UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

Redox Regulation of CLIC1 and CLIC4 Protein-Protein Interactions

by

Olga Maria Magdalena (Mia) Færch

(563776)

Thesis

Submitted in fulfilment of the requirements for the degree

Philosophiae Doctor

in

Molecular and Cell Biology

in the Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa

Supervisor: Prof. H.W. Dirr

February 2021

Declaration

I declare that this thesis is my own, unaided work. It is being submitted for the Degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

Signature  17 February 2021

Abstract

The anionic channels CLIC1 and CLIC4 appear soluble in the cytoplasm and are capable of reversible membrane insertion. The cellular redox state regulates the distribution and function of these proteins. This includes oxidation induced CLIC1 dimerisation, which augments membrane insertion, and *S*-nitrosylation of CLIC4 promoting nuclear import through the binding of a nuclear localisation sequence (NLS) to Impα. In the event of cellular oxidative stress, these processes exhibit a role in tumourigenesis and Aβ-induced neurotoxicity. This study aimed to assess the effect of redox regulation on both CLIC1 dimerisation, as well as whether the CLIC1 and CLIC4 NLSs can bind truncated hImpα isoform 1, omitting the IBB domain (hImpα1ΔIBB), in the event of oxidation induced exposure of the NLS regions. Reduced CLIC1 only appears as a monomer, whereas incubation with 2 mM H_2O_2 results in a mixture of dimeric and monomeric oxidised CLIC1 separable through SEC. Far-UV CD confirmed that all three redox species exhibited isolated α-helical coils. Fluorescence spectroscopy showed a 2 nm hypsochromic shift and 13 % decrease in emission intensity of the CLIC1 oxidised dimer compared to the monomeric forms. This indicates that the two monomeric forms display similar conformations, which differ from that of the oxidised dimer, where Trp35 appears more buried and quenched. Thermal unfolding indicated greater instability of the CLIC1 oxidised dimer compared to the monomeric forms. Finally, LC-MS/MS showed that Cys24 and Cys59 in CLIC1 form a stable intramolecular disulfide bond in the oxidised dimer, while undergoing significant disulfide bond shuffling in the monomeric forms. Peptides of the NLSs of CLIC1 (Pep1) and CLIC4 (Pep4) were designed. Investigating the interactions of Pep1, Pep1_A (Pep1 C/A mutant), Pep1_S (Pep1 C/S mutant) and Pep4 to ImpαΔIBB through molecular docking indicated that the core NLS region (**KKYR**) of these peptides form a similar binding pattern to both truncated mouse Impα isoform 1, omitting the IBB domain (mImpα1ΔIBB), and hImpα1ΔIBB. Fluorescence quenching demonstrated that Pep1_S ($K_d \approx 237$ μM) and Pep4 ($K_d \approx 317$ μM) bind hImpα1ΔIBB weakly. ITC confirmed the weak binding interaction between Pep4 and hImpα1ΔIBB ($K_d \approx 130$ μM), as well as the presence of a proton linked effect. In conclusion, CLIC1 dimerisation appears to stabilise the oxidative conformational changes and both the CLIC1 and CLIC4 NLSs likely bind a different hImpα isoform due to the weak binding detected toward hImpα1ΔIBB.

Aan my geliefde Moeder

Retha

Acknowledgements

I would like to express my gratitude to Prof Heini Dirr for his critical insights and guidance during my studies.

Thank you to my colleagues and friends at the PSFRU, in particular, Dr Sylvia Fanucchi, Kim Wiid, Ashleigh Blane and Roland Worth, for their advice and encouragement. A special thanks to Dr Stoyan Stoychev (Biosciences, CSIR, Pretoria) for his assistance and expertise with the mass spectrometry experiments.

I am also grateful to the National Research Foundation of South Africa and University of the Witwatersrand, Johannesburg, for financial assistance without which, this work would not have been possible.

Finally, my sincere gratitude to my family for their unconditional love and constant support through many sleepless nights. I am grateful to my husband, Cieron Færch, for his continual kindness, patience and love. My brother, Jan-Willem Steeb, who has always been a friend to me and my parents, Retha and Willi-Hans Steeb, for continually encouraging me to accomplish my goals.

"Research is what I'm doing when I don't know what I'm doing." Wernher von Braun

Research Outputs

Conference Presentations

1. University of the Witwatersrand postgraduate symposium, Johannesburg, RSA (2020). *Oral Presentation*. "Redox regulation of CLIC1 and CLIC4". **O.M.M. Færch** S.J.M. Steeb, S.H. Stoychev and H.W. Dirr
2. Molecular Biosciences Research Thrust research day, Johannesburg, RSA (2019). *Oral Presentation*. "Molecular Interactions between Importin-α and the Putative NLSs of CLIC1 and CLIC4". **O.M.M. Færch** and H.W. Dirr
3. University of the Witwatersrand postgraduate symposium, Johannesburg, RSA (2019). *Grad Flash Presentation*. "Molecular Interactions between Importin-α and the Putative NLSs of CLIC1 and CLIC4". **O.M.M. Færch** and H.W. Dirr
4. The Protein Society Symposium, Seattle, USA (2019). *Poster Presentation*. "Impact of Oxidative Stress on the Structural and Chemical Nature of Soluble CLIC1". **O.M.M. Færch**, S.J.M. Steeb, S.H. Stoychev and H.W. Dirr
5. Molecular Biosciences Research Thrust research day, Johannesburg, RSA (2018). *Poster Presentation*. "Impact of Oxidative Stress on the Structural and Chemical Nature of Soluble CLIC1". **O.M.M. Færch**, S.J.M. Steeb, S.H. Stoychev and H.W. Dirr
6. University of the Witwatersrand postgraduate symposium, Johannesburg, RSA (2018). *Poster Presentation*. "Traffic Control in Our Cells: Nuclear Import of CLIC1 via Importin-α". **O.M.M. Færch** and H.W. Dirr
7. EMBO Practical Course in "Quantitative proteomics: Strategies and tools to probe biology", Heidelberg, Germany (2018). *Poster Presentation*. "Impact of Oxidative Stress on the Structural and Chemical Nature of Soluble CLIC1". **O.M.M. Færch**, S.J.M. Steeb, S.H. Stoychev and H.W. Dirr
8. Molecular Biosciences Research Thrust research day, Johannesburg, RSA (2017). *Poster Presentation*. "Impact of Oxidative Stress on the Structural and Chemical Nature of Soluble CLIC1". **O.M.M. Steeb**, S.J.M. Steeb, S.H. Stoychev and H.W. Dirr
9. University of the Witwatersrand postgraduate symposium, Johannesburg, RSA (2017). *Poster Presentation*. "Impact of Oxidative Stress on the Structural and Chemical Nature of Soluble CLIC1". **O.M.M. Steeb**, S.J.M. Steeb, S.H. Stoychev and H.W. Dirr

TABLE OF CONTENTS

Abbreviations

Aβ	Amyloid-β
ADP	Adenosine diphosphate
ARM	Armadillo
ATP	Adenosine triphosphate
ATPase	Adenylpyrophosphatase
CAS	Cellular apoptosis susceptibility protein
CD	Circular dichroism
CFTR	Cystic fibrosis transmembrane-conductance regulator
CLC	Voltage-gated chloride channel
CLIC	Chloride intracellular channel
cNLS	Classical nuclear localisation sequence
DEAE	Diethylaminoethyl
*Dm*CLIC	*Drosophila melanogaster* CLIC
EXC4	*Caenorhabditis elegans* CLIC
GDP	Guanosine diphosphate
Grx	Glutaredoxin
GSH	Glutathione
GSSG	Glutathione disulfide
GST	Glutathione *S*-transferase
GTP	Guanosine triphosphate
HILIC	Hydrophilic interaction liquid chromatography
hImpα1ΔIBB	Truncated human Impα isoform 1 omitting the IBB domain
hnRNP	Heterogeneous nuclear ribonuclear protein
IAA	2-iodoacetamide
IAA94	Indanyloxyacetic acid 94
IBB	Importin β binding
Impα	Importin α
Impβ	Importin β
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
Kapβ	Karyopherin β
LC-MS/MS	Liquid chromatography (LC) tandem mass spectrometry (MS)

mImpα1ΔIBB	Truncated mouse Impα isoform 1 omitting the IBB domain
MRE	Mean residue ellipticity
ncNLS	Non-classical nuclear localisation sequence
NEM	*N*-ethylmaleimide
NES	Nuclear export signal
NLS	Nuclear localisation sequence
NPC	Nuclear pore complex
NTF2	Nuclear transport factor 2
Pep1	Putative CLIC1 NLS peptide
Pep1_A	Putative CLIC1 NLS Ala mutant peptide
Pep1_S	Putative CLIC1 NLS Ser mutant peptide
Pep4	Putative CLIC4 NLS peptide
PFT	Pore-forming toxin
PLIN	Lipid droplet-associated protein
PMSF	Phenylmethylsulfonyl fluoride
PPI	Protein-protein interaction
PTHrP	Parathyroid hormone-related protein
PTM	Post-translational modification
RanGAP	Ran GTPase-activating protein
RanGEF	Ran guanine nucleotide exchange factor
RMSD	Root mean square deviation
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
RP-HPLC	Reversed-phase high-performance liquid chromatography
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SV40	Simian virus 40
$T_{1/2}$	Apparent transition temperature
TCEP	Tris(2-carboxyethyl)phosphine
TMD	Transmembrane domain
TF	Transcription factor
WT	Wild-type

* Amino acids are named according to the IUPAC one and three letter codes

List of Figures

List of Tables

CHAPTER 1. INTRODUCTION

1.1 Protein-Protein Interactions

The interactome within biological systems refers to the complete set of molecular interactions originating from the organisation of modular components. This includes indirect interactions between components of the genome as well as physical interactions between components of the proteome (Sanchez *et al.*, 1999). When investigating the interactome in relation to the proteome, it becomes evident that proteins rarely act in isolation. Rather, groups of two or more polypeptide chains tend to associate with one another to assume quaternary structures. This involves the acquisition of specific physical and chemical contacts known as protein-protein interactions (PPIs). Consequently, multiprotein complexes are generated as required by a biological system or cell, to perform necessary functions. Formation of these complexes is generally tightly regulated to ensure the fine orchestration of biological systems as a whole (De Las Rivas and Fontanillo, 2010). Therefore, it is essential to explore the structural nature of PPIs in order to better understand the working mechanisms of the proteome within interaction networks.

1.1.1 Structures and Physiochemistry

Multiprotein complexes constituted of a single type of protein subunit are classified as homo-oligomeric, whereas a mixture of different protein subunits is classified as hetero-oligomeric (De Las Rivas and Fontanillo, 2010). These two types of complexes tend to exhibit different PPI properties which are essential in accommodating varying biological roles. Homo-oligomers generally take on multimeric states with stable interactions. This is especially seen in seemingly permanent cellular structures which are required to exhibit prolonged stability. An example of a homo-oligomeric complex is the actin microfilaments which act as molecular scaffolding within the cell. Additionally, many homodimers such as cytochrome *c'* display a two-fold symmetry which further limits inter-subunit relationships of a complex. These types of proteins rarely appear in a monomeric state within a cell. Certain hetero-oligomers may also exhibit similar properties to those of homo-oligomers. However, many hetero-oligomers consist of subunits capable of existing independently of one another, only forming non-obligatory contacts in response to the correct external influences. Structural domains within proteins which participate in these transient interactions facilitate molecular communication. This is particularly important in intra- and extracellular signalling events. In both cases of homo- and hetero-oligomeric complexes, subunits must present favourable interface properties to enable PPIs. Understanding interface properties constructed during PPIs provides greater insight into the functionality of the participating protein subunits. These properties include the overall size and shape of the interface areas, conformational changes in subunits as well as physiochemical complementarity between surface residues (Jones and Thornton, 1996).

When the subunits of a multiprotein complex come into contact with one another, certain highly specific physical contacts must form for complex association to occur. Such processes commonly necessitate conformational changes within the respective subunits to allow shape complementarity at the interface. These changes may take place in regard to individual amino acid side-chains, certain segments of the main-chain (e.g. hinged loops) or even entire domains. Depending on the conformational changes, complex interfaces may adopt a planar shape, a more globular shape or exhibit protruding domains. A compilation of the major domains involved in PPIs, in addition to the corresponding interacting components and examples, are given in Table 1 (Berridge, 2014). The extension of appropriate domains may assist in complex formation through the provision of amino acids displaying complementary shapes and charges. Docking between domains will either cause an increase or decrease in the overall flexibility of the complex, with the latter being associated with some form of energetic price to be paid (Jones and Thornton, 1997; De Las Rivas and Fontanillo, 2010).

Electrostatic complementarity at the interface can assist with ensuring that favourable energetics are obtained (Jones and Thornton, 1996). The composition of PPI interfaces generally represents that of the protein surface as oppose to the protein core. This is interesting, considering that these same interfaces tend to contain an increased amount of hydrophobic residues, especially aromatic residues (Yan *et al.*, 2008). Therefore, electrostatic forces may drive the assembly of protein subunits through the association of hydrophobic patches as well as polar interactions (Jones and Thornton, 1996; De Las Rivas and Fontanillo, 2010). Dependent upon the functional significance of a particular PPI within the cellular environment, this interaction may either be stable or transient in nature.

Stable interactions implicate long-lived contacts between protein subunits, in order to build permanent macromolecular cellular machines which, perform structural or functional roles. The size of such interfaces is large, containing buried surface areas typically greater than 2 500 $Å^2$. These types of interactions are prevalent in homo-oligomers, as is seen in the bacterial haem protein cytochrome *c'*. Certain hetero-oligomers may also display stable interactions, such as the adenosine triphosphate (ATP) synthase complex which consists of eight different subunits within mammals. Transient interactions contain weaker bonds capable of assembling and disassembling in far shorter time-scales than stable interactions (less than 1 sec). Furthermore, the size of such interfaces is smaller, with buried surface areas generally less than 2 000 $Å^2$. The reversibility of these brief interactions is crucial in regulating functionality according to the relevant cellular context and timing. This is particularly important in cellular machinery involved in biological cascades which necessitates continuous turnover and reassembly to initiate specific responses in a short amount of time. Examples of reversible interactions include the initiation of gene expression through the binding of transcription factors (TFs) and activators to promoter regions, as well as the activity of the mitotic checkpoint complex during the cell cycle (Pereira-Leal *et al.*, 2006; De Las Rivas and Fontanillo, 2010).

Table 1: Major protein domains identified to participate in PPIs
Interacting components within each domain are given as well as examples of proteins which utilise this particular domain for functionality.

Domain[1]	Interacting Components[1,2]	Example[1,3]
14-3-3	Identifies a phospho-Ser within a specific sequence	Transfer of phosphorylated TAZ out of the nucleus upon initiation of the hippo signalling pathway
Coiled-coil	Forms homo- or hetero-typic interactions with equivalent coiled-coil domains	Mediates the interaction of BRCA1 to Jun
Calponin homology	Around 100 amino acids arranged into four α-helices	Binding of parvin to actin or integrin-linked kinase
EH	Interacts with the Asn-Pro-Phe motif	Found in the scaffolding protein intersectin
FERM	Encompasses three compact units with basic residues that can interact with PtdIns4,5P_2	Present in proteins located on adhesion complexes like talin
Immunoreceptor Tyr-based activation motifs	Situated on the cytoplasmic domains of different receptors	Fyn phosphorylates this domain in Fc receptor γ chains to allow binding to phospholipase Cγ2
LIM	Comprised of a tandem Cys-rich Zn^{2+}-finger motif	Allows the adapter protein Lasp-1 to bind actin
PDZ	Interacts via a short peptide sequence containing a C-terminal hydrophobic residue	Permits certain scaffolding proteins to construct large macromolecular signalling complexes
Phospho-Tyr-binding	Binds a unique sequence with a phospho-Tyr group	Contained in signalling molecules such as the insulin receptor substrate
Sterile alpha motif	Around 70 amino acids arranged into five helices to form homo- or heterotypic interactions	Oligomerisation of Eph receptors into homodimers
Src homology 2	Identifies a phospho-Tyr within a specific sequence	Mediates the interaction of STATs TFs to Janus kinases
Src homology 3	Binds a poly-Pro motif	Enables the adapter protein Grb2 to interact with GEF SoS upon activation of Ras signalling
WW	Interacts with a Pro-rich sequence	Found in the N-terminus of the adapter protein Fe65

[1] Berridge (2014); [2] Korenbaum and Rivero (2002); [3] Hu and Rong (2002)

Interactions within a protein interface may further be classified as either covalent or non-covalent, the latter of which includes water-mediated interactions. Covalent interactions are the strongest and most stable form of associations and involve the sharing of electrons between adjacent residues from different subunits. Despite the rare occurrence of covalent interactions, these interactions play an important role in post-translational regulation. Examples of these include disulfide bond formation, ubiquitination and SUMOylation. Non-covalent interactions appear in abundance within protein interfaces and may exhibit varying degrees of stability over time. These weak bonds are important in transient PPIs and encompass hydrophobic interactions, Van der Waals forces, ionic interactions and hydrogen bonds (De Las Rivas and Fontanillo, 2010; Westermarck *et al.*, 2013). Given that cavities generally appear at protein interfaces, water molecules can occupy these spaces to ensure tight packing of the subunit residues. This is due to these water molecules constructing hydrogen bond bridges between distal charged groups. In some instances, two water molecules form doubly indirect interactions, which is especially evident in low affinity homologous complexes. Water-mediated interactions may therefore affect the affinity and specificity of complex subunits, contribute to the overall energy of the interface and consequently stabilise cross-recognition between subunits. Although interface water molecules experience constant exchange with the environment, complex cavities maintain the necessary water-mediated interactions. Based on an analysis of various protein complex crystal structures, many water-mediated interactions appear to exhibit conservation between homologous complexes (Janin, 1999).

1.1.2 Cellular Regulation and Significance

Overall cellular and systematic homeostasis relies on the interconnected roles of both singular proteins as well as modular supramolecular complexes coalesced by PPIs. While some proteins display competence in functioning independently, many require the assistance of others to fulfil essential roles. Additionally, proteins may display affinity towards a number of interacting partners and hence are able to participate in a variety of different complexes. This diversity in complex formation further enables a vast array of biological processes to occur from a set list of protein subunits, thus allowing cells to perform many different functions without the necessity to encode for and translate completely unique proteins for each individual task. Furthermore, each polypeptide domain within a complex can undertake independent and discrete biological functions. Given the intricate nature of multiprotein complexes, the assembly and subsequent activity of molecular machines thus require tight regulation at both cellular and systematic levels. This is to ensure that the initiation, progression and termination of required molecular processes take place as determined by the current cellular context. Regulation is achieved through the directing influence of environmental signals on the dynamic physiochemical connections within multiprotein complexes. Such actions may either lead to the activation or inhibition of one or more of the protein subunits within a complex (De Las Rivas and Fontanillo, 2010).

The array of signals distributed within a cell is dependent upon several temporal and spatial components. Environmental factors impacting protein complex organisation include the developmental stage of the organism, tissue type, cell type, subcellular localisation, cellular compartmentalisation, cell nutritional status as well as the current cell cycle stage. The mechanisms in which these factors regulate complex formation will now be discussed. A major driving force in PPIs encompasses both the amount as well as physiochemical modifications of the complex subunits. Protein concentration can be monitored on various levels of synthesis, all the way from gene expression levels to quality control steps, ensuring correct folding, and finally protein degradation rates. Additionally, certain proteins require particular post-translational modifications (PTMs), such as phosphorylation, as well as satisfactory pH levels and ion composition to accommodate favourable PPIs. Another deterministic element important in certain pathways is the amount of competitive and inhibitory interacting partners, as well as the affinity of these to protein subunits. Examples include other protein subunits, nucleic acids, ligands, ions and co-factors. Compartmentalisation and high concentrations of the interacting units are essential in bringing these units into close proximity. This in turn facilitates the speed and selectivity of associations which then results in greater cellular efficiency (De Las Rivas and Fontanillo, 2010). With the assistance of regulatory factors mediating PPI formation, the cell is able to transduce signals, perform muscle contractions and metabolise molecular components in addition to transporting these components across membranes. Conversely, the misregulation of such factors could induce the onset of certain pathological conditions.

Signal Transduction

Signal transduction entails the use of signalling molecules to relay extra- and intracellular signals along and/or inside the interior of cells. Cellular signalling proteins thus participate in transient interactions to facilitate multiple sequential dynamic associations and dissociations between the upstream and downstream proteins of a particular pathway. PPIs involved in signalling cascades include the use of PTMs, upstream and downstream effectors as well as adapter and scaffolding proteins. Many signalling proteins display enzymatic activity and hence are capable of altering substrates via PTMs such as phosphorylation, acetylation and hydroxylation. These effects are either rendered by modifying the chemical nature of amino acids, thus creating or eliminating PPIs, or altering the secondary and tertiary structural folds of domains to expose or shelter these domains. Examples of signalling proteins employing PTMs include 14-3-3 proteins, which identify a phospho-Ser within a specific sequence, and the binding of ubiquitin ligase VHL to the hydroxylated TF HIF-1a. An example of where multiprotein complexes utilise effectors is signalling networks that involve the propagation of growth factor signalling through growth factor receptors. Finally, structural scaffolding and adapter proteins aid in signal transduction by facilitating adequate spatial concentrations of network components in addition to phosphorylation reactions between interacting kinases. Examples include cytoskeletal proteins as well as components involved in ERK pathway activity (Westermarck *et al.*, 2013).

Muscle Contraction

The physiology of muscle contraction involves the adjacent sliding actions between two groups of interdigitating protein filaments, namely the thick myosin filaments and thin actin filaments. The commonly acknowledged cross-bridge hypothesis on muscle contraction dictates that sliding occurs when cross-bridges extending from the myosin filaments towards the actin filaments interact cyclically in a rowing motion. During this process ATP undergoes hydrolysis, thereby allowing the myosin filaments to act as molecular motors during the sliding motion (Rayment *et al.*, 1993). In conjunction with the interactions between myosin and actin filaments, there are numerous other PPIs that play a role in muscle contraction. One such PPI comprises the regulation of lipolysis on the lipid droplet surface of skeletal muscles by lipid droplet-associated proteins (PLINs). Upon lipolytic stimulation in adipocytes, PLIN1 undergoes phosphorylation to release CGI-58 and thereby fully activate ATGL which continues to initiate triglyceride breakdown (MacPherson *et al.*, 2013).

Metabolism and Membrane Transport

The processes involved within the cellular metabolic network encompass the interactions between enzymes in order to synthesise and break down both small molecules and macromolecules. This includes enzyme catalysed reactions within the glycolysis pathway and Krebs cycle. Furthermore, many protein complexes involved in cell metabolism may enter into cellular membranes to enable the passage of ions, small molecules and larger biomolecules, such as proteins, between these membranes. This is essential due to the fact that while small hydrophobic molecules can readily pass through the lipid bilayer along the concentration gradient, charged or larger hydrophobic molecules require a protein channel pore to cross (Hille, 1978; Yeagle, 1989). Additionally, particles that need to move against an already established concentration or voltage gradient of a cell membrane can be facilitated across this membrane by energetically induced transport proteins (Hille, 1978). Even though some transport protein channels contain innate pores within individual subunits, many require the oligomerisation of several different subunits via PPIs to form functional pore complexes (Dutzler *et al.*, 2002).

A prominent example of membrane transporter proteins involved in cell metabolism is adenylpyrophosphatase (ATPase) complexes. This class of enzymes can be sub-divided into four major groups, three which exhibit a rotary motor structure for ion transport and one that is distinct in both subunit composition and the types of ions transported by this group. Ion transport through ATPases is coupled to both the synthesis and/or hydrolysis of ATP. The synthesis of ATP requires the driving force of the H^+ gradient, whereas ATP hydrolysis causes dephosphorylation of the compound into adenosine diphosphate (ADP) and a free phosphate ion. The energy released during this decomposition reaction may then be utilised by the enzymes to energise other biochemical reactions such as ion transport (Pedersen and Carafoli, 1987).

Misregulation of PPIs

While the appropriate regulation of PPIs enables correct cellular functionality, any anomalies in this regard will result in the advent of pathogenic states. This is either due to the promotion of signalling pathways that continues to endorse a particular disease state, or the induction of protein aggregation which is detrimental to the cell. The first instance is prevalent in cancer, where oncogenic signalling cues within cell cycle molecular networks promote tumorigenesis, tumour progression, invasion, and/or metastasis. Therefore, many anticancer strategies strive to develop novel therapeutics which would specifically target these pathways in an attempt to disrupt signal transmission and hence perturb disease progression (Ivanov *et al.*, 2013). Many neurodegenerative disorders arise from the misfolding and continual aggregation of particular proteins, which subsequently lead to the formation of cytotoxic inclusion bodies within the nervous system. These types of disorders include Alzheimer's disease, Parkinson's disease, Huntington's disease, Creutzfeldt–Jakob's disease, frontotemporal lobar degeneration and motoneuron diseases. Each of these disease types can be categorised based upon the shape, size, location and composition of the prevalent protein aggregates. Once again, these particular PPI characteristics associated with these diseases, such as allosteric sites and hotspots, provide promising putative targets for therapeutic development (Jellinger, 2012). Drugs already available on the market that have proven to be effective inhibitors against PPIs, include the anti-HIV drug Maraviroc, which inhibits the CCR5-gp120 interaction, and the cardiovascular drug Titrobifan, which inhibits the glycoprotein IIb/IIIa (Ivanov *et al.*, 2013).

1.2 Intracellular Chloride Channels

Proteins that penetrate and become permanently solubilised within a membrane lipid bilayer are known as integral membrane proteins. These proteins either partially enter the membrane to span across a single lipid leaflet or completely insert into both lipid leaflets. Integral membrane proteins are present throughout all membranes, with localisation depending on the functional needs of the cell. Functions, which often entails the use of PPIs, include cell or organelle adhesion, signal reception, enzyme catalysis and transmembrane transportation (Singer and Nicolson, 1972; Almén *et al.*, 2009). Membrane transporters are transmembrane proteins with a pore that stretches across the entire membrane to regulate the passage of certain molecules. These proteins either act through facilitated diffusion or active transport. Facilitated diffusion permits particles to freely pass through the channels, based on an established concentration or voltage gradient. On the other hand, active transport utilises an added source of energy to push particles against these established gradients. Transporter selectivity arises from exhibiting an internal pore environment favourable to the chemical composition of the molecules of interest. This property allows particular transporters to passage polar and charged ions across the hydrophobic membrane interior (Hille, 1978; Almén *et al.*, 2009).

Intracellular chloride channels are ubiquitously expressed anion transporters. Anion transport impacts on the maintenance of cellular ion homeostasis which in turn influences intracellular pH levels and electrical excitability necessary for cell viability. Thus, by transporting anions across organelle membranes, intracellular chloride channels prompt the electrogenic exchange of cations between the organelle and cytoplasm. This ionic counterbalance system generates a more acidic pH within the organelle than the cytoplasm. Internal acidity is important for organelles, such as the *trans*-Golgi network and endosomes, to correctly process proteins targeted towards membrane insertion or secretion. The varying distribution of ions across membranes further causes a difference in membrane potential which facilitates electrical signalling. Intracellular chloride channels have also been found to take responsibility in transporting *trans*-epithelial fluid. Varying orders of selectivity have been identified among these channels, ranging from demonstrating non-selectivity to greater selectivity towards anions other than chloride (Suh and Yuspa, 2005).

Four main classes of intracellular chloride channels exist according to conformational diversity. These classes are ligand-gated chloride channels (extracellular and intracellular), voltage-gated chloride channels (CLCs), cystic fibrosis transmembrane-conductance regulators (CFTRs), and chloride intracellular channels (CLICs) (Suh and Yuspa, 2005). The precise mode in which each of these channel classes associates with and transports anions is dependent upon the structural conformation and oligomerisation assumed by the individual protein subunits upon membrane association. While some channels contain inherent pores within individual subunits, others require the oligomerisation of several subunits to form a functional pore (Dutzler *et al.*, 2002).

1.2.1 CLIC Protein Family

The CLIC family is a class of intracellular chloride channels highly conserved within chordates. Six distinct vertebrate paralogues exist, identified as CLIC1-6, which generally encode for the expression of seven different proteins. The CLIC5 gene is susceptible to alternate splicing, resulting in the production of either the CLIC5A or CLIC5B protein as required. Splice variance may also be displayed in the CLIC2 and CLIC6 genes. Some deviation in the number of paralogues does, however, occur within different animals. For example, birds and lizards omit CLIC1, whereas teleost fish contain an additional CLIC5 designated as CLIC5L. Other metazoans likewise possess CLIC-like orthologues exhibiting related functional roles, such as *Dm*CLIC in *D. melanogaster* and EXC4 in *C. elegans* (Littler *et al.*, 2010). All CLICs exhibit a molecular mass ranging between 30-40 kDa and contain a single N-terminal transmembrane domain (TMD) necessary for membrane insertion and subsequent anion conductance (Suh and Yuspa, 2005). The structure of the CLIC1 protein (Figure 1) is provided as a representative structural model of the CLIC family. Figure 2 displays the sequence homology between CLIC1-6, *Dm*CLIC and EXC4, as well as the TMD region for each of the respective CLIC species.

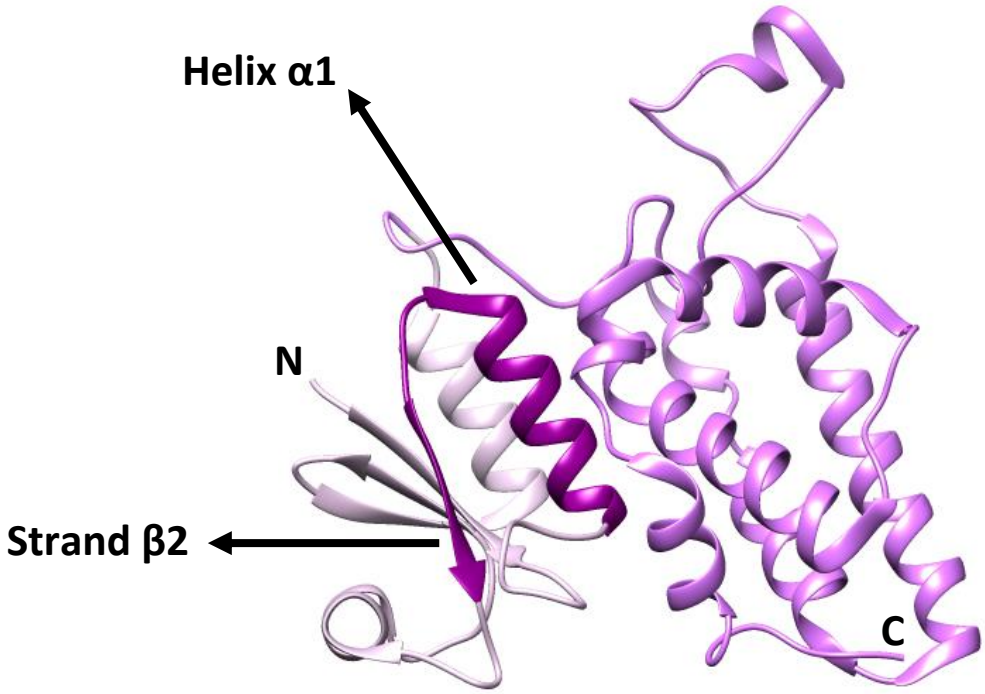


Figure 1: Structural features of vertebrate CLIC1
The TMD (dark purple) is a well conserved structural motif among the CLIC protein family members which emphasises the importance of this structure in membrane insertion and subsequent anion channel functionality. This domain, which reaches from residue Cys24 to Val46 in CLIC1, is situated within the N-terminal domain (lilac) and encompasses most of helix α1 and strand β2. The C-terminal domain (purple) consists wholly of α-helices. PDB file 1K0M was used to generate the cartoon figure using the UCSF Chimera Molecular Graphics System, V1.12.

The N-terminal domain of CLIC family members differs in sequence and overall length (Figure 2). This is particularly true for CLIC5B and CLIC6 which contain an additional N-terminal domain with poorly conserved repetitive sequences. Despite N-terminal variation, CLICs exhibit a thioredoxin fold, which places CLICs within the glutathione *S*-transferase (GST) fold superfamily, and a highly conserved single TMD which plays a fundamental role in the membrane insertion of these proteins (Edwards and Kahl, 2010). While most membrane proteins undergo direct export through the endoplasmic reticulum secretory pathway, CLICs lack a membrane signal sequence and rather interchange between a soluble cytoplasmic globular form and a functional membrane-integrated ion channel. The exact mode of cellular localisation, domain unfolding and membrane auto-insertion of soluble CLICs is currently not entirely known (Suh and Yuspa, 2005; Littler *et al.*, 2010). However, since CLICs display no apparent innate pore structure, the TMD most likely requires the oligomerisation of several domains to form a functional channel. One proposal suggests that CLICs may form a large oligomeric species comprising of approximately six to eight subunits upon membrane insertion (Goodchild *et al.*, 2011). Though, further investigation is necessary to establish the exact number of participating subunits.

