# Investigating a key autophagy protein-protein interaction and its modulation by phosphorylation and an ALS-related mutation

Andrew Brennan, MSci (Hons.)

Thesis submitted to The University of Nottingham for the degree of Doctor of Philosophy

September 2018

# Abstract

Autophagy is a cellular degradation pathway which removes cytoplasmic material including protein aggregates, damaged organelles and pathogens. The interaction of the autophagy adaptor p62 with lipid-anchored hAtg8 receptor proteins is a key step in this process whereby cargo bound to p62 is sequestered to the autophagosomal membrane. Here it is engulfed for transport to the lysosome where it is enzymatically degraded. Recruitment of p62 filaments to the autophagosome membrane also plays an important role in autophagosome expansion. The p62-hAtg8 interaction is mediated by the Atg8-interacting motif (AIM) which is a short sequence present in many autophagy adapter/receptor proteins.

For the biophysical characterisation required in this work, seven hAtg8 proteins were produced and characterised; the six hAtg8 proteins and one phosphomimetic mutant. *In vitro* biophysical characterisation has shown a fivefold range of binding affinities for a WT p62 AIM peptide in its binding to the six hAtg8 proteins, with GABARAPL1 bound with the highest affinity. The binding of GABARAPL2 with the WT p62 AIM peptide has been structurally characterised by NMR chemical shift mapping as an investigation into the entropically driven nature of this interaction, which is unique amongst the hAtg8 proteins. This identified a novel binding patch for the peptide on the GABARAPL2 surface, particulaly in relation to Leu341 of p62.

Regulation of this interaction by phosphorylation has been investigated. Phosphorylation of p62 by TBK1 has been shown to increase the binding affinity of the protein for LC3B. A previously known phosphorylation site on p62 has then been characterised which showed that a pS342 p62 AIM peptide binds to all six hAtg8 proteins with higher affinity than the WT peptide. NMR-based strutural characterisation of the pS342 p62 AIM-LC3B interaction indicated that this interaction occurs in a related but slightly altered manner to the WT p62 AIM-LC3B interaction. The phosphomimetic T50E LC3B mutant has been shown to bind with a weaker affinity to the WT p62 AIM peptide than WT LC3B.

Dysregulation of autophagy has become a key area of research in a number of neu-

rodegenerative disorders including amyotrophic lateral sclerosis and frontotemporal lobar degeneration (ALS-FTLD). Previous work has shown that an ALS-associated L341V mutation of the p62 AIM is defective in recognition of LC3B [1]. This work expands upon this observation and looks at the interaction of L341V p62 with all hAtg8 proteins which produces a more complex picture where different hAtg8 proteins are affected differently by the mutation. However, the average binding affinity and selectivity of the hAtg8 proteins with the p62 AIM is decreased by the mutation which can be clearly associated with ALS-related neurodegeneration.

# Acknowledgements

Firstly, I would like to thank both Prof. Mark Searle and Prof. Rob Layfield for their supervision throughout my project, it has been a priveledge to work with you and to learn from you. This research would have faltered without the help of Dr. Huw Williams who has provided not just invaluable NMR expertise but also constant advice and support throughout my studies. Dr. Jed Long has been a shining light of knowledge for all technical and experimental issues throughout this work and could not receive the credit he deserves. Dr. Dan Scott has provided a wide range of support in both mass spectrometry and binding assays for which I am grateful.

My thanks must also extend to all members of Searle, Panos, Thomas and Loenarz groups who I have had the pleasure of sharing lab and office space with; for their moral support as well as scientific discussions. I would also like to thank all of the technical and support staff in CBS and the School of Chemistry.

Authoring this thesis would not have been possible without the respites of joy Ruth has provided. Finally, I send my thanks to my parents and Andrea for always believing that one day I would finish this and get a real job.

# Contents

# List of Figures

# Abbreviations

**AIM** Atg8-interacting motif
**ALS** Amyotrophic lateral sclerosis
**AMPK** AMP-activated protein kinase
**APS** Ammonium persulphate
**Atg** Autophagy-related
**Atg8** Autophagy-related protein 8
**AU** Absorbance units
**BAG** Bcl-2-associated athanogene
**BSA** Bovine serum albumen
**CD** Circular dichroism
**CMA** Chaperone-mediated autophagy
**CPL** Circularly polarised light
**CSP** Chemical shift perturbation
**CV** Column volume
**DNA** Deoxyribonucleic acid
**E. *coli*** *Escherichia coli*
**EDTA** Ethylene diamine tetraacetic acid
**ESI** Electrospray ionisation
**FPLC** Fast protein liquid chromatography
**FTLD** Frontotemporal lobar degeneration
**FUS** Fused in sarcoma
**GST** Glutathione S-transferase
**hAtg8** Human autophagy-related protein eight
**HDMS** High definition mass spectrometery
**HOPS** Homotypic fusion and protein sorting
**HSQC** Heteronuclear single quantum correlation
**IPTG** Isopropylthiogalactopyranoside
**ITC** Isothermal titration calorimetry
**kb** Kilobase
**kDa** Kilodalton
**KEAP1** Kelch-like ESH-associated protein 1
**KLD** Kinase, ligase, Dpn1
**KO** Knock out
**LB** Lysogeny broth

**MEF** Mouse embryonic fibroblast
**mRNA** Messenger ribonucleic acid
**MRE** Mean residue ellipticity
**MS** Mass spectrometry
**mTORC1** Mammalian target of rapamycin complex 1
**NBR1** Neighbour of BRCA1 gene protein
**NDP52** Nuclear dot protein 52 kDa
**NMR** Nuclear magnetic resonance
**Nrf2** Nuclear factor erythroid 2-related factor 2
**$OD_{595}$** Optical Density at 595 nm
**PB1** Phox and Bem1
**PCR** Polymerase chain reaction
**PDB** Protein data bank
**PE** Phosphatidylethanolamine
**PI3K CIII** class III phosphatidylinositol 3-kinase
**PINK1** PTEN-induced putative kinase 1
**PLEKHM1** Pleckstrin homology and RUN domain containing M1
**ppm** Parts per million
**PTM** post-translational modification
**qTOF** Quadrupole time of flight
**RCF** Relative centrifugal force
**RNA** Ribonucleic acid
**rpm** Revolutions per minute
**SDS-PAGE** Sodium dodecyl sulphate polyacrylamide gel electrophoresis
**siRNA** Short interfering ribonucleic acid
**SOD1** Superoxide dismutase 1
**TBK1** TANK-binding kinase 1
**TDP-43** TAR-DNA binding protein 43
**TEMED** N,N,N,N - tetramethylethylenediamine
**TRAF6** Tumour necrosis factor associated receptor factor 6
**Ub** Ubiquitin
**UBA** Ubiquitin binding domain
**ULK1** unc51-like autophagy activating kinase 1
**UV** Ultraviolet
**VCP** Valosin-containing protein
**1D** One dimensional

**2D** Two dimensional

**3D** Three dimensional

| **Amino acid** | **Three letter code** | **One letter code** |
|---|---|---|
| Alanine | Ala | A |
| Arginine | Arg | R |
| Asparagine | Asn | N |
| Aspartate | Asp | D |
| Cysteine | Cys | C |
| Glutamine | Gln | Q |
| Glutamate | Glu | E |
| Glycine | Gly | G |
| Histidine | His | H |
| Isoleucine | Ile | I |
| Leucine | Leu | L |
| Lysine | Lys | K |
| Methionine | Met | M |
| Phenylalanine | Phe | F |
| Proline | Pro | P |
| Serine | Ser | S |
| Threonine | Thr | T |
| Tryptophan | Trp | W |
| Tyrosine | Tyr | Y |
| Valine | Val | V |

Table 1: Amino acid abbreviations for the twenty common amino acids.

# Chapter 1

# An Introduction To Autophagy

## 1.1 Autophagy

Autophagy (auto=self; phagy=eating) refers to a range of evolutionarily conserved degradation pathways which occur throughout the cytoplasm of most mammalian cells which work to maintain homeostasis. It operates on intracellular material including proteins, organelles and pathogens; both as quality control and in response to cellular emergencies. Autophagy can be considered to work in tandem with proteasomal degradation to maintain proteostasis [4]. It is upregulated as a response to a number of stressors including starvation, hypoxia and osmotic or oxidative stress [5, 6, 7, 8]. Autophagy is unique due to its ability to degrade physically large and chemically diverse substrates [9]. This can be particularly important for large protein aggregates which the proteasome cannot degrade. These degraded products can then be recycled as metabolites to maintain vital cellular processes, which is particularly useful in the case of starvation-induced autophagy.

More recent work has also led to an expanded view of autophagy in which it selectively degrades certain types of cargo and also plays a role in various secretory pathways [4, 10, 11, 12, 13]. The possibility of selective autophagy shows the two sides of the process. On the one hand autophagy can be induced under stress conditions to globally degrade bulk cytosolic components but on the other hand it is continually working to a lesser extent on specific targets which must be removed as they occur to prevent cellular damage. Selective autophagy of these specific targets can be classified according to the type of cargo being degraded. These classifications include aggrephagy (protein aggregates), macrolipidophagy (lipids), mitophagy (mitochondria) and xenophagy (pathogens).

The three distinct types of autophagy are macroautophagy, microautophagy and chaperone-mediated autophagy (CMA). All three pathways result in cargo being delivered to the lysosome for degradation but in different ways. CMA occurs via a specific pathway whereby target proteins are bound to the heat-shock cognate 70 stress protein via a KFERQ motif, estimated to be present in roughly 30% of proteins [14, 15]. This chaperone then directs the protein to the lysosome where it is unfolded and translocated across the membrane. Microautophagy is the direct engulfment of cellular components by the lysosome. Macroautophagy (herein referred to as autophagy) however utilises a specific double membrane compartment known as the autophagosome to recruit and then transport material to the lysosome, and is the focus here.

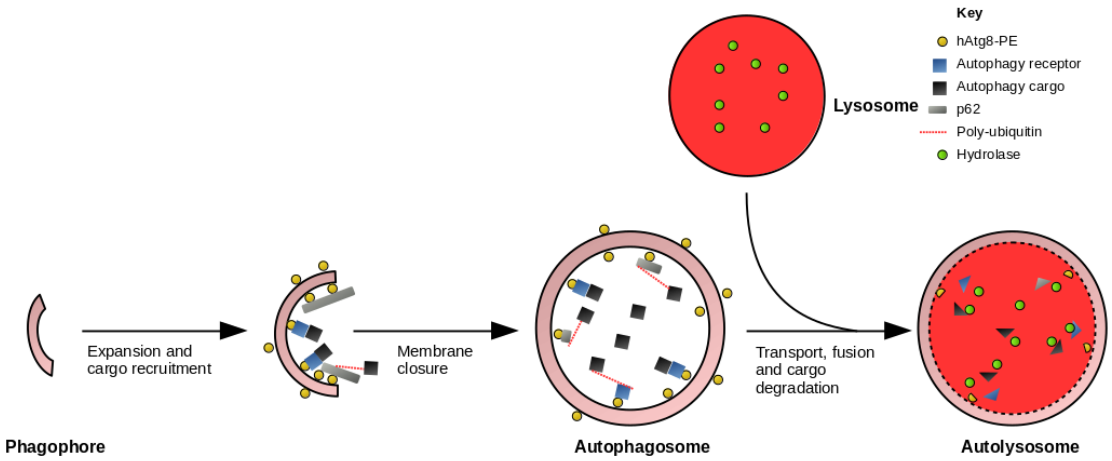



Figure 1.1: Cartoon model showing an overview of autophagy showing the expansion, fusion and maturation of the autophagosome.

The process of autophagy, summarised in Figure 1.1, is initiated by two protein complexes: the unc51-like autophagy activating kinase 1 (ULK1) and the class III phosphatidylinositol 3-kinase (PI3K CIII). ULK1 is regulated by mammalian target of rapamycin complex 1 (mTORC1) inhibition and AMP-activated protein kinase activation. When ULK1 is phosphorylated PI3K CIII moves from the cytoskeleton to the preautophagosomal structure. This nucleates vesicles, which leads to the initial formation of the cup-shaped phagophore [16]. The double-membrane is then expanded by the acquisition of membrane from other organelles. The endoplasmic reticulum is often cited as the source of this membrane as it is frequently observed close to autophagosomes [17]. This process is not fully understood though and the membrane likely also comes from other organelles such as the mitochondria or from

the plasma membrane [18, 19, 20]. Throughout the membrane elongation, material for degradation is either sequestered to the membrane specifically or simply enclosed by the membrane.

The membrane is then sealed to form a mature autophagosome in a process best described as membrane scission [21]. This sealed membrane is then transported using the dynein machinery on microtubules to the lysosome where it fuses to form an autolysosome. The autophagosome will often fuse with a late endosome first before later fusing with a lysosome. This fusion process has also not been fully characterised but a number of protein complexes are known to be involved; namely Rab small GTPases, soluble N-ethylmaleimide-sensitive factor attachment protein receptors and homotypic fusion and protein sorting (HOPS) complexes [22, 23, 24]. In the autolysosome, the cargo is then degraded by several lysosomal hydrolytic enzymes including proteases, lipases and nucleases. This is facilitated by vacuolar ATPase which pumps protons into the autolysosome to drop the pH to $\sim$5, the optimal pH for these enzymes.

The expansion of autophagosomes has not been well characterised and due to the different types of autophagy it is possible that it may form by different mechanisms. Depending upon the context, the autophagosome can be thought of as either forming around the cargo or as forming independently with cargo being directed to the expanding membrane. This may be thought of as part of the difference between bulk and selective autophagy. In the context of starvation it may simply be that the autophagosome forms and whatever happens to be trapped inside is degraded. Whereas, autophagy receptors may bind to selective cargo and recruit the membrane there and elongate the autophagosome around the cargo, particularly in the case of large cargo.

## 1.2 Atg8 and its Human Homologues

Autophagy has been principally investigated in yeast and that work has been subsequently expanded into mamallian studies; in particular the study of human autophagy. In yeast, at least 35 different autophagy-related (Atg) genes have been discovered, predominantly through genetic studies [25]. Atg protein 8 (Atg8) in particular has been shown to play an important role in the maturation and transport of autophagosomes and the recruitment of cargo therein [26]. There are six human Atg8 (hAtg8) homologues (as well as several pseudogenes) which can be split into two subfamilies; the LC3 subfamily which includes LC3A, LC3B and LC3C and the GABARAP subfamily which includes GABARAP, GABARAPL1 and GABARAPL2 (also known as GATE-16) [27, 28]. All mentions of specific Atg8 homologues hereafter

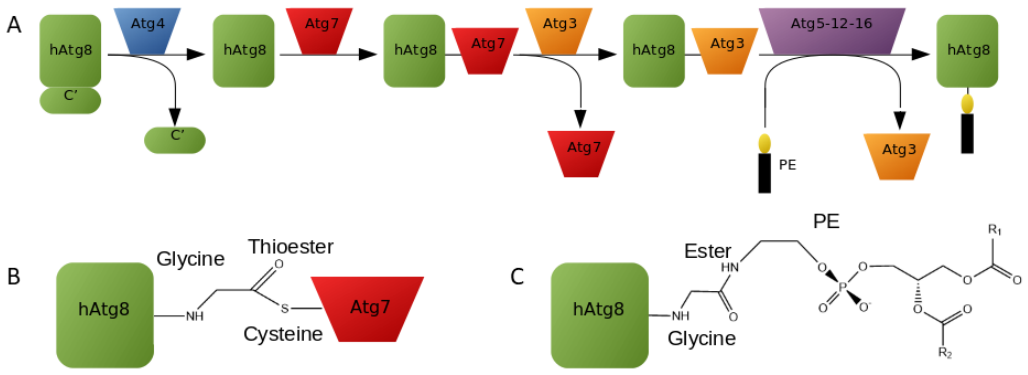



Figure 1.2: Cartoon schematic showing the lipidation reaction of hAtg8 proteins. A: enzymatic reaction scheme showing the lipidation reaction. This shows the removal of a C-terminal section of hAtg8 by Atg4 followed by the activation of hAtg8 by Atg7. The hAtg8 is then transferred to Atg3 before the final lipidation of hAtg8 by the Atg5-12-16 complex; B: structure of the linkage between the C-terminal glycine of hAtg8 proteins and a cysteine of Atg7 via a thioester bond (the structure is analogous to the linkage between hAtg8 and Atg3); C: structure of PE and its linkage to the C-terminal glycine of hAtg8 proteins via an ester bond.

will be in reference to the human protein. Generally, it has been observed that the LC3s mediate the expansion of the autophagosome membrane, whereas GABARAP proteins are involved with the maturation and sealing of the autophagosome and subsequent autophagosome−lysosome fusion [29]. However, these are general rules as, for example, both LC3 and GABARAP proteins have been shown to interact with pleckstrin homology and RUN domain containing M1 (PLEKHM1) in GST-pulldown experiments from HeLa cell lysate [30]. PLEKHM1 was shown in this study to regulate autophagosome-lysosome fusion through the HOPS complex showing a non-typical role of LC3 proteins in this process.

These hAtg8 proteins initially exist in a cytosolic form before lipidation to anchor them to the autophagosomal membrane. This lipidation reaction is described in Figure 1.2. Initially, the cytosolic form is activated by Atg4 which removes a C-terminal tail, leaving a C-terminal glycine residue [31]. This is activated by Atg7 before transfer to Atg3; both steps involve conjugation to the C-terminal hAtg8 glycine carboxyl group by the formation of a thioester bond [32]. Conjugation to a lipid phosphatidylethanolamine (PE) group localised to the autophagosomal membrane is then completed by the Atg12-Atg5-Atg16 complex to give lipidated hAtg8 as well as releasing Atg3 [33, 34]. Atg3 has been shown to have curvature-sensing functionality, whereby the lipidation reaction can only be completed in the presence of a sharply

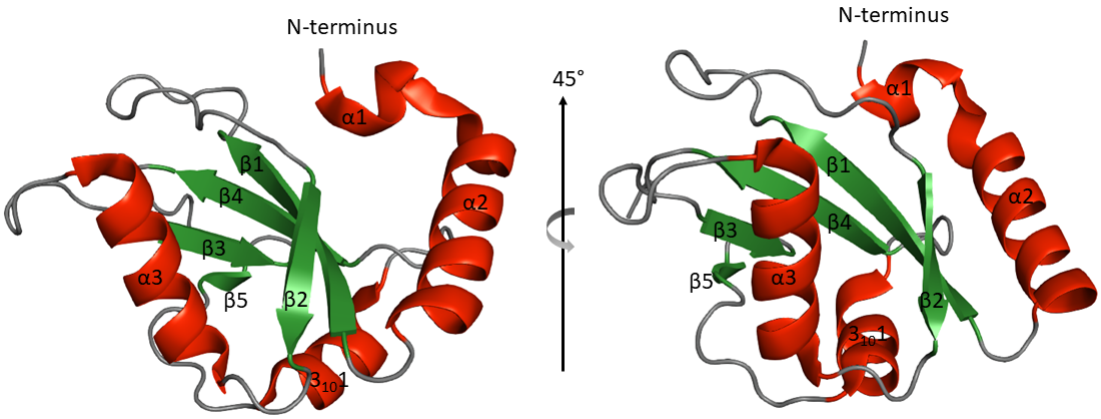



Figure 1.3: Crystal structure of LC3A shown in two orientations highlighting and labeling the secondary structural elements of the hAtg8 proteins. Helices shown in red; $\beta$-strands shown in green; random coil shown in grey (PDB ID: 3WAL).

curved membrane [35]. This results in hAtg8-PE being added at the leading ends of the membrane as it extends, captures cargo and finally fuses; which are key functions of the hAtg8 proteins.

The attachment of hAtg8 to the autophagosome is regulated by the cysteine protease Atg4 which deconjugates them from PE [36]. This allows the maintenance of free hAtg8 levels whilst their removal from autophagosomal membranes has been linked with autophagosome maturation [37]. The conjugation systems previously described are also analogous to the E1, E2, E3 and deubiquitinating enzymes in the ubiquitination post-translational modification (PTM) system.

The Atg8 homologues are formed of a ubiquitin (Ub)-like core fold with an N-terminal $\alpha$-helical extension, shown in Figure 1.3. This N-terminal extension ($\alpha$1 and $\alpha$2) is unique to the hAtg8s amongst Ub-like proteins. The Ub-like core consists of five $\beta$-strands ($\beta$1-$\beta$5) hydrophobically packed around a single $\alpha$-helical turn ($\alpha$3), as well as a short $3_{10}$ helix. This core structure is sometimes referred to as a $\beta$-grasp.

There is a high level of structural similarity among the hAtg8 proteins despite their sequence diversity, as shown in Figure 1.4. There is diversity in the charge and morphology across the surface of the Ub-like core with particularly noticeable differences in the surface charge presented on the N-terminal helices. These helices are basic in all three LC3s, neutral in GABARAPL2 and acidic in GABARAP and GABARAPL1. It is worth noting that the divide of these proteins into two subfamilies can be slightly misleading. There is a high level of sequence identity between LC3A and LC3B and between GABARAP and GABARAPL1; 83% and 87% respectively.

| | LC3B | LC3C | GABARAP | GABARAPL1 | GABARAPL2 |
|---|---|---|---|---|---|
| **LC3A** | 83 | 57 | 31 | 33 | 40 |
| **LC3B** | | 53 | 31 | 32 | 38 |
| **LC3C** | | | 38 | 39 | 40 |
| **GABARAP** | | | | 87 | 58 |
| **GABARAPL1** | | | | | 61 |

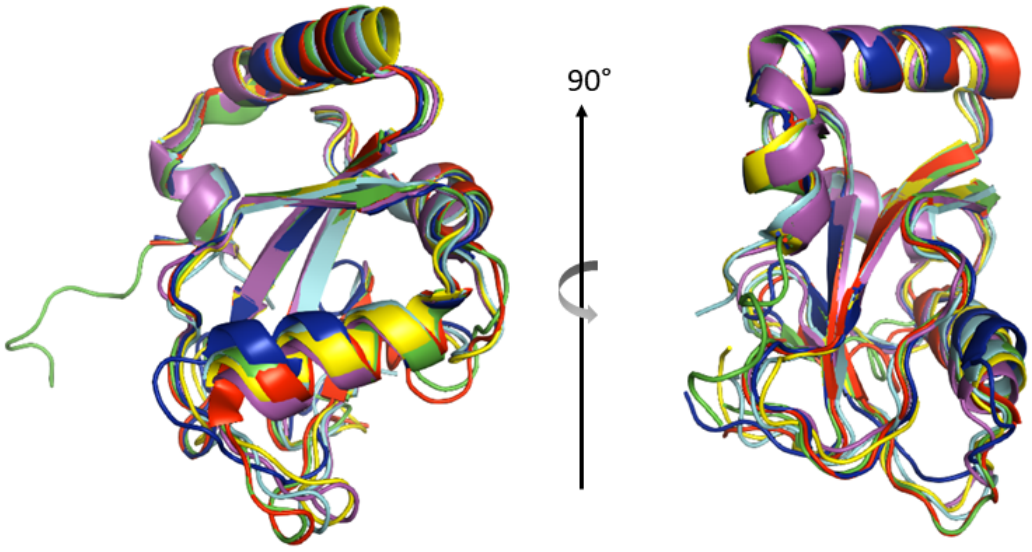



Figure 1.4: Comparison of hAtg8 homologue sequence and structure. Top: Pairwise similarity of hAtg8 proteins shown as a percentage, Bottom: Overlay of crystal structures of hAtg8 proteins showing structural similarity, shown from two angles. Protein (PDB ID) colour; LC3A (3WAL) red; LC3B (3VTU) green; LC3C (3WAM) blue; GABARAP (1GNU) yellow; GABARAPL1 (2R2Q) magenta; GABARAPL2 (4CO7) cyan.

However, there is a much lower sequence similarity between LC3C and GABARAPL2 and their respective subfamilies.

It is presumed that the evolution of diversity from a single Atg8 in yeast to six human homologues is due to the existence of some degree of distinction in the roles of hAtg8 proteins in the autophagic process in humans. There is a growing body of work focused on autophagy-related proteins selectively binding to a single hAtg8 and how this may lead to a specific function [38, 39, 40]. For example, LC3C has been shown to bind with a degree of selectivity, over the other hAtg8 proteins, to nuclear dot protein 52 kDa (NDP52) to promote pathogen-containing autophagosome maturation. This selectivity may be made possible by the low level of sequence similarity of LC3C with the other hAtg8 proteins. However, much is still unknown about how the diversity of hAtg8s may lead to specific functions or conversely how despite the differences they

may operate similarly.

Some tissue specificity in the expression of hAtg8 proteins has been observed [41, 42, 43]. Higher levels of GABARAPL1 and GABARAPL2 messenger RNA (mRNA) are found in the central nervous system. GABARAPL1 is expressed to a higher level in brain regions important in the regulation of somatomotor or endocrine function. The expression of GABARAP on the other hand has been found to be much higher in the endocrine glands themselves. LC3C expression is usually found to be the lowest except in lung tissue where it is particularly high. These differences in expression may be due to these proteins having diverse functions therefore the paralogue which is most required may be expressed more in a particular type of tissue. For example, it can be rationalised that LC3C, which has been shown to be involved in xenophagy, is most highly expressed in the lungs as this is where exposure to pathogens is the greatest.

One recent study which raises key questions about the role of hAtg8 proteins in autophagy was conducted by Nguyen *et al.* [44]. This shows that the hAtg8 proteins are not required for autophagosomes to form and seal in the context of starvation-induced autophagy or drug (oligomycin or antimycin A)-induced mitophagy. This work showed that after a genetic knockout (KO) of all six hAtg8 proteins mature autophagosomes were present but appeared to be smaller than in non-KO cells. This study also implicates GABARAP subfamily proteins as critical for autophagosome-lysosome fusion in these contexts as the KO cells were less able to undergo autophagosome-lysosome fusion. However, restoration of the GABARAP subfamily and not the LC3 led to a return to the WT phenotype. It is important to note that this study does not preclude a vital role for hAtg8 proteins in the selective autophagy of other ubiquitinated cargo including protein aggregates as this context was not studied.

## 1.3 The Autophagy Interacting Motif

Amongst the proteins which interact with hAtg8 proteins are the autophagy receptors and the autophagy adapters. Autophagy receptors bind cargo for degradation; usually through a ubiquitin binding domain but other recognition domains are also utilised [45, 46, 47]. This acts as a bridge between the cargo and the autophagosomal membrane. The process of selective autophagy is mediated by these receptors. Autophagy adapters perform the alternative role of binding to hAtg8 proteins to facilitate the other stages of autophagy; the formation, maturation and transport of the autophagosome and its subsequent fusion with the lysosome. The interaction of

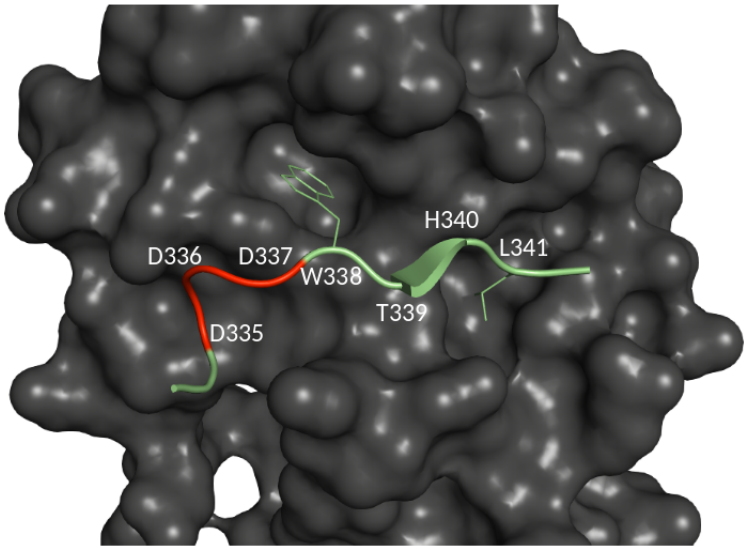



Figure 1.5: Crystal structure of LC3B (grey surface) bound to a short peptide representing the AIM of p62 (red and green cartoon representation). Peptide W338 and L341 side chains are shown with stick representations of the side chains binding into the surface clefts of LC3B. Acidic residues preceding the AIM consensus sequence are shown in red. (PDB ID: 2ZJD).

these receptors and adapters with hAtg8 proteins is mediated by a short amino acid sequence called an Atg8-interacting motif (AIM) [48].

The AIM is typically considered to have the consensus sequence [F/W/Y]-$x_1$-$x_2$-[I/L/V] or more generally [$\theta$]-$x_1$-$x_2$-[$\tau$], with $\theta$ representing a residue with a side chain containing an aromatic group and $\tau$ representing a residue with a hydrophobic side chain [12]. Structural studies have indicated that the first and fourth residues in this sequence generally bind into separate hydrophobic pockets on the hAtg8 surface, these side chains are shown in Figure 1.5 as sticks [49, 50]. It has also been noted that the intermediate residues ($x_1$ and $x_2$) in the motif can have an effect on binding as they can affect hAtg8 binding specificity [51]. Additionally, there is a preference for some acidic residues around the AIM motif particularly the three residues preceding, which can reinforce this hydrophobic interaction with ionic interactions due to the presence of positively charged residues nearby on the hAtg8 surface [38, 52]. For example, the three acidic residues of the p62 AIM preceding the core motif are shown in Figure 1.5.

The binding surface on hAtg8 proteins is characterised by the binding pockets near the exposed $\beta$2-strand, previously referred to as the W-site and the L-site due to their binding of the tryptophan (W) and leucine (L) residues of the p62 AIM.

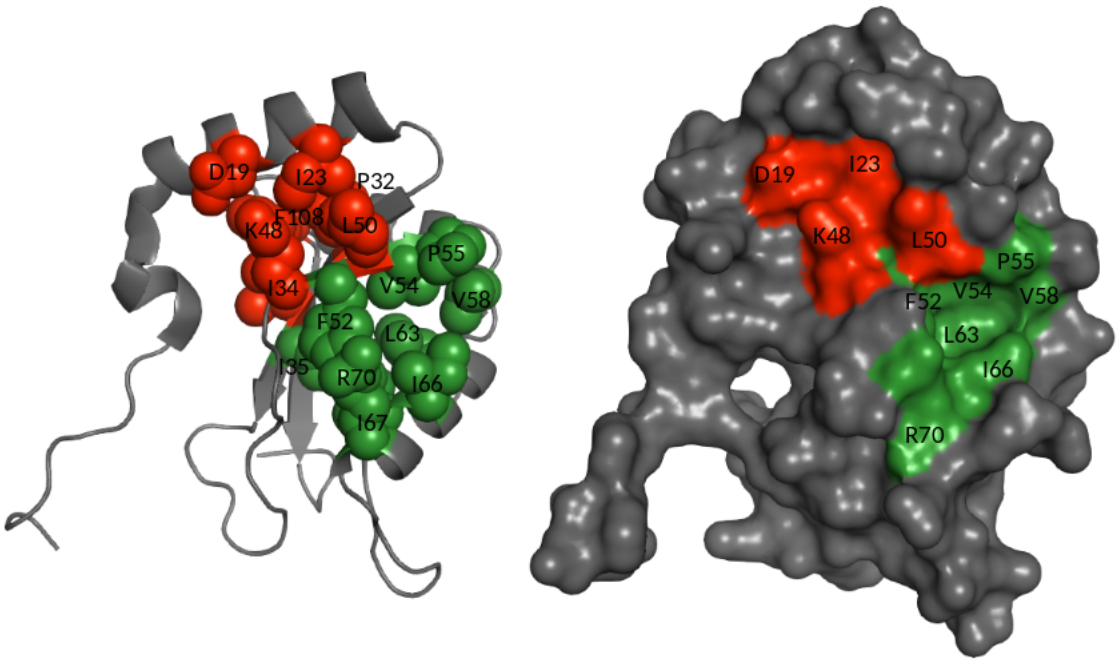



Figure 1.6: Crystal structure of LC3B highlighting the binding pockets for recognition of the AIM; residues which form part of the pockets are labeled. Green surface indicates residues forming the L-site; red surface indicates residues forming the W-site. (PDB ID: 2ZJD).

These predominantly conserved binding pockets are shown on LC3B in Figure 1.6. The W-site is located between $\alpha 2$ and $\beta 2$ and the L-site is located between $\beta 2$ and $\alpha 3$. This mode of binding explains why AIM sequences will selectively bind to hAtg8 proteins over other Ub-like proteins as the N-terminal helix $\alpha 2$ is unique to the hAtg8 family. These binding pockets around the exposed $\beta 2$-strand allow the formation of an intermolecular parallel $\beta$-sheet with the AIM residues adopting an extended $\beta$-conformation. The formation of this conformation may be less defined in full length protein as current structural studies have been performed solely with AIM peptides which are likely to have a greater degree of conformational flexibility.

As well as a high level of structural similarity between the hAtg8 proteins, there is also a high level of similarity in their binding modes with AIM-containing peptides. Comparison of the structure of the GABARAPL1/neighbour of BRCA1 gene protein (NBR1) AIM complex with the LC3B/p62 (over analogous secondary structure elements) and GABARAP/calreticulin (over the structured residues E12-V114) complexes gave RMSD values of 1.85 Å and 1.59 Å respectively [50, 1, 53].

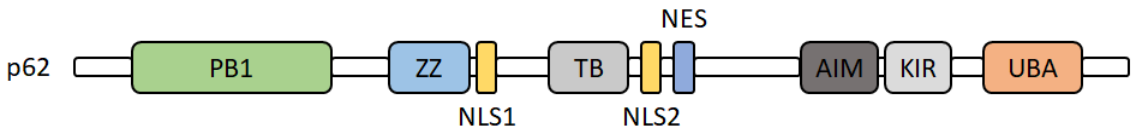



Figure 1.7: Schematic illustrating the domain structure of p62. The PB1 domain facilitates p62 oligomerisation. The ZZ domain is a zinc finger which interacts with RIP which is involved in NF-KB activation. NLS1 and NLS2 are nuclear localisation signal domains. TB is a TRAF6-binding domain which is also involved with NF-KB activation. NES is a nuclear export domain. The AIM binds to Atg8 proteins. The KIR interacts with Keap1 which is involved in a number of signaling pathways. The UBA domain interacts with ubiquitin moieties.

## 1.4 p62 Autophagy Receptor/Adapter

p62, expressed from the gene *SQSTM1*, was the first mammalian protein identified as an autophagy receptor and it is probably the best characterised [11, 49, 54, 55]. It is a multidomain protein comprised of 440 amino acid residues, as illustrated in Figure 1.7, which is involved in a number of important cellular pathways including autophagy.

Three domains mediate the direct role of p62 in autophagy: the N-terminal Phox and Bem1 (PB1) domain, its AIM and the C-terminal ubiquitin binding domain (UBA). These three operate to bring autophagic cargo to the forming autophagosomal membrane. First, a PTM occurs where K63-linked poly-Ub is covalently attached to cargo which has been targeted for degradation by autophagy [56]. The p62 UBA domain selectively binds to this specific poly-Ub linkage type [47, 57]. The p62 AIM then targets this cargo to the forming autophagosome via its interaction with the hAtg8 proteins on the membrane. The crystal structure showing the interaction of p62 with LC3B, mediated by its AIM, was shown in Figure 1.5.

These interactions form the basis of the role of p62 in autophagy but it also has additional functionality. The PB1 domain of p62 has been shown to facilitate protein self-assembly into flexible helical filaments [58, 59]. These filaments have been reported to shorten on binding to K63-linked octa-Ub but not to di-Ub. This indicates that binding to this larger octa-Ub polymer is sterically disrupting the filaments, rather than some conformational change from binding. Binding to LC3B was also shown to not affect filament size in these studies.

This presents two models for the role of p62, both of which presumably occur simultaneously and cooperatively *in vivo*. As autophagosome formation is initiated

and hAtg8 proteins are recruited to the membrane, they bind to p62 filaments which present multiple AIM domains and can form a scaffold for membrane elongation via avid interactions. This is supported by the fact that the genetic KO of all six hAtg8 proteins using CRISPR in HeLa cells results in smaller autophagosomes while p62 is still enclosed in the membrane, suggesting that the filaments may not be able to attach to the membrane and work as a scaffold to elongate the membrane but they can still be engulfed by the autophagosome [44].

On the other hand, with cargo bound the p62 will form smaller filaments or even return to monomers which sequester cargo to the membrane. The smaller filaments would also still be capable of forming avid interactions with both the hAtg8 protein on the membrane and with Ub chains attached to cargo. These avid interactions may be particularly important as the individual binding affinities of these interactions are not particularly strong, with low $\mu$M $K_D$ values [51, 1]. This disruption of the larger filaments may also result in the accessibility of more UBA domains, allowing more cargo to be bound. In this way p62 can therefore be considered as both an autophagy receptor (which recruits cargo to the autophagosome) and an adapter (which facilitates the elongation of the membrane), with both roles simultaneously carried out to some degree. These dual roles for p62 appear to be particularly important in aggrephagy.

Mitophagy is the selective degradation of mitochondria by autophagy in which p62 is postulated to play a role. A common mechanism for this process is the PTEN-induced putative kinase 1 (PINK1)/Parkin pathway whereby PINK1 initially accumulates on the membrane of damaged mitochondria. This leads to the recruitment of the ubiquitinating enzyme Parkin which attaches poly-Ub chains onto outer mitochondrial membrane proteins as a degradation signal. These Ub chains are then recognised by autophagy receptors and the process of autophagy then proceeds as previously described. p62 and optineurin have been shown to be involved as autophagy receptors here.

p62 does appear to operate slightly differently in this selective mitophagy process compared to its role in aggrephagy for example. This may be due to a preference for the use of optineurin as an autophagy receptor in this case. p62 does not appear to be as involved in recruitment of mAtg8 to damaged mitochondria as it is to other ubiquitinated substrates though it does appear to accelerate mitochondrial aggregation [60]. This essentially means that p62 binds to ubiquitinated mitochondria then aggregates through the PB1 domain. This works to hold damaged mitochondria together which facilitates other autophagy receptors recruiting the cargo to the forming

autophagosome. It appears to be playing more of an autophagy adapter role than in other processes.

Optineurin has been identified as a key autophagy receptor involved in this process as engulfment of damaged mitochondria is reduced in HeLa cells short interfering ribonucleic acid (siRNA)-depleted of optineurin [61]. This study also showed some evidence of optineurin and p62 binding to different sites on the mitochondrial surface. Depletion of p62 however did not appear to affect mitochondrial engulfment which indicates that the process may be optimised by aggregation of mitochondria by p62 but that is not essential for optineurin-mediated recruitment.

p62 is also involved in autophagy signaling as well as this direct role in the process; this occurs through two important interactions with mTORC1 and Keap1. Cellular growth, protein synthesis and autophagy are all regulated by mTORC1, through its kinase activity [62]. mTORC1 activation occurs on the surface of the lysosome in an amino acid dependent manner which requires p62 [54, 63]. Raptor is a subunit of the mTORC1 complex to which p62 has been shown to bind, resulting in activation of the complex. This activation comes partly from the interaction of p62 with tumour necrosis factor associated receptor factor 6 (TRAF6) which is an E3 ubiquitinating enzyme [64]. This interaction recruits TRAF6 to mTORC1 where it ubiquitinates the complex with K63-linked poly-Ub which promotes mTORC1 activation. So p62 is required to downregulate autophagy under nutrient rich conditions.

The other important autophagy signaling interaction of p62 is facilitated by kelch-like ESH-associated protein 1 (Keap1) which forms a homodimer that binds to nuclear factor erythroid 2-related factor 2 (Nrf2) under normal cellular conditions. This complex is degraded by the proteasome due to an interaction of Keap1 with Cullin3-based ubiquitin ligases [65]. This interaction is interrupted by oxidative stress or by p62 binding, both of which lead to the release of Nrf2 [66, 67]. The oxidative stress alters the conformation of Keap1 to inhibit binding whereas p62 interacts with Keap1 through its KIR domain to competitively inhibit the interaction. p62 has a phosphorylation site which increases the affinity of this interaction which interestingly has been shown to be phosphorylated by activated mTORC1 [68]. The release of Nrf2 by these two mechanisms allows the protein to translocate to the nucleus to induce expression of a number of genes encoding for anti-oxidant and anti-inflammatory proteins and also p62 itself. In addition to this p62 is also involved in signaling pathways related to NF-$\kappa$B activation, adipogenesis and apoptosis [54, 69, 70].

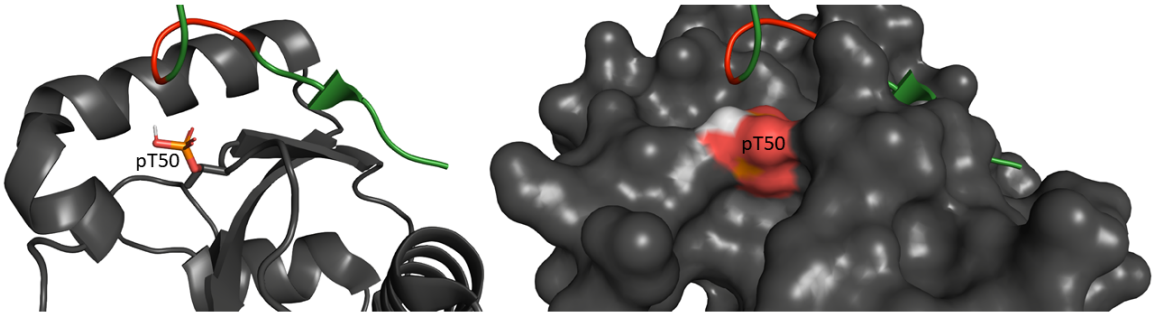



Figure 1.8: Two representations of the crystal structure of LC3B bound to a short peptide representing the AIM of p62 modified to show the LC3B T50 phosphorylation and its proximity to the acidic residues of the AIM which are shown in red (modified from 2ZJD). Left: Cartoon representation of LC3B pT50 showing the phosphorylated side chain as a stick representation; Right: Surface representaiton of LC3B pT50 showing the surface presentation of the phosphate group (red: oxygen, white: hydrogen, orange: phosphate).

## 1.5 Autophagy Regulation

PTMs are a diverse set of chemical operations which can be used to modulate cellular processes [71]. Often this involves attachment of a chemical group to a translated protein to alter the activity of the protein. Moieties which are commonly added as a PTM include phosphate, ubiquitin, acetyl, lipid and methyl groups. These additions can occur reversibly or irreversibly. PTM of autophagy-related proteins is reviewed in Ref [72] and the lipidation of hAtg8 proteins illustrated in Figure 1.2 is a key PTM in the process of autophagy.

Phosphorylation is a particularly common PTM of proteins whereby a protein kinase covalently attaches a phosphate group to an amino acid side chain; most commonly to the hydroxyl group of a serine, threonine or tyrosine. Phosphorylation has become a key area of research in the regulation of autophagy and will be an area of focus within this work. One example of the role of phosphorylation in autophagy is the phosphorylation on S12 of LC3A, which inhibits LC3A recruitment to the autophagosome. This lowers the basal rate of autophagy but metabolic and pathological inducers of autophagy can then result in dephosphorylation and a subsequent increase in LC3A in the autophagosomal membrane, presumably leading to increased autophagic flux [73].

Another phosphorylation site on an LC3 subfamily protein, which will be investigated in this work, is T50 of LC3B which has been shown to be involved in regulating

the fusion between autophagosomes and lysosomes [3]. A modified crystal structure showing the site of this phosphorylation is shown in Figure 1.8. In the work by Wilkinson *et al.*, STK3 and STK4 were shown to phosphorylate LC3B T50 *in vitro*. Subsequent knock down of these kinases in mouse-embryonic fibroblasts (MEF) led to inefficient clearing of group A *streptococcus*, presumably due to the autophagic block in lysosome fusion. A phosphomimetic mutation (T50E) at this site in STK3/STK3 KO MEF was then shown to reverse this blockage in autophagy and result in the restoration of efficient bacterial clearing. As this phosphorylation site appears on the surface of LC3B in close proximity to its AIM binding site it may be the case that the effect on autophagy caused by this phosphorylation is mediated by altering the interaction of LC3B with autophagy adapters.

A significant number of p62 phosphorylation sites are known, throughout the sequence of the protein [46, 68, 74]. Of importance to its role in autophagy, it has been shown to be phosphorylated by casein kinase 2 and TANK-binding kinase 1 (TBK1) at S403 which is part of its UBA domain [74, 75]. This leads to increased affinity for poly-Ub chains and a subsequent enhancement of autophagic clearance of Ub-proteins and Ub-mitochondria.

Another interesting phosphorylation site has been identified on p62, directly adjacent to the AIM, on S342 [76, 77]. In work by Richter *et al.,* p62 was shown to be phosphorylated at S342 significantly more in the presence of WT TBK1 compared to a kinase-deficient mutant (TBK1 K38A) [46]. It has been shown that inhibition of TBK1 leads to inhibition of autophagy, via a number of selective autophagy pathways, so this clearly links this S342 phosphorylation site with autophagy as well as the hypothetical link of close sequential proximity [78]. TBK1 has also been shown to be important in mitophagy and xenophagy [46, 79].

The phosphorylated serine is expected to be involved in the modulation of the ionic interaction between p62 and hAtg8 proteins as it directly follows the hydrophobic residue in the consensus AIM; WTHL[pS] or [$\theta$]-x$_1$-x$_2$-[$\tau$]-[pS]. In the modified crystal structure shown in Figure 1.9 this phosphate group appears in the vicinity of two charged residues on the LC3B surface for example; H57 and R69. The phosphorylation will therefore be hypothesised to increase the affinity of the interaction with LC3B similarly to the acidic residues which precede the AIM. It is worth noting that as the surface charge of the hAtg8 proteins varies it is likely that each hAtg8-p62 interaction will be affected differently, which could clearly be related to functional diversity.

The p62 S342 phosphorylation can be compared to the optineurin S177 phosphorylation, which is also an autophagy receptor believed to be phosphorylated by

TBK1. This phosphorylation site precedes the AIM and phosphorylation there has been shown to increase LC3B binding and promote selective autophagy of *Salmonella enterica* [80]. p62 has also been shown to have a phosphorylation site in this preceding region of the AIM; pS332. So it remains to be determined whether the phosphorylation on the other side of the AIM consensus sequence, as in the case of p62 pS342, will have a similar effect.

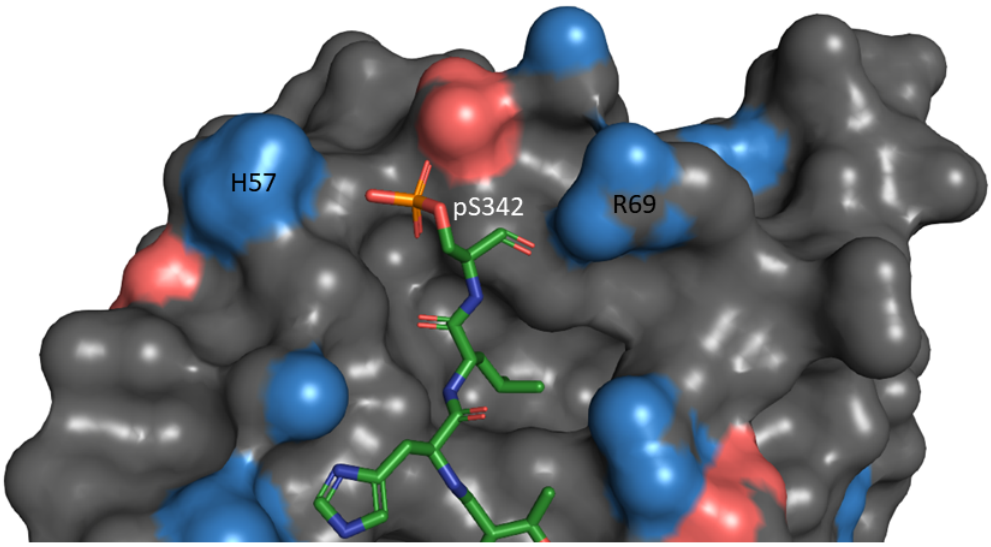



Figure 1.9: Crystal structure of LC3B bound to a short peptide representing the AIM of p62 modified to show pS342 and the proximity of this added phosphate group to charged atoms on the protein surface (modified from 2ZJD). LC3B is shown in grey with the charge shown in red (negatively charged atoms) and blue (positively charged atoms); the peptide is shown in green with oxygen atoms in red, nitrogen atoms in blue and phosphorus atoms in orange.

## 1.6 Neurodegenerative Disease

Neurodegenerative disease is a broad term for a range of disorders which involve degradation and/or death of neurons. Neurodegenerative diseases are commonly characterised by misfolded or aggregated protein or RNA species in neurons and often by the induction of neuronal cell death. Neurons are at particular risk from these diseases as they are post-mitotic and therefore do not undergo mitosis. As such aggregates (sometimes referred to as inclusions) cannot be diluted by cell division. Neurodegenerative diseases can be defined by the particular protein(s) involved in aggregates and in which particular neurons this accumulation is causing the pathology. The type

and position of neurons within the nervous system can usually be correlated with function to some degree so localised damage leads to the differential symptoms of a given condition. Neurodegenerative disease can also be classified as either sporadic (patients have no affected first degree relatives) or familial (patients have affected first degree relatives). However it is usually the case that so-called sporadic cases will have some genetic basis, which indicates the inheritance of risk factors rather than the specifically disease-causing genes of familial cases.

Aggregates can be deleterious to neurons through a range of mechanisms. It is first important to consider the content of the aggregates. Proteins which have aggregated often have important roles so their inclusions may lead to a reduction in functional protein in the cell. For example, TAR-DNA binding protein 43 (TDP-43) plays a role in RNA regulation which could be reduced in cells where significant concentrations of this protein are contained within aggregates; as is the case in amyotrophic lateral sclerosis (ALS) which is discussed later. RNA aggregation results in similar damging processes as mRNA which is aggregated cannot be translated into protein. In addition to this, the autophagic and proteasomal machinery itself will be occupied by these aggregates leading to a reduction in a cells ability to effectively degrade other material [81]. Protein chaperones required for the folding of other proteins may also become entangled in these aggregates making other proteins more likely to misfold and exacerbate the problem [82].

In addition to the loss of protein function it can also be the case that the aggregates develop new activity which can be damaging to the cell. These large structures may, for example, sterically inhibit the function of cells by blocking axonal transport. The projections of neurons are particularly susceptible in this way compared to other cell types due to the relatively large distance they can extend. It has also been suggested that protein aggregates lead to a signaling response which reduces total protein synthesis and that the persistence of these aggregates in neurodegenerative disease results in on-going detrimental deficits in proteins vital for cellular function [83].

### 1.6.1 Link To Autophagy

It has become increasingly clear that there is a link between autophagy and a number of neurodegenerative diseases. Studies in mice have shown removal of important genes involved in autophagy leads to ubiquitin-positive inclusion bodies in neurons which appear to lead to a phenotypic difference in motor neurons [84, 85]. p62 also

appears at reduced levels in the brains of Alzheimer's disease patients, due to reduced transcription from *SQSTM1,* compared with those not suffering from the disease but increased p62 inclusion bodies are observed in the hippocampus of Alzheimer's disease patients [86, 87]. So less p62 is being produced as well as a higher proportion being trapped in aggregates which will clearly affect the capacity of this protein to perform its key functions in the process of autophagy.

It is usually considered that autophagy and proteasomal degradation work together for protein quality control and maintenance of proteostasis, something which appears to be failing in neurodegenerative disease. Both processes recognise poly-Ub proteins but it is generally thought that most of these proteins are degraded by the proteasome with only the larger or insoluble protein aggregates removed via autophagy. Recruitment to a given degradative pathway occurs depending on the type of poly-Ub chain added. The functional redundancy of having both of these systems may be required to ensure this important role is effectively carried out. For example, it has been shown that proteasome inhibition leads to an increase in autophagic activity [88].

One interesting aspect of this duality of degradation pathways is evidence of a switch caused by aging. This comes from studies of the molecular chaperones of the Bcl-2-associated athanogene (BAG) protein family. BAG1 and BAG3 have been shown to regulate proteasomal degradation and autophagy respectively [89, 90]. Informatively, during aging there is a switch in preferential expression from BAG1 to BAG3 which implies older cells are preferentially relying upon autophagy [91]. This goes some way towards explaining the increased importance of autophagy in age-related neurodegenerative disease. However, it is important to also consider that it has been shown that autophagy declines upon aging [92]. This data suggests that both systems decrease during aging but over time there begins to be increased reliance on autophagy to regulate proteostasis.

As well as clear evidence on the role of dysfunctional aggrephagy in neurodegenerative disease, dysfunctional mitophagy has also been implicated. Mitochondria are of particular importance for neuronal function and survival due to the high energy demand of these cells. There is a wide range of evidence linking neurodegeneration with impaired mitochondrial dynamics and function [93, 94, 95, 96]. So the autophagic machinery is required to maintain at least these two types of specific degradation of large substrates (aggrephagy and mitophagy) making dysfunction more likely as the systems decline during aging.

CMA has also been shown to be vulnerable to neurodegenerative disease. For

example, inefficient translocation of disease-related tau mutant fragments into the lysosome, via CMA, has been shown to form damaging Tau aggregates which could contribute to Alzheimer's disease and other tau-related neurodegenerative diseases [97]. CMA has been observed to decline during aging which can clearly combine with other age-related declines to influence neurodegeneration. Age-related CMA decline has been linked to decreasing levels of lysosome-associated membrane protein 2A, an important lysosomal maintenance glycoprotein [98].

### 1.6.2 Amyotrophic Lateral Sclerosis

Amyotrophic lateral sclerosis (ALS) is a subtype of motor neuron disease which results in progressive paralysis from the selective loss of upper and lower motor neurons in the brain and spinal cord. Symptoms of ALS are broadly caused by muscle weakness and can include difficulty controlling speech or swallowing, fatigue in the limbs, muscle cramps and twitches. Symptom onset often varies dramatically amongst patients. Eventually, the muscle weakness will result in paralysis which leads to death by respiratory failure. ALS is generally fatal within 1 to 5 years of onset and is considered to be sporadic in 90 to 95% of cases [99].

Frontotemporal lobar degeneration (FTLD) is a related collection of neurodegenerative diseases typically characterised by the atrophy of the frontal and temporal lobes of the brain. This category of disease includes frontotemporal dementia, progressive supranuclear palsy and corticobasal degeneration. Symptoms typically involve cognitive impairment which leads to a variety of changed behaviours in the patient. These two diseases are believed to have related pathogenic mechanisms and are perhaps best considered as on opposite ends of a spectrum of disease often referred to as ALS-FTLD. There is often an observable overlap in symptoms of these two poles of the continuum as, for example, up to half of ALS patients also experience cognitive impairment of the type more associated with FTLD [100].

Throughout the ALS-FTLD spectrum there is a common pathology associated with the aggregation of TDP-43 and fused in sarcoma (FUS) proteins [101, 102]. Additionally, inclusions of superoxide dismutase 1 (SOD1) and tau proteins are found distinctly at opposite ends of the disease spectrum in ALS and FTLD respectively. All of these protein inclusions can also be characterised by the presence of p62 and Ub as well as other important autophagy receptors such as valosin-containing protein (VCP/p97) [103, 104, 105, 106]. TDP-43 has been reported as the predominant protein species found in inclusions in ALS-FTLD as well as being a secondary pathological

feature in a number of other neurodegenerative diseases [101, 107, 108, 109, 110]. Proteins found in ALS-FTLD inclusions and their normal functions are summarised in Table 1.1. Autophagic dysregulation has been shown in sufferers of ALS-FTLD by virtue of increased numbers of autophagosomes in post mortem analyses of patient tissue [111]. This is indicative of upregulated autophagy or it may be that autophagosomes are not being cleared efficiently.

Mutations which have been associated with ALS can generally be split into those which affect genes linked to RNA processing and those which affect protein degradation. SOD1 is the exception to this which removes superoxide radicals and has been associated with the maintenance of mitochondria. These mutations then either result in a reduction of RNA processing (or protection against oxidative damage) and/or a reduction in the capacity of the machinery to remove protein aggregates and damaged mitochondria. This all amounts to an increased probability of neuron death. The most common mutation observed in ALS is in the C9orf72 gene, where intronic expansion of a hexanucleotide (GGGGCC) repeat is observed [120, 133, 134]. This expanded RNA forms aggregates which have been shown to lead to cell death in neuronal cell lines [135]. The normal function of C9orf72 protein remains poorly defined but it has been implicated in protein degradation by the regulation of endosomal trafficking [136]. Further supporting the role of autophagy in the pathogenesis of ALS-FTLD, a number of mutations believed to have given a genetic predisposition to these diseases are found in key autophagy regulators [137].

TBK1 was previously discussed in Section 1.5 as a kinase involved in the regulation of autophagy and this has also been shown to be mutated in some cases of ALS [138, 139, 140]. A review by Oakes *et al.* collated all known TBK1 mutations from eight genetic studies of patients with ALS-FTLD and found 92 different mutations [141]. These were predicted to produce a mixture of loss-of-expression and loss-of-function results for the protein. They also found that the observed mutations are equally distributed throughout the protein and the majority are missense.

Mutations to p62 have been shown to be related to ALS in a number of studies. One American study reported p62 mutations in 1.76% of cases of familial ALS and 4.37% cases of sporadic ALS [142]. A p62 L341V mutation of particular interest to this work has been found in a patient with late-onset sporadic ALS [143]. This missense mutation changes the hydrophobic residue in the consensus sequence of the p62 AIM. This was initially predicted to be a benign mutation in terms of the proteins interactions with hAtg8 proteins as it is to another hydrophobic residue of similar size so was expected to occupy the hydrophobic pocket (L-site) in the same way. However

| Protein | Pathology | Function |
|---|---|---|
| FUS [112, 113] | ALS-FTLD | DNA repair and RNA processing [114] |
| Optineurin [115, 116] | ALS-FTLD | NF-$\kappa$B and interferon signaling, autophagy and vesicular trafficking [117] |
| p62 [113, 118, 119, 120] | ALS-FTLD | Autophagy, proteasomal targeting, inflammation and apoptosis [121, 122] |
| SOD1 [115, 116] | ALS | Antioxidant enzyme [123] |
| Tau [116, 124] | FTLD | Microtubule stabilisation [125] |
| TDP-43 [112, 113, 116, 119, 120] | ALS-FTLD | RNA processing [126] |
| Ubiquilin [120] | ALS-FTLD | Regulation of protein degradation [127, 128, 129] |
| Ubiquilin 2 [112, 120] | ALS-FTLD | Regulation of protein degradation [127, 130, 131] |
| Ubiquitin [112, 113, 115, 118, 119] | ALS-FTLD | Targeting to pathways including degradation, cellular transport or alteration in activity [132] |

Table 1.1: Protein species which are found in ALS-FTLD inclusions and their non-pathological functions.

it was shown that a L341V p62 AIM peptide had a roughly 3-fold reduction in binding to LC3B *in vitro* [1]. It was also shown that L341V p62 was less incorporated into autophagic vesicles in two neuronal cell lines (NSC-34 and Neuro-2a). The mutated residue is highlighted in Figure 1.10 which aids in the structural rationalisation of this decrease in affinity. Leucine is a larger hydrophobic residue and it can be seen that this allows it to orient the two terminal carbon atoms of the side chain into a cleft on the LC3B surface. Due to its different size and branching this is not possible for valine and instead there is likely steric clash with the $\beta$-branched sidechain which requires more significant remodeling of the binding site; associated with the reduction in affinity. This mutation shows a small change in the autophagic process leading to a clear pathogenic phenotype, which is more likely to lead to cell death and disease. p62 mutations associated with ALS-FTLD have also been shown to be detrimental to mitochondria by virtue of decreased membrane potential [96]. These mutations were shown to be linked with increased reactive oxygen species in the cytosol. This shows how the diverse role of p62 in autophagy and other processes makes mutations which affect its function so damaging.

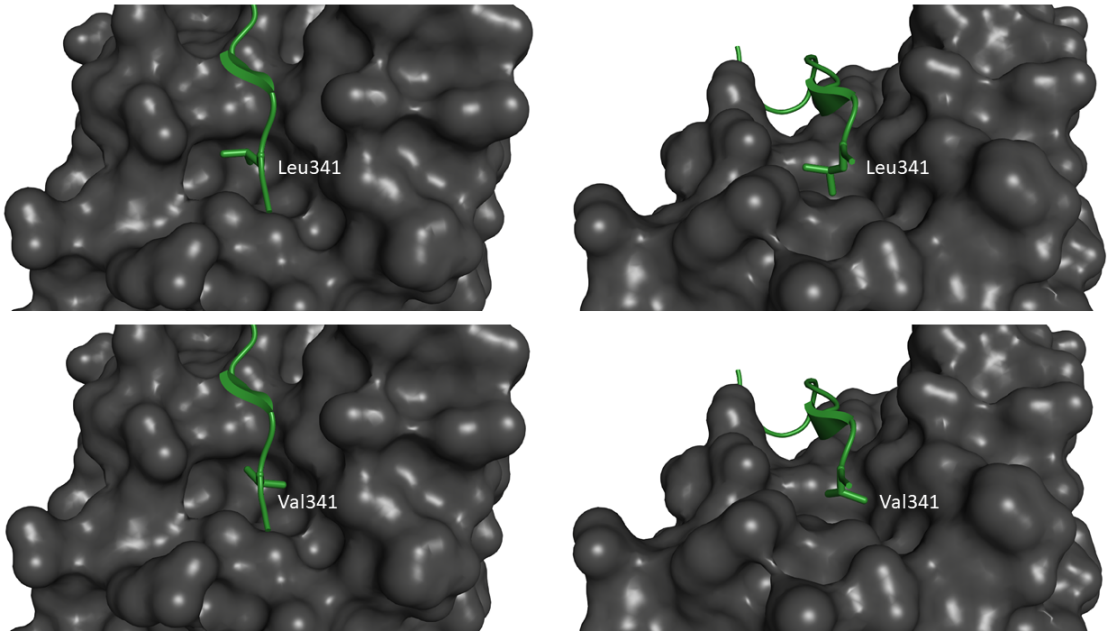



Figure 1.10: Crystal structure of LC3B bound to a short peptide representing the AIM of p62. Top: Two angles of the WT p62 AIM; Bottom: Two angles of the p62 AIM modified to show the L341V ALS-related mutation (modified from 2ZJD).

### 1.6.3 Potential Therapies

There is currently no cure for ALS, most likely due to the complex pathology and causes associated with the disease. Regulation of autophagy has become an interesting therapeutic approach for neurodegenerative disease in general and ALS in particular. mTOR pathway inhibitors, particularly rapamycin, have been used as autophagy inducers in a wide range of research but they induce global, mass degradation which may lead to negative outcomes for cells and their survival if applied clinically. They have also been shown to have an immunosupressing effect. The potential benefits and risks of rapamycin as a treatment for ALS have been shown using a mouse model of ALS (SOD1 G93A). Treatment of this mouse model with rapamycin has been shown to increase neuronal degeneration and reduce lifespan [144]. A similar effect was seen in a another FTLD-ALS mouse model (VCP R155H), where rapamycin treatment worsens the degenerative phenotype [145]. This negative finding in SOD1 G93A mice was supported by Staats *et al.* but they also recognised the off-target immunosuppressing effect of this drug [146]. By suppressing mature lymphocytes in these mice an increased survival rate was observed. This indicates that the increase in autophagy observed during rapamycin treatment does have beneficial effects on this ALS model but the off target interaction with the immune system results in a net negative effect. This clearly opens the door for more specific targeting of autophagy as a therapeutic strategy.

Lithium is another drug that has been investigated for treatment of ALS with mixed results. Work on presymptomatic SOD1 G93A male mice was shown to increase autophagy and extend lifespan [147]. A study to follow up this work used the same SOD1 G93A mutation as a model but looked at female mice from two different genetic backgrounds [148]. Conversely, this study showed no effect in one of the genetic groups and reduced survival in the other. Further study of lithium in this context will be required for its development as a reliable therapeutic.

Another treatment strategy has looked at the role of ataxin-2. This protein appears to be involved in the maturation of stress granules which involve inclusions of TDP-43. Therapeutic reduction in ataxin-2 has been shown, in mice, to reduce TDP-43 inclusions which leads to improved motor function and increased lifespan [149]. This is an interesting example of the possibilities for the treatment of ALS, as counter to the previous examples these positive outcomes were caused by an apparent reduction in autophagy; illustrating the fine balance of this process.

A more promising therapeutic strategy may come from regulation of selective

autophagy rather than this simplistic control of total autophagy. This may occur via control of autophagy receptors and adapters; in terms of both their expression and PTMs. A potential selective therapeutic approach was shown in the case of urolithin A which has been shown to induce mitophagy [150]. This has been shown to prolong *C. elegans* lifespan and to increase exercise capacity in mouse models of age-related decline. Another drug which has been shown to induce mitophagy is p62-mediated mitophagy inducer (PMI) [151]. This has been shown to work by increasing the free concentration of Nrf2 which causes an upregulation in p62 expression; leading to increased mitophagy without observable consequence to mitochondrial function or autophagic flux.

This highlights the variability of the disease and that there is likely no single cure. Further understanding of the widespread causes of ALS-FTLD may lead to targeted therapy whereby drugs can be selected for a specific function to target the particular failings of autophagy in a given case.

## 1.7 Aims and Objectives

The interaction between p62 and the hAtg8 proteins on the autophagosomal membrane is critical in the process of autophagy. This interaction allows p62 filaments to work as scaffolds for the growing autophagosome and allows ubiquitinated cargo to be selectively targeted to the membrane for degradation. This work aims to advance understanding of this protein-protein interaction and how it is regulated and dysregulated.

Regarding hAtg8 proteins, the focus of the literature has predominantly been on the study of LC3B. This protein has been used as a marker for autophagy and has been the focus of more biophysical and structural studies than the other hAtg8 proteins. The divergent evolution of six hAtg8 proteins has been suggested to imply functional differences in these proteins which is supported by their varying sequence homology and some characterisation of selective binding. The overarching view of this work is to expand the focus beyond LC3B or any particular hAtg8 in isolation to investigate these six related proteins simultaneously.

This work will apply a range of biophysical characterisation techniques to study the key protein-protein interactions between the hAtg8 proteins and the AIM of p62. This will investigate the binding interactions of the six hAtg8 proteins with the WT p62 AIM. A previously published backbone NH assignment of GABARAPL2 will first be confirmed before it is utilised to investigate the binding of this protein with the WT

p62 AIM. The initial exploration of the hAtg8-WT p62 AIM interaction will allow comparison for the effects caused by phosphorylation and a disease-related mutant.

The study of phosphorylation as a regulatory mechanism of this interaction will look at a p62 phosphorylation site and an LC3B phosphorylation site. The study of p62 phosphorylation by TBK1 will include a kinase binding assay and will then focus on a particular TBK1 phosphorylation site at S342 of p62. The effect of phosphorylation at this site on binding affinity will be studied in comparison to the WT. Additionally, previous structural work into the interaction of LC3B with the p62 AIM will be furthered by an NMR chemical shift mapping study of the LC3B-phosphorylated peptide interaction. The LC3B phosphorylation site will be studied by production of a phosphomimetic mutant of the protein. This protein will be compared to the WT in terms of its structure and interactions through biophysical characterisation methodologies.

Finally, an ALS-related p62 mutation will be characterised using biophysical techniques to study the strength and selectivity of its interaction with the hAtg8 proteins. This will allow comparison with WT p62 and the subsequent discussion of how the mutation may influence disease progression.

The results pertaining to these aims will be discussed in the following chapters:

- Chapter 3 – The production and purification of the six hAtg8 proteins and their initial characterisation. This will investigate the differences in these proteins as well as their suitability for further study.

- Chapter 4 – Biophysical characterisation of the hAtg8-p62 AIM interaction to produce measures of selectivity.

- Chapter 5 – NMR assignment of GABARAPL2 and the subsequent work on defining a binding interface for its interaction with the WT p62 AIM.

- Chapter 6 – Investigation of the interaction of the hAtg8 proteins with the phosphorylated p62 AIM. This will include a chemical shift mapping study of the binding of a phosphorylated peptide to LC3B.

- Chapter 7 – Production of an LC3B phosphomimetic mutant and subsequent biophysical characterisation of its interaction with the p62 AIM.

- Chapter 8 – Study of the effects of the ALS-related mutant on the hAtg8-p62 AIM interaction.

# Chapter 2

# Materials and Methods

## 2.1 Biological Materials

Glycerol stocks of *E. coli* cell lines BL21 (DE3), C41 (DE3) and DH5$\alpha$ were provided by Dr. Jed Long (School of Chemistry, University of Nottingham). Plasmid DNA from each of the hAtg8 proteins (LC3A, LC3B, LC3C, GABARAP, GABARAPL1 and GABARAPL2) was also provided by Dr. Jed Long. The hAtg8 DNA sequence had been inserted into the pGEX-4T-1 plasmid between the *BamHI* and *XHO1* restriction sites as shown on the plasmid map shown in Figure 2.1. These hAtg8 sequences (corresponding to the protein sequence after $\alpha$-thrombin cleavage) inserted into the pGEX-4T-1 plasmid are shown in Figures 2.2-2.7.

Plasmid DNA corresponding to p62 was provided by Dr. Daniel Scott (School of Life Sciences, University of Nottingham). The p62 DNA sequence had been inserted into the pGEX-4T-1 plasmid between the *BamHI* and *XHO1* restriction sites as shown on the plasmid map in Figure 2.8.

