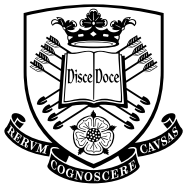

The University Of Sheffield.

Biophysical characterisation and compound screening of disease related pathways in the protein quality control network

By:

Sam Dawes

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

The University of Sheffield

Faculty of Science

School of Chemistry

September 2018

The University Of Sheffield.

University of Sheffield

Biophysical characterisation and compound screening of disease related pathways in the protein quality control network

By Sam Dawes

A thesis submitted in partial
fulfilment of the requirements for
the degree of
Doctor of Philosophy

The University of Sheffield

Faculty of Science

School of Chemistry

Acknowledgements

I would like to thank my supervisors, Professor Beining Chen and Dr. Anastasia Zhuravleva, for their advice and guidance throughout my PhD. I'd also like to thank the latter for help with initial training on how to run NMR experiments and advice on NMR experiment set ups.

Next, I'd like to thank the other members of the Zhuravleva lab, Nik Hurst and Dr, Łukasz Wieteska for their input during group meetings and in the lab. Łukasz Wieteska also produced the NMR spectra in Figure 5.13B.

Dr. James Ault ran and analysed all the mass spectrometry experiments detailed in this thesis. I'd especially like to acknowledge his contribution to the ion mobility mass spectrometry experiments.

Dr. Matthew Baumgartner performed all the initial work to design the compound library in chapter 6. I'd also like to thank various people at Eli Lilly for their advice and support on the work in chapter 6.

I'd like to thank my friends for support during the PhD, but Justin Clarke deserves a special mention for proof reading and advice during writing up.

My parents have provided more love and support than I can begin to thank them for. They've both provided emotional and financial support at times to get me to where I am, and my Dad has provided a great deal of invaluable advice throughout my academic years.

Finally, I'd like to thank Laura. She supported my decision four years ago to leave my job and begin a PhD, despite it meaning she'd be living with a dirty student again. Despite this, in these four years she has become my wife and has been incredibly supportive and understanding when ten more minutes in the lab inevitably turned into another half an hour to an hour. I plan to spend the rest of my life repaying you for this.

Abstract

The majority of secreted proteins are synthesised and folded inside the endoplasmic reticulum (ER). Protein quality control pathways inside the ER ensure aggregation prone and misfolded proteins do not accumulate inside the cell. Breakdown of these pathways leads to numerous diseases, including neurodegenerative diseases, such as Alzheimer's Disease (AD), and various cancers. The unfolded protein response (UPR) is one such pathway that copes with the fluctuating demands inside the ER. Inositol requiring enzyme 1 (IRE1) is the only UPR stress sensor that is conserved throughout all eukaryotes.

Current models suggest that under non-stressed conditions IRE1 is monomeric and associated with the Hsp70 chaperone, BiP, while upon oligomerization of its luminal domain (LD), which is triggered by an interaction with unfolded proteins and/or dissociation from BiP, IRE1 becomes active, which leads to a downstream cascade that results in the upregulation of various elements to aid protein folding and reduction of protein load in the ER.

Several details regarding the activation of IRE1 remain unclear. Conflicting data exists in the literature, with evidence of BiP interacting with IRE1 in a nucleotide dependent manner, via its substrate binding domain, and evidence of a nucleotide independent, non-canonical interaction via its nucleotide binding domain.

In this thesis we have sub-cloned and expressed several constructs of IRE1-LD, characterised the conformational landscape of IRE1-LD in the presence and absence of BiP. We have demonstrated that BiP reduces IRE1 oligomers in a canonical manner. In addition, we have characterised the process of disulphide bond formation, which has previously been shown to be important for prolonged activation of IRE1. We have demonstrated that the nucleotide binding domain of BiP impedes this process.

Furthermore, through the use of 19F NMR, we have identified several compounds that bind to amyloid beta (Aβ) oligomers, widely regarded as the toxic species in AD pathogenesis.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

A – Adenine

A280 – Absorbance at 280 nm

Aβ – Amyloid beta

AD – Alzheimer's Disease

ADP – Adenosine diphosphate

ALS – Amyotrophic lateral sclerosis

ASK1 – Apoptosis signalling kinase 1

ATF6 – Activating transcription factor 6

ATP – Adenosine triphosphate

BAK – Bcl-2 homologous antagonist killer

BAX – Bcl-2 associated X protein

Bcl-2 - Beta cell lymphoma 2

BEST - Band-selective Excitation Short-Transient

BIM – Bcl-2 like protein 11

BiP – Immunoglobulin binding protein

bp – Base pairs

BSA – Bovine serum albumin

CCS – Collision cross section

CD – Circular dichrosim

CPMG – Carr-Purcell-Meiboom-Gill

Da – Daltons

dCTP – Deoxycytosine triphosphate

dGTP – Deoxyguanine triphosphate

DLS – Dynamic light scattering

DNA – Deoxyribonucleic acid

DTT - Dithiothreitol

E. coli – Escherichia coli

EDTA - Ethylenediaminetetraacetic acid

ER – Endoplasmic reticulum

ESI – Electron spray ionisation

EtBr – Ethidium bromide

EtOH - Ethanol

FITC - Fluorescein isothiocyanate

GROMACS - Groningen machine for chemical simulations

HEPES - 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HMK – HEPES, $MgCl_2$ and KCl (see 2.20.1)

HMNa – HEPES, $MgCl_2$ and NaCl (see 2.20.1)

HMQC – Heteronuclear multiple quantum coherence

Hsp – Heat shock protein

HSQC- Heteronuclear single quantum coherence

IM-MS – Ion mobility mass spectrometry

IMAC – Immobilised metal affinity chromatography

INEPT – Insensitive nuclei enhanced by polarisation transfer

IPTG - Isopropyl β-D-1-thiogalactopyranoside

IR - Infrared

IRE1 – Inositol requiring enzyme 1

ITC – Isothermal titration calorimetry

JNK - c-Jun-N-Terminal Kinase

LB - Lysogeny broth

LC – Liquid Chromatography

LD – Luminal domain

MALDI – Matrix assisted laser desorption ionisation

MALLS – Multi angle laser light scattering

MD - Molecular Dynamics

MeCN – Acetonitrile

MHC – Major histocompatibility complex

mRNA – Messenger ribonucleic acid

MS – Mass spectrometry

MST – Microscale thermophoresis

MTSL - (1-oxyl-2,2,5,5-tetramethylpyrroline-3- methyl)methanethiosulfonate

MW – Molecular weight

NBD – Nucleotide binding domain

NMR – Nuclear magnetic resonance

OD400 – Optical density (or absorbance) at 400 nm

OD600 – Optical density (or absorbance) at 600 nm

PBS – Phosphate buffered saline

PCR – Polymerase chain reaction

PDB – Protein data bank

PDIA6 – Protein disulphide isomerase A6

PERK - Protein kinase R-like endoplasmic reticulum kinase

PONDR – Prediction of natural disordered regions

PUMA – p53 regulated modulator of apoptosis

R_h – Hydration radius

RIDD – Regulated IRE1 dependent decay

SDS-PAGE – Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SEC – Size exclusion chromatography

SIM – Selected ion monitoring

SOFAST - band-Selective Optimized Flip-Angle Short-Transient

TAE – Tris acetate EDTA

TCEP - Tris(2-carboxyethyl)phosphine

TEV – Tobacco etch virus

ThT – Thioflavin T

TNF – Tumour necrosis factor alpha

TRAF2- TNF receptor associated factor 2

TRIS - Tris(hydroxymethyl)aminomethane

TROSY - Transverse relaxation optimized spectroscopy

UPR – Unfolded protein response

UV – Ultraviolet

v/v – volume to volume ratio

XBP1 – X-box binding protein 1

1 Introduction

1.1 Protein quality control in the endoplasmic reticulum

The endoplasmic reticulum is the primary organelle in the cell responsible for the vast majority of the synthesis, folding and maturation of secreted proteins. For proteins to perform their designated task, it is often necessary for long, nascent peptide chains to form complex folds over a fluctuating energy landscape. Molecular chaperones interact with unfolded and partially folded intermediates to ensure proteins fold in the correct manner, so they are able to execute their function. Despite this, there are occasions where a misfolded protein cannot be brought back onto the correct folding pathway, under which circumstances the components of the misfolded protein are degraded. This protein quality control pathway ensures that aggregation prone misfolded proteins do not accumulate inside the cell. Unsurprisingly there are numerous diseases associated with the breakdown of this pathway, such as Parkinson's Disease, amyotrophic lateral sclerosis (ALS), Alzheimer's Disease (AD) and numerous cancers (Chen, B. et al., 2011).

1.1.1 The unfolded protein response

The unfolded protein response (UPR) provides a mechanism to control protein load inside the endoplasmic reticulum (ER), by signalling from the ER to the rest of the cell of upon influx of unfolded and misfolded proteins, a condition also termed as ER stress. The response to ER stress is an upregulation of molecular chaperones and repression of translation. At the centre of this network are three proteins; protein kinase R (PKR)-like endoplasmic reticulum kinase (PERK), activating transcription factor 6 (ATF6) and inositol-requiring enzyme 1 (IRE1), each of which have their own distinct pathways (Chakrabarti et al., 2011; Walter and Ron, 2011; Korennykh, A. and Walter, 2012).

IRE1, PERK and ATF6 are all transmembrane proteins with a region in the ER lumen that senses stress and controls activation, and a cytoplasmic region that signal their respective downstream effectors. These three proteins effectively act as communication 'hubs' between the ER and various other organelles in the cell (Figure 1.1).

The activation of these three branches and their subsequent downstream effects is part of the UPR. The result of this is an increase in concentration of chaperones in the ER to cope

with the accumulation of improperly folded proteins and a decrease in protein load in the ER in order to restore the ER back to a non-stressed state (Walter and Ron, 2011).

1.1.2 IRE1

IRE1 is the most conserved of these complexes and is present in all eukaryotes, while PERK and ATF6 are only present in higher-level eukaryotes, such as mammals (Chen, Y.N. and Brandizzi, 2013). In higher eukaryotes IRE1 is divided in to two subforms, IRE1α, which is ubiquitously expressed in tissues and IRE1β, which is only present in the gut (Bertolotti et al., 2001). While there are two subforms, this thesis will focus on the more widely expressed subform, IRE1α, and IRE1 will refer to IRE1α from henceforth.

IRE1 activates when the luminal domain (LD) senses ER stress, which leads to dimerization and oligomerization. This leads to activation of a kinase domain, a domain rearrangement and formation of an active RNase domain in the cytoplasmic region of the protein, which excises a 26-nucleotide intron from the mRNA for X-box binding protein 1 (XBP1) (Chen, Y.N. and Brandizzi, 2013). The spliced mRNA for XBP1 forms a more stable transcription factor which upregulates numerous molecular ER chaperones to increase the folding capacity of the ER (Yoshida et al., 2001). In addition to pro-survival pathways such as this, pro-apoptotic pathways are observed with prolonged activation of IRE1, an example of which is phosphorylation of the c-Jun-N-Terminal Kinase (JNK), with further downstream effects leading to apoptosis (Urano et al., 2000).

IRE1 also targets protein mRNA independently of the XBP1 splicing pathway via a process called regulated IRE1 dependent decay (RIDD). Upon splicing, unlike XBP1, which is then ligated, the targets of RIDD are degraded by exoribonucleases (Hollien and Weissman, 2006). This pathway acts in tandem with the XBP1 pathway by reducing the load of proteins in the ER.

1.1.3 PERK

The mechanism of activation of PERK is similar to that of IRE1; activation is driven by the formation of a homodimer and oligomers in the luminal domain, with the kinase sub-domain in the cytosolic region then undergoing trans-autophosphorylation. The downstream effect of this is the phosphorylation of eiF2α, which then attenuates translation, thereby reducing the folding load in the ER (Walter and Ron, 2011). Much like IRE1, PERK also signals for

apoptosis. Further downstream targets of phosphorylated eiF2α initiate the C/EBP homologous protein (CHOP) apoptotic pathway (Harding et al., 2000).

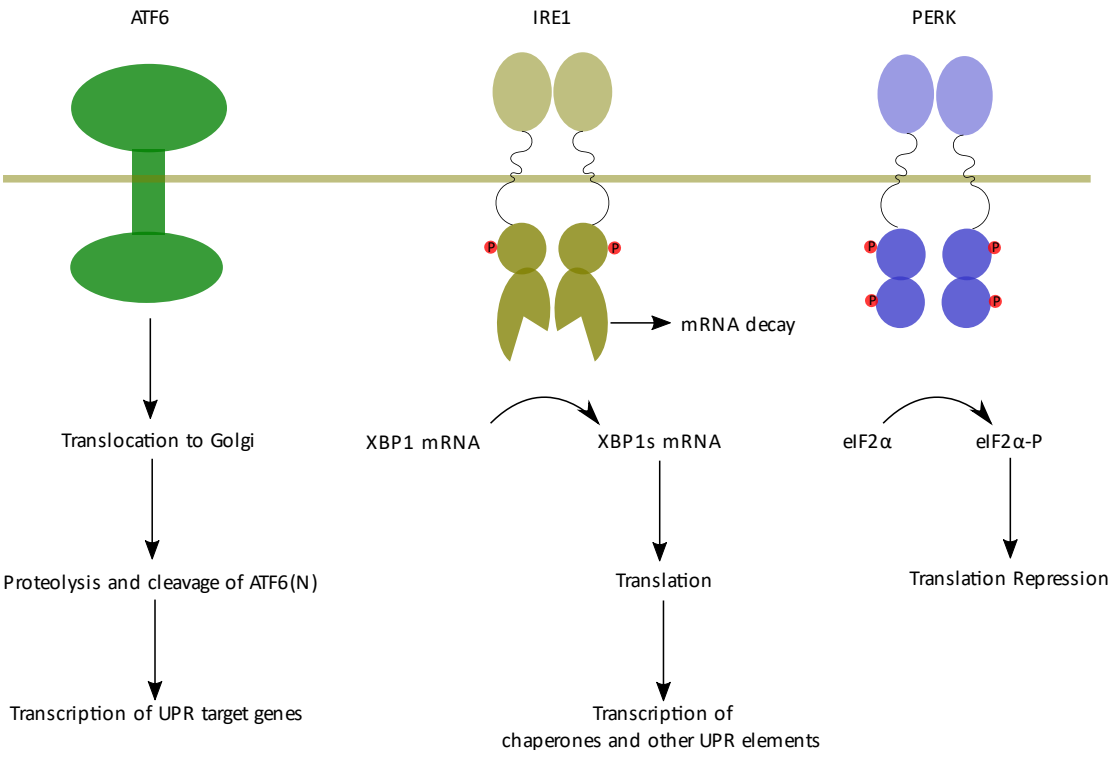

Figure 1.1

(Left) Mechanism of ATF6 activation: ATF6 is relocated to the Golgi body under ER stress, cleaved by S1P and S2P proteases which releases ATF6(N), a transcription factor which upregulates UPR target genes (Centre) Mechanism of IRE1 activation: Under ER stress IRE1 forms a homodimer (and subsequently oligomers) at its luminal domain; the cytoplasmic region undergoes trans-autophosphorylation, which then activates its RNase domain. An intron from XBP1 mRNA is spliced resulting in the formation of XBP1s, a more stable transcription factor which upregulates various chaperones. IRE1 activation also results in RIDD, a pathway that degrades the mRNA of various secretory proteins. (Right) Mechanism of PERK activation: Under ER stress PERK forms a homodimer at its luminal domain, the cytoplasmic region then undergoes trans-autophosphorylation, which then results in the phosphorylation of eIF2α. This results in translational attenuation. Reduction of levels of eIF2 results in an increase of levels of ATF4 that activates the CHOP apoptotic pathway.

1.1.4 ATF6

ATF6 resides in the ER as a transmembrane protein under non-stressed conditions however, under ER stress it translocates to the Golgi body. There it is processed by two different proteases, site-1-protease and site-2-protease (S1P, S2P), which remove the luminal domain fragment. The cytosolic component is then relocated to the nucleus where it acts as a transcription factor that upregulates chaperones such as the ER Hsp70 chaperone, BiP, and glucose-regulated protein 94 (GRP94), a member of the Hsp90 family, and transcription factors such as XBP1 to increase the folding capacity of the ER (Chakrabarti et al., 2011).

This work will focus on IRE1 as this is conserved in all eukaryotes, from yeasts to humans, with an emphasis on its role in disease and the structural mechanisms of activation since comprehension of the latter is required to understand the former and the interactions with the vast network of signalling proteins.

1.2 The misregulation of IRE1 in disease

The pro-survival and pro-apoptotic responses of IRE1 render it critical to cell fate. Pro-survival pathways are utilised by certain cancers and pro-apoptotic pathways are seen in type-2 diabetes and autoimmune diseases. Due to its role as a cell fate regulator, IRE1 has an involvement in numerous disease pathways (Wang and Kaufman, 2012; Hetz et al., 2013). While its role in some pathways is reasonably well understood, its role in other, more complex pathways is generally poorly comprehended. This section will review the role of IRE1 in some of these disease pathways.

1.2.1 Neurodegenerative Disease

Neurodegenerative diseases are characterised by a loss of brain function, due to numerous causes. Protein misfolding is prevalent in many neurodegenerative conditions; for example, improper processing and folding of the amyloid protein causes the formation of amyloid fibrils; an accumulation of insoluble and degradation resistant fibres (Rambaran and Serpell, 2008).

Alzheimer's disease (AD) is characterised by deposits of amyloid beta (Aβ), neurofibrillary tangles of phosphorylated tau and neuronal death. Improper proteolytic processing of amyloid precursor protein (APP), a type I transmembrane protein localised in the ER, leads to the formation Aβ in the ER lumen (Figure 1.2). While APP is ubiquitously expressed in all

tissues, increased levels of expression are seen in brain tissues. Eight isoforms of APP are formed by a variety of splicing pathways, although currently there is no definitive evidence of the function of these isoforms (Placido et al., 2014).

The wider unfolded protein response has been linked with AD in numerous studies (Cornejo and Hetz, 2013; Halliday and Mallucci, 2014). Increased levels of the Hsp70 chaperone, BiP, which is upregulated under ER stress, are observed in the temporal cortex and hippocampus of patients with AD. In addition, significantly increased levels of active PERK are localised in the hippocampus of AD patients. Inhibition of PERK has been shown in mice to improve the prognosis of frontotemporal dementia (Radford et al., 2015).

Considering the link between the UPR and neurodegeneration, it is unsurprising that IRE1 is linked with various neurodegenerative diseases. While earlier work demonstrated a correlation between active IRE1 and AD, such as increased levels of phosphorylated IRE1 in neurons of AD subjects, when compared to levels in control subjects (Hoozemans et al., 2009), a causative link had not been established. More recently, a direct link between IRE1 activation and AD has been demonstrated, with a direct correlation between IRE1 activation and increasing levels of Aβ and APP, which in turn leads to a poorer disease prognosis in mouse models (Duran-Aniotz et al., 2017).

There are numerous examples of IRE1 activation in AD pathologies. Immunostaining of hippocampi of subjects with AD showed phosphorylated, thus activated, IRE1 granules morphologically similar to those of granulovacuolar degeneration (GVD), a characteristic pathological hallmark of AD (Woodard, 1962) that is characterised by the presence of electron dense granules surrounded by a clear zone that are mainly found in hippocampal pyramidal neurons (Okamoto et al., 1991).

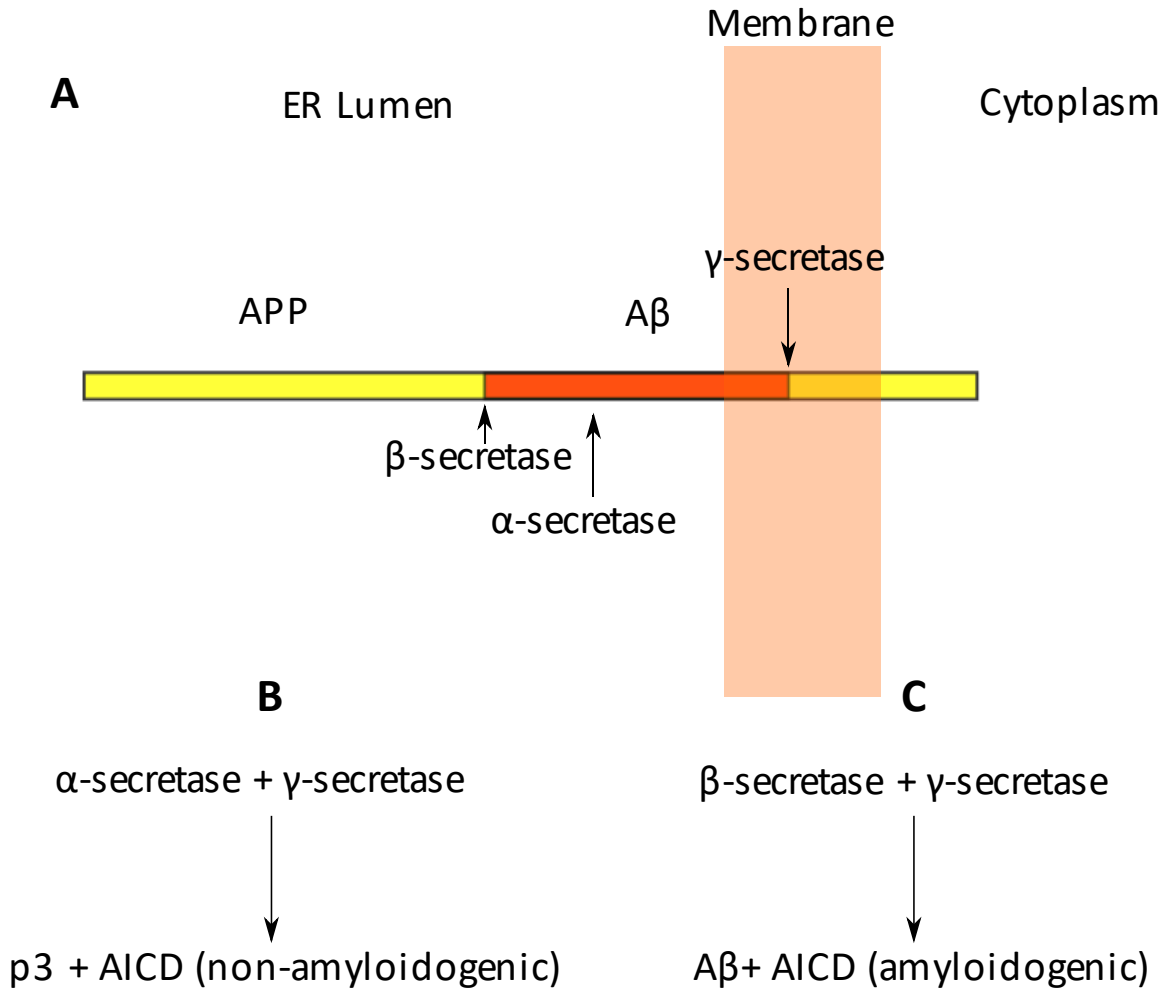

Figure 1.2

Processing of APP. A) Schematic diagram of APP and the sequence of cleavage sites highlighted below. B) The non-amyloidogenic pathway involves cleavage by α-secretase, thereby cleaving amyloidogenic Aβ peptide, followed by γ-secretase cleavage. The role of the cleaved proteins, p3 and AICD is not fully understood. C) The amyloidogenic pathway proceeds with β-secretase, which releases APPsβ. Cleavage by γ-secretase then leads to release of the Aβ peptide.

While AD generally affects those over the age of 65 (Qiu et al., 2009), Familial AD (FAD), a form of AD that is caused by genetic factors and usually observed earlier in life. Specifically, these genetic factors involve mutations of the presenillin proteins, PS1 and PS2, which are catalytic components of the γ-secretase pathway (De Strooper et al., 2012) that is involved in

proteolytic processing of APP. The mutant form of PS1 has been shown to disrupt IRE1 activation and lead to increased levels of cell death in SK-N-SH cells (Katayama et al., 2001). Intriguingly, it has been observed that mutant PS1 associates with IRE1 (Katayama et al., 1999).

Hyperphosphorylation of the intrinsically disordered, cytosol localised, tau protein, is also linked to neurodegenerative disease progression (Bloom, 2014). Experimental evidence suggests a correlation between the presence of active IRE1 and the presence of phosphorylated tau protein in the hippocampus of patients with frontotemporal lobar degeneration with accumulation of tau protein (Nijholt et al., 2012). Further association of ER stress and tau phosphorylation is exhibited by evidence of $A\beta_{1-42}$ activating ER stress and, via glycogen synthase kinase (GSK3β), causing tau phosphorylation (Resende et al., 2008).

In addition to AD, IRE1 has also been implicated in the pathology of other neurodegenerative diseases. **Amyotrophic lateral sclerosis (ALS)** is a disease pathologically characterised by muscle atrophy and weakness and spasticity (Robberecht and Philips, 2013)**.** Various ER stress markers have been detected in patients suffering from ALS (Ilieva et al., 2007) suggesting activation of IRE1 in ALS pathology. The cytoprotective effect of XBP1 possibly aids the progression of ALS with XBP1 and IRE1 knockout models showing decreased levels of mutant superoxide dismutase-1 (SOD-1) aggregation, a key cause of ALS. Instead, increased levels of autophagy-related degradation of SOD-1 mutants were observed in deficient cells (Hetz et al., 2009).

The prion protein (PrP) is predominantly expressed in the nervous system. While little is known about the function of the cellular form of PrP, PrP^{C}, its misfolded isoform, PrP^{Sc}, has been implicated in various transmissible spongiform encephalopathies (TSEs) such as Creutzfeldt-Jakob disease (CJD) (Weissmann, 2004). PrP^{C} that has been present in cells that have undergone ER stress has been shown to be susceptible to conversion to $PrP^{Sc.}$ Furthermore, under ER stress PrP is mistranslocated to the cytoplasm, rather than remaining membrane bound, and accumulates as insoluble aggregates in primary neurons. Overexpression of XBP1s has been shown to reduce this accumulation (Orsi et al., 2006). In addition, increased levels of XBP1 protein reduces the formation of PrP^{C} aggregates when compared to IRE1 and XBP1 knockout cells (Hetz et al., 2007), suggesting a further role of IRE1 in prion disease.

In addition to neurodegenerative disorders, high levels of prion proteins have been observed in numerous cancer cells resistant to treatments, for example a positive correlation between increased levels of PrP mRNA in breast tumour tissues and a poorer prognosis. XBP1s has been shown to increase gene expression of *PRNP*, the gene that codes for PrP, therefore showing a possible link between IRE1 activation and poorer prognosis in breast cancer (Dery et al., 2013).

1.2.2 Cancer

By their definition, cancer cells possess an increased ability to proliferate over healthy cells. Mutations of IRE1 that reduce the pro-apoptotic signals and increase pro-survival signals are present in various cancer cells (Greenman et al., 2007). Melanoma cells have been shown to be resistant to the pro-apoptotic response of ER stress. Rather than attenuating as in healthy cells, IRE1 and ATF6 have been shown to be constitutively active in melanoma cells undergoing ER stress. In addition to this, knockdown of IRE1 and reduction of levels of XBP1s by addition of salicylaldehyde were both shown to increase levels of apoptosis in melanoma cells (Tay et al., 2014). This link is further established in multiple myeloma with enforced expression of XBP1s in B cells in mice models shown to activate multiple myeloma pathways (Carrasco et al., 2007).

IRE1 has also been shown to play a role in tumour growth and proliferation in U87 cells. Expression of a dominant negative IRE1 construct showed a decrease in cell proliferation, which can be attributed to the observation that IRE1 splices the mRNA for SPARC, an extracellular matrix protein whose overexpression is seen to correlate with disease progression in certain cancers, however in glioblastoma has been shown to impede cancer progression in an RIDD manner (Dejeans et al., 2012).

In addition to a direct role in tumour growth in glioma cells, IRE1 has been shown to confer resistance to therapeutic agents, for example the cytotoxicity induced by OSU-03012, a glioblastoma treatment as IRE1 and XBP1 knockdown cells showed an increased susceptibility to OSU-03012 toxicity (Booth et al., 2014).

In summary, the cytoprotective effect that downstream effectors of IRE1 confer upon cells has been shown to be prevalent in numerous cancer cells, making IRE1 a potential target for therapeutic regulation.

1.2.3 Diabetes

As mentioned previously, the role of pancreatic β-cells in producing insulin means activation of the UPR is dependent on blood glucose levels. When the ER cannot cope with the demands of insulin production, the cell undergoes apoptosis. In type II diabetes, the increased demand for insulin results in prolonged activation of the UPR, and subsequently the loss of β-cell function, meaning insulin is no longer produced (Scheuner and Kaufman, 2008).

Overexpressed IRE1 shows non-specific splicing, with insulin mRNA seemingly degraded by IRE1 in COS-7 cells (Lipson et al., 2008). Whether this pathway is independent of the RNase activity of IRE1 or whether non-specific RNA degradation plays a role in apoptotic signalling of IRE1 remains to be seen.

To summarise, both pro-apoptotic and pro-survival pathways of IRE1 have both been shown to be prevalent in a variety of diseases. In order to regulate IRE1, a thorough understanding of the mechanisms of activation and signalling of IRE1 is required to further understand the following points:

- Structural and mechanistic details of the upstream activation of IRE1
- Identify druggable sites of IRE1 that can be exploited as therapeutic targets for chemical regulation

1.3 Upstream activation of IRE1 via its luminal domain

Unsurprisingly, in order to carry out its various functions, IRE1 is a complex transmembrane multidomain protein that spans across the ER membrane. The luminal domain (LD) detects stress and is linked to the single α-helix transmembrane region by a long intrinsically disordered linker. A second intrinsically disordered linker connects the transmembrane region to the cytoplasmic domain (CD) that in turn, can be subdivided into a kinase and RNase domains (Chen, Y.N. and Brandizzi, 2013).

In the absence of ER stress IRE1 is presumably monomeric and inactive. In the presence of unfolded proteins, the monomeric form of IRE1 forms a dimer at its luminal domain, which brings the cytoplasmic domains into close proximity. The kinase domain then undergoes trans-autophosphorylation, which leads to a domain rearrangement and activation of the

RNase domain that splices an intron from XBP1 mRNA in humans and HAC1 mRNA in yeast (Walter and Ron, 2011) (Figure 1.3).

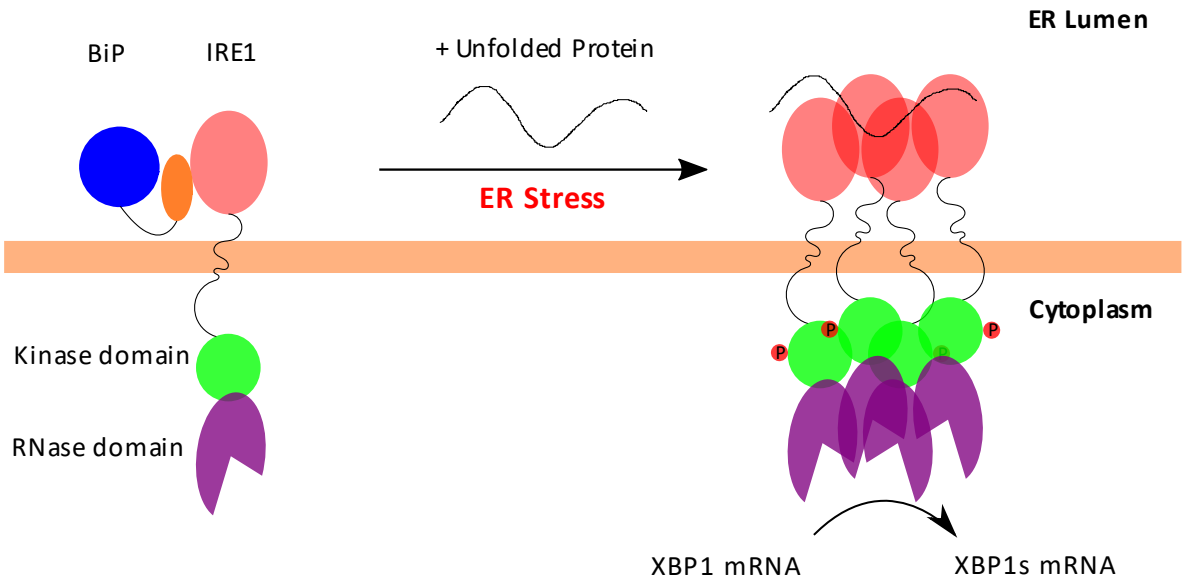

Figure 1.3

Simplified model of the IRE1 activation: The luminal domain drives homodimer and oligomer formation, trans-autophosphorylation occurs at the kinase domain that leads to a conformation change that activates the RNase domain. IRE1-LD is bound to the Hsp70 chaperone BiP (see section 1.4) under non-stressed conditions.

1.3.1 Upstream activation in the yeast homolog of IRE1

As IRE1 is conserved throughout eukaryotes, a substantial amount of earlier experimental work has been carried out in yeast models to elucidate the mechanisms of activation and downstream signalling pathways. Furthermore, early structural studies were carried out in yeast systems (Credle et al., 2005).

The crystal structure of yIRE1 provided useful insights into the mechanism of activation of IRE1, namely that IRE1 dimerises via an anti-parallel beta sheet interface and contains a major histocompatibility complex (MHC)-like binding groove that was hypothesised to bind to unfolded proteins (Figure 1.4). Mutational analysis of the MHC-like binding pocket shows that mutation of M229, F285 and Y301 to alanine results in a significant reduction in yeast IRE1 activity (Credle et al., 2005), suggesting a functional role for this binding pocket. More detailed analysis of the interaction of IRE1 with a mutated, and subsequently constitutively misfolded, form of carboxypeptidase Y (CPY) demonstrated the LD binds in a cooperative manner, suggesting a dimer/oligomerisation event upon binding. Further investigation with a signal peptide, ΔEspP indicated that peptide binding induced oligomerisation, with the authors hypothesising this is driven by a conformational change upon peptide binding. Further mutational analysis of the binding groove using the mutations mentioned previously resulted in a reduction in mRNA splicing and binding affinity (Gardner, B.M. and Walter, 2011).

Following dimer formation, the luminal domain proceeds to form oligomers. While no structure of the luminal domain in its oligomeric state currently exists, an interface has been proposed. It is hypothesised that the luminal domain oligomerises via interactions at an interface opposite the dimer interface (Figure 1.4), with a mutation of W426 (located on the oligomerization interface), to alanine showing a significant reduction in IRE1 activity assays, suggesting that this is part of a functional oligomer interface (Credle et al., 2005). Furthermore, substrate triggered oligomeric species of IRE1 were no longer observed for this mutant, however peptide binding was demonstrated to trigger dimerization, confirming this residue is important for oligomerisation and peptides can induce dimerization of IRE1 (Gardner, B.M. and Walter, 2011).

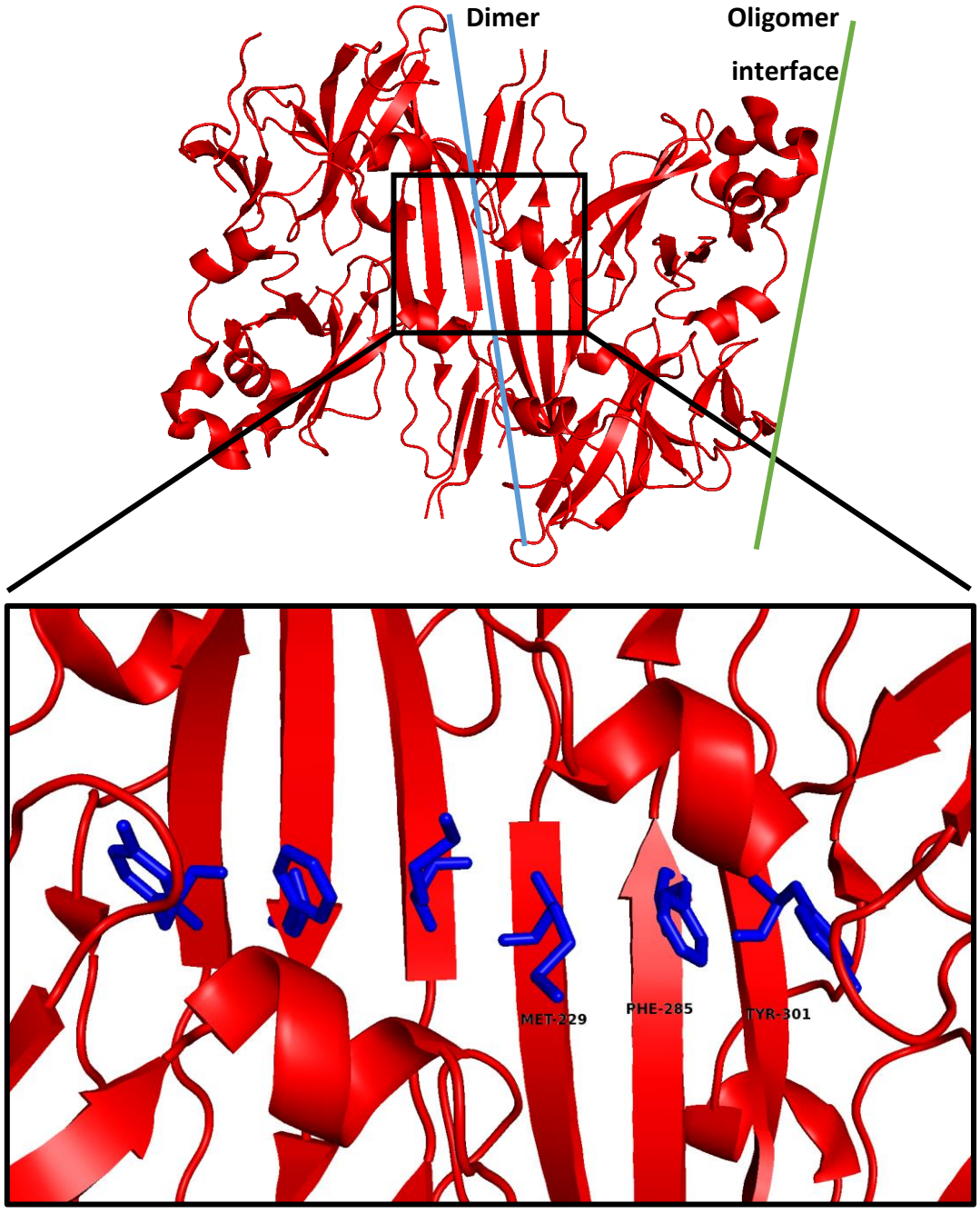

Figure 1.4

Above: Crystal structure of the luminal domain of yeast IRE1 (PDB ID: 2BE1) with the dimeric interface and oligomeric interface indicated by blue and green lines respectively. Below: A zoom of the MHC like binding groove with residues that have been demonstrated to be critical for unfolded substrate binding shown as blue sticks. (Credle et al., 2005). Figure prepared with Pymol (version 1.8)

1.3.2 The mechanism of upstream activation in human IRE1

While the mechanism of upstream activation of yeast IRE1 is relatively well understood, far less is known about the same mechanism in humans, despite some conservation of sequence and structure. Alignment of the structures of the yeast (Credle et al., 2005) and human (Zhou et al., 2006) luminal domains of IRE1 shows a conserved dimerization interface and a potential binding groove, however away from this, few structural elements are conserved (Figure 1.5), notably the well resolved oligomeric interface that is imperative for activation of yeast IRE1 is unresolved in the human crystal structure. Notably, approximately 35% of the hIRE1-LD crystal structure is unresolved, suggesting the human homolog is more flexible than its yeast counterpart.

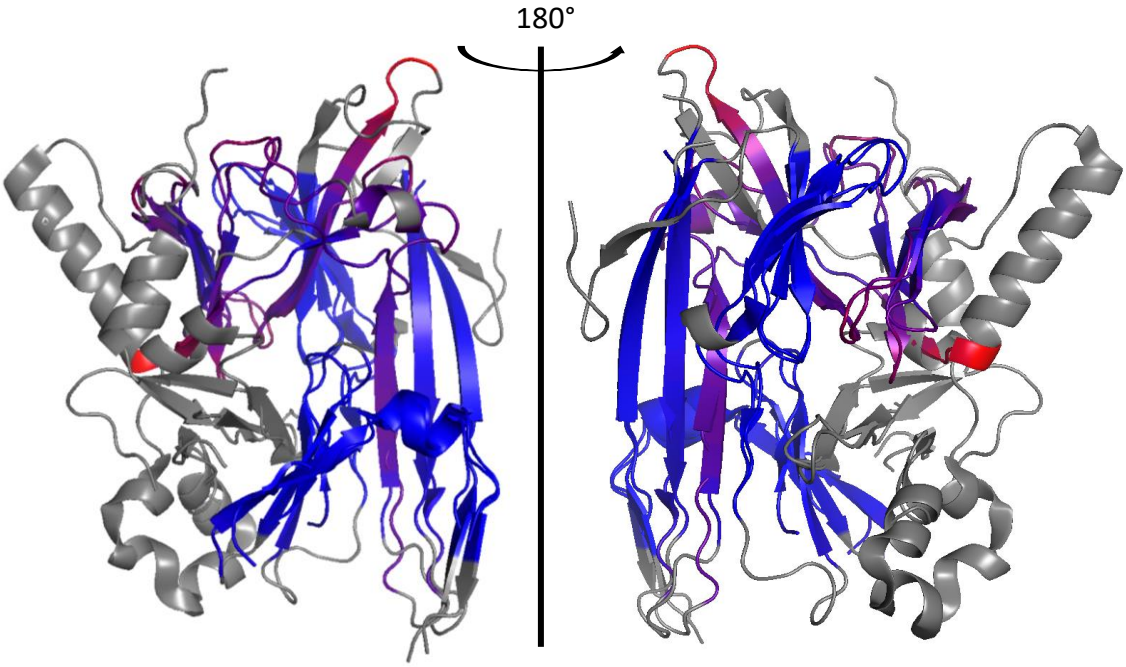

Figure 1.5

Luminal domain of human and yeast IRE1 (2HZ6 and 2BE1). Minimum pairwise RMSD is indicated in blue with maximum in red. Unaligned residues are in grey. Good alignment is observed around the dimer interface while the oligomer interface is not well conserved. Figure prepared using PyMol (version 1.8)

The sole crystal structure of the hIRE1 homolog is in its monomeric state however a dimer interface was inferred by crystal contacts, overlaying the yeast crystal structure and

mutagenesis of residues on the beta sheet interface (Zhou et al., 2006). While the MHC-like groove that's present in the yeast structure is present in the human structure, the authors of the study suggested that the MHC like groove in the human structure was too small to bind unfolded proteins (Zhou et al., 2006), however recent data conflicts with this assertion, as peptide binding to the MHC-like groove has been shown to trigger a conformational change that leads to oligomerisation (Karagoz et al., 2017). Interestingly, only the tyrosine furthermost from the dimerisation interface (Y179 in the human isoform, Y301 in yeast) on the beta-sheet floor of the MHC-like groove is conserved (MFY in yeast and KYY in humans) (Figure 1.6). This coupled with the apparent allosteric effect upon peptide binding, suggested for both yeast (Gardner, B.M. and Walter, 2011) and human (Karagoz et al., 2017) forms of IRE1, suggests that residues other than those identified on the beta-sheet floor are involved in binding to unfolded proteins.

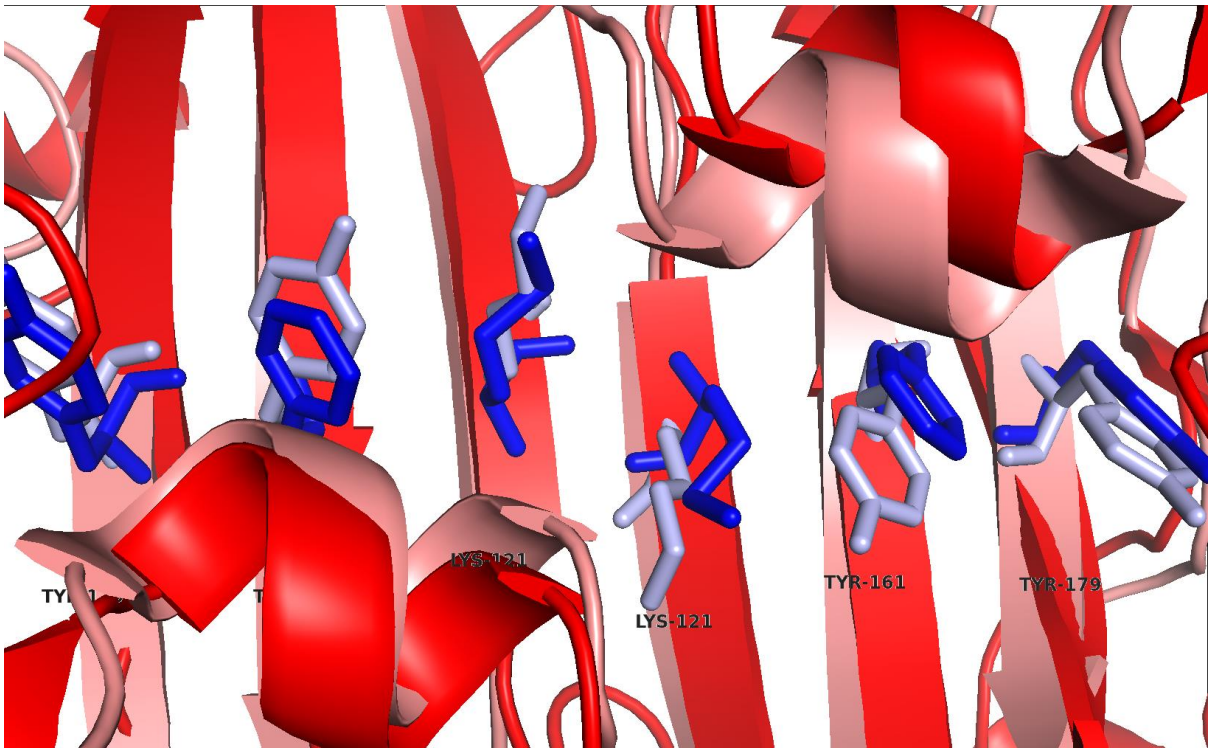

Figure 1.6

The MHC-like groove in hIRE1-LD in pink, aligned with the same region in yIRE1-LD. The residues in the human homolog are represented as pale blue sticks. Only Y179 is conserved between human and yeast homologs. Figure prepared using Pymol (version 1.8)

Unlike the dimer interface, the oligomerization interface of the luminal domain of IRE1 is not conserved between human and yeast (Figure 1.5), with the vast majority of that interface unresolved in the human IRE1 crystal structure. Additionally, present in the unresolved regions of the IRE1-LD crystal structure, are two of three cysteine residues (C148 and C332) whose position are not conserved in yeast. Residues C148 and C332 have been shown to form intermolecular disulphide bonds however, these are not essential for dimerization (Liu et al., 2003). C148 and C332 are present in long (22 and 50 residue respectively) flexible loop regions (Figure 1.7).

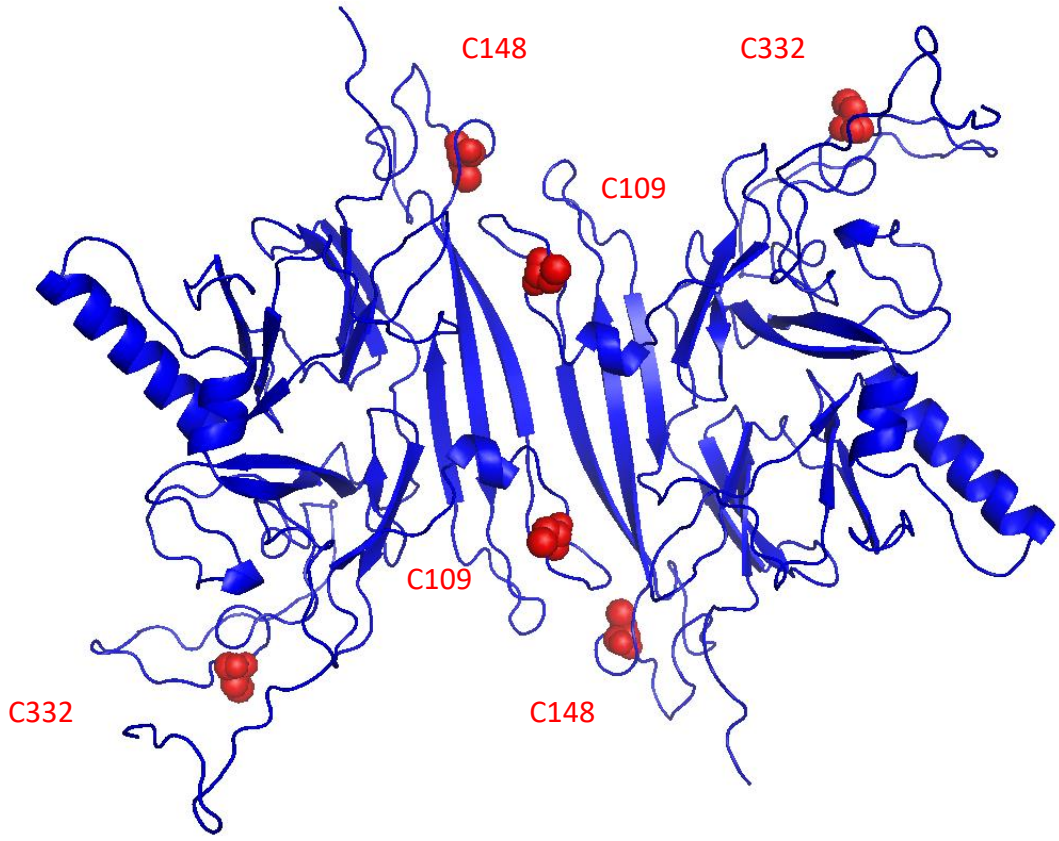

Figure 1.7

Model of luminal domain of IRE1 containing rebuilt loops. Cys residues are labelled and highlighted as red spheres. Model built using crystal structure PDB ID: 2HZ6. Figure prepared using PyMol (version 1.8)

1.4 Regulation of upstream activation of IRE1 by ER resident proteins

1.4.1 Immunoglobulin binding protein (BiP)

As previously indicated, IRE1-LD interacts with the Hsp70 chaperone, BiP. BiP's primary purpose is to act as a molecular chaperone inside the ER. It is a multidomain protein that contains a nucleotide binding domain, with ATPase activity, and a substrate binding domain, which binds to aggregation prone regions of unfolded proteins. There is a significant degree of allostery between these domains, with the ATP binding leading to a 'domain docked' state which the substrate binding domain docks with the nucleotide binding domain and BiP has a low affinity for substrate. Upon ATP hydrolysis, BiP undergoes a conformational change and

adopts a 'domain undocked' conformation where both domains are independent of each other and BiP has a high affinity for substrate (Mayer, 2013) (Figure 1.8). In addition, BiP interacts with a whole host of co-chaperones (Kampinga and Craig, 2010), namely a number of Hsp40s that modulate the ATPase rate of BiP and target it towards substrates, and nucleotide exchange factors (NEFs) that exchange bound ADP for ATP (Behnke et al., 2015). All this highlights the complexity of BiP, prior to introducing IRE1.

IRE1 has been shown to bind to BiP under non-stressed conditions *in vivo*. Upon ER stress, the association of IRE1 and BiP is no longer observed (Bertolotti et al., 2000). While this is well established, the signals that trigger the dissociation of BiP from IRE1 are not currently known. One current model is that BiP represses IRE1 activation, and under ER stress BiP dissociates from IRE1, which allows IRE1 to dimerise and oligomerise (Chen, Y.N. and Brandizzi, 2013).

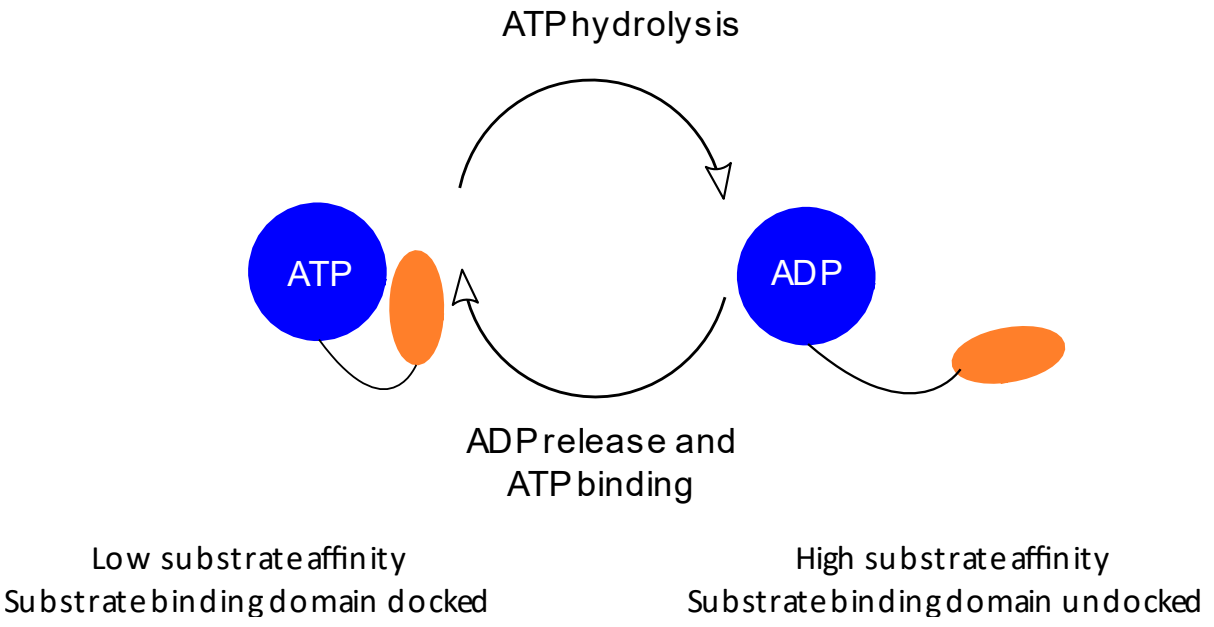

Figure 1.8

The chaperone cycle of the BiP, When BiP is ATP bound, its nucleotide binding domain (blue) is docked with its substrate binding domain (orange) and has a low affinity for substrates. Upon ATP hydrolysis, BiP undergoes a conformational change where its nucleotide binding domain and substrate binding domain are no longer docked. This ADP bound state possesses a high affinity for substrates. This cycle is completed, and subsequently continues upon ADP release and ATP binding (Mayer, 2013).

Expression of an IRE1 mutant containing a deletion of residues 390-433 resulted in a reduced the association of BiP, and subsequently this mutant also showed increased baseline levels of XBP1 mRNA splicing (Oikawa et al., 2009). This region is part of the long intrinsically disordered linker that links the luminal domain to the transmembrane region. Deletion of a similar region in yeast (residues 475-526, corresponding to residues 417-443 in human) did not cause an increase in baseline levels of HAC1 mRNA splicing (Kimata et al., 2004), possibly suggesting a difference in activation mechanisms for yeast and human IRE1, however truncating the linker could lead to steric occlusion of a binding site in the core luminal domain of IRE1 due to the proximity of the ER membrane. This is highlighted in more recent literature data that suggests BiP binds to the core luminal domain of IRE1, rather than the linker (Amin-Wetzel et al., 2017).

The association of IRE1 and BiP and their interaction is a matter of much debate in the literature. Liu et al. suggested the peptide-binding domain of BiP binds to IRE1 by using a series of truncated mutants of BiP and observing whether BiP associaties with IRE1 via immunoprecipitation experiments (Liu et al., 2003). More recently, the association between IRE1 and BiP was shown the be mediated by the Hsp40 co-chaperone, ERdj4, with the interaction suggested to be canonical as the association was ablated by the well characterised ATPase deficient mutant (T229G) and substrate binding deficient mutant (V461F). ERdj4 contains a conserved J-domain that stimulates ATPase activity of BiP, as mentioned previously, and a C-terminal targeting domain that is hypothesised to target BiP towards IRE1-LD. (Amin-Wetzel et al., 2017).

Interestingly, the nucleotide binding domain of BiP has also been shown to interact with IRE1. The binding affinity, measured by MST of full length BiP and the BiP NBD were demonstrated to be similar, while the isolated substrate binding domain no longer interacted. In the same study it was further suggested that BiP binding was not affected by the presence of ATP, ADP or AMPPNP. This conflicting model suggests BiP binds to IRE1 through its nucleotide binding domain and release is allosterically triggered by binding of unfolded protein substrates to the peptide binding domain of BiP (Carrara et al., 2015).

This highlights the need for clarification about the nature of the interaction between IRE1 and BiP. While it is possible that IRE1 interacts with both domains, the canonical and non-canonical models conflict regarding the effect of nucleotides on this interaction. Further studies are required to ascertain the nature of this important interaction.

1.4.2 Sigma-1-receptor (S1R)

While the hypothesis that IRE1-LD interacts with the BiP-NBD maybe controversial, previous literature data has shown that it also interacts with a sigma-1-receptor (S1R) (Ortega-Roldan et al., 2013), a protein associated with ER protein quality control (Hayashi, 2015). Interestingly S1R has also been shown to associate with IRE1-LD during times of early stages of ER stress and stabilise the presence of IRE1 at the mitochondria associated membrane (MAM). Furthermore, the same study demonstrated that this complex was stabilised and IRE1 activation was prolonged upon generation of reactive oxidative species (ROS) (Mori et al., 2013).

1.4.3 Hsp47

In addition to BiP, the luminal domain of IRE1 has several other interacting partners. Only a small number of these have been demonstrated to regulate its activity. Hsp47 is another ER localised chaperone that is critical for collagen formation (Nagata, 1996). This was recently shown to interact with IRE1, specifically with the active form of the protein and enhance IRE1 activity rather than repress activity like BiP. In addition, Hsp47 reduces BiP's association with IRE1, possibly suggesting competition between the two proteins to retain IRE1 in an active and inactive state respectively. It is noteworthy that the *in vitro* studies monitoring the competition between these two interactions were not performed in the presence of nucleotide, suggesting this competition is nucleotide independent (Sepulveda et al., 2018).

1.4.4 Protein disulphide isomerase A6

Protein disulphide isomerases (PDIs) are a family of proteins that reduce or rearrange disulphide bonds in the oxidising environment of the ER and are important for ensuring correct folding of client proteins (Wilkinson and Gilbert, 2004). Protein Disulphide Isomerase A6 (PDIA6) is an oxidoreductase that reduces disulphide compounds and is involved in aggregation of platelets (Jordan, P.A. et al., 2005). Overexpression of PDIA6 has been implicated in lung cancer and resistance pathways in various cancers (Gao et al., 2016) (Kaiser et al., 2007) (Ramos et al., 2015) (Tufo et al., 2014). As previously mentioned, IRE1 contains

three solvent accessible cysteine residues. PDIA6 interacts strongly with one of these cysteine residues, C148 and reduce disulphide bonds in IRE1-LD (Groenendyk et al., 2014). This interaction is critical to halting XBP1 signalling (Eletto et al., 2014), indicating a crucial role for disulphide bond formation and regulating IRE1 activity (Figure 1.9). This conclusion is further supported by PDIA6 knockout experiments in *C. elegans*, the result of which was hyperactivation of the UPR (Eletto et al., 2014). In order to fully elucidate the mechanism of upstream activation of IRE1 in humans, a far greater understanding of the role of these cysteine residues is required. Currently information is limited due to the residues most important for disulphide bond formation, C148 and C332, being unresolved in the crystal structure and a lack of studies.

While a function for these various binding partners has been established, there are several other binding partners that have been identified for IRE1, for which a function has not yet been established.

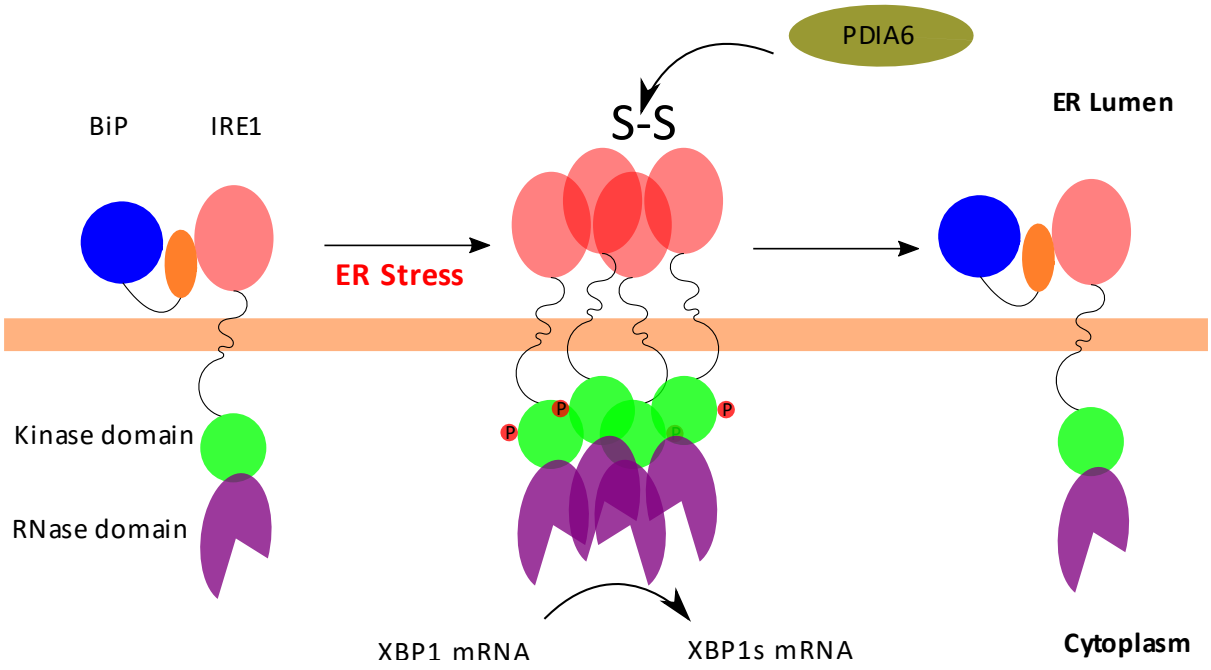

Figure 1.9

The proposed regulation of disulphide linked IRE1-LD oligomers by PDIA6. Upon activation, IRE1 forms intermolecular disulphides at its luminal domain that requires PDIA6 to reduce IRE1 back to its inactive form (Eletto et al., 2014).

All this highlights the importance of the luminal domain of IRE1 in controlling activation, and thus downstream signalling of the whole protein. While there are areas of understanding, such as identified binding partners and downstream effects of these, the nature of some of these interactions and structural and molecular details of the activation pathway of IRE1 are currently lacking. Specifically, further information about the process of disulphide bond formation and its role in regulation of IRE1 activation and the nature of BiP's interaction with IRE1 need to be determined to fully understand the pathway leading to IRE1 activation.

1.5 Cytoplasmic domain of IRE1 and its downstream signalling pathways

1.5.1 Structural details of the cytoplasmic domain of IRE1

While the luminal domain senses ER stress, the cytoplasmic domain regulates downstream signalling pathways. Following activation of the luminal domain, the kinase domain is thought to undergo trans-autophosphorylation, which subsequently regulates the activation of the RNase domain. Downstream signalling pathways are then regulated by IRE1's RNase activity, in addition to a complex pathway of interacting proteins that will be discussed later (Chen, Y.N. and Brandizzi, 2013). Both these domains contain well conserved features observed in other proteins; the kinase domain of IRE1 contain well conserved features observed in numerous other kinases and the RNase domain displays a high level of conservation with RNaseL (Lee et al., 2008).

The crystal structures of the cytoplasmic domains of IRE1 from various species have been crystallised in either a 'face-to-face' conformation or a 'back-to-back' conformation (Figure 1.10). The proximity of the two RNase domains is critical for binding of the mRNA stem loop and RNase splicing. The mammalian form of IRE1 has been crystallised in both a face-to-face (RNase inactive) and back-to-back (RNase active) conformations, named due to the position of the phosphorylation loops of the kinase domain. In the face-to-face conformation, phosphorylation loops are facing, which in turn means due to the sizeable distance between the two RNase domains, the intron from XBP1 mRNA is unable to be spliced. While it is unknown whether this is a biologically active conformation, this conformation has been observed on multiple occasions (Ali et al., 2011; Sanches et al., 2014). Current models suggest this conformation is related to trans-autophosphorylation due to the proximities of the phosphorylation loops and as all constructs have been treated with phosphatases prior to crystallisation. The back-to-back conformation, where phosphorylation loops are no longer

facing, contains classic elements of an active kinase, namely, the conserved DFG motif faces in towards the active site and completion of the helix-αC spine (Figure 1.11) (Endicott et al., 2012). This domain rearrangement leads to RNase domains on each protomer becoming proximal and forming an active RNase subunit. Interestingly, the apo form of IRE1 was found to crystallise in this conformation.

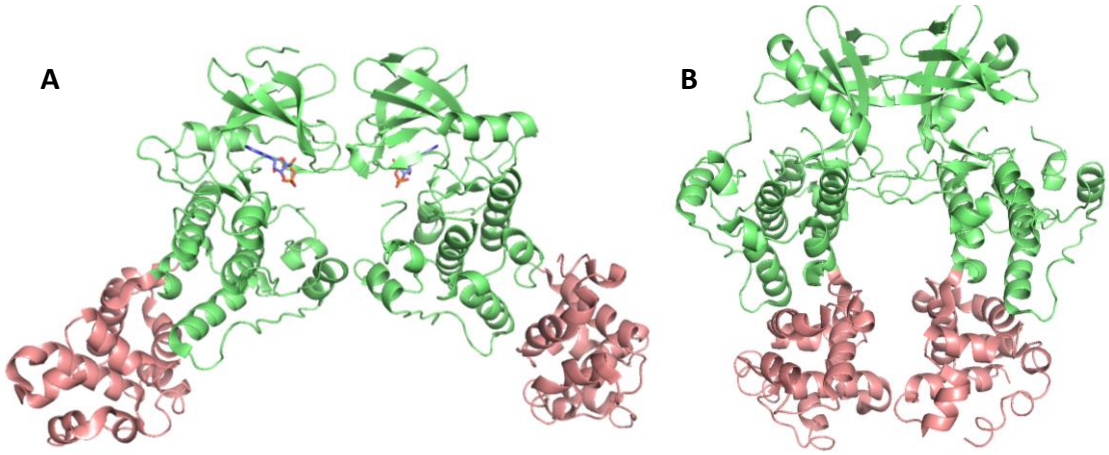

Figure 1.10

A) 'Face-to-face' dimer (trans-autophosphorylation active) crystal structure of human IRE1 and B) 'back-to-back' dimer (RNase active) crystal structure of yeast IRE1. Both structures with ADP bound (represented by sticks) and kinase domain is coloured lime green and RNase domain is coloured pink PDB ID: 3P23 (face to face) (Ali et al., 2011) and 4Z7G (back to back) (Lee et al., 2008). Figure prepared using PyMol (version 1.8)

1.5.2 Comparison between yeast and human homologs of the cytoplasmic domain

While recent crystal structures of the cytoplasmic domain of mammalian IRE1 have been solved, as with the luminal domain of IRE1, the majority of initial studies were performed in yeast. Crystal structures of the yeast cytoplasmic domain are seen in a back-to-back conformation (Lee et al., 2008), both as dimers and oligomers. In this conformation the proximity of the RNase domains from each protomer is in close enough vicinity to splice the intron from HAC1 mRNA. Splicing is proposed to occur via H1061 and Y1043 acting as a general acid-general base pair with recognition occurring via R1056 and N1057 (Lee et al., 2008).