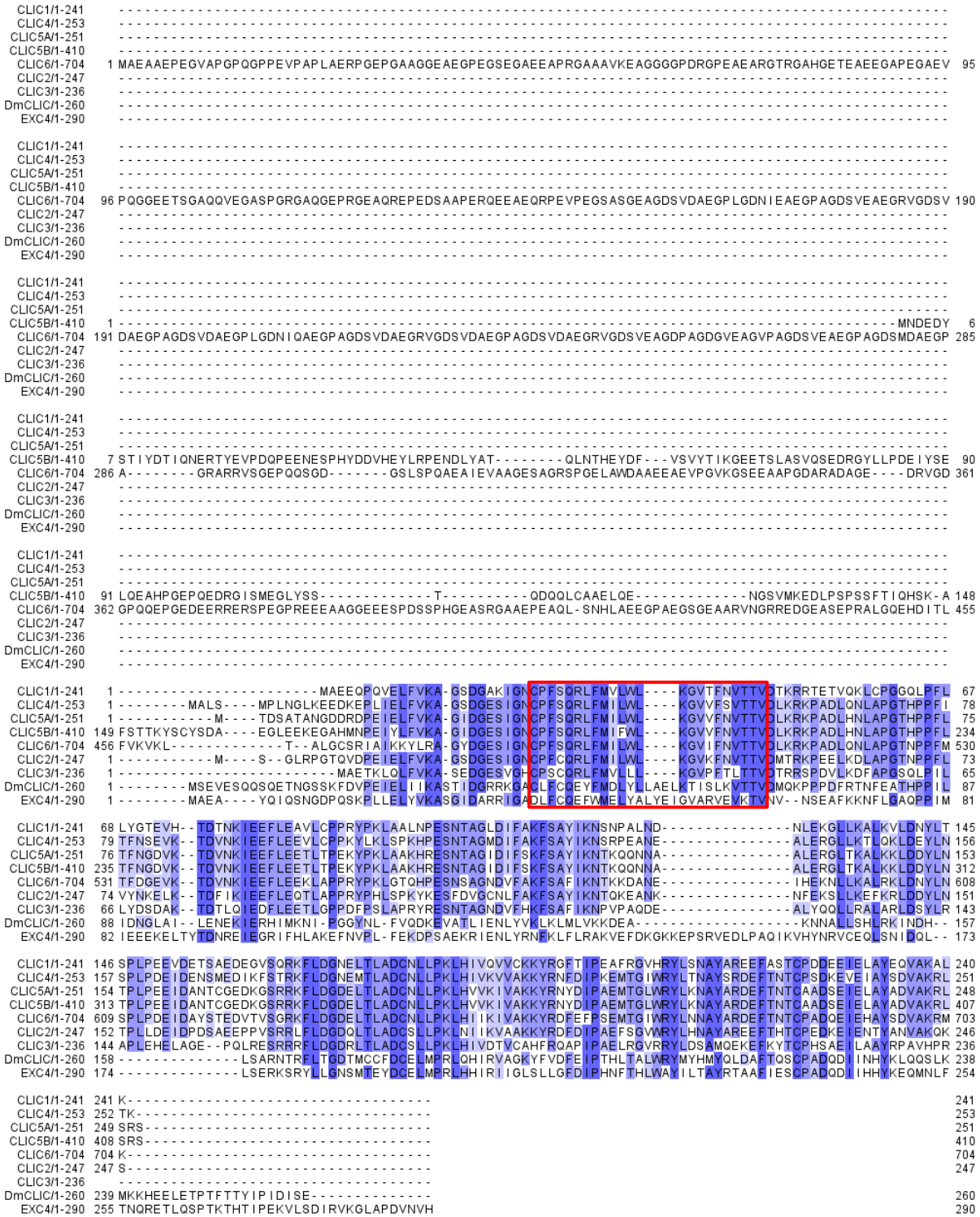


Figure 2: Structure-based sequence alignment of related CLICs

High preservation appears across all seven vertebrate CLICs as well as the invertebrate homologues present in *D. melanogaster* and *C. elegans*. The TMD sequence (red box) which is comprised of residues Cys24-Val46 in CLIC1 displays strong conservation. The sequence alignment was performed using Clustal Omega (Madeira *et al.*, 2019) and formatted with Jalview 2.11.0 to show percentage identity (Waterhouse *et al.*, 2009).

For CLICs to transform from a soluble cytoplasmic form to an oligomeric membrane channel, significant conformational changes around the TMDs would have to transpire. An example of this is CLIC1, where the N-terminus undergoes significant structural rearrangement to permit membrane insertion. During this process four β-strands dissolve while helix α1 stretches out to the C-terminus by two α-helical turns. As a result, two proximal Cys residues which may be disulfide bonded (Cys24 and Cys59), separate from one another. Cys24 migrates towards the *trans*-face of the membrane, while Cys59 remains on the *cis*-face. These modifications seem to suggest that CLICs experience energetically favourable structural rearrangements at the N-terminus to facilitate membrane insertion during channel formation (Singh, 2010). The unusual dimorphic nature of CLICs leads to this family being included in the same class as pore-forming toxins (PFTs). PFTs are generally cytotoxic bacterial proteins which undergo extensive conformational changes to produce unregulated pores within target cell membranes. Both CLICs and PFTs require tight transitional regulation between the soluble and membrane-bound forms of the proteins to elicit the desired functions (Edwards and Kahl, 2010; Littler *et al.*, 2010; Singh, 2010).

The C-terminus of CLICs are generally well conserved with a proximal 220 amino acid predominantly α-helical domain demonstrating further structural homology to the GST superfamily (Edwards and Kahl, 2010). The homological relationship of this highly conserved functional region poses significance when predicting additional functional roles of the CLIC family. Phylogenetic studies have determined that CLICs exhibit probable enzymatic function, which likely includes a glutathione (GSH)-related co-factor (Littler *et al.*, 2010).

1.2.2 Expression and Subcellular Distribution

CLIC family members display differential expression in numerous cell types with certain regulatory factors further directing the subcellular distribution of these proteins. Final cellular localisation corresponds to the functional role these proteins are required to perform (Suh and Yuspa, 2005; Singh, 2010). Similar to other chloride intracellular channel classes, CLICs demonstrate non-channel roles in addition to anion-selective membrane channelling. The dimorphic nature of CLICs allows for the multi-functionality of this protein family by permitting the transition between membrane channels and soluble proteins capable of adhering to structural and signalling molecules (Edwards and Kahl, 2010; Littler *et al.*, 2010; Nguyen *et al.*, 2016). As a result, CLICs show involvement in directing the cell cycle, differentiation, and motility as well as intracellular scaffolding, cell signalling, p53-mediated apoptosis, skeletal muscle and brain formation, tubulogenesis, angiogenesis, bone resorption and amyloid-β (Aβ) induced neurotoxicity (Edwards and Kahl, 2010; Littler *et al.*, 2010; Singh, 2010). An overview of the expression and subcellular distribution of each vertebrate CLIC family member, in addition to corresponding functions and probable disease linkage, are given in Table 2.

Table 2: Localisation and functionality of vertebrate CLIC family members
The molecular mass, tissue expression and subcellular distribution for each family member are given in addition to known associated functions and probable disease linkage.

Protein	Molecular Mass (kDa)[1,2]	Tissue Expression[2,3]	Subcellular Distribution[1,2,3]	Function[2,3,4]	Disease[2,3,5,6]
CLIC1	27	Ubiquitous with low expression in brain, skeletal muscle	Nuclear membrane, nucleoplasm, intracellular membranes, plasma membrane, cytoplasm	Ion channel, cell cycle, Aβ induced neurotoxicity, osteoblast differentiation	Cancer, Alzheimer's disease
CLIC2	28	Fetal liver, adult skeletal muscle	Interact with Sedlin, not widely substantiated	Modulates activity of RyR	Unknown
CLIC3	24	Placenta, pancreas, kidney, lung, heart, low amounts in skeletal muscle	Primarily nucleus, plasma membrane	Ion channel, cell growth	Unknown
CLIC4	29	Placenta, testis, kidney, liver, lung, heart, brain, skeletal muscle	Nucleus, mitochondria, *trans*-Golgi network, ER, large dense-core secretory vesicles, plasma membrane, intracellular membranes, cytosol, cytoskeleton	Ion channel, acidification, cell division, angiogenesis, apoptosis	Cancer
CLIC5A	28	Placenta, kidney, lung, heart, skeletal muscles	Primarily cytoplasm, cytoskeleton associated	Ion channel, coupling to the cytoskeleton, stimulate actin polymerisation	Unknown
CLIC5B	49	Kidney cortex, heart, skeletal muscles	Intracellular membranes, secretory vesicles, plasma membrane, cytoplasm	Ion channel, ion absorption and secretion, formation of stereocilia, development of the organ of corti, bone resorption	Unknown
CLIC6	71 (two isoforms)	Placenta, pancreas, kidney, choroid plexus, gastric mucosa, liver, brain	Plasma membrane, primarily cytoplasm	Ion channel, water transport, interaction with dopamine receptors	Unknown

[1] Suh *et al.* (2005); [2] Suh and Yuspa (2005); [3] Singh (2010); [4] Edwards and Kahl (2010); [5] Peretti *et al.* (2015); [6] Nguyen *et al.* (2016)

While the exact mode of CLIC localisation, unfolding, oligomerisation and membrane insertion remains unclear, certain environmental regulators may either stimulate non-channel roles or functional membrane channel formation. These regulators can either exist as intracellular environmental changes or external stimuli. Environmental regulators identified to influence CLIC expression, subcellular distribution and function include phosphorylation events, lipid content, membrane polarisation, pH levels and cytoplasmic oxidation (Suh and Yuspa, 2005; Singh and Ashley, 2006; Edwards and Kahl, 2010; Singh, 2010). Ultimately, the translocation and functional response of a CLIC protein caused by a particular regulator depends on the CLIC family member involved as well as the cell type in which the protein resides (Peretti *et al.*, 2015).

Consensus phosphorylation sites previously identified to exist within the CLIC protein family include casein kinase II, protein kinase A and C, Tyr phosphorylation and src-homology domain type-2 binding domain. The phosphorylation of CLICs may either trigger membrane insertion and subsequent anion transport activity, or the modulation of phosphorylation pathways. When modulating phosphorylation pathways, CLICs may either directly phosphorylate substrates or indirectly associate with kinases, phosphatases or other scaffolding proteins binding to the kinases or phosphatases (Suh and Yuspa, 2005; Edwards and Kahl, 2010; Singh, 2010).

CLIC channel formation, insertion and function also demonstrate dependency on overall membrane lipid composition and the resultant membrane polarity. Interestingly, CLIC1 monomers display the potential to insert into membranes exhibiting different lipid compositions (Singh and Ashley, 2006). However, the construction of functional channels only takes place in distinct planar bilayers constituted of phospholipids phosphatidyl ethanolamine and phosphatidyl serine, as well as sterol cholesterol, in a ratio of 4:1:1 (mol/mol) (Singh and Ashley, 2006; Valenzuela *et al.*, 2013).

Membrane polarisation and acidification regulate anion transport by CLICs in order to either maintain a resting potential or induce excitability in neuronal cells. ATPases generate an internal proton gradient to establish a positive electrical potential which in turn promotes acidification. While numerous cellular functions rely on membrane polarisation and acidification, the resultant large positive electrical charge hinders any further proton transport or pH reduction (Suh and Yuspa, 2005; Edwards and Kahl, 2010). To counterbalance this, the more acidic internal pH stimulates less stability and greater flexibility in CLICs to allow these proteins to reach a membrane-competent pre-insertion form (Fanucchi *et al.*, 2008; Stoychev *et al.*, 2009). Upon membrane insertion and channel formation, the CLICs offset the high electrical charge by facilitating the passive diffusion of anions into the internal environment. As a result, ATPases can transport more protons to upkeep the optimal cellular pH (Suh and Yuspa, 2005; Edwards and Kahl, 2010).

1.2.3 Redox Regulation of CLICs

Another factor influencing CLIC distribution and functionality is the redox state of the cell. Eukaryotic cells employ membranes to create defined compartmentalised environmental conditions to facilitate particular functions. This allows organelles and the external environment to maintain distinct non-equilibrium redox potentials for central thiol/disulfide redox couples, thioredoxin-1, thioredoxin-2, GSH/glutathione disulfide (GSSG) and Cys/CySS. In the case of the cytoplasm, the GSH/GSSG redox potential experiences vigilant regulation according to functional requirements. The cytoplasm also acts as a buffer for anomalous reactive oxygen species (ROS) production between the various compartments due to the general lack of metabolic oxidases in the cytoplasm. The low cytoplasmic levels of ROS and reactive nitrogen species (RNS) further allow these species to act as secondary messengers dependent on NADPH oxidases and NO synthases. These chemically reactive species relay redox information between cell compartments by using strong oxidising potentials to modify redox-sensitive protein motifs, such as Cys residues and metal co-factors. As a result, ROS and RNS can regulate compartmental oxidation conditions which in turn impact redox state altering processes such as cell replication (Go and Jones, 2008; Averaimo *et al.*, 2010; Baxter *et al.*, 2014; Huang and Sikes, 2014).

As aforementioned, CLICs fall within the larger GST superfamily which consists of stable globular proteins exhibiting enzymatic functionality. Structurally, GSTs are composed of two distinct moieties which are an N-terminal thioredoxin fold and a wholly α-helical C-terminal domain. Thioredoxin folds are present within numerous large redox-active protein families and catalyse disulfide bond formation and isomerisation. These folds centre upon a Cys redox-active site which generally is the first residue of a CXXC motif in an N-terminal α-helix. The GST Omega class contains an active site similar to that in glutaredoxins with GSH-binding sites that aid in cellular redox balance upkeep. This glutaredoxin (Grx)-like active site in Omega GSTs is not found in any other classic GSTs and displays a motif of CP(F/Y)(C/S). Given the structural homology between CLICs and Omega GSTs, vertebrate CLIC proteins display a similar motif at the N-terminus of helix α1 which is CP(F/S)(S/C) and contains a GSH-binding site which may permit subcellular localisation and GSH-dependent redox control (Littler *et al.*, 2010; Nguyen *et al.*, 2016).

Of the three conserved Cys residues in CLIC1 (positions 24, 178 and 223), the Grx-like active site is centred on the highly conserved Cys24. Cys24 functions as the redox-active site residue within the TMD and is localised within close proximity to the pore region of the inserted channel. A single point mutation of Cys24 to an Ala residue disrupts the electrophysiological characteristics of the channels and hence reduces conductance. However, the loss of Cys24 only results in the abrogation of *trans*-redox sensitivity and not the absolute inhibition of channel formation and anion transport (Littler *et al.*, 2010; Singh, 2010; Nguyen *et al.*, 2016).

Given that the GSH containing cytoplasmic environment is reducing and that Cys residues predominantly display a pK_a of 8.0, all CLIC Cys residues probably remain protonated under physiological conditions (Cumming *et al.*, 2004; Nguyen *et al.*, 2016). However, high ROS concentrations induced by oxidative stress or insertion of the proteins into certain intracellular organelles may modify these redox-sensitive residues (Cumming *et al.*, 2004; Singh and Ashley, 2006; Go and Jones, 2008). The specific modes of residue alterations include oxidation, nitrosylation, Met sulfoxide formation and carbonylation (Phang, 2012). Oxidative modifications of some CLICs have previously been shown to chemically and structurally alter these proteins and as such affect overall cellular activity (Singh and Ashley, 2006; Averaimo *et al.*, 2010; Peretti *et al.*, 2015; Nguyen *et al.*, 2016).

Oxidative stress occurs when the rate of ROS production surpasses the cellular capacity to detoxify reactive intermediates or mend the resultant damage. Such an imbalance causes the generated ROS to non-specifically damage crucial cellular components such as DNA, lipids and proteins. Examples include superoxide radicals ($O^{\cdot}{}_2^-$), hydroxyl radicals ($\cdot OH$), peroxyl radicals ($ROO^{\cdot}$), and alkoxyl radicals ($RO^{\cdot}$) as well as some non-radical compounds such as hydrogen peroxide (H_2O_2), the singlet oxygen ($^1O_2$), ozone (O_3), and hypochlorous acid ($HOCl^-$) (Romero-Puertas and Sandalio, 2016). One ROS commonly present within cells is H_2O_2. Under normal physiological conditions, H_2O_2 concentrations range between 10–400 μM. However, high H_2O_2 concentrations ranging from 1-10 mM may occur within cells with comparable levels being generated through the acute oxidative burst in neutrophils while phagocytosing bacteria (Cumming *et al.*, 2004; Averaimo *et al.*, 2010; Baxter *et al.*, 2014). In the occurrence of a redox shift towards severe oxidative stress conditions, cells display overexpression of CLIC1, subsequently leading to hyperactivity of the protein.

Under strongly oxidising conditions, Cys24 and Cys59 of CLIC1 form an intramolecular disulfide bond that exposes a flat hydrophobic patch. CLIC1 subunits then dimerise through hydrophobic contacts to conceal the hydrophobic patches (Singh and Ashley, 2006; Nguyen *et al.*, 2016). However, this property appears unique to CLIC1 since Cys59 is not conserved among other CLICs. This disulfide bond in CLIC1 would also presumably inhibit membrane insertion since reconfiguration of the TMD during insertion propels Cys24 through to the *trans*-face while Cys59 remains on the *cis*-face (Singh and Ashley, 2006, Littler et al., 2010). Despite this, CLIC1 displays chronic functional expression under oxidative stress. Therefore, it would appear that while oxidation may not be the driving force towards membrane insertion, dimerisation of CLIC1 does not prevent the increase in membrane insertion and subsequent functionality during oxidative stress. Given the role that the CLICs play in cell cycle control and Aβ induced neurotoxicity, hyperactivation of the CLIC1 may result in tumourigenesis and the development of Alzheimer's disease (Singh and Ashley, 2006; Averaimo *et al.*, 2010; Peretti *et al.*, 2015; Nguyen *et al.*, 2016).

CLIC1 aids cell cycle control through the regulation of chloride conductance in response to oxidative fluctuations at specific stages. At cytoplasmic oxidative peaks, the protein channel displays transient expression on the plasma membrane. These events display prominence during the G2/M phase in which current density becomes doubled compared to the G1/S resting phase. Oxidative stress conditions would thus promote chloride permeability and conductance due to the chronic functional expression of membrane-inserted CLIC1. Consequently, the rate of cell division and anti-apoptosis signalling is altered to enable cellular transformation, tumour development, proliferation and metastasis. The upregulation of CLIC1 chloride conductance possibly also sustains ROS production by NADPH oxidase to further assist progression through the cell cycle during tumourigenesis. Moreover, specific CLIC1 blockers, such as indanyloxyacetic acid 94 (IAA94), appear to supress CLIC1 chloride conductance and hence reduce intracellular ROS production. This prevents the facilitation of cell volume regulation and inhibits cells in the G2/M stage, which subsequently impedes cell division during mitosis and diminishes migration and invasion (Suh and Yuspa, 2005; Averaimo *et al.*, 2010; Peretti *et al.*, 2015; Nguyen *et al.*, 2016).

Progression of Alzheimer's disease has also been linked to CLIC1 due to overexpression of the protein inducing microglial hyper activation. Microglial cells are central nervous system macrophages that protect against injury and intruding pathogens. Alzheimer's disease produces Aβ deposits which induce a neuroprotective response from the microglia. These cells then experience activation and undergo various morphological and electrophysiological alterations which promote active proliferation, migration, phagocytosis to eliminate intruders as well as increased production and secretion of neurotoxic molecules (Averaimo *et al.*, 2010; Nguyen *et al.*, 2016).

Microglia neurotoxins that assist in neurodegeneration include nitrites such as NO as well as $O^{\cdot}{}_2^-$ and other related ROS. Activated membrane NADPH oxidases begin by transferring electrons from NADPH to two extracellular O_2 in order to produce two $O^{\cdot}{}_2^-$, a $NADP^+$ and a proton. Superoxide dismutase (SOD) then catalyses the dismutation of two $O^{\cdot}{}_2^-$, thus first producing O_2 followed by the addition of two protons to create H_2O_2. Finally, catalase (CAT) decomposes two H_2O_2 molecules to generate two H_2O and an O_2 (Averaimo *et al.*, 2010; Romero-Puertas and Sandalio, 2016). The resultant heightening of oxidative stress by Aβ leads to CLIC1 overexpression in addition to altering the redox state of the CLIC1 Cys24 active site (Figure 3). This causes a metamorphic transition in the protein to form membrane channels, thereby increasing anion conductance. The overexpression of CLIC1 can then further assist Aβ induction of ROS generation in the hyperactivated microglia. Aβ may also directly upregulate channel activity by increasing the open probability and mean open time of CLIC1. Reduction in CLIC1 functionality by IAA94 can stimulate the phagocytosis of Aβ by the microglia, thus impairing neurotoxin production and secretion (Singh and Ashley, 2006; Averaimo *et al.*, 2010; Correani *et al.*, 2015; Nguyen *et al.*, 2016).

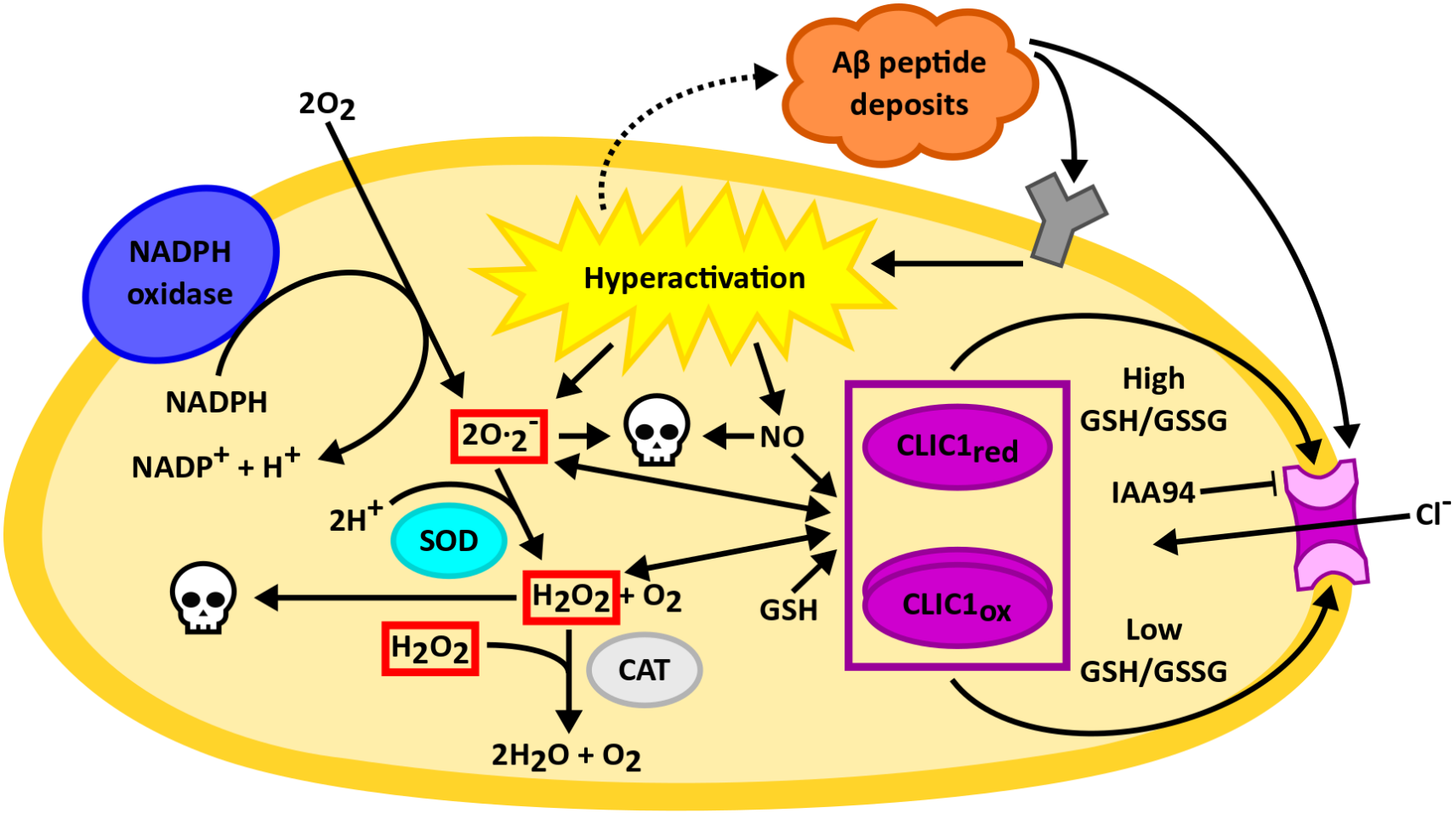


Figure 3: Redox regulation of CLIC1 structure and channel activity
Under physiological conditions, CLIC1 is reduced and mostly remains as a cytoplasmic monomer. Either reduced monomer or oxidised dimer inserts depending on the GSH/GSSG redox potential. CLIC1 oxidation by ROS (red boxes) or *S*-nitrosylation by NO may promote membrane insertion, whereas IAA94 blocks channel activity. Production of ROS from extracellular O_2 is catalysed by NADPH oxidase, SOD and CAT. During neurodegeneration, microglia recognise Aβ deposits bound to surface receptors and undergo hyperactivation to produce neurotoxins (ROS and NO), overexpress CLIC1 and eliminate deposits through phagocytosis. Aβ deposits and neurotoxins may also directly enable CLIC1 channel activity. An increase in CLIC1 activity further promotes ROS production.

1.2.4 Nuclear Import of CLIC1 and CLIC4

The redox state of the cell also seems to impact on CLIC4 structure and functionality with ROS and RNS most likely influencing nuclear translocation (Suh *et al.*, 2004). Previous studies show that *S*-nitrosylation of a CLIC4 Cys residue by NO enhances association to the nuclear import proteins importin α (Impα) and Ran, which in turn induce CLIC4 nuclear import. These associations most likely occur due to *S*-nitrosylation induced conformational changes resulting in CLIC4 destabilisation and unfolding. CLIC4 contains four Cys residues at positions 35 (corresponding to the highly conserved Cys24 in CLIC1), 100, 189 and 234. Despite the burial of the Cys234 thiol side-chain in an interdomain interface, *S*-nitrosylation of this Cys appears to trigger the structural change needed for CLIC4 nuclear import. Cys234 exists between helix 8 and helix 9 as well as close to the helical CLIC4 nuclear localisation sequence (NLS) (Figure 4A) (Malik *et al.*, 2010). An NLS is a signalling sequence which directs the import of cargo proteins to the nucleus via a superfamily of soluble transport receptor factors collectively called karyopherin β (Kapβ) (Mosammaparast and Pemberton, 2004; Lange *et al.*, 2007). Therefore, *S*-nitrosylation of Cys234 probably destabilises the helical CLIC4 NLS thus exposing the sequence to interact with the Kapβ family member Impα (Malik *et al.*, 2010).

The wild-type (WT) NLS established for CLIC4 is 199-KVVA**KKYR**-206 (NLS residues in bold and underlined) with mutations of the basic residues to 199-TVVAITYG-206 preventing nuclear import (Suh *et al.*, 2004). A crystal structure of mouse Impα isoform 1, omitting the autoinhibitory domain (mImpα1ΔIBB), bound to a peptide bearing the CLIC4 NLS has been determined (PDB 3OQS). This structure reveals that the NLS peptide binds mImpα1ΔIBB in an elongated conformation (Figure 4B). Therefore, the native helical region containing the NLS would need to unfold to facilitate binding and subsequent nuclear import. No interactions occur between the N-terminal KVV motif and mImpα1ΔIBB, implying that these residues play a small role in NLS recognition. Conversely, the basic residues in the **KKYR** motif act central to binding the acidic binding site of mImpα1ΔIBB. While the Tyr residue of the **KKYR** motif seems an unusual fit for the acidic binding site, this residue side-chain favourably forms hydrogen bonds with the mImpα1ΔIBB backbone (Mynott *et al.*, 2011).

The CLIC4 NLS exhibits high conservation in all vertebrate CLICs, with the exception of CLIC3 (Figure 4C). Interestingly, only CLIC1, CLIC3 and CLIC4 display presence in the cell nucleus (Table 2). Thus, CLIC1 probably follows a similar nuclear import pathway to CLIC4 where structural destabilisation of the protein through Cys *S*-nitrosylation is required. In the case of CLIC1, Cys59 and Cys89 have been identified as the two potential *S*-nitrosylation sites (Figure 4D) (Phang, 2012). Before determining whether *S*-nitrosylation of either Cys residue is involved in nuclear import, investigation is required to determine whether an elongated peptide bearing the putative CLIC1 NLS does in fact bind to the nuclear importer Impα.

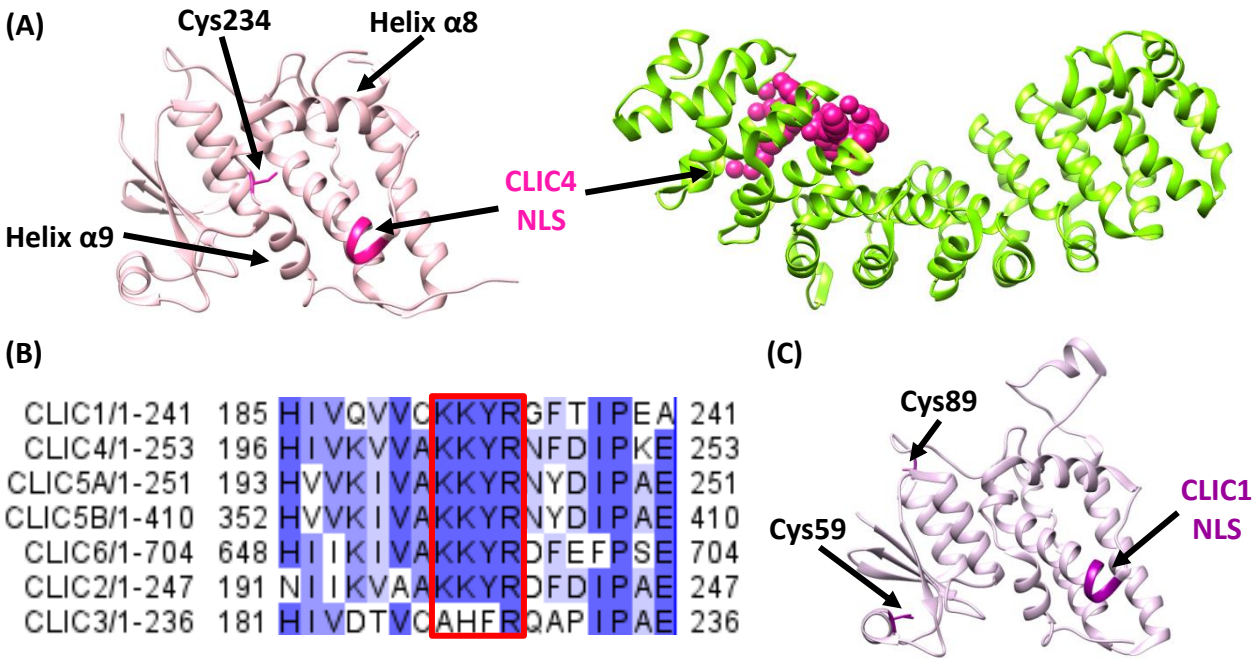


Figure 4: CLIC4 NLS binding and conservation in other CLICs
(A) The CLIC4 NLS (magenta) relative to Cys234 (PDB 2AHE) and binding of an elongated peptide bearing this NLS to mImpα1ΔIBB (light green) (PDB 3OQS). **(B)** The vertebrate CLIC NLS (red box) is high conserved with the exception of CLIC3. The sequence alignment was performed using Clustal Omega (Madeira *et al.*, 2019) and formatted with Jalview 2.11.0 to show percentage identity (Waterhouse *et al.*, 2009). **(C)** The putative NLS (dark purple) of CLIC1 relative to the two Cys residues capable of *S*-nitrosylation (PDB 1K0M). Cartoon protein figures were generated using the UCSF Chimera Molecular Graphics system, V1.12.

1.3 Karyopherins

Cellular structures within eukaryotic cells compartmentalise protein synthesis by transcribing DNA into RNA in the nucleus and then translating RNA into proteins via cytoplasmic ribosomes. However, some of these synthesised proteins (e.g. polymerases and histones) require re-entry into the nucleus to initiate functionality. Eukaryotic cells thus employ large tube-like protein assemblies to span the nuclear envelope double membrane to mediate nucleocytoplasmic transport. These macromolecular assemblies are known as nuclear pore complexes (NPCs) and range from approximately 60 MDa in yeast to 125 MDa in mammalian cells. While NPCs permit the passive diffusion of ions, small molecules as well as small proteins (< 20 kDa), these complexes only permit the trafficking of macromolecules containing apt signals. Cargo proteins may contain one of two types of signals which, when present, determine the directionality of nuclear translocation. An NLS encodes for nuclear import whereas a nuclear export sequence (NES) encodes for nuclear export. As aforementioned, the Kapβ superfamily is central in identifying these signals in order to mediate nucleocytoplasmic pathways. These proteins typically act as either import receptors (i.e. importins) or export receptors (i.e. exportins), while some may act as bidirectional transporters. Figure 5 provides a summary of the Kapβ proteins present in yeast (14) and mammalian cells (20) as well as the directionality of these transporters (Chook and Blobel, 2001; Mosammaparast and Pemberton, 2004; Lange *et al.*, 2007).

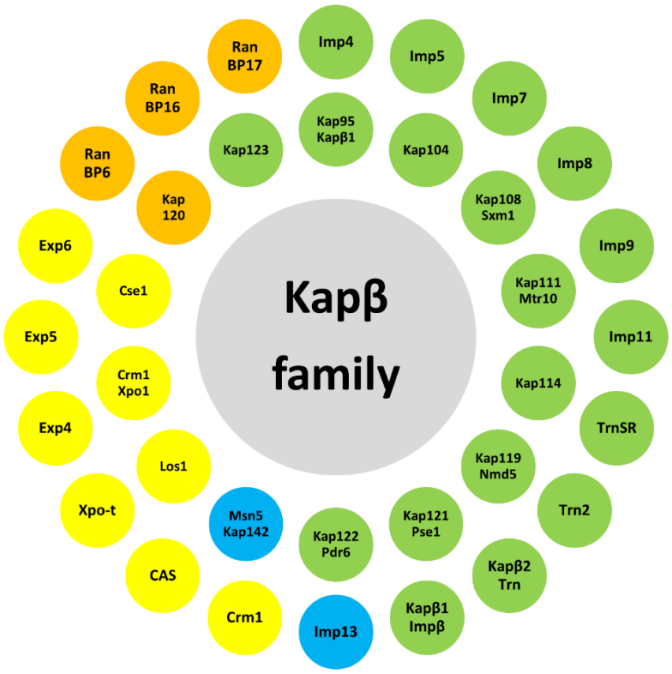


Figure 5: Kapβ family in yeast and mammalian cells
The inner ring of circles represents Kapβs in yeast while the outer ring represents those in mammalian cells with alternative names for certain Kapβs shown. Importins (●), bidirectional transporters (●), exportins (●) and Kapβs with no known function (●) are shown (adapted from Mosammaparast and Pemberton, 2004).