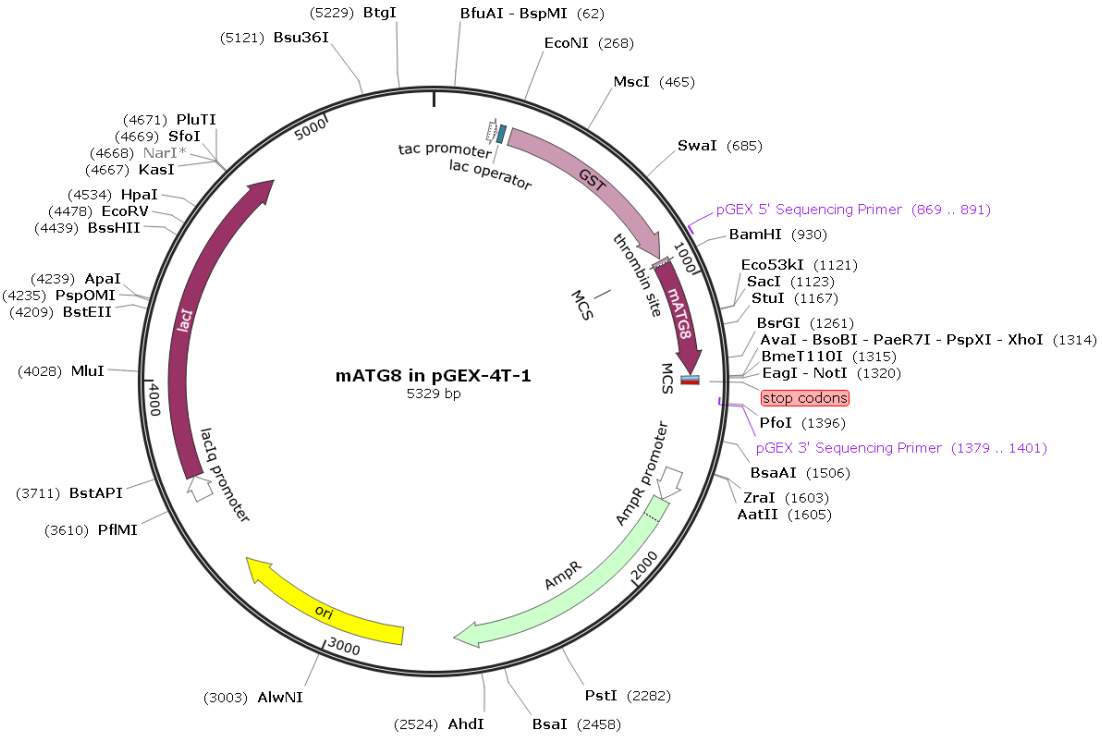



Figure 2.1: Annotated plasmid map showing hAtg8 DNA inserted into pGEX-4T-1 and other important features of the plasmid.

**LC3A:**

```
 G   S   M   P   S   D   R   P   F   K   Q   R   R   S   F   A   D
GGA TCC ATG CCC TCA GAC CGG CCT TTC AAG CAG CGG CGG AGC TTC GCC GAC
 R   C   K   E   V   Q   Q   I   R   D   Q   H   P   S   K   I   P
CGC TGT AAG GAG GTA CAG CAG ATC CGC GAC CAG CAC CCC AGC AAA ATC CCG
 V   I   I   E   R   Y   K   G   E   K   Q   L   P   V   L   S   K
GTG ATC ATC GAG CGC TAC AAG GGT GAG AAG CAG CTG CCC GTC CTG GAC AAG
 T   K   F   L   V   P   D   H   V   N   M   S   E   L   V   K   I
ACC AAG TTT TTG GTC CCG GAC CAT GTC AAC ATG AGC GAG TTG GTC AAG ATC
 I   R   R   R   L   Q   L   N   P   T   Q   A   F   F   L   L   V
ATC CGG CGC CGC CTG CAG CTG AAC CCC ACG CAG GCC TTC TTC CTG CTG GTG
 N   Q   H   S   M   V   S   V   S   T   P   I   A   D   I   Y   E
AAC CAG CAC AGC ATG GTG AGT GTG TCC ACG CCC ATC GCG GAC ATC TAC GAG
 Q   E   K   D   E   D   G   F   L   Y   M   V   Y   A   S   Q   E
CAG GAG AAA GAC GAG GAC GGC TTC CTC TAT ATG GTC TAC GCC TCC CAG GAA
 T   F   G   F
ACC TTC GGC TTC TGA
```

Figure 2.2: Protein and DNA sequence of LC3A protein used in this work after thrombin cleavage.

**LC3B:**

```
 G   S   M   P   S   E   K   T   F   K   Q   R   R   T   F   E   Q
GGA TCC ATG CCG TCG GAG AAG ACC TTC AAG CAG CGC CGC ACC TTC GAA CAA
 R   V   E   D   V   R   L   I   R   E   Q   H   P   T   K   I   P
AGA GTA GAA GAT GTC CGA CTT ATT CGA GAG CAG CAT CCA ACC AAA ATC CCG
 V   I   I   E   R   Y   K   G   E   K   Q   L   P   V   L   D   K
GTG ATA ATA GAA CGA TAC AAG GGT GAG AAG CAG CTT CCT GTT CTG GAT AAA
 T   K   F   L   V   P   D   H   V   N   M   S   E   L   I   K   I
ACA AAG TTC CTT GTA CCT GAC CAT GTC AAC ATG AGT GAG CTC ATC AAG ATA
 I   R   R   R   L   Q   L   N   A   N   Q   A   F   F   L   L   V
ATT AGA AGG CGC TTA CAG CTC AAT GCT AAT CAG GCC TTC TTC CTG TTG GTG
 N   G   H   S   M   V   S   V   S   T   P   I   S   E   V   Y   E
AAC GGA CAC AGC ATG GTC AGC GTC TCC ACA CCA ATC TCA GAG GTG TAT GAG
 S   E   K   D   E   D   G   F   L   Y   M   V   Y   A   S   Q   E
AGT GAG AAA GAT GAA GAT GGA TTC CTG TAC ATG GTC TAT GCT TCC CAG GAG
 T   F   G   M   K   L   S   V
ACG TTC GGG ATG AAA TTG TCA GTG TAA
```

Figure 2.3: Protein and DNA sequence of LC3B protein used in this work after thrombin cleavage.

**LC3C:**

```
 G   S   M   P   P   P   Q   K   I   P   S   V   R   P   F   K   Q
GGA TCC ATG CCG CCT CCA CAG AAA ATC CCA AGC GTC AGA CCC TTC AAG CAG
 R   K   S   L   A   I   R   Q   E   E   V   A   G   I   R   A   K
AGG AAA AGC TTG GCA ATC AGA CAA GAG GAA GTT GCT GGA ATC CGG GCA AAG
 F   P   N   K   I   P   V   V   V   E   R   Y   P   R   E   T   F
TTC CCC AAC AAA ATC CCG GTG GTA GTG GAG CGC TAC CCC AGG GAG ACG TTC
 L   P   P   L   D   K   T   K   F   L   V   P   Q   E   L   T   M
CTG CCC CCG CTG GAC AAA ACC AAG TTC CTG GTC CCG CAG GAG CTG ACC ATG
 T   Q   F   L   S   I   I   R   S   R   M   V   L   R   A   T   E
ACC CAG TTC CTC AGC ATC ATC CGG AGC CGC ATG GTC CTG AGA GCC ACG GAA
 A   F   Y   L   L   V   N   N   K   S   L   V   S   M   S   A   T
GCC TTT TAC TTG CTG GTG AAC AAC AAG AGC CTG GTC AGC ATG AGC GCA ACC
 M   A   E   I   Y   R   D   Y   K   D   E   D   G   F   V   Y   M
ATG GCA GAG ATC TAC AGA GAC TAC AAG GAT GAG GAT GGC TTC GTG TAC ATG
 T   Y   A   S   Q   E   T   F   G   C   L   E   S   A   A   P   R
ACC TAC GCC TCC CAG GAG ACA TTT GGC TGC CTG GAG TCA GCA GCC CCC AGG
 D   G   S   S   L   E   D   R   P   C   N   P   L
GAT GGG AGC AGC CTT GAG GAC AGA CCC TGC AAT CCT CTC TGA
```

Figure 2.4: Protein and DNA sequence of LC3C protein used in this work after thrombin cleavage.

**GABARAP:**

```
 M   K   F   V   Y   K   E   E   H   P   F   E   K   R   R   S   E
ATG AAG TTC GTG TAC AAA GAA GAG CAT CCG TTC GAG AAG CGC CGC TCT GAG
 G   E   K   I   R   K   K   Y   P   D   R   V   P   V   I   V   E
GGC GAG AAG ATC CGA AAG AAA TAC CCG GAC CGG GTG CCG GTG ATA GTA GAA
 K   A   P   K   A   R   I   G   D   L   D   K   K   K   Y   L   V
AAG GCT CCC AAA GCT CGG ATA GGA GAC CTG GAC AAA AAG AAA TAC CTG GTG
 P   S   D   L   T   V   G   Q   F   Y   F   L   I   R   K   R   I
CCT TCT GAT CTC ACA GTT GGT CAG TTC TAC TTC TTG ATC CGG AAG CGA ATT
 H   L   R   A   E   D   A   L   F   F   F   V   N   N   V   I   P
CAT CTC CGA GCT GAG GAT GCC TTG TTT TTC TTT GTC AAC AAT GTC ATT CCA
 P   T   S   A   T   M   G   Q   L   Y   Q   E   H   H   E   E   D
CCC ACC AGT GCC ACA ATG GGT CAG CTG TAC CAG GAA CAC CAT GAA GAA GAC
 F   F   L   Y   I   A   Y   S   D   E   S   V   Y   G   L
TTC TTT CTC TAC ATT GCC TAC AGT GAC GAA AGT GTC TAC GGT CTG TGA
```

Figure 2.5: Protein and DNA sequence of GABARAP protein used in this work after thrombin cleavage.

**GABARAPL1:**

```
 G   S   M   K   F   Q   Y   K   E   D   H   P   F   E   Y   R   K
GGA TCC ATG AAG TTC CAG TAC AAG GAG GAC CAT CCC TTT GAG TAT CGG AAA
 K   E   G   E   K   I   R   K   K   Y   P   D   R   V   P   V   I
AAG GAA GGA GAA AAG ATC CGG AAG AAA TAT CCG GAC AGG GTC CCC GTG ATT
 V   E   K   A   P   K   A   R   V   P   D   L   D   K   R   K   Y
GTA GAG AAG GCT CCA AAA GCC AGG GTG CCT GAT CTG GAC AAG AGG AAG TAC
 L   V   P   S   D   L   T   V   G   Q   F   Y   F   L   I   R   K
CTA GTG CCC TCT GAC CTT ACT GTT GGC CAG TTC TAC TTC TTA ATC CGG AAG
 R   I   H   L   R   P   E   D   A   L   F   F   F   V   N   N   T
AGA ATC CAC CTG AGA CCT GAG GAC GCC TTA TTC TTC TTT GTC AAC AAC ACC
 I   P   P   T   S   A   T   M   G   Q   L   Y   E   D   N   H   E
ATC CCT CCC ACC AGT GCT ACC ATG GGC CAA CTG TAT GAG GAC AAT CAT GAG
 E   D   Y   F   L   Y   V   A   Y   S   D   E   S   V   Y   G   K
GAA GAC TAT TTT CTG TAT GTG GCC TAC AGT GAT GAG AGT GTC TAT GGG AAA

TGA
```

Figure 2.6: Protein and DNA sequence of GABARAPL1 protein used in this work after thrombin cleavage.

**GABARAPL2:**

```
 G   S   M   K   W   M   F   K   E   D   H   S   L   E   H   R   C
GGA TCC ATG AAG TGG ATG TTC AAG GAG GAC CAC TCG CTG GAA CAC AGA TGC
 V   E   S   A   K   I   R   A   K   Y   P   D   R   V   P   V   I
GTG GAG TCC GCG AAG ATT CGA GCG AAA TAT CCC GAC AGG GTT CCG GTG ATT
 V   E   K   V   S   G   S   Q   I   V   D   I   D   K   R   K   Y
GTG GAA AAG GTC TCA GGC TCT CAG ATT GTT GAC ATT GAC AAA CGG AAG TAC
 L   V   P   S   D   I   T   V   A   Q   F   M   W   I   I   R   K
TTG GTT CCA TCT GAT ATC ACT GTG GCT CAG TTC ATG TGG ATC ATC AGG AAA
 R   I   Q   L   P   S   E   K   A   I   F   L   F   V   D   K   T
AGG ATT CAG CTT CCT TCT GAA AAG GCG ATC TTC CTG TTT GTG GAT AAG ACA
 V   P   Q   S   S   L   T   M   G   Q   L   Y   E   K   E   K   D
GTC CCA CAG TCC AGC CTA ACT ATG GGA CAG CTT TAC GAG AAG GAA AAA GAT
 E   D   G   F   L   Y   V   A   Y   S   G   E   N   T   F   G   F
GAA GAT GGA TTC TTA TAT GTG GCC TAC AGC GGA GAG AAC ACT TTT GGC TTC

TGA
```

Figure 2.7: Protein and DNA sequence of GABARAPL2 protein used in this work after thrombin cleavage.

**p62:**

```
 M   A   S   L   T   V   K   A   Y   L   L   G   K   E   D   A   A
ATG GCG AGC CTG ACC GTG AAA GCG TAT CTG CTG GGC AAA GAA GAT GCG GCG
 R   E   I   R   R   F   S   F   C   C   S   P   E   P   E   A   E
CGC GAA ATT CGC CGC TTT AGC TTT TGC TGC AGC CCG GAA CCG GAA GCG GAA
 A   E   A   A   A   G   P   G   P   C   E   R   L   L   S   R   V
GCG GAA GCG GCG GCG GGC CCG GGC CCG TGC GAA CGC CTG CTG AGC CGC GTG
 A   A   L   F   P   A   L   R   P   G   G   F   Q   A   H   Y   R
GCG GCG CTG TTT CCG GCG CTG CGC CCG GGC GGC TTT CAG GCG CAT TAT CGC
 D   E   D   G   D   L   V   A   F   S   S   D   E   E   L   T   M
GAT GAA GAT GGC GAT CTG GTG GCG TTT AGC AGC GAT GAA GAA CTG ACC ATG
 A   M   S   Y   V   K   D   D   I   F   R   I   Y   I   K   E   K
GCG ATG AGC TAT GTG AAA GAT GAT ATT TTT CGC ATT TAT ATT AAA GAA AAA
 K   E   C   R   R   D   H   R   P   P   C   A   Q   E   A   P   R
AAA GAA TGC CGC CGC GAT CAT CGC CCG CCG TGC GCG CAG GAA GCG CCG CGC
 N   M   V   H   P   N   V   I   C   D   G   C   N   G   P   V   V
AAC ATG GTG CAT CCG AAC GTG ATT TGC GAT GGC TGC AAC GGC CCG GTG GTG
 G   T   R   Y   K   C   S   V   C   P   D   Y   D   L   C   S   V
GGC ACC CGC TAT AAA TGC AGC GTG TGC CCG GAT TAT GAT CTG TGC AGC GTG
 C   E   G   K   G   L   H   R   G   H   T   K   L   A   F   P   S
TGC GAA GGC AAA GGC CTG CAT CGC GGC CAT ACC AAA CTG GCG TTT CCG AGC
 P   F   G   H   L   S   E   G   F   S   H   S   R   W   L   R   K
CCG TTT GGC CAT CTG AGC GAA GGC TTT AGC CAT AGC CGC TGG CTG CGC AAA
 V   K   H   G   H   F   G   W   P   G   W   E   M   G   P   P   G
GTG AAA CAT GGC CAT TTT GGC TGG CCG GGC TGG GAA ATG GGC CCG CCG GGC
 N   W   S   P   R   P   P   R   A   G   E   A   R   P   G   P   T
AAC TGG AGC CCG CGC CCG CCG CGC GCG GGC GAA GCG CGC CCG GGC CCG ACC
 A   E   S   A   S   G   P   S   E   D   P   S   V   N   F   L   K
GCG GAA AGC GCG AGC GGC CCG AGC GAA GAT CCG AGC GTG AAC TTT CTG AAA
 N   V   G   E   S   V   A   A   A   L   S   P   L   G   I   E   V
AAC GTG GGC GAA AGC GTG GCG GCG GCG CTG AGC CCG CTG GGC ATT GAA GTG
 D   I   D   V   E   H   G   G   K   R   S   R   L   T   P   V   S
GAT ATT GAT GTG GAA CAT GGC GGC AAA CGC AGC CGC CTG ACC CCG GTG AGC
 P   E   S   S   S   T   E   E   K   S   S   S   Q   P   S   S   C
CCG GAA AGC AGC AGC ACC GAA GAA AAA AGC AGC AGC CAG CCG AGC AGC TGC
 C   S   D   P   S   K   P   G   G   N   V   E   G   A   T   Q   S
TGC AGC GAT CCG AGC AAA CCG GGC GGC AAC GTG GAA GGC GCG ACC CAG AGC
 L   A   E   Q   M   R   K   I   A   L   E   S   E   G   R   P   E
CTG GCG GAA CAG ATG CGC AAA ATT GCG CTG GAA AGC GAA GGC CGC CCG GAA
 E   Q   M   E   S   D   N   C   S   G   G   D   D   D   W   T   H
GAA CAG ATG GAA AGC GAT AAC TGC AGC GGC GGC GAT GAT GAT TGG ACC CAT
 L   S   S   K   E   V   D   P   S   T   G   E   L   Q   S   L   Q
CTG AGC AGC AAA GAA GTG GAT CCG AGC ACC GGC GAA CTG CAG AGC CTG CAG
 M   P   E   S   E   G   P   S   S   L   D   P   S   Q   E   G   P
ATG CCG GAA AGC GAA GGC CCG AGC AGC CTG GAT CCG AGC CAG GAA GGC CCG
 T   G   L   K   E   A   A   L   Y   P   H   L   P   P   E   A   D
ACC GGC CTG AAA GAA GCG GCG CTG TAT CCG CAT CTG CCG CCG GAA GCG GAT
 P   R   L   I   E   S   L   S   Q   M   L   S   M   G   F   S   D
CCG CGC CTG ATT GAA AGC CTG AGC CAG ATG CTG AGC ATG GGC TTT AGC GAT
 E   G   G   W   L   T   R   L   L   Q   T   K   N   Y   D   I   G
GAA GGC GGC TGG CTG ACC CGC CTG CTG CAG ACC AAA AAC TAT GAT ATT GGC
 A   A   L   D   T   I   Q   Y   S   K   H   P   P   P   L
GCG GCG CTG GAT ACC ATT CAG TAT AGC AAA CAT CCG CCG CCG CTG TGA
```

Figure 2.8: Protein and DNA sequence of p62 used in this work after thrombin cleavage.

The WT, pS342 and L341V p62 AIM peptides used here were purchased from Genosphere Biotechnologies and Peptide Synthetics. The peptide sequences are listed in Figure 2.9. TBK1 Recombinant Human Protein was purchased from Invitrogen.

**WT p62 AIM:**
`SGGDDDWTHLSS`
**pS342 p62 AIM:**
`SGGDDDWTHL[pS]S`
**L341V p62 AIM:**
`SGGDDDWTHVSS`

Figure 2.9: Peptide sequences of the p62 AIM peptides used in this work.

| Solution | Concentration |
|---|---|
| 15% Resolving Buffer | 375 mM Tris, 15% acrylamide/methylene bisacrylamide (37.5:1 ratio) solution (v/v), 0.1% SDS (w/v), pH 8.8 |
| 20% Resolving Buffer | 375 mM Tris, 20% acrylamide/methylene bisacrylamide (37.5:1 ratio) solution (v/v), 0.1% SDS (w/v), pH 8.8 |
| 2x Gel Loading Dye | 100 mM Tris, 4% SDS (w/v), 20% glycerol (v/v), 0.2% Bromophenol Blue (w/v), 200 mM DTT, pH 6.8 |
| Blocking Buffer | 5 % powdered milk (w/v, Marvel), 20 mM Tris, 150 mM NaCl, 0.1% Triton X - 100 (v/v), pH 7.5 |
| CD Buffer | 10 mM potassium phosphate buffer, pH 7.0 |
| Cleavage Buffer | 10 mM Tris, 150 mM NaCl, 25 mM $CaCl_2$, (1 mM DTT for LC3A, LC3C and GABARAPL2), pH 8.4 |
| Desalt/MS Buffer | 25 mM ammonium acetate |
| Gel Filtration Buffer | 20 mM Tris, 150 mM NaCl, pH 7.5 |
| Ion Exchange Buffer A | 5 mM potassium phosphate buffer, (1 mM DTT for LC3A, LC3C and GABARAPL2), pH 7.0 |
| Ion Exchange Buffer B | 5 mM potassium phosphate buffer, 1 M NaCl, (1 mM DTT for LC3A, LC3C and GABARAPL2), pH 7.0 |
| ITC Buffer | 25 mM potassium phosphate buffer, 150 mM NaCl, pH 7.0 |
| Kinase Reaction Buffer | 50 mM Tris, 1 mM $MgCl_2$, 1mM ATP, 1mM DTT, pH 7.5 |
| Lysis Buffer | 10 mM Tris, 150 mM NaCl, (1 mM DTT for LC3A, LC3C and GABARAPL2), pH 7.5 |
| NMR Buffer | 25 mM potassium phosphate buffer, 20 mM NaCl, 5% (v/v) D2O, 0.02% (w/v) sodium azide, (1 mM TCEP for GABARAPL2), pH 7.0 |
| SDS Running Buffer | 25 mM Tris, 250 mM Glycine, 0.1% SDS (w/v) |
| Stacking Buffer | 125 mM Tris, 4% acrylamide/methylene bisacrylamide (37.5:1 ratio) solution (v/v), 0.1% SDS (w/v), pH 6.8 |
| Washing Buffer | 20 mM Tris, 150 mM NaCl, 0.1% Triton X - 100 (v/v), pH 7.5 |

Table 2.1: Composition of buffers used throughout this work. All buffers were made using milliQ water.

## 2.2 Chemicals

All chemicals were purchased from Sigma Aldrich or Fisher Scientific unless otherwise stated. Buffers for use throughout this work were produced according to the compositions described in Table 2.1.

## 2.3 Molecular Biology Techniques

### 2.3.1 Agar Plates

Bacteria were plated onto lysogeny broth (LB) Agar plates to produce single clonal colonies. LB Agar (40 g/L) was suspended in milliQ water before sterilisation. This was cooled and stored at room temperature until required. The solid LB Agar was then gently heated in a microwave until molten. After cooling to around 40°C ampicillin (100 $\mu$g/mL) was added. The molten agar was then poured into disposable Petri dishes. Plates which were not initially used were stored at 4°C until required.

### 2.3.2 Chemically Competent *E. coli* Cells

Aliquots of chemically competent *E.coli* cells were prepared to allow genetic transformation of plasmid DNA. A small-scale culture of the appropriate *E. coli* strain was diluted 1:100 in fresh LB media (50 mL) and incubated with shaking (37°C, 180 rpm) until the $OD_{595}$ reached ~0.5. These cultures are carried out without the presence of an antibiotic. Cultures were centrifuged (3,000 xg, 4°C) for 10 minutes before resuspension in ice-cold $MgCl_2$ (0.1M, 15 mL) solution. Cultures were again centrifuged (3,000 xg, 4°C) for 10 minutes before resuspension in ice-cold $CaCl_2$ (0.1 M, 15 mL) solution and incubation (on ice) for 45 minutes. Cultures were again centrifuged (3,000 xg, 4 °C) for 10 minutes before a final resuspension in ice-cold $CaCl_2$ and glycerol (0.1 M and 30% v/v respectively, 5 mL) solution. Cells were then split into 100 $\mu$l aliquots before storage at -85°C.

### 2.3.3 DNA Sequencing

Purified plasmid DNA was sent to the DNA Sequencing Facility, University of Nottingham Medical School, QMC where a 3130xl ABI PRISM Genetic Analyser was used. The pGEX-reverse primer (CCG GGA GCT GCA TGT GTC AGA GG) was used to sequence the insert DNA in the pGEX-4T-1 plasmid. The pGEX-forward primer (GGC TGG CAA GCC ACG TTT GGT G) was required to confirm the sequence of the GST-tag in the plasmid. Results were visualised and confirmed using Chromas 2.21.0, an example of this visualisation is shown in Figure 2.10.

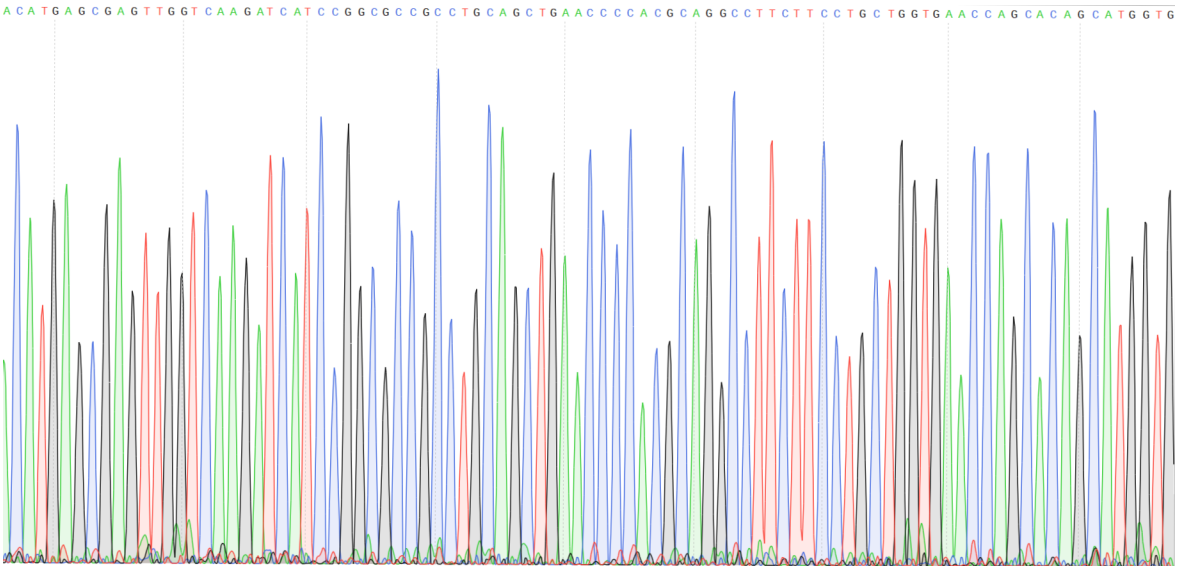


Figure 2.10: Visualisation of DNA sequencing chromatograms using Chromas software.

### 2.3.4 DNA Transformation

Genetic transformations of a DNA plasmid into various *E. coli* cells were required throughout this work. To do this ∼100 ng of purified plasmid DNA was added to a 100 $\mu$l aliquot of chemically competent cells which was then incubated (on ice) for 5 minutes. The sample was heat shocked (42°C) for 1 minute before further incubation (on ice) for 15 minutes. 800 $\mu$l of LB was added before incubation with shaking (37°C, 180 rpm) for one hour. 75 $\mu$l of this culture was then streaked on an agar plate and incubated (37°C) overnight.

### 2.3.5 *E coli.* cell culture

#### 2.3.5.1 Small-scale Cultures

Small scale overnight cultures were required for both protein over-expression and plasmid DNA purification. These cultures proceeded by preparation of 15 or 5 mL freshly sterilised LB media (25 g/L) for protein and DNA respectively, to which ampicillin was added (100 $\mu$g/mL). A small scraping of cells from the desired glycerol stock or a single colony from an agar plate was then added before the culture was incubated with shaking (37°C, 180 rpm) overnight.

#### 2.3.5.2 Protein Over-expression

An overnight culture of *E. coli* containing the relevant plasmid was used to inoculate freshly autoclaved LB media (25 g/L) at a dilution factor of 1:1000. This culture was then incubated with shaking (37°C, 180 rpm) until the $OD_{595}$ of the solution was ~0.7 AU. Protein over-expression was induced by the addition of IPTG (1 mL, 1 M) before incubation with shaking (20°C, 180 rpm) overnight. Cells were then harvested from the culture by centrifugation (3000 xg, 4°C) for 30 minutes. Cell pellets were stored at -80°C.

#### 2.3.5.3 Isotopically Labelled Protein Over-expression

M9 minimal media is composed of the chemical components listed in Table 2.2. 1L batches of this media were prepared and autoclaved. The chemicals listed in Table 2.3, sterilised by filtration, were then supplemented to this autoclaved media. Overnight cultures of *E. coli* cells possessing the required plasmid DNA were centrifuged (1,000 xg, 4°C) for 10 minutes to collect a cell pellet. The cell pellet was then resuspended in freshly prepared minimal media (20 mL) before incubation with shaking (37°C, 180 rpm) for 30 minutes. This was added to the supplemented minimal media (1 L) before incubation with shaking (37°C, 180 rpm) until the $OD_{595}$ of the solution was ~0.8 AU. Over-expression was induced by the addition of IPTG (1 mL, 1 M) before incubation with shaking (30°C, 180 rpm) overnight. Cells were then harvested from the culture by centrifugation (3000 xg, 4°C) for 30 minutes. Cell pellets were stored at -80°C.

| **Component** | **Concentration (g/L)** |
|---|---|
| Disodium hydrogen orthophosphate | 6 |
| Potassium dihydrogen orthophosphate | 3 |
| Sodium chloride | 0.5 |
| Magnesium sulphate | 0.3 |
| Calcium chloride dihydrate | 0.015 |

Table 2.2: Composition of minimal media before autoclaving.

| Component | Concentration (mg/L) |
|---|---|
| Iron chloride hexahydrate | 400 |
| Zinc chloride | 80 |
| Ammonium molybdate tetrahydrate | 20 |
| Copper chloride dihydrate | 20 |
| Manganese chloride tetrahydrate | 20 |
| Sodium tetraborate decahydrate | 20 |
| Biotin | 10 |
| Thiamine | 10 |
| MEM vitamins solution 100X | Varying, according to the manufacturer |
| d-Glucose-$^{13/12}C_6$ | 2000/4000 ($^{13}C_6/^{12}C_6$) |
| Ammonium-$^{15}N$ chloride | 1000 |

Table 2.3: Supplements which are sterilised by filtration before addition to minimal media.

### 2.3.6 Glycerol Stocks

Cells grown in small-scale cultures from single colonies on agar plates were stored in glycerol (25%, v/v) at -85°C.

### 2.3.7 Plasmid DNA Purification

5 mL overnight cultures of DH5$\alpha$ *E. coli* cells containing the required plasmid DNA were prepared. A Sigma GenElute Plasmid MiniPrep Kit was then used to extract plasmid DNA from these pellets according to manufacturer's instructions.

### 2.3.8 Site-directed Mutagenesis

Mutation of plasmid DNA was carried out using the NEB Q5 site-directed mutagenesis kit. This involved a modified version of the supplied protocol which involves making up the reaction mixture in Table 2.4 and using the polymerase chain reaction (PCR) conditions specified in Table 2.5. All primers were purchased from Sigma-Aldrich.

| Reagent | Volume / $\mu$l (Concentration) |
|---|---|
| Q5 Hot Start High-Fidelity 2x Master Mix | 12.5 |
| Forward Primer | 1.25 (10 $\mu$M) |
| Reverse Primer | 1.25 (10 $\mu$M) |
| Template DNA | 1 (25 ng/$\mu$l) |
| Nuclease-free $H_2O$ | 9 |

Table 2.4: Mutagenesis reaction mixture

| Step | Temperature | Time |
|---|---|---|
| Stage I: Initial Denaturation | 98°C | 30 seconds |
| Stage II (25 cycles): Denaturating | 98°C | 10 seconds |
| Annealing | 61-72°C (depending on primer) | 30 seconds |
| Extension | 72°C | 4 minutes |
| Stage III: Final Extension | 72°C | 5 minutes |
| Stage IV: Hold | 4°C | Hold |

Table 2.5: Thermocycler conditions for mutagenesis reaction

The Q5 kit includes a kinase, ligase, *Dpn1* (KLD) treatment enzyme mix to remove template DNA and recircularise PCR product at room temperature. Reagents were assembled according to Table 2.6, mixed and incubated (room temperature)‘ for 10 minutes. This DNA sample was then stored at 4°C until it was transformed into DH5$\alpha$ chemically competent *E. coli* cells.

| Reagent | Volume / $\mu L$ |
|---|---|
| PCR product | 1 |
| 2x KLD Reaction Buffer | 5 |
| 10x KLD Enzyme Mix | 1 |
| Nuclease-free water | 3 |

Table 2.6: KLD Reaction Mixture

### 2.3.9 Sterilisation

All glassware, plastic consumables (which were not supplied sterile), growth media and buffers were sterilised before use by autoclaving (121°C, 15 minutes) using a Rodwell Ensign Autoclave. Heat sensitive reagents (ampicillin, ammonium chloride, etc.) were sterilised separately. These were dissolved in milliQ, filtered with a 0.22 $\mu$m pore size filter and added when required, without autoclaving.

## 2.4 Biochemical Techniques

### 2.4.1 Kinase Binding Assays

Cell pellets from *E. coli* cultures (10 mL) overexpressing the GST-tagged p62 protein were produced as described in section 2.3.5.2. These cell pellets were resuspended in Washing Buffer (1 mL) and cOmplete™ EDTA-free Protease Inhibitor Cocktail (100 $\mu$l, Roche). The cells were then lysed by sonication. This proceeded at 4°C with three repeats of alternating 10 second sonication and 10 second rest intervals. Cell debris was removed from the sample by centrifugation (13000 xg for 20 minutes).

This lysate was then incubated in a 1:1 ratio with Glutathione Sepharose 4B beads (100 $\mu$L beads used for each condition to be investigated, GE Healthcare) with rotation (4°C) for three hours. Beads were incubated in Washing Buffer with rotation (4°C) for 5 minutes, centrifuged (2000 xg for 1 minute) and the buffer was removed by aspiration This was repeated with Washing Buffer and then again with Kinase Reaction Buffer. Beads were resuspended in Kinase Reaction Buffer and split according to the number of conditions to be investigated. After the split, Kinase Reaction Buffer (500 $\mu$L) and TBK1 (where required, 500 ng) were added (negative control experiments without TBK1 were also carried out). These reactions were incubated with shaking (37°C) for one hour. Beads were incubated in Washing Buffer with rotation (4°C) for 5 minutes, centrifuged (2000 xg for 1 minute) and the buffer was removed by aspiration; this was carried out twice. Beads were resuspended in Washing Buffer (500 $\mu$L) and the hAtg8 protein of interest (2 $\mu$g) was added. These samples were incubated with shaking (37°C) for one hour. Beads were incubated in Washing Buffer with rotation (4°C) for 5 minutes, centrifuged (2000 xg for 1 minute) and the buffer was removed by aspiration; this was carried out three times. Beads were then resuspended in a 2:1 mixture of water:2x Gel Loading Dye (50 $\mu$L total for each condition). 10 $\mu$L of this sample was then used in each SDS-PAGE gel lane for detection by Western Blot.

### 2.4.2 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) was used to separate proteins within a sample by molecular weight for visualisation during purifications or binding assays. SDS-PAGE gels were poured using the Bio-Rad Mini-Protean 3 Gel casting System according to manufacturer's instructions. Ammonium

persulfate (40 $\mu$L, 1.1% w/v) then N,N,N,N-Tetramethylethylenediamine (TEMED, 10 $\mu$L, 0.3% v/v) were added to the appropriate percentage Resolving Buffer (15% or 20%, 3.5 mL, Table 2.1) before pouring into the casting system. After this was set, Ammonium persulfate (30 $\mu$L, 1.1% w/v) and TEMED (10 $\mu$L, 0.3% v/v) were added to Stacking Buffer (2.5 mL, Table 2.1) and poured on top of the Resolving Gel in the casting system.

Samples were prepared by adding 2x Gel Loading Dye (Table 2.1) and heating (95°C) for 5 minutes. Novex Sharp Pre-Stained Protein Standard (3.5-260 kDa) was used as a standard to allow estimation of the molecular weight of a given band. The electrophoresis was carried out at 180V for 50 minutes (15% gel) or 200V for 70 minutes (20% gel) using a BioRad PowerPac Basic and a Bio-Rad Mini-Protean 3 cell containing SDS Running Buffer (Table 2.1). Gels were stained using GelCode Blue Safe Protein Stain according to manufacturer's instructions.

### 2.4.3 Western Blot

Proteins were detected as part of binding assays using western blotting. SDS-PAGE gels were run as described in section 2.4.2 but they were not stained. Instead, they were transferred onto a 0.45 $\mu$m nitrocellulose membrane (GE) using a Fisherbrand Electrophoresis Unit, assembled according to manufacturers instructions. This transfer occurred at 40 mA (room temperature) overnight. The membrane was then incubated in Blocking Buffer with rocking (room temperature) for 2 hours before incubation in the primary antibody with rocking (varying concentration in Blocking Buffer, 4°C) overnight. The membrane was submerged in Washing Buffer (3x 5 minute intervals) before incubation in the secondary antibody with rocking (varying concentration in Blocking Buffer, room temperature) for 1 hour. The membrane was then submerged in Washing Buffer (3x 5 minute intervals) before incubation with ECL Western Blotting Substrate (room temperature, Pierce) for one minute. Bands were visualised using X-ray film (Pierce).

## 2.5 Protein Purification

### 2.5.1 Affinity Chromatography

The GST fusion proteins produced in this work were initially purified using affinity chromatography. Cells pellets from protein over-expression (1L of original culture) were thawed at room temperature and then resuspended in Lysis Buffer (Table 2.1,

4 mL) with the addition of DNase 1 (10 mg/mL, 100 $\mu$L) and cOmplete™ EDTA-free Protease Inhibitor Cocktail (100 $\mu$l, Roche). The cells were then lysed by sonication. This proceeded at 4°C for 10 minutes at 10 microns with alternating 30 second sonication and 30 second rest intervals. After sonication, the cell debris was removed from the sample by centrifugation (30000 xg for 30 minutes then 35000 xg for 30 minutes, both at 4°C). The supernatant was then incubated with rotation (4°C, 30 rpm) with Glutathione Sepharose 4B beads (GE Healthcare) which had been equilibrated in Lysis Buffer (Table 2.1). After incubation, the beads were washed with Lysis Buffer (Table 2.1) before equilibration into Cleavage Buffer (Table 2.1). The column was incubated with rotation (4°C, 30 rpm) overnight with human $\alpha$-thrombin (15 Units/Litre of original culture) in Cleavage Buffer (Table 2.1). Cleaved protein was then eluted from the beads, by gravity, in Cleavage Buffer (Table 2.1, 4 mL).

### 2.5.2 Fast Protein Liquid Chromatography

Two FPLC systems were used for the purification of proteins over-expressed in *E coli.* culture; an ÄKTA Prime and ÄKTA Start system (GE Healthcare). Fractions containing the desired protein were identified using the ultraviolet (UV) detector of the FPLC system measuring absorbance at 280 nm and by SDS-PAGE. These fractions were combined, frozen and lyophilised. Lyophilised protein was then stored at -20°C until required.

#### 2.5.2.1 Cationic Exchange

A 5 mL HiTrap SP HP anionic exchange column (GE Healthcare) attached to the ÄKTA Start was equilibrated in Ion Exchange Buffer A (Table 2.1). Protein sample from the affinity chromatography step was diluted in milliQ water until its conductivity was below that of Ion Exchange Buffer A ($\sim$5 mS/cm). This was loaded onto the column at 2 mL/min before washing with 25 mL Ion Exchange Buffer A (Table 2.1). The bound protein was then eluted using a 0-40% linear gradient of Ion Exchange Buffer B (Table 2.1) over 300 mL, collected in 10 mL fractions.

#### 2.5.2.2 Gel Filtration

A Superdex 75 Gel Filtration column (Amersham Biosciences) attached to the ÄKTA Prime was used after equilibration in Gel Filtration Buffer (Table 2.1). Protein samples were injected and run through the column at 3 mL/min and collected in 10 mL fractions.

#### 2.5.2.3 Desalt

A (5 x 5 mL) HiTrap Desalt column (Amersham Biosciences) attached to the ÄKTA Prime was equilibrated in Desalt/MS Buffer (Table 2.1). Lyophilised protein samples were resuspended in Desalt/MS buffer and were run through the column at 2 mL/min. Fractions were collected manually.

## 2.6 Biophysical Techniques

### 2.6.1 Circular Dichroism (CD) Spectroscopy

An Applied Photophysics Chirascan-Plus fitted with a Quantum Northwest temperature controller (JASCO UK) was used to make CD measurements. Protein samples (10 $\mu$M) were made up in CD Buffer (Table 2.1) and loaded into a 1 mm pathlength quartz cuvette (Hellma Analytics).

#### 2.6.1.1 Units

CD experiments produce a measurement of ellipticity (deg) which can be converted to the standardised measure of mean residue ellipticity (MRE), with units of deg cm$^2$ dmol$^{-1}$ residue$^{-1}$,using equation 2.1. MRE is a measure of the average ellipticity of individual residues which is useful for the comparison of proteins of varying size or samples of different concentration.

$$MRE = \frac{\theta \times 10^6}{C.N.l} \tag{2.1}$$

$\theta$ is the ellipticity with the blank subtracted (mdeg); C is the protein concentration (M); N is the number of backbone amide bonds (number of residues - 1); l is the cell path length (cm).

#### 2.6.1.2 Protein Secondary Structure

CDNN software can be used to estimate the proportions of different secondary structure elements present in the protein being investigated. To achieve this, five scans were performed at room temperature by scanning wavelengths between 190-260 nm with a bandwidth of 1 nm. Data were sampled at a rate of 0.5 s$^{-1}$. The collected data were averaged and baseline corrected using a blank measurement before use in CDNN. This software utilises three neural nets which have been trained on different databases of protein CD spectra with known structure. All three neural nets are used

for deconvolution of the secondary structure and the best fit was then selected. This selection occurs based on the fit which puts the total secondary structure content closest to 100% whilst using the widest range of possible wavelengths.

#### 2.6.1.3 Thermal Denaturation Experiments

Thermal denaturation of protein samples was investigated by collecting CD spectra between 190-260 nm at 5°C intervals as the sample was heated between 5°C and 80°C at a rate of 1°C/min. Sample temperature was determined with the ESHU temperature probe placed inside the sample, in the cuvette.

### 2.6.2 Mass Spectrometry

The utility of mass spectrometry (MS) in protein research has continually expanded, particularly since the development of electrospray ionisation (ESI) in the late 1980s [152, 153]. This "soft" ionisation technique allows proteins to enter the gas phase, *in vacuo,* without fragmentation. Furthermore, conditions may be optimised to allow the protein to retain a native-like fold upon desolvation via ESI.

The actual folded state of a protein which has undergone ESI is the subject of debate as there is no evidence to definitively confirm the protein remains structurally identical to the solution phase; hence the use of the term native-like. The removal of water as a solvent may have a number of effects on structure depending on the protein. The effect of hydrophobic packing will be removed due to the lack of solvent and any electrostatic interactions at the surface between polar or charged side chains and water molecules will be removed.

A key piece of evidence supporting the retention of protein fold during ESI comes from the observed charge states. Acid-induced denaturation of a protein was observed by Konermann *et al.* by monitoring the charge state distributions observed in ESI MS [154]. As charge is added at the protein surface during ESI, unfolded proteins produce higher charge state ions due to their larger surface area. This indicates the proteins detected with lower charge state distributions must be folded. In addition, proteins can be shown to have similar binding affinities using ESI titration techniques to solution phase determination of affinity, supporting the retention of native-like protein folding [155].

Regardless of the specifics of this ionisation technique, intermolecular interactions between proteins can survive this process as peaks corresponding to the various charge states of protein-protein complexes can be observed [156, 157, 158]. The relative ratios

of bound and free species present in the solution can be maintained in the gas phase.

One key advantage to the use of MS for studies of protein-protein interactions comes from the ability to investigate the interactions of multiple proteins in the same sample. Mass differences between proteins make it possible to deconvolute the peaks which arise from each individual protein and protein-protein complex. This allows proteins to be studied in a more biologically relevant context where there are multiple competing binding interactions present within the solution.

#### 2.6.2.1 Data Acquisition

Initial experiments to confirm protein identity and purity were performed under denaturing conditions using a Bruker Ultraflex III Mass Spectrometer using matrix-assisted laser dissociation ionisation (MALDI) and a time-of-flight mass analyser. MALDI was also used to check for degradation of protein samples after storage. Samples (1 mg/mL in milliQ) were prepared as a 2:1 Super-DHB matrix:protein mixture.

Native MS experiments were performed between two spectrometers. Initial experiments were performed on a SYNAPT High Definition Mass Spectrometry System (Waters) using electrospray ionisation and a quadrupole time of flight (qTOF) mass analyser. Samples (varying concentration) were dissolved in Desalt/MS Buffer (Table 2.1) and were injected into the sample chamber by direct infusion at a flow rate of 5 $\mu$L/min using a syringe pump and a 100 $\mu$L Hamilton Syringe. The mass spectrometer was operated in positive ion mode under the following optimized conditions: capillary voltage, 2.5 kV; cone voltage, 40 V; trap CE, 8 V; transfer CE, 5 V; backing pressure, $\sim$3.8 mbar; trap pressure, 2.1 $\times$ 10$-$2. Spectra were acquired between 500 to 4000 m/z for 2 min. Instrument control as well as data processing was performed using MassLynx$^{\text{TM}}$ 4.1 software.

The majority of experiments were then performed on a Bruker Impact II Mass Spectrometer using electrospray ionisation and a quadrupole time of flight (qTOF) mass analyser. Samples were injected into the sample chamber by direct infusion at a flow rate of 2 $\mu$L/min using a syringe pump and a 500 $\mu$L Hamilton Syringe. The mass spectrometer was operated in both positive and negative ion mode, depending on the constituents of the sample. Positive ion mode was used under the following optimized conditions: end plate offset, 500 V; capillary voltage, 3.5 kV; nebulizer pressure, 0.6 bar; dry gas pressure, 4 L/min; temperature, 150$^{\text{o}}$C. Negative ion mode was used under the following optimized conditions: end plate offset, 400 V; capillary voltage, 3 kV; nebulizer pressure, 0.6 bar; dry gas pressure, 4 L/min; temperature,

150°C. Spectra were acquired between 500 to 3000 m/z for 1 min. Instrument control was performed and data processing was performed using Bruker software.

#### 2.6.2.2 Data Analysis

Data from competition experiments were analysed by addition of the sum of intensities of peaks corresponding to each species present in the sample; protein only, peptide only or protein-peptide complex. This allowed the calculation of the ratio of bound to unbound protein or peptide.

### 2.6.3 Isothermal Titration Calorimetry (ITC)

ITC allows the key thermodynamic parameters of molecular interactions to be quantified [159]. This technique has expanded in both utility and popularity and is now regularly used to define the binding characteristics of protein-protein interactions, as it was used here [160, 161].

In this method, peptide aliquots were sequentially injected into a protein sample and the heat change due to their interaction was detected. The calorimeter operates by maintaning the sample cell at the same temperature as a reference cell. The parameter that is monitored therefore is the power applied to maintain this reference temperature in the sample cell. This results in power spikes upon each injection of peptide into the sample cell. These spikes are integrated and this allows the production of a binding curve where the energy per mole of injectant is plotted against the molar ratio of the two proteins in the cell.

Based on this binding curve the number of binding sites on the protein (n), the binding constant (K), the enthalpy change ($\Delta H$), the entropy change ($\Delta S$) and the free energy change ($\Delta G$) can be fitted. These values define the strength and nature of a protein-protein interaction [162].

ITC experiments were performed in ITC Buffer (Table 2.1) at 30°C (unless otherwise stated) on a MicroCal VP-ITC (Malvern). The reference power was set to 5 $\mu$cal/sec and the syringe stirring speed was 300 rpm. In each experiment there was an initial sample equilibration time of 1 hour. 10 $\mu$L aliquots of peptide stock solution were sequentially injected into the ITC cell (containing the protein sample) and the heat pulse was measured. Blank experiments were also carried out by injection of the peptide sample into ITC buffer alone; to allow subtraction of the heat of dilution of the peptide sample. All experiments were carried out in duplicate with the derived parameters averaged after calculation.

#### 2.6.3.1 Sample Preparation

Particular care must be taken during the preparation of ITC samples to prevent air bubbles in the cell. As such, ITC buffer was degassed before being used to resuspend protein or peptide samples. To ensure sample buffer homogeneity, both peptide and protein samples were dialysed against the same buffer using a 1kDa MWCO mini dialysis kit (GE Healthcare). The protein samples were dialysed overnight whereas peptides were dialysed for 2 hours, both at 4°C. Sample concentration was then determined using UV spectroscopy before samples were diluted in ITC buffer to the desired concentration. Despite the presence of free solvent exposed thiols from cysteine residues of LC3A, LC3C and GABARAPL2, no reducing agents were used in these samples due to potential issues with baseline alteration.

#### 2.6.3.2 Data Analysis

ITC titration data were analysed using MicroCal Analysis software to determine the thermodynamic parameters of binding using a global fit standard non-linear least-squares regression analysis. Initially the blank experimental data were subtracted before fitting of the data.

The equation used by this software is initially derived by considering the effect of the increasing volume in the cell after each injection [163]. Due to the total-fill cell the extra volume injected is outside of the calorimetrically sensed region. So the number of moles of both the protein and the peptide in the active volume of the cell is diluted on each injection giving a slightly lower concentration than the hypothetical bulk. A quadratic equation can be derived by combining two equations; one related to the binding constant and the number of occupied and unoccupied sites, and the other the relationship between the bulk and free ligand concentration. Solving this quadratic and substituting an equation for the total heat content of the solution gives equation 2.2. This allows for the determination of heat content in terms of the number of sites (n), the binding constant (K) and the enthalpy change ($\Delta H$).

$$Q = \frac{nM_t\Delta HV_o}{2}\left[1 + \frac{X_t}{nM_t} + \frac{1}{nKM_t} - \sqrt{\left(1 + \frac{X_t}{nM_t} + \frac{1}{nKM_t}\right)^2 - \frac{4X_t}{nM_t}}\right] \tag{2.2}$$

$M_t$ is the bulk concentration of protein; $V_o$ is the active cell voume; $X_t$ is the bulk concentration of peptide. This equation is then used to consider the change in

heat from injection $i$-1 to $i$, which is the heat effect that will be observed for the $i$th injection. The volume displaced by a given injection is considered to only contribute 50% of a heat effect as the volume remaining in the cell to give an additional term in the final calculation; equation 2.3.

$$\Delta Q(i) \;=\; Q(i) \;+\; \frac{dV_i}{V_o}\left[\frac{Q(i) \;+\; Q(i-1)}{2}\right] \;-\; Q(i-1) \tag{2.3}$$

The fitting of experimental data using equation 2.3 involves i) initial estimation of thermodynamic parameters (n, K and $\Delta H$); ii) inputting thermodynamic parameter values into equation 2.3 for each injection; iii) comparison of modelled data with experimentally observed heat effect by calculation of the sum of squared residuals; iv) alteration of thermodynamic parameters using standard Marquadt methods v) repeated iterations of steps ii-iv until no further significant reduction in the sum of the squared residuals occurs. From these fitted thermodynamic parameters the entropy change ($\Delta S$) may also be calculated by first determining the free energy ($\Delta G$) with equation 2.4 which can then be input into equation 2.5; along with the other known parameters.

$$\Delta G \;=\; -RTlnK \tag{2.4}$$

$$\Delta S \;=\; \frac{(\Delta H \;-\; \Delta G)}{T} \tag{2.5}$$

### 2.6.4 Protein Nuclear Magnetic Resonance (NMR) Spectroscopy

Proteins were first analysed by NMR in the 1950s and 60s [164, 165, 166]. 2D NMR experiments were then published in the 1970s and this advancement along with technical improvements to spectrometers began to allow the assignment of protein spectra [167].

This work was extended by the incorporation of NMR-active isotopes, mainly $^{13}$C and $^{15}$N, which allowed coupling constants to be more reliably detected. $^{13}$C and $^{15}$N atoms are detected at larger ppm ranges than $^{1}$H atoms so spreading resonances into more dimensions with larger ranges led to vastly improved resolution of signals.

Triple resonance experiments were then developed which produce more resolved spectra which contain more information [168]. This increase in information comes at the cost of decreased sensitivity compared with 2D experiments.

NMR experiments were carried out on a Bruker AV(III)800 (Bruker UK Limited, Coventry) spectrometer with a QCI cryoprobe. Standard Bruker pulse sequences were used at 298 K. 1D solvent suppression was carried out using excitation sculpting with gradients whereas all multidimensional experiments used Watergate solvent suppression. Quadrature detection of indirect dimensions was achieved using States-TPPI [169].

#### 2.6.4.1 Data Acquisition

NMR samples were prepared by resuspension of lyophilised protein in NMR Buffer (Table 2.1) before centrifugation (13,000 xg, room temperature) for 10 minutes. Data acquisition parameters were selected based on the protein concentration of the sample and the sensitivity of a given experiment, these are outlined in Table 2.7.

#### 2.6.4.2 Data Analysis

Data processing was carried out using Topspin 3.5 and spectra were then analysed using CcpNmr Analysis 2.4.2 [170]. All data were zero filled by at least a factor of 2 but a larger factor was used in some cases. The degree of zero filling can be considered as a balance between apparent resolution and file size, as it can be particularly limiting to use multiple large 3D datasets simultaneously. An exponential window function was used for processing 1D experiments whereas a shifted squared sine bell window function was used for multidimensional experiments. The offset was optimised for each experiment individually. For some experiments, forward linear prediction of complex data along particular heteronuclear dimensions was employed. This was optimised for each experiment individually by considering whether there is an apparent increase in resolution and whether this outweighs the increase in file size and signal-to-noise.

Backbone N-H assignment of GABARAPL2 ($^{1}$H, $^{15}$N for 52% of non-prolyl residues) was carried out via a series of 3D experiments to confirm a published assignment [171]. The assignment was achieved here using a combination of 2D and 3D experiments ($^{1}$H-$^{15}$N-HSQC, HNCO, HN(CA)CO, HNCA, HNCACB and HN(CO)CACB) [172, 173, 174, 175]. Despite full backbone assignment not being achieved, as all assignment work confirmed the previous work the remaining assignment of Ma et al was used in the data analysis for titration experiments.

The $^{1}$H-$^{15}$N-HSQC was the key experiment used in the protein backbone assignment here as this experiment shows the backbone N-H signals which were to be assigned along with some distinguishable side chain N-H shifts. These non-backbone

| | Scans | Spectral Width | Transmitter frequency | fid size |
|---|---|---|---|---|
| $^{1}$H 1D | 64 | 20 | 4.7 | 32768 |
| $^{31}$P 1D | 64 | 15.018 | 4.7 | 32768 |
| $^{1}$H-$^{15}$N-HSQC | 160 | F2:13, F1:28 | F2:4.7, F1:118 | 2048x256 |
| $^{1}$H-$^{15}$N-HSQC (titration) | 64 | F2:16, F1:40 | F2:4.7, F1:115 | 2048x128 |
| HNCO | 32 | F3:13, F2:28, F1:22 | F3:4.7, F2:118, F1:176 | 2048x40x128 |
| HN(CA)CO | 32 | F3:13, F2:28, F1:22 | F3:4.7, F2:118, F1:176 | 2048x40x128 |
| CBCANH | 64 | F3:13, F2:28, F1:22 | F3:4.7, F2:118, F1:176 | 2048x36x128 |
| CBCA(CO)NH | 64 | F3:13, F2:28, F1:22 | F3:4.7, F2:118, F1:176 | 2048x36x128 |

Table 2.7: Selected data acquisition parameters for NMR experiments.

N-H signals came predominantly from tryptophan N$\epsilon$-H$\epsilon$, asparagine N$\delta$-H$\delta_2$ and glutamine N$\epsilon$-H$\epsilon_2$ groups. To carry out the backbone assignment, all peaks in the $^{1}$H-$^{15}$N-HSQC were picked and assigned an arbitrary number. Asparagine and glutamine side chain N-H signals were identified as they result in two peaks at the same nitrogen but differing hydrogen shifts in a particular region of the spectrum. Tryptophan side chain N-H signals are less clear however they typically have a high $^{1}$H ppm value resulting in them being in an otherwise clear region of the spectrum. $^{13}$C chemical shifts for CO, C$\alpha$ and C$\beta$ nuclei were then assigned to each $^{1}$H-$^{15}$N-HSQC peak from the various 3D experiments, representing a defined spin system with an arbitrary number.

Backbone assignment was then carried out using standard sequential assignment methods by using pairs of related experiments; in this case these were HNCO/HN(CA)-CO and HNCACB/HN(CO)CACB experiments. The magnetisation transfer of these experiments are illustrated in Figures 2.11 and 2.12. Here we see that in each pair of experiments one can be used to identify the $i$-1 $^{13}$C peaks of a given spin system and the other gives the peaks corresponding to both the $i$-1 and $i$ residue. This allows the sequential walking through the sequence by correlation of these $^{13}$C shifts.

The sequential assignment was then carried out using an iterative method. Initially non-ambigious sequential links were found using HNCACB/HN(CO)CACB experiments. These were considered tentative links. Sequential links were then found in the HNCO/HN(CA)CO experiments. Any initial tentative links which were supported by the HNCO/HN(CA)CO experiments could then be considered confirmed links. By alternating between scans through each experiment type, assignment of some links led to previously ambigious links becoming clear. This was iterated until no further sequential assignment could be achieved with confidence.

The next step involved identifying the distinguishing characteristics of some residues which allow assignment of amino acid type to a spin system. For example, glycine has a characteristic C$\alpha$ chemical shift and no C$\beta$ shifts. Comparison of a string of sequentially linked residues where some spin systems have been assigned an amino acid type to the sequence of the protein then allowed assignment to specific residues in the protein sequence.

Titration experiments with unlabelled peptide were conducted at 298 K by collecting $^{1}$H-$^{15}$N-HSQC spectra of $^{15}$N-hAtg8 proteins. CcpNmr Analysis was used to follow shift changes during titration experiments. This allows the magnitude of peak shift in the $^{1}$H-$^{15}$N-HSQC to be calculated and the chemical shift perturbation (CSP) to be plotted for each residue and used for further analysis. CSP values for each residue were calculated using equation 2.6 [176].

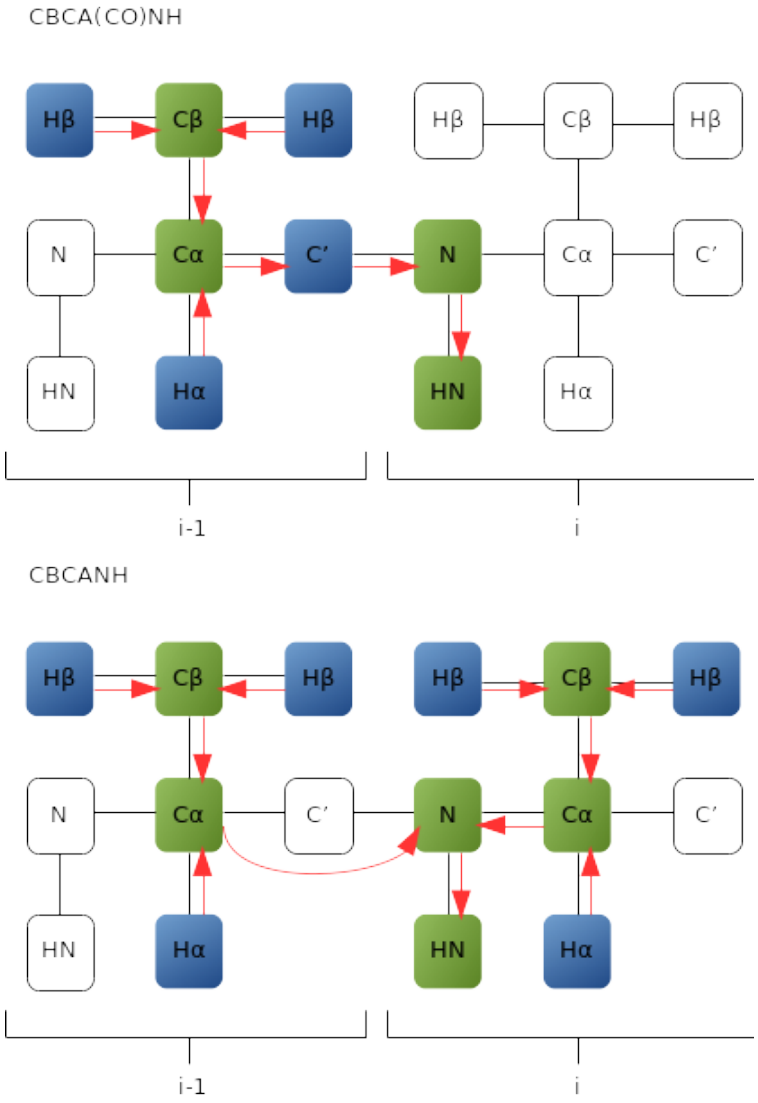



Figure 2.11: Schematic illustrating the magnetisation transfer during HNCACB (top panel) and HN(CO)CACB. Arrows indicate flow of magnetisation. The shifts of nuclei in green boxes are observed whereas the magnetisation simply flows through nuclei in blue boxes. In both experiments the amide proton is used for detection.

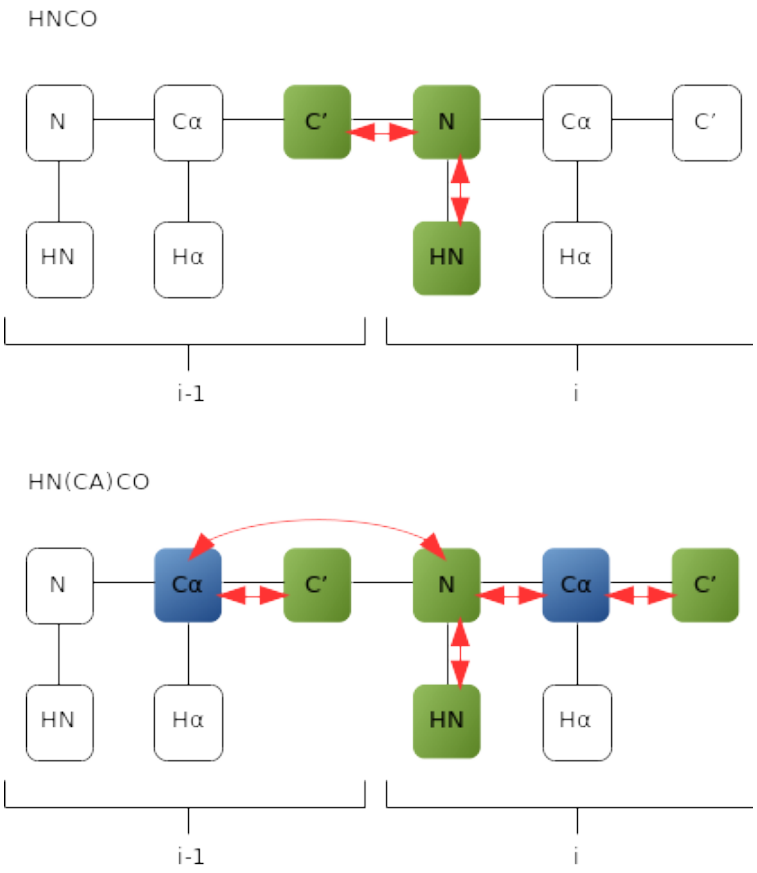



Figure 2.12: Schematic illustrating the magnetisation transfer during HNCO (top panel) and HN(CA)CO. Arrows indicate flow of magnetisation. The shifts of nuclei in green boxes are observed whereas the magnetisation simply flows through nuclei in blue boxes. In these experiments the magnetisation starts on amide proton before transfer to subsequent nuclei before returning to the amide proton for detection.

$$\Delta\delta_{HSQC} = \left[\frac{1}{2}\left[\Delta\delta_H^2 + \left(0.14 \times \Delta\delta_N^2\right)\right]\right]^{\frac{1}{2}} \quad (2.6)$$

$\Delta\delta_{HSQC}$ is the CSP (ppm); $\Delta\delta_H$ is the shift in the $^1$H dimension (ppm); $\Delta\delta_N$ is the shift in the $^{15}$N dimension (ppm).

In the cases where peaks could not be clearly followed through the titration to their end point, due to slow exchange behaviour, a minimum chemical shift procedure was used [177, 178]. This simply involves identifying the closest unassigned peak in the bound $^1$H-$^{15}$N-HSQC spectrum and assuming that this is the minimum shift possible for the given peak. Shifts were then considered to be significant if the CSP was greater than a threshold value determined by Equation 2.7.

$$Threshold \geq \mu + x\sigma \quad (2.7)$$

$\mu$ is the population mean and $\sigma$ is the standard deviation. In this equation $x$=2 for the highest significance threshold and $x$=1 for the lower significance threshold. CSP values from the final titration point were used for this analysis.

CSPs observed can also be used to determine the binding affinity of an interaction using the CcpNmr analysis software. This produces a plot of the CSP at a given titration point versus the peptide:protein ratio. The plotted values are then fit to Equation 2.8 to determine a $K_D$ value.

$$y = A\left[(B + x) - \left((B + x)^2 - 4x\right)^{\frac{1}{2}}\right] \quad (2.8)$$

$$A = \frac{Maximum\,CSP}{2}$$

$$B = \frac{1 + K_D}{[Protein]}$$

$$x = \frac{[Peptide]}{[Protein]}$$

$$y = CSP$$

### 2.6.5 UV Spectroscopy

Both DNA and protein concentrations were determined by measurement of UV absorbance at characteristic wavelengths. A NanoDrop ND-1000 Spectrophotometer was used to measure absorbance of DNA samples at 260 and 280 nm and protein samples at 280 nm. Protein extinction coefficients were estimated, based on their sequence, using the online ExPASy Prot Param server [179].

UV spectroscopy was also used to provide a measure of DNA purity by looking at the ratio of absorbances at various wavelengths. The 260/280 nm ratio was the primary ratio used to investigate purity, with a ratio above 1.8 deemed acceptable. Additionally, a 260/230 nm ratio in the range of 2.0-2.2 was expected for a pure DNA sample. DNA samples which did not meet these criteria were discarded.

# Chapter 3

# Protein Expression and Purification

## 3.1 Protein Constructs

The protein-protein interaction studies in this work required the production of the six hAtg8 proteins. Plasmid DNA samples were supplied by Dr. Jed Long which correspond to the full length hAtg8 proteins in their unprocessed form. The desired DNA sequence which was translated into the protein of interest had been cloned into the pGEX-4T-1 plasmid. The choice of this construct is discussed here as well as the implications of that choice.

Due to the selected pGEX-4T-1 plasmid, the hAtg8 proteins were expressed with an N-terminal glutathione S-transferase (GST)-tag followed by a thrombin cleavage site before the sequence of the protein of interest. Thrombin recognises the amino acid sequence LVPRGS and selectively cleaves the amide bond between the arginine and glycine residues. Therefore, the proteins after thrombin cleavage to remove the tag have an N-terminal extension of Gly-Ser. This small extension was deemed acceptable due to the utility of the GST-tag for protein purification and because it was unlikely to cause significant changes in the structure or function of the proteins.

In addition to this N-terminal extension, the residues at the C-terminus of these constructs may be considered. The first step in the lipidation reaction of hAtg8 proteins, which leads to their addition to the autophagosomal membrane, involves the removal of a C-terminal stretch of the protein. This exposes a glycine residue with a free carboxylate group which is the site where the protein is eventually tethered to the PE class of phospholipids. The hAtg8 constructs produced and used in this work

retained the C-terminal residues after this glycine. hAtg8 proteins at this stage in the autophagic process are sometimes referred to as pro-LC3.

The potential issue from using this construct is that it may not represent the form of the protein most relevant to the interaction with p62, which is the focus of this thesis. The lipidated form of the hAtg8 proteins attached to the membrane would be the most relevant as this is the context in which the hAtg8-p62 interaction occurs during autophagy *in vivo*. However, there are factors which can be considered to support the use of a full length, non-lipidated construct within this work.

Based upon the LC3B-p62 AIM structure shown in Figure and other known hAtg8-AIM structures, the region of the hAtg8 proteins which interacts with the AIM of p62 is distant from these extra C-terminal residues so their presence or absence may be irrelevant. It is possible however that the lipidation of hAtg8 may affect the overall structure of the protein. If this is the case though then having some bulk attached to this glycine may mimic the bulk of the PE. This may allow the C-terminal tail on the pro-LC3 construct used here to be representative of the phospholipid attached to the protein in vivo. This bulk would not be present if a construct with the free C-terminal glycine were utilised.

Finally, characterisation of the interaction strengths of membrane bound proteins was considered likely to produce less informative results than the untethered protein used here as the same techniques could not be used on both systems. For example, the biophysical characterisation of free proteins can be done using solution NMR whereas it is likely that solid state NMR would be required for the membrane bound protein. One of the key differences between these techniques is the lower sensitivity and resolution of solid state compared with solution NMR. Additionally, there are significant challenges involved in the use of ITC in the study of membrane bound protein. As it has been previously shown that the difference in binding affinity between the LC3B-WT p62 AIM and LC3B-L341V p62 AIM interactions is threefold, it was expected that other relevant differences in this study may be equally small and thus the sensitive techniques used here are required to observe these potentially slight differences. The following two sections detail the expression and purification of the hAtg8 constructs required for this biophysical characterisation.

## 3.2 Protein Over-expression

These proteins were produced using a conventional *E. coli* over-expression methodology due to its ease of use. The suitability of this system was initially tested and

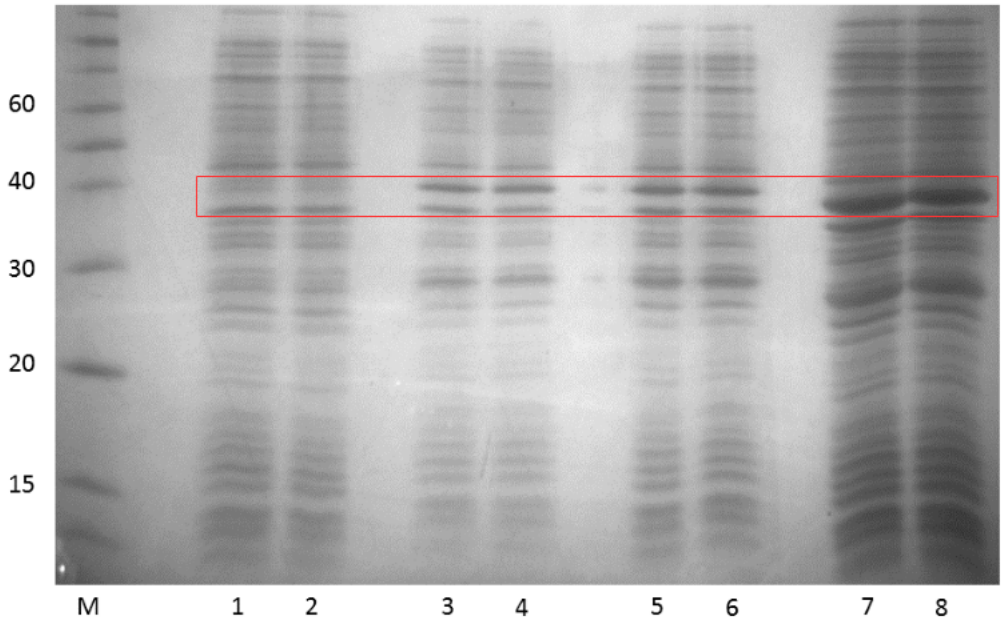



Figure 3.1: SDS-PAGE gel showing the expression test of LC3B highlighting the protein band that appears over time which corresponds to the protein of interest. This shows the solubility of the protein as the bands in the total and soluble lanes of each time point appear to have an equal density. It also illustrates the tightly controlled expression of this protein as no protein is observed before addition of IPTG. Lane loading: M) Markers 1) Uninduced total lysate 2) Uninduced soluble lysate 3) Total lysate 2 hours after induction 4) Soluble lysate 2 hours after induction 5) Total lysate 4 hours after induction 6) Soluble lysate 4 hours after induction 7) Total lysate after overnight induction 8) Soluble lysate hours after overnight induction.

the results are described here. The supplied stocks of the pGEX-4T-1 plasmid with hAtg8 insertions were confirmed to contain the desired DNA insert by sequencing. Plasmid DNA from each construct was then transformed into BL21 (DE3) chemically competent E. coli cells. Small scale expression tests of these constructs were carried out in a range of conditions.

