

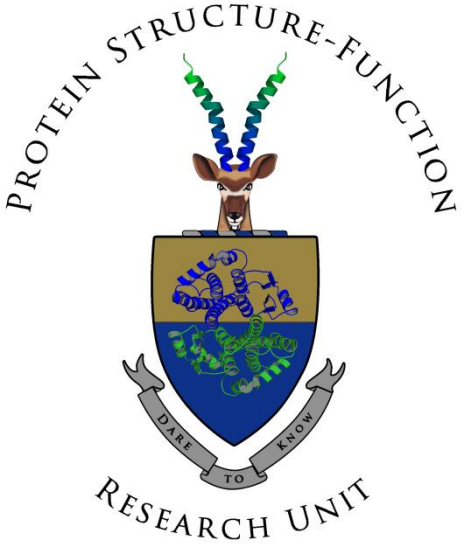
https://dspace.org/

Examining the effect of pH on the structure and stability of CLIC1 with E228L and E85L CLIC1 variants

Cross, Megan Olivia

2013-08-01
http://hdl.handle.net/10539/12956

Downloaded from a DSpace repository. https://dspace.org/



Examining the effect of pH on the structure and stability of CLIC1 with E228L and E85L CLIC1 variants

Megan Olivia Cross

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science.

Johannesburg, 2013

Declaration

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



Megan Olivia Cross
1 May 2013

Research Outputs

2011 - 2012

1. Publications

Achilonu, I., Fanucchi, S., Cross, M., Fernandes, M. and Dirr, H. W. (2012) Role of Individual Histidines in the pH-Dependent Global Stability of Human Chloride Intracellular Channel 1. *Biochemistry* **51,** 995–1004.

2. Conference presentations

Name	Date	Location	Title	Authors
SASBMB Congress	29 Jan - 1 Feb 2012	Drakensberg, South Africa	Recombinant CLIC1 and Sedlin do not interact: is there something missing? (Poster)	M. Cross H.W. Dirr S. Fanucchi
Wits Cross-Faculty Postgraduate Symposium	19 & 22 Oct 2012	University of the Witwatersrand, South Africa	Breaking glutamate interactions involved in pH-induced priming of CLIC1: Structure and stability of an E228L CLIC1 mutant (Poster)	M. Cross H.W. Dirr S. Fanucchi
Structural Biology in the Bio Economy	2 - 4 Dec 2012	University of Cape Town, South Africa	Breaking interactions for CLIC1 membrane competence: Structure and Stability of E85L and E228L CLIC1 variants (Poster)	M. Cross H.W. Dirr S. Fanucchi
Molecular Biosciences Research Thrust Research Day	5-Dec-12	University of the Witwatersrand, South Africa	Breaking interactions for CLIC1 membrane competence: Structure and Stability of E85L and E228L CLIC1 variants (Poster)	M. Cross H.W. Dirr S. Fanucchi

This work is dedicated to my parents, Tom and Carol Cross, who put up with me and my ramblings and loved me regardless,

and to Nana, who prayed for me every day. Twice.

The most all-penetrating spirit before which will open the possibility of tilting not tables, but planets, is the spirit of free human inquiry. Believe only in that.

Dmitri Mendeleev

Abstract

The chloride intracellular channel CLIC1 is an anion channel protein that has been implicated in a number of physiological processes. It is fascinating in that it is synthesised as a soluble monomer that is able to reversibly bind membranes without the aid of a membrane-targeting tag or receptor. CLIC1 membrane binding is promoted by low pH and involves separation of the N- and C-domains and subsequent refolding of the N-domain, which traverses the membrane as an α-helix. At the low pH of a membrane surface, pH 5.5, soluble CLIC1 demonstrates decreased conformational stability and forms a partially unfolded intermediate state under mild denaturing conditions. In this study, these pH-effects are proposed to occur as a result of low pH-induced protonation of two conserved glutamate residues, Glu85 and Glu228. Both are involved in domain-maintaining interactions and are proposed to form part of an electrostatic network of pH-sensitive residues. At low pH, protonation of these glutamates would break their electrostatic interactions, allowing separation of the domains. To investigate this possibility, Glu228 and Glu85 were mutated to leucine residues. Each variant protein was then investigated at pH 7.0 and pH 5.5 and results were compared to the wild-type. Secondary and tertiary structures were examined using far-UV circular dichroism and fluorescence spectroscopy, respectively. Conformational flexibility was investigated with limited thermolysin proteolysis. Stability was studied using thermal and urea-induced equilibrium unfolding. The unfolding intermediate state was detected using ANS binding and its structure was characterised. While neither residue substitution caused global structural perturbations, both destabilised the structure and promoted intermediate formation at pH 5.5. This was particularly evident for the E85L variant, which also formed a significant intermediate population at pH 7.0. It was concluded that the interactions of Glu228 and Glu85 are involved in maintaining the CLIC1 native state. Additionally, the lack of pH-dependence of intermediate formation in the E85L variant suggested that Glu85 is likely to function as a pH-sensor. It is thus involved in the 'priming' of the CLIC1 structure for the conformational changes that may lead to membrane binding.

Acknowledgements

Copious thanks are due to my supervisor, Dr Sylvia Fanucchi, for her insight, infinite patience and unwavering support, which kept me going.

I thank my co-supervisor, Prof. Heini Dirr for constantly asking questions, for his guidance and for teaching me to not be “so obsessive”.

Thanks are also due to Dr Manuel Fernandes for his masterful operation of the Proteum diffractometer and data collection wizardry.

Thank you to my friends, colleagues and fellow hermits in the Protein Structure-Function Research Unit, particularly Bradley Peter, Helen Letseka and Gavin Owen, for elucidating discussions, ineffable humour and late-night experimentals. Special thanks are due to Dr Derryn Legg-E’Silva for her encouragement and invaluable advice.

I am also extremely grateful to Prof. Heini Dirr, the University of the Witwatersrand and the National Research Foundation for the financial support that enabled this work.

Table of Contents

Declaration..... i
Research Outputs..... ii
Abstract..... iv
Acknowledgements..... v
List of figures..... viii
List of tables..... x
Abbreviations..... xi

CHAPTER 1. INTRODUCTION
1.1. Protein folding: an energetic balance..... 1
1.2. Forces maintaining the folded protein structure..... 4
1.3. Multi-state proteins: from water-soluble to membrane-bound..... 6
1.4. The effect of pH on protein structure and function..... 8
1.5. The CLIC protein family..... 9
1.5.1. Chloride intracellular channel 1..... 11
1.5.2. CLIC1 membrane insertion..... 14
1.5.3. The pH-response of CLIC1..... 15
1.6. Objectives..... 19

CHAPTER 2. MATERIALS AND METHODS
2.1. Materials..... 20
2.2 . Methods..... 20
2.2.1. Site-directed mutagenesis by PCR..... 20
2.2.2. Induction trials and protein production..... 23
2.2.3. Purification of wild-type, E228L CLIC1 and E85L CLIC1..... 24
2.2.4. Sodium dodecyl sulphate polyacrylamide gel electrophoresis..... 25
2.2.5. Protein concentration determination..... 26
2.2.6. Structural characterisation..... 27
2.2.6.1. Far-UV circular dichroism spectroscopy..... 27
2.2.6.2. Fluorescence spectroscopy..... 27
2.2.6.3. pH-dependence of secondary and local tertiary structure..... 30
2.2.7. X-ray crystallography..... 30
2.2.7.1. Crystallisation of E228L CLIC1..... 30
2.2.7.2. Data collection and phasing..... 32
2.2.7.3. Refinement and validation..... 33
2.2.8. Conformational stability..... 33
2.2.8.1 Thermal unfolding..... 34
2.2.8.2. Reversibility studies..... 34
2.2.8.3. Urea-induced equilibrium unfolding..... 35
2.2.8.4. Thermolysin proteolysis..... 36

CHAPTER 3. RESULTS
3.1. Creation and purification..... 38
3.2. Structures of wild-type, E228L CLIC1 and E85L CLIC1..... 38
3.2.1. Secondary structure..... 38
3.2.2. Tertiary structure..... 43
3.2.3. pH-Dependence of structure..... 43
3.2.4. Conformational flexibility..... 43
3.2.5. The E228L CLIC1 crystal structure..... 46
3.2.5.1. E228L CLIC1 and E85L CLIC1 crystals..... 46
3.2.5.2. E228L CLIC1 structural model..... 51

3.3. Stability of wild-type, E228L CLIC1 and E85L CLIC1..... 56
3.3.1. Thermal stability..... 56
3.3.2. Unfolding equilibration time..... 58
3.3.3. Reversibility of folding..... 58
3.3.4. Equilibrium unfolding..... 62
3.3.5. The unfolding intermediate..... 68

CHAPTER 4. DISCUSSION
4.1. Structural roles played by Glu228 and Glu85..... 76
4.2. Substitution of Glu228 or Glu85 reduces conformational stability and promotes intermediate formation..... 77
4.3. Primed for conformational change: the CLIC1 intermediate state..... 79
4.4. Glu228: a team player..... 81
4.5. Glu85: a pH-sensor..... 81
4.6. Conclusion..... 82

CHAPTER 5. REFERENCES..... 84

List of figures

Figure number	Title	Page no.
1	Representative protein folding funnels	2
2	Sequence-based alignment of the available human CLIC proteins and invertebrate CLIC-like proteins	10
3	Structure of soluble monomeric CLIC1	12
4	Comparison of the reduced CLIC1 monomer and a subunit of the oxidised CLIC1 homodimer	13
5	Conserved glutamate residues that form H-bonds or salt bridges in soluble CLIC1	18
6	Successful generation of the E228L and E85L mutations	39
7	Production of the GST-E228L-CLIC1 and GST-E85L-CLIC1 fusion proteins in *E. coli* BL21(DE3) pLysS cells	40
8	Purification of wild type, E228L CLIC1 and E85L CLIC1	41
9	Secondary structures of wild-type, E228L CLIC1and E85L CLIC1 at pH 7.0 and pH 5.5	42
10	Intrinsic fluorescence of wild-type, E228L CLIC1 and E85L CLIC1 at pH 7.0 and pH 5.5	44
11	pH-dependence of the secondary and tertiary structures of wild-type, E228L CLIC1 and E85L CLIC1	45
12	Limited thermolysin proteolysis of wild-type, E228L CLIC1 and E85L CLIC1 at pH 7.0 and pH 5.5	47
13	Phase diagram of protein crystallisation as a balance between protein and precipitant concentration	48
14	E228L CLIC1 crystals	50
15	Confirmation of the proteinaceous nature of E228L CLIC1 crystals	50
16	Crystal contacts in the E228L CLIC1 asymmetric unit	52
17	Overlay of the polypeptide backbones of the E228L CLIC1 and wild-type CLIC1 structures	54
18	Electron density for residue 228	55
19	Thermal unfolding of wild-type, E228L CLIC1 and E85L CLIC1	57
20	Unfolding of E228L CLIC1 and E85L CLIC1 over time	59
21	Unfolding and refolding of E228L CLIC1 at pH 7.0 and pH 5.5	60
22	Unfolding and refolding of E85L CLIC1 at pH 7.0 and pH 5.5	61

23	Pulse proteolysis of unfolding CLIC1	63
24	Equilibrium unfolding of wild-type, E228L CLIC1 and E85L CLIC1	64
25	Comparison of urea-induced equilibrium unfolding curves of wild-type, E228L CLIC1 and E85L CLIC1	66
26	Intensity averaged emission wavelength for CLIC1 unfolding	67
27	ANS binding to wild-type, E228L CLIC1 and E85L CLIC1 at pH 7.0 and pH 5.5	69
28	Structure of the intermediate state	70
29	Stern-Volmer plots of wild-type, E228L CLIC1 and E85L CLIC1	72
30	Conformational flexibility of the intermediate state	75
31	Proposed mechanism for pH-induced ‘priming’ of CLIC1 for conformational change via glutamate pH-sensors	83

List of tables

Table number	Title	Page no.
1	Locations, interactions and conservation of the glutamate residues in human CLIC1	17
2	Suppliers used to obtain non-standard laboratory materials	21
3	Data collection and refinement statistics for E228L CLIC1	53
4	K_{sv} values for dynamic quenching of tryptophan fluorescence in the native and intermediate states of the various CLIC1 proteins	73

Abbreviations

$[\Theta]_{xxx}$	mean residue ellipticity at xxx nm
Å	Angström
A.U.	arbitrary units
ANS	8-anilino-1-napthalene sulphonate
ASU	asymmetric unit of a crystal
CD	circular dichroism
CLIC	chloride intracellular channel
C_m	denaturant concentration at which the native and unfolded species are equally populated
DEAE	diethylaminoethyl
$\Delta G°$	change in standard Gibbs free energy
$\Delta H°_{conf}$	change in standard enthalpy due to bond formation within a protein molecule
$\Delta H°_{solv}$	change in standard enthalpy due to dissociation of bonds between the protein and the solvent
$\Delta S°_{conf}$	change in standard entropy of the protein molecule upon folding
$\Delta S°_{solv}$	change in standard entropy of the solvent
DTT	dithiothreitol
ε_{xxx}	molar extinction coefficient at xxx nm
E. coli	*Escherichia coli*
E228L	a substitution of glutamate 228 with leucine
E85L	a substitution of glutamate 85 with leucine
EDTA	ethylenediaminetetraaceticacid
Far-UV	region of light of wavelength 180 - 250 nm
FRET	Förster resonance energy transfer
Fxxx	fluorescence intensity at xxx nm

GSH	reduced glutathione
GST	glutathione transferase
GST-E228L CLIC1	A fusion protein of GST and the E228L CLIC1 variant
GST-E85L CLIC1	A fusion protein of GST and the E85L CLIC1 variant
H-bond	hydrogen bond
I	intermediate state of a protein
IAEW	intensity averaged emission wavelength
IPTG	isopropyl β-D-1-thiogalactopyranoside
kDa	kilodalton
K_{sv}	The Stern-Volmer quenching constant
λ_{max}	wavelength of maximum fluorescence emission
mdeg	millidegrees
min	minutes
m-value	the dependence of unfolding on denaturant concentration
Mw	molecular weight
$n \rightarrow \pi^*$	transition of electrons from a nonbonding to an antibonding pi atomic orbital
N	the folded native state of a protein
OD_{xxx}	optical density at xxx nm
PDB	the Protein Data Bank
PEG	polyethylene glycol
PFT	pore-forming toxin
$\pi \rightarrow \pi^*$	transition of electrons from a bonding to an antibonding pi atomic orbital
p*I*	isoelectric point
pK_a	pH at which an ionisable group will bind or release a proton

rmsd	root mean square deviation
θ_{xxx}	ellipticity at xxx nm
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
TEMED	N,N,N´,N´-tetramethylethylenediamine
T_m	temperature at the midpoint of thermal unfolding
TMR	transmembrane region
Tris-HCl	buffering system consisting of tris(hydroxymethyl)aminomethane (Tris) and HCl
T_{unf}	temperature at which unfolding is initiated
U	unfolded state of the protein
WT	wild-type
YT	yeast tryptone

The one- and three-letter amino acid naming codes are used in accordance with the IUPAC-IUBMB naming system.

CHAPTER 1. INTRODUCTION

1.1. Protein folding: an energetic balance

Proteins are responsible for a number of processes in the cell. Their functions are extremely diverse and intrinsically linked to their structures. Proteins are synthesised by ribosomes, as linear polymers of amino acids which are then able to fold into the functional 3-dimensional structures that form an integral part of cellular functioning. Much work has therefore focussed on the folding process and the elucidation of how the protein is maintained in its functional (or native) state.

The 'blueprint' for the native state of a protein is contained within its primary structure: the order and identities of the amino acids that it is made up of (Anfinsen, 1973). *Ab initio* prediction of the native state of a protein, however, remains difficult and many sequence-to-structure prediction algorithms that are currently available require a known structure with which the target shares homology (Schwede *et al.*, 2003; Kelley and Sternberg, 2009). This is due to the diverse conformational states available to a nascent polypeptide chain (Levinthal, 1968; Privalov, 1979). Despite the large number of available conformations, most proteins are able to fold rapidly. This enigma is known as Levinthal's Paradox (Levinthal, 1968) and suggests that folding is not merely a stochastic search of all the available conformational states, but rather a directed, hierarchical process (Baldwin, 1995). Many varied models of folding have been proposed, including, but not limited to, the diffusion-collision (Karplus and Weaver, 1979), nucleation-condensation (Fersht, 1997) and hierarchic condensation (Lesk and Rose, 1981; Maity *et al.*, 2005) models. None of these are mutually exclusive and most agree with Anfinsen's proposal (1973) of the native state as an energy minimum on the protein folding landscape.

The folding landscape of a protein is an energetic one, which correlates the number of conformations available to a folding protein, its conformational entropy, with the stabilising free energy associated with each conformation. If these are represented as horizontal and vertical coordinates, the landscape takes on a funnel shape, with the native state forming the energy minimum at the bottom of the funnel (Figure 1). A large number of states are available to an unfolded protein, thus the top of the funnel is wide. As the protein folds, the number of states it can assume decreases, as does its free energy, leading to a narrowing of the funnel as the native state is neared (Bryngelson *et al.*, 1995; Dill and Chan, 1997; Dobson, 2004).

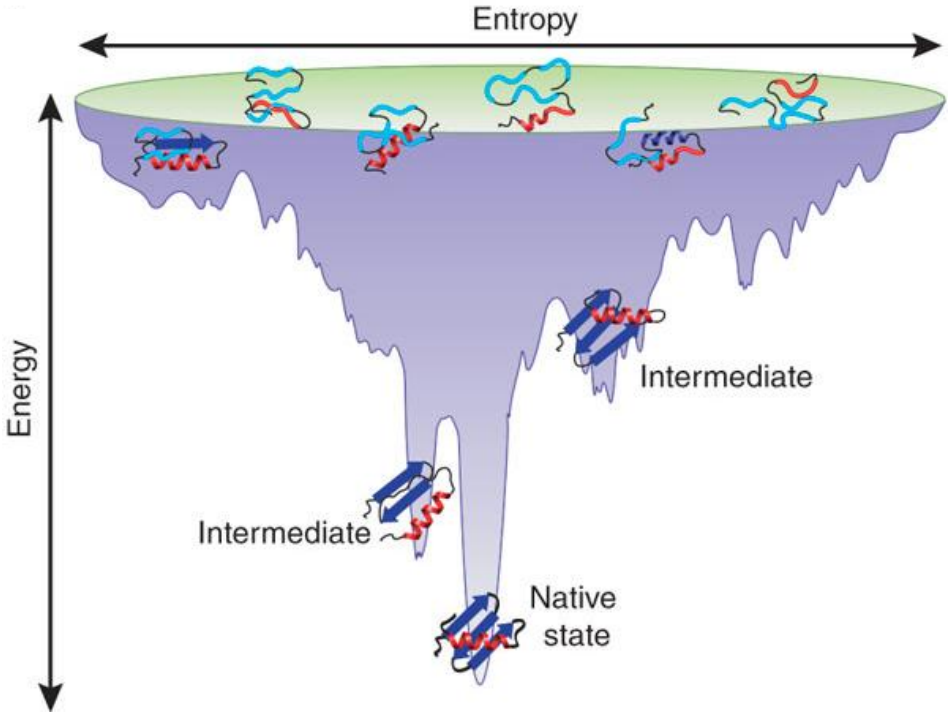


Figure 1. Representative protein folding funnels

The width at the top of the folding funnel corresponds to the conformational entropy and thus, number of possible conformational states. The vertical height of the funnel represents the free energy, which is lowest in the native state. Energy minima on the funnel landscape correspond to folding intermediate states. Folding chains tend towards the lowest energy state, which in turn, reduces the number of possible conformations until a single native state is reached. Image is taken from Bartlett and Radford, 2009.

This so-called 'folding funnel' is useful in explaining the enigma of Levinthal's paradox in that it demonstrates how multiple conformations of an unfolded protein may still fold to a single native state. Troughs and dips on the funnel represent folding intermediates such as the molten globule state, an on-pathway folding intermediate of several proteins (Jamin and Baldwin, 1998; Maki *et al.*, 1999) believe to be a common state in all proteins (Fink, 1995). The molten globule is characterised by (1) defined secondary structure, (2) decreased packing in tertiary structure, (3) increased compactness of the structure and (4) the exposure of hydrophobic residues to the solvent (Arai and Kuwajima, 2000). While the molten globule and similar intermediates may increase the rate of folding (Reviewed in Tsytlonok and Itzhaki, 2012), others represent 'kinetic traps' in which the protein may become misfolded if it is unable to overcome the energy barriers that form the walls of the trough. These misfolded intermediates are often associated with disease (Herczenik and Gebbink, 2008). The ruggedness of the energy landscape appears to increase with the length of the polypeptide chain, suggesting that the folding of larger proteins to a thermodynamically stable state is highly complex (Dobson *et al*., 1998).

The thermodynamic stability of the native state of a protein can be described in terms of its Gibbs free energy ($\Delta G°$), which is a combination of forces that both favour and oppose folding. Under constant pressure:

$$\Delta G° = \Delta H° - T\Delta S° \quad (1)$$

$\Delta H°$ is the enthalpy, $\Delta S°$ is the entropy and T is the Kelvin temperature. For folding to be spontaneous, the reaction must be exergonic and thus, $\Delta G°$ must be negative. The enthalpy term is the heat associated with the breaking and formation of bonds and is influenced by changes in the association of the protein with solvent ($\Delta H°_{solv}$), as well as the formation of bonds that stabilise the folded structure ($\Delta H°_{conf}$) (Reviewed in Dill, 1990. See Section 1.2.). The entropy term, $\Delta S°$, as mentioned above, denotes the ability of the protein chain to explore all available conformations ($\Delta S°_{conf}$). It also includes a contribution by the solvent ($\Delta S°_{solv}$), which becomes more disordered when the chain folds and hydrophobic residues are buried in the core of the structure. Upon folding, the loss of interactions with the solvent ($+\Delta H°_{solv}$) and confinement of the protein chain in a single conformation ($-\Delta S°_{conf}$) oppose folding. This is compensated for by the formation of bonds within the folded protein chain ($-H°_{conf}$) and the increased entropy of the solvent ($+\Delta S°_{solv}$), both of which

promote folding. Thus, spontaneous folding of the protein chain is made possible through an imbalance of these forces such that $\Delta G°$ is negative:

$$\Delta G° = (+\Delta H°_{conf} - \Delta H°_{solv}) - T\,(-\Delta S°_{conf} + \Delta S°_{solv}) \quad (2)$$

The native (N) and unfolded (U) states of a protein exist in constant equilibrium. Under native conditions, the unfolded population is not favoured and is hence undetectable. This equilibrium can be shifted, however, as the difference between the free energies of the native and unfolded states is low and calculated to be ~ 8 - 40 kJ/mol (Pace 1975; Privalov, 1979). The native and unfolded states are separated by an energy barrier known as the transition state. If a shift in the N ↔ U equilibrium is to be achieved, the energy peak of the transition state must be overcome (Onuchic and Wolynes, 2004).

1.2. Forces maintaining the folded protein structure

The amino acids of a protein are connected by strong covalent peptide bonds. Additional covalent bonds linking amino acids are seen in the form of disulphide bonds. These form in cysteine-containing proteins in non-reducing environments (Reviewed in Creighton, 1988). The majority of structure-stabilising forces, however, are non-covalent. These vary in specificity from the more general hydrophobic interactions that occur between non-polar side chains, to electrostatic salt bridges formed between specific amino acids.

Hydrophobic interactions

Hydrophobic residues tend to cluster together in the core of a protein structure in order to minimise their contact with the polar solvent. This greatly reduces the associated entropic cost and is proposed to be the driving force of protein folding, known as the 'hydrophobic effect' (Kauzmann, 1959; Dill, 1990; Lins and Brasseur, 1995). The sequestering of hydrophobic residues through close van der Waals-like interactions (see below) contributes an average of approximately 250 kJ/mol to conformational stability (Dill *et al.*, 1989; Pace *et al*., 1996). Recently, hydrophobic interactions were proposed to be the dominant force in the stability of proteins, as an estimated 60.4% of the overall conformational stability is contributed to by hydrophobic packing (Pace *et al.*, 2011). This is perhaps why on average 81% of hydrophobic residues are found buried in the core of the protein (Lesser and Rose, 1990), which is generally well-packed with very few cavities (Richards, 1977; Chen and Stites, 2001).

Van der Waals forces

Van der Waals interactions are electrostatic and involve fixed or induced dipoles. For this reason, they occur most frequently between residues with the most polarisable side chains. As these are residues with methyl or methylene groups, van der Waals forces tend to occur between the non-aromatic hydrophobic residues (Petsko and Ringe, 2004). They are potentially the weakest of the protein-stabilising interactions with an energetic contribution of 8 – 16 kJ/mol (Lins and Brasseur, 1995; Petsko and Ringe, 2004) and are highly distance-dependent: the strength of the interaction is inversely proportional to the distance between the interacting nuclei raised to the sixth power (Pace, 2001; Petsko and Ringe, 2004). Their role in protein stability appears to be similar to that of the H-bonds discussed below, in that it is their combined energetic contribution that promotes protein stability.

Hydrogen bonds

Hydrogen bonds (H-bonds) are electrostatic (dipole-dipole) interactions between a hydrogen atom bound to an electronegative donor atom (usually N, F or O), and a lone pair of electrons on an electronegative acceptor atom. H-bonds, like van der Waals forces, are directional and tend to be in the range of 8 – 40 kJ/mol (Pauling, 1960) though this value decreases when H-bonds distort away from linearity. The merit of H-bonds lies not in their bond strength, but in their large numbers, which ensure that contacts are maintained but allow for flexibility through the breaking and reforming of local H-bonds in mobile regions (Jeffrey, 1997). The most prominent role of H-bonds is in the formation of the secondary structural elements (Pauling and Corey, 1951; Pauling *et al*., 1951) as 80% of the residues in a protein form either α-helices or β-strands (Stickle *et al.,* 1992). H-bonds were initially proposed to be the driving force of protein folding (Mirsky and Pauling, 1936). While this has since been proven unlikely, they are still considered one of the two fundamental factors in protein stability (Dill, 1990; Myers and Pace, 1996).

Salt bridges

Salt bridge interactions occur between oppositely charged side chains and may result in H-bonds when they are shorter than 3.5 Å. In proteins, the positively charged histidine, lysine and arginine side chains are able to form salt bridges with the anionic glutamate and aspartate side chains (Dill, 1990). Salt bridge interactions may be classified as either direct or indirect. Indirect interactions are ‘background interactions’ occurring between a charged residue and a non-ionisable group, for example a bond between a charged side chain and permanent dipole of the peptide bond. Indirect

interactions also include changes to the orientation of the solvent in order to solvate exposed charges. Direct salt bridges are those between two charged residues (Bosshard *et al.,* 2004). Such coulombic interactions are distance- and geometry-dependent and the energy (U) of the interaction between the two charges (q_i and q_j) is dictated by Coulomb's law:

$$U = \frac{k\, q_i q_j}{D\, r_{ij}} \qquad (3)$$

where k is the Coulombic proportionality constant, D is the dielectric constant and r_{ij} is the distance between the two charges (Bosshard *et al.,* 2004). The upper limit for a salt bridge interaction was proposed to be 4 Å (Jeffrey, 1997), though the average distance is 2.8 Å (Petsko and Ringe, 2004).

Salt bridges contribute a stabilising free energy of approximately 160 kJ/mol (Honig *et al.,* 1986), though this is decreased in interactions that are more solvent-exposed where solvent shielding occurs (Kumar and Nussinov, 1999). Given the pH-sensitivity of the ionisation states of the charged amino acids, salt bridge interactions may vary with pH. The stabilising effect of a salt bridge-forming residue is seen to correlate with the overall difference in the pK_a values of the residues in the native and unfolded states and the nature of the charge on the side chain (+ or -) (Bosshard *et al.,* 2004). The effect of pH on salt bridge interactions is well reported and pH-induced breakage of these interactions often facilitates protein function (Atilgan *et al.,* 2011; Bertrand *et al.,* 2012; Roy *et al.,* 2012; Yokoyama *et al.,* 2012).

1.3. Multi-state proteins: from water-soluble to membrane-bound

The presence of extensive networks of relatively weak bonds in proteins facilitates structural plasticity and allows many proteins to change their structures in order to function. Many of these are proteins that are able to partition from the soluble to the membrane phase including the pore-forming toxins (PFTs), Bcl family of apoptotic proteins and the chloride intracellular channels (CLICs).

Pore-forming toxins have been colloquially termed "hole punchers" as they kill cells by creating pores in membranes either to grant access to other toxic factors, or to facilitate leakage (Reviewed in Tilley and Saibil, 2006). PFTs can be classified according to the secondary structure of the pores they form as either α- or β-PFTs. As their name suggests, the β-PFTs are those that form β-sheeted pore structures. These include α-hemolysin, anthrax toxin and the cholesterol-dependent cytosolins (CDCs), perfringolysin and pneumolysin (Tilley and Saibil, 2006). A central trend in the membrane

association of the β-PFTs appears to be conformational change in response to low pH. This seen in low pH-induced oligomerisation of the membrane-binding segment of anthrax toxin, PA_{63}, (Koehler and Collier, 1991; Lacy *et al.,* 2004), the refolding of α-helices to β-hairpin structures in the membrane binding of the CDCs and the separation of the domains of pneumolysin to "swing out" the transmembrane region for binding (Tilley *et al.,* 2005).

The α-PFTs have α-helical membrane domains and include the colicins, diphtheria toxin, the Cry δ-endotoxins and exotoxin A (Tilley and Saibil, 2006). At low pH, the pore-forming domain of colicin A is destabilised and membrane association is promoted via a molten globule state (Van der Goot *et al.,* 1991; Muga *et al.,* 1993). While molten globules have been implicated in membrane association, as seen in colicin A (Van der Goot *et al.*, 1991), they do not appear to be a requirement for membrane association (Thuduppathy *et al.,* 2006a). The membrane binding of diphtheria toxin is similar to that of colicin A in that it has a similar pH effect (Sandvig and Olsnes, 1980) and also occurs via a molten globule state (Choe *et al.,* 1992). Here, low pH is proposed to trigger the exposure of the hydrophobic membrane-binding domain (Sandvig *et al.,* 1986) and neutralise a number of negatively charged residues, making the binding interface more hydrophobic. This promotes insertion into the membrane by lowering the associated energetic cost. Similar pH-induced mechanisms have been proposed for both tetanus toxin (Boquet and Duflot, 1982) and exotoxin A (Mere *et al.,* 2005), however these mechanisms are not limited to the PFTs.

The Bcl-2 protein family shares high homology with the α-PFTS mentioned above and includes both promoters of cell survival (Bcl-x_L and Bcl-2) and cell death (Bax, Bak and Bid) (Tilley and Saibil, 2006). Of these, Bcl-x_L (Minn *et al.*, 1997), Bax (Antonsson *et al.,* 1997), Bcl-2 (Schendel *et al.*, 1997) and Bid (Schendel *et al.,* 1997), have thus far been shown to have pore-forming activity. Like many PFTs, membrane insertion of Bcl-x_L is promoted by low pH (Thuduppathy *et al.*, 2006b). This is proposed to be via an electrostatic mechanism, and to involve two conserved titratable residues, an aspartate and a glutamate (Thuduppathy *et al.*,2006a). Acidic pH also stabilises the formation of Bcl-x_L dimers. As this is a reported prerequisite for Bax channel formation (Antonsson *et al.,* 2001), it is possible that pH may also promote membrane association by Bax.

Analysis of the above examples of proteins that transit from soluble to membrane-associated forms highlights two central trends: conformational change and a role for acidic pH. The surface of a membrane is proposed to be approximately 2 pH units more acidic than the cytosol, due to the

attraction of protons by the anionic phospholipid groups of the membrane (Vaz *et al.,* 1978). The resultant pH gradient regulates the membrane binding and functions of the above-mentioned soluble-to-membrane proteins and serves as single example of the important role played by pH in protein functionality.

1.4. The effect of pH on protein structure and function

The structure and conformational stability of a protein is highly dependent upon the pH of the environment. Variations in pH may cause changes to the protonation states of the ionisable amino acid side chains as well as the N- and C-terminal groups of the protein, potentially resulting in the disruption or formation of structure-stabilising interactions (Reviewed by Pace *et al.,* 2009). While bonding alterations may lead to protein aggregation at the extremes of the pH scale, changes observed within the pH range that is physiologically-relevant for a particular protein are often tied to its function. Thus pH has been shown to be involved in the regulation of a diverse array of processes, including enzyme activation, membrane association, protein-protein interactions, cell migration and apoptosis (Srivastava *et al.,* 2007).

pH-induced protonation or deprotonation most commonly occurs on histidine residues due to the relatively neutral theoretical pK_a (6.5 (Lide, 1991; Srivastava *et al.,* 2007)) of the imidazole side chain (Gerchman *et al.,* 1993; Srivastava *et al.,* 2007; Fritz *et al.,* 2008; Achilonu *et al.,* 2012). However, pH-effects have been reported for all of the ionisable residues, with glutamate and aspartate second to histidine in frequency (Alexov and Gunner, 1997; Di Russo *et al.,* 2012; Kloda *et al.,* 2006; Roy *et al.,* 2012). This diversity is possible because of the significant effect that the environment of a titratable group may have on its pK_a (Pace *et al.,* 2009; Di Russo *et al.,* 2012). Thus, given an appropriate environment, any of the ionisable amino acids may function as so-called 'pH-sensor' residues.

The loss or acquisition of a proton changes the charge and H-bonding potential of a pH-sensing residue and may thereby enable a protein to function. Enzyme activation is made possible by changes to the ionisation states of active site residues (Rebek, 1990), which results in the highly-specific pH optima of enzymes. Protonation changes may also enable a protein to function by disrupting H-bond and salt bridge interactions that maintain it in a non-functional form. Breakage of such bonds decreases conformational stability (Vishnivetskiy *et al.,* 2012) and thereby, induces conformational changes. Proteins that spontaneously alter their structures with pH variation include

many of the PFTs discussed previously (Section 1.3), the profurin zymogen (Feliciangeli *et al.*, 2006), the envelope protein involved in flavivirus-endosome fusion (Fritz *et al.*, 2008) as well as the NO-binder, nitrophorin 4 (Di Russo *et al.*, 2012). More indirect facilitation of conformational change is seen in mammalian cofilin, which is responsible for severing actin filaments. Here, histidine protonation breaks a His-Asp salt bridge but does not cause significant global conformational changes. Since the activation of cofilin requires both histidine deprotonation and phosphorylation, it has been suggested that the breakage of the His-Asp bond 'primes' the protein for the conformational changes induced upon phosphorylation (Hawkins *et al.*, 1993; Pope *et al.*, 2004; Srivastava *et al.*, 2007). A similar 'priming' for conformational change has recently been proposed (Stoychev *et al.*, 2008; Legg-E'Silva, 2012) for a member of the chloride intracellular channel (CLIC) protein family, which demonstrates increased membrane association and anion transport activity at low pH.

