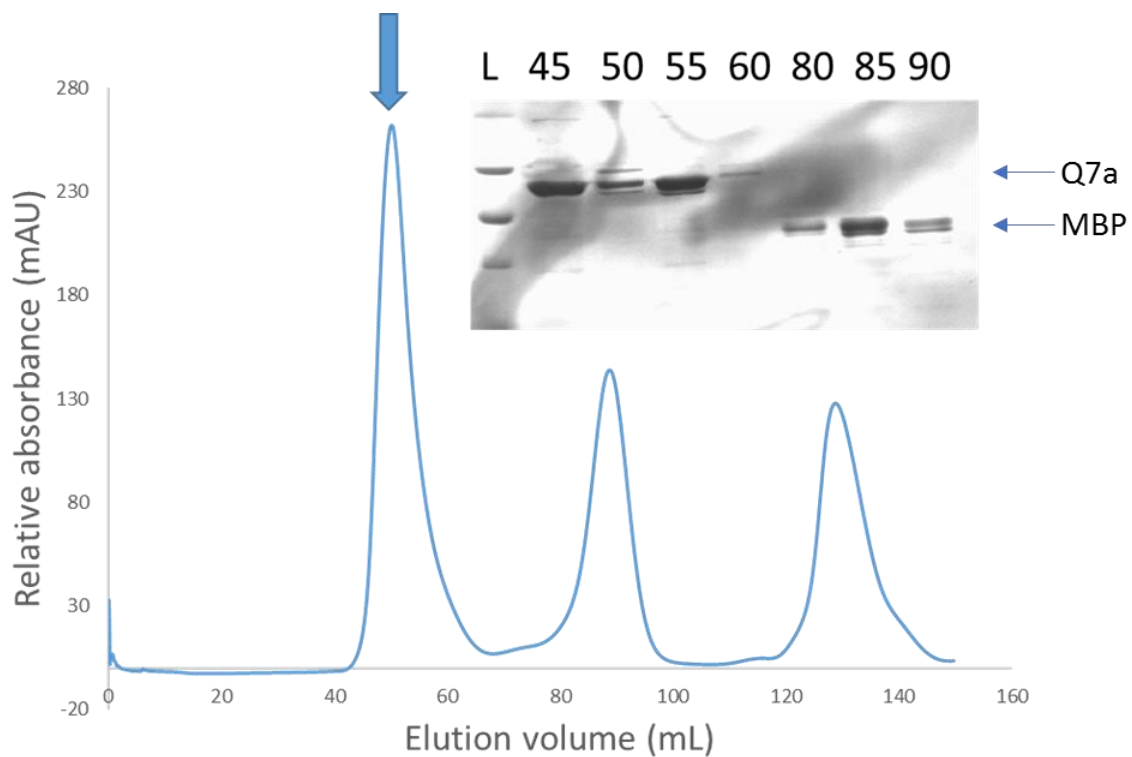


Figure 3.19 Analysis of Q7a following purification using size-exclusion chromatography. Elution fractions from Ni²⁺-NTA purification were collected and further purified via size-exclusion chromatography over a HiLoad 16/60 Superdex 200 prep-grade column using FPLC. The peak containing Q7a is indicated by a blue arrow. The fractions of that peak were analyzed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining.

Expression and purification of Q7b – The results showed the over-expression of Q7a and the successful purification of Q7b using IMAC (Figure 3.20). The predominant band in the elution lane of SDS-PAGE was a protein with the approximate molecular weight of 60 kDa which matches the expected molecular weight Q7b, which is 63.2 kDa. Unlike Q7a, SDS-PAGE gel of the Q7b purification did not show significant amount of MBP or Coq7 (37-217). This may further confirm our hypothesis that Q7a exhibits unspecific cleavage at the flexible disordered segment located on the N-terminus of Coq7 (1-217) as Q7b does not have this segment.

In addition, the initial IMAC purification profile SDS-PAGE gel showed significant amount of Q7a left in the pellet. As a result, the sonication time was increased from 2 minutes to 4 minutes to further lyse the cells to release more Q7b into the lysate. In order to remove the residual contaminants, size-exclusion chromatography step was added to the purification method of Q7a.

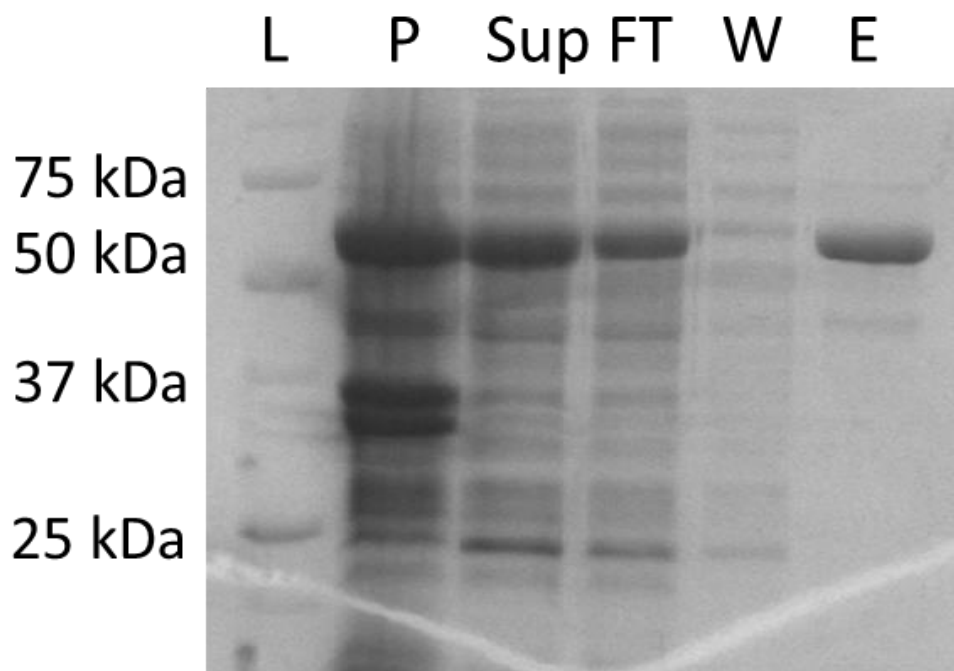


Figure 3.20 A representative profile of Q7b purification using histidine-mediated immobilized metal affinity chromatography.

Using the histidine-tag engineered into the construct, Q7b was extracted from the supernatant of the cell lysate using Ni^{+2} -NTA chromatography. Throughout the purification process, 10 μL of sample were collected and analysed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining. From left to right, the lanes represent: (L) ladder; (P) pellet; (Sup) supernatant; (FT) flow-through; (W) wash; and (E) Elution. The elution fraction was collected for further purification.

The results from the size-exclusion chromatography step showed the elution of Q7b shortly after the void volume around 40 mL (Figure 3.21). Unlike Q7a, SDS-PAGE gel of the fractions from the peak containing Q7b shows 2 contaminants. Using mass spectrometry, contaminant 1 and contaminant 2 were identified to be Q7b and an *E. coli* chaperone protein HtpG. From further investigation, contaminant 2, HtpG, was found to have ATPase activity. In order to remove this protein from the Q7b sample, further purification steps should be explored.

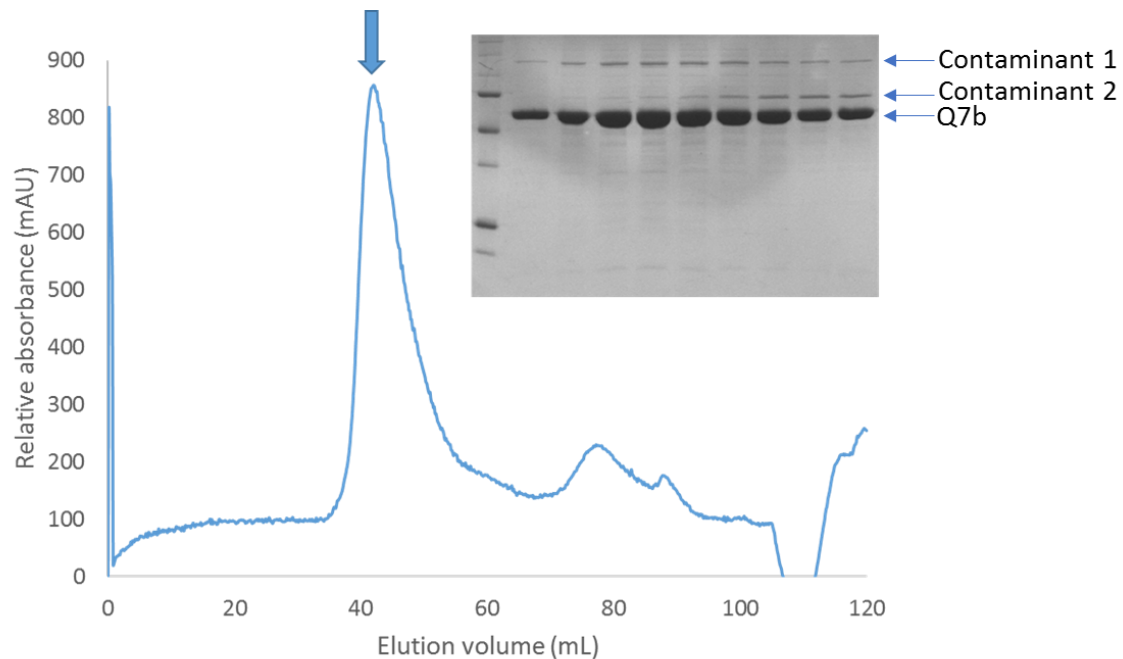


Figure 3.21 Analysis of Q7b following purification using size-exclusion chromatography. Elution fractions from Ni⁺²-NTA purification were collected and further purified via size-exclusion chromatography over a HiLoad 16/60 Superdex 200 prep-grade column using FPLC. The peak containing Q7b is indicated by a blue arrow. The fractions of that peak were analyzed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining.

Another interesting fact from the mass spectrometry is the presence of Q7b dimers. Adding more reducing reagent to the protein loading dye resulted in disappearance of the Q7b dimer band in the subsequent SDS-PAGE gels. However, during buffer optimization, when various reducing reagents including 10 mM β -mercaptoethanol, 2 mM Dithiothreitol (DTT), and 0.5 mM tris(2-carboxyethyl) phosphine (TCEP) were added to the purification buffers, the size-exclusion chromatogram remained to show elution of Q7b shortly after the void volume (Figure 3.22). As a result, although there is a cysteine residue on the outer surface of Coq7 that may form a disulfide bond with a neighbouring cysteine residue of another Coq7 molecule, there may be other factors contributing to the oligomerization of Q7b.

The best purification of Q7b yielded 46.8 mg of protein from 1 L of cells.

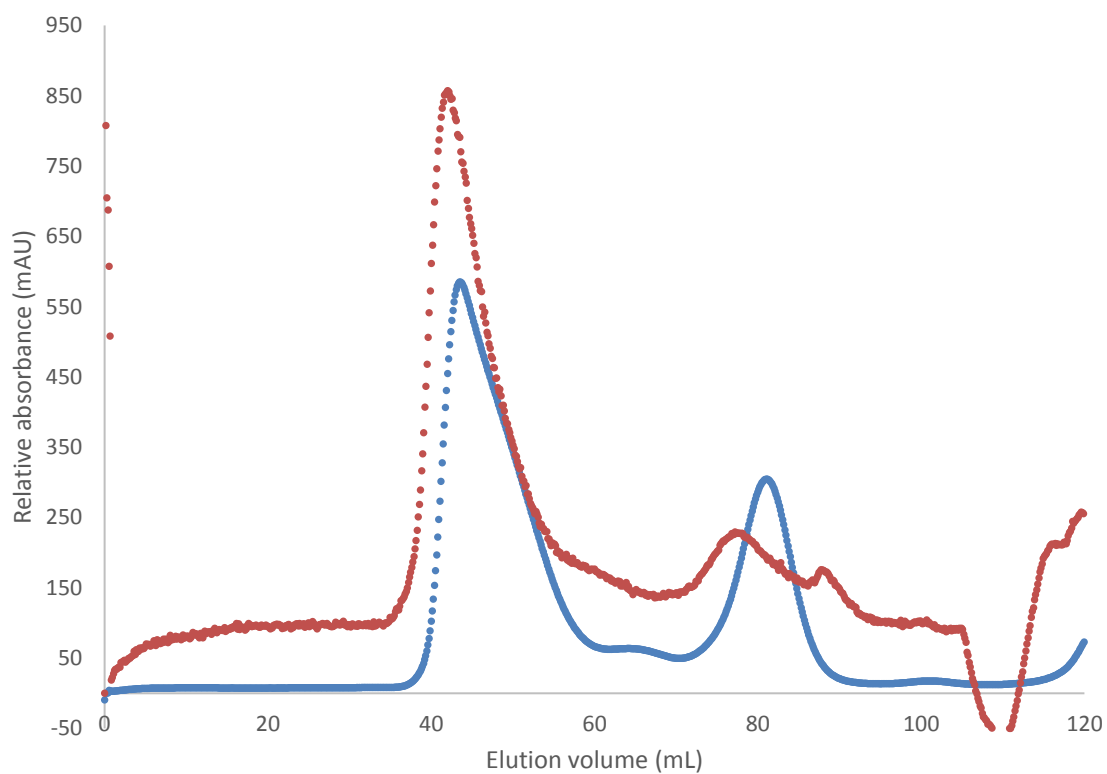


Figure 3.22 Overlay of the size-exclusion chromatograms of Q7b purified with and without reducing reagents.

Q7b purified with reducing reagent, TCEP, is shown in blue and Q7b purified without any reducing reagent is shown in red.

Cleaving the MBP-tag off from the Q protein constructs - When the ATPase activity assays and kinase activity assays were performed using Coq proteins as the substrate for ADCK3, it was suggested that perhaps the MBP tag may have caused structural hindrance to inhibit the ADCK3 activity. This was suggested as MBP is bigger in size compared to the Coq proteins fused to them. As a result, further experiments were done to cleave the MBP off and purify the resulting Coq proteins. When the MBP tag was cleaved from the Q5b construct, the resulting Coq5 (43-327) precipitated during the downstream purification steps. In order to purify Coq (43-327) following TEV protease cleavage, the purification buffer conditions need to be optimized.

When the purified Q7b was subjected to TEV cleavage, initially Coq7 (37-217) was soluble; however, upon concentrating the fraction containing Coq7 (37-217), heavy precipitation formed. As a result, the purification buffer was further optimized and it was found that adding 0.1 % Triton X-100 helps to solubilize Coq7 (37-217). Thus, Coq7 (37-217) was further purified with size-exclusion chromatography and was shown to elute from the Superdex 200 column around 65 mL (Figure 3.23). According to the manufacturer's guidelines, proteins that elute from this column around 65 mL have the molecular weight of 158 kDa. The predicted molecular weight of Coq7 (37-217) is 20.5 kDa and thus, the results suggest that Coq7 (37-217) oligomerizes into a structure consisting of approximately 8 monomers. In addition, the SDS-PAGE analysis of the peak observed in size-exclusion chromatography showed a single protein band in every lane indicating the high purity of the sample; however, according to the ladder, the molecular weight of the proteins of that band is smaller than 20 kDa. The expected molecular weight of Coq7 (37-217) is approximately 20.5 kDa.

The best purification of the cleaved Q7b yielded 11.3 mg of protein from 1 L of cells.