Figure 3.1 shows an expression test for LC3B as an example where the protein was well expressed after IPTG induction (1 mM) at 18°C overnight, which were the optimum conditions found for both yield and solubility of the protein. This SDS-PAGE gel shows a clear expression band which appears over time and runs at a similar rate through the gel to the 40 kDa marker. This was expected for the GST-tagged LC3B fusion protein which has a predicted mass of 40998.6 Da. This also shows that the protein of interest was not being expressed at observable concentrations before IPTG induction i.e. there is no 'leaky expression'. It was also clear that overnight expression produces the highest quantity of protein, compared to the other time points tested, as this appeared to have the most intense protein band.

By comparison of the amount of a protein present in the total and soluble fractions, the solubility of a protein under the given conditions can be evaluated. The total fraction contains all of the cell lysate whereas the soluble fraction is collected after centrifugation of this lysate to remove larger debris. Therefore, if a more intense band is observed in the total fraction than for the soluble fraction there must be some protein involved in larger insoluble aggregates which are removed by centrifugation. Reduction of insoluble protein aggregates will allow for increased protein yield. As there appeared to be equal amounts of the protein of interest in the total and soluble fractions when run on an SDS-PAGE gel it can be surmised that expression of insoluble protein was negligible. This data is representative of the expression tests for all six hAtg8 proteins and as such these expression conditions were used for all six hAtg8 proteins produced in this work.

## 3.3 Protein Purification

Purification of the proteins from cell lysate was a three-step chromatographic process; affinity chromatography then ion exchange (IEX) chromatography and finally size exclusion (SEX) chromatography. Glutathione Sepharose 4B beads were used for the affinity chromatography step. These are cross-linked agarose beads presenting a glutathione ligand which can bind with high affinity to GST. This allows the selective capture of GST-tagged protein from crude cell lysate.

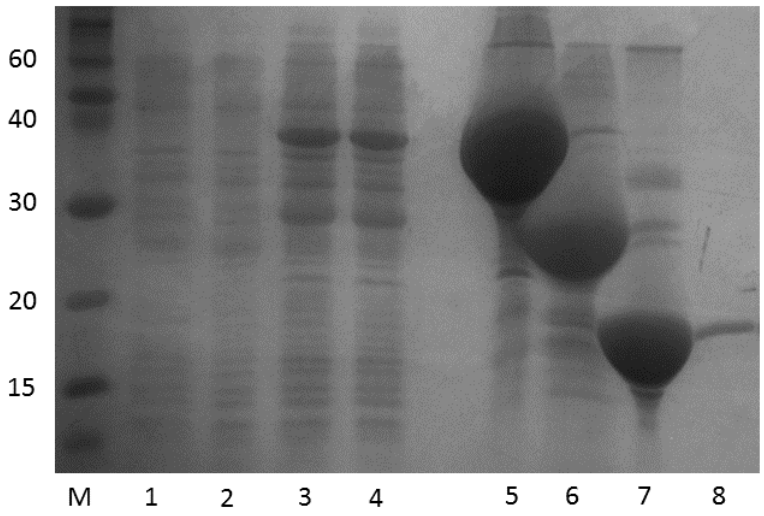



Figure 3.2: SDS PAGE gel showing the contents of the sample at various stages of the purification of LC3A. Lanes 1-4 illustrate the expression of the protein before incubation with Glutathione Sepharose 4B beads. The proteins which bind to the beads are shown in lane 5. The beads are then incubated with thrombin which cleaves off the GST tag, leaving it bound to the beads as shown in labe 6. The sample eluted from this affinity column and which is applied to the IEX column is shown in lane 7. The final purified protein sample is shown in lane 8.

Figure 3.2 shows the purification steps for GST-tagged LC3A as an example which is representative of the process for all six hAtg8 proteins. This shows the protein content of samples at the various stages throughout the purification, separated and visualised on an SDS-PAGE gel. Lanes one and two show cell lysates from before the over-expression of protein was induced, highlighting the lack of 'leaky expression'. The GST-LC3A fusion protein can then be seen in the cell lysate in lanes three and four after IPTG induction, lane three is the total fraction and lane four is the soluble fraction. This prominent band in lanes three and four runs close to the 40 kDa protein marker, as expected due to the predicted GST-tagged protein mass of 40583.03 Da.

This cell lysate is then incubated with Glutathione Sepharose 4B beads and subsequently washed, which concentrates the protein and removes substantial impurities, as shown in lane five. Overnight incubation with thrombin then cleaves the majority of the bound protein leaving the tag on the beads, shown in lane six. This allows the desired protein to be eluted from the beads. The two largest bands in lane six correspond to the GST-LC3A protein which has not been cleaved and the GST tag on its own after cleavage which remains bound to the beads. Lane seven shows the sample eluted from the beads after cleavage, this is a substantially purer protein sample but

| Protein | Theoretical pI |
|---|---|
| **LC3A** | 8.75 |
| **LC3B** | 8.90 |
| **LC3C** | 9.15 |
| **GABARAP** | 8.75 |
| **GABARAPL1** | 8.69 |
| **GABARAPL2** | 8.00 |

Table 3.1: Theoretical pI values for the hAtg8 proteins calculated using the EXPASY Prot Param server from their sequence. This value is useful in determination of protein charge in relation to the pH of the buffering solution.

some impurities are still present.

Fast protein liquid chromatography (FPLC) was then used for the two remaining chromatographic steps in the purification. In this technique, an aqueous buffer mobile phase is passed over a solid resin matrix stationary phase inside a plastic column. The residence time of sample components through the column and subsequently the separation of sample components is affected by several properties. The two key protein properties utilised for separation in this work are size and charge.

IEX chromatography allows the separation of proteins within a sample on the basis of their charge. Protein charge is affected by the pH of the buffer they are in as proteins in different protonation states have varying charges. For this technique the solid resin matrix carries a charge due to functionalisation by a variety of ligands. In this purification the resin was functionalised with sulfopropyl (SP) groups to produce a negatively charged matrix which bound positively charged species. In Ion Exchange Buffers A and B (see Table 2.1), which have a pH of 7, proteins with an isoelectric point (pI) above 7 were positively charged and therefore bound to the negatively charged stationary phase.

The estimated pI values of the hAtg8 proteins are shown in Table 3.1, predicted based on their sequence using the EXPASY Prot Param server. The values for the six hAtg8 proteins are all above 7 so an SP column was used for them all. In addition to ensuring the pI value is above the pH of the buffer to produce a positively charged protein, the pI must also be significantly different from the pH. This is important as proteins in buffers with a pH close to their pI value will have reduced solubility in water due to a lack of surface charge. Typically, the pH and pI values are different by at least one to ensure there is a charge present as the calculation of pI based on sequence is prone to error and this ensures significant difference.

The protein sample eluted from the Glutathione Sepharose 4B beads after cleavage

| Protein | Theoretical Mass | Measured Mass |
|---|---|---|
| **LC3A** | 14416.6 | 14363.9 |
| **LC3B** | 14832.2 | 14834.7 |
| **LC3C** | 16995.7 | 16921.5 |
| **GABARAP** | 14062.2 | 14023.9 |
| **GABARAPL1** | 14188.2 | 14117.4 |
| **GABARAPL2** | 13810.0 | 13739.9 |

Table 3.2: Summary of the theoretical and measured masses of the hAtg8 proteins.

was applied to the SP column and then washed. All of this was done in buffer with low conductivity to ensure protein binding to the functionalised beads. Protein bound to the matrix was then eluted used a salt gradient as the increasing ionic strength of the mobile phase decreases the interaction between protein and the stationary phase. Lane eight in Figure 3.2 shows a pure LC3A sample after gradient elution from an SP column, where fractions have been combined and concentrated.

SEX chromatography separates proteins on the basis of their size by utilising a porous resin stationary phase. The separation occurs as smaller molecules can enter the pores in the resin whereas larger molecules are more likely to be excluded. Smaller molecules therefore have a longer residency time in the column as they travel a longer distance and as such are eluted later than larger molecules. The principle of SEX was used here for buffer exchange as the protein in a sample flows through with a shorter residency time than the smaller buffer and salt molecules in the sample, exchanging the protein into the buffer which has been used for the mobile phase. An ammonium acetate buffer (Desalt/MS Buffer, see Table 2.1) was used as the mobile phase here as this allows all buffer to be removed by repeated lyophilisation. This occurs due to the volatility of ammonium acetate. This produced a protein sample, devoid of buffer or salt, as a solid which can be resuspended into any desired buffer for the experiments to follow in this thesis.

Matrix-assisted laser dissociation ionisation (MALDI) MS was used to confirm protein identity and purity after the purification process described above [180]. These spectra are shown for all six hAtg8 proteins, after resuspension in water, in Appendix Figures A.1-A.6. These spectra show a single pure protein in varying charge states; with 1+, 2+ and 3+ states observed depending on the protein. The observed mass of these proteins showed some difference from the expected mass as shown in Table 3.2. However, they are within the error of the machine and still represent the correct protein as supported by sequencing of the plasmid DNA and subsequent tracking of the protein by SDS PAGE gel. This was also supported by further MS work to

follow, which gives a more accurate measurement of protein mass. MALDI MS also confirmed the low salt concentration of these samples as there were few salt adduct peaks observed. The proteins were therefore considered to be of sufficient purity for use in further experiments.

## 3.4 CD and NMR Spectroscopy

Circular dichroism (CD) spectroscopy was used here to investigate the folding of the hAtg8 proteins after lyophilisation and resuspension. The technique is based on the differential absorption of circularly polarised light (CPL) by optically active, chiral molecules. CD has been used extensively in the study of proteins, as well as other biological macromolecules. Proteins are optically active as 19 of the 20 commonly occurring natural amino acids possess a chiral centre at their $C_\alpha$.

More importantly here, the secondary structure of a protein leads to further asymmetry which will lead to further differential absorbance of CPL. This occurs because proteins undergo two key quantum transitions upon absorption of light by the amide group; n→pi* and pi→pi*. These transitions are dependent on the $\varphi$ and $\psi$ angles present in the protein which are themselves indicative of particular secondary structure. This means the CD spectrum of a protein is not simply the sum of its individual residues. For this reason, CD spectra recorded in the far-UV (190-260 nm) can be used to estimate the amount of the various secondary structural elements which are present in a folded protein. The characteristic spectra of "pure" secondary structural elements are shown in Figure 3.3.

Figure 3.4 shows the CD spectra of the six hAtg8 proteins after resuspension of lyophilised sample in CD Buffer (see Table 2.1). Visual inspection of the line shapes can be useful here. Initially, it was clear that all proteins investigated have some non-random coil secondary structure present due to differential absorbances occurring at wavelengths of light other than the minimum at around 200 nm typical of a random coil. This can be considered indicative of folded protein samples.

There was variation between the proteins in the line shape of their spectra which is indicative of some degree of differing structure. The line shape for each protein is more related to those for proteins in the same subfamily than to those in the other subfamily. There is also some variation within each subfamily. The LC3 proteins all show a similar line shape but all three appear to show some differences. Within the GABARAP subfamily, GABARAP and GABARAPL1 have a similar line shape with GABARAPL2 showing more substantial differences.

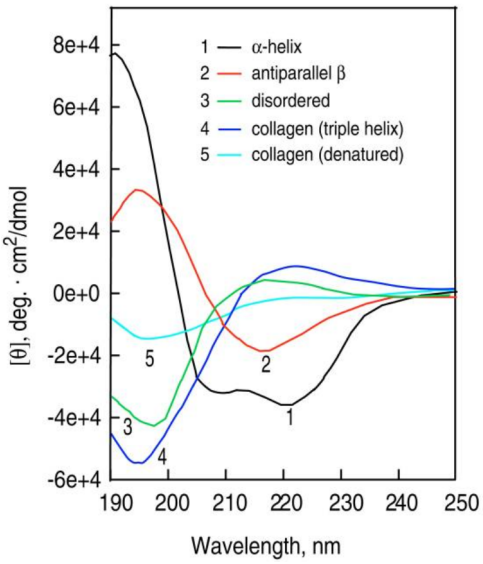



Figure 3.3: CD spectra as reference of the absorbances which correspond to each of the "pure" secondary structural elements taken from Ref. [2].

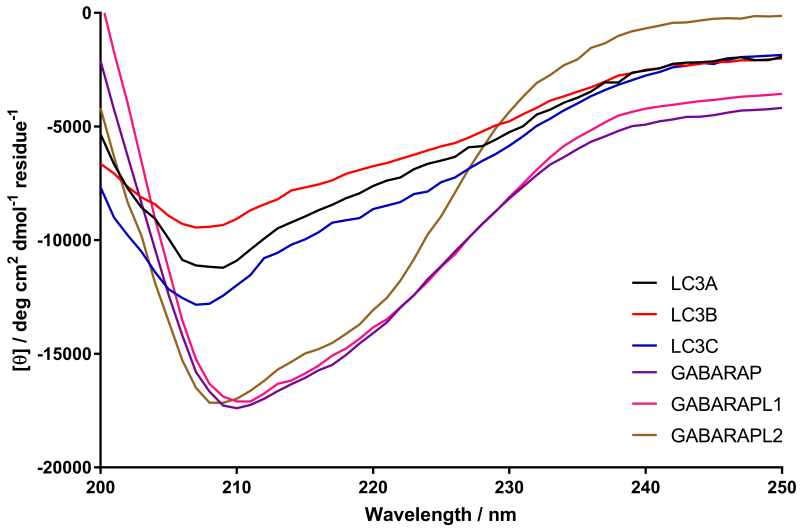



Figure 3.4: CD spectra of the hAtg8 proteins to show that some structural fold is present for all proteins. This also highlights the similarities within each of the subfamilies in terms of line shape and therefore secondary structure.

| | LC3A | | LC3B | | LC3C | |
|---|---|---|---|---|---|---|
| | **CDNN** | **Predicted** | **CDNN** | **Predicted** | **CDNN** | **Predicted** |
| **$\alpha$-helix** | 40% | 31% | 29% | 29% | 19% | 25% |
| **$\beta$-strand** | 17% | 20% | 20% | 21% | 18% | 20% |
| | **GABARAP** | | **GABARAPL1** | | **GABARAPL2** | |
| | **CDNN** | **Predicted** | **CDNN** | **Predicted** | **CDNN** | **Predicted** |
| **$\alpha$-helix** | 31% | 32% | 33% | 34% | 28% | 33% |
| **$\beta$-strand** | 24% | 22% | 20% | 23% | 23% | 22% |

Table 3.3: Proportion of secondary structural elements present in the hAtg8 proteins as calculated from CD data by CDNN compared to the expected values based on structures from the PDB.

Estimation of the proportion of each secondary structural element present in a protein was carried out from this data using CDNN software. This is done using neural nets which have been trained using machine learning against three different databases of CD spectra with corresponding structures from X-ray crystallography or NMR experiments. The results of this calculation for each hAtg8 protein and the predicted values based on their structures from the PDB are shown in Table 3.3. For each protein, the proportion of helix and beta sheet was roughly similar to that present in the crystal structure. There are some clear differences in these CD-derived and structure-derived values though. For example, the largest discrepancies occur for LC3A. These CDNN results roughly confirm what was observed by simply looking at the line shape of the spectra in terms of inter- and intra-subfamily variation.

CD spectroscopy is also useful for the investigation of protein stability as the spectrum can be measured at different environmental conditions such as temperature, pH or denaturant concentration. This allows the denaturation of the protein structure to be monitored as conditions are incrementally changed. Temperature was used as the measure of stability in this work. Figure 3.5 shows the CD thermal denaturation experiments for the hAtg8 proteins. A full scan CD spectrum was collected roughly every 5°C as the sample was heated from 5 to 75°C. In general, these spectra show that as the temperature was increased the magnitude of the various peaks decreased as well as some changes in the line shape, indicative of some structural transition of the protein. As with the room temperature line shapes of the spectra, the character of the melt was most similar among proteins in the same subfamily. For example, the shift in line shape for GABARAP and GABARAPL1 throughout the melt appears to be of a similar character.

The melting transition of the proteins can be further characterised by plotting the

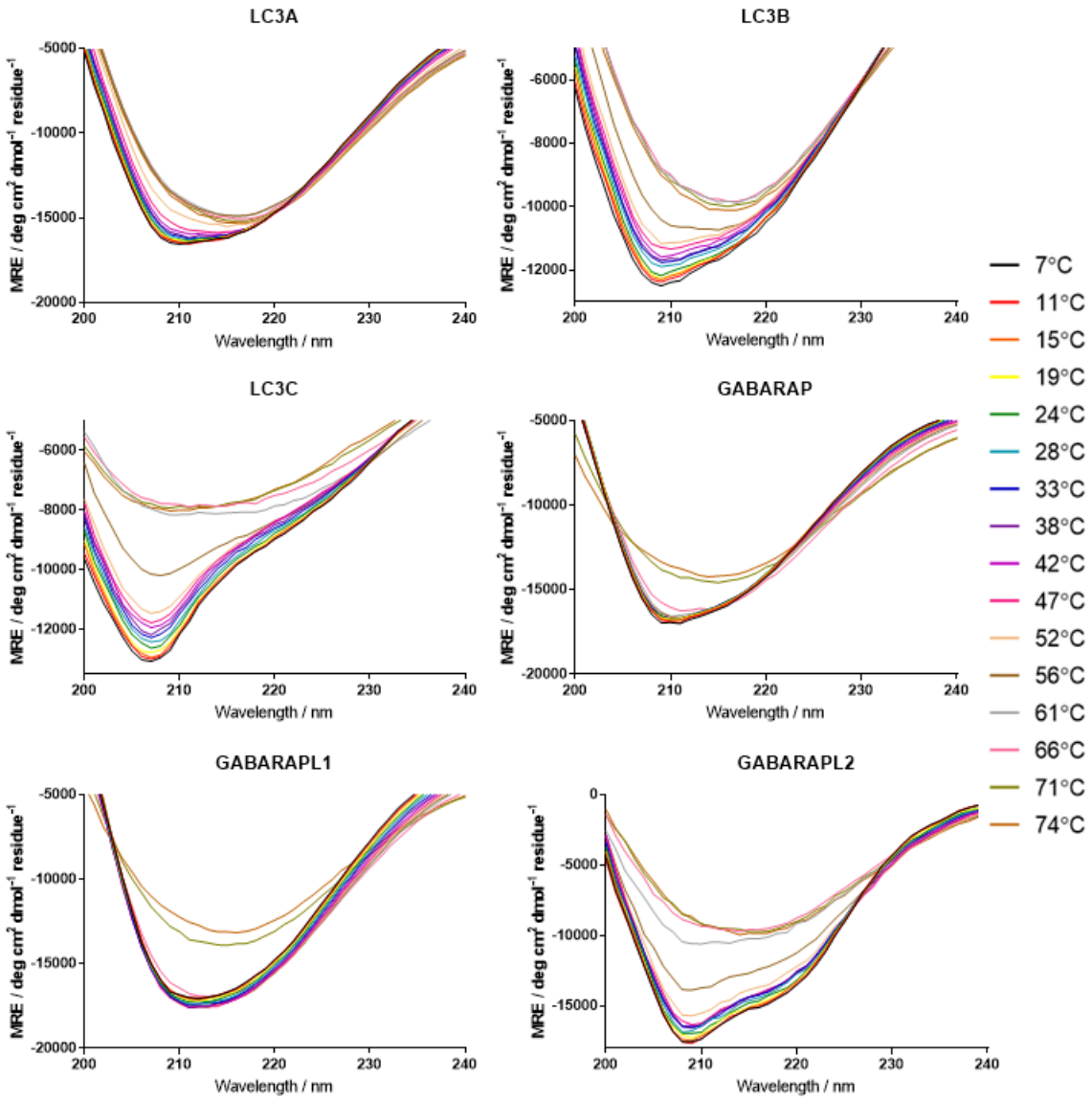



Figure 3.5: CD spectra of the hAtg8 proteins at varying temperature throughout a thermal denaturation experiment. This allowed comparison of the melting transition of the proteins.

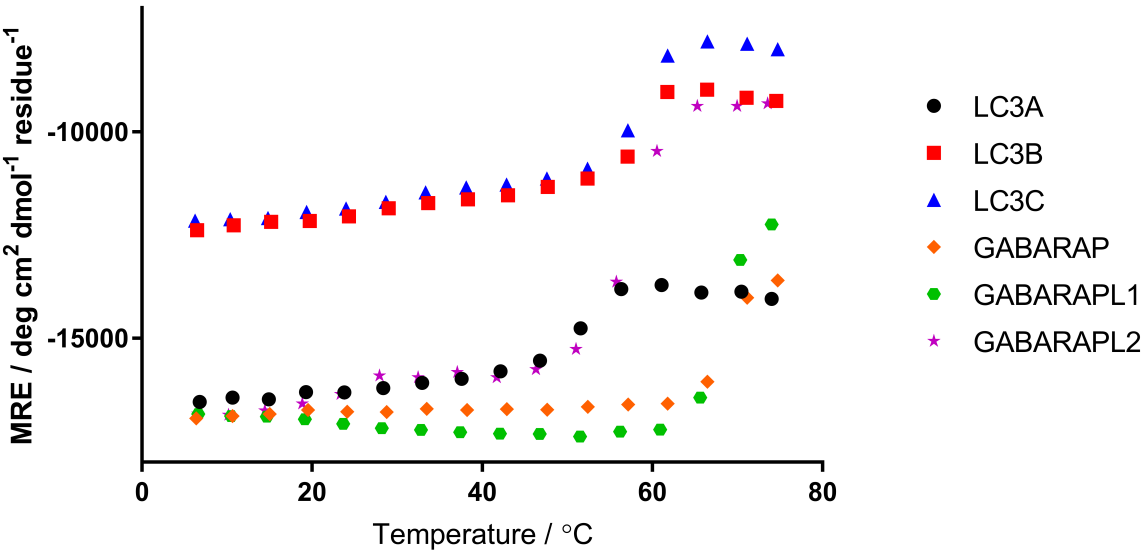



Figure 3.6: Melt curves of the hAtg8 proteins derived from CD thermal denaturation experiments which allow calculation of $T_m$ values.

mean residue ellipticity (MRE) value for a selected wavelength over the course of the experiment. These melt curves can be seen in Figure 3.6 by following the MRE value for each protein at 210 nm. This wavelength was selected to follow the transition as it is the wavelength at which the MRE shifts the most for these proteins so this provides the clearest picture of the melt. Again, these curves are best considered in their respective subfamilies. The melting transitions observed for the LC3 subfamily show highly similar melts for LC3B and LC3C with that for LC3A occurring at a lower temperature. The $T_m$ values for LC3A, LC3B and LC3C were determined by fitting of the curves as 51.4, 57.2 and 57.2°C respectively. For the GABARAP subfamily it is GABARAP and GABARAPL1 which have the most similar melt curves with the melting transition for GABARAPL2 occurring at a lower temperature. The $T_m$ values for GABARAP, GABARAPL1 and GABARAPL2 were determined by fitting of the curves as 67.1, 69.8 and 56.0°C respectively.

NMR spectroscopy was also a useful tool for this initial protein characterisation. 1D $^1$H NMR spectra of proteins suffer from a lack of resolution due to the low ppm range of the technique and the substantial number of signals which are detected. Despite this lack of resolution, 1D spectra may be used for some preliminary assessment of the structure of proteins. Folded proteins will typically show well dispersed peaks, this is particularly noticeable in the amide proton region of the spectrum that is typically from around 6 to 9 ppm in a $^1$H spectrum. There may also be some upfield

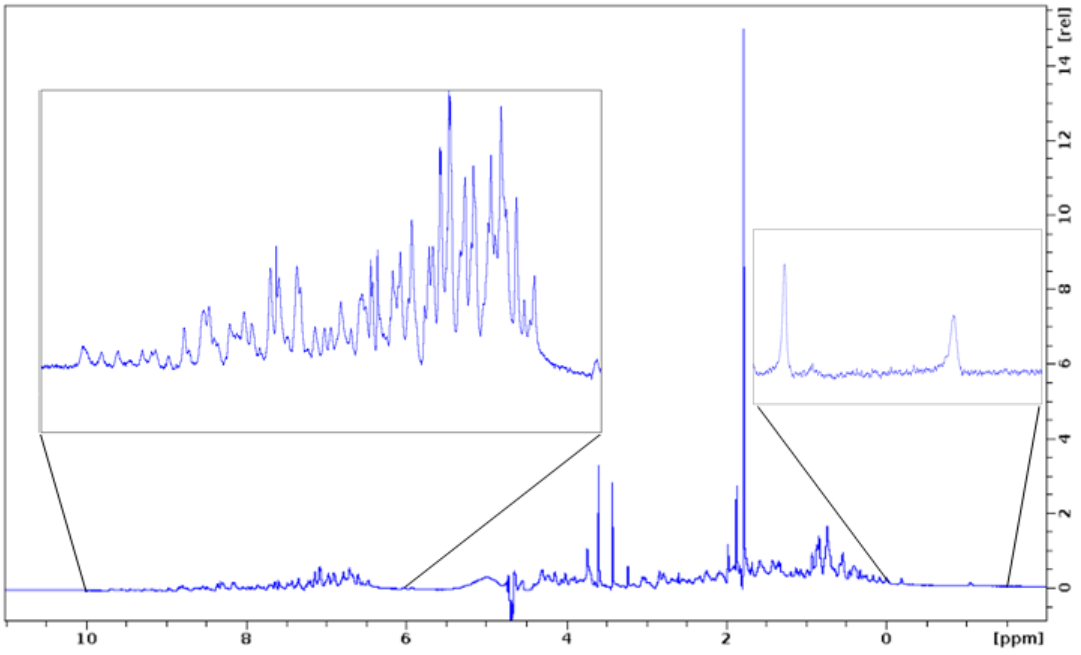



Figure 3.7: $^{1}$H NMR spectrum of LC3A which shows indications of folded protein structure. This folding is implied by upfield shifted methyl signals and a broad distribution of the amide proton peaks which are highlighted by the zoomed-in sections of the spectra.

shifted methyl groups which appear at ppm values below 0. This occurs due to the packing of the protein which can result in ring current effects on methyl groups. So 1D $^{1}$H NMR is a useful initial test for protein folding.

Figure 3.7 shows the 1D $^{1}$H NMR spectrum of LC3A when resuspended in NMR Buffer (see Table 2.1) as an example spectrum with features representative though differing from those observed for the remaining hAtg8 proteins. The spectrum shows a broad distribution of peaks in the amide region and there are two upfield shifted methyl peaks occurring at -0.19 and -1.05 ppm. This suggests the sample contains folded protein. The other spectra show a distribution of amide peaks and upfield shifted methyl peaks and as such are also considered to be folded. The spectra for the other five hAtg8 proteins are shown in Appendix Figures A.7-A.11.

## 3.5 Discussion

The six hAtg8 proteins to be studied were shown to express well in the *E. coli* BL21 (DE3) system; with high yield and no issues with solubility under the stated

conditions. A purification method involving affinity, IEX and SEX chromatography was shown to effectively purify these proteins from cell lysate, as confirmed by MALDI MS experiments.

The hAtg8 proteins produced CD spectroscopy line shapes indicative of the presence of varying secondary structural elements, implying they are folded. The variability in line shape and therefore calculated secondary structure was broadly similar amongst proteins in the same subfamily. The $^{1}$H NMR spectra of the proteins also support the assertion that these proteins are folded.

The utility of the CDNN software for secondary structure calculation was perhaps questionable here. The data analysis process is open to interpretation and therefore open to error and bias. As each of the three different neural nets produce often substantially different values it is not always clear which is the true representation of the protein. Calculation of the proportion of secondary structure present should not be based purely on data from this technique but it can be used to suggest the folding state of a protein. In general the values produced by CDNN can be seen as close enough to the predicted values to indicate these proteins are natively folded after lyophilisation and resuspension. Variation in this data may be due to error in the technique or potentially due to actual variations between the structure as observed as a solid (X-ray crystallography) or in solution (CD spectroscopy).

The values for the Tm calculated using this technique can be compared to Ub as these proteins have a Ub core as part of there structure. Ub had been previously shown to have a $T_m$ of 64°C and the $T_m$ for the hAtg8 proteins as measured here ranged from 51.4 to 69.8°C [181]. These variations will be representative of the effect of the two N-terminal helices of hAtg8 proteins and the differences in the particular residues of the proteins despite them producing similar overall folded structure.

The totality of this CD data and the structures from x-ray crystallography show that although the differences in hAtg8 sequence produce proteins with a similar overall fold, the variety in sequence leads to small alterations in the secondary structure elements and stability. These small variations in structure are likely to contribute to the presumed difference in roles of the protein by modulation of their interactions.

In conclusion, the proteins produced here were pure in terms of both the lack of other proteins and buffer components. The proteins were stored as lyophilised samples which were resuspended when required. The proteins were considered to be folded to their native state when resuspended after lyophilisation, according to CD and NMR studies. As such the samples produced as described in this chapter were utilised for the further biophysical analysis of their interactions in the following chapters.

# Chapter 4

# Biophysical Characterisation of the hAtg8-p62 AIM Interaction

## 4.1 Introduction

The divergent evolution of six hAtg8 proteins from the single Atg8 protein found in yeast has often been considered to imply some functional specialisation of the hAtg8 proteins [182, 183]. As Atg8 in yeast is involved in a number of separate steps within the overall process of autophagy it is likely that hAtg8 proteins have evolved some specialisation for these roles [184, 183, 29, 52]. Due to the differing amino acid sequences and therefore the surface variation at the binding patch, these proteins interact with different autophagy receptors and adapters with varying affinity [40, 50, 51, 185]. Linking particular hAtg8 proteins to specialised roles within autophagy through the discovery of selective interactions is a prominent avenue of research.

The hAtg8-p62 interaction is central to the selective recruitment of cargo to the autophagosome as well as allowing p62 to function as a scaffold for membrane expansion. The work in this chapter determined the binding affinities for each of the hAtg8 proteins interacting with the WT p62 AIM peptide; to investigate their relative binding selectivity. The thermodynamics of these interactions were also used to gain some insight into the differences amongst the protein-peptide interactions. Biophysical techniques were selected to determine relative and absolute binding affinities of these interactions.

## 4.2 Investigating p62 binding preference among the hAtg8 proteins

Experimental use of full length p62 presents issues in terms of expression and solubility which arise due to the oligomerisation of the protein via its PB1 domain [59]. As such the full length protein was not used in this study, instead a peptide was synthesised which is representative of the p62 AIM. A number of AIM sequences from various autophagy related proteins are shown in Figure 4.1. This sequence alignment highlights the core motif and the prevelance of the acidic residues preceding this motif. These acidic residues are mostly present in the previous three amino acid residues though some AIM domains do not present any. From this overview of the AIM sequences it was possible to select a p62 AIM construct which contained all of the conserved AIM components.

| | | | |
|---|---|---|---|
| ATG4B | 1 | ---MDAATLTYDTLRFAEFEDFPE | 21 |
| Calreticulin | 190 | QVESGSLEDDWDFLPPKKIKDPDA | 213 |
| DVL2 | 434 | ESGLEVRDRMWLKITIPNAFLGSD | 457 |
| FUNDC1 | 8 | PQDYESDDDSYEVLDLTEYARRHQ | 31 |
| FYCO1 | 1270 | TDYRPPDDAVFDIITDEELCQIQE | 1293 |
| JMY | 3 | FALEETLESDWVAVRPHVFDEREK | 26 |
| KBTBD6 | 658 | GSSSSLSDDFWVRVAPQ------- | 674 |
| OPTN | 168 | LNSSGSSEDSFVEIRMAEGEAEGS | 191 |
| p62 | 328 | SDNCSGGDDDWTHLSSKEVDPSTG | 351 |
| PLEKHM1 | 625 | QKVRPQQEDEWVNVQYPDQPEEPP | 348 |
| NBR1 | 722 | QSQSSASSEDYIIILPECFDTSRP | 745 |
| NDP52 | 123 | FQFRPENEEDILVVTTQGEVEEIE | 146 |
| NIX/BNIPL3 | 26 | LPPPAGLNSSWVELPMNSSNGNDN | 49 |
| ULK1 | 347 | SKDSSCDTDDFVMVPAQFPGDLVA | 370 |

Figure 4.1: Alignment of AIM sequences from selected proteins. Residues in the aromatic position are highlighted in green, residues in the hydrophobic position are highlighted in yellow and acidic residues preceding the motif are shown in red.

The peptide construct was selected to include the core $[\theta]$-$x_1$-$x_2$-$[\tau]$ motif of the AIM as well as a number of surrounding residues which may be involved in the interaction; particularly a number of preceding acidic residues. This construct was also selected to include the serine residue at S342 which has been shown to be phosphorylated by TBK1 to facilitate comparative work with this site. As such the p62 AIM peptide sequence used incorporated the residues from S332 to S343; SGGDDDWTHLSS. The selected sequence of the WT p62 AIM peptide was shown in Figure 2.9. A peptide

which has been extended at the C-terminus beyond the residues used here was suggested to show no additional significant binding interactions to LC3B in an NMR study [1].

### 4.2.1 Mass spectrometry

Native-like ESI MS was utilised to investigate the relevant protein-peptide interactions in this work. In the spectra produced by this technique, proteins were observed across a range of charge states at different m/z values so these peaks must be summed to measure the amount of a protein in solution. The summed intensity of peaks resulting from proteins and their complexes can be quantified to produce the ratio of bound to free protein in the solution. This ratio can then be compared among proteins in the solution to investigate binding selectivity. In this chapter, the potential selectivity of the p62 AIM among the hAtg8 proteins was investigated. These experiments involved equimolar mixtures of hAtg8 proteins being titrated with the WT p62 AIM peptide.

Unfortunately, using all six hAtg8 proteins in a single set of titration experiments was not possible due to congestion in the spectrum leading to overlapping peaks. Peaks with similar m/z values became indistinguishable when the sample contained all six proteins and therefore quantification of the individual species became challenging. The example which most clearly illustrates this is the comparison of LC3B and GABARAPL2. GABARAPL2 in its 6+ charge state has a predicted m/z value of 2301.83 and the complex of LC3B bound to the p62 WT AIM peptide in its 7+ charge state has a predicted m/z value of 2301.17. As these charge states are both detected, the intensities of these peaks are summed and deconvolution is not possible.

The resutling peak which occurs due to the detection of ions from both species is highlighted in the spectrum shown in Figure 4.2. This is the ESI MS spectrum of an equimolar mixture of LC3B, GABARAPL2 and the WT p62 AIM peptide (all at 10 $\mu$M) which shows the most intense peak in the spectrum at an m/z value corresponding to these two ions. Therefore, quantification of the sum of intensities for each species could not be performed.

There are also other examples of ions with similar m/z values, though not as similar as this, which result in peaks where deconvolution of intensities would be challenging. To solve this problem the two hAtg8 subfamilies, LC3 and GABARAP, were titrated in separate experiments as these produced spectra which could be readily quantified.

Figure 4.3 shows two titration points from this experiment for the LC3 subfamily. This shows the three LC3 proteins in the left-hand spectrum (all at 10 $\mu$M) and

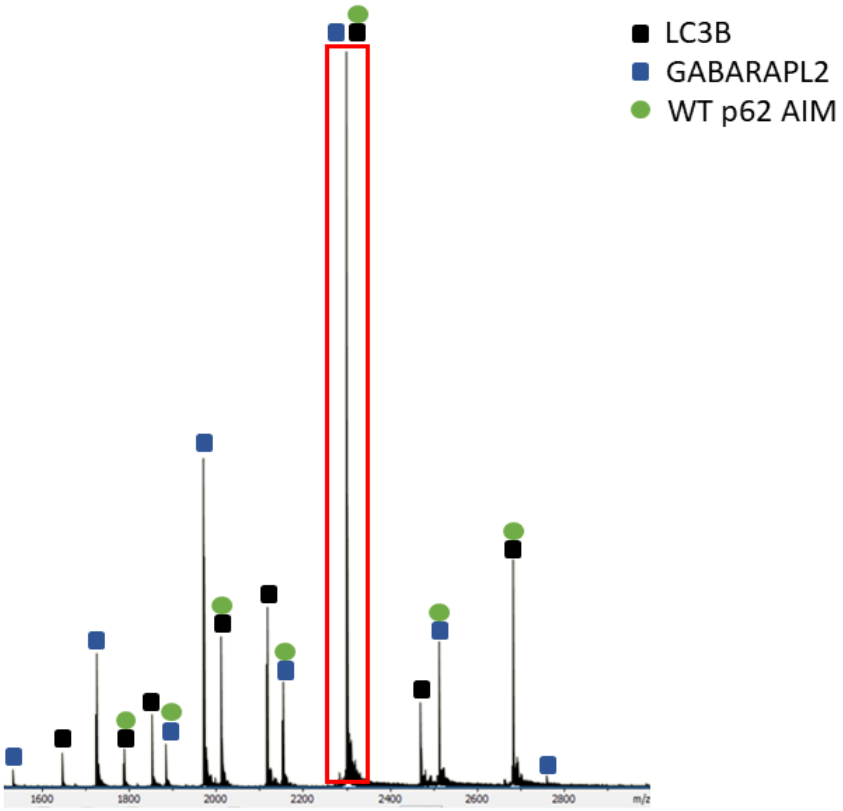



Figure 4.2: ESI MS spectrum of a sample containing LC3B (10 $\mu$M), GABARAPL2 (10 $\mu$M) and the WT p62 AIM peptide (10 $\mu$M). The peak framed by a red box corresponds to two separate species (GABARAPL2 6+ and LC3B-WT p62 AIM 7+) which occur at similar m/z values and their intensity is therefore summed.

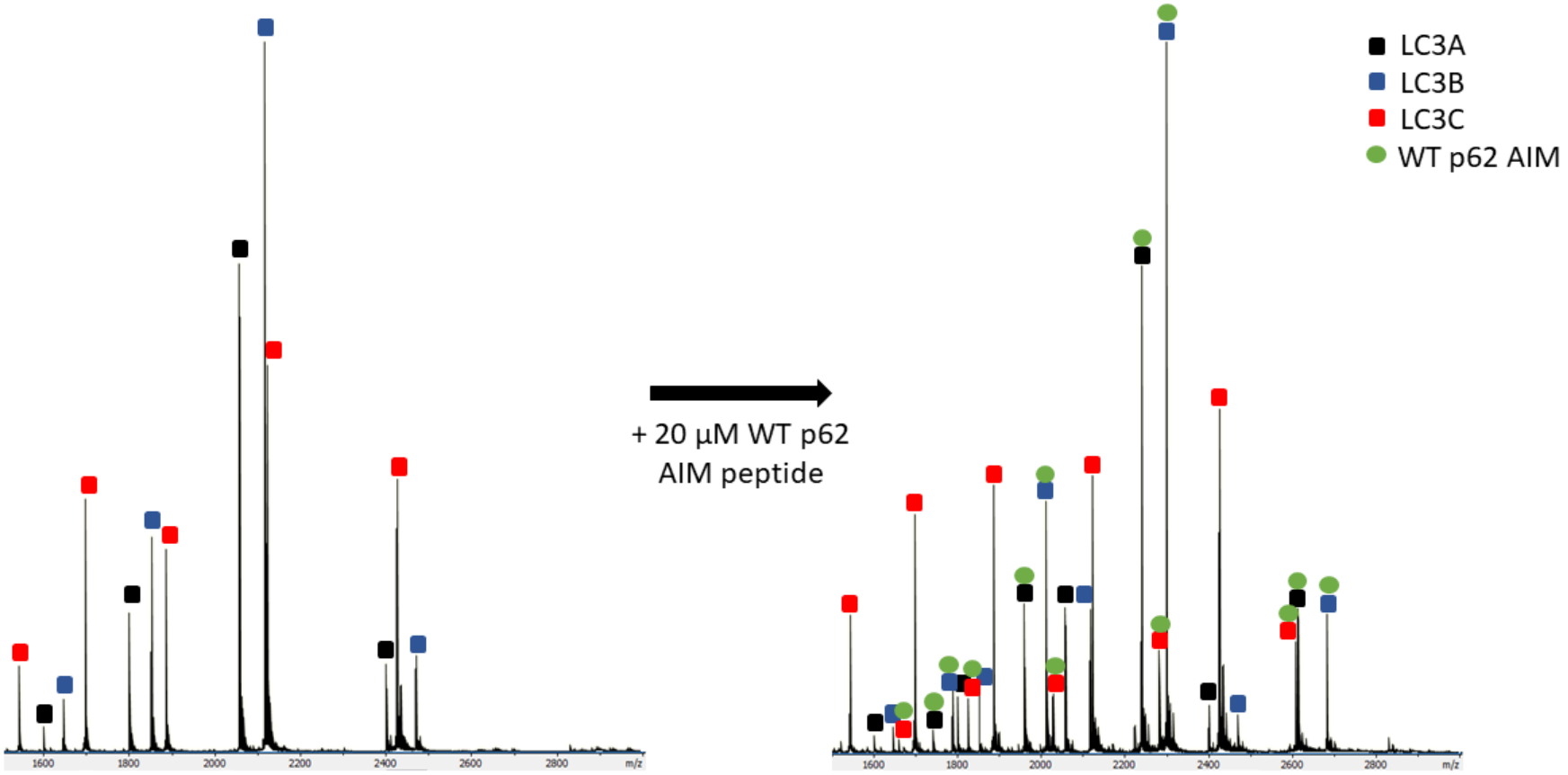



Figure 4.3: ESI MS spectra showing the competition experiment for the LC3 subfamily binding to the WT p62 AIM peptide. Left-hand spectrum: Sample containing the three LC3 subfamily proteins (10 $\mu$M) which shows the proteins in their various charge states; right-hand spectrum: sample containing the three LC3 subfamily proteins (10 $\mu$M) and the WT p62 AIM peptide (20 $\mu$M) which shows the proteins and the protein-peptide complexes in their various charge states. Peaks are labelled according to the particular protein or protein-peptide complex they correspond to as shown in the legend.

then the right-hand spectrum is from a sample with the same concentration of LC3 proteins after the addition of WT p62 AIM peptide (20 $\mu$M). Three observations which were made from these spectra were that the most intense peak in the spectrum after the addition of peptide corresponds to the LC3B-AIM peptide complex. It was also observed that the unbound LC3C peaks remain at a higher intensity than those for LC3A and LC3B. Finally, the peaks observed in this spectrum correspond only to 1:1 protein:peptide complexes; no multicomponent species were observed.

Figure 4.4 then shows a section of the same spectra between 2350 and 2800 m/z from the LC3 protein only and the 20 $\mu$M peptide samples to highlight the change upon addition of the peptide. From this the relative ratio of the unbound proteins can be observed as well as the appearance of the bound peaks and their relative intensities. This shows that the peaks corresponding to unbound LC3A and LC3B decreased relative to LC3C upon addition of the peptide. This was most pronounced for LC3B which showed a greater than threefold reduction in intensity. However, interpreting these experiments with multiple proteins and their complexes in multiple charge states by eye is clearly challenging.

To aid in visualisation and interpretation of these results the intensities of the peaks for each charge state of a protein or protein-peptide complex were summed and a ratio of bound to free hAtg8 was calculated. By plotting these values for each titration point, as shown in Figure 4.5A, a clearer picture of the results emerged. This showed that at each titration point the ratio of bound to free LC3B was higher than for the other LC3 proteins. The calculated ratio for LC3C was significantly lower and the LC3A ratio was intermediate between these two though considerably closer to the values determined for LC3B. The clearest difference in these calculated ratios occured for the final titration points where LC3B had a bound to free ratio of 4.3, LC3A had a ratio of 3.2 and LC3C had a ratio of 0.3. LC3B was therefore considered to be selectively targeted to some degree by the WT p62 AIM peptide among the LC3 subfamily, although the degree of selectivity between LC3B and LC3A may not be large.

The results from the equivalent experiment on the GABARAP subfamily, carried out and analysed by the same methods as for the LC3 subfamily, are shown in Figure 4.5B. Spectra from the beginning and an intermediate point of this titration are also shown in Figure 4.6. As for the LC3 subfamily, no multicomponent species were observed. This data has shown that the ratio of bound GABARAPL1 to unbound was the highest at all titration points, indicating some level of binding preference. In a similar pattern to the LC3s, GABARAP appeared to be bound slightly less but there

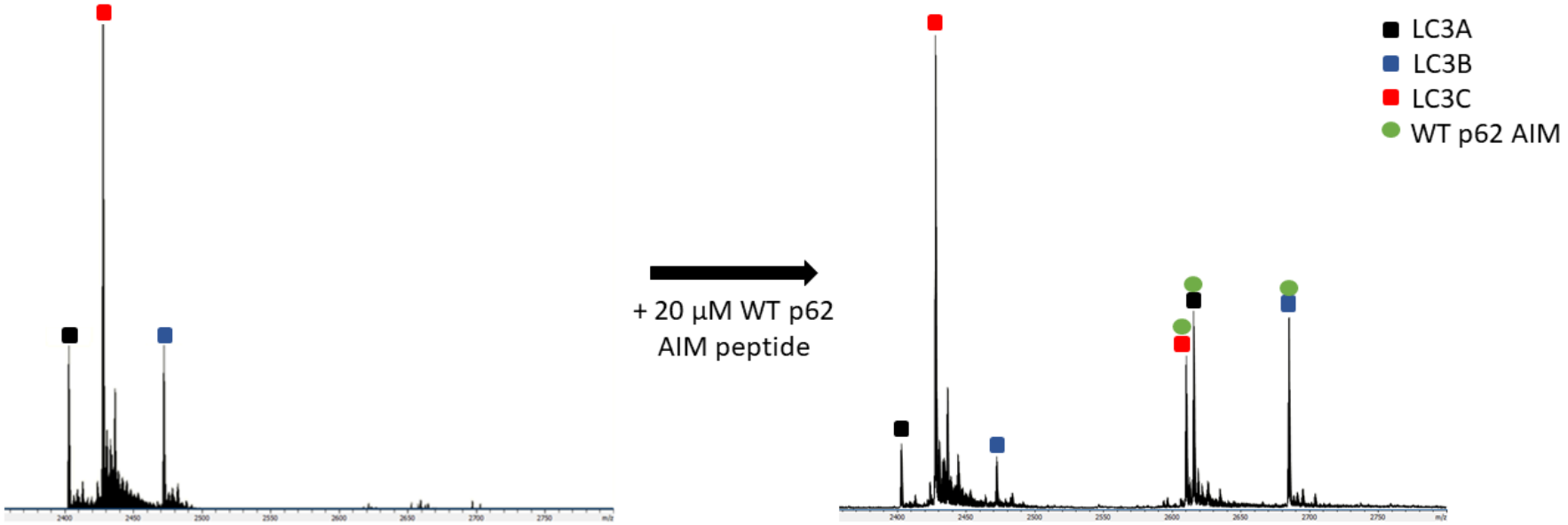



Figure 4.4: ESI MS spectra showing the competition experiment for the LC3 subfamily binding to the WT p62 AIM peptide for a small range of m/z values (2350-2800 m/z). Left-hand spectrum: Sample containing the three LC3 subfamily proteins (10 $\mu$M) which shows the proteins in various charge states; right-hand spectrum: sample containing the three LC3 subfamily proteins (10 $\mu$M) and the WT p62 AIM peptide (20 $\mu$M) which shows the proteins and the protein-peptide complexes. Peaks are labelled according to the particular protein or protein-peptide complex they correspond to as shown in the legend.

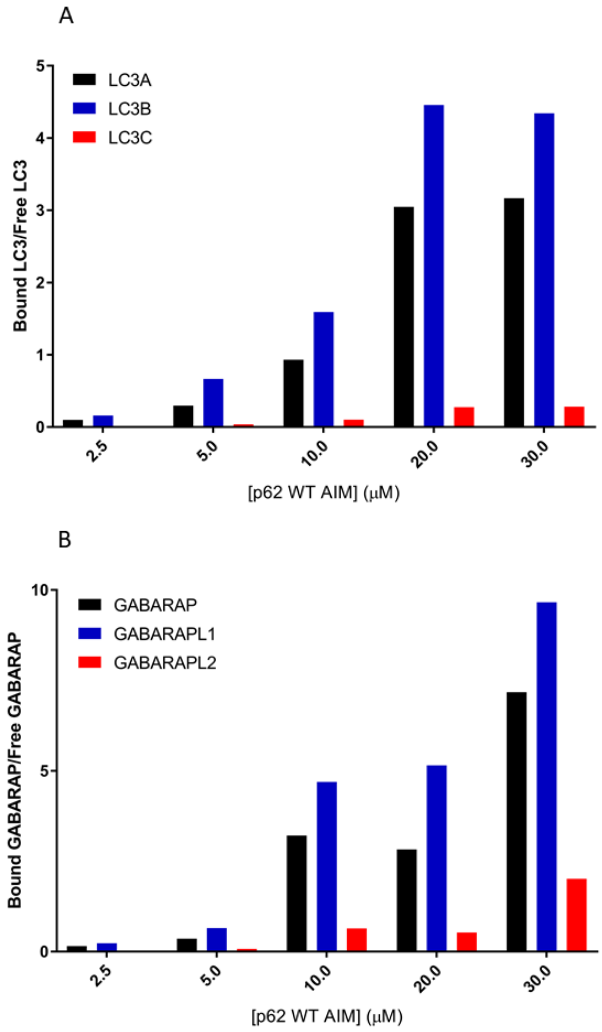



Figure 4.5: Graphs showing the results of the ESI-MS competition experiments as the p62 WT AIM is titrated into an equimolar sample of the hAtg8 proteins to compare binding within each subfamily. Top graph (A) is for the LC3 subfamily and the bottom graph (B) is for the GABARAP subfamily, showing a preference for LC3B and GABARAPL1 respectively.

did not appear to be a large difference in selectivity between the two. GABARAPL2 was observed to have the lowest ratio of the GABARAP subfamily to the WT p62 AIM peptide. The ratios for the proteins at the highest peptide concentration tested illustrate this; GABARAP had a ratio of 7.2, GABARAPL1 had a ratio of 9.7 and GABARAPL2 had a ratio of 2.0.

A further ESI MS competition experiment of this type was carried out to compare the relative affinities of these two subfamilies for the WT peptide. In this titration series the WT p62 AIM peptide was titrated into an equimolar mixture of LC3A and GABARAP. The spectra of two titration points from this experiment are shown in Figure 4.7. The selection of these hAtg8 proteins as representative of their group was arbitrary but they may be considered representative as the intermediate protein in their subfamily in terms of selectivity. Figure 4.9A shows the results from this titration plotted on a graph. As the peptide concentration was increased the ratio of bound to free LC3A increased more than the ratio for GABARAP. This was particularly clear at the 30 µM titration point where the ratio for LC3A was more than double the ratio for GABARAP. This indicated that the WT p62 AIM peptide preferentially bound to LC3A over GABARAP.

To investigate the absolute preference of the WT p62 AIM peptide, a competition experiment was carried out in which an equimolar mixture of LC3B and GABARAPL1 was titrated with the AIM peptide. These are the hAtg8 proteins which were preferentially bound within their respective subfamily so the protein preferred here is preferentially bound to the WT p62 AIM among all of the hAtg8 proteins. Spectra of two titration points from this experiment are shown in Figure 4.8. Figure 4.9B then shows that at every peptide concentration there was a higher ratio of bound to unbound GABARAPL1 than the ratio for LC3B. The highest peptide concentration produced the largest difference in bound to free ratio for these proteins as the value for LC3B was 8.3 and the value for GABARAPL1 was 14.1 This indicated a binding preference of the WT peptide for GABARAPL1 among all of the hAtg8 proteins according to these ESI MS competition experiments.

### 4.2.2 Isothermal titration calorimetry

ITC was used here to further investigate the binding interaction of the WT p62 AIM peptide with all six hAtg8 proteins. Example data for the interaction of GABARAP with the peptide is shown in Figure 4.10. The fitted thermodynamic parameters from these experiments are displayed in Table 4.1 and the raw applied

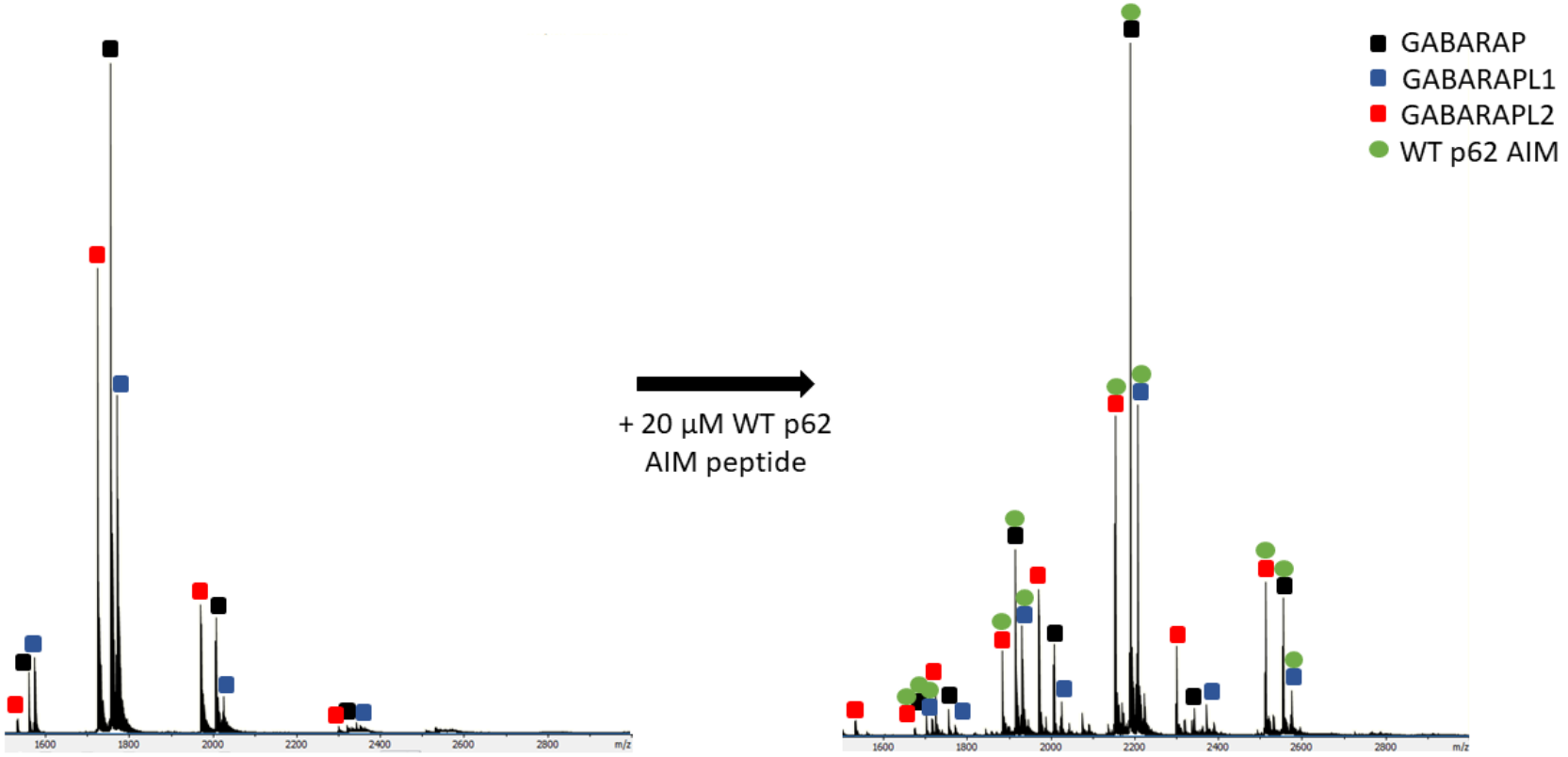



Figure 4.6: ESI MS spectra showing the competition experiment for the GABARAP subfamily binding to the WT p62 AIM peptide. Left-hand spectrum: Sample containing the three GABARAP subfamily proteins (10 $\mu$M) which shows the proteins in their various charge states; right-hand spectrum: sample containing the three GABARAP subfamily proteins (10 $\mu$M) and the WT p62 AIM peptide (20 $\mu$M) which shows the proteins and the protein-peptide complexes in their various charge states. Peaks are labelled according to the particular protein or protein-peptide complex they correspond to as shown in the legend.

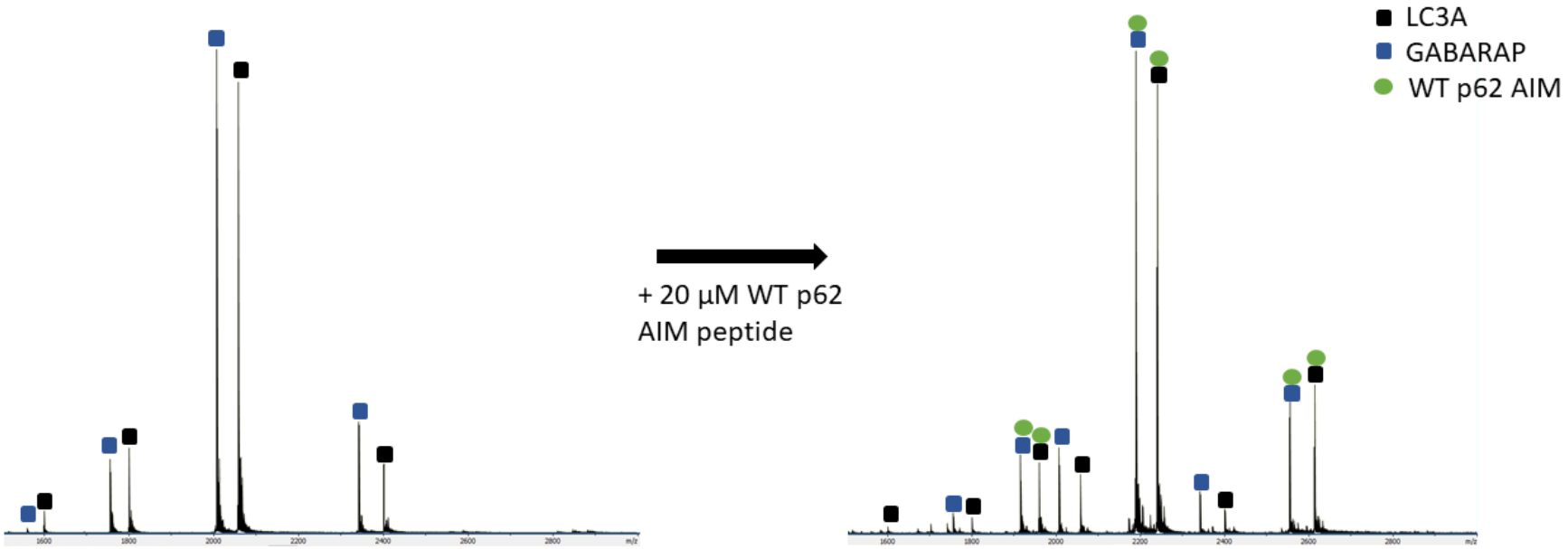



Figure 4.7: ESI MS spectra showing the competition experiment for LC3A and GABARAP binding to the WT p62 AIM peptide. Left-hand spectrum: Sample containing LC3A (10 $\mu$M) and GABARAP (10 $\mu$M) which shows the proteins in their various charge states; right-hand spectrum: sample containing LC3A (10 $\mu$M) and GABARAP (10 $\mu$M) and the WT p62 AIM peptide (20 $\mu$M) which shows the proteins and the protein-peptide complexes in their various charge states. Peaks are labelled according to the particular protein or protein-peptide complex they correspond to as shown in the legend.

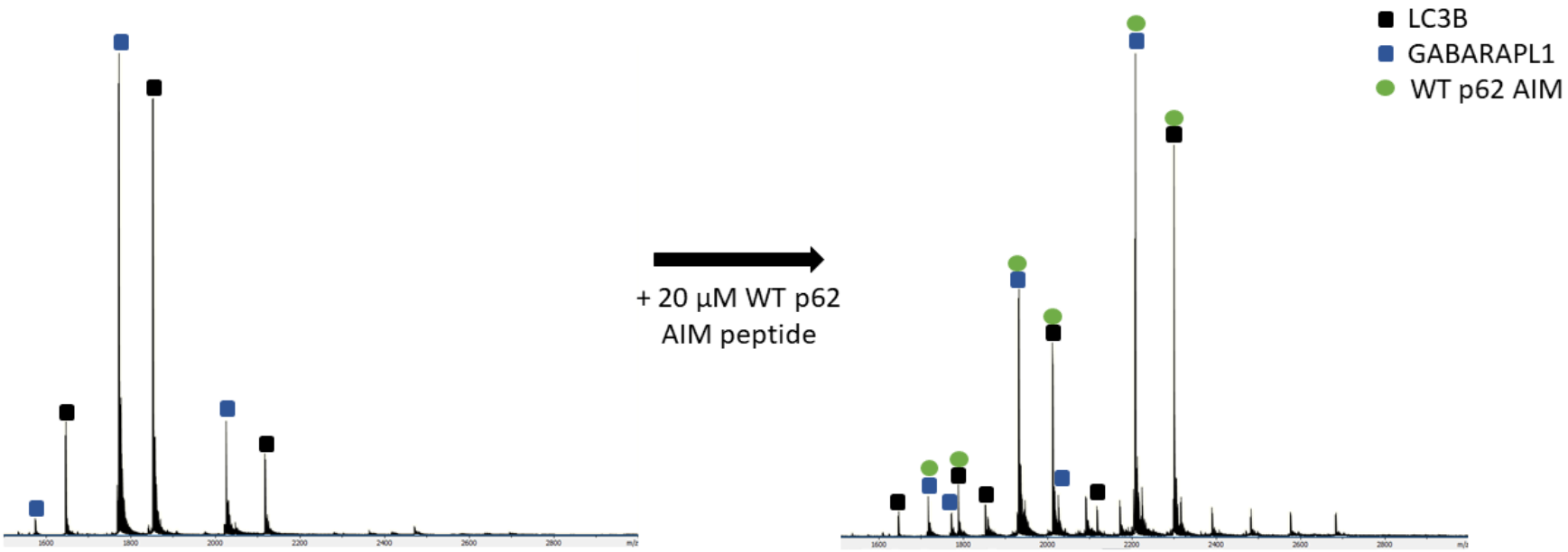



Figure 4.8: ESI MS spectra showing the competition experiment for LC3B and GABARAPL1 binding to the WT p62 AIM peptide. Left-hand spectrum: Sample containing LC3B (10 $\mu$M) and GABARAPL1 (10 $\mu$M) which shows the proteins in their various charge states; right-hand spectrum: sample containing LC3B (10 $\mu$M) and GABARAPL1 (10 $\mu$M) and the WT p62 AIM peptide (20 $\mu$M) which shows the proteins and the protein-peptide complexes in their various charge states. Peaks are labelled according to the particular protein or protein-peptide complex they correspond to as shown in the legend.

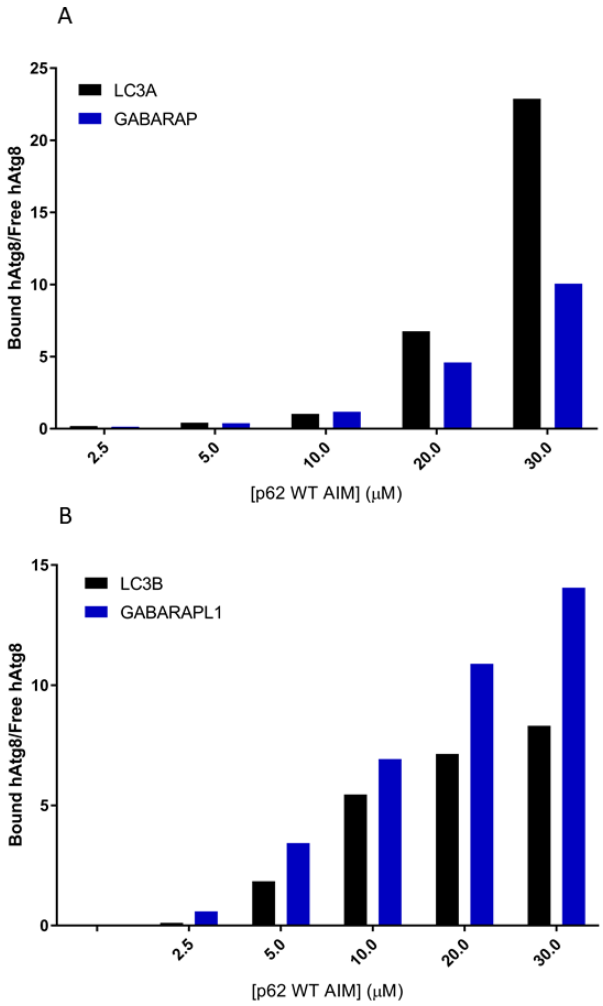



Figure 4.9: Graphs showing the results of the ESI-MS competition experiments as the p62 WT AIM peptide is titrated into an equimolar mixture of the hAtg8 proteins to compare members of the different subfamilies. Top graph (A) is for the LC3A/GABARAP competition experiment and the bottom graph (B) is for the LC3B/GABARAPL1 competition experiment; showing a preference for LC3A and GABARAPL1 respectively.

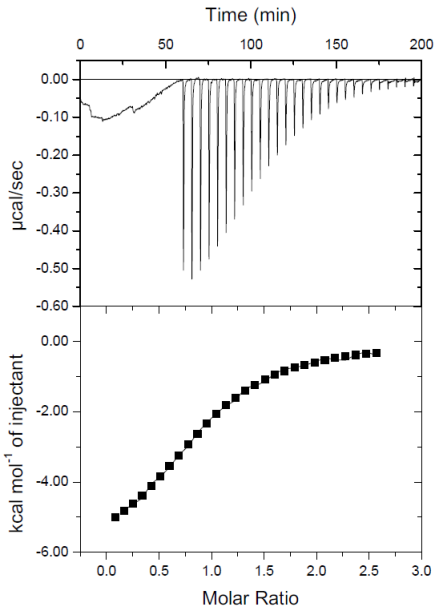



Figure 4.10: ITC data for the titration of the WT p62 AIM peptide into GABARAP. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

power data and derived binding curves used to determine each of these values are shown in Appendix Figures A.14-A.18.

The binding affinities calculated from these experiments indicate that the GABARAPL1-WT p62 AIM peptide interaction has the strongest binding affinity although the LC3B-WT p62 AIM peptide interaction affinity has a similar value and there is overlap in the calculated error so this cannot be conclusively determined as a significant difference. However, there is a significant difference in the affinity of GABARAPL1 binding to the peptide compared to the other four hAtg8 proteins. The LC3C-WT p62 AIM peptide interaction has the weakest affinity however the fit for this data has a higher error so this value has a large degree of uncertainty. There was a fivefold difference in $K_D$ between the strongest and weakest interaction determined from these experiments with smaller but in some cases significant differences between the intermediate proteins.

The thermodynamic characteristics of these hAtg8-p62 AIM interactions can then be compared, as shown in Figure 4.11. This shows the relative contributions of the enthalpic ($\Delta H$) and entropic (-T$\Delta S$) components to the overall binding ($\Delta G$) of these proteins. It can be seen that there are some similarities among proteins which have previously shown similarities in sequence homology and in their CD

| Protein | n | $K_D$ ($\mu$M) | $\Delta H$ (kJ mol$^{-1}$) | $\Delta S$ (J mol$^{-1}$K$^{-1}$) | $\Delta G$ (kJ mol$^{-1}$) |
|---|---|---|---|---|---|
| **LC3A** | 1.05±0.02 | 4.55±0.30 | -74.97±1.57 | -14.37 | -30.96 |
| **LC3B** | 0.92±0.03 | 3.83±0.41 | -44.85±1.83 | -4.39 | -31.41 |
| **LC3C** | 0.82±0.13 | 16.84±4.08 | -15.01±3.52 | 4.14 | -27.69 |
| **GABARAP** | 0.92±0.01 | 5.20±0.31 | -23.97±0.34 | 1.22 | -30.23 |
| **GABARAPL1** | 0.86±0.02 | 3.23±0.45 | -23.49±0.86 | 2.74 | -31.87 |
| **GABARAPL2** | 1.07±0.06 | 6.85±1.02 | -10.15±0.69 | 6.46 | -29.94 |

Table 4.1: ITC-derived thermodynamic parameters for the binding of the WT p62 AIM peptide to each of the hAtg8 proteins.

spectroscopy profile; there are the clearest similarities between LC3A/LC3B and GABARAP/GABARAPL1.

The enthalpic contributation to the LC3A and LC3B interactions with the WT peptide is significantly higher than for the other proteins, with LC3A showing the highest enthalpic contribution. These proteins also uniquely show an unfavourable entropic component, with LC3A again showing the largest contribution. The most favourable enthalpy and most unfavourable entropy of LC3A compensate for each other resulting in an affinity that is close to the other hAtg8 affinities determined here. The remaining four hAtg8 proteins (LC3C, GABARAP, GABARAPL1 and GABARAPL2) show similar profiles other than some differences in magnitiude which result in the observed differences in $\Delta G$. These protein-peptide interactions are dominated by the enthalpic term in all cases other than for GABARAPL2 which is dominated by the entropic term. Despite this largely favourable entropy term for GABARAPL2, this interaction has the smallest enthalpic contribution resulting in the second lowest $\Delta G$ determined here.