1.5. The CLIC protein family

The chloride intracellular channel proteins are atypical anion channel proteins. Unlike classical chloride channels (Reviewed by Jentsch *et al.*, 2002), they are synthesised as soluble cytosolic proteins, which are able to reversibly bind membranes without the aid of membrane-binding tags or receptors (Tulk *et al.*, 2000; Tulk *et al.*, 2002, Singh and Ashley, 2007; Goodchild *et al.*, 2009). The CLICs are classified as part of the glutathione transferase (GST) superfamily as they share high structural similarity with the GSTs, in particular the omega-class GSTs. The key difference, however, is that the CLICs are monomeric, while most GSTs are dimeric (Harrop *et al.*, 2001). Seven CLIC1 orthologues exist in humans: CLIC1 – CLIC6, with CLIC5 having two splice variants (A and B). The CLICs show high sequence similarity of between 46 – 76% (Cromer *et al.*, 2002) (Figure 2). They have thus been proposed to originate from a single ancestor, postulated by Littler *et al.* (2010) to be the chordate *Ciona intestinalis* as it has a single CLIC protein that has 45% homology with the human CLICs. CLIC-like proteins have also been observed in invertebrates in the form of dmCLIC (*Drosophila melanogaster*) and EXC-4 and EXL1 (*Caenorhabditis elegans*), though these show lower sequence identity (35%) with their vertebrate equivalents (Littler *et al.*, 2010).

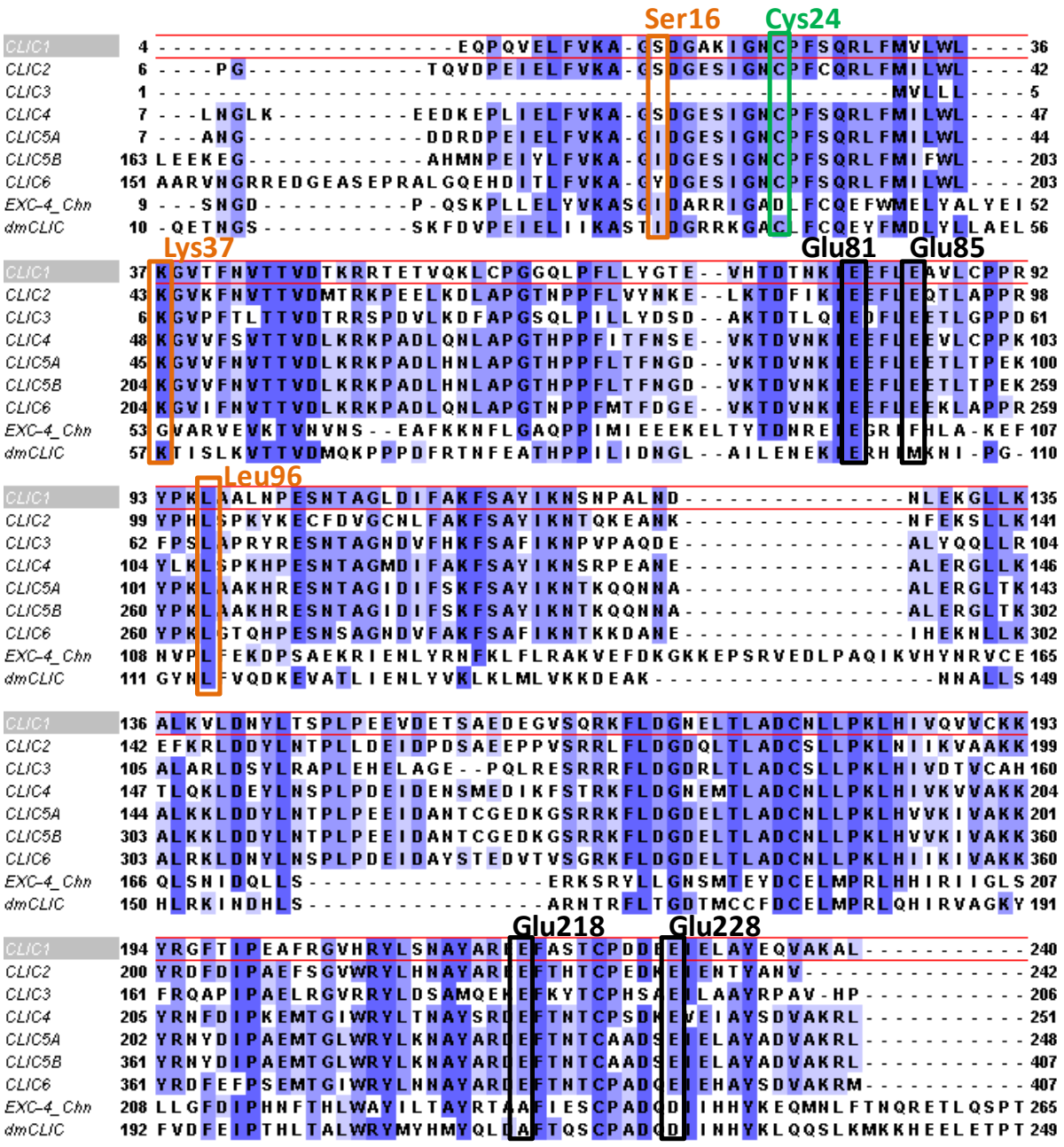


Figure 2. Sequence-based alignment of the available human CLIC proteins and invertebrate CLIC-like proteins

Amino acid sequences of human CLIC1 - 6 and the CLIC-like proteins from *Drosophila melanogaster* (dmCLIC) and *Caenorhabditis elegans* (EXC-4) were aligned using T-Coffee v8.93 (http://www.ebi.ac.uk/Tools/services/rest/tcoffee) with an alignment score of 87. Sequences are shaded according to identity: residues shown in darker blue blocks are most conserved. The CLIC1 sequence is highlighted with the red block. Residues of interest are highlighted with blocks following the numbering of the CLIC1 sequence: Glu85 and Glu228 are in black, Ser16, Lys37 and Leu96 are in orange and Cys 24 is in green.

The cellular distribution of the CLICs is broad (Duncan *et al.,* 1997; Valenzuela *et al.,* 1997; Tulk and Edwards, 1998; Chuang *et al.,* 1999) and they have been implicated in a wide range of processes (Reviewed in Littler *et al.,* 2010; Singh, 2010), though their key physiological role remains elusive. Attempts to uncover the physiological roles of the CLICs using knock-out mice have not been successful (Qiu *et al.,* 2009). This, it is suggested, is due to a high level of redundancy amongst the CLICs, where the loss of one is compensated for by another (Littler *et al.,* 2010).

1.5.1. Chloride intracellular channel 1

The anion channel activity of chloride intracellular channel 1 (CLIC1) has been implicated in a number of processes including cell cycle control (Valenzuela *et al.,* 2000), the oxidative stress response (Averaimo *et al.,* 2010) and β-amyloid toxicity (Novarino *et al.,* 2004). While these implicated functions are different from those of the other CLICs (Reviewed in Littler *et al.,* 2010), what truly sets CLIC1 apart is the structural metamorphism of the soluble state. This makes it part of an emerging class of 'metamorphic' proteins (Murzin, 2008), each of which has multiple stable soluble forms. Under reducing conditions, CLIC1 exists as a monomer with a thioredoxin fold (Figure 3). Upon oxidation, the protein forms a homodimer that is exclusively α-helical (Littler *et al.,* 2004) (Figure 4). The significance of this dimeric state remains unclear however, as it is has not been reported for other members of the CLIC family. The CLIC1 dimer interface is hydrophobic (Littler *et al.,* 2004) which has led to the proposal that, in the presence of membranes, the hydrophobic interface would be used for membrane association instead of dimerisation. While oxidative conditions have been shown to promote membrane binding (Goodchild *et al.,* 2009), they are not a prerequisite. In the absence of oxidising agents, recombinantly-expressed CLIC1 is able to bind and form anion channels *in vitro* that demonstrate ion channel activity equal to that measured *in vivo* (Warton *et al,* 2002). Thus, this study focuses on the monomeric form of soluble CLIC1.

CLIC1 consists of 241 amino acids and has a molecular weight of 26.9 kDa when in the soluble monomeric form (Littler *et al.,* 2004). The protein consists of two domains: an all α-helical C-domain and a thioredoxin N-domain (Figure 3). The N-domain spans a 4-stranded mixed β-sheet (strands β1, β2, β3 and β4) and three α-helices (α1, α2, α3). The C-domain contains α-helices α4-10 as well as an anionic "foot loop" that is a characteristic feature of all CLIC proteins.

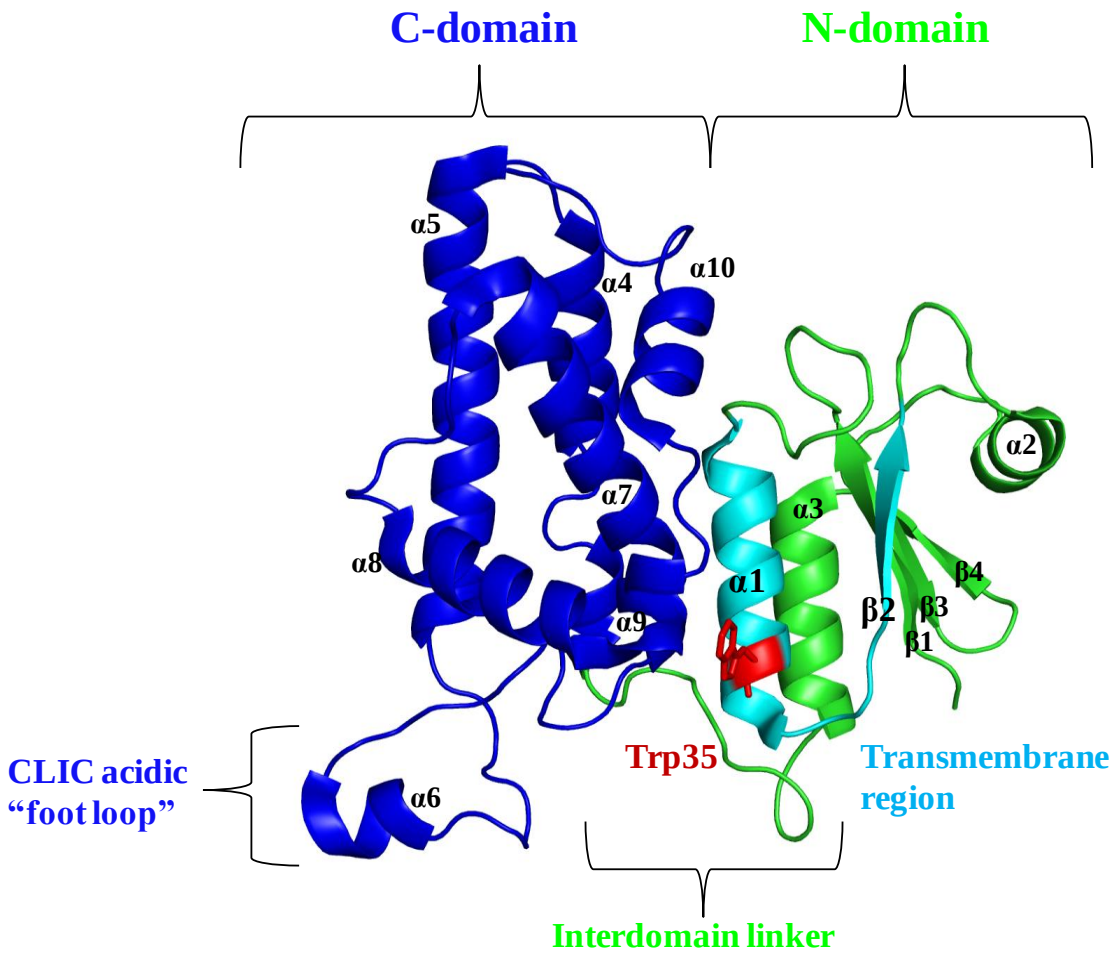


Figure 3. Structure of soluble monomeric CLIC1

The α-helical C-domain is in blue with the thioredoxin N-domain in green. The transmembrane region in the N-domain is shown in cyan. The single tryptophan residue, Trp35, is shown in red. Image created from PDB code 1K0M (Harrop *et al.*, 2001) using the PyMol Molecular Graphics System, version 1.5 Schrödinger, LLC.

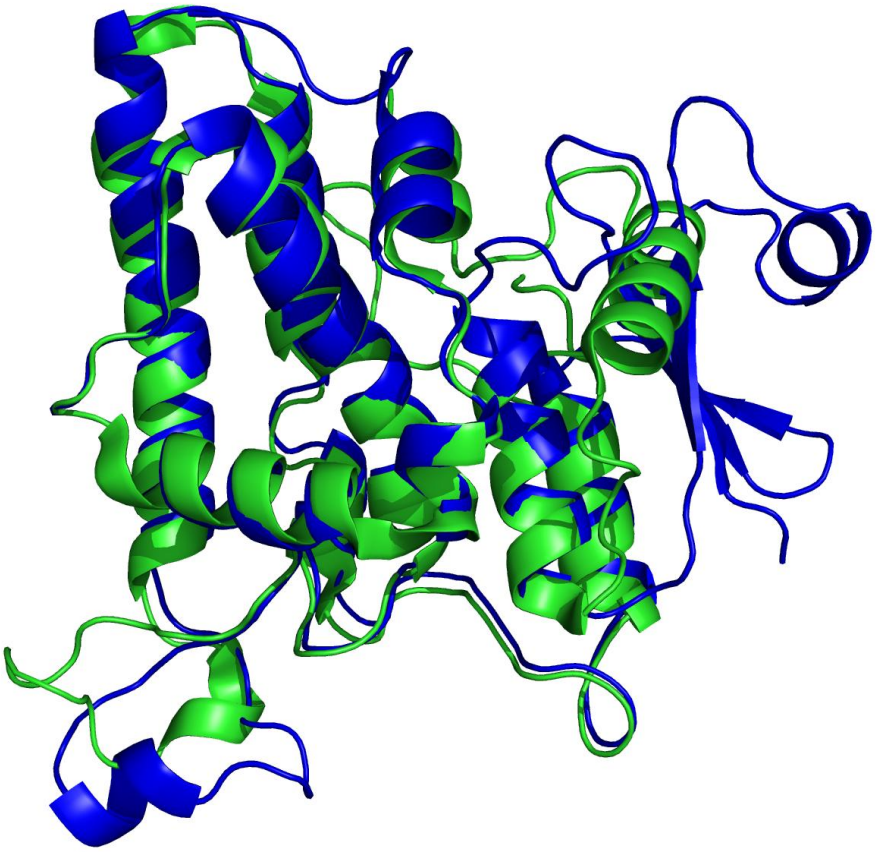


Figure 4. Comparison of the reduced CLIC1 monomer and a subunit of the oxidised CLIC1 homodimer

Monomeric CLIC1, formed under reducing conditions is shown in blue (PDB code 1K0M, Harrop *et al.*, 2001). Homodimeric CLIC1 forms under oxidative conditions. One of the two monomeric subunits is shown in green (PDB code 1RK4, Littler *et al.*, 2004). Upon oxidation, the thioredoxin N-domain of the monomer becomes all α-helical. The two structures overlay with a backbone rmsd of 7.7 Å.

The similarity of the CLICs and GSTs has led to the question of whether CLICs have enzymatic activity. As yet, none has been reported. However, CLICs are able to bind glutathione weakly in a putative active site that is similar to that of the GSTs in that it is found in the N-domain and centres on a cysteine residue (Cys24 in CLIC1) (Harrop *et al.,* 2001; Littler *et al.,* 2004). Cys24 is conserved in most of the human CLIC proteins and has been shown to be essential for ion channel activity (Littler *et al.,* 2004). CLIC1 contains a single tryptophan, Trp35, found at the domain interface (Figure 3), and forms part of the putative transmembrane region.

The transmembrane region (TMR) of CLIC1 is found in the N-domain from Cys24-Val46, spanning α-helix 1 and β-strand 2 (Berry and Hobert, 2006). This was additionally confirmed by the membrane association of a truncated form of the N-domain of CLIC4, which does not require the C-domain for binding (Singh and Ashley, 2007). The TMR region demonstrates sufficient hydrophobicity and α-helical propensity, as predicted with the AGADIR helix-coil algorithm (Lacroix *et al.,* 1998; Nathaniel, 2007; Stoychev *et al.,* 2009). FLAG-labelling (Tonini *et al.,* 2000), proteinase K digestion (Duncan *et al.,* 1997) and FRET-based modelling studies (Goodchild *et al.,* 2011) have also indicated that the N-domain inserts into the membrane such that Cys24 falls on the *trans* side. The presence of an additional KRR motif (residues 49 – 51) proposed to anchor the TMR in the membrane through electrostatic interactions with the anionic phospholipids (Valenzuela *et al.,* 1997) further supports that the TMR has been correctly identified.

1.5.2. CLIC1 membrane insertion

The structure of membrane-bound CLIC1 has yet to be solved. FRET modelling studies by Goodchild *et al.* (2010 and 2011) have indicated that the protein makes a single pass across the membrane and would need to oligomerise to form a functional channel. Indeed, the CLIC1 channel was initially proposed to be a tetramer (Warton *et al.,* 2002), though more recent work suggests the structure may be of higher order, potentially an octamer (Goodchild *et al.,* 2011). The work of Goodchild *et al.* (2010 and 2011) also indicated that large-scale structural rearrangements are necessary for CLIC1 to bind membranes. These rearrangements include separation of the two domains and refolding of the N-domain. These conformational changes are proposed to be facilitated by low pH (Tulk *et al.,* 2002; Warton *et al.,* 2002; Fanucchi *et al.,* 2008; Legg-E'Silva *et al.,* 2012).

1.5.3. The pH-response of CLIC1

As has been observed for the PFTs and Bcl proteins, low pH promotes membrane insertion and anion channel activity of CLIC1 (Tulk *et al.*, 2002; Warton *et al.*, 2002). The low pH (~ 5.5) at the surface of a membrane does not alter the structure of soluble CLIC1. Rather the conformational stability of the soluble protein is significantly reduced (Fanucchi *et al.*, 2008). Hydrogen-deuterium exchange mass spectrometry experiments have demonstrated that this destabilisation coincides with an increase in structural flexibility, specifically in the N-domain (Stoychev *et al.*, 2009). This apparent relaxation of the structure has been proposed to represent a lowering of the energy barrier to conformational change, effectively 'priming' the structure and leaving it more susceptible to undergo the structural rearrangements required for membrane binding (Fanucchi *et al.*, 2008; Stoychev *et al.*, 2009; Legg-E'Silva, 2012). Additionally, low pH induces the formation of a partially-unfolded intermediate state under mild denaturing conditions (Fanucchi *et al.*, 2008). This intermediate is unlikely to be a molten globule as it has well-defined tertiary structure. It does, however, seem to play a similar role to the molten globule states of other soluble-to-membrane proteins as it has been suggested to represent a 'pre-form' of the membrane-inserting structure (Fanucchi *et al.*, 2008; Adamson, 2009).

The changes to the stability of CLIC1 at low pH are proposed to be facilitated by a network of pH-sensitive residues that maintain the domain interface (Fanucchi *et al.*, 2008). Protonation of these residues at low pH would break the side chain H-bond and salt bridge interactions they form at pH 7.0 and thereby, decrease conformational stability and increase flexibility (Fanucchi *et al.*, 2008; Stoychev *et al.*, 2009). The theoretical side chain pK_a values of histidine and glutamate are 6.5 and 4.5, respectively (Lide, 1991; Srivastava *et al.*, 2007). Given that pK_a values vary with the local environment, the pK_a values of histidine and glutamate side chains may fall within the relevant cytosol-to-membrane pH range (~ 7.4 – 5.5). It is these residues that are the most likely to be involved in pH-induced changes to soluble CLIC1.

Recent work has demonstrated that two of the three histidine residues found in CLIC1 function as 'pH-sensors' and contribute to the decreased stability of the protein at pH 5.5 (Achilonu *et al.*, 2012). Substitution of His74 and His185 with alanine and phenylalanine, respectively, enabled the variant proteins to mimic pH 5.5 unfolding behaviour at pH 7.0 (Achilonu *et al.*, 2012). CLIC1 contains 21 glutamate residues in total (Table 1). Of these, only four are both conserved across the human CLIC family and involved in H-bonds or salt bridges. These are Glu81, Glu85, Glu218 and

Glu228 (Figures 2 and 5). Glu81 and its salt bridge partner Arg29 have been thoroughly investigated and found to be involved in the pH-response of soluble CLIC1 (Legg-E'Silva *et al.*, 2012). Glu218 forms a salt bridge with Arg216 in the C-domain (Figure 5). As yet, little is known about the functions of the C-domain and membrane association by a truncated CLIC4 variant has suggested that the N-domain is able to bind membranes in the absence of the C-domain (Singh and Ashley, 2007). The interaction involving Glu218 thus seems unlikely to contribute to separation of the two domains or refolding of the N-domain for membrane binding.

An interaction that is likely to be involved, however, is the H-bond formed between Glu228 and Ser16 (Figure 5). Glu228 is found in α-helix 10 of the C-domain, while Ser16 falls within the linker region between β-strand 1 and α-helix 1 in the N-domain. As this bond connects the two domains, it is likely to be broken in the membrane-inserted form of CLIC1 where the domains are proposed to be separated (Goodchild *et al.,* 2011). Glu85 in α-helix 3 forms a salt bridge with Lys37 of α-helix 1 of the transmembrane region (Figure 5). This bond possibly maintains the integrity of the N-domain in the soluble form of CLIC1. Additionally, the side chain of Glu85 forms an H-bond with the backbone amide group of Leu96. Leu96 is found in the interdomain linker and is conserved across the CLIC family and some GSTS (Figure 2, Stoychev, 2008). Thus Glu85 may serve a dual role, maintaining the integrity of the N-domain interface by connecting α1 and α3 and also anchoring the N-domain to the linker via Leu96. Thus, Glu228 and Glu85 are the focus of this study, given their possible respective roles in interdomain connectivity and maintaining the integrity of the N-domain component of the domain interface.

Table 1. Locations, interactions and conservation of the glutamate residues in human CLIC1

Glutamate residue number	Location		H-bonds or salt bridges	Conserved? (exceptions)	Conserved in ALL human CLICs
	Domain	Secondary structure			
9	N	β1	-	(CLIC6, CLIC5B)	
53	N	α2	-	NO	
72	N	β5	-	(CLIC3/5A/5B, dmCLIC, EXC-4)	
81	N	α3	R29: s-b	YES	x
82	N	α3	-	(dmCLIC, EXC-4)	x
85	N	α3	K37: s-b, L96: bb H-b	(dmCLIC, EXC-4)	x
102	C	α4	-	(EXC)	x
130	C	α6	K31: bb H-b	(CLIC3, dmCLIC, EXC-4)	
150	C	foot loop	E158: s-b	(CLIC3, dmCLIC, EXC-4)	
154	C	foot loop	S156: bb H-b	NO	
158	C	foot loop	E150: s-b	(CLIC3/5A/5B)	
160	C	foot loop	-	NO	
172	C	loop	-	(CLIC3/2, dmCLIC, EXC-4)	
201	C	loop	I199: bb H-b, R204: bb H-b to sc	NO	
217	C	α10	S221: bb H-b, A220: bb H-b	(CLIC3, EXC-4)	
218	C	α10	R216: s-b, T222: bb H-b	(dmCLIC, EXC-4)	x
227	C	α11	-	NO	
228	C	α11	S16: s-b, N225: bb H-b, A232: bb H-b	(dmCLIC, EXC-4)	x
230	C	α11	Y194: bb H-b to sc	(CLIC3, dmCLIC, EXC-4)	
234	C	loop	L231: bb H-b	NO	

Glutamate residues and their bonds in the human CLICs, dmCLIC1 (*D. Melanogaster)* and EXC-4 (*C. Elegans*) were investigated with Swiss-PDB Viewer (Guex and Peitsch, 1997). Secondary structure is indicated as α-helical (α) or β-stranded (β) and numbered according to position in the sequence. H-bonds and salt bridges are indicated with the bonding partner first, followed by the type of bond. Abbreviations: bb – backbone, H-b - hydrogen bond, s-b – salt bridge, sc – side chain.

Conservation of residues was determined using T-coffee v 8.93 (http://www.ebi.ac.uk/Tools/services/rest/tcoffee). Where a residue is partially conserved, the exceptions are listed. Otherwise, conservation across all the human CLICs and CLIC-like proteins is listed as "YES" and where no overlap is seen, conservation is listed as "NO". Residues that are both conserved across the human CLICs and involved in H-bond or salt bridge interactions are highlighted with rectangles.

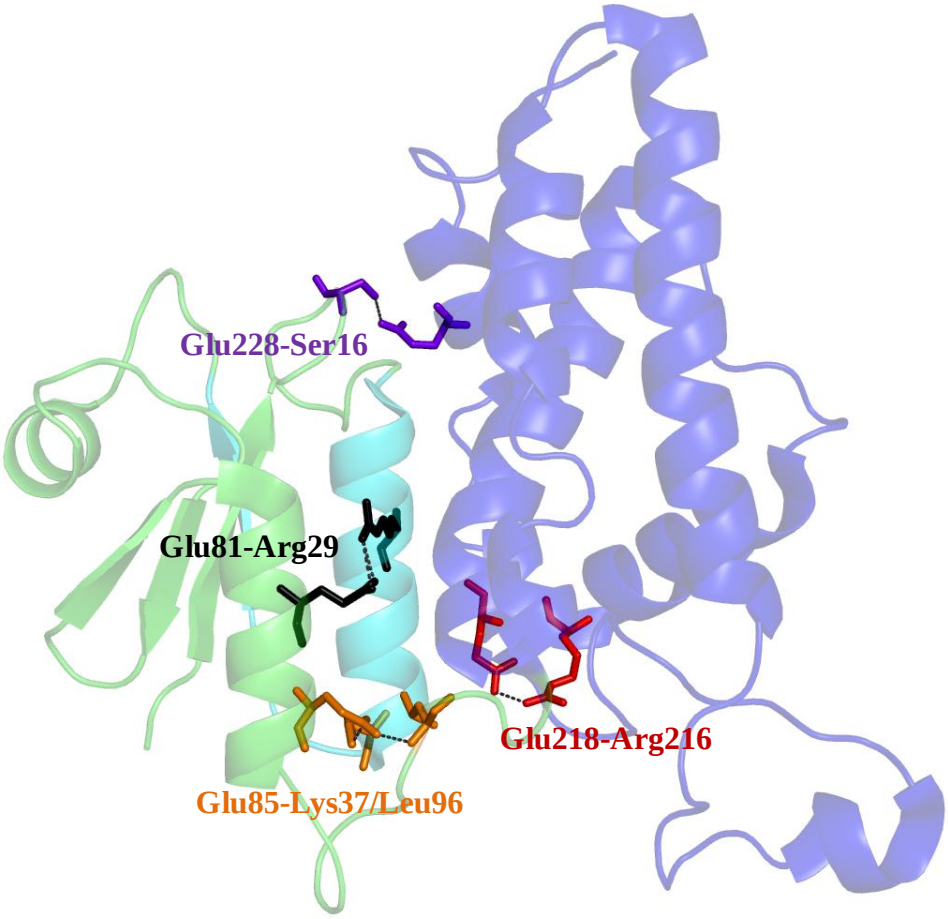


Figure 5. Conserved glutamate residues that form H-bonds or salt bridges in soluble CLIC1

The α-helical C-domain is in blue with the thioredoxin N-domain in green. The transmembrane region is in the N-domain and is shown in cyan. Glu228 and Ser16 form a H-bond and are in purple. Glu81 and Arg29 are shown in black and form a salt bridge. Glu218 and Arg216 are shown in red and form a salt bridge. Glu85 and its salt bridge partner Lys37 (N-domain), as well as its H-bond partner Leu96 (linker) are shown in orange. Image created from PDB code 1K0M (Harrop *et al.*, 2001) using the PyMol Molecular Graphics System, version 1.5 Schrödinger, LLC.

1.6. Objectives

Low pH is suggested to render soluble CLIC1 more susceptible to conformational change, and thereby possibly promote its association with membranes. This is proposed to be through protonation of conserved glutamate and histidine residues. Protonation potentially breaks their interactions and thereby reduces the conformational stability and rigidity of soluble CLIC1. For CLIC1 to assume its membrane-bound form, the N- and C-domains must separate and the N-domain must refold. This study focuses on glutamate residues, which form interactions that are likely to be broken in the structure of membrane-bound CLIC1 or in the soluble structure at a pH of 5.5 as the protein would encounter at a membrane surface. Glu228 was selected for its potential role in maintaining interdomain connectivity and Glu85 was chosen for its proposed function in anchoring the TMR and N-domain in the soluble form of CLIC1. The objective of this study was to investigate the potential roles of Glu228 and Glu85 in pH-induced changes to the structure and stability of soluble CLIC1.

To establish whether Glu228 and Glu85 act as 'pH-sensors', variant CLIC1 proteins were created in which glutamate was replaced with leucine. While both methionine and leucine occupy similar volumes to glutamate, they differ in polarity and the use of a less hydrophobic methionine would theoretically alter the polarity of the substitution site to a lesser degree. Recent work on an E81M CLIC1 variant, however, demonstrated the formation of water-mediated bonds between Met81 and Arg29, effectively preventing the total removal of interactions at the salt bridge being investigated (Legg-E'Silva, 2011). Leucine was thus chosen for its hydrophobicity which would potentially exclude water and prevent formation of similar 'artifical' interactions. The leucine side chain is also pH-insensitive and does not form H-bonds. The E228L CLIC1 and E85L CLIC1 variants would therefore lack the bonds present in the wild type protein and any differences in the behaviour of the proteins could be attributed to the breakage of the bonds in which the glutamate residues are involved. Thus, this study aimed to characterise the structures and conformational stabilities of the CLIC1 variants at the two pHs likely to be encountered by CLIC1 *in vivo* - pH 7.0 and pH 5.5 – and compare them to the wild type protein. Structure was examined using far-UV circular dichroism spectroscopy, fluorescence spectroscopy, limited thermolysin proteolysis and X-ray crystallography. Conformational stability was analysed by thermal and urea-induced equilibrium unfolding.

CHAPTER 2. MATERIALS AND METHODS

2.1. Materials

Wild-type CLIC1 is encoded as an N-terminal glutathione-S-transferase fusion protein (GST-CLIC1) by an insert cloned into the pGEX-4T-1 vector, renamed pGEX-CLIC1. The resulting CLIC1 is linked to the GST tag by a thrombin cleavage site (LVPR$^{\downarrow}$GS). Following cleavage, the C-terminal glycine and serine residues of the cleavage site remain at the N-terminus of CLIC1. The plasmid was a gift from Samuel N. Breit, St Vincent's Hospital, University of New South Wales, Australia. Non-standard laboratory items and chemicals were ordered as listed in Table 2.

2.2 Methods

2.2.1 Site-directed mutagenesis by PCR

Glycerol stocks of *E. coli* XL-1 Blue cells transformed with the pGEX-CLIC1 plasmid encoding wild-type CLIC1 were inoculated into 25 ml 2x YT broth (1.6% (w/v) tryptone, 0.5% (w/v) yeast extract, 10% (w/v) NaCl) containing 100 μg/ml ampicillin. As the pGEX-4T-1 plasmid encodes a β-lactamase which confers ampicillin resistance, the antibiotic is used to select for cells carrying the plasmid of interest. Cultures were grown overnight (16 hours) at 37 °C with shaking at 230 rpm. Plasmid was isolated using the Fermentas GeneJET™ Plasmid MiniPrep kit which is based on the alkaline-lysis technique. Briefly 1.5 – 3.0 ml of cells were harvested by centrifugation and lysed using NaOH and SDS, which serve to increase the pH and denature both chromosomal DNA and proteins. Subsequent addition of potassium acetate neutralises the pH, allowing plasmid DNA to reanneal while larger chromosomal DNA and proteins are precipitated. Plasmid DNA was then bound to a silica matrix, washed and eluted with autoclaved, deionised water. Isolated plasmid was quantified using a NanoDrop ND-1000.

The polymerase chain reaction (PCR) uses temperature cycling to control repeated rounds of DNA replication. Site-directed mutagenesis by PCR uses primers complimentary to the regions of cDNA flanking the nucleotides to be mutated, but containing the desired new sequence at their centres. As the primers themselves are used as the templates for the regions with which they base pair, the mutation is incorporated as the PCR reaction proceeds. To create the E228L and E85L variants, the corresponding GAG codons encoding glutamate had to be mutated to TTG and CTG, respectively, both of which encode leucine.

Table 2. Suppliers used to obtain non-standard laboratory materials

Material	Source	Location
Mutagenesis primers	Inqaba Biotec	Pretoria, South Africa
GeneJET™ Plasmid MiniPrep kit	Fermentas	Ontario, Canada
QuikChange Lightning™ Site-Directed Mutagenesis kit	Fermentas	Ontario, Canada
Escherichia coli BL21(DE3) pLysS cells	Stratagene	La Jolla, USA
Escherichia coli XL120 Gold cells	Fermentas	Ontario, Canada
isopropyl β-D-thiogalactopyranoside (IPTG)	Sigma-Aldrich	St Louis, Missouri, USA
Unstained protein molecular weight marker	Fermentas	Ontario, Canada
dithiothreitol (DTT)	Melford Laboratories Ltd.	Suffolk, UK
reduced glutathione (GSH)	Sigma-Aldrich	St Louis, Missouri, USA
1-anilino-8-napthalene sulphonate (ANS)	Sigma-Aldrich	St Louis, Missouri, USA
ultrapure urea	Merck	Darmstadt, Germany
Amicon Ultra-15 10K centrifugal concentration devices	Merck	Darmstadt, Germany
thermolysin from *Bacillus thermoproteolyticus*	Sigma-Aldrich	St Louis, Missouri, USA
thrombin from bovine plasma	Sigma-Aldrich	St Louis, Missouri, USA
Crystallization Extension Kit for Proteins	Sigma-Aldrich	St Louis, Missouri, USA
24-well VDX crystallisation plates and slides	Hampton Research	Aliso Viejo, California, USA
Hampton Index kit for protein crystallisation	Hampton Research	Aliso Viejo, California, USA
IZIT dye	Hampton Research	Aliso Viejo, California, USA

Nucleotide primers for E228L and E85L were designed using PrimerX (http://www.bioinformatics.org/primerx/index.htm) and synthesised by Inqaba Biotec (Pretoria, South Africa). Mutant codons are underlined.

E228L primers:

Forward: 5'- CTGTCCAGATGATGAGTTGATCGAGCTCGCCTATG - 3'

Reverse: 5'- CATAGGCGAGCTCGATCAACTCATCATCTGGACAG - 3'

E85L primers:

Forward: 5' - CAAGATTGAGGAATTTCTGCTGGCAGTGCTGTGCCCTCCCAG - 3'

Reverse: 5' - CTGGGAGGGCACAGCACTGCCAGCAGAAATTCCTCAATCTTG - 3'

The lyophilised forward and reverse primers were dissolved in autoclaved deioinised water to a concentration of 100 µM and diluted to working concentrations of 125 ng/µl each. A Fermentas QuikChange Lightning™ Site-Directed Mutagenesis kit was used. PCR components were mixed according to the kit instructions, using 100 ng/µl template DNA (pGEX-CLIC1). Following PCR, wild-type template DNA was removed by the addition of 2 µl of *DpnI* and incubation at 37 °C for 5 minutes. *DpnI* is a methylation sensitive endonuclease that cleaves methylated DNA. As the template plasmid was isolated from bacterial cells, it is methylated, and hence cleaved while mutant plasmid (pGEX-CLIC1-E228L/E85L) produced by PCR is not.