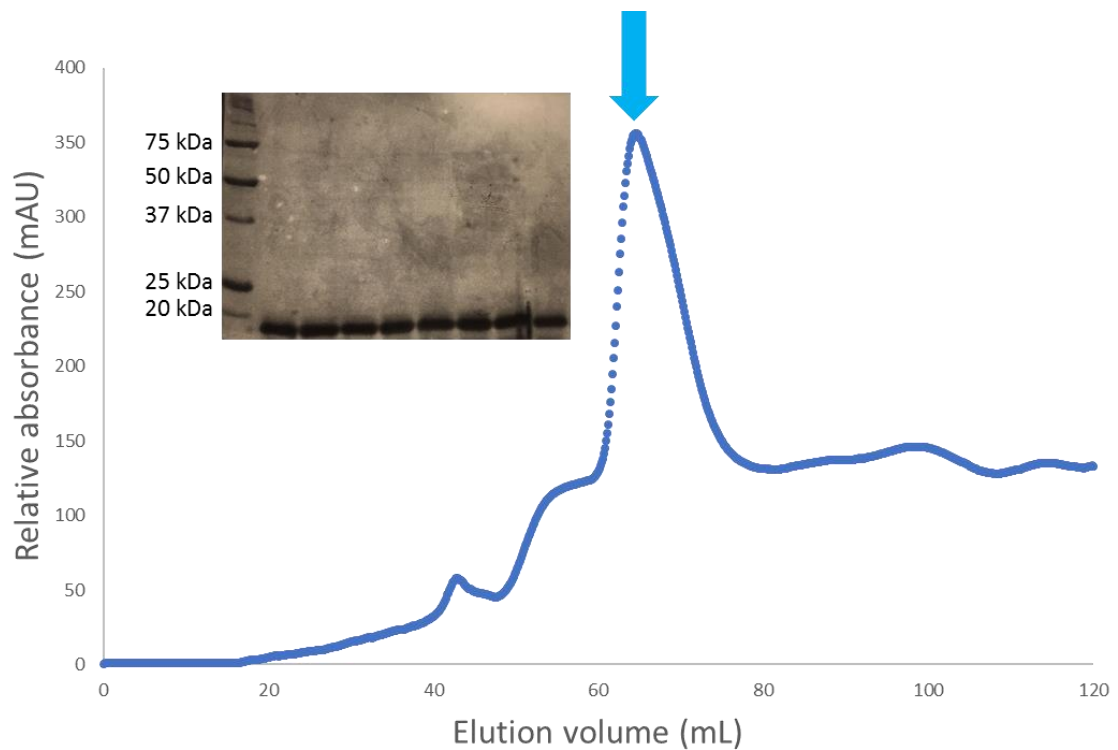


Figure 3.23 Analysis of Coq7 (37-217) following purification using size-exclusion chromatography.

After incubating purified Q7b with TEV protease overnight, the resulting sample was run on Ni²⁺-NTA column and the flow-through containing Coq7 (37-217) was collected. The fractions containing Coq7 (37-217) were pooled and purified via size-exclusion chromatography over a HiLoad 16/60 Superdex 200 prep-grade column using FPLC. The peak containing Coq7 (37-217) is indicated by a blue arrow. The fractions of that peak were analyzed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining.

The oligomerization of Coq7 may be explained by the oligomerization observed with a protein called *Bacteroides fragilis* DSP-like protein, BfDSPL because the structure of BfDSPL was used by PHYRE2 as the model to predict the structure of Coq7. In a study done by Gauss *et al.*, BfDPSL was shown to dimerize in solution and at higher concentrations, BfDPSL was observed to form spherically shaped particles with diameter of approximately 95Å [52]. With their observation of these spheres in addition to the crystal structure of BfDPSL, it was suggested that BfDPSL forms a dodecamer that creates a negatively charged central cavity.

Although the only similarity between BfDPSL and Coq7 is the overall structure, the important observation from BfDPSL that can be applied to Coq7 is that these bundles of helices can combine to create a quaternary structure to increase their stability. For example, like BfDPSL, Coq7 can form oligomers to bury its hydrophobic side of the bundle of helices (Figure 3.24). Perhaps the combination of eight Coq7 forms the most stable quaternary structure and that may be what is observed in our size-exclusion analysis. However, this form of quaternary structure may not exist in its native environment. This is so because, as previously mentioned, Coq7 is a membrane associating protein. Therefore, the hydrophobic surface of the protein needs to be free from interacting with any proteins and associate directly with the membrane.

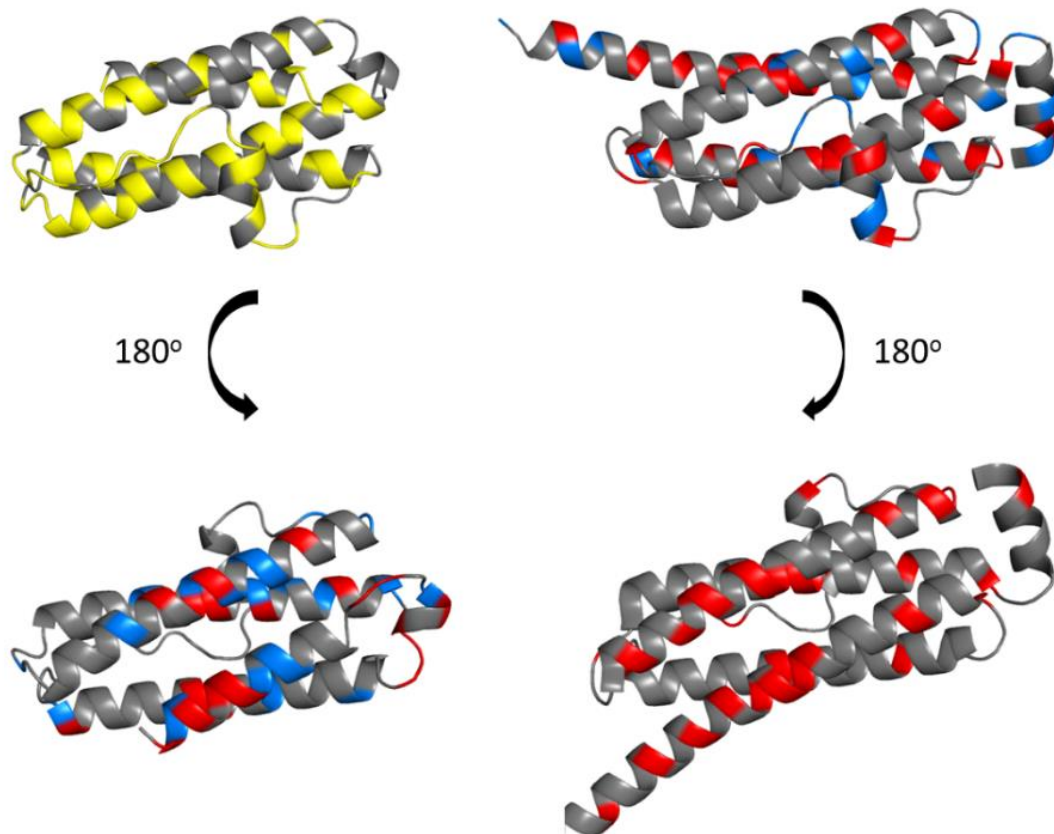


Figure 3.24 Surface charge analysis of Coq7 and BfDSPL.

The structure on the left shows the structure of Coq7 as predicted by PHYRE2 and the structure on the right shows the structure of BfDSPL (PDB:2VZB). PHYRE2 used the structure of BfDSPL as a model when predicted the structure of Coq7. For both proteins, opposite sides of their structure are shown and their surface charges are highlighted with color. Non-polar, aliphatic residues are shown in yellow, positively charged residues are shown in blue, and negatively charged residues are shown in red. For the purpose of highlighting the surface charges, only applicable charges were shown in colour. For both proteins, one side is shown to have a well-balanced ratio of negative and positive charges that would be exposed to the solvents. In addition, Coq7 is shown to have a hydrophobic surface and BfDSPL is shown to have a negatively charged surface.

The small discrepancy observed between the expected molecular weight of Coq7 (37-217) and the observed molecular weight of our protein sample on the SDS-PAGE may also be due to the high hydrophobic amino acid content of Coq7. Out of the 217 residues, 52% are hydrophobic residues. From a study done by Rath *et al.*, it was observed that membrane proteins migrate faster on SDS-PAGE gels compared to non-membrane associating proteins of the same size [53]. From their results, it was suggested that, although 1 SDS molecule typically binds to every 2 amino acids of a protein, SDS may bind to proteins at a higher ratio if the protein has a high hydrophobic amino acid content. To explain, SDS is a detergent molecule that has a hydrocarbon chain attached to a charged head group. In SDS-PAGE, SDS is used to saturate the protein so that the hydrocarbon tail would interact with the protein while the negatively charged sulfate group is exposed to the solvent allowing for the migration of protein to the cathode. If there are significantly more hydrophobic residues available to interact with the hydrocarbon tail of SDS molecules, then more SDS will be able to bind to the protein increasing the degree of the negative charge compared to proteins with lower hydrophobic residue content. Thus, even with the same molecular weight, those proteins with a more overall negative charge would migrate faster toward the cathode.

3.7 The protein folding of Q5b and Q7a

In order to confirm that our proteins were folded, circular dichroism spectroscopy was performed with Q5b and Q7a (Figure 3.25). The results showed protein folding; unfortunately, our buffer contained 300 mM NaCl which led to limited data collection in the lower spectrum wavelength. As a result, we were unable to determine the secondary structure composition; however, the overall structure composition could be compared between Q5b and Q7a. For example, Q5b has a higher ellipticity at 195 nm compared to Q7a. Ellipticity observed at this region results from combination of antiparallel β -sheets and α -helices. With regards to the predicted structures of Coq5 and Coq7, it makes sense for the data to suggest that Q5b has a higher β -sheet and α -helix composition.

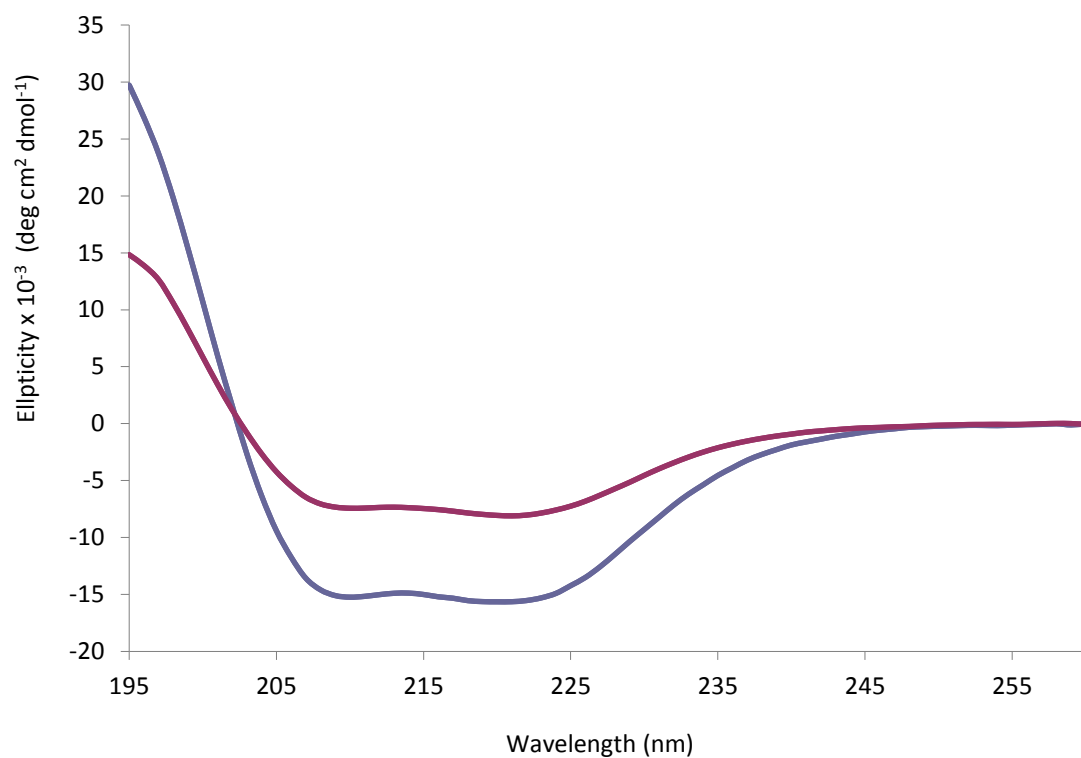


Figure 3.25 Circular dichroism spectra for Q5b and Q7a.

The blue spectrum corresponds to Q5b and the purple spectrum corresponds to Q7a. All spectra are averages of six scans.

3.8 Methyltransferase activity of Coq5

In order to assess the methyltransferase activity of Coq5, Q5b construct consisting of N-terminal MBP tag fusion protein was used. In addition to the protein being assayed for, its substrate is required to assess its activity. Unfortunately, the actual substrate of Coq5 demethyl-demethoxy-hydroquinone (DDMQ₆H₂) is not commercially available, most likely due to its instability. As a result, analogs of DDMQ₆H₂ were sought and 2-methoxyhydroquinone was chosen as our substrate (Figure 3.26).

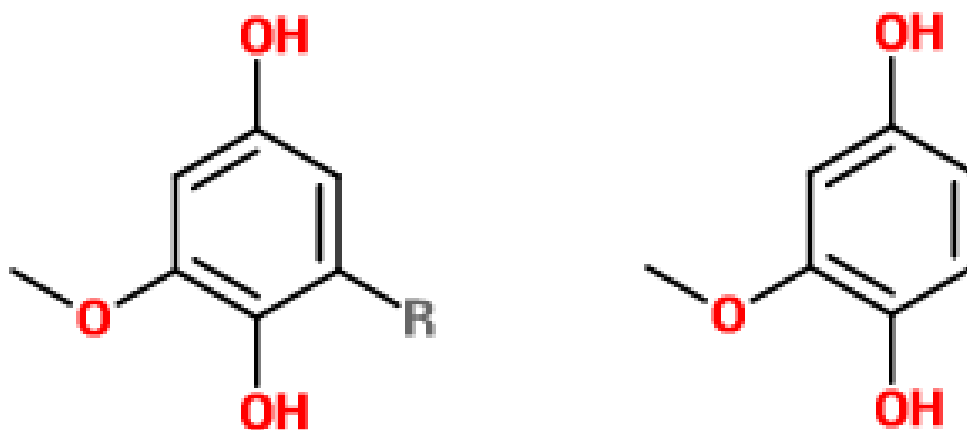


Figure 3.26 The molecular structure of Coq5 substrate and its analog.

The structure on the left is the putative substrate of Coq5, dimethyl-demethoxy-hydroquinone, and the structure on the right is the analog, 2-methoxyhydroquinone, that was used as the methyl group donor for Q5b. The R group on the structure shown on the left is the polyisoprene tail.

Initially when the colorimetric kit was used, it was difficult to differentiate the background noise signal from the actual signal from the methyltransferase activity itself. The high background noise resulted from auto-oxidation of the substrate, 2-methoxyhydroquinone, that starts immediately after it is solubilized into solution. When 2-methoxyhydroquinone auto-oxidized, the solution changed its colour to a dark brown, heavily affecting the absorbance reading. The observed colour change was confirmed to be an oxidizing reaction when the substrate was solubilized in a solution containing a reducing reagent, β -mercaptoethanol (β -ME). The substrate solution containing β -ME did not result in a colour change. However, the user's

manual for the methyltransferase assay kit noted that any addition of reducing reagents will inhibit the downstream reactions that is responsible for the colorimetric signal. As a result, an alternate method of inhibiting the auto-oxidation reaction from occurring was sought.

In order to inhibit the auto-oxidation reaction of 2-methoxyhydroquinone, it was solubilized in water that had been de-gassed in a chamber saturated with Argon. Subsequently, when methyltransferase reaction was performed under Argon, no signal was observed. Initially, it was thought that our Coq5 construct was inactive; however, upon further investigation, it was found that oxygen is required by the master mix of the assay kit to induce the colour change in the downstream reaction. Thus, the fluorometric assay kit was recommended by the manufacturer to overcome this obstacle.

When the methyltransferase assay was performed with the fluorometric assay kit instead, the results showed that auto-oxidation reaction was still detected in the new wavelength range; however, with the fluorometric assay kit, the signal resulting from the reaction could be detected even with the background noise cause by the auto-oxidation reaction. As a result, the reaction conditions were optimized to reduce the background noise as much as possible. Then, in order to confirm that our Coq5 sample was active and that the product formation is due to Q5b, the assay was performed with varying concentrations of Coq5. This would confirm that the increase in the observed product formation rate was the result of increase in our enzyme concentration.

The results showed logarithmic reaction curves where the rate at which the product forms decrease as it reaches the maximum concentration of the product (Figure 3.27). The product concentration increased until a maximum concentration of approximately 4.5 μM was reached. Even for the negative control reaction that contains no enzyme, there was a logarithmic increase in the signal, most likely due to the auto-oxidation of the substrate. For the reactions that did contain Coq5 enzymes, the results showed that reactions with higher enzyme concentrations reached the plateau faster. As well, for the most part, the results showed that at time zero, the

reaction had already proceeded creating a difference in the concentration of the product formed. However, the results also showed overlap between the reactions containing 1, 2, and 3 μM Q5b. Although the experiment was done in triplicates, there may have been pipetting errors as there are many components of the experiment that is time sensitive. Fortunately, by comparing the results of reaction containing 0 μM enzyme and 4 or 5 μM Q5b, it can be confirmed that Q5b was able to use SAM as its source for methyl groups for its function.