These thermodynamic parameters for the interaction of each hAtg8 protein with the peptide are plotted in Figure 4.12. The plotting of these points produced a linear trend showing the compensation effect between these parameters. The proteins which produced values which fall below the fitted line are the proteins with the strongest binding affinities and those above the line produced the weakest binding affinities.

The 'n' parameter from fitting this data deviated away from the expected value of

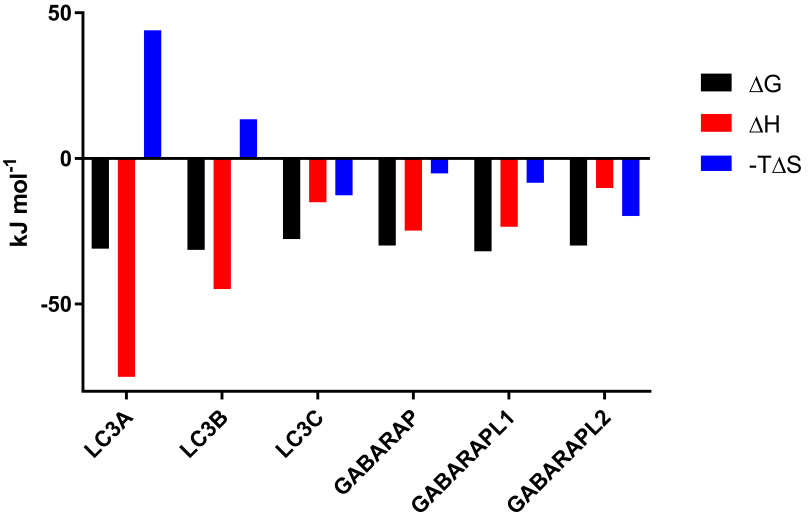



Figure 4.11: Graph comparing ITC-derived thermodynamic parameters for the binding of the WT p62 AIM peptide to each of the hAtg8 proteins to illustrate the differences in the enthalpic and entropic contributions to the Gibbs free energy.

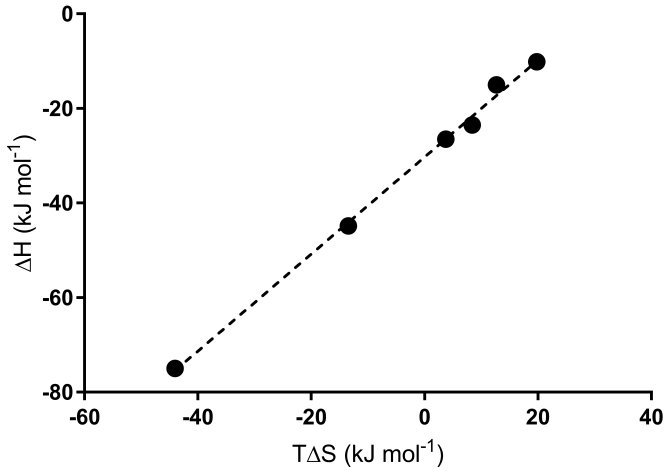



Figure 4.12: Enthalpy-entropy compensation effect demonstrated by ITC-derived thermodynamic parameters.

| Protein | n | $K_D$ ($\mu$M) | $\Delta H$ (kJ mol$^{-1}$) |
|---|---|---|---|
| **LC3A** | 1 | 5.21±0.31 | -79.50±1.12 |
| **LC3B** | 1 | 3.15±0.34 | -40.68±0.94 |
| **LC3C** | 1 | 14.22±2.58 | -12.24±0.91 |
| **GABARAP** | 1 | 4.91±0.36 | -21.42±0.46 |
| **GABARAPL1** | 1 | 2.65±0.54 | -20.77±0.79 |
| **GABARAPL2** | 1 | 7.81±0.86 | -10.95±0.37 |

Table 4.2: ITC-derived thermodynamic parameters for the binding of the WT p62 AIM peptide to each of the hAtg8 proteins. This derivation was carried out by fixing the value of 'n' to one.

one, to some degree, for all of these interactions. It was possible to perform a fitting analysis from the same data whilst fixing the value of 'n' to one. The results of this analysis are shown in Table 4.2. This change in the fitting method resulted in a small change to the derived parameters. The value of the $K_D$ decreased if the 'n' value was previously below one and vice versa.

Binding selectivity is a central concern in this thesis so it should be noted that both fitting procedures produce the same order of binding preference other than for LC3A and GABARAP. Allowing the value of 'n' to float in the fitting process produced data that indicated that the WT p62 AIM peptide preferentially bound to LC3A over GABARAP. On the other hand, fixing the value of 'n' to one led to $K_D$ values which indicated no significant difference in binding between them. The same binding curve data for the GABARAP-WT p62 AIM interaction as shown before but with 'n' fixed during fitting is shown in Figure 4.13. This shows that the fitted line deviates away from the plotted data and therefore the fit doesn't accurately describe the data in this case.

## 4.3 Discussion

### MS Competition Experiments

ESI MS competition experiments, as used here, present a number of possible pitfalls so potential issues with this methodology must first be considered. The ratio of bound to free protein has been used in ESI MS experiments to reliably reproduce $K_D$ values from ITC experiments so this is deemed to be an acceptable measurement of affinity [155]. The difference here is the presence of multiple sample components which may be ionised to differing degrees and complicate in-sample comparisons. Indeed,

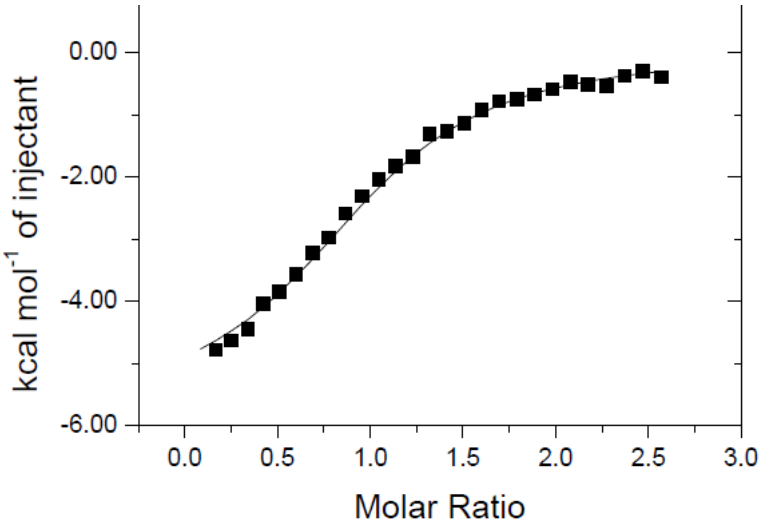



Figure 4.13: ITC data for the titration of the p62 WT AIM peptide into GABARAP. This shows the plotted binding curve from the integration of applied power raw data and a line showing the fit of this data with 'n' fixed to one during the fitting.

ESI MS of equimolar mixtures of the hAtg8 proteins produce spectra with differences in summed intensities for each protein. However, similar ratios of intensities were observed repeatedly from separate sample preparations. This confirms that the differences in intensity occur due to the ESI process rather than differences in the concentrations of the proteins.

Differences in peak intensity from equimolar mixtures of proteins have also been observed elsewhere in the literature [157]. This study cited the differing surfaces and desolvation behaviour as key factors in the intensity difference. The availability of surface sites for charge to be incorporated and their solvent accessibility may lead to proteins with differing ionisation efficiencies. It can also be the case that proteins are not fully desolvated into individual ions for detection and that this can vary from protein to protein. Despite the structural similarity of the hAtg8 proteins there is substantial sequence difference which would support variations in ionisable sites and desolvation behaviour leading to differences in peak intensity among equimolar proteins.

Maintenance of the solution phase equilibrium of bound to free protein throughout the ESI and detection process is central to the accurate representation of the binding interactions. A number of potential sources of error which may lead to differences between the solution phase ratio of bound to free protein to the detected ratio were highlighted in a paper by Kitova *et al.* [158]. This paper suggests that this ratio is likely to be maintained in cases where there is a large size difference between the two

binding partners such that the size of the complex is roughly the same as the size of the protein. This approximation is suggested to hold when the size of the complex is less than 110% of the size of the protein. All of the hAtg8-p62 AIM complexes studied here are between 108 and 109% of the size of the protein only. Other factors considered by Kitova et al. are in-source dissociation, non-specific ligand-protein binding and ESI-induced changes to pH and temperature. Due to the analysis of these proteins under the same experimental conditions, these factors should not produce significant differences between the proteins.

The similar binding affinities of these interactions should result in similar levels of in-source dissociation which should be relatively low due to the important hydrophobic interactions known to be involved in this binding. Non-specific ligand-protein binding may alter the detected ratio of bound to free protein but this is assumed to be low due to the low protein concentrations used here. Finally, any ESI-induced changes to pH and temperature will be uniform to the proteins in the same sample. The errors which may persist in this technique are assumed to be small and similar among the proteins.

These ESI-MS competition experiments indicated that the WT p62 AIM peptide used in this work preferentially binds to GABARAPL1 amongst all six hAtg8 proteins. The full set of experiments conducted indicated the following order of binding preference: GABARAPL1 > LC3B > LC3A > GABARAP > LC3C/GABARAPL2. The preference of the p62 AIM peptide for binding to LC3C or GABARAPL2 was not investigated with these experiments.

## ITC

The quality of fitting for the ITC experiments should be commented upon before discussing any implications from the data. The 'n' parameter determined from ITC data fitting is often a good waypoint for initial assessment of the quality of the fit as the value should always make biological sense i.e. it should represent a known binding conformation. In this study all of the protein-peptide interactions were expected to occur at a 1:1 ratio which was shown in the MS studies and has been previously shown in NMR/X-ray structures. As such the 'n' value was expected to be one. This test indicated a good fit for these experiments as all 'n' values occured within 20% of the expected value; a common benchmark. However, it should be noted that the experiments with GABARAPL1 and particularly LC3C had a low 'n' value which may indicate incorrect protein/peptide concentrations or an error in the fitting process.

There is some debate regarding the fitting of the 'n' parameter in ITC studies as it may be preferential to fix the 'n' value if the binding ratio is known [186].

For the data presented here, the fitting process did not appear to significantly affect many of the fitted parameters. The $K_D$ values produced were similar whether the 'n' value was fixed or could float during fitting. Selectivity was a key factor in this study so the ordering of binding preference was compared using either fitting method. This ordering was the same regardless of fitting method other than for LC3A and GABARAP. The ESI MS competition experiment which compared the binding of the peptide to LC3A and GABARAP indicated a binding preference for LC3A. Floating 'n' during the fitting process produced the same result whereas fixing 'n' led to non-statistically significant differences in values so this data supports allowing the 'n' parameter to float during fitting. Changes to the $\Delta H$ values produced between the two fitting methods were significant though small.

Visual comparison of the fitted line from both fitting methods and how well it represented the data showed that floating the 'n' value produced a more representative line in the case of GABARAP. This was quantified using the $\chi^2$/DoF value produced from each fitting as this is the value that is minimised to determine the quality of fit throughout the iterative fitting process of data analysis. This value was 2459 for the fit when 'n' was allowed to float and 8480 when 'n' was fixed; indicating that floating 'n' produces a better fit. The GABARAP ITC data was selected as an example in this chapter arbitrarily so it should be noted that for some of the hAtg8 proteins the fitting of the line to the data was not noticeably different between the two fitting processes though in no case did the fitting appear better for fixing n; as inspected visually and by comparing $\chi^2$/DoF values.

Based upon this it was determined that allowing 'n' to float in the fitting process was the more appropriate methodology for this work as it more accurately represents the data. Additionally, allowing the 'n' value to float produces a measure that represents the quality of the fit quantitatively rather than relying on visual inspection.

ITC-derived $K_D$ values (from the floating 'n' fitting process) can then be compared with the relative selectivities observed in the ESI-MS competition experiments. These experiments appear to be in relative agreement as for example the WT p62 AIM peptide bound selectively to GABARAPL1 in the MS studies and this interaction produced the lowest $K_D$ value via ITC, although the value was not lower than for LC3B by a statistically significant margin. The ordering of these $K_D$ values then fits with the ordering produced by the ESI MS studies. The combined MS and ITC experiments indicate the following order of binding preference: GABARAPL1 >

| Study | $\Delta H$ (kJ mol$^{-1}$) | $\Delta S$ (J mol$^{-1}$ K$^{-1}$) | $K_D$ ($\mu$M) |
|---|---|---|---|
| This work | -44.85±1.83 | -4.39 | 3.83±0.41 |
| Novak et al. [187] | -43.93±0.54 | -3.60 | 1.47±0.08 |
| Rogov et al. [188] | -39.24±0.71 | -2.24 | 2.01±0.06 |

Table 4.3: Comparison of ITC-derived parameters for the interaction between a p62 AIM peptide and LC3B.

LC3B > LC3A > GABARAP > GABARAPL2 > LC3C.

Additionally it is worth noting that the ITC-derived $K_D$ value for the LC3B-p62 WT AIM peptide interaction determined here (3.83±0.41 $\mu$M) is in agreement with previous work by Goode *et al.* which showed this interaction to have a $K_D$ value of 3.2±1.1 $\mu$M [1]. This previous ITC study used a peptide corresponding to p62 residues 332 to 351 rather than the construct of residues 332 to 343 used here which indicates that the extension of the peptide in their work does not produce a significant effect on binding affinity.

Furthermore, two studies have reported ITC-derived thermodynamics data for the interaction betwen a p62 AIM peptide and LC3B. These are compared to the parameters derived here in Table 4.3. Although there are some differences in the produced the data, the magnitude and sign of these values are in agreement.

## Avidity Effects in AIM Binding

The combined data in this chapter shows a degree of binding selectivity is present for the p62 AIM peptide amongst the hAtg8 proteins. There is a fivefold difference in $K_D$ between GABARAPL1 and LC3C, with some significant differences also present within the proteins with intermediate affinities. The smaller differences in affinity presented by this data should be considered within their biological context where a p62 filament presenting multiple AIMs may bind avidly to multiple hAtg8 proteins anchored to the autophagosomal membrane, as described in Section 1.4. This avid binding may amplify these small differences in binding affinity in the cellular context.

The avidity effect on binding between hAtg8 proteins on the autophagosome membrane and the multiple AIM binding sites on a single p62 filament is considered in Figure 4.14. This shows a simplified version of this system considering just two hAtg8 proteins binding to two sites on a filament but this is representative of the effect on a larger scale. This diagram illustrates that first a single hAtg8 interaction must be made and then subsequent interactions may follow. Although simultaneously binding is possible, this scenario is more likely due to the flexibility of both components.

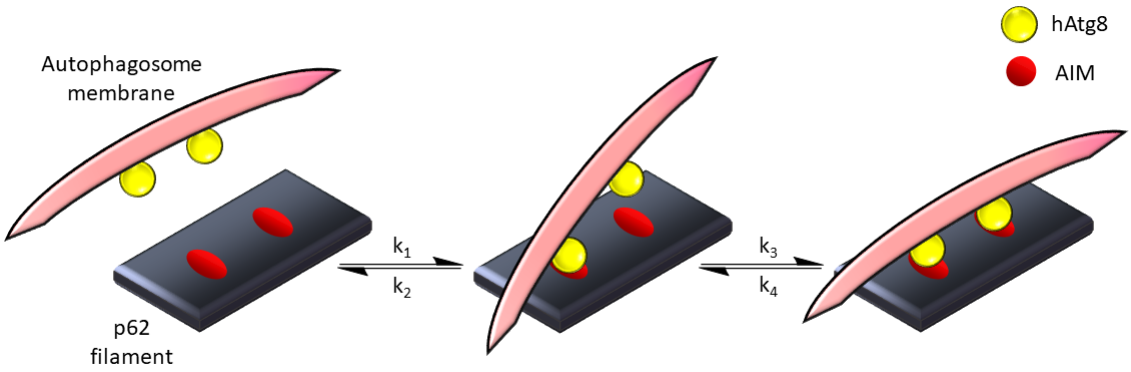



Figure 4.14: Simplified schematic to show avid binding of hAtg8 proteins on the autophagosome membrane to p62 filaments. This illustrates the initial binding of one hAtg8 protein to one AIM on the p62 filament. This then allows avid binding of the second hAtg8 protein to another AIM on the filament. This second interaction has a higher binding affinity than the first interaction due to the increased local concentration of binding sites from being held in proximity to the p62.

The key point of this diagram is that rate constant $k_3$ is larger than $k_1$ due to avidity. Upon binding to the first site the local concentration of binding sites for the second hAtg8 is higher than for a hAtg8 protein free in solution. Essentially, the second hAtg8 is being held in close proximity to the binding site making it more likely to bind.

This simplified version shows the hAtg8 being held near the binding site which may not be the case but the two components must simply diffuse until the hAtg8 finds any neighbouring AIM site, of which there may be many in proximity, before the first hAtg8 dissociates. It may also be important to consider residence time as this is often increased upon avid binding. If $k_3$ is faster than $k_2$ it becomes less likely that there will be full dissociation as the second site binds back on before the full complex can break apart. Therefore an increase in binding affinity may lead to increased residency time as both $k_1$ and $k_3$ become faster whilst $k_2$ and $k_4$ become slower.

## Binding Selectivity

The preference for GABARAPL1 here is supported by previous work which showed the binding affinity of the NBR1 AIM binding to GABARAPL1 was increased by changing the Tyr at the first position in the AIM to a Trp [50]. This supports the strong binding interaction observed here for the p62 AIM which has a Trp at the first position. A study has shown that proteasomal inhibition led to selective induction of

GABARAPL1 and p62 in neuroblastoma cells which supports the importance of this interaction [189]. This indicates that selective autophagy of proteins is compensating for decreased proteasomal degradation and this is achieved through induction of the two proteins with the highest affinity interaction observed here. Increasing levels of these two proteins is the most efficient way to increase selective cargo recruitment.

Recent work using peptide arrays to study the interaction between AIM peptides and the hAtg8 proteins produced a mixture of data that supports this work and some that does not [51]. Rogov *et al.* showed that LC3B binds to the p62 AIM with a higher affinity than LC3C which was also demonstrated here.

They showed that there was no significant difference in binding between GABARAP and GABARAPL2 in terms of the interaction with the p62 AIM and although the difference observed here is considered significant, it is very small. Additionally the peptide array experiments also indicated that GABARAP binds to the p62 AIM with higher affinity than LC3B which is not what has been observed in this thesis as LC3B has been shown to bind with a higher affinity than GABARAP; although the difference can only just be considered significant. Discrepencies between the data in their work and that presented here are therefore present but small.

It may be the case that the high throughput nature of the peptide array may have produced an anomolous result. Similarly, the difference in result observed may come from an artifact of immobilising the peptide. However, due to the small differences in affinity being studied in this work the margins of error are slight and it is not clear which study is accurate.

ITC-derived thermodynamic parameters may be useful when considering this selective binding as they can be informative about differences in binding mode among the proteins. The enthalpy-entropy compensation effect observed for the hAtg8 proteins has resulted in $\Delta G$ values which had a range of 4.18 kJ mol$^{-1}$, despite a range of 64.83 kJ mol$^{-1}$ for the $\Delta H$ values and a range of 63.80 kJ mol$^{-1}$ for the T$\Delta S$ values. This indicates that these interactions have diverse contributions that produce similar affinities.

It may be the case that the interaction of the peptide with LC3A and LC3B occurs via a different binding mode than for the other hAtg8 proteins as these have larger enthalpy terms and are the only two to have unfavourable entropic terms. The proteins with interactions that produce favourable entropic terms likely have significant contributions to their binding from the hydrophobic effect in order to offset the unfavourable entropy inherent in the conformational limitation of a bimolecular interaction. Of these other four hAtg8 proteins GABARAPL1 does not appear to

show any particular defining difference in its enthalpic and entropic contributions to lead to its higher affinity. This may mean that it binds in a similar fashion but the surface charge and topology of GABARAPL1 is simply more optimised to facilitate binding than for the other proteins.

By inspection of the surface charge and topology of these proteins shown in Figure 4.15 it is perhaps unsurprising that there is some diversity in these proteins in terms of their interaction with the peptide. The W- and L-sites of LC3B are indicated and whilst it is possible to speculate about the analogous sites on the other five proteins, it is less clear. The protein with the most similar surface to LC3B is LC3A; as would be predicted based on their sequence homology. Even for these two structurally and sequentially similar proteins though there are some differences in the distribution of charged residues and the topology around the W- and L-sites. The similar sites of LC3A and LC3B in some way produce binding interactions with the peptide that are entropically unfavourable. Perhaps a tight fit of the AIM residues in these pockets is entropically unfavourable as they can sit in a limited number of conformations but they are tightly held enthalpically. The binding pockets on the other hAtg8 proteins may be a less tight fit but perhaps they can accomodate different conformations of the peptide as they have lower enthalpic contributions but favourable entropic contributions to the overall binding. Binding of a peptide will also alter the solvation layer of local water molecules around the binding site which affects the thermodynamic profile of the interactions. The exact nature of the causes of these thermodynamic differences and the relative contribution of desolvation effects versus configurational effects was unclear.

Consideration of different binding modes may also provide a degree of explanation for the low affinity of the LC3C-p62 WT AIM peptide interaction. As discussed in Section 1.2, NDP52 has been shown to bind selectively to LC3C amongst the hAtg8 proteins [40]. In this report von Muhlinen et al. suggest that this interaction is mediated by a noncanonical AIM formed from a Leu-Val-Val sequence which interacts with the unique binding patch of LC3C. The low sequence homology of LC3C compared to the other hAtg8 proteins was highlighted as the cause of this unique binding patch. Therefore, it follows that the binding patch of LC3C which is selectively bound by a noncanonical AIM would bind with a lower affinity to a canonical AIM, as is the case for the p62 motif.

The data presented in this chapter has shown that:

- Allowing the 'n' value to float during fitting of ITC data produces a fit which better represents the data in this case.

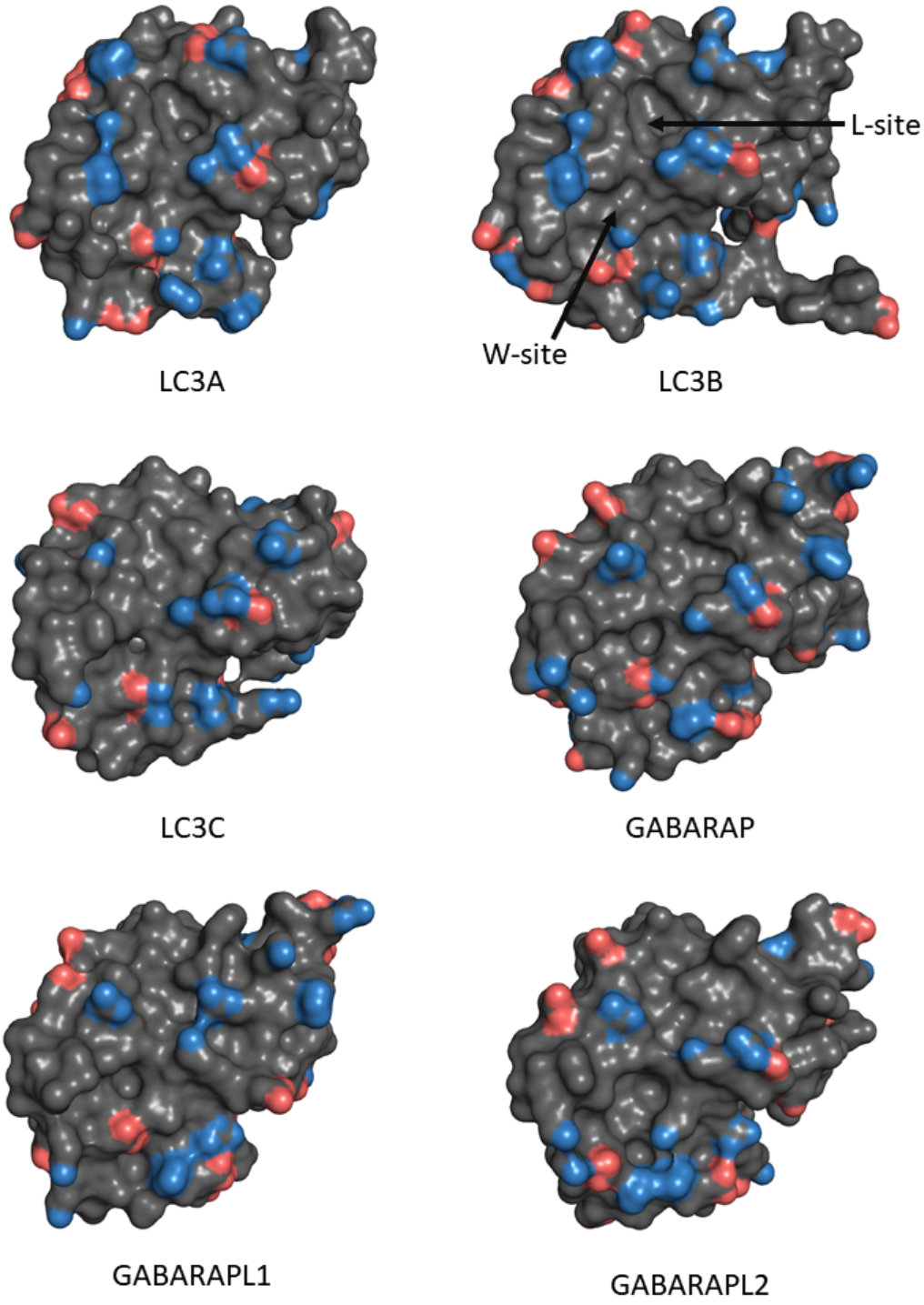



Figure 4.15: Aligned crystal structures of the hAtg8 proteins showing their surface topology and charge. Negatively charged atoms are shown in red and positively charged atoms are shown in blue. The known W- and L-sites on LC3B are also highlighted to indicate the binding patch of this protein and therefore the expected binding patch of the remaining hAtg8 proteins.

- There is a small but measureable degree of binding selectivity amongst the hAtg8 proteins in their interaction with a WT p62 AIM peptide.
- The WT p62 AIM preferentially binds to GABARAPL1 and binds to LC3C with the lowest affinity.

The quantification of the binding affinities in this chapter facilitated work in the following chapters as it allowed the effect of p62 phosphorylation (Chapter 6), LC3B phosphorylation (Chapter 7) and an ALS-related p62 mutation (Chapter 8) to be determined by comparison. The following chapter though focuses on an interaction studied in this chapter between GABARAPL2 and the WT p62 AIM peptide.

# Chapter 5

# NMR Assignment and Characterisation of GABARAPL2

## 5.1 Introduction

Multidimensional, heteronuclear NMR experiments were used for assignment and subsequent characterisation of the GABARAPL2-WT p62 AIM peptide interaction. The characterisation of the thermodynamics of this interaction, shown in Chapter 4, demonstrated that this binding was unique as it is entropically driven with a low enthalpic contribution to binding compared to the other hAtg8 proteins. The binding affinity of this interaction was also shown to be significantly weaker than for LC3A, LC3B, GABARAP and GABARAPL1. Therefore, further study to elucidate how the specific surface charge and topology of this protein causes these differences was warranted.

## 5.2 Backbone Amide Assignment

To allow 3D assignment experiments $^{13}$C,$^{15}$N-GABARAPL2 was over-expressed using a isoptope labeling protocol and subsequently purified as described in Section 2.3 and 2.5. An assignment of the GABARAPL2 $^{1}$H-$^{15}$N-HSQC spectrum had previously been published which identified peaks corresponding to 91% of the non-proline backbone amide resonances [171]. A $^{1}$H-$^{15}$N-HSQC spectrum was acquired and compared to this previous spectrum. The chemical shift RMSD was determined between the signals in the published $^{1}$H-$^{15}$N-HSQC spectrum and the spectrum produced here. This gave an RMSD of 0.10 ppm in the $^{1}$H dimension and 0.09 ppm in the $^{15}$N dimen-

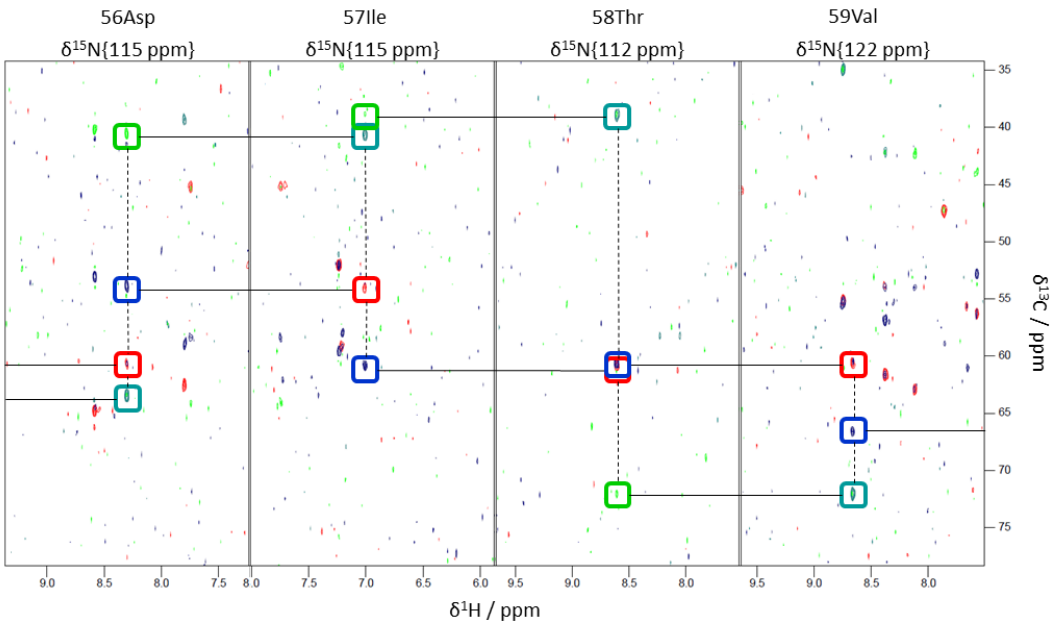



Figure 5.1: Strips of the overlaid HNCACB (blue and green) and HN(CO)CACB (red and torquoise) spectra of GABARAPL2. The strips are taken at different $^{15}$N chemical shifts as indicated. Allignment of peaks in the $^{13}$C dimension from the two spectra at different $^{1}$H and $^{15}$N amide resonances implies connectivity between these amides in the backbone.

sion. Visual comparison of the spectra also confirmed the overlapping nature of the spectra which along with the low RMSD values justified the transfer of the assignment from the work of Ma *et al.* to the spectrum collected here.

The validity of this assignment was also confirmed in the work presented here. Two primary sets of 3D NMR experiments were utilised for the assignment procedure for GABARAPL2; namely HNCO/HN(CA)CO and HNCACB/HN(CO)CACB. The assignment was carried out using the standard sequential assignment method described in Section 2.6.4. An example of the assignment process is highlighted in Figure 5.1 where strips showing an overlay of both the HNCACB and HN(CO)CACB spectra of GABARAPL2 are shown. These strips were taken at the $^{1}$H and $^{15}$N chemical shifts of the NH resonance; corresponding to a single peak on the $^{1}$H-$^{15}$N-HSQC spectrum. From this it can be seen that at these resonances, there were peaks at different $^{13}$C frequencies from the CA and CB atoms in that particular amino acid residue (HNCACB) or the preceding residue (HNCACB and HN(CO)CACB).

One pair of CA and CB signals were observed per strip, as expected, in the HN(CO)CACB experiment which correspond to the atoms in the $i-1$ residue. The HNCACB experiment can produce signals for the CA and CB atoms on the $i$ and

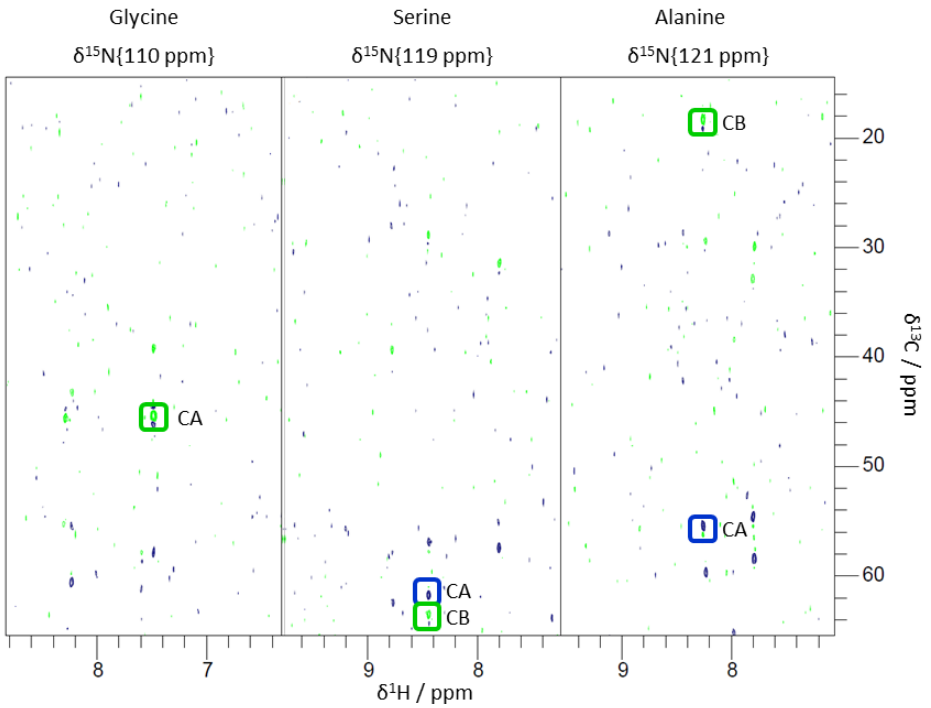



Figure 5.2: Strips from the HNCACB spectrum of GABARAPL2 highlighting the unique CA and CB signals of glycine (left), serine (middle) and alanine (right). Boxed peaks correspond to $i$ CA and CB signals.

$i-1$ residues of a given NH. However, Figure 5.1 shows that the HNCACB spectrum often only produced one pair of CA and CB resonances per NH signal. This absence of signals did not affect the ability to assign the NH backbone peaks from the data collected as the required information was still present in the 3D spectra.

As an example of how the sequential assignment was performed from these strips, the link between 56Asp and 57Ile will be considered. At the same $^{1}$H and $^{15}$N chemical shifts of the 56Asp signal in the $^{1}$H-$^{15}$N-HSQC, the HNCACB experiment produced two peaks at different $^{13}$C chemical shifts which corresponded to the CA and CB of the $i$ residue. They were coloured differently in the spectrum to indicate their opposite phasing which occured due to the nature of the experiment. In these experiments signals are multiplicity edited which means peaks from atoms in a CH or $CH_3$ group have the opposite phasing from atoms in a $CH_2$ group.

At the $^{1}$H and $^{15}$N chemical shifts of the 57Ile signal in the $^{1}$H-$^{15}$N-HSQC, the HN(CO)CACB experiment produced two peaks with matching $^{13}$C chemical shifts to the CA and CB shifts from the 56Asp in the HNCACB experiment. In this way the two signals from the $^{1}$H-$^{15}$N-HSQC were sequentially linked.

In order to assign residues to particular $^{1}$H-$^{15}$N-HSQC peaks, characteristic res-

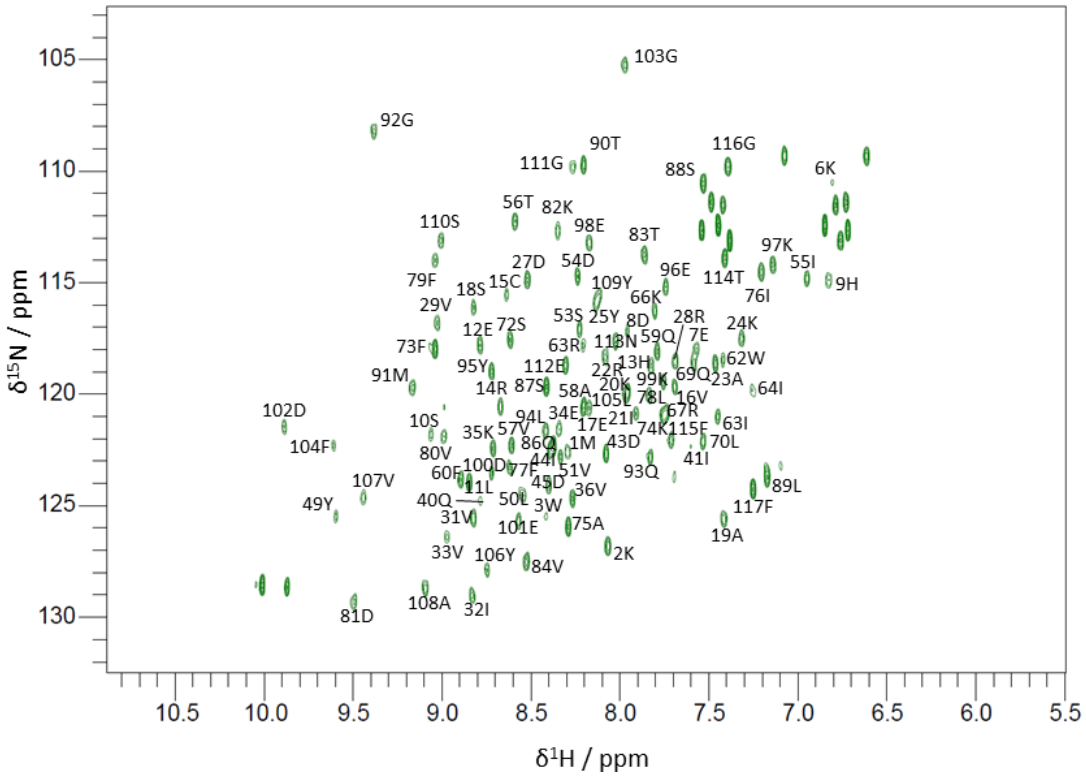



Figure 5.3: $^{1}$H-$^{15}$N-HSQC of $^{15}$N-GABARAPL2 showing the assignment of backbone amide NH signals.

onances for the CA and CB peaks were identified. Figure 5.2 shows some examples of characteristic CA and CB shifts from the HNCACB spectrum of GABARAPL2. This shows a glycine residue which does not have a CB atom and has a $^{13}$CA chemical shift that is particularly upfield. The CA signal for glycine also has opposite phasing to the other CA signals shown as it is a $CH_2$ rather than a CH. Serine (and also threonine) signals are unique as the $^{13}$CB chemical shift is more downfield than the $^{13}$CA shift; the opposite of all other naturally occuring amino acid. The defining characteristic of alanine can be less clear than the previous two examples though it is typically distinguishing. Alanine is the only amino acid residue with a methyl CB which results in a chemical shift for this atom at a significantly lower ppm value than for other $^{13}$CB signals.

The sequential linking of $^{1}$H-$^{15}$N-HSQC signals was combined with this information on amino acid type to perform the assignment. This allowed a sequentially linked string of known and unknown amino acids to be compared to the known sequence of the protein until an unambiguous match was found. After all possible assignments were made from the data collected here, the assignment was compared to the previ-

| Amino Acid Type | Total | Assigned | % assigned | Assigned by Ma et al. | % assigned by Ma et al. |
|---|---|---|---|---|---|
| **Ala** | 5 | 5 | 100 | 5 | 100 |
| **Arg** | 6 | 3 | 50 | 5 | 83 |
| **Asn** | 1 | 1 | 100 | 1 | 100 |
| **Asp** | 8 | 3 | 38 | 8 | 100 |
| **Cys** | 1 | 0 | 0 | 1 | 100 |
| **Gln** | 5 | 3 | 60 | 5 | 100 |
| **Glu** | 9 | 6 | 67 | 9 | 100 |
| **Gly** | 6 | 4 | 67 | 4 | 67 |
| **His** | 2 | 0 | 0 | 2 | 100 |
| **Ile** | 9 | 7 | 78 | 8 | 89 |
| **Leu** | 7 | 3 | 43 | 7 | 100 |
| **Lys** | 12 | 6 | 50 | 11 | 92 |
| **Met** | 4 | 2 | 50 | 3 | 75 |
| **Phe** | 7 | 6 | 86 | 6 | 86 |
| **Pro** | 5 | 0 | 0 | 0 | 0 |
| **Ser** | 10 | 4 | 40 | 8 | 80 |
| **Thr** | 4 | 4 | 100 | 4 | 100 |
| **Trp** | 2 | 1 | 50 | 2 | 100 |
| **Tyr** | 5 | 3 | 60 | 5 | 100 |
| **Val** | 11 | 4 | 37 | 10 | 91 |
| **All Residues** | 119 | 65 | 55 | 104 | 87 |
| **All Residues - Pro** | 114 | 65 | 57 | 104 | 91 |

Table 5.1: Overview of the GABARAPL2 amide backbone assignment by amino acid type indicating that 91% of assignable residues have been identified.

ously published work and it was observed that all assigned peaks were in agreement with the previous work. The assigned $^{1}$H-$^{15}$N-HSQC spectrum is shown in Figure 5.3 and the level of assignment achieved is described in Table 5.1.

The assigned residues have been mapped onto the structure of GABARAPL2 in Figure 5.4 to show the coverage of the assignment. This illustrates that 10 of the 15 unassigned residues corresponded to random coil sections of the protein structure with the other five occuring in positions with defined secondary structure. Four of the 10 unassigned random coil residues and one of the five unassigned structured residues were prolines which cannot be assigned by this methodology.

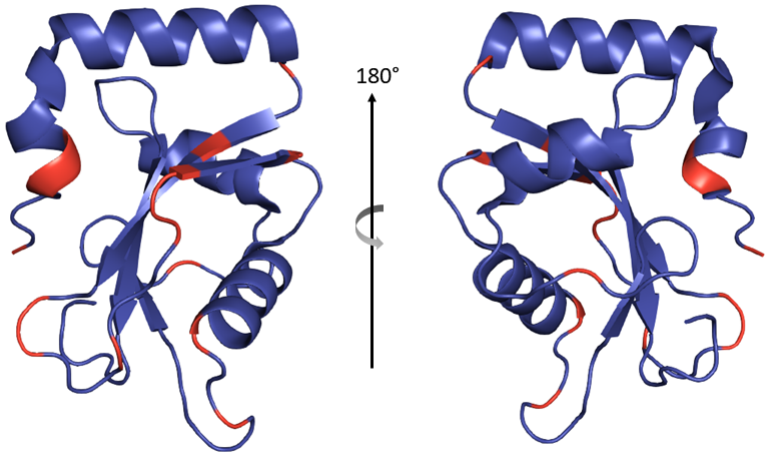



Figure 5.4: Crystal structure of GABARAPL2 coloured to illustrate the assignment of backbone amide residues in the $^{1}$H-$^{15}$N-HSQC spectra. This is shown as a cartoon schematic from two angles and assigned residues are coloured in blue and unassigned residues are coloured red.

## 5.3 GABARAPL2-WT p62 AIM Binding Study

Isotopically-labeled $^{15}$N-GABARAPL2 was over-expressed and purified as described in Section 2.3 and 2.5 for use in this binding study. A $^{1}$H-$^{15}$N-HSQC NMR titration experiment was used to investigate the binding of the WT p62 AIM peptide to $^{15}$N-GABARAPL2. This type of study allowed chemical shift mapping to attempt to define the site of peptide binding on the protein through chemical shift perturbation (CSP) analysis.

Titration of the WT p62 AIM peptide into a $^{15}$N-labeled sample of GABARAPL2 was conducted incrementally at constant GABARAPL2 concentration until a 4:1 peptide to protein ratio was reached; with a $^{1}$H-$^{15}$N-HSQC spectrum collected at each point. Spectra of the unbound and bound $^{15}$N-GABARAPL2 protein are shown in Figure 5.5 and the full spectral overlay of each titration point produced throughout the experiment is shown in Appendix Figure A.12.

The majority of assigned signals (73%) in the $^{1}$H-$^{15}$N-HSQC spectrum underwent fast exchange or intermediate exchange with signals which could be followed through the titration. This allowed the bound form of these peaks to be assigned by following assigned peaks through the titration points. Calculation of CSP values and determination of significance was performed as described in Section 2.6.4. The results of this analysis are shown in Figure 5.6 which shows the CSP values for each amino acid residue which could be determined and which of these are above each level of

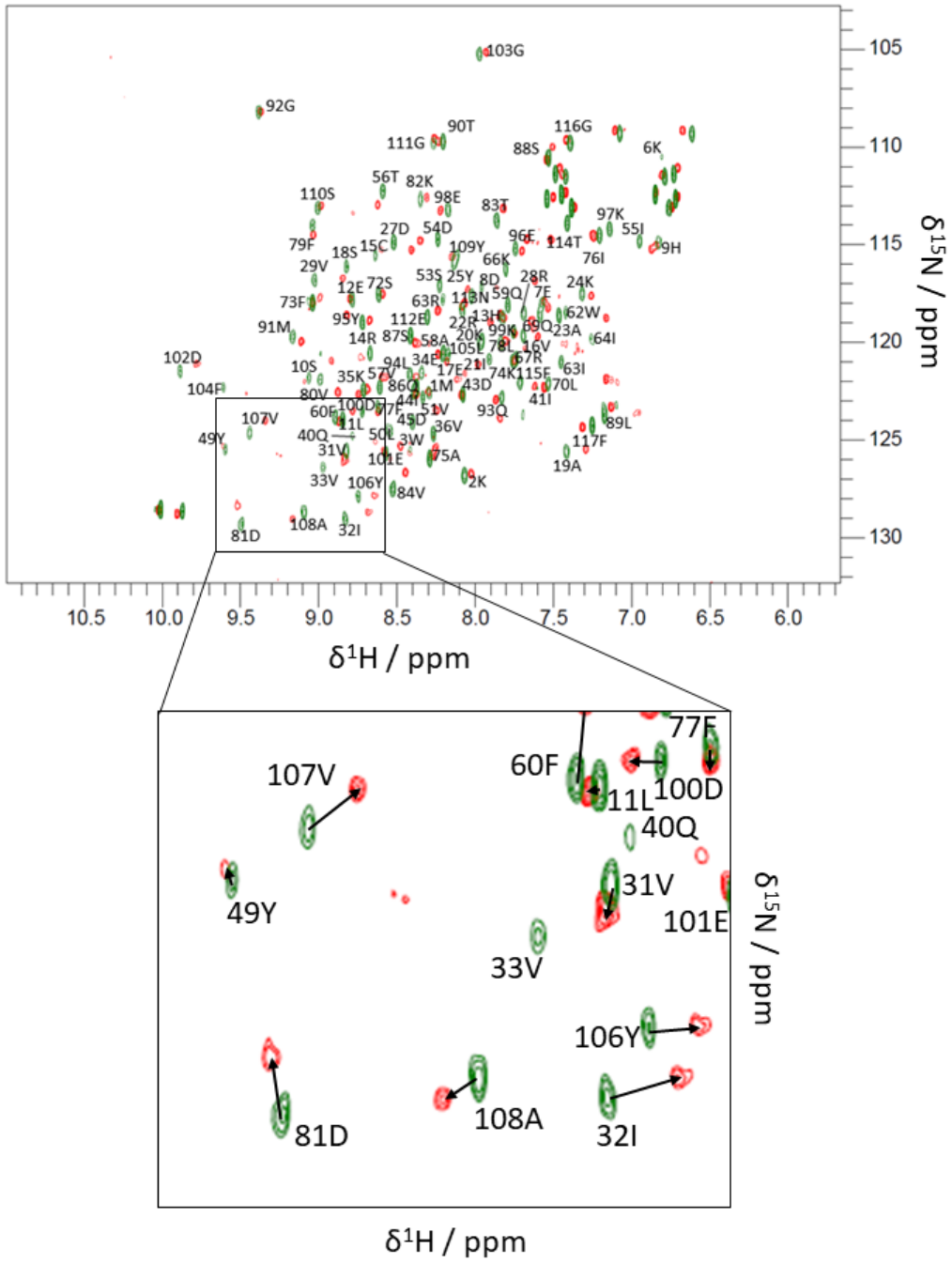



Figure 5.5: Spectral overlay of the $^{1}$H-$^{15}$N-HSQC spectra of unbound GABARAPL2 (green) and 4:1 WT p62 AIM-GABARAPL2 (red), including an insert showing the detail of a smaller region. The assignment of the unbound spectrum has been annotated and the corresponding peak in the bound spectrum is shown by the arrow.

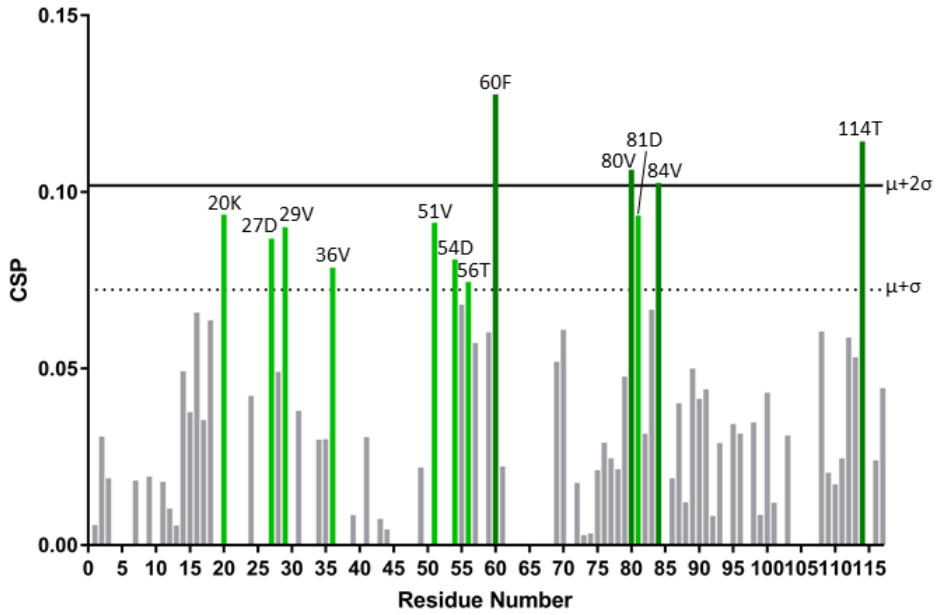



Figure 5.6: Graph illustrating the CSP values which could be determined due to fast exchange for the binding of $^{15}$N-GABARAPL2 to the WT p62 AIM peptide. The lines indicate one or two standard deviations ($\sigma$) above the mean ($\mu$). Green bars denote residues with significant CSP values.

significance.

There were 28 assigned peaks in the $^{1}$H-$^{15}$N-HSQC spectrum which could not be followed during the titration as they exhibited slow exchange behaviour or intermediate exchange with signals which could not be followed through the titration. There were then 15 additional peaks observed in the 4:1 WT p62 AIM peptide:$^{15}$N-GABARAPL2 $^{1}$H-$^{15}$N-HSQC spectrum which were unassigned. Unfortunately, attempts to perform 3D assignment experiments on the bound WT p62 AIM-$^{13}$C,$^{15}$N-GABARAPL2 were unsuccessful due to protein degradation.

Figure 5.7 highlights the amino acid residues with significantly shifting fast or intermediate exchange peaks with peaks with CSP values one or two standard deviations above the mean coloured. This structure also highlights the residues with signals that exhibited slow exchange behaviour or intermediate exchange with signals which could not be followed through the titration; as these are conisdered to be the most significantly perturbed residues. Additionally, it illustrates the amino acid residues which had not been assigned to a peak in the unbound $^{1}$H-$^{15}$N-HSQC spectrum and thus their significance could not be investigated.

Estimation of the binding affinity of this interaction was possible using Equation 2.8. This was carried out using the fitting function within CCPNMR Analysis. The

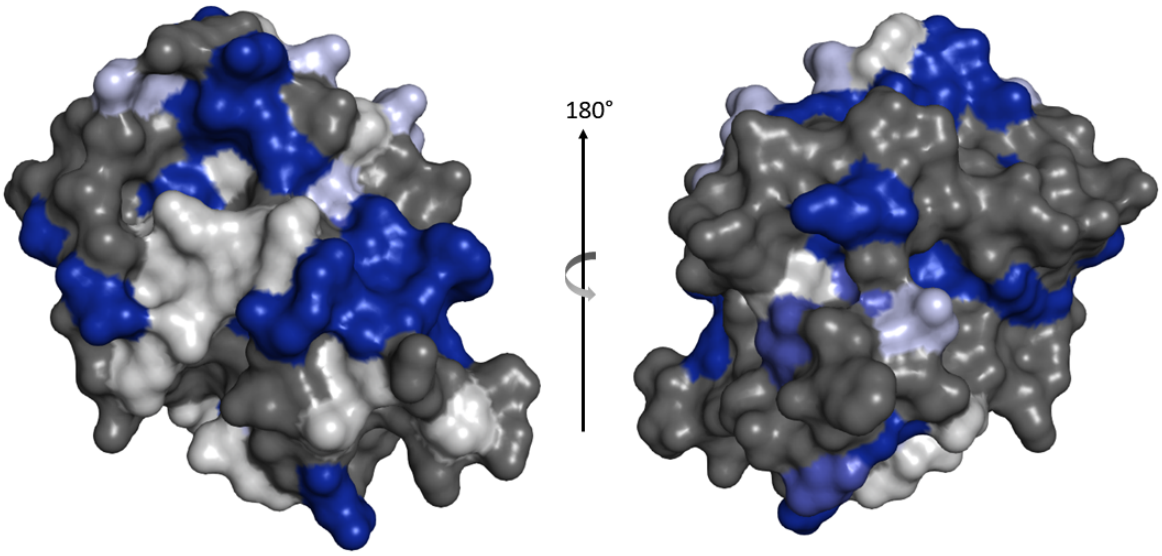



Figure 5.7: Crystal structure of GABARAPL2 which illustrates potential binding sites for the WT p62 AIM peptide based upon the NMR chemical shift mapping experiment. This shows the residues with significantly shifted signals (dark blue shows slow exchanging residues; medium and light blue correspond to the CSP values two and one standard deviations above the mean respectively), the residues without significantly shifted signals (grey) and the residues which were not assigned to peaks in the $^{1}$H-$^{15}$N-HSQC spectrum (white).

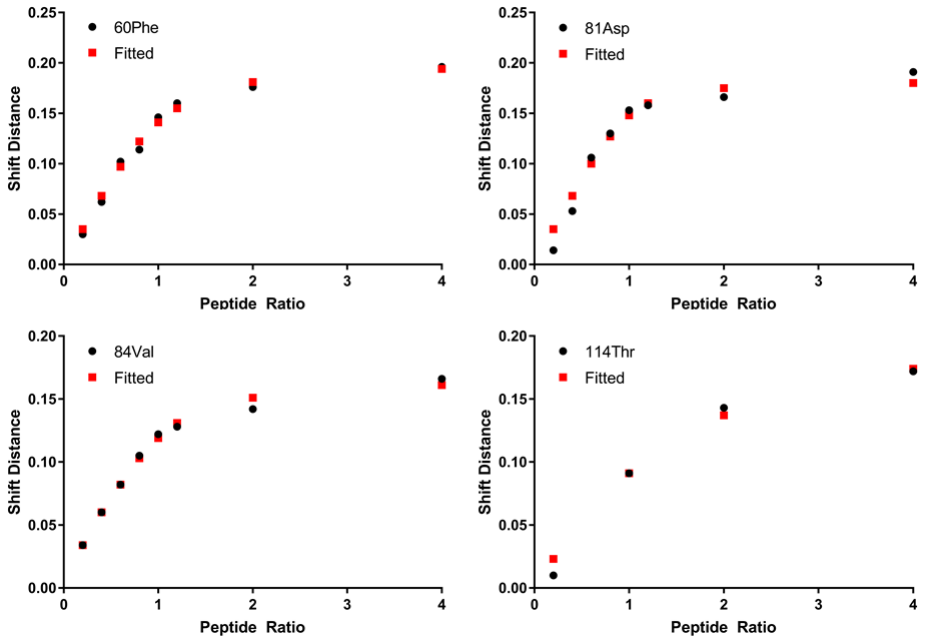



Figure 5.8: Fitting of the CSP data for the four significant shifts determined for the binding of the WT p62 AIM peptide to GABARAPL2.

| Peak | Determined $K_D$ ($\mu$M) | Error ($\mu$M) |
|---|---|---|
| Phe60 | 6.44 | 1.74 |
| Asp81 | 2.33 | 1.98 |
| Val84 | 5.93 | 1.65 |
| Asn113 | 6.14 | 2.68 |

Table 5.2: NMR-derived $K_D$ values for the interaction between GABARAPL2 and the WT p62 AIM peptide. This shows the values determined for the residues with the most significant CSP values.

CSP values for each signal at each ligand concentration were plotted and these data were fit to the equation to determine a $K_D$ value for each peak. The fitting of the data for the four fast exchanging peaks with the most significant CSP values is shown in Figure 5.8 and the determined $K_D$ values for these residues are stated in Table 5.2. These values produced an average estimate of the $K_D$ as 5.2±1.4 $\mu$M which is consistent with with the ITC-derived value of 6.85±1.02 $\mu$M.

The HADDOCK web server was then utilised to dock the WT p62 AIM peptide to GABARAPL2 according to this chemical shift mapping [190, 191]. Trp338 and Leu341 were defined as the active residues on the peptide for the docking runs. Active residues are defined within the web server to aid in docking and are considered to be of central importance for the interaction. As such they are restrained to be part of the interface or if they are removed from the interface the structure incurs a scoring penalty. GABARAPL2 residues with peaks which exhibited slow exchange behaviour during the titration were defined as the active residues as these represent the residues with the largest CSPs. A different set of parameters were also utilised in a second docking run which limited the active residues to those on the surface around the $\beta$2-strand; where the p62 AIM is known to bind to hAtg8 proteins in the previously studied examples [50, 1, 53]. The second docking run (with the shortened list of active residues) produced a cluster of structures with a significantly more favourable HADDOCK score. This is the scoring function the software operates to optimise which is based on a linear combination of calculated structural energies and therefore more negative values are favaourable. The first run produced a cluster of structures with a top HADDOCK score of -49.6±2.4 which was comprised of 15 structures whereas the second docking run produced a cluster of structure with a top score of -105.0±3.1 which was comprised of 75 structures.

The best scoring structure from the top cluster on the second run, i.e. the lowest energy structure produced, is shown in Figure 5.9 with the GABARAPL2 surface coloured according to the chemical shift mapping experiment. In this model, the sig-

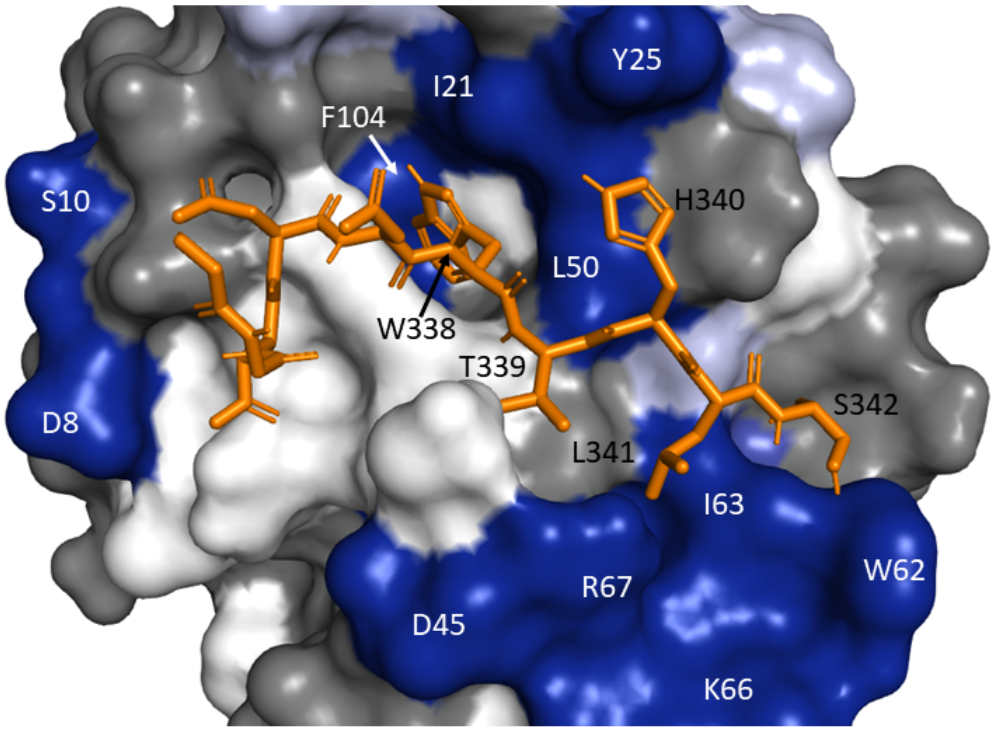



Figure 5.9: Structure of the GABARAPL2-WT p62 AIM peptide structure produced by the HADDOCK web server. This shows the peptide (orange) and the GABARAPL2 residues with significantly shifted signals (dark blue shows slow exchanging residues; medium and light blue correspond to the CSP values two and one standard deviations above the mean respectively), the residues without significantly shifted signals (grey) and the residues which were not assigned to peaks in the $^{1}$H-$^{15}$N-HSQC spectrum (white). Residues annotated in white correspond to the slow exchanging residues on GABARAPL2 and annotations in black correspond to peptide residues.

nificantly shifted Ile21, Leu50 and Phe104 residues as well as a number of unassigned residues form a hydrophobic binding pocket for the Trp side chain. Ile63 forms a hydrophobic patch on the GABARAPL2 surface which the Leu341 sidechainn can pack against. The area around this binding site, comprised of Asp45, Trp62, Lys66 and Arg67 has also been significantly shifted.

## 5.4 Discussion

### Assignment

The $^{1}$H-$^{15}$N-HSQC spectrum can be considered as the fingerprint of a protein under a given set of conditions and was used as the basis of the backbone assignment

in this work. The HNCACB/HN(CO)CACB pair of experiments was preferred for sequential protein assignment. Sequential links are more reliably assigned from these experiments as there are two peaks which can be correlated (CA and CB) rather than the one in the HNCO/HN(CA)CO experiments (CO). In addition to this there is a greater dispersion in chemical shifts among the CA and CB signals compared to the dispersion of CO signals; making non-linked, coincidental overlap of signals less likely. The HNCO/HN(CA)CO pair of experiments provide an additional useful confirmation of sequential linkages found in the HNCACB/HN(CO)CACB experiments though, as this provides three resonances which can be correlated: CO, CA and CB.

Peaks in the 3D spectra were first paired with peaks in the $^{1}$H-$^{15}$N-HSQC spectrum due to their matching $^{1}$H and $^{15}$N chemical shifts. However, many peaks in the $^{1}$H-$^{15}$N-HSQC spectrum did not have corresponding observable peaks in the 3D spectra and this precluded them from assignment. This was explained in some cases by their location in random coil regions of the protein structure. These residues are more likely to be conformationally flexible which results in the nucleus experiencing a range of chemical environments. The flexibility leads to an averaging of these environments so the signal is broadened which reduces signal intensity at the peak (though the intesnity of the full peak remains the same and is simply spread). Five of the 15 unassigned residues were prolines which cannot be assigned by this methodology as they do not produce signals in the $^{1}$H-$^{15}$N-HSQC spectrum due to their distinctive ring structure which does not have a backbone NH.

The signals from the $i-1$ residue CA and CB atoms in a HNCACB experiment are often lower in intensity than for the $i$ residue so the intensity of these signals may be too low to be observed here due to the relatively low concentration of protein (100 $\mu$M) in the sample. Additionally, as the HNCACB signals occured at different $^{13}$C chemical shifts from the HN(CO)CACB signals for the same NH it can be deduced that the HNCACB signals were for the $i$ residue. Therefore in the cases where only one set of CA and CB signals were observed in the HNCACB spectrum, these signals were assigned to the $i$ residue.

All of the peaks assigned in the $^{1}$H-$^{15}$N-HSQC of this work correlated with the previous assignment by Ma *et al.* with no additional assignment achieved. The RMSD values between the peaks in the two spectra were 0.10 ppm in the $^{1}$H dimension and 0.09 ppm in the $^{15}$N dimension which illustrates a clear overlap of the peaks. This assignment confirmation and clear overlay of signals allowed the transfer of the previous assignment onto the spectrum collected here.

## Chemical Shift Mapping

NMR titrations for chemical shift mapping are useful as the changing chemical environment of the backbone amides upon peptide binding results in movement of the $^{1}$H-$^{15}$N-HSQC peaks which is proportional to involvement in the binding interaction. In this way the chemical environment of the backbone amide of a particular amino acid residue is used as a proxy for changes to the residue as a whole.

When considering the movement of peaks throughout the titration, the exchange regime played a crucial role. Exchange behaviour is discussed in terms of fast, intermediate or slow exchange which can be considered in terms of the association and dissociation rates of two molecules and the chemical shifts of the bound and free signals. A signal exhibits slow exchange behaviour when the rate of exchange between bound and free protein, in Hz, is significantly lower than the difference, also in Hz, between the chemical shifts of the free and bound signals. As such two peaks are observed in the spectrum which correspond to the bound or unbound form. The intensities of these peaks are weighted by the ratio of bound to free protein in the sample.

Conversely, a signal exhibits fast exchange behaviour when the rate of exchange between the two states is significantly higher than the difference between the chemical shifts of the free and bound signals, both in Hz. This results in signals which occur at chemical shifts that are a weighted average of the chemical shifts from the bound and unbound states. Depending on the nature of intermediate exchanging peaks, they can fall into either side of the fast/slow dichotomy as the rate of exchange between the two states is roughly the same as the difference in their chemical shifts. Often the peaks can be followed as for fast exchanging peaks however this exchange regime is also associated with changes in linewidth. Signals are typically observed to broaden then as the signal is closer to the bound state there is a sharpening of the signal. As the peak broadens it may be the case that the signal becomes indistinguishable from the noise which, in some cases, prevents the signal from being unambigiously followed throughout the titration. It is therefore possible to see signals with a mixture of exchange regimes from the same binding event; as the magnitude of chemical shift perturbation varies between signals whilst the physical rate of exchange remains the same.

Signals exhibiting behaviour corresponding to each exchange regime are illustrated in Figure 5.10 at various titration points. This shows the fast exchange behaviour of Lys2, the intermediate behaviour of Ala108 and the slow exchange behaviour of Ile32

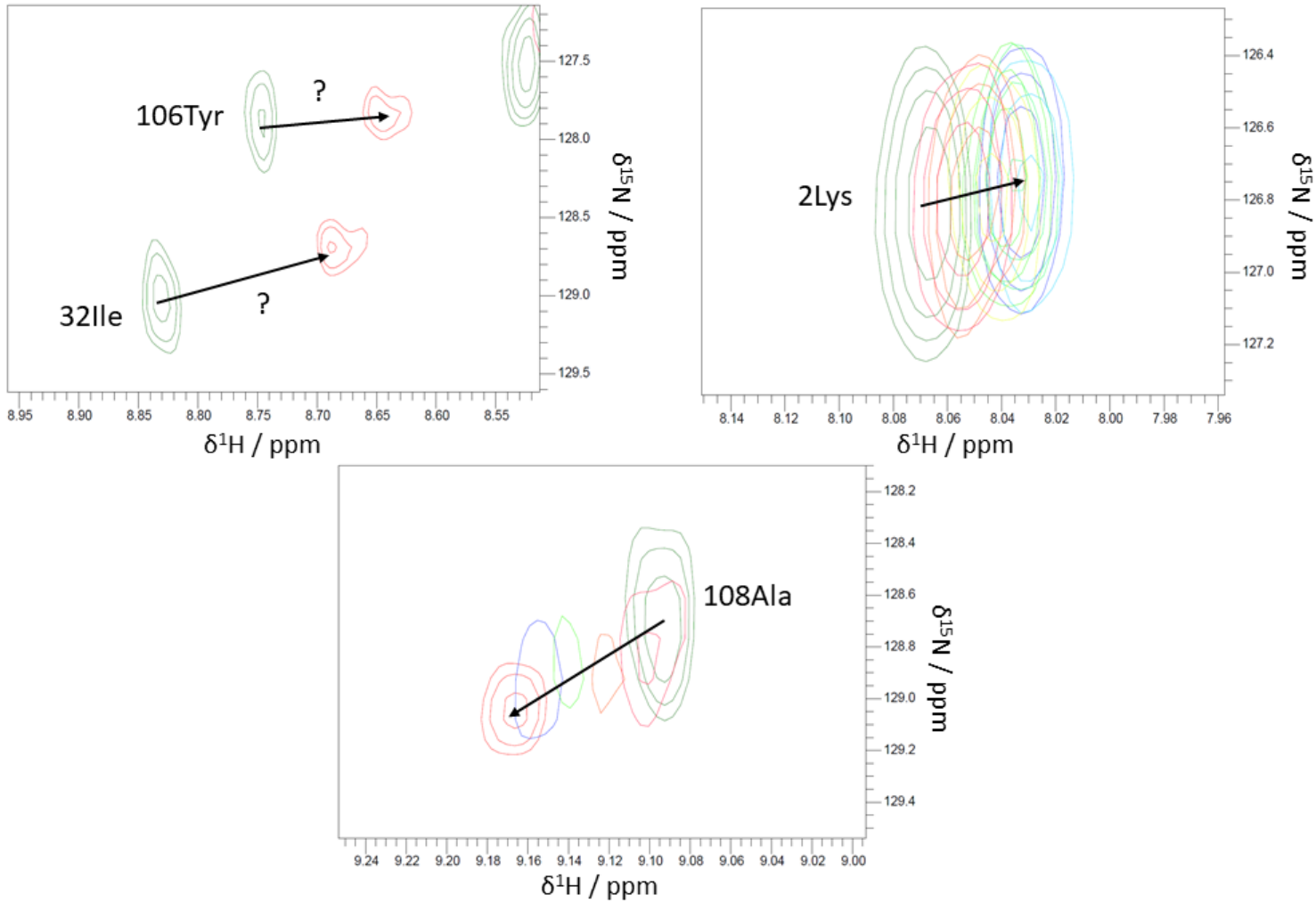



Figure 5.10: $^{1}$H-$^{15}$N-HSQC spectra of GABARAPL2 illustrating the differing exchange regimes for different NH signals throughout the titration with the WT p62 AIM peptide. 32Ile and 106Tyr are in slow exchange, 2Lys is in fast exchange and 108Ala is in intermediate exchange.

and Tyr106. This illustrates that if the peaks are in fast exchange it is possible to follow peaks throughout the titration as the chemical shifts are altered incrementally. This allows unambigous assignment of the bound fast exchanging peaks as the assignment is simply propagated through the signals at each titration point. This is not the case for peaks in the slow exchange regime which decrease in intensity whilst the intensity of the corresponding bound peak increases elsewhere, as the titration progresses. This was observed for Ile32 and Tyr106 and although there are bound peaks in close proximity it is not possible to assign them on that basis alone. It should also be noted that for these two signals, upon the addition of the first aliquot of peptide (0.2:1) no peak could be observed in the region and no peaks were observed nearby until the final titration point was reached. This prevented any attempts to correlate the unbound peaks with bound peaks. In the case of Ala108, which is in intermediate exchange, it was possible to follow the signal as it moved throughout the titration but the changes to the peak linewidth can be observed.