The resulting pGEX-CLIC1-E228L/E85L plasmids were propagated by transformation of 45 µl aliquots of ultracompetent *E.coli* XL10 Gold cells (Stratagene, La Jolla, USA) with 2 µl of the appropriate PCR reaction using heat shock. Although the exact mechanism of heat shock is unclear, it is believed that the increased temperature puts the cells under stress, resulting in the creation of micropores in the plasma membrane through which plasmid is taken up (Panja *et al.*, 2008). The cells were incubated with plasmid on ice for 30 minutes, heat-shocked at 42 °C for 43 s and returned to ice for a further 2 minutes. Thereafter, 900 µl of SOC media (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 10% (w/v) NaCl and 0.02 M glucose) was added and the cells were incubated at 37 °C for 60 minutes to allow cellular growth and stabilisation. Cells were then spread on amp^+-LB agar plates and incubated at 37 °C overnight. Colonies were selected and cultured overnight at 37 °C in 25 ml 2x YT broth. The modified plasmids were then isolated using the Fermentas GeneJET™ Plasmid MiniPrep kit and sent to Inqaba Biotec (Pretoria, South Africa) for nucleotide sequencing, which confirmed successful modification of the pGEX-CLIC1 plasmid to pGEX-CLIC1-E228L and pGEX-CLIC1-E85L, encoding the E228L and E85L variant CLIC1 proteins, respectively.

2.2.2. Induction trials and protein production

E. coli BL21(DE3) pLysS expression cells were transformed separately with the pGEX-CLIC1-E228L/E85L plasmids using heat shock (see procedure detailed above). In the pGEX-CLIC1 plasmid, expression of the CLIC1 gene is under the control of a *tac* promoter. Overexpression may be induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG), which displaces the *tac* repressor (De Boer *et al.,* 1983). The temperature, time and [IPTG] for in optimal production of wild-type CLIC1 are well established: addition of 1 mM IPTG followed by incubation for 5 - 6 hours at 37 °C (Valenzuela *et al.,* 1997; Fanucchi *et al.,* 2008). To verify whether these conditions are similarly appropriate for E228L CLIC1 and E85L CLIC1 production, brief induction trials were carried out, testing expression at 20 °C and 37 °C. Overnight liquid cultures of transformed *E. coli* BL21(DE3) pLysS expression cells were grown at 37 °C and used to inoculate 25 ml fresh 2xYT broth ($^1/_{50}$ dilution). The cultures were grown at 37 °C until the OD_{600} = 0.5. At this point, cultures for overexpression at 20 °C were cooled at 4 °C until the mid-log growth phase (OD_{600} = 0.6) was reached. Overexpression was then induced by the addition of 1 M IPTG to a final concentration of 1 mM IPTG. Cultures were incubated at either 20 °C or 37 °C with shaking at 230 rpm. After either 6 hours or 16 hours (overnight), 1 ml samples were removed, centrifuged briefly and prepared for SDS-PAGE (see Section 2.2.4) as follows: pelleted cells were resuspended in 50 µl STE buffer (10 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA and 1 mM DTT with 10 µg/ml lysozyme, 10 mM $MgCl_2$ and 10 µg/ml DNase), sonicated and centrifuged to separate the soluble and insoluble fractions of the lysate. The pellet was resuspended in 50 µl of STE buffer. An additional 50 µl of SDS-PAGE reducing sample buffer (Section 2.2.4.) was added to both the pellet and supernatant. The samples were boiled for 5 minutes at 95 °C and analysed by SDS-PAGE. The ideal temperature and time for the production of both E228L CLIC1 and E85L CLIC1 was found to be 20 °C with overnight incubation.

The determined optimum overexpression conditions were therafter applied to larger volumes of culture to produce wild-type, E228L CLIC1 and E85L CLIC1 in quantities sufficient for purification and subsequent analyses. For each protein, overnight cultures of *E. coli* BL21(DE3) pLysS expression cells were grown at 37 °C and used to inoculate 4 L of fresh 2x YT broth containing 100 µg/ml ampicillin ($^1/_{50}$ dilution). The cultures were grown at 37 °C until OD_{600} = 0.6 and overexpression was induced with the addition of IPTG to a final concentration of 1 mM. In the case of wild-type CLIC1, the cultures were then incubated at 37 °C for 6 hours. E228L CLIC1 and E85L CLIC1 expression, in contrast, required incubation at 20 °C for 16 hours. After appropriate

incubation time, cells were harvested by centrifugation (20 minutes, 4221 x g, 10 °C). Cells were resuspended in 10 ml STE buffer per 1 L culture and stored at -20 °C.

2.2.3 Purification of wild-type, E228L CLIC1 and E85L CLIC1

Cells were thawed at 4 °C on a rotator and lysed by sonication on ice (three 1 minute pulses separated by 1 minute incubations on ice) (Microson XL-2000 sonicator). The soluble and insoluble cell fractions were separated by centrifugation (27 000 x g, 30 minutes, 10 °C).

The soluble fraction was added to a 25 ml GSH-agarose column pre-equilibrated with equilibration buffer (10 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 0.02% (w/v) NaN_3 and 1 mM DTT pH 8.0). GSH-agarose contains a glutathione (GSH) molecule ligated to agarose beads via an oxirane group (Amersham Biosciences GST Gene Fusion Handbook, 2002, *www.gelifesciences.com/aptrix/upp00919.nsf/.../18115758.pdf*). The affinity of the GST-tag for GSH ensures that the GST-CLIC1 fusion protein binds to the column while contaminants are eluted. After binding, the column was washed with 1 L of 1 M NaCl and 2 L of equilibration buffer. The column was then equilibrated with thrombin cleavage buffer (20 mM Tris-HCl, 150 mM NaCl, 2.5 mM $CaCl_2$ pH 8.4). The GST and CLIC1 moieties of the fusion protein are linked by a thrombin cleavage sequence (LVPR$^{\downarrow}$GS). To separate the two proteins, 20 units of bovine thrombin per 1 L of culture were added to the column, which was incubated on a rotator at 20 °C for 16 hours to permit cleavage. DEAE equilibration buffer (20 mM Tris-HCl, 0.02% (w/v) NaN_3, pH 6.5) was then passed through the column to thoroughly remove CLIC1 and thrombin.

Anion exchange chromatography uses a cationic resin to capture negatively charged macromolecules and was used to separate CLIC1 from thrombin. The mixture of CLIC1 and thrombin was applied to a diethylaminoethyl (DEAE) Sepharose column pre-equilibrated with DEAE equilibration buffer. CLIC1 and thrombin have theoretical p*I*s of 5.0 and 8.0, respectively (estimated with the ExPaSy ProtPARAM tool (Gasteiger *et al.*, 2005)). In the equilibration buffer (pH 6.5), CLIC1 is negatively charged and hence, is able to bind to the positively charged DEAE resin. Thrombin, by contrast, has a net positive charge at this pH, would not bind and was eluted from the column upon further addition of DEAE equilibration buffer. CLIC1 was eluted with buffer containing 300 mM NaCl. The salt outcompetes the protein for interactions with the resin due to the negative charge of the Cl^- ion. SDS-PAGE (Section 2.4.4.) was used to assess the purity of the recovered CLIC1. The buffer was exchanged by dialysis. Unless otherwise specified, all structural

and stability work was conducted in 50 mM sodium phosphate, 0.02% (w/v) NaN_3 with 1 mM DTT at either pH 7.0 or pH 5.5.

When protein was purified for crystallography, an additional purification step was included to ensure purity. Size exclusion chromatography utilises a porous resin to separate proteins according to their hydrodynamic radii (Porath and Flodin, 1959). The pores of the resin are sized to allow smaller molecules to pass through the resin itself. In contrast, larger species move between resin particles and thus, pass through a smaller volume. They are hence eluted first and thereby, separated from their smaller molecular counterparts (Lathe and Ruthven, 1956). Size exclusion chromatography was used to separate CLIC1 from any remaining contaminants, which may have been present at concentrations unlikely to be detected by SDS-PAGE and hence, unlikely to affect spectroscopic work. These contaminants could, however, prevent crystallisation by reducing the monodispersity of the stock protein solution. Following anion exchange, CLIC1 was exchanged into crystallisation buffer (0.1 M Tris-HCl pH 6.5, 0.02% (w/v) NaN_3 with 5 mM DTT, in MilliQ deionised water) by dialysis. The protein was concentrated by ultrafiltration to a volume of ~ 3 ml using an Amicon Ultra-15 10K centrifugal concentration device (Merck, Darmstadt, Germany). A 180 ml Sephacryl S-200 size exclusion column connected to an *ÄKTAprime*™ plus system (GE Healthcare, Pittsburgh, USA) was equilibrated with crystallisation buffer and the concentrated CLIC1 was loaded onto the column at a rate of 1 ml/min. The purity of the recovered CLIC1 protein was verified by SDS-PAGE (Section 2.4.4.) before it was used in crystallisation trials.

2.2.4. Sodium dodecyl sulphate polyacrylamide gel electrophoresis

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is an electrophoresis technique utilising the migration of charged molecules in an electric field as a tool to determine both size (molecular mass) and purity of protein samples. Sodium dodecyl sulphate (SDS) is an anionic detergent able to denature proteins and bind to them in a ratio of ~ 1.4 g SDS per gram of protein. SDS treatment gives protein molecules a uniform charge: mass ratio such that they migrate according to their molecular masses when electrophoresed through a cross-linked acrylamide matrix (Boyer, 2006).

SDS-PAGE was performed according to the method developed by Laemmli (1970). All SDS-PAGE analyses were conducted using a 15 % separating gel (15% (w/v) acrylamide, 1.35% (w/v) bis-acrylamide, 0.25 M Tris-HCl pH 8.8, 0.1% (w/v) SDS, 0.05% (w/v) ammonium persulphate, 0.1%

(v/v) TEMED) and a 4 % stacking gel (4% (w/v) acrylamide, 0.36% (w/v) bis-acrylamide, 0.05 M Tris-HCl pH 6.8, 0.01% (w/v) SDS, 0.05% (w/v) ammonium persulphate, 0.2% (v/v) TEMED). Samples for analysis were diluted with an equal volume of reducing sample buffer (125 mM Tris-HCl pH 6.5, 4% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) β-mercaptoethanol and 3.5 μg/ml bromophenol blue) and boiled at 95 °C for five minutes. Unstained protein molecular weight marker (Fermentas, Ontario, Canada) was used and contains several proteins of known molecular mass: β-galactosidase (116.2 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), lactate dehydrogenase (35 kDa), REase Bsp981 (25 kDa), β-lactoglobulin (18.4 kDa) and lysozyme (14.4 kDa). Electrophoretic separation of protein samples was carried out at 160 V for approximately 1 hour, after which gels were stained with Coomassie Brilliant Blue R-250 for a minimum of 2 hours. Gels were destained with a 1:4:5 mixture of acetic acid, distilled water and methanol.

2.2.5. Protein concentration determination

The concentration of purified protein was calculated using a dilution series and the Beer-Lambert Law:

$$A = \varepsilon c l \quad (4)$$

Where A is the absorbance (arbitrary units), c is the concentration (M), l is the pathlength (cm) and ε is the molar extinction coefficient (M^{-1} cm^{-1}) of the molecule of interest at a particular wavelength (for protein λ = 280 nm). The ε_{280} used for CLIC1 in this study was calculated by taking into account the number of absorbing chromophores present in the structure and their extinction coefficients (Perkins, 1986):

$$\begin{aligned} \varepsilon_{280} &= 5550\ \Sigma\text{tryptophan} + 1340\ \Sigma\text{tyrosine} + 150\ \Sigma\text{cysteine} \quad (5) \\ &= 5550\ (1) + 1340\ (8) + 150\ (6) \\ &= 17\ 170\ M^{-1}\ cm^{-1} \end{aligned}$$

To calculate protein concentration, a series of dilutions was set up (ranging from ¼ - $^1/_{10}$) and the absorbance of each was measured at 280 nm, 260 nm and 340 nm using a Jasco V-630 UV-Vis spectrophotometer. The use of 260 nm and 340 nm allows for the detection of DNA contamination – if any – and protein aggregates, respectively. The measured absorbance of protein was corrected for both buffer and aggregate contributions using the following formula:

$$\text{Corrected } A_{280\text{protein}} = A_{280\text{protein}} - A_{280\text{buffer}} - [(A_{340\ \text{protein}} - A_{340\text{buffer}})(A_{280\text{buffer}}/A_{340\text{buffer}})] \quad (6)$$

(McIntyre, 2007).

2.2.6. Structural characterisation

2.2.6.1. Far-UV circular dichroism spectroscopy

Circular dichroism (CD) is the phenomenon by which left and right circularly polarised light is differentially absorbed by a chiral medium (Yang, 1996; Kelly *et al.*, 2005). In the far-UV range (190 - 250 nm), proteins produce CD spectra characteristic of their secondary structure. This occurs as a result of charge transitions between π-orbitals of the C, O and N atoms in amide bonds that form the polypeptide backbone. As these are strongly dependent on backbone geometry (φ, ψ and ω angles), they directly reflect the secondary structure, particularly in the case of α-helices where geometry is generally repetitive (Whitmore and Wallace, 2008). In predominantly α-helical proteins, CD spectra show troughs at 222 nm and 208 nm (a n → π* transition) and a peak around 190 nm (π → π* transition). In contrast, β-sheet structures show a minimum around 218 nm (n → π*) and a maximum at 195 nm (π → π*) (Bell, 1981).

Far-UV CD spectra were recorded for 2 μM wild-type, E228L CLIC1 and E85L CLIC1 at both pH 7.0 and pH 5.5 (5 mM sodium phosphate, 0.02% (w/v) NaN_3, 1 mM DTT) using a Jasco J-810 spectropolarimeter. A cell of 2 mm path length was used, with band width, data pitch, and response time set to 0.5 nm, 0.2 nm, and 0.5 s, respectively. All spectra were measured as an average of 10 accumulative scans and the resultant CD signal in millidegrees was converted to the more commonly used mean residue ellipticity ([Θ] ($deg.cm^2.dmol^{-1}$)) using the equation:

$$[\Theta] = \frac{100.\theta}{c.n.l} \qquad (7)$$

Where θ is the measured ellipticity (mdeg) at a particular wavelength, c is the sample concentration (mM), n is the number of residues in the protein and l is the pathlength (cm) of the cuvette. For all far-UV analyses: n = 243, l = 0.2 cm and c = 0.001 mM, 0.002 mM or 0.005 mM. All spectra were recorded in triplicate, corrected for buffer ellipticity, smoothed in SigmaPlot® v11.0 (SPSS Science, Chicago, Illinois USA) using the negative exponential method and averaged to give a single spectrum for each protein.

2.2.6.2. Fluorescence spectroscopy

The conjugated ring structures of the aromatic amino acids allow them to be excited by light of a particular wavelength and release the absorbed energy in the form of fluorescence at a longer

wavelength (a phenomenon known as Stokes shift) (Lakowicz, 1999). In the case of tryptophan which is maximally excited with light of 295 nm, the wavelength of emission is highly dependent on the polarity of the local environment of the indole side chain. This allows tryptophan fluorescence to report on packing interactions and thereby, local tertiary structure (Bell, 1981). While tyrosine (excitable at 280 nm) lacks the environmental sensitivity of tryptophan, it is of interest as its emission spectrum overlaps with the excitation spectrum of tryptophan. When the tyrosine residues of a protein that also contains tryptophan are excited at 280 nm, emission occurs at the tryptophan emission wavelength. This is due to Förster resonance energy transfer (FRET), where energy emitted by tyrosine is absorbed by, and excites, tryptophan due to the overlap between the emission and excitation spectra of the two fluorophores. The excitation of tyrosine residues is particularly useful in proteins such as CLIC1 which contain only a single tryptophan and give a weak fluorescence signal when excited at 295 nm. Excitation of CLIC1's eight tyrosine residues at 280 nm produces a stronger signal that remains largely dependent on the Trp35 environment, due to FRET and hence remains useful in monitoring local tertiary packing interactions. Trp35 is located at the CLIC1 domain interface and within the putative transmembrane region (TMR) (Harrop *et al.,* 2001). As separation of the domains and refolding of the TMR are believed to precede membrane insertion, fluorescence of Trp35 is an excellent probe for monitoring changes that may enable the soluble structure to assume a membrane-competent state.

To detect changes in packing about Trp35 potentially caused by the E228L or E85L substitutions, wild type and variant CLIC1 proteins were excited at 280 nm at pH 7.0 and pH 5.5. Emission spectra from 280 – 450 nm were collected at 20 °C using a Perkin Elmer LS50B Luminescence Spectrometer. Slit widths of 5 nm were used with a scanning speed of 300 nm/min. Protein concentration was set at 2 μM. All spectra were recorded in triplicate, each as an average of 5 accumulations, and subsequently corrected for buffer fluorescence, smoothed and averaged to give a single spectrum.

The solvent exposure of tryptophan residues in a protein structure can be examined through dynamic quenching of tryptophan fluorescence. Acrylamide is a dynamic quencher and decreases fluorescence intensity by absorbing the energy of the fluorophore's excited state through non-radiative energy transfer (Lackowitz, 1999). Dynamic quenching thus relies upon an interaction between the quencher (acrylamide) and the excited state of the fluorophore (Möller and Denicola, 2002). For this to happen, the quencher must diffuse through the structure and thus, the relative

accessibility of the tryptophan residue to solvent is reflected in the decrease of fluorescence intensity upon quencher addition. This can be represented by a modified Stern-Volmer Plot:

$$F_0/F = 1 + K_{SV}[Q] \quad (8)$$

where F_0 and F are the peak fluorescence intensities in the absence and presence of quencher, respectively, [Q] is the concentration of quencher and K_{SV} is the Stern-Volmer quenching constant. A Stern-Volmer plot under conditions of dynamic quenching is a linear function whose gradient is equal to K_{SV} (Lackowitz, 1999; Möller and Denicola, 2002). From this, the degree of solvent exposure can be estimated as a greater K_{SV} value indicates a more exposed fluorophore.

The Trp35 solvent accessibilities in the native and intermediate states of the three CLIC1 proteins at pH 7.0 and pH 5.5 were determined using acrylamide quenching. A 1 M acrylamide stock was added to triplicate solutions of 2 μM protein to yield a range of 10 quencher concentrations from 0.0 - 0.5 M. Following a 30 minute incubation period, Trp35 was excited at 295 nm with slit widths of 8 nm. Accumulative emission spectra were measured in triplicate from 295 - 450 nm with a scan speed of 400 nm/min. The fluorescence intensity at 345 nm for each quencher concentration was used to generate Stern-Volmer plots and the K_{SV} values of the wild type and the two mutant CLIC1 proteins were determined.

The extrinsic, amphipathic flurophore, 8-anilino-1-napthalene sulphonate (ANS) was included in structural studies as it binds hydrophobic patches exposed on the surface of a protein. Like tryptophan, the wavelength of its fluorescence depends largely on the polarity of its environment. ANS excited at 390 nm typically fluoresces at approximately 520 nm in a polar environment. However, when found in a more non-polar environment, bound to exposed hydrophobic patches on a protein, for example, its emission is blue shifted according to the hydrophobicity of its proteinaceous environment (Slavik, 1982). Accompanying changes in fluorescence intensity may also occur and offer information on the number of ANS binding sites (Rosen and Weber, 1969). For this reason, ANS is a particularly useful probe for the detection of partially-unfolded intermediate states, which have solvent-exposed non-polar residues.

ANS was used to detect hydrophobic patches exposed by the CLIC1 proteins during unfolding from a native state to an unfolded state (Section 2.2.8.3.). ANS was prepared in the appropriate analysis buffer and its concentration was confirmed using a serial dilution with ANS absorbance at 350 nm (ε_{350} = 4950 $M^{-1}cm^{-1}$, Weber and Young, 1964). ANS was added in molar excess (200 μM) to triplicate samples of 2 μM CLIC1 at pH 7.0 and pH 5.5 which were then incubated at 20 °C for 1

hour to allow binding and equilibration. The samples were excited at 390 nm and 3 cumulative ANS fluorescence spectra were measured for each from 390 - 550 nm with slit widths of 5 nm and a scanning speed of 400 nm/min. To correct for free ANS, additional samples containing only buffer, ANS and, if appropriate, urea were measured and subtracted. Aggregation was monitored by Rayleigh scatter following excitation at 340 nm with slit widths of 5 nm.

2.2.6.3. pH-dependence of secondary and local tertiary structure

As pH plays a critical role in CLIC1 function, the structures of wild-type CLIC1, E228L CLIC1 and E85L CLIC1 were examined across a pH range from 4.0 – 8.0 in 0.5 unit increments. As CLIC1 is a predominantly α-helical protein, the ellipticity at 222 nm (θ_{222}) was used to monitor secondary structure. Similarly, intrinsic tryptophan fluorescence at 345 nm (F345), following excitation at 280 nm, was used to monitor local tertiary structure.

CLIC1 was diluted to a final concentration of 2 μM in either 17 mM sodium phosphate (for pH 6.0 – 8.0) or 50 mM sodium acetate (for pH 4.0 - 5.5), both containing 1 mM DTT, and allowed to equilibrate. Buffers were chosen for their ability to buffer at the required pHs and concentrations were adjusted such that ionic strength remained constant. An equilibration time of 20 s was determined to be sufficient by monitoring the Rayleigh scatter of fluorescence excitation at 280 nm and emission intensity at 345 nm over time. Following equilibration, the θ_{222} and F345 nm were measured in triplicate for a period of 20 s and each was averaged to give a single value. All data were corrected for buffer, triplicates were averaged and ellipticity was converted into mean residue ellipticity ($[\Theta]_{222}$) as described in Section 2.2.6.1.

2.2.7. X-ray crystallography

2.2.7.1. Crystallisation of E228L CLIC1

Protein crystallisation involves the controlled precipitation of protein from a supersaturated solution such that protein molecules form an ordered array of 3-dimensional regularly repeating units: a crystal. Protein solubility is dependent upon many factors, most prominently pH, ionic strength, temperature and the concentrations of both protein and added precipitants (McPherson, 1990). These factors must be carefully balanced in the solution in which crystals are grown - the

crystallisation cocktail - to ensure the growth of crystals that diffract X-rays and produce a diffraction pattern of sufficient quality for structure determination (Bernal and Crowfoot, 1934).

For crystal trials, E228L CLIC1 was prepared at a stock concentration of 10 mg/ml. The hanging drop method of crystallisation was chosen for crystal growth as it simplifies crystal harvesting. Hanging drop involves suspension of drops of a protein/precipitant mixture on a glass cover slide above a well containing precipitant. Wells are sealed to allow for equilibrium to be established between the drops and precipitant solution in the well. Ideally, water diffuses from the protein droplets to the well solution, resulting in an increase in the droplet protein concentration to the point that precipitates -ideally crystalline in nature- begin to form (McPherson, 1990). Crystallisation trials were set up using solutions from a Hampton Research Index kit (Hampton Research, Aliso Viejo, California, USA), which cover a range of crystallisation conditions previously successful for the crystallisation of a variety of proteins. Trials were set up in 24-well VDX plates (Hampton Research, Aliso Viejo, California, USA). Five hundred microlitres of crystallisation cocktail was added to wells and 2, 3 and 4 µl of stock protein was added to an equal volume of crystallisation cocktail on a glass coverslip. The coverslips were flipped and placed over the wells and the plates were incubated at 20 °C.

Initial crystallisation trials identified two conditions suitable for E228L CLIC1: (A) 0.2 M sodium formate, 20% (w/v) polyethylene glycol (PEG) 3350 and (B) 0.2 M sodium chloride, 0.1 M HEPES pH 7.5, 25% (w/v) PEG 3350. Further refinement of these conditions by changing to the protein (7.5, 5.0 and 2.5 mg/ml) and precipitant (20, 17.5 and 25% (w/v)) concentrations revealed that lower protein concentrations reduced twinning and produced larger individual crystals. E228L CLIC1 crystals used for data collection were grown with 5 mg/ml protein in 0.2 M sodium chloride, 0.1 M HEPES pH 7.5, 25% (w/v) PEG 3350 at 20°C within 48 hours of trial setup. The proteinaceous nature of the crystals was confirmed by reducing SDS-PAGE of E228L CLIC1 crystals on a 15% gel as well as by the addition of IZIT dye (Hampton Research, Aliso Viejo, California, USA). IZIT penetrates protein crystals via the large solvent channels they contain, turning them dark blue. IZIT cannot penetrate more densely packed salt crystals and their colour remains unchanged.

Crystal trials were repeated as described above to identify conditions suitable for E85L CLIC1 crystallisation. The full Hampton Index kit was used to set up initial trials with 10 mg/ml protein. This concentration was both increased to 20 mg/ml and lowered to 2.5 mg/ml and the trials were

repeated. The Crystallization Extension Kit for Proteins (Sigma-Aldrich, St Louis, Missouri, USA) was then used to set up additional trials using the recommended protein concentration of 10 mg/ml.

2.2.7.2. Data collection and phasing

Individual E228L crystals were captured using a cryoloop, briefly soaked in the crystallisation cocktail and flash cooled in liquid nitrogen. The crystallisation cocktail contains 25% PEG 3350, which functions as a cryoprotectant and prevents freezing of the solvent component of the crystal, which could cause the crystal to crack. Flash cooling is necessary as it preserves the structural integrity of the crystal and minimises the radiation damage caused by X-ray exposure (Henderson 1990; Garman, 1999). For this reason, the crystal was kept at 113 K in a stream of N_2 gas for the duration of data collection.

X-ray diffraction, data collection and indexing were performed by Dr. Manuel Fernandes, University of the Witwatersrand, Johannesburg, South Africa. Diffraction data were collected with a Bruker X8 Proteum system with a Microstar copper rotating-anode generator, and an Oxford Cryostream Plus cooling system. Detection was with a PLATINUM 135 CCD detector with Montel 200 optics. Data were processed using the SAINT and APEX programs (Bruker AXS Inc, Wisconsin, USA), which were used to assign h, k and l coordinates to the reflections. For further processing the data were imported into the CCP4 Program Suite v. 6.3.0 using Combat (Collaborative Computational Project Number 4, 1994). Reflection intensities were converted to structure factor amplitudes using Truncate (Collaborative Computational Project Number 4, 1994) which also marked 5% of the reflections for later R_{free} calculations. CCP4 Matthews_coef (Collaborative Computational Project Number 4, 1994) was used to estimate the number of molecules per asymmetric unit.

Initial phases were determined by molecular replacement (MR). In this process, a model structure, structurally similar to the protein structure to be solved, is placed in the target unit cell following rotational and translational optimisation. Phases calculated for this structural model are then combined with experimental structure factor amplitudes of the target (Rossman, 1990). Molecular replacement was performed using Phaser (McCoy *et al.,* 2007) with the wild type CLIC1 structure (PDB code 1K0M, Harrop *et al.,* 2001) as the phasing model.

2.2.7.3. Refinement and validation

Refinement of the structure serves to produce the highest quality model possible that complies with both experimental diffraction data as well as geometric and spatial Parameters known for proteins. Refinement was performed in successive rounds using Refmac5 (Murshudov *et al.*, 1997) and COOT (Emsley and Cowtan, 2004) for reciprocal refinement and real space optimisation, respectively. The progression of refinement was tracked by decreases in the *R*-factor and R_{free}. The *R*-factor describes the agreement between the observed structure factor amplitudes and a set calculated for the structural model. The R_{free} is calculated by the same formula as *R*-factor except that it uses 5% of reflections not used in refinement. The R_{free} is thus largely an objective parameter in assessing the quality of the structural model (Brünger, 1992). The backbone geometry of the model was assessed using Ramachandran plots, which graphically analyse the ψ and φ torsion angles of each residue. Residues with torsion angles outside of the accepted limits may indicate improbable backbone geometries (Ramakrishnan and Ramachandran, 1965). The fit of the modelled residues within the electron density and probability of side chain rotamers were checked using validation tools included within COOT. Poor electron density resulting from low resolution prevented the localisation of solvent molecules. Model validation was performed using MolProbity, which detects improbable bond lengths and angles as well as Ramachandran deviations and steric clashes (Davis *et al*., 2007; Chen *et al*., 2010). When the model was of sufficient quality, it was submitted to the RCSB Protein Data Bank (http://www.pdb.org/pdb/home/home.do) and has been made available under the accession code 4IQA.

2.2.8. Conformational stability

The conformational stability of a protein represents its energetic ability to remain in the native state and is described by the difference in the Gibbs free energies of the native and unfolded states (ΔG). Environmental changes such as increased temperature or addition of a chemical denaturant may shift the equilibrium between the folded and unfolded states towards the unfolded state (O'Sullivan and Tompson, 1890). Determining the equilibrium constant enables both the Gibbs free energy of folding as well as the *m*-value to be calculated. The latter describes the dependence of the free energy of unfolding on the denaturant concentration (Santoro and Bolen, 1988; Shortle, 1995). Unfolding studies are useful in comparing the relative stabilities of wild-type and variant proteins and thereby, in determining the effect of point mutations on native structural integrity.

2.2.8.1. Thermal unfolding

Protein thermostability has been attributed to a wide range of factors (Vogt *et al.*, 1997; Vogt and Argos, 1997; Kumar *et al*., 2000), though H-bonds, salt bridges as well as van der Waals packing of the protein's hydrophobic core appear to dominate. For Variant proteins that have lost a stabilising feature, as is true for both E228L (Glu228-Ser16 H-bond) and E85L (Glu85-Lys37 salt bridge) CLIC1, thermal unfolding provides a relatively quick and simple measure of whether the removal of these interactions has impacted significantly on the thermal stability of the variant proteins.

To examine the thermal stability of wild-type, E228L CLIC1 and E85L CLIC1 at both pH 7.0 and pH 5.5, 2 μM protein was unfolded by increasing the temperature in a reaction vessel from 20 °C to 80 °C at a rate of 1 °C/min using a Jasco PTC-4235/L Peltier type temperature control system. Unfolding was monitored using far-UV circular dichroism to measure ellipticity at 222 nm (θ_{222}). Measurements were performed with a data pitch of 0.5 °C, 2 nm bandwidth, response of 0.5 s, and standard sensitivity. Once 80 °C was reached, a delay of 60 s was imposed before the temperature was decreased to 20 °C at the same rate to allow refolding, if possible. No refolding was possible – a likely consequence of aggregation induced at high temperatures. Unfolding data were fitted to a 2-state sigmoid model using SigmaPlot® v11.0 (SPSS Science, Chicago, Illinois USA). Fits were then normalised and plotted as the fraction of unfolded protein versus temperature. As thermodynamic parameters could not be calculated, fits were generated for qualitative analysis of the data.

2.2.8.2. Reversibility studies

Before urea-induced unfolding could be performed, the amount of time required for each of the variant CLIC1 proteins to be unfolded and to reach equilibrium was determined. This serves to minimise aggregation that may result from prolonged exposure of the partially-unfolded states to urea. The equilibration time for wild type CLIC1 was overnight (16 hours) as described by Legg-E'Silva (2011). Variant protein equilibration times were determined by monitoring the intrinsic tryptophan fluorescence (at 345 nm) and far-UV CD (at 222 nm) following manual mixing of 2 μM CLIC1 with 8 M urea, with a dead time of 20 s. Samples of native folded (0 M urea) and fully unfolded (8 M urea, incubated overnight) CLIC1 at the same concentration were used to establish baseline values for the two conformational states.

To calculate thermodynamic Δ*G* and *m*-values from unfolding data, two criteria must be met: the sample must be at equilibrium between the folded and unfolded states, and the reversibility of unfolding along a single pathway must have been established (Pace, 1986). Wild-type CLIC1 has previously been shown to refold fully following chemical denaturation with urea at both pH 7.0 and pH 5.5 (McIntyre, 2007). To investigate whether the same is true for the E228L and E85L CLIC1 variants, 10 μM protein was unfolded at both pH 7.0 and pH 5.5 in 8 M urea for 30 minutes. The unfolded protein was then diluted 10 x with buffer containing decreasing concentrations of urea (from 0.8 – 8.0 M), and allowed to refold for 2 hours. As a control, 1 μM protein samples were unfolded using urea from 0.8 – 8.0 M for 30 minutes. The ellipticity at 222 nm and fluorescence at 345 nm (excitation at 280 nm) of each sample were then measured for a period of 20 s, averaged and plotted as a function of denaturant concentration. As an additional probe to detect the presence of partially folded intermediates, 200 μM ANS was added to both unfolding and refolding samples and ANS spectra were recorded as described in Section 2.2.6.2.

2.2.8.3. Urea-induced equilibrium unfolding

Urea for unfolding was prepared as a 10 M stock in the appropriate analysis buffer using the solvent-urea mass ratios described by Pace (1986). Urea concentration was confirmed by measurement of the refractive index of the denaturant and calculation using the equation:

$$[\text{urea}] = 117.66(\Delta N) + 29.753(\Delta N)^2 + 185.56(\Delta N)^3 \quad (9)$$

ΔN is the difference in the refractive indices of the prepared denaturant and buffer.

Urea was stored at -20 °C and discarded if not used within a week to avoid potential modification of the protein amino groups by cyanate and ammonium ions that result from the decomposition of urea over time (Pace, 1986).

To determine the stability of wild-type CLIC1, E228L CLIC1 and E85L CLIC1 against chemical denaturants, equilibrium unfolding was performed at pH 7.0 and pH 5.5 in triplicate using a range of 41 urea concentrations between 0 M and 8 M. Concentrated urea stock was mixed with appropriate volumes of protein and buffer such that the final protein concentration was 2 μM. The solutions were incubated at 20 ºC to allow equilibrium to be reached (overnight for the wild type; 30 minutes for E228L CLIC1 and E85L CLIC1). The intrinsic fluorescence spectrum and far-UV CD at 222 nm of each sample were then recorded with the settings outlined in Sections 2.2.6.1 and 2.2.6.2, respectively. The exposure of non-polar residues upon unfolding was detected by ANS

binding (Section 2.2.6.2). To monitor the formation of aggregates, samples were additionally excited at 340 nm and Rayleigh scatter determined at the same wavelength with slit widths of 2.5 nm. Fluorescence spectra were used to calculate the intensity averaged emission wavelength (IAEW) using the equation:

$$IAEW = \Sigma\lambda_i.F_i. / \Sigma F_i \quad (10)$$

where λ_i is the wavelength and F_i is the fluorescence intensity.

2.2.8.4. Thermolysin proteolysis

Thermolysin (EC 3.4.24.27) is a thermostable metalloendoproteinase (Ohta *et al.*, 1966) able to cleave peptide bonds on the N-terminal side of bulky hydrophobic residues (Leu, Phe, Ile, Val in the P1' position) (Morihara *et al.*, 1968). It has recently been used in an equilibrium unfolding technique known as pulse proteolysis. The duration of the cleavage 'pulse' is set such that only unfolded protein will be cleaved (Park and Marqusee, 2005) whereupon the reaction is quenched and the protein solution is analysed by SDS-PAGE. The intensity of the native protein band can be analysed by densitometry. As the protein unfolds, exposing more cleavage sites, a greater amount of protein is cleaved, leading to a corresponding decrease in the intensity of the native band. This decrease can be plotted as a function of denaturant concentration and serves as an additional probe for unfolding (Na and Park, 2008).