Various attempts were made to optimize the reaction scheme to slow down the initial reaction so that enzyme kinetics of Coq5 could be determined. For one thing, the concentration of the enzyme and the concentration of the substrates were adjusted but resulted in poor fluorescence readings or showed higher activity in substrate only reaction. Another parameter that was changed were the pre-incubation temperature and reaction temperature. The enzyme, the plate, and the master mix were pre-incubated on ice, at room temperature, and at 37 °C to see if lower initial temperature would be able to slow down the initial reaction. Also, the temperature of the reaction in the fluorescence plate reader were also set to either 15 °C, 25 °C, 30 °C, and 37 °C to see if lower temperatures would be able to slow down the initial reaction. All combinations of these temperatures, except the preincubation and reaction temperatures at 37 °C, were not able to produce any significant results. Instead, some results showed higher methyltransferase activity in the substrate only reaction or no difference in activity between the substrate only and enzyme containing reactions. The fact that Coq5 requires 37 °C to be active may actually be significant as the physiological temperature in which Coq5 is naturally active is 37 °C. In addition, various parameters regarding fluorescence microscopy were also adjusted but failed to optimize the reaction curves.

In order to verify that Q5b has methyltransferase activity, a different method of assaying for methyltransferase activity can be used. For example, the method used by Baba *et al.* to show the methyltransferase activity of yeast Coq5 may be more accurate [38]. As previously mentioned

in the introduction chapter, Baba *et al.* used substrate analogs that retained a lipid tail. After incubating this analog with the crude yeast mitochondrial protein and S-adenosyl-L-[methyl-³H]-methionine, the product could be extracted from the reaction using its lipid-like characteristics. Then, the using HPLC, the molecules are separated and the amount of radioactive methyl group incorporated could be assayed. The only change to their method would be to use our Q5b proteins rather than the crude mitochondrial proteins. This method would allow for more accurate detection as the change in colour of the quinone substrate would not affect the end results.

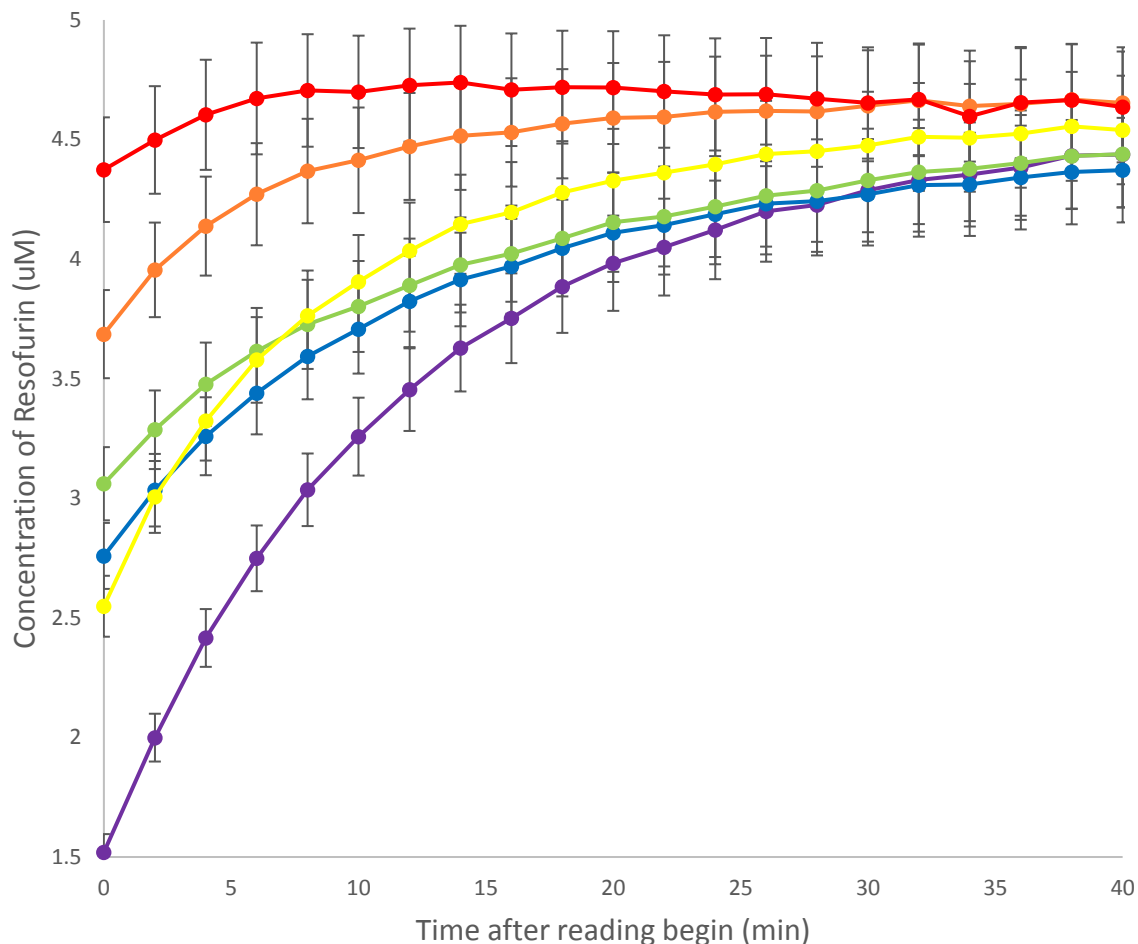


Figure 3.27 Methyltransferase activity of Q5b.

The N-terminal MBP fusion Coq5 (43-327) construct was used to assess its methyltransferase activity on its substrate analog 2-methoxyhydroquinone. The assay measured the fluorescence emitted at 590 nm by the product, resofurin, after exciting the reaction mixture at 535 nm. The first reading of the reactions were taken at time “0” but this does not reflect the time when the reaction started. Then, using the linear equation obtained from the standard curve, the concentrations of resofurin were determined and plotted against the reaction time. Various concentrations of Q5b were used and labelled with colours: (Red) 5 uM Q5b, (Orange) 4 uM Q5b, (Yellow) 3 uM Q5b, (Green) 2 uM Q5b, (Blue) 1 uM Q5b, and (Purple) 0 uM Q5b. The reactions were carried out in triplicates and their averages are shown. Error bars represent standard error of the mean.

3.9 Crystallization trials of Coq5 and Coq7

Immediately after setting up the screening trays, trays containing both Q5b and Q7b proceeded to show precipitate formation. As a result, another set of screening trays were set up with a decreased protein concentration from 10 mg/mL to 5 mg/mL. To date, no sign of crystals was observed for Q5b. Perhaps new set of screening trays should be set up with a different protein concentration. The lack of crystal hits for Q5b may be due to the MBP that remains to be fused to the construct. Aforementioned, the purification buffers must be optimized to be able to purify stable Coq5 (43-327) proteins for crystallization purposes. Although it is case-dependent, Coq5 (43-327) may have a higher chance of being able to form crystals compared to the MBP fusion construct. The high sequence identity between human Coq5 (43-327) and the yeast Coq5 that was crystallized suggests that the two structures will be highly homologous. Although the exact conditions that will induce the crystallization of the human Coq5 (43-327) cannot be predicted, it is likely that proteins like Coq5 (43-327) will be able to form crystals.

In terms of Q7b, some initial hits were observed in the screening trays (Figure 3.28); however, no signs of crystals appeared when these screening conditions were used to set up expansion trays. In order to make sure that these crystal hits are reproducible, screening trays using the same conditions should be set up once more. Also, perhaps new screening trays should be set up with the protein concentration 7.5 mg/mL as there are not many crystals appearing in the wells at once. In addition, screening trays should be set up using Coq7 (37-217), cleaved and purified from Q7b.

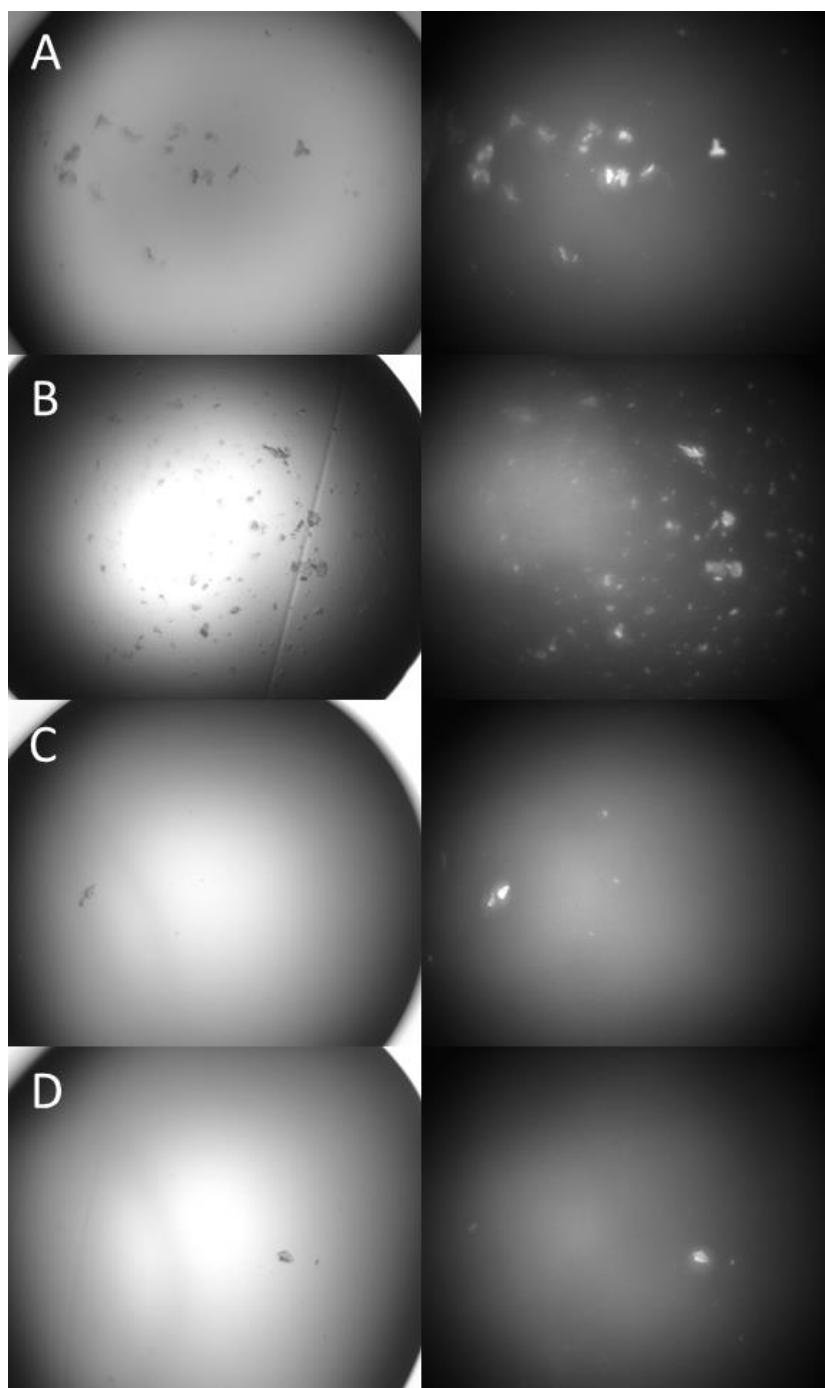


Figure 3.28 Examples of the initial hits for Q7b in screening trays.

0.4 μL of 5 mg/mL Q7b was mixed with 0.4 μL of the screening conditions on a sitting drop 96 well plate. The screening plates were incubated at room temperature or at 4 $^{\circ}\text{C}$. The images on the left pane show crystals in visible light and the images on the right shows their appearance under UV light. The screening conditions for the observed images are: (A) 0.1 M Tris-HCl pH 8.5 and 2.0 M ammonium phosphate; (B) 0.1 M HEPES pH 7.5, 0.05 M Cadmium sulfate, and 1.0 M Sodium acetate; (C) 2.4 M Sodium malonate pH 7.0; and (D) 0.1 M HEPES sodium salt pH 7.5, 0.2 M Magnesium chloride, and 30% (v/v) PEG 400.

3.10 ATPase activity of ADCK3

Previous studies have shown Mg^{2+} - dependent ATPase activity of ADCK3 which provided strong support for the prediction that ADCK3 is an active atypical kinase [48]; however, no functional studies have directly shown the interaction between ADCK3 and its substrate proteins. Since several Coq proteins are considered to be potential substrates of ADCK3 and may interact with it, it would be interesting to compare the ATPase activity of ADCK3 in the presence and absence of the Coq proteins. Thus, to observe for any changes in ADCK3 activity in the presence of its potential substrates, the ATPase activity of ADCK3 (245-647) was assessed.

Prior to assessing and comparing the ATPase activity of ADCK3 in the absence and presence of Coq proteins, our MBP-ADCK3 (245-647) construct was tested to determine whether it has ATPase activity. In previous studies, a shorter construct, MBP-ADCK3 (250-523) was used [48]. In addition, in order to determine the optimal concentration of our enzyme to be used for the experiment, we tested the activity of ADCK3 at various concentrations (Figure 3.29). The results showed increase in free phosphate concentration in the reaction with time. When linear line of best fit was placed with the y-intercept set to 0, the resulting equations for 1 μM and 2 μM enzyme reactions were $y = 0.0224x$ and $y = 0.0489x$, respectively. From these equations, the reaction rate could be determined to be 4.5×10^{-6} μ mole ATP hydrolysed/min and 9.8×10^{-5} μ mole of ATP hydrolysed/min for 1 μM and 2 μM enzyme containing reactions, respectively. These rate values follow our hypothesis since the rate of 2 μM reaction is approximately double the rate of 1 μM reaction. Since there was twice the amount of enzyme in the 2 μM reaction, the rate determined from that reaction should be double the rate of the 1 μM reaction. Overall, the specific ATPase activity of ADCK3 averaged between the two data was determined to be 0.0234 mole of ATP/min/mole of ADCK3. In addition, it was determined that approximately 10 μM of ADCK3 should be used to generate reliable absorbance measurements.

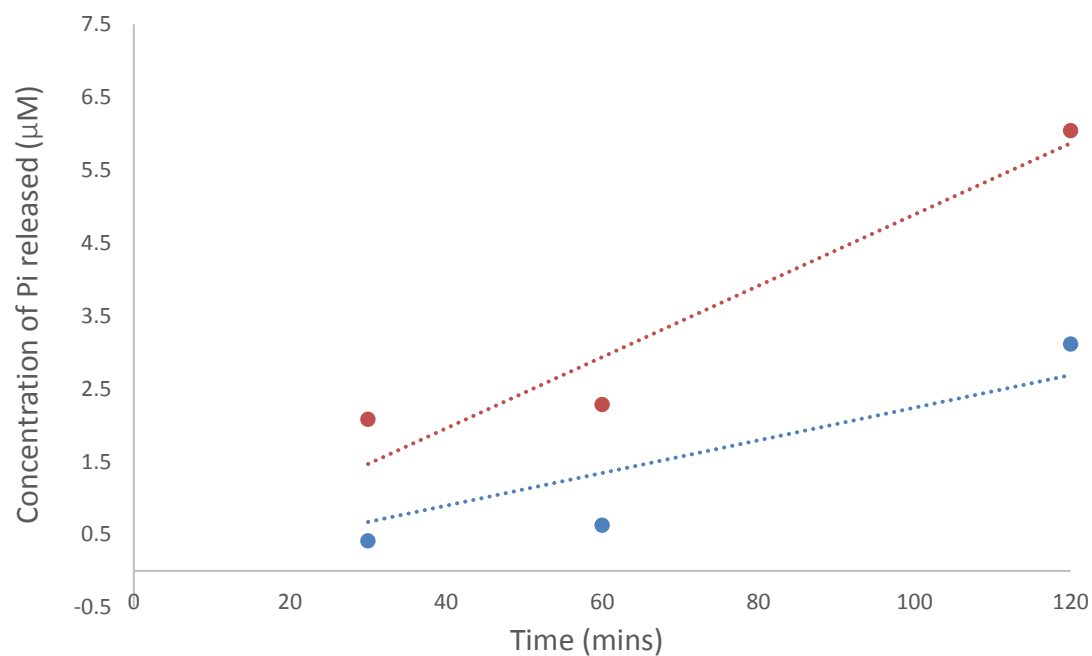


Figure 3.29 ATPase activity of MBP-ADCK3 (245-647).