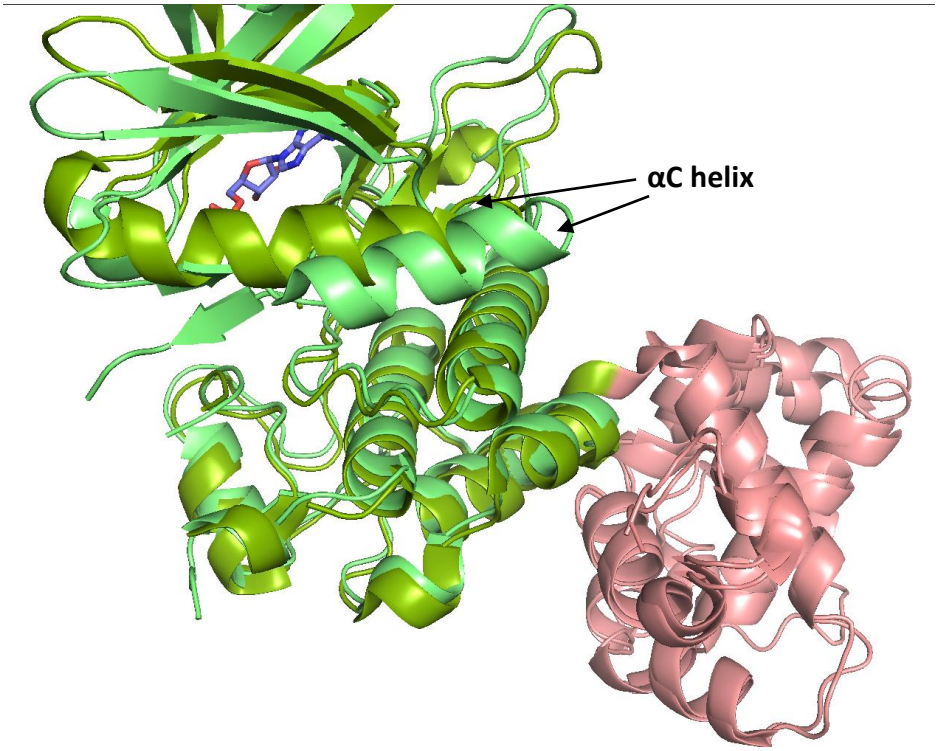

Figure 1.11

The incomplete alpha-C helix of the kinase domain of IRE1 in the face-to-face conformation is hypothesised to prevent the RNase active 'back-to-back' dimer (lime green) whereas in the apo state the completed alpha-C helix adopts a different conformation which allows back-to-back dimer formation (dark green). PDB ID 3P23 (lime green) and 4Z7G (dark green) are used to illustrate inactive and active RNase conformations respectively. Figure prepared using Pymol (version 1.8)

When compared to the luminal domain, a greater proportion of the structural elements present in the cytoplasmic domain are conserved between human and yeast structures (Figure 1.12). Despite this, subtle differences may have a huge impact on activation of RNase activity. Differences in the nucleotide-binding pocket are demonstrated by more potent activation of RNase activity of yeast IRE1 by ADP rather than ATP (Korennykh, A.V. et al., 2011), while in murine IRE1, ADP has been shown to inhibit RNase activity (Sanches et al., 2014).

As mentioned during the discussion of the luminal domain, a structure of oligomers in human IRE1 currently remains elusive. In yeast, a 32 amino acid linker between the kinase domain and transmembrane domain (residues 641-672) is crucial to oligomer formation of cytoplasmic domain constructs, with constructs lacking this linker unable to form oligomeric structures (Korennykh, A.V. et al., 2009). Interestingly, this linker region is not conserved in mammalian IRE1.

A construct of the human IRE1 cytoplasmic domain including a 43 amino acid linker region (residues 528-570) was shown to form oligomers by analytical ultracentrifugation however repeat experiments without the linker were not performed, therefore it is not known whether this linker is crucial in human constructs. Further support for the model of IRE1 oligomer formation was provided by the cooperative activation profile of ^{32}P-labelled stem loop oligoribonucleotide modelled on XBP1 mRNA. While a dimer would have cleaved with a Hill coefficient n=2, the labelled oligoribonucleotide was cleaved with a Hill coefficient of n=3.4, suggesting the formation of higher order oligomers (Li et al., 2010).

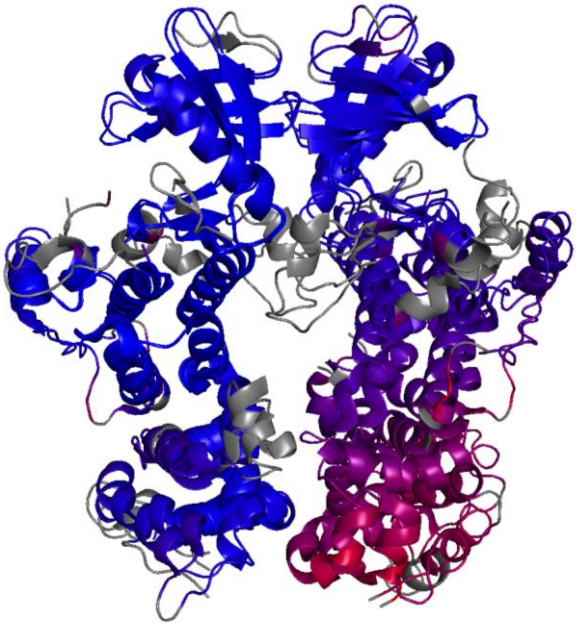

Figure 1.12

Cytoplasmic domain of the RNase active conformation of human and yeast IRE1 (4Z7G and 2RIO). Minimum pairwise RMSD is indicated in blue with maximum in red. Unaligned residues are in grey. Significantly more of the structure is aligned in the cytoplasmic domain compared to the luminal domain. Figure prepared using PyMol (version 1.8)

While there are currently several published structures of the cytoplasmic domain of both yeast and human IRE1, more understanding of the mechanism of activation of the cytoplasmic domain is required to regulate its function. If current models are correct, and the face-to-face dimer precedes phosphorylation and thus subsequently followed by trans-autophosphorylation and a conformational change to an RNase active back-to-back dimer, further understanding is required of phosphorylation steps and potential allosteric effects leading to a conformational change and formation of the back-to-back dimer.

1.5.3 Regulation of downstream activity of the cytoplasmic domain

The kinase domain of IRE1 is crucial for activating the RNase activity. Mutations that affect the kinase region have a knock-on effect and disrupt the RNase activity of IRE1 (Tirasophon et al., 1998). The kinase activity of IRE1 is very complex, with no single phosphorylation providing an on/off activation of the RNase activity. Instead the various phosphorylation sites each serve different purposes. Several phosphorylation sites from a construct of human IRE1 expressed in insect cells have been identified, along with their effect on RNase activity. Phosphorylation in the region linking the transmembrane region and the cytoplasmic domain, specifically S551 and S562, do not induce XBP1 splicing much beyond the levels of dephosphorylated IRE1. Further phosphorylation at S724, S726 and T973 (located in the RNase domain) provided a marked increase in XBP1 splicing, with a further phosphorylation at S729 providing further increases still (Prischi et al., 2014).

In addition to the role of the kinase domain, human IRE1 has been shown to interact with numerous proteins (Figure 1.13), forming part of what is commonly termed the ‘UPRosome’. This massive network of protein complexes and downstream signalling cascades ultimately plays a vital role in determining cell fate.

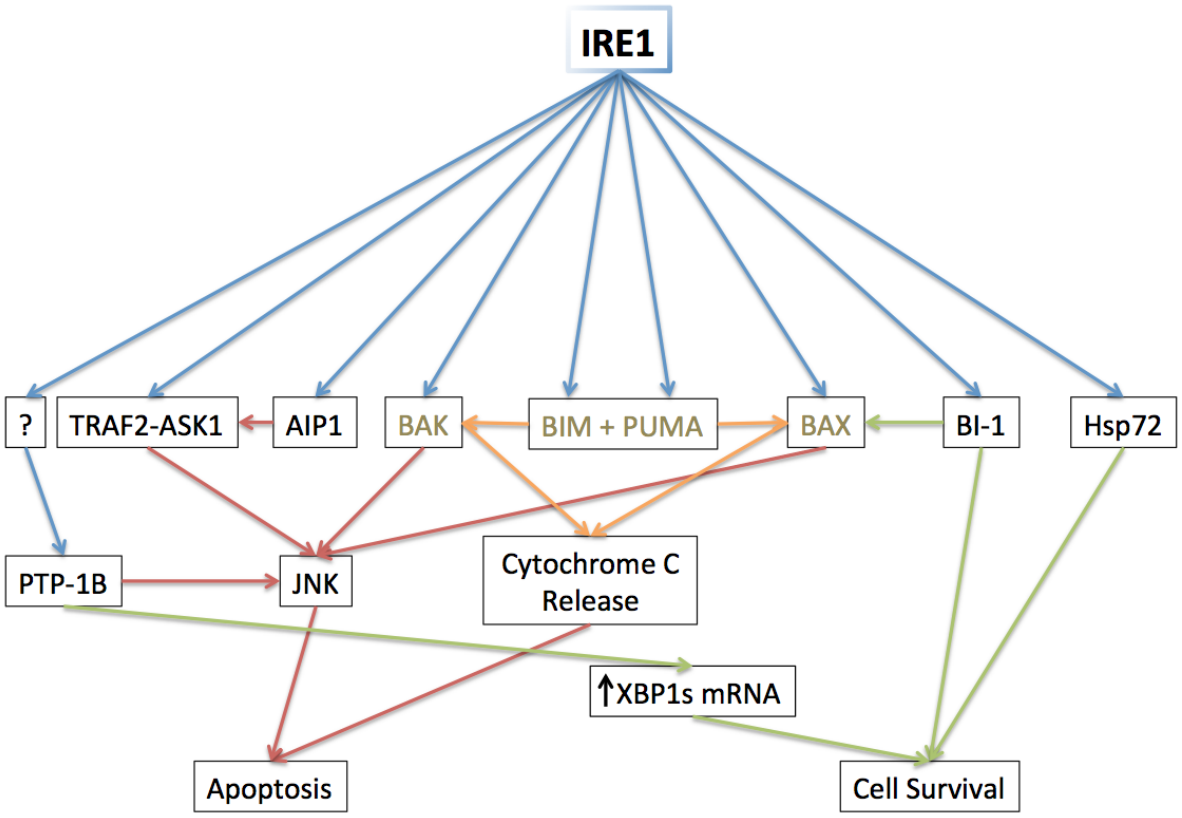

Figure 1.13

Various downstream signalling pathways of IRE1 interacting proteins. Pro-apoptotic pathways are red and pro-survival pathways are green. BCL-2 family of proteins are in gold and their interactions are in orange.

One of the more comprehended complexes is the IRE1-TRAF2-ASK1 complex, which is formed under ER stress. This activates c-Jun N-Terminal Kinase (JNK), which then activates downstream apoptotic signals (Nishitoh et al., 2002). Mutation studies suggest that this pathway is independent of the RNase activity of IRE1 (Urano et al., 2000).

ASK1-interacting protein 1 (AIP1) has also been shown to interact directly with IRE1 with the IRE1-AIP1 complex found in immunoprecipitates of endothelial cells undergoing ER stress. AIP1 has also been shown to facilitate the dimerization of IRE1 with AIP1 knockout experiments showing a reduction of levels of an IRE1α-IRE1β dimer (Luo et al., 2008). AIP1 directly interacts with ASK1 by facilitating dephosphorylation at S967, which leads to

activation and increased ASK1-JNK signalling (Zhang, R. et al., 2003). This could suggest that ASK1 needs to be dephosphorylated to enhance the dimerization of IRE1.

BAX is a pro-apoptotic member of the BCL-2 family of proteins, which are known to regulate cell death activity (Lomonosova and Chinnadurai, 2008). BAX and BAK have been shown to directly interact with the cytoplasmic domain of IRE1, with an increased affinity under ER stress. Double knockout experiments suggest that BAX and BAK are involved in both XBP1 splicing and JNK phosphorylation/apoptosis pathways (Hetz et al., 2006). BAX inhibitor-1 (BI-1) seems to regulate the inactivation of IRE1 with knockout experiments showing prolonged XBP1 splicing and increased levels of cell death (Lisbona et al., 2009).

Further members of the BCL-2 family have also been shown to interact with IRE1. BIM and PUMA are members of the BH3-only subgroup of the BCL-2 family and are classed as BH3 activators. BH3 activators interact with BAX and BAK, and subsequently trigger cytochrome c release and apoptosis (Lomonosova and Chinnadurai, 2008). Both BIM and PUMA have been shown to interact with IRE1 in mouse embryonic fibroblasts undergoing ER stress with knockout experiments showing a decrease in levels of XBP1 mRNA splicing (Rodriguez et al., 2012). Given the cytoprotective effect XBP1s has on cells, it would be assumed that knockout of pro-apoptotic proteins would have a positive effect on XBP1 splicing. These interactions, along with those previous mentioned with BCL-2 proteins, demonstrate the fine balance between pro-survival and apoptotic signals around the UPRosome, as well as the complexity regarding signalling.

Contrarily to the IRE1-TRAF2-ASK1 complex and the BCL-2 proteins, Hsp72 has been shown to complex with IRE1 and enhance cell survival. In this case, increased levels of Hsp72 increase the rate of mRNA splicing, thus enhance the levels XBP1s mRNA. While Hsp72 has previously been shown to have a pro-survival effect in cells, by repressing the apoptotic effect of ASK1, these results demonstrate a direct link between Hsp72 and IRE1. Increased XBP1s mRNA levels were shown to directly correlate with a reduction in cell death as Hsp72 no longer has a pro-survival effect on IRE1 RNase deficient mutants. Observation of the same effect upon the introduction of shRNA targeting XBP1s further confirmed this link (Gupta et al., 2010). This is particularly interesting when considering that elevated levels of Hsp72 are seen in many cancer cells suggesting a viable survival pathway (Ciocca and Calderwood, 2005).

Protein-tyrosine phosphatase 1B (PTP-1B), like other phosphatases, dephosphorylates phosphorylated residues, in this case specifically tyrosine residues (Tonks, 2006). PTP-1B plays a role in IRE1 signalling, specifically in JNK activation, thus as expected, PTP-1B knock out cells show an increased resistance to ER stress induced apoptosis. XBP1s mRNA levels are also reduced in knockout experiments, although BiP and CHOP mRNA levels remained consistent with those seen in WT cells. Further interest in the interaction of PTP-1B and IRE1 is created, as IRE1 is not phosphorylated at a tyrosine residue, suggesting that PTP-1B interacts with an intermediate, rather than directly with IRE1 (Gu et al., 2004).While by no means exhaustive, this list of interacting proteins provides an insight into the balance between pro-survival and pro-apoptotic pathways of IRE1.

1.6 Conclusions and Project Aims

In this chapter, we have highlighted the importance of IRE1, both for its role in regulating protein quality control elements and in determining cell fate. It is clear that IRE1 is a complex system, with a vast network of binding partners which affect downstream cell fate decisions, thus detailed characterisation of upstream activation and downstream responses are beyond the scope of a single thesis. While there are gaps in knowledge in both pathways, there is less conservation between yeast and human homologs of the luminal domain of IRE1 and fewer conserved elements in the luminal domain, when compared to the cytoplasmic domain (i.e. kinase elements, RNase L like domain).

While there is evidence of membrane induced activation of yIRE1, (Halbleib et al., 2017), we will specifically focus on the luminal domain activation pathway, which is complex and lacks clarity, specifically regarding the role of BiP and its effect on the conformational landscape of IRE1-LD, and the role of regulation of IRE1 activity via disulphide bond formation in its luminal domain. In addition, there is a lack of biophysical data on the individual steps of IRE1 activation, more of which would make it possible to ascertain the effects of individual elements in the activation pathway of IRE1. This project will aim to elucidate these gaps in knowledge using the following *in vitro* strategies:

- Using construct design, we will express and purify a number of constructs of IRE1-LD at yields suitable to perform detailed biophysical characterisation and structural studies.

- Ascertain structural and mechanistic insights into the process of IRE1-LD dimerisation, peptide induced oligomerisation, and disulphide bond formation through the use of biochemical, biophysical, structural and computational methods.
- Elucidate the effect of BiP on the conformational landscape of IRE1-LD using well characterised mutants of the former and comparing data to that previously ascertained for IRE1-LD in isolation.

As IRE1-LD is a complex system that undergoes numerous conformational changes, we will require a range of techniques to fully investigate its conformational landscape. In the next section we will detail the methods we will use to achieve this.

1.7 Methods used for characterisation of protein quality control systems

1.7.1 Mass Spectrometry (MS)

Mass spectrometry traditionally provided a sensitive method for determining the molecular mass of molecules. In order to be analysed, molecules need to be transferred to the gas phase, ionised using an ionisation source, separated by a mass analyser and then detected. It has only been possible to analyse large biomolecules in the past 30 years, primarily due to the development of ionisation techniques such as electron spray ionisation (ESI) and matrix assisted laser desorption ionisation (MALDI). This has made it possible to analyse the molecular weight of proteins in their denatured state (Mann et al., 2001).

1.7.1.1 Native Mass Spectrometry

In recent years, the development of softer ionisation techniques such as, ESI, has allowed analysis of intact non-covalent complexes, a technique often referred to as native mass-spectrometry (native-MS). While this does not offer structural insights into protein complexes, this can allow elucidation of the stoichiometry of large protein complexes. In addition, native mass-spectrometry is extremely sensitive and allows analysis of large protein complexes, such as the 20S proteasome (Loo et al., 2005), in contrast with structural techniques such as NMR where sensitivity issues arise with increasing size of proteins (Heck, 2008). Native-MS does suffer from some drawbacks however; due to the necessity to transfer ions to the gas phase hydrophobic interactions often do not survive this transfer (Bich et al., 2010) (Yin et al., 2008), however it is still possible to observe polar and electrostatic interactions (Loo, 1997) (Daniel et al., 2002).

1.7.1.2 Ion mobility mass spectrometry

Ion mobility mass spectrometry (IM-MS) builds upon the principles of native mass spectrometry, by allowing the elucidation of low-level structural detail of intact molecular complexes. There are several common types of instrumentation used for IM-MS; drift time ion mobility spectrometry (DTIMS) is the oldest of these. Ions are introduced to a drift cell filled with an inert buffer gas (most commonly helium), and a static electric field is applied to ensure the ion traverses the drift cell. Ions are separated by their rotationally averaged cross section (often referred to as collision cross section (CCS) or Ω) due to collisions with the buffer gas, and thus reach the detector at different drift times. Using the Mason-Schamp equation

(Equation 1.1, where Ω is the rotationally averaged collision cross section, *ze* is the charge state of the analyte ion, k_B is the Boltzmann constant, *T* is the temperature, μ is the reduced mass of the buffer gas and analyte ion, *P* is the pressure of the drift gas, *N* is the number density of the drift gas, t_d is the drift time, *E* is the electric field and L is the length of the drift cell) it is possible to calculate a CCS, which can be compared with values using other structural and/or computational techniques (Pagel et al., 2013; Baumketner et al., 2006).

$$\Omega = \frac{\sqrt{18\pi}}{16} \frac{ze}{\sqrt{k_B T}} \frac{1}{\sqrt{\mu}} \frac{760}{P} \frac{T}{273.2} \frac{1}{N} \frac{t_d E}{L}$$

Equation 1.1

Travelling wave ion mobility spectrometry (TWIMS) builds upon the principles of DTIMS. Unlike DTIMS, TWIMS utilises a series of ring electrodes with oppositely phased radio frequency voltages that allows the creation of a travelling wave, whereupon ions of differing mobilities travel along the wave at different rates. This allows greater separation of complex mixtures and improved sensitivity when compared to DTIMS (Shvartsburg and Smith, 2008; Pringle et al., 2007). As a static electric field is no longer applied, it is no longer possible to calculate a CCS based on the Mason-Schamp equation, thus calibration standards are often used to determine the CCS when using TWIMS (Lanucara et al., 2014) (Figure 1.14).

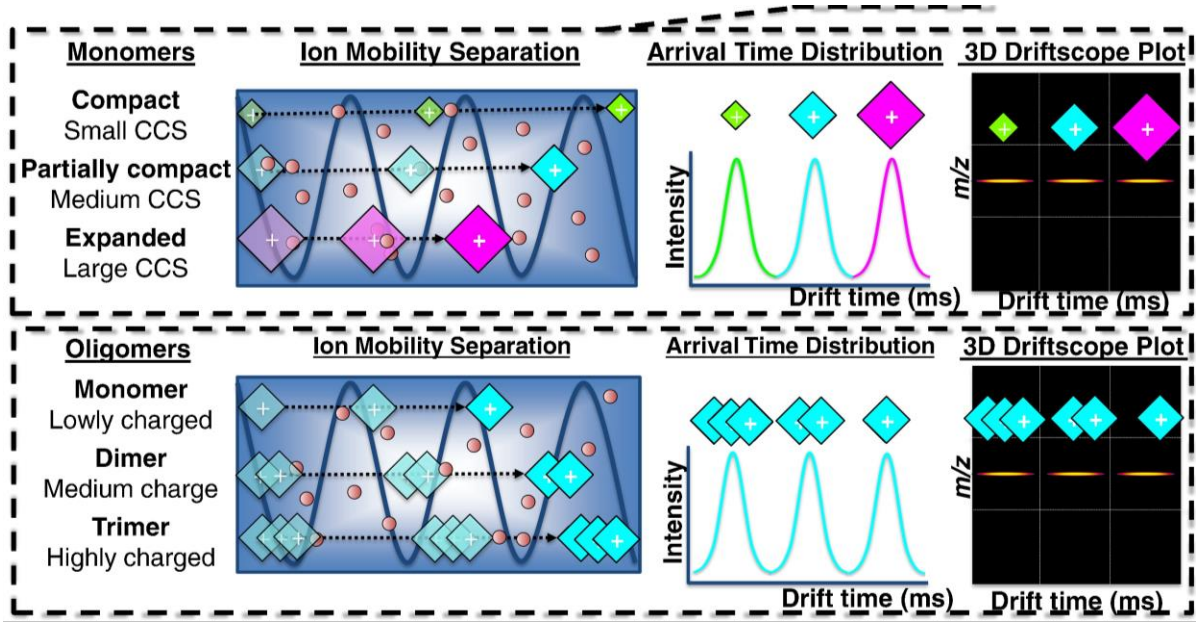

Figure 1.14

Schematic of a TWIMS. Small compact ions traverse the drift cell faster than less compact ions of an equivalent mass, due to additional collisions with the buffer gas. In addition to separation by shape, oligomers can be separated out based on their charge, with larger oligomers generally possessing a higher charge state than monomers. Figure adapted from (Woods et al., 2013) and licensed under CC BY 3.0.

1.7.2 Molecular Dynamics (MD) Simulations

Classic structural techniques such as X-ray crystallography can provide an excellent, high resolution snapshot of a molecule however the reality is that proteins sample numerous conformations over time.

MD simulations rely on Newtonian laws of motions to predict the motion of atoms over the time course of a simulation, based on bonded, non-bonded and electrostatic interactions. The behaviour of each atom is dictated by force fields, models which dictate the how atoms react to forces. These aren't without their drawbacks, with problems known to occur with long simulations due the inaccuracy of the forcefields (Dror et al., 2012).

The use of MD simulations on biomolecular systems can provide various molecular insights into numerous processes that occur such as atomistic detail about what drives conformational changes between different states and investigations into the mechanisms of protein folding (Karplus and McCammon, 2002).

1.7.3 Nuclear Magnetic Resonance (NMR)

NMR provides an opportunity to observe individual nuclei of complexes, thus can provide powerful insights into the structure and dynamics of proteins. While the nuclei which we can observe are limited, through the introduction of NMR active isotopes into media during bacterial growth and protein expression, it is possible to acquire spectra for biomolecules.

All nuclei possess an intrinsic property called spin (I). Those nuclei where I≠0 can be analysed by a technique called nuclear magnetic resonance, although for analyses in this thesis, only nuclei where I=½ were analysed. Under normal conditions, the nuclear magnetisation vector (M) averages to 0, however when placed inside a magnet, this vector averages out to a value proportional to the size of the magnetic field (B_0) and the gyromagnetic ratio (γ) of the relevant nuclei. This vector can then be manipulated by means of radiofrequency pulses. When an element of the vector is perpendicular to the magnetic field, it rotates at its Larmor frequency (ω_0), it creates an alternating current, much like a magnet inside a coil that subsequently (Equation 1.2). A fourier transform can then be applied to this signal to produce an NMR spectrum. The effect of shielding of nuclei from the magnetic field (B_0) leads to variations in the Larmor frequency of identical nuclei, which causes different chemical shifts in NMR spectra. As nuclei are extremely sensitive to their chemical environment, residue specific information about binding events and conformational changes can be obtained both in solution and as solids. This work will focus on solution state NMR.

$$\omega_0 = -\gamma B_0$$

Equation 1.2

As protons are naturally abundant and have a larger gyromagnetic ratio than other nuclei naturally occurring in proteins, they are more sensitive to detection by NMR, thus are typically observed during an NMR experiment. If we were to perform an experiment where we simply observed all protons in a typical protein, it would be impossible to distinguish between signals due to the sheer number of protons and those present in similar or identical chemical

environments, which would be exhibited as overlapped residues. As nuclei influence nearby neighbours, an effect often termed as 'coupled', it is possible to exploit this property in order to only observe nuclei which are coupled to other nuclei which possess a nuclear spin value (typically ^{15}N or ^{13}C in proteins), a technique referred to as polarisation transfer. It is also possible to perform several experiments with varying delay times between polarisation transfer back to the proton, to separate out these proton signals as a function of the chemical shift of the transferred nuclei, thus producing a 2D-NMR spectrum. The basic building block of 2D NMR experiments is insensitive nuclei enhanced by polarization transfer (INEPT) (Morris and Freeman, 1979), which allows polarisation transfer (also referred to as magnetisation transfer) to coupled nuclei.

1.7.3.1 Protein NMR

While it is possible to analyse small molecules using 1D NMR, for large macromolecules and proteins, spectral complexity means it is necessary to perform multidimensional NMR experiments. We will discuss some of these strategies in this section.

1.7.3.1.1 Amide NMR

As the protein backbone is made up of amide bonds, these provide an obvious probe to study protein structure and dynamics. Heteronuclear single quantum correlation (HSQC) spectra can be typically acquired for proteins of <20-30 kDa (Clore and Gronenborn, 1997), however for complexes larger than this, alternative strategies are required due to signal broadening caused by increased rotational correlation times.

Transverse relaxation optimisation spectroscopy (TROSY) based methods exploits a slow relaxing element of the multiplet formed in decoupled spectra. Through the pulse sequence for TROSY based experiments, it is possible to only select this line, thus produce spectra with sharp lines, negating the signal broadening effects observed in standard HSQC based experiments of larger proteins (Tzakos et al., 2006) (Figure 1.15).

It is possible to ascertain information about binding sites, if the spectrum is assigned, upon binding of ligands and proteins through chemical shift perturbation analysis, due to changes in the chemical environment of a nuclei (Marion, 2013).

1.7.3.1.2 Deuteration

As protons are the main source for dipole-dipole signal relaxation in NMR experiments, a typical strategy is to replace these with deuterons (Sattler and Fesik, 1996). This is typically achieved by replacing water in labelling media with deuterium, which leads to incorporation of deuterons throughout the protein. Upon introducing the protein to a water based buffer, the deuterons in the amide backbone exchange back to protons, thus it is still possible to perform amide NMR experiments. While this approach is expensive, it can provide a significant increase in signal (Gardner, K.H. and Kay, 1998).

1.7.3.1.3 Methyl NMR

While TROSY and deuteration substantially increase the size of proteins that it is possible to observe by NMR, signal overlap can become an issue with larger systems.

Methyl groups are an attractive target for NMR for several reasons. The intrinsic rotational properties of methyl groups produce sharp signals in NMR spectra as they relax independently to the rest of the protein, and as they are hydrophobic in nature, they are often buried in the structure of a protein, thus provide a probe for structural and dynamic changes (Tugarinov and Kay, 2005). Through the use of isotope labelled precursors, it is possible to introduce selectively isotope labelled methyl groups of certain amino acids (Tugarinov et al., 2006) (Ayala et al., 2009). Isoleucine, leucine, valine and alanine groups are often labelled together due to the signal dispersion between the isotope labelled methyl groups of these amino acids (Godoy-Ruiz et al., 2010).

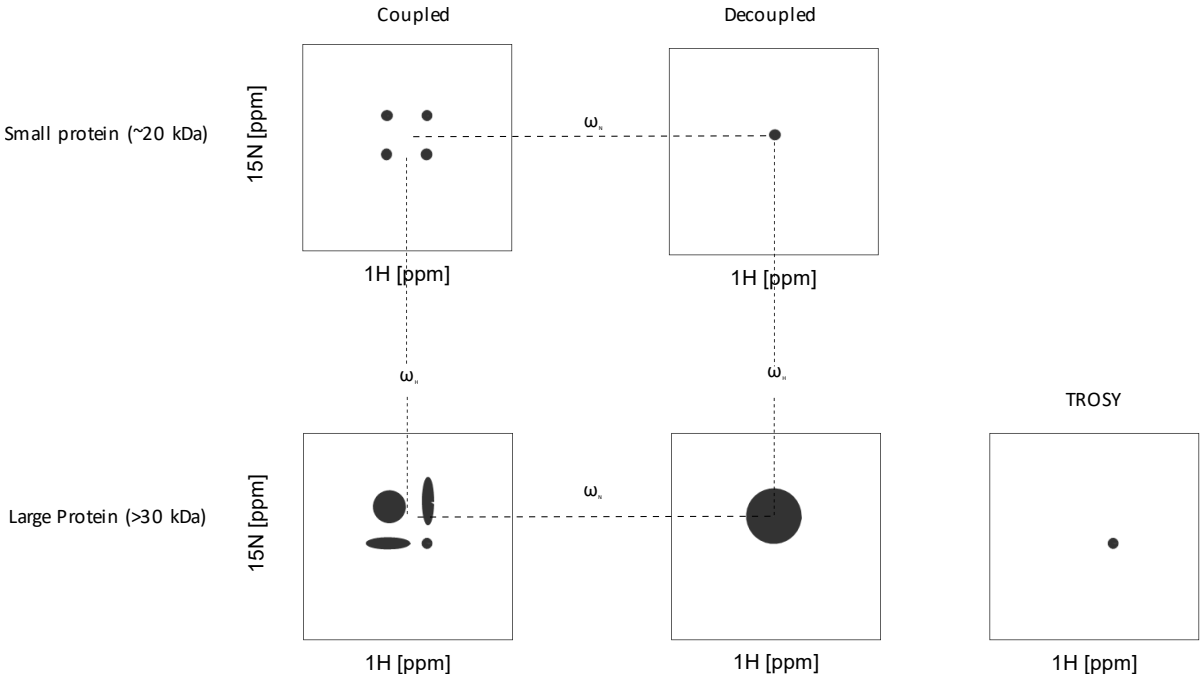

Figure 1.15

Schematic of various coupled and decoupled experiments. For small proteins the four elements of a coupled amide spectrum relax at similar rates, leading to two sets of doublet peaks (separated by $^{1}J_{NH}$) of similar intensities. Decoupling of this spectra leads to a single sharp peak of increased intensity, centred at ω_H and ω_N. For large proteins, the four elements of a coupled amide spectrum relax at different rates, leading to two sets of doublet peaks of differing intensities. Decoupling of this spectra leads to a broad peak centred at ω_H and ω_N. It is possible to select the slow relaxing element of these doublets (the sharp peak), thus produce a spectra with sharp lines for larger proteins. The sharpness of this peak increases with increasing magnetic field strength, with an optimal field strength of ~ 1GHz (Tzakos et al., 2006).

1.7.3.1.4 Dynamics in protein NMR

While it would be assumed all backbone residues, excluding prolines, should be visible in an amide NMR spectrum of a medium sized protein, this is often not the case (Kleckner and Foster, 2011). If chemical exchange occurs between nuclei during an NMR experiment, this will not be the case. Exchange that occurs in the µs-ms time scale can result in three possible

outcomes, dependent on the rate of exchange (k_{ex}) and the difference in chemical shift between the states $\Delta\nu$; if the rate of chemical exchange is fast, ($k_{ex} >> \Delta\nu$) an averaging effect for these nuclei will occur, proportional to the relative populations of these states, thus the peak that appears in the NMR spectra will be an average of these two states, whereas exchange that occurs is slow ($k_{ex} << \Delta\nu$), this results in two peaks for this one residue, with their peak volume proportional to the population of each state. If exchange occurs in on intermediate timescale ($k_{ex} \sim \Delta\nu$), a peak appears between these resonances, which leads to signal broadening. This can potentially render peaks invisible to NMR, thus, for example, a protein containing 100 residues, excluding prolines, may contain far fewer or more peaks, if significant chemical exchange is occurring between nuclei (Kleckner and Foster, 2011). This effect is not unique to amide NMR and also occurs with methyl NMR (Karagoz et al., 2017).

The use of pulse sequences such as Carr-Purcell-Meiboom-Gill (Carr and Purcell, 1954; Meiboom and Gill, 1958) can ensure resonances undergoing chemical exchange are broadened and are often termed relaxation based experiments. These are useful for obtaining information about dynamics in proteins (Massi et al., 2005) (Mittermaier and Kay, 2006), but can also be employed in ligand binding experiments (Moschen et al., 2016) which will be discussed below.

1.7.3.2 ^{19}F NMR and its applications in drug screening

1D ^{19}F NMR can be particularly powerful in drug screening applications and has been utilised in recent years on a number of occasions (Jordan, J.B. et al., 2012; Dalvit et al., 2003b). ^{19}F is 100% abundant, NMR active and possesses a gyromagnetic ratio comparable to that of ^{1}H, thus it is possible to perform inexpensive and sensitive experiments using this nucleus. By either replacing nuclei in existing ligands with 19F, or limiting library searches to compounds with ^{19}F present, it is possible to perform ligand observed experiments these against large protein complexes. This negates the need to isotopically label your drug target. In addition, ^{19}F NMR can be advantageous over proton based methods, which are subject to interference from commonly used solvents, such as DMSO and H_2O and prone to a greater degree of signal overlap, due to its more limited chemical shift range.

The intrinsic properties of the ^{19}F nucleus render it a useful tool for monitoring binding to large protein complexes. As there is generally a large chemical shift anisotropy element to T2

relaxation of ^{19}F nuclei possess, significant line broadening is observed at high magnetic fields and in large complexes. This property becomes advantageous when screening small molecules against large protein complexes as significant line broadening is observed upon binding of the small molecule, thus acts as a powerful reporter of binding (Dalvit, 2007).

While resonances may broaden into the noise upon binding, it can be advantageous to apply a CPMG sequence (also termed relaxation edited sequences), especially for fast exchanging complexes. For these experiments, ligand is added in excess of the target protein, thus only a small fraction (dependent on the K_d) of the total ligand population is bound, with the lifetime of this complex dictated by the k_{off} rate. By employing long CPMG times it is possible to remove signals from compounds in fast exchange, thus detect weak binding compounds from a drug screen (Hajduk et al., 1997), which can be particularly useful for fragment based screens where strong interactions are not expected.

1.7.4 Size Exclusion Chromatography (SEC)

Chromatographic techniques provide means of separating materials based on a number of parameters. Size exclusion chromatography separates elements out by size, or more specifically hydration radius, by means of a matrix, often silica or carbohydrate based, that contains pores of varying size. Smaller molecules are able to travel through these pores while molecules of a larger hydration radius are excluded from passing through these pores, and as such their path through the column is shorter, thus are eluted from the column sooner (Fekete et al., 2014). Classically for proteins, folded protein standards of a known molecular weight are run through the column as a calibration. A curve of the retention volume vs. the molecular weight of these known standards is constructed and used to elucidate the molecular weight of a protein of unknown molecular weight, which often, for folded proteins provides an accurate estimate of molecular weight within approximately 10%. For unfolded proteins, whose hydration radius is usually larger than those of globular proteins of a comparable molecular weight, this method of determining molecular weight is less reliable, thus more costly detection methods such as multi-angle-laser-light-scattering (MALLS) are used to determine an accurate molecular weight. SEC is limited by the molecular weight range of the column, usually ranging from 10s-100s of kDa.

1.7.5 Dynamic Light Scattering (DLS)

While most light scattering techniques are interested in the angle and intensity of scattering, DLS measured the correlation of scattered light intensity between a start time (t) and specified delay period (τ). Small particles have a greater degree of Brownian motion when compared to larger particles, thus the correlation of scattered light between t and τ is quicker to decay. For homogenous samples the correlation plot would be expected to fit to a one-phase exponential decay at a decay rate of Γ. Taking the magnitude of the scattering vector, q, it is possible to calculate the translational diffusion constant, D_t, (Equation 1.4), and subsequently, it is then possible to use the Stokes-Einstein equation (Equation 1.3) to calculate a hydration radius, R_h, in a solvent of viscosity η, assuming a spherical model (Stetefeld et al., 2016). The size of species in more complex mixtures can be ascertained using more complex algorithms, such as cumulants based models and regularisation algorithms (Hassan et al., 2015).

It is important to note that as larger particles scatter light more intensely than smaller particles, therefore intensity-based measurements do not accurately reflect the quantity of a particular particle size present. In addition, as inferred earlier, DLS is not a direct measure of molecular mass, instead it is inferred based on molecular motions. While DLS does not resolve species as accurately as techniques such as SEC, it can be used to determine the presence of large particulates and can determine the molecular weight of complexes that may dissociate when run by SEC.

$$R_h = \frac{k_B T}{6\pi\eta D_t}$$

Equation 1.3

$$D_t = \frac{\Gamma}{q^2}$$

Equation 1.4

1.7.6 Quantitative binding techniques

1.7.6.1 Microscale thermophoresis (MST)

MST is a technique that exploits the difference in thermophoretic properties of a molecule between a bound and unbound state. This involves fluorescently labelling one of the binding

partners and adding a small quantity to a dilution series of the other binding partner. These mixtures are then placed into small capillary tubes and an adjustable infrared laser is applied to an area of this tube, which increases the temperature at this point of the capillary tube. Throughout the experiment the fluorescence of the same area is measured. As the heat of the infrared laser is applied, the labelled molecule thermophoresis away from this area, resulting in a reduced fluorescence reading, and when the laser is turned off, molecules diffuse back towards this area, resulting in an increase.

The difference between these two measurements is recorded, with bound and unbound forms of the fluorescent molecule behaving differently (Figure 1.16). Typically, a smaller binding partner/ligand is fluorescently labelled due to a theoretically larger change in thermophoresis upon binding of a larger binding partner/receptor. This technique is extremely sensitive and can provide a K_d for interactions over a wide range of binding affinities (Wienken et al., 2010; Jerabek-Willemsen et al., 2014). Additionally, MST commonly relies on fluorescent labelling, thus a label free method should be used to confirm binding.

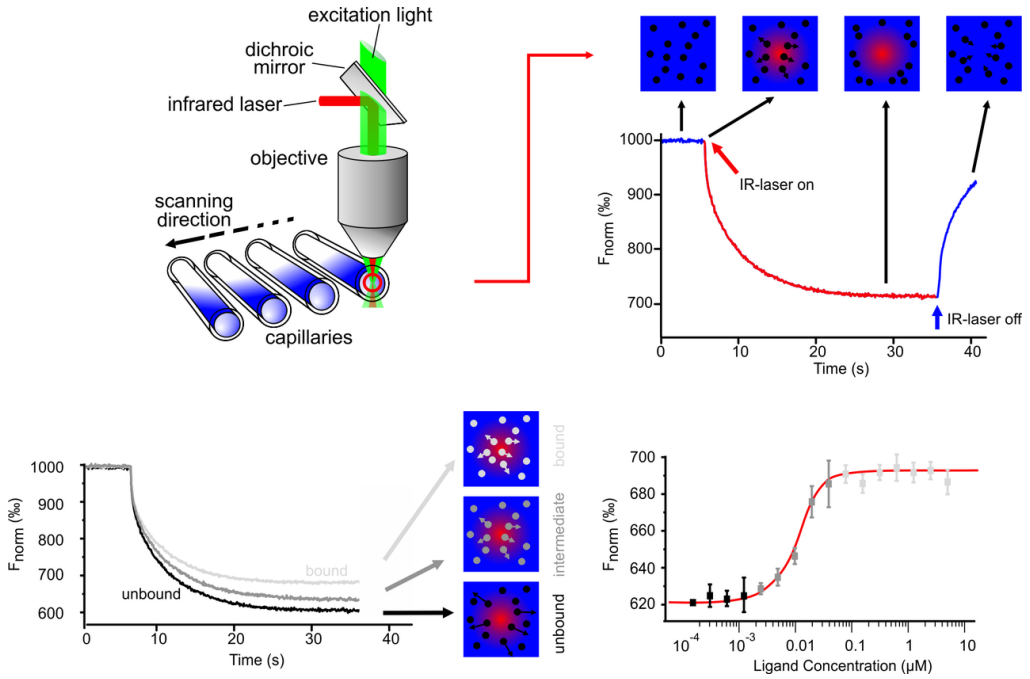

Figure 1.16

Schematic of MST experiments. Fluorescence is monitored at a point on a capillary tube where an infrared (IR) laser is applied. An initial fluorescence reading is obtained prior to application of the (IR) laser. Upon application of the IR laser, molecules diffuse away from the point of application. The IR laser is then turned off which allows molecules to diffuse back

towards the point of IR laser application. The rate of diffusion is dependent upon the shape and size of the fluorescently labelled molecule, thus differences emerging from binding can be monitored and quantified. Image from (Jerabek-Willemsen et al., 2014) and licensed under CC BY 3.0

1.7.6.2 Fluorescence Polarisation

As molecules of different sizes tumble at different rates in solution, this property can be exploited with fluorescently labelled molecules. Plane polarised filters can detect the fluorescence for parallel and perpendicular to the direction of excitation. These fluorescence readings can be subtracted to produce a fluorescence polarisation reading that vary upon binding (Pollard, 2010). Typically small ligands are labelled when binding to large proteins as the expectation would be the difference in fluorescence polarisation would be larger than if the reverse were to be performed.

1.7.6.3 Isothermal calorimetry

ITC can theoretically provide the greatest amount of information about an interaction when compared to other quantitative binding techniques. A protein of introduce is introduced into a sample cell, which is surrounded by a reference cell, which is kept at a constant temperature. A ligand or binding partner is injected into the sample cell at specified quantities. The power required to maintain a constant temperature is then measured, which can then be converted to the enthalpy change upon injection of sample. The Gibbs free energy can then be calculated using the following equation … the entropy contribution can thus be calculated, providing all thermodynamic parameters for the interaction. While ITC is a powerful technique, it does require large quantities of both ligand and protein and typically can only be used for interactions ranging from nM to mM range (Freyer and Lewis, 2008).

2 Materials and Methods

2.1 Sub-cloning

The gene sequence for residues 24-450 of hIRE1 was amplified by PCR, with LIC specific overhangs, from a pBS vector containing the full sequence of hIRE1 (donated by Prof. David Ron) and purified by with an Invotrogen PCR purification kit, as per the manufacturers protocol. An empty pET vector containing the gene sequence for an N-terminal 6xHis-TEV tag (Vector 1B: Addgene) was linearised with the SspI restriction enzyme by incubating for one hour at 37°C. The linearised vector was purified by excising the linear form of the vector with a sterile blade from a 1% agarose gel and purifying using a QIAquick gel extraction kit, as per the manufacturers protocol. Vector and insert overhangs were created by treating both elements individually with T4 Polymerase and either 2.5 mM dGTP or dCTP respectively, at room temperature for 30 minutes. The reaction was terminated by heating at 75°C for 20 minutes. A 1:1 ratio of vector:insert was mixed and 1 µl of this mixture was transformed into DH5α *Escherichia coli*. Six colonies were picked and sequenced to confirm the correct orientation of the insert.

2.2 Recombinant Protein Expression and Purification

2.2.1 Expression of BiP and all IRE1-LD constructs

IRE1-LD and all its constructs were expressed using the expression vector (and variations thereof) detailed in section 2.1. Hamster BiP was expressed with a non-cleavable His-Tag using the protocol detailed in section 2.2.1.1. The His-tag has been demonstrated not to affect BiP activity (Wei and Hendershot, 1995). FL-BiP (1-641) and BiP-NBD (1-413) constructs were used alongside T229G and V461F mutants of FL-BiP. Similar expression protocols were used for both BiP and IRE1-LD constructs. For some IRE1 constructs however, some modification was required. The two basic protocols are described below and the protocols used for each constructs are detailed in Table 2.1.

2.2.1.1 37°C expression

All precultures and cultures contained 50 µg/ml of filter sterilised kanamycin sulphate. The desired expression vector containing the protein of interest was transformed into BL21 (DE3) *Escherichia coli*. A single colony from the resulting plate was used to inoculate 5 ml of LB media and was grown overnight at 30°C with shaking. This 5 ml preculture was used to

inoculate 50 ml of LB media to produce a pre-culture of OD 0.1 and was grown at 37°C with shaking at 250 rpm until the OD was ~1 (~2 hours). The entirety 50 ml pre-culture was then used to inoculate 500 ml of LB media and grown for approximately 2 hours under the same conditions, until the culture reached ~0.8-1. The culture was then induced with 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and grown for 4-5 hours.

2.2.1.2 20°C expression

All precultures and cultures contained 50 µg/ml of filter sterilised kanamycin sulphate. The desired expression vector containing the protein of interest was transformed into BL21 (DE3) *Escherichia coli*. A single colony from the resulting plate was used to inoculate 5 ml of LB media and was grown at 37°C with shaking until the OD ~ 1 (~5 hours). The entirety of the 5 ml pre-culture was used to inoculate 50 ml of LB media to produce a pre-culture of OD 0.1 and was grown at 37°C with shaking at 250 rpm until the OD was ~1 (~2 hours). The entirety of the 50 ml pre-culture was then used to inoculate 500 ml of LB media and grown for approximately 2 hours under the same conditions, until the culture reached ~0.8-1. The temperature of the incubator was dropped to 20°C and the culture was induced with 1 mM IPTG and grown overnight.

37°C expression protocol	**20°C expression protocol**
FL-IRE1-LD (WT)	IRE1 cLD (WT)
IRE1-LD (C109S)	IRE1-LD (D123P)
IRE1-LD (C148S)	
IRE1-LD (C332S)	
BiP (all constructs)	

Table 2.1

Table showing expression protocols used for each protein/construct used

2.2.1.3 Expression of ^{15}N labelled IRE1-LD

Protein expression was carried out as per the specified expression protocol, with the following adjustments. Following the 5 ml preculture, M9 media was used instead of LB media.

2.2.1.4 Expression of U{^{15}N, ^{2}H ^{12}C}, $^{13}CH_3$-Ile$^{\delta1}$ and U{^{15}N, ^{2}H ^{12}C}, $^{13}CH_3$-Ile$^{\delta1}$, $^{13}CH_3$-Ala$^{\beta}$, $^{13}CH_3$-Leu$^{\delta}$ and $^{13}CH_3$-Val$^{\gamma}$ labelled BiP

Expression of U{^{15}N, ^{2}H ^{12}C}, $^{13}CH_3$-Ile$^{\delta1}$ and U{^{15}N, ^{2}H ^{12}C}, $^{13}CH_3$-Ile$^{\delta1}$, $^{13}CH_3$-Ala$^{\beta}$, $^{13}CH_3$-Leu$^{\delta}$ and $^{13}CH_3$-Val$^{\gamma}$ labelled samples was performed using a previously published protocol (Tugarinov et al., 2006) with minor modifications. A single colony of BL21 (DE3) *E.coli* containing the desired plasmid was used to inoculate 5 ml of LB media and grown at 37°C with shaking at 250 RPM for until OD ~1. 20 ml of LB media was then added and the culture was grown for ~1 hour. The cells were pelleted from the resulting culture and all LB media was thoroughly removed. The cells were resuspended in 50 ml of M9 media in D_2O (OD ~0.1) and incubated overnight at 30°C. The resulting culture was then used to inoculate 200 ml of M9 media in D_2O and incubated at 37°C until OD ~0.7-0.8. The desired precursors were then added to the culture, which was incubated for a further 1 hour and then induced with 1 mM IPTG. This was then incubated for 6-7 hours, after which the cells were harvested.

2.2.1.4.1 Labelling Precursors

Isoleucine only: Alpha-ketobutyric acid (Methyl-^{13}C, 99%; 3,3-D2, 98%, Cambridge Isotopes Laboratories) (70 mg/L of M9 media, dissolved in D_2O)

Isoleucine, Leucine, Valine and Alanine: Alpha-ketobutyric acid, sodium salt (Methyl-^{13}C, 99%; 3,3-D2, 98%, Cambridge Isotopes Laboratories) (60 mg/L of M9 media, dissolved in D_2O), Succinate-D_4 (Cambridge Isotopes Laboratories) (2.5g/L of M9 media, dissolved in D_2O), Alpha-ketoisovaleric acid, sodium salt (3-methyl-^{13}C; 3-3-D_2, Cambridge Isotopes Laboratories) (120mg/L of M9 media, dissolved in D_2O) and [2-D,3-^{13}C]-L-alanine (Cambridge Isotopes Laboratories) (800mg/L of M9 media, dissolved in D_2O).

2.2.2 Cell harvesting and purification

Cells were harvested by centrifugation at 5000 RPM for 30 minutes with low rotor acceleration and deceleration. The cell pellet was resuspended in binding buffer (20 mM HEPES, 400 mM NaCl, pH 8.0) and frozen at -80°C. Resuspended cell pellets were thawed at room temperature with running water and, upon thawing, were incubated on ice for 30

minutes with protease inhibitor (cOmplete EDTA Free, Roche) and 5 mg of lysozyme. Cells were then lysed by sonication for 4 minutes of total sonication time using a six seconds on, ten seconds off cycle. The cell lysate was then collected by centrifugation at 20,000 g for 30 minutes and filtered using a 0.45 µM filter. The filtered cell lysate was purified using immobilised metal affinity chromatography (IMAC) with a Ni^{2+} column (GE Healthcare Histrap FF 1ml).

2.2.2.1 IRE1-LD purification (all constructs)

A 1 ml HisTrap FF column (GE Healthcare Life Sciences) was conditioned with binding buffer (20 mM HEPES, 400 mM NaCl, pH 8.0) before filtered cell lysate addition. Following cell lysate addition, the column was washed with 10 volumes of wash buffer (20 mM HEPES, 400 mM NaCl, 20 mM imidazole, pH 8.0) to remove non-specifically bound proteins, followed by an elution step where a gradient was applied over 10 column volumes from 20 mM imidazole to 500 mM imidazole. Columns were then washed with 10 ml of 1M imidazole. Samples were dialysed into 20 mM HEPES, 100 mM KCl, 5 mM $MgCl_2$, 0.02% NaN_3, pH 7.6 (HMK).

2.2.2.2 BiP purification (all constructs)

BiP constructs were purified using the protocol detailed in (Wieteska et al., 2017). A 1 ml HisTrap FF column (GE Healthcare Life Sciences) was conditioned with binding buffer (20 mM HEPES, 400 mM NaCl, pH 8.0) before filtered cell lysate addition. Following cell lysate addition, the column was washed with 10 column volumes of wash buffer (20 mM HEPES, 400 mM NaCl, 40 mM imidazole, pH 8.0) to remove non-specifically bound proteins, followed by an elution step of 500 mM imidazole. Columns were then washed with 10 ml of 1M imidazole. Samples were dialysed into 20 mM HEPES, 100 mM KCl, 5 mM $MgCl_2$, 0.02% NaN_3, pH 7.6 (HMK).

2.2.3 Expression and purification of recombinant TEV protease

To ensure production of soluble TEV tobacco etch virus (TEV) protease, TEV was expressed as a 6xHis-MBP-TEV fusion protein with a S219V mutation as described in (Tropea et al., 2009). A plasmid containing the gene sequence for a fusion protein containing 6xHis-MBP-TEV was transformed into BL21(DE3) and plated onto a LB-agar plate with carbenicillin resistance. Three colonies were used to inoculate 100 ml of LB media containing carbenicillin and the resulting cultures were grown overnight at 30°C with shaking.

50 ml of the 100ml pre-cultures was used to inoculate 2 x 1 L of LB media containing carbenicillin and the resulting cultures were grown at 37°C until OD600 = 0.6. 20 mM of proline and 300 mM of NaCl was then added to each flask and the temperature was reduced to 30°C. After 30 minutes the protein was induced by addition of 0.5 mM IPTG and grown for 4 hours. Cells were then harvested by centrifugation at 4000 rpm for 10 minutes and resuspended in filtered 50 mM NaH_2PO_4 200 mM NaCl, 25 mM imidazole, 10% glycerol, pH 8.0 (IMAC A buffer). Cells were then flash frozen and stored at -80°C prior to purification.

2.2.3.1 Purification of recombinant TEV protease

Frozen cells were thawed at room temperature with the aid of running water. Once thawed, cells were kept on ice. Cells were then lysed by sonication for 4 minutes of total sonication time using a six seconds on, ten seconds off cycle. The cell lysate was then collected by centrifugation at 20,000 g for 30 minutes and filtered using a 0.45 µM filter. Following equilibration of the Ni^{2+} column (GE Healthcare Histrap FF 1ml), the filtered cell lysate was applied to the IMAC column. Following application of the cell lysate, the column was washed with 100 ml of IMAC A buffer, then a gradient was applied with 50 mM NaH_2PO_4 200 mM NaCl, 250 mM imidazole, 10% glycerol, pH 8.0 (IMAC B) from 0-100% of IMAC B over 400 ml. Next, 100 ml of IMAC B was used to wash the column, followed by 100 ml of IMAC A.

Fractions containing TEV protease were pooled and diluted with 50 mM NaH_2PO_4 200 mM NaCl, 10% glycerol, pH 8.0 (dilution buffer) so that the imidazole concentration was <35 mM. The pooled fractions were reloaded onto the IMAC column and run using the same protocol described above with a gradient of 200 ml rather than 400 ml. The resulting fractions were pooled and EDTA and DTT were added to a final concentration of 2 mM and 5 mM respectively. The pooled fractions were then washed with 50 mM NaH_2PO_4 100 mM NaCl, 10% glycerol, pH 8.0 and concentrated using an 10 MWCO Amicon until imidazole was no longer present in the flow through. The concentrated protein was centrifuged, and the supernatant was removed. The protein concentration was calculated by measuring the A280 nm and using the Beer-Lambert equation with an extinction coefficient of 32290 M^{-1} cm^{-1}. The resulting solution was stored in 50% glycerol and stored at – 20°C.

2.3 Preparation of Aβ oligomers

16 mg of lyophilised $A\beta_{1-42}$ (DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA) or scrambled Aβ (AIAEGDSHVLKEGAYMEIFDVQGHVFGGKIFRVVDLGSHNVA) was reconstituted in 3545 µl of 1,1,1,3,3,3-hexafluoroisopropanol (HFIP) and mixed by vortexing. The reconstituted protein was allowed to stand at room temperature for 30 minutes until the solution became clear. 110 µl aliquots of the protein were dispensed into 1.5 ml LoBind Eppendorf tubes, centrifuged at maximum speed for five minutes and then allowed to dry under nitrogen to form a peptide film. These peptide films were stored in a desiccator jar at -20°C.

Prior to use, two Eppendorf tubes containing peptide films were removed from -20°C and equilibrated at room temperature for 30 minutes. Each peptide film was then redissolved in 250 µl of DMSO (500 µl total) and vortexed to ensure the film was fully dissolved. The films were then buffer exchanged into Dulbecco's PBS buffer (ThermoFisher Scientific) by passing through a 1 ml HiTrap column, pre-equilibrated with PBS buffer. Aβ1-42 was then eluted in ~2 ml fractions with each fraction quantified by the Bradford assay. The protein was then diluted to 50 µM and allowed to oligomerise for 1 hour with gentle shaking at room temperature. After 1 hour the size of the oligomeric species were observed by DLS.

2.4 Preparation of peptide solutions

Soluble ΔEspP (MKKHKRILALCFLGLLQSSYSAAKKKK) (Gardner, B.M. and Walter, 2011) and P2 (HTFPAVL) (Marcinowski et al., 2011) were purchased from Biomatik and resolubilised to a final concentration of 1 mM in HMK buffer.

2.5 Circular Dichroism

0.2 mg/ml µM of protein in 50 mM Na_2HPO4, 100 mM NaCl, pH 8.0 was scanned from 180 – 260 nm at 1 nm intervals at 20°C. Spectra were acquired in duplicate and secondary structural elements were determined using K2D3 (Louis-Jeune et al., 2012).

2.6 Mass-spectrometry

2.6.1 Peptide mass fingerprinting

Purified FL-IRE1-LD was run in a reducing buffer by SDS-PAGE and stained with Coomassie blue. A single band was excised using a sterile razor blade and stored in a sterile Eppendorf tube in 20% EtOH. The gel slice was submitted to the mass spectrometry facility where further

analysis was carried out by Dr. James Ault. The gel was resolubilised and subjected to a trypsin digest and analysed by ESI-MS. The m/z ratio of the resulting fragments were used to search the Swiss-Prot database (Boeckmann et al., 2003) containing fragmentation patterns of known proteins and our 6xHis-TEV-IRE1-LD construct sequence.

2.6.2 Molecular mass determination

The molecular mass of FL-IRE1-LD was determined by ESI-MS. 15 µM of IRE1-LD in HMK buffer in the presence and absence of 5 mM DTT was submitted to the mass spectrometry where analysis was performed by Dr. James Ault in the mass spectrometry facility. 1 % v/v formic acid and 50 % v/v acetonitrile was added to the sample prior to analysis

2.6.3 Native-MS

FL-IRE1-LD was buffer exchanged into 200 mM ammonium acetate (pH 7.6) in the presence and absence of 5 mM DTT. Samples were then submitted to the mass spectrometry facility where they were analysed by ESI-MS by Dr. James Ault

2.6.4 Ion mobility mass spectrometry

FL-IRE1-LD samples were prepared as for Native-MS samples and submitted to the mass spectrometry facility where all further analysis and calibration was carried out by Dr. James Ault. ESI-TWIMS-MS measurements were performed on 15 µM FL-IRE1-LD using a Synapt HDMS mass spectrometer (Waters Corp) in helium buffer gas. Collision cross sections were calculated by linear regression against a calibration curve, constructed using avidin, conavalin A, alcohol dehydrogenase and pyruvate kinase.

2.6.5 Affinity mass spectrometry

Affinity mass spectrometry methods were based on methods previously reported in the literature (Waters et al., 2008). 1 µM of each lead compound in PBS puffer was added to either side of the dialysis membrane of a Rapid Equilibrium Device (RED) (Thermo Scientific) with an 8.0 kDa molecular weight cut off. To a separate RED, 1 µM of the lead compound in the presence of 50 µM Aβ preparation was added inside the membrane and 1 µM of lead compound was added outside the membrane. The devices were equilibrated for 1 hour, after which, 500 µl was taken from outside the dialysis membrane. 50 µl was of the sample was analysed by LC-MS using an Agilent 1200 HPLC system, interfaced to an ESI MSD quadrupole mass spectrometry (Agilent Technologies). Samples were separated using a Gemini C18 (50 x

2.0 mm, 3 µm pore size). A linear gradient elution using 10 mM ammonium hydrocarbonate (pH 9) (solvent A) and acetonitrile (solvent B) was applied over three minutes, ranging from 10-100% of solvent B at a flow rate of 1 ml/min, with a flow rate of 250 µl/min applied to the MS detector and 750 µl/min to the waste reservoir. Affinity mass spectrometry samples were run in selected ion monitoring (SIM) mode, and the peak position was compared to the same compound dissolved in a 1:1 ratio of H_2O:MeCN at 50 µM, run in full spectrum mode. The resulting chromatograms were analysed by Chemstation software (Agilent Technologies).

2.7 Molecular Dynamic Simulations

Missing components of the hIRE1 crystal structure (residues 24-450) PDB ID:2HZ6 (Zhou et al., 2006) were rebuilt using the I-TASSER web server (Zhang, Y., 2008). Prior to the simulation, a with a CHARMM27 all-atom forcefield with CMAP (version 2.0) was applied to all atoms and water was added to a cubic box using the TIP3P model. Sodium and chloride ions were added to the box to neutralise the charge and the structure was subject to a steepest decent energy minimisation step, followed by two restrained 100 ps simulations to equilibrate the temperature and pressure of the system. The energy minimised structure was then subject to a 500 ns all atom MD simulation using GROMACS (Berendsen et al., 1995).

2.7.1 Molecular Dymamics Simulation Analysis

The analysis of all MD simulations was carried out in GROMACS. Overall deviation from the starting structure was measured by root mean square deviation (RMSD) and the root mean square fluctuation (RMSF) were calculated using the cartesian coordinates of the Cα atoms.

2.7.2 Collision Cross Section Analysis

Collision cross section analysis of the 500 ns MD trajectory was performed using IMPACT (Marklund et al., 2015) with projection approximation values stated.

2.8 NMR

2.8.1 Protein NMR

All samples were run in HMK buffer (with 5 mM DTT or 1 mM TCEP where stated) and supplemented with 1% v/v of 100 mM AEBSF and 4% v/v D_2O. All samples were run on a Bruker 950 MHz spectrometer with a cryoprobe, except where stated. Methyl-SOFAST-HMQC of BiP NBD was performed at 298K (Schanda and Brutscher, 2005) with a delay of 0.3 seconds

between scans. (^{15}N, ^{1}H)-BEST-TROSY experiments of IRE1-LD in results chapter one were performed at 298K or 288K (Schulte-Herbruggen and Sorensen, 2000; Lescop et al., 2007) with a delay of 0.4 seconds between scans. (Tugarinov et al., 2006). All spectra were processed using NMRPipe (Delaglio et al., 1995) and analysed using CCPN Analysis v2.4.2 (Vranken et al., 2005).

2.8.1.1 Signal to noise analysis

For analysis of peak intensities, the peak height, calculated using the parabolic method, was divided by the spectrum noise, calculated using methods implemented in CCPN Analysis v2.4.2 (Vranken et al., 2005).

2.8.1.2 Chemical shift perturbation analysis

Chemical shift perturbations (CSPs) were calculated using isotope weighted shift differences in CCPN Analysis v2.4.2 (Vranken et al., 2005) using Equation 2.1.

$$CSP = \sqrt{\Delta H^2 + (\Delta N \times 0.154)^2}$$

Equation 2.1

2.8.1.3 MTSL labelling of IRE1-LD

To 50 µM of FL-IRE1-LD in HMK buffer, 5 mM of DTT was added. After incubation for >30 minutes at room temperature, 0.5 ml of FL-IRE1-LD was buffer exchanged into HMK buffer using a pre-equilibrated NAP-5 column (GE Healthcare Life Sciences) as per the manufacturer's protocol. Immediately following elution, 50 µl of a 40 mM solution of *(1-oxyl-2,2,5,5-tetramethylpyrroline-3- methyl)methanethiosulfonate* (MTSL) (Berliner et al., 1982) dissolved in MeCN was added. The solution was incubated overnight at +4°C, following which, excess MTSL was removed by dialysis. Labelling was confirmed by mass spectrometry.

2.8.2 Compound library screening by ^{19}F NMR

All spectra were acquired at 600 MHz on a Bruker spectrometer with a QCI cryoprobe. Standard 1D ^{1}H and ^{19}F spectra were acquired for reference spectra of all compound. Identification of hit compounds was performed using a standard ^{19}F pulse sequence with proton decoupling during acquisition, a 3 second delay between scans and a 400 ms CPMG time. ^{1}H CPMG experiments were performed with a 3 second delay between scans and a 300 ms CPMG time.

2.8.2.1 Sample preparation

10 mM of each small molecule arrived pre-dissolved in DMSO in a 96 well plate. Samples for analysis were prepared using a PAL sample preparation robot by diluting each small molecule to 50 µM in Dulbecco's PBS buffer (ThermoFisher), before adding the sample to a rack of 96 NMR tubes prior to analysis. Experiments were then carried out using a SampleJet robot. For experiments performed in the presence of Aβ oligomers, the sample containing the small molecule was split, with half analysed in the presence of 1 µM (concentration based on initial concentration of Aβ monomer, following solubilisation of peptide film) of the preparation, and half in the absence. Spectra were processed, and noise and peak height values were calculated and compared between both samples using Mestrenova.

2.8.2.2 Mixture preparation

The contents of mixtures of compounds were initially prepared using NMRMix (Stark et al., 2016), however to avoid multiple NMR experiments due to the wide chemical shift range of the ^{19}F nucleus, some manual adjustment of mixtures was performed.

2.8.2.3 Initial compound screening

Initial 1D ^{1}H and ^{19}F spectra of 50 µM each compound was acquired sequentially using a SampleJet robot (Bruker). Compounds whose spectra did not produce a signal to noise ratio of >10 or contained unexpected peaks, indicative of degradation or incorrect compound assignment, were discarded from the library.

2.8.2.4 Screening in the presence of Aβ preparation

1D ^{19}F spectra with a CPMG sequence were acquired of mixtures in the presence and absence of 1 µM Aβ preparation.

2.8.2.5 Competition assays with bexarotene

50 µM of each compound was added to a mixture containing an equimolar concentration of bexarotene and 1 µM Aβ preparation. An equivalent volume of DMSO was added to control samples in the absence of bexarotene.

2.9 Size Exclusion Chromatography

150 µl of IRE1 in HMK buffer in the presence and absence of 5 mM DTT was injected onto a Superdex 200 10/300 GL column (GE Healthcare Life Sciences). Samples were eluted at a flow

rate of 0.35 ml/min with the flow through monitored by UV at 280 nm. All runs were performed at +4°C.

2.9.1 Preparation of molecular weight calibration curve

Ovalbumin, conalbumin, aldolase, ferritin and thyroglobulin protein standards, purchased as part of a HMW Gel Filtration Calibration Kit (GE Healthcare Life Sciences), were individually resolubilised in HMK buffer to a concentration of 20 mg/ml. 150 µl of a mixture containing 4 mg/ml ovalbumin, 3mg/ml conalbumin, 4 mg/ml aldolase, 0.3 mg/ml ferritin and 5 mg/ml thyroglobulin was injected onto a Superdex 200 10/300 GL column (GE Healthcare Life Sciences). The retention volume was monitored UV at 280 nm, which was used to construct a calibration curve (Figure 2.1.)