Wild-type, E228L CLIC1 and E85L CLIC1 were unfolded at a concentration of 10 μM at both pH 7.0 (20 mM Tris-HCl, 50 mM NaCl, 10 mM $CaCl_2$ with 1 mM DTT) and pH 5.5 (16 mM sodium acetate, 50 mM NaCl, 10 mM $CaCl_2$ with 1 mM DTT) using a range of urea concentrations between 0.0 and 6.5 M in a final volume of 100 μl. After equilibration at 20 °C, 2 μl of 10 mg/ml thermolysin was added and cleavage was allowed to proceed for 60 s. To quench the reaction, 50 μl of the reaction mixture was transferred to a fresh tube containing 5 μl of 50 mM Na_2.EDTA. An equal volume of reducing sample buffer was added to the quenched reaction and the samples were boiled for analysis by SDS-PAGE (Section 2.2.4). As a control, buffer was added in place of thermolysin to protein in the native state (0 M urea).

Wheras pulse proteolysis is useful in probing unfolding, limited proteolysis with thermolysin may be used to probe the inherent flexibility of protein structures. Previous work has shown a correlation between the structural sites at which limited proteolysis occurs and their flexibility (reviewed in Fontana *et al.*, 2004). This correlation is due to the inherent ability of more flexible regions to

undergo the conformational changes necessary for binding to the active site of the proteinase (Schechter & Berger, 1967; Hubbard *et al.*, 1994). A larger number of flexible or partially unfolded regions in a protein thus increases the likelihood of cleavage. By varying the time of exposure to the proteinase, more or less protein will be cleaved according to how dynamic and flexible the structure is.

Limited proteolysis was performed to assess the relative flexibilities of the native and intermediate states of the three CLIC1 proteins at both pH 7.0 and pH 5.5. For the native states, no urea was added. For the intermediate states urea was added at concentrations described in Section 2.2.8.3. Proteolysis was performed as described above with the exception of the cleavage times which were varied: 20 s, 40 s, 60 s, 2 minutes, 5 minutes, 10 minutes, 20 minutes, 40 minutes and 60 minutes. A control was prepared as described above.

Following SDS-PAGE and staining with Coomassie Brilliant Blue R-250, gels from both pulse and limited proteolysis were destained and analysed densitometrically analysis using the LabWorks Analysis Software v.4.5 (UVP LLC., Upland, USA).

CHAPTER 3. RESULTS

3.1. Creation and purification

Sanger sequencing performed by Inqaba Biotec (Pretoria, South Africa) (Figure 6) confirmed mutation of the codons for glutamate at positions 228 and 85 to codons for leucine. Optimal conditions for soluble production of the GST-E228L-CLIC1 and GST-E85L-CLIC1 fusion proteins in *E. coli* BL21(DE3) pLysS cells were found to be induction with 1 mM IPTG and incubation at 20 °C for 16 hours (Figure 7).

The variant proteins and wild type CLIC1, were successfully isolated from expression cells and purity was confirmed by SDS-PAGE (Figure 8). Wild-type, E228L CLIC1 and E85L CLIC1 migrate at slightly higher molecular weights than the 26.9 kDa predicted for CLIC1. Wild-type and E85L CLIC1 migrate at a calculated molecular weight of 31 kDa, and E228L CLIC1 migrates at 29 kDa. CLIC1 is an acidic protein that carries an overall negative charge above pH 5. This makes the protein less likely to bind anionic SDS in a charge:mass ratio consistent with other less negative proteins and may explain the anomalous migration of the CLIC1 proteins. A difference in the migration of E228L CLIC1 is also evident. The E228L CLIC1 sequence does not contain thrombin cleavage sites (ExPaSy PeptideCutter http://web.expasy.org/cgi-in/peptide_cutter/peptidecutter.pl) and it is thus unlikely that migration is altered by truncation of the variant by thrombin. E228L CLIC1 has lost a relatively solvent-exposed, negatively-charged glutamate side chain compared to the buried glutamate lost by E85L CLIC1. As such, it may be able to bind a greater quantity of SDS than the wild type and hence, is able to migrate at a calculated molecular weight closer to that predicted for CLIC1.

3.2. Structures of wild-type, E228L CLIC1 and E85L CLIC1

3.2.1. Secondary structure

The secondary structures of the three CLIC1 proteins were examined using far-UV circular dichroism at pH 7.0 and pH 5.5. The spectra of all three proteins at both pHs show troughs at 222 nm and 208 nm (Figure 9), characteristic of α-helical proteins. At pH 7.0, the spectra peak at 190 nm as expected. At pH 5.5, the peak is slightly red-shifted, possibly due to noise introduced by the buffer at lower pH. The 208 nm trough is slightly deeper than that at 222 nm, reflecting the presence of the β-sheet structure found in the thioredoxin-like N-domain (Venyaminov and Yang, 1996). As the four spectra overlay with minimal difference between them, replacing Glu228 and Glu85 with leucine does not result in any significant changes to the secondary structure of CLIC1.

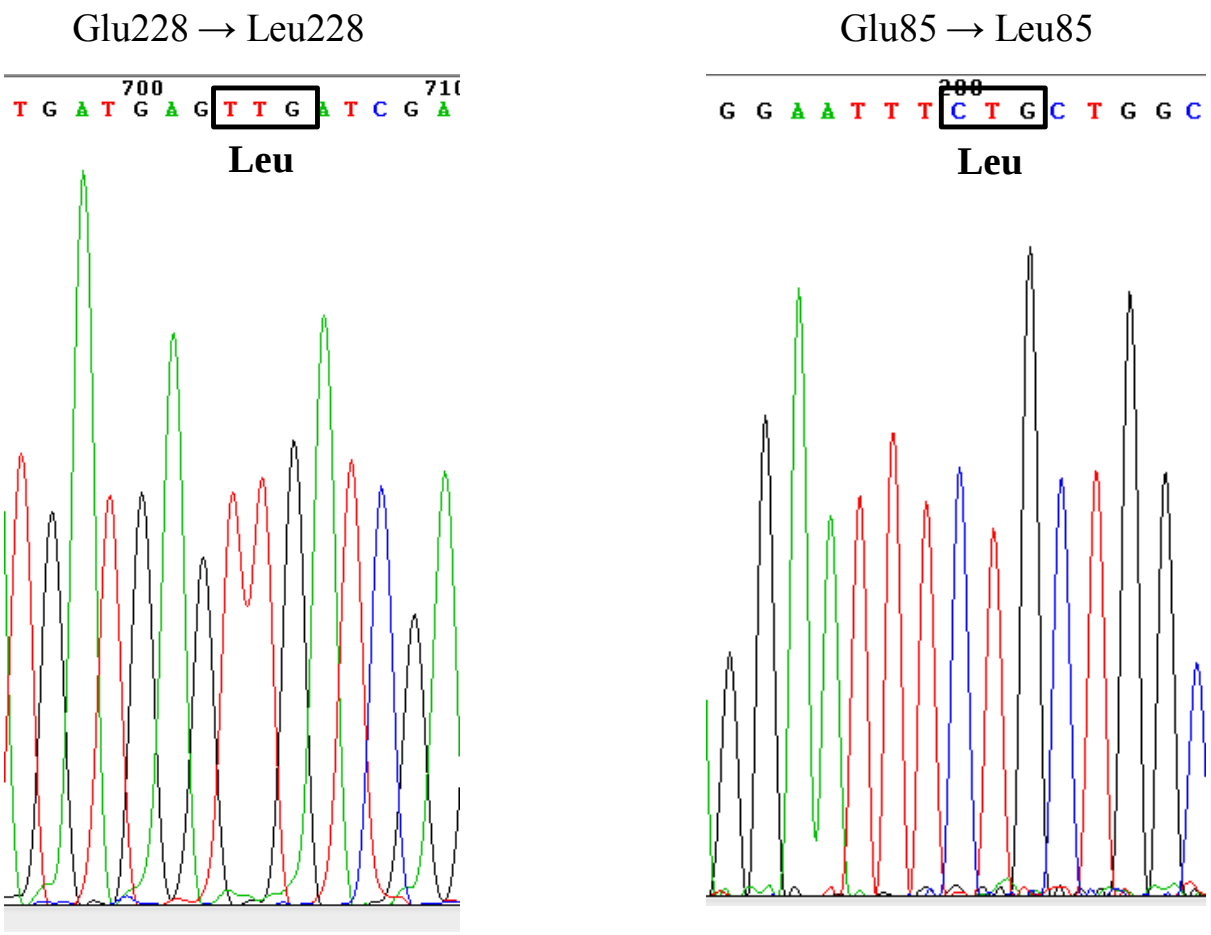


Figure 6. Successful generation of the E228L and E85L mutations

Sequencing data provided by Inqaba Biotec (Pretoria, South Africa) confirmed successful mutation of the GAG codons for glutamates 228 and 85 to a TTG codon encoding leucine 228 and a CTG codon for leucine 85. Full length sequencing data was analysed using Chromas Lite version 2.01 (http://www.technelysium.com.au/chromas_lite.html, Technelysium Pty. Ltd., Helensvale, Australia). The relevant stretches of sequence containing the mutations are shown.

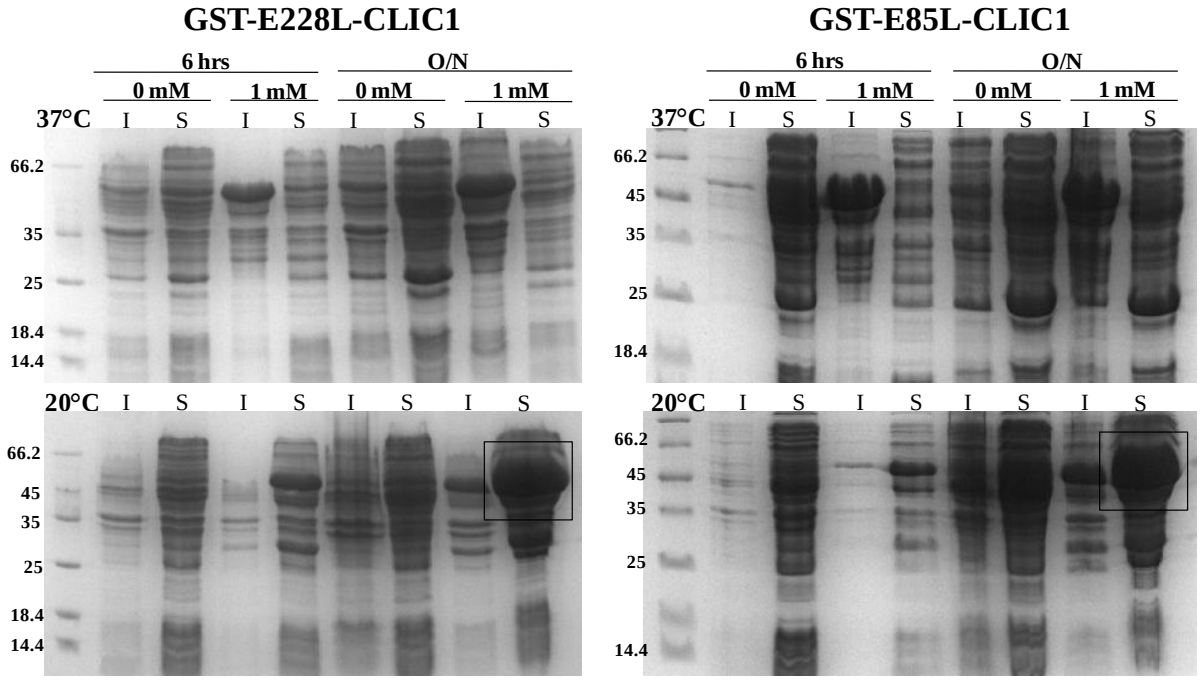


Figure 7. Production of the GST-E228L-CLIC1 and GST-E85L-CLIC1 fusion proteins in *E. coli* BL21(DE3) pLysS cells

15% SDS-PAGE gels depicting production of the 53 kDa variant CLIC1 fusion proteins at different temperatures (37 °C or 20 °C) and for different incubation periods (6 hours or 16 hours overnight (O/N)). Expression was induced with 1 mM IPTG.No IPTG (0 mM) was added to control cultures. The insoluble (I) and soluble (S) cell fractions for each condition are shown. Optimual conditions for expression are 1 mM IPTG and overnight incubation at 20 °C for both proteins, where the greatest amount of the GST-E228L-CLIC1 and GST-E85L-CLIC1 fusion proteins is present in the soluble fraction (indicated above with blocks).

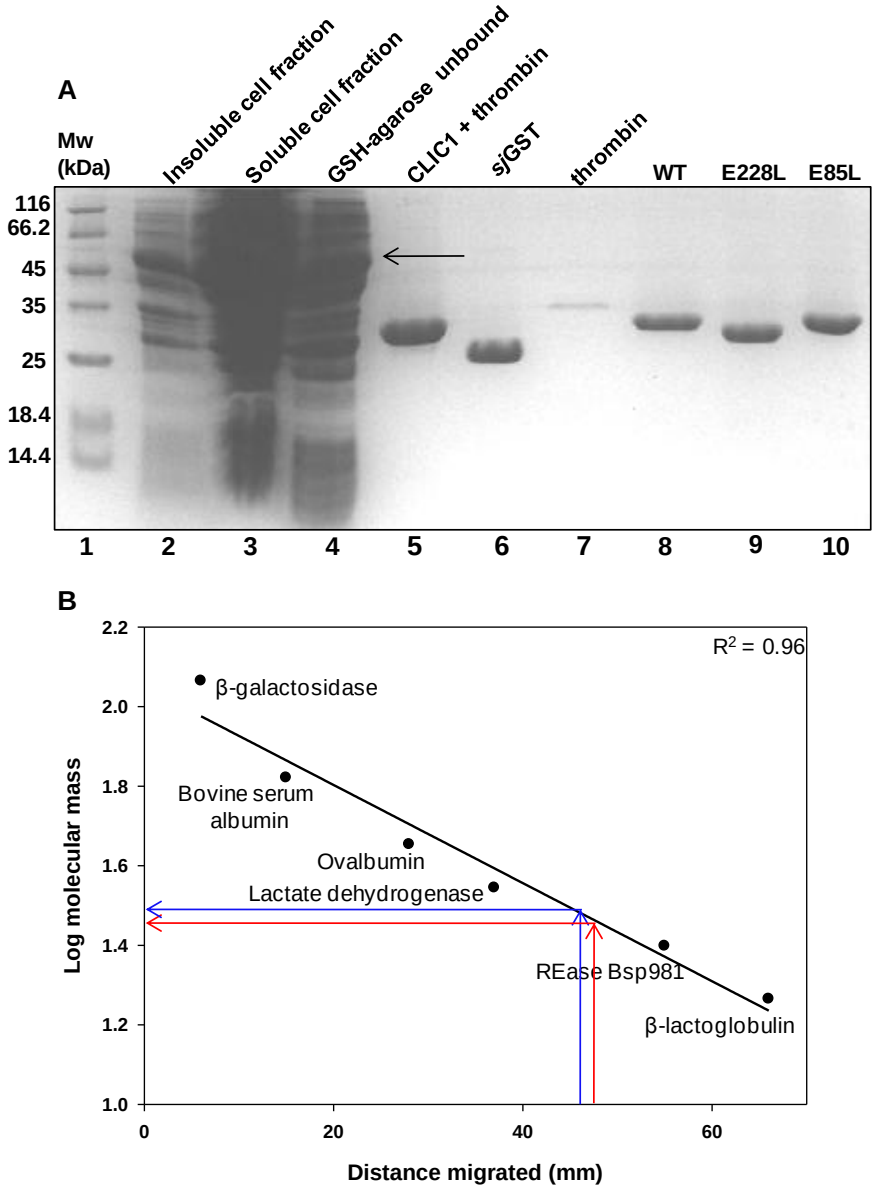


Figure 8. Purification of wild-type, E228L CLIC1 and E85L CLIC1

(A) SDS-PAGE analysis of CLIC1 purification. Following overexpression in *E.coli* BL21(DE3) pLysS cells, the insoluble (Lane 2) and soluble (Lane 3) cell fractions were separated and the soluble fraction containing the GST-CLIC1 fusion protein (indicated with an arrow) was applied to a GSH-agarose column. Unbound protein was collected and contained less GST-CLIC1 fusion protein (Lane 4). Following thrombin cleavage, CLIC1 and thrombin were collected from the column (Lane 5). The GST tag (Lane 6) was subsequently eluted with 10 mM reduced GSH. The CLIC1 and thrombin mixture was separated using a DEAE anion exchange column. Thrombin did not bind (Lane 7) and CLIC1 was eluted with 300 mM NaCl. Lanes 8, 9 and 10 show purified wild-type, E228L CLIC1 and E85L CLIC1, respectively. (B) Estimation of CLIC1 molecular weight. Migration of proteins in the molecular weight marker were used to construct a standard curve based on their molecular masses and the experimental migration distances. The arrows indicate the migration distances of wild-type and E85L CLIC1 (red) and E228L CLIC1 (blue). Molecular weights were calculated to be ~ 31 kDa for wild-type and E85L CLIC1, and ~ 29 kDa for E228L CLIC1.

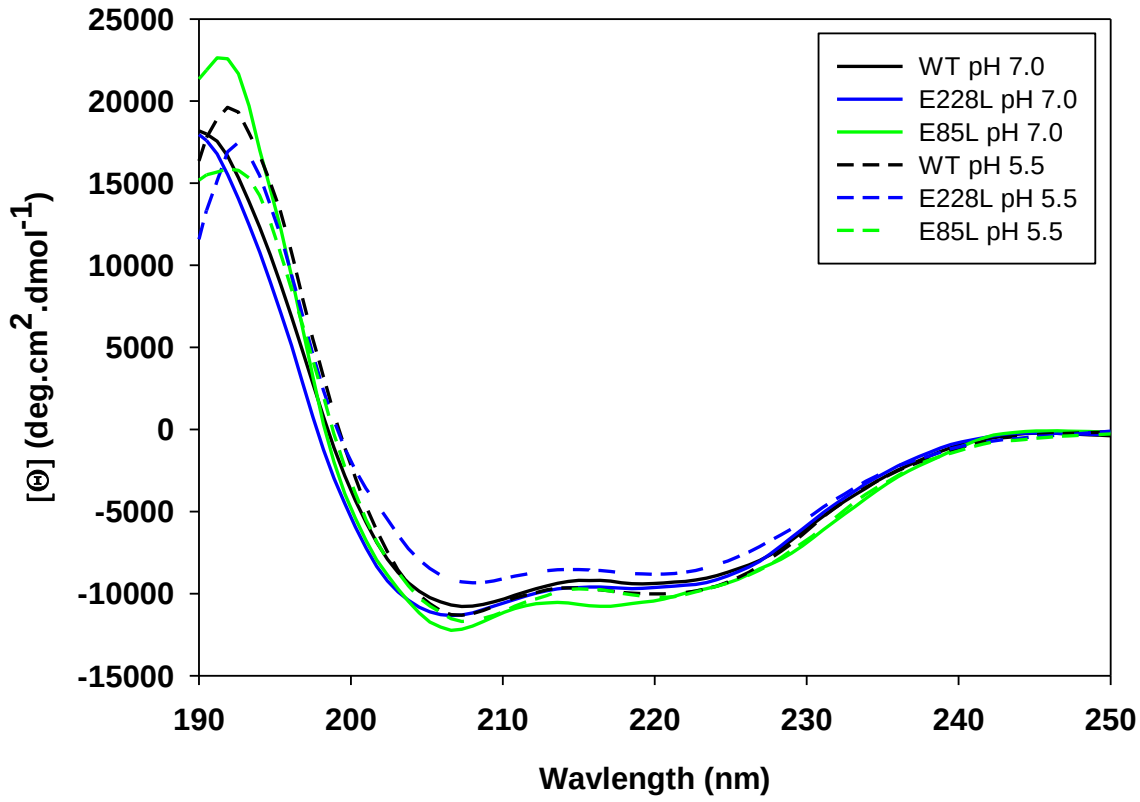


Figure 9. Secondary structures of wild-type, E228L CLIC1 and E85L CLIC1 at pH 7.0 and pH 5.5

The secondary structures of wild-type, E228L CLIC1 and E85L CLIC1 were analysed using far-UV circular dichroism spectra of 2 μM protein in 5 mM sodium phosphate, 0.02% NaN_3 and 1 mM DTT at pH 7.0 and pH 5.5. Spectra were measured in triplicate, corrected for buffer, converted from millidegrees to mean residue ellipticity ([Θ]), averaged and smoothed. All spectra show minima at 222 nm and 208 nm indicating predominantly α-helical structures that do not vary with pH.

3.2.2. Tertiary structure

The packing of side chains around Trp35 in wild-type, E228L CLIC1 and E85L CLIC1 at pH 7.0 and pH 5.5 was examined by intrinsic fluorescence spectroscopy. Spectra (Figure 10) peak between 341 - 344 nm, very near to the 345 nm peak reported in previous CLIC1 work (McIntyre, 2007; Legg-E'Silva, 2011). The fluorescence intensities of the spectra differ. This is most likely due to the highly sensitive nature of fluorescence, as minor discrepancies between replicates may result in large differences in fluorescence intensity. Additionally, intensity tends to vary between experiments conducted on different days. As the spectrum of each CLIC1 protein was measured immediately following purification, the different intensities observed cannot be interpreted in terms of structural changes. Unlike intensity, the maximum emission wavelength (λ_{max}) of fluorescence varies only with environmental polarity changes around tryptophan. It is hence unlikely to vary between experiments and is a more consistent reporter of any changes induced in the variant CLIC1 structures. The λ_{max} values obtained for wild-type, E228L CLIC1 and E85L CLIC1 (Figure R10) do not differ significantly between the proteins or at different pHs. Thus, the local tertiary structure at Trp35 has not been altered in the variants and does not change significantly with pH.

3.2.3. pH-Dependence of structure

The secondary and local tertiary structure of the CLIC1 proteins was monitored between pH pH 4 and 8, by monitoring the mean residue ellipticity at 222 nm ($[\Theta_{222}]$), fluorescence intensity at 345 nm (F345) and fluorescence λ_{max} at each pH. Aggregation was monitored using Rayleigh scatter at 340 nm (F340). Neither secondary nor tertiary structure is altered with pH, as the $[\Theta]_{222}$ and F345 values obtained for all three proteins overlay and remain constant over the pH range tested (Figure 11). While the λ_{max} values of E228L CLIC1 show an increase at lower pHs, a corresponding rise in the F340 is seen (Figure 11), suggesting that the observed difference is likely due to aggregation.

3.2.4. Conformational flexibility

To investigate changes to the flexibility of CLIC1 induced by replacing the Glu228 and Glu85 residues with a leucine, limited proteolysis was performed on the native states of the three CLIC1 proteins at both pH 7.0 and pH 5.5. Due to minor differences in the staining intensity of the individual gels, the measured band intensities are arbitrary and are only comparable to other intensities measured for the same experiment.

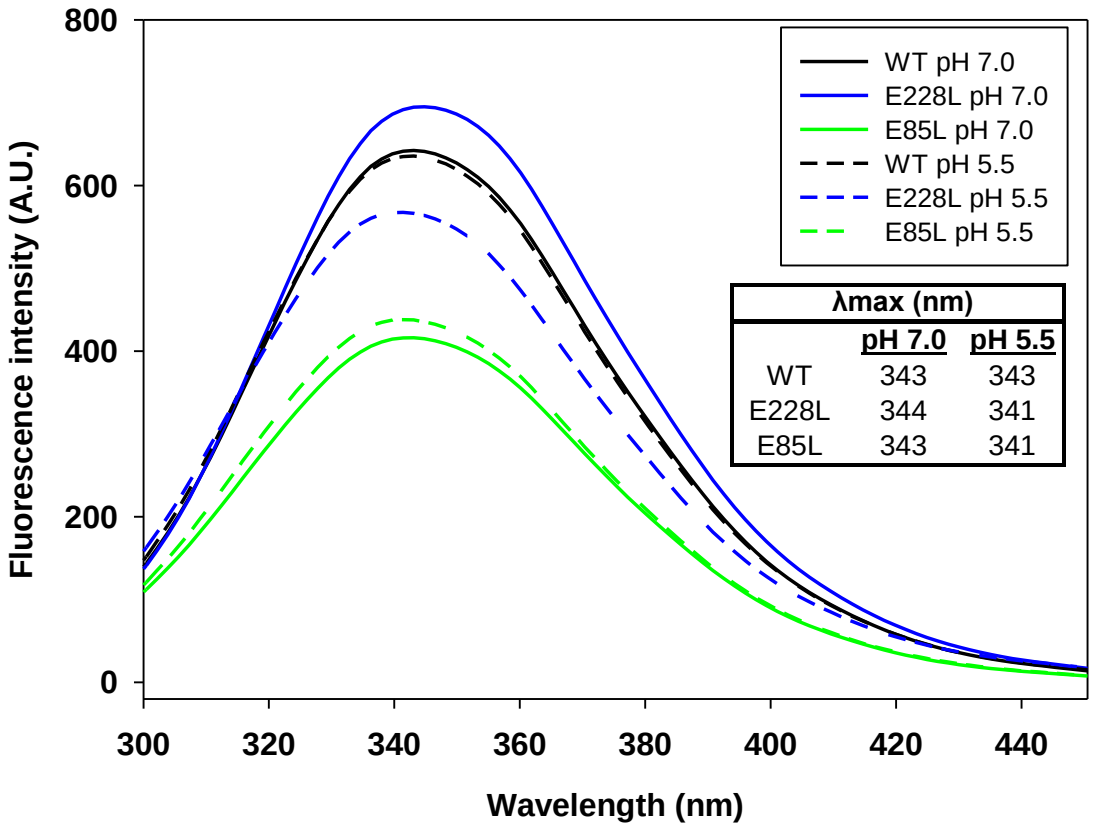


Figure 10. Intrinsic fluorescence of wild-type, E228L CLIC1 and E85L CLIC1 at pH 7.0 and pH 5.5

Tertiary packing of residues about Trp35 was examined using intrinsic fluorescence of the Trp35 residue. For all spectra, 2 μM protein in 5 mM sodium phosphate, 1 mM DTT and 0.02% NaN_3 (pH 7.0 or pH 5.5) was used and excitation was at 280 nm. Spectra were measured in triplicate, smoothed and averaged. The wavelengths of maximum emission (λ_{max}) (inset) of the three proteins fall within the range 341 - 343 nm, indicating that no significant changes to packing at the domain interface have been induced by the substitutions and that no such changes occur at low pH.

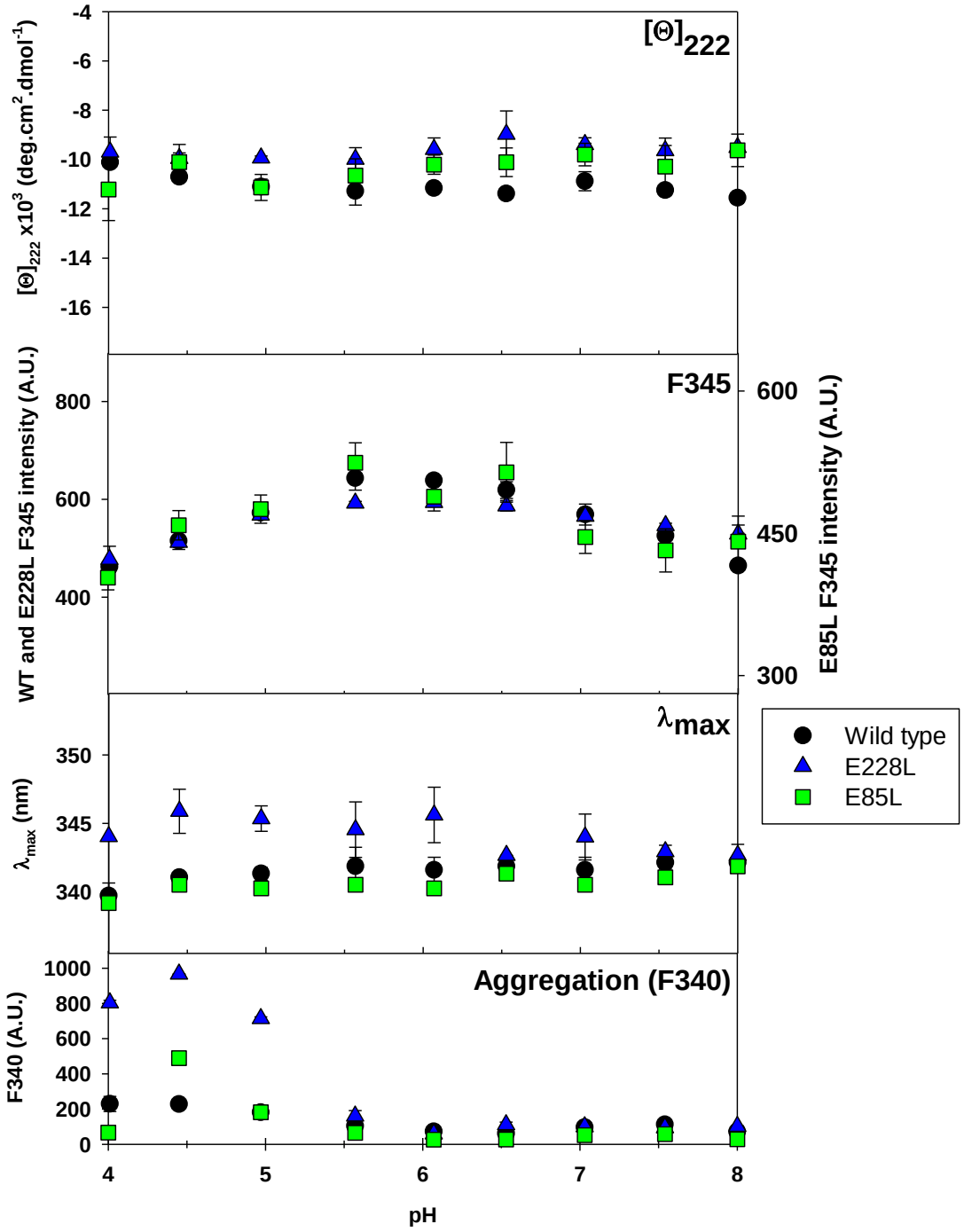


Figure 11. pH-dependence of the secondary and tertiary structures of wild-type, E228L CLIC1 and E85L CLIC1

pH-induced changes to the secondary and tertiary structure of 2 μM protein samples were measured using far-UV circular dichroism at 222 nm ($[\Theta]_{222}$) and intrinsic tryptophan fluorescence (F345 and λ_{max}). $[\Theta]_{222}$ and F345 were measured over a period of 20 s and averaged. For maximum emission wavelength (λ_{max}), tryptophan fluorescence spectra (300 – 380 nm) were measured and smoothed to enable λ_{max} identification. Aggregation of the proteins was monitored by fluorescence Rayleigh scatter at 340 nm (F340). All measurements were performed in triplicate at 20 °C in buffers of standard ionic strength (I = 0.5 M): either 50 mM sodium acetate (pH 4.0 – 5.5) or 17 mM sodium phosphate (pH 6.0 – 8.0).

Overall, limited proteolysis gels and plots of the corresponding densitometry data (Figure 12) show no change in the intensities of the protein bands at pH 7.0 as the length of proteolysis is increased. This suggests that, at pH 7.0, the three CLIC1 structures are rigid and relatively inflexible. In contrast, data for low pH show an eventual decrease in the intensity of the native band of each protein with time. At 40 minutes, the wild type and E85L CLIC1 bands visibly decrease in intensity. This occurs for E228L CLIC1 at 20 minutes. As the native bands of all three proteins are eventually cleaved at pH 5.5 but not at pH 7.0, it is likely that the CLIC1 protein structures become more flexible and/or dynamic at low pH. This observation agrees with previously published data from hydrogen-exchange mass spectrometry of CLIC1 (Stoychev *et al.*, 2009). The shorter cleavage time observed for E228L CLIC1 suggests that it is more flexible than wild-type CLIC1. Removal of the Glu228-Ser16 interdomain bond thus appears to have granted the domains greater freedom of movement.

3.2.5. The E228L CLIC1 crystal structure

3.2.5.1. E228L CLIC1 and E85L CLIC1 crystals

Protein crystals grow as a result of a careful balance that is established between factors that influence its solubility. In an ideal vapour diffusion experiment, such as hanging drop, the diffusion of water out of the supersaturated protein-containing droplet results in an increase in the protein concentration. If the protein and precipitants are correctly optimised, this results in the solution entering a ‘nucleation zone’ (Figure 13), where small crystal nuclei form. The growth of these nuclei decreases th protein concentration to such an extent in the droplet that it enters a ‘stable growth zone’ limiting the number of crystal nuclei. Here, the protein concentration remains sufficiently high for the protein to precipitate from the solution in an ordered manner, resulting in the growth of a limited number of microcrystals. If the protein concentration is too high or the precipitant is too harsh, multiple microcrystals or, at worst, amorphous aggregates form instead of large crystals. Similarly, low protein concentrations or overly mild precipitants cause the protein to remain soluble, resulting in clear droplets (McPherson, 1990).

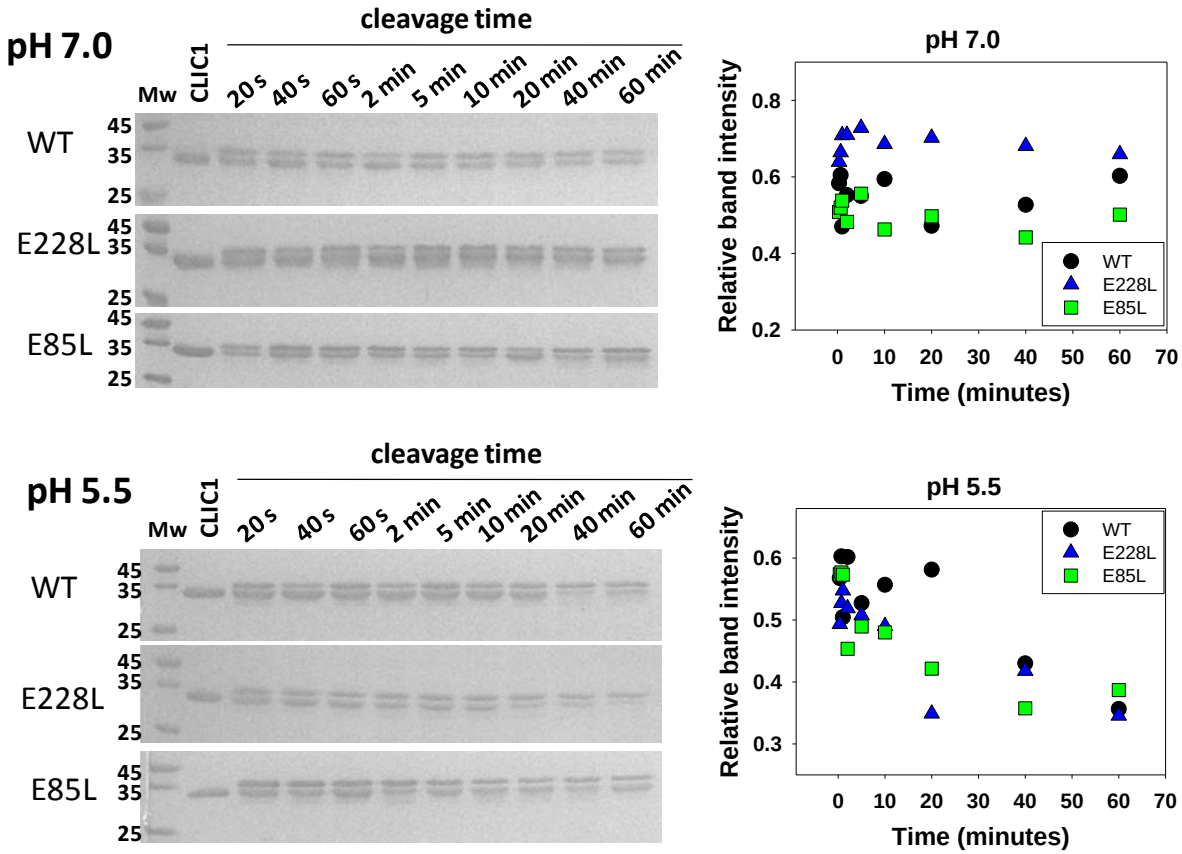


Figure 12. Limited thermolysin proteolysis of wild-type, E228L CLIC1 and E85L CLIC1 at pH 7.0 and pH 5.5

The native states of wild-type, E228L CLIC1 and E85L CLIC1 at both pH 7.0 and pH 5.5 were subjected to proteolysis by thermolysin for increasing periods of time at 20 °C. Samples were analysed on 15% SDS-PAGE gels and stained with Coomassie Brilliant Blue R-250. In all gels, the top band is thermolysin (Mw 34.6 kDa) and the lower band, if visible, is CLIC1 (Mw 26.9 kDa). The first lane is the molecular mass marker. The CLIC1 control, in which buffer was added in place of thermolysin is in the second lane. The remaining lanes show samples with increasing times of exposure to thermolysin. The relative band intensities of the CLIC1 bands at each pH were measured by densitometry and plotted. Intensities are arbitrary and are not comparable between proteins. The band intensities of the native states decrease at pH 5.5 but not at pH 7.0, indicating increased structural flexibility at low pH.