Despite members of the evolutionarily conserved Kapβ family exhibiting low sequence identity (~20 %), all members lie within a close molecular mass range of 95–145 kDa and exhibit isoelectric points between 4.0 and 5.0. Furthermore, it is predicted that all Kapβs almost exclusively consist of multiple (~20) tandem helix–loop–helix structures. This structural motif is termed HEAT repeats, named after the proteins in which the motif was first discovered. Kapβs target transport substrates to the NPC by forming specific PPIs with both the cargo and nucleoporins in a particular sequence of events. Since members of the Kapβ family may encompass multiple distinct cargo-binding sites, these transporters are capable of recognising a variety of signalling sequences (Chook and Blobel, 2001; Mosammaparast and Pemberton, 2004). Therefore, Kapβs display the potential for diverse functionality with the resultant possibility of overlapping functions between members. The cell in turn benefits by differentially regulating pathways in response to temporally controlled events, increasing transport capacity and utilising functional redundancy to protect against external stress factors (Wozniak *et al.*, 1998). Furthermore, a number of Kapβs can either bind directly to cargo or via an adapter protein. While modelling studies demonstrate that adapters reduce transport efficiency, adapters may broaden cargo specificity and aid in creating a greater nuclear cytoplasmic gradient (Mosammaparast and Pemberton, 2004; Lange *et al.*, 2007).

Since the nuclear translocation functionality of Kapβs is essential in ensuring proper signal transduction in numerous cellular pathways, any misregulation of Kapβ expression and distribution, or endogenous nuclear transport inhibitors, may result in pathogenesis. Alterations in Kapβ driven nuclear translocation have been linked to cancer due to aberrant physiological levels as well as the irregular distribution of key cellular regulators, including proto-oncogenes and tumour suppressor genes. Examples of Kapβs linked to tumourigenesis include the overexpression of XPO1, misregulation of CSE1L PTMs resulting in incorrect localisation, as well as the reduced expression or absence of endogenous nuclear transport inhibitors CC3/TIP30 and ARH1/NOEY2 (Stelma *et al.*, 2016; Cagatay and Chook, 2018).

1.3.1 Nuclear Localisation Sequences

As aforementioned, the NLS of a cargo protein is a short sequence that acts as an autonomous signal specifying nuclear import via cytosolic importins. Generally, these sequences predominantly consist of positively charged Lys or Arg residues exposed on the protein surface. Despite the importance of basic residues in NLSs, previous studies demonstrated that in some instances neutral and even acidic residues can have a vital impact on NLS functionality. Furthermore, the affinity of a cargo NLS to the respective importin, in addition to the concentration of this importin, determines the overall effectiveness of selective nuclear accumulation of the cargo. While NLSs share certain common features, these sequences can take on an assortment of different sequence formats. However, a diverse set of cargos may still share the same NLS notwithstanding the variability of NLSs and cargo (Makkerh *et al.*, 1996; Lange *et al.*, 2007). Despite this large sequence variability, NLSs are broadly classified based upon sequence identity as either classical or non-classical (Wozniak *et al.*, 1998).

The best characterised macromolecular nucleocytoplasmic transport system is the classical nuclear import pathway (see Section 1.3.3). This system involves the nuclear import of cargo proteins containing a short sequence rich in basic residues called a classical NLS (cNLS). In this prototypical case, the adapter Impα recognises and directly binds the cargo cNLS. Importin β (Impβ) is then recruited by the flexible N-terminal Impβ binding (IBB) domain of Impα (Chook and Blobel, 2001; Lange *et al.*, 2007). Upon completing the transportation complex, Impβ mediates the interactions between the complex and the cytoplasmically exposed nucleoporins of the NPC to permit nuclear import (Wozniak *et al.*, 1998).

Even though all cNLSs follow the same nuclear transport pathway, this group can be further sub-categorised as either monopartite or bipartite. The main structural difference between these is that monopartite cNLSs contain one small cluster of basic residues, whereas bipartite cNLSs contain two such clusters segregated by a short spacer (Lange *et al.*, 2007). The first discovered NLS was the monopartite cNLS present in the simian virus 40 (SV40) large T-antigen and exhibits the sequence 126-P**KKKRK**V-132 (NLS residues in bold and underlined) (Kalderon *et al.*, 1984). Many other cNLSs have since been discovered utilising this single basic residue cluster as a prototype. Later structural and thermodynamic studies determined that a monopartite cNLS requires the first position to be occupied by a Lys, followed by two basic residues in the second and fourth positions. Consequently, a loose consensus sequence of K(K/R)X(K/R) was proposed (Chelsky *et al.*, 1989). An investigation of nucleoplasmin led to the discovery of the bipartite cNLS. This protein contains the sequence 155-**KR**PAATKKAGQA**KKKK**-170 which encompasses two clusters of basic residues detached by a mutation-tolerant spacer. Even though one of these clusters is comparable to the first discovered NLS, both clusters are essential for nuclear transport. Consequently, all bipartite cNLSs are expected to display a sequence exhibiting two basic residue clusters separated by a 10 residue spacer (Makkerh *et al.*, 1996; Lange *et al.*, 2007).

Numerous cargo proteins have since been identified to contain NLSs distinct from the classical sequences and therefore do not follow the classical nuclear import pathway. These proteins rather encode for non-classical NLSs (ncNLSs) which form direct interactions with different Impβ homologues, as opposed to interactions via an adapter protein. The ncNLSs can take on an array of amino acid sequences which do not correlate with that found in the cNLSs (Wozniak *et al.*, 1998; Chook and Blobel, 2001). One example of an ncNLS is the M9 sequence present within the heterogeneous nuclear ribonuclear protein (hnRNP) A1. This ncNLS is made up of a 38 amino acid sequence and functions as both an NLS and NES. The M9 sequence first facilitates the nuclear export of mRNA-bound hnRNP A1. Following the release of the mRNA sequence, mammalian Kapβ2 (transportin) recognises the M9 sequence of the cytoplasmic hnRNP A1 and consequently promotes nuclear import in an energy dependent manner. Transportin also appears to mediate the nuclear import of a number of other hnRNPs (Wozniak *et al.*, 1998). Another protein that experiences nuclear import via an ncNLS is parathyroid hormone-related protein (PTHrP). Residues 66–94 of PTHrP interacts with a high affinity to an Impβ domain that is separate to the Impα IBB domain binding site (Cingolani *et al.*, 2002).

1.3.2 Impβ and Impα

Impβ is a nuclear importer capable of transferring protein cargo into the nucleus either directly or via an adapter mediator protein. The adapter ordinarily associated to Impβ is the related Impα protein (Mosammaparast and Pemberton, 2004; Lange *et al.*, 2007). Impβ consists of 19 HEAT repeats that arrange into an overall compact superhelix, with a diameter of 50 Å and pitch of 50 Å. Each repeat unit consists of either two or three helices. In a two-helix repeat, helix A is found on the convex side of the repeat while helix B is on the concave side. However, a short helix C can join the helix B from one repeat to the helix A of the next repeat to form a three-helix repeat. Interestingly, the 8th HEAT repeat within the Impβ structure contains an acidic loop (Chook and Blobel, 2001). A noticeable variance exists in the structural determinants from cargos that can be recognised by Impβ. To accommodate a diverse set of cargos, while maintaining ligand specificity, Impβ exposes large interaction interfaces along with conformational plasticity upon cargo binding (Chook and Blobel, 2001; Mosammaparast and Pemberton, 2004). The latter involves the reorientation of helical repeat units in relation to one another as opposed to structural alterations within the actual units. Several substrates that undergo nuclear import via a direct interaction with Impβ include HIV Rev, T-cell Tyr phosphatase, GAL4, PTHrP and cyclin B1. PTHrP has been shown to bind the central region (residues 380–643) of Impβ, while cyclin B1 binds the N-terminus. In the case of Impα facilitated transport, cargo proteins require a cNLS and thus form part of the classical nuclear import pathway (Chook and Blobel, 2001).

Impα assumes a similar cylindrical conformation to Impβ with an overall superhelical structure constructed of helical repeat motifs within the same family as HEAT repeats, known as armadillo (ARM) repeats. The full Impα consists of ten ARM repeats, reaching a final length of 100 Å and diameter of 30 Å. A single ARM repeat consists of ~40 residues organised into three helices denoted as H1, H2 and H3. The relative positioning between each successive repeat is associated with an approximate 9 Å translation and 30° rotation. When individual ARM and HEAT repeats are superimposed, the similarities become evident with helices H3/B lining the concave side of the proteins and helices H1-H2/A producing the convex spinal side. However, some variation lies in the curvature between Impα and Impβ due to different orientations between adjacent repeats. ARM repeats form a smooth, elongated shape in Impα, as a result of regular 30° inter-repeat rotations, while HEAT repeats bend Impβ into defined coils and arches due to large 30–60° interspersed rotations amidst 10–25° rotations between repeats (Chook and Blobel, 2001; Lange *et al.*, 2007).

The structural plasticity of both Impβ and Impα facilitate the essential PPIs during the assembly and disassembly of the cargo nuclear transport complex (Figure 6). Upon heterodimerisation, Impβ forms a compact superhelix with 13 of the 19 HEAT repeats wrapping around the N-terminal IBB domain of Impα (Figure 6A). During this interaction, residues 11–23 of the IBB domain stretch outwards while the remainder of the domain assumes a long α-helical structure. This permits the prolonged region to associate with the concave surface of HEAT repeats 7-11, including the acidic loop in the 8th repeat. The binding interface between the two importins includes highly conserved hydrophobic and acidic Impβ residues as well as basic Impα residues. While the aliphatic region of the basic residues in Impα bind to Trp and other hydrophobic Impβ residues, the guanidinium and amino functional groups of the former form salt bridges with the acidic residues of the latter. The remaining HEAT repeats 12-19 wrap around the α-helix of Impα by 270° (Chook and Blobel, 2001; Mosammaparast and Pemberton, 2004).

The slightly curving and extended Impα houses two cNLS-binding sites, namely the major pocket and the minor pocket, within the superficial groove of the concave face (Figure 6B). These pockets facilitate the physiochemical requirements for protein cargo cNLSs binding. The major pocket is proximal to the N-terminus and binds both monopartite cNLSs and the longer cluster of basic residues in bipartite cNLSs. The minor pocket lies nearer the C-terminus and binds the shorter cluster of basic residues in bipartite cNLSs. Both pockets consist of conserved, solvent exposed Trp residues in addition to a set of invariant Asn residues four residues downstream. Upon binding, the prolonged conformation of the cNLS lies anti-parallel to the Impα main-chain. The conserved Trp residues in Impα then position the main Lys side-chains of the cNLS between the stacked hydrophobic indole side-chains of these residues. This enables the Lys residues to form salt bridges with the negatively charged residues coating the binding groove. Furthermore, the Impα Asn residues proceed to interact with the main-chain of the cNLS (Lange *et al.*, 2007).

The flexible Impα IBB domain also facilitates cargo dissociation. When the IBB domain is not bound to Impβ, residues 44–54 interact with an autoinhibitory sequence within Impα. This sequence is present in the major pocket of Impα and consists of KRR residues, mimicking a basic cNLS. Blockage of the major pocket causes an autoinhibited state, since neither monopartite nor bipartite cNLSs may access the major pocket. Autoinhibition is a Ran dependent pathway (Figure 6C) (Chook and Blobel, 2001; Lange *et al.*, 2007).

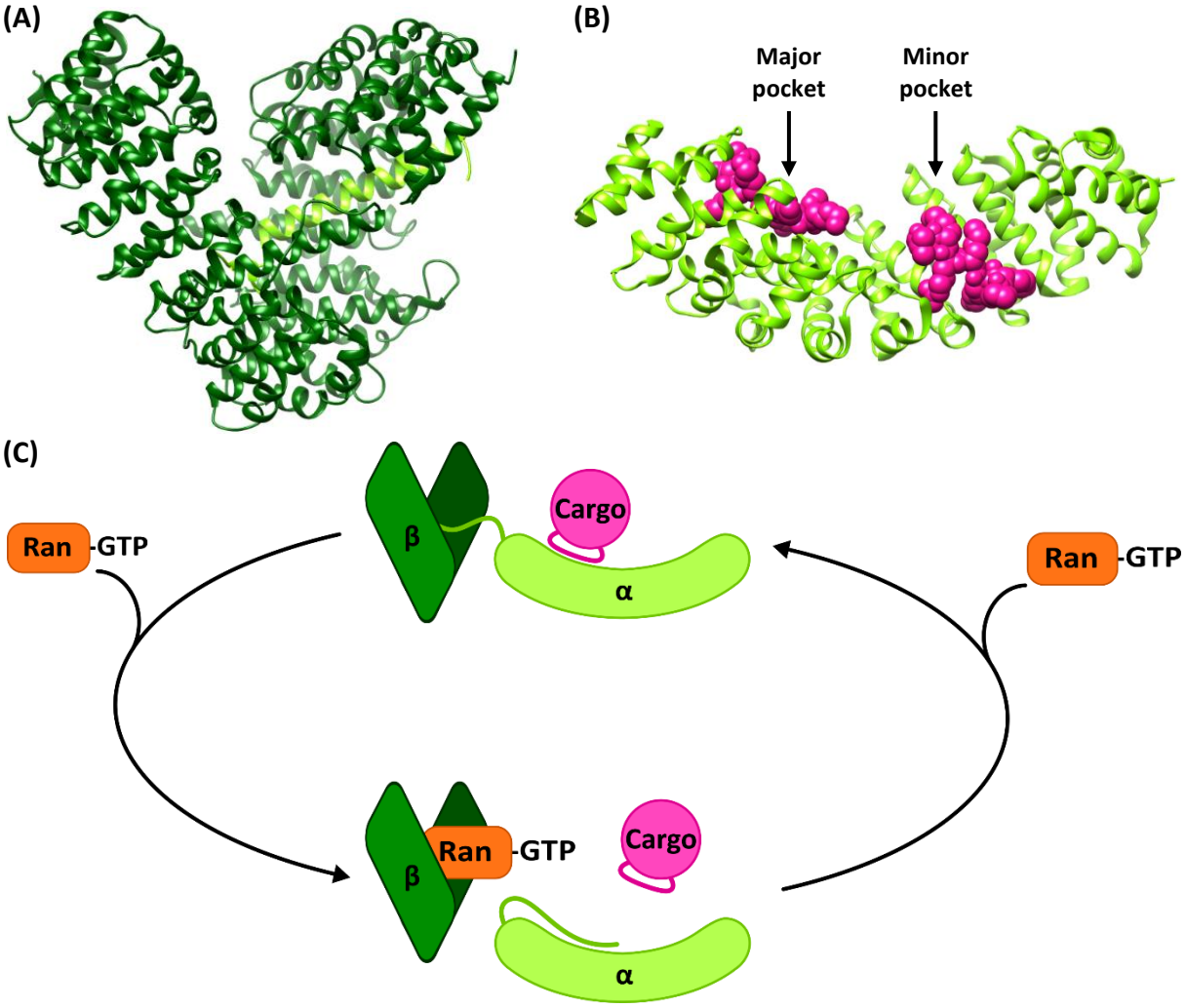


Figure 6: Impβ and Impα configurations in the classical nuclear transport pathway
(A) The extended N-terminal Impα IBB domain (light green) protruding the Impβ HEAT repeat structure (forest green). PDB file 1QGK was used to generate the cartoon figure using the UCSF Chimera Molecular Graphics System, V1.12. **(B)** Two SV40 peptides (magenta) bound to the major and minor pockets of mImpαΔIBB (light green). PDB file 1EJL was used to generate the cartoon figure using the UCSF Chimera Molecular Graphics System, V1.12. **(C)** Impβ binding the autoinhibitory Impα IBB domain in the cytoplasm frees the domain from the major pocket. A cNLS can then bind Impα with a high affinity. Following nuclear import, Impβ interacts with RanGTP, freeing the domain to compete for the major pocket. A decrease in cNLS affinity causes the dissociation of the import cargo from Impα in the nucleus.

1.3.3 Classical Nuclear Import Pathway

The classical nuclear import pathway (Figure 7), as previously mentioned, involves the unique interactions that take place between Impβ and cNLS containing cargo proteins mediated by the adapter Impα (Lange *et al.*, 2007). The energetic requirements and directional regulation of the importins involved in this pathway are provided by a small Ras family GTPase known as Ran. This regulatory factor cycles between a GTP- (guanosine triphosphate) and a GDP- (guanosine diphosphate) bound state, with cellular compartmentalisation inducing a concentration gradient of these two forms between the nucleus and cytoplasm (Wozniak *et al.*, 1998; Mosammaparast and Pemberton, 2004; Lange *et al.*, 2007). Ran GTPase-activating protein (RanGAP) catalyses nucleotide hydrolysis and causes the accumulation of RanGDP by the cytoplasmic fibers of the NPC. Conversely, Ran guanine nucleotide exchange factor (RanGEF) catalyses nucleotide exchange and results in the accumulation of RanGTP at the chromatin and NPC nuclear basket (Chook and Blobel, 2001; Lange *et al.*, 2007). These varying states of Ran are responsible for either the assembly or disassembly of key complexes in this transportation pathway. Therefore, Ran both fulfils the energetic requirement of this pathway and acts as a molecular switch in directing the vectorial movement of cargo between the nucleus and cytoplasm (Wozniak *et al.*, 1998; Chook and Blobel, 2001).

To initiate the classical nuclear import pathway, the IBB domain of Impα binds Impβ which in turn frees the Impα binding pockets to recruit a cNLS containing cargo protein within the cytoplasm. Impβ then docks the transportation complex to the nucleoporins of the NPC to induce nucleocytoplasmic translocation (Chook and Blobel, 2001; Lange *et al.*, 2007). The presence of RanGDP in the cytoplasm stabilises the assembled complex by the cytoplasmic fibers of the NPC (Wozniak *et al.*, 1998; Lange *et al.*, 2007). Following nuclear import, RanGTP interacts with the N-terminal domain of Impβ thus triggering a significant conformational change which decreases the affinity of Impβ for the Impα•cargo complex. This in turn results in the release of the Impα•cargo complex from Impβ in the nucleus (Wozniak *et al.*, 1998; Chook and Blobel, 2001; Mosammaparast and Pemberton, 2004; Lange *et al.*, 2007). Thereafter, the newly formed Impβ•RanGTP complex is exported from the nucleus into the cytoplasm after which RanGAP hydrolyses the GTP back to GDP in the cytoplasm. This in turn alters the structural conformational of Ran resulting in the dissociation of Impβ due to a reduction in binding affinity (Chook and Blobel, 2001; Lange *et al.*, 2007). Impα is then recycled back to the cytoplasm by associating with RanGTP complexed to a nuclear export receptor known as mammalian cellular apoptosis susceptibility protein (CAS, CSE1 in yeast). Similarly, the hydrolysis of the GTP linked to Ran releases Impα in the cytoplasm. The concentration gradient is finally restored with Ran-GDP returning to the nucleus via nuclear transport factor 2 (NTF2), where RanGEF recycles the GDP back to GTP (Lange *et al.*, 2007).

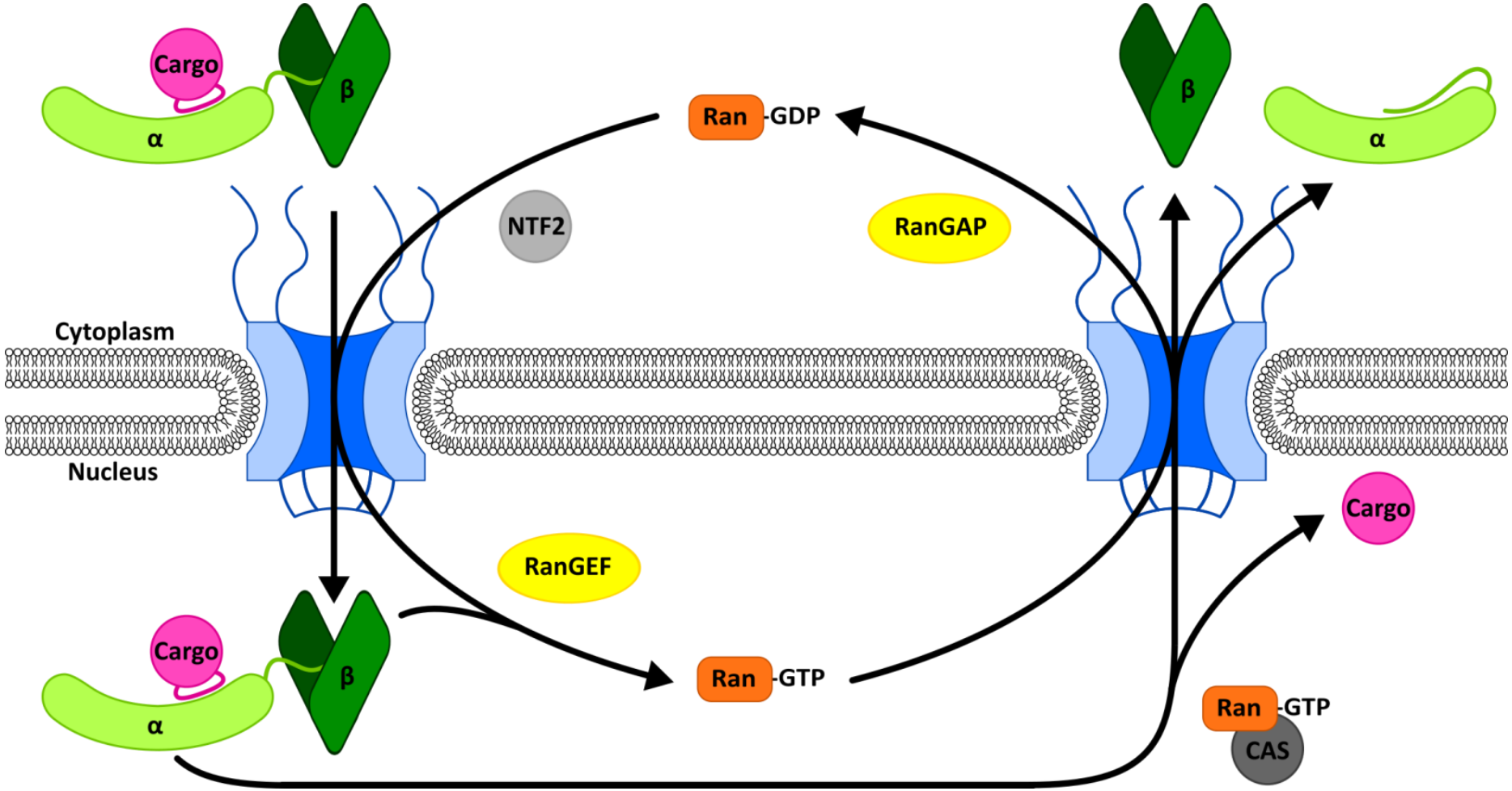


Figure 7: The classical nuclear import pathway
In the cytoplasm, the import cargo cNLS is recognised and bound by Impα followed by the addition of Impβ. Impβ facilitates the transport of the complex through the NPC (blue) into the nucleus. NTF2 assists the re-entry of Ran-GDP into the nucleus where RanGEF exchanges GDP to GTP. Ran-GTP then interacts with Impβ resulting in the dissociation of the Impα•cargo complex. Impβ is consequently exported to the cytoplasm and RanGAP hydrolyses the GTP to GDP. RanGTP complexed CAS exports Impα from the nucleus and releases Impα through the hydrolysis of GTP.

1.4 Aim and Objectives

The involvement of CLIC1 and CLIC4 in oxidative stress responses, including cell cycle control and Aβ-induced neurotoxicity, highlights the importance of developing an understanding of how cellular redox states regulate the PPIs of these proteins. Hence a deeper investigation of the effect of high oxidative levels on CLIC1 structure and dimerisation is required. Furthermore, while other studies illustrated the impact of CLIC4 nuclear presence on tumourigenesis, more work is required to characterise the respective binding interactions of CLIC1 and CLIC4 to the karyopherin mediator protein, Impα. This study therefore aimed to first determine the effect of redox regulation on the dimerisation of CLIC1 by structurally and chemically characterising the reduced and oxidised forms of the protein. Thereafter, peptides were designed of the putative CLIC1 (Pep1) and CLIC4 (Pep4) NLSs to establish whether the NLS regions may bind truncated hImpα (70–529) isoform 1, omitting the IBB domain (hImpα1ΔIBB), in the event of exposure due to oxidative effects.

The specific objectives to achieve this overall aim were to:

- Overexpress and purify full-length WT CLIC1, full-length WT CLIC4 and WT hImpα1ΔIBB
- Generate oxidised CLIC1 and separate the dimeric and monomeric forms through size exclusion chromatography (SEC)
- Perform secondary and tertiary structural analysis of all protein species through far-UV circular dichroism (CD) and fluorescence spectroscopy
- Establish the effects of oxidation on CLIC1 structural and chemical modifications by assessing thermal stability and disulfide bond formation through liquid chromatography tandem mass spectrometry (LC-MS/MS)
- Perform molecular docking to investigate important binding interactions that occur between ImpαΔIBB (mImpαΔ1IBB and hImpα1ΔIBB) and the putative CLIC NLSs (WT Pep1 and Pep4 as well as Pep1 Cys mutants Pep1_A and Pep1_S)
- Design and chemically synthesise peptides Pep1_S and Pep4
- Establish the binding affinity of Pep1_S and Pep4 to hImpα1ΔIBB, respectively, through fluorescence quenching
- Determine the thermodynamic parameters and proton linkage effects involved specifically in the binding of Pep4 to hImpα1ΔIBB through the use of isothermal titration calorimetry (ITC)

CHAPTER 2. EXPERIMENTAL PROCEDURE

2.1 Materials

His-Tag fusion proteins of full-length WT CLIC1, full-length WT CLIC4 and WT hImpα1ΔIBB were each inserted into respective pET-28a plasmids by GenScript (Piscataway, NJ, USA). Competent *E. coli* T7 cells obtained from New England Biolabs (Ipswich, MA, USA) were then transformed with the respective plasmids. The GeneJET® Plasmid Miniprep Kit was purchased from Thermo Fisher Scientific (Waltham, MA, USA). Isopropyl-β-D-1-thiogalactopyranoside (IPTG) was obtained from Melford Laboratories Ltd. (Ipswich, UK). DNase was purchased from Roche Diagnostics (Mannheim, Germany). Kanamycin, phenylmethylsulfonyl fluoride (PMSF), tris(2-carboxyethyl)phosphine (TCEP), thrombin from human plasma (1 000 units) and purification columns were supplied by Sigma-Aldrich Co. (St. Louis, MO, USA). H_2O_2 was obtained from Merck KGaA (Darmstadt, Germany). The putative CLIC1 and CLIC4 NLS peptides were synthesised by GLBiochem (Shanghai, China). All other reagents used were of standard analytical grade.

2.2 Plasmid Preparation

The same expression construct design was utilised for full-length WT CLIC1, full-length WT CLIC4 and WT hImpα1ΔIBB (RAG cohort 1, Impα 1, 70-529), henceforth referred to as CLIC1, CLIC4 and hImpα1ΔIBB, respectively. Each expression construct was designed by Dr I.A. Achilonu (University of the Witwatersrand, Johannesburg, South Africa) to be codon optimised for *E. coli* and contain an N-terminal His-Tag and thrombin cleavage site. The insert DNA sequences were each synthesised and sub-cloned into a pET-28a plasmid by GenScript (Piscataway, NJ, USA). Each plasmid was then transformed into competent *E. coli* T7 cells and expressed overnight. Thereafter, these plasmids were extracted and purified using the GeneJET® Plasmid Miniprep Kit (Thermo Fisher Scientific, Waltham, MA, USA). The resulting plasmid DNA was sent to Inqaba Biotechnical Industries (Pty) Ltd (Pretoria, South Africa) for sequencing to verify each gene insert.

2.3 Protein Overexpression and Purification

Optimal induction conditions for hImpα1ΔIBB were established via induction trials. Late-log growth phase *E. coli* T7 cells were inoculated into 2×YT media, containing 30 $\mu g \cdot ml^{-1}$ kanamycin, which then grew at 37 °C until the cells reached the mid-log growth phase (OD600 ~0.6). Thereafter, one subset of cell culture was taken to incubate at 20 °C while the other subset remained at 37 °C. Cultures from each subset were induced for overexpression using various final IPTG concentrations (0.2-1 mM). Following induction, samples were taken from the cultures of each subset at 0, 2, 4, 8 and 16 hrs. The pellet and supernatant of each sample from both subsets were resolved using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (see Section 2.3.2). The resulting gels were analysed using ImageJ 1.52a by means of densitometry (Abràmoff *et al.*, 2004).

Based on the induction trials, hImpα1ΔIBB was overexpressed in *E. coli* T7 cells by first growing the cells in 2×YT media containing 30 μg·ml^{-1} kanamycin at 37 °C. Upon reaching the mid-log growth phase, the cells were induced with a final IPTG concentration of 0.2 mM and incubated at 20 °C for 6 hrs. Through centrifugation, the cells were harvested and the resultant pellets resuspended in a 20 mM Tris buffer pH 8.0, containing 500 mM NaCl, 1 mM PMSF and 0.01 % (w/v) NaN_3. Resuspended cells were stored at -20 °C until needed. The same overexpression and harvesting protocols were followed for CLIC1 and CLIC4, with the exception of a 16 hr incubation period following induction. The same cell lysis and purification protocols were followed for all three proteins. The cells were thawed by rotation at 30 °C after which 1 μl of 1 M $MgCl_2$, 1 μl of 10 mg·ml^{-1} DNase and 10 μl of 10 mg·ml^{-1} lysozyme per millilitre of cell culture was added. The cells were sonicated on ice for three cycles of 30 sec pulses at medium intensity with a 1 min rest period between cycles. The cell lysate was then centrifuged at 12 000 rpm for 20 mins at 4 °C to remove cell debris and obtain the supernatant containing the soluble protein fraction.

A 5 ml Ni^{2+} charged affinity column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) was equilibrated with a 20 mM Tris buffer pH 8.0, containing 500 mM NaCl, 30 mM imidazole and 0.01 % (w/v) NaN_3. After adding the cell lysate supernatant, the column was washed with the same equilibration buffer to remove all unbound proteins. A step elution using a 20 mM Tris buffer pH 8.0, containing 500 mM NaCl, 500 mM imidazole and 0.01 % (w/v) NaN_3, then eluted the bound fusion protein of interest. After collecting this protein, the absorbance at 280 nm was recorded and then dialysed against a 50 mM sodium phosphate buffer pH 7.4, containing 30 mM NaCl and 0.01 % (w/v) NaN_3, overnight at 4 °C. To remove the His-Tag, 1 μl of a 1 U/ml thrombin stock solution was added per millilitre of pure fusion protein for 4 hrs at 20 °C. The dialysis buffer was used to equilibrate the Ni^{2+} column linked in succession to a 5 ml diethylaminoethyl (DEAE) column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). Upon adding the cleaved protein, the Ni^{2+} column binds the His-Tag while the DEAE column binds the cleaved protein, thus causing the thrombin to be washed out. Finally, a step elution using a 50 mM sodium phosphate buffer pH 7.4, containing 500 mM NaCl and 0.01 % (w/v) NaN_3, eluted the cleaved protein which was collected and the absorbance at 280 nm recorded. The final purified protein was concentrated and dialysed against a 50 mM sodium phosphate buffer pH 7.4 for storage and kept at -80 °C.

2.3.1 Protein Concentration Determination

All reported protein concentrations refer to subunit/monomer concentrations. Concentrations were determined spectrophotometrically at 280 nm using the Beer-Lambert Law:

$$A_{280} = \varepsilon C l, \tag{1}$$

where A_{280} is the absorbance at 280 nm (AU), ε the molar extinction coefficient ($M^{-1}\cdot cm^{-1}$), C the protein concentration (M) and *l* the cuvette path length (cm).

Protein molar extinction coefficients at 280 nm were calculated by the ProtParam tool on the ExPASy Server (Gasteiger *et al.*, 2005). The molar extinction coefficients for CLIC1, CLIC4 and hImpα1ΔIBB were found to be 17 420 $M^{-1} \cdot cm^{-1}$, 21 430 $M^{-1} \cdot cm^{-1}$ and 50 420 $M^{-1} \cdot cm^{-1}$, respectively.

2.3.2 SDS-PAGE

SDS-PAGE was utilised to assess protein overexpression during the induction trial of hImpα1ΔIBB, as well as the relative purity and molecular mass of each protein following purification. All gels used consisted of a 4 % stacking gel and 12.5 % separating gel (Laemmli, 1970). Preparation of samples entailed incubation in a 0.125 M Tris SDS-PAGE reducing buffer pH 6.8, containing 4 % (w/v) SDS, 20 % (v/v) glycerol, 5 % (v/v) β-mercaptoethanol and 0.02 % (w/v) bromophenol blue, for 5 mins at 95 °C. The BLUeye Prestained Protein Ladder from Merck KGaA (Darmstadt, Germany), containing a set of known standards, was loaded onto each gel to determine the molecular mass of the protein samples. Electrophoresis took place for approximately 1 hr at 180 V. Upon completion, the gels were first immersed in a staining solution containing Coomassie Blue R250 which was then replaced with a destaining solution consisting of 40 % (v/v) methanol and 10 % (v/v) acetic acid. Destaining took place until the background appeared clear.