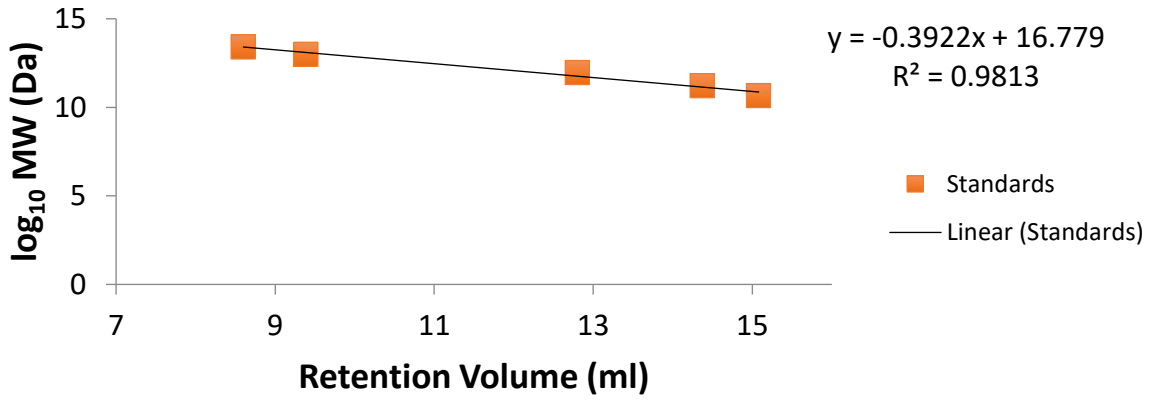

Figure 2.1

Molecular weight calibration curve prepared from eluted protein standards.

2.9.2 Estimation of hydration radius of folded standards

The hydration radius (R_h) of the folded protein standards was estimated using , where N is the number of residues in the protein (Marsh and Forman-Kay, 2010).

$$R_h = 4.92N^{0.285}$$

Equation 2.2

Following estimation of the molecular weight of the protein standards, a calibration curve was prepared to estimate the hydration radius of eluted proteins.

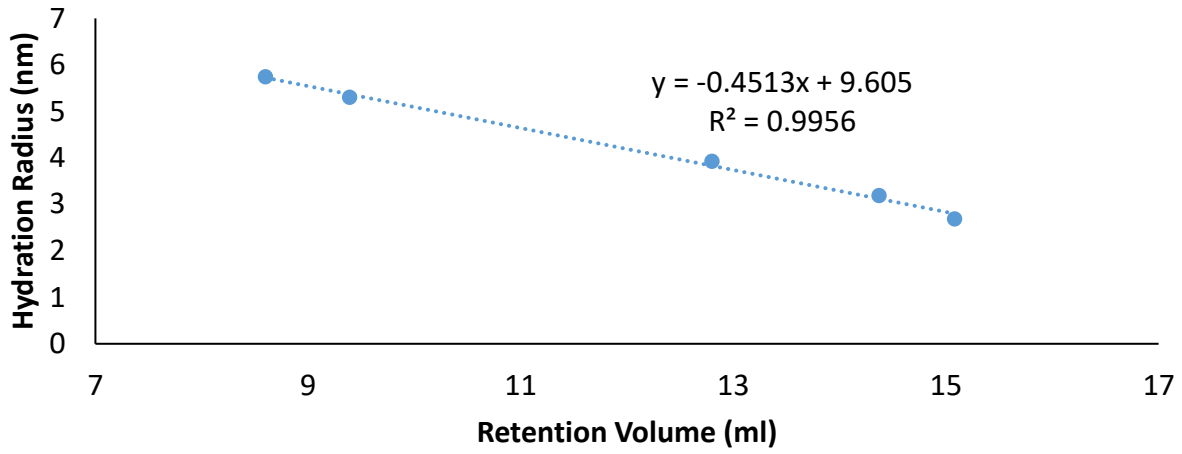

Figure 2.2

Hydration radius calibration curve based on the estimated hydration of the folded protein standards

2.9.3 Estimation of the molecular weight of FL-IRE1-LD

As FL-IRE1-LD contains both structured and disordered regions (Zhou et al., 2006), it was not possible to estimate the molecular mass of the protein using either folded or unfolded calibration standards. We therefore calculated the mass of FL-IRE1-LD based on equation 2.2, where MW_{folded} is the estimated molecular mass of FL-IRE1-LD based on the folded protein calibration curve, $MW_{folded\ D123P}$ is the estimated molecular mass of IRE1-LD (D123P), based on the same calibration curve, and MW_{D123P} is the mass of monomeric IRE1-LD (D123P).

$$MW_{FL-IRE1-LD} = \frac{MW_{folded}}{MW_{folded\ D123P}} \times MW_{D123P}$$

Equation 2.3

2.9.4 Estimation of eluted concentration of FL-IRE1-LD

One of the fundamental properties observed during any chromatographic procedure is that of band broadening, thus the concentration of a molecule upon elution from a column is less than that upon injection. As FL-IRE1-LD was eluted primarily as one peak, we estimated the concentration of FL-IRE1-LD based on equation 2.3, where C_o is the concentration upon

injection, V_0 is the injection volume, V_{peak} is the volume of FL-IRE1-LD eluted from the column and $C_{elution}$ is the eluted concentration of FL-IRE1-LD.

$$C_{elution} = \frac{V_0}{V_{elution}} \times C_0$$

Equation 2.4

2.10 Dynamic Light Scattering

2.10.1 IRE1-LD oligomerisation

250 µl of 0.22 µM filtered 5 µM IRE1-LD (+ 10 µM of Esp peptide where stated) diluted in 20 mM HEPES, 100 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP pH 7.6, was injected into a Wyatt miniDawnTreos® system (equipped with an additional DLS detector). Data was analysed using Astra 6.0.3® software. Three minutes of data following sample injection was used for analysis by the software. This was preceded and proceeded by a five minute baseline of 0.22 µM filtered and degassed buffer. Following each sample injection, the flow cell was flushed with 0.22 µM filtered and degassed 1M nitric acid and Milli-Q grade H_2O. Correlation curves were analysed using the Astra 6.0.3 software by regularization methods (Hassan et al., 2015). Hydration radii calculated assume a spherical model and are stated ±SD of the peak. Where peaks were not resolved at the baseline, peaks were separated at the minima between peaks.

2.10.2 Aβ oligomers

100 µl of the Aβ preparation was added to a 96 well plate and analysed using a Wyatt DynaPro plate reader. Correlation curves were analysed using the Dynamics software by regularization methods (Hassan et al., 2015).

2.11 Microscale Thermophoresis MST

All MST experiments were performed on the NanoTemper Monolith NT 1.15. MST power was 40% and LED power was 100%. 25 µM FL-IRE1-LD in HMK buffer (without NaN_3) was labelled in a 1:2 molar ratio with fluorescein isothiocyanate dissolved in DMSO (IRE1:FITC) by incubation on ice for 1 hour. Excess dye was removed by size exclusion (NAP-5, GE Healthcare). Experiments were carried out in HMK buffer. 1 µM of IRE1 was pipetted at an equal volume into a dilution series of unlabelled binding partner consisting of sixteen twofold dilutions. The thermophoresis and T-Jump values were plotted and binding curves were calculated using NanoTemper Analysis 1.2.101 using the nonlinear solution of the law of mass

action. Dissociation constants reported were an average of three independent experiments ±SEM.

2.12 Isothermal Titration Calorimetry

All experiments were performed using the Microcal ITC200 system. 0.4 mM of ATP or ADP was titrated into the cell containing 40 µM of BiP T229G in matched HMK buffer or HMNa buffer. The heat upon dilution was determined by titrating matched buffer into 40 µM BiP T229G and was subtracted from all curves. Data was fitted using the non-linear least squares curve fitting algorithm in MicroCal Origin (OriginLab, Northampton, MA). All experiments were performed at 25°C and data reported is an average of three repeats ± standard deviation.

2.13 Rate of disulphide formation by SDS-PAGE

5 mM DTT was added to 25 µM protein sample in degassed HMK buffer and incubated at room temperature for at least 30 minutes. 500 µl of protein sample was added to a NAP-5 column pre-equilibrated with HMK buffer, and eluted to with 1 ml of the same buffer. Gel samples were taken at specified time points, relative to addition of protein to the NAP-5 column. Bands were quantified using ImageJ (Schneider et al., 2012). The percentage of monomer in the non-reduced sample was calculated relative to the intensity of the monomeric band in the reduced lane. Results stated are an average of three repeats ± standard deviation. A one-way ANOVA was performed between all gels to assess if there was any statistically significant difference between all samples with follow up tests comparing WT protein to all other samples performed using the Dunnett method. Adjusted p values are stated.

2.14 Fluorescence Polarisation

Protein samples were labelled as previously described for MST experiments. 50 nM of labelled protein sample was added to an equal volume dilution series of unlabelled binding partner consisting of twelve twofold dilutions. Parallel and perpendicular fluorescence readings were analysed using a BMG POLARstar OPTIMA plate reader and polarisation readings were calculated based on equation 2.4.

$$P = \frac{F_{parallel} - F_{perpendicular}}{F_{parallel} + F_{perpendicular}}$$

Equation 2.5

2.15 OD400 Assay

To a 384 well plate, 20 µM of IRE1-LD, 40 µM of BiP (various constructs), 40 µM ΔEspP and 10 mM ATP was added to a total volume of 90 µl. For experiments in the absence of BiP, ΔEspP and/or ATP, an equivalent volume of HMK buffer was added. The plate was covered and left to incubate at room temperature for 1 hour, after which the OD400 was analysed using a HIDEX sense plate reader. Values reported are an average of five independent readings ± standard deviation and adjusted p-values stated were calculated by a one-way ANOVA with multiple comparisons follow up tests comparing IRE1-LD + ΔEspP to all other samples performed using the Dunnett method.

2.16 ATPase Assay

A modified protocol of the previously published malachite green ATPase assay was used (Chang et al., 2008) to determine the ATPase rate of BiP. A 2:1:1:2 ratio of malachite green (0.081% w/v in d.H_2O), poly(vinyl alcohol) (2.3% w/v in d.H_2O), ammonium molybdate tetrahydrate (5.7% w/v in 6M HCl) and d.H_2O was mixed to prepare the malachite reagent.

To a 96 well plate, 5 µM of FL-BiP (WT), 2.5 mM of ATP was added (final volume 25 µl) and mixed. Autohydrolysis, protein blank, and buffer blanks were prepared in the absence of FL-BiP (WT), ATP and both FL-BiP (WT) and ATP respectively. A phosphate calibration curve was also prepared by preparing 10 twofold dilutions of 0.5 mM Na_2HPO_4 dissolved in HMK or HMNa in triplicate. The plate was covered and incubated at 37°C for 1 hour. After incubation, 80 µl of the malachite reagent and 10 µl of sodium citrate (34% w/v in d.H_2O) was added and mixed. The plate was then covered and incubated at 37°C for a further 15 minutes, after which the A620 was measured using a HIDEX sense plate reader.

ATPase rates stated were determined by calculating the molar concentration of phosphate produced by an equal molar concentration of FL-BiP (WT) sample and dividing by 60 to produce an ATPase rate. All values stated are from five independent readings ± standard deviation and p-values stated are from an unpaired two tail t-test.

2.17 General Bacterial Methods

2.17.1 Sterilisation

All solutions were sterilised by autoclaving at 121°C for 20 minutes or by filtration through a 0.22 µm filter into a sterile vessel.

2.17.2 LB plate preparation

A solution of 25g/L of Lysogeny Broth media and 15 g/L of agar was sterilised by autoclave and allowed to cool until the vessel was ~60°C (warm to the touch). A suitable amount of a 1000x solution of 100 mg/ml of carbenicillin or 50 mg/ml of kanamycin sulphate was added to the solution and the solution was poured into a sterile petri dish and allowed to solidify.

2.17.3 LB media preparation

A 25g/L solution of Lysogeny Broth media was prepared and sterilised by autoclave and allowed to cool until <60°C.

2.17.4 M9 media preparation

6.5g/L of Na_2HPO_4, 3.0g/L of KH_2PO_4, 0.5g/L of NaCl, 1.0g/L of $^{15}NH_4Cl$, 2g/L of glucose, 2.5g/L of LB media, 120 mg/L of $MgSO_4$, 11 mg/L of $CaCl_2$ and 10 mg/L of biotin and thiamine were dissolved in d.H_2O and sterilised by filtration. For U{^{15}N, ^{2}H ^{12}C}, $^{13}CH_3$-Ile$^{\delta1}$ and U{^{15}N, ^{2}H ^{12}C}, $^{13}CH_3$-Ile$^{\delta1}$, $^{13}CH_3$-Ala$^{\beta}$, $^{13}CH_3$-Leu$^{\delta}$ and $^{13}CH_3$-Val$^{\gamma}$ labelling, glucose was replaced with D-glucose (1,2,3,4,5,6,6-D7, 98%) (Cambridge Isotopes Laboratories), LB media was replaced with 1g Celtone base powder (D,97%+; ^{15}N, 98%+) (Cambridge Isotopes Laboratories) and d.H_2O was replaced with D_2O (.

2.18 Transformation of E. coli strains

A 50 µl aliquot of DH5α or BL21 (DE3) was removed from -80°C storage and thawed on ice for ~10 minutes. A specified amount of plasmid DNA or PCR product was added to the *E. coli* aliquot and incubated on ice for 20 minutes. Cells were then removed from ice and heat shocked at 42°C for 45 seconds and returned to ice. 500 µl of LB media was added to the transformed *E. coli* cells and subsequently, was incubated with shaking at 37°C for 1 hour to grow out antibiotic resistant proteins. Cells were then harvested by centrifuging at 3000 rpm for 2 minutes and 400 µl of the supernatant was removed. The remaining supernatant was used to resuspend the cells and the resulting culture was pipetted onto a LB-agar plate,

containing a suitable antibiotic, and streaked with a sterile plate spreader. Plates were then incubated overnight at 37°C.

2.19 General DNA methods

2.19.1 Agarose Gel

1% agarose gels were prepared by adding 1 g of UltraPure Agarose to 100 ml of TAE buffer. Solutions were dissolved by microwaving and the solution was allowed to cool. 2 µl of 10 mg/ml ethidium bromide was added to the agarose solution and mixed. This solution was then poured into a casting tray with a comb in place and left until solid. TAE buffer was then added to the gel tray until covering the agarose gel, followed by 2 µl of EtBr. A mixture containing 1 µl of DNA, 8 µl of d.H_2O and 1 µl of 10 x loading buffer (New England Biolabs) was added to each well with a mixture containing 1 µl 1 kb DNA ladder and 5 µl of 6 x loading buffer (New England Biolabs). Gels were run at a constant voltage of 100 V for 1 hour and, for analytical gels, analysed using a UV gel dock.

2.19.2 Plasmid purification

A single colony of the DH5α cells containing the desired plasmid was used to inoculate 5 ml of LB media containing antibiotics and grown overnight at 30°C. The cells from the resulting culture were harvested by centrifugation at 10,000g for 10 minutes and all LB media was thoroughly removed. Plasmid was purified from the resulting cells using an Invitrogen Mini-Prep kit, as per the manufacturer's protocol.

2.19.3 DNA quantification

DNA concentration was measured by pipetting 1 µl of DNA onto a Nanodrop, following blanking with TE buffer, with the A260 value used to calculate the DNA concentration.

2.19.4 Mutagenesis

The basic mutagenesis protocol that was followed was as follows:

5 µl of 10 x Pfu Ultra HF Polymerase Buffer (Agilent)

1 µl of dsDNA template (~50-100 ng)

2 µl of primer mix (1 µl of forward and 1 µl of reverse primers, both reconstituted to a concentration of 1 µg/µl, mixed with 8 µl of sterile dd.H_2O)

1.5 µl of QuikSolution Reagent

39.5 µl of sterile dd.H_2O

1 µl of Pfu Ultra HF polymerase was then added and the solution was mixed with a pipette tip.

The resulting mixture was then transferred to a thermocycler and run using the following protocol, with annealing times adjusted depending on the primers used:

Initial denaturing step: 95°C – 2 minutes

Cycling:

Melting: 95°C – 30 seconds

Annealing: Dependent on primers – 30 seconds

Extension: 72°C – 6 minutes

(Above cycling steps repeated for 30 total cycles)

Final extension:

72°C – 10 minutes

Following PCR, the thermocycler was cooled to 4°C. 1 µl of DpnI was added to the reaction mixture to digest the parental DNA and incubated at 37°C for 1 hour. 1-10 µl of the reaction mixture was then transformed into DH5α. All mutants were then confirmed by sequencing.

2.19.5 Sequencing

All sanger sequencing was performed by Source Bioscience. Plasmids and primers were sent at volumes and concentrations specified on the Source Bioscience website and sequenced using sequencing primers detailed in the appendix, or T7 promoter (forward) and T7 terminator (reverse) primers provided by Source Bioscience.

2.20 General Protein Methods

2.20.1 Buffers

HMK – 20 mM HEPES, 100 mM KCl, 5 mM $MgCl_2$, 0.02% NaN_3 pH 7.6

HMNa – 50 mM HEPES, 75 mM NaCl, 10 mM $MgCl_2$, 0.02% NaN_3, pH 7.5

2.20.2 Dialysis

Proteins were dialysed using Pur-a-Lyzer Maxi dialysis devices (Sigma Aldrich) for 24 hours with at least three changes of dialysis buffer.

2.20.3 SDS-PAGE

To 900 µl of 4x Laemmli sample buffer, either 100 µl of beta-mercaptoethanol for reducing gel samples or 100µl of 10 mM N-ethylmaleimide for non-reducing gel samples, was added. Samples were added to the resulting mixture in a 3:1 ratio of sample buffer:protein and boiled on a heat block for 1 minute. A molecular weight marker and the gel samples were then applied to an SDS-PAGE gel (Bio-Rad Mini-Protean TGX Stain Free Pre-Cast Gels) and run at 180V until the gel front reached the bottom of the gel. Gels were then coomassie stained (Bio-Rad Bio-Safe Coomassie) as per the manufacturer's protocol or imaged using UV.

2.20.4 Protein Quantification

All IRE1-LD and BiP constructs were quantified by diluting 20 µl of protein in 380 µl buffer containing 6M guandine hydrochloride, measuring the absorbance by UV spectroscopy at 280 nm and applying the Beer-Lambert equation. An estimated extinction coefficient at 280 nm, calculated using the Expasy web server (Gasteiger et al., 2003), of 66935 M^{-1} cm^{-1} for IRE1-LD constructs, 29005 M^{-1} cm^{-1} for FL-BiP constructs and 17420 M^{-1} cm^{-1} for BiP-NBD constructs, was used to calculate the concentration.

TEV was quantified by direct addition to a cuvette and measuring the absorbance at 280 nm by UV spectrometry. An extinction coefficient of 32290 M^{-1} cm^{-1} was used to calculate the concentration using the Beer-Lambert equation.

For Aβ quantification, the Bradford assay was performed by preparing BSA standards in duplicate at 2 mg/ml, 1 mg/ml, 0.75 mg/ml, 0.5 mg/ml. 0.25 mg/ml and 0.125 and 0.025 mg/ml. To 360µl of Bradford reagent, 10 µl of each of these standards, PBS buffer and the collected Aβ fractions were added. The absorbance at 595 nm was measured and the concentration of Aβ calculated.

2.20.5 TEV Cleavage

To an A280 ratio of 1:100 TEV:IRE1-LD, 0.5mM EDTA and 1mM DTT was added. The reaction was incubated overnight at +4°C where the resulting sample was passed through an Amicon Concentrator of MWCO 30 kDa to concentrate the solution. Several washes were performed

with 20 mM HEPES, 400 mM NaCl, pH 8.0 to remove residual EDTA and DTT, prior to addition to HisTrap column to remove any residual His-Tag in solution.

3 Construct design and recombinant protein expression and purification of the IRE1-LD

As we are focussing on the activation of IRE1, it is necessary to perform a divide and conquer approach to perform structural, biophysical and biochemical characterisation of the initial steps of activation. This requires the design of several constructs of the luminal domain of IRE1, in addition to the use of well characterised mutants. As we will be performing a variety of *in vitro* methods, many of which require large quantities of protein and isotope labelling of IRE1, it would be cost effective to produce protein recombinantly using an *E. coli* expression system. The work in this chapter focusses on these preliminary steps to ensure we can characterise IRE1 activation in later chapters.

3.1 Construct Design and Sub-Cloning

A bacterial expression system was decided upon to express IRE1-LD as this would provide a cost-effective method to produce a high yield of recombinant protein. Previous literature work indicates that *E. coli* expression systems are suitable for structural studies of IRE1-LD (Zhou et al., 2006). In addition to this, the effect of post-translational modifications of IRE1-LD have been shown to be insignificant to protein function (Liu et al., 2002).

A divide and conquer approach was decided upon to focus on the initial steps of activation of IRE1 and to avoid complications of expressing a transmembrane protein. Furthermore, IRE1 has been shown to dimerise and oligomerise in the absence of its transmembrane domain and cytoplasmic domain, and in the absence of any membrane (Zhou et al., 2006) (Karagoz et al., 2017), thus it would be possible to study the activation of IRE1 in the absence of the transmembrane or cytoplasmic domain, or any membrane mimetics. It is worth noting that there is evidence that perturbations of the ER membrane can cause activation of the yeast homolog of IRE1 (Halbleib et al., 2017). To minimise the risk of compromising protein function, the vast majority of the luminal domain and the entirety of the linker region connecting the luminal domain to the transmembrane region (Figure 3.1). In addition, previously published data on constructs containing these residues have been shown to be active (Zhou et al., 2006). In order to express hIRE1 in *E.* coli, it was necessary to sub-clone the cDNA for the construct into a bacterial expression vector.

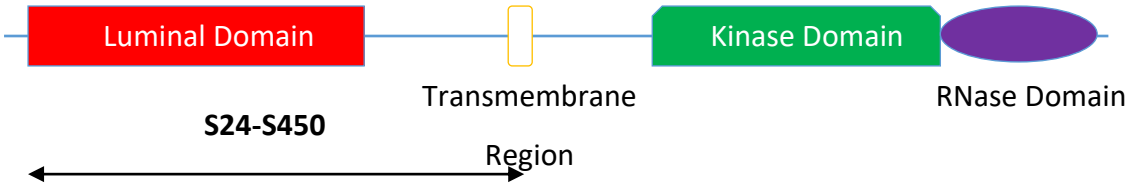
Luminal Domain
Kinase Domain
Transmembrane
RNase Domain
S24-S450
Region

Figure 3.1

A schematic of IRE1. The juxtramembrane connecting the luminal domain to the transmembrane region is shown in blue. The region of our IRE1-LD construct is indicated below the schematic.

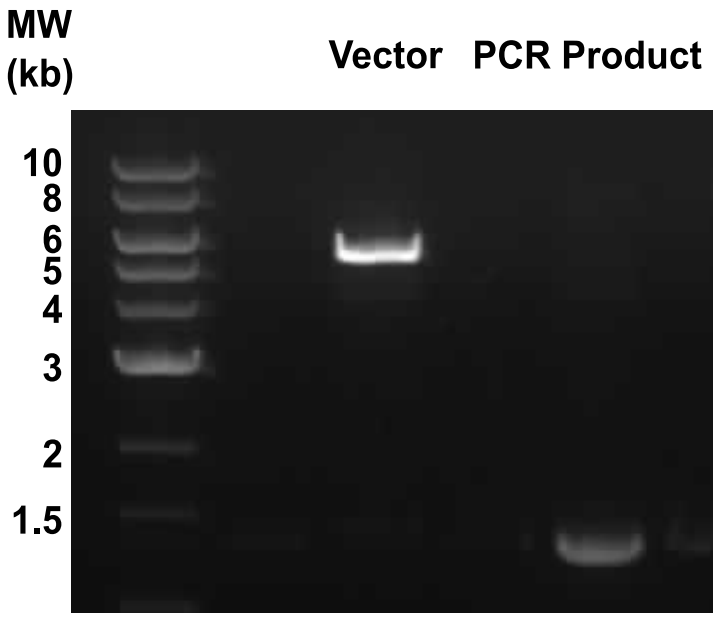

Figure 3.2

1% DNA Agarose gel of purified linearised vector 1B (vector) and insert containing the gene sequence for the IRE1-LD construct (PCR Product). Vector 1B is 5343 bp and the cDNA for IRE1-LD is 1321 bp

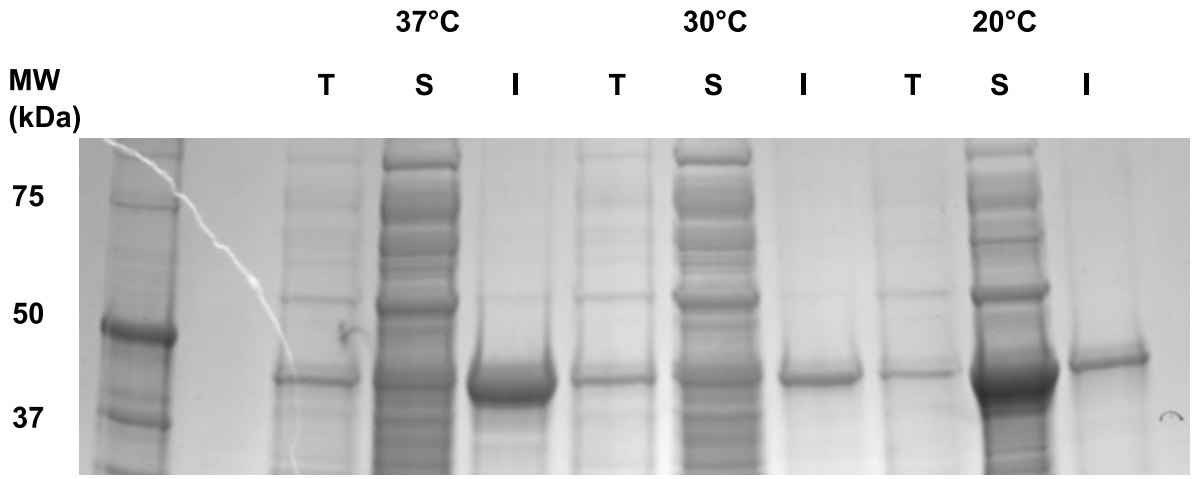

Figure 3.3

SDS-PAGE of products of a solubility test at various expression temperatures (indicated above) of IRE1-LD (D123P). The total protein content (T) and soluble (S) and insoluble (I) fractions are labelled. This indicates that unlike IRE1-LD (WT), IRE1-LD (D123P) is expressed in the insoluble fraction at 37°C

Ligation independent (LIC) (Aslanidis and Dejong, 1990) was used to sub-clone the cDNA for residues 24-450 of hIRE1 (1321 bp) into a pET vector (5343 bp) containing the gene for an N-terminal 6×His-TEV site (Figure 3.2), resulting in expression of a fusion 6×His-TEV-IRE1-LD. Sequencing confirmed the correct orientation of six separate plasmids.

As we will be characterising the steps of activation of IRE1, it was necessary to produce a variety of constructs and mutants to act as controls, and to focus on individual steps of activation. These are summarised in Table 3.1, however, in brief, these include non-dimerising mutant (Zhou et al., 2006), a core luminal domain construct that does not contain the linker region and a variety of cysteine mutants.

3.2 Expression and Purification of Recombinant 6xHis-TEV-IRE1-LD

Once the cDNA for 6xHis-TEV-IRE1-LD was successfully sub-cloned into a pET vector, it was then necessary to optimise the conditions that allowed expression of soluble protein. We initially performed a small scale expression at temperatures of 37°C, 30°C and 20°C and performed solubility tests on the cell pellet to ascertain if the protein was soluble or expressed

in inclusion bodies. Solubility tests of FL-IRE1-LD and the C109S, C148S and C332S mutants indicated that the protein was soluble when expressed at all temperatures, however IRE1 cLD and D123P (Figure 3.3) mutants were only soluble when expressed at 20°C thus it was necessary to express these constructs overnight at 20°C.

Once conditions for protein expression had been ascertained, it was necessary to optimise conditions for purification of FL-IRE1-LD (Figure 3.5). Application of an imidazole gradient indicated FL-IRE1-LD was eluted at an imidazole concentration of approximately 100 mM, thus it was possible to include an extensive wash step to ensure host proteins that bound non-specifically to the IMAC column were removed prior to elution of recombinant 6xHis-TEV-IRE1-LD. The protein was successfully expressed and purified at a yield of around 15-20 mg/L of media (Figure 3.4), which is suitable for biophysical characterisation and structural studies.

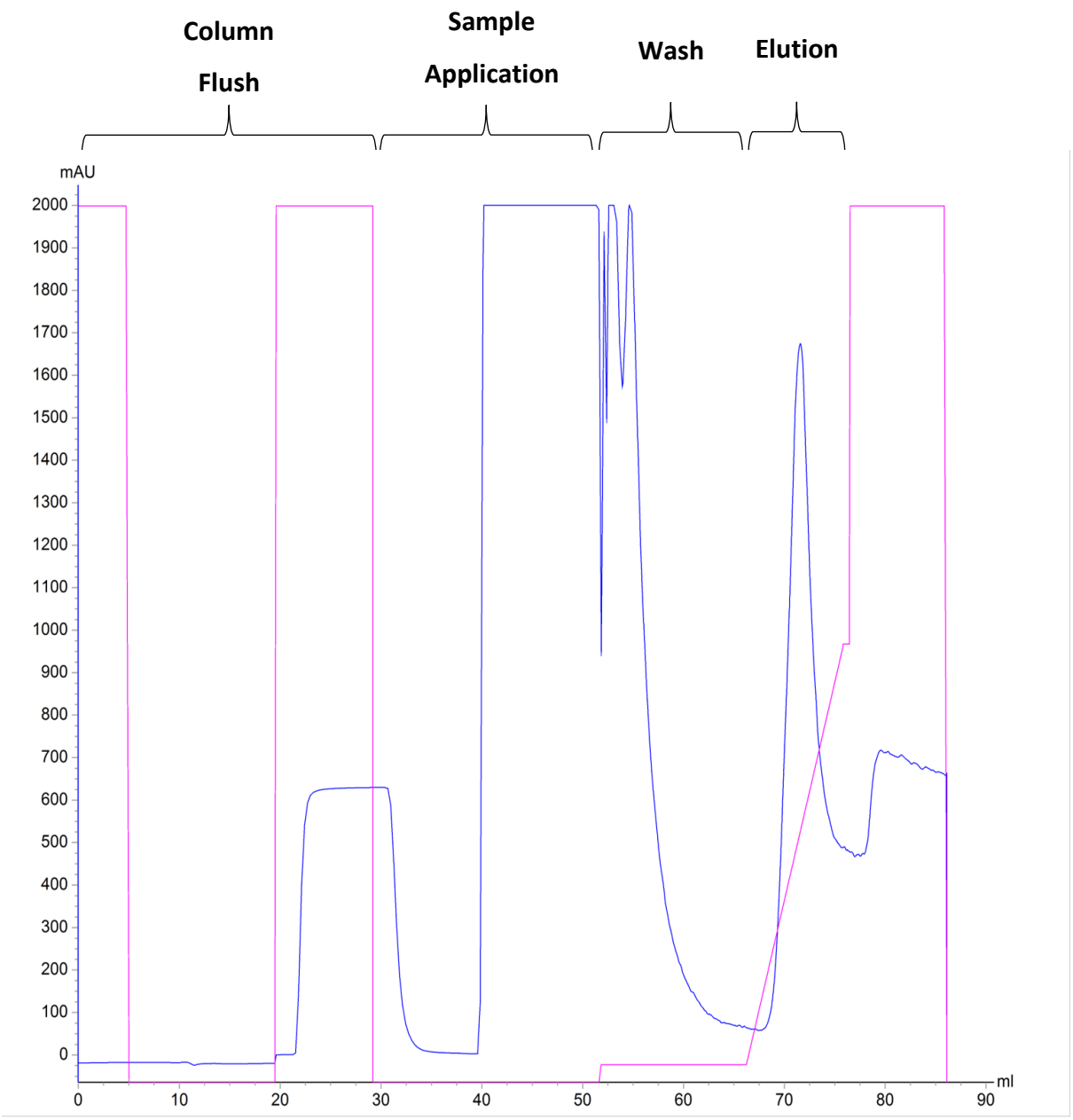

Figure 3.4

Example chromatogram of the IMAC purification of IRE1-LD with the various steps indicated above. The absorbance at 280 nm (blue) and percentage of 1M imidazole (pink) are displayed. A yield of ~15mg/L was obtained for all IRE1-LD constructs.

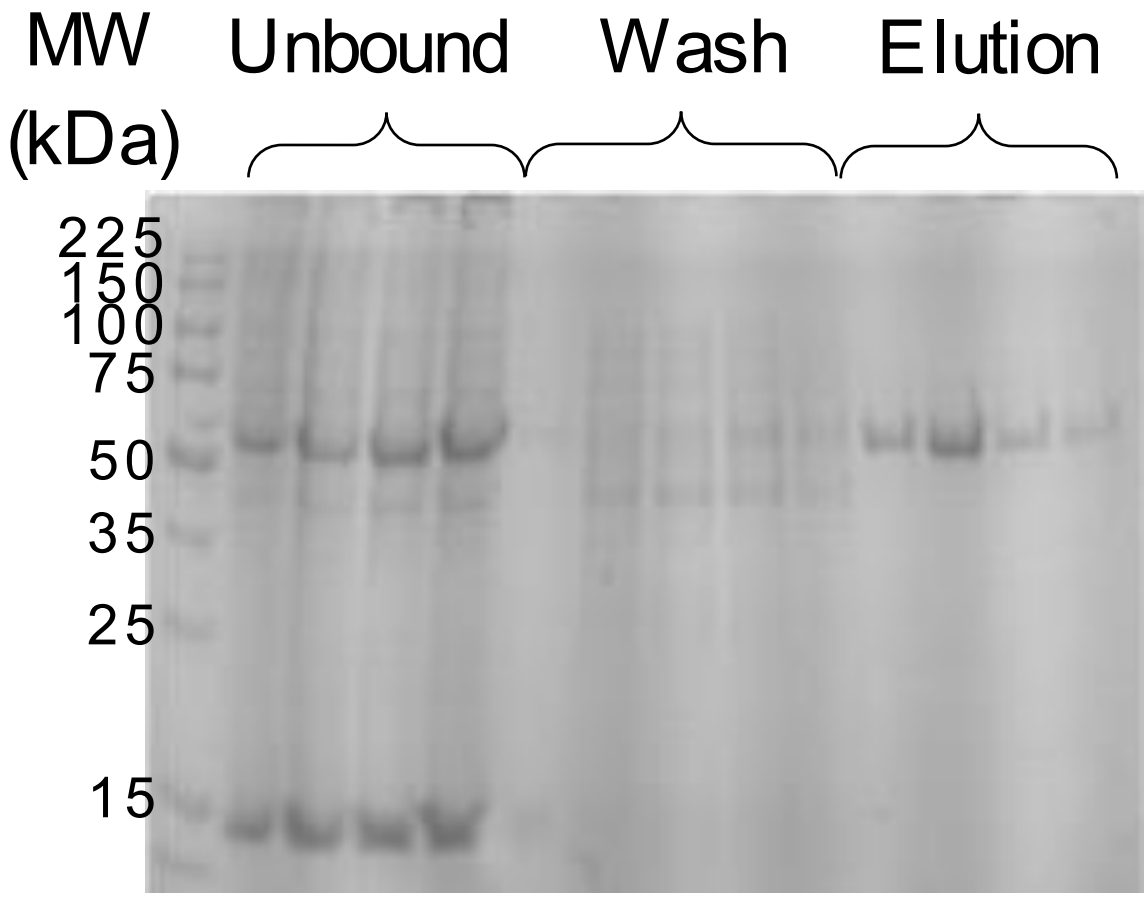

Figure 3.5

SDS-PAGE gel of specified fractions during protein purification protocol optimisation. Recombinant FL-IRE1-LD was still present in the unbound fraction, however only a single band is detectable in the elution fraction.

Construct name	Residues	Function
FL-IRE1-LD	S24-S450	Full length luminal domain construct
IRE1-cLD	S24-V390	Core luminal domain, no linker
IRE1-LD (D123P)	S24-S450	Non-dimerising mutant(Zhou et al., 2006).
IRE1-LD (C109S)	S24-S450	One of three solvent exposed cysteines. Unimportant for disulphide bond formation (Liu et al., 2003).
IRE1-LD (C148S)	S24-S450	One of three solvent exposed cysteines. Important for disulphide bond formation (Liu et al., 2003).
IRE1-LD (C332S)	S24-S450	One of three solvent exposed cysteines. Important for disulphide bond formation (Liu et al., 2003).

Table 3.1

Table of all IRE1-LD constructs used throughout this work with residues numbers and the construct function. The naming will remain consistent throughout this work.

3.3 Identification and initial biophysical characterisation of FL-IRE1-LD

To ensure the correct identity of the purified protein, mass spectrometry was performed on the trypsin digested fragments of the expressed construct (Figure 3.6). Our construct was the top scoring construct (-10lgP = 524.12) from a search of the Swiss-Prot database (Boeckmann et al., 2003), confirming the correct identity of our expressed construct. Following purification of FL-IRE1-LD, in the absence of any established activity assay for IRE1-LD, it was necessary to ensure our construct was correctly folded. To do this, we performed circular dichroism on FL-IRE1-LD (Figure 3.7) and performed a secondary structure prediction (Louis-Jeune et al., 2012) on the resulting data. This indicated that our IRE1-LD construct consisted of 7% alpha helix and 30% beta sheet. This is in good agreement with the previously published X-ray structure

that indicates the protein comprises of 8% alpha helices and 27% beta sheets (Zhou et al., 2006).

In agreement with this observation, our sequence analysis of FL-IRE1-LD using PONDR (Romero et al., 2001) predicts large regions of disorder in the structure of IRE1-LD (Figure 3.8). Unsurprisingly the majority of these intrinsically disordered regions are not resolved in the X-ray structure, including C148 and C332, which have both previously been shown to be important for disulphide bond formation (Liu et al., 2003), while the former has is essential to interact with PDIA6 (Eletto et al., 2014) (Groenendyk et al., 2014). These results show that we have successfully expressed and purified recombinant hIRE1-LD in *E. coli* at a yield suitable for biophysical and structural characterisation. Interestingly, our preliminary analysis of the secondary structure elements of FL-IRE1-LD also suggests it contains a significant level of disorder.

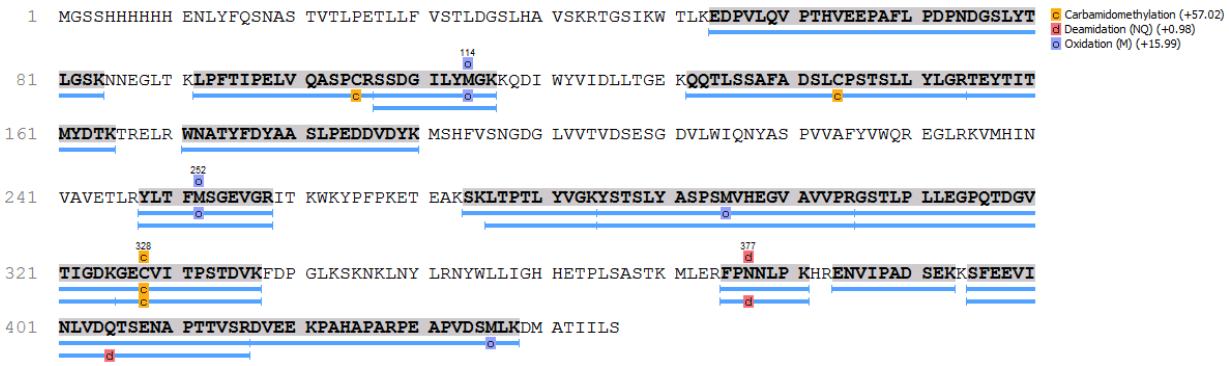

Figure 3.6

ESI-MS analysis of the peptide fragments following trypsin digest was analysed against the SWISS-PROT database (Boeckmann et al., 2003). Matching fragments are highlighted with blue lines.

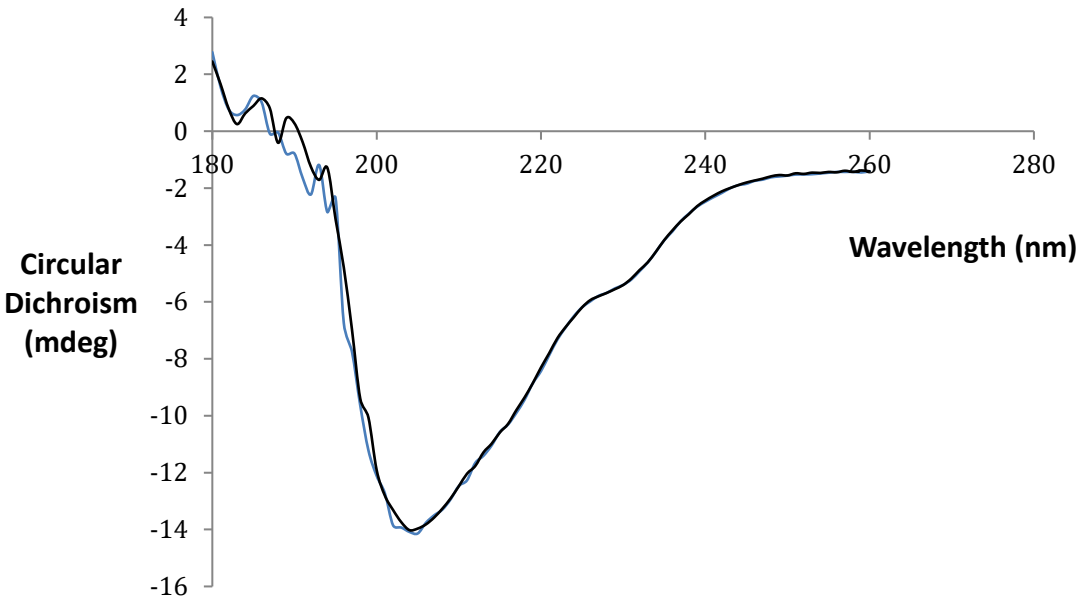

Figure 3.7

CD spectra of 0.2 mg/ml FL-IRE1-LD. Two separate scans are shown. Secondary structure analysis of the spectra using the K2D3 webserver (Louis-Jeune et al., 2012) indicates that our IRE1-LD construct is composed of 7% alpha helix and 30% beta sheet.

Following identification of the secondary structural elements, we then ascertained if our FL-IRE1-LD construct forms disulphide bonds, in agreement with previous literature data (Liu et al., 2003; Eletto et al., 2014; Liu et al., 2002). To do this, denatured mass spectrometry was performed to establish if there was any covalent linkage between FL-IRE1-LD protomers. This indicated that in the presence of a reducing agent (5 mM DTT), FL-IRE1-LD had a molecular mass of 49657.59 Da ±0.57, while in the absence of a reducing agent the majority of FL-IRE1-LD had an average molecular mass of 99314.30 Da ±1.18 (Figure 3.9), indicating that FL-IRE1-LD is covalently linked in the absence of a reducing agent. The expected average molecular mass of our FL-IRE1-LD construct is 49789.38 Da, a mass loss of 131.79 Da which suggests cleavage of the N-terminal methionine (Liao et al., 2004). We then characterised the size of FL-IRE1-LD in solution in the presence and absence of a reducing agent. To do this, we performed SEC in the presence and absence of 5 mM DTT in the running buffer and protein solution (Figure 3.10). This indicated that the Rh of 40 µM FL-IRE1-LD is 4.19 nm in the

presence of DTT and 5.94 nm in the absence of DTT, indicating that larger species are formed upon disulphide bond formation.

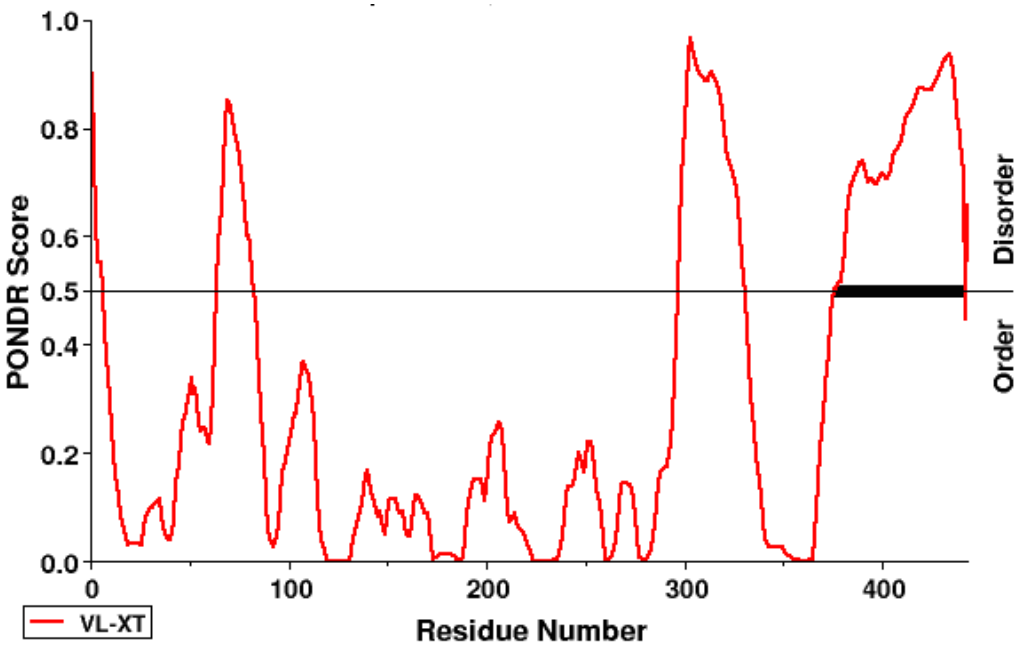

Figure 3.8

PONDR analysis (Romero et al., 2001) of the hIRE1-LD sequence indicates several regions of the protein are predicted to be disordered.

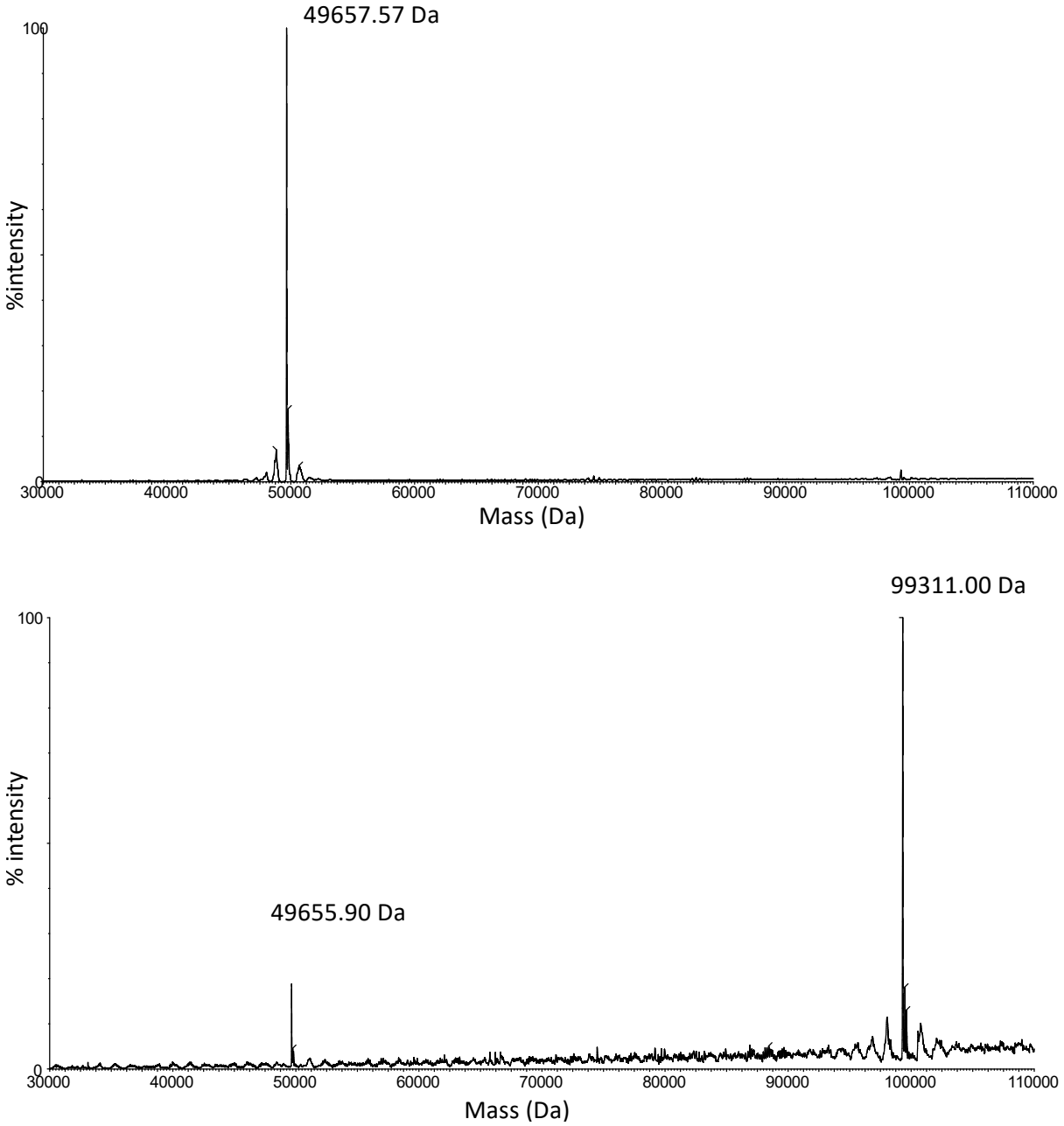

Figure 3.9

Molecular mass data of FL-IRE1-LD deconvoluted from denatured mass spectra of 15 µM FL-IRE1-LD in the presence (above) and absence (below) a 5 mM DTT. This indicates that FL-IRE1-LD forms covalently linked dimers in the absence of a reducing agent. The expected molecular mass of FL-IRE1-LD is 49789 Da, compared to an observed mass of 49657.57 Da, consistent with the removal of the N-Terminal methionine (Liao et al., 2004).

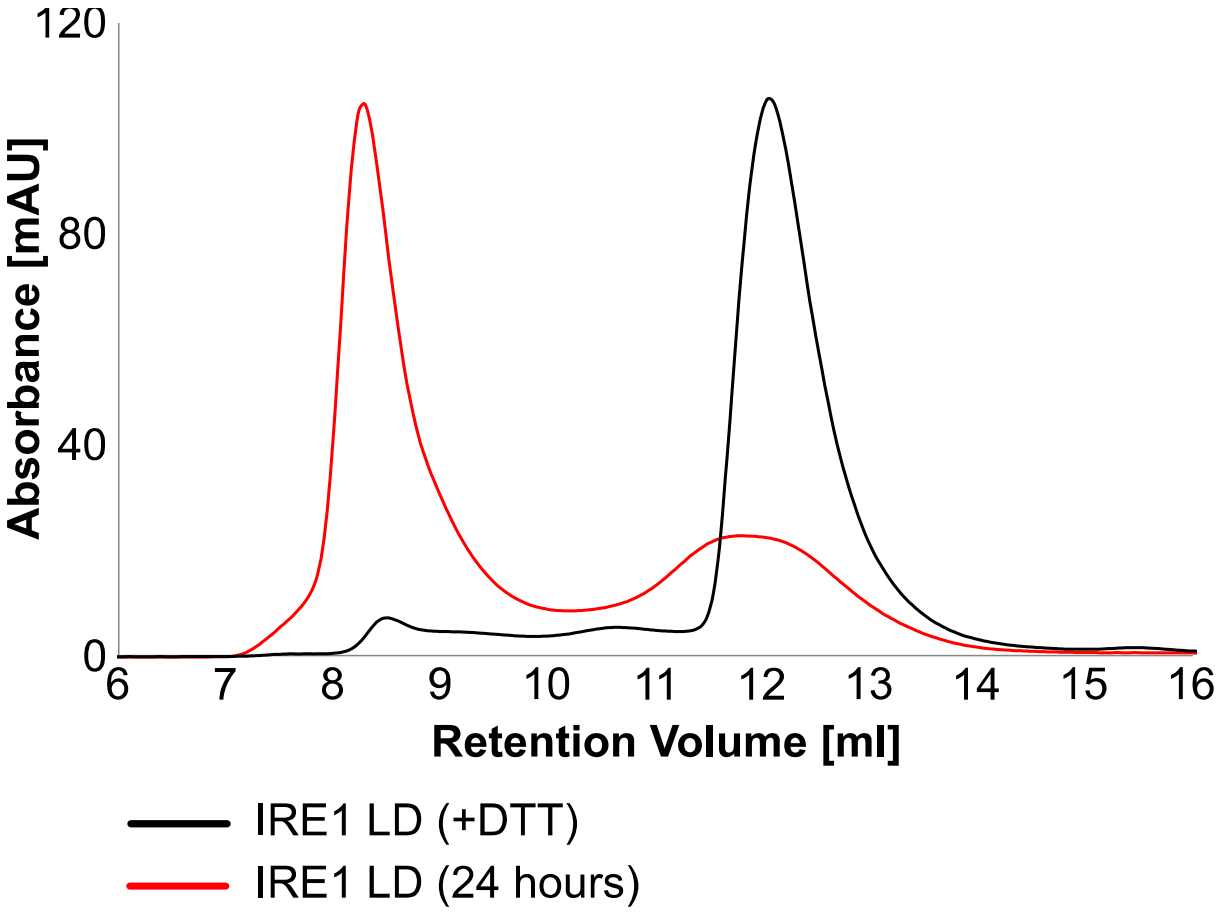

Figure 3.10

SEC chromatogram of 40 µM (injection concentration) FL-IRE1-LD in the presence (black) and absence (red) of 5 mM DTT indicates FL-IRE1-LD forms species of a R_h 5.94 nm after incubation in a non-reducing environment.

To ascertain the distribution of species present in FL-IRE1-LD, we performed native mass spectrometry in the presence (Figure 3.11) and absence (Figure 3.12) of 5 mM DTT. This indicated that a combination of monomeric, dimeric and a small amount of tetrameric species were present in the sample in the presence of a reducing agent, while in the absence of a reducing agent the monomeric species completely disappeared. All this data suggests FL-IRE1-LD is present in different oligomeric states and the formation of disulphide bonds may affect the equilibrium between different oligomeric states.

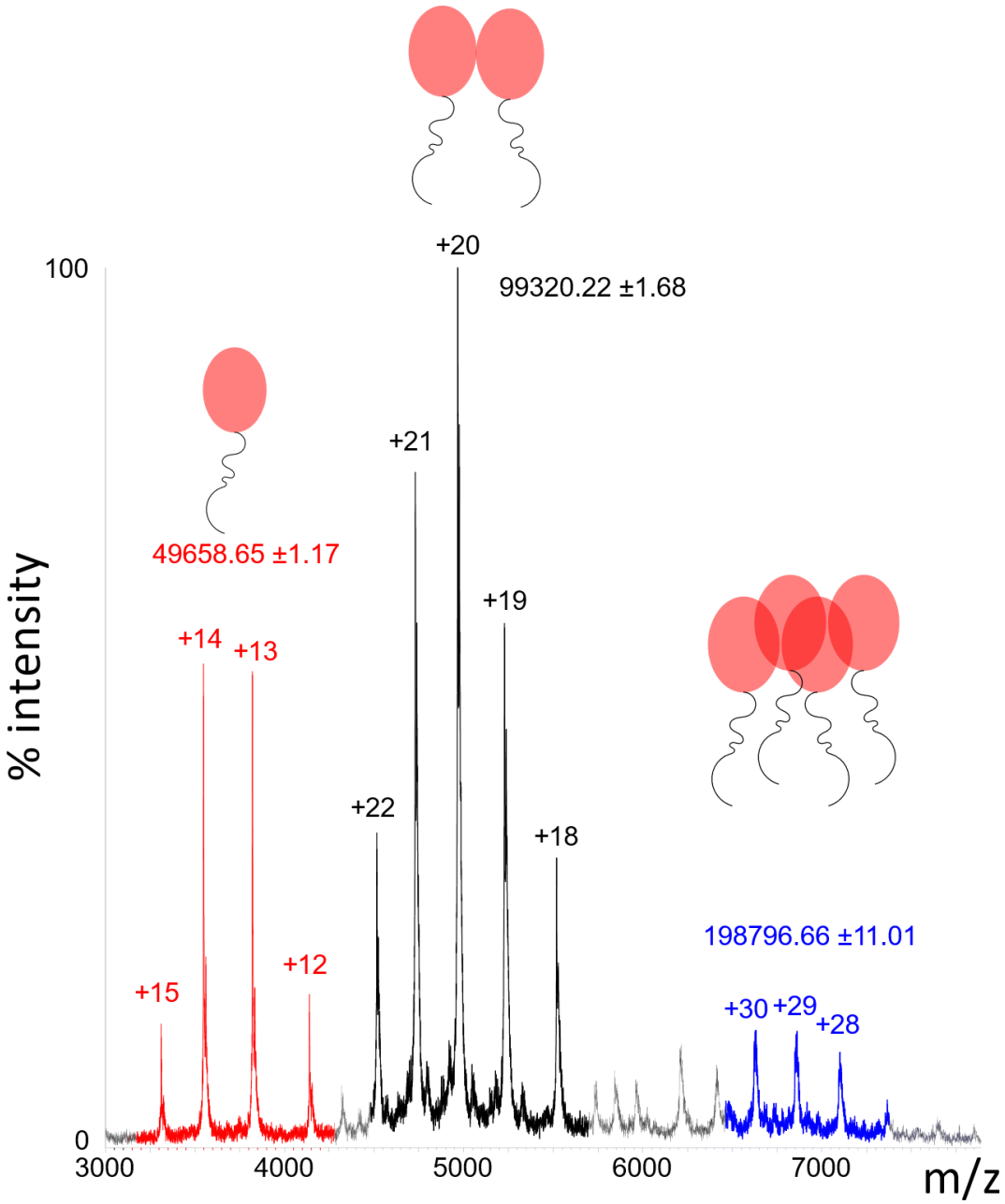

Figure 3.11

Native mass spectrum of 15 µM FL-IRE1-LD in the presence of 5 mM DTT indicates monomeric, dimeric and a small amount of tetrameric species are observable.

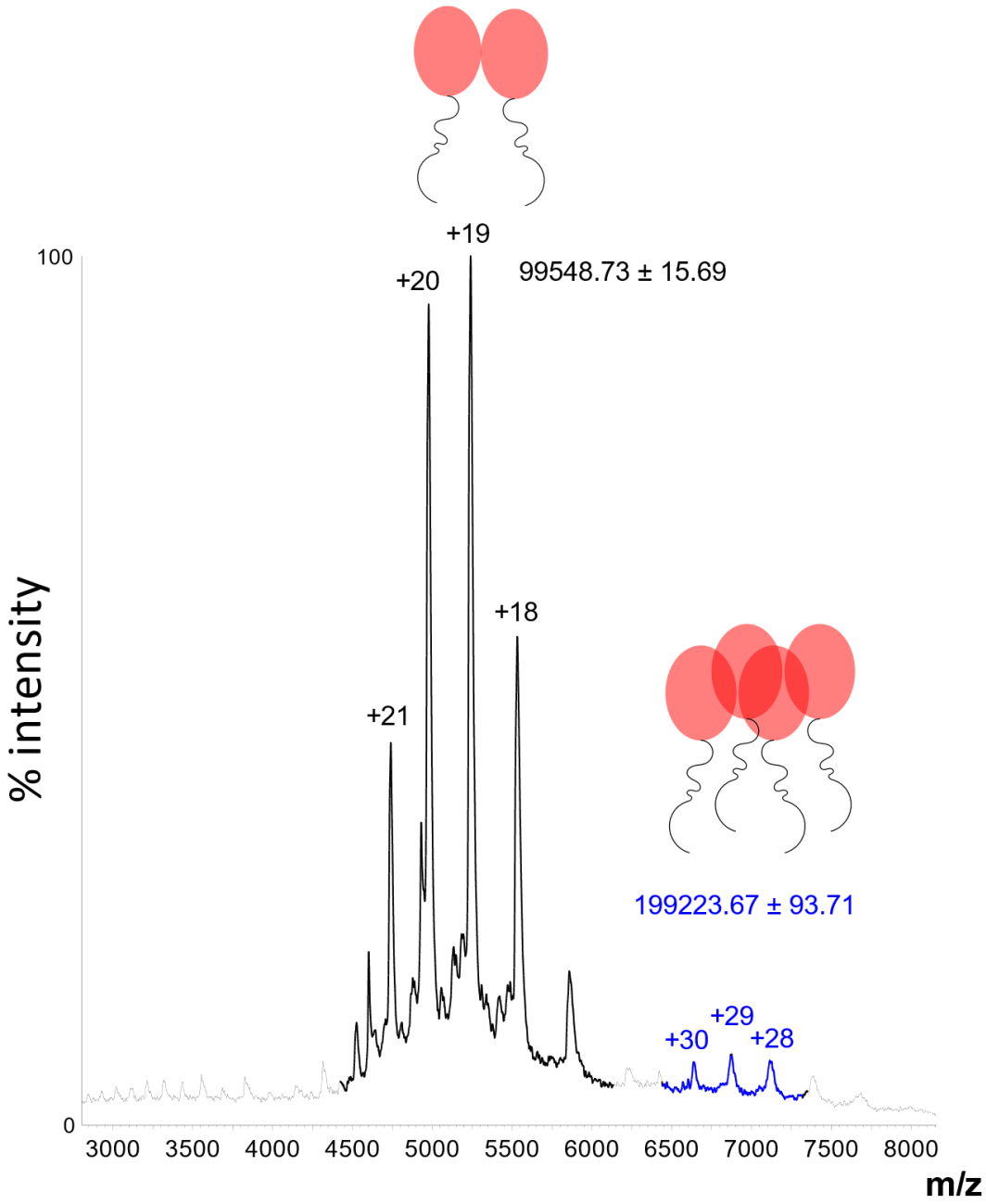

Figure 3.12

Native mass spectrum of 15 µM FL-IRE1-LD in the absence of 5 mM DTT indicates that the monomeric species observed in the presence of DTT are no longer present in the absence of DTT.

3.4 Conclusion

The data in this chapter demonstrates that we have successfully subcloned the gene sequence for our hIRE1-LD construct into a vector that is suitable for expression in *E. coli*. Following this, we have successfully optimised expression and purification protocols for six variants of our

IRE1-LD construct. The yields we have obtained using these protocols mean it will be possible to perform biophysical and structural characterisation of this system in later chapters.

Initial characterisation of our FL-IRE1-LD construct indicates the secondary structure elements are similar to those predicted using the disorder prediction software, PONDR (Romero et al., 2001), and those determined in the crystal structure of hIRE1-LD (Zhou et al., 2006), thus suggesting our construct contains comparable folds to those used in previous studies. Analysis of our data and data from the crystal structure indicate ~60% of IRE1-LD is devoid of secondary structure.

Finally, our initial biophysical characterisation using mass spectrometry and size exclusion suggests that IRE1-LD forms disulphide bonds, in agreement with previous literature data (Liu et al., 2003; Eletto et al., 2014). In addition, the formation of covalent complexes appears to affect the oligomeric state of IRE1-LD, with monomeric species no longer observed in the mass spectrum of FL-IRE1-LD in a non-reducing buffer, and an increased retention volume observed in SEC experiments, when compared to IRE1-LD in a reducing buffer.

4 The conformational landscape of IRE1-LD

The luminal domain of IRE1 is essential for downstream functioning of its cytosolic region. Upstream activation of IRE1 is a complex, multi-step process that requires dimerization and oligomerisation of the luminal domain of IRE1. Previous literature data indicates this process maybe under control of the molecular chaperone, BiP (Bertolotti et al., 2000; Kimata et al., 2004; Todd-Corlett et al., 2007; Amin-Wetzel et al., 2017), and a direct interaction with unfolded proteins (Gardner, B.M. and Walter, 2011) (Karagoz et al., 2017).

IRE1-LD occupies a complex functional landscape that is thought to include monomeric, dimeric and oligomeric forms, in addition to proposed induced conformational changes by unfolded substrates (Karagoz et al., 2017). Furthermore, IRE1-LD has been shown to form covalent complexes, driven by the formation of disulphide bonds (Liu et al., 2003), adding further complexity to the conformational landscape of IRE1-LD. While the majority of this work has been carried out *in vivo*, it is necessary characterise this using isolated elements *in vitro* to properly understand the conformational landscape of IRE1-LD and how elements such as unfolded substrates affect this.

In the previous chapter we performed initial biophysical characterisation of the conformational landscape of IRE1-LD. The work in this chapter will focus on more detailed biophysical characterisation of the functional landscape of IRE1-LD to obtain mechanistic insights into the steps of IRE1 activation.

4.1 The inactive IRE1-LD conformation

4.1.1 SEC characterisation of the IRE1 monomer/dimer transition

As described previously, IRE1-LD activation is a multistep process; under physiological or chemical stresses, inactive (presumably monomeric) IRE1-LD undergoes conformational rearrangements, resulting in an active (presumably oligomeric) conformation. This conformation then triggers downstream activation of IRE1 in the cytoplasm (Zhou et al., 2006; Karagoz et al., 2017).

First, we characterised the inactive IRE1-LD conformation that the protein adopts in the absence of any unfolded substrate and/or molecular chaperone, BiP. To do this, we performed SEC at varying concentrations of recombinant IRE1-LD. For these experiments we

used the FL-IRE1-LD construct and IRE1-LD (D123P) constructs, as the linker maybe important in the activation process of IRE1-LD (Kimata et al., 2004). The SEC retention volume gradually decreased with increasing IRE1-LD injection concentration, suggesting the protein undergoes an oligomeric transition in the µM range (Figure 4.1A), however no changes in peak position were observed for the previously characterised D123P mutant of IRE1-LD (Zhou et al., 2006), indicating this process is likely dimerisation (Figure 4.1B).

Next, we roughly estimated the dimerization constant of IRE1-LD using the peak volumes at different IRE1-LD concentrations. As ~60% of IRE1-LD is intrinsically disordered (see section 3.3), it is not possible to use globular protein calibration standards directly to estimate the size of IRE1-LD oligomers. Instead, we assumed the D123P construct was eluted as pure monomer and calculated the molecular weight based on this principle (see section 2.9 for further details). This analysis suggests that in the absence of substrate and BiP, IRE1-LD oligomerisation occurs in the sub-µM range.