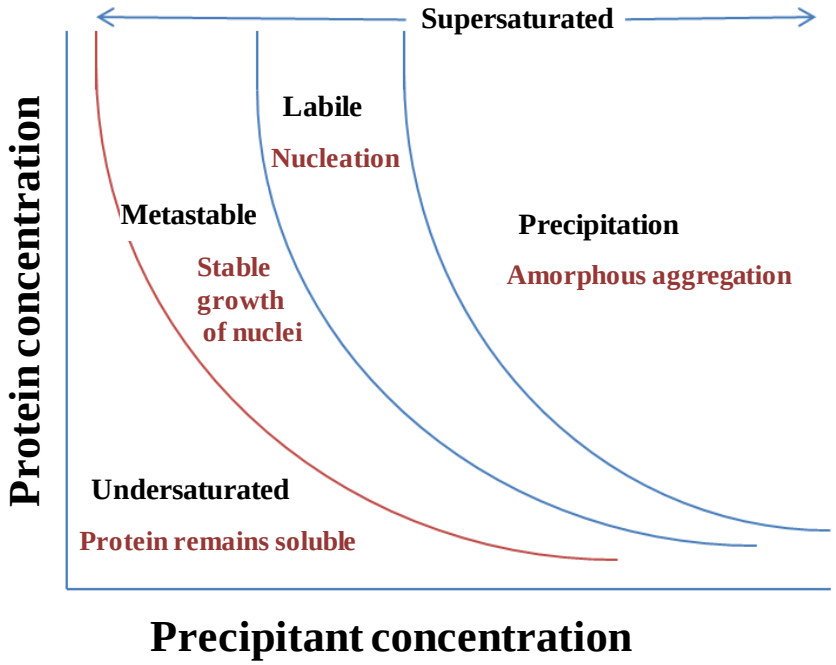


Figure 13. Phase diagram of protein crystallisation as a balance between protein and precipitant concentration

The various 'zones' involved in protein crystallisation are separated by lines, with the red line separating the under- and supersaturated solution. A careful balance of protein and precipitant concentrations will result in a crystallisation solution entering the labile zone, where stable nuclei form. The resulting decrease in protein concentration would bring the solution into the metastable zone, for growth of the nuclei formed. Extremes of either protein or precipitant concentration result in either precipitation or prolonged protein solubility. Figure adapted from McPherson, 1990.

Initial crystallisation trials were set up with 10 mg/ml E85L CLIC1, using the 96 solutions from the Hampton Index Kit. The majority of these resulted in the formation of amorphous aggregates and the remainder remained clear. No microcrystals were observed. The protein concentration was reduced to 2.5 mg/ml and the 96 trials were repeated in the hope that reducing the concentration would decrease precipitation and allow nucleation. Again, most trials resulted in amorphous aggregates. The remaining droplets remained clear, suggesting that the metastable phase had not been reached. The trials were repeated with 4 - 5 mg/ml protein, which again resulted in aggregate formation. This suggested that the conditions contained within the Hampton Index kit were either too harsh, or too moderate to induce stable crystal growth. Trials were therefore reinitiated using the Sigma-Aldrich Crystallization Extension kit. Again, trials either resulted in aggregation or prolonged solubility. No crystals of any size or form were observed. No trend in the pH, buffer or precipitant causing aggregation was observed. Thus, it was concluded that E85L CLIC1 is destabilised to such an extent that any condition that promotes precipitation from the solution is also too harsh to allow crystal growth. Crystallisation of the E85L CLIC1 variant may require the introduction of stabilising mutations (Pautsch *et al.,* 1999; Reviewed in Dale *et al.,* 2003). This, however, could result in changes to the structure and thus, an inaccurate picture of what the E85L structure truly is. Future efforts could utilise seeding of droplets with wild type CLIC1 microcrystals.

Fortunately, the initial crystallisation trials set up for E228L CLIC1 resulted in the formation of microcrystals. A decrease in the protein and precipitant concentrations reduced the number of crystals and improved their size and morphological quality. The optimal crystallisation conditions for E228L CLIC1 were determined to be 5 mg/ml protein in 0.2 M sodium chloride, 0.1 M HEPES pH 7.5, 25% (w/v) PEG 3350 at 20°C. E228L CLIC1 crystals grew as needles within 48 hours of trial setup (Figure 14). The crystals took up the IZIT dye, resulting in a dark blue hue and confirming their proteinaceous nature (Figure 15). This was also confirmed by SDS-PAGE analysis of crystals (Figure 15) in which a clear protein band is seen.



Figure 14. E228L CLIC1 crystals

Representative E228L CLIC1 crystals from two separate wells, both with the same crystallisation conditions. Crystals were grown in 0.2 M sodium chloride, 0.1 M HEPES pH 7.5, 25% (w/v) PEG 3350 at 20 °C using the hanging drop method.

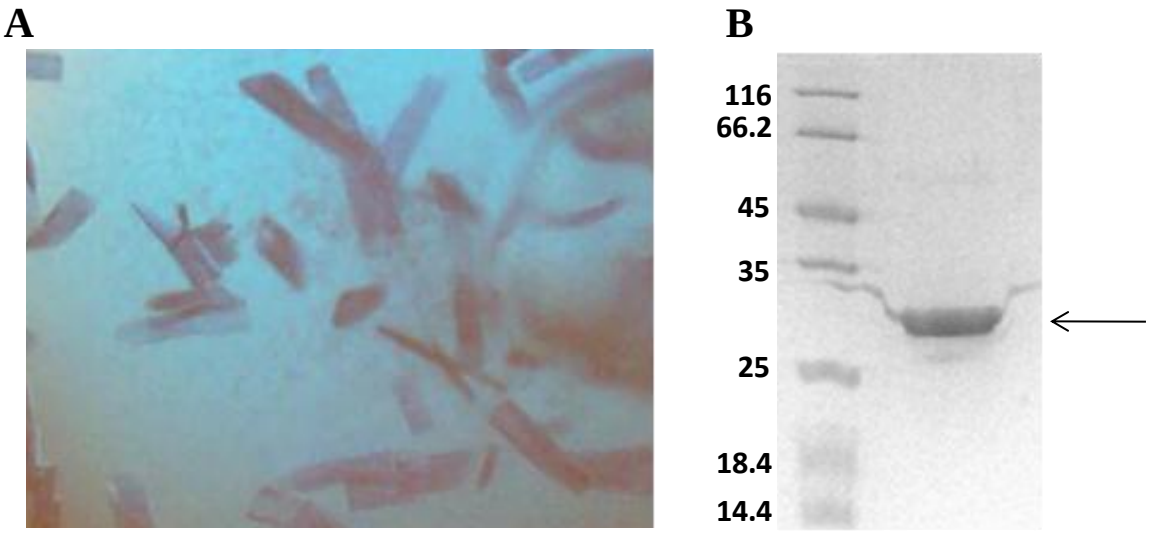


Figure 15. Confirmation of the proteinaceous nature of E228L CLIC1 crystals

To prove the proteinaceous character of the crystals grown, (A) 1 µl of IZIT dye was added to an 8 µl drop containing crystals. IZIT is able to penerate protein crystals and turn them dark blue but cannot diffuse into salt crystals due to the lack of interatomic space in a salt crystal lattice. (B) Crystals were harvested, dissolved in 10 µl of reducing SDS-PAGE sample buffer and run on an SDS-PAGE gel. A clear band, indicated with an arrow, confirms the presence of protein in the extracted crystals.

3.2.5.2. The E228L CLIC1 structural model

The structure of the E228L CLIC1 protein was solved at a resolution of 2.5 Å. Two CLIC1 monomers are present in the asymmetric unit (ASU). Crystal packing contacts in the ASU between the N-domain of the A chain and the C-domain of the B chain are formed by a network of H-bond and salt bridge interactions (Figure 16). The structure was found to have a solvent content of 47.7%, well within the 25 - 67% range first identified by Matthews (1968). However, due to poor electron density, solvent could not be modelled in the final structure. The same was true for the flexible loop region spanning residues 152 - 164, which was excluded due to a lack of sufficient density. The average B-factor (Table 3) is significantly higher than those of other CLIC1 variants (Achilonu *et al.*, 2012; Legg-E'Silva, 2012), which suggests a more dynamic and flexible structure, though this could also be due to the low resolution of the crystal structure

An *R*-factor of 0.25 and an R_{free} of 0.3 (Table 3) were obtained for refinement. As these values differ by less than 0.7, the data have not been over-interpreted (Wlodawer *et al.*, 2008). The rmsd values for bond lengths and angles are 0.0105 Å and 1.399°, respectively, and are below the accepted 0.02 Å and 4° limits. This indicates good agreement of the E228L CLIC1 model with accepted structural geometry (Wlodawer *et al.*, 2008). Additionally, >95% of all residues in the structure are found within favoured and allowed regions in the Ramachandran plot. Thus, despite the regrettable lack of solvent density, the structure is of sufficient quality to allow comparisions to the wild-type CLIC1 structure.

An overlay of the wild-type and E228L CLIC1 structures (Figure 17) indicates that no global changes were introduced to the CLIC1 fold by the mutation. The electron density confirms successful mutation of Glu228 to Leu228 (Figure 18) as no density is present for the more extended glutamate side chain. Introduction of the mutation appears to have also weakened and potentially broken the salt bridge between Asp225 and Lys20. Both side chains are oriented away from the Leu228 residue in the E228L CLIC1 structure and the distance between the two is increased (3.4 Å) compared to in the wild-type structure (2.6 Å) (Figure 17).

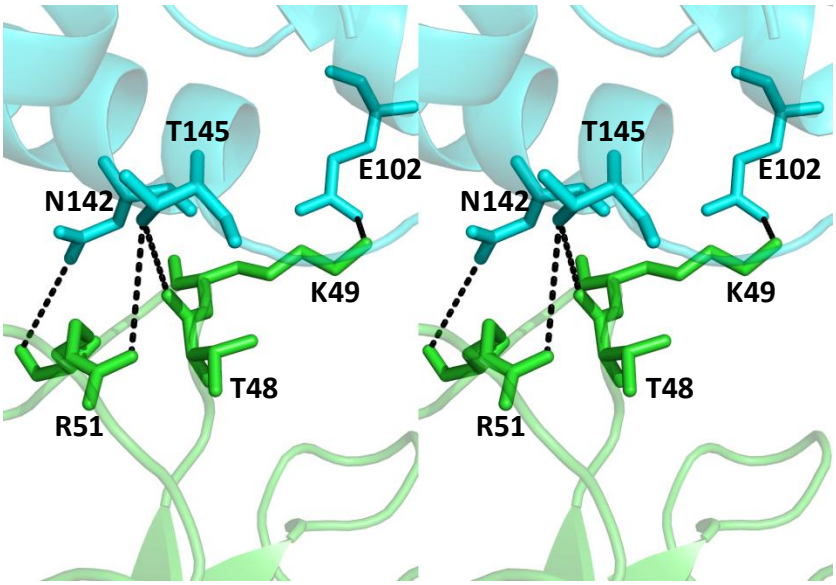


Figure 16. Crystal contacts in the E228L CLIC1 asymmetric unit

The stereo image shows crystal packing contacts between the two monomeric E228L CLIC1 molecules that form the asymmetric unit. The N-domain of the A chain (green) forms a network of H-bonds with the C-domain of the B chain (cyan). Image rendered with the PyMol Molecular Graphics System, version 1.5 Schrödinger, LLC, using PDB file 4IQA.

Table 3. Data collection and refinement statistics for E228L CLIC1

Data collection via CCD	
wavelength (Å)	1.5418
space group	$P\,2_1$
unit cell dimensions (Å)	a = 41.8 b = 71.9 c = 83.1 $\alpha = \gamma$ = 90°, β = 90.4°
resolution (Å)	71.94-2.49 (2.49-2.51)
R_{sym} (%)	0.22
$I/\sigma(I)$	31.77 (4.71)
completeness (%)	99 (40.9)
no. observed reflections	192152 (17255)
asymmetric unit content	2
Matthew's coefficient (Å^3)	2.34
solvent content (%)	47.4
Wilson B-factor (Å)	32.8
Refinement	
resolution (Å)	2.49 - 36.12 (2.49-2.55)
R_{work}/R_{free}	0.25/0.30
no. of reflections	16367 (1086)
completeness (%)	99 (89)
no. protein atoms	3489
rmsd bond lengths (Å)	0.0105
rmsd bond angles (deg)	1.399
B-factor (Å^2)	45.2
Ramachandran Analysis (%)	
favoured	95.2
allowed	4.1
disallowed	0.7

Values in parentheses are those for the highest resolution shell.

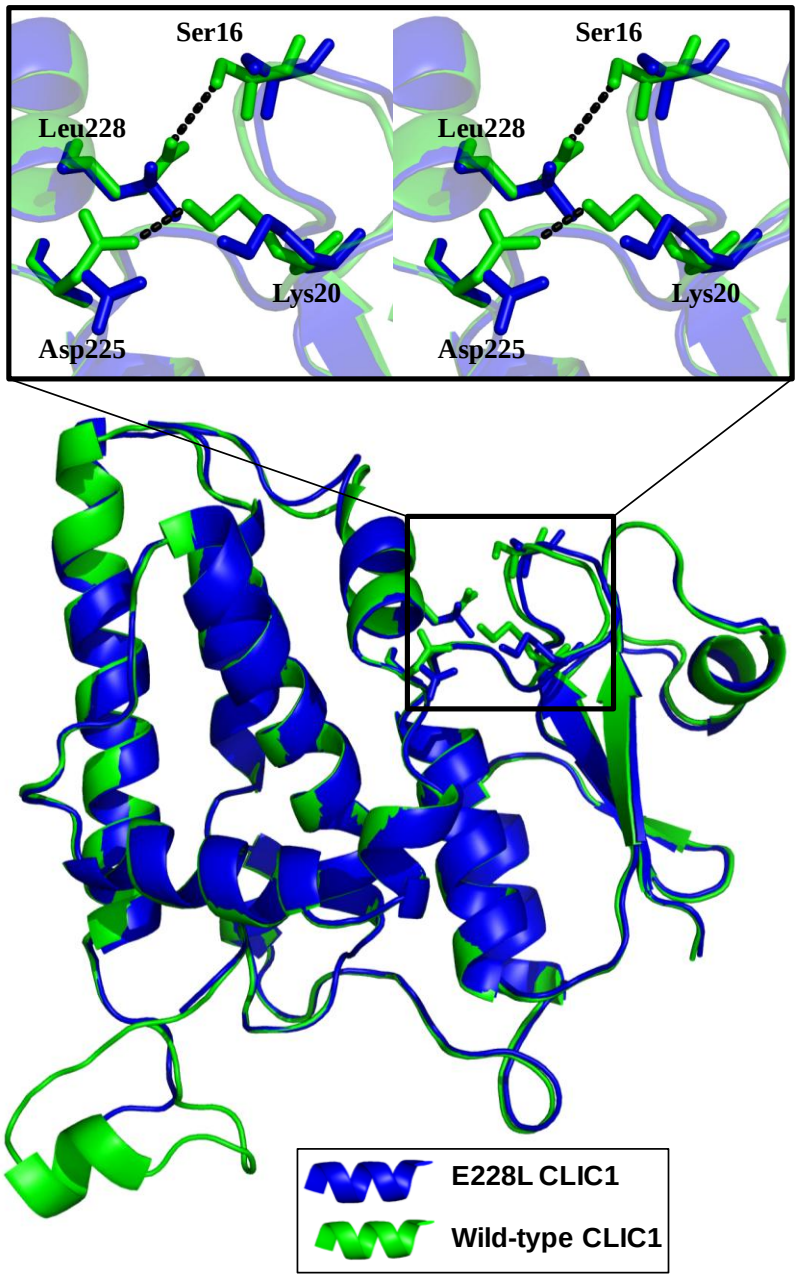


Figure 17. Overlay of the polypeptide backbones of the E228L CLIC1 and wild-type CLIC1 structures

Wild-type (PDB code 1K0M, Harrop *et al.*, 2001) and E228L CLIC1 (PDB code 4IQA) are shown in green and blue, respectively. The insert shows a stereo image of the E228L mutation site with Leu228 in blue and wild-type Glu228 in green. The structures overlay with an rmsd of 0.53 Å. Image generated using the PyMol Molecular Graphics System, version 1.5 Schrödinger, LLC.

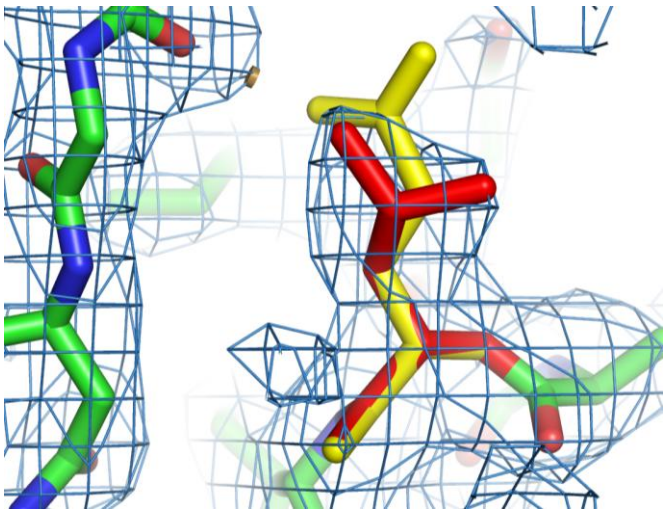


Figure 18. Electron density for residue 228

The Leu228 residue (red) of the E228L CLIC1 crystal structure (PDB code 4IQA) fits comfortably within the electron density map (blue), whereas a glutamate residue from the wild type structure extends beyond the electron density. The image was rendered using the PyMol Molecular Graphics System, Version Schrödinger, LLC, with the electron density map sigma level set to 0.3.

3.3. Stability of wild-type, E228L CLIC1 and E85L CLIC1

3.3.1. Thermal stability

Thermal unfolding of the three CLIC1 proteins was irreversible and hence no thermodynamic parameters could be calculated from data fitting. Additionally, the gradients of the thermal unfolding curves obtained differ (Figure 19). While this indicates a potentially interesting change to the unfolding cooperativity upon mutation of CLIC1, it also prevents the use of the temperature at the unfolding midpoint (T_m) to compare the thermal stabilities. The linear extrapolation method (Greene and Pace, 1974) is used for the calculation of thermodynamic parameters. Unfolding curves with shallower gradients produce greater C_m values which suggest higher conformational stability. However, fitting of the data and calculation of the ΔG values for unfolding has indicated that those proteins with less steep unfolding curves are often of lower stability than those with steep curves (Myers *et al.*, 1995). Comparison of the denaturation midpoints of unfolding curves with different gradients can therefore be misleading. Thus, in this dissertation, thermal stability is inferred from the temperature at which the unfolding process is initiated (T_{unf}). This is taken as the temperature at which an initial loss of secondary structure is observed.

Thermal unfolding data for the wild type and the two variant CLIC1 proteins shows a decrease in T_{unf}, as well as a pH-effect on thermal stability. While the T_{unf} for wild-type CLIC1 at pH 7.0 is ~ 50 °C, it is 10 °C lower for E228L CLIC1 and E85L CLIC1, where T_{unf} is ~ 40 °C. At pH 5.5, the wild type T_{unf} is 46 °C, while the E228L CLIC1 and E85L CLIC1 values are 38 °C and 31 °C, respectively (Figure 19). This indicates that both variant proteins are destabilised relative to the wild type, with the least stable of the three being the E85L variant. E85L CLIC1 also demonstrates a significant loss of unfolding cooperativity at pH 5.5, which indicates a considerable change to the susceptibility of the protein to temperature. Furthermore, the proximity of the unfolding temperatures of E228L CLIC1 and E85L CLIC1 to human physiological temperature (37.4°C) demonstrates significant destabilisation of the CLIC1 fold and suggests that these residues are important for the structural integrity of CLIC1.

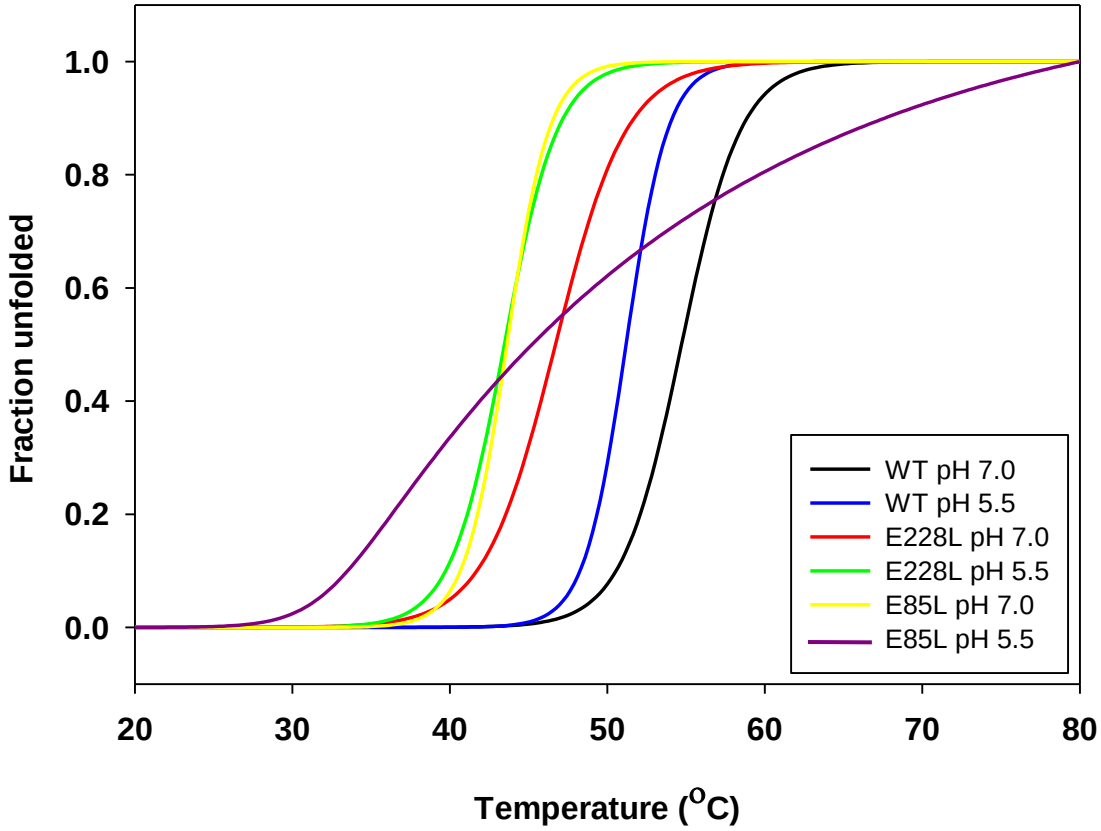


Figure 19. Thermal unfolding of wild-type, E228L CLIC1 and E85L CLIC1

Unfolding of 2 μM CLIC1 monitored at both pH 7.0 and pH 5.5 (50 mM sodium phosphate, 1 mM DTT, 0.02% NaN_3) using the mean residue ellipticity at 222 nm, which reflects the amount of α-helical structure present. Temperature was increased at a rate of 1 °C/min from 20 °C to 80 °C. Mean residue ellipticity at 222 nm was normalised such that 0 represents the native state and 1, the unfolded state. Unfolding data were fitted to a sigmoid using SigmaPlot v11.0 (SPSS Science, Chicago, Illinois USA) and fits are presented for clarity.

The *trend* in the pH-dependence of thermal stability that is observed for the wild type is retained in the variants, in that both variants show lower stability at pH 5.5 than at pH 7.0. However, the *extent* to which pH affects the unfolding temperature is reduced in the variants. While the wild type undergoes a 12% decrease in T_{unf} at lower pH, the E228L and E85L variants show a decrease of 9.5% and 7.8%, respectively. This suggests that some pH-sensitivity is lost in the variants, but may simply be a result of the degree of global destabilisation caused by the removal of the respective glutamate residues. Thus, thermal unfolding data show that the E228L and E85L variant CLIC1 proteins are destabilised relative to the wild type.

3.3.2. Unfolding equilibration time

The time required for wild-type CLIC1 to reach equilibrium between the native and unfolded states upon exposure to urea has been well-established (McIntyre, 2007; Legg-E'Silva, 2011). The equilibration times for E228L CLIC1 and E85L CLIC1 however, required examination using timed unfolding. The timed unfolding traces overlay with the unfolded baselines (Figure 20) of each variant. This indicates that the proteins unfold, losing both secondary and tertiary structure within the 20 s dead time that is required for manual mixing, sample loading and initiation of the measurement. As the wild type takes 10 minutes to unfold in 8 M urea (McIntyre, 2007), this serves to further confirm the destabilisation of the E228L CLIC1 and E85L CLIC1 proteins. The variant proteins were allowed to unfold for 30 minutes to reach equilibrium. Although this is longer than the unfolding period of either protein, this time proved to be practical for sample preparation and neither protein showed aggregation (F340, Figures 21 and 22) with this longer incubation period.

3.3.3. Reversibility of folding

In order for unfolding data to be fitted to an unfolding model and thereby used to calculate thermodynamic parameters, unfolding must have been shown to be reversible (Pace, 1986). Refolding of E228L CLIC1 and E85L CLIC1 was monitored using tryptophan and ANS fluorescence as well as far-UV CD. Overlays of the unfolding and refolding curves of both proteins monitored by CD and fluorescence show recovery of native structure in 0.8 M urea (Figures 21 and 22). However, the CD-monitored unfolding and refolding curves of both variants do not overlay at low denaturant concentrations (2 - 4 M urea), indicating that the two pathways differ.

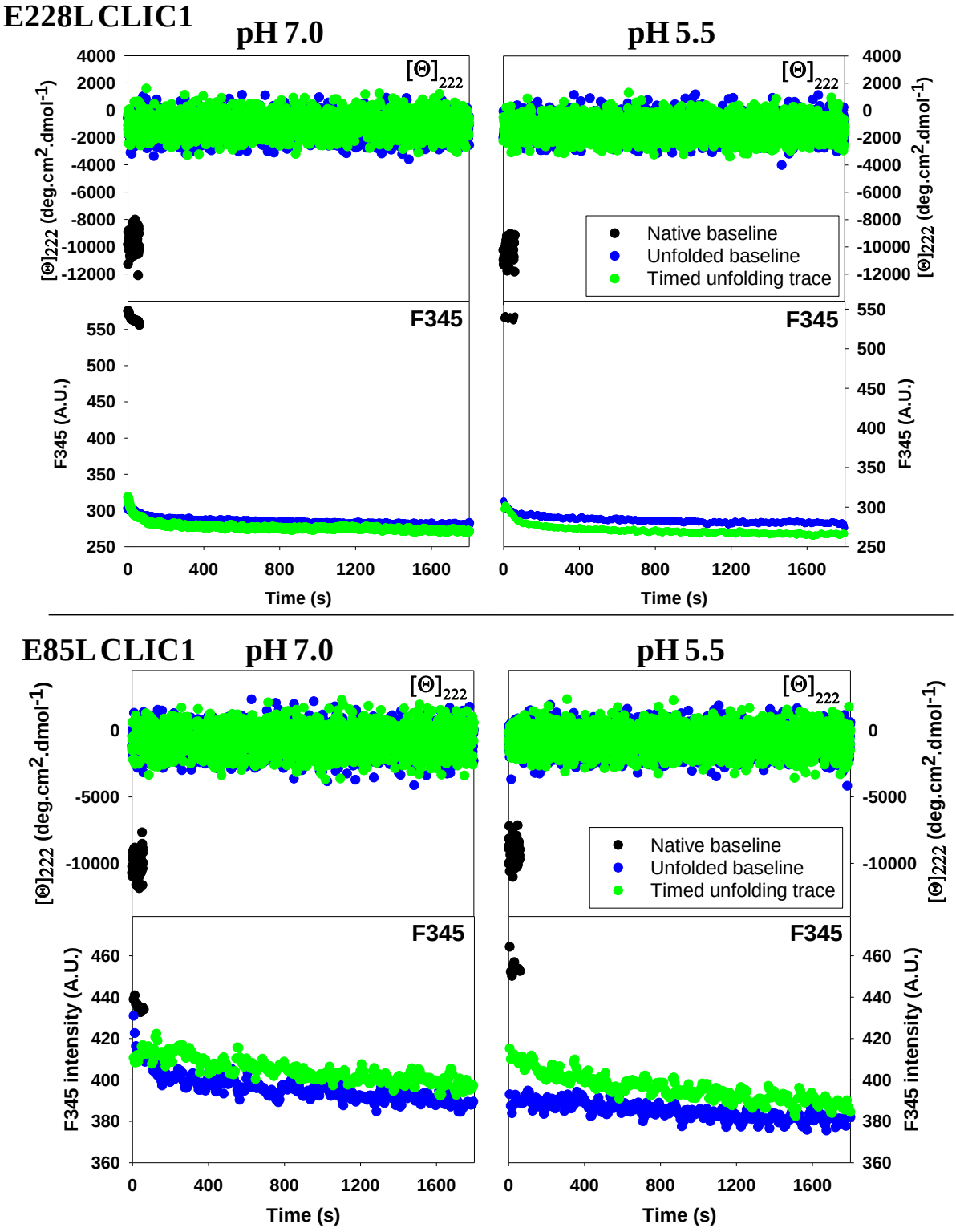


Figure 20. Unfolding of E228L CLIC1 and E85L CLIC1 over time

E228L CLIC1 and E85L CLIC1 at a concentration of 2 μM were unfolded in 8 M urea at both pH 7.0 and pH 5.5 (50 mM sodium phosphate, 0.02% NaN_3, 1 mM DTT). Unfolding was monitored from 20 s after urea was added using far-UV circular dichroism ($[\Theta]_{222}$) and intrinsic fluorescence (F345). No difference was observed between a pre-unfolded control and time-monitored samples, indicating that the proteins unfold fully within 20 s of urea addition.

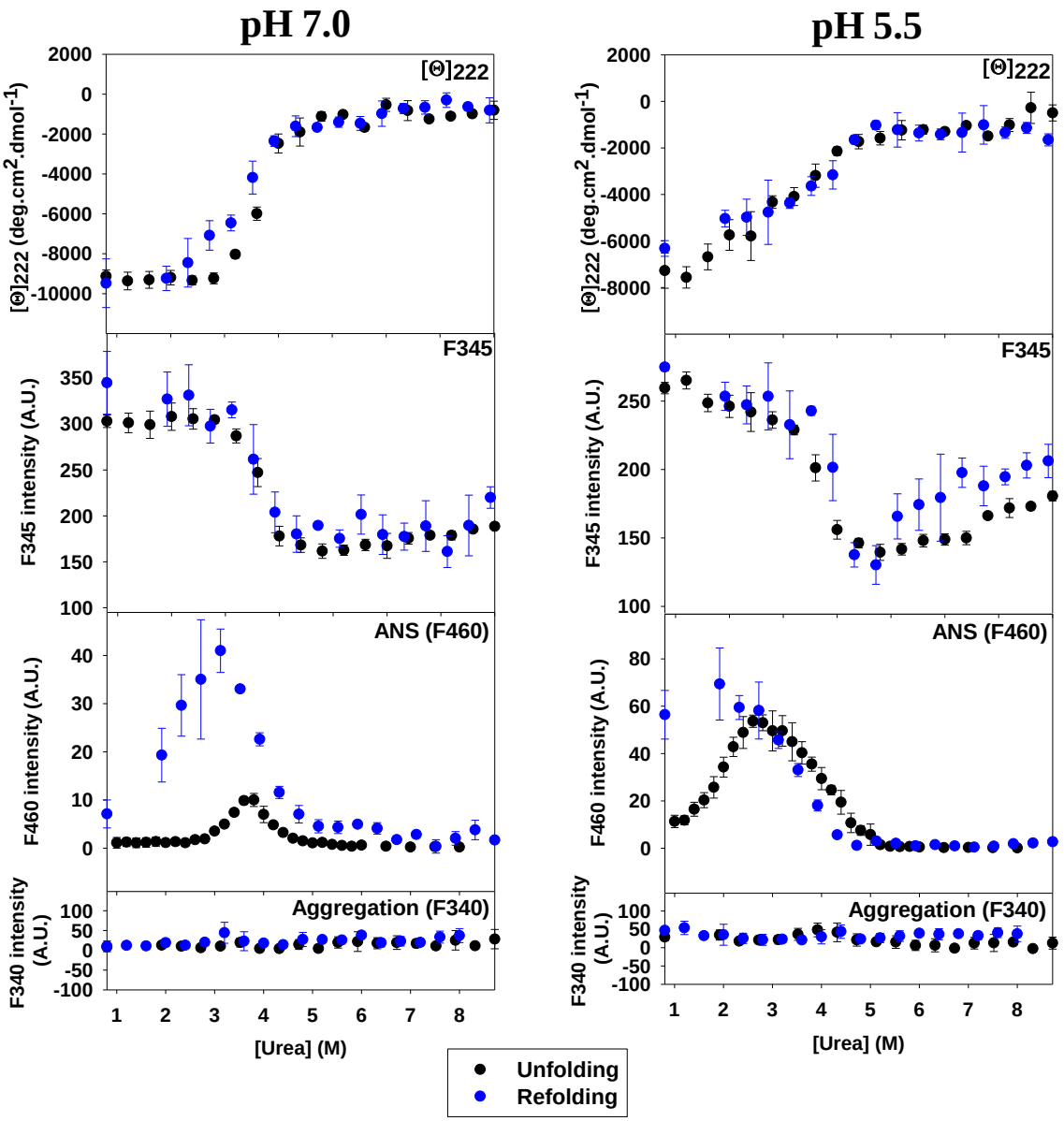


Figure 21. Unfolding and refolding of E228L CLIC1 at pH 7.0 and pH 5.5

Unfolding and refolding of E228L CLIC1 was monitored by far-UV circular dichroism ($[\Theta]_{222}$) and tryptophan (F345) and ANS fluorescence (F460) at 20 °C in 50 mM sodium phosphate, 0.02% NaN_3 with 1 mM DTT at pH 7.0 and pH 5.5. For refolding, 10 µM E228L CLIC1 was unfolded in 8 M urea for 30 minutes, diluted 10 x with buffer and urea to 1 µM and allowed to refold for 2 hours at 20 °C. For unfolding, measurements were done with 1 µM protein that had been allowed to reach equilibrium for 30 minutes. For extrinsic fluorescence, ANS was added in excess (200 µM) and samples were incubated for 1 hour to allow binding. Emission spectra were then measured in triplicate, corrected for buffer and the intensity at 460 nm was used. Protein aggregation was monitored using fluorescence scatter at 340 nm (F340).