The blue data points represent the reaction performed with 1 µM MBP-ADCK3 while the red data points represent the reaction performed with 2 µM MBP-ADCK3. Reactions were carried out in duplicates and their averages are shown. The linear trend lines were placed with the y-intercept set to 0. The linear equations and their R^2 values for reactions of 1 µM and 2µM enzymes are $y = 0.022x$, $R^2 = 0.92$ and $y = 0.049x$, $R^2 = 0.83$, respectively.

In order to assess the effect of the presence of Coq protein on the ATPase activity of ADCK3, ATPase activity assay was performed using Q5b, Q7a, and Q7b as the substrates. The results showed unexpected set of data (Figure 3.30). First, two control reactions containing just Q5b and Q7a resulted in higher ATPase activity compared to ADCK3; however, the reaction containing Q7b did not generate any significant data showing ATPase activity. The ATPase activity observed in both Q5b and Q7a are unexpected as Q5b is a methyltransferase protein and Q7a is a hydroxylase protein. Also, MBP is not known to generate any specific ATPase activity. The lack of activity observed with Q7b suggested that, perhaps if Coq7 were to have ATPase activity, it would require the full length of the protein; however, it is highly unlikely that Coq5 or Coq7 has ATPase activity.

When ADCK3 and Q7a were incubated together, the reaction rate decreased to a similar rate of ADCK3 alone. It is difficult to determine whether the remaining activity observed is the activity from Q7a or ADCK3. It was suggested that perhaps ADCK3 is transferring the γ -phosphate to these substrate proteins rather than releasing them into solution resulting in kinase activity rather than ATPase activity. Even so, it is still difficult to determine the cause for the decrease in the ATPase activity of Q7a alone.

When ADCK3 was incubated with Q7b, the reaction did not generate any significant data indicating ATPase activity. This is interesting because it suggests that ADCK3 lost its ATPase activity in the presence of Q7b. Similar to the observations seen in Q7a reactions, the loss of ATPase activity in Q7b may be that since ADCK3 recognizes its substrate, rather than releasing the γ -phosphate into the solution, it is trying to transfer it on to the substrate proteins; however, Q7b construct lacks the putative phosphorylation sites that exist in the N-terminus of Coq7. Therefore, it is difficult to determine if the ATPase activity of ADCK3 was inhibited by the presence of Q7b.

Furthermore, when Q5b was incubated with ADCK3, the reaction generated absorbance measurements indicating higher reaction rates; however, it is difficult to determine whether this is the increase in the activity of ADCK3 or Q5b. Although the reaction rate of ADCK3 + Q5b was higher than both ADCK3 and Q5b reactions, it was still less than when the individual rates were added.

Repetition of the assay resulted in similar results and so, it was established that, before continuing further with the ATPase activity assay, the possibility of contaminants which may have ATPase activity from the purification of Q5b, Q7a, and Q7b must be considered. The procedure may need to be optimized to rid of any contaminating ATPases. Meanwhile, radioactive kinase assay was carried out using Coq5 and Coq7 as the substrates of ADCK3.

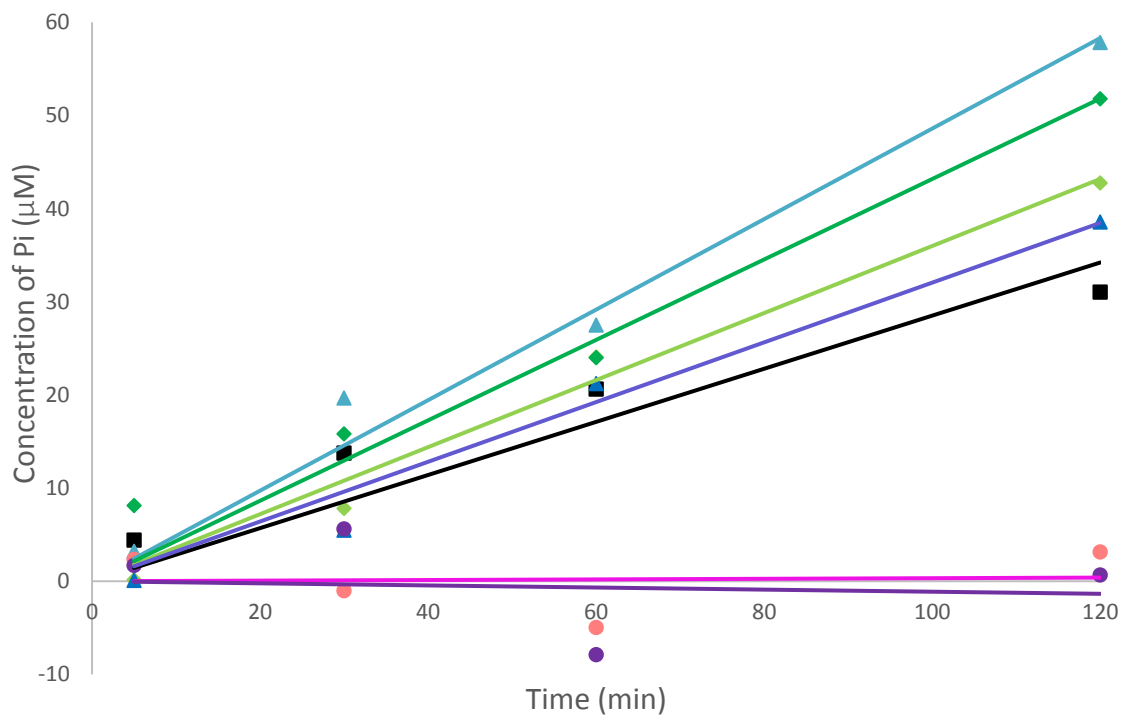


Figure 3.30 The ATPase activity of ADCK3 and its substrates.

The reactions were carried out in triplicates and averages are shown. Some outliers were eliminated from the data. The linear trend lines were placed with the y-intercept set to 0. The data points and their linear line of the individual reactions are shown in different colours: (black) 10 μM ADCK3; (light green) 2 μM Q5b; (dark green) 2 μM Q5b + 10 μM ADCK3; (light blue) 2 μM Q7a; (dark blue) 2 μM Q7a + 10 μM ADCK3; (pink) 2 μM Q7b; and (purple) 2 μM Q7b + 10 μM ADCK3.

3.11 Kinase activity of ADCK3

A preliminary set of radioactive kinase assay was performed using a mixture of Q5b, Q7a, and Q7b as the substrate for ADCK3. The results showed that the phosphorylation signal for the reaction containing just ADCK3 slightly decreased with time (Figure 3.31). On the other hand, the signal for the Q protein mixture slightly increased with time. In addition, the reaction mixture containing both ADCK3 and Q protein mixture did not show an increased phosphorylation signal. Initially, it was thought that lack of kinase activity observed from ADCK3 in the presence of the Q proteins was due to the structural hindrance caused by the bulky MBP tags of all the constructs. Consequently, the assay was repeated with ADCK3 (245-647) without a protein tag; however, the results remained similar.

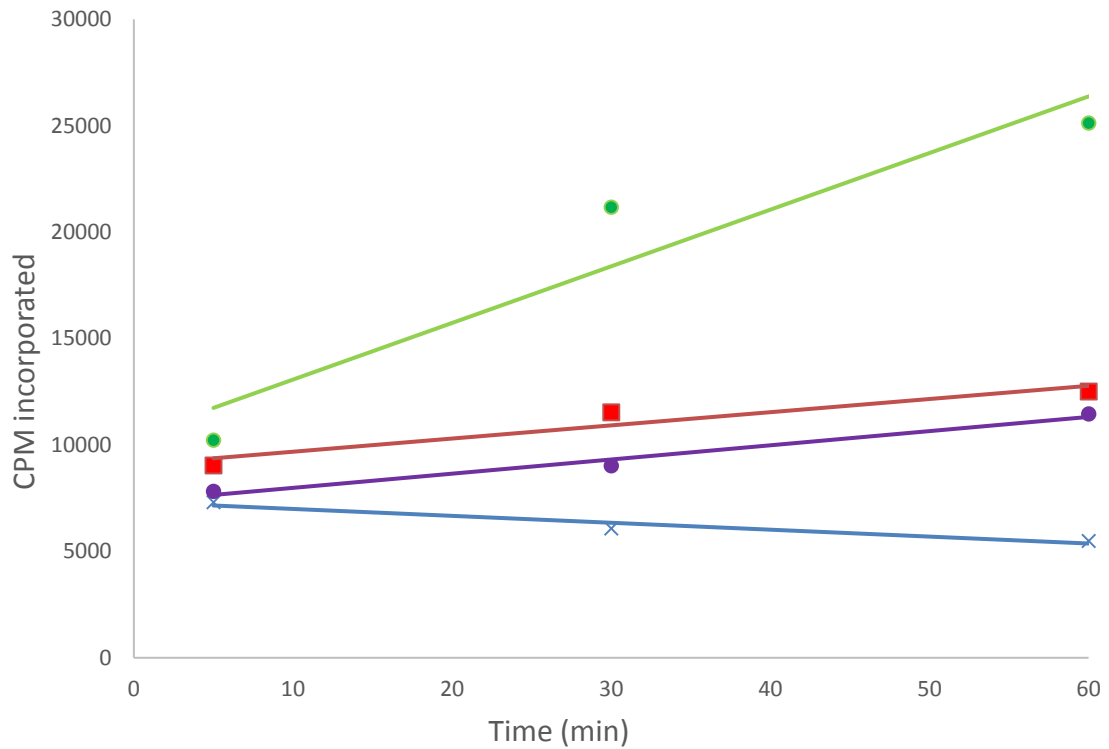


Figure 3.31 The kinase activity of ADCK3 in the absence and presence of Coq proteins. Blue graph corresponds to reaction with only ADCK3, red graph corresponds to reaction with Q protein mix (consisting of Q5b, Q7a, and Q7b), purple graph corresponds to reaction with ADCK3 and Q protein mix, and green graph corresponds to eEF2K protein used as a positive control.

Initially, it was expected for ADCK3 to show increasing phosphorylation activity for both reactions that do and do not contain Coq proteins. For the reaction not containing the substrate mixture, increasing signal was expected as *cis*-autophosphorylation activity have been previously observed; however, upon close investigation, autophosphorylation activity of the wild-type ADCK3 was minuscule compared to ADCK3 with a A339G mutation [41]. Alanine 339 residue is part of the alanine-rich loop of ADCK3 that is in place of a glycine-rich loop found in canonical kinases. This increase in autophosphorylation activity when alanine was mutated to a glycine shows the importance of the glycine residues. Although alanine and glycine are very similar in structure, glycine-rich loop may have the optimal flexibility to allow for enzymatic functions. It is considered to be so as glycine-rich loops are conserved in many different types of enzymes such as kinases, deacetylases, and even in methyltransferases. For instance, in a study done with cAMP-dependent protein kinase, a significant decrease in kinase activity were shown when the glycine residues of the glycine-rich loop were mutated to alanine or serine residues [54]. Specifically, mutation of the first two conserved glycine residues of the GxGxxG sequence resulted in 50-80% reduction in the turnover rate. In addition, in a study done with a histone deacetylase, a significant difference in activity was shown when the glycine-rich loop of the enzyme was mutated to an alanine-rich loop [55]. The study showed that when 4 glycine residues of the loop were mutated to alanine, the enzyme showed to have substantial loss of catalytic efficiency. In addition, the crystal structure of the enzyme harbouring the mutation showed change of a catalytic residue orientation due to the rigidity formed by the substitution of glycine residues to alanine residues. Although this observation does not pertain to a kinase, it shows the structural importance of having glycine residues to allow for optimal efficiency of enzymatic activity.

Above all, when the structure of apo- and holo- ADCK3 were compared, there was a significant shift in the position of the alanine-rich loop. When the two structures were aligned,

there was a slight overlap between the AMP-PNP of the holo-structure and the alanine-rich loop of the apo-structure. As previously mentioned, in order to allow ADCK3 to have this structural change to include the nucleotide, a R611K mutation was made that led to a loosened interaction between the N-lobe and the C-lobe. This allowed for more space in the active site that is located in the cleft between the N-lobe and C-lobe. Once again, the flexibility of the overall structure is demonstrated to be of great importance in binding of a nucleotide which is essential for a kinase activity.

It may be interesting to investigate the structure of ADCK3 harbouring a glycine-rich loop in place of an alanine-rich loop to assess the degree of importance of the flexibility provided by glycine residues for binding of a nucleotide. In contrast, there may be a reason as to why this alanine-rich loop survived through evolution across species. Perhaps it is better to have a diminished ADCK3 activity rather than having highly active ADCK3 that requires a tight regulation.

In terms of the lack of activity observed in the reaction containing ADCK3 and Q protein mixture, similar observations were made when kinase activity assay was done with the yeast homolog, Coq8, with Coq3-7, 9 as the substrates [41]. Although we cannot detect any kinase activity for ADCK3 or Coq8 in the presence of their putative substrates, the fact that these putative substrates were observed to be phosphorylated and dephosphorylated in wild-type and Coq8 knockout yeast cells offer a possible explanation. In the study that showed the phosphorylation patterns on Coq3, coq5, and Coq7, the phosphorylation activity of Coq8 occurred in the yeast cell [34–36]; however, the phosphorylation assay that were done in this thesis and the study done by Stefely *et al*, the enzyme and the substrates were expressed and purified separately [41]. As a result, it was considered that perhaps there is a component missing in our *in vitro* assays that may serve as the activator of our kinases. When Coq8 was active in the wild-type yeast cell, this activator would have coexisted in the cell allowing it to activate Coq8.

In addition, this activator molecule or protein may perhaps release the N-terminal “inhibitory” extension of ADCK3 and Coq8 to allow for them to function. Furthermore, the potential activator may alleviate the constraints caused by the alanine-rich loop.

At the same time, although the phosphorylation pattern disappeared in Coq8 knock-out yeast cells, Coq8 should also be considered to be the activator of another kinase that has not been previously identified. In any case, although kinase activity of ADCK3 was not shown in our work, ADCK3 should not be considered as an inactive kinase before further experiments suggest otherwise. As previously mentioned, although ADCK3 conserves only 5 of the 12 essential kinase subdomains, the 5 domains that are conserved are the catalytic domains of a kinase. The 7 domains that are considered missing are the domains that play roles in stabilizing the overall structure. For example, ADCK3 conserves subdomains I, II, III, VIB, and VII, which are the subdomains that carry the catalytic residues (Figure 1.5). The subdomains that are not conserved by ADCK3, subdomains IV, V, VIA, VIII, IX, X, and XI, play a role in stabilizing the overall structure or have an unknown function. It is perhaps the additional N-terminal extension compensates for the lack of these subdomains in addition to playing a regulatory role. Consequently, in addition to the ATPase activity observed in previous studies, these characteristics suggest that ADCK3 may be harbouring a tightly regulated kinase activity.

3.12 General Discussion

In addition to our results, evidence shown from previous studies can suggest some explanations for the lack of kinase activity observed from ADCK3 in the presence of Coq proteins. For one thing, when the apo- and holo- structures of ADCK3 were compared, both structures retained a cavity in which a nucleotide can enter the active site; however, the space allowed for the nucleotide to bind to the active site was more than doubled for the holo- structure (Figure 3.32). Specifically, the distance between alanine 338 of the alanine-rich loop and serine 492 of the disordered region between $\beta 6$ and $\beta 7$, where the nucleotide was found in the holo- structure, was measured for both apo- and holo- structures. The resulting distance were 3.8 and 10.0 Å for the apo- and holo- structures, respectively. It is unclear as to what induced the structural shift that allowed the nucleotide to enter and bind to the alanine-rich loop in the holo- structure. A possible explanation for this structural shift may be the fact that a mutation was made to the kinase domain of ADCK3 that was co-crystallized with AMP-PNP. The arginine to lysine mutation led to a “loosened” structure where a salt-bridge holding the C-terminal domain together was lost (Figure 3.9B). The 2 hydrogens of ϵ nitrogen and one of the η nitrogen atoms of the arginine 611 residue side chain originally formed hydrogen bonds with the 2 oxygen atoms of the aspartic acid 488 residue side chain; however, when the arginine was mutated to lysine, although the charge of the residue remained positively charged, the shorter side chain length resulted in the loss of hydrogen bond to the aspartic acid 488 side chain. Furthermore, the comparison between the distance from methionine 280 and arginine/lysine 611 showed increased distance in the holo- structure (Figure 3.33). As a result, this further suggest that the structural shift observed was due to the arginine to lysine mutation rather than the binding of the nucleotide.