4.1.2 MST characterisation of IRE1-LD dimerization constant

To obtain an accurate value for the IRE1-LD dimerization constant, we employed microscale thermophoresis (MST). IRE1-LD was labelled with FITC and 0.5 µM FITC labelled FL-IRE1-LD was titrated with 0.003 – 85 µM IRE1-LD (Figure 4.2). The MST data indicated that changes in the IRE1-LD conformational landscape were occurring with a K_d = 0.2 µM ±0.03. As a control, we repeated the experiment with FITC labelled IRE1-LD (D123P) and titrated either unlabelled FL-IRE1-LD (WT) or IRE1-LD (D123P). As expected from our SEC results, no changes in the IRE1-LD conformational landscape was observed for the non-dimerising mutant (Figure 4.1B).

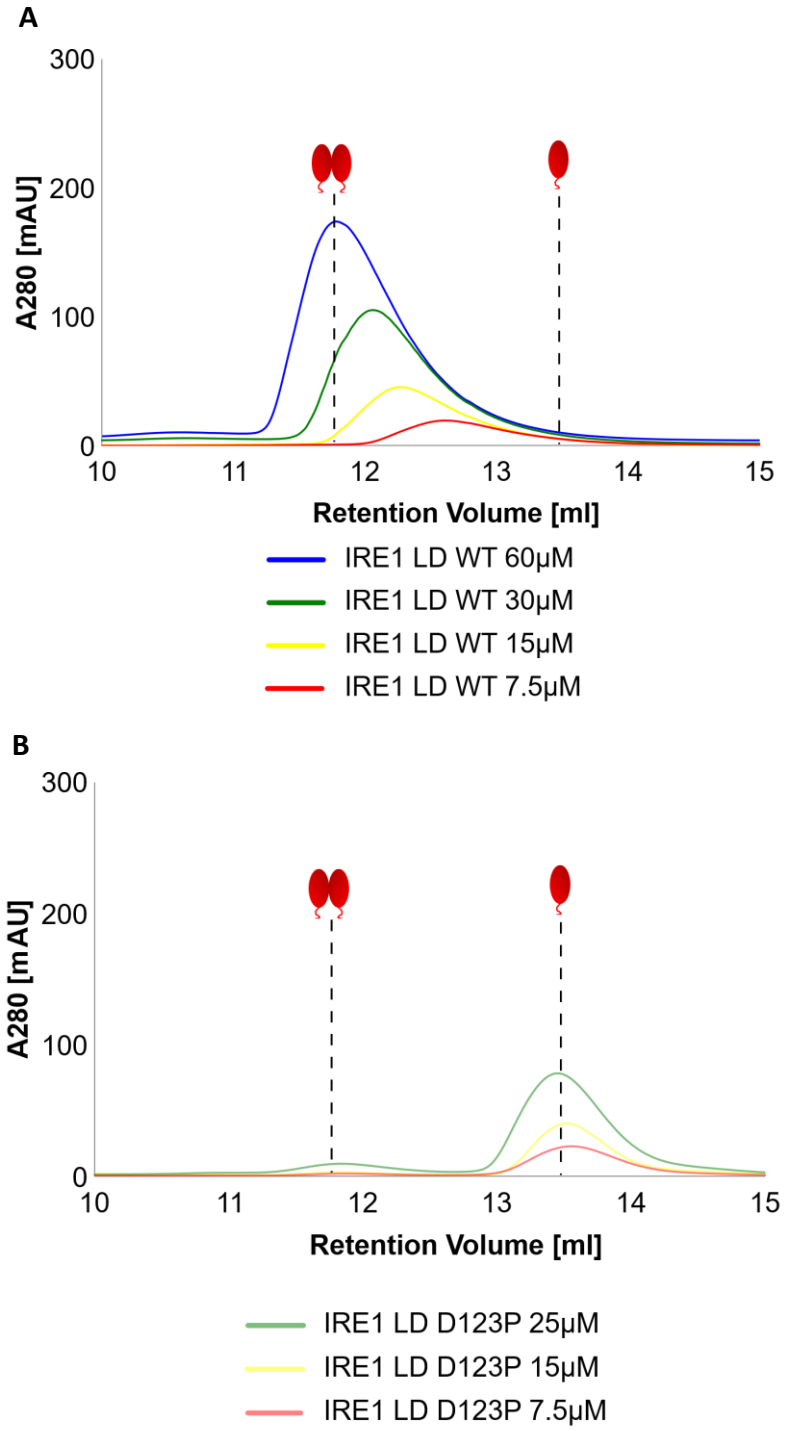

Figure 4.1

Size exclusion chromatogram of IRE1-LD (WT) (A) and IRE1-LD (D123P) (B) at varying injection concentrations. The retention volume of IRE1-LD (WT) reduces with increasing concentration whereas that of IRE1-LD (D123P) remains static.

Injected FL-IRE1-LD concentration (µM)	Dilution factor (eluted volume/injected volume)	Eluted concentration (µM)
60	22.2	2.70
30	25.3	1.19
15	23.4	0.64
7.5	21.4	0.35

Table 4.1

Table of eluted concentration of FL-IRE1-LD vs. the injected concentration.

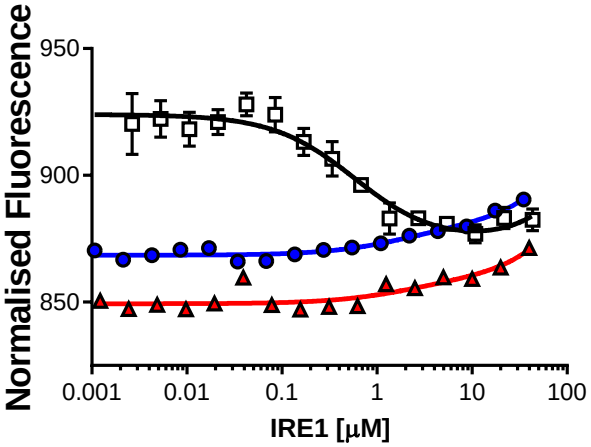

Figure 4.2

Microscale thermophoresis trace of IRE1-LD interaction. FITC labelled IRE1-LD (WT) was titrated into a dilution series of IRE1-LD (WT) (black K_d 0.2μM±0.03). This interaction did not occur upon titration of IRE1-LD (WT) (red) and IRE1-LD (D123P) (blue) into FITC labelled IRE1-LD (D123P).

Finally, to confirm that we were observing the same process by both SEC and MST, we compared the amounts of monomeric IRE1-LD observed by SEC and MST at different FL-IRE1-LD concentrations (Figure 4.3) This shows excellent agreement between MST and SEC experiments, suggesting that IRE1-LD co-exists in equilibrium between its monomeric and dimeric forms with a dimerization constant of (K_d) 0.2 µM ±0.03.

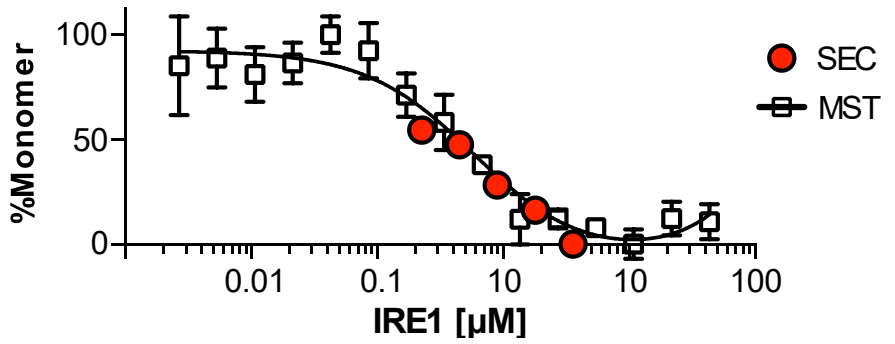

Figure 4.3

Normalised MST data and SEC data overlaid indicates that the interaction observed by MST is the same as that observed by SEC, confirming that IRE1-LD dimerises with a sub-µM affinity.

Karagoz et al. recently demonstrated that some oligomeric protein is present at concentrations >25 µM. While this may explain a slight inflection in the isotherm observed at concentrations >10 µM, it did not interfere with the calculation of a dimerization constant. The authors also demonstrated that at an injection concentration, thus a lower elution concentration, of 62.5 µM, IRE1-LD is eluted as dimeric protein, in agreement with our observation that at an injection concentration of 60 µM IRE1-LD is eluted as a dimer. This data is in contrast with that observed for yIRE1, whose dimerization constant has been estimated at ~10 µM (Credle et al., 2005) in the absence of an unfolded substrate. As previously discussed in section 1.3, sequence and structural elements are well conserved at the dimerization interface, indicating differences between the conformational landscape of hIRE1 and yIRE1.

4.2 Structural insights of the inactive IRE1-LD dimer

While the structure of the presumably active dimeric form of IRE1-LD has been previously suggested, based upon the X-ray structure of the monomeric form of the truncated (S24-V390) hIRE1-LD variant (Zhou et al., 2006), and the dimeric structure of yIRE1-LD (Credle et al., 2005), no structural insights into the dimeric (inactive) full length hIRE1-LD conformation are currently available. Another complication comes from the fact that approximately 35% of the truncated hIRE1-LD construct used to determine the X-ray structure, are unresolved. The disorder predictor software determined these unresolved regions and the juxtamembrane segment truncated in the X-ray structure are intrinsically disordered (see section 3.3).

4.2.1 NMR characterisation of the IRE1-LD intrinsically disordered regions

To obtain further structural insights into the flexible regions of FL-IRE1-LD, we performed NMR characterisation. The protein was isotopically labelled to produce U{^{15}N} and U{^{15}N, ^{2}H}. The secondary structural contents predicted from the CD data (Figure 3.7, section 3.2), which was in excellent agreement with the IRE1-LD X-ray structure, suggested that the protein is correctly folded.

As a ^{15}N-^{1}H BEST-TROSY experiment should produce a peak for every amide group in the protein backbone, we would expect to see ~400 peaks, excluding side chain peaks, in the FL-IRE1-LD spectrum, however only ~100 backbone resonances were observed in the amide spectrum of FL-IRE1-LD (Figure 4.4). The low proton dispersion of the observed resonances suggested that these residues are located in the disordered regions. Addition of 8M urea resulted in complete FL-IRE1-LD unfolding, which resulted in an expected peak pattern for a 49 kDa unfolded protein (Figure 4.5), while in the absence of urea the majority of these peaks disappeared from the spectrum. As the major source of relaxation in protein NMR, which results in line broadening of resonances, emanates from protons, we deuterated FL-IRE1-LD, however this did not result in the appearance of any additional peaks in the amide spectrum of folded FL-IRE1-LD. Furthermore, we attempted experiments at 298K, 288K and 278K, however none of these conditions improved the spectral quality (see appendix for spectra).

To examine whether the monomer/dimer transition is responsible for line broadening in the NMR spectra, we recorded NMR spectra of the non-dimerising mutant, IRE1-LD (D123P) (Figure 4.6). No improvements in the NMR spectra were observed, suggesting that dimerization is not responsible for line broadening. All these results suggested that the folded regions of IRE1-LD are invisible by NMR, apparently due to their enhanced conformational

flexibility and that IRE1-LD has several extended intrinsically disordered regions (IDRs) and that these possess conformational dynamics that are different from the rest of the protein. In addition, it was not possible to change the exchange rate of the structured residues from the intermediate exchange regime by recording the NMR spectrum at 298K, 288K or 278K.

To validate that the peaks we observed in the amide NMR spectrum of FL-IRE1-LD correspond to intrinsically disordered regions, we produced several ^{15}N labelled constructs: a truncated construct that lacked the last 60 intrinsically disordered C-terminal residues (IRE1 cLD) (Figure 4.7) and four single point mutants of FL-IRE1-LD, two of which are at sites present in the folded regions, IRE1-LD (D123P) (Figure 4.6) and IRE1-LD (C109S) (Figure 4.8), and two of which are unresolved in the X-ray structure, IRE1-LD (C148S) (Figure 4.8) and IRE1-LD (C332S) (Figure 4.9). The major changes were observed in the IRE1 cLD spectra where over half of the peaks were no longer observed, while almost no changes in peak position were observed for the remaining peaks. Several peaks were shifted or disappeared upon the C332S substitution, indicating that these resonances emanate from residues either in proximity to, or at the mutation site. No significant changes were observed upon the C109S and C148S mutations indicating folded regions around these residues are not visible in the NMR spectrum. Several peaks were shifted upon the D123P mutation, indicating that that the formation of dimeric IRE1-LD perturbs the chemical environment in the disordered residues of IRE1-LD.

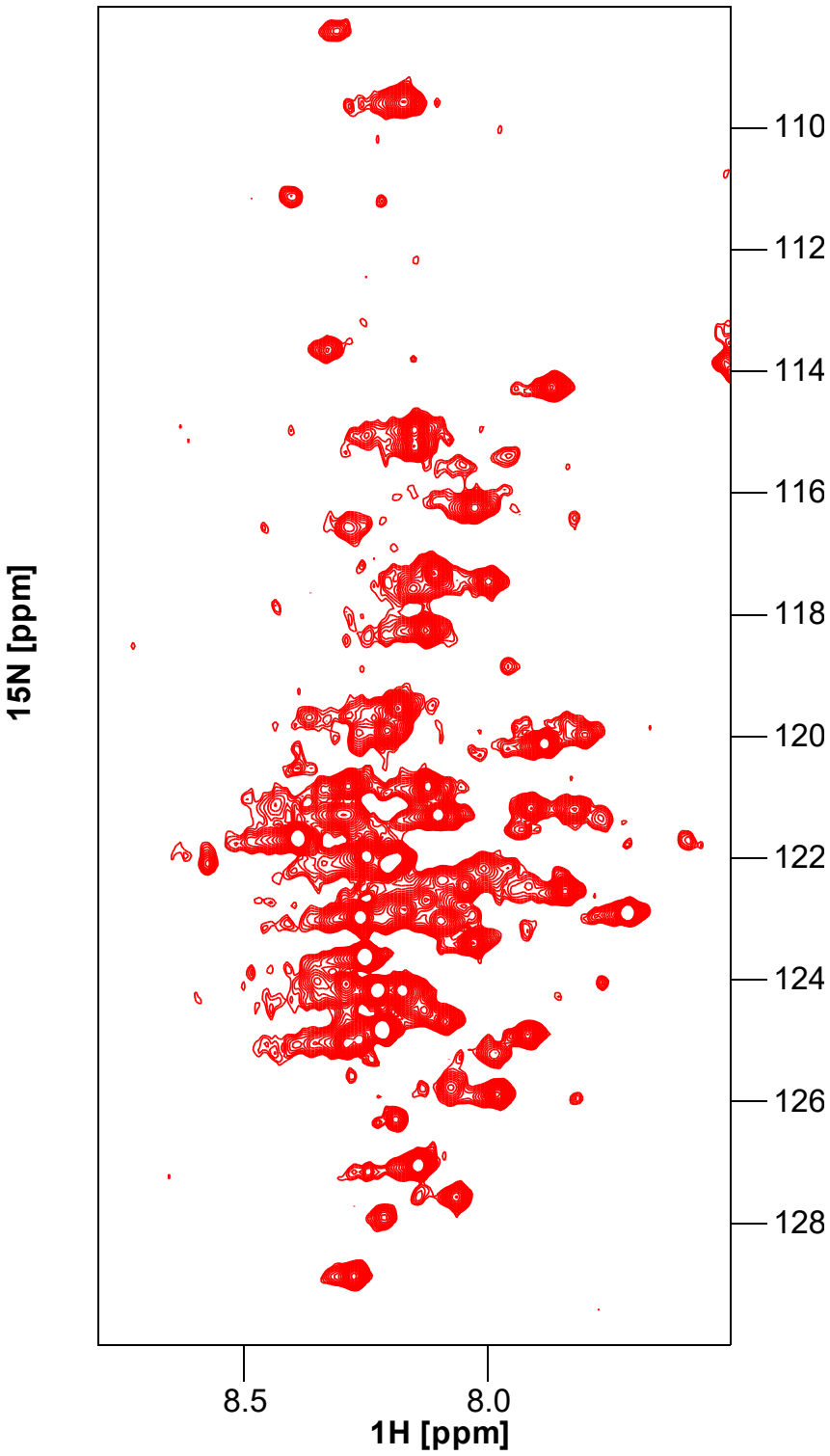

Figure 4.4

^{15}N-^{1}H TROSY of 50 µM FL-IRE1-LD in HMK buffer (+5 mM DTT). ~100 peaks are observed, rather than an expected ~400 peaks, with low dispersion in the proton dimension, suggesting all these peaks belong to the disordered regions of FL-IRE1-LD. Spectrum recorded at 288K.

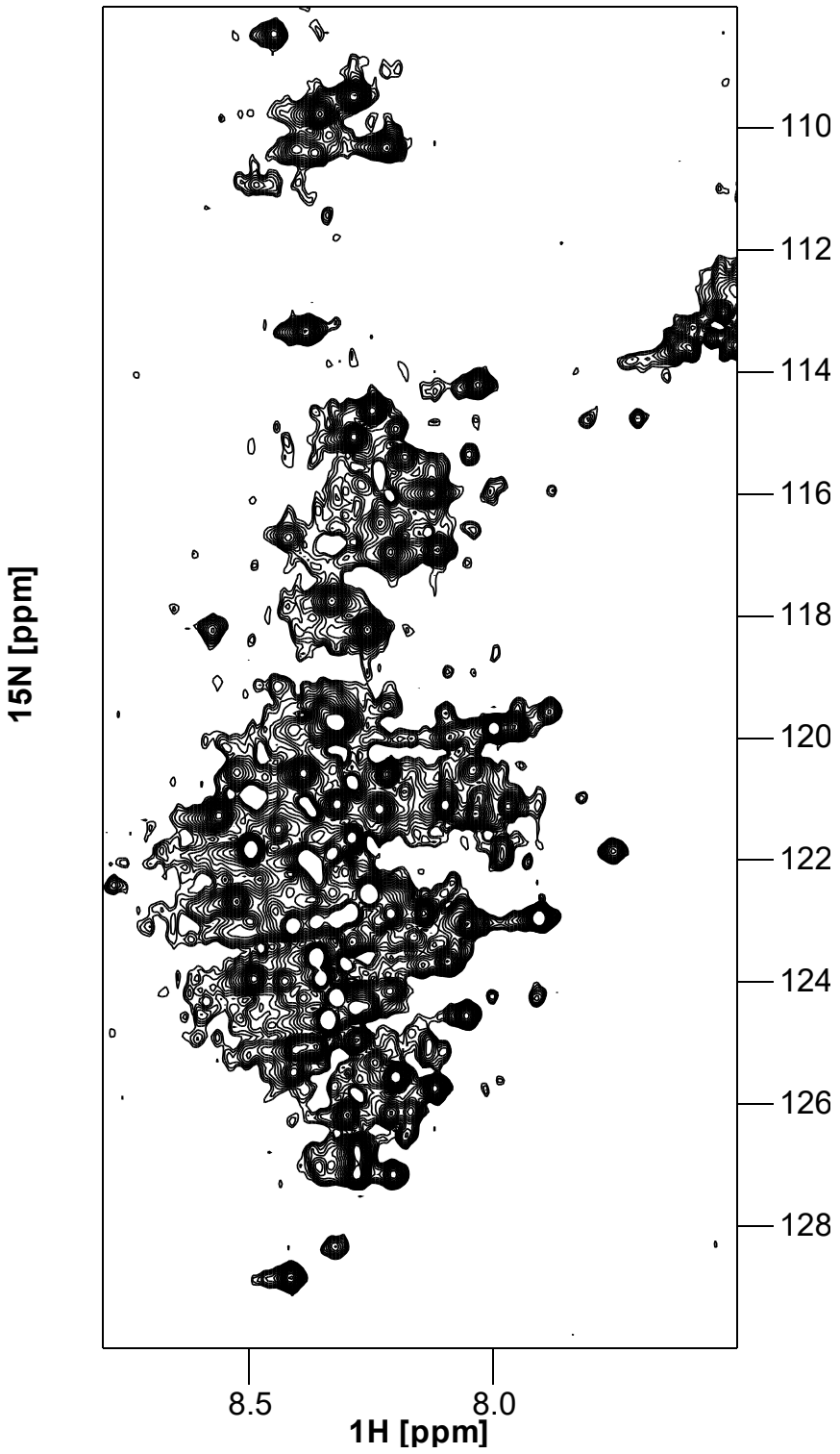

Figure 4.5

^{15}N-^{1}H TROSY of 50 µM FL-IRE1-LD in HMK buffer (+ 5 mM DTT and 8M urea). Numerous additional peaks are observed when compared to Figure 4.4, indicating previously folded regions of the protein are now disordered, thus confirming our construct is not completely unfolded. Spectra recorded at 298K

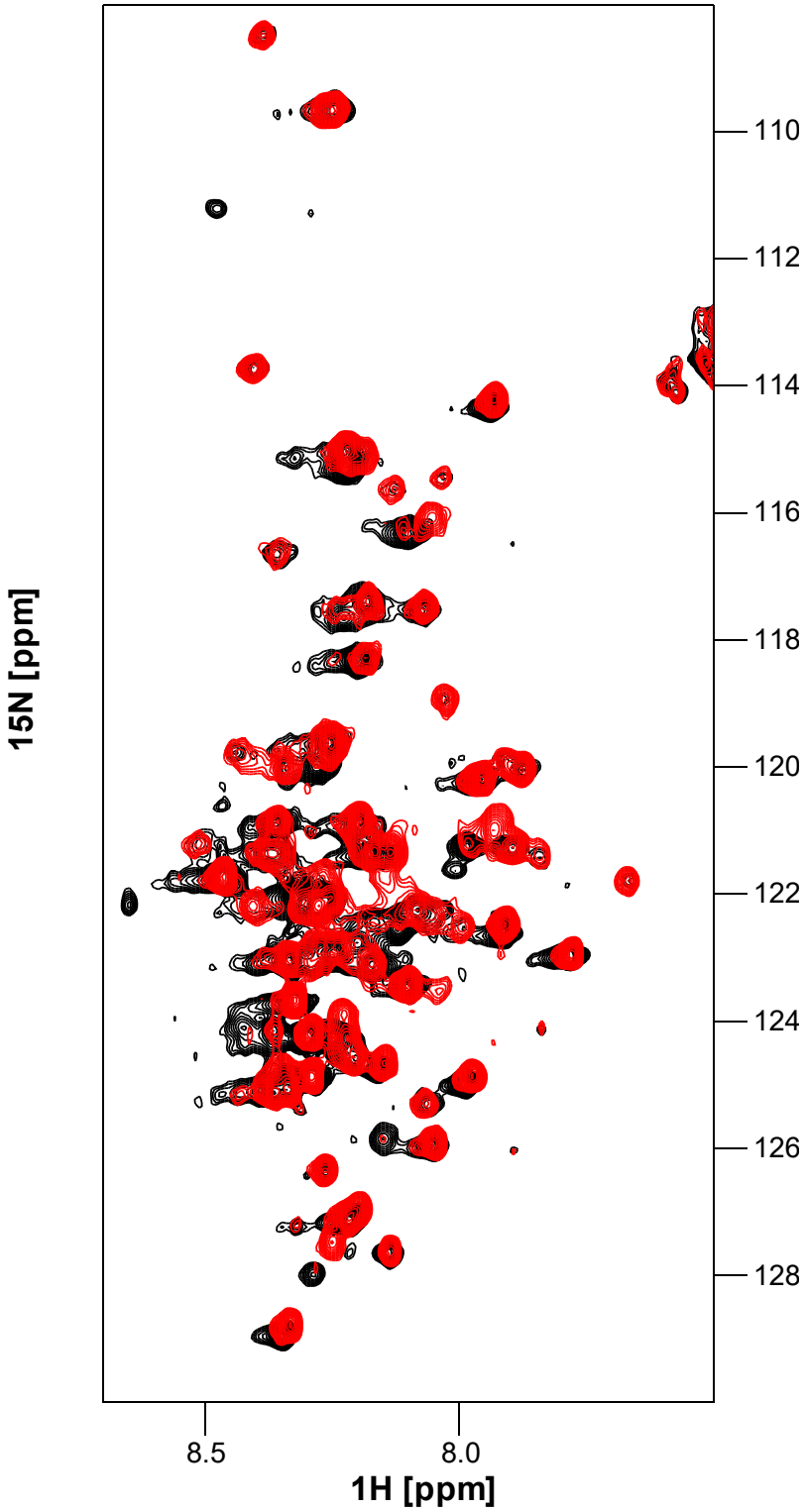

Figure 4.6

^{15}N-^{1}H TROSY of 50 µM FL-IRE1-LD in HMK buffer (+5 mM DTT) (black) overlaid with the spectrum of IRE1-LD (D123P) (red) in the same buffer. This indicates that the line broadening observed in the spectrum of FL-IRE1-LD (WT) is not due to monomer/dimer exchange. Spectrum recorded at 288K

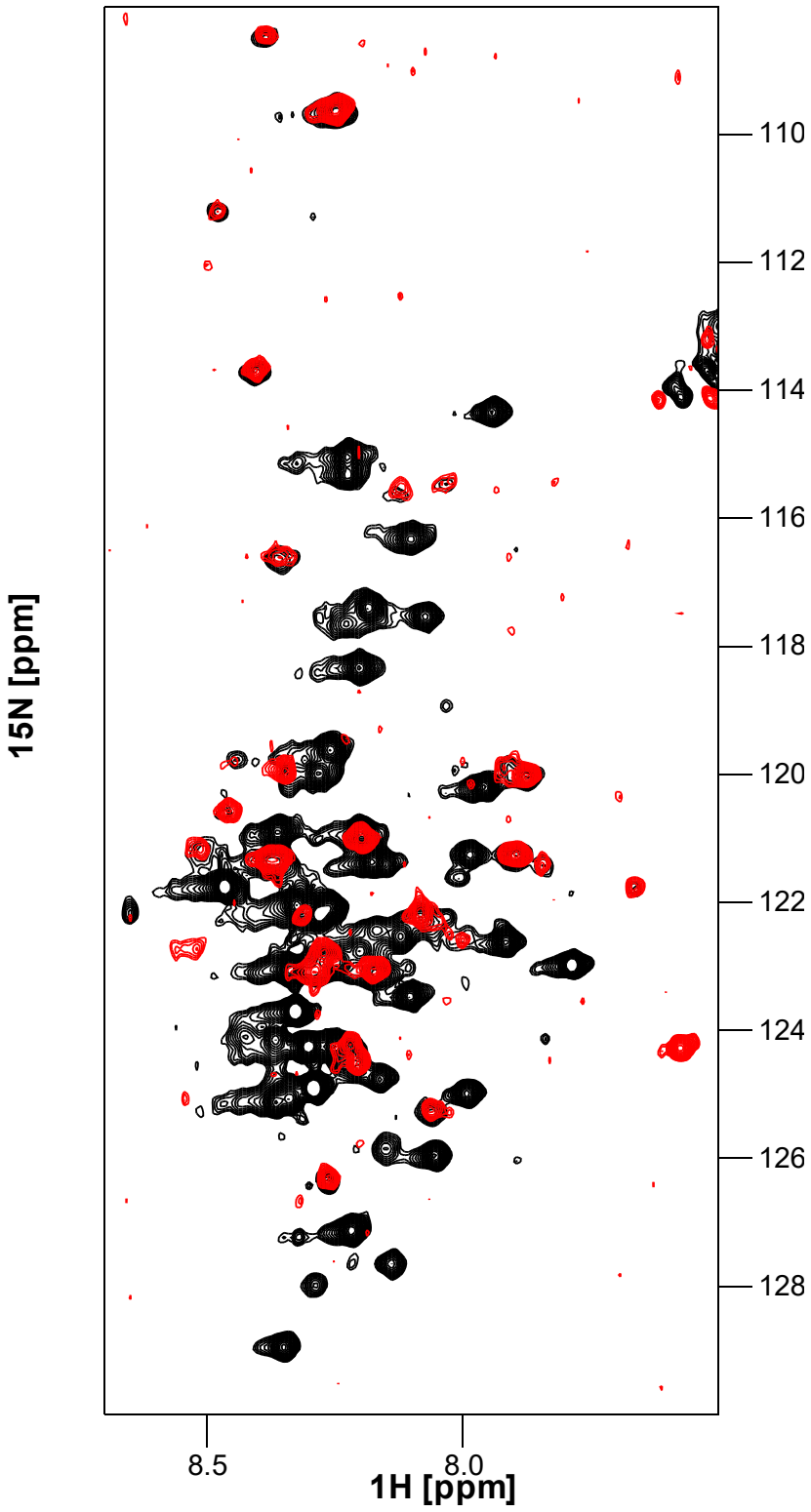

Figure 4.7

^{15}N-^{1}H TROSY of 50 µM FL-IRE1-LD in HMK buffer (+5 mM DTT) (black) overlaid with the spectrum of IRE1 cLD (red) in the same buffer. Good overlap is observed between resonances, indicating these resonances are not perturbed upon truncation of the juxtamembrane linker. Spectrum recorded at 288K

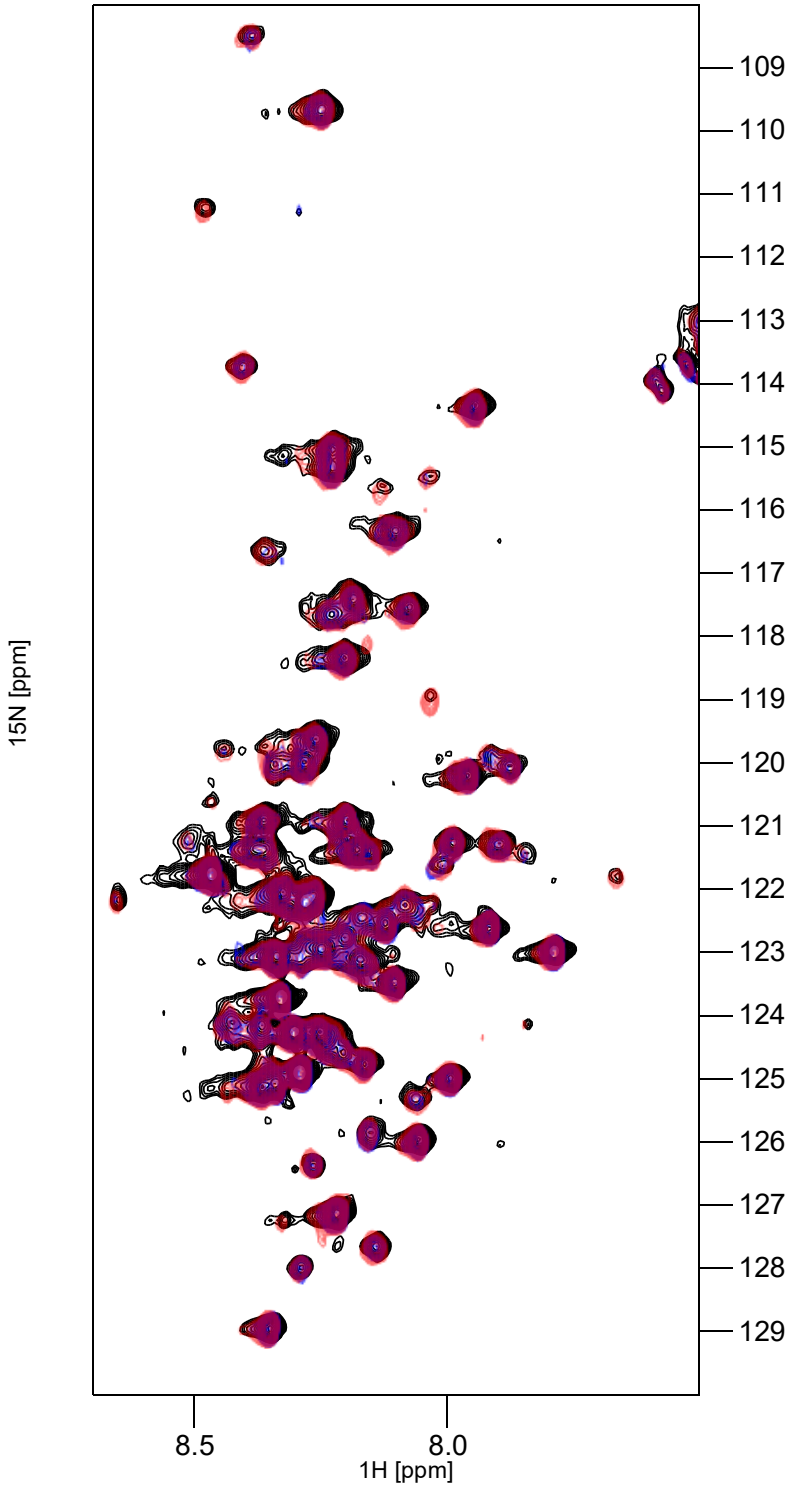

Figure 4.8

^{15}N-^{1}H TROSY of 50 µM FL-IRE1-LD (WT) (black), IRE1-LD (C109S) (blue) and C148S (red) in the presence of 5 mM DTT. No significant changes in peak position between these constructs suggests these residues are not visible in the spectrum. Spectra recorded at 288K.

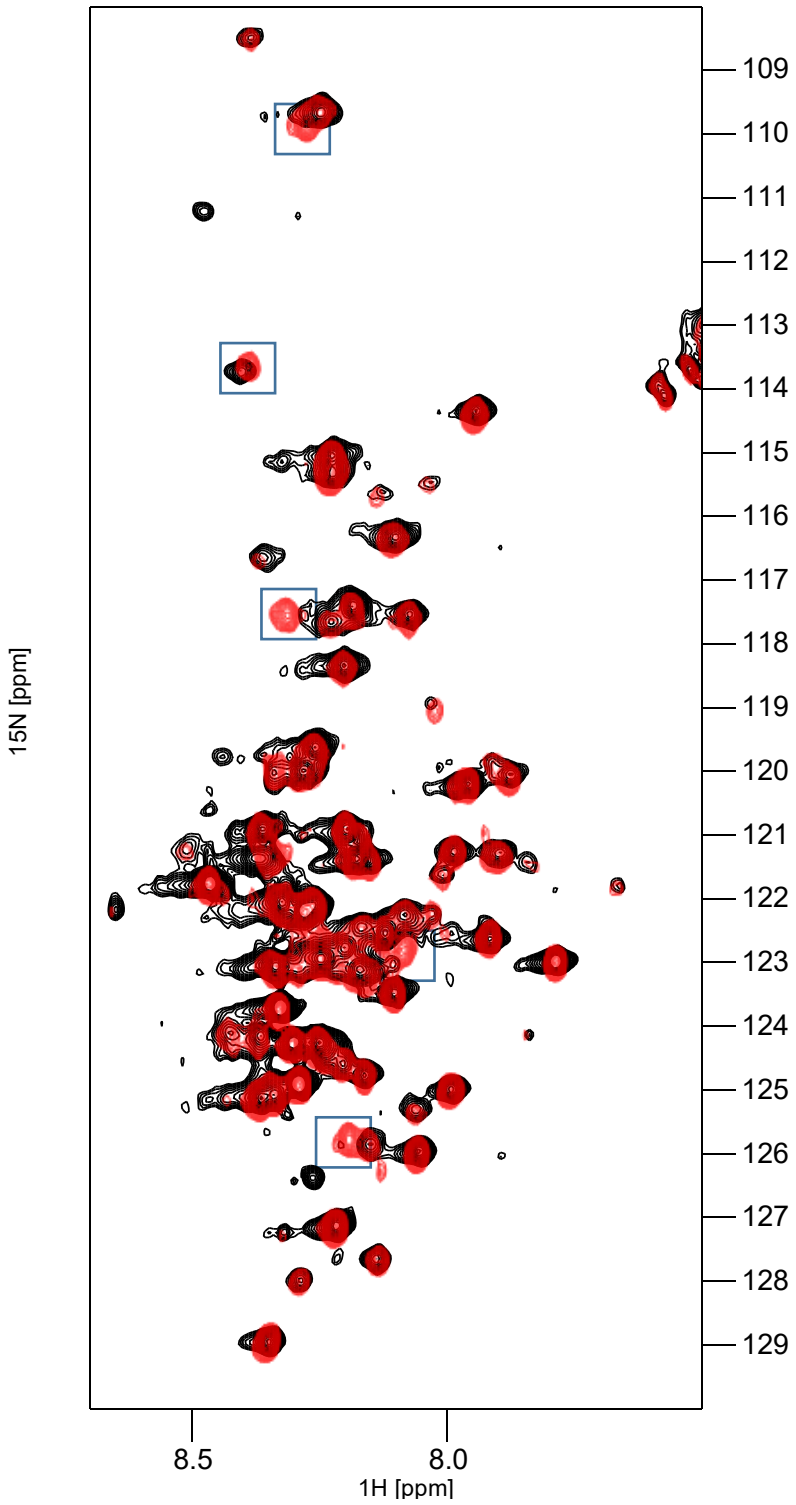

Figure 4.9

^{15}N-^{1}H TROSY of 50 µM FL-IRE1-LD (WT) (black) and IRE1-LD (C332S) (red) in the presence of 5 mM DTT. Five new peaks are present in the spectrum of IRE1-LD (C332S) when compared to FL-IRE1-LD (WT) (highlighted with blue boxes), suggesting residues proximal to C332 are visible in the spectrum of FL-IRE1-LD. Spectra recorded at 288K.

4.2.2 Integrating mass spectrometry with molecular dynamics simulations to characterise the IRE1-LD dimeric conformation

To elucidate whether the inactive IRE1-LD dimer conformation possesses similar individual protomer arrangements to those observed in the X-ray structure of the yeast homolog, we subjected FL-IRE1-LD to ion mobility mass-spectrometry (IM-MS) analysis. The collision cross section of the +20 charge state dimer measured by IM-MS was 5295.07 $Å^2$ ± 207.75 (Figure 4.10).

Following IM-MS analysis, we built a structural model using the crystal structure of hIRE1-LD (PDB ID:2HZ6) as a template and the I-TASSER webserver (Zhang, Y., 2008) to predict the conformation of the C-terminus (residues I391-S450, truncated, in the X-ray structure) and the regions of IRE1-LD unresolved in the X-ray structure, followed by a 500 ns MD simulation (Figure 4.11). The collision cross section values during the MD simulation (50 – 500 ns) were calculated using IMPACT software (Marklund et al., 2015) and averaged, resulting in a CCS of 5258.02 $Å^2$ ±105.62 (Figure 4.10), using the projection approximation method. Previous studies of disordered systems indicates the projection approximation method is in better agreement with experimental data (Smith et al., 2009; Pagel et al., 2013), however we cannot rule out gas phase compaction of our protein for ion mobility measurements, that has previously been observed for other disordered systems (Devine et al., 2017; Jhingree et al., 2017).

The good agreement between IM-MS and MD CCS suggested that the protomers in the hIRE1-LD dimeric conformation are arranged in a similar manner to those observed in yIRE1-LD (Credle et al., 2005) (PDB ID: 2BE1).

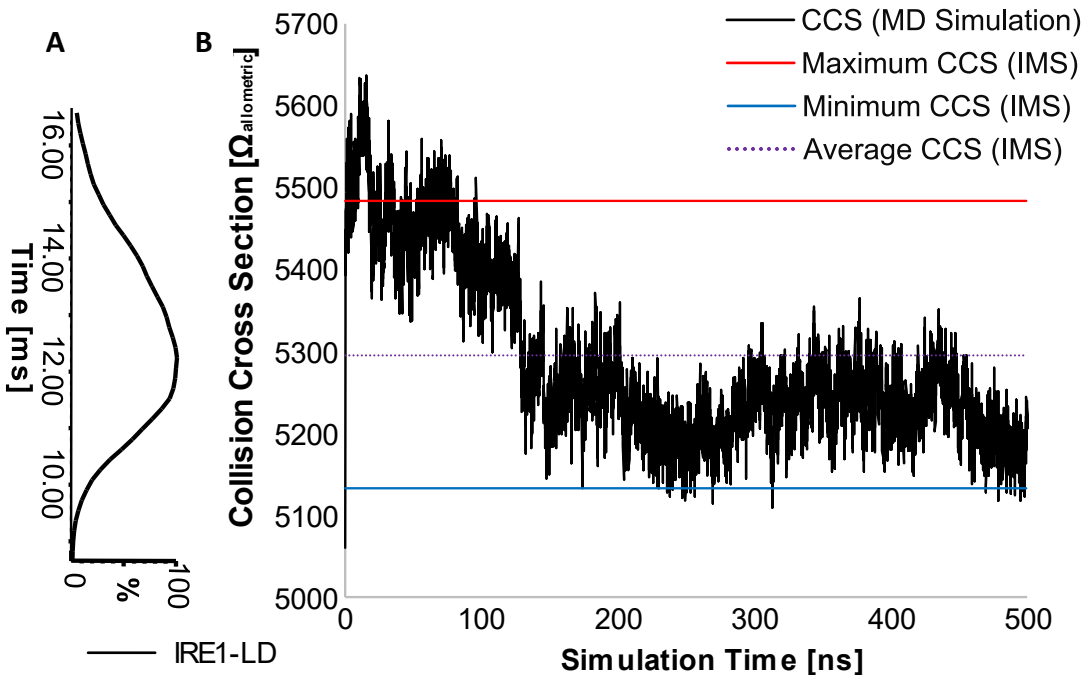

Figure 4.10

(A) The arrival time distribution of the IRE1-LD dimer in the presence of DTT determined by IM-MS. This was used to calculate collision cross section in B, with the average (purple), maximum (red) and minimum (blue) represented. The CCS calculated from the MD simulation is shown in black, which is in good agreement with the IM-MS data (5258.02 $Å^2$ ±105.62 vs 5295.07 $Å^2$ ± 207.75)

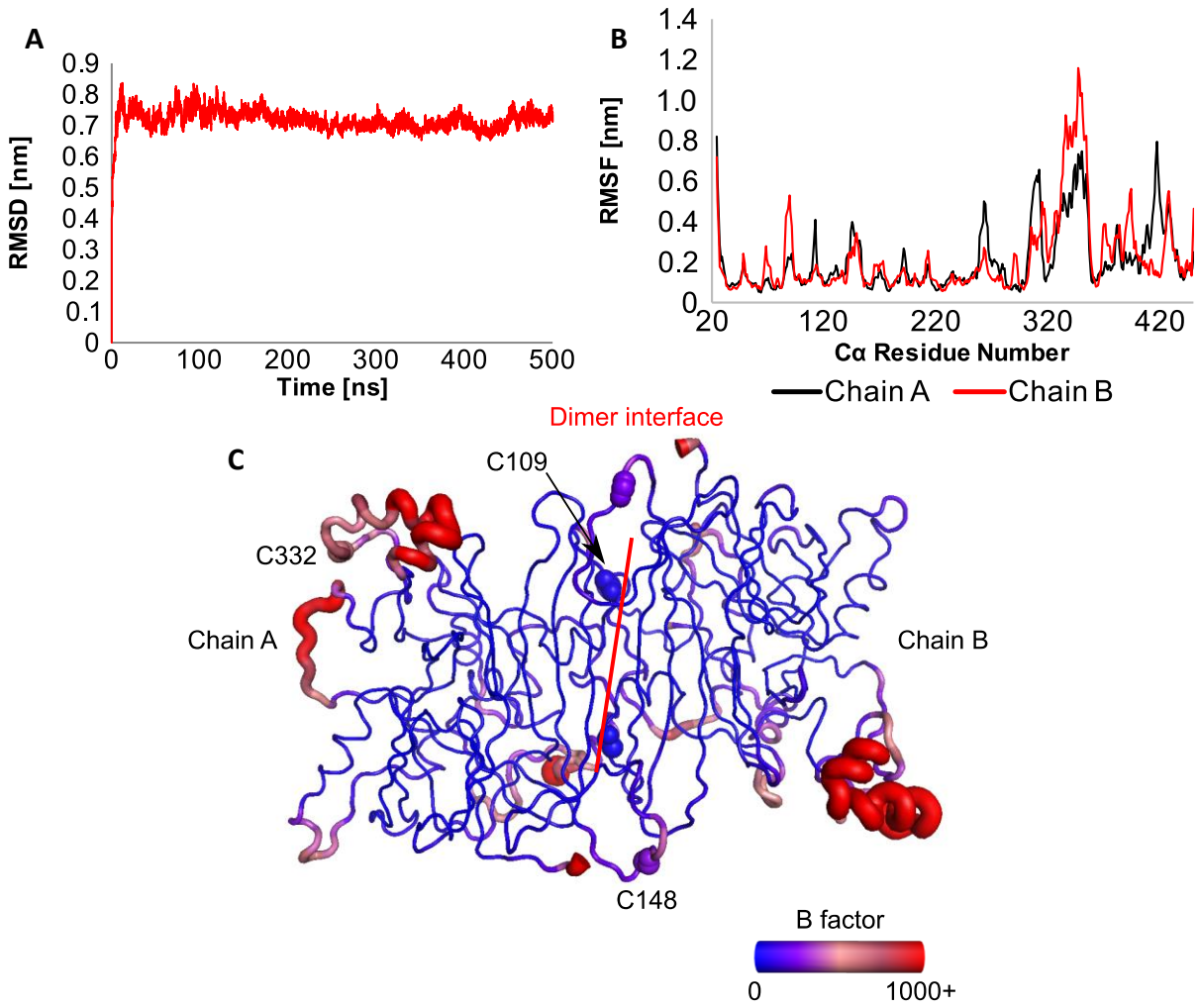

Figure 4.11

(A) The RMSD of the coordinates from the MD trajectory, compared to the energy minimised structure prior to the start of the trajectory. The simulation stabilises after approximately 50 ns. (B) The RMSF of the Cα from each residue throughout the 500 ns MD simulation. A large degree of flexibility is observed in the C-terminal of the protein. (C) The RMSF values from the MD simulation have been converted to B factors and are mapped onto the average coordinates from the MD simulation.

4.3 Oligomerisation of IRE1-LD in the presence of peptide

4.3.1 Substrate dependent oligomerisation of IRE1-LD

Previous studies in yeast indicated that oligomerisation is essential for IRE1 activation (Credle et al., 2005). More recently, there has been evidence that oligomerisation is crucial for IRE1 activation in humans (Karagoz et al., 2017), with these events in yeast and human homologs induced by peptide binding. A signal peptide, ΔEspP, has previously been shown to interact with yIRE1-LD *in vitro* (Gardner, B.M. and Walter, 2011). For human IRE1, it was demonstrated that, while ΔEspP binds to IRE1, it does not induce oligomerisation (Carrara et al., 2015), however this was ascertained by SEC-MALLS, where there is a possibility of dissociation of the complex due to the dilution effect upon passage through the SEC column.

To ascertain if ΔEspP has an effect on the oligomeric state of FL-IRE1-LD, we performed DLS measurements. A significant decrease in the decay rate of the correlation function was observed (15637 s^{-1} vs 958.4 s^{-1}) upon addition of ΔEspP to FL-IRE1-LD (Figure 4.13), which is characteristic of an increase in the particle size. Regularisation of this data indicated an appearance of large molecular weight species (17.3 nm ± 14.9 v 6.08 nm ±2.02 and 404.9 nm ± 268.17), indicating ΔEspP triggers oligomerisation of FL-IRE1-LD (Figure 4.12). As a control, to ensure ΔEspP had not aggregated and was the cause of the large particles, we performed DLS on 10 µM ΔEspP. The correlation plot indicated that the correlation function decayed rapidly, indicating a lack of large particles (Figure 4.14)

To elucidate the effect of ΔEspP on concentrations of IRE1-LD below the K_d ascertained for dimerisation, we performed fluorescence polarisation analysis of 50 nM FITC-labelled FL-IRE1-LD and increasing concentrations of ΔEspP (Figure 4.15). A change in fluorescence polarisation was observed upon increasing concentrations of ΔEspP, which would likely indicate IRE1-LD forms oligomers upon binding of ΔEspP, as we would not anticipate observing a noticeable change in polarisation of FITC FL-IRE1-LD upon binding of a 3 kDa peptide.

Interestingly, this effect was also observed in the non-dimerising IRE1-LD (D123P) mutant, which indicates that either peptide induced oligomerisation does not require dimer formation conformational changes in monomeric IRE1-LD or the non-dimerising mutant can form dimers with a weaker dimerisation constant. To determine if IRE1-LD (D123P) still formed oligomeric species like FL-IRE1-LD (WT), we performed DLS in the presence and absence of 10 µM ΔEspP

(Figure 4.16). This indicated, while some larger particles were formed (65.89 nm ±33.91), these were not comparable in size to those observed in FL-IRE1-LD WT. Unlike FL-IRE1-LD (WT), a significant intensity fraction of IRE1-LD (D123P) remained monomeric/dimeric (4.13 nm±2.16).

Interestingly, upon addition of an increased concentration of protein and peptide (20 µM FL-IRE1-LD, 40 µM ΔEspP), an insoluble precipitate was observed, which could be observed by monitoring the OD400. Taking into account the DLS data, this could indicate that increased concentrations of ΔEspP induce the formation of large oligomers that are no longer soluble. This effect was observable for both the IRE1 cLD and IRE1-LD constructs (Figure 4.17).

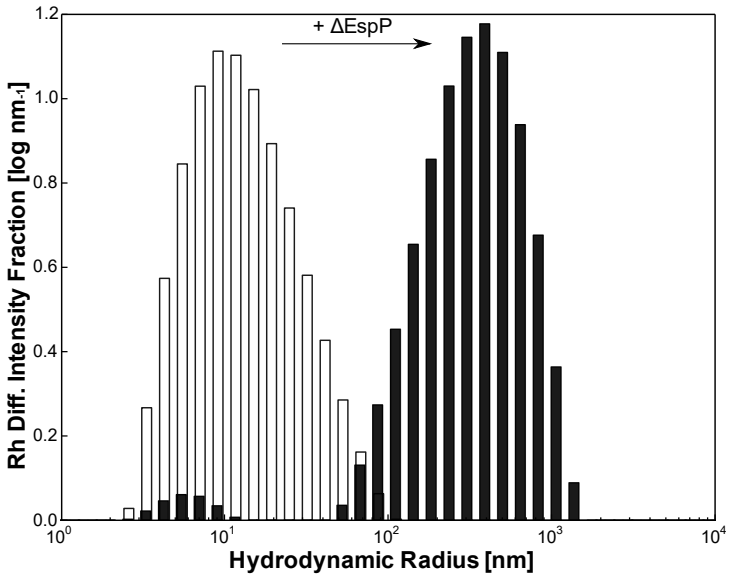

Figure 4.12

Regularisation analysis of 5 µM IRE1-LD in the absence (white) and presence (black) of 10 µM ΔEspP indicates an increase in the particle size upon addition of ΔEspP (17.3 nm ± 14.9 v 6.08 nm ±2.02 and 404.9 nm ± 268.17).

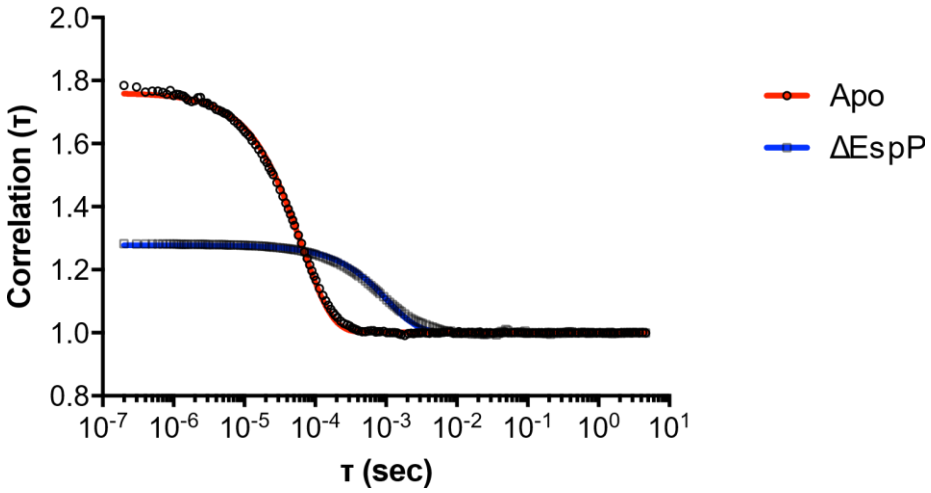

Figure 4.13

One phase exponential decay fit of the correlation function of 5 µM FL-IRE1-LD in the absence (red) and presence (blue) of 10 µM ΔEspP. A reduction in the decay constant (15637 s^{-1} vs 958.4 s^{-1}) is observed upon addition of 10 µM ΔEspP, characteristic of the formation of larger particles.

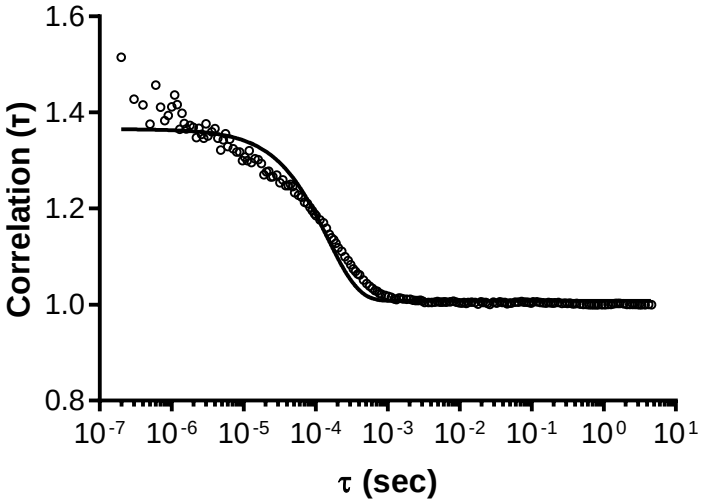

Figure 4.14

One phase exponential decay fit of the correlation function of 10 µM ΔEspP. The correlation function decays rapidly and does not fit well to a one phase exponential decay (R^2=0.9762) indicating the ΔEspP peptide is not aggregating

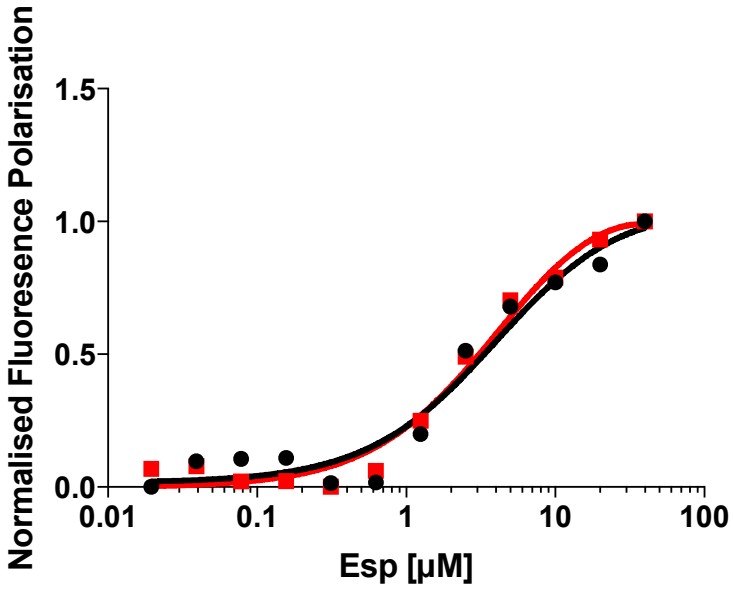

Figure 4.15

Fluorescence polarisation data of 50 nM FITC-labelled FL-IRE1-LD (WT) (black) and IRE1-LD (D123P) (red) upon increasing concentrations of ΔEspP. This indicates IRE1-LD is particularly sensitive to unfolded substrates

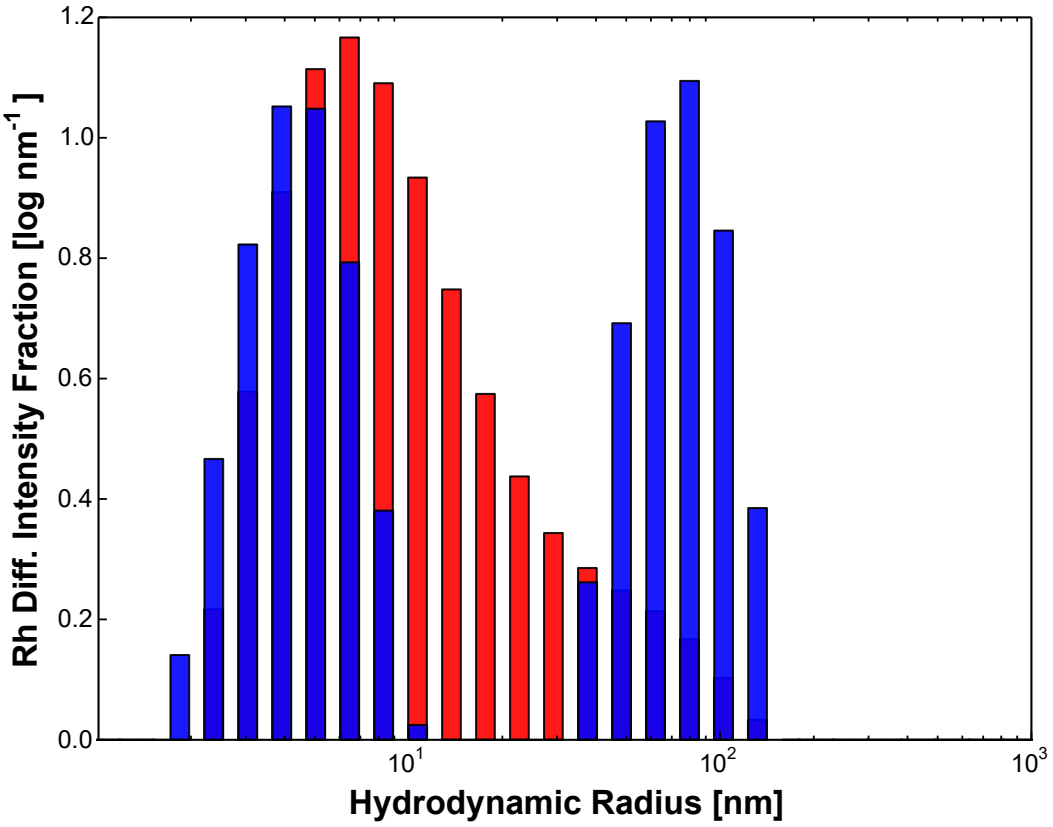

Figure 4.16

Regularisation plot obtained from DLS data of 5 µM IRE1-LD (D123P) in the absence (red) and presence (blue) of 10 µM indicates a change in estimated hydration radius from 15.55 nm ± 19.31 to 4.13 nm±2.16 and 65.89 nm ±33.91. This indicates that larger species of IRE1-LD (D123P are formed upon addition of ΔEspP, however, unlike IRE1-LD (WT) a significant fraction of these are still monomeric/dimeric.

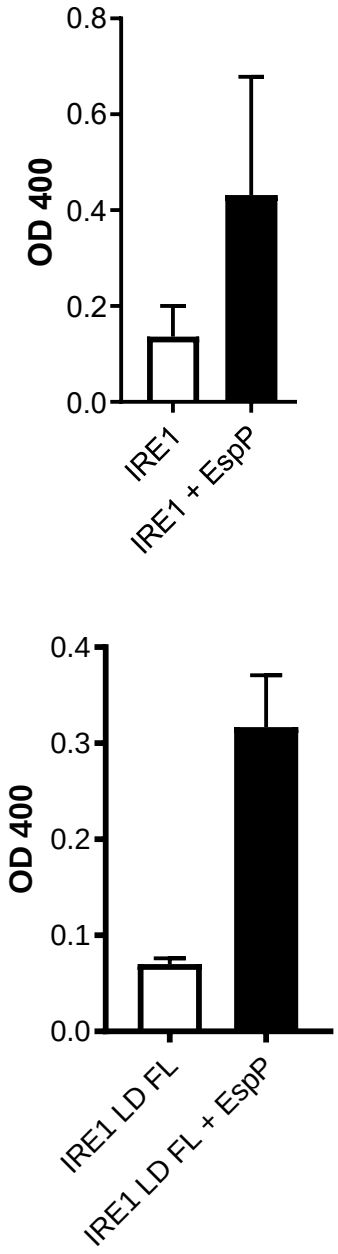

Figure 4.17

(Above) OD400 data of 20 µM IRE1 cLD in the absence and presence of 40 µM ΔEspP. An increase from 0.168±0.061 to 0.431±0.247 was observed upon addition of ΔEspP. Data is an average of five repeats (± standard deviation). (Below) The same effect is observed for our full length construct (average of three repeats).

4.3.2 Structural insights of substrate induced oligomerisation of IRE1-LD

Karagoz et al. proposed that IRE1-LD undergoes a conformational change upon binding to unfolded substrate that triggers oligomerisation. While addition of a high concentration of ΔEspP led to the formation of insoluble particulates, it was possible to negate this effect by using a low stock concentration of ΔEspP. To investigate the conformational changes, we performed an ^{15}N-^{1}H BEST-TROSY on 30 µM FL-IRE1-LD in the presence of an equimolar concentration of ΔEspP (Figure 4.19). No significant changes in peak positions were observed upon peptide addition. As we would expect the larger oligomeric species of FL-IRE1-LD observed by DLS (see section 4.3.1) will be invisible to NMR due to their slow tumbling time, we performed peak height analysis of FL-IRE1-LD in the presence and absence of ΔEspP and compared the peak intensity of 18 well resolved peaks. A 20% reduction in the relative peak intensity (average I(IRE1+ΔEspP)/I(IRE1) of 0.80 ±0.11) suggested that a slight loss of flexibility and/or an increase in tumbling time induced by peptide binding and/or consequent oligomerisation.

To ascertain the effect of individual residues, we compared the peak intensities of individual residues in absence and presence of ΔEspP. A significant reduction (reduction in signal to noise >5) was observed in 7 of the 18 peaks analysed (Figure 4.18). This may indicate the effect of peptide binding or peptide triggered oligomerisation leads to a reduction in flexibility in these residues, while others still retain their flexibility or that these residues are part of the proposed oligomerisation interface (Karagoz et al., 2017) that leads to µs-ms exchange in these residues and increased dipole induced relaxation, both leading to a broadening of these resonances.

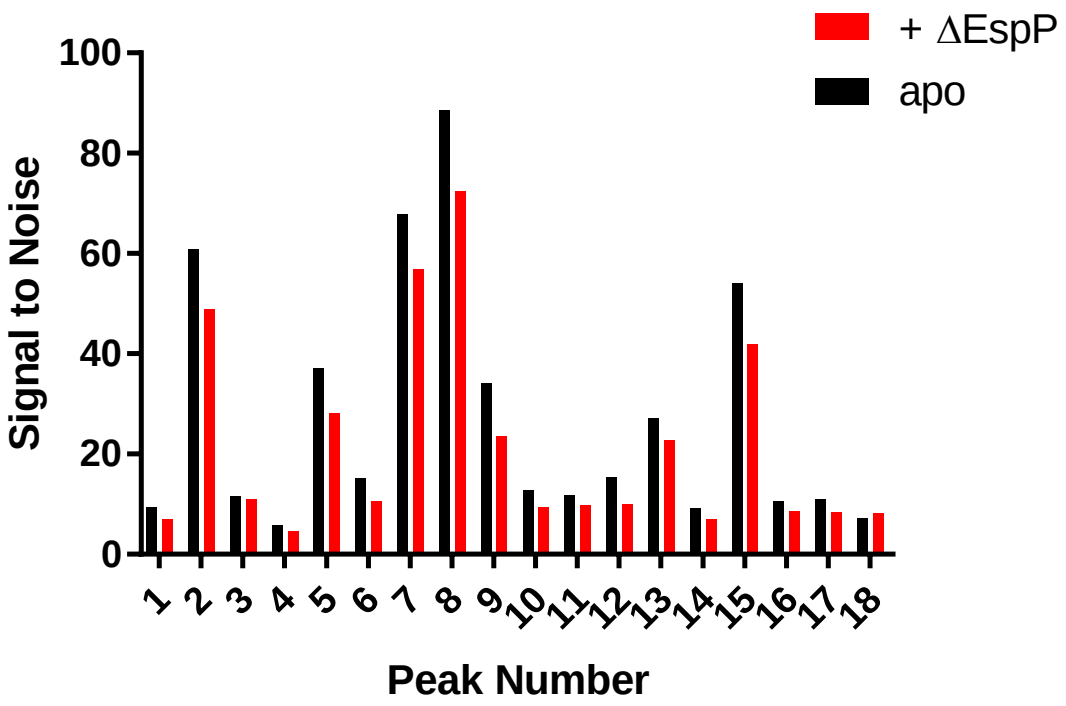

Figure 4.18

Signal to noise ratios of 18 isolated peaks from the spectrum in Figure 4.19. A significant reduction (>5) in the signal to noise ratio is observed for peaks 2, 5, 7, 8, 9, 12 and 15.

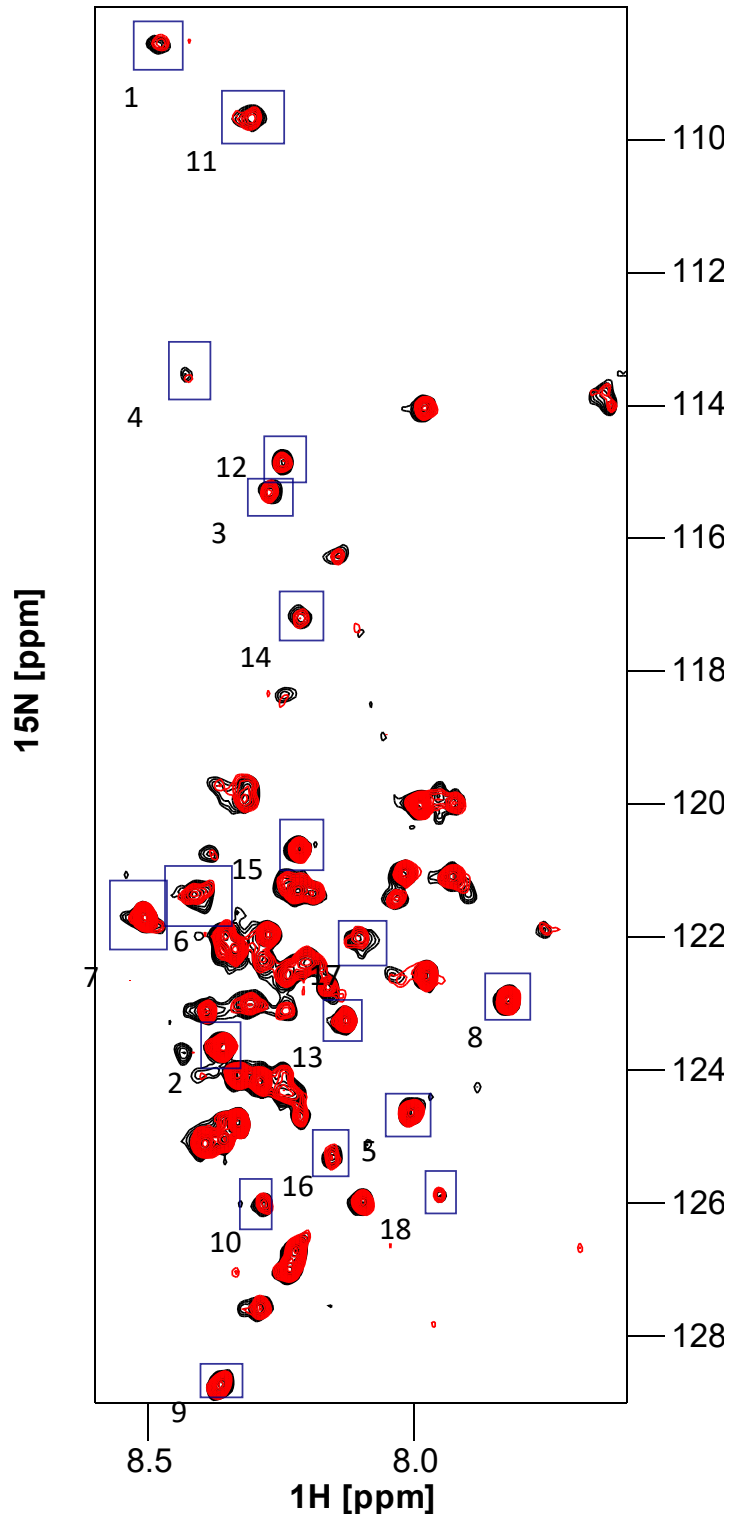

Figure 4.19

^{15}N-^{1}H TROSY of 30 µM of ^{15}N labelled FL-IRE1-LD in the absence (black) and presence (red) of an equimolar concentration of ΔEspP. The peaks used for peak height analysis are highlighted in blue boxes.