2.4 CLIC1 Oxidation

A 20 mM H_2O_2 stock solution was prepared in MilliQ water. A final H_2O_2 concentration of 2 mM was added to 2 $mg \cdot ml^{-1}$ CLIC1 prepared in a 20 mM sodium phosphate buffer pH 7.4, containing 150 mM NaCl. The mixture was incubated for 2 hrs at room temperature directly prior to SEC separation (Goodchild *et al.*, 2011).

2.4.1 SEC

The quaternary structure of CLIC1 in both the reduced and oxidised forms were verified using SEC. The separated oxidised dimer and monomer were used for downstream experiments. A HiLoad™ 16/60 Superdex™ 75 prep grade size exclusion column with a resolution of 3-70 kDa (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) was used. The column was connected to an ÄKTA pump and equilibrated with a 20 mM sodium phosphate buffer pH 7.4, containing 150 mM NaCl at 20 °C. A constant flow rate was maintained at 1 $ml \cdot min^{-1}$. Both the reduced and oxidised CLIC1 samples were at a concentration of 2 $mg \cdot ml^{-1}$. The reduced CLIC1 sample was prepared in a 20 mM sodium phosphate buffer pH 7.4, containing 150 mM NaCl and 1 mM TCEP. A standard curve was constructed through the use of a gel filtration markers kit for protein molecular masses 12-200 kDa (GE Healthcare Bio-Sciences AB, Uppsala, Sweden).

2.5 Protein Structural Characterisation

The overall structural conformations of reduced CLIC1, oxidised CLIC1 (dimer and monomer), CLIC4 and hImpα1ΔIBB were determined. Secondary and tertiary structural characteristics were established via far-UV CD and fluorescence spectroscopy, respectively.

2.5.1 Far-UV CD

Far-UV CD spectra were obtained for reduced CLIC1, oxidised CLIC1 (dimer and monomer), CLIC4 and hImpα1ΔIBB at a concentration of 2 μM in a 50 mM sodium phosphate buffer pH 7.4. A Jasco J-810® spectropolarimeter was used in conjunction with a Peltier temperature controller set at 20 °C. All measurements took place in a quartz cuvette with a pathlength of 2 mm. Parameters were set with a band width of 5 nm, data pitch of 0.5 nm and scan speed of 200 $nm \cdot min^{-1}$. Spectra were recorded using the continuous scanning mode from 195-255 nm with ten accumulations obtained per sample. Triplicate readings for each protein species were averaged and corrected for buffer contribution. Averaged machine readings in millidegree ellipticity were normalised to mean residue ellipticity (MRE $deg \cdot cm^2 \cdot dmol^{-1}$), through the following equation (Woody, 1995):

$$\mathrm{MRE} = \frac{100 \cdot \theta}{C l \mathrm{n}}, \qquad (2)$$

where θ is the measured ellipticity signal at the respective wavelength (mdeg), C the protein concentration (mM), *l* the cuvette path length (cm) and n the total number of residues.

2.5.2 Intrinsic Fluorescence

Fluorescence emission spectra were obtained for reduced CLIC1, oxidised CLIC1 (dimer and monomer), CLIC4 and hImpα1ΔIBB at a concentration of 2 μM in a 50 mM sodium phosphate buffer pH 7.4. A Jasco FP-6300® fluorescence spectrophotometer was used in conjunction with a Peltier temperature controller set at 20 °C. All measurements took place in a quartz cuvette with a pathlength of 10 mm. Parameters were set with an excitation band width of 5 nm, emission band width of 2.5 nm, data pitch of 1 nm and scan speed of 200 $nm \cdot min^{-1}$. Samples were excited at 280 nm and emission spectra recorded from 290-500 nm. Three accumulations were obtained per sample. Triplicate readings for each protein species were averaged and corrected for buffer contribution.

2.6 Oxidised CLIC1 Modifications

Structural and chemical modifications induced by oxidation on CLIC1 were determined. This was done by comparing the thermal stability between reduced CLIC1 and oxidised CLIC1 (dimer and monomer) as well as assessing CLIC1 disulfide bond formation.

2.6.1 Thermal Unfolding

Thermal melts were obtained for reduced CLIC1 and oxidised CLIC1 (dimer and monomer) at a concentration of 20 μM for the monomers and 10 μM for the oxidised dimer in a 50 mM sodium phosphate buffer pH 7.4. A Jasco J-810® spectropolarimeter was used in conjunction with a Peltier temperature controller. The temperature was set to increase from 10-80 °C at a rate of 1 °C·min^{-1}. All measurements took place in a quartz cuvette with a pathlength of 2 mm. Parameters were set with a band width of 2 nm, data pitch of 0.5 °C and response of 0.5 sec. Readings were taken at 222 nm and corrected for buffer contribution. Machine readings in millidegree ellipticity were normalised to relative MRE (see Section 2.5.1).

2.6.2 LC-MS/MS

CLIC1 disulfide bond formation was assessed through a double alkylation experiment analysed via LC-MS/MS (Yen *et al.*, 2000; Klapoetke and Xie, 2017). Reduced and oxidised (dimer and monomer) CLIC1 samples were prepared to a concentration of 0.5 mg·ml^{-1} in triplicate. All experimental procedures took place at pH 4.5. Free Cys residues were labelled with *N*-ethylmaleimide (NEM) by adding 2.5 μl of 100 mM NEM and 4 % SDS, after which the samples were incubated for 1 hr in the dark. MagReSyn® hydrophilic interaction liquid chromatography (HILIC) beads were equilibrated by adding 175 μl of HILIC binding buffer (15 % acetonitrile and 100 mM NH_4Ac) to 25 μl of HILIC beads (20 mg·ml^{-1}). To allow the samples to bind the HILIC beads, 50 μl protein and 50 μl HILIC binding buffer was added to the beads and incubated for 10 mins. Excess NEM was removed by washing the beads with 200 μl HILIC binding buffer twice. To reduce and denature the samples, the beads were incubated for 45 mins with 100 μl HILIC binding buffer, containing 2% SDS and 10 mM DTT. Disulfide bonded Cys residues prior to reduction, were labelled with 2-iodoacetamide (IAA) and further denatured by adding 100 μl HILIC binding buffer, containing 40 mM IAA and 2% SDS, and incubated for 45 mins. Excess IAA and SDS was removed by washing the beads with 200 μl of 95 % acetonitrile twice. Protein digestion was initiated by adding 195 μl of 50 mM ammonium bicarbonate and 10 μl of 0.5 μg·μl^{-1} trypsin stock and incubated for 4 hrs. Following digestion, samples were vacuum dried and stored at -80 °C.

Samples were analysed via an electrospray ionisation ABSciex Triple TOF 6600 MS at the CSIR (Pretoria, South Africa) by Dr S.H. Stoychev who further generated a spectral library through ProteinPilot™ (ABSciex) (Seymour and Hunter, 2015). Peak areas were extracted by Skyline V3.7 to quantify the relative abundance of IAA and NEM labelling for all Cys residues within each sample (MacLean *et al.*, 2010). These relative abundance values were used to calculate the response, which is the proportion of IAA labelling for a specific Cys residue within a sample. The logarithm of the odds for IAA labelling was modelled by Dr S.J.M. Steeb (University of Johannesburg, South Africa) as a linear function of the two factors (Cys residue and oxidation state) to estimate the influence of these predictors on the response.

The response was found to be a continuous proportion, with the observed response variance not constant across the predictor levels. While transformations to ordinary least squares regression are often applied to overcome these challenges, this results in biased estimates and difficulties in interpretation. Based on exploratory plots of the data (Appendix A), several models were compared using AIC, BIC and hypothesis tests to examine the significance of individual parameters included in a model. A beta regression model was chosen and assumes the following. The continuous proportions were generated from a beta distribution and the expected proportion is related to a set of predictors through a linear predictor which is a linear function of unknown parameters. Furthermore, a logit link function is applied to the expected response and is equal to the linear predictor. The model also includes a precision parameter which depends on one or more predictors through a link function. The betareg package, which provides the class of beta regressions in the R system for statistical computing, was used to fit the model: betareg(formula = IAA ~ Cys * Ox | Cys, link = "logit") (Cribari-Neto and Zeileis, 2010).

2.7 CLIC Peptides and hImpα1ΔIBB Interaction Studies

Molecular docking was used to assess and compare the interactions of the CLIC peptides to hImpα1ΔIBB and mImpα1ΔIBB, respectively. Apparent K_d values for both Pep1_S and Pep4 binding hImpα1ΔIBB were established through fluorescence quenching. Thermodynamic parameters involved specifically in Pep4 binding hImpα1ΔIBB were determined via ITC.

2.7.1 Molecular Docking

All preparative steps utilised The PyMOL Molecular Graphics System V1.7.4.5 Edu (Schrödinger, LLC, USA). The apo mImpα1ΔIBB structure was prepared from the 3OQS PDB structure by removing the CLIC4 NLS peptide, VA**KKYR**N (NLS residues in bold and underlined), and solvent. This peptide will henceforth be referred to as truncated Pep4. The apo hImpα1ΔIBB structure was prepared from the 3WPT PDB structure by removing one dimer subunit, the PEG molecules and solvent. The Pep4 and Pep1 sequences were based on the CLIC4 NLS peptide designed by Mynott *et al.* (2011), with each peptide exhibiting five residues prior to and one residue following the NLS. In addition to the WT Pep1 sequence, two mutant sequences replacing the Cys with either an Ala (Pep1_A) or Ser (Pep1_S) were determined, since experimental work would necessitate a Cys mutant to prevent peptide dimerisation. To conserve the expected backbone conformation, the Pep4 structure was generated by elongating the truncated Pep4 3OQS PDB structure to VKVVA**KKYR**N. The Pep1 structures were created by mutating the Pep4 structure to the Pep1 VQVVC**KKYR**G, Pep1_A VQVVA**KKYR**G and Pep1_S VQVVS**KKYR**G structures. Models were generated for mImpα1ΔIBB bound to Pep4 and Pep1, respectively, as well as hImpα1ΔIBB bound to Pep4, Pep1, Pep1_A and Pep1_S, respectively. Peptide structures were positioned relative to where truncated Pep4 is in mImpα1ΔIBB from the 3OQS PDB structure. Each ImpαΔIBB CLIC peptide complex was then submitted to the FlexPepDock server.

The FlexPepDock server executes a high-resolution peptide docking protocol optimised for 'local' docking within the Rosetta framework (Raveh *et al.*, 2010; London *et al.*, 2011). This server requires estimated, coarse-grain input complexes where the initial peptide conformation is within 5 Å RMSD (root mean square deviation) of the native structure. First, the server iteratively optimises the rigid-body alignment of the peptide to the receptor protein. Further peptide refinement proceeds by honing in on the peptide backbone. On-the-fly side-chain optimisation takes place throughout both processes. An optional low-resolution pre-optimisation protocol may also be implemented. A total of 200 models were generated for each complex with 100 created for the high-resolution protocol and 100 created for including the low-resolution pre-optimisation protocol (Appendix B). The Rosetta scores and rmsBB values for the top ten modelled poses of each complex were retrieved for further analysis.

2.7.2 Peptide Synthesis

The Pep1_S and Pep4 structures were synthesised and assessed for purity by GLBiochem (Shanghai, China). Peptides were generated via a solid phase continuous flow system with the N-termini acetylated and C-termini amidated. Purity was determined through reversed-phase high-performance liquid chromatography (RP-HPLC). Lyophilised peptides were stored at -20 °C and dissolved in a suitable assay buffer prior to use. Once solubilised, the peptides were stored at 4 °C. Peptide concentrations and molar extinction coefficients at 280 nm were calculated as was for protein samples (see Section 2.3.1). The molar extinction coefficient determined for both Pep1_S and Pep4 was found to be 1 490 $M^{-1} \cdot cm^{-1}$.

2.7.3 Fluorescence Quenching

The hImpα1ΔIBB protein was prepared at a concentration of 1 μM, in a 50 mM sodium phosphate buffer pH 7.4. Both Pep1 and Pep4 were solubilised in MilliQ water to create 3.9 mM peptide stock solutions. Increments of 5 μl peptide stock solution were added to hImpα1ΔIBB up to a total volume of 45 μl. Fluorescence emission spectra were recorded of the buffer, hImpα1ΔIBB and at each increment of peptide added to either buffer or hImpα1ΔIBB. A Jasco FP-6300® fluorescence spectrophotometer was used in conjunction with a Peltier temperature controller set at 20 °C. All measurements took place in a quartz cuvette with a pathlength of 10 mm. Parameters were set with an excitation band width of 2.5 nm, emission band width of 2.5 nm, data pitch of 0.5 nm and scan speed of 200 $nm \cdot min^{-1}$. Samples were excited at 280 nm, since excitation at 295 nm did not provide emission spectra with high enough intensity values discernible between each peptide increment. Emission spectra were recorded from 290-500 nm. Three accumulations were obtained per sample. Triplicate readings for each sample were averaged and corrected for buffer and peptide contribution. The readings at 350 nm for each increment from 0-45 μl were taken and the percentage quenching calculated. Percentage quenching values for each increment were fitted to a single site ligand binding curve using SigmaPlot V12.0 (Systat Software, San Jose, CA, USA) and the apparent K_d values for each peptide binding hImpα1ΔIBB were determined.

2.7.4 ITC

Binding thermodynamics was only determined for hImpα1ΔIBB interacting with Pep4, since Pep1_S was insoluble in the ITC assay buffers. In addition to binding thermodynamics, possible proton linked effects in the hImpα1ΔIBB Pep4 binding reaction were also evaluated by using buffers with varying ΔH_{ion} while keeping all other parameters constant. Phosphate, PIPES and HEPES buffers were chosen with the ΔH_{ion} at 25 °C for each buffer being 3.60 kJ·mol^{-1}, 11.20 kJ·mol^{-1} and 20.40 kJ·mol^{-1}, respectively. As all experiments took place at 20 °C, the ΔH_{ion} for each buffer was converted using the formula:

$$\Delta H_T = \Delta H_\theta + \Delta C_\theta (T - \theta), \quad (3)$$

where ΔH is the change in enthaply, ΔC is the change in heat capacity and T refers to a temperature in the vicinity ($T \approx$ 274 K to 350 K) of the reference temperature which is θ = 298.15 K (Goldberg *et al.*, 2002). The ΔH_{ion} at 20 °C for the phosphate, PIPES and HEPES buffers were calculated to be 2.45 kJ·mol^{-1}, 11.31 kJ·mol^{-1} and 20.64 kJ·mol^{-1}, respectively. All buffers were at pH 7.4 and the concentrations were determined in such a way to ensure that each buffer exhibited the same ionic strength. Consequently, respective hImpα1ΔIBB stocks were dialysed extensively against either 50 mM phosphate, 37.5 mM PIPES or 50 mM HEPES buffers. Each protein stock was concentrated to ~320 µM and the respective protein dialysis buffers used to solubilise Pep4 to a concentration of ~5.75 mM.

Titration experiments were performed on a TA Instruments Nano ITC at 20 °C with a stir speed of 250 rpm. A total of 50 injections were used with an injection volume of 5 µl. The injection duration for the phosphate buffer titration was 300 sec, while the duration for the PIPES and HEPES buffer titrations were both 500 sec. All titrations were performed in duplicate and heats of saturation were subtracted from the data points prior to data analysis. NITPIC V1.2.7 was used for automated peak assignment, integration of raw data points to create an isotherm and providing statistical error estimates for individual isotherm data points (Keller *et al.*, 2012; Scheuermann and Brautigam, 2015). The resulting isotherm and associated data was then imported into SEDPHAT V15.2b to fit an A + B ↔ AB hetero-association model to the data (Zhao *et al.*, 2015). From the fitted model, SEDPHAT determined the thermodynamic parameters of the interactions. Thermodynamic parameters for each buffer duplicate were averaged. The final thermograms, fitted isotherms and residual plots were visualised using GUSSI 1.4.2 (Brautigam, 2015). The ΔH_{obs} for each buffer was plotted versus the respective buffer ΔH_{ion} and fitted using linear regression. The proton linkage effect of the interaction was subsequently determined through the equation:

$$\Delta H_{obs} = \Delta H_0 + N_{H^+} \Delta H_{ion}, \quad (4)$$

where ΔH_{obs} is the observed enthalpy for each buffer, ΔH_{ion} is the buffer ionisation enthalpy, ΔH_0 is the intercept indicating the true interaction binding enthalpy and N_{H^+} is the slope which signifies the number of protons released or absorbed by the buffer during the reaction (Baker and Murphy, 1996).

CHAPTER 3. RESULTS

3.1 Plasmid Sequence Verification

Sequencing of the respective plasmid DNA in the regions encoding for CLIC1, CLIC4 and hImpα1ΔIBB, indicated that each codon optimised nucleic acid sequence encoded for the correct protein sequence. In addition to no sequence mutations, each protein displayed the presence of an N-terminal His-Tag and thrombin cleavage site (Figure 8).

Figure 8: Sequence identity of CLIC1, CLIC4 and hImpα1ΔIBB
(A) Verified CLIC1 sequence (Inqaba_Seq) displaying the presence of a His-Tag (☐) and thrombin cleavage site (|) required for protein purification, as well as 100 % sequence identity to the aligned query sequence. **(B)** Verified CLIC4 sequence (Inqaba_Seq) displaying the presence of a His-Tag and thrombin cleavage site required for protein purification, as well as 100 % sequence identity to the aligned query sequence. **(C)** Forward verified hImpα1ΔIBB sequence (Inqaba_Seq_Fw) displaying the presence of a His-Tag and thrombin cleavage site required for protein purification. The region where the reverse sequence (Inqaba_Seq_Rev) begins to overlap with the forward (Inqaba_Seq_Fw) is indicated (‖), and vice versa. The overlap of the two verified sequences displays 100 % sequence identity to the aligned query sequence. All sequence alignments were performed using Clustal Omega (Madeira *et al.*, 2019) and formatted with Jalview 2.11.0 (Waterhouse *et al.*, 2009) and Inkscape.

3.2 Protein Overexpression and Purification

The overexpression and purification of CLIC1, CLIC4 and hImpα1ΔIBB, in addition to assessing the relative purity and molecular mass of the resultant protein through SDS-PAGE.

3.2.1 Induction Trials of hImpα1ΔIBB

Prior to hImpα1ΔIBB overexpression, an induction study was performed to determine the ideal temperature, induction time and IPTG concentration needed for optimal induction. The protein was found to be mostly in the insoluble fraction when cells were incubated at 37 °C, regardless of the induction time and added IPTG concentration (data not shown). Optimal induction conditions for hImpα1ΔIBB were found to occur at 20 °C following 8 hrs of induction with 0.2 mM IPTG (Figure 9).

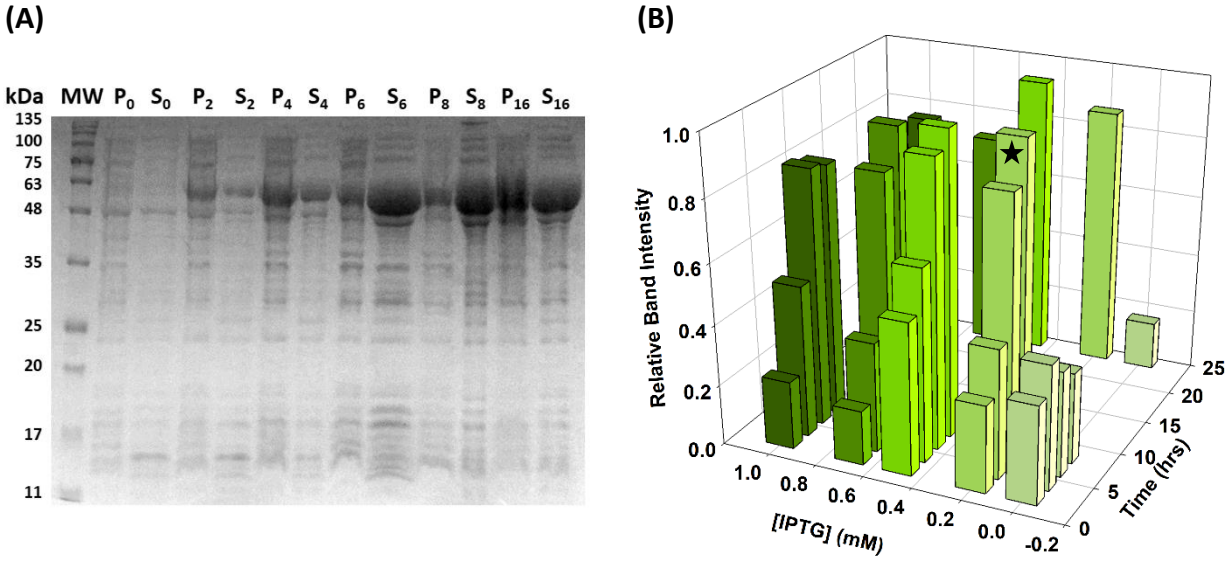


Figure 9: Overexpression and induction study of hImpα1ΔIBB at 20 °C
(A) SDS-polyacrylamide gel (12.5%) showing the pellet (P) and supernatant (S) samples taken at various time points following induction with a final IPTG concentration of 0.2 mM. **(B)** Three-dimensional bar-chart showing the band intensities corresponding to the hImpα1ΔIBB soluble fractions at various time points and IPTG concentrations. The greatest band intensity (★) occurred at 8 hrs following induction with a final IPTG concentration of 0.2 mM.

3.2.2 CLIC1, CLIC4 and hImpα1ΔIBB Purification

Each protein was purified from the respective supernatants of sonicated *E. coli* cells (see Section 2.3). The respective proteins were eluted from the Ni^{2+} charged affinity column via a step elution using a 20 mM Tris buffer pH 8.0, containing 500 mM NaCl, 500 mM imidazole and 0.01 % (w/v) NaN_3 (Figure 10 left panels). Following thrombin cleavage and passage through the Ni^{2+} charged affinity column to remove the His-Tag, the protein of interest was eluted via a step elution from the DEAE column using a 50 mM sodium phosphate buffer pH 7.4, containing 500 mM NaCl and 0.01 % (w/v) NaN3 was used (Figure 10 right panels).

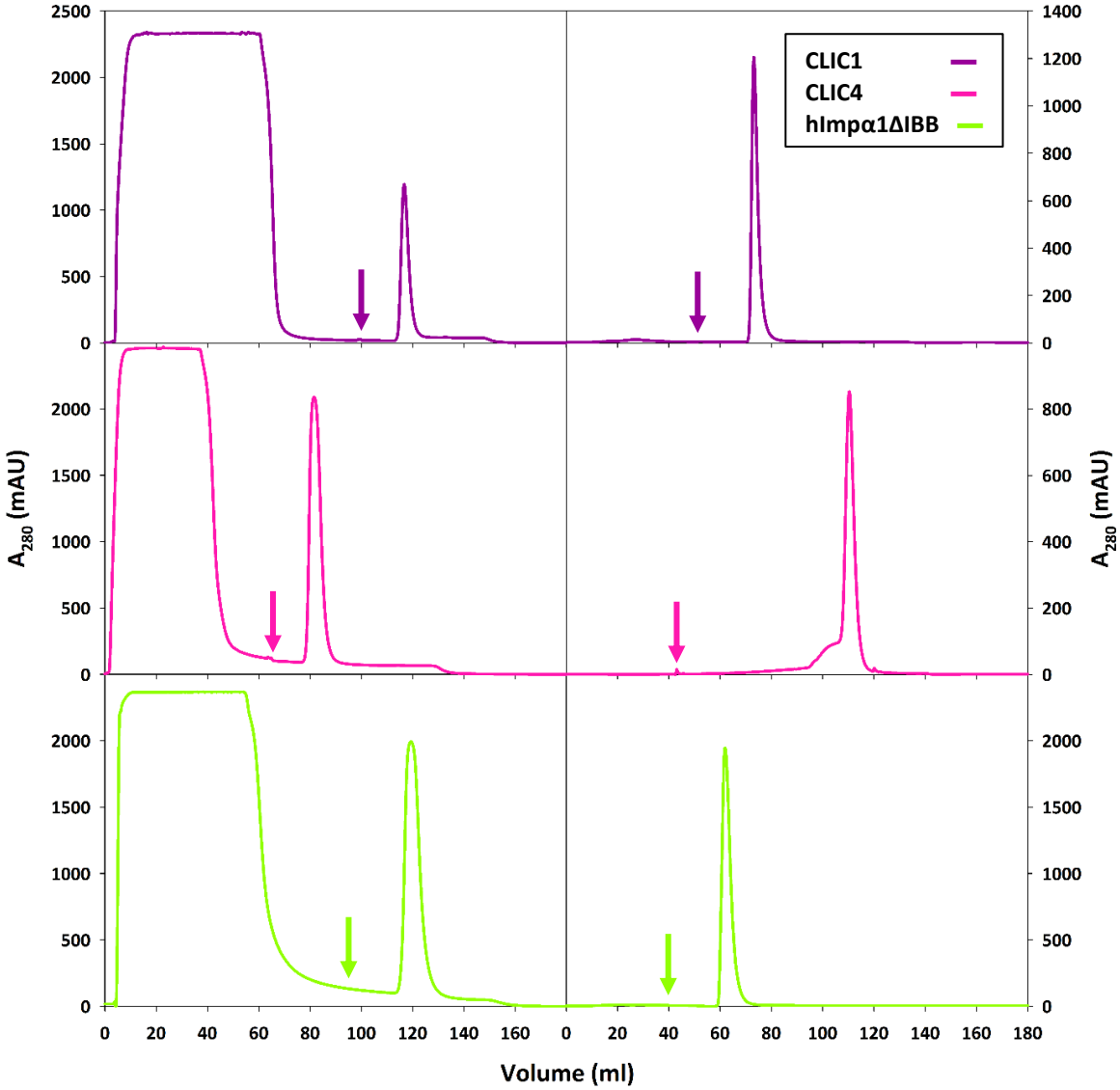


Figure 10: Purification elution profiles of CLIC1, CLIC4 and hImpα1ΔIBB
Left panels: Elution profiles of CLIC1 (—), CLIC4 (—) and hImpα1ΔIBB (—) from the Ni^{2+} charged affinity column showing the absorbance at 280 nm of the effluent with a step elution performed (arrow) using a 20 mM Tris buffer pH 8.0, containing 500 mM NaCl, 500 mM imidazole and 0.01 % (w/v) NaN_3. **Right panels:** Elution profiles of the proteins from the linked Ni^{2+} charged affinity and DEAE columns after 4 hrs of cleavage, showing the absorbance at 280 nm of the effluent with a step elution performed (arrow) using a 50 mM sodium phosphate buffer pH 7.4, containing 500 mM NaCl and 0.01 % (w/v) NaN_3.

Each protein displays a single band during SDS-PAGE analysis alluding to the relative purity of these (Figure 11A). Protein molecular mass was determined by constructing a calibration curve from the molecular mass marker of known standards (Figure 11B). The masses calculated by the ProtParam tool on the ExPASy Server (Gasteiger *et al.*, 2005) for CLIC1, CLIC4 and hImpα1ΔIBB were 27 kDa, 29 kDa and 54 kDa, respectively. While hImpα1ΔIBB appears within the expected mass at 49.9 kDa, CLIC1 and CLIC4 display larger than expected masses of 36.9 kDa and 34.7 kDa, respectively. CLICs do however display slower than expected migration during SDS-PAGE analysis (Tulk *et al.*, 2000).

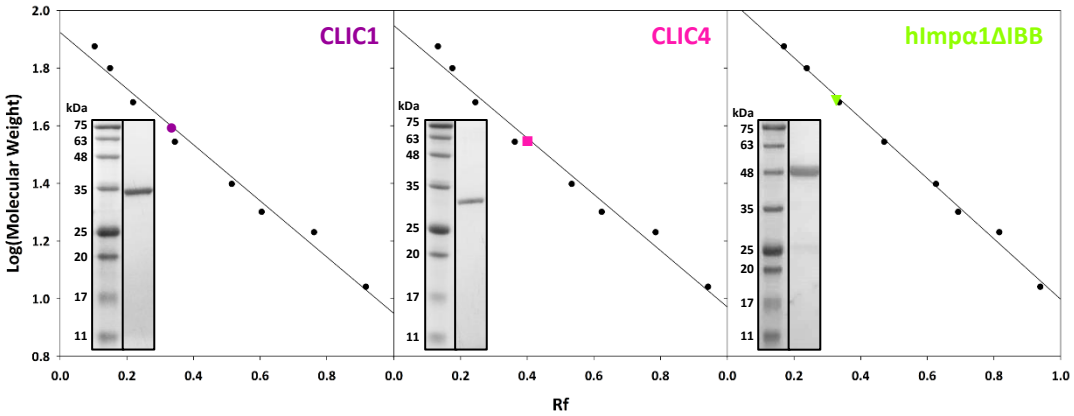


Figure 11: SDS-PAGE analysis of CLIC1, CLIC4 and hImpα1ΔIBB
Calibration curves of the molecular mass markers fitted from the respective SDS-polyacrylamide gel insets (12.5 %) using linear regression. The gels indicate that all three proteins are pure with the calibration curves of known standards showing that CLIC1 (●) is 36.9 kDa, CLIC4 (■) is 34.7 kDa and hImpα1ΔIBB (▼) is 49.9 kDa.

3.3 CLIC1 Oxidation

Oxidised CLIC1 was prepared by incubating the protein with 2 mM H_2O_2 for 2 hrs at room temperature. Thereafter, the oxidised dimeric and monomeric forms of the protein were separated through SEC for downstream experiments.

3.3.1 SEC

The oligomeric states of reduced and oxidised CLIC1 were verified by assessing the hydrodynamic volumes of these species through SEC. Reduced CLIC1 displays a single symmetrical peak at 63.2 ml (Figure 12A) while oxidised CLIC1 displays two significant symmetrical peaks (Figure 12B) at 54.8 ml and 64.5 ml, the latter bearing a shoulder to the left. The elution profile shapes and proportions coincide with the literature (Littler *et al.*, 2004). A calibration curve was constructed from a set of known standards and linear regression analysis performed to calculate the molecular masses of each CLIC1 species (Figure 12C and D). The mass of reduced CLIC1 and the two oxidised species (measured using 2 $mg \cdot ml^{-1}$ protein) were found to be 43 kDa, 98 kDa and 39 kDa, respectively. The reduced and second oxidised peaks elute at a larger than expected monomer mass but are too small to be considered a dimer. The first oxidised peak appears larger than double the monomeric mass. Crystallisation has shown that the oxidised CLIC1 dimer assumes an oblong shape (Littler *et al.*, 2004). This shape would probably result in anomalous behaviour in the column.

3.4 Protein Structural Characterisation

The secondary and tertiary structural conformations of the CLIC1 redox species, CLIC4 and hImpα1ΔIBB were assessed through far-UV CD and fluorescence spectroscopy.

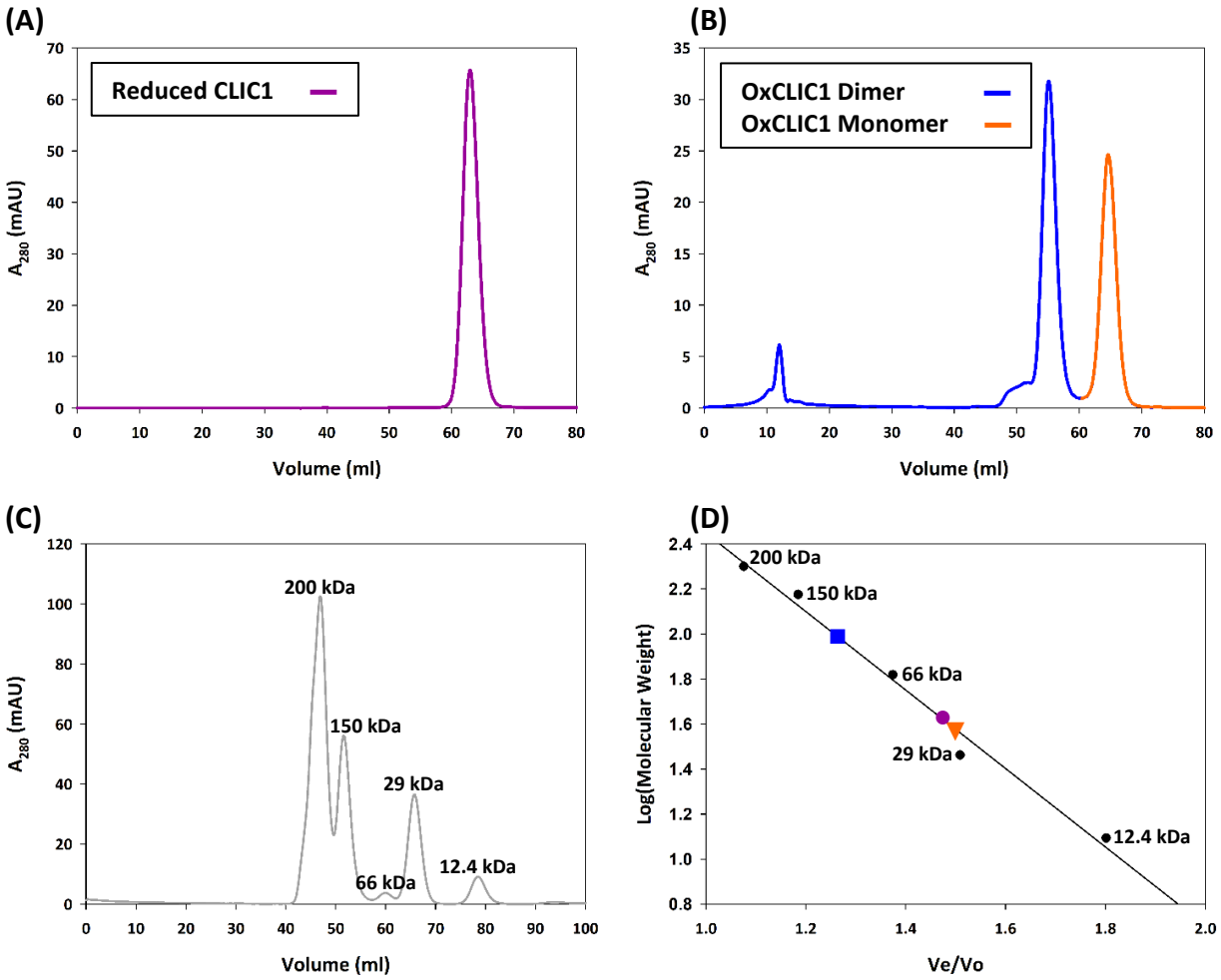


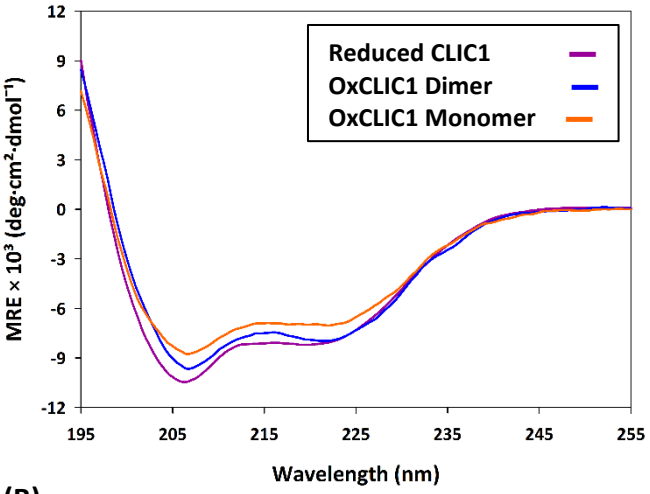
Figure 12: SEC elution profiles of reduced and oxidised CLIC1
(A) Elution profile of reduced CLIC1 monomer (—) when injecting 2 mg·ml^{-1} protein reduced with 1 mM TCEP. **(B)** Elution profiles of the oxidised CLIC1 dimer (—) and monomer (—) when adding 2 mg·ml^{-1} protein oxidised with 2 mM H_2O_2 for 2 hrs. **(C)** Elution profiles of the gel filtration markers kit (—) for protein molecular masses 12-200 kDa. **(D)** Calibration curve of the kit standards fitted using linear regression (R^2 = 0.993). The equation of the fitted line is y = - 1.744x + 4.193. The calibration curve of known standards shows that the reduced CLIC1 monomer (●) is 43 kDa, oxidised CLIC1 dimer (■) is 98 kDa and oxidised CLIC1 monomer (▼) is 39 kDa.