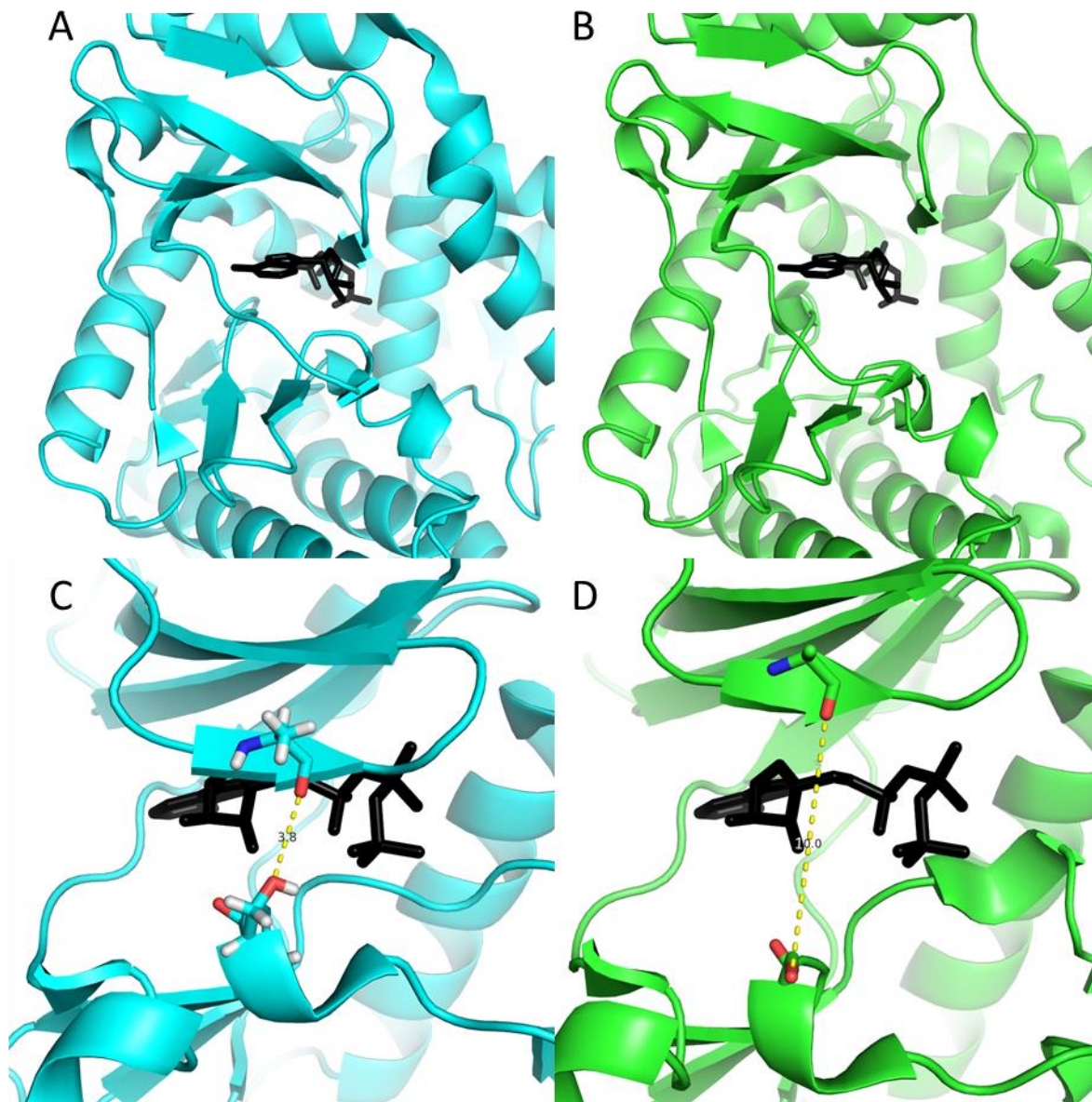


Figure 3.32 The nucleotide binding cavity comparison between apo- and holo- ADCK3 structures.

The apo- and holo- structures were aligned and AMP-PNP binding regions in the apo-structure was analysed. The AMP-PNP molecule is shown as black sticks while the apo- and holo- structures are shown in cyan and green, respectively. The distance between alanine 338 and serine 492 are shown for both apo- and holo- structures to be 3.8 and 10.0 Å, respectively.

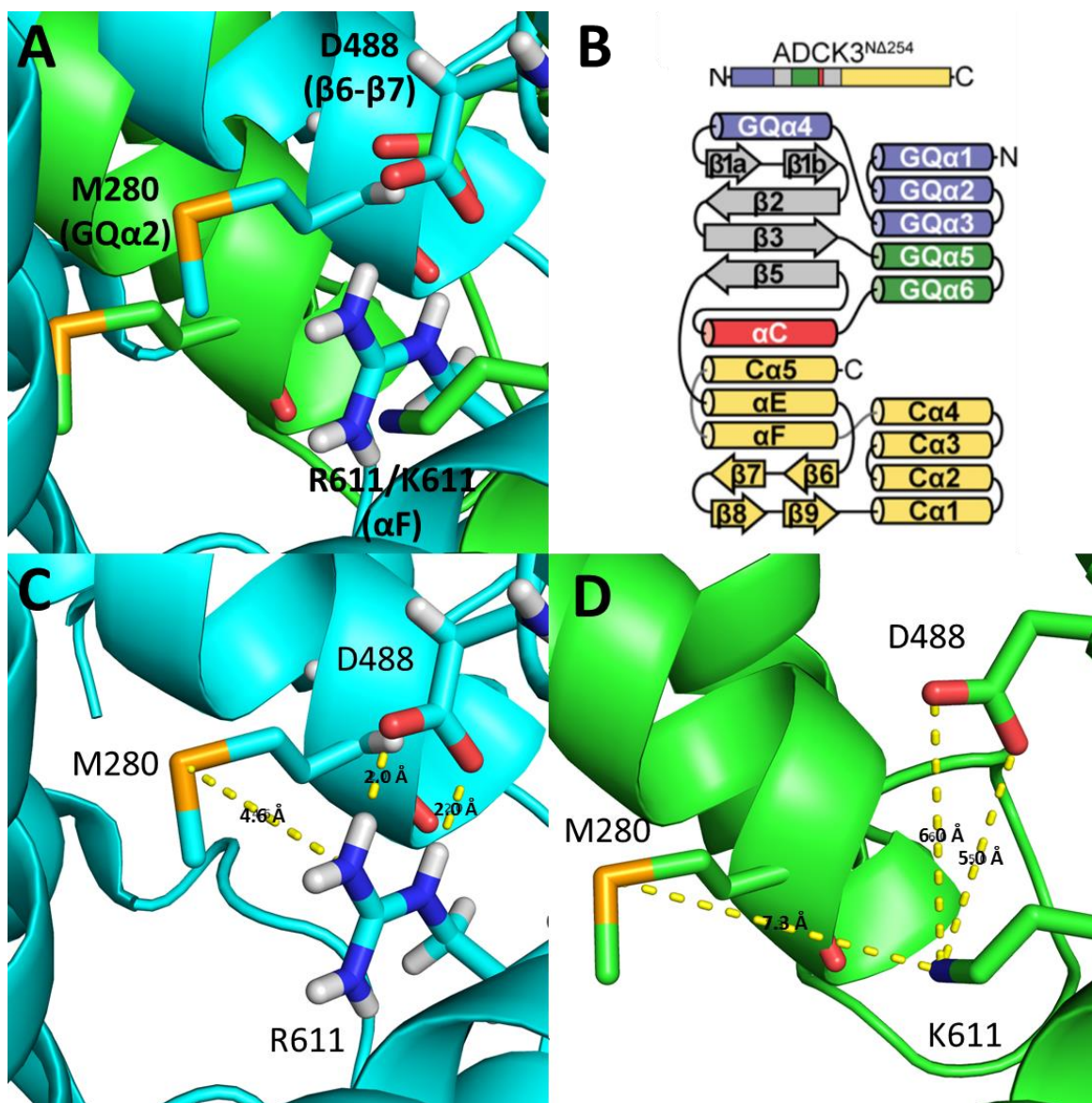


Figure 3.33 The comparison of the wild-type residue arginine 611 and mutated residue lysine 611 and their effect on the structure of ADCK3.

The apo- and holo- structures of ADCK3 are shown in cyan and green colours, respectively. The residues arginine/lysine 611, aspartic acid 488, and methionine 280 are shown as sticks and labelled. (A) Overlay of the apo- and holo- structure showing the structural shift of the GQ α 2 helix. (B) The schematic map of the secondary structure of ADCK3 [14]. (C) The apo-structure of ADCK3 (PDB:4PED) showing wild-type arginine 611 residue forming hydrogen bonds with the aspartic acid 488. (D) The holo-structure of ADCK3 (PDB:5I35) showing the mutated lysine 611 residue not being able to form hydrogen bonds with the aspartic acid 488. The distance between R611/K611 and M280 shows the increase in the distance between N-terminal extension and the C-terminal domain in the holo- structure.

The requirement of structural shift for the binding of a nucleotide in the active site of ADCK3 is further supported by the results of the functional activity assays. For example, increase in ATPase activity of ADCK3 was shown when the alanine-rich loop was mutated to a glycine rich-loop [48]. As well, when the autophosphorylation activity of ADCK3 was observed, the construct harbouring the mutated glycine-rich loop showed higher amount of phosphorylated product [41]. The increase in the flexibility provided by the combination of 3 glycine residues may provide this structural shift required for the binding of nucleotide which led to their functional activity.

In addition to the structural features required for the phosphotransfer activity of ADCK3, further analysis revealed a small pocket that can potentially accommodate a small peptide substrate [41]. This pocket is located between the N-terminal extension and the C-terminal domain and connects to the active site region. It also seems to be accessed with minimal conformation changes to the N-terminal extension. If this pocket were to be the substrate binding region, then the idea of the N-terminal extension being released to accommodate the substrate would need some reconsideration. For instance, the putative phosphorylation sites of yeast Coq7 were predicted to be Ser20, Ser28, and Thr32 [36]. In the small pocket observed in the ADCK3 structure, the chain of 32 amino acids would not be able to fit; however, in the sequence of Coq7, there is a mitochondrial targeting sequence (MTS) that would be cleaved off once Coq7 has reached its target location in the mitochondria. Coq7 mitochondrial targeting sequence is predicted to be 15 amino acids, and so small molecule binding pocket would only have to accommodate for approximately 17 amino acids. Secondary structure prediction software predicts that this chain of amino acids is a highly disordered region which is the characteristic of a phosphorylation site. Perhaps the N-terminal extension is in place to act as a docking system for the disordered region to anchor itself so that the phosphotransfer can occur more fluently. In addition, if this were to be the case, then the lack of kinase activity observed in our assay could be

explained. Since the N-terminus of the Coq proteins were linked to a bigger protein, MBP, there would have been no way for this region to be able to enter the substrate pocket. As a result, aforementioned, the purification method needs to be optimized to be able to stably purify the Coq proteins after cleaving the MBP off its constructs.

Chapter 4

Conclusion and Future Directions

4.1 Conclusion

Currently, although the mild symptoms of CoQ₁₀ deficiency can be treated through CoQ₁₀ supplements, the severe symptoms of the disorders caused by CoQ₁₀ deficiency lacks treatment that can fully reverse the detrimental effects experienced by the patients. In order to develop therapeutics that can prevent or relieve these symptoms, the mechanism of CoQ biosynthesis and its regulation must be understood in detail. This thesis aimed to characterize ADCK3/Coq8A and its three potential substrates (Coq3, Coq5, and Coq7), all involved in the CoQ biosynthetic process.

In this thesis, we have found a new condition that leads to the crystallization of ADCK (245-647) that does not use ammonium sulphate. Using the resulting protein crystals, diffraction data was collected and processed to solve the structure of ADCK3. Attempts were made to co-purify protein substrates of ADCK3 and ADCK4 from human embryonic kidney cell (HEK293) lysate to assess their kinase activity. The results from a similar study done by Stefely *et al.* showed co-purification of Coq3, Coq4, Coq5, Coq6, and Coq7 when FLAG-tagged ADCK3 was purified from HEK293 cells. Various truncations of Coq3, Coq5, and Coq7 were successfully cloned into expression vectors. From the extensive expression trials, the combination of N-terminal MBP fusion and BL21-CodonPlus (DE3)-RIPL *E. coli* cell line led to the expression of both Coq5 and Coq7; however, further optimization of the plasmids must be done for the expression of Coq3 and the full length of Coq5. In addition, the purification methods for Coq5 and Coq7 were optimized to exclude the contaminants that produce background activity in the activity assays. Furthermore, using 2-methoxyhydroquinone, an analog of CoQ precursor, preliminary methyltransferase activity of Coq5 was shown. In addition, both ATPase activity

assay and radioactive kinase assay showed no significant evidence indicating change in activity of ADCK3 when its substrates are present. Further characterization must be done before any solid conclusions can be made.

It is hoped that our work can serve as a foundation for further study that aims to characterize the CoQ biosynthesis occurring in the human mitochondria.

4.2 Future directions

In order to ensure that the conformational change observed between the holo-structure containing AMP-PNP and the apo-structure was in fact due to AMP-PNP, a new condition to crystallize ADCK3 should be explored. The current holo-structure was solved by using a construct harbouring a mutation in the core of the structure. As a result, it cannot be certain as to if the conformational change was induced from the binding of AMP-PNP. It is more likely that the loss of 2 hydrogen bonding interaction between an arginine and an aspartic acid caused loosened structure of ADCK3 that led to the binding of AMP-PNP. To provide a solid foundation for the structural characterization, ADCK3 must be co-crystallized with a nucleotide without a mutation that will cause a structural shift.

On the other hand, conditions must be optimized to successfully express and solubilize the ADCK3 construct that lacks the N-terminal extension, ADCK3 (329-647). This construct can potentially help with co-crystallization of ADCK3 with its substrate proteins. In addition, previous work with ADCK3 construct lacking the N-terminal extension showed both ATPase and kinase activity; however, that construct lacked 2 helices at the C-terminal which is most likely the reason for its heavy degradation. As a result, in order to obtain a more stable construct lacking the N-terminal extension, the construct must be cloned to include the full C-terminal domain. After failure to expression MBP fusion construct of ADCK3 (329-647) in *E. coli*, ADCK3 (329-647) construct was cloned into a vector used for insect cell expression. Perhaps the expression of ADCK3 in insect will yield soluble protein ready to be used for further experiments.

Considering the conditions of the experiments that produced the indirect evidence of kinase activity of Coq8, yeast homolog of ADCK3, the possibility of ADCK3 requiring a cofactor should also be considered. The studies that showed phosphorylation patterns on the putative substrates of Coq8 were done in yeast cells. Meanwhile the studies that were unable to produce the phosphorylation patterns on the putative substrates of both Coq8 and ADCK3 were done after the individual proteins were separately expressed and purified. As a result, it can be suspected that there is a cofactor or a specific physiological condition in the cell that triggers Coq8 and ADCK3 to function.

Furthermore, in addition to providing a phosphorylation target for ADCK3 for the kinase activity assays, full lengths constructs of Coq3 and Coq5 must be expressed and purified. Aforementioned, the Coq3 cDNA may require codon optimization to inhibit the formation of secondary structures within the plasmid. In terms of the full length of Coq5, the last round of mutagenesis may help to prepare the plasmid for efficient transcription and translation. At this time, efforts are being made to express all the Coq proteins using insect cell line.

Aforementioned, in order to confirm the preliminary evidence of Coq5 methyltransferase activity, radioactive methyltransferase assay may be done. The substrate analog used in the study done with the yeast Coq5 may be difficult to obtain and may need to be synthesized; however, it seems to be the best method to overcome the obstacle of using a kit that measured the change in absorbance or fluorescence within a sample.

In addition to solving the structure of ADCK3, structures of human Coq3, Coq5, and Coq7 remain to be solved. In order to do so, buffers must be optimized to be able to purify stable Coq3 and Coq5 after cleaving the MBP tag off. Then, various screening trays should be set up using various concentration of each construct. Also, the screening trays that produced the initial hits observed should be set again in order to test the reproducibility. If reproducible, expansion trays of those specific conditions should be set and optimized.