4.4 Alternative route of IRE1-LD oligomerisation through formation of intramolecular disulphide bonds

4.4.1 IRE1-LD forms covalent oligomers within 24-48 hours

As discussed in the introduction (section 1.4), the formation of intermolecular disulphide bonds has been shown to be important for regulation of for regulation of hIRE1-LD activity, with this process mediated by an ER protein disulphide isomerase, PDIA6 (Eletto et al., 2014). It has also been suggested that C148 and C332 are responsible for covalent oligomerisation of hIRE1-LD oligomerisation (Liu et al., 2003). To characterise this process we performed non-reducing SDS-PAGE analysis of FL-IRE1-LD on samples that were incubated under non-reducing conditions for between 1 and 48 hours (Figure 4.21, Figure 4.22, Figure 4.23). The fraction of IRE1-LD that formed disulphide bonds was calculated as relative band intensities of oligomeric (disulphide bonded) vs. monomeric (non-disulphide bonded) species. Our results indicated that disulphide bond formation is a slow process that takes several hours to complete.

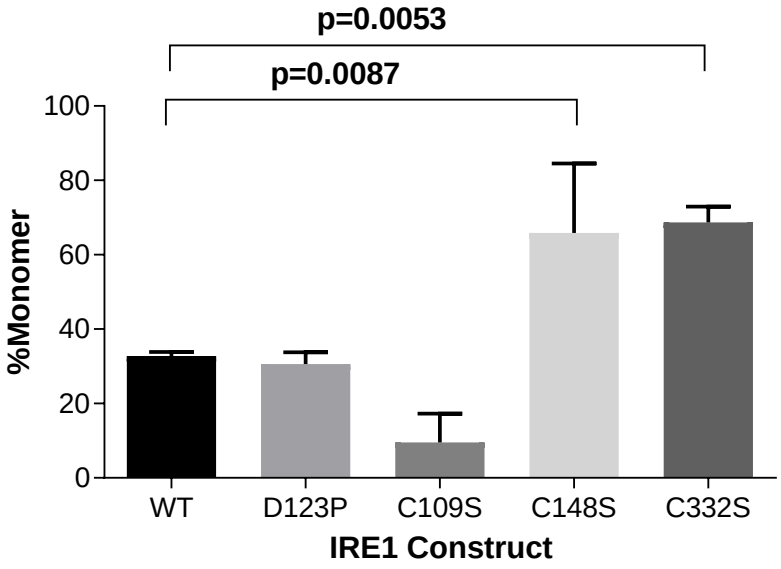

Figure 4.20

The percent band intensity of the non-reduced monomeric band, relative to the reduced monomeric band at 24 hours for the specified constructs indicates C148 and C332 are critical for disulphide bond formation. Adjusted p-values from a one-way ANOVA followed by Dunnet's multiple comparison test (with data from section 5.2) are stated.

To validate that C148 and C332 are involved in disulphide bond formation in hIRE1-LD, we performed single point mutagenesis to produce C109S, C148S and C332S variants of FL-IRE1-LD. In agreement with previous observations (Liu et al., 2003), C148 and C332 are both essential for disulphide bond formation, with a significant reduction in the rate of disulphide bond formation observed with these two mutants (Figure 4.20)

The cysteine residues that are important for disulphide bond formation, C148 and C332, are present on flexible loops of IRE1-LD, therefore it is not possible to ascertain from a structure which conformation drives the process of intermolecular disulphide bond formation. We first asked whether non-covalent dimerization of IRE1-LD is necessary for formation of disulphide bonds by performing non-reducing SDS-PAGE, as in section 4.4.1. Both FL-IRE1-LD (WT) and IRE1-LD (D123P) formed disulphide bonds at a similar rate (32.75% ±1.12 and 30.66% ±3.15 of monomeric band respectively), suggesting that dimerization isn't a pre-requisite to disulphide bond formation.

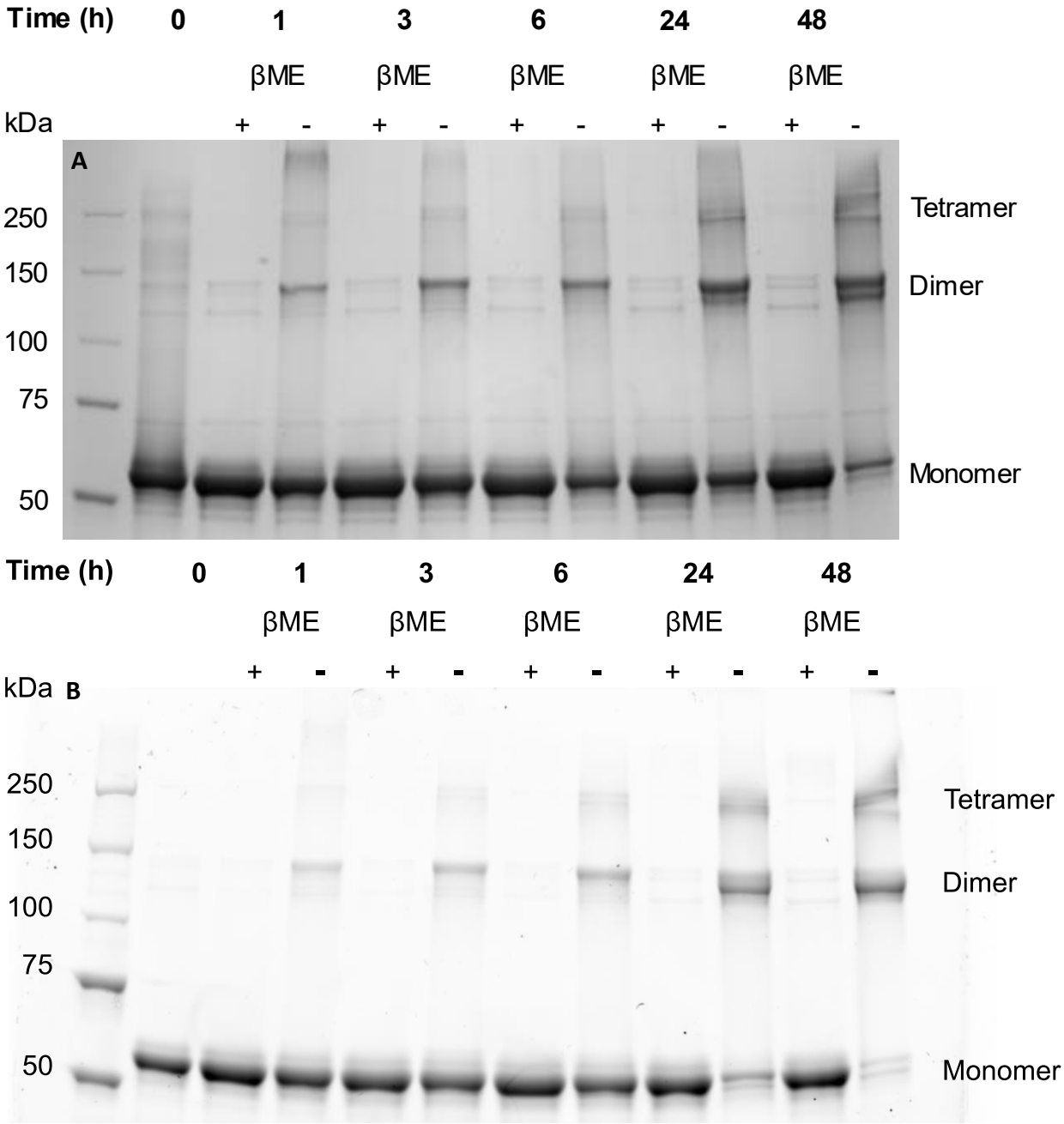

Figure 4.21

SDS-PAGE gels of FL-IRE1-LD (A) WT and (B) D123P, at various timepoints, relative to buffer exchange from a reducing buffer to a non-reducing buffer. Reducing and non-reducing gel samples re shown and monomeric, dimeric and tetrameric band positions are indicated.

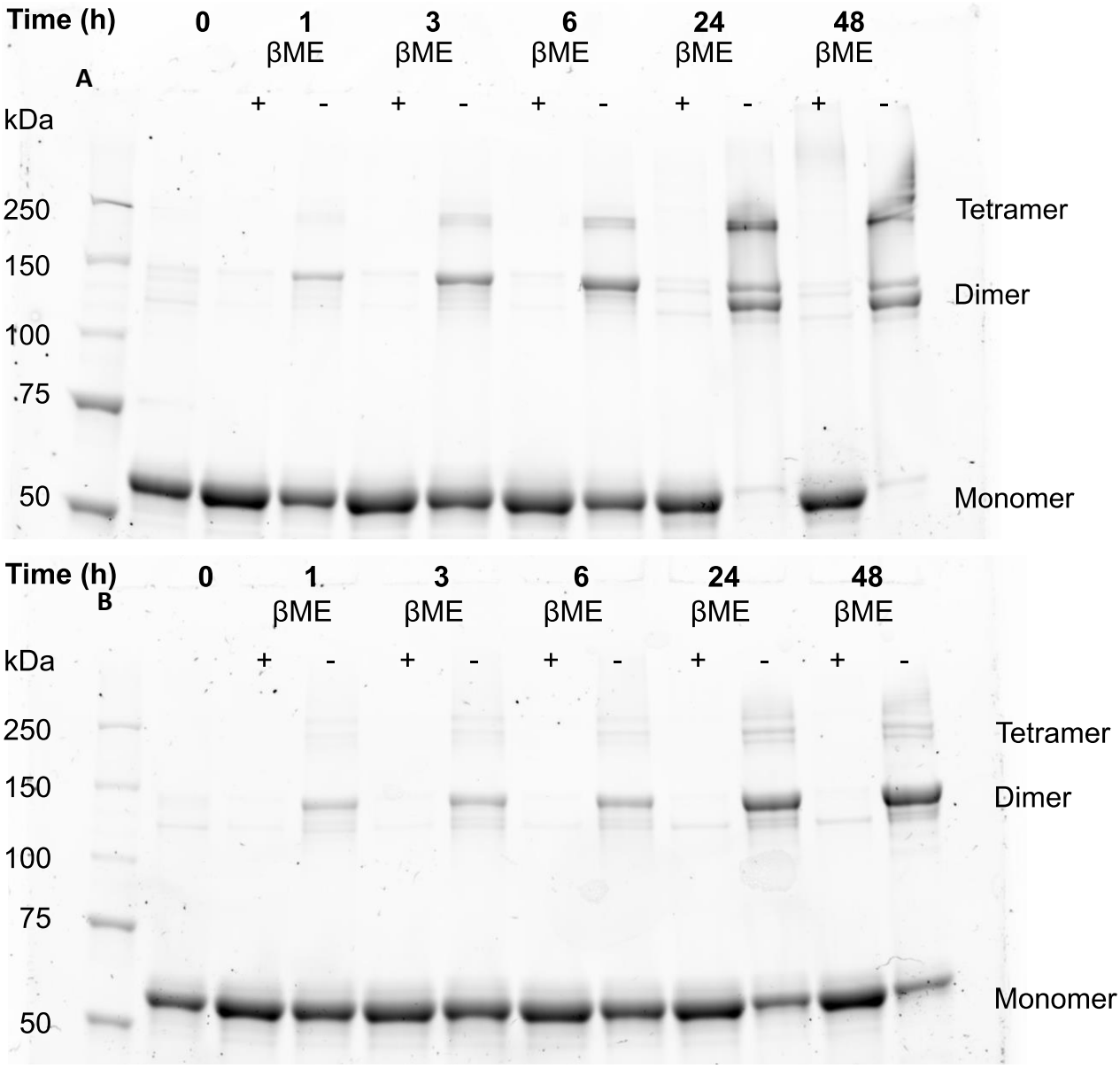

Figure 4.22

SDS-PAGE gels of IRE1-LD (A) C109S and (B) C148S, at various timepoints, relative to buffer exchange from a reducing buffer to a non-reducing buffer. Reducing and non-reducing gel samples re shown and monomeric, dimeric and tetrameric band positions are indicated.

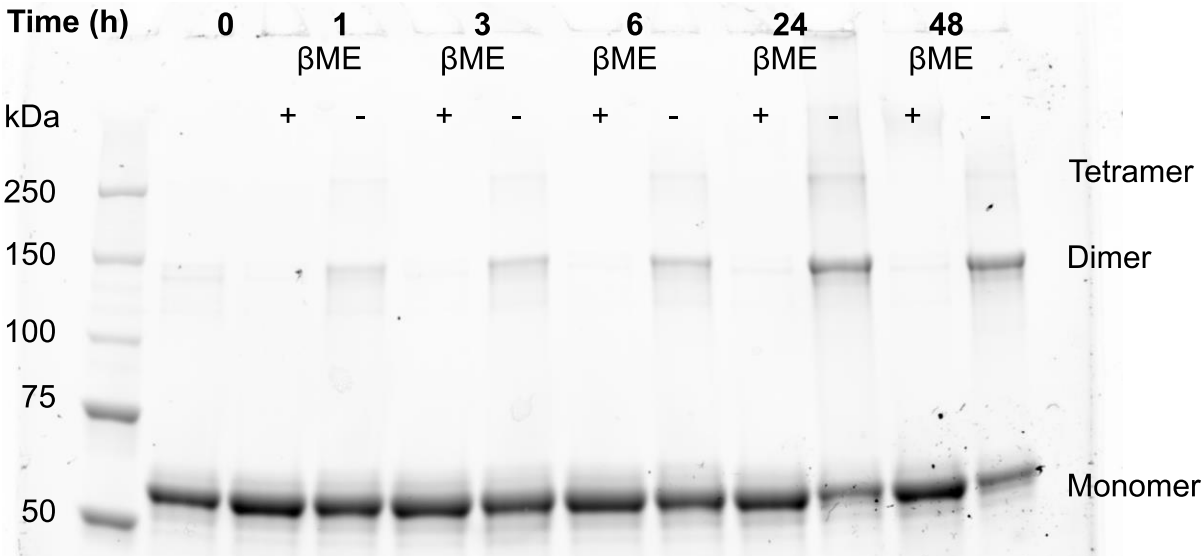

Figure 4.23

SDS-PAGE gels of IRE1-LD C332S, at various timepoints, relative to buffer exchange from a reducing buffer to a non-reducing buffer. Reducing and non-reducing gel samples re shown and monomeric, dimeric and tetrameric band positions are indicated.

As dimerization is not essential for disulphide bond formation, we next questioned whether the structures of the non-covalent and covalent dimers are different. To ascertain if the structural features of these dimers were different, we performed IM-MS and calculated the CCS of the +20 charge state distribution (Figure 4.24). The CCS of the disulphide linked dimer was 5454.07 $Å^2$ ± 220.05, compared to 5295.07 $Å^2$ ± 207.75. Both these values are within the error of each other, meaning it is not possible to conclude that both these structures are different.

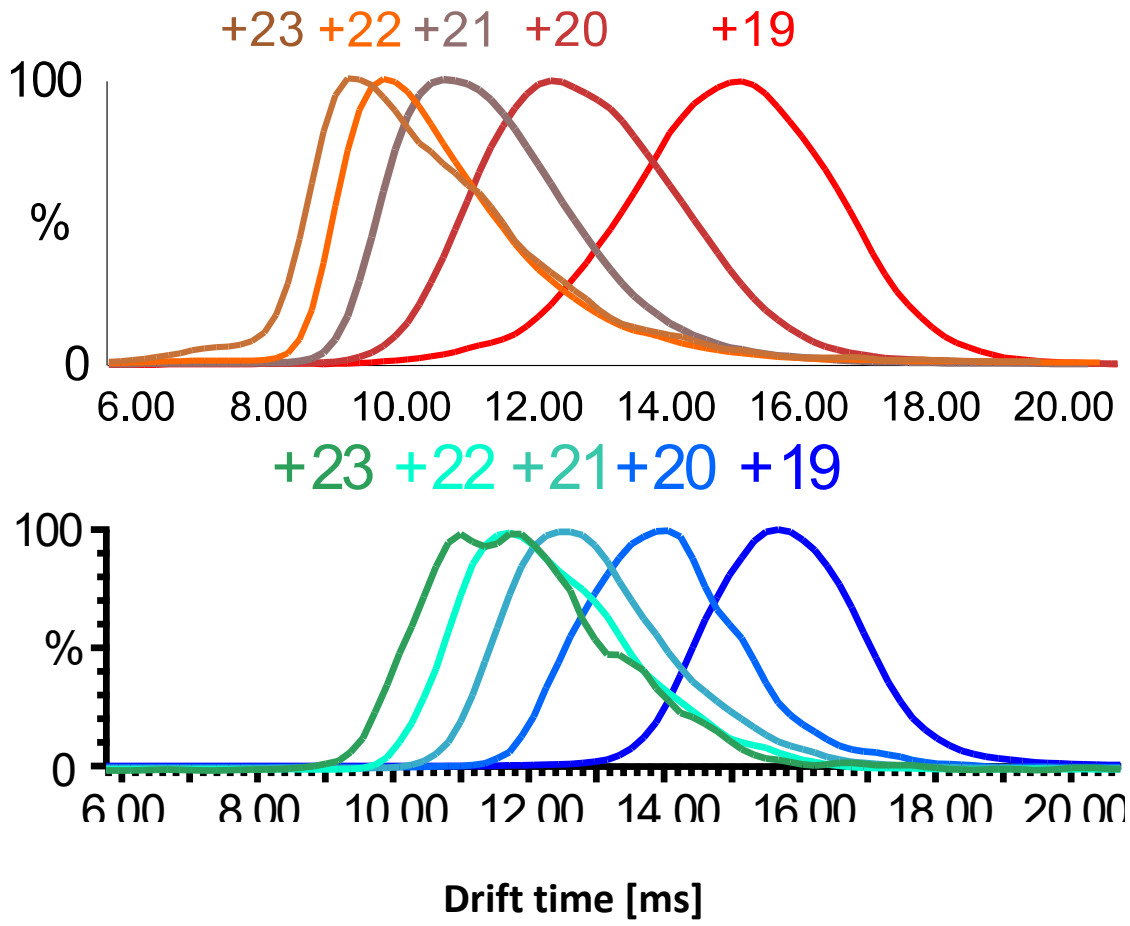

Figure 4.24

Arrival time distribution of various charge states of IRE1-LD in the presence (above) and absence (below) of 5 mM DTT. The +20 charge state was used for calculation of CCS

4.4.2 Structural characterisation of the disulphide linked dimer

To gain structural insights into the process of disulphide bond formation, we performed NMR characterisation for FL-IRE1-LD (WT) (Figure 4.26), the non-dimerising variant, IRE1-LD (D123P) (Figure 4.28) and single point cysteine mutants, C109S, C148S and C332S (Figure 4.27). For each construct, ^{15}N labelled samples were incubated for 24 hours under non-reducing and reducing (in the presence of 5 mM DTT) conditions prior to acquiring NMR spectra. All constructs revealed similar changes between non-reducing and reducing samples. In the non-reducing samples, approximately 20 peaks either disappeared or significantly reduced in intensity, suggesting that the intrinsic flexibility of these residues has decreased. Interestingly, for the C332S construct, the resonances that were perturbed upon this

mutation (Figure 4.9) did not disappear upon disulphide bond formation (Figure 4.27), which further suggests that disulphide bond formation of C332 leads to a reduction in loop flexibility, which leads to disappearance of several peaks in the FL-IRE1-LD (WT) spectrum. Further evidence for this conclusion is demonstrated by the fact that the majority of these peaks do not belong to the C-terminal linker and are apparently located in the intrinsically disordered regions around C148 and C332 (Figure 4.25). Peak height analysis of four peaks in this region of the three cysteine mutants indicated a similar trend to that observed by SDS-PAGE (Figure 4.29). All together, these observations reveal that two cysteine residues C148 and C332, located in the intrinsically disordered regions of FL-IRE1-LD, are involved in covalent dimerization.

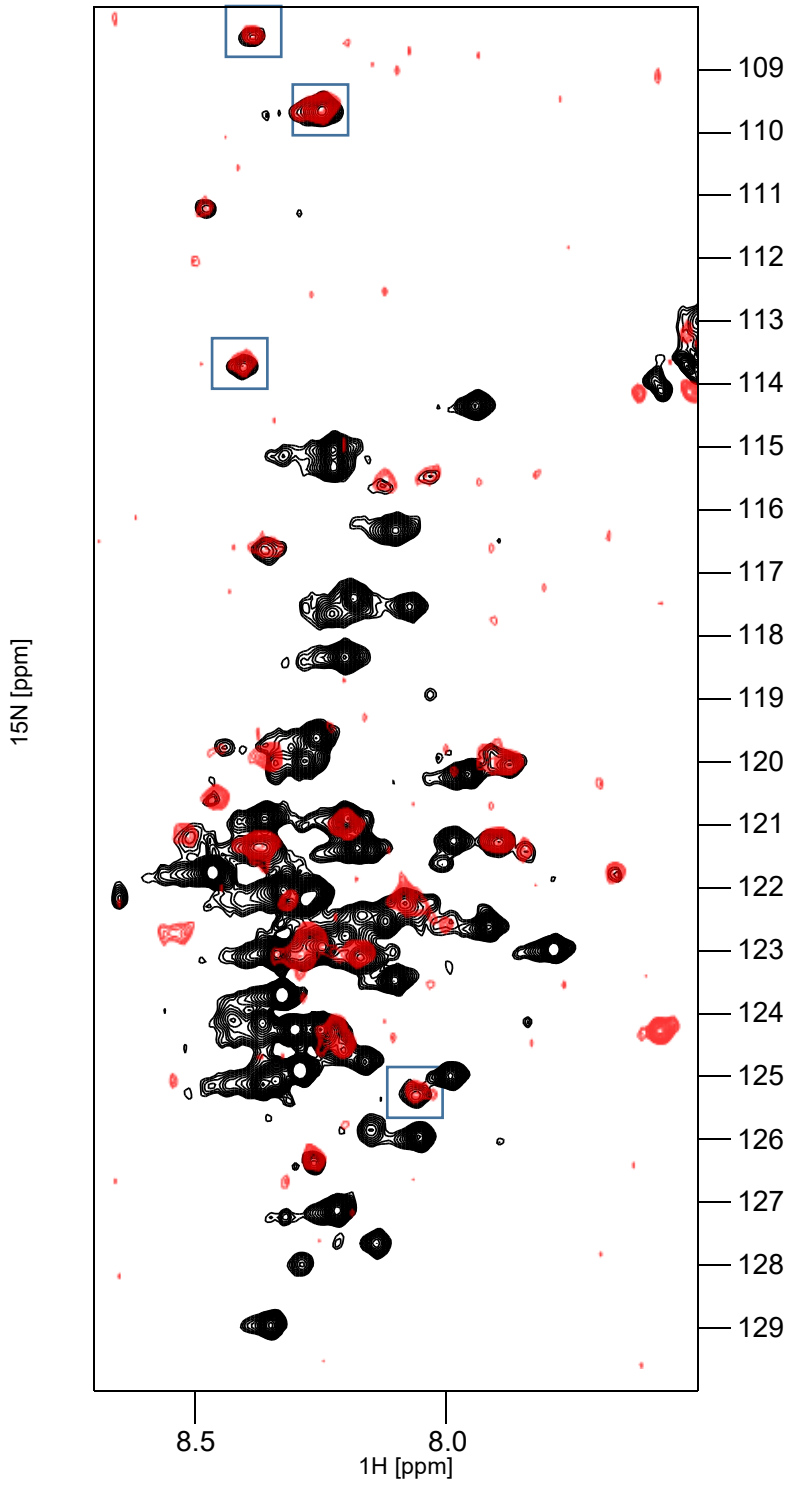

Figure 4.25

^{15}N-^{1}H TROSY spectra of (A) 50 µM FL-IRE1-LD (black) and 30 µM IRE1-cLD (red) in HMK + 5 mM DTT Spectra recorded at 288K. Peaks used for peak height analysis in Figure 4.29 are highlighted with blue boxes. These do note emanate from the juxtamembrane linker

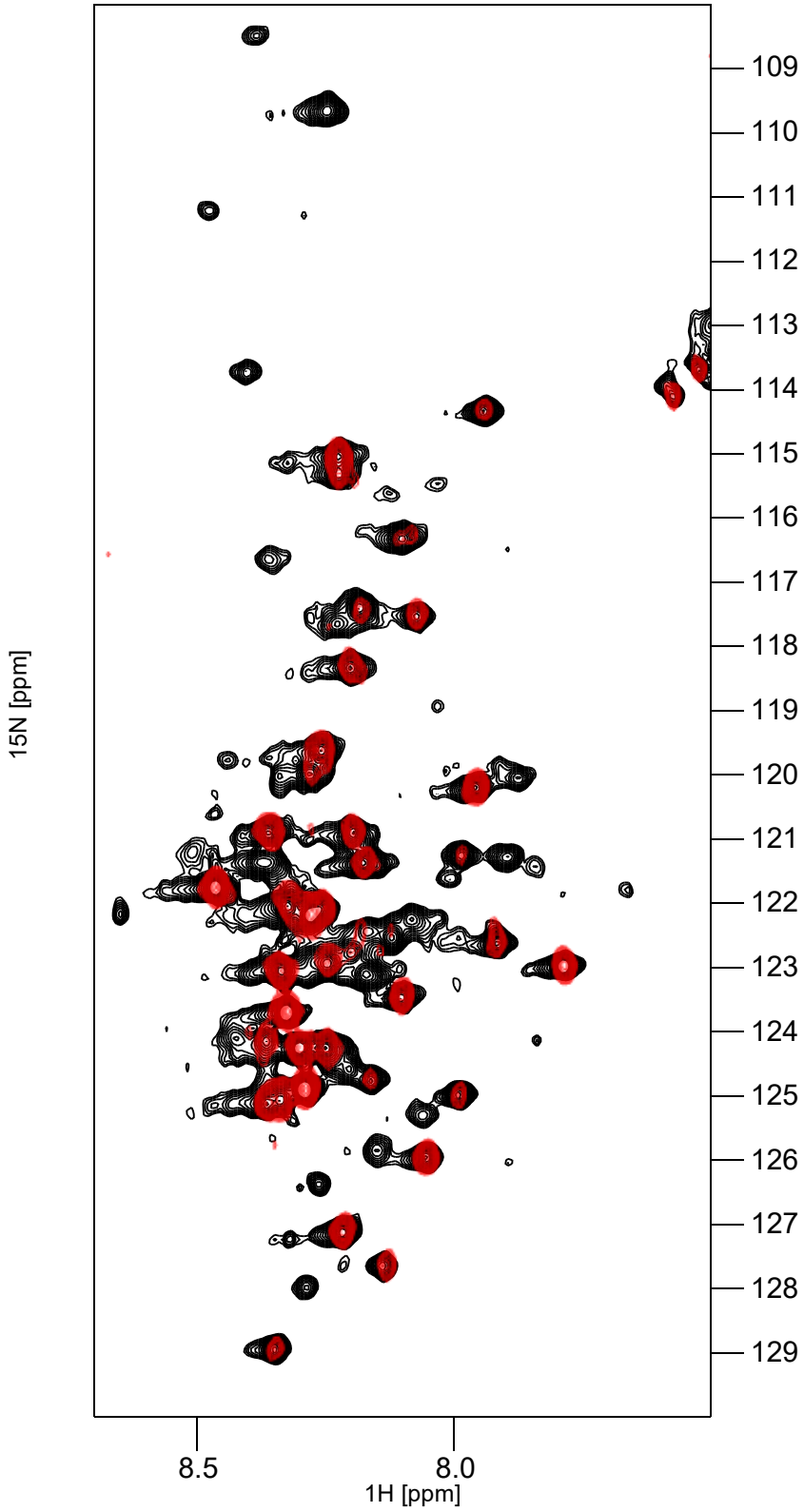

Figure 4.26

^{15}N-^{1}H TROSY spectra of 50 µM FL-IRE1-LD in the presence of 5 mM DTT (black) and 24 hours after removal from a reducing buffer. ~30 peaks disappear or reduce in intensity, suggesting a loss of flexibility in these residues. Spectra recorded at 288K.

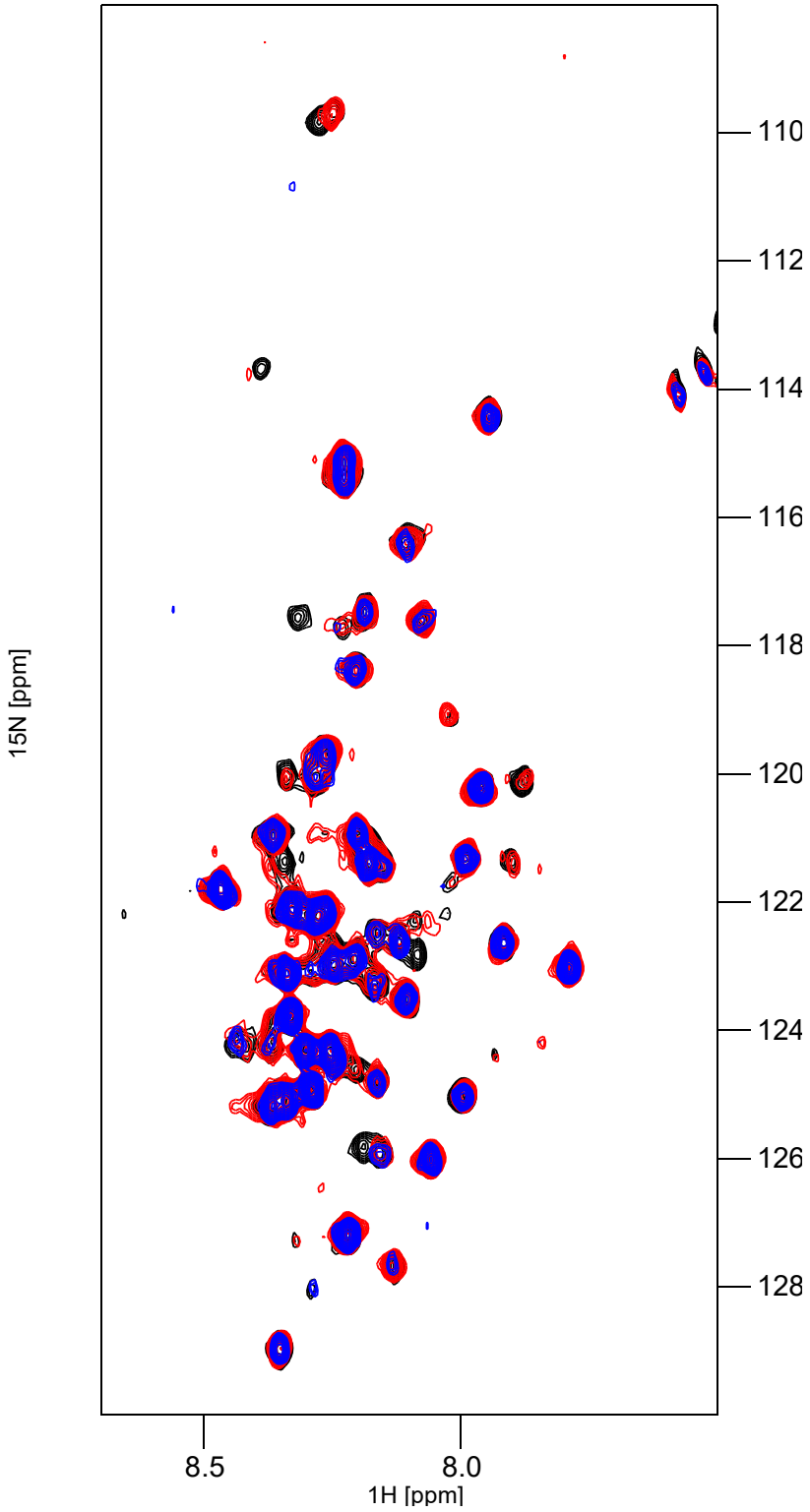

Figure 4.27

^{15}N-^{1}H TROSY of 50 µM IRE1-LD (C109S) (blue), C148S (red) and C332S (black), 24 hours after removal from a reducing buffer. Spectra recorded at 288K.

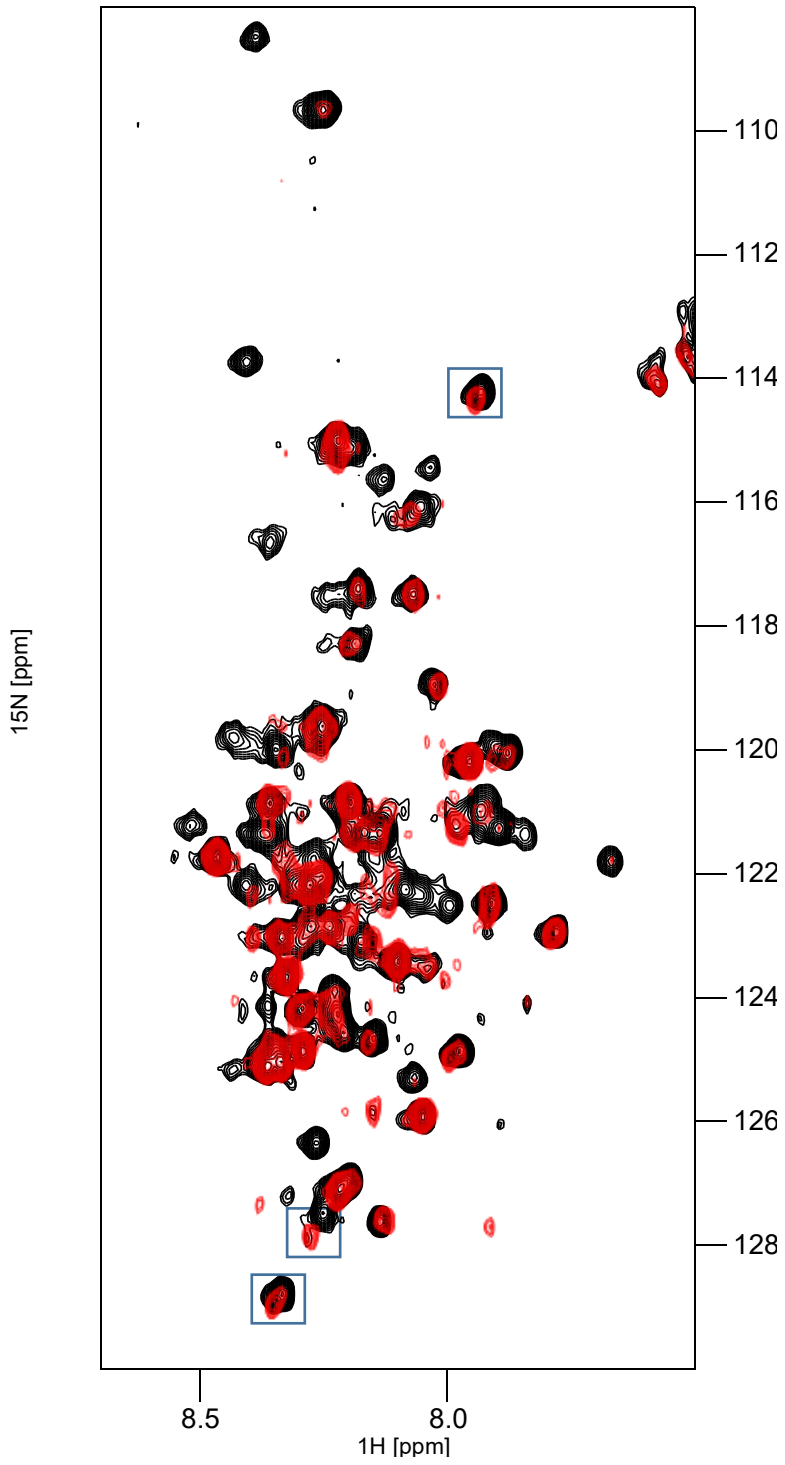

Figure 4.28

^{15}N-^{1}H TROSY of 50 µM IRE1-LD (D123P) in the presence (black) and absence (red) of 5 mM DTT. 3 minor chemical shift changes are observed (highlighted in blue boxes), unlike other constructs. Spectra recorded at 288K

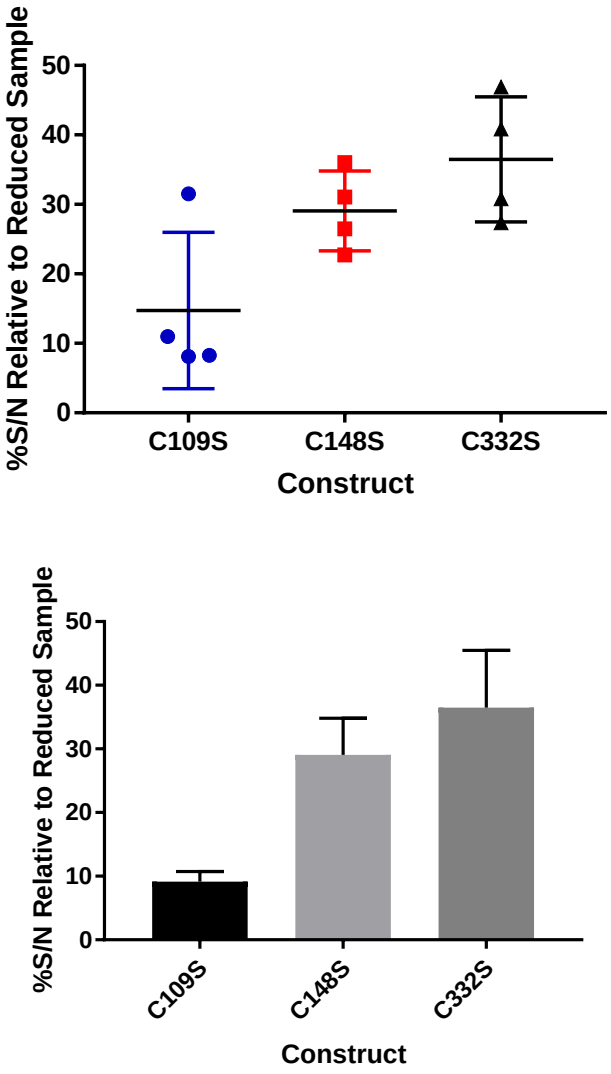

Figure 4.29

(Above) The signal to noise of the non-reduced sample, relative to the reduced sample, of the peaks specified in Figure 4.25. The mean and SD are displayed. One outlier was identified in C109S using the Grubbs method (α=0.05). (Below) The average data from above ±SD, with the outlier removed. Similar trends are observed between NMR

4.5 Conclusions

We have shown, in the absence of substrate, the main species of the IRE1-LD is the dimer at concentrations in excess of the K_d observed for dimerisation. The K_d obtained for this interaction is significantly lower than that previously observed for yIRE1 (~10 μM) (Credle et al., 2005), indicating subtle differences in the activation pathways of these homologs. While the structural model produced for the hIRE1 dimer has not been validated at a single residue

level, it provides some insight into the relative flexibility of the disordered regions around C332 of IRE1-LD, in particular the second loop that contains C332 has a large degree of flexibility, which is further indicated by the sharpness of peaks in this region in the ^{15}N-^{1}H TROSY spectrum.

The functional importance of this region is highlighted by the process of disulphide bond formation, previously shown to be important for protein regulation. The data in this chapter shows this process is slow and that C148 and C332 are important for this process, in agreement with previous literature data (Liu et al., 2003). This leads to a loss of flexibility in the loop regions, demonstrated by a loss of intensity In the NMR spectrum of peaks around this region. While we cannot conclude that the structure of the disulphide linked dimer and non-disulphide linked dimer are significantly different from our IM-MS data, disulphide bond formation does lead to structural changes, as indicated by the loss of intensity of residues in the ^{15}N-^{1}H TROSY upon disulphide bond formation.

The data in this chapter supports a model for substrate triggered oligomerisation with DLS indicating that upon introducing a model peptide, large oligomeric species are formed. IRE1-LD appears to be particularly sensitive to unfolded substrate induced changes, with fluorescence polarisation data indicating that as little as 50 nM of IRE1-LD is required for these conformational changes to occur. Interestingly this also observed with the non-dimerising mutant, possibly suggesting that an unfolded protein can drive oligomer formation of the small population of dimeric IRE1-LD in solution.

Our NMR data indicates that the structured residues of IRE1-LD are undergoing conformational exchange on the intermediate regime. It was still not possible to observe these residues by changing the temperature at which the spectrum was recorded, nor by recording spectra at 30, 50 or 100 µM (see chapter 5.4) of IRE1-LD.

While we have been able to identify regions of the protein through construct design (i.e. residues close to C332 and those emanating from the linker), we did not assign the protein as the 15N-1H TROSY NMR experiments took ~5 hours to complete. This coupled with spectral overlap due the disordered nature of the observed residues and the loss of sensitivity when performing 3D NMR experiments, means it would have likely required a significant amount of NMR time to assign some of these residues. In hindsight, HSQC based experiments could have

been suitable as the spectra we observe contains mainly sharp peaks due to the fact that the flexible residues relax independently of the protein, thus TROSY based experiments could be reducing the intensity of the peaks emanating from these flexible residues.

In conclusion, our data demonstrates a complex conformational landscape, which is shifted by an unfolded substrate. In the absence of unfolded substrate, we have demonstrated that IRE1-LD forms high affinity dimers, in contrast with those observed in yeast. Similar to yeast, IRE1 is sensitive to substrate triggered dimerisation and oligomerisation (Gardner, B.M. and Walter, 2011). Furthermore, we have characterised the process of covalent bond formation which has been shown to lead to prolonged activation of IRE1-LD, which subsequently requires PDIA6 to attenuate its activity (Eletto et al., 2014) (Figure 4.30).

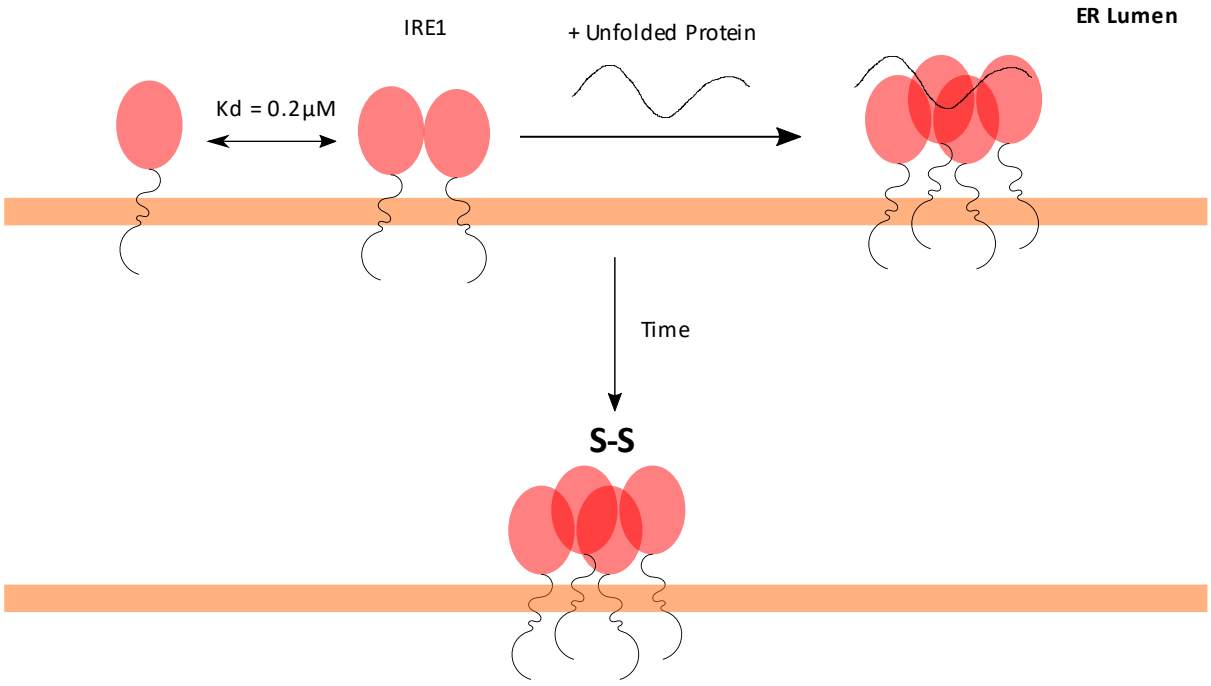

Figure 4.30

The conformational landscape of IRE1-LD involves a concentration dependent dimerization process, a substrate dependent oligomer formation and time dependent disulphide bond formation. The covalent and non-covalent oligomers are likely the active forms of the protein, based on previous literature data (Eletto et al., 2014; Karagoz et al., 2017).

5 The effect of BiP on the conformational landscape of IRE1-LD

There is currently much debate in the literature regarding the interaction of BiP and IRE1-LD (Amin-Wetzel et al., 2017) (Carrara et al., 2015). It is well established that an association between IRE1 and BiP exists under non-stressed conditions that is no longer seen under times of ER stress (Bertolotti et al., 2000). From this observation, and numerous other publications investigating this interaction (Pincus et al., 2010; Kimata et al., 2004; Oikawa et al., 2009), several models have been proposed to explain the nature of the interaction between IRE1-LD and BiP.

The first model suggests that, while IRE1 and BiP are associated under times of ER stress and are no longer associated under stressed conditions, BiP has a limited role in regulating IRE1 activity and the main trigger for activation is a direct interaction between IRE1-LD and unfolded proteins (Pincus et al., 2010), as observed in yeast models. The model reasons that the low cellular concentration of IRE1, and the relative abundance of BiP inside the ER, dictates that the excess of free BiP inside the ER compared to IRE1 would render a BiP regulated model unfeasible.

Contrarily, others suggest that BiP interacts with IRE1 in a canonical manner, via its substrate binding domain and this association renders IRE1 monomeric and inactive, with an assumption that BiP forms a complex with IRE1 monomers and prevents further oligomerisation, thus activation (Bertolotti et al., 2000). More recent data indicates that an active ATPase domain is required to necessary for this interaction and it is mediated by a Hsp40 co-chaperone, ERdj4 (Amin-Wetzel et al., 2017).

An alternative to this model hypothesises that BiP interacts with IRE1 in a non-canonical manner, i.e. via its nucleotide binding domain and independent of the nucleotide bound state of BiP. Evidence for this model is mainly based on *in vitro* binding assays that demonstrate that an interaction between the nucleotide binding domain of BiP (BiP-NBD) and IRE1 is observed that is no longer observed with the isolated substrate binding domain (Carrara et al., 2015). The model then proposes that upon ER stress, BiP interacts with substrates, which triggers its dissociation from IRE1.

In this chapter we will attempt to elucidate the nature of BiP's interaction with IRE1-LD using biophysical techniques and ascertain its effect on the conformational landscape of IRE1-LD that was described in the previous chapter.

5.1 Characterisation of the ATPase cycle of BiP

It has previously been suggested that the functional interaction between IRE1-LD and BiP is via its nucleotide binding domain, and is independent of its nucleotide bound state (Carrara et al., 2015). Despite this, there is conflicting evidence that indicates this interaction is canonical, and that BiP interacts with IRE1-LD via its substrate binding domain (SBD) (Amin-Wetzel et al., 2017). As BiP is a complex system that undergoes conformational changes that are driven by ATP binding and hydrolysis (see section 1.4), we first looked to characterise the ATPase activity of BiP. To do this, we used a malachite green based ATPase assay in the presence of a sodium and potassium based buffer similar to those used in studies that concluded BiP interacts with IRE1 non-canonically and canonically respectively. Any differences in ATPase activity would suggest that the chaperone activity of FL-BiP (WT) is affected by the choice of buffer. Interestingly, the presence of sodium significantly reduced the ATPase rate of BiP by a factor of ~2 (0.237 min^{-1} ±0.042 vs 0.111 min^{-1} ±0.011, p=0.0002) (Figure 5.1), which indicates that sodium affects BiP's canonical interactions.

To investigate the affinity of BiP for ATP and ADP, we performed ITC to determine the affinity of BiP for ATP and ADP. To eliminate any complications emerging from ATP hydrolysis, we used the ATPase deficient mutant, T229G to monitor the binding affinity. Previous indicated that in the presence of potassium BiP has a K_d of 0.80 µM for ATP and 5.73 µM for ADP, indicating that BiP binds ATP preferentially over ADP (Wieteska et al., 2017). Interestingly in the presence of sodium, while the affinity for ADP remains relatively unchanged with a K_d of 3.50 µM ±0.60, a ~ 10x reduction in the affinity for ATP is observed with an observed K_d of 8.33 µM ±0.97 (Figure 5.2). This data is comparable with values previously obtained for isolated BiP-NBD without the interdomain linker, with values of 7.41 µM and 5.27 µM observed for ATP and ADP respectively (Wieteska et al., 2017). This data indicates that the presence of sodium retards the ATPase activity of BiP, thus its ability to act effectively as a chaperone. This finding indicates that a potassium based buffer is necessary to investigate any canonical effect of BiP on IRE1-LD. Interestingly, Carrara et al. have used buffers containing either 75 mM NaCl and 30 mM KCl (for MST experiments) or 75 mM NaCl for pull

down assays. This choice of buffer may have affected their conclusions that the functional interaction between IRE1-LD and BiP is non-canonical.

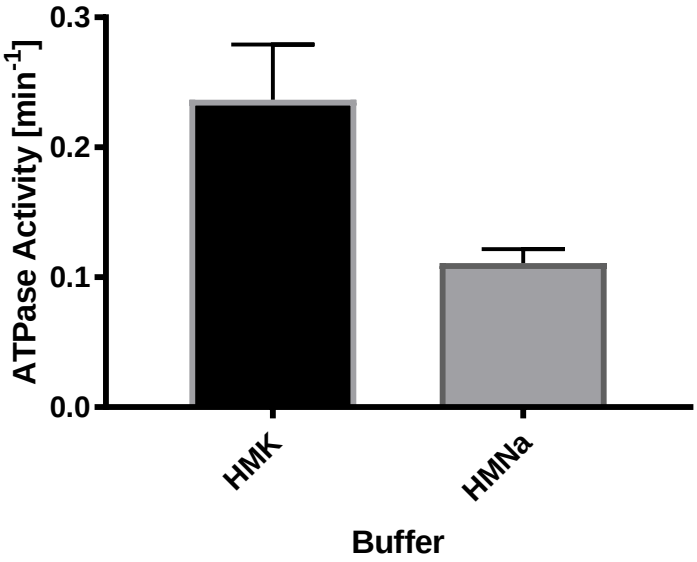

Figure 5.1

ATPase rates of FL-BiP (WT) calculated as an average over 60 minutes. Rates of 0.237 min^{-1} ± 0.042 and 0.111 min^{-1} ± 0.011 were observed in HMK and HMNa buffers respectively (p=0.0002), indicating a significant slow down in the presence of sodium, indicating the chaperoning activity of BiP is significantly reduced in a sodium based buffer.

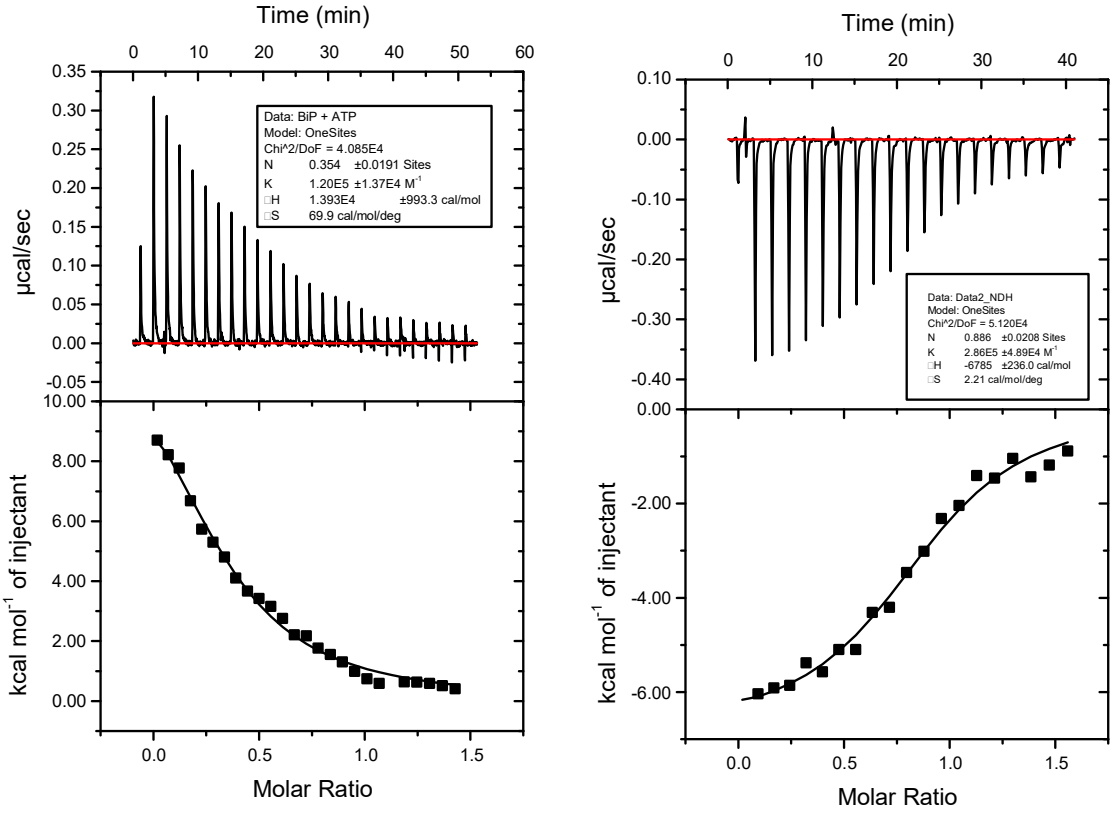

Figure 5.2

ITC thermogram of FL-BiP (T229G) in a sodium based buffer upon addition of (A) ATP and (B) ADP. A Kd of 8.33 µM ± 0.97 and 3.50 µM ± 0.60 were calculated respectively from n=3, compared with values of 0.80 µM ± 0.30 and 5.70 µM ± 1.05 in HMK buffer previously obtained in our lab, indicating that a significant reduction in affinity for ATP is observed in a sodium based-buffer**.**

5.2 BiP T229G has no effect on the substrate independent IRE1-LD dimer

We next ascertained if BiP renders IRE1-LD monomeric by forming a stable complex with IRE1-LD and preventing formation of dimeric IRE1-LD. In the previous chapter we performed detailed characterisation of the process of dimer formation of IRE1-LD, thus we were able to use similar methodology to ascertain the oligomeric state of IRE1-LD upon the addition of BiP. To isolate BiP's high-substrate affinity state, we used a well characterised, ATPase deficient mutant of FL-BiP, T229G and added ADP, which has previously been shown to be able to bind to substrates (Wieteska et al., 2017). We then injected a mixture of 30 µM FL-BiP (T229G), 5 mM ADP and 15 or 30 µM FL-IRE1-LD and monitored the position of the FL-IRE1-LD peak as a reporter of the dimerization state of IRE1. If BiP were to alter the oligomeric state of IRE1-LD, we would expect the retention volume of the IRE1-LD to be altered, and observe a new peak corresponding to a complex between the two proteins, however no significant change was observed in the retention volume of FL-IRE1-LD, when compared to the retention volume in the absence of BiP (Figure 5.3A), suggesting that IRE1-LD does not form a complex with the high substrate affinity state of BiP (Figure 5.3B).

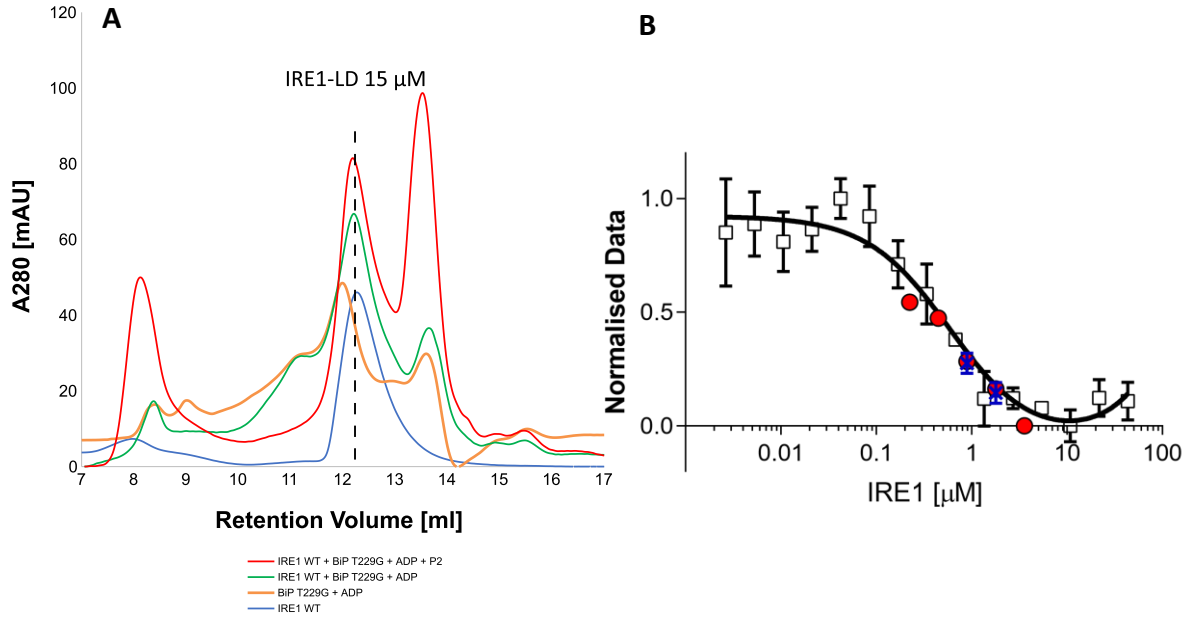

Figure 5.3

(A) SEC Chromatogram of 15µM FL-IRE1-LD (blue) overlaid with FL-IRE1-LD + BiP T229G + ADP (green) and FL-IRE1-LD + FL-BiP (T229G) + ADP + a substrate for BiP, P2 (red). The peak for IRE1-LD does not shift in the presence of BiP in its high substrate affinity state, indicating it does not affect the dimeric state of IRE1-LD. BiP T229G + ADP in the absence of IRE1-LD is shown in orange. (B) Normalised dimerization data of IRE1-LD, as described in 1.1 and shown in Figure 4.3, with additional normalised SEC data in the presence of FL-BiP T229G + ADP plotted as blue crosses. Data plotted is an average of n=2 and is from two injection concentrations of FL-IRE1-LD, 15 µM and 30 µM.

5.3 The BiP NBD retards the rate of disulphide bond formation of IRE1-LD

As it has been suggested that IRE1-LD interacts with BiP-NBD in a non-canonical manner, i.e. independent of the nucleotide bound state of BiP and in a non-chaperone like manner, (Carrara et al., 2015), we elucidated the mechanistic details of this interaction and its effect on IRE1-LD. To confirm the previously reported literature data, we initially performed MST by titrating in varying concentrations of BiP-NBD to FITC labelled IRE1-LD. This indicated that

IRE1-LD and BiP-NBD interact with a K_d = 0.38 µM ± 0.16, confirming the observation that an interaction between these two proteins is observable by MST (Figure 5.4A).

To ascertain that that IRE1-LD still interacted with BiP-NBD in the absence of an affinity tag, we performed the same experiment on IRE1-LD that had been subject to TEV cleavage. This indicated that IRE1-LD still binds to BiP-NBD in the absence of the affinity tag, with a similar affinity (0.38 µM ± 0.16 vs 0.19 ± 0.02). (Figure 5.4C)

As MST indicated that the interaction between IRE1-LD and BiP-NBD has a sub-µM affinity, if a complex is formed between these two proteins, we would expect to be able to observe and isolate it by SEC. Surprisingly, upon injection of a mixture of 40 µM IRE1-LD and 80 µM BiP-NBD, thus at concentrations well in excess of the observed dissociation constant, both proteins eluted separately from the column and a peak corresponding to a complex was not observed indicating that a long-lived complex is not formed between these two compounds (Figure 5.4B).

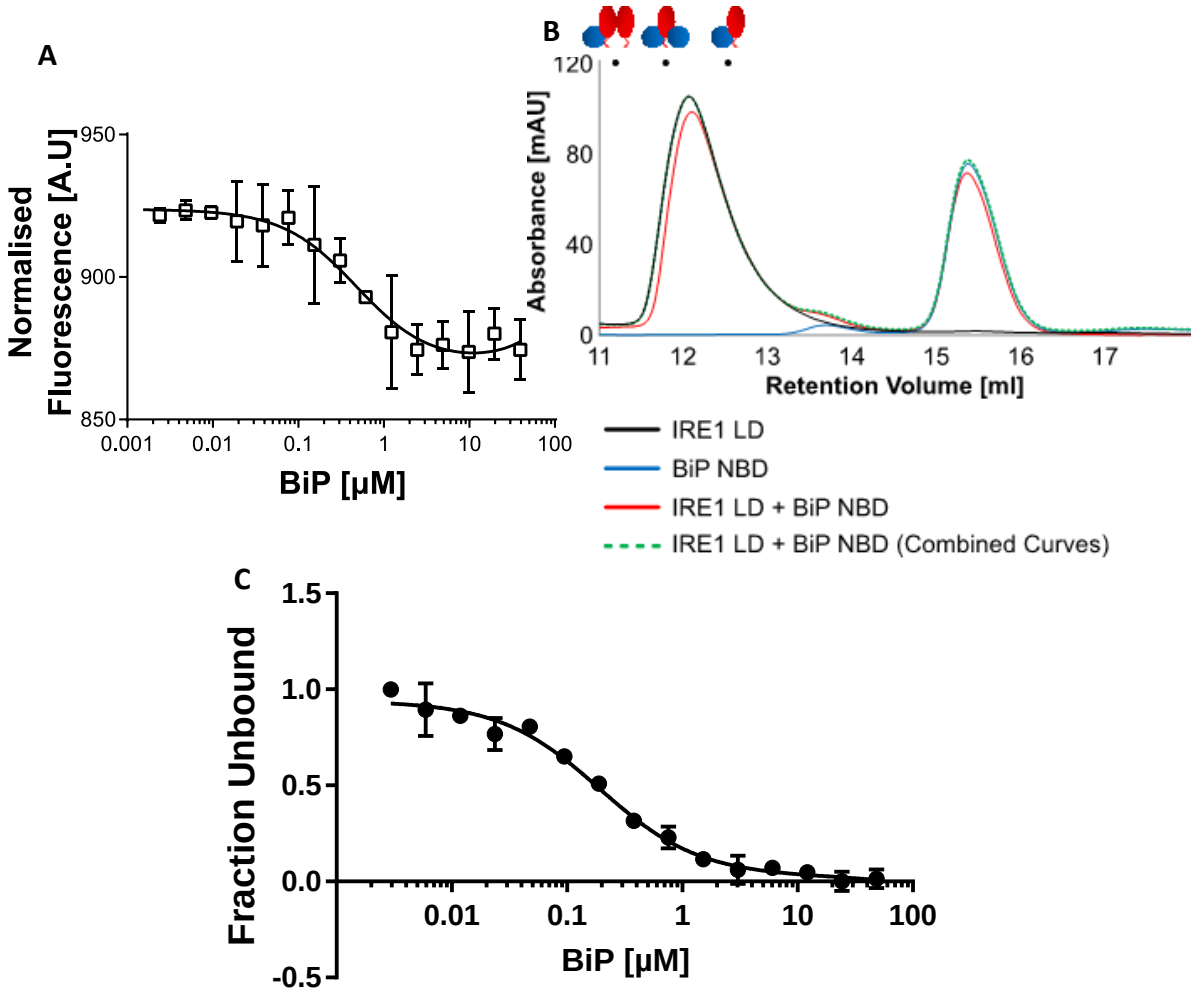

Figure 5.4

(A) Isotherm from MST measurements of FITC-labelled FL-IRE1-LD in the presence of 5 mM DTT. A K_d of 0.38 µM ± 0.16 (n=3) was measured for the sample in a reducing buffer. (B) SEC chromatograms of 40 µM FL-IRE1-LD (black), 80 µM BiP-NBD (blue) and a combination of 40 µM FL-IRE1-LD and 80 µM BiP NBD (red). The combined chromatograms of FL-IRE1-LD and BiP-NBD (green) indicates there is no significant change between the chromatogram of FL-IRE1-LD and BiP-NBD. The expected peak positions of monomeric FL-IRE1-LD (red protomer) and one BiP-NBD protomer (blue protomer), monomeric FL-IRE1-LD and two BiP-NBD protomers and dimeric FL-IRE1-LD and one BiP-NBD protomer are shown above the chromatogram. (C) Fraction of BiP-NBD bound to IRE1-LD in the absence of an affinity tag, as determined by MST. A K_d of 0.19 ± 0.02 (n=2) indicating that the presence of an affinity tag does not have a significant effect on this interaction.