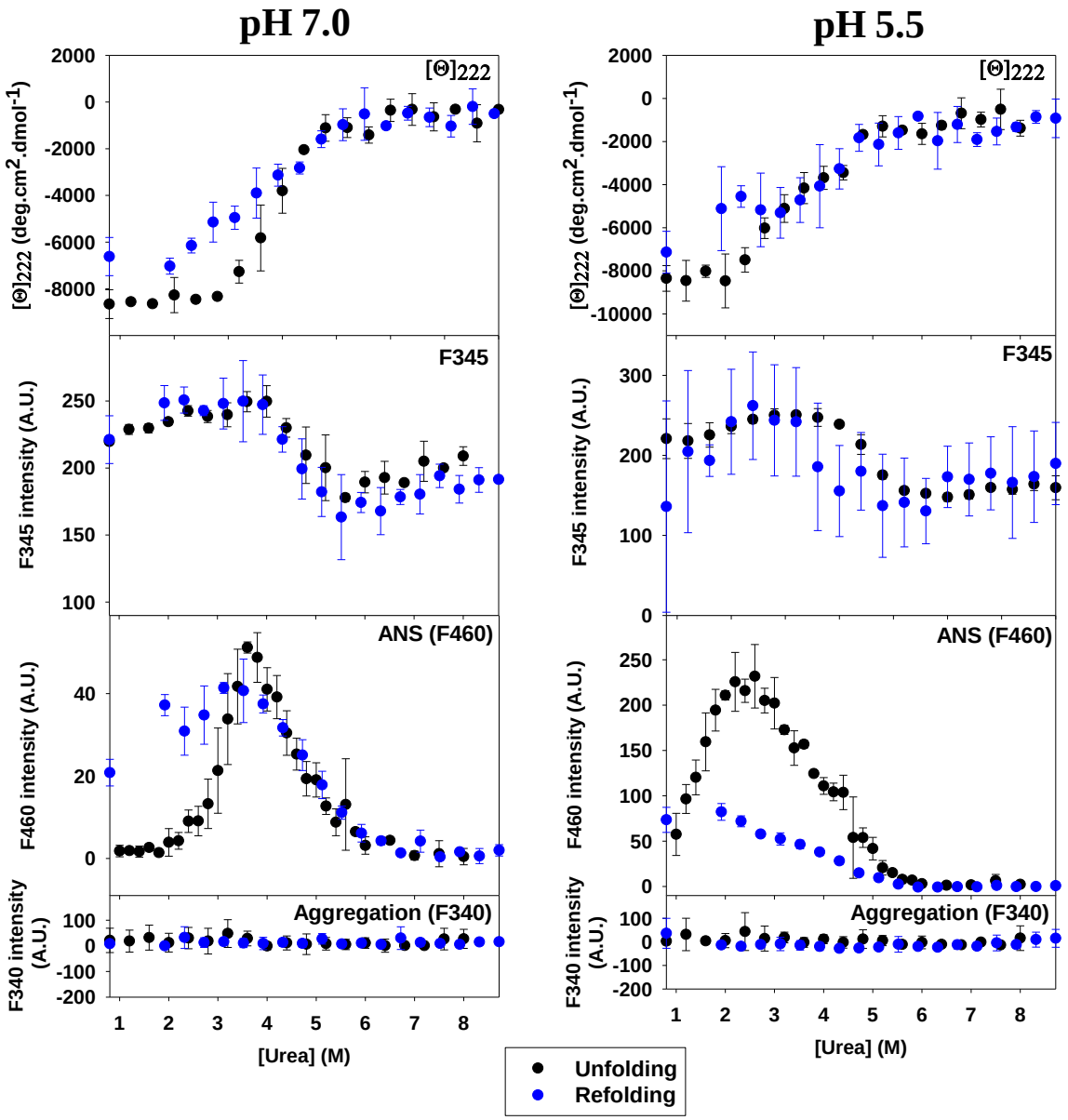


Figure 22. Unfolding and refolding of E85L CLIC1 at pH 7.0 and pH 5.5

Unfolding and refolding of E85L CLIC1 was monitored by far-UV circular dichroism ($[\Theta]_{222}$) and tryptophan (F345) and ANS fluorescence (F460) at 20 °C in 50 mM sodium phosphate, 0.02% NaN_3 with 1 mM DTT at pH 7.0 and pH 5.5. For refolding, 10 µM E85L CLIC1 was unfolded in 8 M urea for 30 minutes, diluted 10 x with buffer and urea to 1 µM and allowed to refold for 2 hours at 20 °C. For unfolding, measurements were done with 1 µM protein that had been allowed to reach equilibrium for 30 minutes. For extrinsic fluorescence, ANS was added in excess (200 µM) and samples were incubated for 1 hour to allow binding. Emission spectra were then measured in triplicate, corrected for buffer and the intensity at 460 nm was used. Protein aggregation was monitored using fluorescence scatter at 340 nm.

This noncoincidence of unfolding and refolding transitions is also seen in the ANS data (F460 in Figures 21 and 22), and is not a consequence of aggregation (Figures 21 and 22, F340). The E228L and E85L substitutions thus appear to have altered the refolding pathway of CLIC1. Unfolding is therefore irreversible and thermodynamic parameters can regrettably not be determined for either variant.

3.3.4. Equilibrium unfolding

Protein stability was examined using urea-induced equilibrium unfolding. Thermolysin pulse proteolysis was performed on wild-type, E228L CLIC1 and E85L CLIC1 unfolded in a range of urea concentrations from 0.5 - 6.5 M. Proteolysis gels and their corresponding unfolding curves at pH 7.0 (Figure 23) show the disappearance of the wild-type CLIC1 band at 4.5 M urea, while the E228L CLIC1 and E85L CLIC1 bands are absent at concentrations above 3.5 M urea. At pH 5.5, the protein band disappears at respectively lower urea concentrations. This decrease suggests that the three CLIC1 proteins are destabilised at pH 5.5.

Additionally, unfolding was monitored spectroscopically by both far-UV CD and intrinsic fluorescence. The use of two independent spectroscopic methods is important as it allows the detection of on-pathway intermediates that may not be revealed by a single probe (Yagawa *et al.*, 2010). Shifting of the λ_{max} values of the fluorescence spectra as the protein unfolds results in differences between unfolding curves that are produced for different emission wavelengths. For example, as the population of unfolded protein increases, the λ_{max} will tend towards that of the unfolded state (356 nm, for CLIC1 (McIntyre, 2007)). This would result in red-shifting of the λ_{max} and a corresponding decrease in intensity at the native λ_{max} (345 nm). Thus, the unfolding curves monitored by F356 and F345 may differ. For this reason, multiple emission wavelengths were monitored. The fluorescence wavelengths used were 310 nm, for tyrosine fluorescence, 345 nm for maximal Trp35 fluorescence in the native state and 356 nm for maximal Trp35 fluorescence in the unfolded state (McIntyre, 2007). These data were then overlaid with mean residue ellipticity at 222 nm (Figure 24).

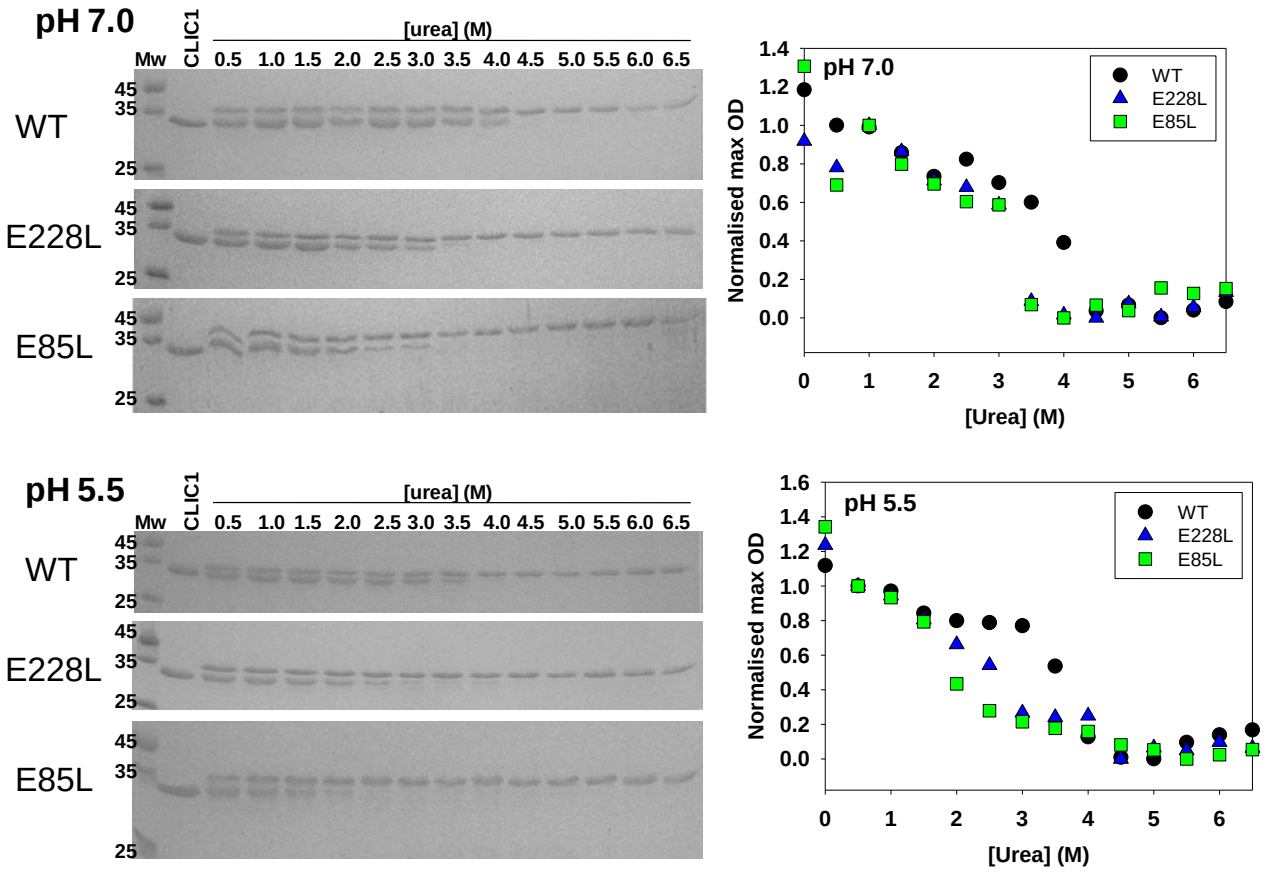


Figure 23. Pulse proteolysis of unfolding CLIC1

Wild-type, E228L CLIC1 and E85L CLIC1 at both pH 7.0 (above) and pH 5.5 (below) were allowed to unfold in increasing concentrations of urea before undergoing pulse proteolysis by thermolysin for a period of 60 s at 20 °C. Following quenching with EDTA, samples were analysed on 15% SDS-PAGE gels and stained with Coomassie Brilliant Blue R-250. In all gels, the top band is thermolysin (Mw 34.6 kDa) and the lower band, if visible, is CLIC1 (Mw 26.9 kDa). The first lane is the molecular mass marker. The CLIC1 control, in which buffer was added in place of thermolysin is in the second lane. The remaining lanes show unfolded samples in various [urea] digested with thermolysin. The maximum optical densities of the CLIC1 bands at each pH were measured by densitometry, normalised and plotted as a function of urea. Disappearance of the CLIC1 band at pH 5.5 occurs at lower [urea] than pH 7.0, indicating destabilisation of all three CLIC1 proteins at low pH. Similarly, the E228L CLIC1 and E85L CLIC1 variants appear to be of lower stability than wild-type CLIC1.

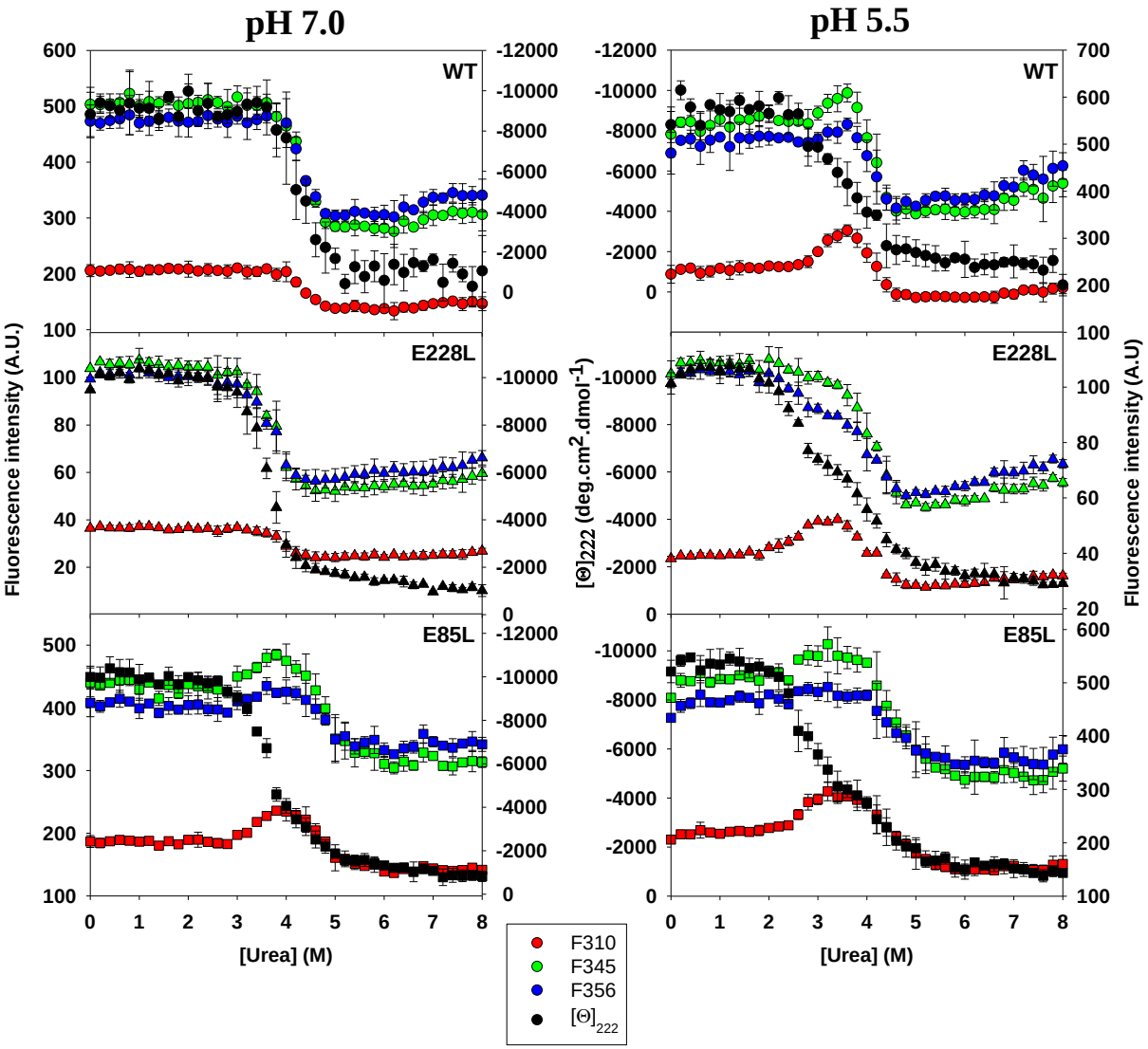


Figure 24. Equilibrium unfolding of wild-type, E228L CLIC1 and E85L CLIC1

Unfolding of 2 μM CLIC1 samples at 20 ºC with 0 - 8 M urea was monitored using intrinsic fluorescence at 310 nm, 345 nm, 356 nm and mean residue ellipticity at 222 nm. Unfolding of CLIC1 was performed in triplicate at both pH 7.0 and pH 5.5 in 50 mM sodium phosphate with 1 mM DTT and 0.02% NaN_3.

As ΔG values to quantify conformational stability could not be determined, the concentration of urea at which unfolding is first observed is used to compare the relative stabilities of the three CLIC1 proteins, as this is most likely to indicate the susceptibility of the native state to urea. Data from the spectroscopic probes is used, as they are more sensitive than the pulse proteolysis technique. Fluorescence data at 310 nm (Figure 24) show an inflection, which suggests the formation of a stable intermediate state at pH 5.5. As this inflection is not seen consistently in the F345- and F356-monitored unfolding curves (Figure 24), the F310-monitored curves were used along with far-UV CD data for comparison of the three CLIC1 proteins (Figure 25). At pH 7.0, wild-type CLIC1 starts unfolding in ~ 3.6 M urea (Figure 25). In contrast, E228L CLIC1 and E85L CLIC1 unfold in ~ 3.2 M urea, again suggesting destabilisation. At pH 5.5, these concentrations are further decreased: wild-type unfolds in ~ 2.6 M urea with E228L CLIC1 and E85L CLIC1 unfolding in ~ 2.2 M and ~ 2.4 M urea, respectively (Figure 25). Thus, both variants show lower chemical stability towards urea than wild-type and all three proteins are destabilised at low pH.

Far-UV CD-monitored unfolding curves at both pH 7.0 and pH 5.5 (Figure 25) are free of additional inflections, suggesting that the three CLIC1 proteins unfold via a 2-state mechanism at both pHs. The same is true for the F310-monitored unfolding curves of wild-type and E228L CLIC1 at pH 7.0. However, the F310 data for E85L CLIC1 at pH 7.0 and all three proteins at pH 5.5 show an inflection, which suggests the formation of an intermediate state and the assumption of a 3-state unfolding pathway.

A 3-state unfolding curve is also seen in the intensity averaged emission wavelength plots (Figure 26), which show inflections and corresponding blue-shifting of the fluorescence wavelength at denaturant concentrations from 2.5 - 4.5 M. Thus, while wild-type and E228L CLIC1 unfold via a 2-state pathway at pH 7.0, they assume a 3-state pathway at low pH. This agrees with previous unfolding data for wild-type, which demonstrates the formation of a partially-unfolded, ANS-binding intermediate state at pH 5.5 in 3.8 M urea (Fanucchi *et al.*, 2008). E85L CLIC1 appears to follow a 3-state unfolding transition at both pHs.

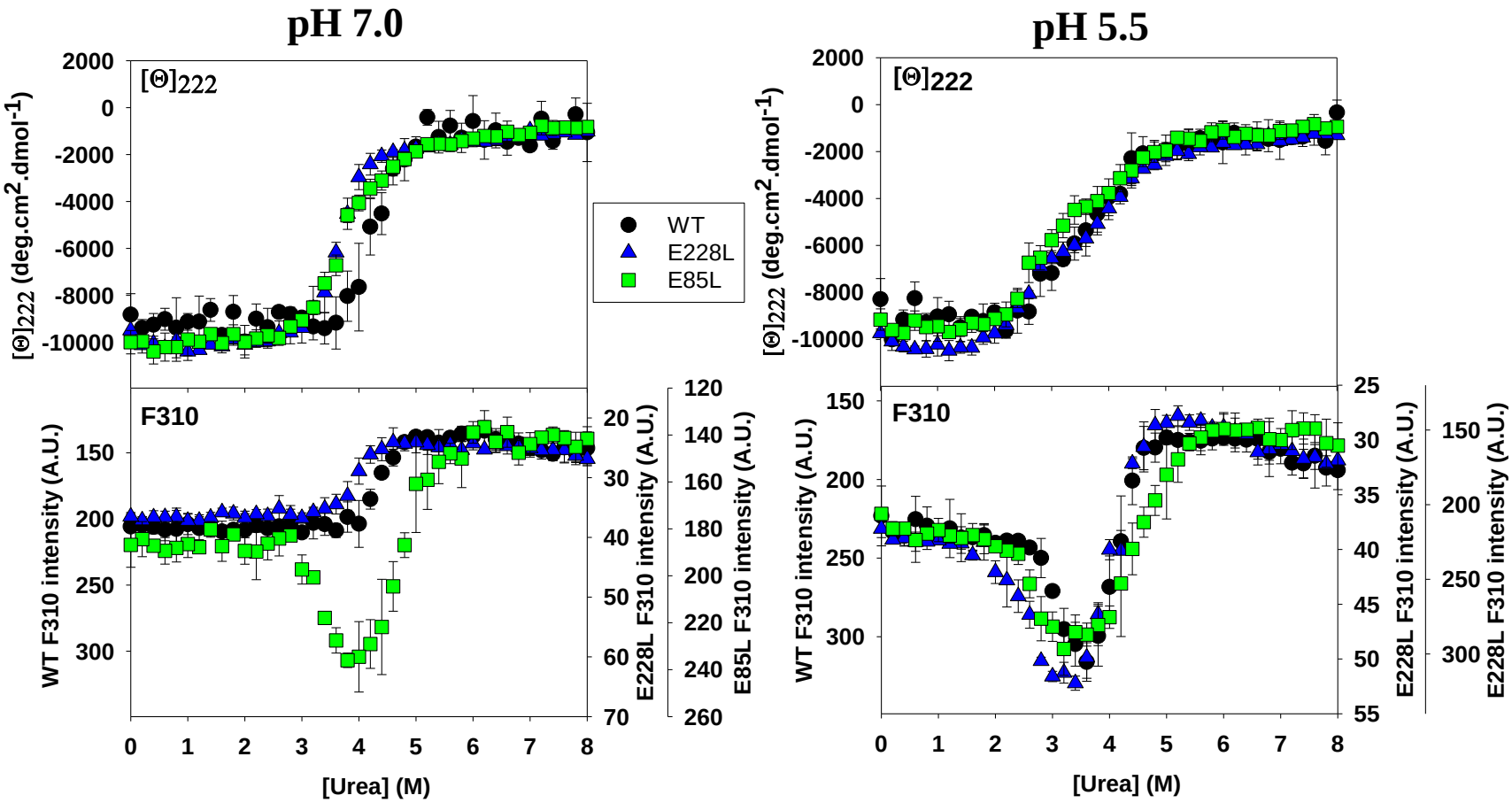


Figure 25. Comparison of urea-induced equilibrium unfolding curves of wild-type, E228L CLIC1 and E85L CLIC1

Unfolding of 2 μM CLIC1 at pH 7.0 and pH 5.5 (50 mM sodium phosphate, 1 mM DTT, 0.02% NaN_3) was monitored using far-UV CD ($[\Theta]_{222}$) and tryptophan fluorescence (F310) after excitation at 280 nm at both pH 7.0 and pH 5.5. All unfolding was done in 50 mM Na_2HPO_4, 1 mM DTT, 0.02% NaN_3 at 20 °C in triplicate.

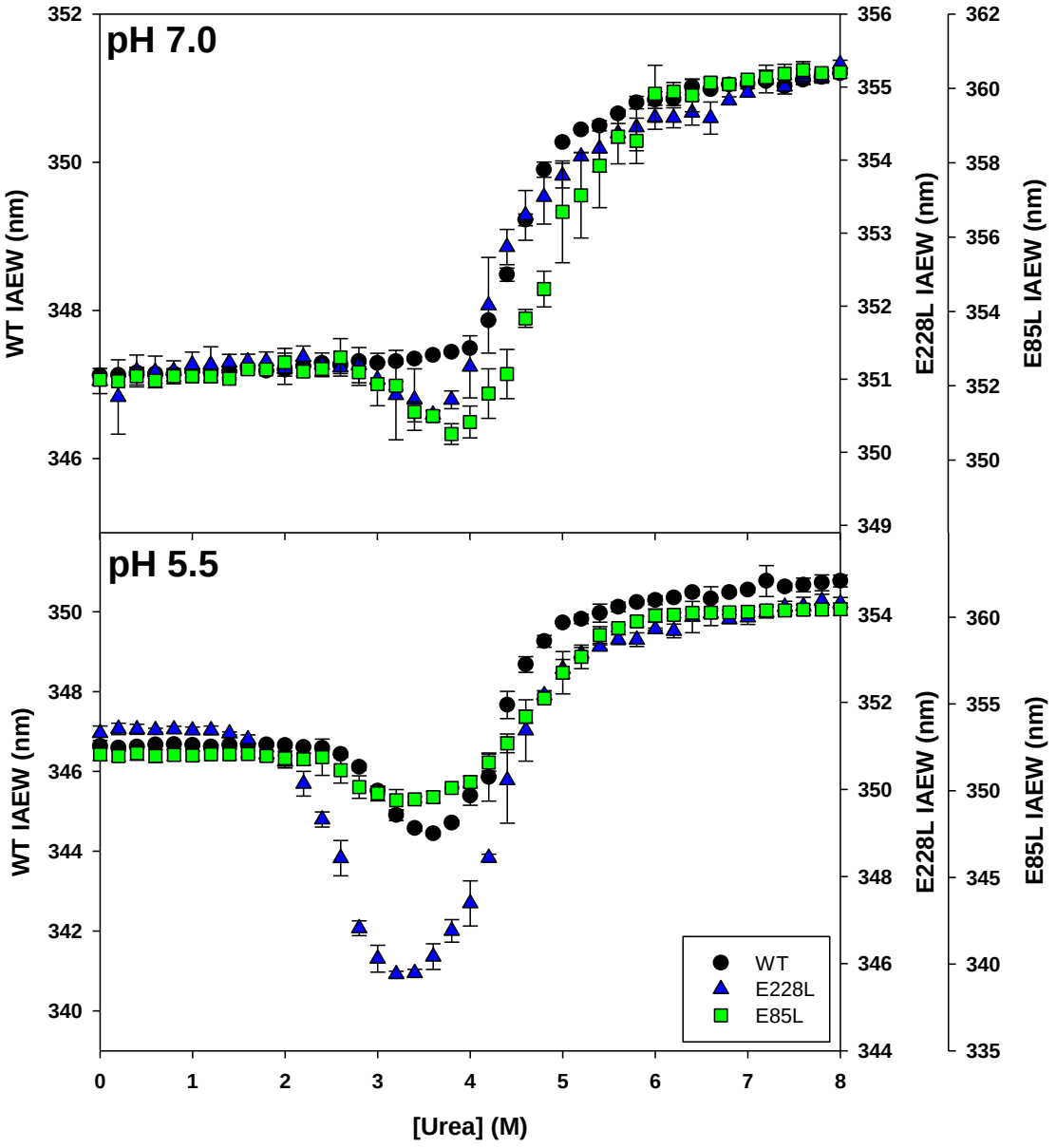


Figure 26. Intensity averaged emission wavelength for CLIC1 unfolding

Intensity averaged emission wavelength (IAEW) was calculated using fluorescence spectra for pH 7.0 and pH 5.5 for 2 μM wild-type, E228L CLIC1 and E85L CLIC1 using the equation: $IAEW = \Sigma\lambda_i.F_i. / \Sigma F_i$ where λ_i is the wavelength and F_i is the fluorescence intensity.

3.3.5. The unfolding intermediate

In order to confirm that the observed inflections in the unfolding data are due to the formation of the partially-unfolded intermediate state that was previously observed for the wild type, ANS was applied as an additional probe. When ANS binds to exposed hydrophobic patches on a protein, the fluorescence λ_{max} is blue-shifted (Semisotnov, *et al.*, 1991). ANS has been seen to fluoresce at 460 nm upon binding to the CLIC1 intermediate state (McIntyre, 2007) and hence the fluorescence intensity at this wavelength was used in the ANS studies.

The observation of an F460 peak under mild denaturing conditions for wild-type at pH 5.5 and not at pH 7.0 (Figure 27) confirms equilibrium unfolding data (Figure 27) and the formation of a partially-unfolded intermediate state. Similarly, both E228L CLIC1 and E85L CLIC1 bind ANS at pH 5.5, though the peak is at lower denaturant concentrations than for wild-type (2.8 M and 2.6 M, urea respectively). This suggests that E228L CLIC1 and E85L CLIC1 require less denaturant to form intermediate. These urea values do not, however, agree entirely with those from the intrinsic fluorescence studies (Figure 25). Given the high sensitivity and global nature of ANS fluorescence, these F460 increases at lower denaturant concentrations may also be due to unfolding and exposure of non-polar residues distal to and thus not detected by Trp35 fluorescence.

Of additional interest are the F460 peaks observed for E228L CLIC1 and E85L CLIC1 at pH 7.0 (Figure 27). E228L CLIC1 binds a relatively small amount of ANS, suggesting the intermediate is formed in populations not large enough for detection by fluorescence (Figure 25). The F460 intensity of E85L CLIC1, however, equals that of wild-type at pH 5.5. This indicates that a significant population of intermediate state is formed by E85L CLIC1 at pH 7.0 and agrees with fluorescence-monitored unfolding curves (Figure 25).

Following the confirmation of intermediate state formation by wild-type and both variant CLIC1 proteins, the structures of the four intermediate states were characterised to confirm that the variant proteins form the same intermediate as wild-type. The far-UV CD, fluorescence and ANS binding spectra of the intermediate state formed by each protein were measured (Figures 27 and 28) and compared to native state. All three proteins show blue-shifting of λ_{max} to ~ 336 nm upon intermediate formation, suggesting that Trp35 moves into a more hydrophobic environment. Far-UV CD spectra of the intermediates indicate that some secondary structure is lost (Figure 28).

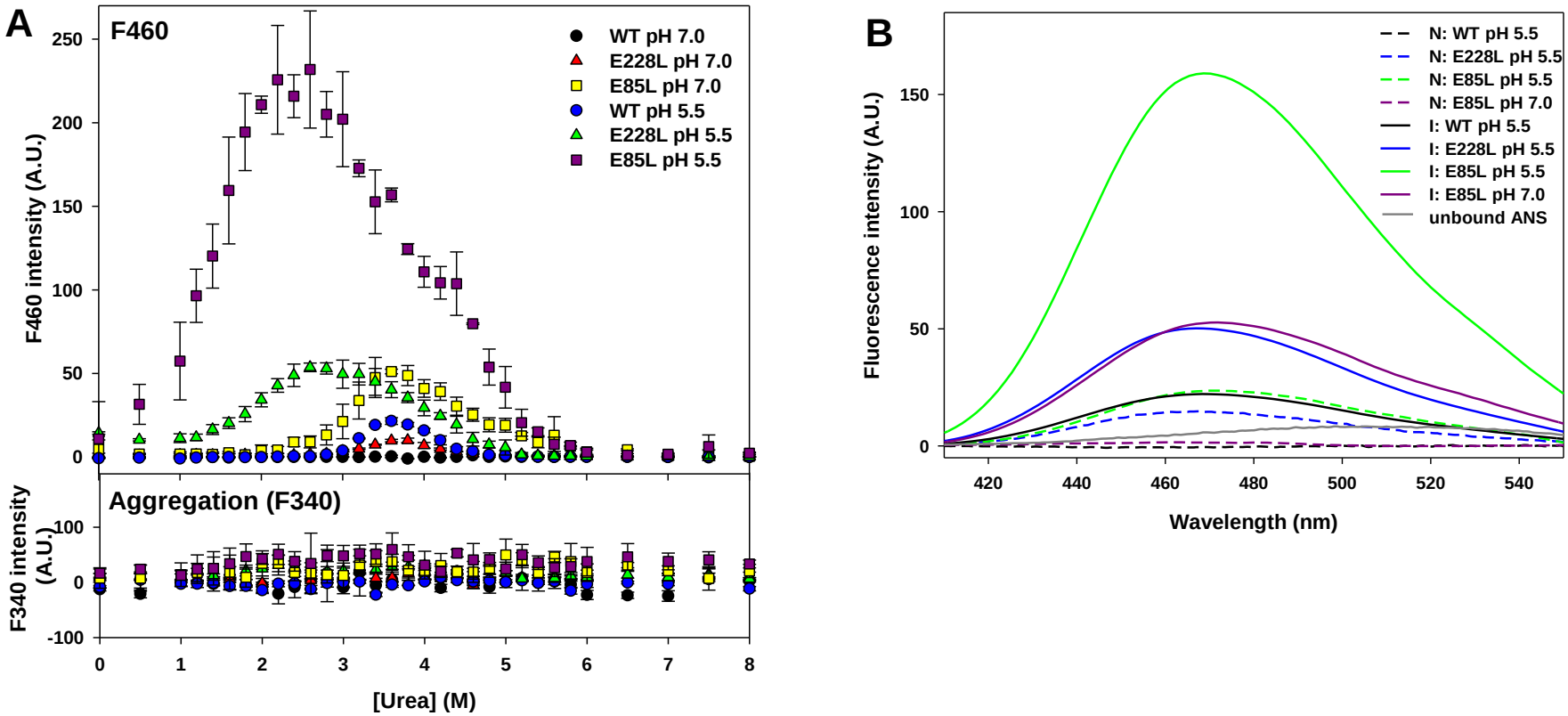


Figure 27. ANS binding to wild-type, E228L CLIC1 and E85L CLIC1 at pH 7.0 and pH 5.5

(A) ANS-monitored equilibrium unfolding shows detection of partially-unfolded intermediate states under mild denaturing conditions. ANS at a final concentration of 200 μM was allowed to bind to 2 μM CLIC1 at pH 7.0 and pH 5.5 (50 mM sodium phosphate, 1 mM DTT, 0.02% NaN_3) for 60 minutes at 20 °C. ANS was excited at 390 nm and fluorescence emission at 460 nm was monitored. Samples were also excited at 340 nm to monitor aggregation. Peaks in F460 observed at pH 5.5 for wild-type, E228L CLIC1 and E85L CLIC1 at 3.6 M, 2.8 M and 2.4 M urea, respectively, suggest intermediate formation. This is also seen for E85L CLIC1 at pH 7.0 in 3.8 M urea. (B) Fluorescence spectra of ANS bound to the native (N) and intermediate (I) states of 2 μM wild-type, E228L CLIC1 and E85L CLIC1. Free, unbound ANS fluoresces maximally at 510 nm. ANS bound to the intermediate states of wild-type, E228L CLIC1 and E85L CLIC1 at pH 5.5 fluoresces at 468 nm, 467 nm and 469 nm, respectively. ANS bound to the E85L CLIC1 intermediate at pH 7.0 peaks at 472 nm, indicating that hydrophobicity of the ANS binding environment is similar for all four intermediate states.