3.4.1 Far-UV CD

The far-UV region ranges from 180-250 nm and encompasses the wavelengths at which the polypeptide backbone amide groups absorb circularly polarised light (Woody, 1995). The polypeptide backbone of a protein may assume an array of alignments which in turn results in the presence of different secondary structural elements. The variation of backbone conformations of these elements causes different features to be evident in a far-UV CD spectra. As a result, the overall secondary structural conformation of the protein can be elucidated by identifying specific far-UV CD spectra characteristics (Greenfield, 2006). Far-UV CD spectra were obtained for reduced CLIC1, oxidised CLIC1 (dimer and monomer), CLIC4 and hImpα1ΔIBB (Figure 13).

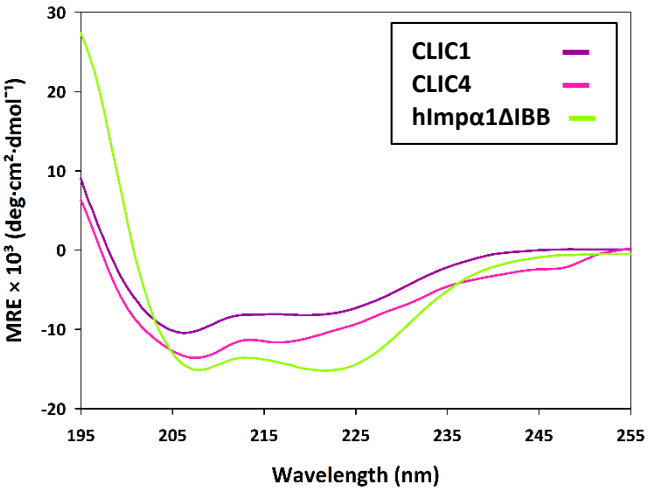
Despite slight variation in intensity, all three CLIC1 redox species exhibit negative bands at 222 nm and 208 nm. This assumes a predominantly α-helical conformation for the CLIC1 species as expected from the reduced (PDB 1K0M) and oxidised CLIC1 (PDB 1RK4) crystal structures (Figure 13A). CLIC4 and hImpα1ΔIBB also display spectra indicative of α-helicity. Furthermore, the ratio between intensity values at 222 nm and 208 nm comments on the type of α-helical structure (Lau *et al.*, 1984). Both reduced CLIC1 and CLIC4 present a ratio close to 0.8 indicating isolated coils, while hImpα1ΔIBB presents a ratio close to 1.0 indicating coiled coils (Figure 13B). These assumptions agree with the crystal structures for CLIC4 (PDB 2AHE) and hImpα1ΔIBB (3WPT).

(A)



Reduced		Structural Element	Oxidised	
Number of Residues	%		Number of Residues	%
137	58	α-helices	138	65
19	8	β-sheets	0	0
79	34	Random Coils	75	35
235	100	Full Length	213	100

(B)



Protein	MRE_{222}:MRE_{208}	Structure Type
CLIC1	0.80	Isolated helices
CLIC4	0.77	Isolated helices
hImpα1ΔIBB	1.01	Coiled coils

Figure 13: Far-UV CD spectra of the CLIC1 redox species, CLIC4 and hImpα1ΔIBB
(A) Spectra of reduced CLIC1 (—) as well as the oxidised dimer (—) and monomer (—). The inset table indicates the percentage of each structural element as determined from the PDB structures 1K0M for reduced CLIC1 and 1RK4 for oxidised CLIC1. **(B) S**pectra of reduced CLIC1 (—), CLIC4 (—) and hImpα1ΔIBB (—). The inset table indicates the predominant secondary structural conformation of each protein based on the ratio of MRE_{222}:MRE_{208}. All protein species were at a concentration of 2 μM in 50 mM sodium phosphate buffer pH 7.4. Each spectrum represents the average obtained from three separate samples at 20 °C. Resultant spectra were corrected for baseline signals.

3.4.2 Intrinsic Fluorescence

The aromatic ring structures of Phe, Tyr and Trp residues contribute to the UV fluorescence of proteins. A wavelength of 280 nm excites both Tyr and Trp, whereas 295 nm selectively excites Trp. Fluorescence emission spectra can comment on the general environment of the Trp residues. A bathochromic shift indicates that a Trp in a hydrophobic core region becomes solvent exposed. Conversely, a hypsochromic shift infers that a solvent exposed Trp becomes buried. Furthermore, a hyperchromic shift dictates an increase in the molar absorptivity while a hypochromic shift indicates a decrease (Vivian and Callis, 2001; Lakowicz, 2006). Fluorescence emission spectra were obtained for reduced CLIC1, oxidised CLIC1 (dimer and monomer), CLIC4 and hImpα1ΔIBB (Figure 14) with excitation at 280 nm.

CLIC1 contains a single Trp residue (Trp35) within the TMD. Both the reduced and oxidised monomers emit maximally at 341 nm whereas the oxidised dimer emits maximally at 339 nm (Figure 14A). This 2 nm hypsochromic shift in the oxidised dimer peak is accompanied by a 13 % decrease in emission intensity compared to the monomeric forms. The observed intensity difference is not the result of different protein concentrations between samples, as these were confirmed to be the same (see Section 2.3.1). These differences indicate that dimerisation most likely causes a conformation change that buries and quenches Trp35. CLIC4 and solvent exposed were found to emit maximally at 338 and 345 nm, respectively (Figure 14B).

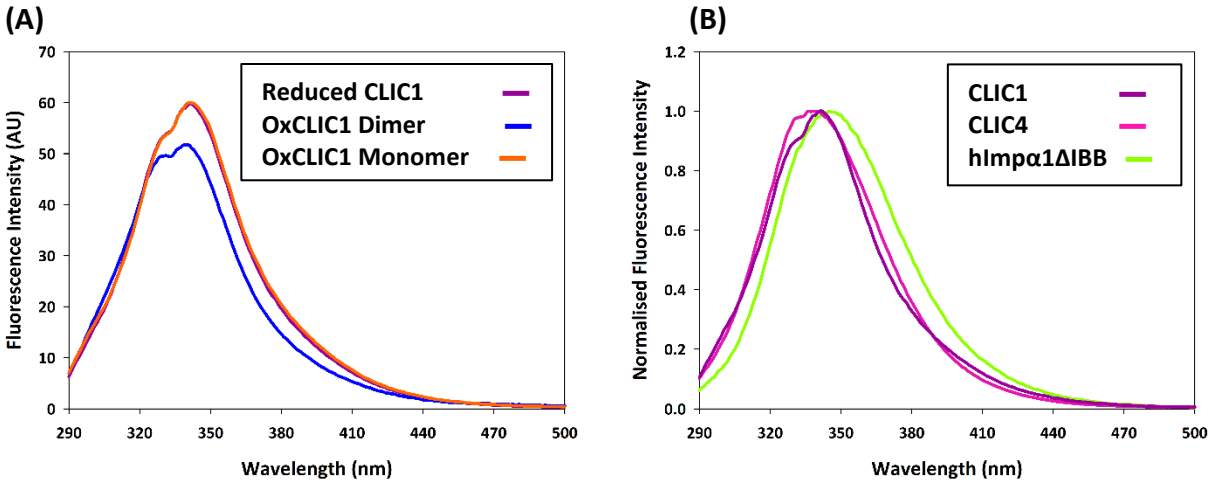


Figure 14: Fluorescence emission spectra of the CLIC1 redox species, CLIC4 and hImpα1ΔIBB
(A) Emission spectra of reduced CLIC1 (—) as well as the oxidised dimer (—) and monomer (—) excited at 280 nm. The spectra indicate that both monomeric species emit maximally at 341 nm with similar intensities. The oxidised dimer emits maximally at 339 nm with a 13 % reduced emission intensity compared to the monomeric species. **(B)** Emission spectra of reduced CLIC1 (—), CLIC4 (—) and hImpα1ΔIBB (—) excited at 280 nm. The spectra indicate that CLIC4 and hImpα1ΔIBB emit maximally at 338 and 345 nm, respectively. All protein species were at a concentration of 2 µM in 50 mM sodium phosphate buffer pH 7.4. Each spectrum represents the average obtained from three separate samples at 20 °C. Resultant spectra were corrected for baseline signals.

3.5 Oxidised CLIC1 Modifications

The effects of oxidation on CLIC1 structural and chemical modifications were assessed through thermal melts and LC-MS/MS.

3.5.1 Thermal Unfolding

Investigating protein unfolding provides details regarding protein stability. While temperature may act as a denaturant, thermal unfolding is irreversible and thus prevents thermodynamic parameters from being obtained. As such, only an apparent transition temperature ($T_{1/2}$) can be obtained. The $T_{1/2}$ depicts the temperature at which half the protein population is in the folded state ($N = N_0/2$), but not necessarily where $N = U = ½$, as determined by a reversible two-state unfolding process in equilibrium (Pace, 1986; Duy and Fitter, 2005). The thermal stability of the CLIC1 redox species were determined through far-UV CD by assessing the relative MRE at 222 nm (MRE_{222}) with a heating rate of 1 °C·min^{-1} (see Section 3.4.1).

The $T_{1/2}$ for each redox species was established by extrapolating the temperature at which the relative MRE_{222} = 0.5 (Figure 15). The $T_{1/2}$ for the reduced monomer, oxidised dimer and oxidised monomer were found to be approximately 54 °C, 42 °C and 50.5 °C, respectively. Therefore, the oxidised monomer appears less stable than the reduced monomer, while the oxidised dimer displays significantly less stability than both monomeric species.

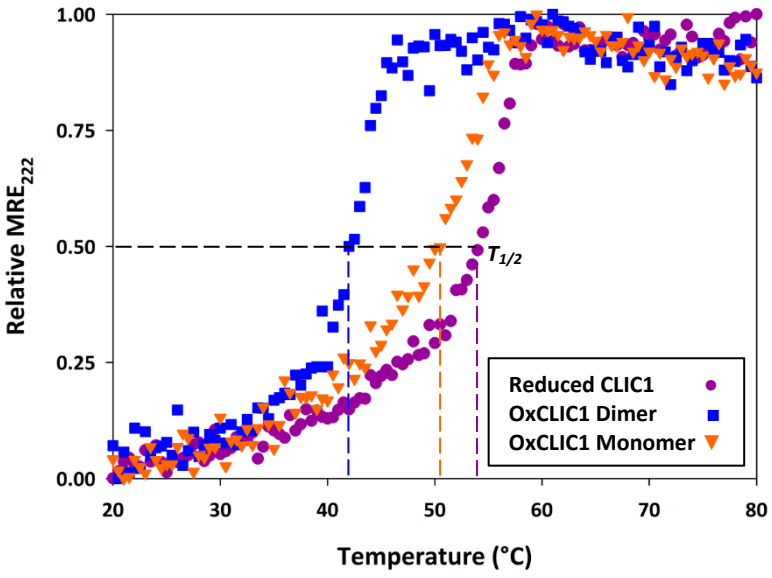


Figure 15: Thermal melts of the CLIC1 redox species
The relative MRE_{222} for reduced CLIC1 (●) as well as the oxidised dimer (■) and monomer (▼) was monitored with an increase in temperature. The $T_{1/2}$ for each species was extrapolated at relative MRE_{222} = 0.5. The $T_{1/2}$ for the reduced monomer is 54 °C, oxidised dimer is 42 °C and oxidised monomer is 50.5 °C. The two monomeric species were at a concentration of 20 µM, while the oxidised dimer was at 10 µM in 50 mM sodium phosphate buffer pH 7.4. Resultant spectra were corrected for baseline signals.

3.5.2 LC-MS/MS

Bottom-up proteomics utilise proteolytic digestion followed by MS analysis to identify and characterise proteins, amino acid sequences and/or PTMs. LC-MS/MS combines the proficiency of LC to physically separate multiple components in a mixture with the very sensitive analytical ability of MS/MS to provide the structural identity of the separated components. This technique can therefore recognise labelled residue sites by atomic mass unit increases in the molecular mass of peptic fragment ions (Pitt, 2009). CLIC1 contains six Cys residues: Cys24, Cys59, Cys89, Cys178, Cys191 and Cys223. To confirm which Cys residues form disulfide bonds within each of the CLIC1 redox species, a double alkylation experiment was performed and analysed via LC-MS/MS (see Section 2.6.2).

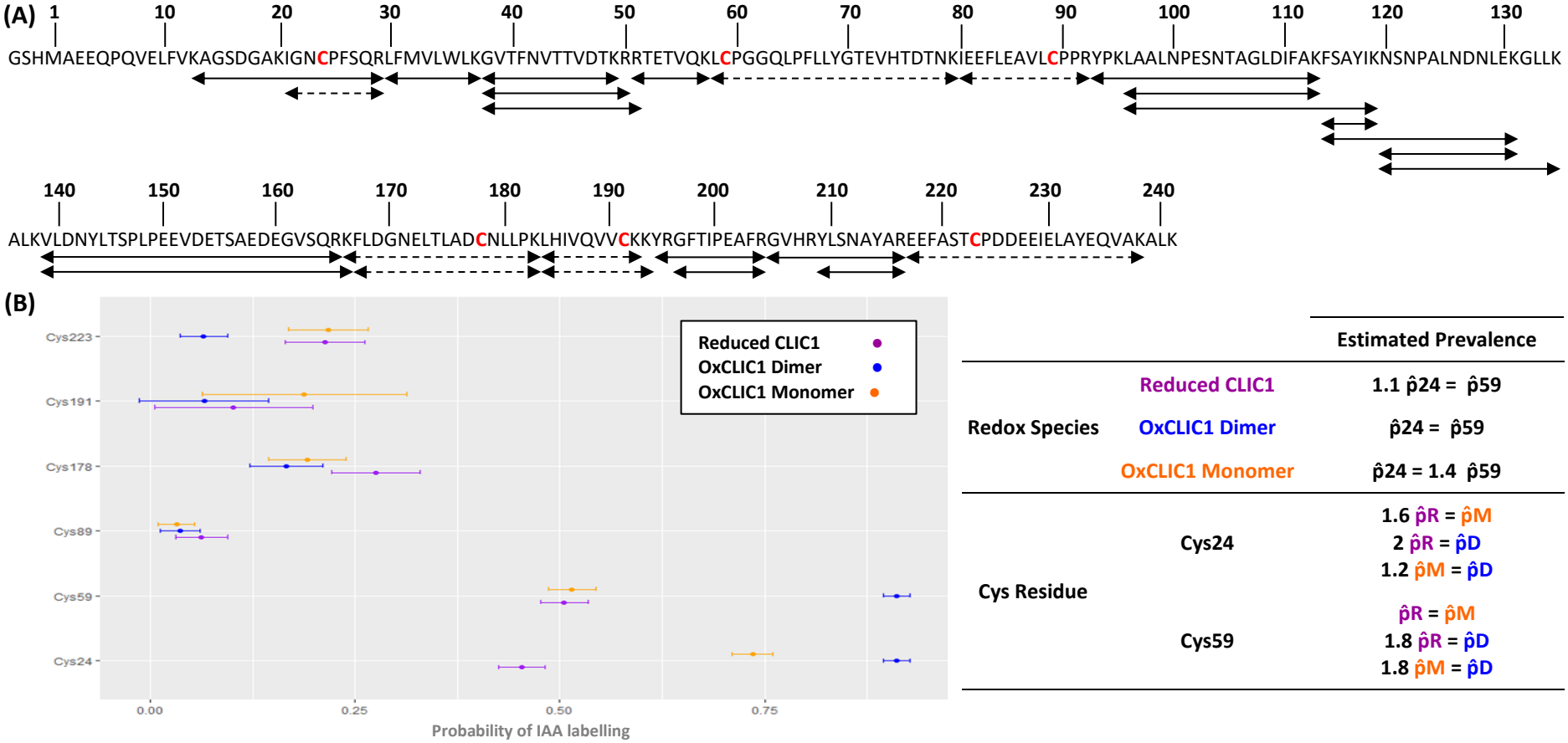
All redox samples were prepared from the same initial protein stock. Despite this fact, each sample replicate is assumed to be independent as using the same protein production methods would result in the same intact folded protein (see Appendix A Figure 4). Each independent sample was subjected to two alkylating agents, namely NEM and IAA. NEM was first added to the samples in order to label free Cys residues. Thereafter, a reducing agent was added, followed by IAA which would label Cys residues that were disulfide bonded prior to the aforementioned reduction step. The labelled intact protein samples underwent proteolytic digestion by trypsin. A total of 27 unique peptides common to all three redox species were identified, covering 92.5 % of the protein sequence and all Cys residues. From these peptides, eight encompassing the Cys residues were chosen for analysis (Figure 16A). All Cys residues from each redox species were found to be alkylated, with the proportion of unlabelled Cys residues appearing negligible.

The proportion of IAA labelling for each Cys residue within each sample was calculated. The two factors, relative location of Cys residues within a folded protein and the oxidation state of the protein, were considered as explanatory variables for disulfide bondage. The observed IAA labelling proportions were all real numbers between 0 and 1, with non-constant sample variance across the levels of the two explanatory variables. It is therefore clear that the response IAA labelling proportion does not have a normal distribution. While transformations to ordinary least squares regression are often applied to overcome these challenges, this results in biased estimates and difficulties in interpretation. Therefore, a beta regression model was chosen which assumes that the continuous proportion has a beta distribution. Furthermore, a logit link function applied to the expected proportion is set equal to a function of the predictors (Cys and oxidation) which is linear in the unknown parameters. The model also includes a precision parameter which depends on one of predictors through a link function (see Appendix A).

The fitted model with interaction terms is given in Appendix A and was used to obtain for each combination of Cys and oxidation levels, the fitted (or estimated) odds of IAA labelling to NEM labelling, the fitted probability of IAA labelling as well as the fitted prevalence of IAA labelling. The expected IAA labelling proportion is considered to be the probability that IAA labelling will take place in the CLIC1 protein. In Figure 16B, each error bar gives an approximate 68 % confidence interval for the probability of IAA labelling (with half width one fitted standard error from the fitted expected IAA labelling proportion) for each combination of Cys and oxidation levels. Thus, one can be 68 % confident that the true but unknown probability of IAA labelling will be between the lower and upper limits of a given confidence interval.

Cys89, Cys178, Cys191 and Cys223 predominantly exhibit NEM labelling within all three redox species (Figure 16B). This indicates that, while these Cys residues may experience some disulfide bond shuffling, these residues do not form significant stable disulfide bonds regardless of the oxidation state. The fitted probability of IAA labelling for Cys24 and Cys59, respectively, are as follows: approximately 0.45 and 0.51 for the reduced monomer, approximately 0.74 and 0.51 for the oxidised monomer and both are approximately 0.91 for the oxidised dimer. The fitted probability of IAA labelling was also used to estimate the prevalence of IAA labelling for Cys24 relative to Cys59 for each redox species, as well as for the redox species relative to one another for Cys24 and Cys59, respectively (Figure 16B table inset).

If only Cys24 and Cys59 formed disulfide bonds, one would expect these residues to have approximately the same estimated prevalence within each redox species. The estimated prevalence for all three species appears within an acceptable range and hence supports that probably only Cys24 and Cys59 form significant disulfide bonds. Comparing the reduced and oxidised monomer, Cys24 IAA labelling is estimated to be 60 % more prevalent in the latter, while Cys59 IAA labelling displays approximately the same prevalence for both. Thus, both Cys residues in the reduced monomer and Cys59 in the oxidised monomer appear to experience the same amount of disulfide bond shuffling, whereas Cys24 in the oxidised monomer undergoes less shuffling. Considering the oxidised dimer, Cys24 IAA labelling is estimated twice as prevalent than in the reduced monomer, but only estimated 20 % more prevalent than in the oxidised monomer. On the other hand, the prevalence of Cys59 IAA labelling is estimated to be 80 % higher in the oxidised dimer than the other two species. The higher prevalence in Cys24 and Cys59 IAA labelling in the oxidised dimer, compared to the reduced monomer, suggests more stable disulfide bonds in the former. While Cys59 in the oxidised dimer is also estimated to be twice as prevalent as in the oxidised monomer, Cys24 displays similar prevalence in both. Therefore, Cys24 in the oxidised monomer displays similar prevalence to Cys24 and Cys59 in the oxidised dimer, with Cys59 in the oxidised monomer exhibiting greater disulfide bond shuffling than all the aforementioned.



		Estimated Prevalence
Redox Species	Reduced CLIC1	1.1 p̂24 = p̂59
	OxCLIC1 Dimer	p̂24 = p̂59
	OxCLIC1 Monomer	p̂24 = 1.4 p̂59
Cys Residue	Cys24	1.6 p̂R = p̂M 2 p̂R = p̂D 1.2 p̂M = p̂D
	Cys59	p̂R = p̂M 1.8 p̂R = p̂D 1.8 p̂M = p̂D

Figure 16: LC-MS/MS analysis of the CLIC1 redox species

(A) CLIC1 sequence with the peptic fragments common to all redox species (**—**). All Cys residues were covered by at least one peptic fragment with those used for analysis indicated (- - -). **(B)** Fitted probability of IAA labelling for each Cys and reduced CLIC1 (●), oxidised dimer (●) and oxidised monomer (●) grouping. Each error bar gives an ~68 % confidence interval for the probability of IAA labelling. Inset table shows estimated prevalence of IAA labelling for Cys24 relative to Cys59 for the redox species, as well as for the redox species relative to one another for Cys24 and Cys59. The estimated probability for reduced CLIC1, oxidised dimer, oxidised monomer, Cys24 and Cys 59 are p̂R, p̂D, p̂M, p̂24 and p̂59, respectively.

3.6 CLIC Peptides and hImpα1ΔIBB Interaction Studies

Particular interactions between the CLIC peptides and hImpα1ΔIBB were assessed through molecular docking, fluorescence quenching and ITC.

3.6.1 Molecular Docking

Interactions between a ligand and receptor protein can be modelled by means of molecular docking. Modelled structures provide information on the preferred orientation of ligand binding, important complex bonds and the strength of association. The ligands designed for this docking study were Pep4 VKVVA**KKYR**N, Pep1 VQVVC**KKYR**G, Pep1_A VQVVA**KKYR**G and Pep1_S VQVVS**KKYR**G (NLS residues in bold and underlined). The receptor proteins consisted of mImpα1ΔIBB and hImpα1ΔIBB. The FlexPepDock server was used to model mImpα1ΔIBB bound to Pep4 and Pep1, respectively, as well as hImpα1ΔIBB bound to Pep4, Pep1, Pep1_A and Pep1_S, respectively. The top ten modelled poses for each complex, based on the Rosetta scores, was retrieved for analysis (Figure 17).

The Rosetta score indicates the total energy for a docked model, while the rmsBB value is the RMSD calculated for the peptide backbone heavy atoms of the docked model relative to the initially submitted structure (see Appendix B Table 1). Considering the Rosetta scores for the models involving Pep4 and Pep1, the scores for the mImpα1ΔIBB complexes range from -772.723 to -768.106, while the scores for the hImpα1ΔIBB complexes range from -717.981 to -709.717. Based on these values, the strength of association to Pep4 and Pep1 appear greater for mImpα1ΔIBB than hImpα1ΔIBB. This may be attributed to slight conformational and sequence differences between the ImpαΔIBB major pockets (Figure 18A). The Rosetta scores for hImpα1ΔIBB bound to Pep1, Pep1_A and Pep1_S, respectively, display a similar range of values, thus signifying similar strengths of association. The rmsBB values for Pep4 and Pep1 range from 1.491 to 2.690 when docked to mImpα1ΔIBB and from 1.880 to 4.881 when docked to hImpα1ΔIBB. Hence, when these peptides are docked to mImpα1ΔIBB, the peptides undergo less conformational changes than when docked to hImpα1ΔIBB. This is interesting as the initial input structures resemble the 3OQS PDB structure of truncated Pep4 (VA**KKYR**N) bound to mImpα1ΔIBB. The rmsBB values for hImpα1ΔIBB bound to Pep1, Pep1_A and Pep1_S, respectively, also fall within a similar range.

The top poses of Pep4 and Pep1 bound to mImpα1ΔIBB were aligned to truncated Pep4 in the 3OQS PDB structure (Figure 18B). The backbones of the sequences appear well aligned with only some side-chains deviating from one another. When comparing the top poses of Pep4 and Pep1 bound to hImpα1ΔIBB, the peptides align form Val1 to Lys6, after which the last four residues (three of which belong to the NLS) deviate (Figure 18C). Lastly, the alignment of the top WT and mutant Pep1 poses bound to hImpα1ΔIBB display good alignment with only the last three residues exhibiting some degree of variation (Figure 18D).

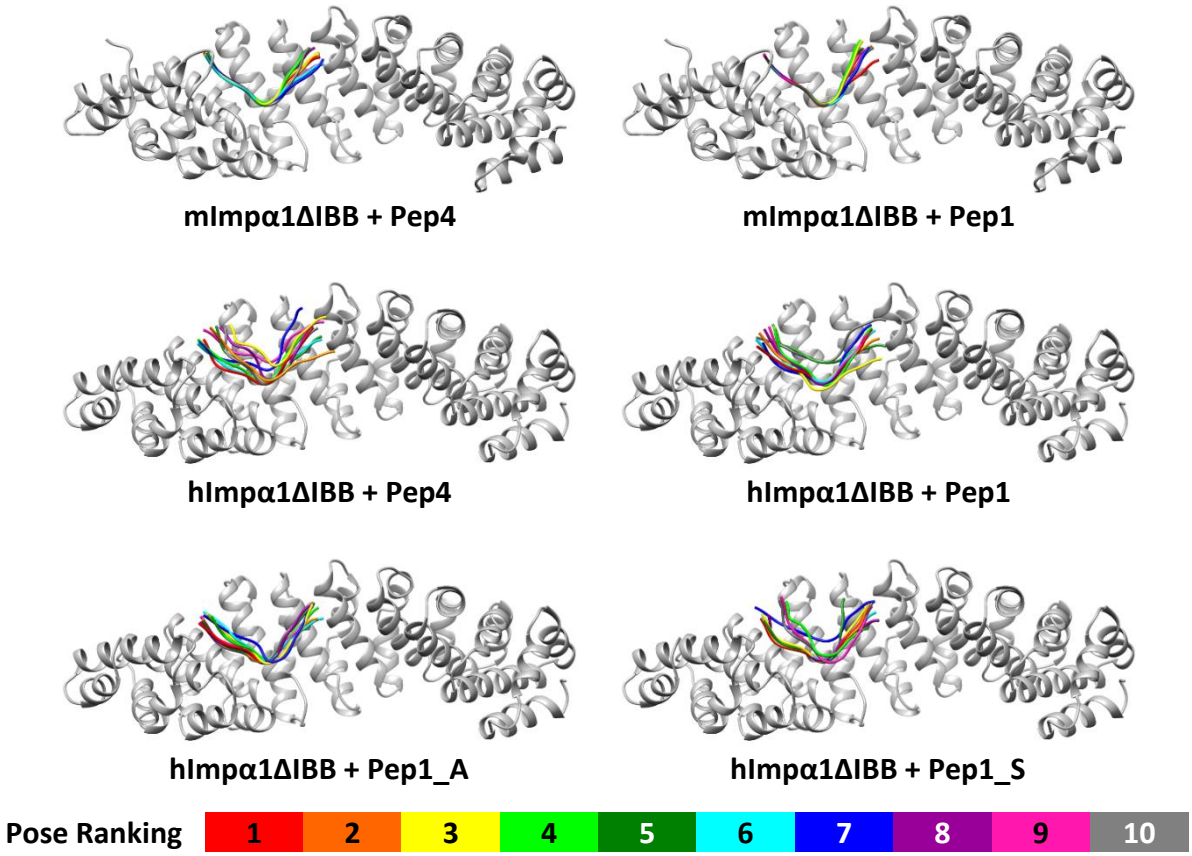


Figure 17: Molecular docking of ImpαΔIBB with the CLIC NLS peptides
Top ten poses of Pep4 and Pep1 docked to mImpα1ΔIBB and top ten poses of Pep4, Pep1, Pep1_A and Pep1_S docked to hImpα1ΔIBB as determined by the Rosetta score. Cartoon protein figures were generated using the UCSF Chimera Molecular Graphics system, V1.12.

The interactions formed by Pep4 and Pep1 to mImpα1ΔIBB in the top docking poses were compared to those formed by truncated Pep4 in the 3OQS PDB structure (Figure 19A). Few interactions appear for the five residues preceding the NLS, with only the Ala/Cys5 backbone establishing a conserved interaction. However, the NLS forms a number of interactions with nine conserved interactions stemming from the three basic residues. While six of these conserved interactions utilise the peptide backbone, three derive from the Lys6 side-chain. The interactions formed by Pep4, Pep1, Pep1_A and Pep1_S to bind hImpα1ΔIBB in the top docking poses were also assessed (Figure 19B). These peptides only form two conserved interactions using the five residues preceding the NLS, including the Ala/Cys/Ser5 backbone interaction conserved when binding mImpα1ΔIBB. As in the case with mImpα1ΔIBB, the peptides form a number of interactions utilising the NLS basic residues, with the Lys6 side-chain interaction being conserved. For interactions from the NLS basic residues, Pep4 shares two bonds with Pep1, one with Pep1_A and interestingly seven with Pep1_S. Furthermore, Pep1 shares three bonds with Pep1_A and four with Pep1_S for the basic NLS residues. While water-mediated interactions could not be assessed, the 3OQS PDB structure indicates the presence of ten such interactions, half of which involve the core NLS. Based on the significant similarity detected in bond formation between the various peptides and the two ImpαΔIBB proteins, these water-mediated interactions are also likely to be conserved.