As SEC indicated a long-lived complex is not formed between IRE1-LD and BiP-NBD, we decided to use NMR to obtain residue specific information about the interaction between these two proteins. If an interaction were to occur between these two proteins, it would be expected that this would be apparent through chemical shift perturbations close to the site of interaction (due to changes in chemical environment) and line broadening (due to the increased rotational correlation time of a complex and chemical exchange between bound and unbound states). As our FL-IRE1-LD construct is 49 kDa as a monomer and BiP NBD is 44 kDa, we would expect a complex between these two compounds to be at least 93 kDa, assuming 1:1 stoichiometry. We therefore performed a ^{13}C-^{1}H SOFAST-HMQC (Schanda and Brutscher, 2005) of U{^{15}N, ^{12}C, ^{2}H}, selectively $^{13}CH_3$-Ile$^{\delta1}$, $^{13}CH_3$-Ala$^{\beta}$, $^{13}CH_3$-Leu$^{\delta}$, $^{13}CH_3$-Val$^{\gamma}$ labelled BiP-NBD in the presence and absence of IRE1-LD. Upon comparing the spectra of 25 µM BiP-NBD in the presence and absence of 25 µM FL-IRE1-LD, it was apparent that no significant changes were observed (Figure 5.6A).

As ^{13}C-^{1}H SOFAST-HMQC of selective methyl labelled only provides a limited number of probes when compared to conventional ^{15}N-^{1}H TROSY methods, we hypothesised that the interaction between IRE1-LD and BiP-NBD only produces very local changes in the BiP-NBD structure. To ascertain this, we labelled FL-IRE1-LD with (1-oxyl-2,2,5,5-tetramethylpyrroline-3- methyl)methanethiosulfonate (MTSL) (Berliner et al., 1982), a thiol reactive, nitroxide based spin-label. Any peak corresponding to a residue within 25 Å of the spin label would be expected to be broadened, due to the increased relaxation properties of the paramagnetic species. The three cysteine residues of IRE1-LD were labelled successfully, as confirmed by mass spectrometry, which indicated a mass increase of 551 Da, when comparing MTSL labelled sample to unlabelled FL-IRE1-LD (Figure 5.7). Again, when comparing spectra of 25 µM BiP-NBD in the presence and absence of 25 µM MTSL labelled FL-IRE1-LD, we were unable to observe any significant reduction in peak intensity, thus indicating that the absence of any chemical shift perturbations and line broadening is not due to local changes at the interaction site (Figure 5.6B).

Finally, as FITC-labelled IRE1-LD is only present at 0.5 µM in MST experiments, we questioned whether BiP-NBD is only able to interact with the monomeric species of IRE1-LD. To do this, we performed a ^{13}C-^{1}H SOFAST-HMQC on 1 µM U{^{15}N, ^{12}C, ^{2}H}, selectively $^{13}CH_3$-Ile$^{\delta1}$, $^{13}CH_3$-Ala$^{\beta}$, $^{13}CH_3$-Leu$^{\delta}$, $^{13}CH_3$-Val$^{\gamma}$ labelled BiP-NBD in the absence and presence of and absence of

MTSL labelled FL-IRE1-LD. Several peaks were still observable at this low concentration, however no obvious line broadening was observed, indicating that the absence of monomeric IRE1-LD was not hindering the formation of a complex between IRE1-LD and BiP (Figure 5.6C).

As a control, to confirm the presence of spin label was not blocking the IRE1-LD BiP-NBD interaction site, we performed MST on FITC and MTSL labelled FL-IRE1-LD. This indicated that the presence of the MTSL label did not prevent the binding of BiP-NBD to IRE1-LD, and, did not produce a significant effect (0.38 µM ±0.16 vs 1.70 µM ± 0.87, p=0.2099) (Figure 5.5).

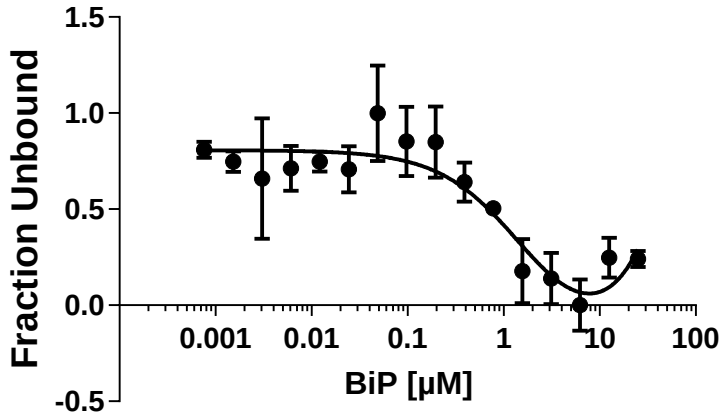

Figure 5.5

Fraction of BiP-NBD bound to FL-IRE1-LD in the presence of a MTSL, as determined by MST. A K_d of 1.70 ± 0.87 (n=3) indicates that BiP-NBD still interacts with MTSL labelled FL-IRE1-LD.

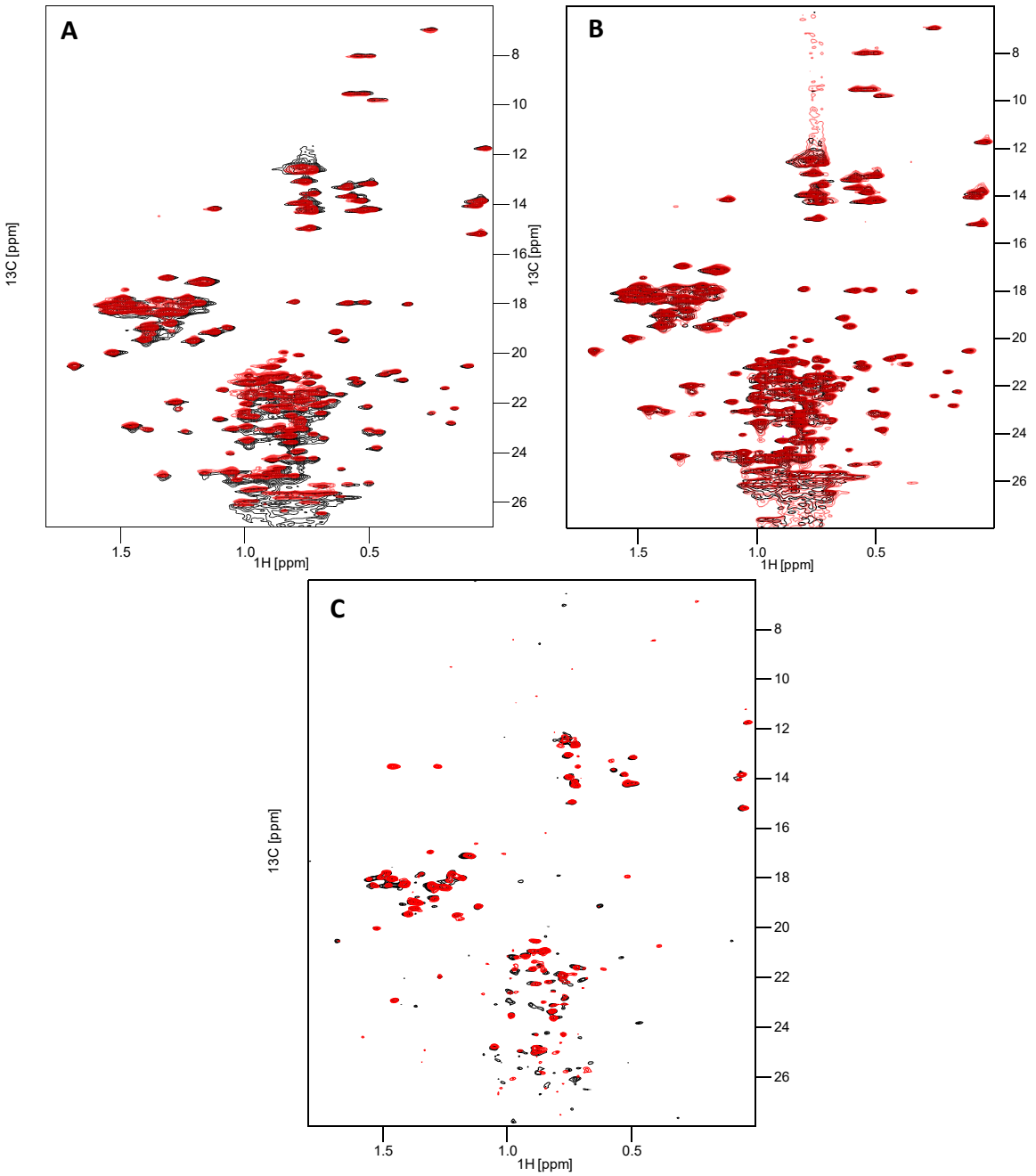

Figure 5.6

$^{13}C^{1}H$-SOFAST-HMQC of U{^{15}N, ^{12}C, ^{2}H}, selectively $^{13}CH_3$-Ile$^{\delta1}$, $^{13}CH_3$-Ala$^{\beta}$, $^{13}CH_3$-Leu$^{\delta}$, $^{13}CH_3$-Val$^{\gamma}$ BiP-NBD at 25 µM (A and B) and 1 µM (C) in the absence (black) and presence (red) of an equimolar concentration of unlabelled IRE1-LD (A) or MTSL labelled IRE1-LD (B and C). No significant chemical shift perturbations or reduction in peak intensity is observed in any spectra, indicating a complex is not formed between BiP-NBD and IRE1-LD. Experiments performed in HMK buffer + 5 mM DTT

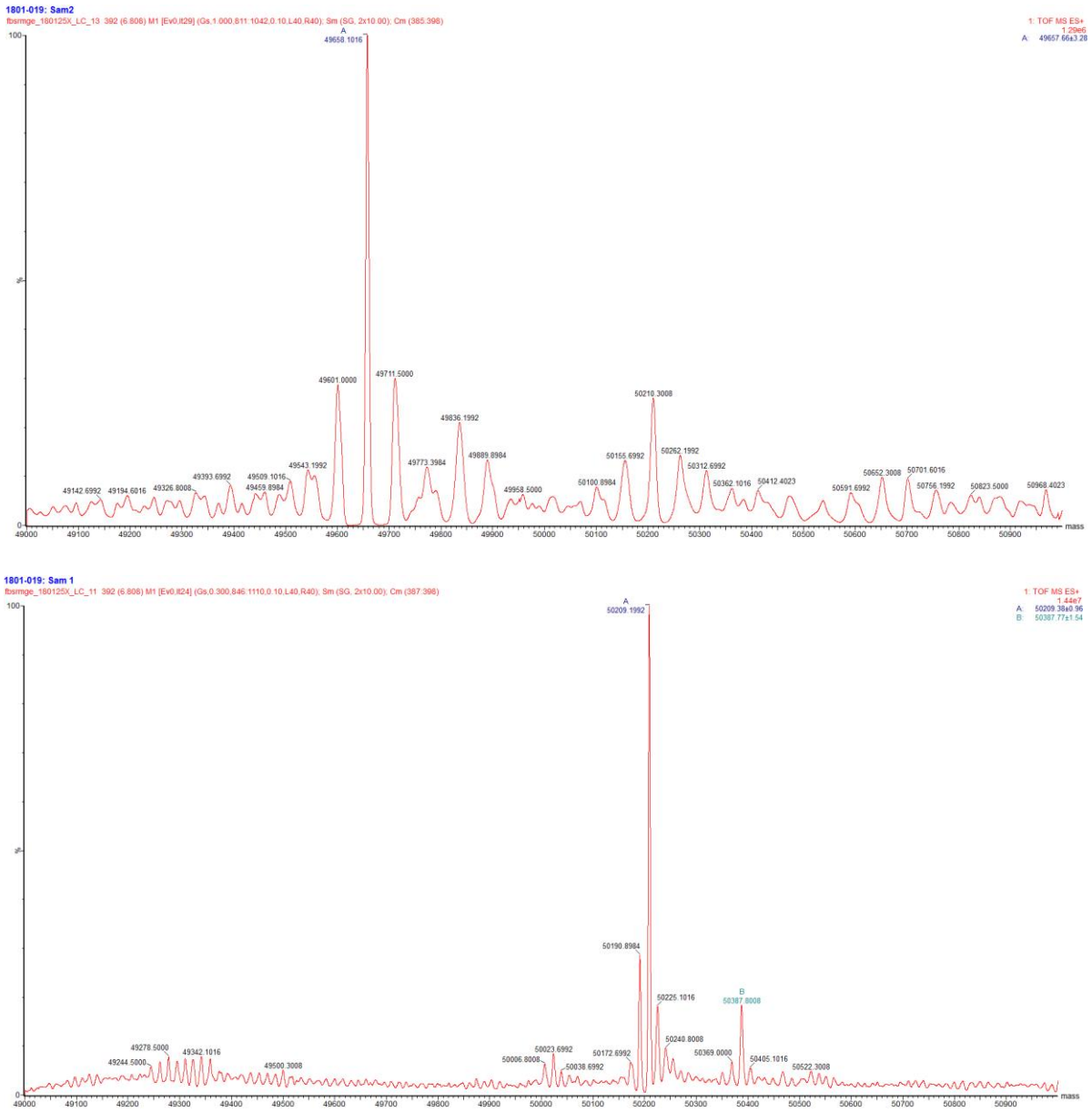

Figure 5.7

Mass spectra of FL-IRE1-LD prior to labelling with MTSL (above) and following labelling with MTSL (below). A mass shift of 551.72 ±4.24 is observed between unlabelled and labelled FL-IRE1-LD, indicating three MTSL molecules have bound to IRE1-LD.

As BiP-NBD did not apparently affect the conformational landscape of the non-covalent interactions of IRE1-LD we previously characterised, we next questioned if BiP-NBD influences conformational landscape of the covalent interactions of IRE1-LD. In the previous chapter we determined that this process was slow, and quantified it using reducing and non-reducing

SDS-PAGE. If BiP-NBD were to affect this process, we would expect to observe a significant change in the amount of monomeric IRE1-LD present in the non-reduced gel samples taken. Upon addition of an equimolar (15 μM) amount of BiP-NBD to FL-IRE1-LD (Figure 5.8C), the rate of disulphide bond formation was significantly reduced (Figure 5.8A), with our data indicating that after 24 hours, the percentage of monomeric IRE1-LD in the presence of BiP-NBD was 66.51% ± 9.68 compared to 32.75% ± 1.12 in the absence. This was comparable with our data for C148 and C332 mutants (65.85% ± 18.67 and 68.74% ± 4.22 respectively) (Figure 4.20), indicating that BiP-NBD is important for retarding the rate of disulphide bond formation.

We then questioned whether disulphide formation of IRE1-LD prevents BiP-NBD binding. To do this we performed MST on FITC-labelled FL-IRE1-LD that was incubated for 24 hours in a non-reducing buffer. The interaction we previously observed between IRE1-LD and BiP-NBD was no longer observable (Figure 5.8B) indicating that the process of disulphide bond formation prevents an interaction between BiP-NBD and IRE1-LD.

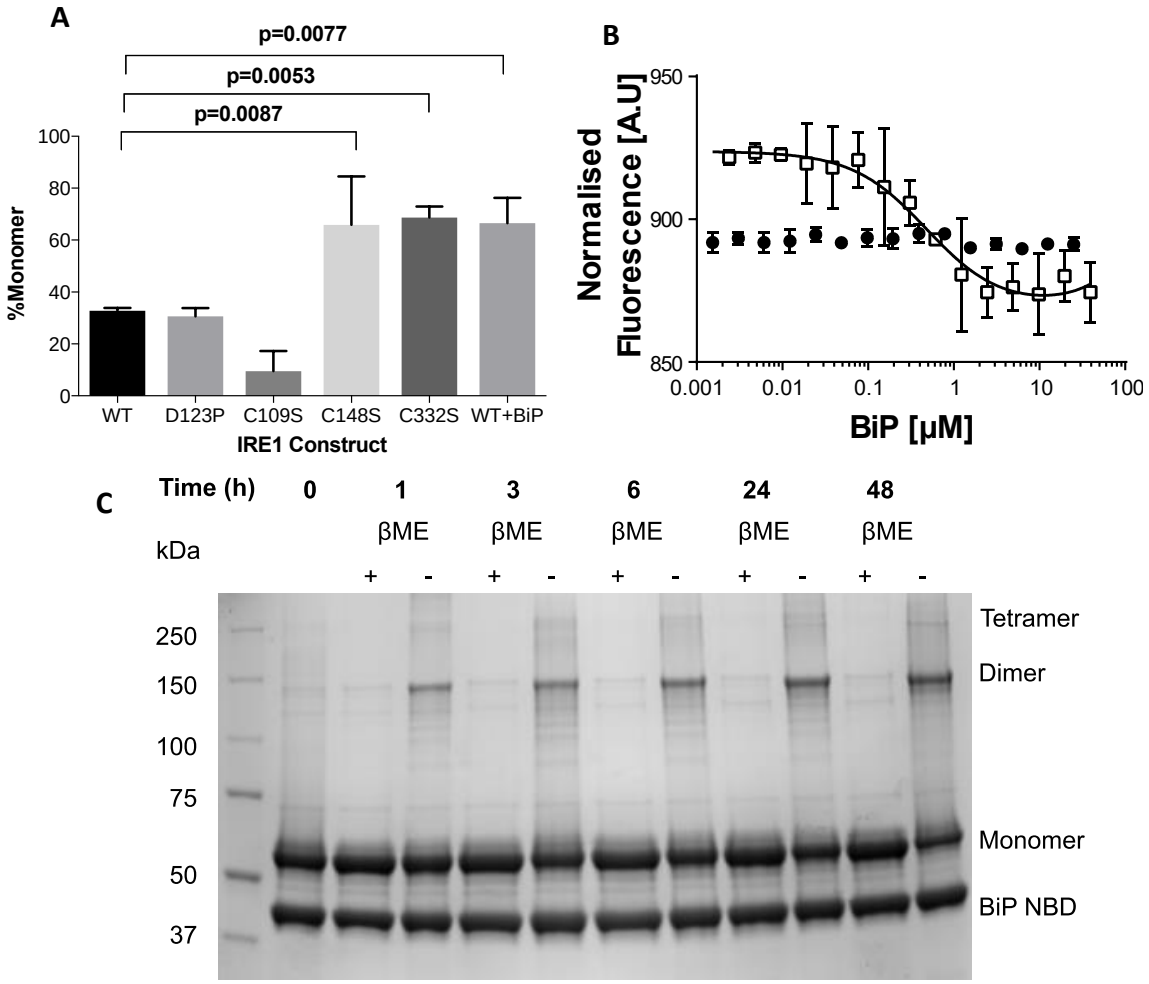

Figure 5.8

(A) The percentage of monomeric IRE1-LD calculated from reducing and non-reducing SDS-PAGE gels after 24 hours. The presence of BiP-NBD retards the rate of disulphide bond formation to a level similar to that observed in C148 and C332 mutants. (B) Isotherm of IRE1-LD FITC in a reducing buffer (white squares) and 24 hours after incubation in a non-reducing buffer (black circles), indicates that disulphide bond formation of IRE1-LD prevents the interaction with BiP-NBD. Isotherm is an average of three repeats. (C) SDS-PAGE of IRE1-LD in the presence of an equimolar concentration of BiP-NBD in a reducing and non-reducing buffer at timepoints relative to buffer exchange from a reducing buffer. Analysis is shown in (A)

5.4 The ATPase dependent de-oligomerisation of IRE1-LD by BiP

As we have established that IRE1-LD does not form a stable complex with BiP and does not appear to prevent dimerization, we asked whether BiP has an effect on the formation of IRE1-LD oligomers. In the previous chapter we characterised the formation of soluble and insoluble oligomers of IRE1 cLD upon addition of ΔEspP. As it has been demonstrated that BiP interacts in a canonical manner with the core luminal domain of IRE1, i.e. not the juxtamembrane linker (Amin-Wetzel et al., 2017), we used our IRE1 cLD construct. To establish whether BiP was able to de-oligomerise IRE1 cLD, we performed the same assay in the presence of various BiP constructs (Figure 5.9).

Upon addition of FL-BiP (WT) and ATP, a significant reduction in the OD400 was observed, suggesting that BiP can de-oligomerise large insoluble IRE1 cLD particulates. To ascertain if this interaction was canonical, we performed the same assay with FL-BiP (WT) (apo), the well-characterised ATPase deficient mutant, T229G (ATP), the substrate binding deficient mutant, V461F (ATP), and isolated BiP-NBD (ATP). While a reduction in the OD400 was observed for the T229G mutant (0.431±0.247, 0.259 ±0.121 for IRE1 cLD + ΔEspP in the absence and presence of BiP T229G + ATP), the only statistically significant reduction observed upon addition of FL-BiP (WT) + ATP (0.168 ±0.026, p=0.0038).

To gain insight into the size of the soluble particulates present, we performed DLS (Figure 5.9). As a characteristic large oligomeric peak is observed upon addition of ΔEspP to IRE1-LD, we were able to monitor to see if various BiP constructs reduced the size of this oligomeric species.

The addition of 10 μM FL-BiP (WT) in the presence of 5 mM ATP reduced the particle sizes observed in the presence of IRE1 cLD and ΔEspP to 5.08 nm ±2.04 and 39.1 nm ±21.26, compared to 6.08 nm ±2.02 and 404.90 nm ±268.17 in the absence of FL-BiP (WT) (Figure 5.10C).

Furthermore, the particle size in the presence of 10 μM FL-BiP (T229G) and 5mM ATP was 4.87 nm ±1.29 and 62.61 nm ±30.19 (Figure 5.10D), 10 μM FL-BiP (WT) apo was 1.43 nm ±0.19, 11.49 nm ±5.58 and 146.52 nm ±83.76 (Figure 5.10E), with a larger intensity fraction emanating from the largest peak for both of these conditions. The OD400 and DLS data indicates that the ATPase activity of BiP is necessary to fully de-oligomerise IRE1 cLD.

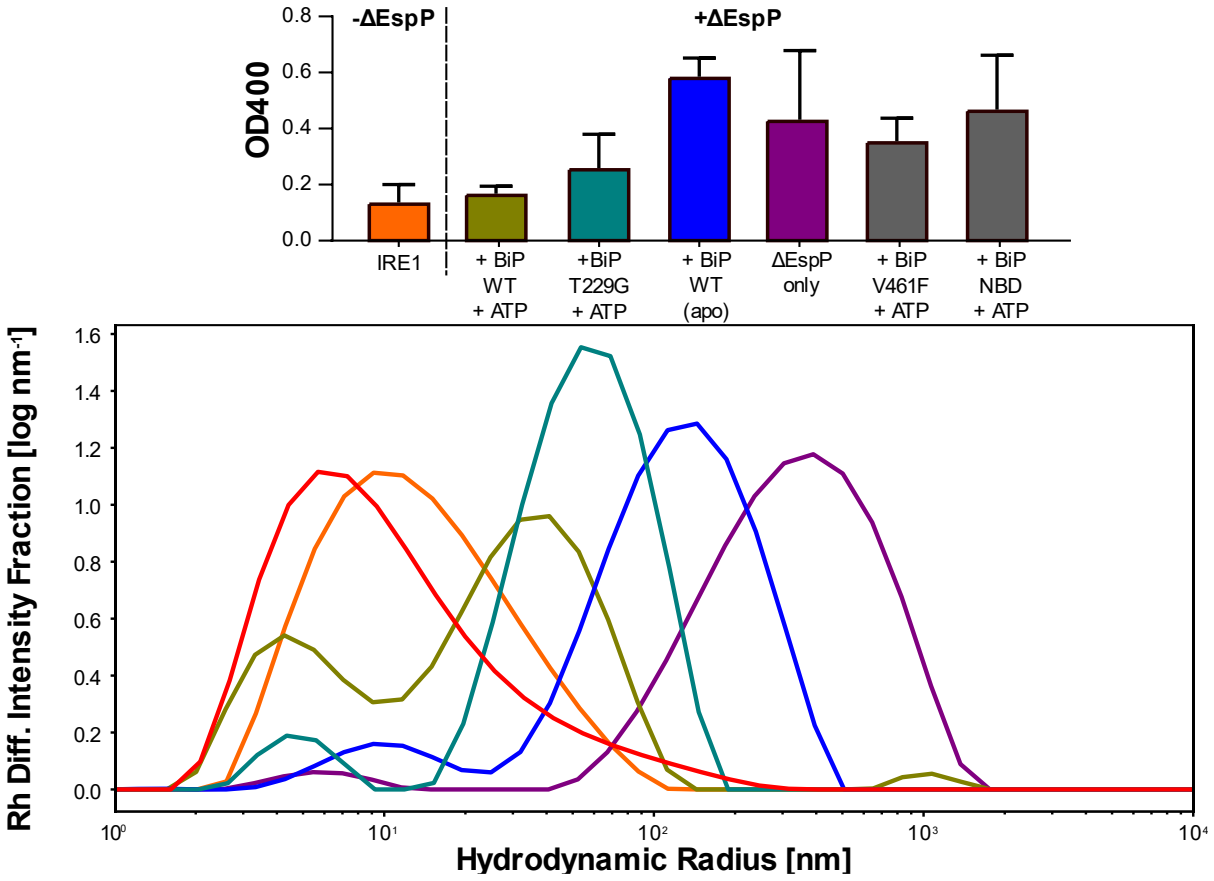

Figure 5.9

OD400 data of 20 µM IRE1 cLD in the absence (orange) and presence of 40 µM ΔEspP (various). Various BiP constructs were at a concentration of 40 µM in solution and ATP was added at a concentration of 10 mM. Overlay of DLS regularisation plots. Colours correspond to those above, with an additional plot representative of IRE1-LD (D123P) (red). IRE1 constructs were added at 5µM, BiP constructs were added at 10 µM, ΔEspP was added at 10 µM and ATP was added at 10 mM. Size of species is shown in Table 5.1.

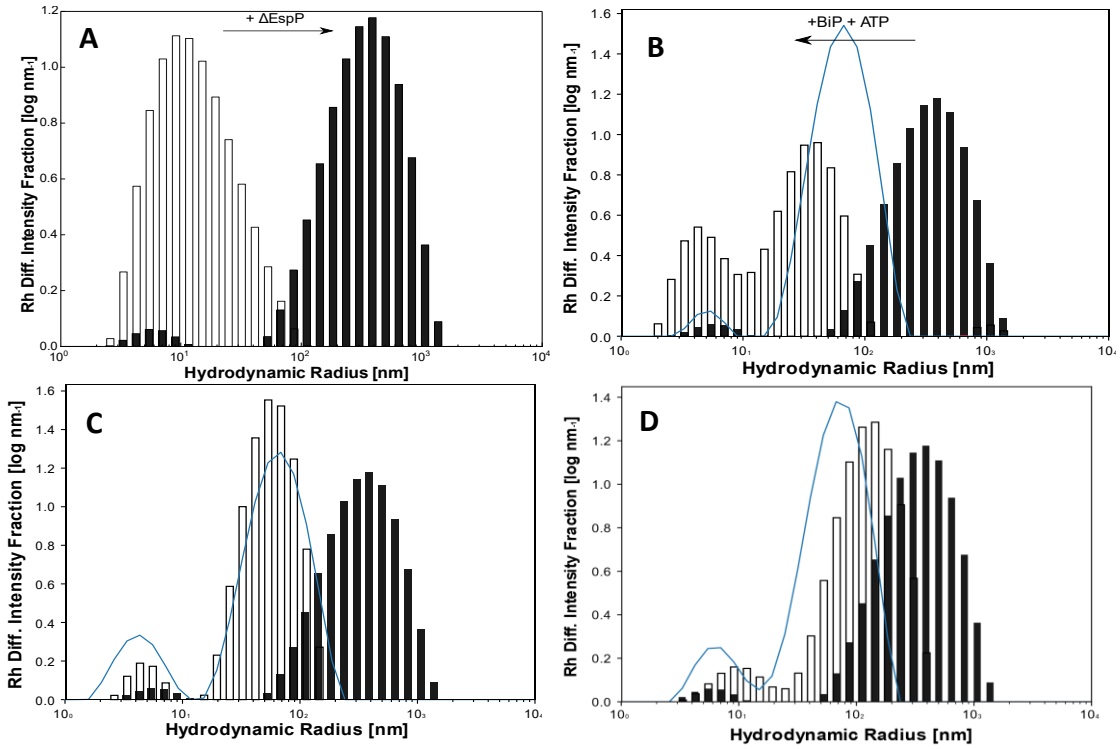

Figure 5.10

DLS regularisation data of (A) 5 µM IRE1-LD in the absence (white) and presence (black) of 10 µM ΔEspP (B) 5 µM IRE1-LD and 10 µM ΔEspP in the absence (black) and presence (white) of 10 µM FL-BiP (WT) and 10 mM ATP. The regularisation plot of FL-BiP (WT) and 10 mM ATP is represented as a blue line. (C) As for (B), with FL-BiP (T229G) instead of (WT). (D) As for (B), in the absence of any nucleotide.

		IRE1 D123P	IRE1 cLD	IRE1 + BiP + ATP	IRE1 + BiP T229G	IRE1 + BiP apo	IRE1 + EspP
Rh (nm)	*Peak 1 Average*	15.55	17.3	5.08	4.87	1.43	6.08
	Peak 1 Maxima	6.46	9.13	4.27	4.35	1.57	5.45
	Peak 1 Std Dev	19.31	14.9	2.04	1.29	0.19	2.02
	Peak 2 Average			39.1	62.61	11.49	404.9
	Peak 2 Maxima			40.94	53.67	9.11	390.13
	Peak 2 Std Dev			21.26	30.19	5.58	268.17
	Peak 3 Average			1046.77		146.52	
	Peak 3 Maxima			1073.56		144.53	
	Peak 3 Std Dev			206.63		83.76	

Table 5.1

Table of calculated Rh values calculated by DLS for varying conditions. The average particle size, peak maxima and standard deviation of each peak has been calculated from each intensity bin.

As we have previously obtained a K_d for IRE1 and BiP-NBD, by MST, we then questioned whether FL-BiP (WT) in the presence of 5 mM ATP affects the observed K_d. Consistent with previously obtained observations (Carrara et al., 2015), the presence of FL-BiP (WT) and 5 mM ATP did not significantly change the binding affinity observed by MST (0.37 µM ± 0.07), which may be due to the lack of IRE1-LD oligomers at this concentration.

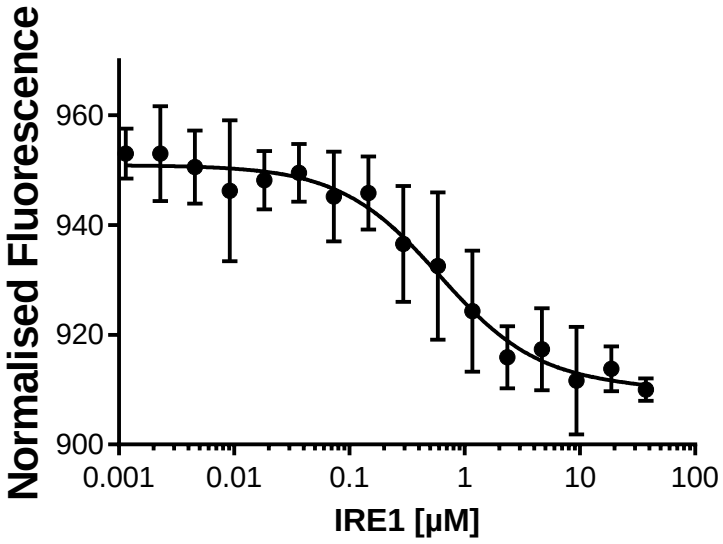

Figure 5.11

MST data of FITC labelled FL-IRE1-LD upon addition of FL BiP (WT) and 5 mM ATP (n=3). A K_d of 0.37 µM ± 0.07 indicates that the observed affinity is not significantly different from that obtained for BiP-NBD (0.38 µM ± 0.16).

To obtain structural information about this process, we performed a ^{15}N-^{1}H BEST-TROSY. As we obtained better spectra quality for our FL-IRE1-LD construct, we acquired several spectra using this construct in the presence of ΔEspP, ΔEspP and FL-BiP (WT) and ΔEspP, FL-BiP (WT) and ATP (Figure 5.12) and compared the peak intensities of 34 isolated peaks to the spectrum of FL-IRE1-LD (Figure 5.13). The presence of FL-BiP (WT) and ΔEspP increased the average peak intensity ratio from 0.812 ±0.144 to 1.170 ±0.255 (p=<0.0001), however no significant chemical shift perturbations were observed upon comparing both sets of spectra. This is in good agreement with the DLS data, as an increase in the peak height is indicative of a decrease

in the rotational correlation time and/or a change in the µs-ms dynamics, Therefore, when considering the DLS data, it is likely that a reduction in the size of the oligomeric species of IRE1-LD is observed upon addition of FL-BiP (apo).

Upon addition of FL-BiP (WT), ΔEspP and ATP, a further increase in the average peak intensity ratio compared to that of FL-IRE1-LD was observed (1.357 ±0.435). This increase is statistically significant when compared to apo FL-BiP (WT) (p=0.0110), which is in good agreement the DLS data, which indicates a decrease in the particle size upon addition of ATP. Interestingly a larger degree of variability in the intensity was observed, suggesting that the presence of FL-BiP (WT) is affecting µs-ms dynamics of certain residues more than others, in particular peaks 21 and 26, whose intensity ratio increased from 1.34 and 1.27 to 2.55 and 2.24 when compared to apo FL-BiP (WT).

In addition, minor chemical shift perturbations were observed, the greatest of which was 0.048 at peak 20 (Figure 5.12). Taking all this data into account, we conclude that while the ATPase deficient and apo forms of BiP do influence the size of large IRE1-LD particles, the chaperoning activity of BiP, driven by its ATPase activity, is necessary to fully reduce IRE1-LD to a similar particle size observed in monomeric/dimeric protein. This effect is also observable in the ^{15}N-^{1}H TROSY spectra of IRE1-LD, in the presence of ΔEspP and BiP, however a greater degree of variability in the peak intensity ratio upon addition of ATP, suggesting that the chaperone activity of BiP has an effect on the µs-ms dynamics of certain residues in IRE1-LD as well.

As literature data indicates IRE1-LD forms oligomers at high concentrations in the absence of substrate (Karagoz et al., 2017), we questioned whether we would be able to observe a similar effect in the ^{15}N-^{1}H TROSY spectrum of FL-IRE1-LD without addition of ΔEspP. The average peak intensity of 12 isolated peaks in the spectra of FL-BiP (apo) and FL-BiP (ATP) increased in a similar manner to that in the presence of ΔEspP (1.282 ±0.193 vs 1.708 ± 1.117 respectively for (I/I(IRE1)) (Figure 5.15).

As BiP is hypothesised to monomerise IRE1-LD, we then compared the spectra of IRE1 in the presence of FL-BiP (+/- ATP) and the non-dimerising mutant, IRE1-LD (D123P). While we did not observe chemical shift perturbations back towards positions observed in the spectrum of IRE1-LD (D123P) (Figure 5.14), comparison of the peak height of the same 12 isolated peaks,

which overlapped in all spectra, revealed that significant increases in peak height (I/I(IRE1)>2) were observable in identical peaks in the spectra of IRE1-LD (D123P) and FL-IRE1-LD + FL-BiP (ATP) (Figure 5.15). This suggests that active BiP affects the µs-ms dynamics of these residues in a similar manner to that observed upon monomerisation of IRE1-LD, although the lack of chemical shift perturbations back towards those observed in the IRE1-LD (D123P) spectrum suggests that active BiP does not render IRE1-LD monomeric. This maybe different in the presence of ERdj4, as this would increase both the ATPase rate of BiP and its affinity for IRE1-LD (Amin-Wetzel et al., 2017).

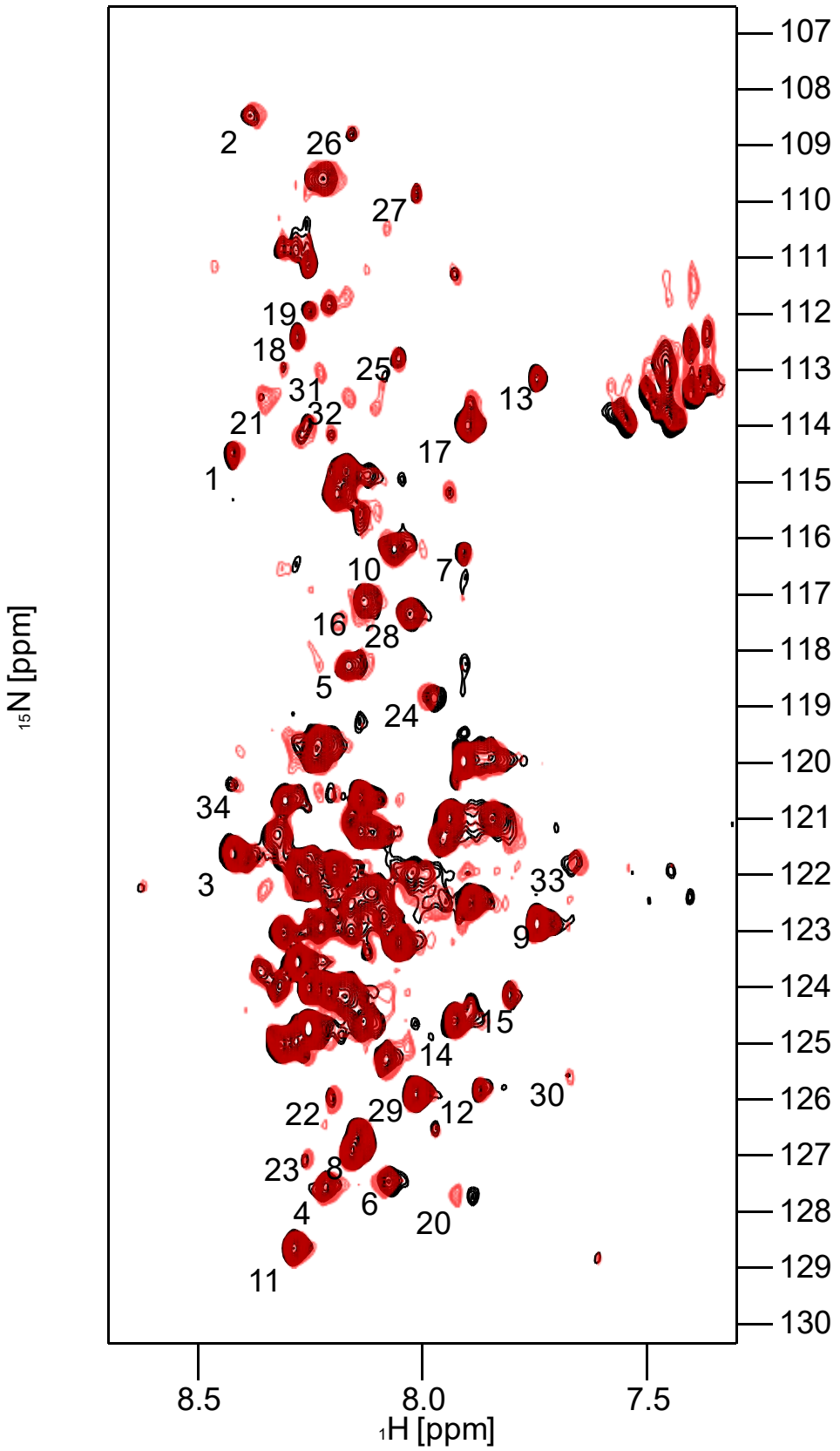

Figure 5.12

^{15}N-^{1}H TROSY of 100 µM IRE1-LD in the absence (black) and presence (red) of 100 µM ΔEspP, 100 µM FL-BiP WT and 40 mM ATP. Peak numbers used for analysis are arbitrary and are to the bottom left of each peak. Spectra recorded at 298K in HMK + 1 mM TCEP

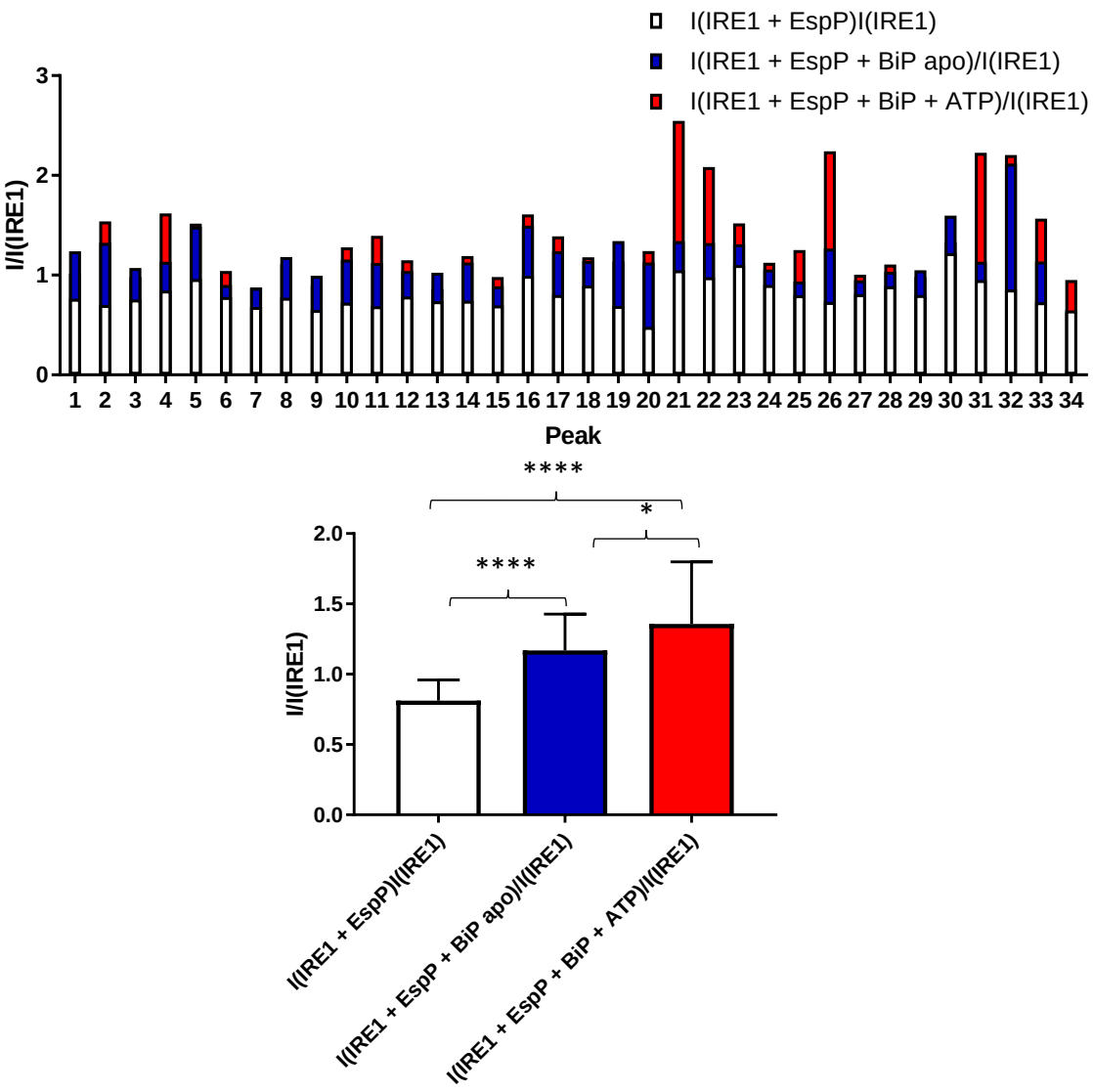

Figure 5.13

(Above) Peak intensity ratio compared to IRE1-LD using peak numbers in Figure 5.11. (Below) Average peak intensity ratio of all 34 peaks under specified conditions. Error bars indicate the standard deviation. One way ANOVA followed by Tukey's multiple comparison indicates significant differences between all data sets (**** - adjusted p value <0.0001, * adjusted p value = 0.0110)

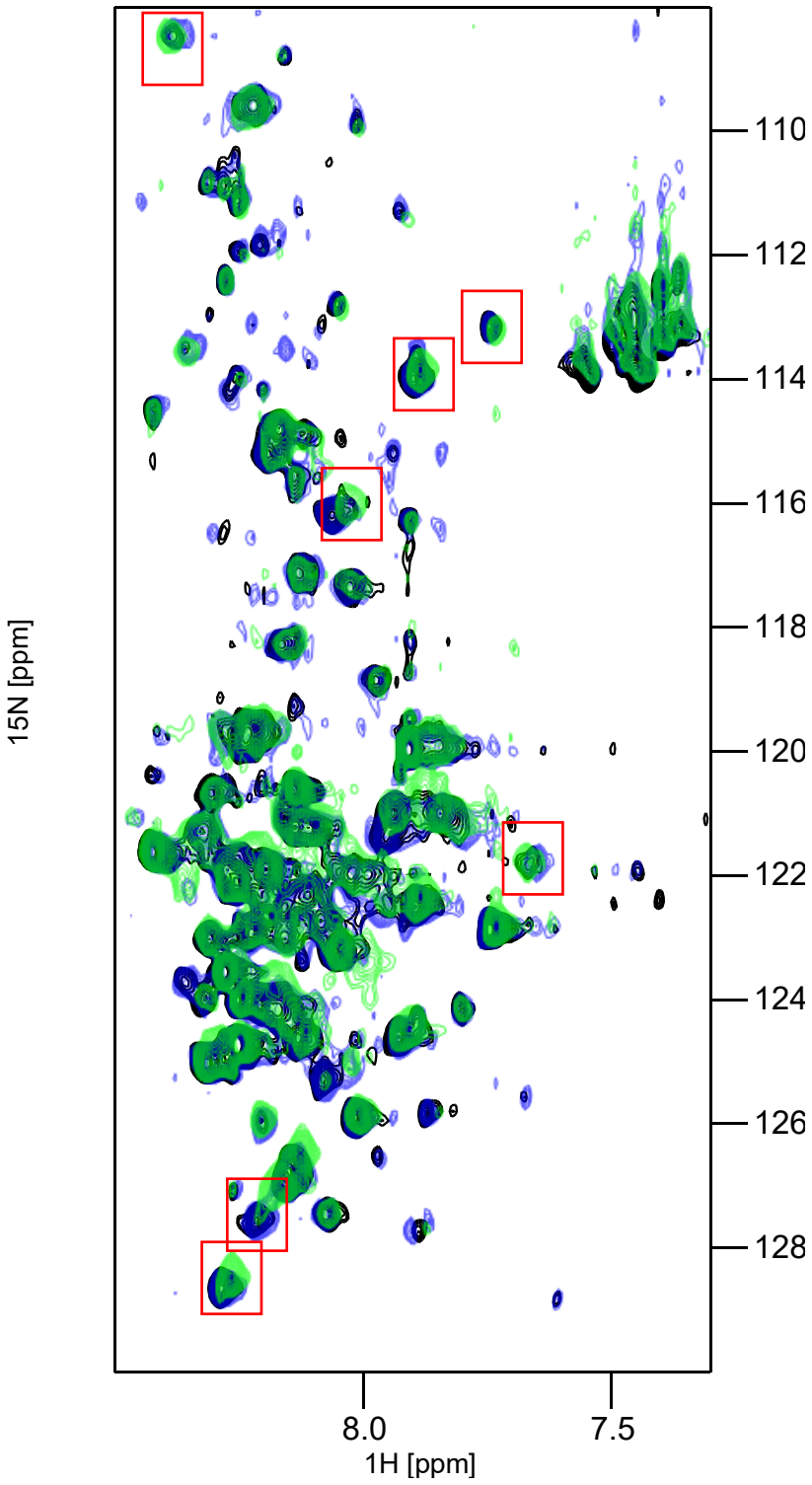

Figure 5.14

^{15}N-^{1}H TROSY of 100 µM FL-IRE1-LD in the absence (black) and presence (blue) of 100 µM FL-BiP WT and 40 mM ATP, overlaid with the spectrum of 100 µM IRE1-LD (D123P) (green) Peak changes between WT and D123P spectra are highlighted in red boxes. Spectra recorded at 298K in HMK + 1 mM TCEP

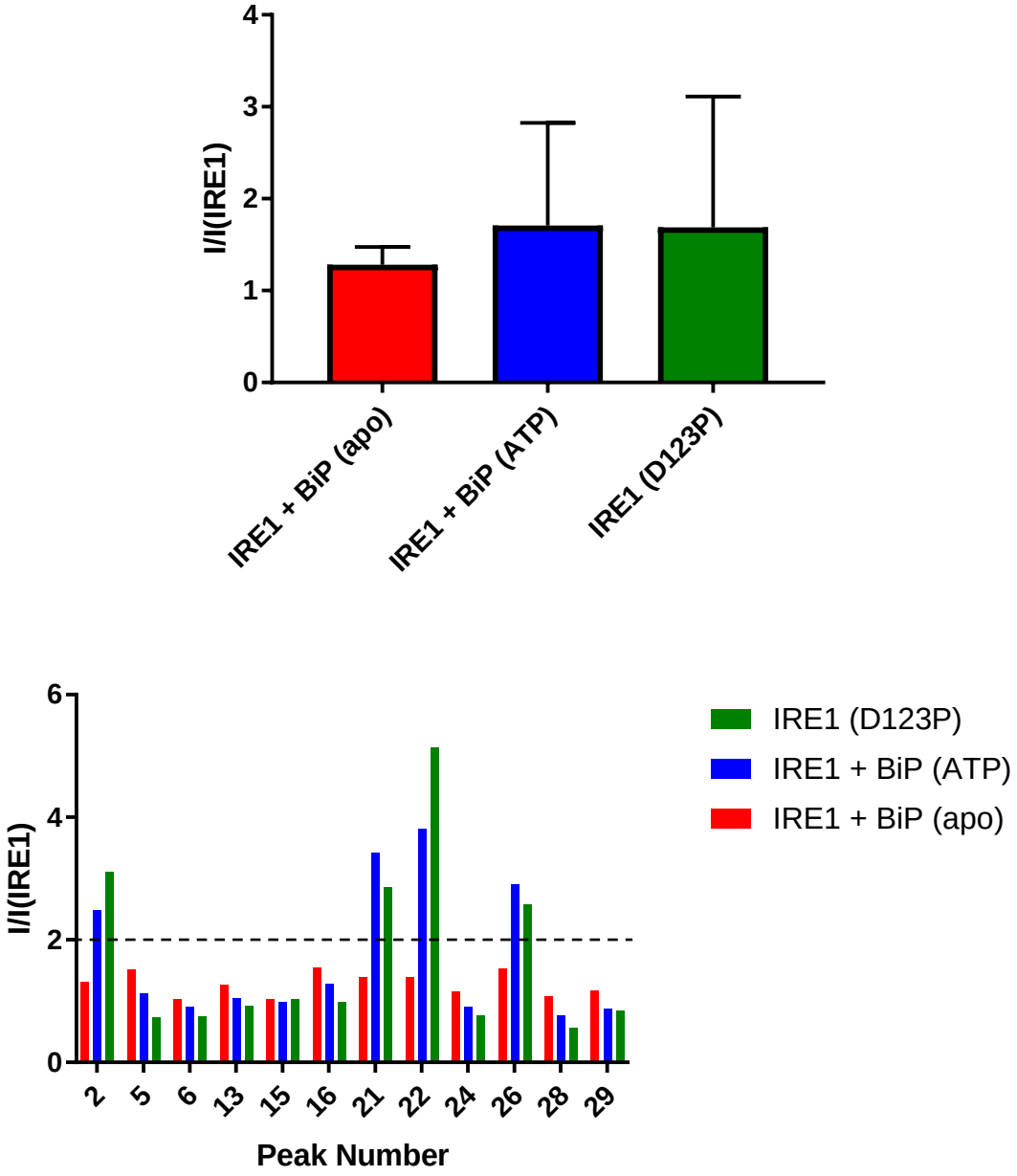

Figure 5.15

(Above) The average relative peak intensity of 12 isolated peaks for the specified spectra. Error bars represent the standard deviation. (Below) The individual relative peak intensities of peaks in the specified spectra. Peak numbers are specified in Figure 5.12.

We then performed a ^{13}C-^{1}H SO-FAST HMQC (Schanda and Brutscher, 2005) of δIle-^{13}C,^{1}H, U-[^{15}N-^{13}C,^{2}H] FL-BiP (WT) in the presence of ATP, IRE1-cLD and ΔEspP to further investigate the nature of this interaction. This indicated, that when compared to FL-BiP (WT) + ΔEspP and ATP, no significant chemical shift perturbations or changes in peak intensity were observed (Figure 5.16). This is indicative of either no interaction or a transient interaction between BiP and a large species, invisible to NMR due to its short lived nature and size. As an additional control, to further confirm that reduction of IRE1-LD oligomers was not merely due to BiP-FL (WT) binding to the ΔEspP peptide, we compared the NMR spectra of BiP-FL (WT) in the presence of ATP and P2, a well characterised model peptide of BiP (Marcinowski et al., 2011), to the previously obtained spectrum in the presence of IRE1-LD, ΔEspP and ATP. If binding was due to binding of ΔEspP to BiP, we would expect to see characteristic peaks of BiP substrate binding, that were previously described for the T229G variant in the ADP bound, high substrate affinity state (Wieteska et al., 2017). Upon overlaying both spectra, the peaks observed in the presence of P2 (Figure 5.16), that are characteristic of substrate binding are not present in the spectrum in the presence of IRE1-cLD and ΔEspP, indicating that BiP does not reduce IRE1-LD oligomers by binding to ΔEspP. Taking other results into account, we conclude that BiP reduces IRE1-LD oligomers in a chaperone dependent manner.

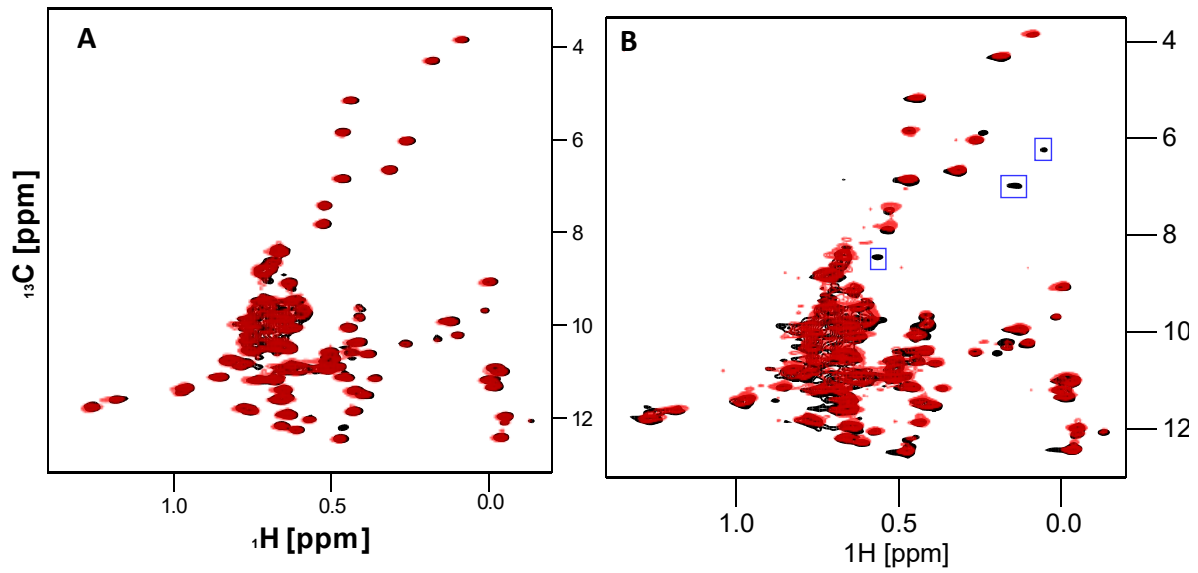

Figure 5.16

^{13}C-^{1}H SOFAST-HMQC of (A) 50 µM FL-BiP (WT) and 40 mM ATP In the absence (black) and presence (red) of 50 µM IRE1-cLD and 100 µM ΔEspP. (B) 50 µM FL-BiP (WT) and 40 mM ATP in the presence of 100 µM P2, a BiP model substrate (Marcinowski et al., 2011) (black) (spectrum acquired by Dr. Lukasz Wieteska at 750 MHz) and 50 µM IRE1-cLD, 50 µM FL-BiP (WT), 100 µM ΔEspP and 40 mM ATP (red). No chemical shift perturbations are observed in (A) and the characteristic peaks indicating substrate binding of a small peptide are not observed in (B), indicating FL-BiP (WT) is not interacting with the ΔEspP peptide.

5.5 Conclusions

In this chapter we have used a combination of structural and biophysical techniques to characterise the canonical and non-canonical interactions between IRE1-LD and BiP. In doing so, we have demonstrated that BiP interacts with IRE1-LD oligomers in a chaperone like manner to reduce them into its inactive monomeric/dimeric form. This process requires a functioning ATPase domain, and an active substrate binding domain, indicating that this process is a canonical interaction. While this model of negative regulation of IRE1 activity by BiP has previously been proposed (Bertolotti et al., 2000; Amin-Wetzel et al., 2017), to our knowledge, this is the first time it has been demonstrated that BiP reduces IRE1-LD oligomers.

This type of behaviour has been characterised between HSP90 and heat shock factor 1 (HSF1), with the former reducing a trimeric form of the latter during times of heat shock (Zou et al., 1998).

While we have not explored the role of the Hsp40 co-chaperone, ERdj4 in this interaction, we would hypothesise that in line with other Hsp40s, and as suggested by the authors (Amin-Wetzel et al., 2017), the C-terminal targeting domain strengthens the interaction between IRE1-LD and BiP, while the J-domain increases the ATPase rate of BiP, thus effectively enhancing the effective ability of BiP to de-oligomerise IRE1-LD.

As discussed earlier, it had previously been proposed that the functional interaction between IRE1-LD and BiP was solely non-canonical (Carrara et al., 2015). While our data does not support these claims, our characterisation of the ATPase activity of BiP in a sodium based buffer indicates that both ATP binding and the ATPase activity of BiP is significantly compromised, when compared to potassium based buffers, in agreement with similar observations in other Hsp70s (Palleros et al., 1993; Feifel et al., 1996). Carrara et al. used buffers which were composed of either 75 mM NaCl and 30 mM KCl for MST studies or just 75 mM NaCl for pull down assays. Our MST data indicates that we did not observe a significant effect with lower concentrations of IRE1-LD, however at higher concentrations, where there is a greater population of IRE1-LD oligomers, for example in pull down assays, the presence of sodium may affect findings.

We have determined a functional role for the non-canonical interaction, in addition to that of the canonical interaction, specifically a retardation of the rate of disulphide bond formation of IRE1-LD, which has previously been shown to be important for prolonging its activation (Eletto et al., 2014). Carrara et al. suggested that the BiP-NBD impeded tetramer formation of IRE1-LD using cross-linking experiments. Our data indicates that BiP-NBD impedes the formation of disulphide linked dimers and tetramers, thus it is plausible that both experiments are observing the same process. This may also offer an explanation as to why FL-BiP (apo) and the T229G variant (ATP) have an apparent effect on the large species observed by DLS, and why FL-BiP (apo) has an effect on the peak height observed in the ^{15}N-^{1}H TROSY spectrum of IRE1-LD, with these interactions governed by the nucleotide binding domain of BiP. Additionally, the OD400 assay and DLS data indicates a reduced concentration of insoluble particulates and smaller particles are observed for IRE1-LD in the presence of

T229G variant and ATP. One possible explanation is that while ATPase activity is reduced for this variant but not abolished (Yang et al., 2015) therefore the residual ATPase activity may reduce the size of the oligomeric species observed.

In the previous chapter we characterised the conformational landscape of IRE1-LD, in this chapter we have demonstrated that BiP has an ATPase dependent effect on the non-covalent monomer/dimer/oligomer landscape, specifically shifting the equilibrium towards inactive conformations of IRE1. In addition, BiP-NBD retards the rate of disulphide bond formation, thus possibly in combination with PDIA6 (Eletto et al., 2014; Eletto et al., 2016; Groenendyk et al., 2014), prevents prolonged activation of IRE1-LD via disulphide bond formation (Figure 5.17).

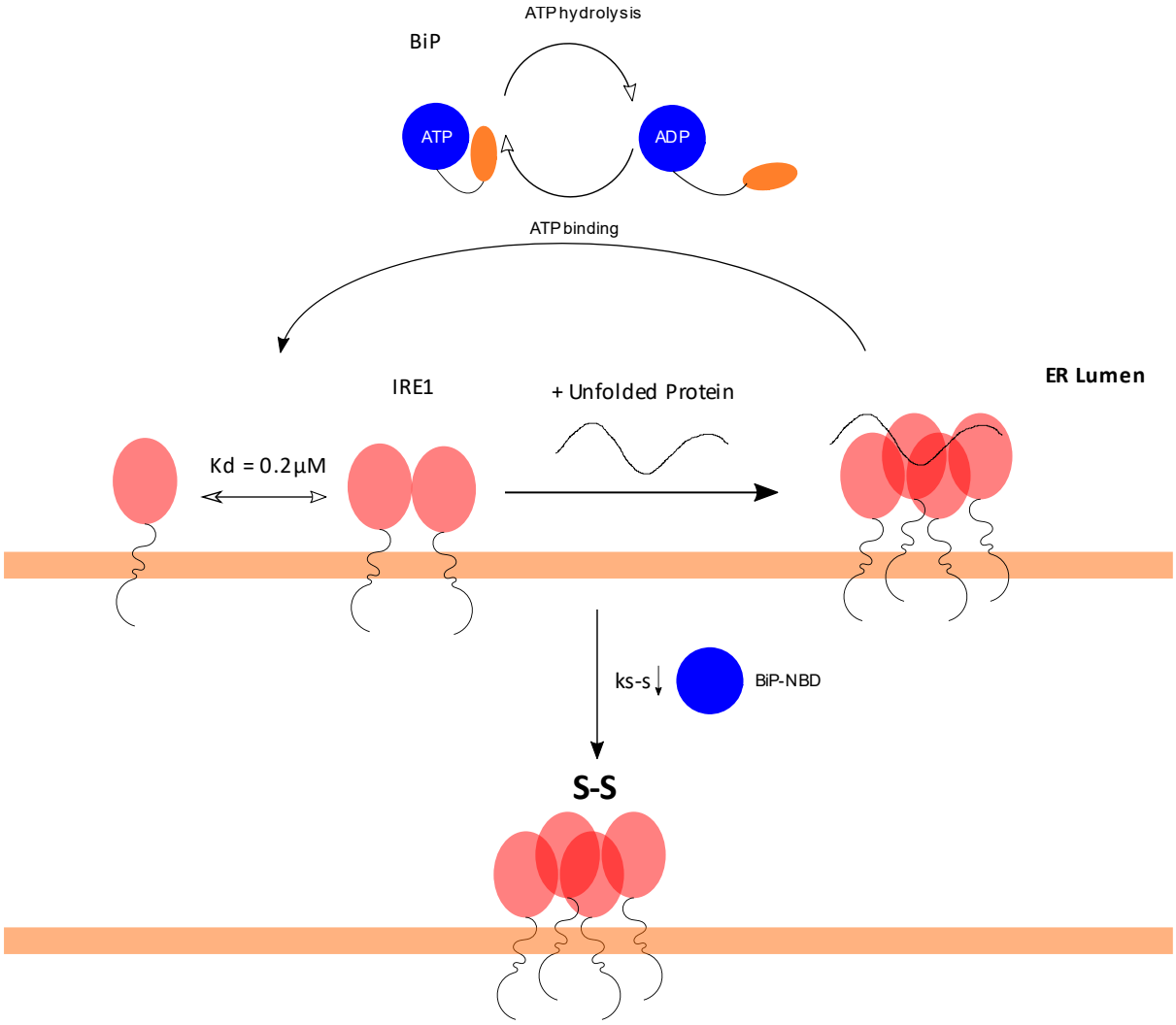

Figure 5.17

Model based on new data: BiP alters the conformational landscape of IRE1-LD by chaperoning IRE1-LD oligomers, thus retaining them in an inactive state. This process requires ATPase activity and an active substrate binding domain. The BiP-NBD reduces the rate of disulphide bond formation of IRE1-LD, thus preventing prolonged activation through disulphide linked IRE1-LD, as previously described in (Eletto et al., 2014).

6 Identification of lead compounds for $A\beta_{1-42}$

The previous chapters in this thesis have focussed on elucidating activation pathways that are important in regulating protein quality control elements, specifically IRE1 and its role in the wider UPR. While this work should provide additional insights into ways to regulate the activation pathway of IRE1, thus potentially provide a means to treat various disease states, this chapter will focus on chemical regulation of a well characterised disease pathway, namely the amyloid cascade pathway and its role in AD.

Neurodegenerative diseases in particular, are of pressing financial and social concern, due to a lack of effective treatments, with options currently limited to symptomatic treatments. Of these, Alzheimer's disease is the most common with over 20 million cases worldwide (Goedert and Spillantini, 2006). AD is characterised by deposits of amyloid beta (Aβ), neurofibrillary tangles of hyperphosphorylated tau and neuronal death. Improper proteolytic processing of amyloid precursor protein (APP), a type I transmembrane protein localised in the ER, leads to the formation Aβ in the ER lumen. While APP is ubiquitously expressed in all tissues, increased levels of expression are seen in brain tissues. Eight isoforms of APP are formed by a variety of splicing pathways, although currently there is no definitive evidence of the function of these isoforms (Placido et al., 2014).

The amyloid cascade hypothesis is currently the most widely accepted model for AD progression. Upon improper processing of the APP, monomeric Aβ oligomerises into a wide range of soluble oligomeric species. Toxic, on-pathway oligomeric species form insoluble fibrils which make up the characteristic deposits in brain tissues (Hardy and Higgins, 1992) (Figure 6.1)**.** While there is debate in the literature regarding the model of AD progression (Herrup, 2015; Selkoe and Hardy, 2016), the oligomeric species is widely considered the toxic species that is pivotal to disease progression (Campioni et al., 2010; Mannini et al., 2014; Arosio et al., 2014).

Several observations strongly support the link between APP and AD. Lifelong overexpression of APP has been shown to be causative for AD. This is further highlighted with the prevalence of AD in Down's syndrome, due to the presence of the APP gene on the chromosome that's duplicated in Down's syndrome, chromosome 21 (Selkoe and Hardy, 2016).

Structural studies of Aβ have understandably been difficult, due to the inhomogeneous structures formed upon oligomerisation and fibrilisation, and inconsistencies between methodology used to produce and analyse these fibrils. This is highlighted by the inhomogeneity of structures of amyloid beta fibrils. While common features such as a stacked parallel beta sheets are observed, variations are observed in contacts and stoichiometry. Two recent structures from independent labs, both derived using solid state NMR, were near identical, perhaps offering some further structural insights (Colvin et al., 2016; Walti et al., 2016).