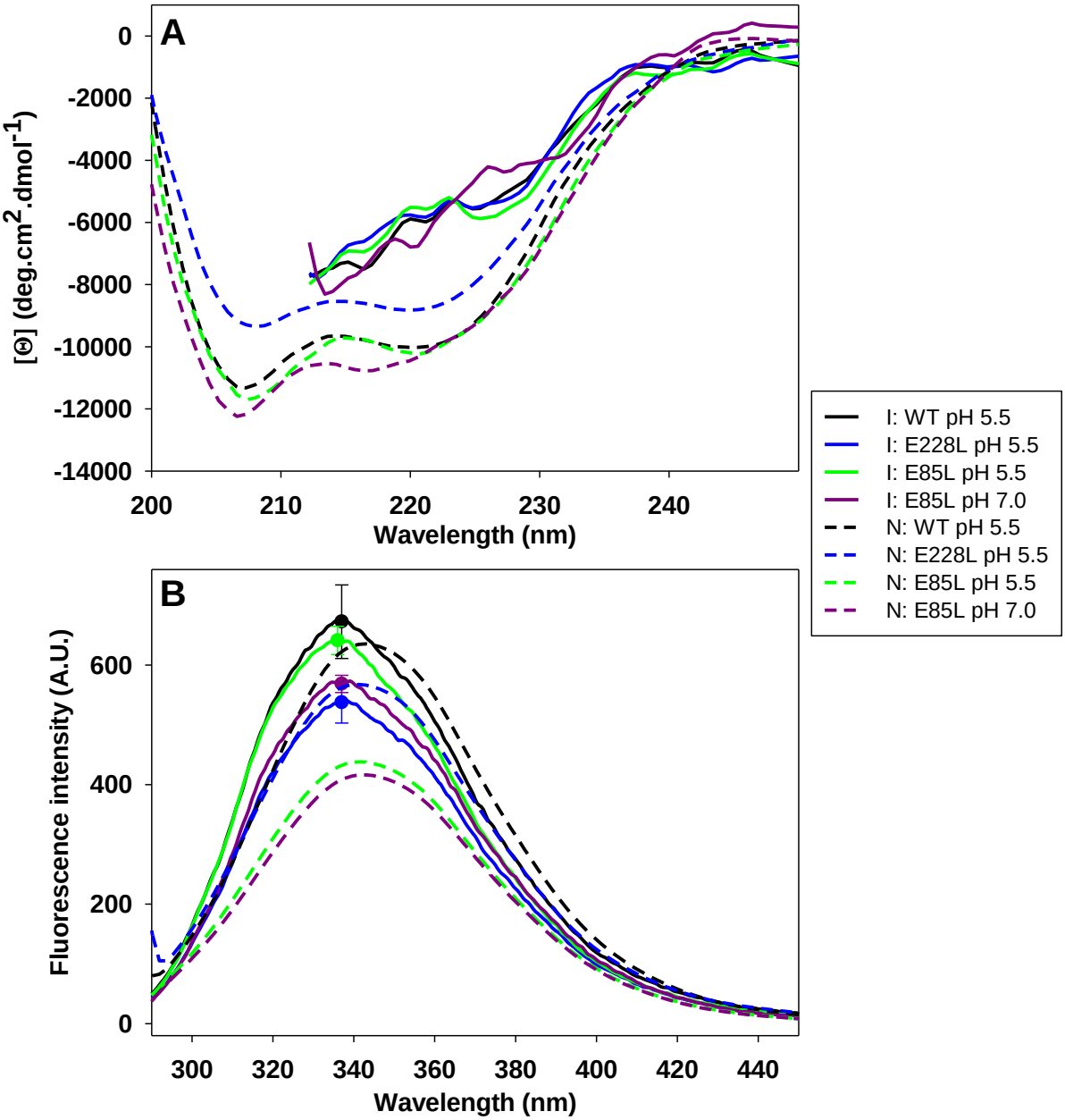


Figure 28. Structure of the intermediate state

The secondary and tertiary structures of the native (N) and intermediate (I) states of wild-type, E228L CLIC1 and E85L CLIC1s were examined using (A) far-UV circular dichroism and (B) intrinsic fluorescence. Spectra of 2 μM CLIC1 were measured in triplicate at 20 °C and averaged. The variations in the fluorescence intensity at λ_{max} values of the intermediate states are shown as individual points with error bars in (B). The legend shown on the right applies to both A and B. The intermediate states formed by wild-type and the two mutant CLIC1 proteins show a decrease in secondary structure relative to the native state and a blue-shifting of the fluorescence peak from 343 to 337 nm, suggesting that the Trp35 probe is in a more hydrophobic environment. The intermediate states formed by the three proteins are similar in both secondary and tertiary structure.

ANS binding spectra for the intermediate states of wild-type, E228L CLIC1 and E85L CLIC1 at pH 5.5 and the E85L CLIC1 intermediate at pH 7.0 show a blue-shifted peak, suggesting that the hydrophobic character of the ANS binding sites is similar in all four structures (Figure 27). E228L CLIC1 and E85L CLIC1 also bind ANS in the native state (Figure 27) and spectra peak at 467 nm and 470 nm, respectively. As these values are the same as those seen for the corresponding intermediate states, they may indicate the formation of a small population of intermediate under native conditions.

Changes to the domain interface upon intermediate formation were investigated using acrylamide quenching of Trp35 fluorescence. Stern-Volmer plots of quenching data (Figure 29) were used to calculate the quenching constant, K_{SV}, of each of the three CLIC1 proteins. This was used to estimate the solvent exposure of the Trp35 residue. K_{SV} values calculated for wild-type in the native state at both pHs (Table 4) agree well with previous data (Fanucchi *et al.,* 2008). The values obtained are similar to those estimated by Eftink and Ghiron (1976) to be the highest possible for a protein-bound tryptophan. Thus, Trp35 is maximally solvent-exposed as predicted by the CLIC1 crystal structure (Harrop *et al.,* 2001). Upon intermediate formation, the wild-type K_{SV} decreases, indicating shielding of Trp35 from the solvent. The same trend is seen for both the variant proteins and their intermediates, indicating that the structural changes that enable intermediate formation are similar in all three CLIC1 proteins.

The K_{SV} values obtained for the intermediate states of wild-type and E228L CLIC1 were 8.29 M^{-1} and 4.93 M^{-1}, respectively. The E85L CLIC1 intermediate states formed at pH 5.5 and pH 7.0 have similar K_{SV} values of 3.76 M^{-1} and 3.61 M^{-1}, respectively. The K_{SV} values for the intermediate states of the variant proteins are lower than wild-type which suggests that Trp35 is more buried in the variant intermediate states than in the wild type. This is unlikely, however, because the Trp35 fluorescence spectra of all the intermediate species peak at the same wavelength (Figure 28). Therefore, the difference in K_{SV} values between wild-type and the variants may be due to the increased population of intermediate present in the variant proteins as shown with ANS (Figure 27). A larger population of intermediate, where the Trp35 is buried, may bias the acrylamide quenching quantum yield data.

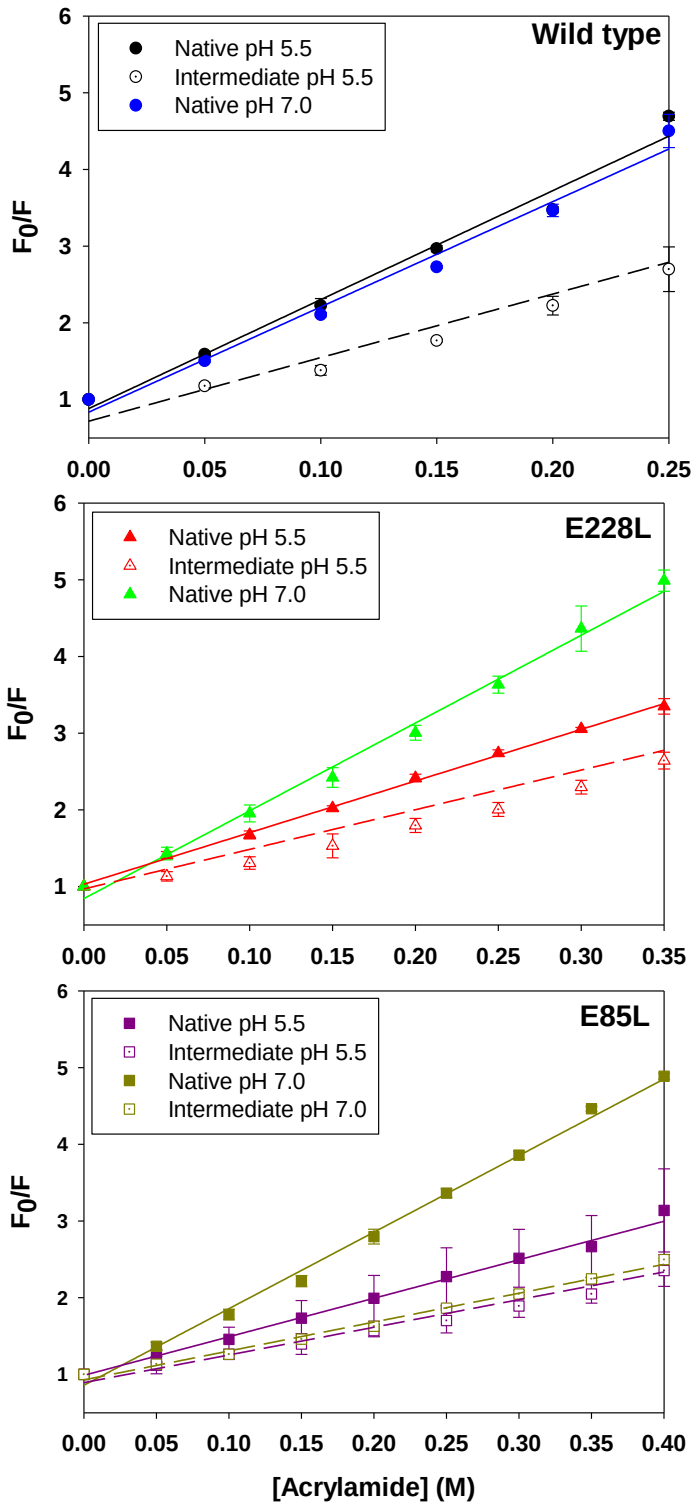


Figure 29. Stern-Volmer plots of wild-type, E228L CLIC1 and E85L CLIC1

F_0/F is the ratio of measured fluorescence intensity in the absence and presence of quencher, respectively. Quenching was performed using acrylamide to quench Trp35 fluorescence of 2 μM CLIC1 protein at pH 7.0 and pH 5.5 (both 50 mM sodium phosphate, 0.02% NaN_3 and 1 mM DTT). Intermediate states were obtained by incubating the proteins in urea to a previously determined concentration: 3.6 M for wild-type and 3.2 M for E228L CLIC1, both at pH 5.5. E85L CLIC1 formed the intermediate state in 3.8 M urea at pH 7.0 and 3.4 M urea at pH 5.5.

Table 4. K_{SV} values for dynamic quenching of tryptophan fluorescence in the native and intermediate states of the various CLIC1 proteins

Native			
CLIC1	**pH**	**K_{sv} (M^{-1})**	**R^2**
WT	7	13.72 (0.87)	0.98
	5.5	14.21 (0.93)	0.98
E228L	7	11.44 (0.29)	0.98
	5.5	6.73 (0.09)	0.99
E85L	7	9.98 (0.14)	0.99
	5.5	5.16 (0.41)	0.86

Intermediate			
CLIC1	**pH**	**K_{sv} (M^{-1})**	**R^2**
WT	5.5	8.29 (0.49)	0.93
E228L	5.5	4.93 (0.17)	0.96
E85L	7	3.76 (0.07)	0.99
	5.5	3.61 (0.16)	0.94

The Stern-Volmer constant, K_{SV} (M^{-1}) expresses the degree of Trp35 exposure to the solvent. Standard errors are shown in parentheses. The observed decreases in the calculated K_{SV} values upon intermediate formation suggest that, in all three CLIC1 variants, formation of the intermediate state shields Trp35 from the solvent and hence, from quenching by acrylamide.

Finally, limited proteolysis was used to examine the flexibility of the intermediate states (Figure 30). The intensity of the wild-type intermediate band does not change with increased incubation time, suggesting that the structure is rigid and compact. In contrast, the bands of the three variant intermediate states decrease in intensity following brief exposure to thermolysin: the E228L CLIC1 intermediate is cleaved after 5 minutes, that of E85L CLIC1 at pH 5.5 after 2 minutes and finally, the pH 7.0 E85L CLIC1 intermediate is cleaved within 20 s. This suggests that the intermediate states formed by the variant proteins are more flexible than that of the wild type, despite being structurally similar.

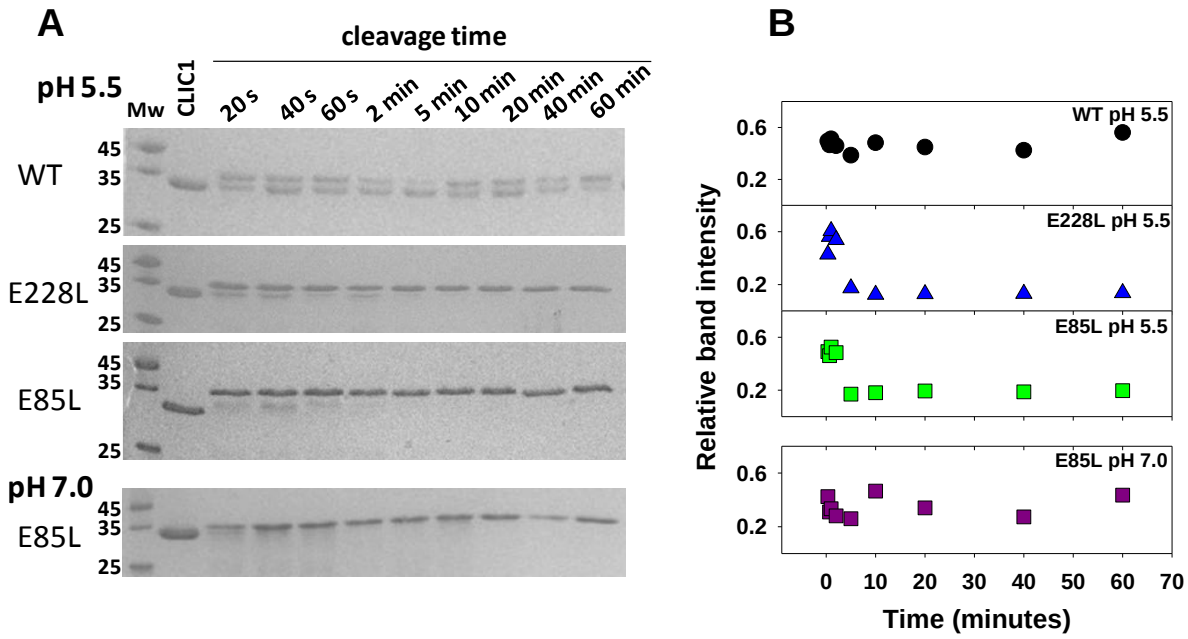


Figure 30. Conformational flexibility of the intermediate state

The intermediate states of wild-type, E228L CLIC1 and E85L CLIC1 at both pH 7.0 and pH 5.5 were subjected to proteolysis by thermolysin for increasing periods of time at 20 °C.
(A) Samples were analysed on 15% SDS-PAGE gels and stained with Coomassie Brilliant Blue R-250. In all gels, the top band is thermolysin (Mw 34.6 kDa) and the lower band, if visible, is CLIC1 (Mw 26.9 kDa). The first lane is the molecular mass marker. The CLIC1 control, in which buffer was added in place of thermolysin is in the second lane. The remaining lanes show samples with increasing times of exposure to thermolysin.
(B) The relative band intensities of the CLIC1 bands at each pH were measured by densitometry and plotted. Intensities are arbitrary and are not comparable between proteins. Intensities in the intermediate states of the variants decrease over time but not in the wild type. This suggests a more rigid intermediate state is formed by wild-type.

CHAPTER 4. DISCUSSION

4.1. Structural roles played by Glu228 and Glu85

Before the possible effect of Glu228 and Glu85 on any pH-induced changes to CLIC1 can be examined, the effect of the E228L and E85L substitutions on the CLIC1 structure must be analysed to determine if they are structurally disruptive. Circular dichroism data (Figure 9) indicates that the secondary structure of CLIC1 remains intact upon substitution of the Glu228 and Glu85 residues with leucine. Similarly, the maximum emission wavelengths of intrinsic fluorescence (Figure 10) do not differ significantly, indicating that the local environment of Trp35 – the domain interface – largely remains unchanged. Hypochromic fluorescence shifts observed for E228L CLIC1 at low pH (Figure 10) are likely a result of minor local reorientations of charged residues that are located in close proximity to Trp35 and capable of quenching tryptophan fluorescence. Such structural rearrangements are possibly facilitated by the increased flexibility observed for the E228L variant at low pH (Figure 12). The crystal structure of E228L CLIC1 confirms these observations as the variant structure overlays with that of wild type, with a backbone rmsd of 0.53 Å (Figure 17). Thus, the global CLIC1 structure has not been altered. However, closer examination of local structure around the site of the E228L substitution indicates some local alterations. Figure 17 shows a reorienting of the Asp225 residue and its salt bridge partner, Lys20, away from the hydrophobic side chain of Leu228. This causes an increase in the distance between the Asp225 and Lys20 R-groups. As the observed distance of 3.4 Å is below the 4 Å cut-off distance for salt bridges (Jeffrey, 1997), this bond presumably still stabilises the structure but is likely to be of significantly lower strength.

Given the proposed role of pH in CLIC1 channel formation (Tulk et al., 2002; Warton *et al.,* 2002), the global structures of the soluble form of wild-type, E228L CLIC1 and E85L CLIC1 were also analysed over an expanded pH range. The secondary and tertiary structures of the three proteins do not differ between pH 4 and 8 (Figure 11). This confirms the non-disruptive nature of the substitutions and suggests that pH changes do not affect membrane insertion through structural changes but rather, through altered conformational stability and flexibility of CLIC1, as proposed by Fanucchi *et al.* (2008) and Stoychev *et al.* (2009).

Wild-type CLIC1 demonstrates increased conformational flexibility at pH 5.5 (Stoychev *et al.* (2009). Analysis of the wild-type crystal structure (Harrop *et al.,* 2001) suggests that the Glu228-Ser16 and Glu85-Lys37/Leu96 interactions may contribute to connectivity between the two

domains and stabilise the N-domain interface, respectively. Loss of these bonds would therefore result in increased flexibility of the CLIC1 structure. The flexibilitiy of the two variant CLIC1 proteins were analysed by limited thermolysin proteolysis at both pH 7.0 and pH 5.5 and compared to the wild type. The band intensities at pH 7.0 remain unchanged as the cleavage time increases, indicating that all three proteins remain structured and relatively rigid at pH 7.0. In contrast, the decrease in the intensities of the CLIC1 bands at pH 5.5 suggests that the CLIC1 proteins are more susceptible to cleavage and hence, are more flexible at low pH. This agrees with observations of Stoychev *et al.* (2009). The E228L CLIC1 protein appears to be of greater flexibility than wild type CLIC1 at pH 5.5, as it is cleaved within a shorter period of time (Figure 12). This observation supports its proposed role of connecting the N- and C-domains in the soluble native state. E85L CLIC1 does not share this increase in flexibility, possibly due to the more local nature of the Glu85-Lys37 bond. While the Glu228-Ser16 bond connects the domains, the Glu85-Lys37 bond maintains the internal structure of the N-domain. Thus, loss of Glu228 would result in increased global flexibility of the CLIC1 domains relative to one another, whereas loss of Glu85, would increase only the local flexibility of the N-domain, which may not be detectable by low-resolution methods such as proteolysis.

Greater flexibility of the CLIC1 structure is suggested to represent increased susceptibility to conformational change (Stoychev *et al.* (2009). The increased flexibility of E228L CLIC1 may therefore indicate that the variant has increased propensity for membrane insertion. Future liposome-binding studies currently in development in our laboratory would allow this hypothesis to be corroborated.

4.2. Substitution of Glu228 or Glu85 reduces conformational stability and promotes intermediate formation

The thermal stability of wild-type CLIC1 is pH-dependent as the temperature at which the protein unfolds decreases with a decrease in pH. This pH-effect is also seen in both CLIC1 variants, indicating that the pH-dependence of thermal unfolding is retained. While both variants show lower thermal stability than wild-type CLIC1, E85L CLIC1 is more strongly destabilised than E228L CLIC1 (Figure 19). This difference in thermal stability between the two variant CLIC1 proteins probably correlates with the number and nature of bonds that have been broken in each variant. In E228L CLIC1, the Glu228-Ser16 H-bond has been broken and the Asp225-Lys20 salt bridge has been weakened as demonstrated by the crystal structure (Figure 17). In E85L CLIC1, the Glu85-Lys37 salt bridge and Glu85-Leu96 H-bond are predicted to have both been broken. In the absence

of a crystal structure, this may be confirmed by the changes in stability and unfolding behaviour observed for the E85L variant (Figure 25) which suggest that the mutation has altered the protein. A strong correlation between the thermal stability and the number of salt bridges in a protein has recently been demonstrated by Chan *et al.* (2011). Thus, while the substitutions of Glu228 and Glu85 both cause the loss of an H-bond, the resultant decrease in thermal stability is greater in E85L CLIC1, as an additional salt bridge interaction has been broken.

In order for unfolding data to be fitted and used to calculate thermodynamic parameters, unfolding must be shown to be reversible (Pace, 1986). While the secondary and tertiary structure of the native states of both E228L CLIC1 and E85L CLIC1 at pH 7.0 and pH 5.5 are recovered upon refolding (Figures 21 and 22), the process demonstrates folding hysteresis at low denaturant concentrations. This is likely due to shifts in the urea concentration at which the ANS-binding intermediate forms (Figures 21 and 22, F340), which alters the refolding pathway. This has also been observed in a green fluorescent protein variant (sfGFP), where hysteresis is attributed to trapping of the refolding protein in a 'native-like intermediate state' (Andrews *et al.*, 2007). Thus, the E228L and E85L substitutions appear to have altered the refolding pathway of CLIC1. Unfolding is therefore irreversible and accurate thermodynamic parameters can regrettably not be determined for either variant.

The decreased stability at low pH that is shown by thermal unfolding (Figure 19) is also seen in urea-induced equilibrium unfolding experiments of wild-type CLIC1. At pH 5.5, it unfolds in at lower urea concentrations than at pH 7.0, indicating destabilisation. This same behaviour is observed for E228L CLIC1 and E85L CLIC1 (Figures 23 and 25). Additionally, at both pH 7.0 and pH 5.5 the variant proteins unfold at lower denaturant concentrations than wild type, indicating destabilisation. This further supports the roles of the two glutamate residues in maintaining the native structure of soluble CLIC1.

Wild-type, E228L CLIC1 and E85L CLIC1 all form an ANS-binding intermediate state at pH 5.5. However, the variant proteins give rise to larger populations of intermediate that are detectible at lower urea concentrations (Figures 25 and 27). This suggests that removal of the Glu228-Ser16 and Glu85-Lys37/Leu96 bonds has lowered the energy barriers to intermediate formation. E228L CLIC1 shows an ANS-binding peak at pH 7.0 (Figure 27), suggesting that the protein is also able to form intermediate at neutral pH. This was, however, not detected by intrinsic fluorescence (Figure 25). Wild-type CLIC1 *is* able to form an intermediate at pH 7.0 at 20 °C, but it is present at such a

low concentration that it is not detectible by spectroscopic probes (ANS included) (Stoychev, 2008) unless the temperature is increased to 37 °C (McIntyre, 2007). Similarly, M32A CLIC1 and H207F CLIC1 variants form intermediate at pH 7.0 and 20 °C but only due to structural destabilisation (Stoychev 2008; Achilonu *et al.*, 2012). The introduction of a hydrophobic leucine residue in E228L CLIC1 may have created a hydrophobic patch to which ANS is able to bind. This is unlikely, however, as E228L CLIC1 does not bind ANS in the native state (0 M urea, Figure 20). Thus, the observed F460 peak is certainly indicative of intermediate formation at pH 7.0, but it is more likely a result of structural destabilisation than the removal of pH-dependence. It was thus excluded from later characterisation of the intermediate states.

Low pH is a requirement for the formation of the intermediate state of wild-type CLIC1. This is not the case for E85L CLIC1 which forms an intermediate state at pH 7.0 that is detectible by ANS binding (Figure 27) and, unlike E228L CLIC1, also by intrinsic fluorescence (Figure 25). Additionally, the intermediate is detected at the same urea concentrations at pH 7.0 as the wild-type intermediate at pH 5.5 (Figure 27). Thus, intermediate formation by E85L CLIC1 appears to be pH-independent, suggesting that Glu85 may function in pH-induced changes to CLIC1.

4.3. Primed for conformational change: the CLIC1 intermediate state

Is the intermediate state formed by CLIC1 functionally significant in the conversion of soluble CLIC1 to its membrane-bound form? Wild-type CLIC1 is able form intermediate at pH 7.0 at 37 °C (McIntyre, 2007) under mild denaturing conditions. The destabilising effect of urea may be mimicked *in vivo* by a combination of the increased temperature and low pH of the membrane surface or possibly its lower dielectric constant (Honig *et al.*, 1986). Thus, if intermediate formation *in vivo* is a possibility, what clues about its relevance can be gleaned from its structure? A comparison of the wild type and variant intermediates is additionally useful because demonstration of similarity between the four intermediates indicates that the E85L and E228L substitutions have merely promoted intermediate formation and not initiated the formation of a new species that may not be functionally significant.

Circular dichroism spectra (Figure 28) of all four intermediates indicate a loss of α-helical structure upon formation of the intermediate state. This may be due to unfolding of the N-domain in preparation for refolding to a potentially membrane-competent form (Goodchild *et al.*, 2009; Goodchild *et al.*, 2011). The four intermediate states show blue-shifting of the λ_{max} of intrinsic fluorescence (Figure 28), which indicates that Trp35 is present in a more hydrophobic environment

in the intermediate. Additionally, changes to the solvent accessibility of Trp35 that were investigated using K_{SV} values from acrylamide quenching confirm the burial of Trp35 in a hydrophobic environment upon intermediate formation (Table 4). This may be a result of structural rearrangement in the monomer or a result of oligomerisation. CLIC1 is known to be oligomeric when in its membrane-bound channel form (Warton *et al.,* 2002; Goodchild *et al.,* 2011), however, it is not yet clear as to whether the protein oligomerises before or after membrane association.

While the intermediate states of wild-type, E228L CLIC1 and E85L CLIC1 are similar in structure, they differ in flexibility. The wild-type intermediate has a relatively rigid structure, but the E228L CLIC1 and E85L CLIC1 intermediate states show increased flexibility (Figure 30). The increased structural dynamics observed for native CLIC1 at pH 5.5 have been proposed to facilitate the conformational changes required for membrane binding (Stoychev *et al.,* 2009). Thus, increased flexibility in the intermediates of the variant proteins may represent further lowering of the energy barriers to structural rearrangement.

Detection of the intermediate states by ANS binding indicates that the four intermediate states have exposed non-polar residues (Figure 27). The agreement between the λ_{max} values of protein-bound ANS indicates that these exposed binding surfaces are of similar hydrophobicity (Figure 27). Such solvent-exposed hydrophobic patches may potentially facilitate membrane binding and suggest that the intermediate is a molten-globule (Semisotnov *et al.,* 1991). While a molten globule intermediate has been observed for a number of soluble-to-membrane proteins, (Van der Goot *et al.,* 2001) (Section 1.3.), it is not a requirement. The Bcl-x_L apoptotic protein is the most fitting example of this as its insertion into membranes is, like CLIC1, dependent upon pH but it does not form a molten globule state (Thuduppathy and Hill, 2006). As discussed briefly in Secion 1.3, the molten globule is characterised by intact secondary but loosened tertiary structure, which results in a greater hydrodynamic volume and increased access by the solvent to hydrophobic residues (Reviewed in Arai and Kuwajima, 2000). The intermediate states of the wild type and variant CLIC1 proteins bind ANS, indicating exposure of non-polar residues (Figure 27). However, secondary structure is lost upon intermediate formation and blue-shifting of intrinsic fluorescence suggests a more tightly-packed tertiary structure (Figure 28). Thus, the CLIC1 intermediate states are stable and partially-unfolded, but they do not fit all the criteria that define the molten globule state. Rather, they are "molten globule-like" (Fanucchi *et al.,*2008).

The partially-unfolded structures of the wild type and variant CLIC1 intermediate states suggest a potential role for the intermediate in CLIC1 membrane binding. Furthermore, the structural similarities of the four intermediate states indicate that replacement either of Glu228 or Glu85 with leucine has not affected the global structure of the CLIC1 intermediate, but rather, promoted its formation. It is thus likely that the energy barrier to intermediate formation is lowered in the variant protein structures. This is particularly true for E85L CLIC1, where the intermediate states formed at both pH 7.0 and pH 5.5 have the same core structure as that of the wild-type intermediate at pH 5.5.

4.4. Glu228: a team player

The replacement of Glu228 with leucine appears to have affected the stability of the CLIC1 fold. The E228L protein is destabilised relative to wild type and shows increased intermediate formation at pH 5.5. The small population of intermediate formed by E228L CLIC1 at pH 7.0 is a result of destabilisation caused by the E228L substitution. Thus, breaking the Glu228-Ser16 H-bond did not remove the pH-dependence of intermediate formation. Glu228 is therefore unlikely to function as an independent pH-sensor. It cannot, however, be eliminated entirely from the pH-response, as breakage of the H-bond, as would occur in wild type at low pH, does promote intermediate formation and destabilise the native state. Thus, if Glu228 is involved in pH-induced changes to the stability of CLIC1, it is likely to be in concert with other pH-sensitive residues. One such multi-residue H-bond network that breaks with pH (and corresponding protonation) was recently observed in transthyretin (Yokoyama *et al.,* 2012). Here, protonation of a histidine residue by acidification breaks a network of H-bond interactions and destabilises the tetrameric structure of the protein.

4.5. Glu85: a pH-sensor

Substitution of Glu85 with leucine affects both the stability of CLIC1 and, unlike E228L CLIC1, the pH-dependence of intermediate formation. E85L CLIC1 is destabilised relative to wild-type, and able to form the ANS-binding intermediate state at both pH 7.0 and pH 5.5. The structures of both of these intermediates are the same as that of wild-type, though they demonstrate increased structural flexibility. This increased flexibility of the intermediate and that of the E85L CLIC1 native state at acidic pH, likely make the protein more susceptible to conformational changes and thus serve to further emphasise how the removal of the Glu85-Lys37 salt bridge facilitates structural rearrangement.

At pH 5.5, the Glu85-Lys37 (and Glu85-Leu96) bonds of wild-type CLIC1 are proposed to break following protonation of the Glu85 side chain. This, in turn, facilitates the separation of α-helices 1 and 3 in the N-domain, effectively freeing α-helix 1 of the transmembrane region. In E85L CLIC1, these bonds are broken regardless of pH. As the protein demonstrates the pH 5.5 behaviour of wild-type at pH 7.0, Glu85 is likely to function as a pH-sensitive residue. A similarly pH-responsive glutamate is seen in calmodulin, where changes to the ionisation state of Glu31 induce conformational changes (Atilgan *et al.*, 2011; Negi *et al.*, 2012).

4.6 Conclusion

The conversion of soluble CLIC1 to its membrane-bound form is promoted by the low pH it encounters at the surface of a membrane. Low pH renders the soluble structure more susceptible to conformational changes and is thereby said to 'prime' soluble CLIC1 for the structural rearrangements required for membrane binding. The effects of low pH appear to be facilitated by a network of electrostatic interactions formed by glutamate (Glu81 (Legg E'-Silva *et al.*, 2012)) and histidine (His74 and His185 (Achilonu *et al.*, 2012)) residues. At low pH, the side chains of these residues are likely to become protonated. This would effectively break the H-bonds and salt bridges in which these residues are involved and hence, reduce the energy barriers that oppose the required conformational changes. In this study, two additional

conserved glutamate residues, Glu228 and Glu85, were selected and the possible function of each in pH-induced changes to soluble CLIC1 was investigated. Glu228 does not function as an independent pH sensor, but may play a minor role in the network of pH-sensitive residues. In contrast, replacement of Glu85 allows CLIC1 to mimic low pH unfolding behaviour at neutral pH. It can thus be said that Glu85 functions in pH-induced changes to CLIC1. The network of residues responsible for 'priming' CLIC1 for conformational change (Figure 31) is therefore likely to include Glu85 – and to a lesser degree, Glu228.

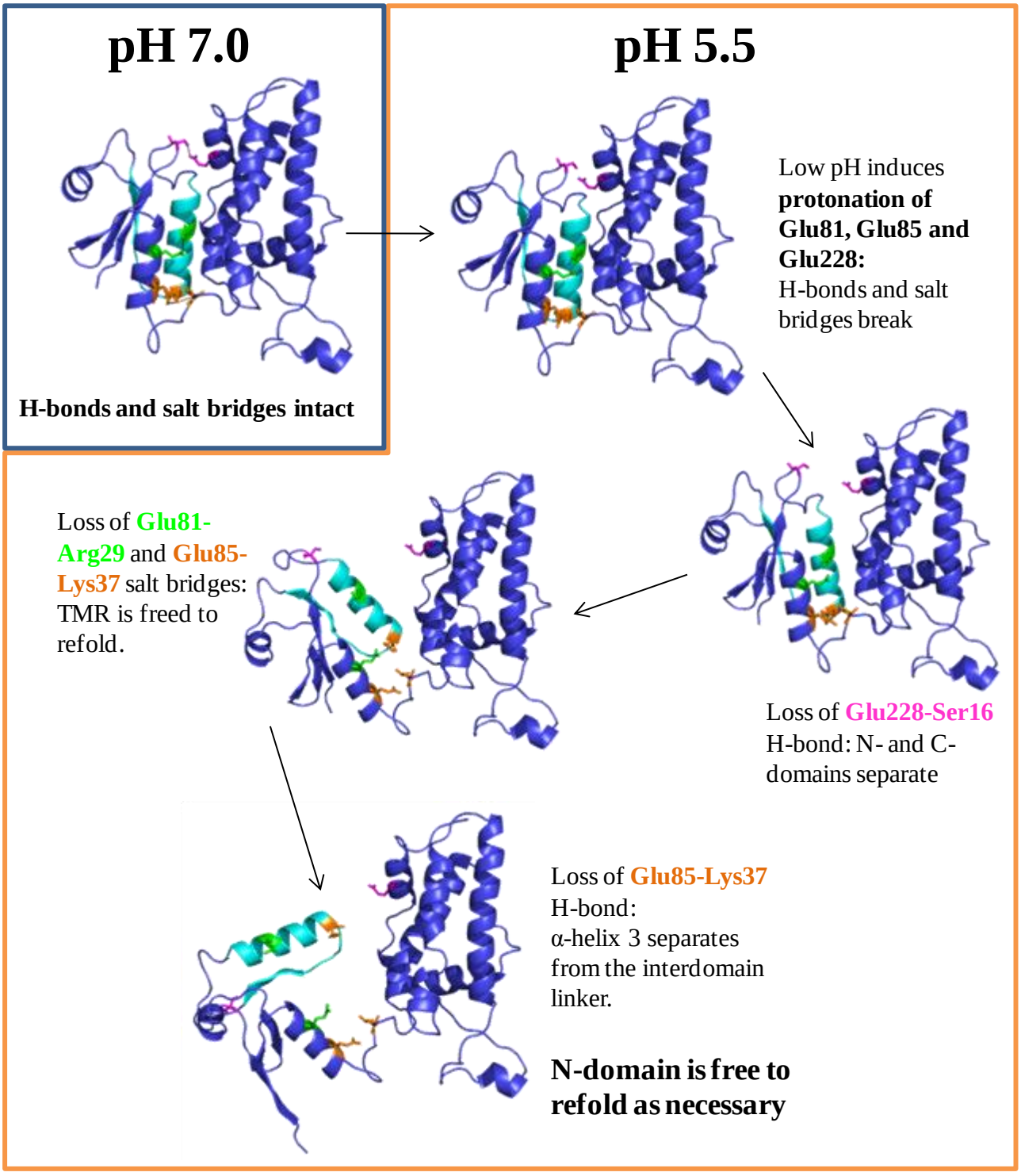


Figure 31. Proposed mechanism for pH-induced 'priming' of CLIC1 for conformational change via glutamate pH-sensors

The transmembrane region (TMR) consists of α-helix 1 and β-strand 2 is shown in cyan, with ionisable residues and their binding partners coloured: Glu228-Ser16 are pink; Glu81-Arg29 are shown in green; Glu85-Lys37/Leu96 are orange. The breakage of bonds is shown sequentially for clarity. *In vivo*, the changes depicted may not occur in the order shown here. Images were rendered using PDB 1K0M (Harrop *et al.*, 2001) and the PyMol Molecular Graphics System, version 1.5 Schrödinger, LLC.