When performing kinase activity assays in the future, it may also be worth to mutate the potential phosphorylation target sites on the substrates to ensure that the signal produced is due to the phosphorylation of the substrate protein. In addition, although it is highly unlikely, perhaps the substrate proteins are phosphorylated in the cell used to express the protein. UbiB, the *E. coli* homolog of ADCK3, may have phosphorylated Coq5 and Coq7, which may have led to the lack of ADCK3 activity observed in our activity assays. Mass spectrometry can be used to identify any post-translational modification done to proteins. If the results determine that the substrate proteins are indeed phosphorylated, then perhaps incubating these samples with a phosphatase can ensure that the proteins are ready to be phosphorylated. Furthermore, a more sensitive method for determining whether the putative substrate is phosphorylated or not would be to do an autoradiogram to visualize it in a SDS-PAGE gel-like form where the phosphorylated protein is fixed in a gel.

Finally, it may be interesting to express the individual members of the CoQ biosynthetic complex and form the complex to observe how the individual proteins are organized together. Although the complex will be too large for solving the structure using X-ray crystallography, it may be possible to do so using cryo-EM. The structural details of the complex will provide many important details about how these proteins interact to stabilize the complex, how the CoQ precursor is passed along from one another, and ultimately, how many proteins are actually found in the complex. In addition, doing a kinase activity with the individual proteins and with the complex and comparing the results may give us a hint as to when these substrates are phosphorylated. At this time, it is uncertain when the substrates are phosphorylated. There could be predictions that say phosphorylation may help with molecular organization of these individual proteins into a complex or predictions that say phosphorylation occurs after the complex has formed to regulate the activity of the substrate proteins.

References

- [1] M.J. Hubbard, P. Cohen, On target with a new mechanism for the regulation of protein phosphorylation, *Trends Biochem. Sci.* 18 (1993) 172–177.
- [2] G. Manning, G.D. Plowman, T. Hunter, S. Sudarsanam, Evolution of protein kinase signaling from yeast to man, *Trends Biochem. Sci.* 27 (2002) 514–520.
doi:10.1016/S0968-0004(02)02179-5.
- [3] P. Lahiry, A. Torkamani, N.J. Schork, R.A. Hegele, Kinase mutations in human disease : interpreting genotype – phenotype relationships, *Nat. Rev. Genet.* 11 (2010) 60–74.
doi:10.1038/nrg2707.
- [4] J.A. Endicott, M.E.M. Noble, L.N. Johnson, The Structural Basis for Control of Eukaryotic Protein Kinases, *Annu. Rev. Biochem.* 81 (2012) 587–613.
doi:10.1146/annurev-biochem-052410-090317.
- [5] S.S. Taylor, A.P. Kornev, Protein kinases: Evolution of dynamic regulatory proteins, *Trends Biochem. Sci.* 36 (2011) 65–77. doi:10.1016/j.tibs.2010.09.006.
- [6] S.K. Hanks, T. Hunter, The eukaryotic protein kinase superfamily : (catalytic) domain structure and classification, *FASEB J.* 9 (1995) 576–596.
- [7] N. Laronde-leblanc, A. Wlodawer, The RIO kinases : An atypical protein kinase family required for ribosome biogenesis and cell cycle progression, *Biochim. Biophys. Acta.* 1754 (2005) 14–24. doi:10.1016/j.bbapap.2005.07.037.
- [8] E. Vanrobays, P.E. Gleizes, C. Bousquet-Antonelli, J. Noaillac-Depeyre, M. Caizergues-Ferrer, J.P. Gélugne, Processing of 20S pre-rRNA to 18S ribosomal RNA in yeast requires Rp10p, an essential non-ribosomal cytoplasmic protein, *EMBO J.* 20 (2001) 4204–4213.
doi:10.1093/emboj/20.15.4204.
- [9] G. Giaever, A.M. Chu, L. Ni, C. Connelly, L. Riles, S. Véronneau, S. Dow, A. Lucau-Danila, K. Anderson, B. André, A.P. Arkin, A. Astromoff, M. El Bakkoury, R. Bangham,

- R. Benito, S. Brachat, S. Campanaro, M. Curtiss, K. Davis, A. Deutschbauer, K.D. Entian, P. Flaherty, F. Foury, D.J. Garfinkel, M. Gerstein, D. Gotte, U. Güldener, J.H. Hegemann, S. Hempel, Z. Herman, D.F. Jaramillo, D.E. Kelly, S.L. Kelly, P. Kötter, D. LaBonte, D.C. Lamb, N. Lan, H. Liang, H. Liao, L. Liu, C. Luo, M. Lussier, R. Mao, P. Menard, S.L. Ooi, J.L. Revuelta, C.J. Roberts, M. Rose, P. Ross-Macdonald, B. Scherens, G. Schimmack, B. Shafer, D.D. Shoemaker, S. Sookhai-Mahadeo, R.K. Storms, J.N. Strathern, G. Valle, M. Voet, G. Volckaert, C. Yun Wang, T.R. Ward, J. Wilhelmy, E.A. Winzeler, Y. Yang, G. Yen, E. Youngman, K. Yu, H. Bussey, J.D. Boeke, M. Snyder, P. Philippsen, R.W. Davis, M. Johnston, Functional profiling of the *Saccharomyces cerevisiae* genome, *Nature*. 418 (2002) 387–391. doi:10.1038/nature00935.
- [10] G. Dong, G. Chakshusmathi, S.L. Wolin, K.M. Reinisch, Structure of the La motif: A winged helix domain mediates RNA binding via a conserved aromatic patch, *EMBO J.* 23 (2004) 1000–1007. doi:10.1038/sj.emboj.7600115.
- [11] M.P. Cicero, S.T. Hubl, C.J. Harrison, O. Littlefield, J. a Hardy, H.C. Nelson, The wing in yeast heat shock transcription factor (HSF) DNA-binding domain is required for full activity., *Nucleic Acids Res.* 29 (2001) 1715–1723. doi:10.1093/nar/29.8.1715.
- [12] K.S. Gajiwala, S.K. Burley, Winged helix proteins, *Curr. Opin. Struct. Biol.* 10 (2000) 110–116. doi:10.1016/S0959-440X(99)00057-3.
- [13] C.H. He, L.X. Xie, C.M. Allan, U.C. Tran, C.F. Clarke, Coenzyme Q supplementation or over-expression of the yeast Coq8 putative kinase stabilizes multi-subunit Coq polypeptide complexes in yeast coq null mutants ☆, *Biochim. Biophys. Acta.* 1841 (2014) 630–644. doi:10.1016/j.bbali.2013.12.017.
- [14] J.A. Stefely, A.G. Reidenbach, A. Ulbrich, K. Oruganty, B.J. Floyd, A. Jochem, J.M. Saunders, I.E. Johnson, C.E. Minogue, R.L. Wrobel, G.E. Barber, D. Lee, S. Li, N. Kannan, J.J. Coon, C.A. Bingman, D.J. Pagliarini, Mitochondrial ADCK3 employs an

- atypical protein kinase-like fold to enable coenzyme Q Biosynthesis, *Mol. Cell.* 57 (2015) 83–94. doi:10.1016/j.molcel.2014.11.002.
- [15] C.M. Quinzii, M. Assoum, C. Lagier-tourenne, M. Tazir, L.C. Lo, N. Drouot, C. Busso, S. Makri, L. Ali-pacha, T. Benhassine, L. Bianchetti, M. Anheim, D.R. Lynch, C. Thibault, C. Tranchant, O. Poch, S. Dimauro, J. Mandel, M.H. Barros, M. Hirano, M. Koenig, ADCK3 , an Ancestral Kinase , Is Mutated in a Form of Recessive Ataxia Associated with Coenzyme Q 10 Deficiency, *Am. J. Hum. Genet.* 82 (2008) 661–672. doi:10.1016/j.ajhg.2007.12.024.
- [16] B. Wheeler, Z. Jia, Preparation and characterization of human ADCK3, a putative atypical kinase, *Protein Expr. Purif.* 108 (2015) 13–17. doi:10.1016/j.pep.2014.12.008.
- [17] R. Horvath, B. Czermin, S. Gulati, S. Demuth, G. Houge, A. Pyle, C. Dineiger, E.L. Blakely, A. Hassani, C. Foley, M. Brodhun, K. Storm, J. Kirschner, G.S. Gorman, H. Lochmu, E. Holinski-feder, R.W. Taylor, P.F. Chinnery, Adult-onset cerebellar ataxia due to mutations in CABC1/ADCK3, *J. Neurol. Neurosurg. Psychiatry.* 83 (2012) 174–178. doi:10.1136/jnnp-2011-301258.
- [18] E. Barca, O. Musumeci, F. Montagnese, S. Marino, F. Granata, D. Nunnari, Cerebellar ataxia and severe muscle CoQ 10 deficiency in a patient with a novel mutation in ADCK3, *Clin. Genet.* 90 (2016) 156–160. doi:10.1111/cge.12742.
- [19] K. Malgireddy, R. Thompson, D. Torres-Russotto, A novel CABC1/ADCK3 mutation in adult-onset cerebellar ataxia, *Park. Relat. Disord.* 33 (2016) 151–152. doi:10.1016/j.parkreldis.2016.10.010.
- [20] O. Hikmat, C. Tzoulis, P.M. Knappskog, S. Johansson, H. Boman, P. Sztromwasser, E. Lien, E. Brodtkorb, D. Ghezzi, L.A. Bindoff, ADCK3 mutations with epilepsy, stroke-like episodes and ataxia: a POLG mimic?, *Eur. J. Neurol.* 23 (2016) 1188–1194. doi:10.1111/ene.13003.

- [21] Y.T. Liu, J. Hersheson, V. Plagnol, K. Fawcett, K.E.C. Duberley, E. Preza, I.P. Hargreaves, A. Chalasani, M. Laurá, N.W. Wood, M.M. Reilly, H. Houlden, Autosomal-recessive cerebellar ataxia caused by a novel ADCK3 mutation that elongates the protein: Clinical, genetic and biochemical characterisation, *J. Neurol. Neurosurg. Psychiatry*. 85 (2014) 493–498. doi:10.1136/jnnp-2013-306483.
- [22] D. Chretien, D. Schlemmer, J. Mollet, O. De Baulny, A. Lombes, N. Boddaert, I. Desguerre, P. De Lonlay, CABC1 Gene Mutations Cause Ubiquinone Deficiency with Cerebellar Ataxia and Seizures, *Am. J. Hum. Genet.* 82 (2008) 623–630. doi:10.1016/j.ajhg.2007.12.022.
- [23] M. Gerards, B. van den Bosch, C. Calis, K. Schoonderwoerd, K. van Engelen, M. Tijssen, R. de Coo, A. van der Kooi, H. Smeets, Nonsense mutations in CABC1/ADCK3 cause progressive cerebellar ataxia and atrophy, *Mitochondrion*. 10 (2010) 510–515. doi:10.1016/j.mito.2010.05.008.
- [24] Y. Wang, J. Hersheson, D. Lopez, M. Hammer, Y. Liu, K.H. Lee, V. Pinto, J. Seinfeld, S. Wiethoff, J. Sun, R. Amouri, F. Hentati, N. Baudry, J. Tran, A.B. Singleton, M. Coutelier, A. Brice, G. Stevanin, A. Durr, X. Bi, H. Houlden, M. Baudry, Defects in the CAPN1 Gene Result in Alterations in Cerebellar Development and Cerebellar Ataxia in Mice and Humans, *Cell Rep.* 16 (2016) 79–91. doi:10.1016/j.celrep.2016.05.044.
- [25] M.J. Acosta, L.V. Fonseca, M.A. Desbats, C. Cerqua, R. Zordan, E. Trevisson, L. Salviati, Coenzyme Q biosynthesis in health and disease, *Biochim. Biophys. Acta*. 1857 (2016) 1079–1085. doi:10.1016/j.bbabbio.2016.03.036.
- [26] M. Forsgren, A. Attersand, S. Lake, J. Grunler, E. Swiezewska, G. Dallner, I. Climent, Isolation and functional expression of human COQ2, a gene encoding a polyprenyl transferase involved in the synthesis of CoQ, *Biochem. J.* 382 (2004) 519–526.
- [27] T. Jonassen, C.F. Clarke, Isolation and Functional Expression of Human COQ3, a Gene

- Encoding a Methyltransferase Required for Ubiquinone Biosynthesis, *J. Biol. Chem.* 275 (2000) 12381–12387.
- [28] Y. Dai, D. Cao, F. Meng, C. Chi, Y. Ren, Y. Chen, C. Zhou, Crystal structures and catalytic mechanism of the C-methyltransferase Coq5 provide insights into a key step of the yeast coenzyme Q synthesis pathway, *Acta Crystallogr.* 70 (2014) 2085–2092. doi:10.1107/S1399004714011559.
- [29] M. Ozeir, U. Mu, H. Webert, R. Lill, M. Fontecave, F. Pierrel, Coenzyme Q Biosynthesis : Coq6 Is Required for the C5-Hydroxylation Reaction and Substrate Analogs Rescue Coq6 Deficiency, *Chem. Biol.* 18 (2011) 1134–1142. doi:10.1016/j.chembiol.2011.07.008.
- [30] B. Marbois, P. Gin, M. Gulmezian, C.F. Clarke, The yeast Coq4 polypeptide organizes a mitochondrial protein complex essential for coenzyme Q biosynthesis, *Biochim. Biophys. Acta.* 1791 (2009) 69–75. doi:10.1016/j.bbaliip.2008.10.006.
- [31] D.C. Lohman, F. Forouhar, E.T. Beebe, M.S. Stefely, C.E. Minogue, A. Ulbrich, J.A. Stefely, S. Sukumar, M. Luna-Sanchez, A. Jochem, S. Lew, J. Seetharaman, R. Xiao, H. Wang, M.S. Westphall, R.L. Wrobel, J.K. Everett, J.C. Mitchell, L.C. Lopez, J.J. Coon, L. Tong, D.J. Pagliarini, Mitochondrial COQ9 is a lipid-binding protein that associates with COQ7 to enable coenzyme Q biosynthesis, *Proc. Natl. Acad. Sci.* 111 (2014) E4697–E4705. doi:10.1073/pnas.1413128111.
- [32] U.C. Tran, C.F. Clarke, Endogenous Synthesis of Coenzyme Q in Eukaryotes, *Mitochondrion.* 27 (2007) 320–331. doi:10.1002/nbm.3066.Non-invasive.
- [33] H. Takahashi, K. Shimoda, Coenzyme Q10 in neurodegenerative disorders: Potential benefit of CoQ10 supplementation for multiple system atrophy, *World J. Neurol.* 4 (2014) 1–6. doi:10.5316/wjn.v4.i1.1.
- [34] L. Xie, E.J. Hsieh, S. Watanabe, C.M. Allan, J.Y. Chen, U.C. Tran, C.F. Clarke, Expression of the Human atypical kinase ADCK3 rescues coenzyme Q biosynthesis and

- phosphorylation of Coq polypeptides in yeast coq8 mutants, *Biochim. Biophys. Acta.* 1811 (2011) 348–360. doi:10.1016/j.bbaliip.2011.01.009.Expression.
- [35] A. Tauche, U. Krause-buchholz, Ubiquinone biosynthesis in *Saccharomyces cerevisiae* : the molecular organization of O -methylase Coq3p depends on Abc1p / Coq8p, *Fed. Eur. Microbiol. Soc.* 8 (2008) 1263–1275. doi:10.1111/j.1567-1364.2008.00436.x.
- [36] A. Martín-Montalvo, I. González-Mariscal, S. Padilla, M. Ballesteros, D.L. Brautigam, P. Navas, C. Santos-Ocaña, Respiratory-induced coenzyme Q biosynthesis is regulated by a phosphorylation cycle of Cat5p/Coq7p, *Biochem. J.* 440 (2011) 107–114. doi:10.1042/BJ20101422.
- [37] A. Martín-montalvo, I. González-mariscal, T. Pomares-viciana, S. Padilla-lópez, M. Ballesteros, L. Vazquez-fonseca, P. Gandolfo, D.L. Brautigam, P. Navas, C. Santos-ocaña, The Phosphatase Ptc7 Induces Coenzyme Q Biosynthesis by Activating the Hydroxylase Coq7 in Yeast, *J. Biol. Chem.* 288 (2013) 28126–28137. doi:10.1074/jbc.M113.474494.
- [38] S.W. Baba, G.I. Belogrudov, J.C. Lee, P.T. Lee, J. Strahan, J.N. Shepherd, C.F. Clarke, Yeast Coq5 C-Methyltransferase Is Required for Stability of Other Polypeptides Involved in Coenzyme Q Biosynthesis, *J. Biol. Chem.* 279 (2004) 10052–10059. doi:10.1074/jbc.M313712200.
- [39] L.X. Xie, M. Ozeir, J.Y. Tang, J.Y. Chen, S.K. Jaquinod, M. Fontecave, C.F. Clarke, F. Pierrel, Overexpression of the Coq8 kinase in *Saccharomyces cerevisiae* coq null mutants allows for accumulation of diagnostic intermediates of the coenzyme Q6 biosynthetic pathway, *J. Biol. Chem.* 287 (2012) 23571–23581. doi:10.1074/jbc.M112.360354.
- [40] D.E. Torres Pazmiño, M. Winkler, A. Glieder, M.W. Fraaije, Monooxygenases as biocatalysts: Classification, mechanistic aspects and biotechnological applications, *J. Biotechnol.* 146 (2010) 9–24. doi:10.1016/j.jbiotec.2010.01.021.
- [41] J.A. Stefely, F. Licitra, L. Laredj, A.G. Reidenbach, Z.A. Kemmerer, A. Grangeray, T.