Aβ not only interacts with itself to undergo oligomerisation and fibrilisation, but also interacts with other aggregation prone proteins, such as tau (Bloom, 2014) and the prion protein (Jarosz-Griffiths et al., 2016). Interactions between Aβ oligomers and the prion protein have been reported in the literature, suggesting a possible role in these proteins in AD progression (Barry et al., 2011; Lauren et al., 2009). Unsurprisingly, with a lack of clear understanding about the disease pathway, there are currently no approved treatments for AD. While the exact pathways are not clear, Aβ is commonly linked with AD, with the oligomeric species widely regarded as the toxic species, therefore this is the obvious target for screening, however these are usually present as part of a homogenous population, therefore it can be difficult to target these individually.

Targeting the toxic oligomeric species can be challenging due to the heterogenous nature of Aβ oligomers, however by monitoring processes that are caused by the oligomeric species, such as fibrilisation, it is possible to identify compounds that target this toxic species. There are several methods that are able to monitor the formation of fibrils, such as the ThT assay and the congo red spectral shift assay, however these assays are not without their drawbacks; the ThT dye has been shown to interact with hydrophobic pockets in non-fibrillar proteins (Biancalana and Koide, 2010), while compounds such as curcumin can induce false negative readings due to an increase ThT induced fluorescence (Hudson et al., 2009), while both dyes have variable affinities for different fibril structures under different conditions (Groenning, 2010).

An alternative to measuring fibrilisation, is to directly monitor a compound to ascertain if it interacts with a heterogeneous population of Aβ. NMR can be utilised in this manner as a high-throughput technique by directly monitoring resonances of the ligand. The use of ^{19}F

NMR can be particularly advantageous over ^{1}H NMR methods, as these produce more complex spectra and have the added complication of interference from most common solvents such as DMSO and H_2O. ^{19}F NMR is becoming increasingly popular in the field of drug screening (Dalvit et al., 2003a; Dalvit, 2007), for these reasons and the fact that binding of a small molecule to a large protein induces an obvious reduction in the peak height due to the intrinsic properties of the ^{19}F nucleus (Dalvit et al., 2003b) (covered in detail in section 1.6.3). Using ^{19}F NMR with in combination with high throughput based sampling, such as the SampleJet and a well-designed and diverse compound library, it can be possible to rapidly screen a large library of compounds to identify lead compounds for a particular target. In this chapter, we employed ^{19}F NMR to screen a large library of compounds against a Aβ preparation, followed by affinity mass spectrometry screening of a small selection of hit compounds to confirm these findings.

6.1 Compound library design (work performed by Dr. Matthew Baumgartner)

Upon oligomerisation, Aβ forms a heterogenous population of oligomeric species. With the lack of structural information about on-pathway oligomers it was not possible to design compounds to target this species specifically, therefore as an alternative approach, we identified a large data set of compounds that have been determined to interact with monomeric, dimeric, oligomeric and fibrillar species of Aβ from the literature, based on methods previously described (Joshi et al., 2016). To reduce the spectral complexity and interference from solvents, we used ^{19}F NMR to identify compounds that interact with Aβ oligomers, however this meant it was essential for the library compounds to contain at least one fluorine atom. A library search of ChEMBL (Gaulton et al., 2012; Bento et al., 2014; Gaulton et al., 2017) and the literature was performed for compounds shown to interact with $Aβ_{1-42}$. 151 compounds were initially identified and, subsequently these were subject to computational fragmentation using an internal Eli Lilly software tool Dicer. A search of compounds containing these substructures was performed on compounds present in the Eli Lilly internal repository that are present public libraries (PubChem (Kim et al., 2016), ChEMBL (Gaulton et al., 2012; Bento et al., 2014; Gaulton et al., 2017), ZINC (Irwin et al., 2012) and Drugbank (Wishart et al., 2018)), to build a large library of compounds. These were then filtered to remove low complexity structures, such as phenyl rings and duplicate structures, and ensure a chemically diverse range of compounds was present in the library. These

resulting compounds were filtered by solubility to produce a library of 971 compounds was then taken forward for experimental analysis (Figure 6.1).

6.2 Suitability of compounds and Aβ$_{1-42}$ preparation

While the compound library had undergone an *in silico* solubility screen, it was necessary to ensure in practise compounds were both soluble in the buffer of choice and that compounds hadn't degraded. To ensure this, 50 µM of each compound was individually screened in PBS buffer by ^{1}H and ^{19}F NMR. Compounds that were either completely insoluble, produced a poor signal to noise ratio (less than 3). The ^{1}H and ^{19}F NMR spectra of each compound was also analysed with aberrant spectra, for example spectra compounds that apparently contained one fluorine atom whose spectra exhibited a number of fluorine resonances, were eliminated from the compound library. In total 317 compounds were eliminated from the library at this stage.

As Aβ$_{1-42}$ preparations vary, it was necessary to validate the *in vitro* preparation to ensure the sample was stable for ample time to perform screening of several compounds, and to ensure that our preparation contained on-pathway oligomers. Stability was monitored using DLS, with the size of the oligomeric preparation monitored at hour intervals, up to and including four hours. As DLS is intensity weighted, i.e. larger species scatter light more intensely, it was necessary to convert the intensity weighted data to mass weighted data to ascertain the relative population of oligomers present in the preparation. This indicated that the Rh of the major species present (determined by mass) was approximately 10 nm at all time points (Figure 6.4, Table 6.1), indicating that the size of the oligomeric species monitored over this time period would remain relatively similar for all experiments. As we have determined species by mass, it is still possible that we have a significant population of smaller Aβ species present in our preparation.

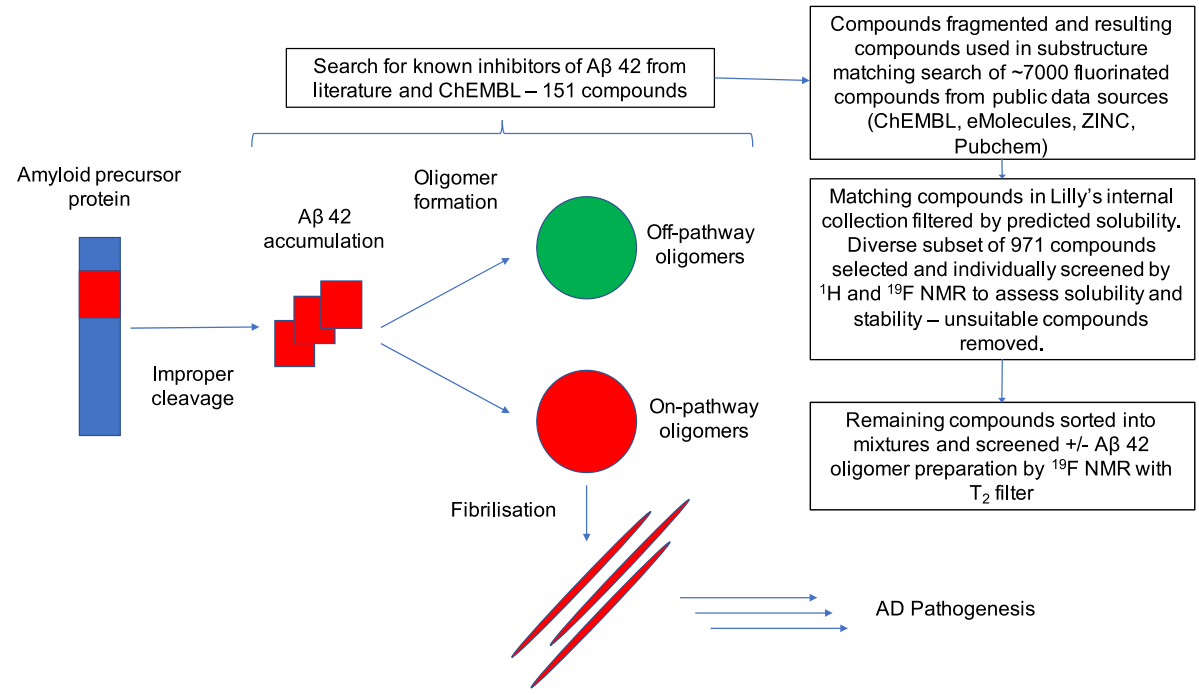

Figure 6.1

The Amyloid Cascade Hypothesis indicates that improper cleavage of the Amyloid Precursor Protein (APP) results in an accumulation of $A\beta_{1\text{-}42}$. This in turn leads to formation of a number of oligomeric species of $A\beta_{1\text{-}42}$. On pathway oligomers proceed to form fibrils that p leads to Alzheimer's disease pathogenesis. Our strategy for selection of a compound library is indicated above, but briefly involves searching for known compounds that interact with $A\beta_{1\text{-}42}$, fragmenting these compounds, searching Lilly's internal database and, after an initial screen of all compounds to assess solubility and stability in aqueous buffer, performing a screen using ^{19}F NMR in mixtures of compounds.

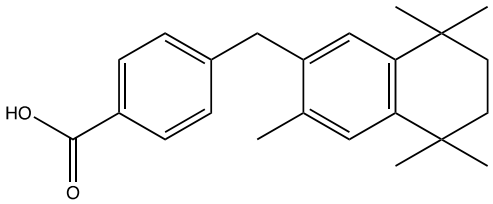

Figure 6.2

The structure of bexarotene, an anticancer compound previously shown to interact with Aβ oligomers (Habchi et al., 2016).

As our *in vitro* preparation will contain several different species, it was then necessary to ascertain if they contained on-pathway $A\beta_{1-42}$ oligomers (Fandrich, 2012). To do this, we took a compound that was recently shown in the literature to bind on-pathway oligomers, the anti-cancer compound, bexarotene (Figure 6.2) (Habchi et al., 2016), and performed ^{1}H NMR with a Carr-Purcell-Meiboom-Gill (CPMG) sequence (Loria et al., 1999), which culminates in an intensity loss of the resonances of any nuclei undergoing chemical exchange, in the presence and absence of our $A\beta_{1-42}$ preparation. Several peaks in the ^{1}H spectrum of bexarotene were observed to reduce in intensity (Figure 6.3) in the presence of the $A\beta_{1-42}$ preparation, thus indicating an interaction between our $A\beta_{1-42}$ preparation and bexarotene. This validated our $A\beta_{1-42}$ preparation as a target for further experiments.

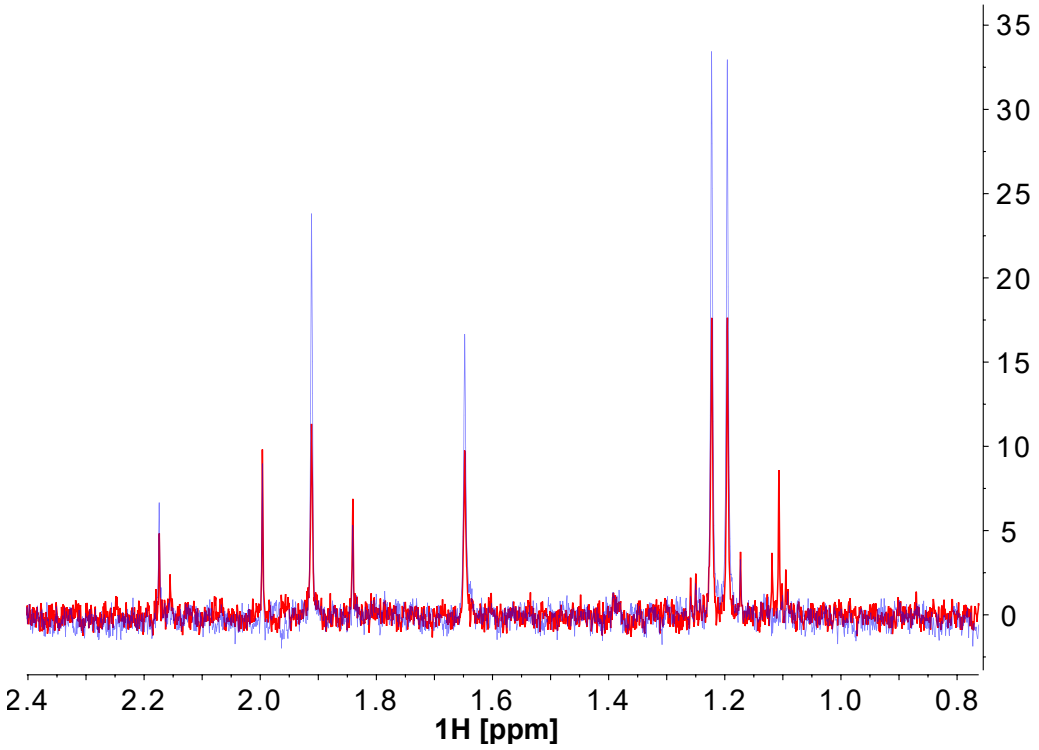

Figure 6.3

Section of ^{1}H NMR spectra of 50 µM bexarotene in the presence (red) and absence (blue) of 1 µM of $A\beta_{1-42}$ preparation (as monomer). Both spectra included a CPMG filter in the pulse sequence that reduces the intensity of resonances undergoing chemical exchange, such as those that are undergoing binding. The reduction of peak height upon the addition of the $A\beta_{1-42}$ preparation indicates an interaction between bexarotene and the preparation.

6.3 Initial identification of compounds binding to Aβ$_{1-42}$ preparation by ^{19}F NMR

One of the advantages of using NMR as a screening method is that the spectrum of a compound provides an easily identifiable signal. Assuming there's sufficient resolution between the chemical shift of peaks emanating from a selection of compounds, it is possible to screen compounds in mixtures, rather than individually, thus increasing the throughput for compound screening.

After initially removing compounds from the library that were not adequately soluble or unstable in aqueous buffer, the remaining 654 compounds were organised into groups, each of which contains six or seven compounds. As the ^{1}H chemical shift range only spans ~ 10 ppm, it is possible to excite all these resonances with a pulse at a single offset frequency, often ~4.5 ppm. ^{19}F resonances span a chemical shift range of ~300 ppm, thus it is not possible to excite all frequencies in the range with a pulse at a single offset frequency. NMRMix (Stark et al., 2016), a program designed to automate mixture preparation of compound screens using ^{1}H NMR, was used as a first pass to ensure that compounds with overlapping resonances were not placed in the same mixtures. Following mixture preparation with NMRMix, mixtures were manually refined to ensure resonances were within 15 ppm from the offset frequency.

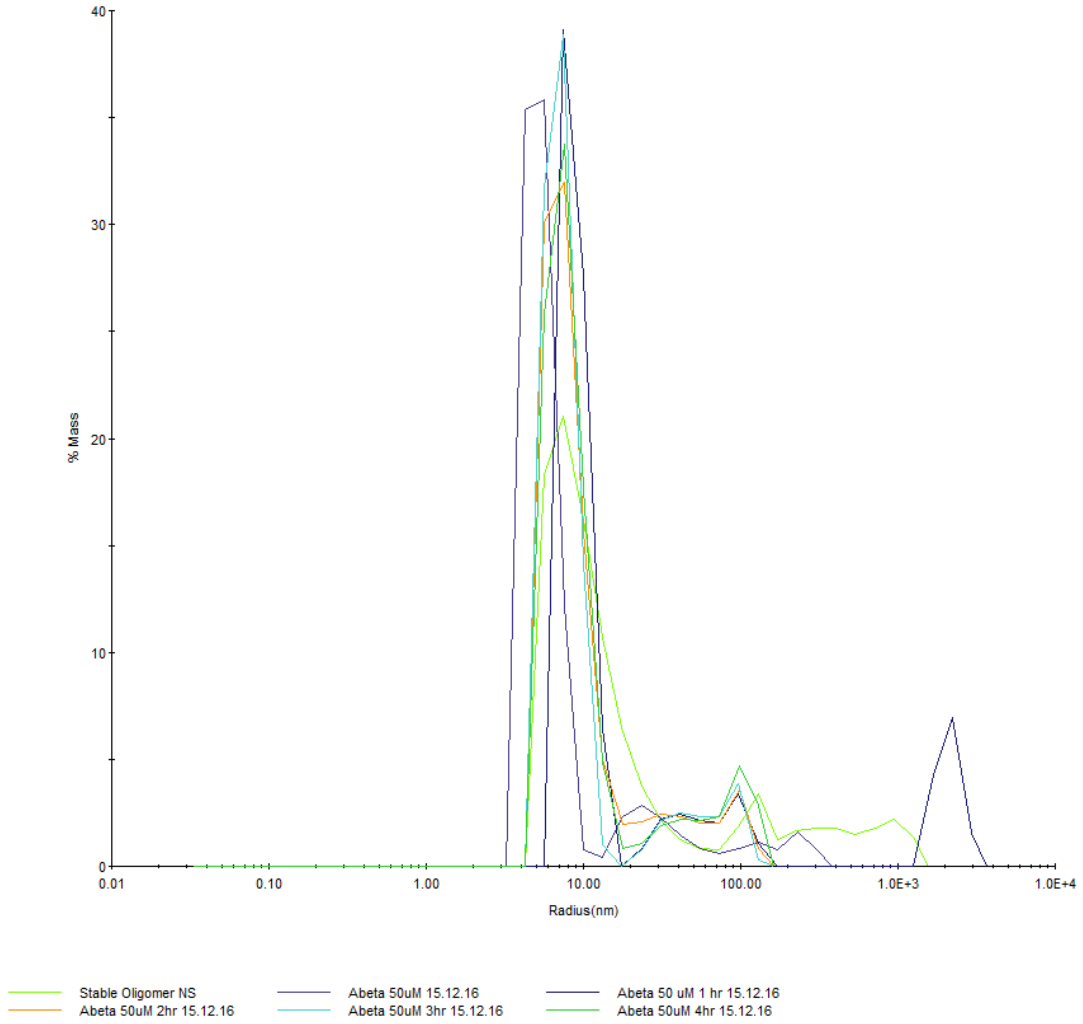

Figure 6.4

$A\beta_{1-42}$ preparations were prepared in PBS. 50 µM of $A\beta_{1-42}$ (concentration as monomer) was analysed by DLS immediately following preparation and measurements were taken at one hour intervals. Regularisation data is shown which indicates the size of the main species does not change significantly over a four hour period, meaning the size of the oligomers present remains relatively consistent throughout all NMR experiments

	Radius (nm) - T= 1 hour	Radius (nm) - T = 4 hours
Peak 1	5.410	7.9
Peak 2	31.750	40.130
Peak 3	131.936	101.028
Peak 4	252.797	N/A

Table 6.1

The calculated size of the amyloid beta oligomers using the regularisation algorithm at T = 1 hour and T = 4 hours are shown. The size of the main species did not change significantly over the time period.

Once mixtures were prepared, ^{19}F CPMG experiments were performed in the presence and absence of our $A\beta_{1-42}$ preparation. Any compounds that produced a reduction in peak height of greater than twice the height of the spectral noise were deemed to be active hits, while compounds that did not fulfil this requirement were removed. In total, 26 compounds in the mixtures exhibited a reduction in peak height of greater than twice the signal to noise upon addition of our Aβ preparation (Figure 6.5). These compounds were screened individually as a control to ensure interactions were not occurring between compounds in the mixtures (Figure 6.6), then taken forward for further analysis.

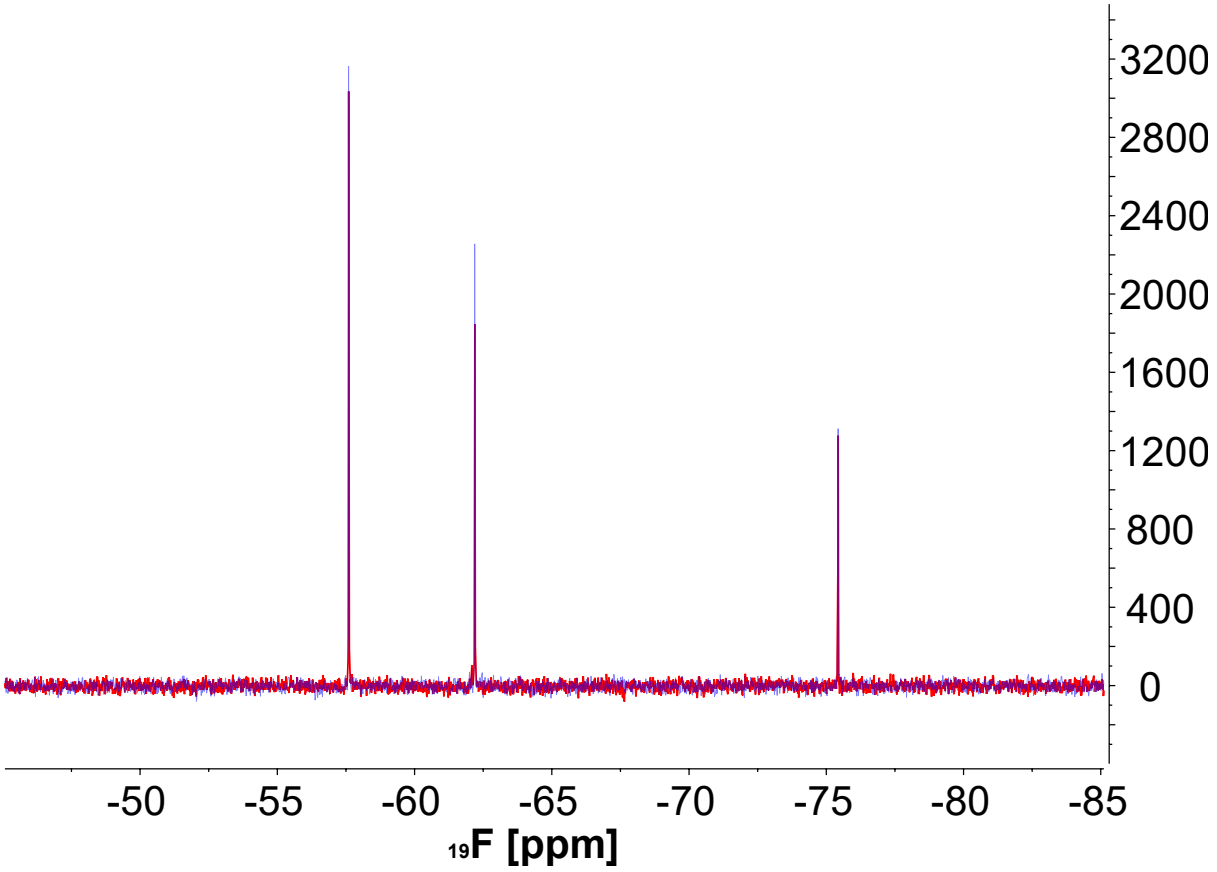

Figure 6.5

A mixture of 50 µM of each compound was screened in the presence (red) and absence (blue) of 1 µM (concentration as monomer) $A\beta_{1\text{-}42}$ preparation by ^{19}F NMR with a CPMG filter. In this example the compound responsible for the resonance in the centre of the spectrum (~62 ppm) was analysed in further experiments while the other two compounds were discarded from the library.

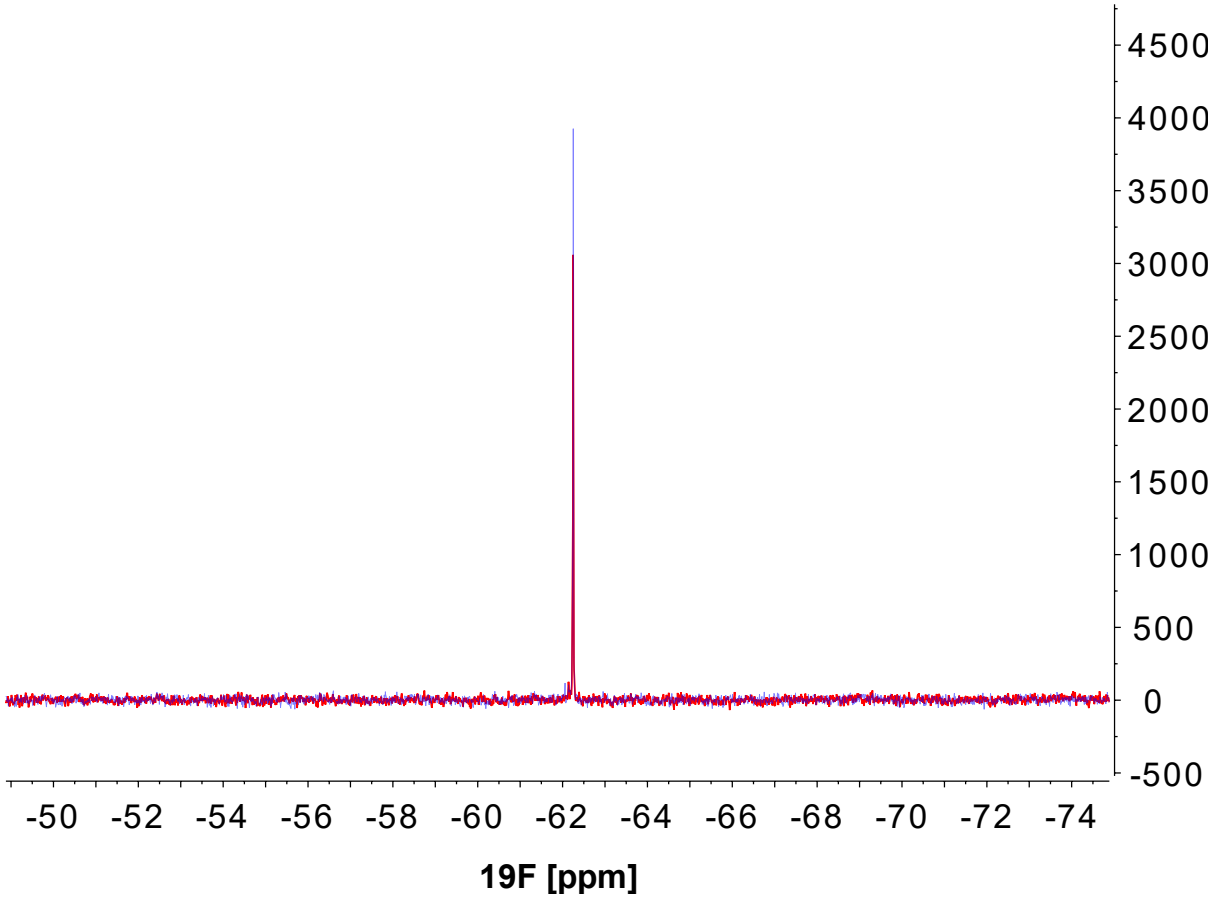

Figure 6.6

^{19}F NMR spectra of 50 µM compound 9 in the presence (red) and absence (blue) of 1 µM (concentration) $A\beta_{1\text{-}42}$ preparation. The reduction in peak height confirms the interaction observed for this compound in the mixtures and rules out any interaction with other compounds.

6.4 Removal of non-specific binding compounds

While we had identified 26 compounds that interacted with our $A\beta_{1\text{-}42}$ preparation, due to its aggregation prone nature, compounds are very likely to bind to Aβ in a non-specific manner. To remove non-specific binding compounds, we employed the use of a scrambled Aβ preparation. This was prepared in the same manner as the $A\beta_{1\text{-}42}$ preparation and formed similar size oligomers, however only compounds with a non-specific binding mode, for example binding to the surface of Aβ oligomers, would interact with this preparation. 14 of the 26 compounds were identified as binding in a non-specific manner (reduction in peak

height > twice the spectral noise), thus were eliminated from the hit list (Figure 6.7), leaving a hit list of 12 compounds (Figure 6.9)

To ascertain if any compounds had a similar binding mode to that of bexarotene, we then carried out a competition-based experiment. While this approach would be further evidence of a specific binding mode, we would not eliminate compounds from the library based on these results as it is feasible that compounds in the library both bind specifically but do not possess the same binding mode of bexarotene. Each of the 12 remaining compounds were analysed as before by ^{19}F NMR, monitoring the peak height of the resonance(s) emanating from the library compound, but on this occasion an equimolar concentration of bexarotene was added. Interestingly, the ^{19}F spectrum of compound 9 showed a significant increase in peak height (>twice the spectral noise) when compared to a control containing compound 9 and DMSO (Figure 6.8), indicating competition between bexarotene and compound 9 for $A\beta_{1-42}$. None of the other compounds displayed any change in peak height upon addition of bexarotene.

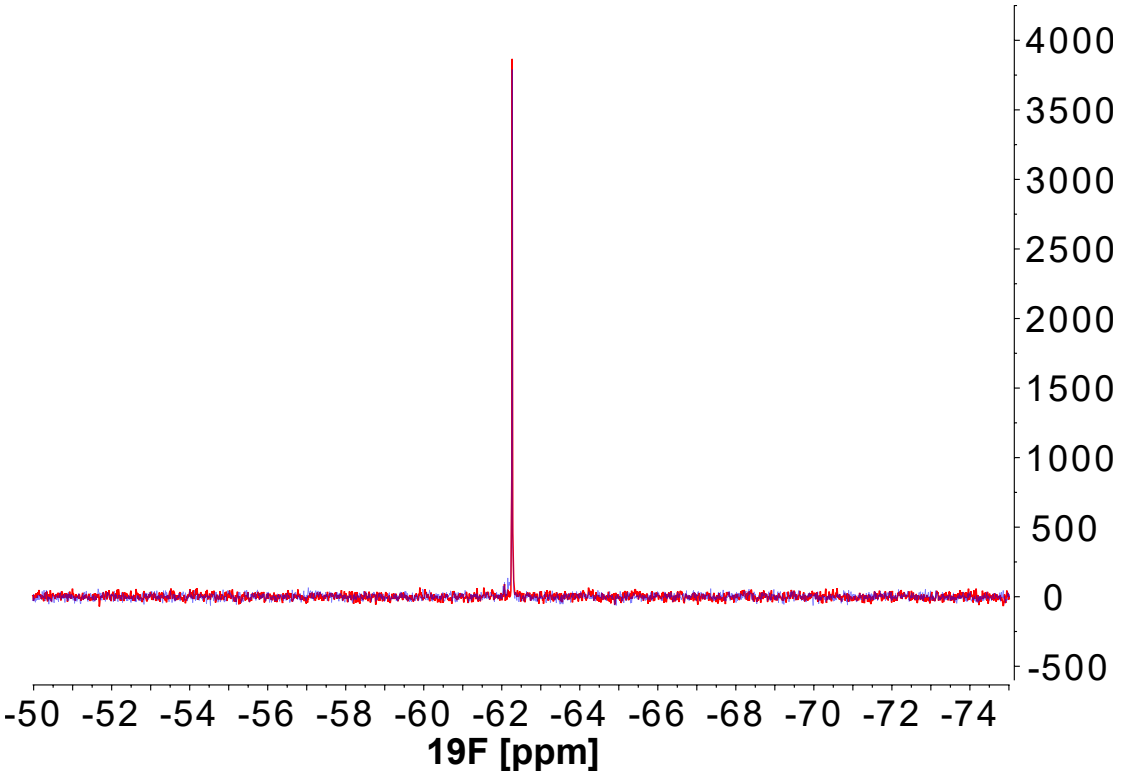

Figure 6.7

1 µM of a scrambled peptide (Aβ) was used to identify any promiscuous binding compounds. The ^{19}F NMR spectra of 50 µM compound 9 in the presence (red) and absence (blue) of the scrambled peptide preparation indicates this compound did not interact with this preparation due to no significant drop in peak height.

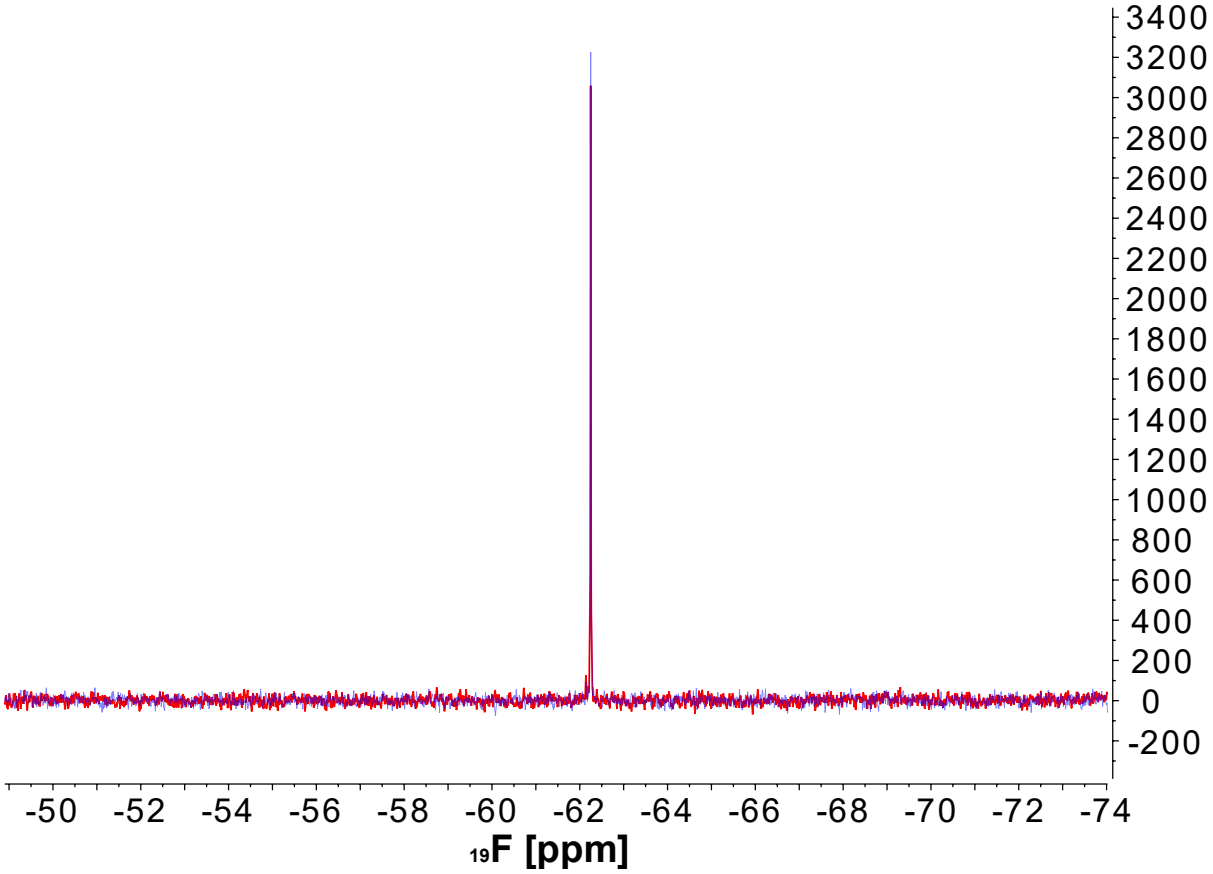

Figure 6.8

An increase in the peak height in the spectrum of 50 µM compound 9 is observed upon addition of an equimolar concentration of bexarotene (blue), when compared to that of compound 9 and 1 µM (concentration as monomer) of the Aβ preparation only (red), indicating competition between bexarotene and compound 9 for our Aβ preparation.

6.5 Affinity Mass Spectrometry

While NMR offers a number of advantages for screening a compound library, like other methods it is still prone to artefacts. Therefore, to confirm binding, we used an orthogonal technique. The use of equilibrium dialysis coupled with elected ion monitoring liquid chromatography mass spectrometry (SIM-LC-MS) offers a sensitive and specific way to quantitatively detect binding of a compound to a target protein (Waters et al., 2008). In brief, equimolar concentrations of compound in an aqueous buffer is dispensed inside and outside the dialysis membrane. If a compound has an affinity for the target protein then an uneven distribution of compound on either side of the dialysis membrane would be expected that

would be manifested in a reduction in the peak area of the chromatogram of the sample taken in from the dialysis device in the presence of the target protein, when compared to that in the absence.

Three compounds (1, 4 and 9) were chosen for analysis by AMS. Compound 1 showed a reduction in peak height with the retention time of both peaks (0.98 mins) identical to the control samples in the absence of the Aβ preparation and almost identical to that analysed in full spectrum scan mode (0.97 min) (Figure 6.10), however this maybe due to a different solvent condition (1:1 H_2O:MeCN). No detectable peak for compound 9 was observed in the presence of $A\beta_{1\text{-}42}$ preparation, however as the peak in the absence of $A\beta_{1\text{-}42}$ preparation (1.04 and 1.05 min) did not overlap completely with the control run in full spectrum scan mode and dissolved in 1:1 H_2O:MeCN, further optimisation is required to confirm this result (Figure 6.11). Compound 4 did not show any change in peak height. While these assays provided some further evidence of an interaction between these compounds and $A\beta_{1\text{-}42}$, they require further optimisation before drawing any significant conclusions from this data.

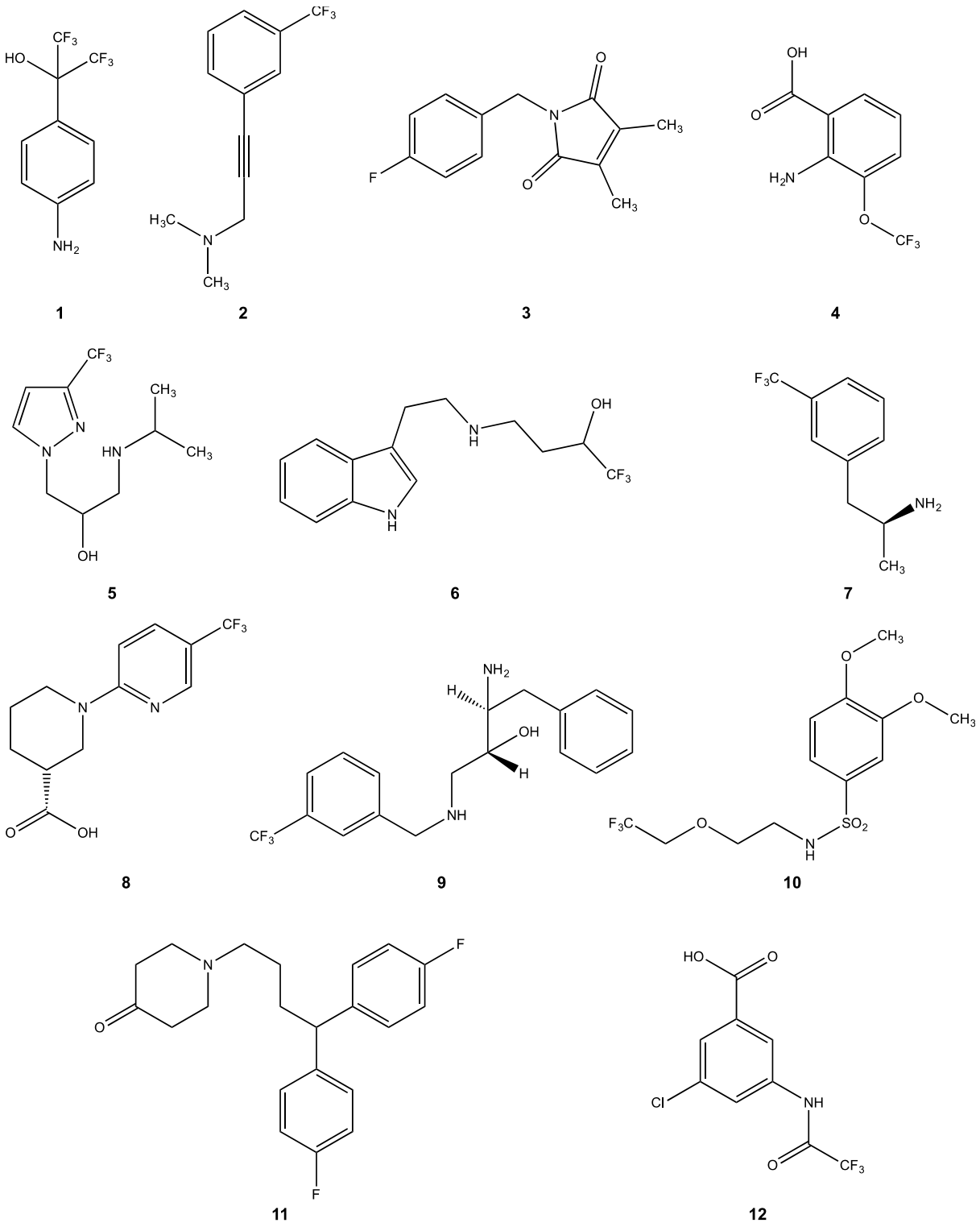

Figure 6.9

The twelve compounds that were identified as binding to our Aβ preparation. All these compounds did not show a significant reduction in peak height upon addition of the scrambled peptide preparation. The ID number from the corresponding databases is shown in table 6.2.

Compound Number	Database ID Number
1	CHEMBL1499171
2	PBCHM3049683
3	PBCHM4680099
4	PBCHM68501019
5	PBCHM4284201
6	CHEMBL1351172
7	PBCHM120765
8	ZINC19735908
9	PBCHM57487213
10	ZINC04653292
11	CHEMBL1673279
12	PBCHM81560982

Table 6.2

Compound number (shown in Figure 6.9) with the corresponding database ID number.

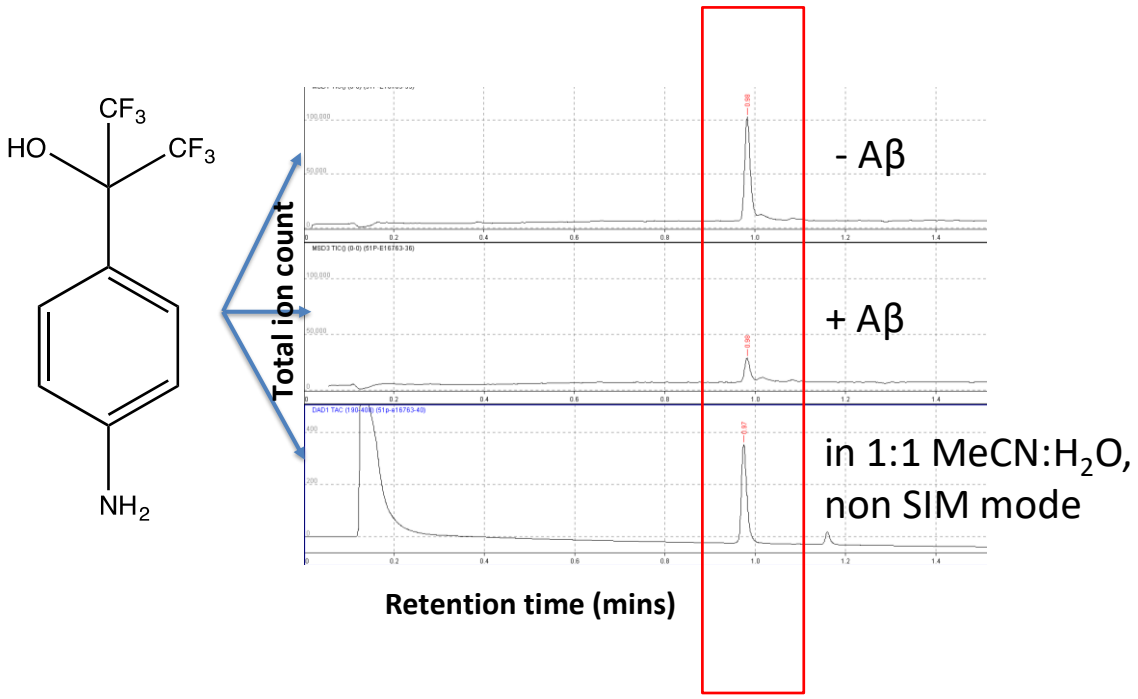

Figure 6.10

Affinity mass spectrometry of compound 1. SIM-LC-MS of compound 1 in the absence of our Aβ preparation (top), compared with that in the presence (middle) indicates a reduction in the concentration compound 1, due to an uneven distribution of compound on either side of the dialysis membrane due to an affinity for the preparation. LC-MS of compound 1 diluted in 1:1 H_2O:MeCN in non SIM mode (bottom) is used as a reference for the position of the peak in the chromatogram.

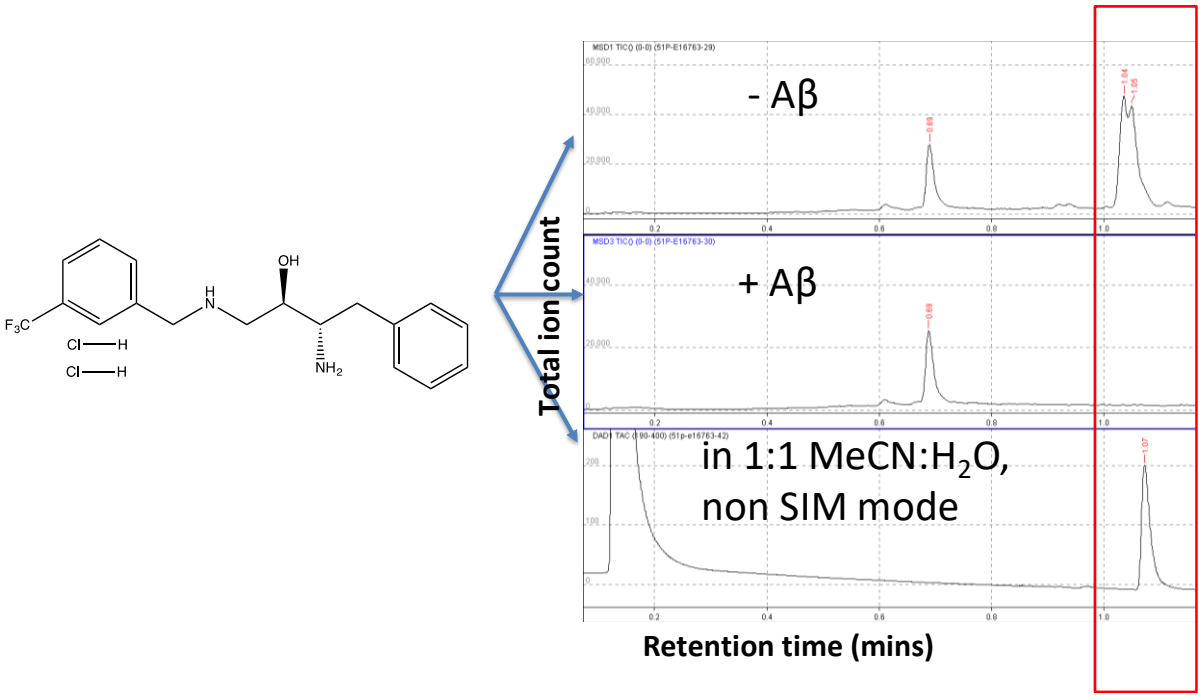

Figure 6.11

Affinity mass spectrometry of compound 9. SIM-LC-MS of compound 9 in the absence of our Aβ preparation (top), compared with that in the presence (middle) indicates a reduction in the concentration compound 9, due to an uneven distribution of compound on either side of the dialysis membrane due to an affinity for the preparation. LC-MS of compound 9 diluted in 1:1 H_2O:MeCN in non SIM mode (bottom) is used as a reference for the position of the peak in the chromatogram.

6.6 Conclusions

Our data displays the potential for a high throughput screen against amyloidogenic targets by NMR. The current method of choice for identifying lead compounds, ThT assays, while providing a demonstrable physiological output, namely a reduction in the rate of fibrilisation, it can be prone to false positives due to a number of reasons (Jameson et al., 2012).

While an interaction by NMR does not necessarily indicate a compound that will produce a physiological effect, it does provide a means of eliminating compounds from large libraries that do not bind or bind non-specifically. It is also worth noting, for a compound to produce a physiological effect, it is necessary for an interaction to occur, thus this methodology is a far more efficient means to screen a large and diverse subset of compounds as a means to perform more detailed testing on any molecules that have been identified to interact with the target molecule, in this case $A\beta_{1-42}$. In addition, as NMR monitors the ligand directly, it is not prone to artefacts in the same manner as dye based assays (Hudson et al., 2009).

As the majority of these compounds (all excluding compounds 9, 10 and 11) possess a molecular weight of <300 Da, there is potential to build upon these as scaffolds (Jhoti et al., 2013) to produce more complex compounds and potentially increase the affinity of these compounds for Aβ. Work carried out after the completion of the work in this chapter has identified two active scaffolds (Figure 6.12), based on compounds 3 and 12, to be active through further biophysical testing. Not only does this further validate the strategy used in this study, it also means it is possible to perform both searches of compound libraries containing these scaffolds, and synthesis of new compounds using these scaffolds, to identify additional compounds that interact with Aβ oligomers. Interestingly, both these scaffolds are either acetanilide or N-benzylacetamide derivatives respectively, both similar starting materials. Further derivatisation of these starting materials could be a useful strategy for future investigation.

It is worth noting, that all of these assays have been performed *in vitro*, therefore it would be worthwhile to perform *in vivo* screens using well established AD models (McColl et al., 2012; Choi et al., 2014) to ascertain that these compounds have a physiological effect.

One other possible drawback of this methodology is that compounds may be binding to off-pathway species. We limited this possibility by forming Aβ oligomers using similar methods

to those that have previously been validated (Stine et al., 2003; Balducci et al., 2010), however we cannot rule out the presence of smaller sized species in our preparation and orthogonal techniques such as *in vivo* screens and ThT assays should be used to further validate these hits.

In conclusion, several promising compounds and scaffolds have been identified against $A\beta_{1-42}$, which can be further investigated both by performing searches of chemical libraries for structural analogues, and organic synthesis using these scaffolds to potentially identify a pharmacophore. The methods employed in this chapter could also be applied to other drug targets as a means to rapidly identify compounds and scaffolds that interact.

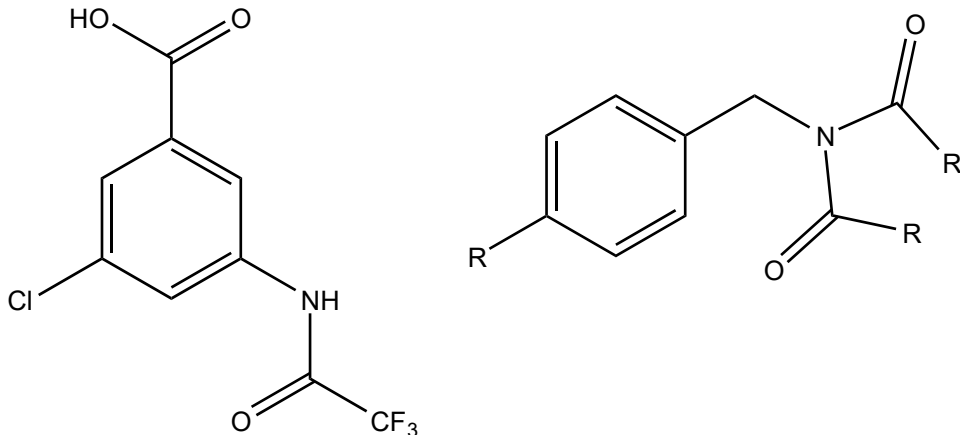

Figure 6.12

Scaffolds that have been confirmed as active, based on compounds 12 and 3 respectively.

7 Conclusions and Future Work

In this thesis, we have investigated two protein quality control pathways, the mechanism of activation of the ER stress sensor, IRE1, and small molecule regulation of Alzheimer's Disease by identifying a series of compounds that interact with Aβ.

7.1 The substrate induced activation pathway of IRE1

The data in this thesis indicates that IRE1-LD forms high affinity dimers in the absence of any substrate, in contrast with the yeast homolog (Credle et al., 2005). Recent data indicates that this is not the active form of the protein (Karagoz et al., 2017), however this step is essential for protein activation (Zhou et al., 2006). By combining ion mobility mass spectrometry and molecular dynamics simulation, we have built a structural model of the IRE1-LD dimer that suggests that it is arranged in a similar manner to its yeast homolog. While this model has not been verified using high resolution methods, it could provide insights into functional hotspots for further mutagenesis studies.

Through the use of NMR we have demonstrated that there is significant chemical exchange on the µs-ms time scale in the structured regions of IRE1-LD, that is not caused by monomer-dimer exchange. Recent data on a non-oligomerising mutant of IRE1-LD indicates that the spectral quality does not improve significantly upon this mutation. We would hypothesise that either conformational changes, possibly between hypothesised open and closed conformations (Karagoz et al., 2017) or exchange between dimeric and oligomeric species may be responsible for the µs-ms exchange observed in the spectrum of IRE1-LD. A very small population of oligomeric species could feasibly lead to the structured residues of IRE1-LD becoming invisible to conventional NMR techniques, thus if the non-oligomerising mutant proposed by Karagoz et al., does not completely ablate IRE1-LD oligomers, structured residues may not be visible due to this exchange. Our data highlights IRE1 is an incredibly dynamic system, with conformational changes in its structured regions leading to exchange broadening, and numerous intrinsically disordered regions, indicated by sharp peaks emanating from residues in its juxtamembrane linker and long loop which contains C332.

Substrate triggered oligomerisation has been shown to be essential for IRE1 activation in both yeast (Credle et al., 2005) and human (Karagoz et al., 2017) homologs. Like the yeast homolog (Gardner, B.M. and Walter, 2011), our data indicates that human IRE1 is sensitive to substrate

triggered dimerisation/oligomerisation. Interestingly, we have observed changes in the fluorescence polarisation and the particle size in the presence of the well characterised non-dimerising mutant, suggesting that substrate can trigger dimer/oligomer formation independently of the dimerization interface.

Structural studies investigating the process of peptide binding and oligomerisation indicates that, as expected with an increase in molecular weight, a reduction in peak height is observed, however this affects certain residues more than others, suggesting a loss of flexibility or a change in µs-ms dynamics in certain residues upon peptide binding. This may be caused by an interaction in the hypothesised oligomerisation interface that is present in residues V307-Y358 (Karagoz et al., 2017). Our data suggests that substrate triggered conformational changes, which leads to activation of IRE1, leads to a loss of flexibility in the disordered regions of IRE1.

While both the T229G variant of BiP in the presence of ATP, and apo BiP both have an effect on the size of IRE1-LD species, our data indicates that an active chaperone is necessary to reduce IRE1-LD complexes to the size seen in the inactive form of the protein. Our NMR data reveals structural insights into this process. The ^{15}N-^{1}H TROSY spectra of IRE1-LD are in good agreement with the DLS data with increases in the average peak intensity consistent with a reduction in the size of IRE1-LD. The analysis of individual peak intensities reveals that these changes are more variable upon addition of BiP and ATP, suggesting changes in µs-ms dynamics of certain residues, in addition to the effect of a change in the overall rotational correlation time of IRE1-LD. This is further highlighted upon comparison of spectra of IRE1-LD in the presence of BiP and ATP and that of IRE1-LD (D123P), which suggests significant changes in the µs-ms dynamics of identical residues are affected, when compared to the spectrum of IRE1-LD (WT). All this indicates that the canonical activity of BiP leads to de-oligomerisation of IRE1-LD, which culminates in increased flexibility of disordered regions of IRE1-LD. Our data is in good agreement with literature data.

We have not explored the effect of the Hsp40 chaperone, ERdj4 in this study, however we would hypothesise that this would strengthen this interaction by directing BiP to IRE1-LD and increasing the ATPase rate of BiP (Kampinga and Craig, 2010), in agreement with the findings observed by the authors of this study (Amin-Wetzel et al., 2017).

7.2 Covalent regulation of IRE1 activation

IRE1-LD has been shown to form active disulphide linked species *in vivo*, that require PDIA6 to attenuate IRE1 activity (Eletto et al., 2014). In this thesis we have performed detailed characterisation of the process of disulphide bond formation. Our data indicates the formation of disulphide linked species is a slow process, which occurs over several hours. Furthermore, our data indicates this does not require dimerization, indicating this is driven either by another process, possibly suggesting an alternative activation pathway for IRE1. We have gained structural insights into the process of disulphide bond formation, with our data demonstrating that upon disulphide bond formation, there is a loss of flexibility in disordered regions of IRE1-LD. Furthermore, our SEC data indicates larger IRE1-LD species are observed. We would hypothesise that under non-stressed conditions, PDIA6 negatively regulates IRE1 by an interaction via C148 (Eletto et al., 2014), whereas upon prolonged stress IRE1 forms disulphides by C148 and C332 which are reduced by PDIA6. Interestingly, PDIA6 is upregulated in times of ER stress (Vekich et al., 2012), which may suggest a negative feedback loop.

As mentioned in the introduction the BiP-NBD has been shown to interact with S1R (Ortega-Roldan et al., 2013), which has also been linked to stabilisation of IRE1-LD at the MAM, with ROS shown to prolong XBP1 signalling (Mori et al., 2013). We have demonstrated that the BiP-NBD impedes the formation of disulphide bonds. In addition, the BiP-NBD no longer binds to disulphide linked oligomers, suggesting BiP may repress this alternative activation pathway, however upon covalent linkage, its role in this pathway is redundant and PDIA6 is required to inactivate IRE1-LD (Eletto et al., 2014). This is further supported by our observation that both BiP T229G and apo BiP reduce the size of species observed upon peptide binding, as demonstrated by DLS and NMR. Taking these results into consideration, and previous data linking IRE1, S1R and the BiP-NBD (Mori et al., 2013; Ortega-Roldan et al., 2013), this could suggest a link between the redox state of IRE1-LD, its prolonged activation and a role for BiP-NBD in impeding this process.

7.3 Promising lead compounds for Aβ identified by ^{19}F NMR

Our screening protocol identified 12 hit compounds that can be taken forward for future studies. We have demonstrated that all these compounds do not bind to scrambled Aβ species, thus suggesting these compounds are not interacting in a non-specific manner,

however this does not completely rule this out. Furthermore, our preliminary data indicates that two of these compounds interact using an orthogonal technique, although this methodology does require further optimisation In addition one of these compounds was shown to compete with bexarotene, the anti-cancer compound that has previously been shown to interact with Aβ oligomers (Habchi et al., 2016).

The majority of the compounds identified are <300 Da, therefore can be classed as fragments, as such either further synthesis or searches of structural analogues of these fragments will likely need to be performed before a potent hit molecule can be determined. With the absence of a structure for on-pathway Aβ oligomers, this makes this prospect more challenging, however it may be possible to identify pore potent compounds through the use of methods like FAXS (Dalvit et al., 2003b), by utilising compounds from our hit list as spy molecules, and competing them off with structural analogues.

This work was part of a larger project that hypothesised that the interaction between PrP and Aβ oligomers leads to AD pathogenesis (Jarosz-Griffiths et al., 2016; Kellett and Hooper, 2009; Lauren et al., 2009). The development of a screening assay to ascertain if any of these compounds interact with PrP could provide some insights into the nature of the Aβ-PrP interaction and may provide some insights into AD pathways.

7.4 Future Work

Recent identification of an oligomerisation deficient mutant of IRE1-LD indicates that it is possible to separate the effects of peptide binding and oligomerisation (Karagoz et al., 2017). Using this mutant would make it possible to perform further NMR analysis on the peptide bound form of IRE1-LD and separate the effects of oligomerisation and conformational changes. While our current data suggests that the disordered residues of IRE1-LD become less flexible upon peptide binding, the use of this mutant would prove a useful control.

As the structured residues of IRE1-LD are not visible in 15N-1H TROSY experiments, it would be useful to identify the cause of the µs-ms exchange. Translational diffusion NMR experiments would be useful to identify the size of the species that the sharp residues are emanating from.

The current crystal structure of hIRE1-LD is hypothesised to be in a 'closed' conformation (Zhou et al., 2006). Further use of the non-oligomerising mutant of IRE1-LD may make it possible to crystallise IRE1-LD in its substrate bound form, thus gain structural insights into the conformational changes. Karagoz et al. hypothesised from their NMR studies that one of the disordered regions of IRE1-LD (V307-Y358) may form part of an oligomerisation interface. Both peak height and relaxation based experiments, such as CPMG, of resonances around C332, which we have identified through mutagenesis, may provide some information into conformational changes in this region upon peptide binding.

While the oligomerisation mutant may prove useful for future work investigating peptide binding, we would expect additional complications in the using this variant in studies investigating the interaction between IRE1-LD and BiP as oligomeric IRE1-LD maybe necessary for this interaction. A possible strategy to elucidate a binding interface would be through paramagnetic relaxation enhancement. This has previously been used to elucidate transient interfaces (Tang et al., 2008b; Tang et al., 2008a) and spin labels have successfully been introduced into the substrate binding domain of the bacterial homolog of BiP, DnaK (Schlecht et al., 2011). A possible drawback of this strategy is that IRE1 species that interact with BiP may already be invisible to conventional NMR methods. For any of these strategies, addition of ERdj4 (Amin-Wetzel et al., 2017) would likely prove a marked effect, and could prove useful in obtaining better quality data, although this would make it necessary to differentiate between an IRE1-LD ERdj4, IRE1-LD BiP and BiP ERdj4 interfaces.

We have carried out some analysis of the effect of sodium on the ATPase activity of BiP. In doing so, we have demonstrated BiP has a significantly compromised ATPase rate, when compared to that in the presence of potassium. Investigation into the effect of sodium on the ability of BiP to reduce IRE1-LD oligomers may provide further clarity on current literature data, using similar methods to those used in this thesis to investigate the canonical IRE1-LD-BiP interaction.

The role of disulphide bond formation in IRE1 is not well understood, despite literature indicating that IRE1-LD interacts with PDIA6 (Groenendyk et al., 2014), that this interaction is functionally important in regulating the XBP1 (Eletto et al., 2014) and RIDD (Eletto et al., 2016) activity of IRE1. Furthermore, recent work linking a possible role of ROS in IRE1 activity (Mori et al., 2013) suggests additional importance regarding the redox state of IRE1-LD.

Characterisation of the process of disulphide bond formation in the presence of PDIA6, and the process of disulphide bond reduction using SDS-PAGE and NMR could provide useful insights into the mechanism of this process *in vivo*. In addition, it would be useful to monitor the effect of BiP-NBD in this process, as BiP is known to interact with PDIA6 (Jessop et al., 2009), thus this may provide further mechanistic insights.

In our compound screen, we identified 12 compounds that interacted with Aβ species. As later work identified two active scaffolds, future work should focus on synthesising compounds based on these scaffolds, identifying structural analogues of these compounds and performing both further biophysical characterisation of these compounds, using a combination of NMR based methods, such as saturation transfer difference (STD) (Meyer and Peters, 2003), and ThT assays and *in vivo* screening methods to ascertain if these induce a physiological effect.

While the compounds identified in our drug screen provides sufficient material for future work, the methodology used in our compound screen can be used for other drug targets. This would only require minor modification of the methodology we used to create a compound library (Joshi et al., 2016), instead searching for known binders of another drug target. Due to the nature of ^{19}F NMR, larger drug targets would be advantageous as these induce a greater effect upon binding.

8 References

Ali, M.M.U., Bagratuni, T., Davenport, E.L., Nowak, P.R., Silva-Santisteban, M.C., Hardcastle, A., McAndrews, C., Rowlands, M.G., Morgan, G.J., Aherne, W., Collins, I., Davies, F.E. and Pearl, L.H. 2011. Structure of the Ire1 autophosphorylation complex and implications for the unfolded protein response. *Embo Journal.* **30**(5), pp.894-905.

Amin-Wetzel, N., Saunders, R.A., Kamphuis, M.J., Rato, C., Preissler, S., Harding, H.P. and Ron, D. 2017. A J-Protein Co-chaperone Recruits BiP to Monomerize IRE1 and Repress the Unfolded Protein Response. *Cell.* **171**(7), pp.1625-+.

Arosio, P., Cukalevski, R., Frohm, B., Knowles, T.P.J. and Linse, S. 2014. Quantification of the Concentration of A beta 42 Propagons during the Lag Phase by an Amyloid Chain Reaction Assay. *Journal of the American Chemical Society.* **136**(1), pp.219-225.

Aslanidis, C. and Dejong, P.J. 1990. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Research.* **18**(20), pp.6069-6074.

Ayala, I., Sounier, R., Use, N., Gans, P. and Boisbouvier, J. 2009. An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. *Journal of Biomolecular Nmr.* **43**(2), pp.111-119.

Balducci, C., Beeg, M., Stravalaci, M., Bastone, A., Sclip, A., Biasini, E., Tapella, L., Colombo, L., Manzoni, C., Borsello, T., Chiesa, R., Gobbi, M., Salmona, M. and Forloni, G. 2010. Synthetic amyloid-beta oligomers impair long-term memory independently of cellular prion protein. *Proceedings of the National Academy of Sciences of the United States of America.* **107**(5), pp.2295-2300.

Barry, A.E., Klyubin, I., Mc Donald, J.M., Mably, A.J., Farrell, M.A., Scott, M., Walsh, D.M. and Rowan, M.J. 2011. Alzheimer's Disease Brain-Derived Amyloid-beta-Mediated Inhibition of LTP In Vivo Is Prevented by Immunotargeting Cellular Prion Protein. *Journal of Neuroscience.* **31**(20), pp.7259-7263.

Baumketner, A., Bernstein, S.L., Wyttenbach, T., Bitan, G., Teplow, D.B., Bowers, M.T. and Shea, J.E. 2006. Amyloid beta-protein monomer structure: A computational and experimental study. *Protein Science.* **15**(3), pp.420-428.

Behnke, J., Feige, M.J. and Hendershot, L.M. 2015. BiP and Its Nucleotide Exchange Factors Grp170 and Sil1: Mechanisms of Action and Biological Functions. *Journal of Molecular Biology.* **427**(7), pp.1589-1608.

Bento, A.P., Gaulton, A., Hersey, A., Bellis, L.J., Chambers, J., Davies, M., Kruger, F.A., Light, Y., Mak, L., McGlinchey, S., Nowotka, M., Papadatos, G., Santos, R. and Overington, J.P. 2014. The ChEMBL bioactivity database: an update. *Nucleic Acids Research.* **42**(D1), pp.D1083-D1090.

Berendsen, H.J.C., Vanderspoel, D. and Vandrunen, R. 1995. GROMACS - A message-passing parallel molecular-dynamics implementation. *Computer Physics Communications.* **91**(1-3), pp.43-56.

Berliner, L.J., Grunwald, J., Hankovszky, H.O. and Hideg, K. 1982. A novel reversible thiol-specific spin label - papain active-site labeling and inhibition. *Analytical Biochemistry.* **119**(2), pp.450-455.

Bertolotti, A., Wang, X.Z., Novoa, I., Jungreis, R., Schlessinger, K., Cho, J.H., West, A.B. and Ron, D. 2001. Increased sensitivity to dextran sodium sulfate colitis in IRE1 beta-deficient mice. *Journal of Clinical Investigation.* **107**(5), pp.585-593.

Bertolotti, A., Zhang, Y.H., Hendershot, L.M., Harding, H.P. and Ron, D. 2000. Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. *Nature Cell Biology.* **2**(6), pp.326-332.

Biancalana, M. and Koide, S. 2010. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochimica Et Biophysica Acta-Proteins and Proteomics.* **1804**(7), pp.1405-1412.

Bich, C., Baer, S., Jecklin, M.C. and Zenobi, R. 2010. Probing the Hydrophobic Effect of Noncovalent Complexes by Mass Spectrometry. *Journal of the American Society for Mass Spectrometry.* **21**(2), pp.286-289.