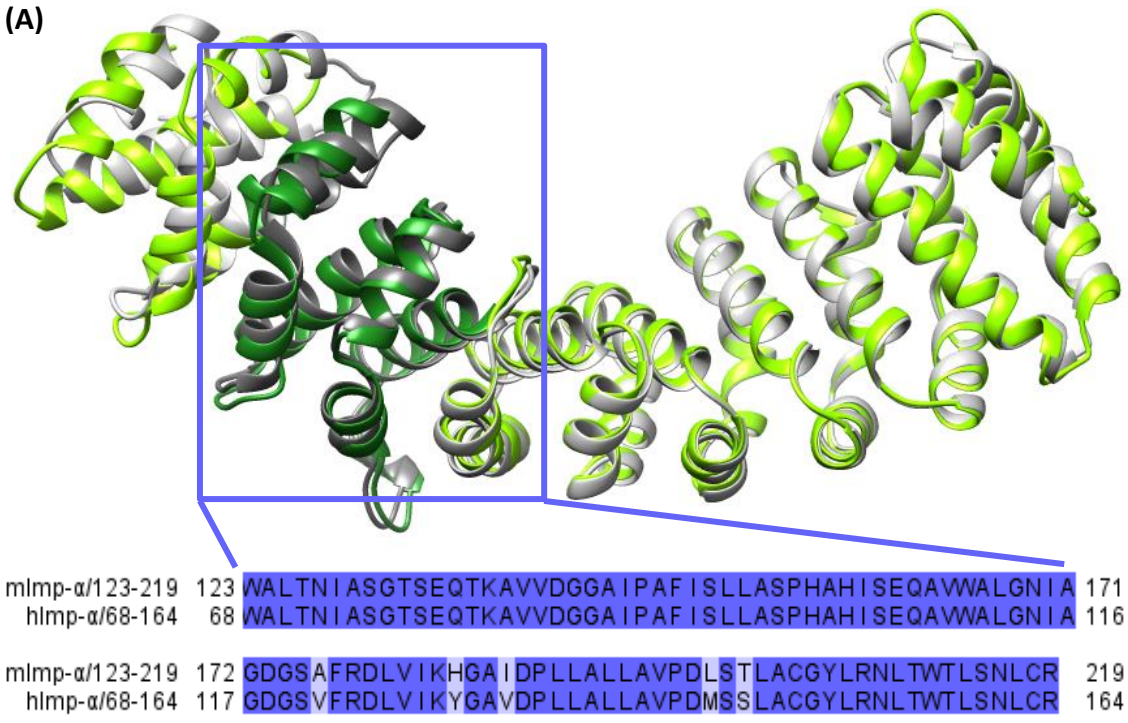
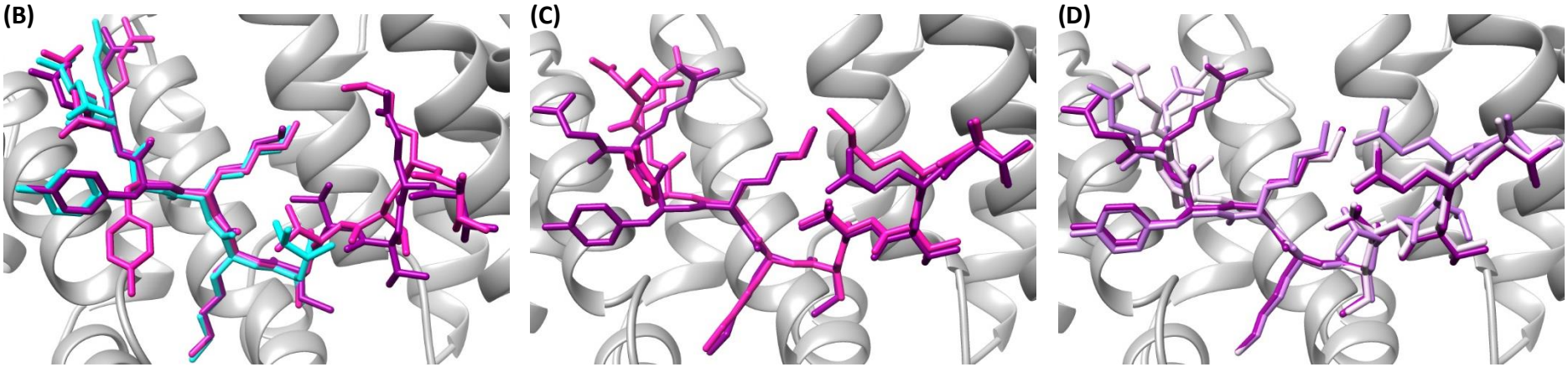


Figure 18: Structural and CLIC NLS binding differences between mImpα1ΔIBB and hImpα1ΔIBB

(A) Conformational and sequence differences between mImpα1ΔIBB (grey) (PDB 3OQS) and hImpα1ΔIBB (green) (PDB 3WPT) are shown. The major binding pocket of Impα for interacting with cargo NLSs is highly conserved. The sequence alignment was performed using Clustal Omega (Madeira et al., 2019) and formatted with Jalview 2.11.0 to show percentage identity (Waterhouse et al., 2009). **(B)** Binding of mImpα1ΔIBB by truncated Pep4 from the 3OQS structure (cyan) as well as the top Pep4 (pink) and Pep1 (dark purple) poses. **(C)** Binding of hImpα1ΔIBB by the top Pep4 (pink) and Pep1 (dark purple) poses. **(D)** Binding of hImpα1ΔIBB by the top Pep1 (dark purple), Pep1_A (purple) and Pep1_S (lilac) poses. Cartoon protein figures were generated using the UCSF Chimera Molecular Graphics system, V1.12.



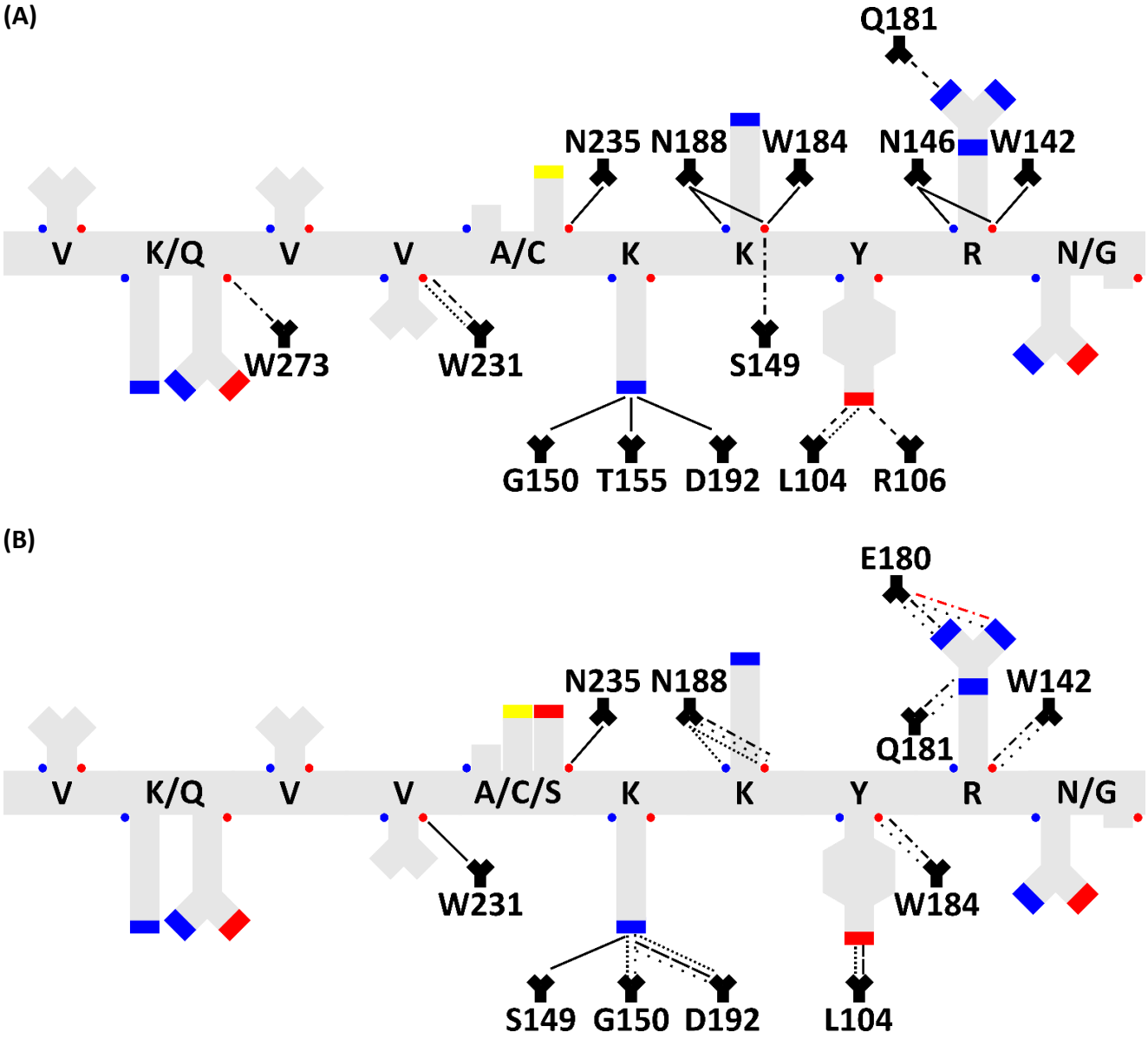


Figure 19: Schematic of the interactions between ImpαΔIBB and the CLIC NLS peptides
(A) Binding between mImpα1ΔIBB and the CLIC NLS sequences truncated Pep4, Pep4 and Pep1. **(B)** Binding between hImpα1ΔIBB and the CLIC NLS sequences Pep4, Pep1, Pep1_A and Pep1_S.

3.6.2 Peptide Synthesis

In order to experimentally assess the binding interactions of the CLIC NLSs with hImpα1ΔIBB, peptides of each CLIC NLS had to be synthesised. Mutation of the Cys residue in Pep1 was essential as to prevent peptide dimerisation. Based on the molecular docking studies, Pep1_S appeared to resemble the binding of Pep1 more closely to hImpα1ΔIBB than Pep1_A. As such, the 10-residue Pep1_S VQVVS**KKYR**G and Pep4 VKVVA**KKYR**N structures (NLS residues in bold and underlined), containing an acetylated N-terminus and amidated C-terminus, were synthesised using a solid phase continuous flow system by GLBiochem (Shanghai, China). Following synthesis, GLBiochem utilised RP-HPLC to confirm that the purity of each peptide was >95 %.

3.6.3 Fluorescence Quenching

The hImpα1ΔIBB protein contains a total of seven Tyr and seven Trp residues. From these aromatic residues, the major pocket houses two Tyr and three Trp residues while the minor pocket only houses two Trp residues. Based on the molecular docking study (see Section 3.6.1), the CLIC NLS peptides form hydrogen bonds with the hImpα1ΔIBB major pocket Trp residues and not the Tyr residues (Figure 19). Therefore, when assessing the binding of these peptides to hImpα1ΔIBB through fluorescence spectroscopy, quenching of Trp142, Trp184 and Trp231 should be observed (Figure 20A). Monitoring this quenching of hImpα1ΔIBB with an increase in peptide concentration would in turn allow apparent K_d values to be determined for hImpα1ΔIBB binding to Pep1_S and Pep4, respectively.

Excitation of the hImpα1ΔIBB CLIC NLS peptide complexes at a wavelength of 295 nm was considered since this would selectively excite the hImpα1ΔIBB Trp residues. Furthermore, this would exclude any emission produced by the peptide Tyr residue which is essential for peptide concentration determination and could be important for hImpα1ΔIBB binding (Mynott *et al.*, 2011). However, an excitation wavelength of 280 nm had to be used instead, since excitation at 295 nm did not provide emission spectra with high enough intensity values discernible between each peptide increment. Adjustments to the excitation and emission band widths did not resolve this problem. The fluorescence emission for each sample increment thus required correction for peptide contribution by deducting the emission of the peptide increment from the corresponding hImpα1ΔIBB peptide complex reading. The corrected spectra of all complex increments for both Pep1_S and Pep4, respectively, exhibited a maximal peak at 314 nm with a steady decrease in intensity as peptide concentration increased (Figure 20B and C). The absence of a wavelength shift in these spectra indicate that peptide binding likely does not alter the solvent exposure of the Trp residues in the hImpα1ΔIBB major pocket. Furthermore, the accompanying hypochromic shift can be attributed to fluorescence quenching induced by peptide binding. Finally, the fitted ligand binding one site saturation curve for each peptide determined that the apparent K_d values for Pep1_S and Pep4 are 231.3 ± 11.0 μM and 308.8 ± 15.1 μM, respectively (Figure 20D).

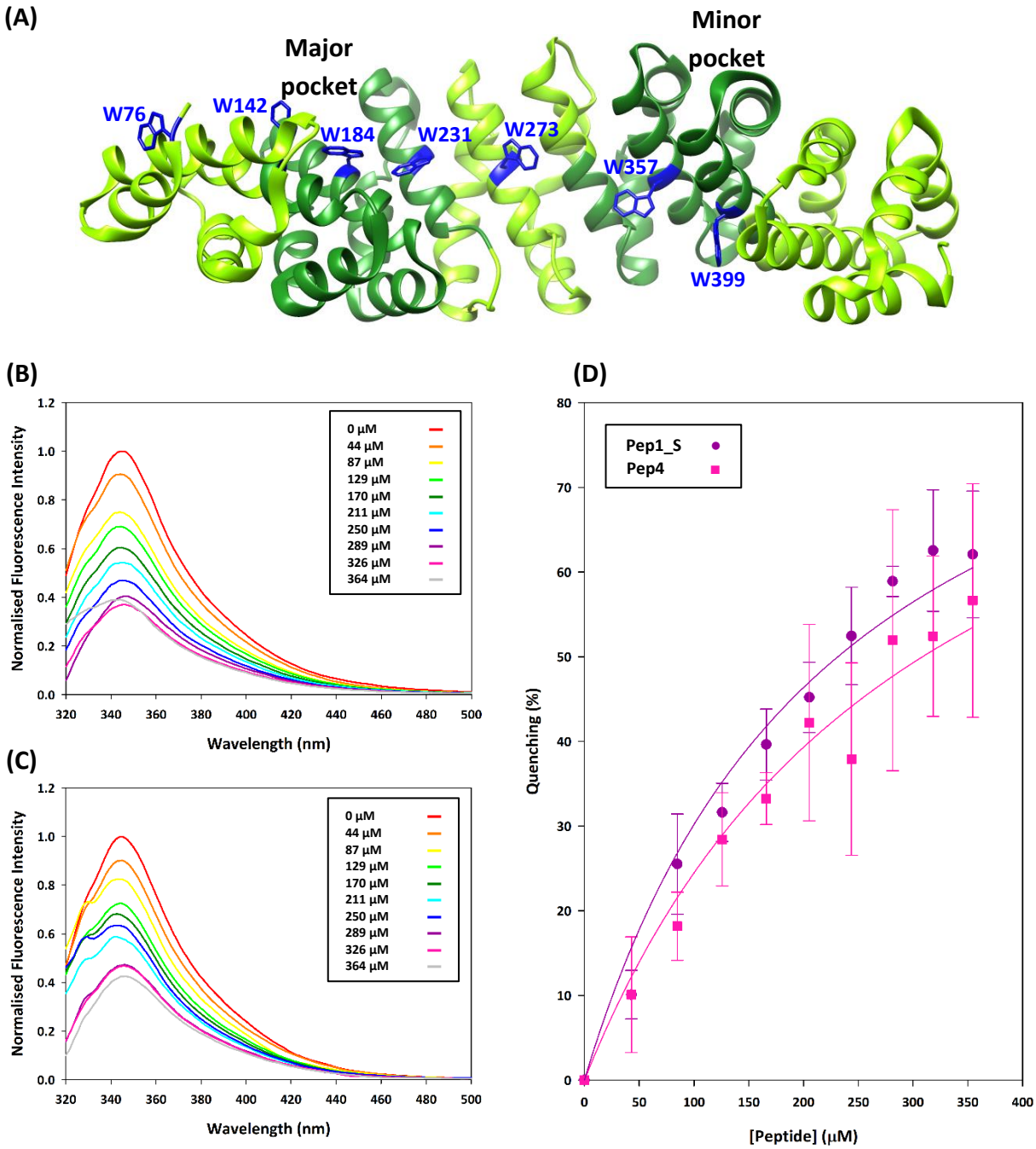


Figure 20: Fluorescence quenching of hImpα1ΔIBB by Pep1_S and Pep4
(A) Localisation of Trp residues (blue) in hImpα1ΔIBB relative to the two binding pockets (forest green). PDB file 3WPT was used to generate the cartoon figure using the UCSF Chimera Molecular Graphics System, V1.12. **(B)** Corrected fluorescence emission spectra of hImpα1ΔIBB with increasing concentrations of Pep1_S excited at 280 nm. All spectra emit maximally at 341 nm. **(C)** Corrected fluorescence emission spectra of hImpα1ΔIBB with increasing concentrations of Pep4 excited at 280 nm. All spectra emit maximally at 341 nm. All protein species were at a concentration of 2 µM in 50 mM sodium phosphate buffer pH 7.4. Each spectrum represents the average obtained from three separate samples at 20 °C and were corrected for peptide and baseline signals. **(D)** Percentage quenching of hImpα1ΔIBB with increasing concentrations of Pep1_S (●) and Pep4 (■), respectively. A ligand binding one site saturation curve was fitted for each peptide. These curved indicate K_d values of 231.3 ± 11.0 µM for Pep1_S (R^2 = 0.978) and 308.8 ± 15.1 µM for Pep4 (R^2 = 0.971).

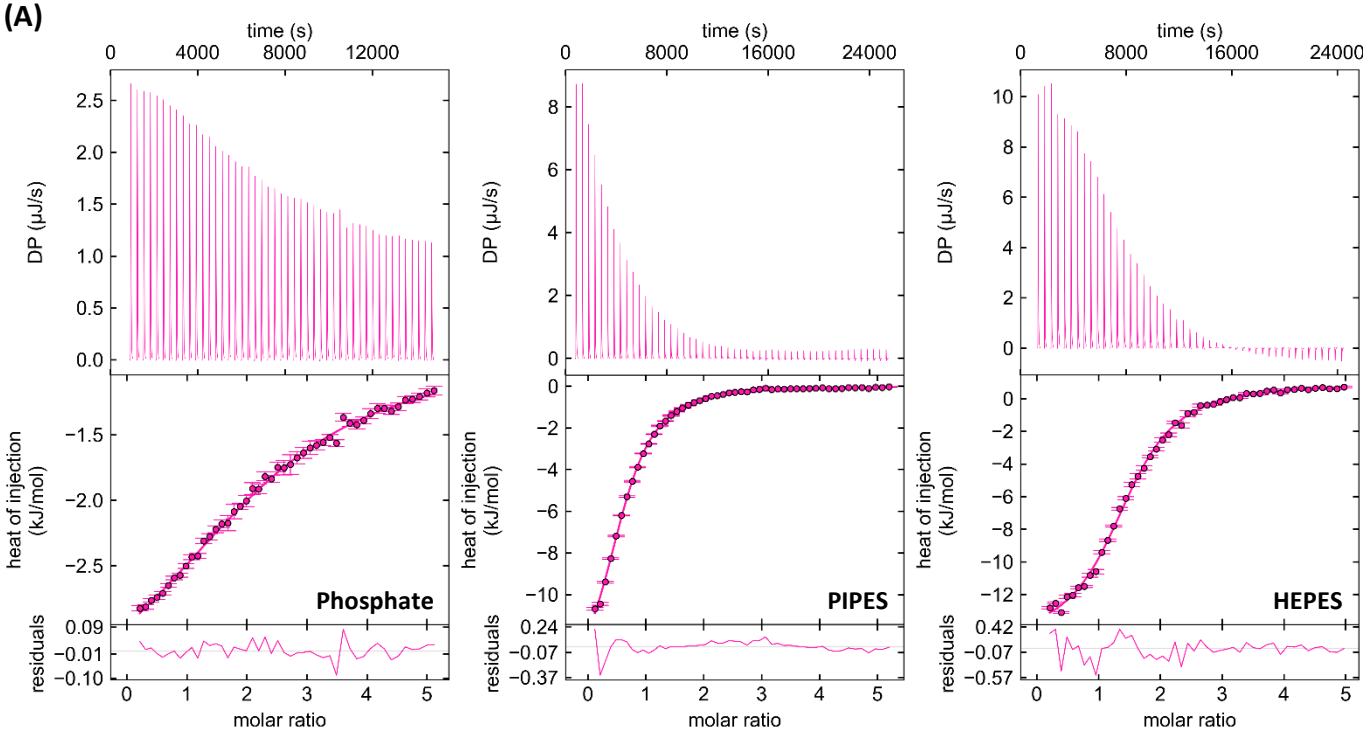
3.6.4 ITC

The thermodynamic parameters of a ligand-protein complex interaction can be elucidated through the use of ITC. An isothermal titration calorimeter consists of two identical cells, known as the reference and sample cells, enclosed by an adiabatic jacket. The reference cell is filled with water or buffer while the sample cell contains the macromolecule of interest. The ligand is titrated in precise aliquots into the sample cell and either results in heat evolved or taken up. Exothermic reactions increase the sample cell temperature, thus causing a decrease in feedback power to the sample cell in order to maintain a constant temperature with the reference cell. Conversely, endothermic reactions decrease the sample cell temperature, thus causing an increase in feedback power to the sample cell. The power required to uphold equal temperatures between the two cells is measured in real time and generates a thermogram. When the heat signal approaches zero, the limiting reactant has reached saturation. Integration of the peaks from the thermogram provide data points which are fitted to an isotherm. The isotherm quantifies the interaction stoichiometry (n), binding affinity (K_d) and enthalpy (ΔH), which in turn are used to calculate the Gibbs free energy (ΔG) and entropy (ΔS) (Duff *et al.*, 2011). Furthermore, determining the ΔH of the reaction in the presence of different buffers can establish whether proton linkage effects occur (Baker and Murphy, 1996).

Only the interaction between Pep4 and hImpα1ΔIBB was assessed due to the insoluble nature of Pep1_S in the required assay buffers. Thermograms, isotherms and residual plots were obtained for the Pep4 and hImpα1ΔIBB interaction in the presence of three different buffers, namely phosphate, PIPES and HEPES (Figure 21A). The resulting isotherms provided the thermodynamic parameters for each respective buffer of which the ΔH_{obs} was used to examine any proton linkage effects (Figure 21B and C).

The K_d values of the interaction within each buffer appear significantly lower than the 308.8 ± 15.1 μM established through fluorescence quenching (Figure 20D). This noteworthy variance is most likely due to ITC being a far more sensitive technique than fluorescence spectroscopy. However, both techniques indicate K_d values within the μM range suggesting weak binding, since strong biomolecular interactions usually fall within the nM range. While the ΔH and ΔS values vary between buffers, ΔH remains negative and ΔS positive for all three buffers. The ΔH and ΔS values also seem to compensate for one another as the ΔG remains constant among the buffers. It is therefore concluded that the Pep4 to hImpα1ΔIBB is an exothermic reaction that is both enthalpically and entropically driven under these experimental conditions (Figure 20B). Proton linkage in the Pep4 hImpα1ΔIBB interactions was determined by plotting the ΔH_{obs} values for each buffer versus the respective ΔH_{ion} values. The slope of the plot (N_{H^+} = -0.908) establishes that a proton linked effect does occur during this interaction by indicating that one proton is taken up by the buffer. Furthermore, the true interaction binding enthalpy (ΔH_0) is determined to be -4.904 kJ·mol^{-1} (Figure 20C).

(A)



(B)

Buffer	K_d (µM)	ΔH (kJ·mol^{-1})	$-T\Delta S$ (kJ·mol^{-1})	ΔG (kJ·mol^{-1})
Phosphate	106.3 ± 4.0	-6.7 ± 0.7	-15.6 ± 0.8	-22.3 ± 0.1
PIPES	67.7 ± 3.7	-15.9 ± 1.0	-7.5 ± 1.1	-23.4 ± 0.1
HEPES	37.4 ± 7.4	-23.3 ± 0.3	-1.6 ± 0.2	-24.9 ± 0.5

(C)

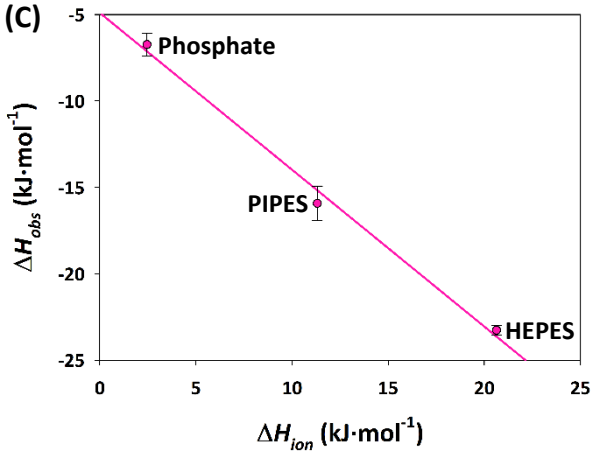


Figure 21: Titration of Pep4 into hImpα1ΔIBB in the presence of various buffers
(A) Representative thermograms, isotherms and residual plots for each Pep4 titration into hImpα1ΔIBB in either a phosphate, PIPES or HEPES buffer. SEDPHAT was used to fit an A + B ↔ AB hetero-association model to the data points for each buffer titration isotherm. **(B)** Summary of the mean thermodynamic parameters obtained for each buffer titration. **(C)** Proton linkage for Pep4 binding to hImpα1ΔIBB at 20 °C and pH 7.4. The observed enthalpy is plotted versus the ionisation enthalpy of the buffer. The buffer used for each value of ΔH_{ion} is labelled. The data points are fitted using linear regression (R^2 = 0.994). The equation of the fitted line is y = - 0.908x - 4.904. The y-axis intercept is ΔH_0 which indicates the true interaction binding enthalpy. The slope is N_{H^+} which signifies the number of protons released by the buffer during the reaction.

CHAPTER 4. DISCUSSION

The literature suggests that the redox regulation of cellular compartments likely impacts on the overall cellular distribution and functionality of the CLIC protein family. While CLICs exhibit varying numbers of Cys residues, three appear conserved and correspond to Cys24, Cys178 and Cys223 in CLIC1. The highly conserved Cys24 acts as a redox-active site and may form a disulfide bond with Cys59 in CLIC1 under strongly oxidising conditions. Despite this disulfide bond appearing to be unique to CLIC1, since Cys59 is unconserved, disulfide bonding induces dimerisation which further promotes membrane insertion (Littler *et al.*, 2010; Singh, 2010; Nguyen *et al.*, 2016). Therefore, understanding the mechanisms underpinning oxidation induced CLIC1 dimerisation may shed light on the role of oxidation in CLIC membrane insertion. The redox regulation of CLICs also take the form of Cys *S*-nitrosylation. The *S*-nitrosylation of Cys234 in CLIC4 is of particular interest, since this event induces conformational changes which permit CLIC4 nuclear translocation by allowing the NLS to bind Impα (Malik *et al.*, 2010). While the nuclear import pathway is established for CLIC4, more work is required to further understand the molecular mechanisms involved in the CLIC4 NLS binding Impα. Also, the question remains whether an exposed CLIC1 NLS can bind Impα and therefore follow the same nuclear import pathway as CLIC4.

4.1 CLIC1 Oxidation Induces Partial Dimerisation

While Cys residues generally remain protonated under physiological conditions, a significant increase in H_2O_2 concentrations under oxidative stress conditions (1-10 mM) may result in modifications of the redox sensitive thiol group (Cumming *et al.*, 2004; Averaimo *et al.*, 2010; Baxter *et al.*, 2014; Nguyen *et al.*, 2016). A study performed by Littler *et al.* (2004) indicated that untreated CLIC1 exists in the monomeric form, while exposing CLIC1 to a relatively high H_2O_2 concentration (2 mM) results in a heterogenous mixture of oxidised dimer and monomer, separable via SEC. To assess the results obtained by Littler *et al.* (2004), SEC was employed to confirm the oligomeric state of CLIC1 in the presence of both reducing and highly oxidising conditions. Under reducing conditions, achieved by the addition of 1 mM TCEP, CLIC1 exists as a single species with an approximate mass of 43 kDa (Figure 12A and D). While the reduced CLIC1 peak exhibits a mass larger than the expected 27 kDa of the monomeric form (see Section 3.2.2), the calculated mass is not large enough to signify a dimeric state. Therefore, reduced CLIC1 exists wholly monomeric with a larger than expected hydrodynamic volume. When treating CLIC1 with 2 mM H_2O_2, the protein elutes as two distinct species with the sizing of the first and second peaks corresponding to 98 kDa and 39 kDa, respectively (Figure 12B and D). The second oxidised peak agrees with the sizing of the reduced peak, indicating that this species also exists as a monomer with a similar hydrodynamic volume. Furthermore, the first oxidised peak displays a mass just over double that of the monomeric species, and thus likely resembles a dimer with a larger than expected hydrodynamic volume. This larger hydrodynamic volume corresponds with the CLIC1 dimer crystal structure (PDB 1RK4) (Littler *et al.*, 2004), which indicates an oblong shape that would most likely present anomalous behaviour in a SEC column.

4.2 CLIC1 Dimerisation Stabilises Oxidative Conformational Changes

The crystal structures of the CLIC1 reduced monomer (PDB 1K0M) and oxidised dimer (PDB 1RK4) indicate several significant structural alterations that provide insight into the process of oxidation induced dimerisation in CLIC1. In the presence of strongly oxidising conditions, Cys24 and Cys59 form an intramolecular disulfide bond resulting in radical N-terminal domain structural rearrangements (Figure 22A). These changes entail the dissolution of four β-strands and subsequent extension of helix α1 out towards the C-terminus by two α-helical turns. Consequently, flat hydrophobic patches become exposed which lead to the reversible transition from a monomeric to a non-covalent dimeric form. Around two thirds of the oxidised CLIC1 form a dimer with the remainder existing as a monomer (Singh and Ashley, 2006; Singh, 2010; Nguyen *et al.*, 2016). However, the available crystal structures only reflect on the static state and not necessarily the state in solution of the reduced monomer and dimer. These structures also do not provide clarity on the conformation of the oxidised monomer. This study therefore aimed to characterise the reduced monomer, as well as the oxidised dimer and monomer, with regards to structure, stability and disulfide bonding in solution.

Following separation of the oxidative species through SEC (see Section 4.1), the resultant samples underwent structural characterisation. Secondary structural analysis confirmed that the reduced monomer and oxidised dimer exhibit a predominantly α-helical structure with the helices forming isolated coils. This is as expected from the respective crystal structures (PDB 1K0M and 1RK4). Examination of the oxidised monomer indicated that this species also shares the secondary structural elements observed in both the reduced monomer and oxidised dimer (Figure 13A).

While the overall secondary structure appears mainly unaffected by oxidation, significant tertiary structural changes transpire during oxidation induced dimerisation. Intrinsic fluorescence spectroscopy was used to probe tertiary structure, since the N-terminal helix α1 contains Trp35 which is the only CLIC1 Trp residue. The emission spectra of the reduced and oxidised monomeric species overlaid, while the oxidised dimer displayed both a hypsochromic shift and decrease in emission intensity relative to the monomeric species (Figure 14A). The hypsochromic shift suggests a reduction in Trp35 solvent exposure, while the quenching effect may be caused by exposure to amino acids capable of quenching Trp. The oxidised dimer crystal structure (PDB 1RK4) indicates that Trp35 is less solvent exposed and exists in close approximation to Glu217 and Glu218 (Figure 22B and C). Glu residues are particularly known quenchers of Trp via excited state electron transfer (Chen and Barkley, 1998). Therefore, the environment of Trp35 in the oxidised monomer seems similar to that in the reduced monomer and hence does not undergo the same N-terminal changes observed in the oxidised dimer.

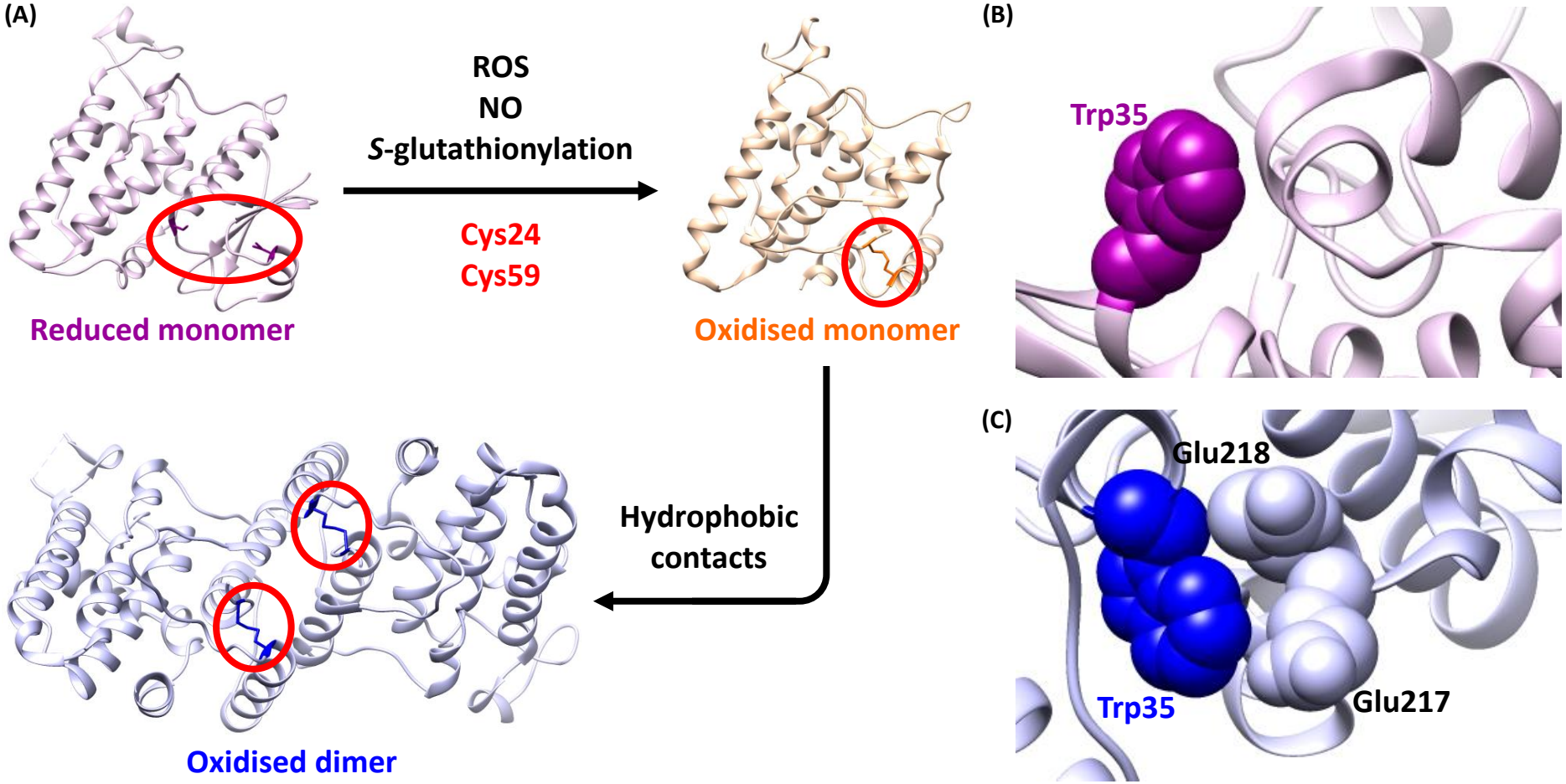


Figure 22: Effects of highly oxidising conditions on CLIC1 structure
(A) Oxidation induced dimerisation of CLIC1. Under highly oxidising conditions, Cys24 and Cys59 in the reduced monomer (PDB 1K0M) form a disulfide bond. This causes a structural rearrangement of the N-terminal domain which results in the exposure of hydrophobic patches. Exposed hydrophobic patches from two oxidised monomers may then form contacts to produce the oxidised dimer (PDB 1RK4) **(B)** Alterations in Trp35 exposure to solvent and intrinsic quenchers. The crystal structures of the CLIC1 reduced monomer (PDB 1K0M) and oxidised dimer (PDB 1RK4) indicate varying environments for Trp35 caused by the significantly different structural conformations of the N-terminal domain for each redox species. Cartoon protein figures were generated using the UCSF Chimera Molecular Graphics system, V1.12.