Due to the rate of association and dissociation of GABARAPL2 and the WT p62 AIM peptide a mixture of exchange regimes were observed. The signals which had small differences in chemical shift between their bound and free chemical shifts exhibited fast exchange behaviour and the signals which had a larger chemical shift difference between their bound and free chemical shifts exhibited slow exchange behaviour.

There were fewer signals in the 4:1 spectrum than in the unbound 0:1 spectrum which indicated that peaks in the bound spectrum had broadened due to peptide binding. This raised the possibility that a higher peptide concentration would have produced a spectrum where these peaks, which had broadened so as to become indistinguishable from the noise, had sharpened and become observable again. However, based upon the graphs of CSPs in Figure 5.8, the binding appeared to be approaching saturation as the curves were leveling off.

## GABARAPL2-WT p62 AIM Structure

By the definition previously discussed, the residues with peaks in slow exchange experienced the largest perturbations upon binding and were therefore used to define the GABARAPL2-WT p62 AIM binding site. These significantly shifted residues on the surface predominantly centre around the known hAtg8 binding patch for AIM peptides around the $\beta$2-strand. This is analogous to the various known structures of the hAtg8-AIM interaction including LC3B-p62, GABARAPL1-NBR1 and GABARAP-

calreticulin [50, 1, 53].However, there was also a distribution of residues with slow exchanging peaks across the protein surface which may have been taken to imply non-specific binding across the surface. However, data from ESI MS and ITC studies in Chapter 4 imply a 1:1 binding ratio which supports a single binding site in addition to analogy with previous structures.

These distributed residues with significant shifts can instead be explained by considering structural shifts of the protein to accomodate binding. Figure 5.11 shows the protein in a simplified ribbon schematic to show secondary structural elements. The distribution of residues with large CSP values indicate that binding to the site around the $\beta$2-strand leads to a shift in structure that propagates through the protein. There are slow exchanging peaks which correspond to neighbouring residues in the $\beta$-sheet which may indicate that these residues were shifting to accomodate binding and this effect then propagates through the hydrogen bonding network within the structure, altering the chemical environment of all shifted residues. The binding of AIM peptides to hAtg8 proteins is characterised by the formation of an intermoleular $\beta$-sheet so the propagation throughout the sheet is consistent with this model. This indicated a significant shift in protein structure to accomodate binding of this peptide. This propagation effect was also previously observed in the p62 AIM binding to LC3B.

For both docking runs, the lowest energy structure was found by binding of the peptide to the site around the $\beta$2-strand; analogous to previous hAtg8-AIM structures. This is the largest continuous patch of significantly shifted residues on the protien surface. For the first docking run which utilised all slow exchanging surface peaks as active residues, this was the only cluster of structures which was output with binding at this site. The remaining clusters of structures involved binding elsewhere with significantly higher energy structures. For the second docking run, all of the clusters of structures involved the peptide binding at this site site. The lowest energy structure produced from this docking experiment was shown in Figure 5.9 which shows the peptide binding around the $\beta$2-strand of GABARAPL2. This shows that during docking, even when binding to any slow exchanging surface residue was possible, the site around $\beta$2 was preferred and the top scoring structure simply indicated the optimised conformation of the peptide at this site.

Caution should be taken when interpreting results from a simplified docking experiment to not rely too heavily on the outputted structure however the structure is useful in visualising the interaction and can be justified with experimental data. The structure produced from HADDOCK web server docking binds in an analogous manner to the interaction of the same peptide with LC3B which allows comparison.

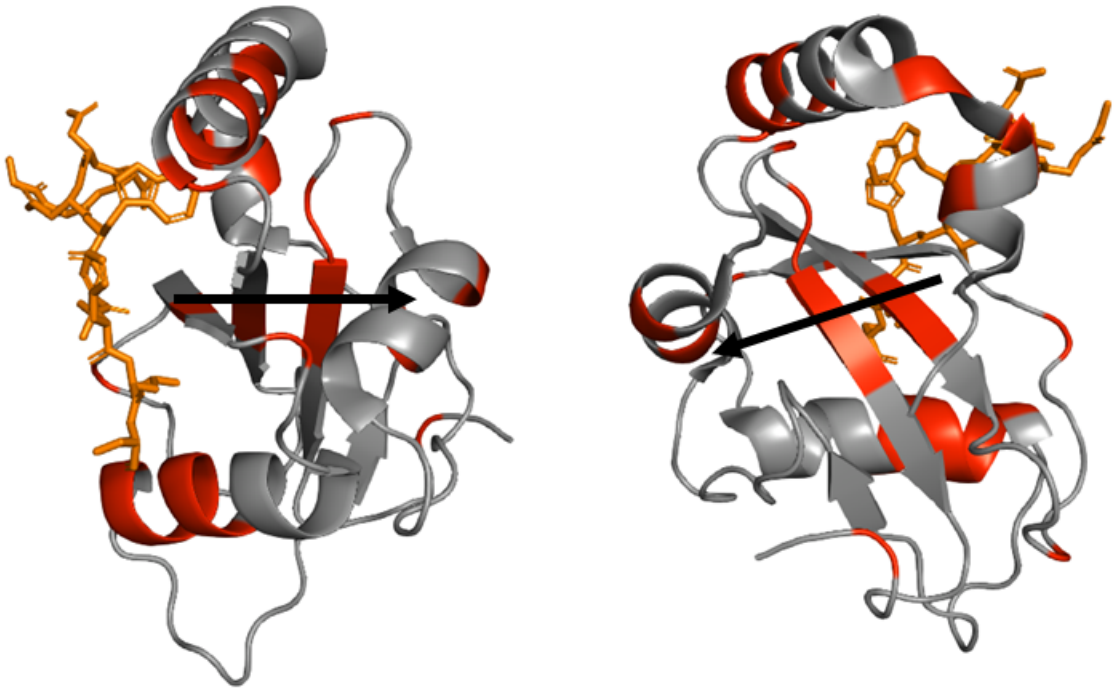


Figure 5.11: Crystal structure of GABARAPL2 from two angles with the residues most perturbed according to NMR by the p62 AIM peptide (shown in orange) binding coloured in red. This illustrates that binding of the peptide to the expected binding site propagates structural shifts through the protein structure, with a black arrow illustrating this propagation through the structure.

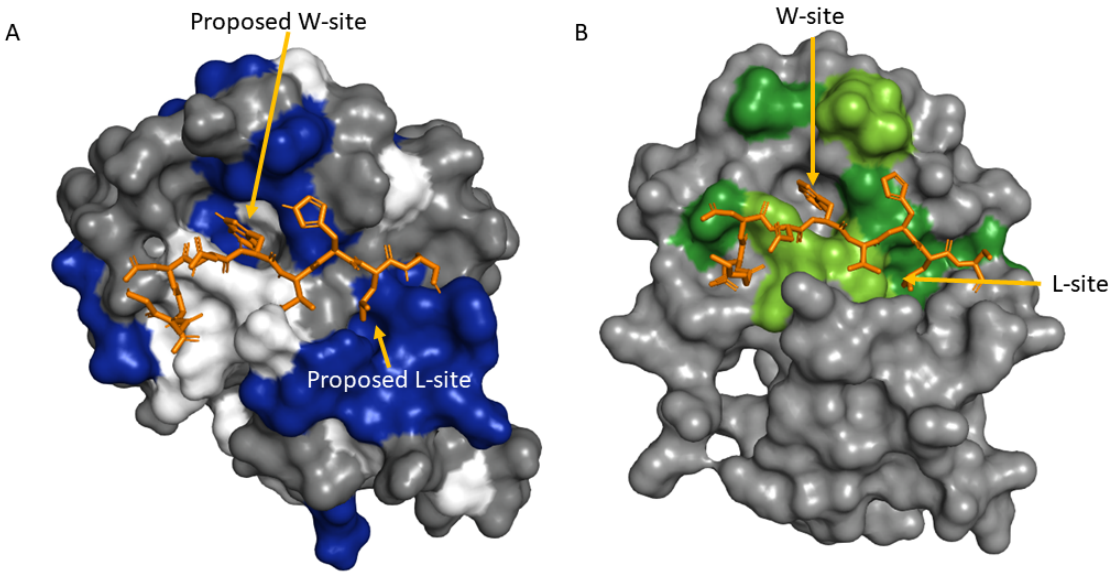



Figure 5.12: NMR data for the binding of the WT p62 AIM peptide to LC3B and GABARAPL2. (A) Structure of GABARAPL2 which illustrates a potential binding site for the WT p62 AIM peptide based upon the NMR chemical shift mapping experiment-based docking run. This shows the residues with significantly shifted signals (dark blue), the residues without significantly shifted signals (grey) and the residues which were not assigned to peaks in the $^{1}$H-$^{15}$N-HSQC spectrum (white). (B) Crystal structure of LC3B bound to the WT p62 AIM peptide coloured according to a previous NMR chemical shift mapping experiment (dark green and light green correspond to CSP values two and one standard deviations above the mean respectively).

Figure 5.12A shows the colour-coded structure of the GABARAPL2-WT p62 AIM peptide structure produced using the HADDOCK web server. The W- and L-sites from this proposed structure are indicated. For comparison Figure 5.12B shows NMR data from the work by Goode *et al.* which illustrates the NMR-derived binding site of the WT p62 AIM peptide on the LC3B surface mapped onto the known crystal structure of the complex. The modelled binding of the peptide to GABARAP2 shows that the Trp338 of the peptide packs into an analogous W-site on the surface which is formed of significantly shifted surface residues and some unassigned residues. The binding of Leu341 onto GABARAPL2 clearly occurs with some variation to the L-site compared with LC3B though. There is no defined cleft on GABARAPL2 for this residue to bind but instead there are a number of significantly shifted residues on the surface which form a proposed surface L-site. This is supported by visual inspection of the unbound crystal structures of these two proteins where there is a defined cleft on the surface of LC3B at the L-site which is not present for GABARAPL2. This degree of preorganisation allows the Leu341 to bind into this hydrophobic cleft which produces a stronger binding interaction.

Figure 5.13 compares the insertion of the Trp338 and Leu341 sidechains of the WT p62 AIM peptide into these W- and L-sites of GABARAPL2 or LC3B. This elucidates a structural rationalisation for the variation in the thermodynamics of these interactions determined in Chapter 4 and shown again in Figure 5.14. In both cases Trp338 of the p62 AIM inserts into a binding cleft between $\alpha 2$ and $\beta 3$. The dominant difference arises when looking at the binding of the Leu341 sidechain. For binding to the LC3B surface this hydrophobic sidechain inserts into a binding cleft between $\alpha 3$ and $\beta 2$ however when binding to GABARAPL2 this insertion into the globular fold does not occur as no cleft exists. The overlay of the two structures in Figure 5.13C particularly illustrates this shift in the binding of Leu341 between the two proteins. The sidechain lies perpendicular to the surface. This results in less tight binding and therefore a weaker affinity for GABARAPL2. The effect of this binding can be further interpreted by looking at the highly favourable enthalpic contribution to binding in the case of LC3B which is compensated to some extent by an unfavourable entropy from this tight hydrophobic binding. The enthalpic contribution to binding with GABARAPL2 is significantly lower as the strong binding interaction cannot occur and the interaction also has a small favourable entropic component. This suggests that the hydrophobic effect of the binding is actually similar in the two interactions but instead this deep binding pocket limits conformational flexibility in the case of LC3B which results in an unfavourable entropy change upon binding.

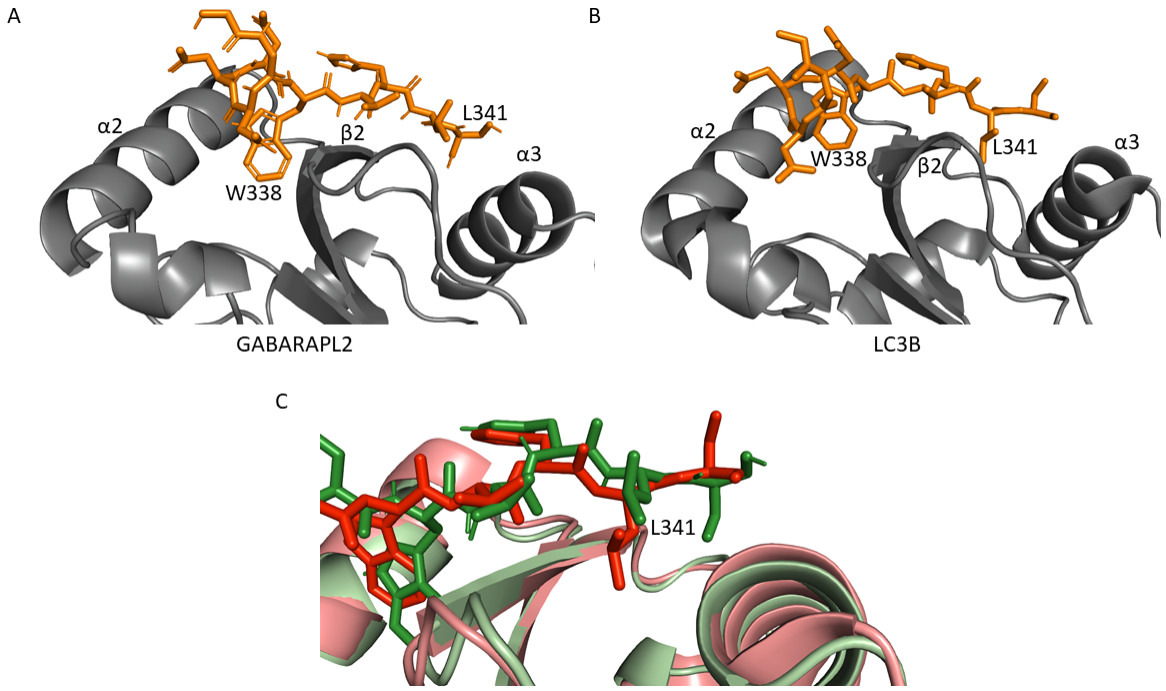



Figure 5.13: Structures of GABARAPL2 (A) and LC3B (B) binding to the WT p62 AIM to illustrate the insertion of Trp338 and Leu341 into the protein fold. (C) is an overlay of these two structures focused on the insertion of Leu341 of p62 (red for binding to LC3B; dark green for binding to GABARAPL2) into the fold of LC3B (pink) or GABARAPL2 (light green).

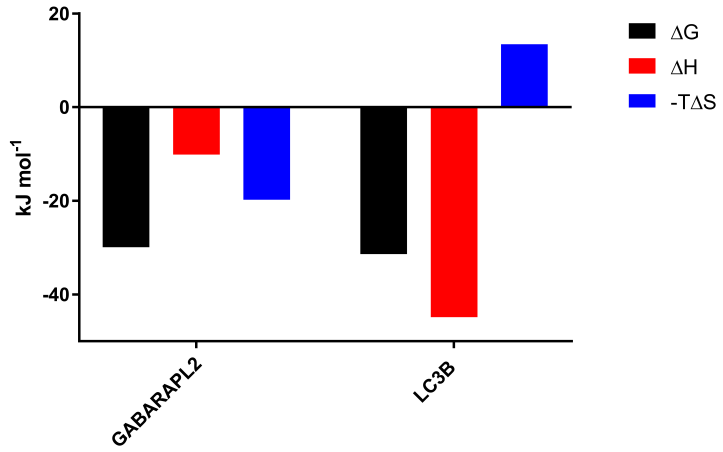



Figure 5.14: Graph comparing ITC-derived thermodynamic parameters for the binding of the WT p62 AIM peptide to GABARAPL2 or LC3B to illustrate the differences in the enthalpic and entropic contributions to the Gibbs free energy.

The data presented in this chapter has shown that:

- The assignment from the previously published $^{1}$H-$^{15}$N-HSQC spectrum of GABARAPL2 could be transferred to the spectrum collected here.
- The WT p62 AIM binds to the site around the $\beta$2-strand of GABARAPL2 which is a related site to the other known hAtg8-AIM interactions.
- The p62 AIM Leu341 residue binding site is significantly different on the GABARAPL2 surface than on the previously investigated LC3B surface despite similar Trp338 binding sites.

Having studied the interaction between the WT p62 AIM peptide and the hAtg8 proteins in this chapter and the previous one, the following two chapters investigate the potential regulation of this important protein-protein interaction by phosphorylation.

# Chapter 6

# Phosphorylation of p62 for Autophagy Regulation

## 6.1 Introduction

After the investigation of p62 binding to the hAtg8 proteins shown in Chapter 4, this and the following chapter will look at the effect of phosphorylation on this key protein-protein interaction in autophagy. TBK1 has been implicated in the process of autophagy and of particular interest in this work are the phosphorylation sites of this kinase on p62. Work presented in this chapter initially explored the effect of TBK1 phosphorylation on full length p62 binding to LC3B; one of the hAtg8 proteins. Phosphorylation of p62 at S342 specifically was then investigated by comparing the interactions of the WT and pS342 p62 AIM peptides with the hAtg8 proteins using related methodologies to those discussed in Chapter 4. Finally, the structural effect of this phosphorylation was studied in the case of LC3B.

## 6.2 TBK1 as a putative kinase

In this experiment, full length GST-tagged p62 was expressed in BL21 (DE3) *E. coli.* cells. Lysate from these cells was incubated with Glutathione Sepharose 4B beads. The tag allowed the p62 to bind to the beads which, after washing, were incubated with either TBK1 or a buffer control. After a wash step, the beads were incubated with LC3B and the protein bound to the beads was detected by western blot. Figure 6.1 shows that there was a more intense band when detecting LC3B on the beads incubated with TBK1 than with a buffer control. This shows that

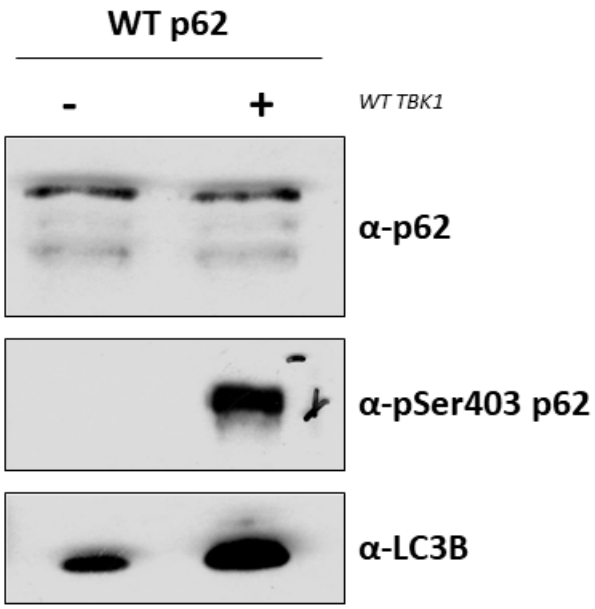



Figure 6.1: Western blot showing the results of a p62-LC3B kinase binding assay. This shows that after a TBK1 incubation step more LC3B remains bound to p62 than in the absence of TBK1. The p62 loading control shows that equal amounts of p62 were initally bound to beads and the pS403 antibody indicates that S403 is only phosphorylated in the presence of TBK1.

incubation of p62 with TBK1 leads to a stronger binding affinity for the interaction of p62 with LC3B.

The amount of p62 detected bound to the beads was the same for the buffer control and the TBK1+ beads as indicated by similar band intensity on the western blot. The phosphorylation of S403 was also used as a control in this experiment to confirm kinase activity. An antibody for the S342 phosphorylation would clearly have been preferred but this was not available. This S403 phosphorylation was only detected in the presence of TBK1 and not for the buffer control.

## 6.3 Mass Spectrometry

Native-like ESI MS was used to investigate the role of phosphorylation on the interaction between the hAtg8 proteins and the p62 AIM. Competition experiments were used to study the binding preference of the hAtg8 proteins for either the WT or pS342 p62 AIM peptides. However, an issue arose in these experiments when studying equimolar mixtures of the WT and pS342 peptides. The previous set of ESI

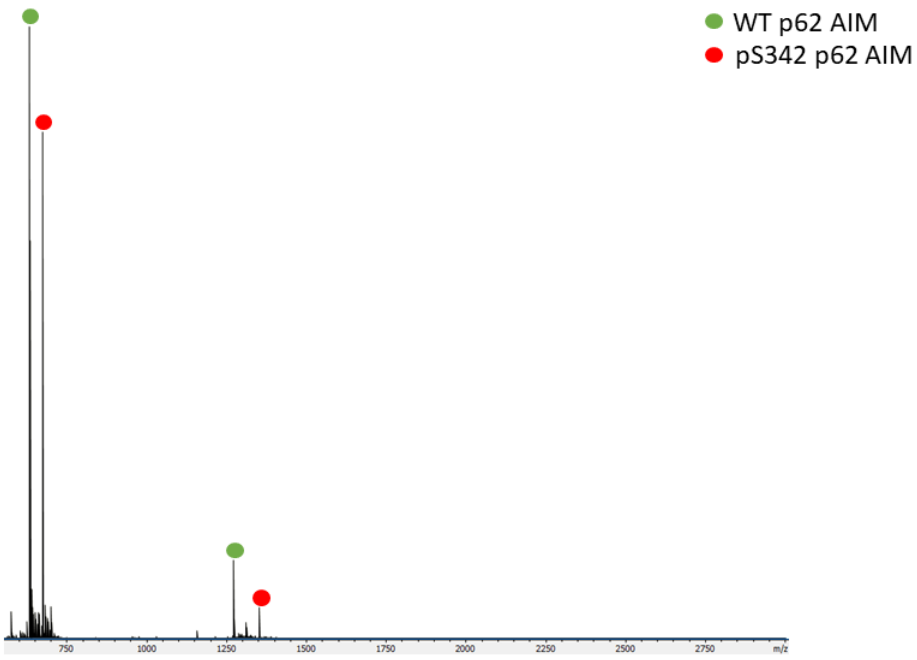



Figure 6.2: ESI MS spectrum of an equimolar mixture of WT and pS342 p62 AIM peptides in positive ion mode.

MS experiments in Chapter 4 were carried out in positive ion mode. Using the same parameters for the experiments required in this chapter however led to significant differences in peak intensity between WT and pS342 p62 AIM peptides when an equimolar sample of the two peptides was injected, as shown in Figure 6.2.

To circumnavigate this issue, the negative ion mode of the spectrometer was utilised. Optimisation of conditions in this mode allowed for relatively similar peak intensities to be observed for the peptides in the equimolar mixture. Unfortunately, this led to a reduction in ionisation efficiency of the hAtg8 proteins so significantly lower peak intensities were observed for these peaks than for the peaks from the injection of an equivalent protein sample in positive ion mode.

Study of the effect of p62 AIM phosphorylation on the individual hAtg8 proteins was carried out by titrating the protein into an equimolar mixture of WT and pS342 p62 AIM peptides. This allowed the selectivity of a given hAtg8 protein for WT or phosphorylated peptide to be determined by comparison of the ratio of free WT to free pS342 peptide or by the ratio of hAtg8 bound to WT compared to pS342 peptide. In this case the quantification of binding preference focused on the ratio of free peptides.

Figure 6.3 shows two spectra from this type of competition experiment for GABARAP,

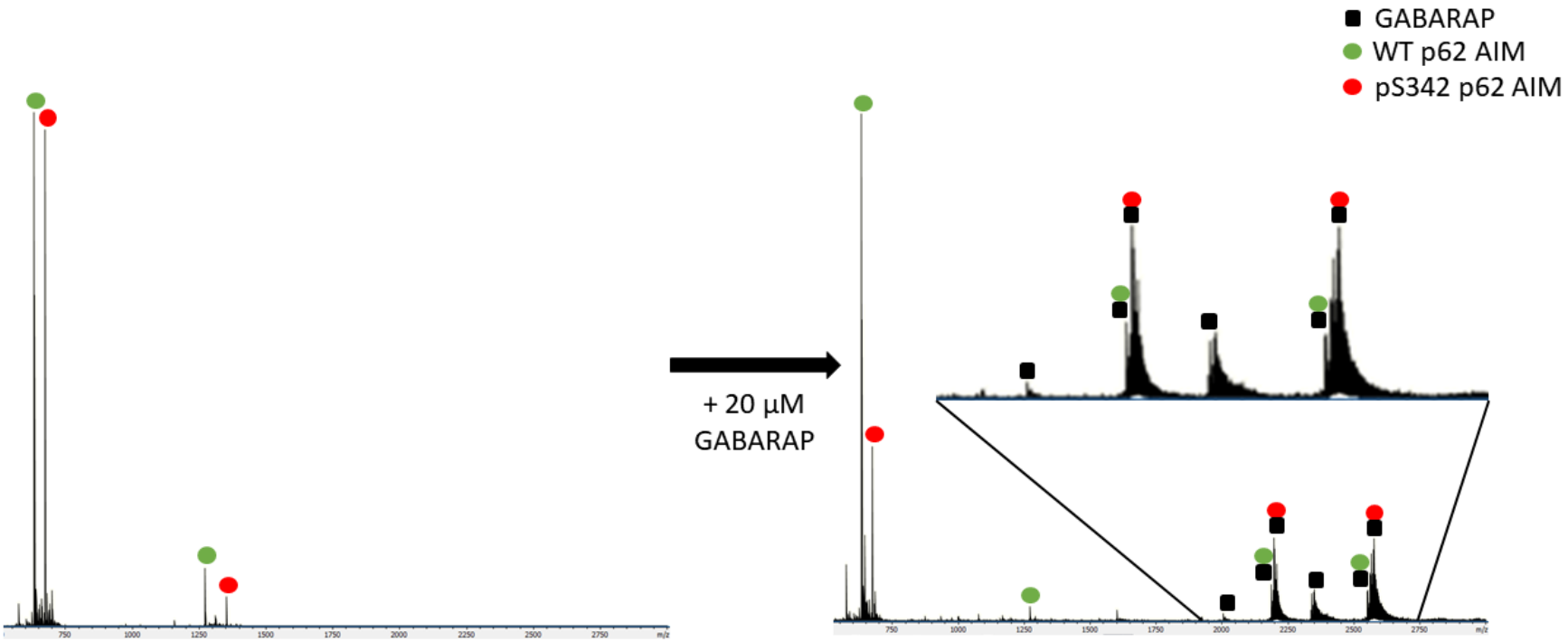



Figure 6.3: ESI MS (negative ion mode) spectra showing the competition experiment for the WT or pS342 p62 AIM peptide binding to GABARAP. Left-hand spectrum: Sample containing the WT and pS342 p62 AIM peptides (10 $\mu$M) which shows the peptides in their two observed charge states; right-hand spectrum: sample containing the WT and pS342 p62 AIM peptides (10 $\mu$M) and GABARAP (20 $\mu$M) which shows the change in relative intensity of the two peptides. Peaks are labelled by the peptide they correspond to, as shown in the legend.

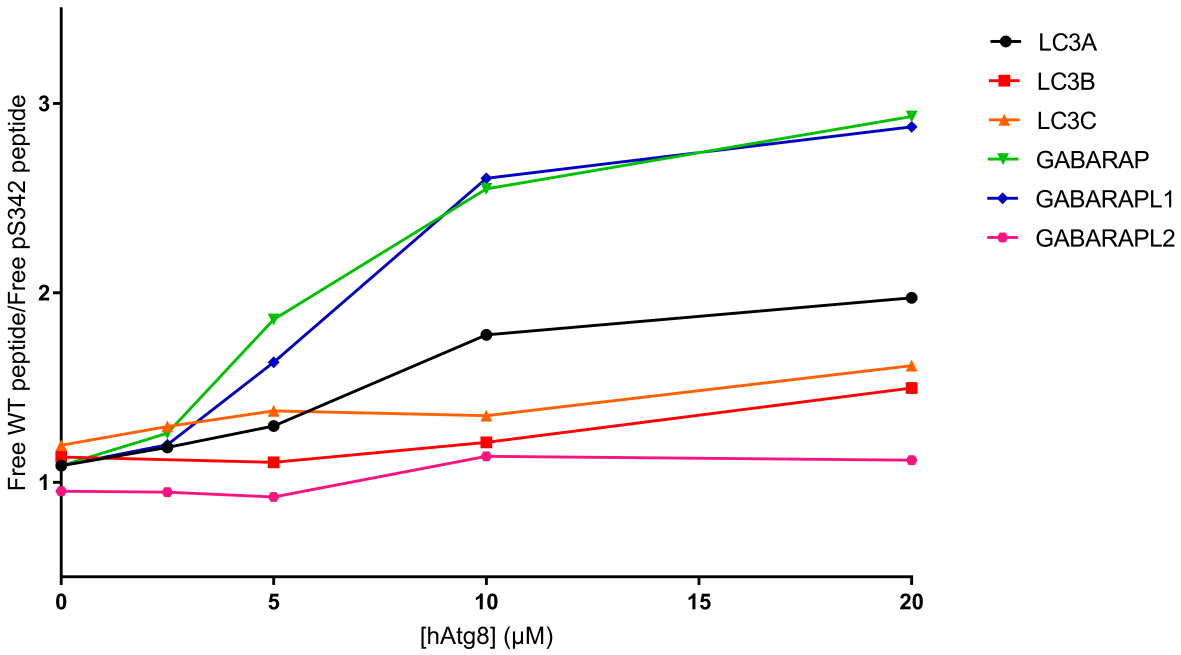



Figure 6.4: Graph showing data from ESI MS competition experiments for each of the six hAtg8 proteins titrated into an equimolar mixture of the WT and pS342 p62 AIM peptides. This shows the ratio of the intensities of the unbound peptides at each hAtg8 concentration and illustrate varying trends in binding preference.

as an example. The left-hand spectrum shows the peptides only (10 $\mu$M of each) with the right-hand spectrum showing the same concentration of peptides after the addition of GABARAP (20 $\mu$M). It can be seen that there was a threefold reduction in the intensity of the pS342 peptide peaks relative to the WT peaks; the 1+ charge state of the phosphorylated peptide at 1355.24 m/z was no longer observable for example. This indicated the preference of GABARAP to bind to the phosphorylated peptide. Peaks corresponding to a mixture of free and bound GABARAP were also observed with peaks corresponding to GABARAP bound to the pS342 peptide occuring at a threefold higher intensity than those for GABARAP bound to the WT.

The ratios of unbound WT to unbound pS342 peptide were plotted at each point throughout the full titration for each of the hAtg8 proteins in Figure 6.4. These lines demonstrated a general trend whereby increasing the hAtg8 concentration led to a higher ratio of free WT peptide to free pS342 peptide. The highest ratio of free WT peptide to free pS342 peptide observed was 2.93 which occured after the addition of 20 $\mu$M GABARAP. This implied that in general the hAtg8 proteins preferentially bound to the pS342 p62 AIM peptide over the WT, though with varying degrees of

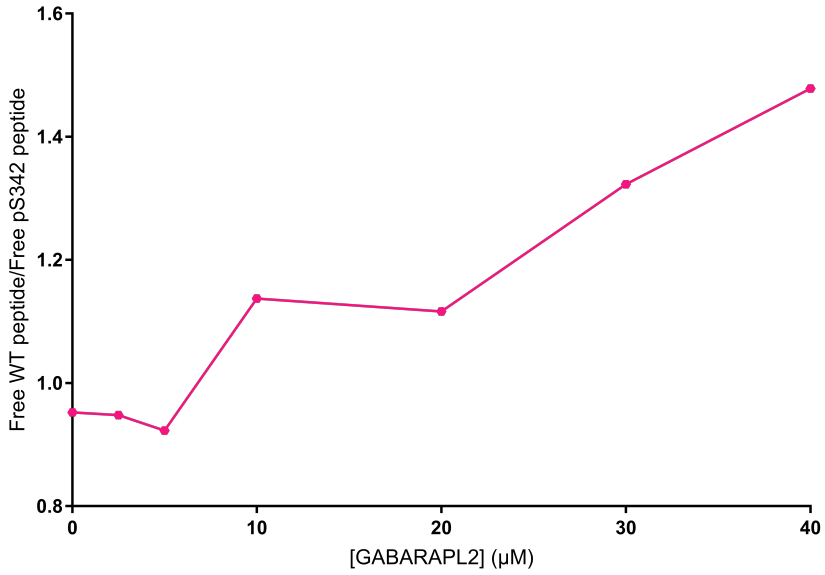



Figure 6.5: Graph showing data from ESI MS competition experiments for GABARAPL2 titrated into an equimolar mixture of the WT and pS342 p62 AIM peptides. This shows the ratio of the intensities of the unbound peptides at each hAtg8 concentration and illustrates a binding preference for the phosphorylated peptide.

selectivity. Although the gradient for these lines was small for LC3B and LC3C, the ratio changed from 1.13 to 1.50 for LC3B and 1.20 to 1.62 for LC3C throughout the titration which indicated a preference. There was no significant trend apparent from the GABARAPL2 experiment so the range of tested concentrations was extended up to 40 $\mu$M. This data is shown in Figure 6.5 and illustrates that GABARAPL2 preferentially interacts with the phosphorylated peptide at higher concentrations.

The use of negative ion mode for these studies prevented binding selectivity competition experiments, as conducted for the WT p62 AIM peptide in Chapter 4, from being carried out. These experiments were not reliable as the peak intensities of the hAtg8 protein and hAtg8-peptide complex peaks were too low. The low intensity resulted in low signal/noise which prevented the peaks from being reliably quantified.

## 6.4 Isothermal Titration Calorimetry

ITC was used to investigate the binding interaction of the pS342 p62 AIM peptide with the hAtg8 proteins. An example of the raw applied power data and the derived binding curve for the interaction of GABARAP with the peptide are shown in Figure 6.6. This data is shown here to illustrate the range of the quality of data in this thesis

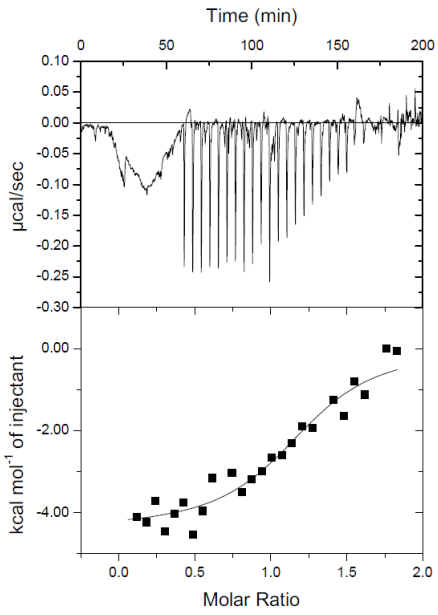



Figure 6.6: ITC data for the titration of the pS342 p62 AIM peptide into GABARAP. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

that allows acceptable integrations of peak area for data analysis. The calculated thermodynamic parameters from these experiments are shown in Table 6.1 and the remaining raw data and binding curves used to determine these values are shown in Appendix Figures A.19-A.23.

The thermodynamic characteristics of these hAtg8-pS342 p62 AIM interactions can then be compared, as shown in Figure 6.7. This shows the relative contributions of the enthalpic ($\Delta H$) and entropic (-T$\Delta S$) components to the overall binding ($\Delta G$) of these proteins. LC3A and LC3B both exhibited the largest enthalpic contributions to binding as well as being the only two proteins which were shown to have an unfavourable entropic component. However, GABARAP and particularly GABARAPL1 have large enthalpies compared to the remaining two proteins; LC3C and GABARAPL2. LC3C and GABARAPL2 are also the only two proteins which presented entropically driven binding interactions with the peptide.

## 6.5 Nuclear Magnetic Resonance Spectroscopy

To further study the role of phosphorylation in the regulation of the hAtg8-p62 interaction, an NMR chemical shift mapping experiment was carried out. Due to

| Protein | n | $K_D$ ($\mu$M) | $\Delta H$ (kJ mol$^{-1}$) |
|---|---|---|---|
| **LC3A** | 0.84±0.01 | 2.11±0.11 | -37.57±0.44 |
| **LC3B** | 0.97±0.01 | 3.08±0.19 | -36.94±0.47 |
| **LC3C** | 0.71±0.10 | 6.33±1.67 | -10.69±2.01 |
| **GABARAP** | 1.21±0.04 | 1.98±0.49 | -18.36±0.85 |
| **GABARAPL1** | 0.96±0.02 | 1.93±0.33 | -22.46±0.52 |
| **GABARAPL2** | 0.77±0.01 | 3.82±0.32 | -11.55±0.28 |

| Protein | $\Delta S$ kJ mol$^{-1}$K$^{-1}$) | $\Delta G$ (kJ mol$^{-1}$) | Affinity increase compared to WT |
|---|---|---|---|
| **LC3A** | -1.52 | -32.93 | 2.15 |
| **LC3B** | -1.62 | -31.99 | 1.25 |
| **LC3C** | 6.38 | -30.22 | 2.66 |
| **GABARAP** | 4.60 | -32.43 | 2.63 |
| **GABARAPL1** | 3.15 | -32.10 | 1.67 |
| **GABARAPL2** | 6.50 | -31.46 | 1.79 |

Table 6.1: ITC-derived thermodynamic parameters for the binding of the pS342 p62 AIM peptide to each of the hAtg8 proteins.

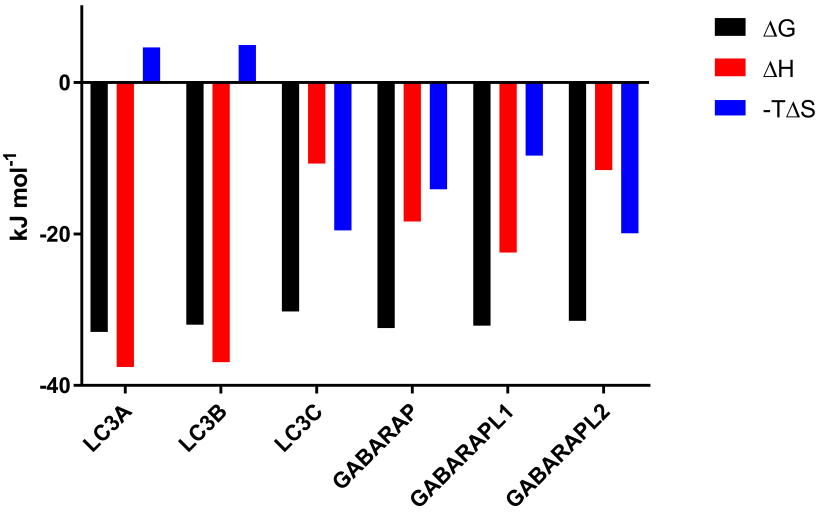



Figure 6.7: Graphs comparing ITC-derived thermodynamic parameters for the binding of the WT (top) and pS342 (bottom) p62 AIM peptide to each of the hAtg8 proteins to illustrate the differences in the enthalpic and entropic contributions to the Gibbs free energy.

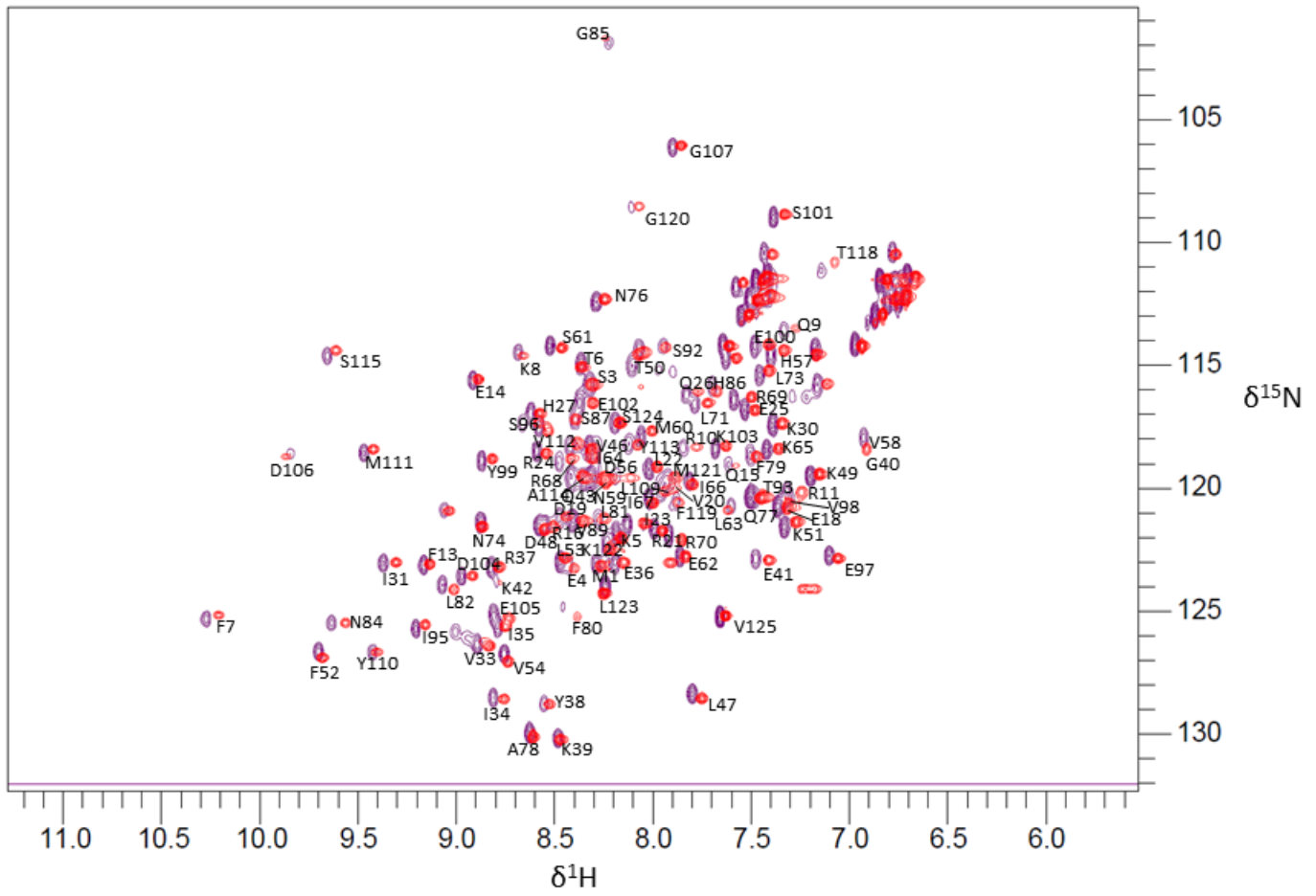



Figure 6.8: Spectral overlay of $^{1}$H-$^{15}$N-HSQC spectra from previously published assignment of LC3B (blue) and the $^{1}$H-$^{15}$N-HSQC collected for this work (red). This illustrates the small shifts in peaks between the two spectra but that the assignment could reliably be copied.

previous NMR work published on LC3B binding to the p62 AIM, the LC3B-p62 AIM interaction was selected for study as this allowed comparison between the WT and phosphorylated peptides. A $^{1}$H-$^{15}$N-HSQC spectrum of isotopically-labelled $^{15}$N-LC3B was first collected. The previous $^{1}$H-$^{15}$N-HSQC backbone assignment of LC3B was then transferred to the spectrum collected here [1].

The previous assignment and the assignment of the $^{1}$H-$^{15}$N-HSQC collected in this work are shown in Figure 6.8. It can be seen that there are some differences in chemical shifts between peaks in the spectra but that the transfer of the assignment is clear in most cases. In those cases where propogation of the assignment from the previous data to the spectrum collected here was ambiguous, the peak was left unassigned. This occured for seven peaks (5.6%). In the published assignment there was only one unassigned non-prolyl residue but the assignment coverage which was unambigously copied here is shown in Table 6.2. The RMSD was determined between the signals in the published $^{1}$H-$^{15}$N-HSQC and the corresponding signals they were assigned to in the $^{1}$H-$^{15}$N-HSQC produced here. This gave an RMSD of 0.04 ppm in the $^{1}$H dimension and 0.10 ppm in the $^{15}$N dimension.

The titration of the pS342 p62 AIM peptide into a $^{15}$N-labeled sample of LC3B was conducted incrementally at constant LC3B concentration until a 2:1 peptide to protein ratio was reached. A $^{1}$H-$^{15}$N-HSQC spectrum was collected at each titration point. The 2:1 sample was used as a final titration point as the binding was saturated, as indicated by the lack of further movement of peaks from the previous titration point (1.2:1). The unbound and bound $^{1}$H-$^{15}$N-HSQC spectra of LC3B are shown in Figure 6.9 which illustrates the chemical shift perturbations (CSPs) of the NH signals upon binding. The full spectral overlay of each titration point produced throughout the experiment is shown in Appendix Figure A.13.

During this titration the peaks were observed to exhibit slow exchange behaviour which prevented the following of peaks through the titration to assign the bound spectrum as for fast exchanging peaks. The previously published chemical shift mapping study by Goode *et al.* produced two assignments of the $^{1}$H-$^{15}$N-HSQC spectra of LC3B when bound to either a WT or L341V p62 AIM peptide. All three spectra of LC3B bound to the differing peptides are similar though there are differences for some peaks. Two clear examples of the differences observed are shown for the peaks corresponding to 52Phe and 51Lys in Figure 6.10 which shows the perturbation of the unbound peak by each of the three peptides. The perturbation caused by the pS342 peptide is more similar to the L341V peptide than the WT. By comparison of the bound LC3B-pS342 p62 AIM $^{1}$H-$^{15}$N-HSQC spectrum to the spectra of LC3B bound

| Amino Acid Type | Total | Assigned | % Assigned |
|---|---|---|---|
| **Ala** | 3 | 3 | 100 |
| **Arg** | 9 | 9 | 100 |
| **Asn** | 4 | 4 | 100 |
| **Asp** | 5 | 4 | 80 |
| **Cys** | 0 | 0 | - |
| **Gln** | 7 | 6 | 85.7 |
| **Glu** | 12 | 12 | 100 |
| **Gly** | 4 | 3 | 75 |
| **His** | 3 | 3 | 100 |
| **Ile** | 8 | 8 | 100 |
| **Leu** | 11 | 11 | 100 |
| **Lys** | 10 | 10 | 100 |
| **Met** | 5 | 4 | 80 |
| **Phe** | 7 | 6 | 85.7 |
| **Pro** | 6 | 0 | 0 |
| **Ser** | 9 | 8 | 88.9 |
| **Thr** | 6 | 5 | 83.3 |
| **Trp** | 0 | 0 | - |
| **Tyr** | 4 | 4 | 100 |
| **Val** | 12 | 11 | 91.7 |
| **All Residues** | 125 | 111 | 94.4 |
| **All Residues - Pro** | 119 | 111 | 93.3 |

Table 6.2: Overview of the LC3B amide backbone assignment by amino acid type indicating that 93.3% of assignable residues have been copied from the previous assignment.

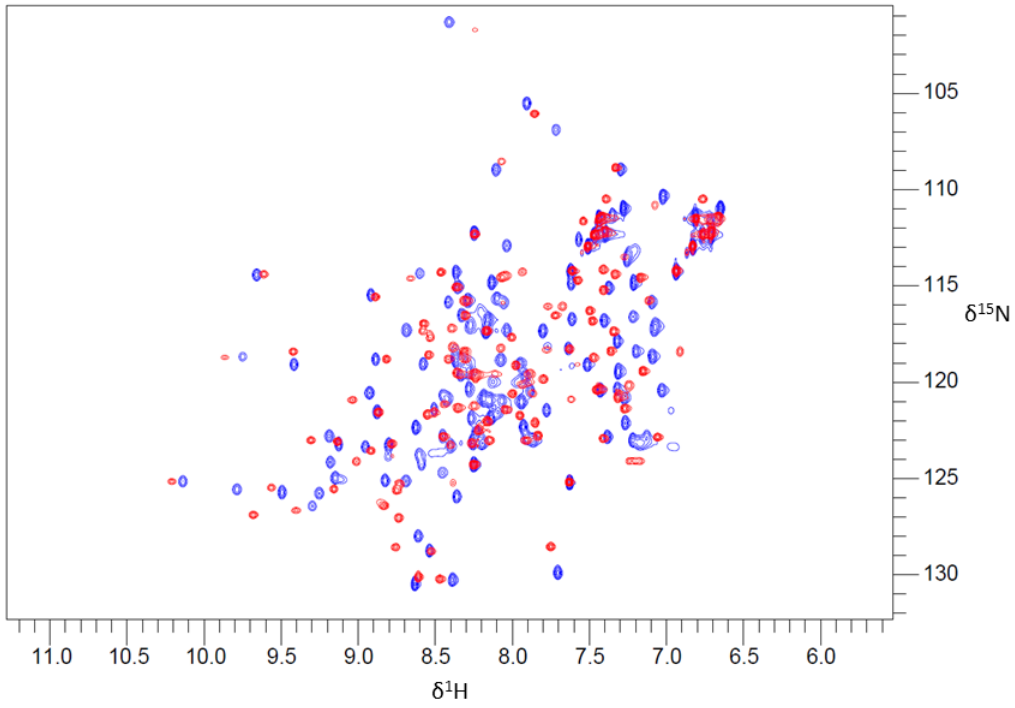



Figure 6.9: Spectral overlay of the $^{1}$H-$^{15}$N-HSQC spectra of unbound LC3B and pS342 p62 AIM-LC3B (2:1).

to the WT or the L341V peptide it can be seen that the pS342-bound spectrum and the L341V-bound spectrum are most similar. This is supported by calculated RMSD values. The RMSD values determined between the L341V-bound signals and the pS342-bound signals were 0.05 ppm in the $^{1}$H dimension and 0.13 ppm in the $^{15}$N dimension. The equivalent values betwen the WT-bound LC3B signals and the pS342-bound signals were calculated as 0.16 ppm in the $^{1}$H dimension and 0.98 ppm in the $^{15}$N dimension. As such the assignment of LC3B bound to the L341V peptide was copied to the pS342-bound $^{1}$H-$^{15}$N-HSQC spectrum produced in this work. Any ambigious peak correlations were left unassigned. The assigned LC3B-pS342 p62 AIM $^{1}$H-$^{15}$N-HSQC spectrum is shown in Figure 6.10 along with the spectrum of LC3B-L341V p62 AIM to illustrate the copying of assignment from this spectrum.

One exception to this similarity between the spectra of LC3B bound to the pS342 or L341V peptides exists for Leu47. Figure 6.11 shows the varying CSPs of the peak corresponding to 47Leu which illustrates a similar shift for the WT and L341V but a larger and differing shift upon binding to the pS342 peptide. This indicated that the backbone amide of this residue experienced a differential change in chemical environment when bound to the phosphorylated peptide. As this residue is now not analogous to either of the bound spectra it had to be assigned using a closest

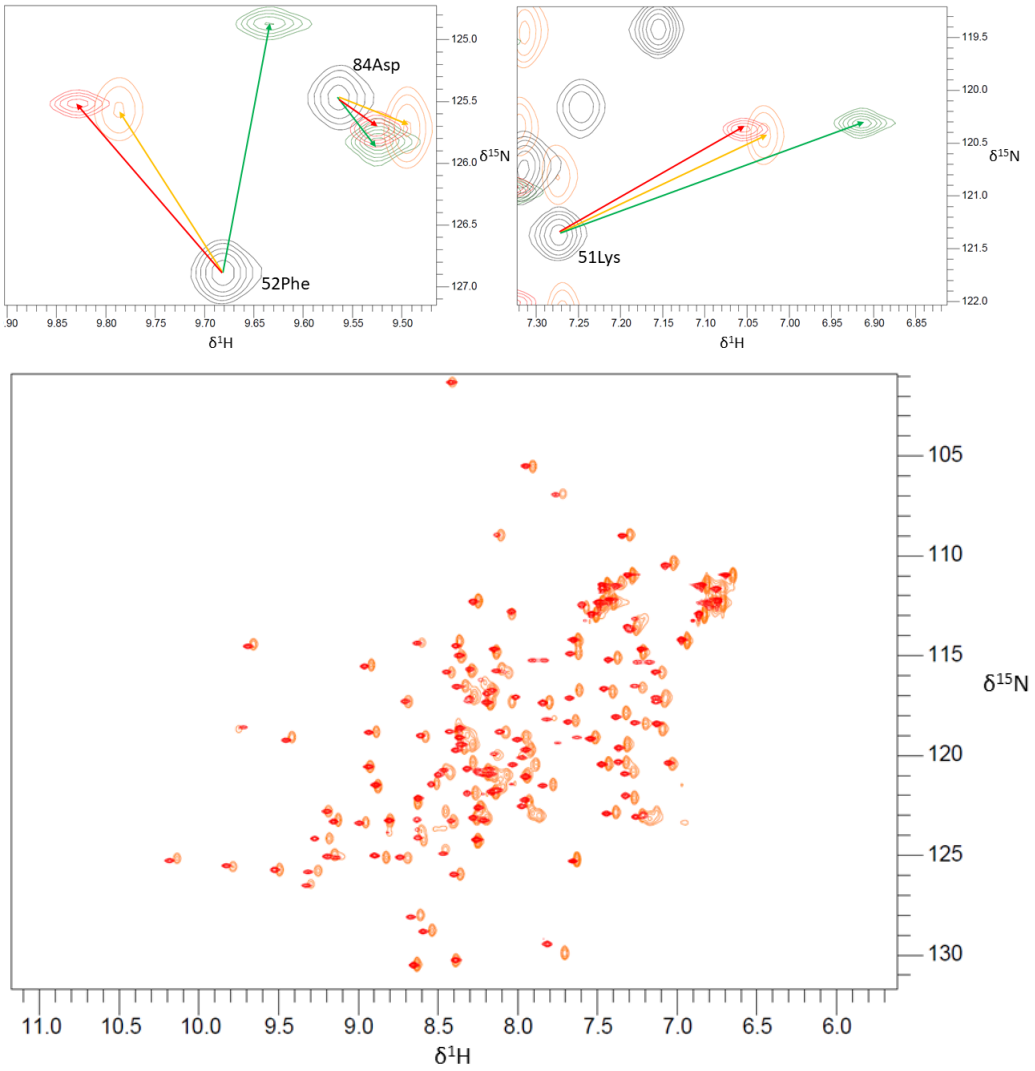



Figure 6.10: $^{1}$H-$^{15}$N HSQC spectra of LC3B overlaid to show the differences in spectra depending on which p62 AIM peptide is bound. This illustrates that the spectrum of LC3B bound to the pS342 p62 AIM peptide is most similar to the spectrum of LC3B bound to the L341V peptide. (Top) Selected regions of the overlaid $^{1}$H-$^{15}$N-HSQC spectra of LC3B showing the protein only (black), the protein bound to the WT peptide (green), the protein bound to the pS342 peptide (orange) or the protein bound to the L341V peptide (red). (Bottom) Spectral overlay of the $^{1}$H-$^{15}$N-HSQC spectra of LC3B bound to the pS342 AIM peptide (orange, produced here) or the L341V peptide (red, previously published).

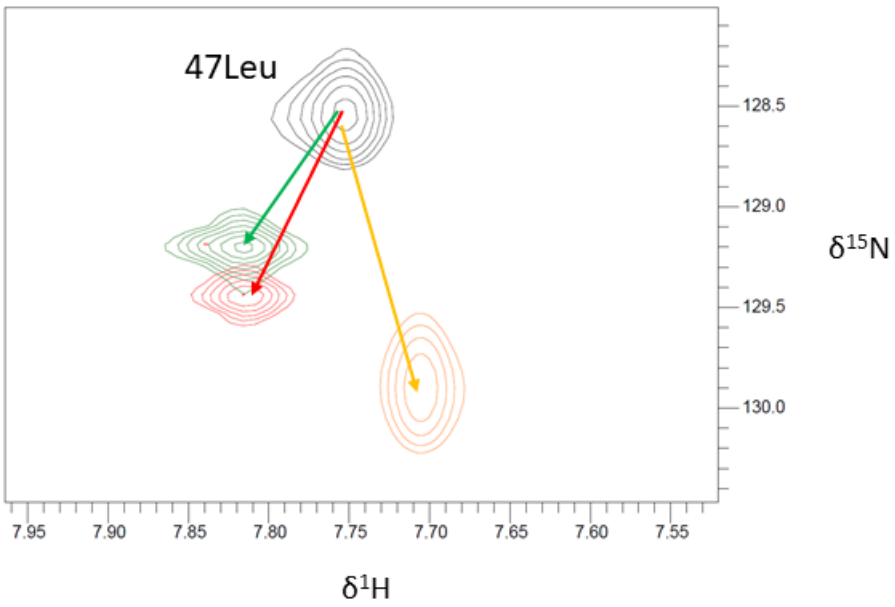



Figure 6.11: Spectral overlay of the $^{1}$H-$^{15}$N-HSQC spectra of LC3B when unbound (black), bound to the WT p62 peptide (green), bound to the pS342 (orange) or bound to the L341V peptide (red).

unassigned procedure whereby the peak must have shifted at least as far as the closest unassigned peak. This CSP is used for this peak as it is considered to be the minimum possible value.

Using the signals which were unambiguously assigned in both the unbound and bound spectra, CSP analysis was performed and the results are shown in Figure 6.12. This graph illustrates the residues which underwent significant CSPs upon binding to the phosphorylated peptide according to two levels of significance. There were 11 backbone NH signals with CSPs with the highest level of significance and 14 NH signals which underwent shifts with some significance. The residues at each of the two levels of significance were then mapped onto the crystal structure of LC3B which is illustrated in Figure 6.13.

This binding study defined a similar binding interface to that observed previously for the WT and L341V peptide however as was previously described the pS342 AIM-bound LC3B spectrum more closely resembled the L341V AIM-bound LC3B spectrum. This was demonstrated further by calculation of $\Delta$CSP values which show the difference in CSP for each residue caused by the different peptides. The average $\Delta$CSP between the WT and phosphorylated peptide was 0.04 and the average value between the phosphorylated peptide and the L341V peptide was 0.02. Significance of $\Delta$CSP values was then shown using the two levels of significance; as for the CSP

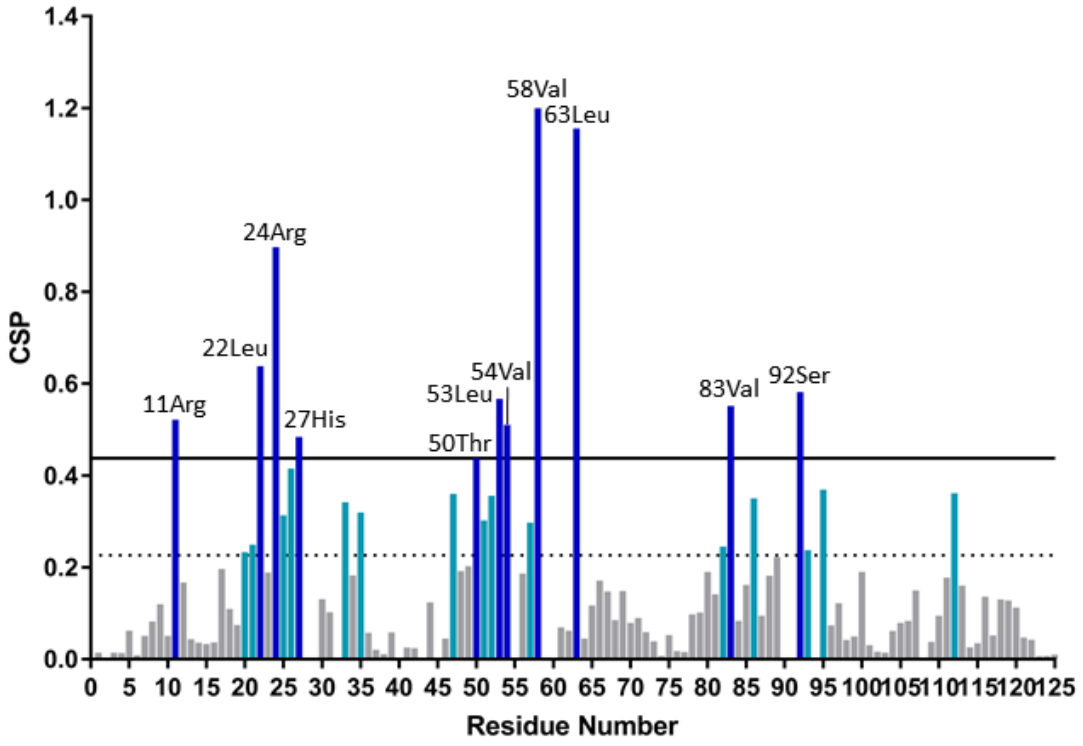



Figure 6.12: Graph illustrating the CSP values for the binding of $^{15}$N-LC3B to the pS342 p62 AIM peptide. The solid line denotes the limit for two standard devations above the mean and the dotted line denotes the limit for one standard devation above the mean.

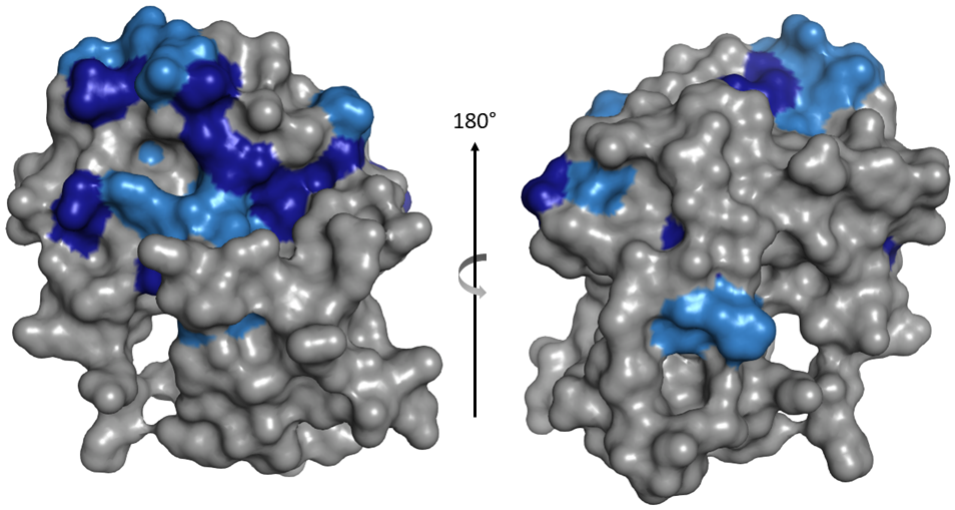



Figure 6.13: Crystal structure of LC3B which highlights the residues with backbone NH signals which were significantly perturbed in the $^{1}$H-$^{15}$N-HSQC spectrum upon addition of the pS342 p62 AIM peptide. This shows the residues with significantly shifted signals which could be followed through the titration (dark blue is most significant; light blue is some significance).

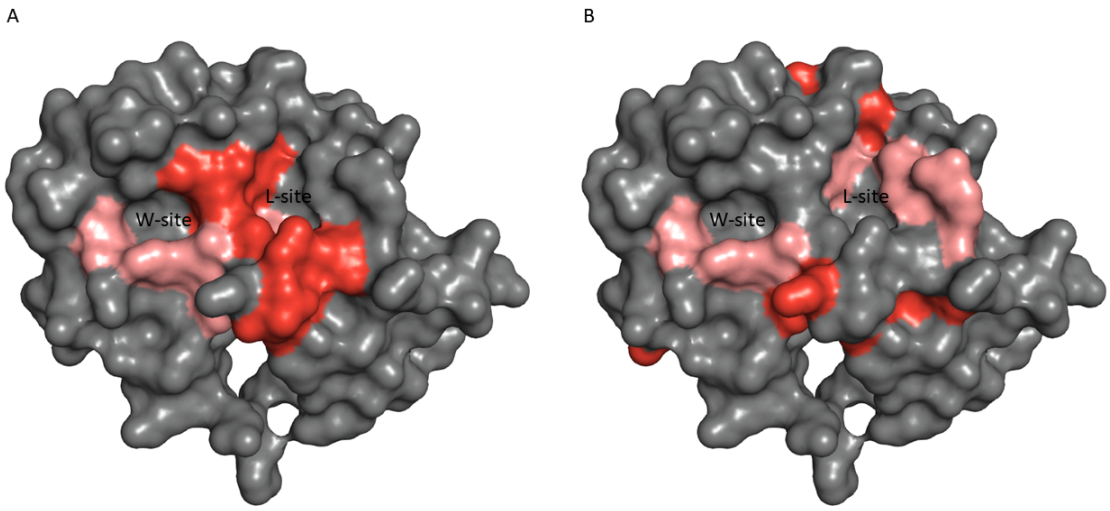



Figure 6.14: Crystal structure of LC3B highlighted according to the degree of significance of the ΔCSP value for each residue by comparing (A) the WT and pS342 peptides and (B) the pS342 and L341V peptides. The residues with the most significant values are coloured red and the residues with some significance are coloured pink.

values themselves. The significant ΔCSP values for the comparison between the phosphorylated peptide and the WT or L341V peptides have been mapped onto the LC3B structure in Figure 6.14. This illustrates the clear differences in LC3B CSP between binding to the WT and pS342 peptides which occur around the core AIM binding region, particularly around the L-site. The significant ΔCSP values between binding to the pS342 and L341V peptides are less pronounced around this binding site.

In addition to this chemical shift mapping experiment it was also possible to utilise $^{31}$P NMR experiments to investigate the pS342 p62 AIM-hAtg8 interaction. The 1D $^{31}$P experiment of the pS342 p62 AIM peptide is shown in Figure 6.15A. There are two peaks present in this spectrum which can be assigned to the phosphorus atom in the phosphate ion from the buffer ($PO_4^{3-}$) at 1.364 ppm and the phosphorus atom in the phosphate group on the peptide at 3.452 ppm. Figure 6.15B shows the region of the spectrum containing the signal from the phosphorus on the peptide. This figure also shows the same signal upon the addition of LC3B or GABARAP to the sample which shows a smaller CSP upon binding to LC3B (0.105 ppm) and a larger CSP upon binding to GABARAP (0.434 ppm). These CSPs indicate the phosphorus on the peptide undergoes a small change in chemical environment upon binding to LC3B and a larger shift in chemical environment upon binding to GABARAP.

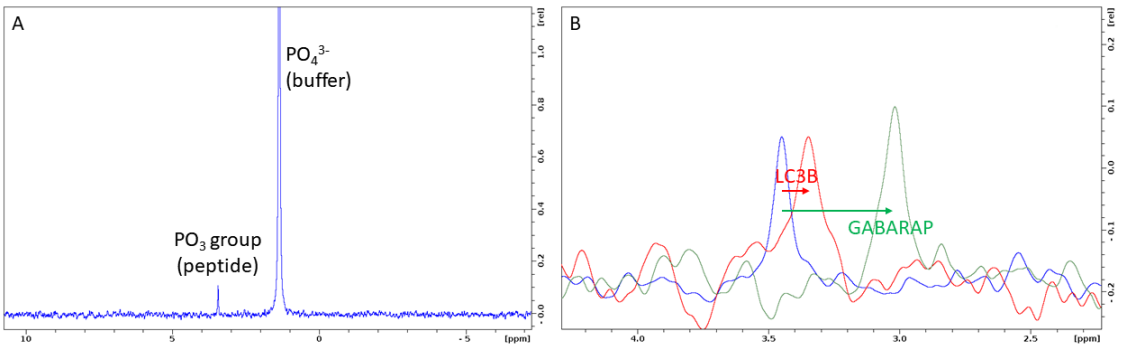



Figure 6.15: $^{31}$P NMR data from the study of the interaction of the pS342 p62 AIM peptide with LC3B and GABARAP. (A) NMR spectrum of the pS342 p62 AIM peptide (B) Selected region of the $^{31}$P NMR spectrum of the pS342 p62 AIM peptide on its own (blue), in the presence of LC3B (red) and in the presence of GABARAP (green).

## 6.6 Discussion

### Kinase Binding Assay

This assay demonstrated an increase in binding affinity for full length p62 binding to LC3B upon incubation with TBK1. This indicates that phosphorylation is modulating the interaction between the p62 AIM and LC3B. As a control, it was shown that incubation of the GST-p62 fusion protein bound to Glutathione Sepharose beads with TBK1 led to phosphorylation at S403 which does not occur in the absence of TBK1. It is assumed therefore that phosphorylation also occurs at the other known TBK1 phosphorylation sites on p62 [46, 68, 74]. However, this does not confirm the role of specific phosphorylation sites in this increased affinity which is the reason the specific phosphorylation site at S342 was studied further in this chapter.

### Biophysical Study of the pS342 p62 AIM-hAtg8 Interaction

Negative ion mode was utilised in the ESI MS competition experiments in this chapter as a result of the difference in peak intensity observed between the WT and pS342 p62 AIM peptides from an equimolar mixture in positive ion mode. The difference in peak intensity between the peptides was assumed to have been caused by the negative charge presented by the phosphate group. Both peptides have three Asp residues which were expected to be negatively charged at pH seven but the phosphate group added additional negative charge to the pS342 peptide. The larger negative charge of this peptide may have reduced the generation of overall positive ions of this

peptide during ionisation which resulted in a 21% lower peak intensity.

As predicted, negative ion mode produced peaks with larger and more similar intensities for the two peptides. Surprisingly, the phosphorylated peptide remained at a lower intensity relative to the WT suggesting some other unknown factor was limiting the ionisation of this peptide. Despite this small discrepency, corresponding to a 8% lower intensity of the pS342 peptide on average, these conditions were considered acceptable for further experiments.