- Jaeg-Ehret, C.E. Minogue, A. Ulbrich, P.D. Hutchins, E.M. Wilkerson, Z. Ruan, D. Aydin, A.S. Hebert, X. Guo, E.C. Freiburger, L. Reutenauer, A. Jochem, M. Chergova, I.E. Johnson, D.C. Lohman, M.J.P. Rush, N.W. Kwiecien, P.K. Singh, A.I. Schlagowski, B.J. Floyd, U. Forsman, P.J. Sindelar, M.S. Westphall, F. Pierrel, J. Zoll, M. Dal Peraro, N. Kannan, C.A. Bingman, J.J. Coon, P. Isope, H. Puccio, D.J. Pagliarini, Cerebellar Ataxia and Coenzyme Q Deficiency through Loss of Unorthodox Kinase Activity, *Mol. Cell.* 63 (2016) 608–620. doi:10.1016/j.molcel.2016.06.030.
- [42] A. Chaikuad, S. Knapp, F. Von Delft, Defined PEG smears as an alternative approach to enhance the search for crystallization conditions and crystal-quality improvement in reduced screens, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 71 (2015) 1627–1639. doi:10.1107/S1399004715007968.
- [43] W. Kabsch, XDS, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 66 (2010) 125–132. doi:10.1107/S0907444909047337.
- [44] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, *J. Appl. Crystallogr.* 40 (2007) 658–674. doi:10.1107/S0021889807021206.
- [45] P.D. Adams, P. V. Afonine, G. Bunkóczy, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 66 (2010) 213–221. doi:10.1107/S0907444909052925.
- [46] P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 60 (2004) 2126–2132. doi:10.1107/S0907444904019158.

- [47] D. Perkins, D. Pappin, D. Creasy, J. Cottrell, Probability-based protein identification by searching sequence database using mass spectrometry data., *Electrophoresis*. 20 (1999) 3551–3567.
- [48] B. Wheeler, FUNCTIONAL CHARACTERIZATION OF HUMAN ADCK3 AND ADCK4 , MITOCHONDRIAL ATYPICAL KINASES, 2015.
- [49] B. Jia, C.O. Jeon, High-throughput recombinant protein expression in *Escherichia coli* : current status and future perspectives, *Open Biol*. 6 (2016) 160196.
doi:10.1098/rsob.160196.
- [50] J.R. Buchan, I. Stansfield, Halting a cellular production line: responses to ribosomal pausing during translation, *Biol. Cell*. 99 (2007) 475–487. doi:10.1042/BC20070037.
- [51] S. Maloy, V. Stewart, R. Taylor, Genetic analysis of pathogenic bacteria: a laboratory manual, Cold Spring Harbor Laboratory Press, New York, 1996.
- [52] G.H. Gauss, M.A. Reott, E.R. Rocha, M.J. Young, T. Douglas, C.J. Smith, Characterization of the *Bacteroides fragilis* bfr Gene Product Identifies a Bacterial DPS-Like Protein and Suggests Evolutionary Links in the Ferritin Superfamily, *J. Bacteriol*. 194 (2012) 15–27. doi:10.1128/JB.05260-11.
- [53] A. Rath, M. Glibowicka, V.G. Nadeau, G. Chen, C.M. Deber, Detergent binding explains anomalous SDS-PAGE migration of membrane proteins, *Proc. Natl. Acad. Sci*. 106 (2009) 1760–1765. doi:10.1073/pnas.0813167106.
- [54] W. Hemmer, M. McGlone, I. Tsigelny, S.S. Taylor, Role of the glycine triad in the ATP-binding site of cAMP-dependent protein kinase, *J. Biol. Chem*. 272 (1997) 16946–16954. doi:10.1074/jbc.272.27.16946.
- [55] N.J. Porter, N.H. Christianson, C. Decroos, D.W. Christianson, Structural and Functional Influence of the Glycine-Rich Loop G302GGGY on the Catalytic Tyrosine of Histone Deacetylase 8, *Biochemistry*. 55 (2016) 6718–6729. doi:10.1021/acs.biochem.6b01014.

- [56] A.I. Jonckheere, J.A.M. Smeitink, R.J.T. Rodenburg, Mitochondrial ATP synthase: architecture, function and pathology, *J. Inherit. Metab. Dis.* 35 (2012) 211–225.
doi:10.1007/s10545-011-9382-9.
- [57] A.E. Senior, S. Nadanaciva, J. Weber, The molecular mechanism of ATP synthesis by F1F0-ATP synthase, *Biochim. Biophys. Acta - Bioenerg.* 1553 (2002) 188–211.
doi:10.1016/S0005-2728(02)00185-8.
- [58] A.N. Stephens, M.A. Khan, X. Roucou, P. Nagley, R.J. Devenish, The molecular neighborhood of subunit 8 of yeast mitochondrial F1F0-ATP synthase probed by cysteine scanning mutagenesis and chemical modification, *J. Biol. Chem.* 278 (2003) 17867–17875. doi:10.1074/jbc.M300967200.
- [59] C. Herrstadt, J.L. Elson, E. Fahy, G. Preston, D.M. Turnbull, C. Anderson, S.S. Ghosh, J.M. Olefsky, M. Flint Beal, R.E. Davis, N. Howell, Reduced-Median-Network Analysis of Complete Mitochondrial DNA Coding-Region Sequences for the Major African, Asian, and European Haplogroups, *Am. J. Hum. Genet.* 70 (2002) 1152–1171.
doi:10.1086/339933.

Appendix A

ATP synthase F₀ subunit 8 as the potential substrate of ADCK3

A.1 Introduction

Previous work done with ADCK3 and ADCK4 revealed that ADCK3 was able to phosphorylate a peptide corresponding to the sequence of ATP synthase F₀ subunit 8 (F₀S8) of *Homo sapiens* [48]. In order to identify this peptide sequence as the potential phosphorylation target of ADCK3, MBP-ADCK3 (329-616) construct was used in phosphorylation assays along with a library of peptides as the substrate. When the initial detectable activity was observed, the amino acid residues of the peptides were substituted at various positions of the peptide in order to determine the sequence that resulted in strongest kinase activity of ADCK3. When the peptide that produced the strongest signal was used as the substrate for a kinase activity assay, MBP-ADCK3 (329-616) was determined phosphorylate the peptide at a rate of 0.0033 min⁻¹ [48]. When the sequence of the peptide that produced the strongest signal was searched using the Basic Local Alignment Search Tool (BLAST), the single top result indicated that sequence is of *Homo sapiens* ATP synthase F₀ subunit 8. Although it is uncertain if F₀S8 is the substrate of ADCK3, the fact that both proteins are found in the inner mitochondria is a great foundation to the study.

ATP synthase is an enzyme found in the inner membrane of the mitochondria that synthesizes most of the ATP molecules that are used by the cell. As seen in Figure A.1, ATP synthase consists of a F₀ subunit embedded into the membrane and a F₁ subunit that is in the matrix of the mitochondria, connected by an axle and a stator. In simple terms, in order to synthesize ATP, F₀ subunit uses the proton gradient created by the electron transport chain to rotate its motor [56,57]. The flow of protons through its structure powers the rotation and in turn, this energy is transmitted to the F₁ subunit through the rotating axle. Then the F₁ subunit converts this mechanical energy into the chemical bond between ADP and inorganic phosphate.

Subunit 8 of the F₀ subunit (F₀S8) is an integral membrane protein that is also known as the A6L subunit (Figure A.1). Although the exact function of this subunit is known, it was shown to be critical in the formation and the stability of the F₀ subunit [58]. In this thesis, the F₀S8 was expressed and purified to be used as a substrate for assessing the kinase activity of ADCK3 and ADCK4.

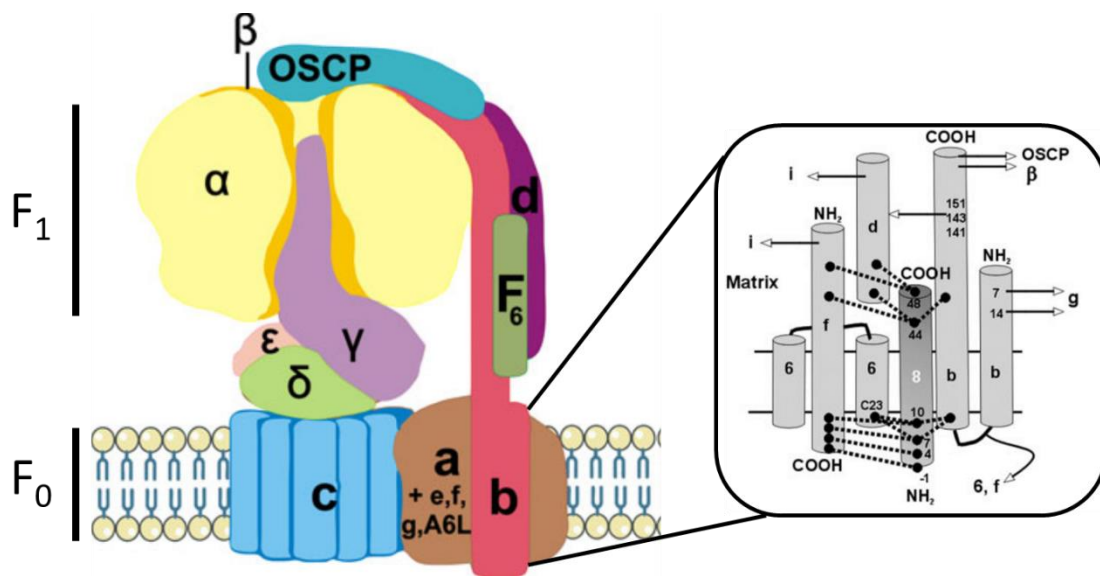


Figure A.1 The schematic structure of mitochondrial ATP synthase. The molecular organization of the smaller subunits that make up the F₀ subunit are shown on the right. Adapted from [56,58].

A.2 Methods

A.2.1 Mutagenesis

Complementary primers were designed so that the DNA base that is being mutated is in the centre of the primers. Using the designed primers, the plasmid containing the F₀S8 gene to be mutated was amplified using Pfu DNA polymerase (Thermo Fisher Scientific). Following the mutagenic PCR, 1 µL of DpnI (New England Biolabs) was added to the reaction mixture allowing for the digestion of the original methylated DNA. After incubating the reaction mixture at 37 °C for 1 hour, the digested mixture was transformed into Top10 chemically competent *E. coli* cells (Thermo Fisher Scientific). Colonies from the transformed plates were used to inoculate 5 mL of LB containing 100 µg/mL ampicillin and grown overnight at 37 °C. Plasmid DNA was purified from the 5 mL culture using the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific). The successful mutation of the gene was verified by DNA sequencing.

A.2.2 Expression and purification

The purified recombinant plasmid was transformed into BL21 (DE3) competent *Escherichia coli* cells (Agilent, Santa Clara, CA). A single colony was used to inoculate 25 mL LB containing 100 µg/mL ampicillin and grown overnight at 37 °C. Subsequently, the overnight culture was used to inoculate 1 L LB containing 100 µg/mL ampicillin. The cells were shaken while being grown at 37 °C until the culture OD₆₀₀ was between 0.6 and 0.8. The culture was cooled to room temperature before protein expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The culture was incubated at 18 °C overnight.

The cells were harvested by centrifugation at 4000 rpm for 30 minutes at 4 °C and then resuspended in lysis buffer A (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 10% glycerol, and 1 mg/mL lysozyme). After 20 minutes of incubation on ice, the resuspended cells were lysed by sonication (2 minutes with 5 seconds on, 15 seconds off, 30% amplitude). The debris of the cells

were then pelleted by centrifugation at 16000 rpm for 30 minutes at 4 °C and the supernatant was retained for further purification.

Purification – The N-terminal His-tag engineered into the recombinant protein was used to purify F₀S8 using immobilized metal affinity chromatography (IMAC). The supernatant of the cell lysate was allowed to flow through 5 mL of nickel-nitrilotriacetic acid (Ni⁺² – NTA) resin (Qiagen, Valencia, CA) that have been pre-equilibrated with 50 mL of buffer A (50 mM Tris-HCl pH 8.0, 300 mM NaCl, and 10% glycerol). Once the supernatant has flown through the resin, 50 mL of buffer A was added to the column and allowed to flow through to wash the resin followed by 25 mL of wash buffer A (buffer A with 10 mM imidazole). To elute our target protein from the resin, 25 mL of elution buffer A (buffer A with 400 mM imidazole) was added to the column. The collected fractions were then assessed using SDS-PAGE.

The protein fraction from the previous purifications were subjected to size-exclusion chromatography through a HiLoad 16/60 Superdex 200 prep-grade column (GE healthcare) or a HiLoad 16/60 Superdex 75 prep-grade column (GE healthcare) using fast-protein liquid chromatography (FPLC). Following SDS-PAGE analysis, fractions containing purified protein were combined and concentrated using 30K and 50K Amicon Ultra centrifugal filters for cleaved and non-cleaved target proteins, respectively. The optimal centrifugation condition to result in stable protein sample was 2500 rpm at 4 °C. The concentration of the protein samples was determined by measuring the absorbance of the diluted protein sample at 280 nm and calculating the concentration based on the protein's predicted extinction coefficient.

The purification of ADCK3 and ADCK4 constructs, 25HT29 and 34HT29, respectively were expressed and purified according to the methods described in the work done by Brody Wheeler [48].

A.2.3 Radioactive kinase assay

The kinase activity of ADCK3 and ADCK4 were assessed using 25HT29 and 34HT29 constructs cloned by Brody Wheeler [48]. The assay was performed using 5 μ M ADCK3 and either 50 μ M of the peptide and 100 μ M of MBP-F₀S8 as the substrate. The buffer used in the reaction consisted of 50 mM Tris-HCl pH 7.4, 1 mM DTT, and 20 mM MgCl₂. In addition to the buffer, the reaction mixture contained 1 mM of cold ATP spiked with 0.1 μ Ci/ μ L of [γ -³²P]-radiolabeled ATP (Perkin Elmer, Waltham, MA). For the second round of kinase activity assay, the cold ATP was spiked with 0.07 μ Ci/ μ L of [γ -³²P]-radiolabeled ATP instead.

The reactions were carried out in duplicates at 25 °C and 30 μ L aliquots were removed and placed on a P81 ion exchange cellulose chromatography paper (Whatman International Ltd., Maidstone, UK), at 5 min, 30 min, and 60 min time points. After the 60-minute time point, the paper squares were dried for 15 minutes before being washed twice in 1.6 % o-phosphoric acid for 15 minutes. Then, the paper squares were placed on a bed of paper towel in a plexiglass chamber over night to dry. Next day, the paper squares were immersed in 5 mL of ScintiVerse™ II cocktail buffer (Thermo Fisher Scientific). The radioactivity of the samples was measured using Beckman LS 9000 scintillation counter (Beckman Coulter, Brea, CA).

A.3 Results and discussion

A.3.1 Mutagenesis

The sequencing results showed successful mutation of Trp57 to a cysteine residue. The original peptide shown to be phosphorylated by ADCK3 (329-616) had a tryptophan residue on the -1 position; however, when analysing the sequence of F₀S8, it indicated that the tryptophan on the -1 position of the protein is the result of a single nucleotide polymorphism found among certain haplogroups [59]. Instead, cysteine was found to be the predominant variant for the -1 position in the F₀S8 protein. As a result, the plasmid harbouring the predominant variant, cysteine 57, was transformed into *E. coli* for its expression and purification.