When comparing the thermal stability of the CLIC1 redox species, the reduced monomer displayed the highest $T_{1/2}$, whereas the oxidised monomer exhibited a slight decrease and the oxidised dimer a significant decrease in $T_{1/2}$ (Figure 15). The stability of the oxidised monomer more closely resembles the stability of the reduced monomer than the oxidised dimer. The tertiary structural analysis, in addition to the thermal unfolding study, thus suggests that while the reduced and oxidised monomer species share a similar conformation, oxidation causes some destabilisation. However, the destabilisation experienced by the oxidised monomer is not to the same extent as observed in the oxidised dimer. The significant destabilisation of the oxidised dimer presents relevance with regards to channel formation. Even though CLIC1 experiences chronic functional expression during oxidative stress, a Cys24-Cys59 disulfide bond presumably hinders membrane insertion. This is due to channels displaying Cys24 on the *trans*-face and Cys59 on the *cis*-face of the membrane (Singh and Ashley, 2006, Littler et al., 2010). Therefore, destabilisation caused by a Cys24-Cys59 disulfide bond may promote the conformation changes needed for channel formation.

The crystal structures available for the reduced monomer (PDB 1K0M) and oxidised dimer (PDB 1RK4) indicate an absence of disulfide bonds in the former and a single Cys24-Cys59 disulfide bond in the latter. However, crystal structures represent static states and not necessarily what occurs within solution. Thiol groups of Cys residues are prone to intra-protein thiol–disulfide exchange reactions known as disulfide bond shuffling. During this phenomenon, a Cys thiolate group attacks one of the disulfide bonds within the same protein. This rearrangement of disulfide bonds does not alter the number of disulfide bonds present within a protein, but merely the location of the bond (Gilbert, 1998).

CLIC1 contains six Cys residues, with Cys24, Cys59 and Cys89 present in the N-terminal domain and Cys178, Cys191 and Cys223 appearing in the C-terminal domain. The potential disulfide bonds formed by each Cys residue within the three redox species, were determined by means of a double alkylation experiment assessed through LC-MS/MS. Free Cys thiol groups first underwent alkylation via NEM, after which the samples were reduced and any previously disulfide bonded residues experienced alkylation via IAA. Statistical modelling determined the fitted probability of IAA labelling for each Cys residue within each redox species (Figure 16B and C). The fitted probability of IAA labelling for Cys89, Cys178, Cys191 and Cys223 generally did not exceed 0.25 in all redox species, thus implying predominant NEM labelling. This indicates that while these Cys residues experience some disulfide bond shuffling, no significant stable disulfide bonds exist regardless of oxidation state. Interestingly, Cys24 and Cys59 displayed a fitted probability of IAA labelling of approximately 0.5 in the reduced and oxidised monomer, and a fitted probability of approximately 0.91 in the oxidised dimer. This suggests that the oxidised dimer forms a stable Cys24-Cys59 disulfide bond, while the monomeric forms undergo significant disulfide bond shuffling with these same residues. This again reiterates that the oxidised monomer more closely resembles the reduced monomer than a subunit from the oxidised dimer.

4.3 The Pep1 and Pep4 Core NLSs Bind ImpαΔIBB with Similar Affinities

All vertebrate CLICs, with the exception of CLIC3, contain a highly conserved monopartite cNLS with the sequence **KKYR** (NLS residues in bold and underlined) which may assist in nuclear import (Figure 4C). However, the only CLIC proteins that appear within the cell nucleus are CLIC1, CLIC3 and CLIC4 (Phang, 2012). The mechanism by which CLIC4 enters the nucleus presumably follows the classical nuclear import pathway. In this pathway the cargo cNLS binds Impβ via the mediator Impα in order to facilitate translocation via the NPCs (Chook and Blobel, 2001; Lange *et al.*, 2007). The interactions formed between a CLIC4 NLS peptide and the major pocket of mImpα1ΔIBB have been characterised through the 3OQS PDB crystal structure (Mynott *et al.*, 2011). CLIC4 nuclear import also appears enhanced when NO induces the *S*-nitrosylation of Cys234. This Cys residue exists between helix 8 and helix 9, which is in close proximity to the helical NLS. *S*-nitrosylation of Cys234 seemingly causes conformational changes which destabilise and unfold CLIC4, subsequently exposing the helical NLS to interact with Impα (Malik *et al.*, 2010). Similar to CLIC4, CLIC1 contains two Cys residues capable of *S*-nitrosylation, namely Cys59 and Cys89 (Figure 4C). The *S*-nitrosylation of these Cys residues may also induce conformational changes that expose the helical NLS to Impα (Phang, 2012). This study aimed to further assess the nuclear import of CLICs by investigating the binding of the CLIC1 and CLIC4 NLSs, with the associated flanking regions, to mImpα1ΔIBB and hImpα1ΔIBB.

Despite the core NLS regions of CLIC4 and CLIC1 sharing the same sequence (**KKYR**), differences exist within the flanking regions (Suh *et al.*, 2004; Littler *et al.*, 2005). The identified NLS regions for CLIC4 and CLIC1 are 198-VKVVA**KKYR**N-207 and 187-VQVVC**KKYR**G-196, respectively. The N-terminal KVVA of CLIC4 displays significant conservation amongst CLIC proteins and contains a basic residue not present in CLIC1 (Figure 4B). Hence, molecular docking was performed to determine whether the amino acid differences in the flanking regions impact the interactions formed with the mImpα1ΔIBB major pocket. Water mediated interactions could not be assessed due to limitations of the docking software. The docking of Pep4 to mImpα1ΔIBB indicated that the peptide forms two bonds with the flanking region not seen in the 3OQS structure of truncated Pep4 (VA**KKYR**N) bound to mImpα1ΔIBB. The two bonds observed entail the backbones of Lys2 and Val4 binding the side-chains of Trp273 and Trp231, respectively. However, these bonds may be of little significance, due to the involvement of backbone and not side-chain functional groups. The docked models of Pep4 and Pep1 bound to the mImpα1ΔIBB major pocket displayed similar Rosetta score values. This indicates that both peptides experience similar strengths of association to mImpα1ΔIBB. Comparing the interactions formed by Pep4 and Pep1 in these models showed that only the backbones of Val4 and Ala/Cys5 form conserved interactions with mImpα1ΔIBB, again signifying non-specific interactions. Assessing the NLS core regions of Pep4 and Pep1 demonstrates that the majority of conserved interactions with mImpα1ΔIBB occur within this region indicating that only the core region is required for signalling nuclear import. Of particular significance is Lys6 which is the only residue that forms three conserved side-chain interactions with mImpα1ΔIBB.

The sequence alignment of the mImpα1ΔIBB and hImpα1ΔIBB major pockets indicates five amino acid variations occurring at positions 195, 203, 206, 218 and 220 (Figure 18A). Since this study aimed to utilise hImpα1ΔIBB for experimental work, molecular docking was also performed to determine whether these amino acid differences may alter the binding of Pep4 and Pep1. The Rosetta scores calculated for Pep4 and Pep1 binding hImpα1ΔIBB displayed values similar to one another and thus both experience similar strengths of association to hImpα1ΔIBB. While the Rosetta scores appear slightly higher when bound to hImpα1ΔIBB than mImpα1ΔIBB, the difference seems negligible. The flanking regions of Pep4 and Pep1 both only formed two bonds with hImpα1ΔIBB, which are the same interactions as when binding to mImpα1ΔIBB. The majority of the interactions between the respective peptides and hImpα1ΔIBB occurred within the core region, with Lys6 forming specific side-chain interactions. Consequently, the CLIC NLS flanking region and differences in the two ImpαΔIBB major pocket sequences do appear significant in the interaction between the CLIC NLSs and ImpαΔIBB. While the interactions formed between the two respective peptides and hImpα1ΔIBB seem fairly similar, the use of fluorescence quenching established the binding affinity of each peptide to hImpα1ΔIBB. This assisted in determining whether the few differences noticed in the molecular docking may impact binding affinity. However, since the Cys residue in Pep1 would cause peptide dimerisation, a suitable mutant would need to be synthesised for experimental work. The Rosetta scores for hImpα1ΔIBB bound to Pep1, Pep1_A and Pep1_S, respectively, displayed a similar range of values, thus signifying similar strengths of association. The Pep1_S mutant most closely resembled the interactions formed between Pep1 and hImpα1ΔIBB and was thus chosen for the fluorescence quenching study.

The indole rings of the seven Trp residues within hImpα1ΔIBB protrude towards the convex side of the protein, which is the side where NLSs can bind. The major pocket houses three of these Trp residues, namely Trp142, Trp184 and Trp231. Molecular docking established that Pep4 and Pep1 interact with these residues upon binding hImpα1ΔIBB (Figure 20A). As such, the binding of either Pep1 or Pep4 is expected to result in fluorescence quenching of hImpα1ΔIBB from which the relative binding affinity can be extrapolated. The addition of increasing concentrations of either Pep1 or Pep4 to hImpα1ΔIBB indicated a systematic decrease in fluorescence intensity with no wavelength shift observed. The decrease in fluorescence intensity signifies the binding of the NLS peptide to a region in hImpα1ΔIBB containing Trp residues. However, the lack of a wavelength shift further dictates no change in Trp solvent exposure and hence no observable tertiary structural change in hImpα1ΔIBB upon peptide binding (Figure 20B and C). Since fluorescence quenching is a low-resolution technique, only apparent K_d values may be attained. However, this technique does allow the determination of comparative values for the binding of the two peptides to hImpα1ΔIBB. The values obtained for Pep1 and Pep4 were 231.3 ± 11.0 μM and 308.8 ± 15.1 μM, respectively (Figure 20D). These values show that both Pep1 and Pep4 bind hImpα1ΔIBB with similar affinities. Based on the fluorescence quenching and molecular docking results, the flanking regions appear to not impact binding interactions and subsequent binding affinity. Hence, the core NLS region of CLIC1 and CLIC4 (**KYYR**) is generally responsible for binding Impα.

4.4 The CLIC NLS Core Displays Weak Binding to hImpα1ΔIBB

While molecular docking elucidated the respective interactions formed by Pep1 and Pep4 to hImpα1ΔIBB, ITC provided further information regarding the thermodynamic parameters of these interactions. Pep1 proved insoluble in the assay buffers, most likely due to containing one less basic residue compared to Pep4. However, given that Pep1 and Pep4 both seem to bind hImpα1ΔIBB with similar binding affinities using an identical core sequence, the thermodynamic parameters of the two peptides likely coincide. The concentrations obtainable for Pep4 and hImpα1ΔIBB also presented limitations. Both experienced solubility constraints whereas hImpα1ΔIBB exhibited a further restriction with regards to the amount that could be expressed and purified. These limitations, in addition to a relatively weak interaction, only allowed a c-value around 10 to be obtained, resulting in non-ideal isotherms which usually lacked clear pre- and post-transition regions. Isotherms displaying c-values ranging from 1 to 1 000 allow the determination of accurate K_d values. However, when dealing with low c-values, prior information regarding the stoichiometry is beneficial, since a fixed n value allows only the K_d and ΔH values to float, thereby preventing problems caused by multiple minima (Wiseman *et al.*, 1989). Although hImpα1ΔIBB contains two binding pockets, monopartite cNLSs generally bind to the major pocket, with only bipartite NLSs binding both the major and minor pockets. The crystal structure of truncated Pep4 bound to hImpα1ΔIBB (PDB 3OQS) further confirms that the CLIC4 NLS only binds the major pocket. As such, all isotherm data was fitted with an A + B ↔ AB hetero-association model using SEDPHAT.

Nuclear import cargos form reversible and transient hetero-oligomeric complexes with Impα and Impβ through non-obligatory contacts in response to particular signalling events or external influences. While the trimeric complex facilitates highly selective and tight interactions in the cytoplasm, the complex switches to a very low affinity state within the nucleus to allow cargo release. This deliberate affinity switch in nuclear import ensures the correct directionality of nuclear translocation (Goldfarb *et al.*, 2004). Since Impα exhibits a physiological concentration of ~1 μM in the cell, cargo NLSs need to form interactions within the nM range to Impα to ensure a strong cytoplasmic complex (Hodel *et al.*, 2001; Riddick and Macara, 2005; Wu *et al.*, 2017). Furthermore, higher eukaryotes exhibit a minimum of seven Impα isoforms to assist the regulation of nuclear import. Despite high sequence and structure conservation, Impα isoforms exhibit very distinct substrate specificity. This permits diverse functionality with the resultant possibility of overlapping roles between isoforms (Pumroy and Cingolani, 2015). This study obtained a K_d value of 106.3 ± 4.0 μM for Pep4 binding hImpα1ΔIBB in a phosphate buffer (Figure 21). This confirms the weak affinity of Pep4 for the general cNLS importer established via fluorescence quenching. Therefore, Pep4 likely exhibits preferential binding to a hImpα isoform with different flexibility properties. A possible contender would be hImpα5 which displays homology to mImpα1 that was used in the 3OQS PDB crystal structure (Mynott *et al.*, 2011; Oka and Yoneda, 2018). Even though Pep4 likely binds optimally to a different isoform partner than hImpα1, the thermodynamic parameters obtained through ITC may still elucidate some information concerning Pep4 binding hImpα in general.

The thermodynamic parameters indicate that the overall reaction occurs spontaneously and is both enthalpically and entropically driven (Figure 21B). The ΔH_{obs} represents the sum of heat energy absorbed and release by many individual bonds that undergo formation and dissolution during a ligand-protein interaction. Bond shuffling occurs due to rearrangements of the ligand, protein receptor binding site and surface solvent molecules. During binding, solvent-protein hydrogen bonds and van der Waals interactions break, while new ligand–protein hydrogen bonds, salt bridges, van der Waals contacts and water-mediated interactions form (Freire *et al.*, 1990; Doyle, 1997; Perozzo *et al.*, 2004). Another factor to consider, which also contributes to the ΔH_{obs} term, is proton linked effects (Baker and Murphy, 1996). The interaction between Pep4 and hImpα1ΔIBB yielded a ΔH_{obs} of -6.7 ± 0.7 kJ·mol^{-1}, indicating an overall exothermic reaction. Assessing the docked structure of Pep4 bound to hImpα1ΔIBB, the Pep4 flanking region forms two hydrogen bonds with hImpα1ΔIBB, while the core region forms six hydrogen bonds and a single salt bridge (Figure 19). One of the flanking region and four of the core region hydrogen bonds appear conserved with the crystal structure of truncated Pep4 bound to mImpα1ΔIBB (PDB 3OQS). This signifies the assumed importance of the three basic CLIC NLSs core residues, in addition to the residue just prior to the core region, in binding Impα. While the docking tool used could not assess water-mediated interactions, the 3OQS crystal structure indicates a significant number of water-mediated interactions between truncated Pep4 and mImpα1ΔIBB. The interaction between Pep4 and hImpα1ΔIBB likely also shares some of these water-mediated interactions that would contribute to the enthalpic component. Protonation effects also contribute towards the ΔH_{obs}. This study showed that Pep4 binding hImpα1ΔIBB does involve the exchange of one proton (N_{H^+} = -0.908) which results in a ΔH_0 equal to -4.904 kJ·mol^{-1} (Figure 20C). Therefore, the reaction remains enthalpically driven despite the added proton linked effect.

The ΔG_{obs} and ΔH_{obs} are used to directly calculate the ΔS_{obs} which describes all other positive and negative free energy driving forces not attributed to the ΔH_{obs}. These driving forces entail the molecular disorder exhibited by the ligand, protein receptor and solvent molecules in the system. Hydration effects function as the main contributing factor towards the ΔS_{obs} in a ligand-protein interaction. Protein receptor binding sites generally expose water-accessible surface areas which experience desolvation when the ligand enters the binding site. This results in a favourable entropic contribution which further assists in shielding hydrophobic residues from the displayed water molecules. However, ligand binding reduces the degrees of freedom experienced by the solvent molecules, as well as the ligand and protein receptor side-chain torsion angles, thus causing an unfavourable entropic contribution (Murphy *et al.*, 1994; Ladbury, 1996; Perozzo *et al.*, 2004). The ΔS_{obs} calculated for Pep4 binding hImpα1ΔIBB was -15.6 ± 0.8 kJ·mol^{-1} indicating an overall larger entropic than enthalpic contribution. Considering the significant flexibility Pep4 would experience in solution, binding to hImpα1ΔIBB would result in a large unfavourable entropic contribution due to the drastic decrease in ligand flexibility. However, this energetic price is likely compensated for by hydration effects, the shielding of hydrophobic residues and stable hydrogen bonds formed between Pep4 and hImpα1ΔIBB (Figure 19).

4.5 Conclusion

Oxidative conditions have proven to be a significant regulatory factor impacting on the PPIs formed by CLIC1 and CLIC4, respectively. By mediating particular PPIs, redox regulation impacts on the overall distribution and function of these proteins. This has demonstrated significance in promoting CLIC1 membrane insertion through oxidation induced dimerisation, as well as the nuclear import of both CLIC1 and CLIC4 via mediator Impα.

Under redox conditions resembling oxidative stress, the majority of CLIC1 undergo dimerisation while the residue remains monomeric. Oxidation causes Cys24 and Cys59 to form a disulfide bond which results in significant N-terminal changes causing the exposure of hydrophobic patches. In the event that the hydrophobic patches of two CLIC1 monomers interact, a dimer forms. Dimerisation prevents disulfide bond shuffling of Cys24 and Cys59, thus maintaining both this single disulfide bond and the resultant oligomeric state. Conversely, in the absence of dimerisation, Cys24 and Cys59 undergo significant disulfide bond shuffling and exhibit a tertiary structure resembling the reduced monomer as opposed to a single dimeric subunit. The oxidised monomer therefore does not maintain the N-terminal changes induced by a stable Cys24-Cys59 disulfide bond. Furthermore, the dimeric form exhibits a decrease in overall thermal stability compared to the monomeric form. This is significant in understanding the role of oxidative stress in the chronic functional expression of CLIC1. Since membrane insertion propels Cys24 to the *trans*-face while maintaining Cys59 on the *cis*-face, dimerisation would presumable hinder membrane insertion (Singh and Ashley, 2006; Littler *et al.*, 2010). This study indicates that a destabilised dimer may aid membrane insertion by readily permitting reconfiguration of the CLIC1 dimer. However, the exact mode of CLIC1 membrane insertion still remains undetermined.

In addition to potentially regulating CLIC1 membrane insertion, redox regulation may also impact the cellular distribution of this protein. Both CLIC1 and CLIC4 appear within the cell nucleus and exhibit a highly conserved NLS. *S*-nitrosylation of Cys234 in CLIC4 induces conformational changes that allow access of the NLS to Impα. The mediator karyopherin Impα may then interact with Impβ to translocate CLIC4 into the nucleus (Suh *et al.*, 2004; Mynott *et al.*, 2011). In order to determine whether CLIC1 follows a similar redox regulated nuclear import pathway, the binding of the CLIC1 NLS to hImpα1ΔIBB was compared to that of CLIC4. While the core sequences remain the same (**KKYR**), there are a few different flanking region residues that may impact on hImpα1ΔIBB binding. Molecular docking and binding studies indicated that mainly the core NLS facilitates weak binding to hImpα1ΔIBB. CLIC1 thus also interacts with Impα through the same sequence as CLIC4 to facilitate the classical nuclear import pathway. However, further investigation is needed to determine which hImpα isoform the CLIC1 and CLIC4 NLSs bind to optimally. After establishing the correct hImpα isoform, analysis may be performed to assess the binding of *S*-nitrosylated full-length CLIC1 and CLIC4 to hImpαΔIBB.

REFERENCES

ABRÀMOFF, M. D., MAGALHÃES, P. J. & RAM, S. J. 2004. Image Processing with ImageJ. *Biophotonics Int.,* 11**,** 36-42.

ALMÉN, M. S., NORDSTRÖM, K. J. V., FREDRIKSSON, R. & SCHIÖTH, H. B. 2009. Mapping the Human Membrane Proteome: A Majority of the Human Membrane Proteins Can Be Classified According to Function and Evolutionary Origin. *BMC Biol.,* 7**,** 1-14.

AVERAIMO, S., MILTON, R. H., DUCHEN, M. R. & MAZZANTI, M. 2010. Chloride Intracellular Channel 1 (CLIC1): Sensor and Effector during Oxidative Stress. *FEBS Lett.,* 584**,** 2076-2084.

BAKER, B. M. & MURPHY, K. P. 1996. Evaluation of Linked Protonation Effects in Protein Binding Reactions Using Isothermal Titration Calorimetry. *Biophys. J.,* 71**,** 2049-2055.

BAXTER, A., MITTLER, R. & SUZUKI, N. 2014. ROS As Key Players in Plant Stress Signalling. *J. Exp. Bot.,* 65**,** 1229-1240.

BERRIDGE, M. J. 2014. Module 6: Spatial and Temporal Aspects of Signalling. *Cell Signalling Biology.* Portland Press.

BRAUTIGAM, C. A. 2015. Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. *Meth. Enzymol.,* 562**,** 109-133.

CAGATAY, T. & CHOOK, Y. M. 2018. Karyopherins in Cancer. *Curr. Opin. Cell Biol.,* 52**,** 30-42.

CHELSKY, D., RALPH, R. & JONAK, G. 1989. Sequence Requirements for Synthetic Peptide-Mediated Translocation to the Nucleus. *Mol. Cell. Biol.,* 9**,** 2487-2492.

CHEN, Y. & BARKLEY, M. D. 1998. Toward Understanding Tryptophan Fluorescence in Proteins. *Biochemistry,* 37**,** 9976-9982.

CHOOK, Y. M. & BLOBEL, G. 2001. Karyopherins and Nuclear Import. *Curr. Opin. Struct. Biol.,* 11**,** 703-715.

CINGOLANI, G., BEDNENKO, J., GILLESPIE, M. T. & GERACE, L. 2002. Molecular Basis for the Recognition of a Nonclassical Nuclear Localization Signal by Importin β. *Mol. Cell,* 10**,** 1345-1353.

CORREANI, V., DI FRANCESCO, L., CERA, I., MIGNOGNA, G., GIORGI, A., MAZZANTI, M., FUMAGALLI, L., FABRIZI, C., MARAS, B. & SCHININÀ, M. E. 2015. Reversible Redox Modifications in Microglia Proteome Challenged by Beta Amyloid. *Mol. Biosyst.,* 11**,** 1584-1593.

CRIBARI-NETO, F. & ZEILEIS, A. 2010. Beta Regression in R. *J. Stat. Softw.,* 34**,** 1-24.

CUMMING, R. C., ANDON, N. L., HAYNES, P. A., PARK, M., FISCHER, W. H. & SCHUBERT, D. 2004. Protein Disulfide Bond Formation in the Cytoplasm during Oxidative Stress. *J. Biol. Chem.,* 279**,** 21749-21758.

DE LAS RIVAS, J. & FONTANILLO, C. 2010. Protein-Protein Interactions Essentials: Key Concepts to Building and Analyzing Interactome Networks. *PLOS Comput. Biol.,* 6**,** e1000807.

DOYLE, M. L. 1997. Characterization of Binding Interactions by Isothermal Titration Calorimetry. *Curr. Opin. Biotechnol.,* 8**,** 31-35.

DUFF, M. R., GRUBBS, J. & HOWELL, E. E. 2011. Isothermal Titration Calorimetry for Measuring Macromolecule-Ligand Affinity. *J. Vis. Exp.***,** e2796.

DUTZLER, R., CAMPBEL, E. B., CADENE, M., CHAIT, B. T. & MACKINNON, R. 2002. X-Ray Structure of a ClC Chloride Channel at 3.0 Å Reveals the Molecular Basis of Anion Selectivity. *Nature,* 415**,** 287-294.

DUY, C. & FITTER, J. 2005. Thermostability of Irreversible Unfolding α-Amylases Analyzed by Unfolding Kinetics. *J. Biol. Chem.,* 280**,** 37360-37365.

EDWARDS, J. C. & KAHL, C. R. 2010. Chloride Channels of Intracellular Membranes. *FEBS Lett.,* 584**,** 2102-2111.

FANUCCHI, S., ADAMSON, R. J. & DIRR, H. W. 2008. Formation of an Unfolding Intermediate State of Soluble Chloride Intracellular Channel Protein CLIC1 at Acidic pH. *Biochemistry,* 47**,** 11674-11681.

FREIRE, E., MAYORGA, O. L. & STRAUME, M. 1990. Isothermal Titration. *Anal. Chem.,* 62**,** 950A-959A.

GASTEIGER, E., HOOGLAND, C., GATTIKER, A., DUVAUD, S., WILKINS, M. R., APPEL, R. D. & BAIROCH, A. 2005. Protein Identification and Analysis Tools on the ExPASy Server. *In:* WALKER, J. M. (ed.) *The Proteomics Protocols Handbook.* Humana Press.

GILBERT, H. F. 1998. Thiol–Disulfide Exchange of Divalent Sulfur. *In:* SINNOTT, M. (ed.) *Comprehensive Biological Catalysis: A Mechanistic Reference.* Academic Press.

GO, Y.-M. & JONES, D. P. 2008. Redox Compartmentalization in Eukaryotic Cells. *Biochim. Biophys. Acta.,* 1780**,** 1273-1290.

GOLDBERG, R. N., KISHORE, N. & LENNEN, R. M. 2002. Thermodynamic Quantities for the Ionization Reactions of Buffers. *J. Phys. Chem. Ref. Data,* 31**,** 231-370.

GOLDFARB, D. S., CORBETT, A. H., MASON, D. A., HARREMAN, M. T. & ADAM, S. A. 2004. Importin α: A Multipurpose Nuclear-Transport Receptor. *Trends Cell Biol.,* 14**,** 505-514.

GOODCHILD, S. C., ANGSTMANN, C. N., BREIT, S. N., CURMI, P. M. G. & BROWN, L. J. 2011. Transmembrane Extension and Oligomerization of the CLIC1 Chloride Intracellular Channel Protein upon Membrane Interaction. *Biochemistry,* 50**,** 10887-10897.

GREENFIELD, N. J. 2006. Using Circular Dichroism Spectra to Estimate Protein Secondary Structure. *Nat. Protoc.,* 1**,** 2876-2890.

HILLE, B. 1978. Ionic Channels in Excitable Membranes. Current Problems and Biophysical Approaches. *Biophys. J.,* 22**,** 283-294.

HODEL, M. R., CORBETT, A. H. & HODEL, A. E. 2001. Dissection of a Nuclear Localization Signal. *J. Biol. Chem.,* 276**,** 1317-1325.

HU, Y.-F. & RONG, L. 2002. JunB Potentiates Function of BRCA1 Activation Domain 1 (AD1) through a Coiled-Coil-Mediated Interaction. *Genes Dev.,* 16**,** 1509-1517.

HUANG, B. K. & SIKES, H. D. 2014. Quantifying Intracellular Hydrogen Peroxide Perturbations in Terms of Concentration. *Redox Biol.,* 2**,** 955-962.

IVANOV, A. A., KHURI, F. R. & FU, H. 2013. Targeting Protein-Protein Interactions as an Anticancer Strategy. *Trends Pharmacol. Sci.,* 34**,** 393-400.

JANIN, J. 1999. Wet and Dry Interfaces: The Role of Solvent in Protein–Protein and Protein–DNA Recognition. *Structure,* 7**,** R277-R279.

JELLINGER, K. A. 2012. Interaction Between Pathogenic Proteins in Neurodegenerative Disorders. *J. Cell. Mol. Med.,* 16**,** 1166-1183.

JONES, S. & THORNTON, J. M. 1996. Principles of Protein-Protein Interactions. *Proc. Natl. Acad. Sci. USA,* 93**,** 13-20.

JONES, S. & THORNTON, J. M. 1997. Analysis of Protein-Protein Interaction Sites Using Surface Patches. *J. Mol. Biol.,* 272**,** 121-132.

KALDERON, D., ROBERTS, B. L., RICHARDSON, W. D. & SMITH, A. E. 1984. A Short Amino Acid Sequence Able to Specify Nuclear Location. *Cell,* 39**,** 499-509.

KELLER, S., VARGAS, C., ZHAO, H., PISZCZEK, G., BRAUTIGAM, C. A. & SCHUCK, P. 2012. High-Precision Isothermal Titration Calorimetry with Automated Peak Shape Analysis. *Anal. Chem.,* 84**,** 5066-5073.

KLAPOETKE, S. & XIE, M. H. 2017. Disulfide Bond Characterization of Human Factor Xa by Mass Spectrometry through Protein-Level Partial Reduction. *J. Pharm. Biomed. Anal.,* 132**,** 238-246.

KORENBAUM, E. & RIVERO, F. 2002. Calponin Homology Domains at a Glance. *J. Cell Sci.,* 115**,** 3543-3545.

LADBURY, J. E. 1996. Just Add Water! The Effect of Water on the Specificity of Protein-Ligand Binding Sites and Its Potential Application to Drug Design. *Chem. Biol.,* 3**,** 973-980.

LAEMMLI, U. K. 1970. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature,* 227**,** 680-685.

LAKOWICZ, J. R. 2006. *Principles of Fluorescence Spectroscopy*, Springer.

LANGE, A., MILLS, R. E., LANGE, C. J., STEWART, M., DEVINE, S. E. & CORBETT, A. H. 2007. Classical Nuclear Localization Signals: Definition, Function, and Interaction with Importin α. *J. Biol. Chem.,* 282**,** 5101-5105.

LAU, S. Y. M., TANEJA, A. K. & HODGES, R. S. 1984. Synthesis of a Model Protein of Defined Secondary and Quaternary Structure. *J. Biol. Chem.,* 259**,** 13253-13261.

LITTLER, D. R., ASSAAD, N. N., HARROP, S. J., BROWN, L. J., PANKHURST, G. J., LUCIANI, P., AGUILAR, M.-I., MAZZANTI, M., BERRYMAN, M. A., BREIT, S. N. & CURMI, P. M. 2005. Crystal Structure of the Soluble Form of the Redox-Regulated Chloride Ion Channel Protein CLIC4. *FEBS J.,* 272**,** 4996-5007.

LITTLER, D. R., HARROP, S. J., FAIRLIE, W. D., BROWN, L. J., PANKHURST, G. J., PANKHURST, S., DEMAERE, M. Z., CAMPBELL, T. J., BAUSKIN, A. R., TONINI, R., MAZZANTI, M., BREIT, S. N. & CURMI, P. M. G. 2004. The Intracellular Chloride Ion Channel Protein CLIC1 Undergoes a Redox-Controlled Structural Transition. *J. Biol. Chem.,* 279**,** 9298-9305.

LITTLER, D. R., HARROP, S. J., GOODCHILD, S. C., PHANG, J. M., MYNOTT, A. V., JIANG, L., VALENZUELA, S. M., MAZZANTI, M., BROWN, L. J., BREIT, S. N. & CURMI, P. M. G. 2010. The Enigma of the CLIC Proteins: Ion Channels, Redox Proteins, Enzymes, Scaffolding Proteins? *FEBS Lett.,* 584**,** 2093-2101.

LONDON, N., RAVEH, B., COHEN, E., FATHI, G. & SCHUELER-FURMAN, O. 2011. Rosetta FlexPepDock Web Server – High Resolution Modeling of Peptide-Protein Interactions. *Nucleic Acids Res.,* 39**,** W249-W253.

MACLEAN, B., TOMAZELA, D. M., SHULMAN, N., CHAMBERS, M., FINNEY, G. L., FREWEN, B., KERN, R., TABB, D. L., LIEBLER, D. C. & MACCOSS, M. J. 2010. Skyline: An Open Source Document Editor for Creating and Analyzing Targeted Proteomics Experiments. *Bioinformatics,* 26**,** 966-968.

MACPHERSON, R. E. K., RAMOS, S. V., VANDENBOOM, R., ROY, B. D. & PETERS, S. J. 2013. Skeletal Muscle PLIN Proteins, ATGL and CGI-58, Interactions at Rest and Following Stimulated Contraction. *Am. J. Physiol. Regul. Integr. Comp. Physiol.,* 304**,** R644-R650.

MADEIRA, F., PARK, Y. M., LEE, J., BUSO, N., GUR, T., MADHUSOODANAN, N., BASUTKAR, P., TIVEY, A. R. N., POTTER, S. C., FINN, R. D. & LOPEZ, R. 2019. The EMBL-EBI Search and Sequence Analysis Tools APIs in 2019. *Nucleic Acids Res.,* 47**,** W636-W641.

MAKKERH, J. P. S., DINGWALL, C. & LASKEY, R. A. 1996. Comparative Mutagenesis of Nuclear Localization Signals Reveals the Importance of Neutral and Acidic Amino Acids. *Curr. Biol.,* 6**,** 1025-1027.

MALIK, M., SHUKLA, A., AMIN, P., NIEDELMAN, W., LEE, J., JIVIDEN, K., PHANG, J. M., DING, J., SUH, K. S., CURMI, P. M. G. & YUSPA, S. H. 2010. S-Nitrosylation Regulates Nuclear Translocation of Chloride Intracellular Channel Protein, CLIC4. *J. Biol. Chem.,* 285**,** 23818-23828.

MOSAMMAPARAST, N. & PEMBERTON, L. F. 2004. Karyopherins: From Nuclear-Transport Mediators to Nuclear-Function Regulators. *Trends Cell Biol.,* 14**,** 547-556.