CHAPTER 5. REFERENCES

Achilonu, I., Fanucchi, S., Cross, M., Fernandes, M. and Dirr, H.W. (2012) Role of Individual Histidines in the pH-Dependent Global Stability of Human Chloride Intracellular Channel 1. *Biochemistry* **51,** 995–1004.

Adamson, R.J. (2009) Preparation of soluble CLIC1 and liposomes. *MSc Dissertation*. University of the Witwatersrand, Molecular and Cell Biology. http://wiredspace.wits.ac.za/handle/10539/7074.

Alexov, E.G. and Gunner, M.R. (1997) Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. *Biophys. J.* **72**, 2075–2093.

Andrews, B.T., Schoenfish, A.R., Roy, M., Waldo, G. and Jennings, P.A. (2007) The rough energy landscape of superfolder GFP is linked to the chromophore. *J.Mol. Biol.* **373**, 476–490.

Anfinsen, C.B. (1973) Principles that govern the folding of protein chains. *Science* **181**, 223-230.

Antonsson, B., Conti, F., Ciavatta, A., Montessuit, S., Lewis, S., Martinou, I., Bernasconi, L., Bernard, A., Mermod, J.J., Mazzei, G., Maundrell, K., Gambale, F., Sadoul, R. and Martinou, J.C. (1997) Inhibition of Bax channel-forming activity by Bcl-2. *Science* **277**, 370–372.

Antonsson, B., Montessuit, S., Sanchez, B. and Martinou, J.-C. (2001) Bax is present as a high molecular weight oligomer/complex in the mitochondrial membrane of apoptotic cells. *J. Biol. Chem.* **276**, 11615-11623.

Arai, M. and Kuwajima, K. (2000) Role of the molten globule state in protein folding. *Adv. Protein Chem.* **53**, 209-282.

Atilgan, A.R., Aykut, A.O. and Atilgan, C. (2011) Subtle pH differences trigger single residue motions for moderating conformations of calmodulin. *J. Chem. Phys.* **135**, 15-22.

Averaimo, S., Milton, R.H., Duchen, M.R. and Mazzanti, M. (2010) Chloride intracellular channel 1 (CLIC1): Sensor and effector during oxidative stress. *FEBS Lett.* **10**, 2076–2084.

Baldwin, R.L. (1995) The nature of protein folding pathways: the classical versus the new view. *J. Biomol. NMR* **5**, 103-109.

Bartlett, A.I. and Radford, S.E. (2009) An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat. Struct. Mol. Biol.* **16,** 582-588.

Bell, J.E. (1981) Spectroscopy in Biochemistry Vol I. CRC Press, Inc., Boca Raton, Florida.

Bernal, J.D. and Crowfoot, D. (1934) X-Ray Photographs of Crystalline Pepsin. *Nature* **133**, 794-795.

Berry, K.L. and Hobert, O. (2006) Mapping Functional Domains of Chloride Intracellular Channel (CLIC) Proteins in Vivo. *J.Mol. Biol.* **359**, 1316–1333.

Bertrand, T., Kothe, M., Liu, J., Dupuy, A., Rak, A., Berne, P.F., Davis, S., Gladysheva, T., Valtre, C., Crenne, J.Y. and Mathieu, M. (2012) The Crystal Structures of TrkA and TrkB Suggest Key Regions for Achieving Selective Inhibition. *J.Mol. Biol.* **423**, 439-453.

Boquet, P. and Duflot, E. (1982) Tetanus toxin fragment forms channels in lipid vesicles at low pH. *Proc. Natl. Acad. Sci.USA* **79**, 7614–7618.

Bosshard, H.R., Marti, D.N. and Jelesarov, I. (2004) Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings. *J. Mol. Recognit.* **17**, 1–16.

Boyer, R. (2006) Biochemistry Laboratory: Modern Theory and Techniques. pp 140, 181-186. Benjamin Cummings, San Francisco, USA.

Brünger, A.T. (1992) The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–474.

Bryngelson, J.D., Onuchic, J.N. and Wolynes, P.G. (1995) Funnels, pathways and the energy landscape of protein folding: a synthesis. *Proteins* **21**, 167-195.

Chan, C.-H., Yu, T.-H. and Wong, K.-B. (2011) Stabilizing Salt-Bridge Enhances Protein Thermostability by Reducing the Heat Capacity Change of Unfolding. *PLoS ONE* **6**, e21624. doi:10.1371/journal.pone.0021624

Chen, J. and Stites, W.E. (2001) Packing Is a Key Selection Factor in the Evolution of Protein Hydrophobic Cores. *Biochemistry* **40**, 15280-15289.

Chen, V.B., Arendall III, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D.* **D66,** 12-21.

Choe, S., Bennett, M.J., Fujii, G., Curmi, P.M., Kanterdjieff, K.A., Collier, R.J. and Eisenberg, D. (1992) The crystal structure of diphtheria toxin. *Nature* **357**, 216-222.

Chuang, J. Z., Milner, T. A., Zhu, M. and Sung, C.-H. (1999) A 29 kDa Intracellular Chloride Channel p64H1 Is Associated with Large Dense-Core Vesicles in Rat Hippocampal Neurons. *J. Neurosci.* **19**, 2919–2928.

Collaborative Computational Project Number 4 (1994) The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr. D.* **D50**, 760-763.

Creighton, T.E. (1988) Disulphide bonds and protein stability. *Bioessays* **8**, 57-63.

Cromer, B.A., Morton, C.J., Board, P.G. and Parker, M.W. (2002) From glutathione transferase to pore in a CLIC. *Eur.Biophys. J.* **31**, 356–364.

Dale, G.E., Oefner, C. and D'Arcy, A. (2003) The protein as a variable in protein crystallization. *J. Struct. Biol.* **142**, 88–97.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall III, W.B., Snoeyink, J., Richardson, J.S. and Richardson, D.C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, 375-383.

De Boer, H.A., Comstock, L.A. and Vasser, M. (1983) The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoters. *Proc. Natl. Acad. Sci. USA* **80**, 21-25.

Di Russo, N.V, Estrin, D.A., Martí, M.A. and Roitberg, A.E. (2012) pH-Dependent Conformational Changes in Proteins and Their Effect on Experimental pK_as: The Case of Nitrophorin 4. *PLoS Computational Biology* **8**, e1002761. doi:10.1371/journal.pcbi.1002761.

Dill, K.A., Alonso, D.O.V. and Hutchinson, K. (1989) Thermal Stabilities of Globular Proteins. *Biochemistry* **28,** 5439-5449.

Dill, K.A. (1990) Dominant forces in protein folding. *Biochemistry* **29,** 7133-7155.

Dill, K.A. and Chan, H.S. (1997) From Levinthal to pathways to funnels. *Nature Struct. Biol.* **4**, 10–19.

Dobson, C.M., Sali, A. and Karplus, M. (1998) Protein folding: a perspective from theory and experiment. *Angew. Chem. Int.Edit.* **37**, 868-893.

Dobson, C.M. (2004) Principles of protein folding, misfolding and aggregation. *Semin.Cell Dev. Biol.* **15**, 3-16.

Duncan, R.R., Westwood, P.K., Boyd, A. and Ashley, R.H. (1997) Rat brain p64H1, Expression of a New Member of the p64 Chloride Channel Protein Family in Endoplasmic Reticulum. *J. Biol. Chem.* **272**, 23880–23886.

Eftink, M.R. and Ghiron, C.A. (1976) Exposure of Tryptophanyl Residues in Proteins. Quantitative Determination by Fluorescence Quenching Studies. *Biochemistry* **15**, 672-680.

Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D.* **D60**, 2126-2132.

Fanucchi, S., Adamson, R.J. and Dirr, H.W. (2008) Formation of an Unfolding Intermediate State of Soluble Chloride Intracellular Channel Protein CLIC1 at Acidic pH. *Biochemistry* **47**, 11674-11681.

Feliciangeli, S.F., Thomas, L., Scott, G.K., Subbian, E., Hung, C.-H., Molloy, S.S., Jean, F., Shinde, U. and Thomas, G. (2006) Identification of a pH Sensor in the Furin Propeptide That Regulates Enzyme Activation. *J. Biol. Chem.* **281**, 16108–16116.

Fersht, A.R. (1997) Nucleation mechanisms in protein folding. *Curr. Opin. Struct. Biol.* **7**, 3–9.

Fink, A. L. (1995) Molten Globules. In: Methods in Molecular Biology, Volume 40: Protein Stability and Folding: Theory and Practice (Ed. Shirley, B. A.). pp 343-30. Humana Press Inc., Totowa, New Jersey, USA.

Fontana, A., de Laureto, P.P., Spolaore, B., Frare, E., Picotti, P. and Zambonin, M. (2004) Probing protein structure by limited proteolysis. *Acta Biochim. Pol.* **51**, 299–321.

Fritz, F., Stiasney, K. and Heinz, F.X. (2008) Identification of specific histidines as pH sensors in flavivirus membrane fusion. *J. Cell. Biol* **183**, 353-361.

Garman, E. (1999) Cool data: quantity and quality. *Acta Crystallogr. D.* **D55,** 1641-1653.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D. and Baroch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. In: The Proteomics Protocols Handbook (Ed. Walker, J.M.). pp 571-607. Humana Press, New York.

Gerchman, Y., Olami, Y., Rimon, A., Taglicht, D., Schuldiner, S. and Padan, E. (1993) Histidine-226 is part of the pH sensor of NhaA, a Na^+/H^+ antiporter in Escherichia coli. *Proc. Natl. Acad. Sci.USA* **90**, 1212–1216.

Goodchild, S.C., Howell, M.W., Cordina, N.M., Littler, D.R., Breit, S.N., Curmi, P.M.G. and Brown, L.J. (2009) Oxidation promotes insertion of the CLIC1 chloride intracellular channel into the membrane. *Eur.Biophys. J.* **39**, 129–138.

Goodchild, S.C., Howell, M.W., Littler, D.R., Mandyam, R.A., Sale, K.L., Mazzanti, M., Breit, S.N., Curmi, P.M.G. and Brown, L.J. (2010) Metamorphic Response of the CLIC1 Chloride Intracellular Ion Channel Protein upon Membrane Interaction. *Biochemistry* **49**, 5278–5289.

Goodchild, S.C., Angstmann, C.N., Breit, S.N., Curmi, P.M.G. and Brown, L.J. (2011) Transmembrane extension and oligomerization of the CLIC1 chloride intracellular channel protein upon membrane interaction. *Biochemistry* **50**, 10887–10897.

Greene, R.F.Jr. and Pace, C.N. (1974) Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, α-chymotrypsin and β-lactoglobulin. *J. Biol. Chem.* **249**, 5388-5393.

Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-Pdb viewer: an environment for comparative protein modelling. *Electrophoresis.* **18**, 2714-2723.

Harrop, S.J., Demaere, M.Z., Fairlie, W.D., Reztsova, T., Valenzuela, S.M., Mazzanti, M., Tonini, R., Ru, M., Jankova, L., Warton, K., Bauskin, A.R., Man, W., Pankhurst, S., Campbell, T.J., Breit, S.N. and Curmi, P.M.G. (2001) Crystal Structure of a Soluble Form of the Intracellular Chloride Ion Channel CLIC1 (NCC27) at 1.4-Å Resolution. *J. Biol. Chem.* **276**, 44993–45000.

Hawkins, M., Pope, B., Maciver, S.K. and Weeds, A.G. (1993) Human actin depolymerizing factor mediates a pH-sensitive destruction of actin filaments. *Biochemistry* **32,** 9985–9993.

Henderson, R. (1990) Cryoprotection of protein crystals against radiation damage in electron and X-ray diffraction. *Proc. Roy. Soc. Lond. B Bio.* **241**, 6-8.

Herczenik, E. and Gebbink, M.F.B.G. (2008) Molecular and cellular aspects of protein misfolding and disease. *FASEB J.* **22**, 2115-2133.

Honig, B.H., Hubbell, W.L. and Flewelling, R.F. (1986) Electrostatic interactions in membranes and proteins. *Annu.Rev. Biophys. Biophys. Chem.* **15**, 163-193.

Hubbard, S.J., Eisenmenger, F. and Thornton, J.M. (1994) Modelling studies of the change in conformation required for cleavage of limited proteolytic sites. *Protein Sci.* **3**, 757–768.

Jamin, M. and Baldwin, R.L. (1998) Two forms of the pH 4 folding intermediate of apomyoglobin. *J.Mol. Biol.* **276,** 491-504.

Jeffrey, G.A. (1997) An introduction to hydrogen bonding. pp 192. Oxford University Press, Inc., New York, USA.

Jentsch, T.J., Stein, V., Weinreich, F. and Zdebik, A.A. (2002) Molecular Structure and Physiological Function of Chloride Channels. *Physiol. Rev.* **82**, 503–568.

Karplus, M. and Weaver, D.L. (1979) Diffusion-collision model for protein folding. *Biopolymers* **18**, 1421-1437.

Kauzmann, W. (1959) Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* **14**, 1-63.

Kelley, L.A. and Sternberg, M.J.E. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protoc.* **4**, 363–371.

Kelly, S.M., Jess, T.J. and Price, N.C. (2005) How to study proteins by circular dichroism. *Biochim. Biophys. Acta* **1751**, 119 – 139.

Kloda, A., Ghazi, A. and Martinac, B. (2006) C-terminal Charged Cluster of MscL, RKKEE, Functions as a pH Sensor. *Biophys. J.***90**, 1992–1998.

Koehler, T.M. and Collier, R.J. (1991) Anthrax toxin protective antigen: low-pH-induced hydrophobicity and channel formation in liposomes. *Mol. Microbiol.* **5**, 1501–1506.

Kumar, S. and Nussinov, R. (1999) Salt bridge stability in monomeric proteins. *J.Mol. Biol.* **293**, 1241–1255.

Kumar, S., Tsai, C.J. and Nussinov, R. (2000) Factors enhancing protein thermostability. *Protein Eng.* **13**, 179–191.

Lacroix, E., Viguera, A.R. and Serrano, L. (1998) Elucidating the folding problem of α-helices: local motifs, long-range electrostatics, ionic strength dependence and prediction of NMR parameters. *J.Mol. Biol.* **284**, 173–191.

Lacy, D.B., Wigelsworth, D.J., Melnyk, R.A. Harrison, S.C. and Collier, R.J. (2004) Structure of heptameric protective antigen bound to an anthrax toxin receptor: a role for receptor in pH-dependent pore formation. *Proc. Natl. Acad. Sci.USA* **101**, 13147-13151.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lakowicz, J.R. (1999) Principles of fluorescence spectroscopy, Protein fluorescence. pp 11-14, 188, 237-249, 445-465. Plenum Publishers, USA.

Lathe, G.H. and Ruthven, C.R. (1956) The separation of substances on the basis of their molecular weights, using columns of starch and water. *Biochem. J.* **60**, 665-674.

Legg-E'silva, D., Achilonu, I., Fanucchi, S., Stoychev, S., Fernandes, M. and Dirr, H.W. (2012) Role of arginine 29 and glutamic acid 81 interactions in the conformational stability of human chloride intracellular channel 1. *Biochemistry* **51**, 7854–62.

Legg-E'Silva, D.A. (2011) The role of electrostatic interactions in the stability and structural integrity of human CLIC1. *PhD thesis.* University of the Witwatersrand, Molecular and Cell Biology. http://wiredspace.wits.ac.za/handle/10539/11336.

Lesk, A.M. and Rose, G.D. (1981) Folding units in globular proteins. *Proc. Natl. Acad. Sci.USA* **78,** 4304–4308.

Lesser, G.J. and Rose, G.D. (1990) Hydrophobicity of amino acid subgroups. *Proteins* **8**, 6-13.

Levinthal, C. (1968) Are there pathways for protein folding? *J. Chem. Phys.* **65**, 44-45.

Lide, D.R. (1991) The Handbook of Chemistry and Physics, 72nd edition. CRC Publishers, Boca Raton, Florida, USA.

Lins, L. and Brasseur, R. (1995) The hydrophobic effect in protein folding. *FASEB J.* **9**, 535–540.

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. and 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. and Curmi, P.M.G. (2010) The enigma of the CLIC proteins: Ion channels, redox proteins, enzymes, scaffolding proteins? *FEBS Lett.* **584**, 2093–2101.

Maity, H., Maity, M., Krishna, M.M., Mayne, L. and Englander, S.W. (2005) Protein folding: the stepwise assembly of foldon units. *Proc. Natl. Acad. Sci.USA* **102**, 4741–4746.

Maki, K., Ikura, T., Hayano, T., Takahashi, N. and Kuwajima, K. (1999) Effects of proline mutations on the folding of staphylococcal nuclease. *Biochemistry* **38**, 2213-2223.

Matthews, B.W. (1968) Solvent content of protein crystals. *J.Mol. Biol.* **33**, 491–497.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) ***Phaser*** crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674.

McIntyre, S. (2007) Effects of the environment on the conformational stability of the chloride intracellular channel protein CLIC1. *PhD thesis*. University of the Witwatersrand, Molecular and Cell Biology. http://wiredspace.wits.ac.za/handle/10539/4853.

McPherson, A. (1990) Current approaches to macromolecular crystallization. *Eur. J. Biochem.* **189**, 1–23.

Mere, J., Morion-Guyot, J., Bonhoure, A., Chiche, L. and Beaumelle, B. (2005) Acid-triggered membrane insertion of Pseudomonas exotoxin A involves an original mechanism based on pH-regulated tryptophan exposure. *J. Biol. Chem.* **280,** 21194-21201.

Minn, A.J., Velez, P., Schendel, S.L., Liang, H., Muchmore, S.W., Fesik, S.W., Fill, M. and Thompson, C.B. (1997) Bcl-x_L forms an ion channel in synthetic lipid membranes. *Nature* **385**, 353-357.

Mirsky, A.E. and Pauling, L. (1936) On the structure of native, denatured, and coagulated proteins. *Proc. Natl. Acad. Sci.USA* **22**, 439–447.

Möller, M. and Denicola, A. (2002) Protein Tryptophan Accessibility Studied by Fluorescence Quenching. *Biochem. Mol. Biol. Edu.* **30**, 175-178.

Morihara, K., Tsuzuki, H. and Oka, T. (1968) Comparison of the specificities of various neutral proteinases from microorganisms. *Arch. Biochem.Biophys.* **123**, 572-588.

Muga, A., Gonzalez-Manas, J.M., Lakey, J.H., Pattus, F. and Surewicz, W.K. (1993) pH-dependent Stability and Membrane Interaction of the Pore-forming Domain of Colicin A. *J. Biol. Chem.* **268**, 1553–1557.

Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. D.* **D53**, 240-255.

Murzin, A.G. (2008) Metamorphic proteins. *Science* **320**, 1725-1726.

Myers, J.K. and Pace, C.N. (1996) Hydrogen bonding stabilizes globular proteins. *Biophys. J.* **71**, 2033–2039.

Myers, J.K., Pace, C.N. and Scholtz, J.M. (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–48.

Na, Y.-R. and Park, C. (2008) Investigating protein unfolding kinetics by pulse proteolysis. *Protein Sci.* **18**, 268–276.

Nathaniel, C. (2007) Structural dynamics and membrane interactions of the chloride intracellular channel protein CLIC1. *PhD thesis*. University of the Witwatersrand, Molecular and Cell Biology. http://wiredspace.wits.ac.za/handle/10539/4561

Negi, S., Atilgan, A.R., and Atilgan, C. (2012) Driving Calmodulin Protein towards Conformational Shift by Changing Ionization States of Select Residues. *J. Phys. Conference Series* **402**, 1-10.

Novarino, G., Fabrizi, C., Tonini, R., Denti, M.A., Malchiodi-albedi, F., Lauro, G.M., Sacchetti, B., Paradisi, S., Ferroni, A., Curmi, P.M., Breit, S.N. and Mazzanti, M. (2004) Involvement of the Intracellular Ion Channel CLIC1 in Microglia-Mediated β-Amyloid-Induced Neurotoxicity. *J. Neurosci.* **24**, 5322–5330.

O'Sullivan, C. and Tompson, F.W. (1890) Invertase: A Contribution to the History of an Enzyme or Unorganized Ferment. *J. Chem. Soc. Trans.* **57**, 834-931.

Ohta, Y., Ogura, Y. and Wada, A. (1966) Thermostable Protease from Thermopbilic Bacteria. Thermostability, physicochemical properties and amino acid composition. *J. Biol. Chem.* **241**, 5919–5925.

Onuchic, J.N. and Wolynes, P.G. (2004) Theory of protein folding. *Curr. Opin. Struct. Biol.* **14**, 70-75.

Pace, C.N. (1975) The stability of globular proteins. *CRC Crit. Rev. Biochem.Mol. Biol.* **3**, 1-43.

Pace, C.N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Method. Enzymol.* **131**, 266-280.

Pace, C.N., Shirley, B.A., McNutt, M., and Gajiwala, K. (1996) Forces contributing to the conformational stability of proteins. *FASEB J.* **10**, 75–83.

Pace, C.N. (2001) Polar group burial contributes more to protein stability than nonpolar group burial. *Biochemistry* **40**, 310-313.

Pace, C.N., Grimsley, G.R. and Scholtz, J.M. (2009) Protein ionizable groups: pK values and their contribution to protein stability and solubility. *J. Biol. Chem.* **284**, 13285–13289.

Pace, C.N., Fu, H., Fryar, K.L., Landua, J., Trevino, S.R., Shirley, B.A., Hendricks, M.M., Iimura, S., Gajiwala, K., Scholtz, J.M. and Grimsley, G.R. (2011) Contribution of hydrophobic interactions to protein stability. *J.Mol. Biol.* **408**, 514–528.

Panja, S., Aich, P., Jana, B. and Basu, T. (2008) How does plasmid DNA penetrate cell membranes in artificial transformation process of Escherichia coli? *Mol. Membr. Biol.* **25**, 411–422.

Park, C. and Marqusee, S. (2005) Pulse proteolysis: A simple method for quantitative determination of protein stability and ligand binding. *Nat. Methods* **2**, 207–212.

Pauling, L. (1960) The Nature of the Chemical Bond, 3rd edition. Cornell University Press, Ithaca, New York.

Pauling, L. and Corey, R.B. (1951) Configurations of Polypeptide Chains With Favored Orientations Around Single Bonds: Two New Pleated Sheets. *Proc. Natl. Acad. Sci.USA* **37**, 729-740.

Pauling, L., Corey, R.B. and Branson, H.R. (1951) The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci.USA* **37**, 205-211.

Pautsch, A., Vogt, J., Model, K., Siebold, C. and Schulz, G.E. (1999) Strategy for membrane protein crystallization exemplified with OmpA and OmpX. *Proteins* **1**, 167–172.

Perkins, S.J. (1986) Protein volumes and hydration effects. *Eur.J. Biochem.* **319**, 749-754.

Petsko, G.A. and Ringe, D. (2004) Protein Structure and Function. pp 10-13, 20. New Science Press, Ltd., London, UK.

Pope, B.J., Zierler-Gould, K.M., Kuhne, R., Weeds, A.G. and Ball, L.J. (2004) Solution structure of human cofilin: actin binding, pH sensitivity, and relationship to actin- depolymerizing factor. *J. Biol. Chem.* **279,** 4840–4848.

Porath, J. and Flodin, P. (1959) Gel filtration: A method for desalting and group separation. *Nature* **183**, 1657-1659.

Privalov, P.L. (1979) Stability of proteins: small globular proteins. *Adv. Protein Chem.* **33**, 167-241.

Qiu, M.R., Jiang, L., Matthaei, M., Schoenwaelder, S.M., Kuffner, T., Mangin, P., Joseph, J.E., Low, J., Connor, D., Valenzuela, S.M., Curmi, P.M., Brown, L.J., Mahaut-Smith, M., Jackson, S.P. and Breit, S.N. (2009) Generation and characterization of mice with null mutation of the chloride intracellular channel 1 gene. *Genesis* **1**, 1–10.

Ramakrishnan, C. and Ramachandran, G.N. (1965) Stereo-chemical criteria for polypeptide and protein chain conformations. II Allowed conformation for a pair of peptide units. *Biophys. J.* **5**, 909–933.

Rebek, J. (1990) On the structure of histidine and its role in enzyme active sites. *Struct. Chem.* **1**, 129-121.

Richards, F.M. (1977) Areas, volumes, packing and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**, 151-176.

Rosen, C. and Weber, G. (1969) Dimer Formation from 1-Anilino-8-naphthalenesulfonate catalyzed by Bovine Serum Albumin. A New Fluorescent Molecule with Exceptional Binding properties. *Biochemistry* **8**, 3915–3920.

Rossman, M.G. (1990) The Molecular Replacement Method. *Acta Crysallogr. A.* **A46**, 73-82.

Roy, S., Choudhury, D., Aich, P., Dattagupta, J.K. and Biswas, S. (2012) The structure of a thermostable mutant of pro-papain reveals its activation mechanism. *Acta Crystallogr. D.* **D68**, 1591–1603.

Sandvig, K. and Olsnes, S. (1980) Diphtheria Toxin Entry into Cells is Facilitated by Low pH. *J. Cell Biol.* **87**, 828–832.

Sandvig, K., Tønnessen, T.I., Sand, O. and Olsnes, S. (1986) Requirement of a Transmembrane pH Gradient for the Entry of Diphtheria Toxin into Cells at low pH. *J. Biol. Chem.* **261**, 11639–11644.

Santoro, M.M. and Bolen, D.W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* **27**, 8063–8068.

Schechter, I. and Berger, A. (1967) On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Com.* **27**, 157–162.

Schendel, S.L., Xie, Z., Montal, M O., Matsuyama, S., Montal, M. and Reed, J.C. (1997) Channel formation by antiapoptotic protein Bcl-2. *Proc. Natl. Acad. Sci.USA* **94**, 5113–5118.

Schwede, T., Kopp, J., Guex, N. and Peitsch, M.C. (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* **31**, 3381-3385.

Semisotnov, G.V., Rodionova, N.A., Razgulyaev, O.I., Uversky, V.N., Gripas, A.F. and Gilmanshin, R.I. (1991) Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers* **31**, 119–128.

Shortle, D. (1995) Staphylococcal nuclease: A showcase of m-value effects. *Adv. Protein Chem.* **46**, 217–247.

Singh, H. (2010) Two decades with dimorphic Chloride Intracellular Channels (CLICs). *FEBS Lett.* **584**, 2112–2121.

Singh, H. and Ashley, R.H. (2007) CLIC4 (p64H1) and its putative transmembrane domain form poorly selective, redox-regulated ion channels. *Mol. Membr. Biol.* **24,** 41-52.

Slavik, J. (1982) Anilinonaphthalene sulfonate as a probe of membrane composition and function. *Biochim. Biophys. Acta* **694**, 1–25.

Srivastava, J., Barber, D.L. and Jacobson, M.P. (2007) Intracellular pH Sensors: Design Principles and Functional Significance. *Physiology* **22**, 30–39.

Stickle, F.D., Presta, G.L., Dill, A.K. and Rose, D.G. (1992) Hydrogen bonding in globular proteins. *J.Mol. Biol.* **226**, 1143-1159.

Stoychev, S.H. (2008) The role of the domain interface in the stability, folding and function of CLIC1, *PhD thesis*. University of the Witwatersrand, Molecular and Cell Biology. http://wiredspace.wits.ac.za.handle.net/10539/5604.

Stoychev, S.H., Nathaniel, C., Fanucchi, S., Brock, M., Li, S., Asmus, K., Woods, V.L.Jr. and 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.

Thuduppathy, G.R., Craig, J.W., Schon, V.K.A. and Hill, R.B. (2006a) Evidence that Membrane Insertion of the Cytosolic Domain of Bcl-x_L is Governed by an Electrostatic Mechanism. *J.Mol. Biol.* **359**, 1045–1058.

Thuduppathy, G.R., Terrones, O., Craig, J.W., Basañez, G. and Hill, R.B. (2006b) The N-Terminal Domain of Bcl-x_L Reversibly Binds Membranes in a pH-Dependent Manner. *Biochemistry* **45**, 14533–14542.

Tilley, S.J. and Saibil, H.R. (2006) The mechanism of pore formation by bacterial toxins. *Curr. Opin. Struct. Biol.* **16**, 230–236.

Tilley, S.J., Orlova, E.V., Gilbert, R.J., Andrew, P.W. and Saibil, H.R. (2005) Structural basis of pore formation by the bacterial toxin pneumolysin. *Cell* **121**, 247-256.

Tonini, R., Ferroni, A., Valenzuela, S.M., Warton, K., Campbell, T.J., Breit, S.N. and Mazzanti, M. (2000) Characterisation of the NCC27 nuclear ion channel: Comparison of its electrophysiological properties on the nuclear and plasma membranes and inhibition of its conductance in transfected CHO-K1 cells by antibody blockade. *FASEB J.* **14**, 1171-1178.

Tsytlonok, M. and Itzhaki, L. S. (2012) The how's and why's of protein folding intermediates. *Archives of Biochemistry and Biophysics*. In press: http://dx.doi.org/10.1016/j.abb.2012.10.006.

Tulk, B.M. and Edwards, J.C. (1998) NCC27, a homolog of intracellular Cl channel p64, is expressed in brush border of renal proximal tubule. *Am. J. Renal Physiol.***274**, 1140–1149.

Tulk, B.M., Schlesinger, P.H., Kapadia, S.A. and Edwards, J.C. (2000) CLIC-1 Functions as a Chloride Channel When Expressed and Purified from Bacteria. *J. Biol. Chem.***275**, 26986–26993.

Tulk, B.M., Kapadia, S. and Edwards, J.C. (2002) CLIC1 inserts from the aqueous phase into phospholipid membranes, where it functions as an anion channel. *Am. J. Cell. Physiol.* **282**, 1103–1112.

Valenzuela, S.M., Martin, D.K., Por, S.B., Robbins, J.M., Warton, K., Bootcov, M.R., Schofield, P.R., Campbell, T.J. and Breit, S.N. (1997) Molecular Cloning and Expression of a Chloride Ion Channel of Cell Nuclei. *J. Biol. Chem.* **272**, 12575-12582.

Valenzuela, S.M., Mazzanti, M., Tonini, R., Qiu, M.R., Warton, K., Musgrove, E.A., Campbell, T.J. and Breit, S.N. (2000) The nuclear chloride ion channel NCC27 is involved in regulation of the cell cycle. *J. Physiol.* **529**, 541–552.

Van der Goot, F.G., Gonzalez-Manas, J.M., Lakey, J.H. and Pattus, F. (1991) A "molten globule" membrane-insertion intermediate of the pore-forming domain of colicin A. *Nature* **354**, 408-410.

Vaz, W.L., Nisksch, A. and Jahnig, F. (1978) Electrostatic interactions at charged lipid membranes. Measurement of surface pH with fluorescence lipoid pH indicators. *Eur. J. Biochem.* **83**, 299-305.

Venyaminov, S.Y. and Yang, J.T. (1996) Determination of protein secondary structure. In: Circular dichroism and the conformational analysis of biomolecules (Ed. Fasman, G.D.) Plenum Press, New York.

Vishnivetskiy, S.A., Qiuyan, C., Palazzo, M.C., Brooks, E.K., Altenbach, C., Iverson, T.M., Hubbell, W.L. and Gurevich, V.V. (2012) Engineering visual arrestin-1 with special functional characteristics. *J. Biol. Chem.* **288**, 3394-3405.

Vogt, G. and Argos, P. (1997) Protein thermal stability: hydrogen bonds or internal packing? *Fold. Des.* **2**, S40–S46.

Vogt, G., Woell, S. and Argos, P. (1997) Protein thermal stability, hydrogen bonds, and ion pairs. *J.Mol. Biol.* **269**, 631–643.

Warton, K., Tonini, R., Fairlie, W.D., Matthews, J.M., Valenzuela, S.M., Qiu, M. R., Wu, W.M., Pankhurst, S., Bauskin, A.R., Harrop, S.J., Campbell, T.J., Curmi, P.M.G., Breit, S.N. and Mazzanti, M. (2002) Recombinant CLIC1 (NCC27) Assembles in Lipid Bilayers via a pH-dependent Two-state Process to Form Chloride Ion Channels with Identical Characteristics to Those Observed in Chinese Hamster Ovary Cells Expressing CLIC1. *J. Biol. Chem.* **277**, 26003–26011.

Weber, G. and Young, L.B. (1964) Fragmentation of Bovine Serum Albumin by Pepsin. I. The Origin of the Acid Expansion of the Albumin Molecule. *J. Biol. Chem.* **239**, 1415-23.

Whitmore, L., and Wallace, B.A. (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* **89**, 392–400.

Wlodawer, A., Minor, W., Dauter, Z. and Jaskolski, M. (2008) Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS Journal* **275**, 1–21.

Yagawa, K., Yamano, K., Oguro, T., Maeda, M., Sato, T., Momose, T., Kawano, S. and Endo, T. (2010) Structural basis for unfolding pathway-dependent stability of proteins: vectorial unfolding versus global unfolding. *Protein Sci.* **19**, 693–702.

Yang, J.T. (1996) Remembrance of Things Past: A Career in Chiroptical Research. In: Circular Dichroism and the Conformational Analysis of Biomolecules (Ed. Fasman, G.). Plenum Press, New York.

Yokoyama, T., Mizuguchi, M., Nabeshima, Y., Kusaka, K., Yamada, T., Hosoya, T., Ohhara, T., Kurihara, K., Tomoyori, K., Tanaka, I. and Niimura, N. (2012) Hydrogen-bond network and pH sensitivity in transthyretin: Neutron crystal structure of human transthyretin. *J. Struct. Biol.***177**, 283–290.