Bloom, G.S. 2014. Amyloid-beta and Tau The Trigger and Bullet in Alzheimer Disease Pathogenesis. *Jama Neurology.* **71**(4), pp.505-508.

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.C., Estreicher, A., Gasteiger, E., Martin, M.J., Michoud, K., O'Donovan, C., Phan, I., Pilbout, S. and Schneider, M. 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Research.* **31**(1), pp.365-370.

Booth, L., Roberts, J.L., Cruickshanks, N., Grant, S., Poklepovic, A. and Dent, P. 2014. Regulation of OSU-03012 Toxicity by ER Stress Proteins and ER Stress-Inducing Drugs. *Molecular Cancer Therapeutics.* **13**(10), pp.2384-2398.

Campioni, S., Mannini, B., Zampagni, M., Pensalfini, A., Parrini, C., Evangelisti, E., Relini, A., Stefani, M., Dobson, C.M., Cecchi, C. and Chiti, F. 2010. A causative link between the structure of aberrant protein oligomers and their toxicity. *Nature Chemical Biology.* **6**(2), pp.140-147.

Carr, H.Y. and Purcell, E.M. 1954. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physical Review.* **94**(3), pp.630-638.

Carrara, M., Prischi, F., Nowak, P.R., Kopp, M.C. and Ali, M.M.U. 2015. Noncanonical binding of BiP ATPase domain to Ire1 and Perk is dissociated by unfolded protein C(H)1 to initiate ER stress signaling. *Elife.* **4**, p16.

Carrasco, D.R., Sukhdeo, K., Protopopova, M., Sinha, R., Enos, M., Carrasco, D.E., Zheng, M., Mani, M., Henderson, J., Pinkus, G.S., Munshi, N., Horner, J., Ivanova, E.V., Protopopov, A., Anderson, K.C., Tonon, G. and DePinho, R.A. 2007. The differentiation and stress response factor XBP-1 drives multiple myeloma pathogenesis. *Cancer Cell.* **11**(4), pp.349-360.

Chakrabarti, A., Chen, A.W. and Varner, J.D. 2011. A Review of the Mammalian Unfolded Protein Response. *Biotechnology and Bioengineering.* **108**(12), pp.2777-2793.

Chang, L., Bertelsen, E.B., Wisen, S., Larsen, E.M., Zuiderweg, E.R.P. and Gestwicki, J.E. 2008. High-throughput screen for small molecules that modulate the ATPase activity of the molecular chaperone DnaK. *Analytical Biochemistry.* **372**(2), pp.167-176.

Chen, B., Retzlaff, M., Roos, T. and Frydman, J. 2011. Cellular Strategies of Protein Quality Control. *Cold Spring Harbor Perspectives in Biology.* **3**(8), p14.

Chen, Y.N. and Brandizzi, F. 2013. IRE1: ER stress sensor and cell fate executor. *Trends in Cell Biology.* **23**(11), pp.547-555.

Choi, S.H., Kim, Y.H., Hebisch, M., Sliwinski, C., Lee, S., D'Avanzo, C., Chen, H.C., Hooli, B., Asselin, C., Muffat, J., Klee, J.B., Zhang, C., Wainger, B.J., Peitz, M., Kovacs, D.M., Woolf, C.J., Wagner, S.L., Tanzi,

R.E. and Kim, D.Y. 2014. A three-dimensional human neural cell culture model of Alzheimer's disease. *Nature.* **515**(7526), pp.274-U293.

Ciocca, D.R. and Calderwood, S.K. 2005. Heat shock proteins in cancer: diagnostic, prognostic, predictive, and treatment implications. *Cell Stress & Chaperones.* **10**(2), pp.86-103.

Clore, G.M. and Gronenborn, A.M. 1997. NMR structures of proteins and protein complexes beyond 20,000 M-r. *Nature Structural Biology.* **4**, pp.849-853.

Colvin, M.T., Silvers, R., Ni, Q.Z., Can, T.V., Sergeyev, I., Rosay, M., Donovan, K.J., Michael, B., Wall, J., Linse, S. and Griffin, R.G. 2016. Atomic Resolution Structure of Monomorphic A beta(42) Amyloid Fibrils. *Journal of the American Chemical Society.* **138**(30), pp.9663-9674.

Cornejo, V.H. and Hetz, C. 2013. The unfolded protein response in Alzheimer's disease. *Seminars in Immunopathology.* **35**(3), pp.277-292.

Credle, J.J., Finer-Moore, J.S., Papa, F.R., Stroud, R.M. and Walter, P. 2005. On the mechanism of sensing unfolded protein in the endoplasmic reticulum. *Proceedings of the National Academy of Sciences of the United States of America.* **102**(52), pp.18773-18784.

Dalvit, C. 2007. Ligand- and substrate-based F-19 NMR screening: Principles and applications to drug discovery. *Progress in Nuclear Magnetic Resonance Spectroscopy.* **51**(4), pp.243-I.

Dalvit, C., Ardini, E., Flocco, M., Fogliatto, G.P., Mongelli, N. and Veronesi, M. 2003a. A general NMR method for rapid, efficient, and reliable biochemical screening. *Journal of the American Chemical Society.* **125**(47), pp.14620-14625.

Dalvit, C., Fagerness, P.E., Hadden, D.T.A., Sarver, R.W. and Stockman, B.J. 2003b. Fluorine-NMR experiments for high-throughput screening: Theoretical aspects, practical considerations, and range of applicability. *Journal of the American Chemical Society.* **125**(25), pp.7696-7703.

Daniel, J.M., Friess, S.D., Rajagopalan, S., Wendt, S. and Zenobi, R. 2002. Quantitative determination of noncovalent binding interactions using soft ionization mass spectrometry. *International Journal of Mass Spectrometry.* **216**(1), pp.1-27.

De Strooper, B., Iwatsubo, T. and Wolfe, M.S. 2012. Presenilins and gamma-Secretase: Structure, Function, and Role in Alzheimer Disease. *Cold Spring Harbor Perspectives in Medicine.* **2**(1), p19.

Dejeans, N., Pluquet, O., Lhomond, S., Grise, F., Bouchecareilh, M., Juin, A., Meynard-Cadars, M., Bidaud-Meynard, A., Gentil, C., Moreau, V., Saltel, F. and Chevet, E. 2012. Autocrine control of glioma cells adhesion and migration through IRE1 alpha-mediated cleavage of SPARC mRNA. *Journal of Cell Science.* **125**(18), pp.4278-4287.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. 1995. NMRPipe - A multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular Nmr.* **6**(3), pp.277-293.

Dery, M.A., Jodoin, J., Ursini-Siegel, J., Aleynikova, O., Ferrario, C., Hassan, S., Basik, M. and LeBlanc, A.C. 2013. Endoplasmic reticulum stress induces PRNP prion protein gene expression in breast cancer. *Breast Cancer Research.* **15**(2), p19.

Devine, P.W.A., Fisher, H.C., Calabrese, A.N., Whelan, F., Higazi, D.R., Potts, J.R., Lowe, D.C., Radford, S.E. and Ashcroft, A.E. 2017. Investigating the Structural Compaction of Biomolecules Upon Transition

to the Gas-Phase Using ESI-TWIMS-MS. *Journal of the American Society for Mass Spectrometry.* **28**(9), pp.1855-1862.

Dror, R.O., Dirks, R.M., Grossman, J.P., Xu, H.F. and Shaw, D.E. 2012. Biomolecular Simulation: A Computational Microscope for Molecular Biology. In: Rees, D.C. ed. *Annual Review of Biophysics, Vol 41.* Palo Alto: Annual Reviews, pp.429-452.

Duran-Aniotz, C., Cornejo, V.H., Espinoza, S., Ardiles, Á.O., Medinas, D.B., Salazar, C., Foley, A., Gajardo, I., Thielen, P., Iwawaki, T., Scheper, W., Soto, C., Palacios, A.G., Hoozemans, J.J.M. and Hetz, C. 2017. IRE1 signaling exacerbates Alzheimer's disease pathogenesis. *Acta Neuropathologica.* pp.1-18.

Eletto, D., Boyle, S. and Argon, Y. 2016. PDIA6 regulates insulin secretion by selectively inhibiting the RIDD activity of IRE1. *Faseb Journal.* **30**(2), pp.653-665.

Eletto, D., Dersh, D., Gidalevitz, T. and Argon, Y. 2014. Protein Disulfide Isomerase A6 Controls the Decay of IRE1 alpha Signaling via Disulfide-Dependent Association. *Molecular Cell.* **53**(4), pp.562-576.

Endicott, J.A., Noble, M.E.M. and Johnson, L.N. 2012. The Structural Basis for Control of Eukaryotic Protein Kinases. In: Kornberg, R.D. ed. *Annual Review of Biochemistry, Vol 81.* Palo Alto: Annual Reviews, pp.587-613.

Fandrich, M. 2012. Oligomeric Intermediates in Amyloid Formation: Structure Determination and Mechanisms of Toxicity. *Journal of Molecular Biology.* **421**(4-5), pp.427-440.

Feifel, B., Sandmeier, E., Schonfeld, H.J. and Christen, P. 1996. Potassium ions and the molecular-chaperone activity of DnaK. *European Journal of Biochemistry.* **237**(1), pp.318-321.

Fekete, S., Beck, A., Veuthey, J.L. and Guillarme, D. 2014. Theory and practice of size exclusion chromatography for the analysis of protein aggregates. *Journal of Pharmaceutical and Biomedical Analysis.* **101**, pp.161-173.

Freyer, M.W. and Lewis, E.A. 2008. Isothermal titration calorimetry: Experimental design, data analysis, and probing Macromolecule/Ligand binding and kinetic interactions. In: Correia, J.J. and Detrich, H.W. eds. *Biophysical Tools for Biologists: Vol 1 in Vitro Techniques.* San Diego: Elsevier Academic Press Inc, pp.79-113.

Gao, H.J., Sun, B., Fu, H.L., Chi, X.M., Wang, F.M., Qi, X.Y., Hu, J. and Shao, S.J. 2016. PDIA6 promotes the proliferation of HeLa cells through activating the Wnt/beta-catenin signaling pathway. *Oncotarget.* **7**(33), pp.53289-53298.

Gardner, B.M. and Walter, P. 2011. Unfolded Proteins Are Ire1-Activating Ligands That Directly Induce the Unfolded Protein Response. *Science.* **333**(6051), pp.1891-1894.

Gardner, K.H. and Kay, L.E. 1998. The use of H-2, C-13, N-15 multidimensional NMR to study the structure and dynamics of proteins. *Annual Review of Biophysics and Biomolecular Structure.* **27**, pp.357-406.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D. and Bairoch, A. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research.* **31**(13), pp.3784-3788.

Gaulton, A., Bellis, L.J., Bento, A.P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B. and Overington, J.P. 2012. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Research.* **40**(D1), pp.D1100-D1107.

Gaulton, A., Hersey, A., Nowotka, M., Bento, A.P., Chambers, J., Mendez, D., Mutowo, P., Atkinson, F., Bellis, L.J., Cibrian-Uhalte, E., Davies, M., Dedman, N., Karlsson, A., Magarinos, M.P., Overington, J.P., Papadatos, G., Smit, I. and Leach, A.R. 2017. The ChEMBL database in 2017. *Nucleic Acids Research.* **45**(D1), pp.D945-D954.

Godoy-Ruiz, R., Guo, C.Y. and Tugarinov, V. 2010. Alanine Methyl Groups as NMR Probes of Molecular Structure and Dynamics in High-Molecular-Weight Proteins. *Journal of the American Chemical Society.* **132**(51), pp.18340-18350.

Goedert, M. and Spillantini, M.G. 2006. A century of Alzheimer's disease. *Science.* **314**(5800), pp.777-781.

Greenman, C., Stephens, P., Smith, R., Dalgliesh, G.L., Hunter, C., Bignell, G., Davies, H., Teague, J., Butler, A., Edkins, S., O'Meara, S., Vastrik, I., Schmidt, E.E., Avis, T., Barthorpe, S., Bhamra, G., Buck, G., Choudhury, B., Clements, J., Cole, J., Dicks, E., Forbes, S., Gray, K., Halliday, K., Harrison, R., Hills, K., Hinton, J., Jenkinson, A., Jones, D., Menzies, A., Mironenko, T., Perry, J., Raine, K., Richardson, D., Shepherd, R., Small, A., Tofts, C., Varian, J., Webb, T., West, S., Widaa, S., Yates, A., Cahill, D.P., Louis, D.N., Goldstraw, P., Nicholson, A.G., Brasseur, F., Looijenga, L., Weber, B.L., Chiew, Y.E., Defazio, A., Greaves, M.F., Green, A.R., Campbell, P., Birney, E., Easton, D.F., Chenevix-Trench, G., Tan, M.H., Khoo, S.K., Teh, B.T., Yuen, S.T., Leung, S.Y., Wooster, R., Futreal, P.A. and Stratton, M.R. 2007. Patterns of somatic mutation in human cancer genomes. *Nature.* **446**(7132), pp.153-158.

Groenendyk, J., Peng, Z.L., Dudek, E., Fan, X., Mizianty, M.J., Dufey, E., Urra, H., Sepulveda, D., Rojas-Rivera, D., Lim, Y., Kim, D.H., Baretta, K., Srikanth, S., Gwack, Y., Ahnn, J., Kaufman, R.J., Lee, S.K., Hetz, C., Kurgan, L. and Michalak, M. 2014. Interplay Between the Oxidoreductase PDIA6 and microRNA-322 Controls the Response to Disrupted Endoplasmic Reticulum Calcium Homeostasis. *Science Signaling.* **7**(329), p14.

Groenning, M. 2010. Binding mode of Thioflavin T and other molecular probes in the context of amyloid fibrils-current status. *Journal of chemical biology.* **3**(1), pp.1-18.

Gu, F., Nguyen, D.T., Stuible, M., Dube, N., Tremblay, M.L. and Chevet, E. 2004. Protein-tyrosine phosphatase 1B potentiates IRE1 signaling during endoplasmic reticulum stress. *Journal of Biological Chemistry.* **279**(48), pp.49689-49693.

Gupta, S., Deepti, A., Deegan, S., Lisbona, F., Hetz, C. and Samali, A. 2010. HSP72 Protects Cells from ER Stress-induced Apoptosis via Enhancement of IRE1 alpha-XBP1 Signaling through a Physical Interaction. *Plos Biology.* **8**(7), p15.

Habchi, J., Arosio, P., Perni, M., Costa, A.R., Yagi-Utsumi, M., Joshi, P., Chia, S., Cohen, S.I.A., Muller, M.B.D., Linse, S., Nollen, E.A.A., Dobson, C.M., Knowles, T.P.J. and Vendruscolo, M. 2016. An anticancer drug suppresses the primary nucleation reaction that initiates the production of the toxic A beta 42 aggregates linked with Alzheimer's disease. *Science Advances.* **2**(2), p13.

Hajduk, P.J., Olejniczak, E.T. and Fesik, S.W. 1997. One-dimensional relaxation- and diffusion-edited NMR methods for screening compounds that bind to macromolecules. *Journal of the American Chemical Society.* **119**(50), pp.12257-12261.

Halbleib, K., Pesek, K., Covino, R., Hofbauer, H.F., Wunnicke, D., Hanelt, I., Hummer, G. and Ernst, R. 2017. Activation of the Unfolded Protein Response by Lipid Bilayer Stress. *Molecular Cell.* **67**(4), pp.673-+.

Halliday, M. and Mallucci, G.R. 2014. Targeting the unfolded protein response in neurodegeneration: A new approach to therapy. *Neuropharmacology.* **76**, pp.169-174.

Harding, H.P., Novoa, I., Zhang, Y.H., Zeng, H.Q., Wek, R., Schapira, M. and Ron, D. 2000. Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Molecular Cell.* **6**(5), pp.1099-1108.

Hardy, J.A. and Higgins, G.A. 1992. Alzheimers-disease - The amyloid cascade hypothesis. *Science.* **256**(5054), pp.184-185.

Hassan, P.A., Rana, S. and Verma, G. 2015. Making Sense of Brownian Motion: Colloid Characterization by Dynamic Light Scattering. *Langmuir.* **31**(1), pp.3-12.

Hayashi, T. 2015. Sigma-1 receptor: The novel intracellular target of neuropsychotherapeutic drugs. *Journal of Pharmacological Sciences.* **127**(1), pp.2-5.

Heck, A.J.R. 2008. Native mass spectrometry: a bridge between interactomics and structural biology. *Nature Methods.* **5**(11), pp.927-933.

Herrup, K. 2015. The case for rejecting the amyloid cascade hypothesis. *Nature Neuroscience.* **18**(6), pp.794-799.

Hetz, C., Bernasconi, P., Fisher, J., Lee, A.H., Bassik, M.C., Antonsson, B., Brandt, G.S., Iwakoshi, N.N., Schinzel, A., Glimcher, L.H. and Korsmeyer, S.J. 2006. Proapoptotic BAX and BAK modulate the unfolded protein response by a direct interaction with IRE1 alpha. *Science.* **312**(5773), pp.572-576.

Hetz, C., Castilla, J. and Soto, C. 2007. Perturbation of endoplasmic reticulum homeostasis facilitates prion replication. *Journal of Biological Chemistry.* **282**(17), pp.12725-12733.

Hetz, C., Chevet, E. and Harding, H.P. 2013. Targeting the unfolded protein response in disease. *Nature Reviews Drug Discovery.* **12**(9), pp.703-719.

Hetz, C., Thielen, P., Matus, S., Nassif, M., Court, F., Kiffin, R., Martinez, G., Cuervo, A.M., Brown, R.H. and Glimcher, L.H. 2009. XBP-1 deficiency in the nervous system protects against amyotrophic lateral sclerosis by increasing autophagy. *Genes & Development.* **23**(19), pp.2294-2306.

Hollien, J. and Weissman, J.S. 2006. Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response. *Science.* **313**(5783), pp.104-107.

Hoozemans, J.J.M., van Haastert, E.S., Nijholt, D.A.T., Rozemuller, A.J.M., Eikelenboom, P. and Scheper, W. 2009. The Unfolded Protein Response Is Activated in Pretangle Neurons in Alzheimer's Disease Hippocampus. *American Journal of Pathology.* **174**(4), pp.1241-1251.

Hudson, S.A., Ecroyd, H., Kee, T.W. and Carver, J.A. 2009. The thioflavin T fluorescence assay for amyloid fibril detection can be biased by the presence of exogenous compounds. *Febs Journal.* **276**(20), pp.5960-5972.

Ilieva, E.V., Ayala, V., Jove, M., Dalfo, E., Cacabelos, D., Povedano, M., Bellmunt, M.J., Ferrer, I., Pamplona, R. and Portero-Otin, M. 2007. Oxidative and endoplasmic reticulum stress interplay in sporadic amyotrophic lateral sclerosis. *Brain.* **130**, pp.3111-3123.

Irwin, J.J., Sterling, T., Mysinger, M.M., Bolstad, E.S. and Coleman, R.G. 2012. ZINC: A Free Tool to Discover Chemistry for Biology. *Journal of Chemical Information and Modeling.* **52**(7), pp.1757-1768.

Jameson, L.P., Smith, N.W. and Dzyuba, S.V. 2012. Dye-Binding Assays for Evaluation of the Effects of Small Molecule Inhibitors on Amyloid (A beta) Self-Assembly. *Acs Chemical Neuroscience.* **3**(11), pp.807-819.

Jarosz-Griffiths, H.H., Noble, E., Rushworth, J.V. and Hooper, N.M. 2016. Amyloid- Receptors: The Good, the Bad, and the Prion Protein. *Journal of Biological Chemistry.* **291**(7), pp.3174-3183.

Jerabek-Willemsen, M., Andre, T., Wanner, R., Roth, H.M., Duhr, S., Baaske, P. and Breitsprecher, D. 2014. MicroScale Thermophoresis: Interaction analysis and beyond. *Journal of Molecular Structure.* **1077**, pp.101-113.

Jessop, C.E., Watkins, R.H., Simmons, J.J., Tasab, M. and Bulleid, N.J. 2009. Protein disulphide isomerase family members show distinct substrate specificity: P5 is targeted to BiP client proteins. *Journal of Cell Science.* **122**(23), pp.4287-4295.

Jhingree, J.R., Bellina, B., Pacholarz, K.J. and Barran, P.E. 2017. Charge Mediated Compaction and Rearrangement of Gas-Phase Proteins: A Case Study Considering Two Proteins at Opposing Ends of the Structure-Disorder Continuum. *Journal of the American Society for Mass Spectrometry.* **28**(7), pp.1450-1461.

Jhoti, H., Williams, G., Rees, D.C. and Murray, C.W. 2013. The 'rule of three' for fragment-based drug discovery: where are we now? *Nature Reviews Drug Discovery.* **12**(8), pp.644-+.

Jordan, J.B., Poppe, L., Xia, X.Y., Cheng, A.C., Sun, Y., Michelsen, K., Eastwood, H., Schnier, P.D., Nixey, T. and Zhong, W.G. 2012. Fragment Based Drug Discovery: Practical Implementation Based on F-19 NMR Spectroscopy. *Journal of Medicinal Chemistry.* **55**(2), pp.678-687.

Jordan, P.A., Stevens, J.M., Hubbard, G.P., Barrett, N.E., Sage, T., Authi, K.S. and Gibbins, J.M. 2005. A role for the thiol isomerase. protein ERP5 in platelet function. *Blood.* **105**(4), pp.1500-1507.

Joshi, P., Chia, S., Habchi, J., Knowles, T.P.J., Dobson, C.M. and Vendruscolo, M. 2016. A Fragment-Based Method of Creating Small-Molecule Libraries to Target the Aggregation of Intrinsically Disordered Proteins. *Acs Combinatorial Science.* **18**(3), pp.144-153.

Kaiser, B.K., Yim, D., Chow, I.T., Gonzalez, S., Dai, Z., Mann, H.H., Strong, R.K., Groh, V. and Spies, T. 2007. Disulphide-isomerase-enabled shedding of tumour-associated NKG2D ligands. *Nature.* **447**(7143), pp.482-U485.

Kampinga, H.H. and Craig, E.A. 2010. `The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature Reviews Molecular Cell Biology.* **11**(8), pp.579-592.

Karagoz, G.E., Acosta-Alvear, D., Nguyen, H.T., Lee, C.P., Chu, F.X. and Walter, P. 2017. An unfolded protein-induced conformational switch activates mammalian IRE1. *Elife.* **6**, p29.

Karplus, M. and McCammon, J.A. 2002. Molecular dynamics simulations of biomolecules. *Nature Structural Biology.* **9**(9), pp.646-652.

Katayama, T., Imaizumi, K., Honda, A., Yoneda, T., Kudo, T., Takeda, M., Mori, K., Rozmahel, R., Fraser, P., St George-Hyslop, P. and Tohyama, M. 2001. Disturbed activation of endoplasmic reticulum stress transducers by familial Alzheimer's disease-linked presenilin-1 mutations. *Journal of Biological Chemistry.* **276**(46), pp.43446-43454.

Katayama, T., Imaizumi, K., Sato, N., Miyoshi, K., Kudo, T., Hitomi, J., Morihara, T., Yoneda, T., Gomi, F., Mori, Y., Nakano, Y., Takeda, J., Tsuda, T., Itoyama, Y., Murayama, O., Takashima, A., St George-Hyslop, P., Takeda, M. and Tohyama, M. 1999. Presenilin-1 mutations downregulate the signalling pathway of the unfolded-protein response. *Nature Cell Biology.* **1**(8), pp.479-485.

Kellett, K.A.B. and Hooper, N.M. 2009. Prion protein and Alzheimer disease. *Prion.* **3**(4), pp.190-194.

Kim, S., Thiessen, P.A., Bolton, E.E., Chen, J., Fu, G., Gindulyte, A., Han, L.Y., He, J.E., He, S.Q., Shoemaker, B.A., Wang, J.Y., Yu, B., Zhang, J. and Bryant, S.H. 2016. PubChem Substance and Compound databases. *Nucleic Acids Research.* **44**(D1), pp.D1202-D1213.

Kimata, Y., Oikawa, D., Shimizu, Y., Ishiwata-Kimata, Y. and Kohno, K. 2004. A role for BiP as an adjustor for the endoplasmic reticulum stress-sensing protein Ire 1. *Journal of Cell Biology.* **167**(3), pp.445-456.

Kleckner, I.R. and Foster, M.P. 2011. An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica Et Biophysica Acta-Proteins and Proteomics.* **1814**(8), pp.942-968.

Korennykh, A. and Walter, P. 2012. Structural Basis of the Unfolded Protein Response. In: Schekman, R. ed. *Annual Review of Cell and Developmental Biology, Vol 28.* Palo Alto: Annual Reviews, pp.251-277.

Korennykh, A.V., Egea, P.F., Korostelev, A.A., Finer-Moore, J., Stroud, R.M., Zhang, C., Shokat, K.M. and Walter, P. 2011. Cofactor-mediated conformational control in the bifunctional kinase/RNase Ire1. *Bmc Biology.* **9**, p16.

Korennykh, A.V., Egea, P.F., Korostelev, A.A., Finer-Moore, J., Zhang, C., Shokat, K.M., Stroud, R.M. and Walter, P. 2009. The unfolded protein response signals through high-order assembly of Ire1. *Nature.* **457**(7230), pp.687-U682.

Lanucara, F., Holman, S.W., Gray, C.J. and Eyers, C.E. 2014. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. *Nature Chemistry.* **6**(4), pp.281-294.

Lauren, J., Gimbel, D.A., Nygaard, H.B., Gilbert, J.W. and Strittmatter, S.M. 2009. Cellular prion protein mediates impairment of synaptic plasticity by amyloid-beta oligomers. *Nature.* **457**(7233), pp.1128-U1184.

Lee, K.P.K., Dey, M., Neculai, D., Cao, C., Dever, T.E. and Sicheri, F. 2008. Structure of the dual enzyme ire1 reveals the basis for catalysis and regulation in nonconventional RNA splicing. *Cell.* **132**(1), pp.89-100.

Lescop, E., Schanda, P. and Brutscher, B. 2007. A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *Journal of Magnetic Resonance.* **187**(1), pp.163-169.

Li, H., Korennykh, A.V., Behrman, S.L. and Walter, P. 2010. Mammalian endoplasmic reticulum stress sensor IRE1 signals by dynamic clustering. *Proceedings of the National Academy of Sciences of the United States of America.* **107**(37), pp.16113-16118.

Liao, Y.D., Jeng, J.C., Wang, C.F., Wang, S.C. and Chang, S.T. 2004. Removal of N-terminal methionine from recombinant proteins by engineered E-coli methionine aminopeptidase. *Protein Science.* **13**(7), pp.1802-1810.

Lipson, K.L., Ghosh, R. and Urano, F. 2008. The Role of IRE1 alpha in the Degradation of Insulin mRNA in Pancreatic beta-Cells. *Plos One.* **3**(2), p7.

Lisbona, F., Rojas-Rivera, D., Thielen, P., Zamorano, S., Todd, D., Martinon, F., Glavic, A., Kress, C., Lin, J.H., Walter, P., Reed, J.C., Glimcher, L.H. and Hetz, C. 2009. BAX Inhibitor-1 Is a Negative Regulator of the ER Stress Sensor IRE1 alpha. *Molecular Cell.* **33**(6), pp.679-691.

Liu, C.Y., Wong, H.N., Schauerte, J.A. and Kaufman, R.J. 2002. The protein kinase/endoribonuclease IRE1 alpha that signals the unfolded protein response has a luminal N-terminal ligand-independent dimerization domain. *Journal of Biological Chemistry.* **277**(21), pp.18346-18356.

Liu, C.Y., Xu, Z.H. and Kaufman, R.J. 2003. Structure and intermolecular interactions of the luminal dimerization domain of human IRE1 alpha. *Journal of Biological Chemistry.* **278**(20), pp.17680-17687.

Lomonosova, E. and Chinnadurai, G. 2008. BH3-only proteins in apoptosis and beyond: an overview. *Oncogene.* **27**, pp.S2-S19.

Loo, J.A. 1997. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrometry Reviews.* **16**(1), pp.1-23.

Loo, J.A., Berhane, B., Kaddis, C.S., Wooding, K.M., Xie, Y.M., Kaufman, S.L. and Chernushevich, I.V. 2005. Electrospray ionization mass spectrometry and ion mobility analysis of the 20S proteasome complex. *Journal of the American Society for Mass Spectrometry.* **16**(7), pp.998-1008.

Loria, J.P., Rance, M. and Palmer, A.G. 1999. A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. *Journal of the American Chemical Society.* **121**(10), pp.2331-2332.

Louis-Jeune, C., Andrade-Navarro, M.A. and Perez-Iratxeta, C. 2012. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins-Structure Function and Bioinformatics.* **80**(2), pp.374-381.

Luo, D.H., He, Y., Zhang, H.F., Yu, L.Y., Chen, H., Xu, Z., Tang, S.B., Urano, F. and Min, W. 2008. AIP1 is critical in transducing IRE1-mediated endoplasmic reticulum stress response. *Journal of Biological Chemistry.* **283**(18), pp.11905-11912.

Mann, M., Hendrickson, R.C. and Pandey, A. 2001. Analysis of proteins and proteomes by mass spectrometry. *Annual Review of Biochemistry.* **70**, pp.437-473.

Mannini, B., Mulvihill, E., Sgromo, C., Cascella, R., Khodarahmi, R., Ramazzotti, M., Dobson, C.M., Cecchi, C. and Chiti, F. 2014. Toxicity of Protein Oligomers Is Rationalized by a Function Combining Size and Surface Hydrophobicity. *Acs Chemical Biology.* **9**(10), pp.2309-2317.

Marcinowski, M., Holler, M., Feige, M.J., Baerend, D., Lamb, D.C. and Buchner, J. 2011. Substrate discrimination of the chaperone BiP by autonomous and cochaperone-regulated conformational transitions. *Nature Structural & Molecular Biology.* **18**(2), pp.150-U210.

Marion, D. 2013. An Introduction to Biological NMR Spectroscopy. *Molecular & Cellular Proteomics.* **12**(11), pp.3006-3025.

Marklund, E.G., Degiacomi, M.T., Robinson, C.V., Baldwin, A.J. and Benesch, J.L.P. 2015. Collision Cross Sections for Structural Proteomics. *Structure.* **23**(4), pp.791-799.

Marsh, J.A. and Forman-Kay, J.D. 2010. Sequence Determinants of Compaction in Intrinsically Disordered Proteins. *Biophysical Journal.* **98**(10), pp.2383-2390.

Massi, F., Grey, M.J. and Palmer, A.G. 2005. Microsecond timescale backbone conformational dynamics in ubiquitin studied with NMR R-1p relaxation experiments. *Protein Science.* **14**(3), pp.735-742.

Mayer, M.P. 2013. Hsp70 chaperone dynamics and molecular mechanism. *Trends in Biochemical Sciences.* **38**(10), pp.507-514.

McColl, G., Roberts, B.R., Pukala, T.L., Kenche, V.B., Roberts, C.M., Link, C.D., Ryan, T.M., Masters, C.L., Barnham, K.J., Bush, A.I. and Cherny, R.A. 2012. Utility of an improved model of amyloid-beta (A beta(1-42)) toxicity in Caenorhabditis elegans for drug screening for Alzheimer's disease. *Molecular Neurodegeneration.* **7**, p9.

Meiboom, S. and Gill, D. 1958. Modified spin-echo method for measuring nuclear relaxation times. *Review of Scientific Instruments.* **29**(8), pp.688-691.

Meyer, B. and Peters, T. 2003. NMR Spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angewandte Chemie-International Edition.* **42**(8), pp.864-890.

Mittermaier, A. and Kay, L.E. 2006. Review - New tools provide new insights in NMR studies of protein dynamics. *Science.* **312**(5771), pp.224-228.

Mori, T., Hayashi, T., Hayashi, E. and Su, T.P. 2013. Sigma-1 Receptor Chaperone at the ER-Mitochondrion Interface Mediates the Mitochondrion-ER-Nucleus Signaling for Cellular Survival. *Plos One.* **8**(10), p13.

Morris, G.A. and Freeman, R. 1979. Enhancement of nuclear magnetic-resonance signals by polarization transfer. *Journal of the American Chemical Society.* **101**(3), pp.760-762.

Moschen, T., Grutsch, S., Juen, M.A., Wunderlich, C.H., Kreutz, C. and Tollinger, M. 2016. Measurement of Ligand-Target Residence Times by H-1 Relaxation Dispersion NMR Spectroscopy. *Journal of Medicinal Chemistry.* **59**(23), pp.10788-10793.

Nagata, K. 1996. Hsp47: A collagen-specific molecular chaperone. *Trends in Biochemical Sciences.* **21**(1), pp.23-26.

Nijholt, D.A.T., van Haastert, E.S., Rozemuller, A.J.M., Scheper, W. and Hoozemans, J.J.M. 2012. The unfolded protein response is associated with early tau pathology in the hippocampus of tauopathies. *Journal of Pathology.* **226**(5), pp.693-702.

Nishitoh, H., Matsuzawa, A., Tobiume, K., Saegusa, K., Takeda, K., Inoue, K., Hori, S., Kakizuka, A. and Ichijo, H. 2002. ASK1 is essential for endoplasmic reticulum stress-induced neuronal cell death triggered by expanded polyglutamine repeats. *Genes & Development.* **16**(11), pp.1345-1355.

Oikawa, D., Kimata, Y., Kohno, K. and Iwawaki, T. 2009. Activation of mammalian IRE1 alpha upon ER stress depends on dissociation of BiP rather than on direct interaction with unfolded proteins. *Experimental Cell Research.* **315**(15), pp.2496-2504.

Okamoto, K., Hirai, S., Iizuka, T., Yanagisawa, T. and Watanabe, M. 1991. Reexamination of granulovacuolar degeneration. *Acta Neuropathologica.* **82**(5), pp.340-345.

Orsi, A., Fioriti, L., Chiesa, R. and Sitia, R. 2006. Conditions of endoplasmic reticulum stress favor the accumulation of cytosolic prion protein. *Journal of Biological Chemistry.* **281**(41), pp.30431-30438.

Ortega-Roldan, J.L., Ossa, F. and Schnell, J.R. 2013. Characterization of the Human Sigma-1 Receptor Chaperone Domain Structure and Binding Immunoglobulin Protein (BiP) Interactions. *Journal of Biological Chemistry.* **288**(29), pp.21448-21457.

Pagel, K., Natan, E., Hall, Z., Fersht, A.R. and Robinson, C.V. 2013. Intrinsically Disordered p53 and Its Complexes Populate Compact Conformations in the Gas Phase. *Angewandte Chemie-International Edition.* **52**(1), pp.361-365.

Palleros, D.R., Reid, K.L., Shi, L., Welch, W.J. and Fink, A.L. 1993. ATP-induced protein HSP70 complex dissociation requires K+ but not ATP hydrolysis. *Nature.* **365**(6447), pp.664-666.

Pincus, D., Chevalier, M.W., Aragon, T., van Anken, E., Vidal, S.E., El-Samad, H. and Walter, P. 2010. BiP Binding to the ER-Stress Sensor Ire1 Tunes the Homeostatic Behavior of the Unfolded Protein Response. *Plos Biology.* **8**(7), p14.

Placido, A.I., Pereira, C.M.F., Duarte, A.I., Candeias, E., Correia, S.C., Santos, R.X., Carvalho, C., Cardoso, S., Oliveira, C.R. and Moreira, P.I. 2014. The role of endoplasmic reticulum in amyloid precursor protein processing and trafficking: Implications for Alzheimer's disease. *Biochimica Et Biophysica Acta-Molecular Basis of Disease.* **1842**(9), pp.1444-1453.

Pollard, T.D. 2010. A Guide to Simple and Informative Binding Assays. *Molecular Biology of the Cell.* **21**(23), pp.4061-4067.

Pringle, S.D., Giles, K., Wildgoose, J.L., Williams, J.P., Slade, S.E., Thalassinos, K., Bateman, R.H., Bowers, M.T. and Scrivens, J.H. 2007. An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *International Journal of Mass Spectrometry.* **261**(1), pp.1-12.

Prischi, F., Nowak, P.R., Carrara, M. and Ali, M.M.U. 2014. Phosphoregulation of Ire1 RNase splicing activity. *Nature Communications.* **5**, p10.

Qiu, C., Kivipelto, M. and von Strauss, E. 2009. Epidemiology of Alzheimer's disease: occurrence, determinants, and strategies toward intervention. *Dialogues in clinical neuroscience.* **11**(2), pp.111-128.

Radford, H., Moreno, J.A., Verity, N., Halliday, M. and Mallucci, G.R. 2015. PERK inhibition prevents tau-mediated neurodegeneration in a mouse model of frontotemporal dementia. *Acta Neuropathologica.* **130**(5), pp.633-642.

Rambaran, R.N. and Serpell, L.C. 2008. Amyloid fibrils Abnormal protein assembly. *Prion.* **2**(3), pp.112-117.

Ramos, F.S., Serino, L.T.R., Carvalho, C.M.S., Lima, R.S., Urban, C.A., Cavalli, I.J. and Ribeiro, E. 2015. PDIA3 and PDIA6 gene expression as an aggressiveness marker in primary ductal breast cancer. *Genetics and Molecular Research.* **14**(2), pp.6960-6967.

Resende, R., Ferreiro, E., Pereira, C. and Oliveira, C.R. 2008. ER stress is involved in A beta-induced GSK-3 beta activation and tau phosphorylation. *Journal of Neuroscience Research.* **86**(9), pp.2091-2099.

Robberecht, W. and Philips, T. 2013. The changing scene of amyotrophic lateral sclerosis. *Nature Reviews Neuroscience.* **14**(4), pp.248-264.

Rodriguez, D.A., Zamorano, S., Lisbona, F., Rojas-Rivera, D., Urra, H., Cubillos-Ruiz, J.R., Armisen, R., Henriquez, D.R., Cheng, E.H., Letek, M., Vaisar, T., Irrazabal, T., Gonzalez-Billault, C., Letai, A., Pimentel-Muinos, F.X., Kroemer, G. and Hetz, C. 2012. BH3-only proteins are part of a regulatory network that control the sustained signalling of the unfolded protein response sensor IRE1 alpha. *Embo Journal.* **31**(10), pp.2322-2335.

Romero, P., Obradovic, Z., Li, X.H., Garner, E.C., Brown, C.J. and Dunker, A.K. 2001. Sequence complexity of disordered protein. *Proteins-Structure Function and Genetics.* **42**(1), pp.38-48.

Sanches, M., Duffy, N.M., Talukdar, M., Thevakumaran, N., Chiovitti, D., Canny, M.D., Lee, K., Kurinov, I., Uehling, D., Al-awar, R., Poda, G., Prakesch, M., Wilson, B., Tam, V., Schweitzer, C., Toro, A., Lucas, J.L., Vuga, D., Lehmann, L., Durocher, D., Zeng, Q.P., Patterson, J.B. and Sicheri, F. 2014. Structure and mechanism of action of the hydroxy-aryl-aldehyde class of IRE1 endoribonuclease inhibitors. *Nature Communications.* **5**, p16.

Sattler, M. and Fesik, S.W. 1996. Use of deuterium labeling in NMR: Overcoming a sizeable problem. *Structure.* **4**(11), pp.1245-1249.

Schanda, P. and Brutscher, B. 2005. Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *Journal of the American Chemical Society.* **127**(22), pp.8014-8015.

Scheuner, D. and Kaufman, R.J. 2008. The unfolded protein response: A pathway that links insulin demand with beta-cell failure and diabetes. *Endocrine Reviews.* **29**(3), pp.317-333.

Schlecht, R., Erbse, A.H., Bukau, B. and Mayer, M.P. 2011. Mechanics of Hsp70 chaperones enables differential interaction with client proteins. *Nature Structural & Molecular Biology.* **18**(3), pp.345-U135.

Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods.* **9**(7), pp.671-675.

Schulte-Herbruggen, T. and Sorensen, O.W. 2000. Clean TROSY: Compensation for relaxation-induced artifacts. *Journal of Magnetic Resonance.* **144**(1), pp.123-128.

Selkoe, D.J. and Hardy, J. 2016. The amyloid hypothesis of Alzheimer's disease at 25years. *Embo Molecular Medicine.* **8**(6), pp.595-608.

Sepulveda, D., Rojas-Rivera, D., Rodriguez, D.A., Groenendyk, J., Kohler, A., Lebeaupin, C., Ito, S., Urra, H., Carreras-Sureda, A., Hazari, Y., Vasseur-Cognet, M., Ali, M.M.U., Chevet, E., Campos, G., Godoy, P., Vaisar, T., Bailly-Maitre, B., Nagata, K., Michalak, M., Sierralta, J. and Hetz, C. 2018. Interactome Screening Identifies the ER Luminal Chaperone Hsp47 as a Regulator of the Unfolded Protein Response Transducer IRE1 alpha. *Molecular Cell.* **69**(2), pp.238-+.

Shvartsburg, A.A. and Smith, R.D. 2008. Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Analytical Chemistry.* **80**(24), pp.9689-9699.

Smith, D.P., Knapman, T.W., Campuzano, I., Malham, R.W., Berryman, J.T., Radford, S.E. and Ashcroft, A.E. 2009. Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies. *European Journal of Mass Spectrometry.* **15**(2), pp.113-130.

Stark, J.L., Eghbalnia, H.R., Lee, W., Westler, W.M. and Markley, J.L. 2016. NMRmix: A Tool for the Optimization of Compound Mixtures in 1D H-1 NMR Ligand Affinity Screens. *Journal of Proteome Research.* **15**(4), pp.1360-1368.

Stetefeld, J., McKenna, S.A. and Patel, T.R. 2016. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophysical reviews.* **8**(4), pp.409-427.

Stine, W.B., Dahlgren, K.N., Krafft, G.A. and LaDu, M.J. 2003. In vitro characterization of conditions for amyloid-beta peptide oligomerization and fibrillogenesis. *Journal of Biological Chemistry.* **278**(13), pp.11612-11622.

Tang, C., Ghirlando, R. and Clore, G.M. 2008a. Visualization of transient ultra-weak protein self-association in solution using paramagnetic relaxation enhancement. *Journal of the American Chemical Society.* **130**(12), pp.4048-4056.

Tang, C., Louis, J.M., Aniana, A., Suh, J.Y. and Clore, G.M. 2008b. Visualizing transient events in amino-terminal autoprocessing of HIV-1 protease. *Nature.* **455**(7213), pp.693-U692.

Tay, K.H., Luan, Q., Croft, A., Jiang, C.C., Jin, L., Zhang, X.D. and Tseng, H.Y. 2014. Sustained IRE1 and ATF6 signaling is important for survival of melanoma cells undergoing ER stress. *Cellular Signalling.* **26**(2), pp.287-294.

Tirasophon, W., Welihinda, A.A. and Kaufman, R.J. 1998. A stress response pathway from the endoplasmic reticulum to the nucleus requires a novel bifunctional protein kinase/endoribonuclease (Ire1p) in mammalian cells. *Genes & Development.* **12**(12), pp.1812-1824.

Todd-Corlett, A., Jones, E., Seghers, C. and Gething, M.J. 2007. Lobe IB of the ATPase domain of Kar2p/BiP interacts with Ire1p to negatively regulate the unfolded protein response in Saccharomyces cerevisiae. *Journal of Molecular Biology.* **367**(3), pp.770-787.

Tonks, N.K. 2006. Protein tyrosine phosphatases: from genes, to function, to disease. *Nature Reviews Molecular Cell Biology.* **7**(11), pp.833-846.

Tropea, J.E., Cherry, S. and Waugh, D.S. 2009. Expression and Purification of Soluble His(6)-Tagged TEV Protease. In: Doyle, S.A. ed. *Methods in Molecular Biology.* Humana Press Inc, 999 Riverview Dr, Ste 208, Totowa, Nj 07512-1165 USA, pp.297-307.

Tufo, G., Jones, A.W.E., Wang, Z., Hamelin, J., Tajeddine, N., Esposti, D.D., Martel, C., Boursier, C., Gallerne, C., Migdal, C., Lemaire, C., Szabadkai, G., Lemoine, A., Kroemer, G. and Brenner, C. 2014. The protein disulfide isomerases PDIA4 and PDIA6 mediate resistance to cisplatin-induced cell death in lung adenocarcinoma. *Cell Death and Differentiation.* **21**(5), pp.685-695.

Tugarinov, V., Kanelis, V. and Kay, L.E. 2006. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nature Protocols.* **1**(2), pp.749-754.

Tugarinov, V. and Kay, L.E. 2005. Methyl groups as probes of structure and dynamics in NMR studies of high-molecular-weight proteins. *Chembiochem.* **6**(9), pp.1567-+.

Tzakos, A.G., Grace, C.R.R., Lukavsky, P.J. and Riek, R. 2006. NMR techniques for very large proteins and RNAs in solution. *Annual Review of Biophysics and Biomolecular Structure.* Palo Alto: Annual Reviews, pp.319-342.

Urano, F., Wang, X.Z., Bertolotti, A., Zhang, Y.H., Chung, P., Harding, H.P. and Ron, D. 2000. Coupling of stress in the ER to activation of JNK protein kinases by transmembrane protein kinase IRE1. *Science.* **287**(5453), pp.664-666.

Vekich, J.A., Belmont, P.J., Thuerauf, D.J. and Glembotski, C.C. 2012. Protein disulfide isomerase-associated 6 is an ATF6-inducible ER stress response protein that protects cardiac myocytes from ischemia/reperfusion-mediated cell death. *Journal of Molecular and Cellular Cardiology.* **53**(2), pp.259-267.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, P., Ulrich, E.L., Markley, J.L., Ionides, J. and Laue, E.D. 2005. The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins-Structure Function and Bioinformatics.* **59**(4), pp.687-696.

Walter, P. and Ron, D. 2011. The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. *Science.* **334**(6059), pp.1081-1086.

Walti, M.A., Ravotti, F., Arai, H., Glabe, C.G., Wall, J.S., Bockmann, A., Guntert, P., Meier, B.H. and Riek, R. 2016. Atomic-resolution structure of a disease-relevant A beta(1-42) amyloid fibril. *Proceedings of the National Academy of Sciences of the United States of America.* **113**(34), pp.E4976-E4984.

Wang, S.Y. and Kaufman, R.J. 2012. The impact of the unfolded protein response on human disease. *Journal of Cell Biology.* **197**(7), pp.857-867.

Waters, N.J., Jones, R., Williams, G. and Sohal, B. 2008. Validation of a rapid equilibrium dialysis approach for the measurement of plasma protein binding. *Journal of Pharmaceutical Sciences.* **97**(10), pp.4586-4595.

Wei, J.Y. and Hendershot, L.M. 1995. Characterization of the nucleotide-binding properties and ATPase activity of recombinant hamster BiP purified from bacteria. *Journal of Biological Chemistry.* **270**(44), pp.26670-26676.

Weissmann, C. 2004. The state of the prion. *Nature Reviews Microbiology.* **2**(11), pp.861-871.

Wienken, C.J., Baaske, P., Rothbauer, U., Braun, D. and Duhr, S. 2010. Protein-binding assays in biological liquids using microscale thermophoresis. *Nature Communications.* **1**, p7.

Wieteska, L., Shahidi, S. and Zhuravleva, A. 2017. Allosteric fine-tuning of the conformational equilibrium poises the chaperone BiP for post-translational regulation. *Elife.* **6**, p20.

Wilkinson, B. and Gilbert, H.F. 2004. Protein disulfide isomerase. *Biochimica Et Biophysica Acta-Proteins and Proteomics.* **1699**(1-2), pp.35-44.

Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y.F., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., Pon, A., Knox, C. and Wilson, M. 2018. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Research.* **46**(D1), pp.D1074-D1082.

Woodard, J.S. 1962. Clinicopathologic Significance of Granulovaculoar Degeneration in Alzheimers Disease. *Journal of Neuropathology and Experimental Neurology.* **21**(1), pp.85-&.

Woods, L.A., Radford, S.E. and Ashcroft, A.E. 2013. Advances in ion mobility spectrometry-mass spectrometry reveal key insights into amyloid assembly. *Biochimica Et Biophysica Acta-Proteins and Proteomics.* **1834**(6), pp.1257-1268.

Yang, J., Nune, M., Zong, Y.N., Zhou, L. and Liu, Q.L. 2015. Close and Allosteric Opening of the Polypeptide-Binding Site in a Human Hsp70 Chaperone BiP. *Structure.* **23**(12), pp.2191-2203.

Yin, S., Xie, Y.M. and Loo, J.A. 2008. Mass spectrometry of protein-ligand complexes: Enhanced gas-phase stability of ribonuclease-nucleotide complexes. *Journal of the American Society for Mass Spectrometry.* **19**(8), pp.1199-1208.

Yoshida, H., Matsui, T., Yamamoto, A., Okada, T. and Mori, K. 2001. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell.* **107**(7), pp.881-891.

Zhang, R., He, X.R., Liu, W.M., Lu, M., Hsieh, J.T. and Min, W. 2003. AIP1 mediates TNF-alpha-induced ASK1 activation by facilitating dissociation of ASK1 from its inhibitor 14-3-3. *Journal of Clinical Investigation.* **111**(12), pp.1933-1943.

Zhang, Y. 2008. I-TASSER server for protein 3D structure prediction. *Bmc Bioinformatics.* **9**, p8.

Zhou, J.H., Liu, C.Y., Back, S.H., Clark, R.L., Peisach, D., Xu, Z.H. and Kaufman, R.J. 2006. The crystal structure of human IRE1 luminal domain reveals a conserved dimerization interface required for activation of the unfolded protein response. *Proceedings of the National Academy of Sciences of the United States of America.* **103**(39), pp.14343-14348.

Zou, J.Y., Guo, Y.L., Guettouche, T., Smith, D.F. and Voellmy, R. 1998. Repression of heat shock transcription factor HSF1 activation by HSP90 (HSP90 complex) that forms a stress-sensitive complex with HSF1. *Cell.* **94**(4), pp.471-480.

9 Appendix

9.1 Primers (all 5' → 3')

9.1.1 Subcloning primers (LIC overhangs highlighted in bold)

Forward Primer

TACTTCCAATCCAATGCA AGC ACA GTG ACG CTT CCT

Reverse Primer

TTATCCACTTCCAATGTTATTA GCT CAG GAT GAT GGT AGC CAT G

9.1.2 Mutagenesis primers

C109S (forward)

GGTGCAGGCATCCCCAAGCCGAAGTTCA

C109S (reverse)

TGAACTTCGGCTTGGGGATGCCTGCACC

C148S (forward)

CCTTTGCAGATAGTCTCAGCCCATCAACCTCTCTT

C148S (reverse)

AAGAGAGGTTGATGGGCTGAGACTATCTGCAAAGG

C332S (forward)

GGACAAGGGGGAGAGTGTGATCACGCC

C332S (reverse)

GGCGTGATCACACTCTCCCCCTTGTCC

I391* (forward)

TACCCAAACATCGGGAAAATGTGTAGCCTGCTGATTCAGAGAAAAAGAG

I391* (reverse)

CTCTTTTTCTCTGAATCAGCAGGCTACACATTTTCCCGATGTTTGGGTA

D123P (forward)

AGGTCAATAACATACCAGATGGGCTGCTTTTTACCCATGTAGAG

D123P (reverse)

CTCTACATGGGTAAAAAGCAGCCCATCTGGTATGTTATTGACCT

9.1.3 Sequencing primers

IRE1 (forward)

GGGGACGTCCTGTGGATCC

IRE1 (reverse)

GTGGAAGTACCCGTTCCCCAAGG

T7 Promoter (Forward)

TAATACGACTCACTATAGGG

T7 Terminator (Reverse)

GCTAGTTATTGCTCAGCGG

9.2 Molecular Dynamics Simulation Parameters

9.2.1 Em.mdp (Energy minimisation parameters, pre-simulation)

```
; minim.mdp - used as input into grompp to generate em.tpr

; Parameters describing what to do, when to stop and what to save

integrator      = steep         ; Algorithm (steep = steepest descent minimization)

emtol           = 1000.0        ; Stop minimization when the maximum force < 1000.0 kJ/mol/nm

emstep          = 0.01          ; Energy step size

nsteps          = 50000         ; Maximum number of (minimization) steps to perform
```

; Parameters describing how to find the neighbors of each atom and how to calculate the interactions

nstlist = 1 ; Frequency to update the neighbor list and long range forces

ns_type = grid ; Method to determine neighbor list (simple, grid)

rlist = 1.0 ; Cut-off for making neighbor list (short range forces)

coulombtype = PME ; Treatment of long range electrostatic interactions

rcoulomb = 1.0 ; Short-range electrostatic cut-off

rvdw = 1.0 ; Short-range Van der Waals cut-off

pbc = xyz ; Periodic Boundary Conditions (yes/no)

9.2.2 ions.mdp – Parameter file for the addition of counter ions to the system

; ions.mdp - used as input into grompp to generate ions.tpr

; Parameters describing what to do, when to stop and what to save

integrator = steep ; Algorithm (steep = steepest descent minimization)

emtol = 1000.0 ; Stop minimization when the maximum force < 1000.0 kJ/mol/nm

emstep = 0.01 ; Energy step size

nsteps = 50000 ; Maximum number of (minimization) steps to perform

; Parameters describing how to find the neighbors of each atom and how to calculate the interactions

nstlist = 1 ; Frequency to update the neighbor list and long range forces

ns_type = grid ; Method to determine neighbor list (simple, grid)

rlist = 1.2 ; Cut-off for making neighbor list (short range forces)

coulombtype = PME ; Treatment of long range electrostatic interactions

```
rcoulomb       = 1.2          ; Short-range electrostatic cut-off

rvdw           = 1.2          ; Short-range Van der Waals cut-off

pbc            = xyz          ; Periodic Boundary Conditions (yes/no)
```

9.2.3 md.mdp – Parameters for full MD simulation

```
; Run parameters

integrator     = md           ; leap-frog integrator

nsteps         = 250000000            ; 2 * 2500000 = 500ns

dt             = 0.002        ; 2 fs

; Output control

nstxout        = 500          ; save coordinates every 1 ps

nstvout        = 0            ; save velocities every 1 ps

nstxtcout      = 0

nstfout        = 0

nstenergy      = 100          ; save energies every 0.2 ps

nstlog         = 500          ; update log file every 1 ps

; Bond parameters

continuation    = yes          ; Restarting after NVT

constraint_algorithm = lincs    ; holonomic constraints

constraints     = all-bonds     ; all bonds (even heavy atom-H bonds) constrained

lincs_iter      = 1            ; accuracy of LINCS

lincs_order     = 4            ; also related to accuracy

; Neighborsearching

ns_type        = grid         ; search neighboring grid cells
```

```
nstlist         = 10          ; 20 fs
rlist           = 1.0         ; short-range neighborlist cutoff (in nm)
rcoulomb        = 1.0         ; short-range electrostatic cutoff (in nm)
rvdw            = 1.0         ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype     = PME         ; Particle Mesh Ewald for long-range electrostatics
pme_order       = 4           ; cubic interpolation
fourierspacing  = 0.16        ; grid spacing for FFT
; Temperature coupling is on
tcoupl          = V-rescale   ; modified Berendsen thermostat
tc-grps         = Protein Non-Protein   ; two coupling groups - more accurate
tau_t           = 0.1   0.1   ; time constant, in ps
ref_t           = 300   300   ; reference temperature, one for each group, in K
; Pressure coupling is on
pcoupl          = Parrinello-Rahman   ; Pressure coupling on in NPT
pcoupltype      = isotropic   ; uniform scaling of box vectors
tau_p           = 2.0         ; time constant, in ps
ref_p           = 1.0         ; reference pressure, in bar
compressibility = 4.5e-5      ; isothermal compressibility of water, bar^-1
refcoord_scaling = com
; Periodic boundary conditions
pbc             = xyz         ; 3-D PBC
; Dispersion correction
```

```
DispCorr       = EnerPres     ; account for cut-off vdW scheme
; Velocity generation
gen_vel        = no           ; Velocity generation is off
```

9.2.4 pr.mdp – Parameters for 100 ps restrained simulation to stabilise temperature (pre-simulation)

```
define         = -DPOSRES     ; position restrain the protein
; Run parameters
integrator     = md           ; leap-frog integrator
nsteps         = 50000        ; 2 * 50000 = 100 ps
dt             = 0.002        ; 2 fs
; Output control
nstxout        = 100          ; save coordinates every 0.2 ps
nstvout        = 100          ; save velocities every 0.2 ps
nstenergy      = 100          ; save energies every 0.2 ps
nstlog         = 100          ; update log file every 0.2 ps
; Bond parameters
continuation    = no          ; first dynamics run
constraint_algorithm = lincs    ; holonomic constraints
constraints     = all-bonds     ; all bonds (even heavy atom-H bonds) constrained
lincs_iter      = 1           ; accuracy of LINCS
lincs_order     = 4           ; also related to accuracy
; Neighborsearching
ns_type        = grid         ; search neighboring grid cells
nstlist        = 5            ; 10 fs
```

```
rlist         = 1.0         ; short-range neighborlist cutoff (in nm)
rcoulomb      = 1.0         ; short-range electrostatic cutoff (in nm)
rvdw          = 1.0         ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype   = PME         ; Particle Mesh Ewald for long-range electrostatics
pme_order     = 4           ; cubic interpolation
fourierspacing  = 0.16      ; grid spacing for FFT
; Temperature coupling is on
tcoupl        = V-rescale   ; modified Berendsen thermostat
tc-grps       = Protein Non-Protein   ; two coupling groups - more accurate
tau_t         = 0.1   0.1   ; time constant, in ps
ref_t         = 300   300   ; reference temperature, one for each group, in K
; Pressure coupling is off
pcoupl        = no          ; no pressure coupling in NVT
; Periodic boundary conditions
pbc           = xyz         ; 3-D PBC
; Dispersion correction
DispCorr      = EnerPres    ; account for cut-off vdW scheme
; Velocity generation
gen_vel       = yes         ; assign velocities from Maxwell distribution
gen_temp      = 300         ; temperature for Maxwell distribution
gen_seed      = -1          ; generate a random seed
```

9.2.5 pr_P.mdp - Parameters for 100 ps restrained simulation to stabilise pressure (pre-simulation)

```
define          = -DPOSRES      ; position restrain the protein

; Run parameters

integrator      = md            ; leap-frog integrator

nsteps          = 50000         ; 2 * 50000 = 100 ps

dt              = 0.002         ; 2 fs

; Output control

nstxout         = 100           ; save coordinates every 0.2 ps

nstvout         = 100           ; save velocities every 0.2 ps

nstenergy       = 100           ; save energies every 0.2 ps

nstlog          = 100           ; update log file every 0.2 ps

; Bond parameters

continuation    = yes           ; Restarting after NVT

constraint_algorithm = lincs    ; holonomic constraints

constraints     = all-bonds     ; all bonds (even heavy atom-H bonds) constrained

lincs_iter      = 1             ; accuracy of LINCS

lincs_order     = 4             ; also related to accuracy

; Neighborsearching

ns_type         = grid          ; search neighboring grid cells

nstlist         = 5             ; 10 fs

rlist           = 1.0           ; short-range neighborlist cutoff (in nm)

rcoulomb        = 1.0           ; short-range electrostatic cutoff (in nm)

rvdw            = 1.0           ; short-range van der Waals cutoff (in nm)
```

```
; Electrostatics
coulombtype     = PME          ; Particle Mesh Ewald for long-range electrostatics
pme_order       = 4            ; cubic interpolation
fourierspacing  = 0.16         ; grid spacing for FFT
; Temperature coupling is on
tcoupl          = V-rescale    ; modified Berendsen thermostat
tc-grps         = Protein Non-Protein   ; two coupling groups - more accurate
tau_t           = 0.1   0.1    ; time constant, in ps
ref_t           = 300   300    ; reference temperature, one for each group, in K
; Pressure coupling is on
pcoupl          = Parrinello-Rahman    ; Pressure coupling on in NPT
pcoupltype      = isotropic    ; uniform scaling of box vectors
tau_p           = 2.0          ; time constant, in ps
ref_p           = 1.0          ; reference pressure, in bar
compressibility = 4.5e-5       ; isothermal compressibility of water, bar^-1
refcoord_scaling = com
; Periodic boundary conditions
pbc             = xyz          ; 3-D PBC
; Dispersion correction
DispCorr        = EnerPres     ; account for cut-off vdW scheme
; Velocity generation
gen_vel         = no           ; Velocity generation is off
```

9.3 NMR Experiment Parameters

9.3.1 ^{13}C ^{1}H HMQC of U{^{15}N, ^{12}C, ^{2}H}, selectively $^{13}CH_3$-Ile$^{\delta1}$. $^{13}CH_3$-Ala$^{\beta}$, $^{13}CH_3$-Leu$^{\delta}$, $^{13}CH_3$-Val$^{\gamma}$ 25 µM BiP-NBD (Chapter 4)

	1H	13C
Total number of points	1280	512
Number of valid points	616	256
Spectral width (Hz)	12335.526	2986.858

9.3.2 ^{13}C ^{1}H HMQC of U{^{15}N, ^{12}C, ^{2}H}, selectively $^{13}CH_3$-Ile$^{\delta1}$. $^{13}CH_3$-Ala$^{\beta}$, $^{13}CH_3$-Leu$^{\delta}$, $^{13}CH_3$-Val$^{\gamma}$ 1 µM BiP-NBD (Chapter 4)

	1H	13C
Total number of points	1280	512
Number of valid points	616	256
Spectral Width (Hz)	12335.526	2986.858

9.3.3 ^{13}C-^{1}H HMQC of U{^{15}N, ^{12}C, ^{2}H}, selectively $^{13}CH_3$-Ile$^{\delta1}$ FL-BiP (WT) (Chapter 5)

	1H	13C
Total number of points	1280	300
Number of valid points	616	150
Spectral Width (Hz)	12335.526	2386.972

9.3.4 ^{15}N-^{1}H TROSY of 50 µM IRE1-LD WT (+/- 8M Urea) and IRE1-LD (D123P) (Chapter 4)

	1H	15N
Total number of points	1792	290
Number of valid points	811	145
Spectral Width (Hz)	11432.927	3467.406

9.3.5 ^{15}N-^{1}H TROSY of IRE1 cLD, IRE1-LD (C109S), (C148S) and (C332S) (Chapter 4)

	1H	15N
Total number of points	1792	200
Number of valid points	811	100
Spectral Width (Hz)	11432.927	3467.406

9.3.6 ^{15}N-^{1}H TROSY of IRE1-LD and FL-BiP (Chapter 5)

	1H	15N
Total number of points	1792	180
Number of valid points	811	90
Spectral Width (Hz)	11432.927	3467.406

9.3.7 ^{19}F for mixture screening (Chapter 6)

	19F
Number of scans	128
Spectral Width (Hz)	39682.5

9.3.8 ^{19}F for individual compound screening (Chapter 6)

	19F
Number of scans	128
Spectral Width (Hz)	39682.5

9.3.9 ^{1}H for bexarotene (Chapter 6)

	1H
Number of scans	512
Spectral Width (Hz)	8370.56

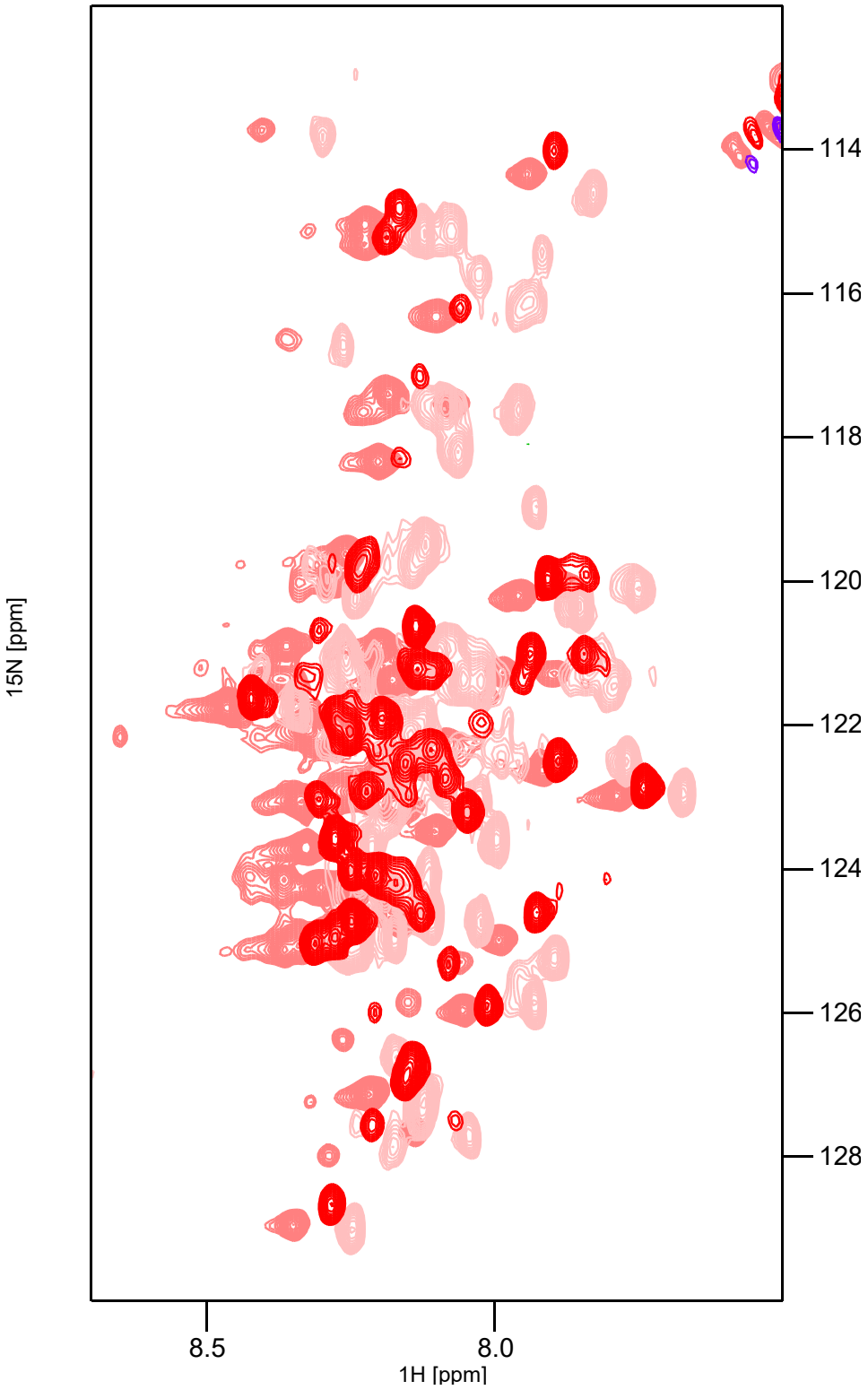

Figure 9.1

15N-1H TROSY spectra of 50 µM IRE1-LD (FL) at 298K, 288K and 278K (red, dark pink and light pink respectively)