MURPHY, K. P., XIE, D., THOMPSON, K. S., AMZEL, L. M. & FREIRE, E. 1994. Entropy in Biological Binding Processes: Estimation of Translational Entropy Loss. *Proteins,* 18**,** 63-67.

MYNOTT, A. V., HARROP, S. J., BROWN, L. J., BREIT, S. N., KOBE, B. & CURMI, P. M. G. 2011. Crystal Structure of Importin-α Bound to a Peptide Bearing the Nuclear Localisation Signal from Chloride Intracellular Channel Protein 4. *FEBS J.,* 278**,** 1662-1675.

NGUYEN, H., LAUTO, A., SHANU, A. & MYERS, S. J. 2016. Chloride Intracellular Channel Protein 1 and Its Role in Neurodegenerative Disorders and Cancerous Tumors. *Biochem. Anal. Biochem.,* 5**,** 1-6.

OKA, M. & YONEDA, Y. 2018. Importin α: Functions as a Nuclear Transport Factor and Beyond. *Proc. Jpn. Acad. Ser. B,* 94**,** 259-274.

PACE, C. N. 1986. Determination and Analysis of Urea and Guanidine Hydrochloride Denaturation Curves. *Methods Enzymol.,* 131**,** 266-280.

PEDERSEN, P. L. & CARAFOLI, E. 1987. Ion Motive ATPases. I. Ubiquity, Properties, and Significance to Cell Function. *Trends Biochem. Sci.,* 12**,** 146-150.

PEREIRA-LEAL, J. B., LEVY, E. D. & TEICHMANN, S. A. 2006. The Origins and Evolution of Functional Modules: Lessons from Protein Complexes. *Phil. Trans. R. Soc. B,* 361**,** 507-517.

PERETTI, M., ANGELINI, M., SAVALLI, N., FLORIO, T., YUSPA, S. H. & MAZZANTI, M. 2015. Chloride Channels in Cancer: Focus on Chloride Intracellular Channel 1 and 4 (CLIC1 and CLIC4) Proteins in Tumor Development and as Novel Therapeutic Targets. *Biochim. Biophys. Acta,* 1848**,** 2523-2531.

PEROZZO, R., FOLKERS, G. & SCAPOZZA, L. 2004. Thermodynamics of Protein–Ligand Interactions: History, Presence, and Future Aspects. *J. Recept. Signal Transduct. Res.,* 24**,** 1-52.

PHANG, J. M. 2012. *Crystallographic Studies of CLIC Proteins and Ezrin.* PhD, The University of New South Wales.

PITT, J. J. 2009. Principles and Applications of Liquid Chromatography- Mass Spectrometry in Clinical Biochemistry. *Clin. Biochem. Rev.,* 30**,** 19-34.

PUMROY, R. A. & CINGOLANI, G. 2015. Diversification of Importin-α Isoforms in Cellular Trafficking and Disease States. *Biochem J.,* 466**,** 13-28.

RAVEH, B., LONDON, N. & SCHUELER-FURMAN, O. 2010. Sub-Angstrom Modeling of Complexes Between Flexible Peptides and Globular Proteins. *Proteins,* 78**,** 2029-2040.

RAYMENT, I., HOLDEN, H. M., WHITTAKER, M., YOHN, C. B., LORENZ, M., HOLMES, K. C. & MILLIGAN, R. A. 1993. Structure of the Actin-Myosin Complex and Its Implications for Muscle Contraction. *Science,* 261**,** 58-65.

RIDDICK, G. & MACARA, I. G. 2005. A Systems Analysis of Importin-α-β Mediated Nuclear Protein Import. *J. Cell Biol.,* 168**,** 1027-1038.

ROMERO-PUERTAS, M. C. & SANDALIO, L. M. 2016. Nitric Oxide Level is Self-Regulating and Also Regulates Its ROS Partners. *Front. Plant Sci.,* 7**,** 1-5.

SANCHEZ, C., LACHAIZE, C., JANODY, F., BELLON, B., RÖDER, L., EUZENAT, J., RECHENMANN, F. & JACQ, B. 1999. Grasping at Molecular Interactions and

Genetic Networks in Drosophila Melanogaster Using FlyNets, an Internet Database. *Nucleic Acids Res.,* 27**,** 89-94.
SCHEUERMANN, T. H. & BRAUTIGAM, C. A. 2015. High-Precision, Automated Integration of Multiple Isothermal Titration Calorimetric Thermograms: New Features of NITPIC. *Methods,* 76**,** 87-98.
SEYMOUR, S. L. & HUNTER, C. J. 2015. ProteinPilot™ Software Overview. High Quality, In-Depth Protein Identification and Protein Expression Analysis. *USA: AB Sciex Pte. Ltd.***,** 1-5.
SINGER, S. J. & NICOLSON, G. L. 1972. The Fluid Mosaic Model of the Structure of Cell Membranes. *Science,* 175**,** 720-731.
SINGH, H. 2010. Two Decades with Dimorphic Chloride Intracellular Channels (CLICs). *FEBS Lett.,* 584**,** 2112-2121.
SINGH, H. & ASHLEY, R. H. 2006. Redox Regulation of CLIC1 by Cysteine Residues Associated with the Putative Channel Pore. *Biophys. J.,* 90**,** 1628-1638.
STELMA, T., CHI, A., VAN DER WATT, P. J., VERRICO, A., LAVIA, P. & LEANER, V. D. 2016. Targeting Nuclear Transporters in Cancer: Diagnostic, Prognostic and Therapeutic Potential. *IUBMB Life,* 68**,** 268-280.
STOYCHEV, S. H., NATHANIEL, C., FANUCCHI, S., BROCK, M., LI, S., ASMUS, K., WOODS, V. L. & DIRR, H. W. 2009. Structural Dynamics of Soluble Chloride Intracellular Channel Protein CLIC1 Examined by Amide Hydrogen-Deuterium Exchange Mass Spectrometry. *Biochemistry,* 48**,** 8413-8421.
SUH, K. S., MUTOH, M., GERDES, M. & YUSPA, S. H. 2005. CLIC4, an Intracellular Chloride Channel Protein, is a Novel Molecular Target for Cancer Therapy. *J. Investig. Dermatol. Symp. Proc.,* 10**,** 105-109.
SUH, K. S., MUTOH, M., NAGASHIMA, K., FERNANDEZ-SALAS, E., EDWARDS, L. E., HAYES, D. D., CRUTCHLEY, J. M., MARIN, K. G., DUMONT, R. A., LEVY, J. M., CHENG, C., GARFIELD, S. & YUSPA, S. H. 2004. The Organellular Chloride Channel Protein CLIC4/mtCLIC Translocates to the Nucleus in Response to Cellular Stress and Accelerates Apoptosis. *J. Biol. Chem.,* 279**,** 4632-4641.
SUH, K. S. & YUSPA, S. H. 2005. Intracellular Chloride Channels: Critical Mediators of Cell Viability and Potential Targets for Cancer Therapy. *Curr. Pharm. Des.,* 11**,** 2753-2764.
TULK, B. M., SCHLESINGER, P. H., KAPADIA, S. A. & EDWARDS, J. C. 2000. CLIC-1 Functions as a Chloride Channel when Expressed and Purified from Bacteria. *J. Biol. Chem.,* 275**,** 26986-26993.
VALENZUELA, S. M., ALKHAMICI, H., BROWN, L. J., ALMOND, O. C., GOODCHILD, S. C., CARNE, S., CURMI, P. M. G., HOLT, S. A. & CORNELL, B. A. 2013. Regulation of the Membrane Insertion and Conductance Activity of the Metamorphic Chloride Intracellular Channel Protein CLIC1 by Cholesterol. *PLOS ONE,* 8**,** e56948.
VIVIAN, J. T. & CALLIS, P. R. 2001. Mechanisms of Tryptophan Fluorescence Shifts in Proteins. *Biophys. J.,* 80**,** 2093–2109.
WATERHOUSE, A. M., PROCTER, J. B., MARTIN, D. M. A., CLAMP, M. & BARTON, G. J. 2009. Jalview Version 2 – A Multiple Sequence Alignment Editor and Analysis Workbench. *Bioinformatics,* 25**,** 1189-1191.
WESTERMARCK, J., IVASKA, J. & CORTHALS, G. L. 2013. Identification of Protein Interactions Involved in Cellular Signaling. *Mol. Cell. Proteomics,* 12**,** 1752-1763.
WISEMAN, T., WILLISTON, S., BRANDTS, J. F. & LIN, L.-N. 1989. Rapid Measurement of Binding Constants and Heats of Binding Using a New Titration Calorimeter. *Anal. Biochem.,* 179**,** 131-137.

WOODY, R. W. 1995. Circular Dichroism. *Methods Enzymol.,* 246**,** 34–71.
WOZNIAK, R. W., ROUT, M. P. & AITCHISON, J. D. 1998. Karyopherins and Kissing Cousins. *Trends Cell Biol.,* 8**,** 184-188.
WU, W., SANKHALA, R. S., FLORIO, T. J., ZHOU, L., NGUYEN, N. L. T., LOKAREDDY, R. K., CINGOLANI, G. & PANTE, N. 2017. Synergy of Two Low-Affinity NLSs Determines the High Avidity of Influenza A Virus Nucleoprotein NP for Human Importin α Isoforms. *Sci. Rep.,* 7**,** 1-16.
YAN, C., WU, F., JERNIGAN, R. L., DOBBS, D. & HONAVAR, V. 2008. Characterization of Protein-Protein Interfaces. *Protein J.,* 27**,** 59-70.
YEAGLE, P. L. 1989. Lipid Regulation of Cell Membrane Structure and Function. *FASEB J.,* 3**,** 1833-1842.
YEN, T.-Y., JOSHI, R. K., YAN, H., SETO, N. O. L., PALCIC, M. M. & MACHER, B. A. 2000. Characterization of Cysteine Residues and Disulfide Bonds in Proteins by Liquid Chromatography/Electrospray Ionization Tandem Mass Spectrometry. *J. Mass Spectrom.,* 35**,** 990–1002.
ZHAO, H., PISZCZEK, G. & SCHUCK, P. 2015. SEDPHAT – A Platform for Global ITC Analysis and Global Multi-Method Analysis of Molecular Interactions. *Methods,* 76**,** 137-148.

APPENDIX

List of Appendices

Appendix A: Probability of IAA labelling

(Dr S.J.M. Steeb and O.M.M. Færch)

1. Aim

The aim was to model the expected proportion of IAA labelling as a function of two factors, the Cys residue involved and the oxidation state of the protein. The expected proportion of IAA labelling is also the probability of IAA labelling for the CLIC1 protein.

2. Data

The dependent random variable or response, IAA labelling proportion, is a continuous proportion in the interval (0, 1), with no 0 or 1 observations. The independent explanatory variable or predictor, Cys residue, is a factor with six levels, namely Cys24, Cys59, Cys89, Cys178, Cys191 and Cys223. The predictor, protein oxidation state, is a factor with three levels, namely Red (reduced CLIC1), OxD (oxidised CLIC1 dimer) and OxM (oxidised CLIC1 monomer). A sample of 64 IAA labelling proportions were obtained from the 18 covariate patterns (combinations of factor levels) with between 2 and 6 replicates per covariate pattern. The IAA labelling proportions are assumed to be independent. All redox samples were prepared from the same initial protein stock. However, each sample replicate is assumed to be independent as using the same protein production methods would result in the same intact folded protein.

3. Terms and Notation

IAA: IAA labelling proportion

Cys: factor Cys residue

Ox: factor protein oxidation state

p = E(IAA): expected proportion of IAA labelling for a given combination of Cys and Ox levels, which is the corresponding probability of IAA labelling for the CLIC1 protein

Fitted value: refers to an estimate obtained from the fitted model

4. Exploratory Plots

Exploratory plots (Figures 1 and 2) were generated in order to assist in establishing which model should be used to fit to the data.

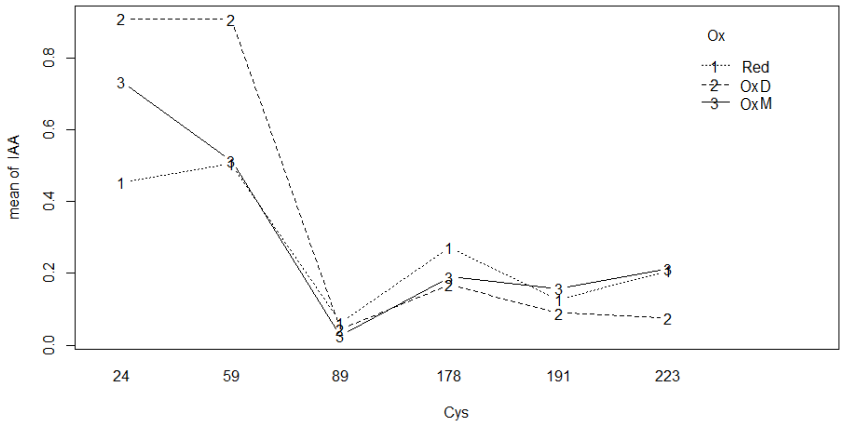


Figure 1: Mean of IAA for each Cys under each Ox
The means for the proportion of IAA labelling was determined for each Cys residue under each oxidation state.

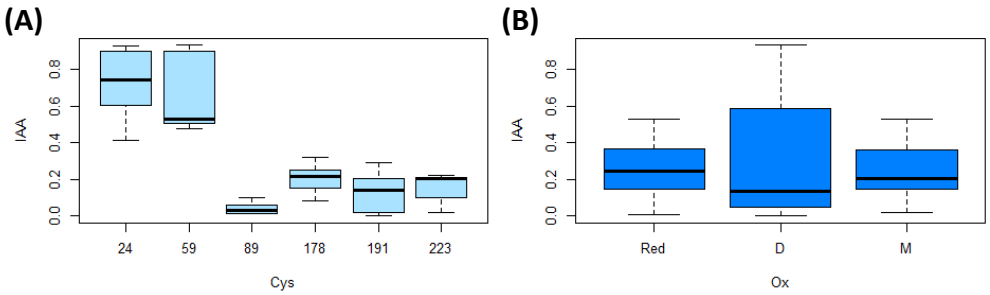


Figure 2: Box plots of IAA vs Cys and Ox
(A) IAA labelling proportion means and distribution for each Cys residue. **(B)** IAA labelling proportion means and distribution for each oxidation state.

Figure 1 suggests an effect on IAA due to Cys, with the observed mean IAA for Cys24 and Cys59 larger than the values for the remaining four Cys levels. Oxidation appears to affect the mean IAA only for Cys24 and Cys59. The lines that cross indicate that interaction between the factors may be present. Therefore, this plot suggests a model with interaction terms. Figure 2 indicates that IAA is heteroscedastic, i.e. the variance of IAA depends on the predictor levels.

5. Regression of IAA on Cys and Ox

The response IAA is a continuous proportion, with values between zero and one, and the observed response variance is not constant across the levels of the predictors (Figure 2). Transformations to overcome these problems so that ordinary least squares regression can be applied, are often used, but can lead to biased estimates and difficulties in interpretation. An alternative method, beta regression, was applied.

The assumptions made are:

- The IAA proportions were generated independently from a beta distribution but are not identically distributed
- E(IAA) = p is related to the predictors Cys and Ox through a linear predictor which is a linear function of unknown parameters

A logit link function is used, therefore:

$$logit(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 Cys59 + \cdots + \beta_{17} Cys223{:}\, OxM \qquad (1)$$

The model also includes a precision parameter which depends on one or more predictors through a link function with Cys24 and reduced CLIC1 as the reference levels. This approach naturally incorporates features such as heteroscedasticity or skewness which are commonly observed in proportions.

The `betareg` package provides the class of beta regressions in the `R` system for statistical computing and was used to fit a model. Based on the exploratory plots of the data (Figures 1 and 2), several models were compared using AIC, BIC and hypothesis tests to test the significance of individual parameters included in a model. The following model with interaction terms was chosen:

```
betareg(formula = IAA ~ Cys * Ox | Cys, link = "logit")


Standardised weighted residuals 2:

    Min      1Q  Median      3Q     Max

-2.6799 -0.8442  0.0096  0.8977  2.1700
```

```
Coefficients (mean model with logit link):
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.18812    0.08023  -2.345 0.019046 *
Cys59        0.20974    0.11468   1.829 0.067414 .
Cys89       -2.51866    0.37094  -6.790 1.12e-11 ***
Cys178      -0.78083    0.15610  -5.002 5.67e-07 ***
Cys191      -1.99567    0.46585  -4.284 1.84e-05 ***
Cys223      -1.12001    0.21837  -5.129 2.91e-07 ***
OxD          2.51080    0.13934  18.019  < 2e-16 ***
OxM          1.21098    0.10908  11.101  < 2e-16 ***
Cys59:OxD   -0.20521    0.19956  -1.028 0.303790
Cys89:OxD   -3.08423    0.52652  -5.858 4.69e-09 ***
Cys178:OxD  -3.15658    0.23295 -13.550  < 2e-16 ***
Cys191:OxD  -2.98653    0.59426  -5.026 5.02e-07 ***
Cys223:OxD  -3.86908    0.36320 -10.653  < 2e-16 ***
Cys59:OxM   -1.17294    0.15196  -7.718 1.18e-14 ***
Cys89:OxM   -1.92638    0.53321  -3.613 0.000303 ***
Cys178:OxM  -1.68317    0.21219  -7.932 2.15e-15 ***
Cys191:OxM  -0.49057    0.54735  -0.896 0.370107
Cys223:OxM  -1.18745    0.28293  -4.197 2.70e-05 ***

Phi coefficients (precision model with log link):
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  5.74437    0.49969  11.496  < 2e-16 ***
Cys59       -0.05096    0.70661  -0.072   0.9425
Cys89       -1.66978    0.71935  -2.321   0.0203 *
Cys178      -1.51723    0.61128  -2.482   0.0131 *
Cys191      -3.57554    0.62523  -5.719 1.07e-08 ***
Cys223      -1.48740    0.70718  -2.103   0.0354 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Type of estimator: ML (maximum likelihood)

Log-likelihood:   119 on 24 Df

Pseudo R-squared: 0.7681

Number of iterations: 38 (BFGS) + 2 (Fisher scoring)
```

6. Residual Analysis

The raw response residuals (observed proportion – fitted proportion) are not normally distributed. Therefore, `betareg` provides several types of transformed residuals that can be investigated to determine the adequacy of the fitted model. The default and recommended residuals provided are standardised and weighted in such a way that if the model is a good fit, these residuals are independently normal (0, 1) distributed.

6.1 Normality

The null hypothesis states that the sweight2 residuals have a normal distribution.

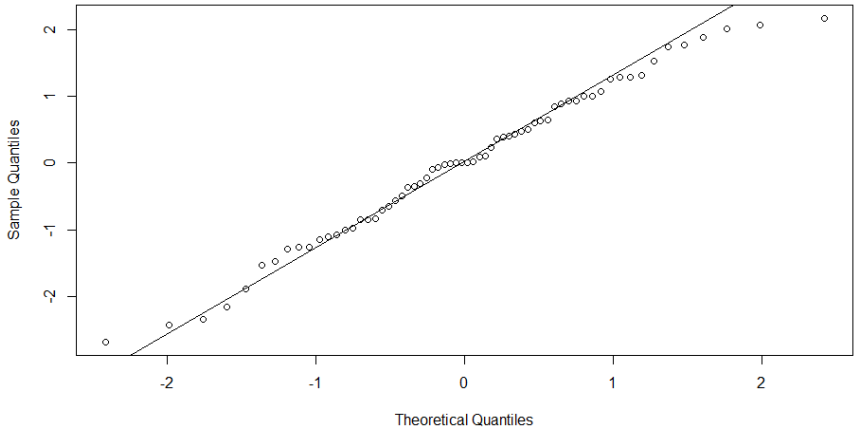


Figure 3: Normal Q-Q plot for sweighted2 residuals
Sample quantiles versus the theoretical quantiles.

Do not reject the null hypothesis that the sweight2 residuals have a normal distribution:

```
Shapiro-Wilk normality test

data:  res

W = 0.98403, p-value = 0.5769
```

Both the plot (Figure 3) and hypothesis test support the assumption that the sweight2 residuals are normally distributed.

6.2 Residual plots

The residual analysis also aimed to assess for serial dependence, that the predictors captured the variation in the data and that the sweight2 residuals are homoscedastic.

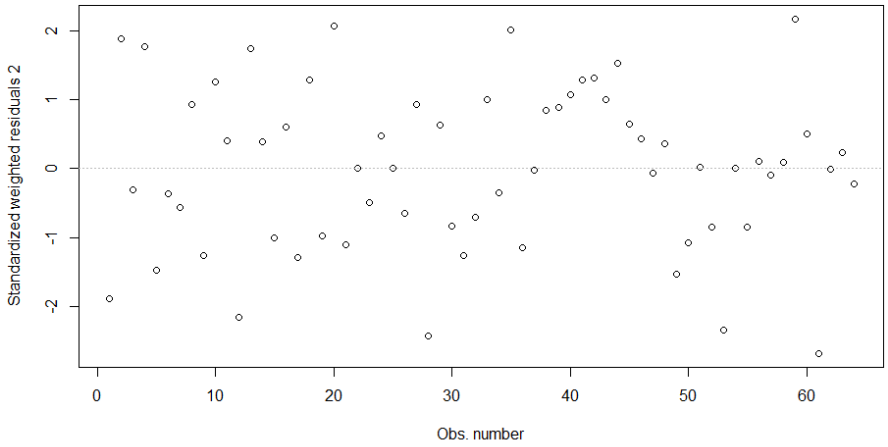


Figure 4: Sweighted2 residuals versus the indices of observation
Standardised weighted 2 residuals plotted versus the observation number.

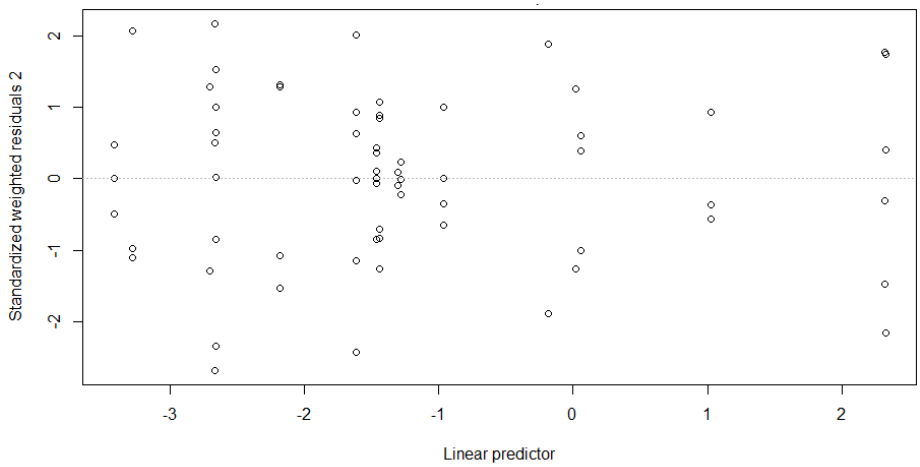


Figure 5: Sweighted2 residuals versus the linear predictor
Standardised weighted 2 residuals plotted versus the linear predictor value.

The sweight2 residuals lie in the interval [-2.6799, 2.1700] with median 0.0096 and the hypothesis that the mean is 0 is not rejected (p-value approximately 1 from a two-sided t-test). No patterns that suggest serial dependence is present, which supports the assumption that the responses are independent (Figure 4). Furthermore, no systematic patterns are evident in either location or spread and therefore indicates that the predictors captured the variation in the data and that the sweight2 residuals are homoscedastic (Figure 5).

6.3 Conclusion

This residual analysis suggests that the model is an adequate fit to the data.

7. Estimates

Fitted log(p/(1-p)) (linear predictor)：

	Red	OxD	OxM
Cys24	-0.18811735	2.32268594	1.02285796
Cys59	0.02162259	2.32721211	0.05965831
Cys89	-2.70678104	-3.28021239	-3.42218969
Cys178	-0.96894572	-1.61472243	-1.44114209
Cys191	-2.18378644	-2.65951267	-1.46338384
Cys223	-1.30812683	-2.66640047	-1.28460292

Fitted odds (p/(1-p) = exp(linear predictor)):

	Red	OxD	OxM
Cys24	0.82851748	10.20304189	2.78113178
Cys59	1.02185805	10.24932767	1.06147379
Cys89	0.06675133	0.03762027	0.03264088
Cys178	0.37948291	0.19894588	0.23665732
Cys191	0.11261431	0.06998232	0.23145175
Cys223	0.27032595	0.06950195	0.27676046

Fitted expected IAA labelling proportions (with corresponding fitted standard errors):

	Red	OxD	OxM
Cys24	0.45310886	0.91073853	0.73552892
	(0.02811790)	(0.01610497)	(0.02491266)
Cys59	0.50540544	0.91110580	0.51491016
	(0.02896718)	(0.01648856)	(0.02895599)
Cys89	0.06257441	0.03625629	0.03160913
	(0.03131257)	(0.02416710)	(0.02261953)
Cys178	0.27509069	0.16593400	0.19136855
	(0.05355759)	(0.04461790)	(0.04717934)
Cys191	0.10121595	0.06540512	0.18795032
	(0.09660356)	(0.07918784	(0.12512779)
Cys 223	0.21280046	0.06498534	0.21676772
	(0.04837084)	(0.02913210)	(0.04869647)

The fitted probability of IAA labelling was also plotted with an error bars for each combination of Cys and oxidation levels (Figure 6).

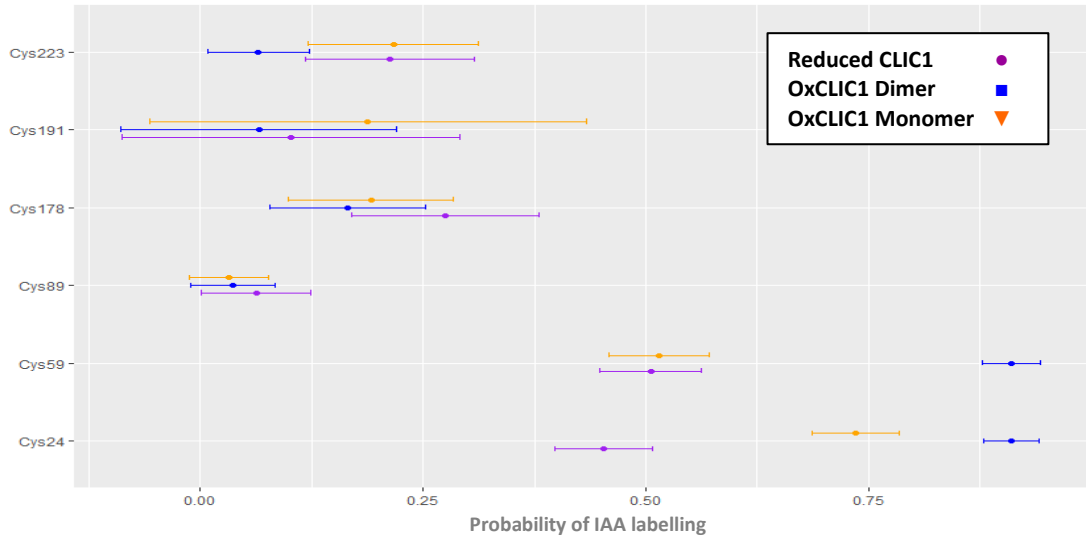


Figure 6: LC-MS/MS analysis of the CLIC1 redox species
Fitted probability of IAA labelling for each Cys residue and reduced CLIC1 (●), oxidised dimer (■) and oxidised monomer (▼) combination. Each error bar gives an approximate 95 % confidence interval for the probability of IAA labelling.

Appendix B: Rosetta scores and rmsBB values

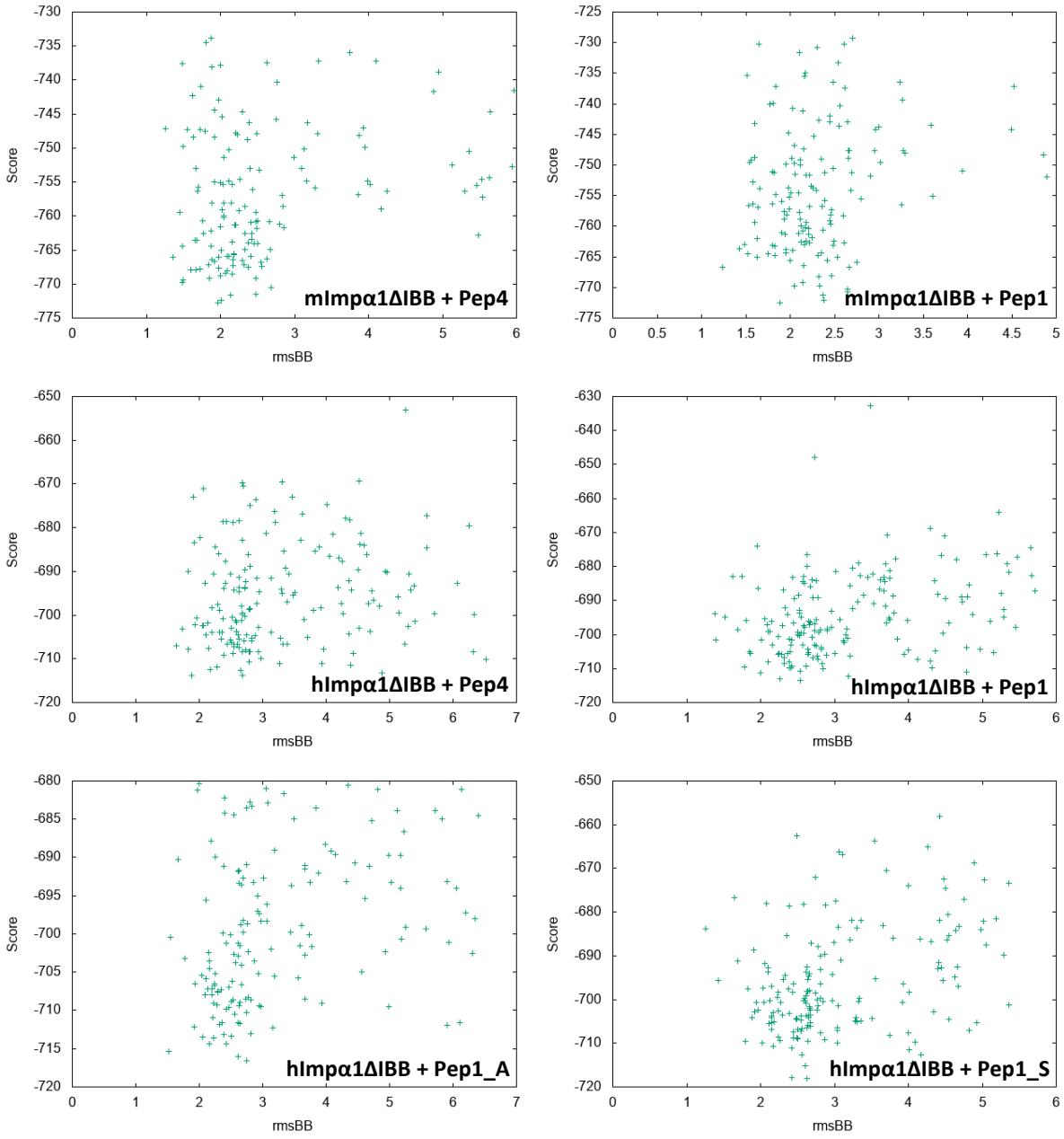


Figure 1: Rosetta scores and rmsBB values for all docked models
A total of 200 models were generated for each ImpαΔIBB and CLIC NLS peptide complex with 100 created for the high-resolution protocol and 100 created for including the low-resolution pre-optimisation protocol. The Rosetta scores and rmsBB values for all the models from each complex have been plotted.

Table 1: Modelling scores for the top ten poses for each ImpαΔIBB CLIC peptide complex
Rosetta scores and rmsBB values shown for mImpα1ΔIBB bound to Pep4 and Pep1, respectively, and hImpα1ΔIBB bound to Pep4, Pep1, Pep1_A and Pep1_S, respectively.

Pose Rank	mImpα1ΔIBB + Pep4		mImpα1ΔIBB + Pep1		hImpα1ΔIBB + Pep4		hImpα1ΔIBB + Pep1		hImpα1ΔIBB + Pep1_A		hImpα1ΔIBB + Pep1_S	
	Rosetta score	rmsBB	Rosetta score	rmsBB	Rosetta score	rmsBB	Rosetta score	rmsBB	Rosetta score	rmsBB	Rosetta score	rmsBB
1	-772.723	1.961	-772.483	1.884	-713.836	2.679	-713.623	2.537	-716.587	2.741	-717.981	2.630
2	-772.433	2.013	-772.107	2.381	-713.829	1.880	-712.982	2.260	-715.978	2.618	-717.790	2.419
3	-771.642	2.136	-771.337	2.371	-713.247	4.881	-712.315	3.190	-715.333	1.519	-715.220	2.601
4	-771.512	2.485	-770.747	2.644	-712.674	2.647	-711.405	1.995	-714.381	2.155	-714.661	4.060
5	-770.527	2.690	-770.340	2.652	-712.450	2.176	-711.118	4.787	-714.345	2.428	-712.661	4.173
6	-769.775	1.491	-769.803	2.330	-711.874	2.284	-710.737	2.535	-713.630	2.228	-712.625	2.560
7	-769.468	1.496	-769.785	2.047	-711.582	4.392	-710.164	2.840	-713.518	2.055	-711.563	4.014
8	-769.202	2.477	-769.221	2.146	-711.119	3.264	-709.833	2.391	-713.317	2.503	-711.019	2.423
9	-769.141	1.846	-768.212	2.320	-711.109	3.931	-709.802	2.410	-713.200	2.392	-710.632	2.171
10	-768.753	2.002	-768.106	2.470	-710.730	2.786	-709.717	4.316	-713.021	2.811	-709.981	2.020