A.3.2 Expression and purification

The results showed the over-expression of MBP-F₀S8 and the purification of MBP-F₀S8 via IMAC (Figure A.2). There were 2 predominant bands in the elution lanes of SDS-PAGE with very similar molecular weights of approximately 45 kDa and 48 kDa. These molecular weights match the expected molecular weight of His-tagged MBP and His-MBP- F₀S8 of 42.5 kDa and 45.4 kDa, respectively. As shown in Figure A.2, the 2 bands show approximately equal thickness suggesting that proper quantification of His-MBP- F₀S8 cannot be determined from this sample. As a result, size-exclusion chromatography step was added to the purification method of His-MBP- F₀S8; however, the results from size-exclusion chromatography using both Superdex 200 and Superdex 75 columns show a single elution peak containing both proteins (Figure A.3). As a result, since the MBP proteins will not affect the kinase activity of ADCK3, the protein sample containing both proteins were chosen to be used as the substrate.

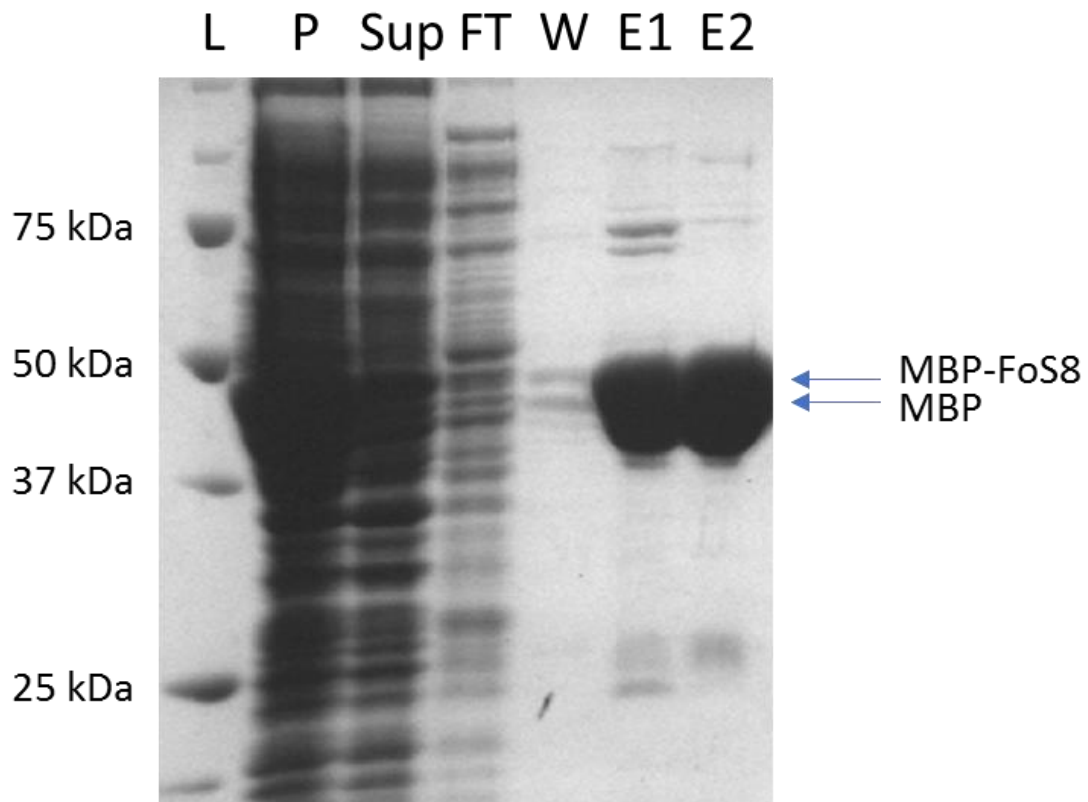


Figure A.2 A representative profile of MBP-FoS8 purification using histidine-mediated immobilized metal affinity chromatography.

Using the histidine-tag engineered into the construct, MBP-FoS8 was extracted from the supernatant of the cell lysate using Ni⁺²-NTA column. Throughout the purification process, 10 μ L of the fraction were collected and analysed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining. From left to right, the lanes represent: (L) ladder; (P) pellet; (Sup) supernatant; (FT) flow-through; (W) wash; (E1) elution 1; and (E2) elution 2. The elution fractions were pooled for further purification.

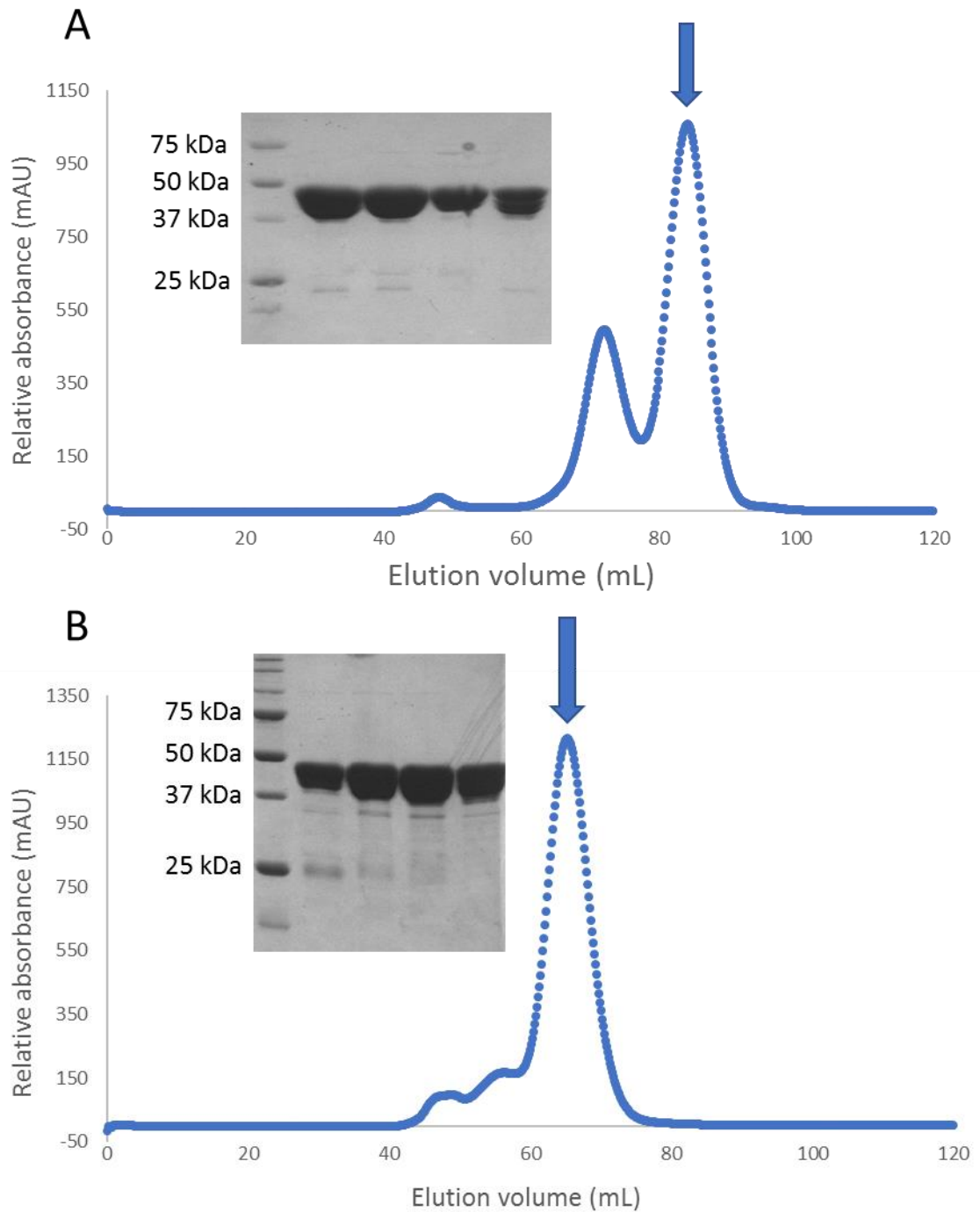


Figure A.3 Analysis of MBP-FoS8 following purification using size-exclusion chromatography.

Elution fractions from Ni²⁺-NTA were collected and further purified via size-exclusion chromatography over: (A) a HiLoad16/60 Superdex 200 prep-grade column and (B) a HiLoad16/60 Superdex 75 prep-grade column using FPLC. The peak containing MBP-FoS8 is indicated by a blue arrow. The fractions of that peak were analyzed using SDS-PAGE with 12.5% polyacrylamide gel and visualized by Coomassie Blue staining.

A.3.3 Radioactive kinase assay

Although the assay was performed in duplicates, the results showed considerable difference between the duplicates. In addition, processing the data after filtering out the outliers resulted in no specific trends (Figure A.4). To be specific, when the linear line of best fit was plotted for each reaction, the R^2 values of the four reactions were 6.4%, 0.2%, 57.6% and 73.5%. Interestingly, the reactions with the lowest R^2 values were the reactions containing the peptide as the substrate for both ADCK3 and ADCK4. Interestingly, for all reactions, the results showed increase in phosphorylated product from 5-minute time point to the 30-minute time point. Then, at the 60-minute time point for both ADCK3 and ADCK4 reactions containing MBP-F₀S8, the amount of phosphorylated product remained the same as the 30-minute time point. In contrast, the results showed that both ADCK3 and ADCK4 reactions containing the peptide had decreased amount of phosphorylated product at the 60-minute time point. In addition, at the 5-minute time point, the reactions containing the peptides have significantly higher values than the values of the reactions containing the MBP-F₀S8.

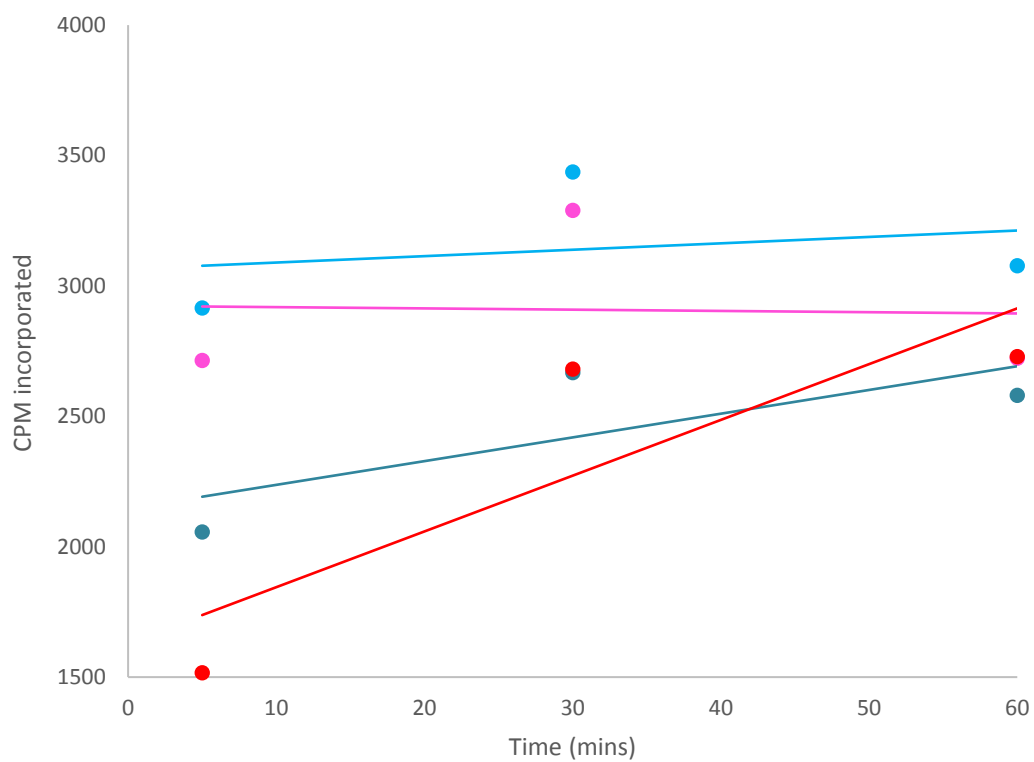


Figure A.4 Radioactive kinase activity of ADCK3 and ADCK4 with peptide segment and full length of human ATP synthase F₀S₈ as the substrate.

The individual reactions are represented with different colours: (Light blue) ADCK3 with peptide substrate; (Dark blue) ADCK3 with MBP-F₀S₈; (Pink) ADCK3 with peptide substrate; and (Red) ADCK4 with MBP-F₀S₈.

There are couple of possible explanations that can explain the results shown in the assay; however, the results of this assay cannot be confirmed to be valid for many reasons as well. First, During the assay, precipitation formation was observed. The protein that formed precipitation was predicted to be ADCK3 and ADCK4 precipitating in the reaction because the 25HT29 and 34HT29 constructs used have been consistently unstable. Both proteins have shown to be highly unstable and degrade in the buffers they were purified in. This instability can be explained by analysis of the ADCK3 structure. As shown in Figure A.5, the residues not included in the 25HT29 construct are located in the centre of the kinase. Lack of these segments would disrupt the overall stability of the kinase. Consequently, the results of this assay may not be valid as it is difficult to determine if there were any kinase left in the reaction to function properly. Second, the discrepancy observed between the duplicates show inconsistency especially since the duplicates were from the same reaction tube. In addition, compared to the activity of other kinases, the “activity” of ADCK3 or ADCK4 shown by the results are minuscule and may suggest that there, in fact, are no activity.

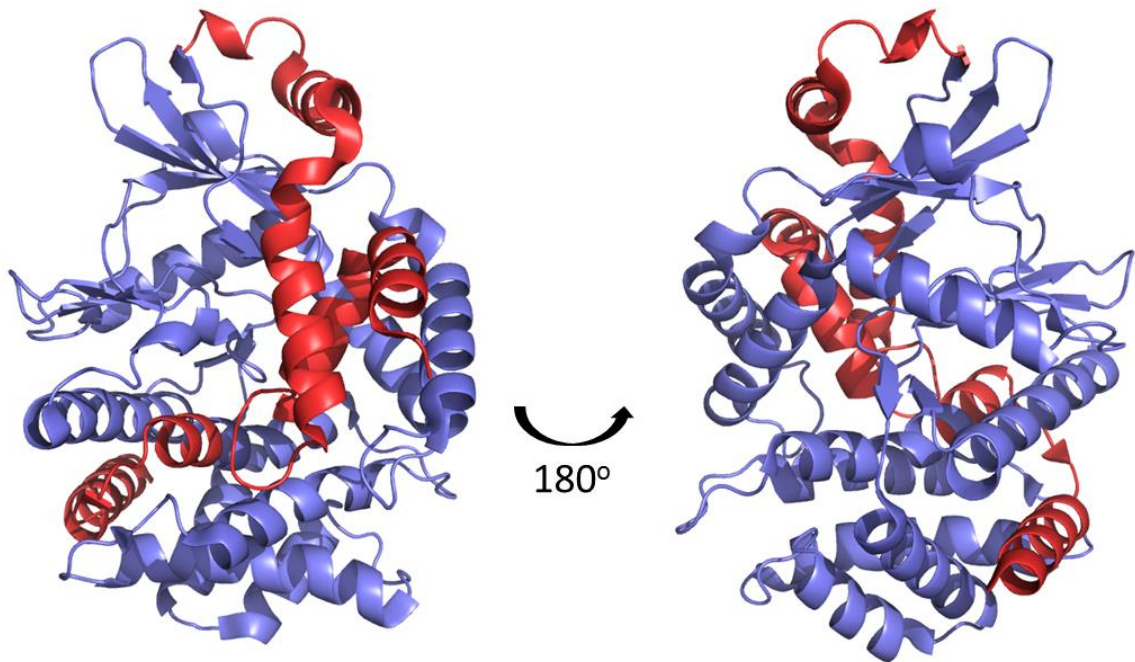


Figure A.5 The Structure of ADCK3.

The structure of ADCK3 (258-647) (PDB:4PED) is shown with residues 258-328 and 614-644 highlighted in red. The residues highlighted in red are not included in the 25HT29 construct used for the kinase activity assay.

A.4

Conclusion

Although the F₀S8 protein was successfully expressed as a MBP fusion construct, the purification method was unsuccessful in separate the fusion construct from the construct that was incompletely expressed. It may be difficult to separate the two proteins because the fusion construct is only longer by 41 amino acids, which corresponds to the increase in the molecular weight by 4864 Da. Furthermore, the results of the kinase activity assay performed with ADCK3 and ADCK4 showed incomprehensible results which may have been caused by the precipitation of the kinase proteins. In order to assess the kinase activity of ADCK3 and ADCK4, a more soluble construct must be cloned for its expression and purification. Thus, ADCK3 (245-647) and ADCK3 (329-647) were cloned to be used for the subsequent experiments.