Quantitative data analysis for these experiments was then considered. Despite the desalting procedure for these proteins some metal ions remained in the sample, possibly due to binding to the protein. Accordingly, some ion adducts to the protein peaks were observed, as illustrated in Figure 6.3. Quantitative analysis of these peaks was considered less reliable as some of the ion adducts of the GABARAP-WT p62 AIM peaks contribute to the intensity of the GABARAP-pS342 p62 AIM peaks due to peak overlap and therefore the relative contributions to the peak intensity cannot be deconvoluted. In addition to this, as negative ion mode was required for these experiments the signal to noise ratio of the hAtg8 proteins was much lower which hindered analysis, particularly at low hAtg8 concentrations. So whilst the protein peaks qualitatively implied a preference for the phosphorylated peptide, analysis of the unbound peptide peaks allowed a clearer quantitative analysis of this preference.

Overall, the MS data indicated that the hAtg8 proteins bound preferentially to the phosphorylated p62 AIM peptide over the WT; to some degree. The calculated ratio of free WT to free phosphorylated peptide increased as all six hAtg8 proteins were titrated into the sample.

To allow quantification of the observed binding preference, ITC was utilised. The 'n' parameter derived from this fit was expected to be one due to the 1:1 binding observed in the MS studies. There is some deviation from this value for the six interactions investigated but the fits were generally deemed to be representative of the data due to the low $\chi^2$/DoF. By averaging of the binding preferences, it was determined that the six hAtg8 proteins bound with a two-fold preference to the phosphorylated peptide over the WT.

Comparison of the MS ESI and ITC data illustrated an interesting constraint within the MS competition methodology used here. There were some discrepencies observed between the apparent order of binding preference from the MS and ITC experiments. This was a result of the concentration range used in these experiments. The shape of the ESI MS binding curve within a given range of hAtg8 concentrations is influenced by the binding affinity of the protein to both peptides being investigated.

The maximum ratio determined is indicative of the strength of binding preference but the concentration of titrant at which this point occurs will vary between proteins depending on the affinities. Therefore, although LC3C has a strong binding preference for the phosphorylated peptide, this was not observed at the concentrations tested in the ESI experiment. On the other end of the scale, GABARAPL1 has a relatively small preference for the phosphorylated peptide but due to its higher affinities for the peptides, the difference was observed. Therefore, this type of ESI MS competition experiment with a smaller concentration range should only be utilised to confirm the presence of a binding preference and not to quantify it.

## Structural Characterisation using NMR

The $^{1}$H-$^{15}$N-HSQC spectrum of LC3B was used as the basis of the chemical shift mapping experiment as it serves as a fingerprint of the protein which shifts in proportion to perturbations of the chemical environment. The assignment of the spectrum of LC3B from previous work in this lab was readily transfered onto the collected spectrum, as indicated by the low RMSD values between the two sets of peaks. Previously collected 3D experimental data was also reanalysed to confirm the assignment used in this work.

Comparison of the spectra of LC3B bound to each of the three peptides (WT, pS342 or L341V p62 AIM) showed that in most cases there was little difference between the peaks for each. This indicated a similar binding interaction for each of the peptides and allowed for unambigious transfer of the assignment from the previous bound spectra to the bound spectrum collected here. In the small number of cases where there was a difference in the CSP between the different peptides, the pS342 and L341V peptides typically produced more similar CSPs than for the WT. This was supported by lower RMSD values and lower average $\Delta$CSP values when comparing the pS342-bound LC3B spectrum to the L341V-bound spectrum than the WT.

The binding site for the pS342 peptide on the LC3B surface according to CSP analysis is similar to that of the WT peptide, as illustrated in Figure 6.16. Binding of the phosphorylated peptide to LC3B has been modelled by alteration of the WT p62 AIM peptide-LC3B crystal structure to include the phosphate group. This figure also shows the residues with significant CSPs. The area of significant shifts is more extensive upon binding to the phosphorylated peptide than the WT but the core binding surface remains perturbed with a similar pattern of significance. One difference that has been highlighted is the significant shift of His57 upon binding to

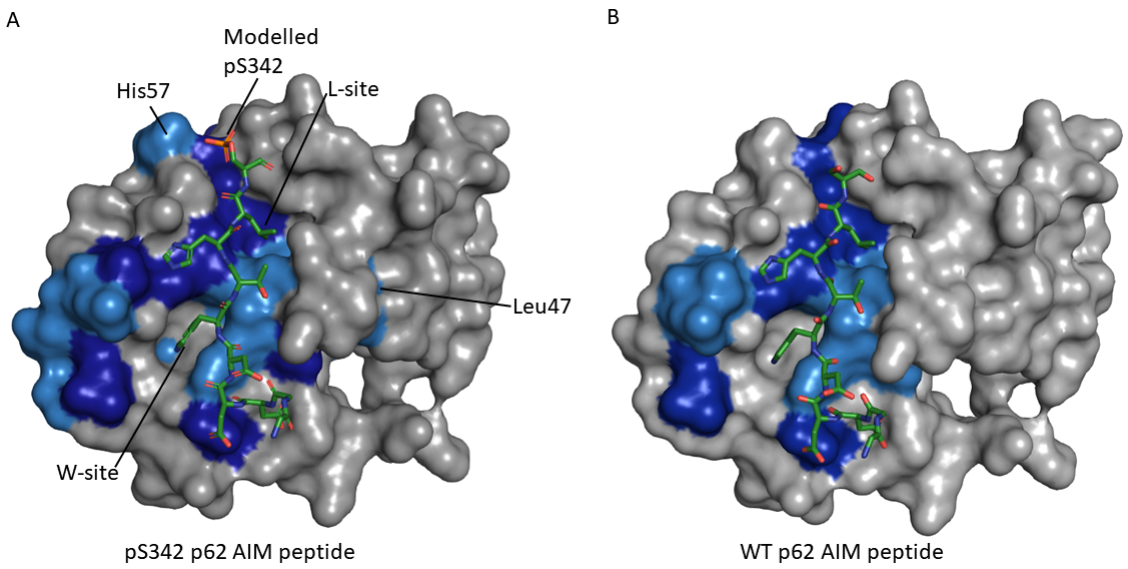



Figure 6.16: Crystal structure of LC3B which highlights the residues with backbone NH signals which were significantly perturbed in the $^{1}$H-$^{15}$N-HSQC spectrum upon addition of the pS342 p62 AIM peptide (left) or the WT peptide (right). The crystal structure of the WT p62 AIM peptide binding to LC3B has been modified to include the phosphate group on the Ser342 sidechain. This shows the residues with significantly shifted signals which could be followed through the titration (dark blue is for the most significant; light blue is for some significance).

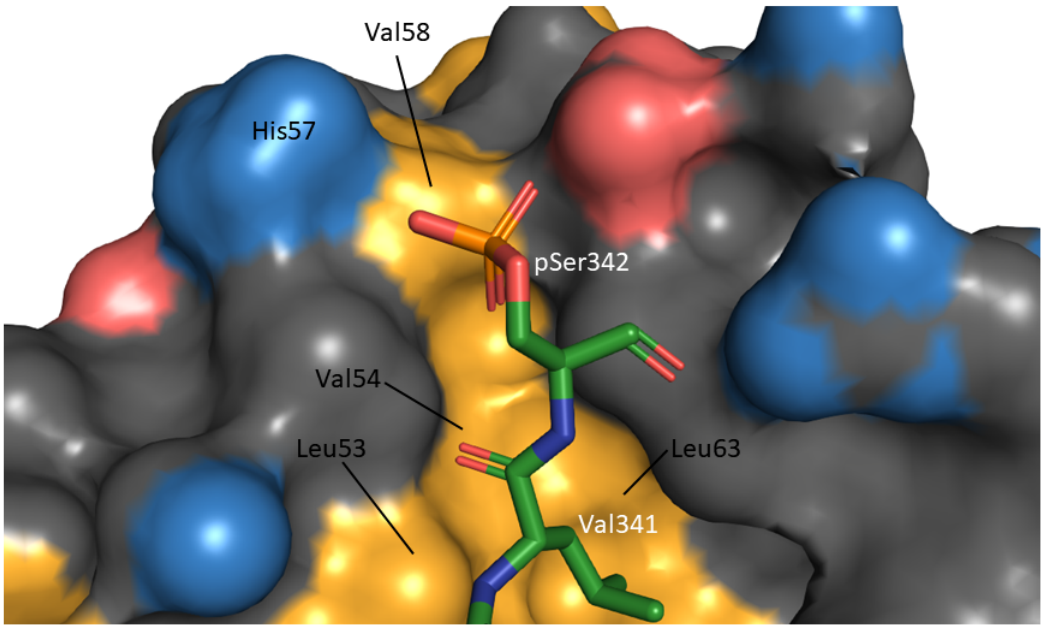



Figure 6.17: Structure of LC3B bound to the pS342 p62 AIM peptide highlighting the binding region around the pS342 residue; annotated residues have shown significant CSP in the NMR titration experiment. This shows the surface of the residues with significantly shifted NMR signals in yellow and any charged atoms on the surface are shown in red (negative) and blue (positive). The peptide structure is modelled on the WT p62 AIM peptide structure modified to include the phosphate group and shown in green other than oxygen atoms (red), nitrogen atoms (blue) and phosporous atoms (orange).

the phosphorylated peptide which did not shift significantly upon binding to the WT peptide. This residue is located on the surface of LC3B in close proximity to p62 S342 based upon the crystal structure of binding to the WT as shown in Figure 6.17. This therefore raises the possibility of an electrostatic interaction between the positively charged side chain of the histidine and the negatively charged phosphate group on the peptide. Leu47 was identified as the one residue which underwent a different CSP upon binding to the phosphorylated peptide than binding to either the WT or L341V peptides. This residue is located outside the expected core AIM binding site and therefore implies a differential structural rearrangement upon binding to the phosphorylated peptide.

The $\Delta$CSP values produced by comparison of LC3B binding to the WT or pS342 p62 AIM peptide indicated that the main differences in these peptides interactions with the protein surface occur most significantly in the region around the L-site. This can be rationalised as the phosphorylated serine is directly adjacent to the leucine

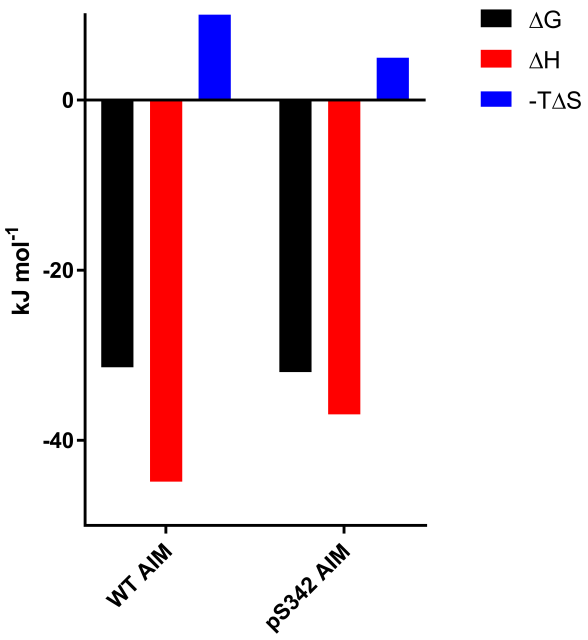



Figure 6.18: Graph comparing ITC-derived thermodynamic parameters for the binding of the WT and pS342 p62 AIM peptides to LC3B to illustrate the differences in the enthalpic and entropic contributions to the Gibbs free energy from phosphorylating the peptide.

that forms this key hydrophobic interaction so clearly this region of the LC3B peptide would be more significantly perturbed. It is likely that the pS342 peptide binds in a slightly different mode than the WT peptide which prevents the Leu341 of the peptide binding tightly into the LC3B L-site, as was the case for the reduced affinity binding of the L341V peptide. The phosphorylation at S342 leads to an alteration in this binding site that produces a stronger affinity compared to the WT peptide whereas the L341V mutation leads to alteration in this binding site that abbrogates affinity.

When considering the effect of phosphorylation of p62 in binding to LC3B we may also consider the change in thermodynamic properties as determined by ITC. These differences are shown in Figure 6.18 and indicate that phosphorylating actually reduces the enthalpic contribution to binding but leads to a less unfavourable entropic component. This taken alongside the NMR data may indicate a binding mode which does not involve strong electrostatic interactions between the phosphate group and LC3B. Instead this phosphate group may sterically guide the binding of these two species in some way. This may be relationalised from the structure of the WT p62 AIM peptide binding to LC3B in which the side chain of this serine is directed away from the LC3B binding site and into free solution. There may be some conformational

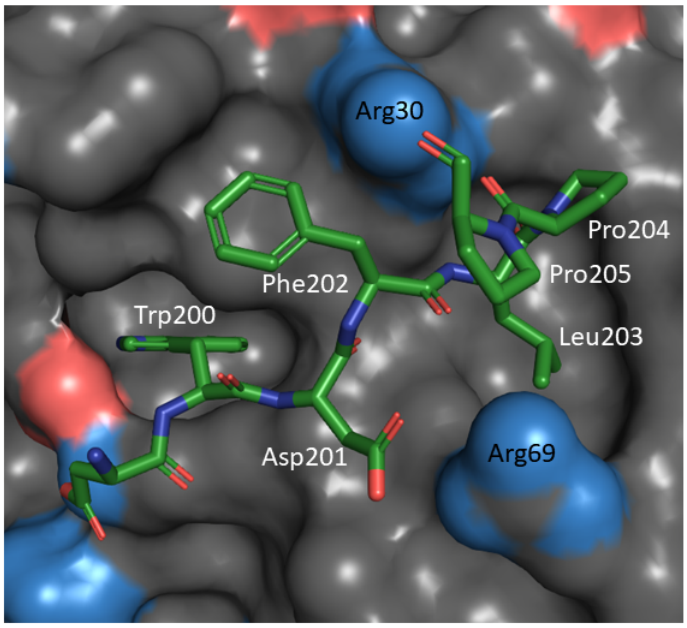



Figure 6.19: Crystal structure of GABARAP bound to a peptide corresponding to the AIM of calreticulin to show the charged residues around the leucine binding cleft (L-site). The GABARAP surface is shown in grey other than charged atoms which are shown in blue (positive) and red (negative). Residue annotations are shown in white for the calreticulin peptide and black for the GABARAP surface.

flexibility though which permits transient interactions between the His57 sidechain and the phosphate group that produced CSPs of their corresponding NMR signals in the $^{1}$H-$^{15}$N-HSQC and the $^{31}$P spectra respectively. This is supported by the more favourable entropy term for the pS342 peptide binding determined by ITC.

The titration of GABARAP here serves as an example that the CSP of the signal upon binding to LC3B is small and that the phosphorylation of the peptide leads to an increase in affinity in the GABARAP subfamily via an interaction with the protein which perturbs this phosphate more significantly. This implies a different binding mode for these proteins which has not been further studied here. There is one previously published structure of GABARAP bound to a calreticulin AIM peptide which is shown in Figure 6.19 [53]. Structural comparisons are limited in their utility here as the calreticulin peptide has a proline in the position analogous to Ser342 of p62 which leads to a peptide backbone angle unlikely in the case of the pS342 p62 AIM peptide. Additionally, whilst calreticulin has a Trp in the same position as for p62, the $x_1$ and $x_2$ residues are Asp-Phe for calreticulin but Thr-His for p62 which will clearly differentiate their binding due to differences in sidechain size and charge. However, around the analogous L-site on the GABARAP surface there are

two areas of charge due to the side chains of Arg30 and Arg69 which may interact with the phosphorylated sidechain of Ser342 when binding to the p62 AIM to result in the more significant perturbation of the phosphate $^{31}P$ signal upon titration of GABARAP than LC3B.

## Biological Context

Taken in totality, the evidence presented implies that phosphorylation at S342 of p62 leads to an increase in p62 recruitment to the autophagosome through the increased hAtg8-p62 AIM affinity. This presumably allows p62 to perform its important roles in cargo recruitment and autophagosome expansion to a greater extent. TBK1 has been shown to phosphorylate at this site and has been shown in this work to increase the p62-LC3B affinity [46]. TBK1 has been previously shown to contribute to pathogen and mitochondrial clearance as well as affecting autophagosome maturation [192, 78]. These are the two key functions played by recruitment of p62 to the autophagosome; as a scaffold for membrane expansion and as a recptor for selective cargo recruitment. TBK1 can therefore be said to increase specific autophagy and autophagosome maturation by phosphorylation at S342 of p62, as well as roles at other sites.

The data presented in this chapter has shown that:

- The binding affinity of full length p62 with LC3B is increased by phosphorylation with TBK1.
- All six hAtg8 proteins preferentially bind to a pS342 p62 AIM peptide over the WT, with varying degrees of preference.
- Binding of the pS342 p62 AIM peptide to LC3B occurs in the same binding site as for the WT peptide however there are significant differences in the structural perturbation caused by binding around the L-site due to the bulk and charge of the phosphate group.

Phosphorylation of p62 has been studied in terms of its effect on hAtg8 binding in this chapter. In the following chapter phosphorylation on the other side of this important protein-protein interaction will be investigated by looking at a phosphorylation site on one of the hAtg8 proteins; LC3B.

# Chapter 7

# Phosphorylation of LC3B for Autophagy Regulation

## 7.1 Introduction

Phosphorylation of LC3B has been shown at a number of sites in high throughput studies. Thr50 was found to be phosphorylated in this way and this site has subsequently been investigated through a range of techniques which illustrated the importance of this site in autophagosome-lysosome fusion. LC3B phosphorylation at Thr50 was studied here using a phosphomimetic T50E mutant. The structures of phosphothreonine and glutamate are shown in Figure 7.2 which shows the related but differing size and charges of these residues. The utility of phosphomimetic mutations can be debated though as there are clear differences between a phosphate group and the side chain carboxylic acid of a glutamate amino acid. The $pK_a$ values of the two groups are different, they have different charges and can form hydrogen bonds in differing orientations and strengths. Therefore, evidence to support the use of a phosphomimetic mutation instead of incorporation of a phosphate must be presented before conclusions can be drawn from the use of a phosphomimetic in biophysical studies.

Knockout (KO) of STK3 and STK4 kinases was shown to result in a loss of function in cells; namely a block in autophagosome-lysosome fusion [3]. These kinases were shown to phosphorylate LC3B at T50. Expression of LC3B T50E was shown to recover activity lost by this knockdown which suggests that in this case the phosphomimetic is an appropriate approximation to the phosphothreonine. This is illustrated in Figure 7.1 taken from that work which shows that STK3/4 KO leads to an increased number

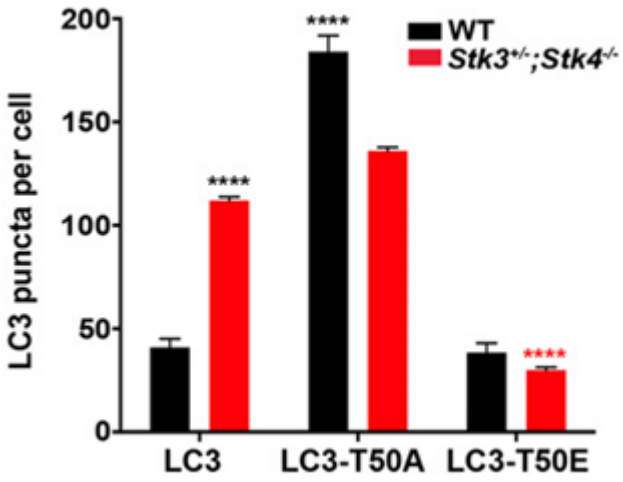



Figure 7.1: Graph published by Wilkinson et al. illustrating autophagy impairment due to loss of the STK3/4 kinases or from the LC3B T50A mutation and the restoration of autophagy by the LC3B T50E mutation. LC3 puncta were quantified using confocal fluorescence microscopy as the protein was tagged with GFP. Increased numbers of LC3 puncta in a cell indicates a blockage in autophagy. This graph was taken from Ref [3].

of LC3 puncta per cell. This is indicative of a blockage in autophagic flux as the LC3 accumulated without being degraded after fusion with the lysosome. The T50A mutant is phospho-null and this therefore showed that prevention of phosphorylation at this site produced a related increase in LC3 puncta. The data from the T50E mutant therefore showed that the glutamate at this site allowed reuglar autophagic flux to be maintained regardless of the STK3/4 KO. This supports the use of the phosphomimetic in these biophysical studies to understand the mechanism of action of the phosphorylation.

Mutation of Thr50 to a Glu residue has been modelled on the LC3B-p62 AIM peptide crystal structure in Figure 7.3; showing the location of the incorporated charge on the surface. This was considered to be representative of the same change in surface charge caused by the phosphorylation, which was modelled previously in Figure 1.8. This shows the negative charge which has been added to the surface due to the carboxylate group of glutamate which is considered a mimetic of the charge from the phosphate group. This is in close proximity, based on the structure of the binding of the WT p62 AIM peptide, to the negatively charged aspartate residues of the p62 AIM. In particular this area of introduced charge could shield the electrostatic effects of Asp335 and Asp337 in their interactions with positive charge on the LC3B sur-

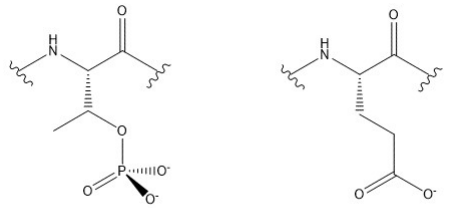



Figure 7.2: Chemical structures of phosphothreonine and glutamate amino acid residues as part of a protein structure to illustrate their similar size and charge despite some differences.

| Primer (5'-3') | Sequence |
|---|---|
| LC3B T50E Forward | TCTGGATAAAGAAAAGTTCCTTGTACCTG |
| LC3B T50E Reverse | ACAGGAAGCTGCTTCTCA |

Table 7.1: Primer sequences for the site-directed mutagenesis of LC3B.

face. Based on these structures it was predicted that this phosphomimetic mutation would decrease affinity for peptides which include acidic amino acid residues in these positions preceding the core AIM motif, as is the case for p62.

## 7.2 Mutagenesis

Site-directed mutagenesis was used to substitute bases within the LC3B insert in the pGEX-4T-1 vector for the T50E mutation, using the primers in Table 7.1. Colonies were then picked and cultured for plasmid DNA purification. The resulting DNA was sent for sequencing to confirm the substitution. Once confirmed, the DNA of this new LC3B T50E construct was transformed into BL21 (DE3) chemically competent *E. coli* cells.

## 7.3 Protein Over-expression

Over-expression of this new protein construct was tested using the same methodology and conditions as for the WT protein. Small scale expression tests showed that this construct expressed well after IPTG induction (1 mM) at 18$^{o}$C overnight, as shown in Figure 7.4. This SDS-PAGE gel shows a clear expression band in the samples collected after overnight induction which runs at a similar rate to the 40 kDa

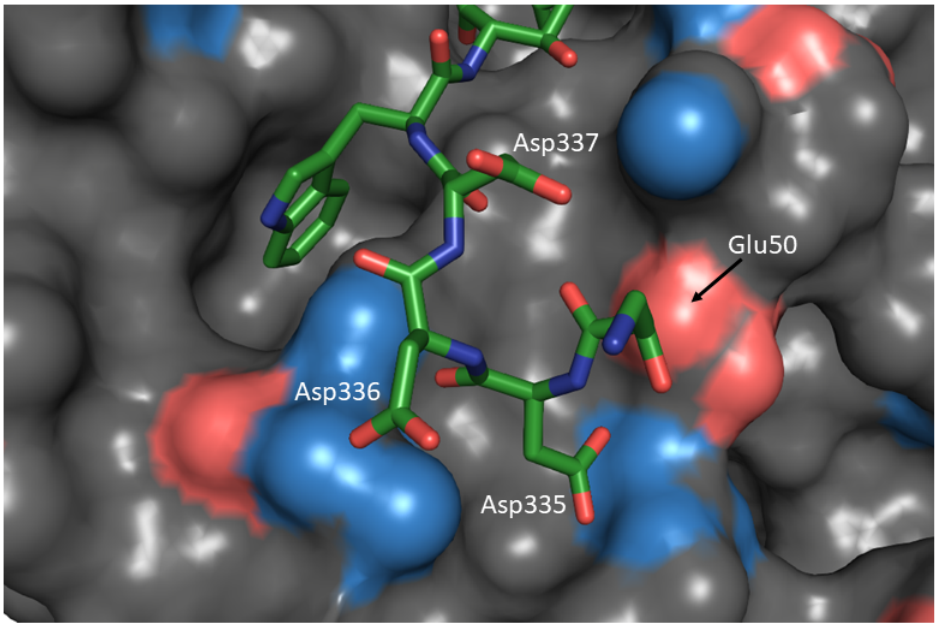



Figure 7.3: Crystal structure of LC3B bound to a peptide representing the AIM of p62 modified to include the LC3B T50E phosphomimetic mutation (modified from 2ZJD). This illustrates the predicted disruption of electrostatic interactions in this area upon the introduction of the negatively charged sidechain of GLu50. The LC3B surface is coloured according to charge: atoms with positive charge in blue, atoms with negative charge in red and the remaining surface is grey. The p62 peptide is coloured green with nitrogen atoms in blue and oxygen atoms in red.

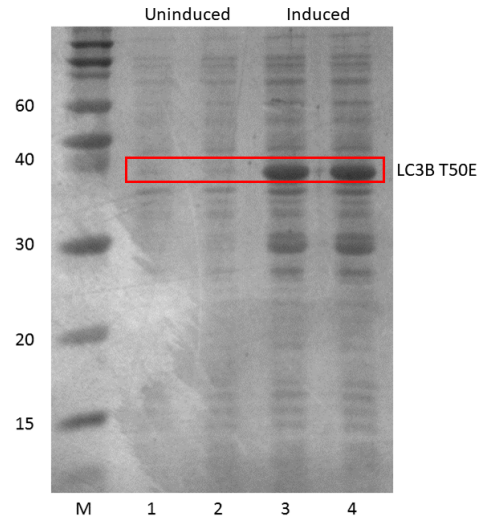



Figure 7.4: SDS PAGE gel showing the expression test of LC3B T50E. The red box highlights the relevant protein band. Lane loading: M) Markers 1) Uninduced total lysate 2) Uninduced soluble lysate 3) Total lysate after overnight induction 4) Soluble lysate hours after overnight induction.

marker. This was expected for the GST-tagged LC3B T50E construct which has a predicted protein mass of 41026.6 Da. The gel also shows no clearly observable 'leaky expression' of the protein before induction. As there appears to be equal amounts of protein in the total and soluble fractions after overnight induction; it can also be surmised that expression of insoluble protein is minimal.

## 7.4 Protein Purification

LC3B T50E has been purified using the same three-step chromatographic method as for the WT protein. This involved an initial affinity chromatography step followed by IEX and then SEX chromatography. Protein samples from throughout this process were separated and visualised on SDS-PAGE gels which are shown in Figure 7.5. The affinity chromatography step utilised Glutathione Sepharose 4B beads to capture the GST-tagged protein from the cell lysate shown in lane one. The GST-tagged protein bound to the beads is shown in lane two and the beads after thrombin cleavage are shown in lane three. The protein sample eluted from the beads is shown in lane four which shows the desired LC3B T50E protein which appears at a slightly higher molecular weight than the 15 kDa marker as well as some other protein contaminants. LC3B T50E has an expected protein mass of 14716.04.

As for the WT, the remaining two chromatographic steps utilise FPLC. In order

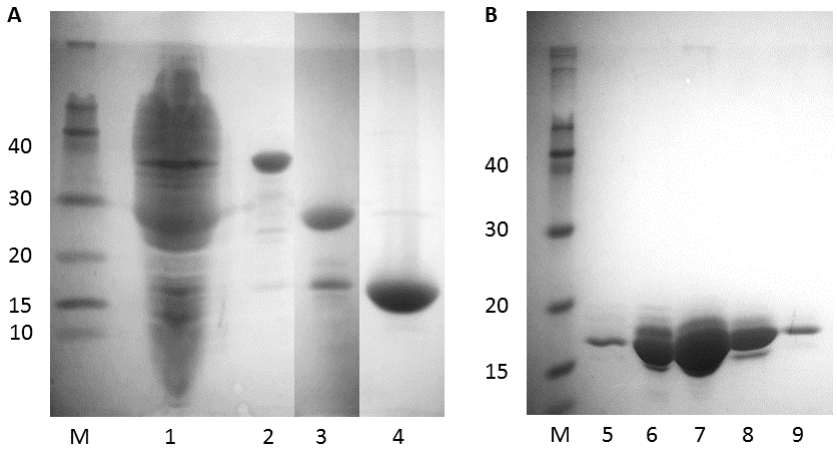



Figure 7.5: SDS PAGE gels showing purification of LC3B T50E from cell lysate. Left-hand gel (A) shows the affinity chromatography step, cropped from a single gel. Lane loading: M) Markers 1) Soluble cell lysate 2) Glutathione sepharose beads after incubation with cell lysate 3) Glutathione sepharose beads after overnight incubation with thrombin 4) Elution from beads after cleavage. The right-hand gel (B) shows pure protein fractions eluted from the IEX chromatography step. Lane loading: M) Markers 5-9) IEX elution fractions.

for IEX chromatography to be carried out the theoretical pI of the protein was first calculated, based on its sequence using the EXPASY Prot Param server, to be 7.97. Based upon this value the LC3B T50E protein was predicted to be positively charged in Ion Exchange Buffers A and B (see Table 2.1) which have a pH of 7. As such the protein sample eluted from the Glutathione Sepharose 4B beads was applied to an SP column, as for the WT. After a washing step in Ion Exchange Buffer A, protein bound to the matrix was eluted using a salt gradient. An SDS PAGE gel of eluted fractions is shown in Figure 7.5 which shows the pure protein.

The SEX chromatography step was then used to exchange the protein into the Desalt/MS Buffer (see Table 2.1). The ammonium acetate in this solution was then removed from the sample by repeated lyophilisation due to its volatility. This produced a protein sample, devoid of any other sample components such as buffer or salt, as a solid which was resuspended into the desired buffers for the following experiments in this chapter.

ESI MS was used to confirm the protein identity and purity. The protein was resuspended in Desalt/MS Buffer and the spectrum produced from this sample is shown in Figure 7.6A. The deconvoluted mass from this experiment was 14860.32±0.55 which is close to the expected mass of 14860.16. This also shows a relatively small amount of salt adduct peaks and no other components in the sample. This protein sample

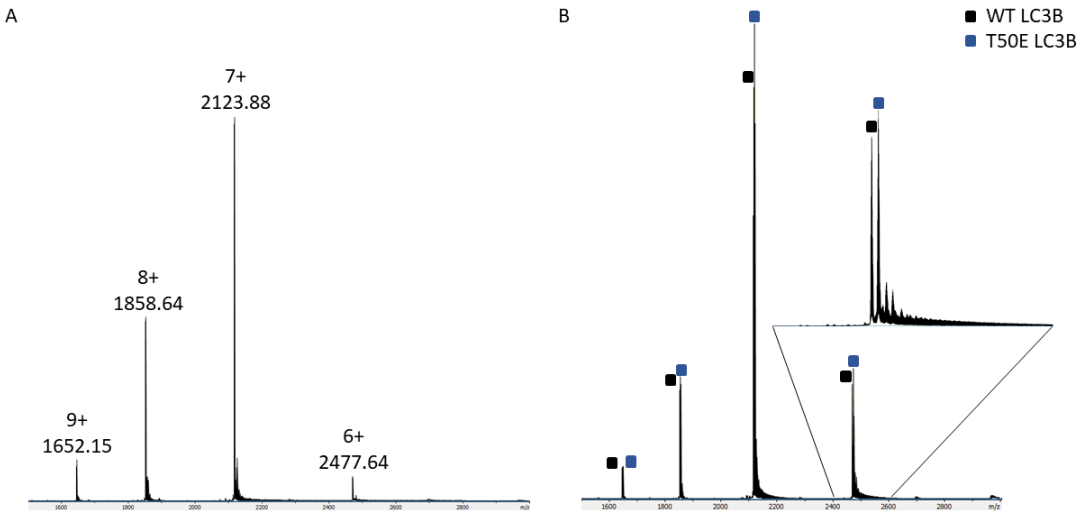



Figure 7.6: ESI MS spectrum of LC3B T50E to confirm sample identity and purity. (A) ESI MS of LC3B T50E (10 $\mu$M) showing the m/z values of the various charge states of the protein; (B) ESI MS of a mixture of WT and T50E LC3B (both at 10 $\mu$M) with an insert to illustrate the distinction between peaks.

was therefore considered to be of sufficient purity to support its use in further experiments. Figure 7.6B shows the ESI MS spectrum of an equimolar mixture of WT and T50E LC3B to show the presence of two separate proteins which confirms the T50E mutation.

## 7.5 Initial protein characterisation

CD spectroscopy was used as an initial test of protein folding after the mutation and also to allow comparison with the WT protein. LC3B T50E was resuspended in CD Buffer (see Table 2.1) and the CD spectrum of this sample is shown in Figure 7.7. There are clear indications of secondary structural elements present in this spectrum as signified by the occurence of differential absorbances occurring at wavelengths of light other than the minimum at around 200 nm typical of a random coil. This can be considered indicative of a folded protein.

The similarity of the T50E spectrum with that of the WT, also shown in Figure 7.7, indicated that the two proteins share similar proportions of secondary structural elements and therefore overall structure. Estimation of the proportion of each secondary structural element present in LC3B T50E was carried out from this spectrum using CDNN software. The results from this fitting process are shown in Table 7.2

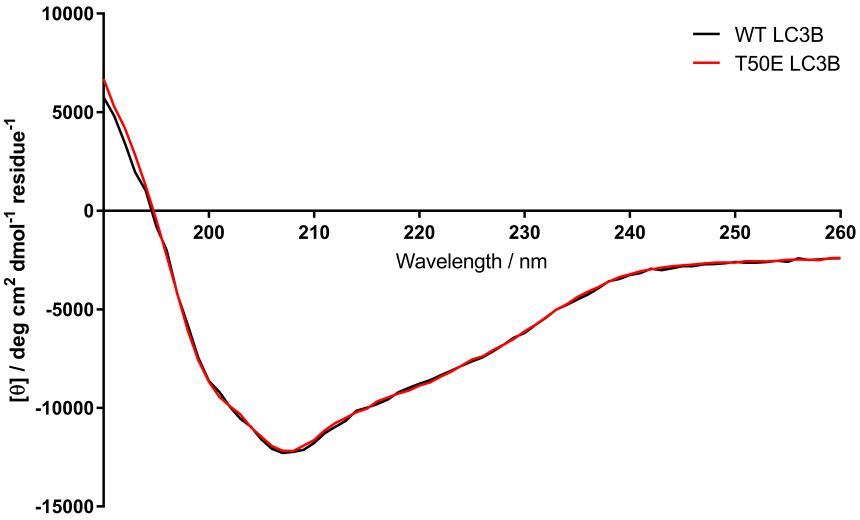



Figure 7.7: Overlay of CD spectra from WT LC3B (black) and T50E LC3B (red) to show the similarity of these spectra which indicated a similar protein fold.

| | **T50E Calculated** | **WT Calculated** | **WT Predicted** |
|---|---|---|---|
| **$\alpha$-helix** | 33% | 29% | 29% |
| **$\beta$-strand** | 25% | 20% | 21% |

Table 7.2: Proportion of secondary structural elements present in the WT and T50E LC3B proteins as calculated from CD data by CDNN compared to the expected values based on structures from the PDB.

along with the previously determined values for LC3B and the predicted values for comparison. This data indicates there was a slightly different protein fold present for LC3B T50E than was observed for the WT though these differences may be considered within the error of this analysis.

The melting transition of LC3B T50E was also investigated using CD spectroscopy as a further point of comparison with the WT protein. The determined melt curves for the WT and T50E proteins are shown in Figure 7.8 by following the MRE for each protein at 210 nm. This produced a $T_m$ value of 60.7$^{o}$C for LC3B T50E with the value for the WT being evaluated at 57.2$^{o}$C.

1D $^{1}$H NMR was utilised here as a further tool for investigation of protein folding. LC3B T50E was resuspended in NMR Buffer (see Table 2.1) and the spectrum produced from this sample is shown in Figure 7.9. As before the dispersion of amide protons and the presence of upshifted methyl peaks are a good indication of protein folding.

This technique is also useful in the study of a protein after mutagenesis as it allows

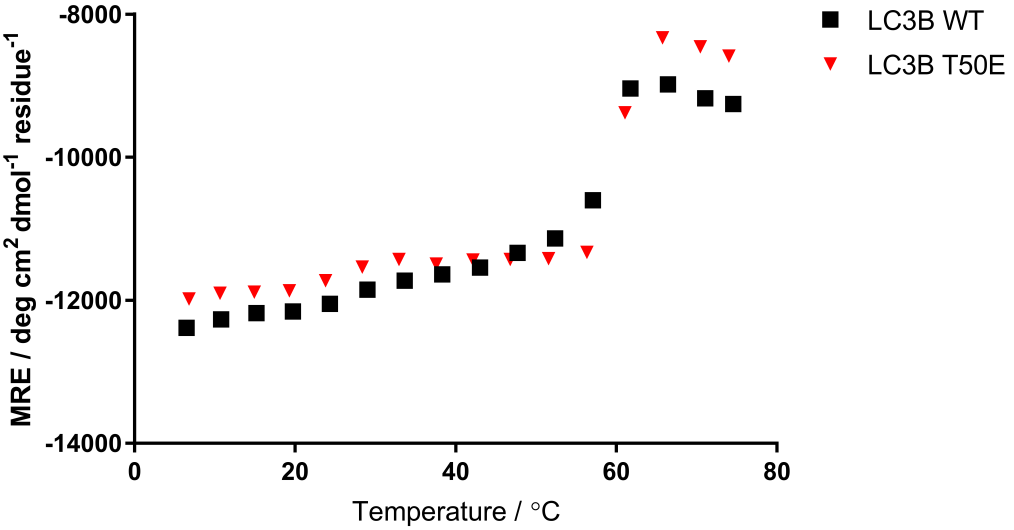



Figure 7.8: Melt curves of the WT and T50E LC3B proteins derived from CD thermal denaturation experiments which allow calculation of $T_m$ values.

a comparison of the new protein construct with the WT. Overlaying and comparison of the 1D $^1$H NMR spectra of the WT and T50E proteins has been utilised as an initial investigation into the effect of this mutation. Also shown in Figure 7.9 is the 1D $^1$H spectrum of the WT LC3B protein in order to make this comparison. Although these two proteins produce relatively similar spectra there are some clear differences. The upfield shifted methyl peaks from both spectra provide the easiest point of analysis where there is clearly a difference in chemical shift which is indicative of different packing of these methyls in the hydrophobic core of the protein. There were also observable shifts in the amide region of the spectrum but the shifts in the spectra are present throughout.

## 7.6 Effect on p62 binding

### 7.6.1 Mass Spectrometry

A native-like ESI MS competition experiment was initially used to investigate the effect of the LC3B T50E mutation on p62 AIM binding. In this experiment the WT p62 AIM peptide was titrated into an equimolar mixture of LC3B WT and LC3B T50E and the intensities of peaks corresponding to free and bound protein were monitored.

The initial spectrum before the addition of the peptide and the spectrum at a

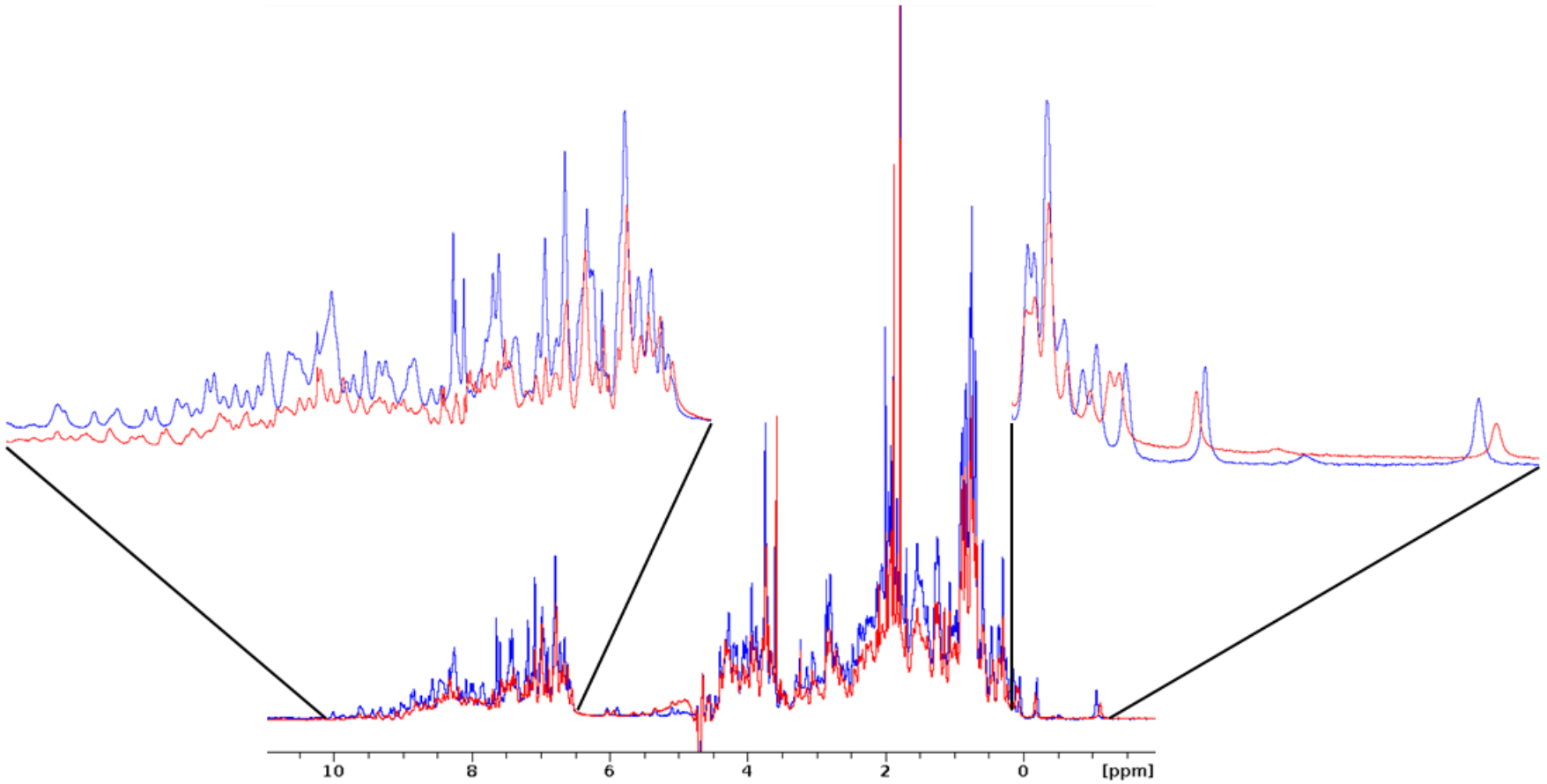



Figure 7.9: $^{1}$H NMR spectrum of WT LC3B (red) and T50E LC3B (blue) to assess protein folding and allow comparison. LC3B T50E protein folding is implied by upfield shifted methyl signals and a broad distribution of the amide proton peaks. There are some clear differences in the two spectra indicating some differences in protein structure between the WT and T50E proteins.

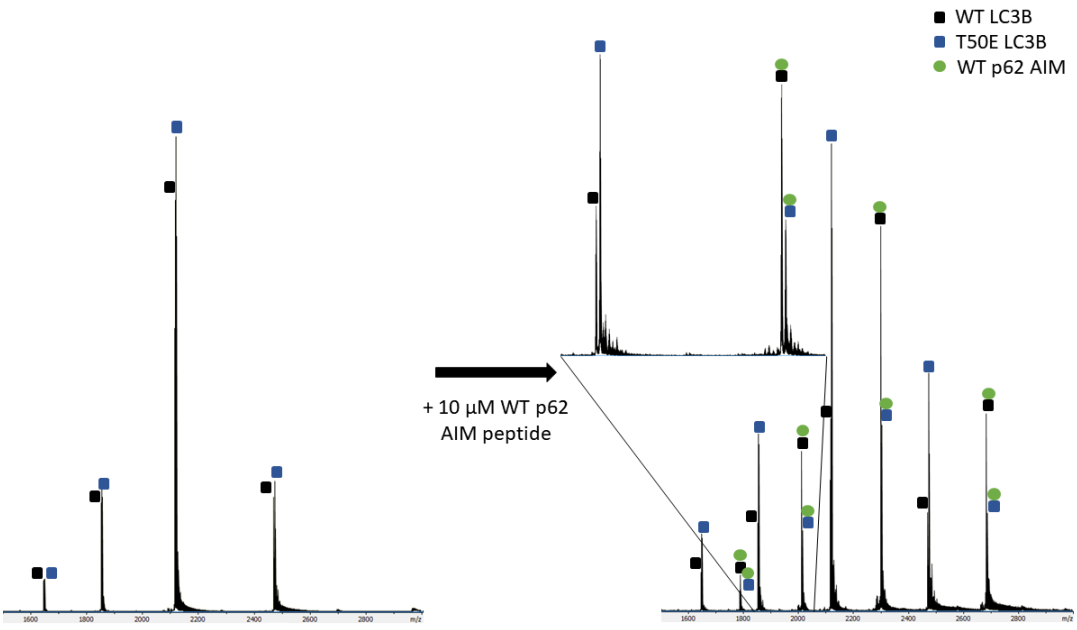



Figure 7.10: ESI MS spectra showing the competition experiment for the WT p62 AIM peptide binding to either WT LC3B or T50E LC3B. Left-hand spectrum: Sample containing LC3B and LC3B T50E (10 $\mu$M) which shows the proteins in their observed charge states; right-hand spectrum: sample containing LC3B and LC3B T50E (10 $\mu$M) and the WT p62 AIM peptide (10 $\mu$M) which shows a mixture of peaks corresponding to bound and free protein. Peaks are labelled by the peptide they correspond to, as shown in the legend.

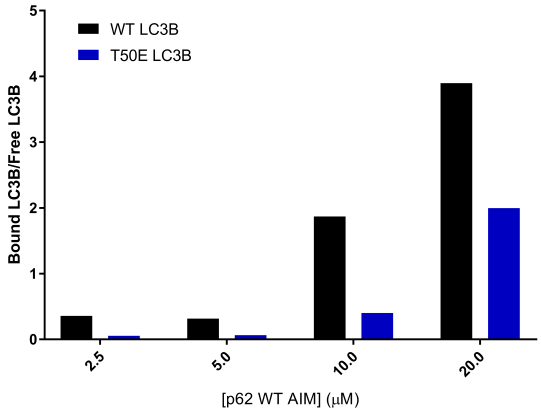


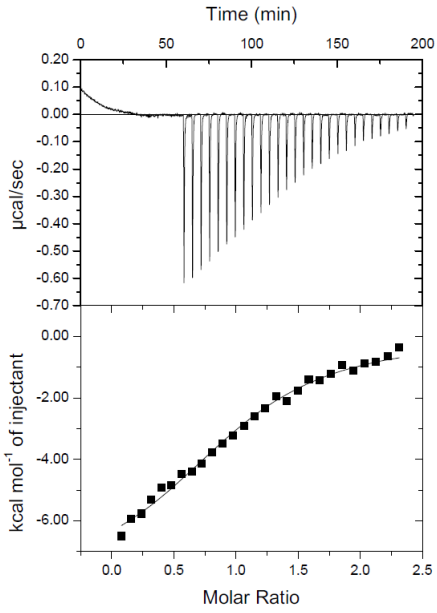

Figure 7.11: Graphs showing data from the ESI MS competition experiment for the WT p62 AIM peptide titrated into an equimolar mixture of WT LC3B and T50E LC3B. This shows the ratio of bound to free protein at each peptide concentration and illustrates a binding preference for LC3B over the phosphomimetic LC3B T50E protein.

subsequent titration point in this experiment are shown in Figure 7.10. The spectrum for the 10 $\mu$M titration point shows that the intensity of the unbound T50E peaks were higher than the intensity of the unbound WT peaks and subsequently the bound T50E peaks had a lower intensity than the bound WT peaks.

The ratios of bound to unbound protein at each titration point have been determined and are displayed in Figure 7.11. This shows that at each titration point the ratio of bound to free WT LC3B is higher than for the T50E protein. At the 10 $\mu$M titration point, the WT protein had a ratio 4.6 times greater than for the T50E mutant. This indicates that the WT p62 AIM peptide is preferentially binding to WT LC3B over the phosphomimetic T50E mutant.

### 7.6.2 Isothermal Titration Calorimetry

ITC was used to investigate the interaction of LC3B T50E with the WT p62 AIM peptide. To achieve this peptide aliquots were sequentially injected into a protein sample and the heat change due to their interaction was detected. This data is shown as applied power in Figure 7.12 which also shows the integrals of the heat spikes plotted and fitted. The calculated thermodynamic parameters from this experiment are also tabulated in this figure. The 'n' value produced from this fit varies from the expected value by 3% and the errors for the other parameters are $\leq$10% different from the fitted value. Additionally, the fits in all cases produced low $\chi^2$/DoF values so the




| Protein | n | $K_D$ ($\mu$M) | $\Delta H$ (kJ mol$^{-1}$) |
|---|---|---|---|
| **LC3B T50E** | 1.03±0.03 | 7.69±0.77 | -33.38±1.36 |
| **LC3B WT** | 0.92±0.03 | 3.83±0.41 | -44.85±1.83 |

| Protein | $\Delta S$ (J mol$^{-1}$K$^{-1}$) | $\Delta G$ (kJ mol$^{-1}$) |
|---|---|---|
| **LC3B T50E** | -1.21 | -29.68 |
| **LC3B WT** | -4.39 | -31.41 |

Figure 7.12: ITC data for the titration of the p62 WT AIM peptide into LC3B T50E. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel. The table shows the ITC-derived thermodynamic parameters produced from fitting this data..

fit can be considered satisfactory. The $K_D$ value determined here of 7.69±0.77 $\mu$M can be compared with that for LC3B WT (3.83±0.41 $\mu$M) to support the preferential binding of this peptide to the LC3B WT over the T50E mutant.

Figure 7.13 shows the enthalpic and entropic contributions to the binding interaction for the WT p62 AIM peptide binding to WT and T50E LC3B. It can be seen that the reduction in binding affinity was driven by a decrease in enthalpy despite some compensation from the entropy becoming less unfavourable.

## 7.7 Discussion

After succesful mutagenesis, the phosphomimetic T50E LC3B mutant was over-expressed using the same *E. coli* expression system as was used for the WT; discussed

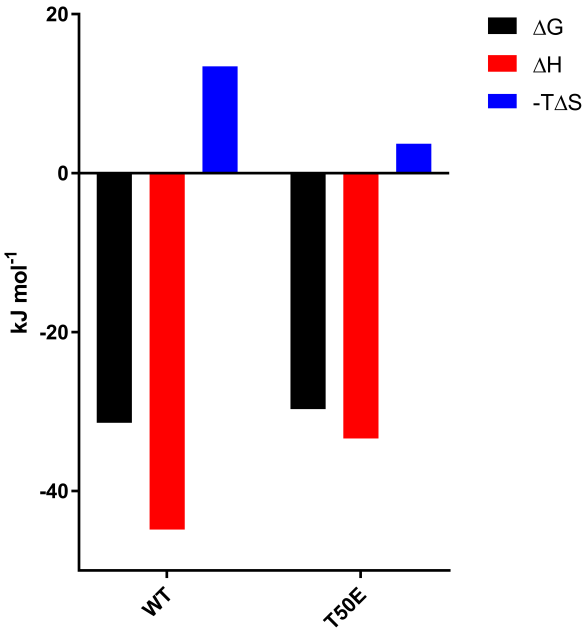



Figure 7.13: Graph comparing ITC-derived thermodynamic parameters for the binding of the WT p62 AIM peptide to WT or T50E LC3B to illustrate the differences in the enthalpic and entropic contributions to the Gibbs free energy.

in Chapter 3. Despite having a lower predicted pI value than the WT protein, the T50E mutant was also able to bind to an SP IEX column and therefore the same purifcation methodology was used as for the WT.

MALDI MS was previously used to confirm protein identity however due to the substantial differences between observed and predicted mass in those experiments this method was not used to confirm the protein mass here. As there is only a small difference in mass between the WT and T50E LC3B proteins, MALDI MS would have been unable to confirm the different mass of the mutated protein. Construct identity and purity was instead confirmed using ESI MS which showed the protein to have the expected mass and to be sufficiently pure for further study. The over-expression, purification, lyophilisation and subsequent resuspension of LC3B T50E was deemed to produce folded protein based upon the CD and NMR spectra produced after these processes. These spectra showed characteristic signals of folded protein such as differential absorbance at wavelengths other than 210 nm in the CD spectrum and the presence of upfield shifted methyl signals and dispersed amide peaks in the $^{1}$H NMR spectrum.

These experiments also allowed the structural comparison of the WT and T50E LC3B constructs. The CD spectra of the two proteins appeared near identical in

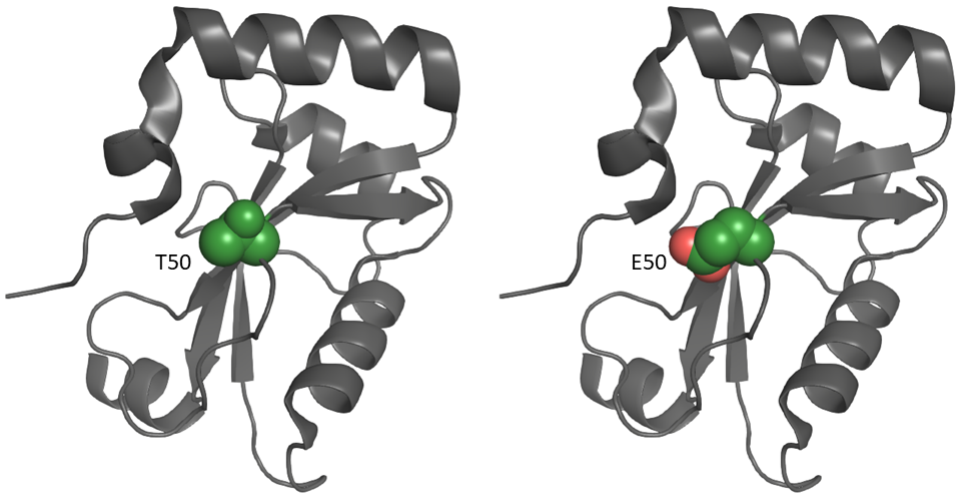



Figure 7.14: Crystal structure of LC3B showing the bulk of either threonine (left) or glutamate (right) at position 50 in the structure. The two structures are shown in cartoon representation other than the highlighted residues whose side chains are shown as spheres. The sidechains are coloured green other than the negatively charged oxygen atoms on the glutamate sidechain.

terms of both lineshape and magnitude which would indicate minimal perturbation to protein secondary structure elements from the mutation. The $T_m$ values produced for the WT and T50E LC3B proteins from the CD thermal denaturation experiments are similar but the T50E protein was determined to be more stable by 3.5$^{o}$C. Determination of the $T_m$ from this data suffers from a sampling problem which leads to a binding curve that is not well defined. These values should therefore be considered as estimates that show the proteins have a similar $T_m$value. The errors and biases of analysis using CDNN were discussed in Section 4.3 and the values produced from this analysis here should be considered within the range of error; indicating similar proportions of secondary structure for the WT and T50E proteins.

The WT and T50E LC3B proteins produced similar 1D $^{1}$H NMR spectra with some differences. Throughout the spectra there appeared to be similar peaks with related peak shapes and sizes which have been shifted along the spectra slightly or in some cases the line width is altered. The peak shapes of the upfield shifted methyls for example were very similar but simply with slightly different chemical shifts. The peak at -1.15 ppm shifted to -1.08 ppm and the peak at -0.181 ppm shifted to -0.195 ppm for example. This indicated that the packing of these methyls, which produces the upfield shifts, has been perturbed due to this mutation.

There is therefore a potential contradiction within these data whereby CD spectroscopy indicates the WT and mutant protein have the same structure but NMR shows that some of the signals experience a different chemical environment due to the mutation. However, these data can be explained by the sensitive nature of NMR experiments whereby small perturbations in the protein fold are too slight to affect the CD spectrum as they do not significantly alter the proportions of secondary structure elements present but alter the chemical environment of some atoms sufficiently for NMR detection. Overall, this indicated that these proteins have a similar structure but the incorporation of a larger, charged amino acid at position 50 has perturbed the structure to some unknown though small degree. Figure 7.14 shows the bulk of Thr and Glu within the context of their positions in LC3B which supports the proposition that this mutation would shift the structure of the protein due to sterics as well as charge. It might be that this bulk shifts the N-terminal helices due to their proximity. As this phosphomimetic mutation has been shown to reconstitute activity from the STK3 and STK4 KO, the assumption is that this small shift in structure mimics the effect on the structure of phosphorylation at T50. Therefore, further study of this structural shift was not pursued here.

Both MS and ITC experiments support the assertion that the WT p62 AIM peptide preferentially bound to WT LC3B over the T50E phosphomimetic mutant. The ITC-derived $K_D$ values indicated that the interaction of the peptide was twice as strong for WT as for T50E LC3B. Additionally, the phosphomimetic mutation produced a species with a lower charge than for a phosphate group so the true effect from phosphorylation at this site may be greater than observed here.

LC3B has been shown to interact with a high affinity to the p62 AIM and this interaction is thought to be important for recruiting p62 filaments to the autophagosome membrane to allow them to function as a scaffold for membrane expansion and for the selective recruitment of autophagy cargo [59, 49, 74, 189]. Phosphorylation at T50 also links LC3B with autophagosome-lysosome fusion, presumably by increasing the affinity of LC3B for some autophagy adapter protein or proteins which has not yet been determined [3].

The hypothesis therefore is that STK3/4 phosphorylation at T50 operates as a switch for LC3B between these two functions. As this phosphorylation is needed for fusion to occur optimally then it can be assumed that WT LC3B has a reduced affinity for whichever autophagy adapter protein or proteins that mediate this role whilst non-phosphorylated LC3B interacts strongly with the p62 AIM. At some point STK3/4 phosphorylation is increased, by some spatiotemporal upregulation or selec-

tivity, leading to a decreased affinity of LC3B for p62 and presumably an increased affinity for the autophagy adapter protein or proteins.

The final piece to confirm this would be the identification of an autophagy adapter which mediates autophagosome-lysosome fusion through binding to phosphorylated LC3B. One potential candidate for this role is PLEKHM1 which has been shown to be involved in autophagosome-lysosome fusion [30]. This study by McEwan et al. showed that PLEKHM1 immunoprecipitated LC3B from cell lysate through binding at its AIM domain. This work also demonstrated that genetic loss of the PLEKHM1 impeded autophagic flux at the point of autophagosome-lysosome fusion. Therefore, PLEKHM1 may be the autophagy adapter involved in this switching of LC3B binding but further work is required.

The data presented in this chapter has shown that:

- The LC3B T50E phosphomimetic mutant protein was produced.
- Structural comparison between the WT and T50E LC3B structures indicates a small shift in structure due to the mutation.
- The WT p62 AIM peptide binds preferentially to the WT LC3B protein over the phosphomimetic T50E LC3B.

This chapter and the previous one have investigated two modes by which this important protein-protein interaction can be regulated to control autophagic flux. The following chapter however will look at how this interaction is dysregulated by a mutation and how that dysregulation may contribute to disease.

# Chapter 8

# Autophagic Dysregulation by a Disease-related Mutant

## 8.1 Introduction

The ALS-related L341V p62 mutation studied in this work had been considered benign before its more thorough study both by biophysical techniques and live cell imaging studies. The structures of leucine and valine residues are shown in Figure 8.1. This shows the key differences in these residues in terms of both size and shape despite them both being medium-sized hydrophobic amino acids.

The previous study of the interaction between the L341V peptide and LC3B suggested that the $\beta$-branching of valine, as opposed to the $\gamma$-branching of leucine, precludes its ability to tightly bind into the L-site on the LC3B surface.

The NMR titration data in that study indicated that the residues which form the L-site on LC3B were differentially perturbed upon binding to the L341V peptide than for the WT. Differences in perturbation indicate differences in binding mode and in this case it indicated more significant reorganisation of the protein to accomodate the L341V peptide due to steric clashes.

This chapter presents results from an investigation into the binding interactions of the L341V p62 AIM peptide with all six hAtg8 proteins. MS and ITC were utilised to study the L341V peptide as this allowed direct comparison to the results collected using the WT peptide in Chapter 4. This work has determined affinities and relative selectivity for each of the hAtg8 proteins binding to the L341V p62 AIM peptide and also produced a measure of the thermodynamic properties of these interactions.

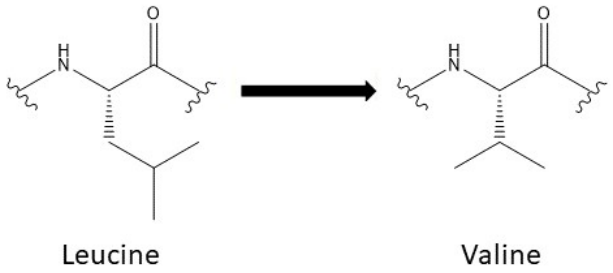



Figure 8.1: Chemical structures of leucine and valine amino acid residues as part of a protein structure to illustrate the change in size and shape which the L341V mutation represenets.

## 8.2 Effect of the p62 L341V mutation

### 8.2.1 Mass Spectrometry

Native-like ESI MS competition experiments were initially used to investigate the effect of the L341V p62 mutation. A related competition experiment setup to that used to investigate the effect of phosphorylation in Chapter 6 was used here. In these titration experiments each hAtg8 protein was separately titrated into an equimolar mixture of WT and L341V p62 AIM peptide and the relative ratio of the intensity of peaks corresponding to the unbound peptides was monitored.

The spectra for the initial equimolar mixture of the peptides and a subsequent titration point for GABARAP, as an example, are shown in Figure 8.2. This shows the reduction in the relative intensity of the two peaks corresponding to the WT peptide compared to the L341V peptide upon addition of GABARAP. This implies that GABARAP is selectively binding to the WT peptide over the L341V peptide. Peaks for the free and bound GABARAP protein can also be observed however due to the similar m/z values there was not a clear distinction between the peaks for GABARAP bound to the WT or L341V peptide.

This work was then expanded to explore the effect of the mutation on binding to each of the hAtg8 proteins through a full titration experiment. Graphs showing the data from the titration of each hAtg8 into an equimolar mixture of WT and L341V peptide are shown in Figure 8.3. These graphs show the ratio of free WT peptide to free L341V peptide at varying hAtg8 concentration. Some data points were removed as they were clear outliers.

The general trend followed by LC3A, LC3B, GABARAP and GABARAPL1 shows

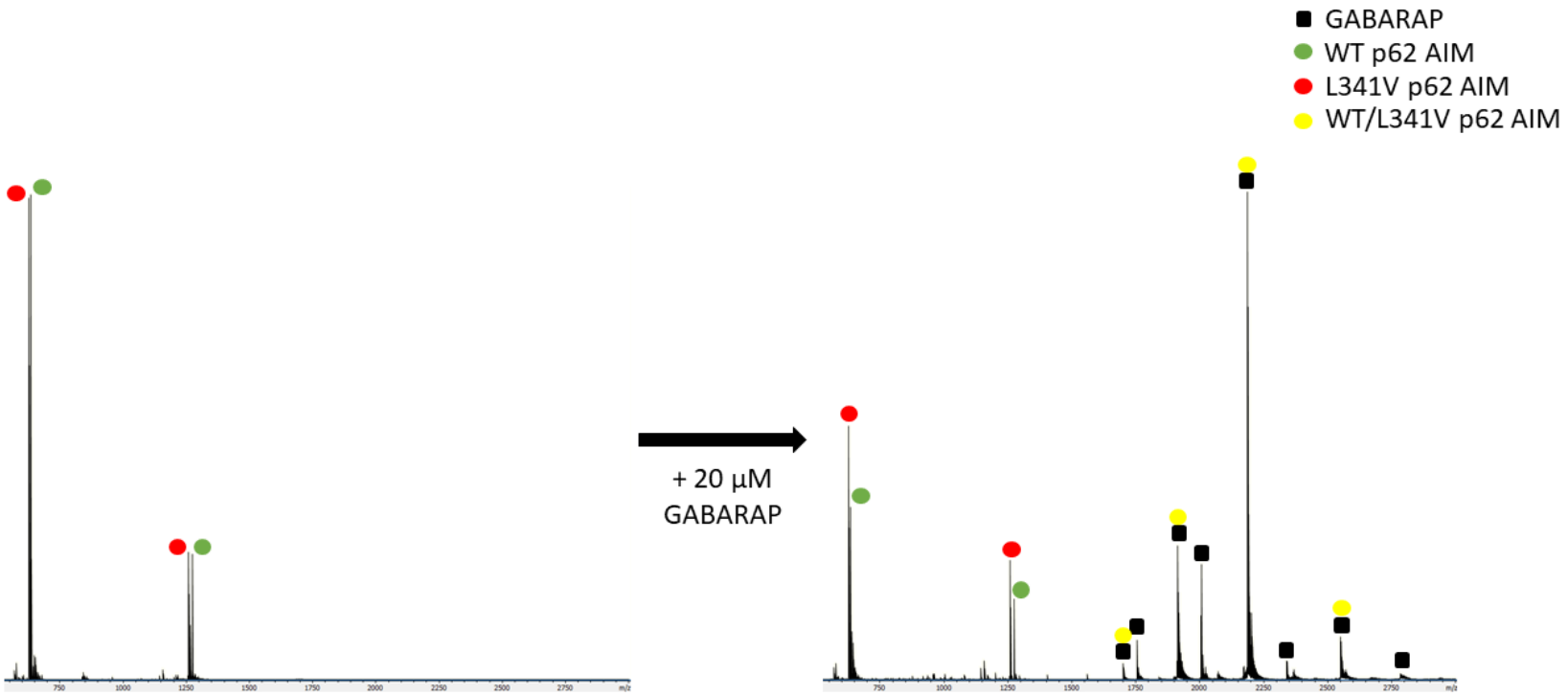



Figure 8.2: ESI MS spectra showing the competition experiment for the WT or L341V p62 AIM peptide binding to GABARAP. Left-hand spectrum: Sample containing the WT and L341V p62 AIM peptides (10 $\mu$M) which shows the peptides in their two observed charge states; right-hand spectrum: sample containing the WT and L341V p62 AIM peptides (10 $\mu$M) and GABARAP (30 $\mu$M) which shows the change in relative intensity of the two peptides. Peaks are labelled by the protein, peptide or protein-peptide complex they correspond to, as shown in the legend. Due to the overlap of the two protein-peptide complex peaks they are simply labeled as binding to both p62 AIM peptides.

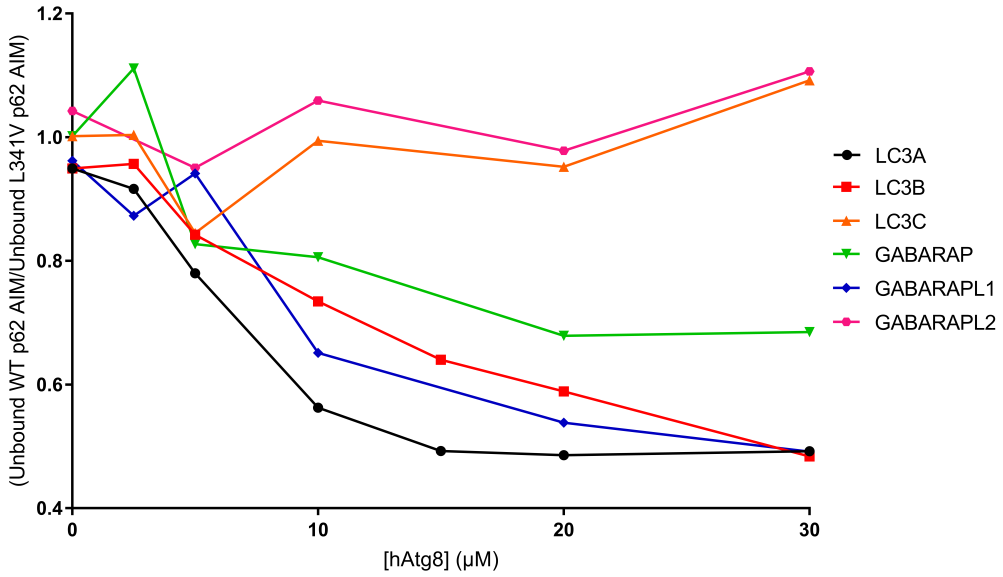



Figure 8.3: Graph showing data from ESI MS competition experiments for each of the six hAtg8 proteins titrated into an equimolar mixture of the WT and L341V p62 AIM peptides. This shows the ratio of the intensities of the unbound peptides at each hAtg8 concentration and illustrate varying trends in binding preference.

that as the hAtg8 concentration was increased the ratio of unbound WT peptide to unbound L341V peptide decreased. This is indicative of these hAtg8 proteins binding preferentially to the WT p62 AIM peptide over the L341V peptide.

The data collected from the LC3C or GABARAPL2 titrations show only a weak trend in binding preference among the peptides. During this titration the ratio of unbound WT peptide to unbound L341V peptide remained relatively unchanged which was indicative of a similar binding affinity for the interaction of these proteins with either of the two peptides.

In addition to studying these alterations in affinity to a single hAtg8 it was also possible to look at changes in selectivity amongst the proteins resulting from the L341V p62 mutation. These selectivity experiments were carried out using the same methodology as for the WT peptide competition experiments (Section 4.2.1) in which the peptide is titrated into an equimolar mixture of either the LC3 or GABARAP subfamily. The experiments were again split into the two subfamilies due to the congestion and subsequent peak overlap of attempting to investigate all six proteins together. Additionally, this allowed for comparison of the results with those for the WT peptide.

Figure 8.4 shows the spectra for two points during the titration of the L341V p62

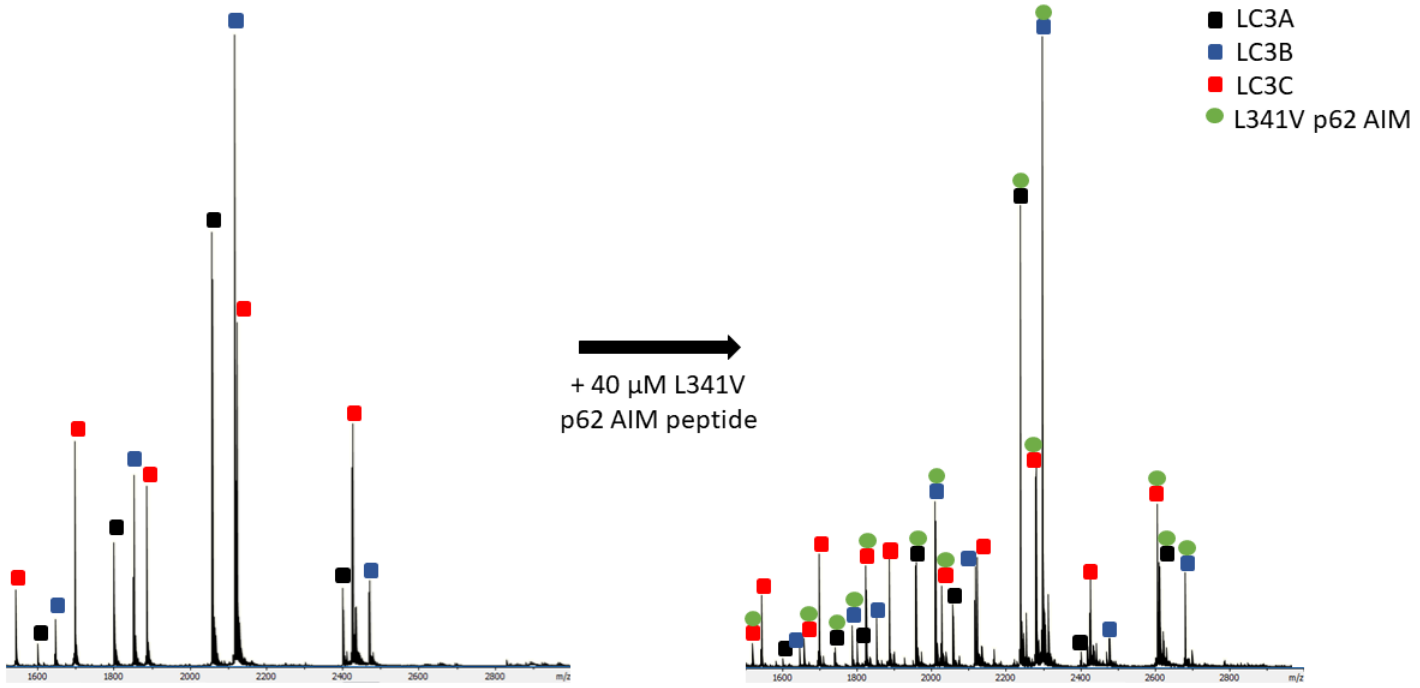



Figure 8.4: ESI MS spectra showing the competition experiment for the LC3 subfamily binding to the L341V p62 AIM peptide. Left-hand spectrum: Sample containing the three LC3 subfamily proteins (10 $\mu$M) which shows the proteins in their various charge states; right-hand spectrum: sample containing the three LC3 subfamily proteins (10 $\mu$M) and the L341V p62 AIM peptide (40 $\mu$M) which shows the proteins and the protein-peptide complexes in their various charge states. Peaks are labelled according to the particular protein or protein-peptide complex they correspond to as shown in the legend.

AIM peptide into an equimolar mixture of the LC3 subfamily. Figure 8.6A then shows the graphed results from this experiment where the ratio of bound to free protein was plotted for each peptide concentration. LC3A and LC3B appeared to bind to the peptide with a similar affinity as the determined ratios for these proteins were similar at each titration points.The final titration point indicated a slight preference for LC3A. LC3C was shown to bind with a lower affinity as the ratios determined for this protein were substantially less than for LC3A and LC3B.

The results of the equivalent competition experiment for the GABARAP subfamily are shown in Figure 8.6B and the spectra for two titration points in this experiment are shown in Figure 8.5. This set of titration experiments shows the L341V peptide interacting with GABARAP and GABARAPL1 with a similar affinity. There may have been a small potential preference for GABARAPL1 as it had a slightly higher bound to free ratio at each titration point. GABARAPL2 appears to have the lowest affinity for this peptide due to its significantly lower bound to free protein ratios.

A further competition experiment was carried out to compare the relative affinities of these two subfamilies for the L341V p62 AIM peptide. In this series, the peptide was titrated into an equimolar mixture of LC3A and GABARAP. As for the WT peptide, the selection of these hAtg8 proteins as representative of their group was arbitrary.

Figure 8.8 shows the graphed results from this titration and Figure 8.7 shows the spectra for two titration points. As the peptide concentration was increased, the relative level of bound GABARAP increased compared to LC3A. For the final titration point these ratios were 4.3 for LC3A and 6.7 for GABARAP which indicated a binding preference for GABARAP. This is a switch in binding preference from the WT peptide which preferentially interacted with LC3A over GABARAP (Section 4.2.1).

### 8.2.2 Isothermal Titration Calorimetry

ITC was used here to investigate the interaction of the hAtg8 proteins with the L341V p62 AIM peptide. Example data from the experiment into the interaction of GABARAP with the L341V p62 AIM peptide is shown in Figure 8.9. The fitted thermodynamic parameters from all six hAtg8 proteins are shown in Table 8.1 and the other five binding curves used to determine these parameters are shown in Appendix Figures A.24-A.28. This table also shows a measure of comparison between the $K_D$ value produced in these experiments and those for binding to the WT peptide, determined in Section 4.2.2.

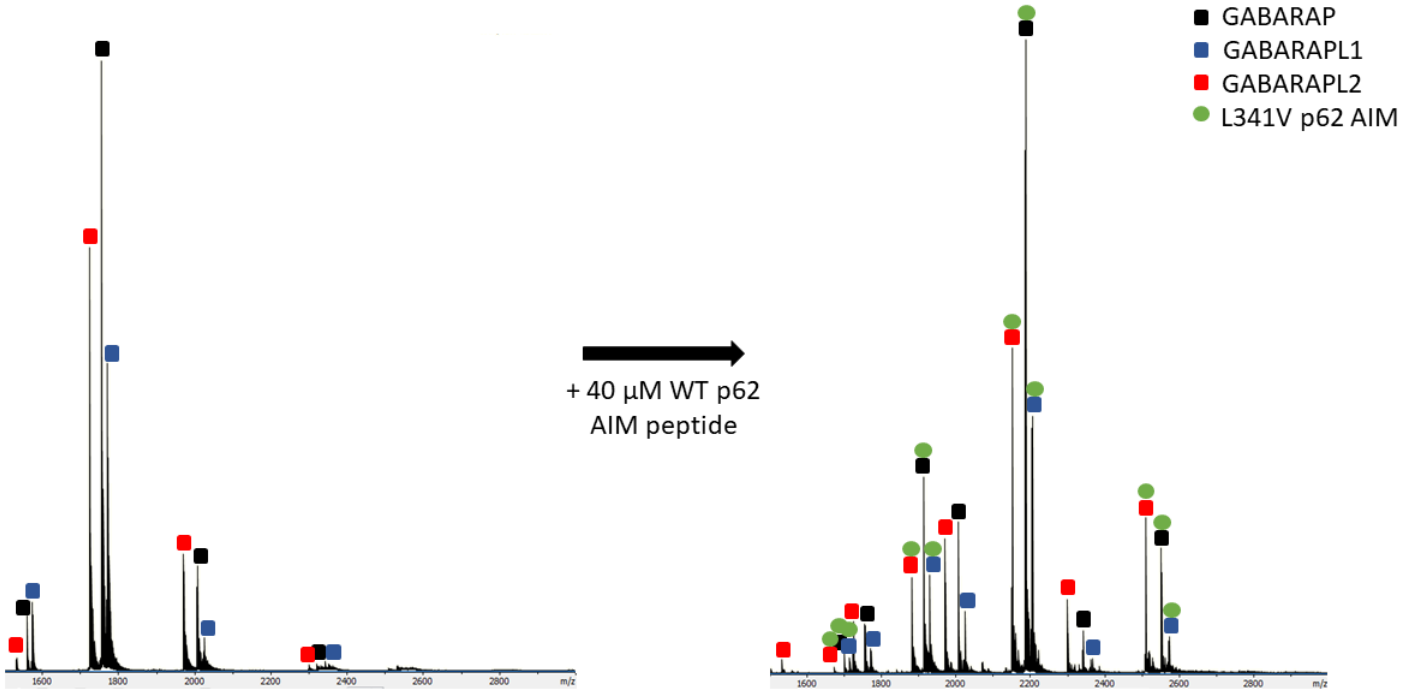



Figure 8.5: ESI MS spectra showing the competition experiment for the GABARAP subfamily binding to the L341V p62 AIM peptide. Left-hand spectrum: Sample containing the three GABARAP subfamily proteins (10 $\mu$M) which shows the proteins in their various charge states; right-hand spectrum: sample containing the three GABARAP subfamily proteins (10 $\mu$M) and the L341V p62 AIM peptide (40 $\mu$M) which shows the proteins and the protein-peptide complexes in their various charge states. Peaks are labelled according to the particular protein or protein-peptide complex they correspond to as shown in the legend.

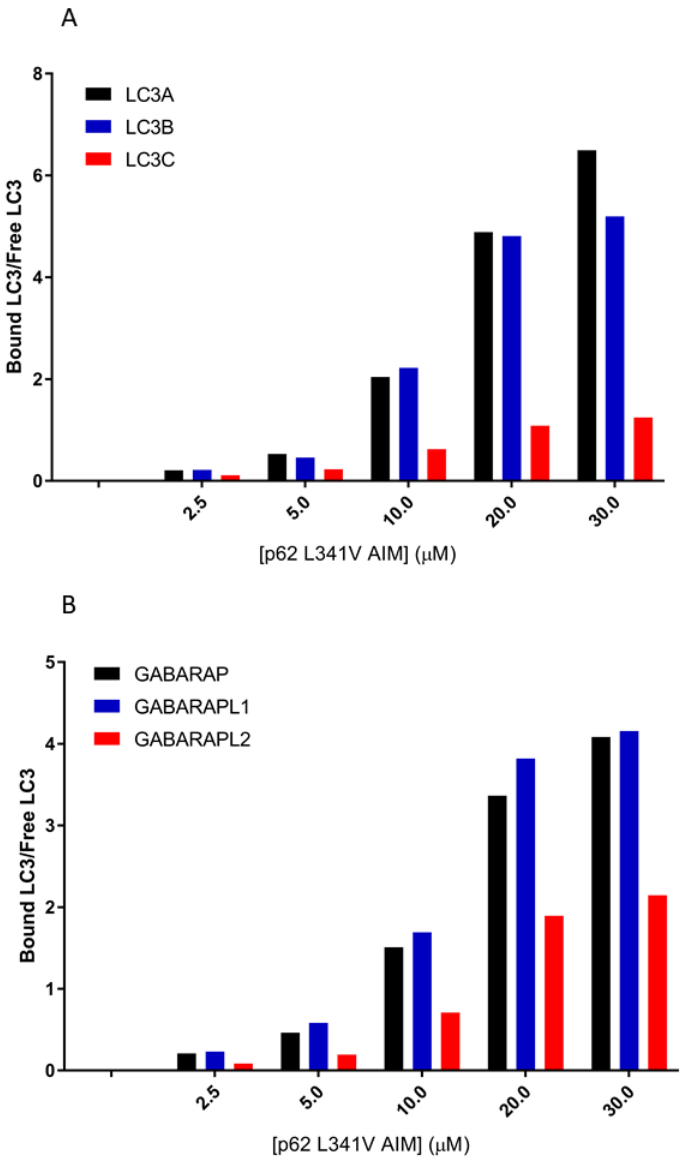



Figure 8.6: Graphs showing the results of the ESI-MS competition experiments as the p62 WT AIM is titrated into an equimolar sample of the hAtg8 proteins to compare binding within each subfamily. Top graph (A) is for the LC3 subfamily and the bottom graph (B) is for the GABARAP subfamily. These graphs show a lower degree of selectivity for binding to the L341V p62 AIM peptide than was observed for the WT peptide.

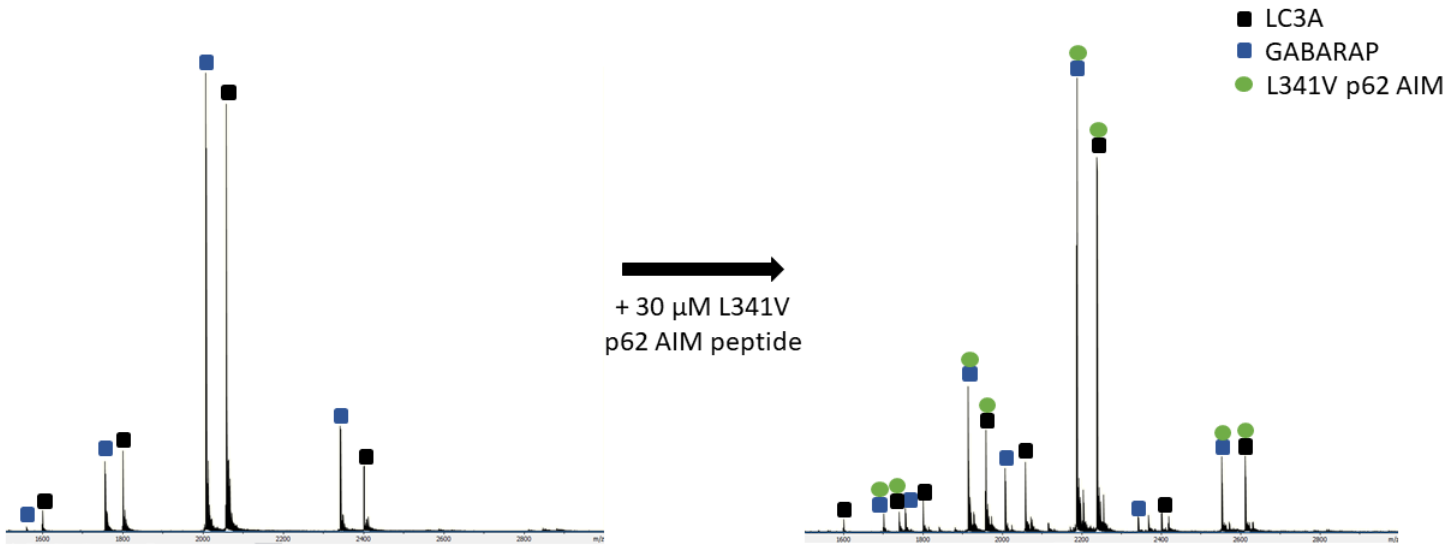



Figure 8.7: ESI MS spectra showing the competition experiment for LC3A and GABARAP binding to the L341V p62 AIM peptide. Left-hand spectrum: Sample containing LC3A (10 $\mu$M) and GABARAP (10 $\mu$M) which shows the proteins in their various charge states; right-hand spectrum: sample containing LC3A (10 $\mu$M) and GABARAP (10 $\mu$M) and the L341V p62 AIM peptide (30 $\mu$M) which shows the proteins and the protein-peptide complexes in their various charge states. Peaks are labelled according to the particular protein or protein-peptide complex they correspond to as shown in the legend.

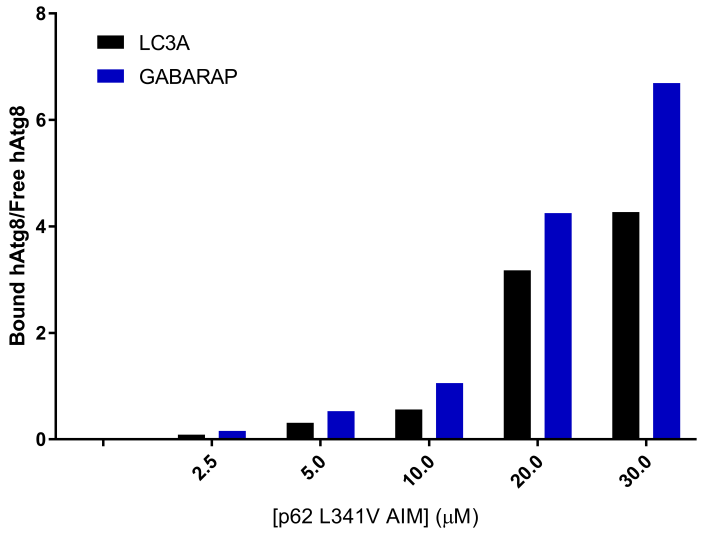



Figure 8.8: Graph showing the results of the ESI-MS competition experiment as the L341V p62 AIM peptide is titrated into an equimolar sample of LC3A and GABARAP (10 $\mu$) which shows a preference in binding for GABARAP.

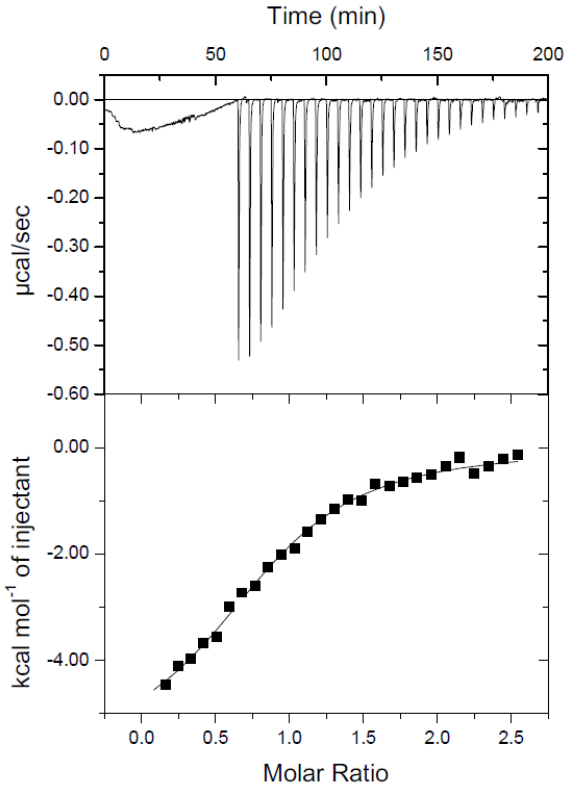



Figure 8.9: ITC data for the titration of the L341V p62 AIM peptide into GABARAP. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

| Protein | n | $K_D$ ($\mu$M) | $\Delta H$ (kJ mol$^{-1}$) | $\Delta S$ (J mol$^{-1}$ K$^{-1}$) | $\Delta G$ (kJ mol$^{-1}$) | Affinity reduction compared to WT |
|---|---|---|---|---|---|---|
| **LC3A** | 0.87±0.08 | 8.70±1.84 | -26.37±3.50 | 0.98 | -29.36 | 1.91 |
| **LC3B** | 1.14±0.01 | 8.47±0.61 | -16.20±0.33 | 4.31 | -29.39 | 2.21 |
| **LC3C** | 0.90±0.02 | 11.40±0.95 | -22.52±0.98 | 2.01 | -28.69 | 0.68 |
| **GABARAP** | 0.87±0.02 | 6.41±0.52 | -24.19±0.87 | 1.94 | -30.13 | 1.26 |
| **GABARAPL1** | 0.98±0.02 | 5.78±0.50 | -22.46±0.74 | 2.59 | -30.40 | 1.79 |
| **GABARAPL2** | 1.03±0.04 | 6.99±0.94 | -16.75±0.98 | 4.31 | -29.94 | 1.02 |

Table 8.1: ITC-derived thermodynamic parameters for the binding of the L341V p62 AIM peptide to each of the hAtg8 proteins.

The interaction between GABARAPL1 and the L341V peptide was shown to have the lowest $K_D$ value of the hAtg8 proteins which indicates the strongest binding interaction; as was the case for the WT peptide. However, GABARAPL1 was not shown to bind with a significantly higher affinity than GABARAP in these experiments. LC3C also remained the protein with the weakest binding interaction with the peptide after the mutation. A two-fold difference in binding affinity was determined between the GABARAPL1 and LC3C interactions with the L341V peptide.

The general trend observed from this data was a reduction in binding affinity for the interactions of the hAtg8 proteins with the L341V p62 AIM peptide compared to the WT. The two exceptions to this were for LC3C which produced a 1.5-fold stronger binding interaction to the L341V peptide and GABARAPL2 which did not significantly change. The ITC-derived $K_D$ value for the interaction between LC3B and the p62 AIM was the most strongly affected by this mutation.

The thermodynamic parameters produced by the fitting of this data for each hAtg8 protein are shown in Figure 8.10. This shows the relative contributions of the enthalpic ($\Delta H$) and entropic (-T$\Delta S$) components to the overall binding ($\Delta G$) of these proteins. Both the enthalpic and entropic terms were favourable, to varying degrees, for all six interactions studied.

All of the interactions were enthalpically driven, with the interaction of LC3A producing the highest enthalpic contribution in its interaction with the peptide. LC3A also had the smallest entropic contribution to its overall affinity. Although enthalpically driven, the entropic components for LC3B and GABARAPL2 were almost the same magnitude as their enthalpy components, with LC3B having the largest entropy component of the hAtg8 proteins.

## 8.3 Discussion

### Biophysical Study of the L341V p62 AIM-hAtg8 Interaction

The ESI MS competition experiments for the individual hAtg8 proteins indicated that LC3A, LC3B, GABARAP and GABARAPL1 show a binding preference for the WT peptide over the L341V peptide by varying factors up to a twofold preference. LC3C and GABARAPL2 did not show a clear preference for either peptide at the concentrations investigated.

As well as the effect on the specific hAtg8 proteins from the L341V mutation, the relative binding selectivity was also investigated. The LC3 subfamily titration

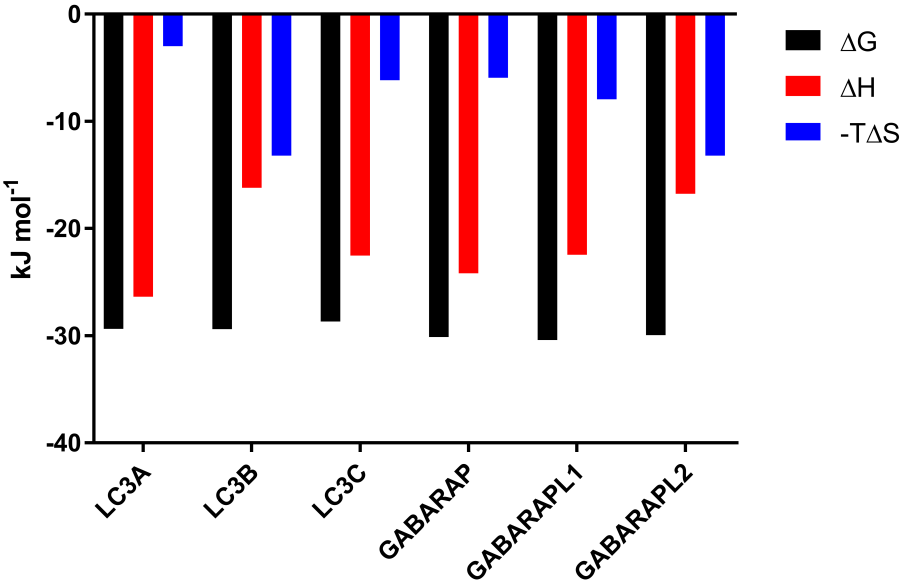



Figure 8.10: Graph comparing ITC-derived thermodynamic parameters for the binding of the L341V p62 AIM peptide to each of the hAtg8 proteins; to illustrate the differences in the enthalpic and entropic contributions to the Gibbs free energy change.

did not distinguish between LC3A and LC3B in their affinity for the L341V peptide however it did indicate that LC3C was bound with the lowest affinity. At the 20 $\mu$M titration point, there was a fourfold difference between the ratios of LC3A/LC3B and LC3C. The GABARAP subfamily titration indicated the following order of binding preference: GABARAPL1 $\geq$ GABARAP $>$ GABARAPL2. The difference between GABARAP and GABARAPL1 was not significant however the ratios for these two proteins were twofold higher than for GABARAPL2. Finally the L341V peptide was shown to preferentially bind to GABARAP over LC3A with a 1.5-fold difference in the bound ratios. These ESI-MS competition experiments therefore indicated that the L341V p62 AIM peptide used in this work preferentially bound to GABARAPL1 amongst all six hAtg8 proteins.

This ESI MS data suggested the following order of binding preference: GABARAP-L1 $>$ GABARAP $>$ LC3A $\sim$ LC3B $>$ LC3C. GABARAPL2 was shown to have a lower affinity than GABARAP and GABARAPL1 however its affinity in relation to the LC3 subfamily was not determined in these MS experiments. In both subfamilies, the ratios of bound to free protein had more similar values when binding to the L341V peptide than for the WT which implies a reduction in binding selectivity.

Inspection of the fitted 'n' values was used to initially determine the quality of

fitting for the ITC data. These interactions are expected to occur at a ratio of 1:1 based on the structure of the proteins and the absence of any multicomponent complexes observed in the MS experiments. All six fits have produced 'n' values within 20% of the expected value which supported the suitability of the fit.

As discussed in Section 4.3, the fitting procedure could be carried out with the 'n' value fixed to one but allowing the 'n' value to float appeared to produce the best results in this study based upon the lower $\chi^2$/DoF values produced. However, one point which should be noted is that fixing the n value to one produced lower $K_D$ values if the 'n' value was previously fitted below one and vice versa. This could be relevant when considering the $K_D$ values of LC3B and GABARAP as they produced fitted 'n' values the furthest from one.

Binding affinities determined from this ITC data were in agreement with the ordering of selectivity produced by the ESI MS competition experiments. There was significant overlap in the margins of error determined for the ITC-derived $K_D$ values which prevented many clear conclusions about selectivity from being drawn. The $K_D$ value of GABARAPL1 binding to the L341V peptide was not shown to be significantly lower than the value for GABARAP. GABARAPL1 was shown to have a significantly stronger binding interaction for the L341V peptide than the other four hAtg8 proteins by this method. LC3C was also significantly shown to have the lowest binding affinity.

Taken together the MS and ITC data indicate the following order of binding preference: GABARAPL1 > GABARAP > GABARAPL2 > LC3B > LC3A > LC3C. However, the differences between some of these have not been shown with statistical significance so this is therefore an estimate guided by both sets of experiments.

A previous investigation into LC3B binding to an L341V peptide produced data that indicated a three-fold reduction in binding affinity for the L341V peptide compared to the WT. This is close to the value of 2.21 suggested here which lends some validity to the results posited here. The difference in the magnitude of preference may possibly be attributed to the longer peptide construct used in that previous work, although it may simply be from errors inherent in determination of binding affinity by ITC.

The data presented here shows that the L341V mutation leads to an overall reduction in binding affinity between the p62 AIM and the six hAtg8 proteins. This can be shown by averaging the affinity change factor for each of the six proteins shown in Table 8.1. This averaging gives a factor of 1.48 times reduction in binding affinity for the L341V peptide compared to the WT between all six hAtg8 proteins.

## Structural Explanations of Binding Selectivity

Calculated thermodynamic properties of these interactions may aid to some extent in the understanding of this shift in binding. The clear shift in contribution of the thermodynamic parameters to the overall affinity observed for LC3A and LC3B suggests a significant shift in how these proteins interact with the L341V peptide compared to the WT. For both, there is a substantial decrease in the enthalpic contribution which leads to the decrease in affinity despite some compensation by the more favourable entropic component.

The change in binding mode for LC3B has been previously suggested through NMR chemical shift mapping which showed differences in chemical shift perturbations for amino acids at the L-site on the protein surface. This change of binding mode implied a more substantial reorganisation of the LC3B upon binding and the similarities between LC3A and LC3B in terms of sequence and thermodynamic properties may indicate that this is also the case for LC3A. This indicates that the valine cannot pack as tightly into the L-site resulting in a weaker binding enthalpy; with some entropy compensation. The altered binding mode induced by the L341V mutation may increase steric strain within the molecules which would also reduce the magnititude of $\Delta H$.

Interestingly, there is no large change in thermodynamic parameters for GABARAP and GABARAPL1 but just a small reduction in both the enthalpic and entropic terms. This may indicate that these proteins are still binding at the same site for the L341V peptide as they do for the WT, just with a reduction in the strength of the binding due to small changes in the steric interactions.

So despite significant reductions in affinity from the same mutation, the change in affinity for LC3A/LC3B and GABARAP/GABARAPL1 may be mediated by differing mechanisms. Figure 8.11 shows the differences in binding interface between LC3B and GABARAPL1. This highlights some clear differences which suggest that different binding modes would be possible for these proteins and that mutations to the peptide binding to the surface may be expected to affect them differently.

Based on the ITC-derived binding affinities, LC3C preferentially binds to the L341V p62 AIM peptide over the WT. This is driven by a significantly larger enthalpic contribution to the overall binding though there is also a less favourable entropic contribution. The preference of LC3C for valine over leucine at the hydrophobic fourth position in the p62 AIM can be rationalised by considering the non-canonical AIM of NDP52 which preferentially binds to LC3C. NDP52 binds to LC3C via an

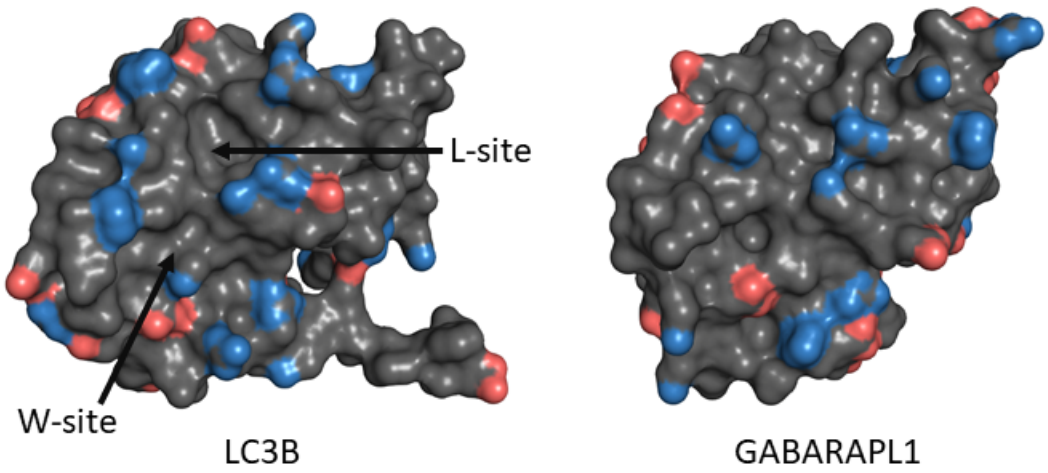



Figure 8.11: Aligned crystal structures of LC3B and GABARAPL1 showing their differing surface topology and charge around the AIM binding site. Negatively charged atoms are shown in red and positively charged atoms are shown in blue. The known W- and L-sites on LC3B are also highlighted to indicate the binding patch of this protein and therefore the expected binding patch of GABARAPL1.

Ile-Leu-Val-Val motif rather than the Trp-Thr-His-Leu sequence of the p62 AIM so the L341V mutation makes the p62 AIM more sequentially similar to the NDP52 non-canonical AIM.

Figure 8.12 shows the crystal structure of LC3C in complex with a peptide corresponding to the NDP52 AIM. An alignment of this structure with the LC3B-p62 AIM peptide structure has also been included to show a theoretical binding site for the WT p62 AIM on LC3C by analogy to LC3B. Both peptides adopt a $\beta$-strand structure at the same position on the surface in order to form intramolecular $\beta$-sheet interactions with the $\beta$2-strand of the hAtg8 protein. The peptides then diverge in their binding away from this point to the sites on the surface which best accomodate the side chains of their specific residues.

This modelled alignment results in steric clash between Leu341 of the p62 AIM and the surface of LC3C so clearly the peptide must bind in a different conformation. It is possible that mutation of Leu341 of p62 to a Val will allow this residue to bind into the pocket on LC3C in the same way as the Val at the fourth position on the NDP52 AIM, as shown in the structure. The proposed binding of the p62 AIM Val341 into the same binding pocket as Val136 of NDP52 is supported by the increased enthalpic and decreased entropic contribution to the binding affinity as this would occur due to a tighter hydrophobic binding which limits conformational flexibility.

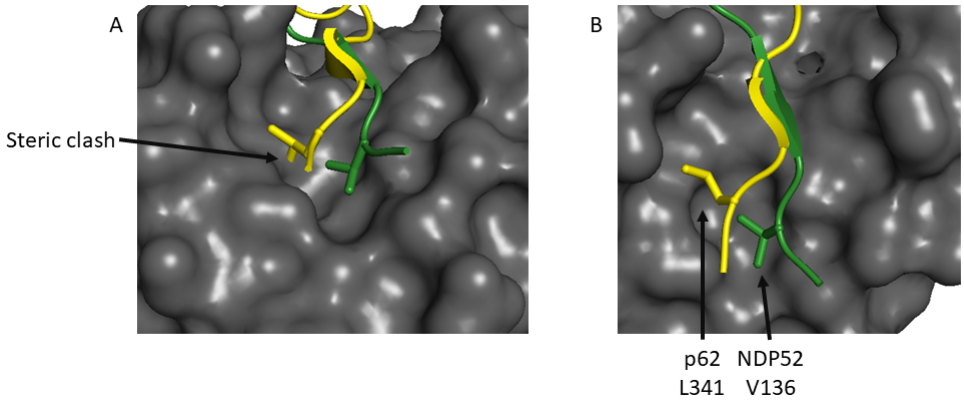



Figure 8.12: Crystal structure of LC3C (grey) in complex with NDP52 (green) with the LC3B-p62 AIM complex (yellow) aligned, shown at two different angles. The peptides are shown in cartoon representation other than the L341 of p62 and V136 of NDP52 which have side chains shown in stick representations. This shows the p62 AIM sterically clashing with the LC3C surface (A) and suggests that the L341V p62 AIM mutation may allow the peptide to bind in a similar mode to the NDP52 AIM.

ESI MS and ITC indicated that GABARAPL2 had no preference in binding between the WT or L341V peptide. There is a switch in thermodynamic contributions between the two interactions though as the interaction was entropically driven to the WT but it is enthalpically driven for the L341V. There is some unknown entropy/enthalpy compensation effect from this mutation that results in the same binding affinity for either peptide. Structural study of GABARAPL2 binding to the WT p62 AIM peptide in Chapter 5 was relatively unsuccessful which prevented that work being extended to study the structural cause of this entropy/enthalpy compensation.

In addition to the change in binding affinity, the data also indicated a reduction in binding selectivity for the L341V peptide amongst the hAtg8 proteins compared to the WT. This was observed as the binding curves in the subfamily ESI MS competition experiments were less differentiated than for the WT peptide.

Furthermore, the range of binding affinities observed decreased from 13.61 $\mu$M for the WT to 5.62 $\mu$M for the L341V peptide. There was also a fivefold difference in $K_D$ between the weakest and strongest binders for the WT and only a twofold difference for the L341V. This is made clear by the inability to distinguish significant differences between most of the hAtg8 binding affinities for the L341V peptide.

The reduction in selectivity observed here is mediated by the abbrogation of the strongest interactions and the increase in binding affinity for the weakest interaction. There are two subtle but interesting differences in the results for the LC3 subfamily

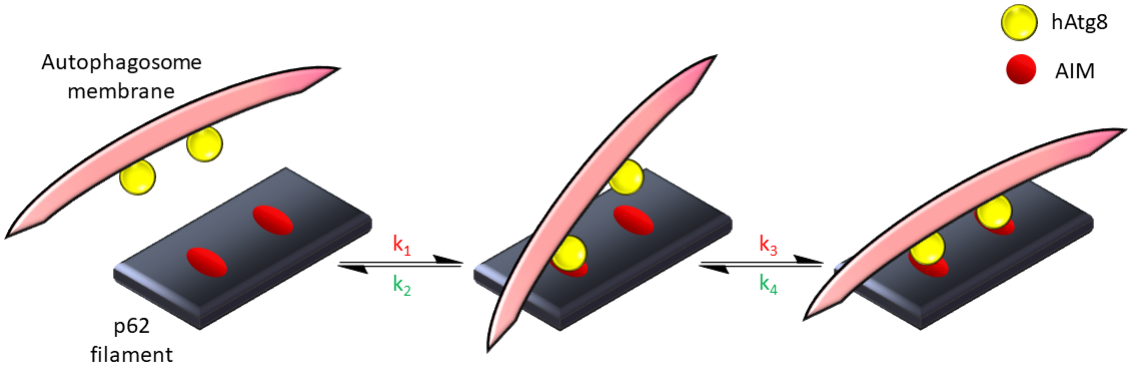



Figure 8.13: Simplified schematic to show avid binding of hAtg8 proteins on the autophagosome membrane to p62 filaments. This illustrates the initial binding of one hAtg8 protein to one AIM on the p62 filament. This then allows avid binding of the second hAtg8 protein to another AIM on the filament. This second interaction has a higher binding affinity than the first interaction due to the increased local concentration of binding sites from being held in proximity to the p62. Rate constants which are decreased by the L341V p62 mutation are shown in red and those increased by the mutation are shown in green.

ESI MS competition experiments that show this. Firstly, there appeared to be some selectivity for LC3B over LC3A in the WT experiments whereas for the L341V results there appeared to be no significant difference in binding. Secondly, LC3C binds with a higher relative affinity, compared to the other LC3 subfamily proteins, to the L341V peptide than to the WT. The differences in binding selectivity for the GABARAP subfamily from these experiments are much more subtle and less clear than for the LC3 subfamily.

ITC data may then be considered which indicated that LC3A, LC3B and GABARAPL1 had the three highest binding affinities to the WT peptide and these are those most reduced by the L341V mutation. LC3C produced the weakest binding affinity to the WT peptide but its binding affinity increased by a factor of 1.48 upon the introduction of the L341V mutation.

## Biological Context of the ALS-related L341V p62 mutation

These changes to binding affinity and selectivity must be considered in their biological context whereby multiple hAtg8 proteins can interact with a single p62 filament which allows for avidity effects. The decrease in affinity of a single interaction can be thought of as decreasing avid binding in two ways. Either the p62 filament is less

likely to bind to the first hAtg8 which is the first step required for avid interactions or when this first interaction does occur it will dissociate at a faster rate which gives less time for a subsequent binding event to occur. Some combination of these two effects will occur to reduce avid binding.

This can be thought of in terms of the rate constants shown in Figure 8.13 which gives a simplified schematic of avid binding. This decrease in affinity results in a lower $k_1$ rate constant which means the protein is less likely to bind in the first place which is required for avidity. In Section 4.3 a discussion was presented regarding how avid interactions led to increased residence times especially in the case where $k_3$ is faster than $k_2$. The decease in interaction affinity due to the L341V mutation studied here is expected to decrease $k_3$ and increase $k_2$ which will reduce residency time.

Therefore, a small shift in affinity may be amplified to a more significant effect through avidity. It is presumed that this will result in a reduction in selective recruitment of cargo to the autophagosome through this pathway, which will presumably decrease autophagic flux. The overall reduction in p62 degradation by the lysosome has previously been shown in cell based assays and this was assumed to indicate a reduction in recruitment of p62 and therefore cargo to the autophagosome.

The role of p62 as a scaffold to the expanding autophagosome membrane should also be considered. The genetic knockout of hAtg8 proteins has been shown to lead to smaller autophagosomes and this may be because of abbrogated recruitment of the p62 scaffold [44]. The reduction in hAtg8-p62 binding may also cause a similar effect though this has not been shown.

The data presented here shows the contribution from each individual hAtg8 protein to the overall effect observed in cells. This can be clearly associated with the symptoms of ALS as this mutation will result in a reduction in aggrephagy which would allow the inclusions associated with the disease to accumulate and damage the motor neurons through the various possible mechanisms discussed in Section 1.6.

GABARAPL1 is the hAtg8 protein which is most highly expressed within the central nervous system of humans and rats (CNS) [193, 194]. This along with the binding preference of the p62 AIM for this homologue may suggest a key role for the protein in selective autophagy in neurons. Therefore the affinity of this GABARAPL1-p62 AIM interaction decreasing by a factor of almost a half could clearly have a strong impact on dysregulation of autophagy in the CNS which is associated with ALS.

Further dysregulation of autophagy likely occurs through the reduction in binding selectivity produced by this mutation. The effect of this reduction in selectivity has not been investigated here but it is possible to speculate how this may impact the

overall process of autophagy. Binding selectivity of these proteins to p62 and to the other autophagy receptors and adapters is assumed to have evolved to produce an equilibrium of binding interactions which produces the appropriate rate of basal autophagic flux. The complexity of this system suggests that changes may perturb the equilibrium of the process.

For example, the increased relative binding of LC3C to the L341V p62 AIM may result in a lower concentration of free LC3C on the membrane which would preclude this protein from interactions with other proteins which could mediate other specialised roles.

The data presented in this chapter has shown that:

- The L341V p62 mutation resulted in an overall reduction in affinity of p62 for the hAtg8 family but there are differences in the effect for each protein.
- The mutation also produced an overall reduction in binding selectivity of p62 amongst the hAtg8 proteins.
- These two effects can be clearly linked to dysregulation of autophagy in ALS and it has been suggested how disease progression may be affected in patients with this mutation.

# Chapter 9

# Discussion

## 9.1 Characterisation of the hAtg8-WT p62 AIM interaction

The work presented in this thesis has further characterised the protein-protein interaction between p62 and the six hAtg8 proteins through a range of biophysical techniques. Binding selectivity has been characterised for the WT p62 AIM peptide amongst the six hAtg8 proteins using ESI MS and ITC. The ITC-derived $K_D$ values for these interactions, which were supported by the MS data, are shown in Table 9.1. A fivefold difference in the $K_D$ values was observed between the strongest and weakest interactions with the WT p62 AIM peptide. Although the differences in some cases are not significant, the ESI MS and ITC experiments indicated that the ordering of binding preference for the WT p62 AIM was: GABARAPL1 > LC3B > LC3A > GABARAP > GABARAPL2 > LC3C.

| | **$K_D$ ($\mu$M)** | | |
|---|---|---|---|
| **Protein** | WT | pS342 | L341V |
| LC3A | 4.55±0.30 | 2.11±0.11 | 8.70±1.84 |
| LC3B | 3.83±0.41 | 3.08±0.19 | 8.47±0.61 |
| LC3B T50E | 7.69±0.77 | - | - |
| LC3C | 16.84±4.08 | 6.33±1.67 | 11.40±0.95 |
| GABARAP | 5.20±0.31 | 1.98±0.49 | 6.41±0.52 |
| GABARAPL1 | 3.23±0.45 | 1.93±0.33 | 5.78±0.50 |
| GABARAPL2 | 6.85±1.02 | 3.82±0.32 | 6.99±0.94 |

Table 9.1: ITC-derived thermodynamic parameters for the binding of the p62 AIM peptides to each of the hAtg8 proteins.

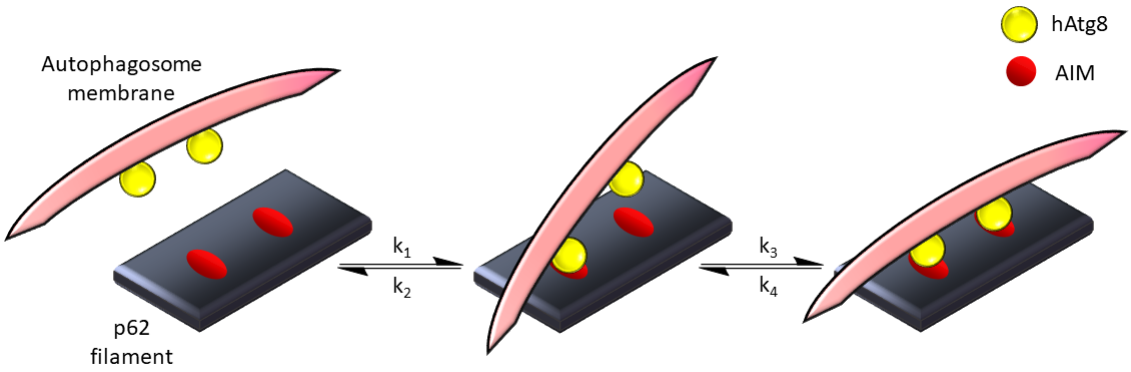



Figure 9.1: Simplified schematic to show avid binding of hAtg8 proteins on the autophagosome membrane to p62 filaments. This illustrates the initial binding of one hAtg8 protein to one AIM on the p62 filament. This then allows avid binding of the second hAtg8 protein to another AIM on the filament. This second interaction has a higher binding affinity than the first interaction due to the increased local concentration of binding sites from being held in proximity to the p62.

The relatively small degree of binding selectivity shown in these experiments is amplified *in vivo* by the multi-dentate interaction found in the biological assembly. The hAtg8 proteins on the autophagosome membrane are capable of an avid binding interaction with multiple AIM domains on a p62 filament which results in a synergistic effect that leads to longer residence times. This is illustrated in the simplified schematic in Figure 9.1. To place this avidity effect in context, the interaction of tandem ubiquitin binding domains with di-Ub can be considered. In the work of Sim *et al.* the inidividual ubiquitin interacting motifs (UIM1 and UIM2) of Rap80 bind to K63-linked di-Ub with an affinity of 230 and 470 $\mu$M, for UIM1 and UIM2 respectively [195]. The tandem-linked UIM1-UIM2 construct binds to K63-linked di-Ub with an affinity of 22 $\mu$M due to the avidity effect of this bi-dentate interaction. A related example was shown in the work of Scott *et al.* which showed a different tandem ubiquitin binding domain construct bound to monoubiquitin with a 10.9 $\mu$M affinity but bound to K48-linked di-Ub with a 1.3 $\mu$M affinity [196]. This amplification of affinity is proportional to the binding affinity so the differences in selectivity will be larger when avid binding occurs.

Whilst selectivity has been shown, even LC3C which has the lowest affinity for the AIM peptide has been shown to be bound to a measureable degree during the ESI MS competition experiments. Additionally, a fivefold difference in $K_D$ is not especially large in a biological context where significantly larger binding preferences are often observed. It is likely therefore that p62 will bind to all six hAtg8 proteins present

on the autophagosomal membrane as the process occurs *in vivo*. This will result in an overall hAtg8-p62 affinity that is modulated by the relative levels of each hAtg8 anchored to the membrane.

hAtg8 protein expression levels and the rate of the lipidation reaction will control the amount of each hAtg8 anchored to the autophagosome. Expression levels of the hAtg8 proteins has been studied and there has been selective expression based on the tissue in question however the quantification of relative lipidation rates or the resulting amount of anchored hAtg8 proteins have not yet been performed [41, 42, 43]. It may be that regulation of these processes allow for different amounts of each protein to be lipidated, which controls the overall affinity for p62 binding. This will allow spatiotemporal regulation of the levels of each hAtg8 protein on the membrane and the effect this has on autophagy. For example, more GABARAPL1 may be lipidated on the membrane in tissues where a higher level of p62 recruitment is required for its functions in membrane elongation and selective cargo recruitment. Regardless of the level of each hAtg8 protein on the autophagosome, there will be proportionately less LC3C and GABARAPL2 bound to p62 which may allow these proteins to be specialised for interactions with other autophagy receptors and adapters which may influence other steps within autophagy. Low percentages of sequence homology between these proteins and the other hAtg8 proteins also suggests specialisation. This specialisation has been shown by the discovery of a protein which selectively binds to LC3C but it has not yet been shown for GABARAPL2 [40].

GABARAPL2 presents an interesting case in its binding to the WT p62 AIM peptide due to its unique thermodynamic profile. The ITC-derived thermodynamic parameters indicated that GABARAPL2 was the only hAtg8 protein which was entropically driven in its binding and the enthalpic contribution to binding was also the lowest, as shown in Figure 9.2. This warranted further structural study by NMR that produced data which indicated a novel binding interaction; shown in Figure 9.3. The clear differences between previously detailed structures of hAtg8-AIM interactions and the binding patch defined by the chemical shift mapping NMR study in this thesis indicated that GABARAPL2 does not have an L-site analogous to that observed for the other known interactions. The GABARAPL2 L-site is a hydrophobic patch on the protein surface formed predominantly by an isoleucine sidechain, as opposed to the binding cleft observed for the other hAtg8-p62 interactions. This results in an interaction with a weaker binding affinity than for the other hAtg8-p62 AIM interactions previously studied [50, 1, 53]. Structural determination of this interaction using X-ray crystallography or further NMR experiments would be an interesting further

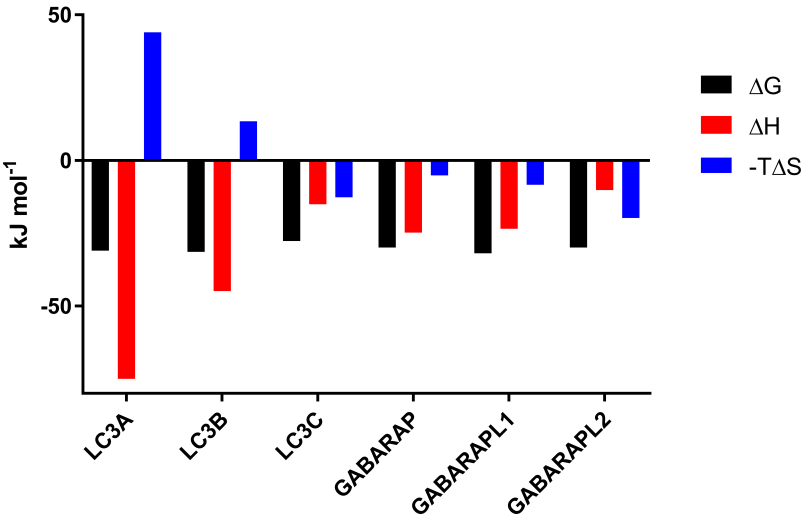



Figure 9.2: Graph comparing ITC-derived thermodynamic parameters for the binding of the WT p62 AIM peptide to each of the hAtg8 proteins to illustrate the differences in the enthalpic and entropic contributions to the Gibbs free energy.

avenue of reasearch, as would structural study of the strongest interaction of the WT p62 AIM with GABARAPL1.

## 9.2 Regulation

Phosphorylation of full length p62 has been shown to increase its affinity for LC3B via a kinase binding assay. Although phosphorylation at the AIM could not be observed due to a lack of appropriate antibodies, TBK1 target sites at S332 and S342 of p62 are known and assumed to be phosphorylated in this experiment to cause the observed effect. This effect has been shown to be caused at least in part by the previously known p62 phosphorylation site at S342 by the biophysical characterisation shown in this thesis.

The chemical shift mapping experiment for the interaction of LC3B and the pS342 p62 AIM peptide suggested a binding patch for this peptide similar to the binding of the WT but with some interesting differences. The NMR data and proposed binding site are shown in Figure 9.4. The phosphorylation results in a binding interaction that is less enthalpically driven but with a less unfavourable entropy as shown in Figure 9.5. The phosphorus on the peptide undergoes a small change in chemical environment upon binding with LC3B. As the His57 signal in the $^{1}$H-$^{15}$N-HSQC of LC3B is significantly shifted for binding to the phosphorylated peptide but not the

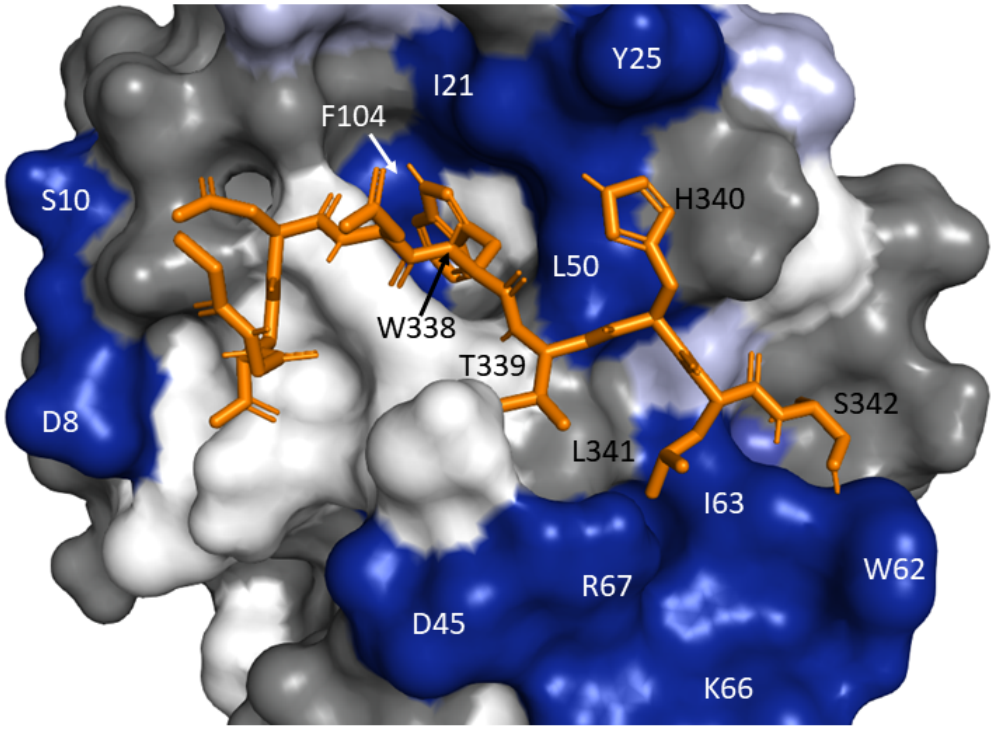



Figure 9.3: Structure of the GABARAPL2-WT p62 AIM peptide structure produced by the HADDOCK web server. This shows the peptide (orange) and the GABARAPL2 residues with significantly shifted signals (dark blue shows slow exchanging residues; medium and light blue correspond to the CSP values two and one standard deviations above the mean respectively), the residues without significantly shifted signals (grey) and the residues which were not assigned to peaks in the $^{1}$H-$^{15}$N-HSQC spectrum (white). Residues annotated in white correspond to the slow exchanging residues on GABARAPL2 and annotations in black correspond to peptide residues.

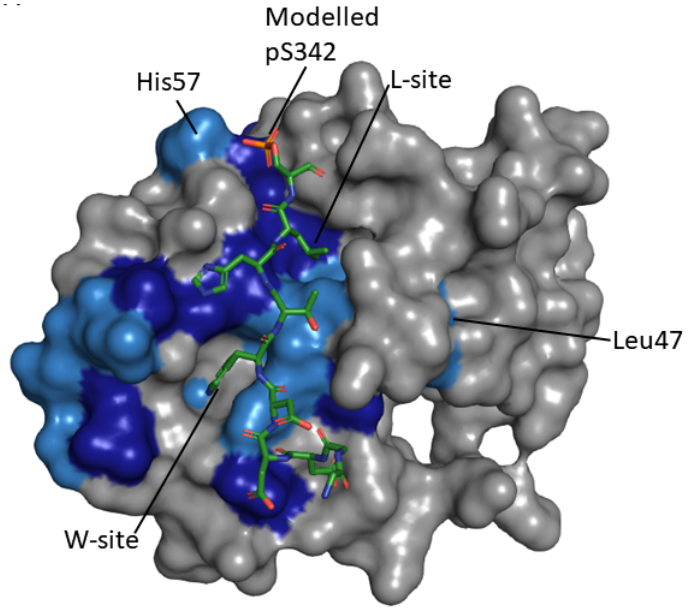



Figure 9.4: Crystal structure of LC3B which highlights the residues with backbone NH signals which were significantly perturbed in the $^{1}$H-$^{15}$N-HSQC spectrum upon addition of the pS342 p62 AIM peptide. The crystal structure of the WT p62 AIM peptide binding to LC3B has been modified to include the phosphate group on the Ser342 sidechain. This shows the residues with significantly shifted signals which could be followed through the titration (dark blue is for the most significant; light blue is for some significance).

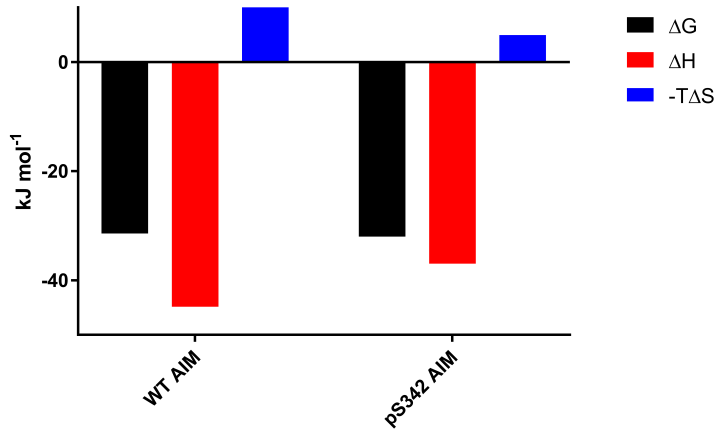



Figure 9.5: Graph comparing ITC-derived thermodynamic parameters for the binding of the WT or pS342 p62 AIM peptide to LC3B to illustrate the differences in the enthalpic and entropic contributions to the Gibbs free energy.

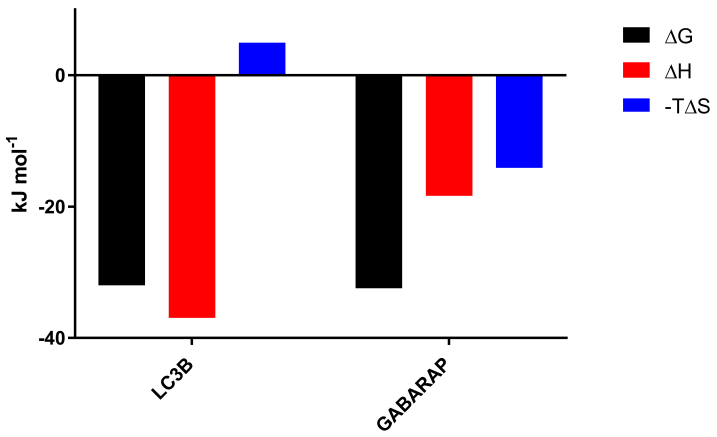



Figure 9.6: Graph comparing ITC-derived thermodynamic parameters for the binding of the pS342 p62 AIM peptide to LC3B or GABARAP to illustrate the differences in the enthalpic and entropic contributions to the Gibbs free energy.

WT, it is assumed that there is some extra electrostatic interaction between this residue and the phosphate group on the peptide. However this may be weak or transient as it does not significantly shift the signal for the phosphorus atom.

The remaining five hAtg8 proteins have also been shown to preferentially bind to the phosphorylated p62 AIM peptide over the WT by ESI MS and ITC; though to varying degrees. The ITC-derived binding affinities for these interactions are shown in Table 9.1 where they can be compared to values from binding to the WT peptide. These values indicate that the average binding affinity between the hAtg8 proteins and the p62 AIM is twice as strong upon phosphorylation at S342 which should again be considered in the context of avid binding where the differences in affinity will be amplified.

The role of the phosphate group in the binding of the pS342 p62 AIM to GABARAP was also studied using NMR. Binding to GABARAP resulted in a larger CSP of the $^{31}$P atom on the phosphorylated peptide than for LC3B which suggests different modes of binding where the phosphate is differentially perturbed. There may be some correlation between the larger CSP for GABARAP and its larger increase in binding to the phosphorylated peptide. ITC-derived thermodyanmic parameters, as shown in Figure 9.6, illustrate a clearly different profile between the two proteins. The NMR and ITC data together then indicate that despite both proteins preferentially binding to the pS342 p62 AIM peptide over the WT, this is mediated by differing binding mechanisms with the phosphate group being more involved in the binding to the GABARAP surface.

All of the data collected in relation to p62 phosphorylation suggests that TBK1 upregulation in the context of autophagy works, at least in part, by increasing the binding of p62 to the autophagosomal membrane. Increases in this interaction will increase autophagic flux by providing more p62 as a scaffold to aid in autophagosome membrane elongation and it will also increase the amount of selectively bound autophagy cargo being targeted to the membrane for degradation.

Structural characterisation of LC3B T50E indicated that the mutated protein was folded and that there were some subtle structural difference between this protein and the WT. The lineshape and intensity seen in the CD spectra of the WT and T50E proteins were similar which indicated the same proportions of secondary structural elements were present. There were some noticeable differences in the $^{1}$H NMR spectra of the two proteins and these have been attributed to small shifts in the internal packing rather than significant structural differences. As this phosphomimetic protein has been shown to perform the function of the phosphorylated protein *in vivo*, it is assumed that this structural change from the mutation also occurs due to phosphorylation [3].

The WT p62 AIM peptide has been shown, by ESI MS and ITC, to preferentially interact with WT LC3B over the phosphomimetic T50E mutant. This phosphorylation site has therefore been suggested as a biochemical switch between different roles of LC3B. Hypothetically, non-phosphorylated LC3B plays an important role in recruiting p62 to the autophagosome membrane but this interaction is abbrogated upon phosphorylation which facilitates the proteins role in autophagosome-lysosome fusion. It is unclear how this role in autophagosome-lysosome fusion is mediated as significant further work is required.

## 9.3 Dysregulation

The effect of the ALS-related L341V mutation of p62 has been studied in terms of the binding of a L341V p62 AIM peptide to the six hAtg8 proteins. Previously, this mutation has been shown to reduce the binding of a p62 AIM peptide to LC3B but this work shows a more complicated picture across the hAtg8 family. The ITC-derived $K_D$ values for these interactions, which are shown in Table 9.1 and supported by ESI MS studies, indicate that the interactions of LC3A, LC3B, GABARAP and GABARAPL1 with the p62 AIM are weakened due to this disease-related mutation. Interestingly, the interaction with GABARAPL2 is not affected by the mutation and LC3C actually preferentially binds to the L341V peptide over the WT. These effects

on the individual hAtg8 proteins result in an overall reduction in the binding affinity between the hAtg8 proteins and the p62 AIM as well as resulting in a decrease in binding selectivity. This data is supported by previous living cell-based studies where this mutation led to a decrease in p62 degradation by acidic vesicles [1]. The mutation therefore has been shown to decrease p62 recruitment to the autophagosome where it is engulfed and degraded. Therefore, as there is less p62 being degraded there will also be a reduced amount of ubiquitinated cargo bound to the p62 degraded. This results in a decrease in the selective degradation of protein aggregates and damaged organelles that results in damage to neurons which increases the likelihood of cell death. This leads to neurodegeneration that produces the symptoms associated with ALS. The cellular effect of the decrease in binding selectivity has not been demonstrated however it could clearly have some effect on the equilibrium of binding interactions involved in the process of autophagy which could also promote ALS-related neurodegeneration.

The work discussed in this thesis expands previous reasearch into the importance of Leu at the fourth AIM position to produce a higher binding affinity interaction [197, 198]. This work shows that the minimal change of the L341V p62 mutation reduces binding affinity for four of the six hAtg8 proteins. For LC3A and LC3B the large decrease in binding to the L341V peptide compared to the WT is produced by a substantial reduction in binding enthalpy and a switch from an unfavourable entropic component to a slightly favourable one. The reduction in affinity for GABARAP and GABARAPL1 in binding to the L341V peptide compared to the WT arises due to smaller changes in both the enthalpy and entropy terms. This indicates that the abbrogation in binding affinity occurs due to different binding mechanisms which is clearly rationailsed by the differences in the thermodynamics of these proteins when binding to the WT peptide. This suggests that the Leu341 of the AIM binds tightly into the LC3B (and presumably LC3A) L-site and the large enthalpic contribution is lost when the leucine is mutated to a valine. For GABARAP and GABARAPL1 though any binding of this leucine into a hydrophobic pocket is weaker so the enthalpic contribution to binding is not altered so significantly though there is still a reduction in affinity.

Despite the overall reduction in affinity caused between the hAtg8 proteins and the p62 AIM, GABARAPL2 was shown to be unaffected by this mutation. GABARAPL2 does not have a defined L-site on its surface in its unbound crystal structure and the data collected in this work indicates that Leu341 of the p62 AIM binds to a hydrophobic patch on the surface rather than a hydrophobic pocket into the globular structure, as shown in the structure produced using the HADDOCK web server,

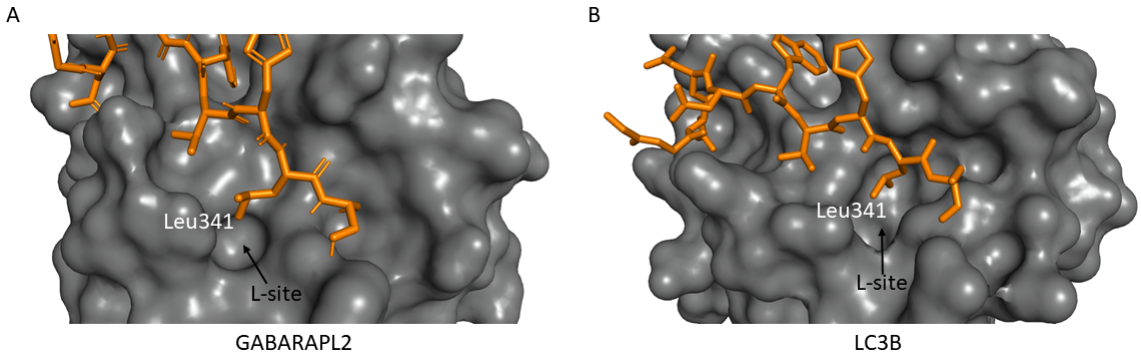



Figure 9.7: Comparison of Leu341 of the WT p62 AIM peptide binding on the surface of (A) GABARAPL2 or (B) LC3B. This indicates the Leu341 sidechain binds into a deeper hydrophobic pocket on the LC3B surface than on GABARAPL2.

based upon NMR data which is shown in Figure 9.7A. For this reason GABARAPL2 presents no binding preference for Leu or Val at this site as presumably either can bind to the surface with equal affinity. The deeper binding cleft of LC3B, shown in Figure 9.7B, which can accomodate the Leu residue leads to a stronger binding interaction but changes to this residue require significant rearrangement to accomodate binding.

This is not the only ALS-related mutation that can be associated with dysregulated autophagy. Some ALS-patients have been shown to possess a mutated form of the autophagy-related kinase TBK1 [138, 139, 140, 141]. As has been demonstrated within this work, TBK1 phosphorylation of p62 likely plays an important role in the regulation of autophagy whereby increased TBK1 activity will lead to an increase in autophagic flux. In an analysis by Oakes *et al.*, 32 of 88 ALS-related TBK1 mutations investigated were found in the kinase domain which would likely affect the role of TBK1 in autophagy but there were also loss-of-expression mutations and mutations elsewhere in the protein could also disrupt function. The study by Freischmidt *et al.* showed four of the five missense mutations tested (one from each TBK1 domain) produced functional deficits in TBK1 activity. This illustrates another avenue by which ALS-related mutations dysregulate autophagy and cause disease however further study into which TBK1 mutations are relevant to disease and how this is mediated are required. The work in this thesis could be extended in the vein of disruption of autophagy regulation to investigate how the L341V p62 mutation affects the regulatory role of phosphorylation by the study of a p62 AIM peptide with both the L341V mutation and the S342 phosphorylation and how it interacts with the hAtg8 proteins.

## 9.4 Conclusion

In conclusion, this study has produced an in-depth biophysical characterisation of the p62-hAtg8 interaction. Characterisation of the six hAtg8 proteins in addition to one hAtg8 with a phosphomimetic mutation binding to three different p62 AIM peptides has facilitated an understanding of how this important interaction is regulated and dysregulated. Specifically this has been concentrated on the role of phosphorylation and a disease-related mutation. Initial production, purification and characterisation of the required proteins confirmed samples were the correct protein and that they were folded before futher characterisation was started.

The binding selectivity of the WT p62 AIM among the six hAtg8 proteins has been elucidated, whereby the peptide representing this AIM preferentially bound to GABARAPL1 and there was a fivefold range in preference between this and LC3C which bound with the lowest affinity. Structural characterisation of the GABARAPL2-WT p62 AIM peptide indicated a similar binding site for this interaction to the previously studied hAtg8-AIM interactions with some clearly observable differences in the binding conformation.

Phosphorylation of p62 by TBK1 has been shown to increase the affinity of the protein for LC3B. A phosphorylation site of TBK1 on p62 at S342 has been studied specifically which indicated that phosphorylation there resulted in an increase in the binding affinity of the peptide for all six hAtg8 proteins, to varying degrees. The binding of the pS342 p62 AIM peptide to LC3B was further studied by NMR chemical shift mapping which indicated binding at the same patch on LC3B as for the WT however with some small differences to the binding site around the phosphorylated serine. Conversely, a phosphomimetic mutation on LC3B at T50 results in a decreased binding affinity for the p62 AIM compared to WT LC3B.

Finally, the ALS-related L341V p62 mutation when compared to the WT has been shown to result in a decrease in binding affinity for the hAtg8 proteins on average. The six hAtg8 proteins are affected differently though with LC3A, LC3B, GABARAP and GABARAPL1 producing weaker binding affinities with the L341V p62 AIM peptide compared to the WT, GABARAPL2 showing no binding preference and LC3C preferentially binding to the L341V p62 AIM peptide. These alterations to the individual hAtg8 binding affinities lead to a reduction in binding selectivity of the p62 AIM amongst the hAtg8 proteins. This reduction in both affinity and selectivity from this disease-related mutation can be clearly associated with dysregulation of autophagy and the resulting ALS-related neurodegeneration experienced by patients

with this mutation.

# Bibliography


[1] Alice Goode, Kevin Butler, Jed Long, James Cavey, Daniel Scott, Barry Shaw, Jill Sollenberger, Christopher Gell, Terje Johansen, Neil J. Oldham, Mark S. Searle, and Robert Layfield. Defective recognition of LC3b by mutant SQSTM1/p62 implicates impairment of autophagy as a pathogenic mechanism in ALS-FTLD. *Autophagy*, 12(7):1094–1104, 2016.

[2] Norma J. Greenfield. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc*, 1(6):2876–2890, 2006.

[3] Deepti S. Wilkinson, Jinel S Jariwala, Ericka Anderson, Koyel Mitra, Jill Meisenhelder, Jessica T. Chang, Trey Ideker, Tony Hunter, Victor Nizet, Andrew Dillin, and Malene Hansen. Phosphorylation of LC3 by the Hippo kinases STK3/STK4 is essential for autophagy. *Molecular cell*, 57(1):55–68, 2015.

[4] Claudine Kraft, Matthias Peter, and Kay Hofmann. Selective autophagy: ubiquitin-mediated recognition and beyond. *Nature Cell Biology*, 12(9):836–841, 2010.

[5] Libin Shang, She Chen, Fenghe Du, Shen Li, Liping Zhao, and Xiaodong Wang. Nutrient starvation elicits an acute autophagic response mediated by Ulk1 dephosphorylation and its subsequent dissociation from AMPK. *Proc Natl Acad Sci U S A*, 108(12):4788–4793, 2011.

[6] Grégory Bellot, Raquel Garcia-Medina, Pierre Gounon, Johanna Chiche, Danièle Roux, Jacques Pouysségur, and Nathalie M. Mazure. Hypoxia-Induced Autophagy Is Mediated through Hypoxia-Inducible Factor Induction of BNIP3 and BNIP3l via Their BH3 Domains. *Mol. Cell. Biol.*, 29(10):2570–2581, 2009.

[7] Monika Plescher, Aurelio A. Teleman, and Constantinos Demetriades. TSC2 mediates hyperosmotic stress-induced inactivation of mTORC1. *Scientific Reports*, 5:13828, 2015.

[8] Meghan B. Azad, Yongqiang Chen, and Spencer B. Gibson. Regulation of Autophagy by Reactive Oxygen Species (ROS): Implications for Cancer Progression and Treatment. *Antioxidants & Redox Signaling*, 11(4):777–790, 2008.

[9] Serhiy Pankiv, Terje Høyvarde Clausen, Trond Lamark, Andreas Brech, Jack-Ansgar Bruun, Heidi Outzen, Aud Øvervatn, Geir Bjørkøy, and Terje Johansen. p62/SQSTM1 binds directly to Atg8/LC3 to facilitate degradation of ubiquitinated protein aggregates by autophagy. *J. Biol. Chem.*, 282(33):24131–24145, 2007.

[10] Damián Gatica, Vikramjit Lahiri, and Daniel J. Klionsky. Cargo recognition and degradation by selective autophagy. *Nature Cell Biology*, 20(3):233–242, 2018.

[11] Terje Johansen and Trond Lamark. Selective autophagy mediated by autophagic adapter proteins. *Autophagy*, 7(3):279–296, 2011.

[12] Nobuo N. Noda, Yoshinori Ohsumi, and Fuyuhiko Inagaki. Atg8-family interacting motif crucial for selective autophagy. *FEBS Letters*, 584(7):1379–1385, 2010.

[13] Marisa Ponpuak, Michael A. Mandell, Tomonori Kimura, Santosh Chauhan, Cédric Cleyrat, and Vojo Deretic. Secretory autophagy. *Curr. Opin. Cell Biol.*, 35:106–116, 2015.

[14] Esperanza Arias and Ana Maria Cuervo. Chaperone-mediated autophagy in protein quality control. *Current Opinion in Cell Biology*, 23(2):184–189, 2011.

[15] H. L. Chiang, Terlecky, C. P. Plant, and J. F. Dice. A role for a 70-kilodalton heat shock protein in lysosomal degradation of intracellular proteins. *Science*, 246(4928):382–385, 1989.

[16] James H. Hurley and Lindsey N. Young. Mechanisms of Autophagy Initiation. *Annual Review of Biochemistry*, 2017.

[17] Elizabeth L. Axe, Simon A. Walker, Maria Manifava, Priya Chandra, H. Llewelyn Roderick, Anja Habermann, Gareth Griffiths, and Nicholas T. Ktistakis. Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum. *The Journal of Cell Biology*, 182(4):685–701, 2008.

[18] Sharon A. Tooze and Tamotsu Yoshimori. The origin of the autophagosomal membrane. *Nature Cell Biology*, 12(9):831–835, 2010.

[19] Dale W. Hailey, Angelika S. Rambold, Prasanna Satpute-Krishnan, Kasturi Mitra, Rachid Sougrat, Peter K. Kim, and Jennifer Lippincott-Schwartz. Mitochondria Supply Membranes for Autophagosome Biogenesis during Starvation. *Cell*, 141(4):656–667, 2010.

[20] Brinda Ravikumar, Kevin Moreau, Luca Jahreiss, Claudia Puri, and David C. Rubinsztein. Plasma membrane contributes to the formation of pre-autophagosomal structures. *Nature Cell Biology*, 12(8):747–757, 2010.

[21] Roland L. Knorr, Reinhard Lipowsky, and Rumiana Dimova. Autophagosome closure requires membrane scission. *Autophagy*, 11(11):2134–2137, 2015.

[22] Carla F. Bento, Claudia Puri, Kevin Moreau, and David C. Rubinsztein. The role of membrane-trafficking small GTPases in the regulation of autophagy. *J Cell Sci*, 126(5):1059–1069, 2013.

[23] Kevin Moreau, Maurizio Renna, and David C. Rubinsztein. Connections between SNAREs and autophagy. *Trends in Biochemical Sciences*, 38(2):57–63, 2013.

[24] Peidu Jiang, Taki Nishimura, Yuriko Sakamaki, Eisuke Itakura, Tomohisa Hatta, Tohru Natsume, and Noboru Mizushima. The HOPS complex mediates autophagosome–lysosome fusion through interaction with syntaxin 17. *Molecular Biology of the Cell*, 25(8):1327–1337, 2014.

[25] Noboru Mizushima, Tamotsu Yoshimori, and Yoshinori Ohsumi. The role of Atg proteins in autophagosome formation. *Annual Review of Cell and Developmental Biology*, 27:107–132, 2011.

[26] Tomer Shpilka, Hilla Weidberg, Shmuel Pietrokovski, and Zvulun Elazar. Atg8: an autophagy-related ubiquitin-like protein family. *Genome Biology*, 12(7):226, 2011.

[27] Hua He, Yongjun Dang, Fangyan Dai, Zekun Guo, Jiaxue Wu, Xinyu She, Yuan Pei, Yongjing Chen, Wenhai Ling, Chaoqun Wu, Shouyuan Zhao, Jun O. Liu, and Long Yu. Post-translational Modifications of Three Members of the Human MAP1lc3 Family and Detection of a Novel Type of Modification for MAP1lc3b. *J. Biol. Chem.*, 278(31):29278–29287, 2003.

[28] Yurong Xin, Long Yu, Zheng Chen, Lihua Zheng, Qiang Fu, Jianmin Jiang, Pingzhao Zhang, Romu Gong, and Shouyuan Zhao. Cloning, Expression Patterns, and Chromosome Localization of Three Human and Two Mouse Homologues of GABAA Receptor-Associated Protein. *Genomics*, 74(3):408–413, 2001.

[29] Hilla Weidberg, Elena Shvets, Tomer Shpilka, Frida Shimron, Vera Shinder, and Zvulun Elazar. LC3 and GATE-16/GABARAP subfamilies are both essential yet act differently in autophagosome biogenesis. *The EMBO Journal*, 29(11):1792–1802, 2010.

[30] David G. McEwan, Doris Popovic, Andrea Gubas, Seigo Terawaki, Hironori Suzuki, Daniela Stadel, Fraser P. Coxon, Diana Miranda de Stegmann, Sagar Bhogaraju, Karthik Maddi, Anja Kirchof, Evelina Gatti, Miep H. Helfrich, Soichi Wakatsuki, Christian Behrends, Philippe Pierre, and Ivan Dikic. PLEKHM1 Regulates Autophagosome-Lysosome Fusion through HOPS Complex and LC3/GABARAP Proteins. *Molecular Cell*, 57(1):39–54, 2015.

[31] Yukiko Kabeya, Noboru Mizushima, Akitsugu Yamamoto, Satsuki Oshitani-Okamoto, Yoshinori Ohsumi, and Tamotsu Yoshimori. LC3, GABARAP and GATE16 localize to autophagosomal membrane depending on form-II formation. *Journal of Cell Science*, 117(13):2805–2812, 2004.

[32] Yoshinobu Ichimura, Takayoshi Kirisako, Toshifumi Takao, Yoshinori Satomi, Yasutsugu Shimonishi, Naotada Ishihara, Noboru Mizushima, Isei Tanida, Eiki Kominami, Mariko Ohsumi, Takeshi Noda, and Yoshinori Ohsumi. A ubiquitin-like system mediates protein lipidation. *Nature*, 408(6811):488–492, 2000.

[33] Takao Hanada, Nobuo N. Noda, Yoshinori Satomi, Yoshinobu Ichimura, Yuko Fujioka, Toshifumi Takao, Fuyuhiko Inagaki, and Yoshinori Ohsumi. The Atg12-Atg5 Conjugate Has a Novel E3-like Activity for Protein Lipidation in Autophagy. *Journal of Biological Chemistry*, 282(52):37298–37302, 2007.

[34] Takayoshi Kirisako, Misuzu Baba, Naotada Ishihara, Kouichi Miyazawa, Mariko Ohsumi, Tamotsu Yoshimori, Takeshi Noda, and Yoshinori Ohsumi. Formation Process of Autophagosome Is Traced with Apg8/Aut7p in Yeast. *The Journal of Cell Biology*, 147(2):435–446, 1999.

[35] Sangeeta Nath, Julia Dancourt, Vladimir Shteyn, Gabriella Puente, Wendy M. Fong, Shanta Nag, Joerg Bewersdorf, Ai Yamamoto, Bruno Antonny, and

Thomas J. Melia. Lipidation of the LC3/GABARAP family of autophagy proteins relies upon a membrane curvature-sensing domain in Atg3. *Nature cell biology*, 16(5):415–424, 2014.

[36] Zhong-Qiu Yu, Tao Ni, Bing Hong, Hai-Yan Wang, Fen-Jun Jiang, Shenshen Zou, Yong Chen, Xi-Long Zheng, Daniel J. Klionsky, Yongheng Liang, and Zhiping Xie. Dual roles of Atg8-PE deconjugation by Atg4 in autophagy. *Autophagy*, 8(6):883–892, 2012.

[37] Fulvio Reggiori and Christian Ungermann. Autophagosome Maturation and Fusion. *Journal of Molecular Biology*, 429(4):486–496, 2017.

[38] Hallvard L. Olsvik, Trond Lamark, Kenji Takagi, Kenneth Bowitz Larsen, Gry Evjen, Aud Øvervatn, Tsunehiro Mizushima, and Terje Johansen. FYCO1 Contains a C-terminally Extended, LC3a/B-preferring LC3-interacting Region (LIR) Motif Required for Efficient Maturation of Autophagosomes during Basal Autophagy. *J. Biol. Chem.*, 290(49):29361–29374, 2015.

[39] Justin Joachim and Sharon A. Tooze. GABARAP activates ULK1 and traffics from the centrosome dependent on Golgi partners WAC and GOLGA2/GM130. *Autophagy*, 12(5):892–893, 2016.

[40] Natalia von Muhlinen, Masato Akutsu, Benjamin J. Ravenhill, Ágnes Foeglein, Stuart Bloor, Trevor J. Rutherford, Stefan M.V. Freund, David Komander, and Felix Randow. LC3c, Bound Selectively by a Noncanonical LIR Motif in NDP52, Is Required for Antibacterial Autophagy. *Molecular Cell*, 48(3):329–342, 2012.

[41] Fabrice Tolle, Pierre-Yves Risold, Virginie Mansuy-Schlick, Emilie Rossi, Michaël Boyer-Guittaut, Annick Fraichard, and Michèle Jouvenot. Specific regional distribution of gec1 mRNAs in adult rat central nervous system. *Brain Research*, 1210:103–115, 2008.

[42] Y. Sagiv, A. Legesse-Miller, A. Porat, and Z. Elazar. GATE-16, a membrane transport modulator, interacts with NSF and the Golgi v-SNARE GOS-28. *The EMBO journal*, 19(7):1494–1504, 2000.

[43] Yuchen Feng, Zhiyuan Yao, and Daniel J. Klionsky. How to control self-digestion: transcriptional, post-transcriptional, and post-translational regulation of autophagy. *Trends in Cell Biology*, 25(6):354–363, 2015.

[44] Thanh Ngoc Nguyen, Benjamin Scott Padman, Joanne Usher, Viola Oorschot, Georg Ramm, and Michael Lazarou. Atg8 family LC3/GABARAP proteins are crucial for autophagosome–lysosome fusion but not autophagosome formation during PINK1/Parkin mitophagy and starvation. *J Cell Biol*, 2016.

[45] Charleen T. Chu et al. Cardiolipin externalization to the outer mitochondrial membrane acts as an elimination signal for mitophagy in neuronal cells. *Nature Cell Biology*, 15(10):1197–1205, 2013.

[46] Benjamin Richter, Danielle A. Sliter, Lina Herhaus, Alexandra Stolz, Chunxin Wang, Petra Beli, Gabriele Zaffagnini, Philipp Wild, Sascha Martens, Sebastian A. Wagner, Richard J. Youle, and Ivan Dikic. Phosphorylation of OPTN by TBK1 enhances its binding to Ub chains and promotes selective autophagy of damaged mitochondria. *Proceedings of the National Academy of Sciences*, 113(15):4039–4044, 2016.

[47] M. Lamar Seibenhener, Jeganathan Ramesh Babu, Thangiah Geetha, Hing C. Wong, N. Rama Krishna, and Marie W. Wooten. Sequestosome 1/p62 Is a Polyubiquitin Chain Binding Protein Involved in Ubiquitin Proteasome Degradation. *Molecular and Cellular Biology*, 24(18):8055–8068, 2004.

[48] Åsa Birna Birgisdottir, Trond Lamark, and Terje Johansen. The LIR motif - crucial for selective autophagy. *Journal of Cell Science*, 126(Pt 15):3237–3247, 2013.

[49] Yoshinobu Ichimura, Taichi Kumanomidou, Yu-shin Sou, Tsunehiro Mizushima, Junji Ezaki, Takashi Ueno, Eiki Kominami, Takashi Yamane, Keiji Tanaka, and Masaaki Komatsu. Structural basis for sorting mechanism of p62 in selective autophagy. *The Journal of Biological Chemistry*, 283(33):22847–22857, 2008.

[50] Alexis Rozenknop, Vladimir V. Rogov, Natalia Yu Rogova, Frank Löhr, Peter Güntert, Ivan Dikic, and Volker Dötsch. Characterization of the interaction of GABARAPL-1 with the LIR motif of NBR1. *Journal of Molecular Biology*, 410(3):477–487, 2011.

[51] Vladimir V. Rogov, Alexandra Stolz, Arvind C. Ravichandran, Diana O. Rios-Szwed, Hironori Suzuki, Andreas Kniss, Frank Löhr, Soichi Wakatsuki, Volker Dötsch, Ivan Dikic, Renwick CJ Dobson, and David G. McEwan. Structural and functional analysis of the GABARAP interaction motif (GIM). *EMBO reports*, 18(8):1382–1396, 2017.

[52] Endalkachew Ashenafi Alemu, Trond Lamark, Knut Martin Torgersen, Aasa Birna Birgisdottir, Kenneth Bowitz Larsen, Ashish Jain, Hallvard Olsvik, Aud Øvervatn, Vladimir Kirkin, and Terje Johansen. ATG8 Family Proteins Act as Scaffolds for Assembly of the ULK Complex. *J. Biol. Chem.*, 287(47):39275–39290, 2012.

[53] Y Thielmann, OH Weiergraber, J Mohrluder, and D Willbold. Structural framework of the GABARAP-calreticulin interface-implications for substrate binding to endoplasmic reticulum chaperones. *FEBS J.*, 276(4):1140–1152, 2009.

[54] Jorge Moscat and Maria T. Diaz-Meco. p62 at the Crossroads of Autophagy, Apoptosis, and Cancer. *Cell*, 137(6):1001–1004, 2009.

[55] Jorge Moscat and Maria T. Diaz-Meco. p62: a versatile multitasker takes on cancer. *Trends in biochemical sciences*, 37(6):230–236, 2012.

[56] Jeanne M. M. Tan, Esther S. P. Wong, Donald S. Kirkpatrick, Olga Pletnikova, Han Seok Ko, Shiam-Peng Tay, Michelle W. L. Ho, Juan Troncoso, Steven P. Gygi, Michael K. Lee, Valina L. Dawson, Ted M. Dawson, and Kah-Leong Lim. Lysine 63-linked ubiquitination promotes the formation and autophagic clearance of protein inclusions associated with neurodegenerative diseases. *Hum Mol Genet*, 17(3):431–439, 2008.

[57] Jed Long, Thomas P. Garner, Maya J. Pandya, C. Jeremy Craven, Ping Chen, Barry Shaw, Michael P. Williamson, Robert Layfield, and Mark S. Searle. Dimerisation of the UBA domain of p62 inhibits ubiquitin binding and regulates NF-kappaB signalling. *Journal of Molecular Biology*, 396(1):178–194, 2010.

[58] Tomohide Saio, Masashi Yokochi, and Fuyuhiko Inagaki. The NMR structure of the p62 PB1 domain, a key protein in autophagy and NF-$\kappa$B signaling pathway. *Journal of Biomolecular NMR*, 45(3):335, 2009.

[59] Rodolfo Ciuffa, Trond Lamark, Abul K. Tarafder, Audrey Guesdon, Sofia Rybina, Wim J. H. Hagen, Terje Johansen, and Carsten Sachse. The Selective Autophagy Receptor p62 Forms a Flexible Filamentous Helical Scaffold. *Cell Reports*, 11(5):748–758, 2015.

[60] Okatsu Kei, Saisho Keiko, Shimanuki Midori, Nakada Kazuto, Shitara Hiroshi, Sou Yu-shin, Kimura Mayumi, Sato Shigeto, Hattori Nobutaka, Komatsu

Masaaki, Tanaka Keiji, and Matsuda Noriyuki. p62/SQSTM1 cooperates with Parkin for perinuclear clustering of depolarized mitochondria. *Genes to Cells*, 15(8):887–900, 2010.

[61] Yvette C. Wong and Erika L. F. Holzbaur. Optineurin is an autophagy receptor for damaged mitochondria in parkin-mediated mitophagy that is disrupted by an ALS-linked mutation. *PNAS*, 111(42):E4439–E4448, 2014.

[62] Yasemin Sancak, Liron Bar-Peled, Roberto Zoncu, Andrew L. Markhard, Shigeyuki Nada, and David M. Sabatini. Ragulator-Rag Complex Targets mTORC1 to the Lysosomal Surface and Is Necessary for Its Activation by Amino Acids. *Cell*, 141(2):290–303, 2010.

[63] Angeles Duran, Ramars Amanchy, Juan F. Linares, Jayashree Joshi, Shadi Abu-Baker, Aleksey Porollo, Malene Hansen, Jorge Moscat, and Maria T. Diaz-Meco. p62 Is a Key Regulator of Nutrient Sensing in the mTORC1 Pathway. *Molecular Cell*, 44(1):134–146, 2011.

[64] Juan F. Linares, Angeles Duran, Tomoko Yajima, Manolis Pasparakis, Jorge Moscat, and Maria T. Diaz-Meco. K63 Polyubiquitination and Activation of mTOR by the p62-TRAF6 Complex in Nutrient-Activated Cells. *Molecular Cell*, 51(3):283–296, 2013.

[65] Taguchi Keiko, Motohashi Hozumi, and Yamamoto Masayuki. Molecular mechanisms of the Keap1–Nrf2 pathway in stress response and cancer evolution. *Genes to Cells*, 16(2):123–140, 2011.

[66] Alexandria Lau, Xiao-Jun Wang, Fei Zhao, Nicole F. Villeneuve, Tongde Wu, Tao Jiang, Zheng Sun, Eileen White, and Donna D. Zhang. A Noncanonical Mechanism of Nrf2 Activation by Autophagy Deficiency: Direct Interaction between Keap1 and p62. *Mol. Cell. Biol.*, 30(13):3275–3285, 2010.

[67] Masaaki Komatsu et al. The selective autophagy substrate p62 activates the stress responsive transcription factor Nrf2 through inactivation of Keap1. *Nature Cell Biology*, 12(3):213–223, 2010.

[68] Yoshinobu Ichimura, Satoshi Waguri, Yu-shin Sou, Shun Kageyama, Jun Hasegawa, Ryosuke Ishimura, Tetsuya Saito, Yinjie Yang, Tsuguka Kouno, Toshiaki Fukutomi, Takayuki Hoshii, Atsushi Hirao, Kenji Takagi, Tsunehiro Mizushima, Hozumi Motohashi, Myung-Shik Lee, Tamotsu Yoshimori, Keiji

Tanaka, Masayuki Yamamoto, and Masaaki Komatsu. Phosphorylation of p62 Activates the Keap1-Nrf2 Pathway during Selective Autophagy. *Molecular Cell*, 51(5):618–631, 2013.

[69] Jorge Moscat and Maria T. Diaz-Meco. Feedback on fat: p62-mTORC1-autophagy connections. *Cell*, 147(4):724–727, 2011.

[70] Angeles Duran, Juan F. Linares, Anita S. Galvez, Kathryn Wikenheiser, Juana M. Flores, Maria T. Diaz-Meco, and Jorge Moscat. The Signaling Adaptor p62 Is an Important NF-$\kappa$B Mediator in Tumorigenesis. *Cancer Cell*, 13(4):343–354, 2008.

[71] Yonathan Lissanu Deribe, Tony Pawson, and Ivan Dikic. Post-translational modifications in signal integration. *Nature Structural & Molecular Biology*, 17(6):666–672, 2010.

[72] Yangchun Xie, Rui Kang, Xiaofang Sun, Meizuo Zhong, Jin Huang, Daniel J. Klionsky, and Daolin Tang. Posttranslational modification of autophagy-related proteins in macroautophagy. *Autophagy*, 11(1):28–45, 2015.

[73] Salvatore J. Cherra, Scott M. Kulich, Guy Uechi, Manimalha Balasubramani, John Mountzouris, Billy W. Day, and Charleen T. Chu. Regulation of the autophagy protein LC3 by phosphorylation. *The Journal of Cell Biology*, 190(4):533–539, 2010.

[74] Gen Matsumoto, Koji Wada, Misako Okuno, Masaru Kurosawa, and Nobuyuki Nukina. Serine 403 Phosphorylation of p62/SQSTM1 Regulates Selective Autophagic Clearance of Ubiquitinated Proteins. *Molecular Cell*, 44(2):279–289, 2011.

[75] Gen Matsumoto, Tomomi Shimogori, Nobutaka Hattori, and Nobuyuki Nukina. TBK1 controls autophagosomal engulfment of polyubiquitinated mitochondria through p62/SQSTM1 phosphorylation. *Hum Mol Genet*, 24(15):4429–4442, 2015.

[76] Florence R. Depontieu, Jie Qian, Angela L. Zarling, Tracee L. McMiller, Theresa M. Salay, Andrew Norris, A. Michelle English, Jeffrey Shabanowitz, Victor H. Engelhard, Donald F. Hunt, and Suzanne L. Topalian. Identification

of tumor-associated, MHC class II-restricted phosphopeptides as targets for immunotherapy. *Proceedings of the National Academy of Sciences of the United States of America*, 106(29):12073–12078, 2009.

[77] Philipp Mertins et al. Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature*, 534(7605):55–62, 2016.

[78] Shenghong Yang, Yu Imamura, Russell W. Jenkins, Israel Cañadas, Shunsuke Kitajima, Amir Aref, Arthur Brannon, Eiji Oki, Adam Castoreno, Zehua Zhu, Tran Thai, Jacob Reibel, Zhirong Qian, Shuji Ogino, Kwok K. Wong, Hideo Baba, Alec C. Kimmelman, Marina Pasca Di Magliano, and David A. Barbie. Autophagy Inhibition Dysregulates TBK1 Signaling and Promotes Pancreatic Inflammation. *Cancer Immunol Res*, 4(6):520–530, 2016.

[79] Manohar Pilli, John Arko-Mensah, Marisa Ponpuak, Esteban Roberts, Sharon Master, Michael Mandell, Nicolas Dupont, Wojciech Ornatowski, Shanya Jiang, Steven Bradfute, Jack-Ansgar Bruun, Tom Egil Hansen, Terje Johansen, and Vojo Deretic. TBK-1 promotes autophagy-mediated antimicrobial defense by controlling autophagosome maturation. *Immunity*, 37(2):223–234, 2012.

[80] Philipp Wild, Hesso Farhan, David G. McEwan, Sebastian Wagner, Vladimir V. Rogov, Nathan R. Brady, Benjamin Richter, Jelena Korac, Oliver Waidmann, Chunaram Choudhary, Volker Dötsch, Dirk Bumann, and Ivan Dikic. Phosphorylation of the autophagy receptor optineurin restricts Salmonella growth. *Science (New York, N.Y.)*, 333(6039):228–233, 2011.

[81] Tiffany A. Thibaudeau, Raymond T. Anderson, and David M. Smith. A common mechanism of proteasome impairment by neurodegenerative disease-associated oligomers. *Nature Communications*, 9(1):1097, 2018.

[82] Christopher J. Cummings, Michael A. Mancini, Barbara Antalffy, Donald B. DeFranco, Harry T. Orr, and Huda Y. Zoghbi. Chaperone suppression of aggregation and altered subcellular proteasome localization imply protein misfolding in SCA1. *Nature Genetics*, 19(2):148–154, 1998.

[83] Mark Halliday et al. Partial restoration of protein synthesis rates by the small molecule ISRIB prevents neurodegeneration without pancreatic toxicity. *Cell Death & Disease*, 6(3):e1672, 2015.

[84] Masaaki Komatsu, Satoshi Waguri, Tomoki Chiba, Shigeo Murata, Jun-ichi Iwata, Isei Tanida, Takashi Ueno, Masato Koike, Yasuo Uchiyama, Eiki Kominami, and Keiji Tanaka. Loss of autophagy in the central nervous system causes neurodegeneration in mice. *Nature*, 441(7095):880–884, 2006.

[85] Taichi Hara, Kenji Nakamura, Makoto Matsui, Akitsugu Yamamoto, Yohko Nakahara, Rika Suzuki-Migishima, Minesuke Yokoyama, Kenji Mishima, Ichiro Saito, Hideyuki Okano, and Noboru Mizushima. Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. *Nature*, 441(7095):885–889, 2006.

[86] Yifeng Du, Michael C Wooten, and Marie W. Wooten. Oxidative damage to the promoter region of SQSTM1/p62 is common to neurodegenerative disease. *Neurobiol Dis*, 35(2):302–310, 2009.

[87] Babu Jeganathan Ramesh, Geetha Thangiah, and Wooten Marie W. Sequestosome 1/p62 shuttles polyubiquitinated tau for proteasomal degradation. *Journal of Neurochemistry*, 94(1):192–203, 2005.

[88] Wen-Xing Ding and Xiao-Ming Yin. Sorting, recognition and activation of the misfolded protein degradation pathways through macroautophagy and the proteasome. *Autophagy*, 4(2):141–150, 2008.

[89] David Bimston, Jaewhan Song, David Winchester, Shinichi Takayama, John C. Reed, and Richard I. Morimoto. BAG-1, a negative regulator of Hsp70 chaperone activity, uncouples nucleotide hydrolysis from substrate release. *The EMBO Journal*, 17(23):6871–6878, 1998.

[90] Serena Carra, Samuel J. Seguin, Herman Lambert, and Jacques Landry. HspB8 Chaperone Activity toward Poly(Q)-containing Proteins Depends on Its Association with Bag3, a Stimulator of Macroautophagy. *J. Biol. Chem.*, 283(3):1437–1444, 2008.

[91] Martin Gamerdinger, Parvana Hajieva, A. Murat Kaya, Uwe Wolfrum, F. Ulrich Hartl, and Christian Behl. Protein quality control during aging involves recruitment of the macroautophagy pathway by BAG3. *The EMBO Journal*, 28(7):889–901, 2009.

[92] Alessandra Del Roso, Simona Vittorini, Gabriella Cavallini, Alessio Donati, Zina Gori, Matilde Masini, Maria Pollera, and Ettore Bergamini. Ageing-related

changes in the in vivo function of rat liver macroautophagy and proteolysis. *Experimental Gerontology*, 38(5):519–527, 2003.

[93] Hitoshi Kikuchi, Akiko Furuta, Ken-ichi Nishioka, Satoshi O. Suzuki, Yusaku Nakabeppu, and Toru Iwaki. Impairment of mitochondrial DNA repair enzymes against accumulation of 8-oxo-guanine in the spinal motor neurons of amyotrophic lateral sclerosis. *Acta Neuropathol*, 103(4):408–414, 2002.

[94] Carlo Rodolfo, Silvia Campello, and Francesco Cecconi. Mitophagy in neurodegenerative diseases. *Neurochemistry International*, 2017.

[95] Wenzhang Wang, Li Li, Wen-Lang Lin, Dennis W. Dickson, Leonard Petrucelli, Teng Zhang, and Xinglong Wang. The ALS disease-associated mutant TDP-43 impairs mitochondrial dynamics and function in motor neurons. *Hum Mol Genet*, 22(23):4706–4719, 2013.

[96] Fernando Bartolome, Noemi Esteras, Angeles Martin-Requero, Claire Boutoleau-Bretonniere, Martine Vercelletto, Audrey Gabelle, Isabelle Le Ber, Tadashi Honda, Albena T. Dinkova-Kostova, John Hardy, Eva Carro, and Andrey Y. Abramov. Pathogenic p62/SQSTM1 mutations impair energy metabolism through limitation of mitochondrial substrates. *Scientific Reports*, 7(1):1666, 2017.

[97] Yipeng Wang, Marta Martinez-Vicente, Ulrike Krüger, Susmita Kaushik, Esther Wong, Eva-Maria Mandelkow, Ana Maria Cuervo, and Eckhard Mandelkow. Tau fragmentation, aggregation and clearance: the dual role of lysosomal processing. *Hum Mol Genet*, 18(21):4153–4170, 2009.

[98] Cong Zhang and Ana Maria Cuervo. Restoration of chaperone-mediated autophagy in aging liver improves cellular maintenance and hepatic function. *Nat. Med.*, 14(9):959–965, 2008.

[99] Don W. Cleveland and Jeffrey D. Rothstein. From charcot to lou gehrig: deciphering selective motor neuron death in als. *Nature Reviews Neuroscience*, 2(11):806–819, 2001.

[100] Catherine Lomen-Hoerth, Thomas Anderson, and Bruce Miller. The overlap of amyotrophic lateral sclerosis and frontotemporal dementia. *Neurology*, 59(7):1077–1079, 2002.

[101] Manuela Neumann et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*, 314(5796):130–133, 2006.

[102] Manuela Neumann, Rosa Rademakers, Sigrun Roeber, Matt Baker, Hans A. Kretzschmar, and Ian R. A. Mackenzie. A new subtype of frontotemporal lobar degeneration with FUS pathology. *Brain*, 132(Pt 11):2922–2931, 2009.

[103] Jozsef Gal, Anna-Lena Ström, Renee Kilty, Fujian Zhang, and Haining Zhu. p62 Accumulates and Enhances Aggregate Formation in Model Systems of Familial Amyotrophic Lateral Sclerosis. *J. Biol. Chem.*, 282(15):11068–11077, 2007.

[104] Yuji Mizuno, Masakuni Amari, Masamitsu Takatama, Hitoshi Aizawa, Ban Mihara, and Koichi Okamoto. Immunoreactivities of p62, an ubiqutin-binding protein, in the spinal anterior horn cells of patients with amyotrophic lateral sclerosis. *Journal of the Neurological Sciences*, 249(1):13–18, 2006.

[105] Kasey L. Jackson, Wen-Lang Lin, Sumitra Miriyala, Robert D. Dayton, Manikandan Panchatcharam, Kevin J. McCarthy, Monica Castanedes-Casey, Dennis W. Dickson, and Ronald L. Klein. p62 Pathology Model in the Rat Substantia Nigra with Filamentous Inclusions and Progressive Neurodegeneration. *PLOS ONE*, 12(1):e0169291, 2017.

[106] Cezary Wójcik, Mihiro Yano, and George N. DeMartino. RNA interference of valosin-containing protein (VCP/p97) reveals multiple cellular roles linked to ubiquitin/proteasome-dependent proteolysis. *Journal of Cell Science*, 117(Pt 2):281–292, 2004.

[107] Jean-Baptiste Chanson, Andoni Echaniz-Laguna, Thomas Vogel, Michel Mohr, Aurélien Benoilid, Georges Kaltenbach, and Michèle Kiesmann. TDP43-positive intraneuronal inclusions in a patient with motor neuron disease and Parkinson's disease. *Neurodegener Dis*, 7(4):260–264, 2010.

[108] Todd J. Cohen, Virginia M. Y. Lee, and John Q. Trojanowski. TDP-43 functions and pathogenic mechanisms implicated in TDP-43 proteinopathies. *Trends Mol Med*, 17(11):659–667, 2011.

[109] Catalina Amador-Ortiz, Wen-Lang Lin, Zeshan Ahmed, David Personett, Peter Davies, Ranjan Duara, Neill R. Graff-Radford, Michael L. Hutton, and Dennis W. Dickson. TDP-43 immunoreactivity in hippocampal sclerosis and Alzheimer's disease. *Ann. Neurol.*, 61(5):435–445, 2007.

[110] Yvonne Davidson, Hanan Amin, Thomas Kelley, Jing Shi, Jinzhou Tian, Ravindran Kumaran, Tammaryn Lashley, Andrew J. Lees, Daniel DuPlessis, David Neary, Julie Snowden, Haruhiko Akiyama, Tetsuaki Arai, Masato Hasegawa, Rina Bandopadhyay, Steve Sikkink, Stuart Pickering-Brown, and David M. A. Mann. TDP-43 in ubiquitinated inclusions in the inferior olives in frontotemporal lobar degeneration and in other neurodegenerative diseases: a degenerative process distinct from normal ageing. *Acta Neuropathol.*, 118(3):359–369, 2009.

[111] Shoichi Sasaki. Autophagy in spinal cord motor neurons in sporadic amyotrophic lateral sclerosis. *Journal of Neuropathology and Experimental Neurology*, 70(5):349–359, 2011.

[112] Kelly L. Williams, Sadaf T. Warraich, Shu Yang, Jennifer A. Solski, Ruvini Fernando, Guy A. Rouleau, Garth A. Nicholson, and Ian P. Blair. UBQLN2/ubiquilin 2 mutation and pathology in familial amyotrophic lateral sclerosis. *Neurobiol. Aging*, 33(10):2527.e3–10, 2012.

[113] Han-Xiang Deng, Hong Zhai, Eileen H. Bigio, Jianhua Yan, Faisal Fecto, Kaouther Ajroud, Manjari Mishra, Senda Ajroud-Driss, Scott Heller, Robert Sufit, Nailah Siddique, Enrico Mugnaini, and Teepu Siddique. FUS-immunoreactive inclusions are a common feature in sporadic and non-SOD1 familial amyotrophic lateral sclerosis. *Ann Neurol*, 67(6):739–748, 2010.

[114] A. D. Efimova, R. K. Ovchinnikov, A. Yu Roman, A. V. Maltsev, V. V. Grigoriev, E. A. Kovrazhkina, and V. I. Skvortsova. The FUS protein: Physiological functions and a role in amyotrophic lateral sclerosis. *Mol. Biol. (Mosk.)*, 51(3):387–399, 2017.

[115] Han-Xiang Deng, Eileen H. Bigio, Hong Zhai, Faisal Fecto, Kaouther Ajroud, Yong Shi, Jianhua Yan, Manjari Mishra, Senda Ajroud-Driss, Scott Heller, Robert Sufit, Nailah Siddique, Enrico Mugnaini, and Teepu Siddique. Differential involvement of optineurin in amyotrophic lateral sclerosis with or without SOD1 mutations. *Arch. Neurol.*, 68(8):1057–1061, 2011.

[116] Tibor Hortobágyi, Claire Troakes, Agnes L. Nishimura, Caroline Vance, John C. van Swieten, Harro Seelaar, Andrew King, Safa Al-Sarraj, Boris Rogelj, and Christopher E. Shaw. Optineurin inclusions occur in a minority of TDP-43 positive ALS and FTLD-TDP cases and are rarely observed in other neurodegenerative disorders. *Acta Neuropathol.*, 121(4):519–527, 2011.

[117] Karolina Slowicka, Lars Vereecke, and Geert van Loo. Cellular Functions of Optineurin in Health and Disease. *Trends in Immunology*, 37(9):621–633, 2016.

[118] Tetsuaki Arai, Takashi Nonaka, Masato Hasegawa, Haruhiko Akiyama, Mari Yoshida, Yosio Hashizume, Kuniaki Tsuchiya, Tatsuro Oda, and Kenji Ikeda. Neuronal and glial inclusions in frontotemporal dementia with or without motor neuron disease are immunopositive for p62. *Neuroscience Letters*, 342(1):41–44, 2003.

[119] Andrew King, Satomi Maekawa, Istvan Bodi, Claire Troakes, and Safa Al-Sarraj. Ubiquitinated, p62 immunopositive cerebellar cortical neuronal inclusions are evident across the spectrum of TDP-43 proteinopathies but are only rarely additionally immunopositive for phosphorylation-dependent TDP-43. *Neuropathology*, 31(3):239–249, 2011.

[120] Johannes Brettschneider et al. Pattern of ubiquilin pathology in ALS and FTLD indicates presence of C9orf72 hexanucleotide expansion. *Acta Neuropathol*, 123(6):825–839, 2012.

[121] Min Long, Xing Li, Li Li, Matthew Dodson, Donna D. Zhang, and Hongting Zheng. Multifunctional p62 Effects Underlie Diverse Metabolic Diseases. *Trends in Endocrinology & Metabolism*, 28(11):818–830, 2017.

[122] Katsuragi Yoshinori, Ichimura Yoshinobu, and Komatsu Masaaki. p62/SQSTM1 functions as a signaling hub and an autophagy adaptor. *The FEBS Journal*, 282(24):4672–4678, 2015.

[123] Rosie K. A. Bunton-Stasyshyn, Rachele A. Saccon, Pietro Fratta, and Elizabeth M. C. Fisher. SOD1 Function and Its Implications for Amyotrophic Lateral Sclerosis Pathology: New and Renascent Themes. *Neuroscientist*, 21(5):519–529, 2015.

[124] Dennis W. Dickson, Naomi Kouri, Melissa E. Murray, and Keith A. Josephs. Neuropathology of frontotemporal lobar degeneration-tau (FTLD-tau). *J. Mol. Neurosci.*, 45(3):384–389, 2011.

[125] Jesús Avila, José J. Lucas, Mar Pérez, and Félix Hernández. Role of Tau Protein in Both Physiological and Pathological Conditions. *Physiological Reviews*, 84(2):361–384, 2004.

[126] Chantelle F. Sephton, Basar Cenik, Bercin Kutluk Cenik, Joachim Herz, and Gang Yu. TDP-43 in CNS development and function: clues to TDP-43-associated neurodegeneration. *Biol Chem*, 393(7):589–594, 2012.

[127] Elsa-Noah N'Diaye, Kimberly K. Kajihara, Ivy Hsieh, Hiroshi Morisaki, Jayanta Debnath, and Eric J. Brown. PLIC proteins or ubiquilins regulate autophagy-dependent cell survival during nutrient starvation. *EMBO Rep.*, 10(2):173–179, 2009.

[128] Han Seok Ko, Takashi Uehara, Kazuhiro Tsuruma, and Yasuyuki Nomura. Ubiquilin interacts with ubiquitylated proteins and proteasome through its ubiquitin-associated and ubiquitin-like domains. *FEBS Lett.*, 566(1-3):110–114, 2004.

[129] Precious J. Lim, Rebecca Danner, Jing Liang, Howard Doong, Christine Harman, Deepa Srinivasan, Cara Rothenberg, Hongmin Wang, Yihong Ye, Shengyun Fang, and Mervyn J. Monteiro. Ubiquilin and p97/VCP bind erasin, forming a complex involved in ERAD. *J. Cell Biol.*, 187(2):201–217, 2009.

[130] Maurits Kleijnen et al. The hPLIC proteins may provide a link between the ubiquitination machinery and the proteasome. *Mol. Cell*, 6(2):409–419, 2000.

[131] Yuxing Xia, Linda H. Yan, Bo Huang, Mujun Liu, Xionghao Liu, and Cao Huang. Pathogenic mutation of UBQLN2 impairs its interaction with UBXD8 and disrupts endoplasmic reticulum-associated protein degradation. *J. Neurochem.*, 129(1):99–106, 2014.

[132] David Komander and Michael Rape. The Ubiquitin Code. *Annual Review of Biochemistry*, 81(1):203–229, 2012.

[133] Mariely DeJesus-Hernandez et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9orf72 causes chromosome 9p-linked FTD and ALS. *Neuron*, 72(2):245–256, 2011.

[134] Alan Renton et al. A hexanucleotide repeat expansion in C9orf72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*, 72(2):257–268, 2011.

[135] Youn-Bok Lee et al. Hexanucleotide Repeats in ALS/FTD Form Length-Dependent RNA Foci, Sequester RNA Binding Proteins, and Are Neurotoxic. *Cell Reports*, 5(5):1178–1186, 2013.

[136] Manal A. Farg, Vinod Sundaramoorthy, Jessica M. Sultana, Shu Yang, Rachel A.K. Atkinson, Vita Levina, Mark A. Halloran, Paul A. Gleeson, Ian P. Blair, Kai Y. Soo, Anna E. King, and Julie D. Atkin. C9orf72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Hum Mol Genet*, 23(13):3579–3595, 2014.

[137] Veronika Majcher, Alice Goode, Victoria James, and Robert Layfield. Autophagy receptor defects and ALS-FTLD. *Molecular and Cellular Neurosciences*, 66(Pt A):43–52, 2015.

[138] Elizabeth Cirulli et al. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science*, 347(6229):1436–1441, 2015.

[139] Axel Freischmidt et al. Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. *Nature Neuroscience*, 18(5):631–636, 2015.

[140] Genki Tohnai et al. Frequency and characteristics of the TBK1 gene variants in Japanese patients with sporadic amyotrophic lateral sclerosis. *Neurobiology of Aging*, 64:158.e15–158.e19, 2018.

[141] James Oakes, Maria Davies, and Mark Collins. TBK1: a new player in ALS linking autophagy and neuroinflammation. *Molecular Brain*, 10, 2017.

[142] Faisal Fecto, Jianhua Yan, S. Pavan Vemula, Erdong Liu, Yi Yang, Wenjie Chen, Jian Guo Zheng, Yong Shi, Nailah Siddique, Hasan Arrat, Sandra Donkervoort, Senda Ajroud-Driss, Robert L. Sufit, Scott L. Heller, Han-Xiang Deng, and Teepu Siddique. SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. *Arch. Neurol.*, 68(11):1440–1446, 2011.

[143] YongPing Chen, Zhen-Zhen Zheng, XuePing Chen, Rui Huang, Yuan Yang, LiXing Yuan, Lei Pan, Shinji Hadano, and Hui-Fang Shang. SQSTM1 mutations in Han Chinese populations with sporadic amyotrophic lateral sclerosis. *Neurobiology of Aging*, 35(3):726.e7–726.e9, 2014.

[144] Xiaojie Zhang, Liang Li, Sheng Chen, Dehua Yang, Yi Wang, Xin Zhang, Zheng Wang, and Weidong Le. Rapamycin treatment augments motor neuron degeneration in SOD1g93a mouse model of amyotrophic lateral sclerosis. *Autophagy*, 7(4):412–425, 2011.

[145] James K. Ching and Conrad C. Weihl. Rapamycin-induced autophagy aggravates pathology and weakness in a mouse model of VCP-associated myopathy. *Autophagy*, 9(5):799–800, 2013.

[146] Kim A. Staats, Sara Hernandez, Susann Schönefeldt, André Bento-Abreu, James Dooley, Philip Van Damme, Adrian Liston, Wim Robberecht, and Ludo Van Den Bosch. Rapamycin increases survival in ALS mice lacking mature lymphocytes. *Molecular Neurodegeneration*, 8:31, 2013.

[147] Francesco Fornai, Patrizia Longone, Luisa Cafaro, Olga Kastsiuchenka, Michela Ferrucci, Maria Laura Manca, Gloria Lazzeri, Alida Spalloni, Natascia Bellio, Paola Lenzi, Nicola Modugno, Gabriele Siciliano, Ciro Isidoro, Luigi Murri, Stefano Ruggieri, and Antonio Paparelli. Lithium delays progression of amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci. U.S.A.*, 105(6):2052–2057, 2008.

[148] Chiara Pizzasegola, Ilaria Caron, Cristina Daleno, Anna Ronchi, Claudio Minoia, Maria Teresa Carrì, and Caterina Bendotti. Treatment with lithium carbonate does not improve disease progression in two different strains of SOD1 mutant mice. *Amyotroph Lateral Scler*, 10(4):221–228, 2009.

[149] Lindsay A. Becker, Brenda Huang, Gregor Bieri, Rosanna Ma, David A. Knowles, Paymaan Jafar-Nejad, James Messing, Hong Joo Kim, Armand Soriano, Georg Auburger, Stefan M. Pulst, J. Paul Taylor, Frank Rigo, and Aaron D. Gitler. Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice. *Nature*, 544(7650):367–371, 2017.

[150] Dongryeol Ryu, Laurent Mouchiroud, Pénélope A. Andreux, Elena Katsyuba, Norman Moullan, Amandine A. Nicolet-dit Félix, Evan G. Williams, Pooja Jha, Giuseppe Lo Sasso, Damien Huzard, Patrick Aebischer, Carmen Sandi, Chris Rinsch, and Johan Auwerx. Urolithin A induces mitophagy and prolongs lifespan in *C. elegans* and increases muscle function in rodents. *Nature Medicine*, 22(8):879–888, 2016.

[151] Daniel A. East, Francesca Fagiani, James Crosby, Nikolaos D. Georgakopoulos, Hélène Bertrand, Marjolein Schaap, Adrian Fowkes, Geoff Wells, and Michelangelo Campanella. PMI: A $\delta\psi$m Independent Pharmacological Regulator of Mitophagy. *Chemistry & Biology*, 21(11):1585–1596, 2014.

[152] John. Fenn et al. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246(4926):64–71, 1989.

[153] Shibdas Banerjee and Shyamalava Mazumdar. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *Int J Anal Chem*, 2012, 2012.

[154] Lars Konermann et al. Acid-induced denaturation of myoglobin studied by time-resolved electrospray ionization mass spectrometry. *Biochemistry*, 36(21):6448–6454, 1997.

[155] Jürg M. Daniel, Sebastian D. Friess, Sudha Rajagopalan, Silke Wendt, and Renato Zenobi. Quantitative determination of noncovalent binding interactions using soft ionization mass spectrometry. *International Journal of Mass Spectrometry*, 216(1):1–27, 2002.

[156] Noam Kirshenbaum, Izhak Michaelevski, and Michal Sharon. Analyzing Large Protein Complexes by Structural Mass Spectrometry. *J Vis Exp*, (40), 2010.

[157] Mark C. Kuprowski, Brian L. Boys, and Lars Konermann. Analysis of Protein Mixtures by Electrospray Mass Spectrometry: Effects of Conformation and Desolvation Behavior on the Signal Intensities of Hemoglobin Subunits. *Journal of the American Society for Mass Spectrometry*, 18(7):1279–1285, 2007.

[158] Elena N. Kitova, Amr El-Hawiet, Paul D. Schnier, and John S. Klassen. Reliable Determinations of Protein-Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet? *J. Am. Soc. Mass Spectrom.*, 23(3):431–441, 2012.

[159] J. J. Christensen, R. M. Izatt, and L. D. Hansen. New Precision Thermometric Titration Calorimeter. *Review of Scientific Instruments*, 36(6):779–783, 1965.

[160] Yi Liang. Applications of isothermal titration calorimetry in protein science. *Acta Biochimica et Biophysica Sinica*, 40(7):565–576, 2008.

[161] Alan Cooper and Christopher M. Johnson. Introduction to Microcalorimetry and Biomolecular Energetics. In *Microscopy, Optical Spectroscopy, and Macroscopic Techniques*, Methods in Molecular Biology, pages 109–124. Humana Press, 1994.

[162] Xing Du, Yi Li, Yuan-Ling Xia, Shi-Meng Ai, Jing Liang, Peng Sang, Xing-Lai Ji, and Shu-Qun Liu. Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Int J Mol Sci*, 17(2), 2016.

[163] Michael M. Pierce, C. S. Raman, and Barry T. Nall. Isothermal Titration Calorimetry of Protein–Protein Interactions. *Methods*, 19(2):213–221, 1999.

[164] Martin Saunders and Arnold Wishnia. Nuclear Magnetic Resonance Spectra of Proteins. *Annals of the New York Academy of Sciences*, 70(4):870–874, 1958.

[165] Arthur Kowalsky. Nuclear Magnetic Resonance Studies of Proteins. *Journal of Biological Chemistry*, 237(6):1807–1819, 1962.

[166] Oleg Jardetzky and Christine D. Jardetzky. Proton Magnetic Resonance Spectra of Amino Acids. *J. Biol. Chem.*, 233(2):383–387, 1958.

[167] K. Nagayama, K. Wüthrich, P. Bachmann, and R. R. Ernst. Two-dimensional J-resolved 1h n.m.r. spectroscopy for studies of biological macromolecules. *Biochemical and Biophysical Research Communications*, 78(1):99–105, 1977.

[168] Gaetano T Montelione and Gerhard Wagner. Conformation-independent sequential NMR connections in isotope-enriched polypeptides by 1h?13c?15n triple-resonance experiments. *Journal of Magnetic Resonance (1969)*, 87(1):183–188, 1990.

[169] Dominique Marion, Mitsuhiko Ikura, Rolf Tschudin, and Ad Bax. Rapid recording of 2d NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *Journal of Magnetic Resonance (1969)*, 85(2):393–399, 1989.

[170] Wim F. Vranken, Wayne Boucher, Tim J. Stevens, Rasmus H. Fogh, Anne Pajon, Miguel Llinas, Eldon L. Ulrich, John L. Markley, John Ionides, and Ernest D. Laue. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*, 59(4):687–696, 2005.

[171] Peixiang Ma, Oliver Schillinger, Melanie Schwarten, Justin Lecher, Rudolf Hartmann, Matthias Stoldt, Jeannine Mohrlüder, Olujide Olubiyi, Birgit Strodel, Dieter Willbold, and Oliver H. Weiergräber. Conformational Polymorphism in Autophagy-Related Protein GATE-16. *Biochemistry*, 54(35):5469–5479, 2015.

[172] Stephan Grzesiek and Ad Bax. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *Journal of Magnetic Resonance (1969)*, 99(1):201–207, 1992.

[173] Stephan Grzesiek and Ad Bax. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *Journal of the American Chemical Society*, 114(16):6291–6293, 1992.

[174] Lewis E Kay, Mitsuhiko Ikura, Rolf Tschudin, and Ad Bax. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance (1969)*, 89(3):496–514, 1990.

[175] Robert T Clubb, V Thanabal, and Gerhard Wagner. A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue 1hn, 15n, and 13c' chemical shifts in 15n?13c-labelled proteins. *Journal of Magnetic Resonance (1969)*, 97(1):213–217, 1992.

[176] Mike P. Williamson. Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc*, 73:1–16, 2013.

[177] Bennett T. Farmer Ii, Keith L. Constantine, Valentina Goldfarb, Mark S. Friedrichs, Michael Wittekind, Joseph Yanchunas Jr, James G. Robertson, and Luciano Mueller. Localizing the NADP+ binding site on the MurB enzyme by NMR. *Nature Structural & Molecular Biology*, 3(12):995–997, 1996.

[178] Holger Lüttgen, Rudolf Robelek, René Mühlberger, Tammo Diercks, Stephan C. Schuster, Peter Köhler, Horst Kessler, Adelbert Bacher, and Gerald Richter. Transcriptional Regulation by Antitermination. Interaction of RNA with NusB Protein and NusB/NusE Protein Complex of Escherichia coli. *Journal of Molecular Biology*, 316(4):875–885, 2002.

[179] John M. Walker, editor. *The Proteomics Protocols Handbook*. Humana Press, Totowa, NJ, 2005.

[180] Michael. Karas and Franz. Hillenkamp. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.*, 60(20):2299–2301, 1988.

[181] Hoi Sung Chung, Ali Shandiz, Tobin R. Sosnick, and Andrei Tokmakoff. Probing the folding transition state of ubiquitin mutants by temperature-jump induced downhill unfolding. *Biochemistry*, 47(52):13870–13877, 2008.

[182] Valerie Wood et al. The genome sequence of Schizosaccharomyces pombe. *Nature*, 415(6874):871–880, 2002.

[183] You-Kyung Lee and Jin-A Lee. Role of the mammalian ATG8/LC3 family in autophagy: differential and compensatory roles in the spatiotemporal regulation of autophagy. *BMB Rep*, 49(8):424–430, 2016.

[184] Moran Rawet Slobodkin and Zvulun Elazar. The Atg8 family: multifunctional ubiquitin-like key regulators of autophagy. *Essays In Biochemistry*, 55:51–64, 2013.

[185] Katarzyna Zientara-Rytter and Suresh Subramani. AIM/LIR-based fluorescent sensors-new tools to monitor mAtg8 functions. *Autophagy*, pages 1–5, 2018.

[186] Samuel A. Kantonen, Niel M. Henriksen, and Michael K. Gilson. Evaluation and Minimization of Uncertainty in ITC Binding Measurements. *Biochim Biophys Acta*, 1861(2):485–498, 2017.

[187] Ivana Novak, Vladimir Kirkin, David G. McEwan, Ji Zhang, Philipp Wild, Alexis Rozenknop, Vladimir Rogov, Frank Löhr, Doris Popovic, Angelo Occhipinti, Andreas S. Reichert, Janos Terzic, Volker Dötsch, Paul A. Ney, and Ivan Dikic. Nix is a selective autophagy receptor for mitochondrial clearance. *EMBO Rep.*, 11(1):45–51, 2010.

[188] Vladimir V. Rogov, Hironori Suzuki, Mija Marinković, Verena Lang, Ryuichi Kato, Masato Kawasaki, Maja Buljubašić, Matilda Šprung, Natalia Rogova, Soichi Wakatsuki, Anne Hamacher-Brady, Volker Dötsch, Ivan Dikic, Nathan R. Brady, and Ivana Novak. Phosphorylation of the mitochondrial autophagy receptor Nix enhances its interaction with LC3 proteins. *Scientific Reports*, 7(1):1131, 2017.

[189] Zhe Sha, Helena M. Schnell, Kerstin Ruoff, and Alfred Goldberg. Rapid induction of p62 and GABARAPL1 upon proteasome inhibition promotes survival before autophagy activation. *J Cell Biol*, 217(5):1757–1776, 2018.

[190] Gydo van Zundert et al. The HADDOCK2.2 webserver: User-friendly integrative modeling of biomolecular complexes. *Journal of Molecular Biology*, 428(428):720–725, 2015.

[191] Wassenaar et al. WeNMR: Structural biology on the grid. *J. Grid. Comp.*, 10(10):743–767, 2012.

[192] Jin-Mi Heo, Alban Ordureau, Joao A. Paulo, Jesse Rinehart, and J. Wade Harper. The PINK1-PARKIN Mitochondrial Ubiquitylation Pathway Drives a Program of OPTN/NDP52 Recruitment and TBK1 Activation to Promote Mitophagy. *Mol. Cell*, 60(1):7–20, 2015.

[193] Virginie Mansuy-Schlick, Fabrice Tolle, Régis Delage-Mourroux, Annick Fraichard, Pierre-Yves Risold, and Michèle Jouvenot. Specific distribution of gabarap, gec1/gabarap Like 1, gate16/gabarap Like 2, lc3 messenger RNAs in rat brain areas by quantitative real-time PCR. *Brain Research*, 1073-1074:83–87, 2006.

[194] Christophe Nemos, Virginie Mansuy, Sandrine Vernier-Magnin, Annick Fraichard, Michèle Jouvenot, and Régis Delage-Mourroux. Expression of gec1/GABARAPL1 versus GABARAP mRNAs in human: predominance of gec1/GABARAPL1 in the central nervous system. *Molecular Brain Research*, 119(2):216–219, 2003.

[195] J.J. Sims and R.E. Cogen. Linkage-specific avidity defines the lysine 63-linked polyubiquitin binding preference of Rap80. *Molecular Cell*, 33, 2009.

[196] Daniel Scott, Tom P. Garner, Jed Long, Jo Strachan, Sharad C. Mistry, Andrew R. Bottrill, David J. Tooth, Mark S. Searle, Neil J. Oldham, and Rob Layfield. Mass spectrometry insights into a tandem ubiquitin-binding domain hybrid engineered for the selective recognition of unanchored polyubiquitin. *Proteomics*, 16(14):1961–1969, 2016.

[197] Noda et al. Structural basis of target recognition by Atg8/LC3 during selective autophagy. *Genes to Cells*, 12, 2008.

[198] Yamaguchi et al. Autophagy-related protein 8 (Atg8) family interacting motif in Atg3 mediates the Atg3-Atg8 interaction and is crucial for the cytoplasm-to-vacuole targeting pathway. *Journal of Biological Chemistry*, 285, 2010.

# Chapter 10

# Appendix

## 10.1 MS

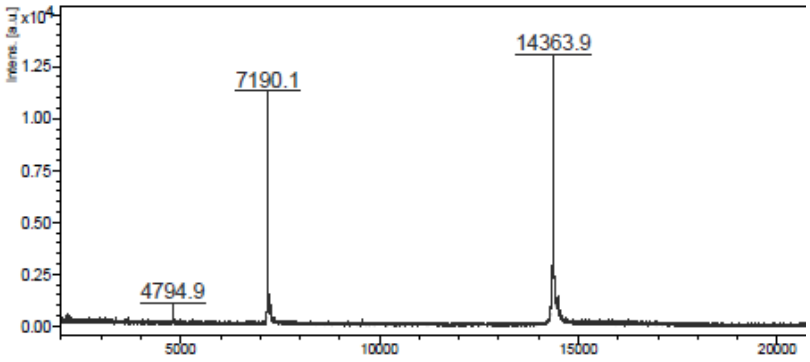



Figure A.1: MALDI TOF mass spectrum of LC3A (1 mg/ml).

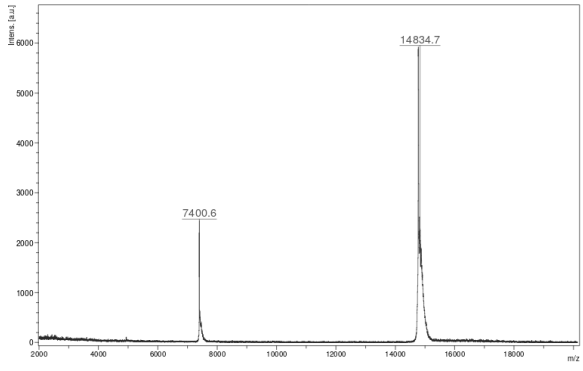



Figure A.2: MALDI TOF mass spectrum of LC3B (1 mg/ml).

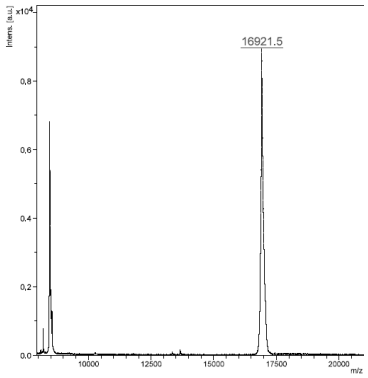



Figure A.3: MALDI TOF mass spectrum of LC3C (1 mg/ml).

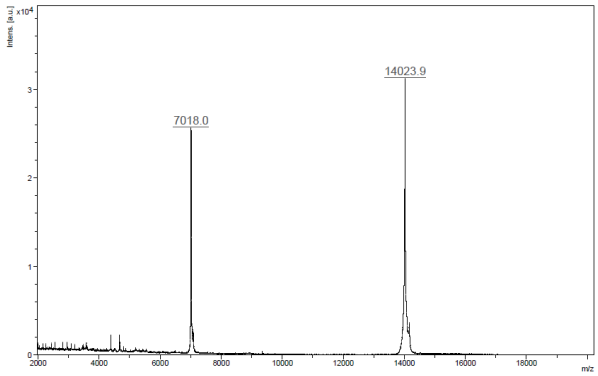



Figure A.4: MALDI TOF mass spectrum of GABARAP (1 mg/ml).

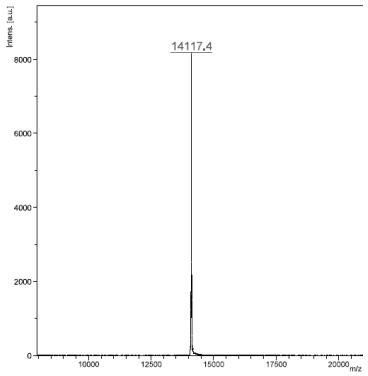



Figure A.5: MALDI TOF mass spectrum of GABARAPL1 (1 mg/ml).

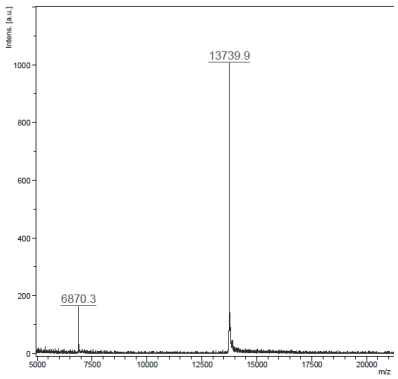



Figure A.6: MALDI TOF mass spectrum of GABARAPL2 (1 mg/ml).

## 10.2 NMR

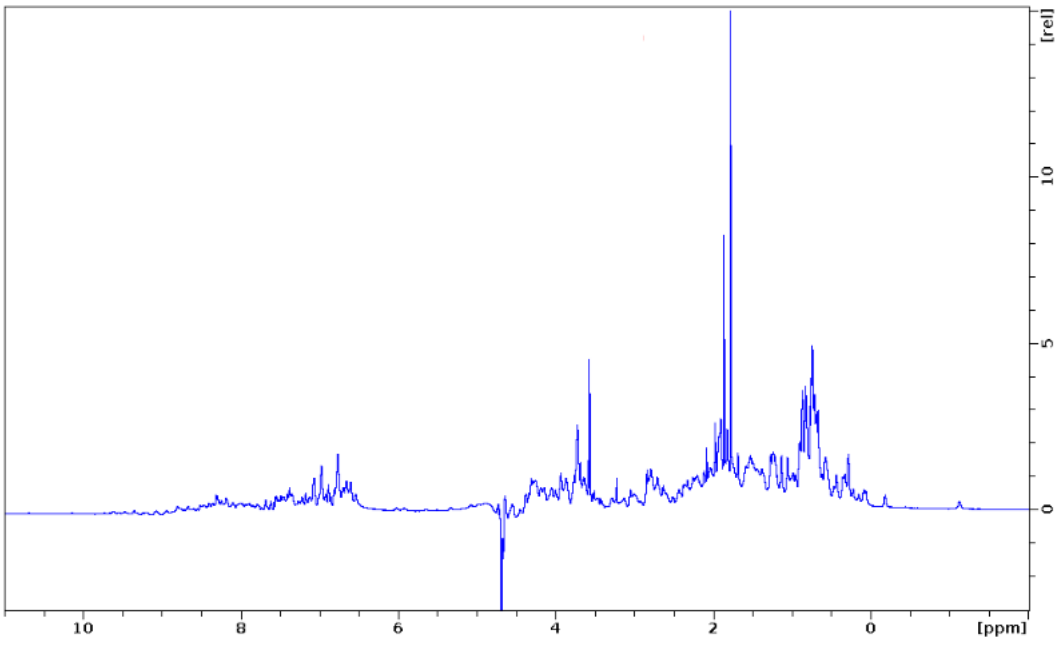



Figure A.7: $^{1}$H NMR spectra of LC3B which show indications of folded protein.

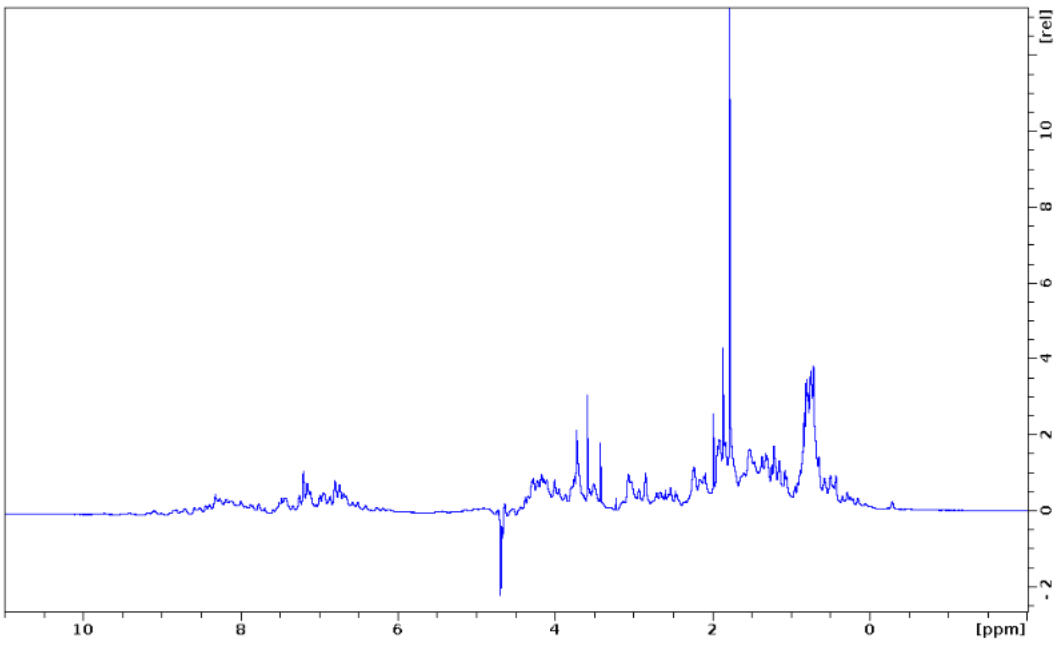



Figure A.8: $^{1}$H NMR spectra of LC3C which show indications of folded protein.

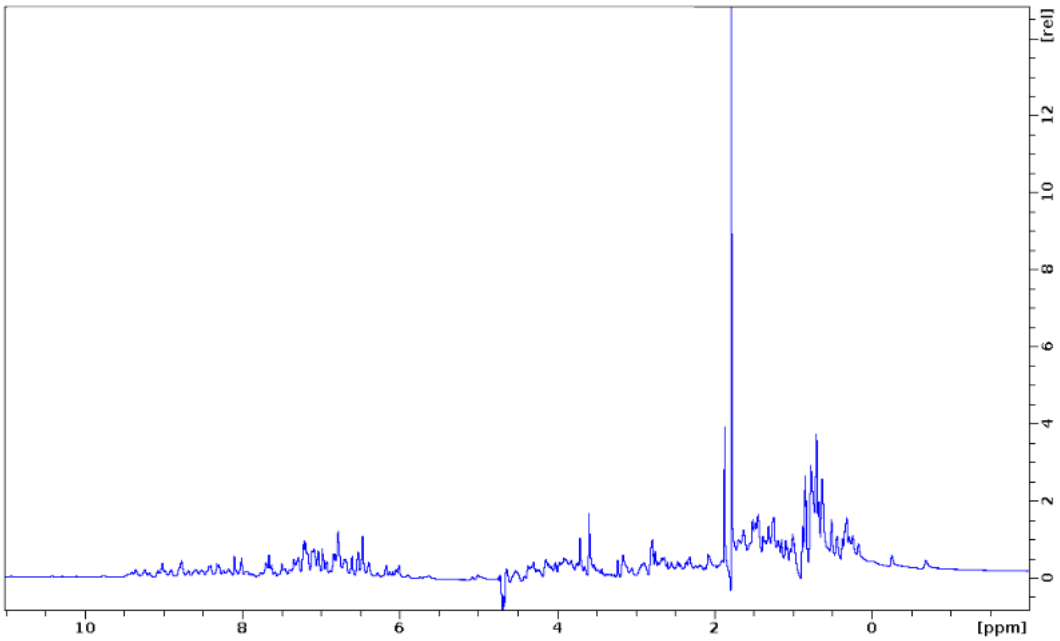



Figure A.9: $^{1}$H NMR spectra of GABARAP which show indications of folded protein.

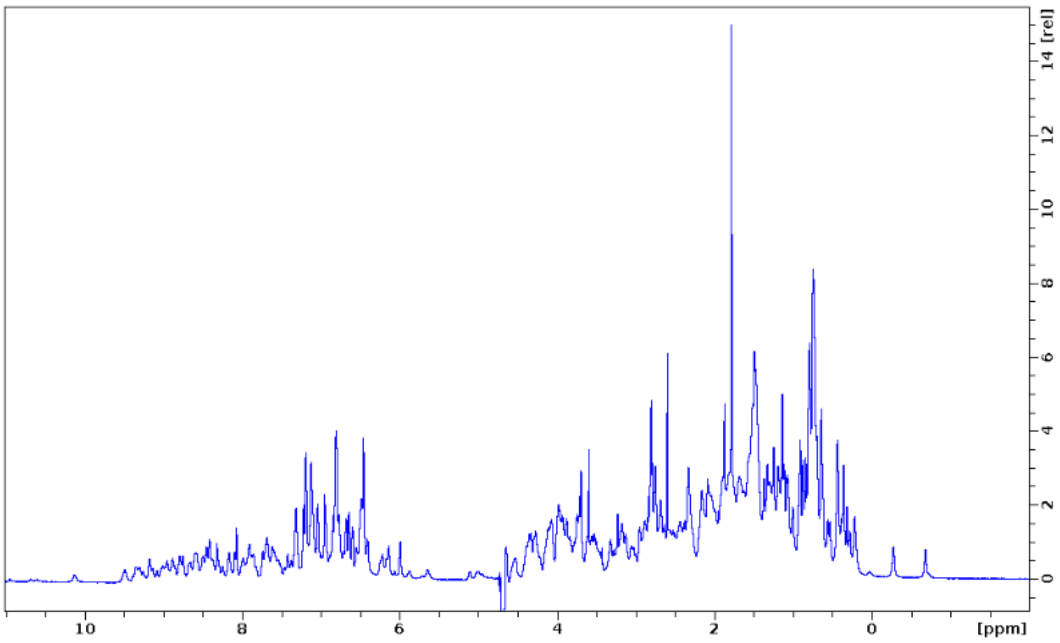



Figure A.10: $^{1}$H NMR spectra of GABARAPL1 which show indications of folded protein.

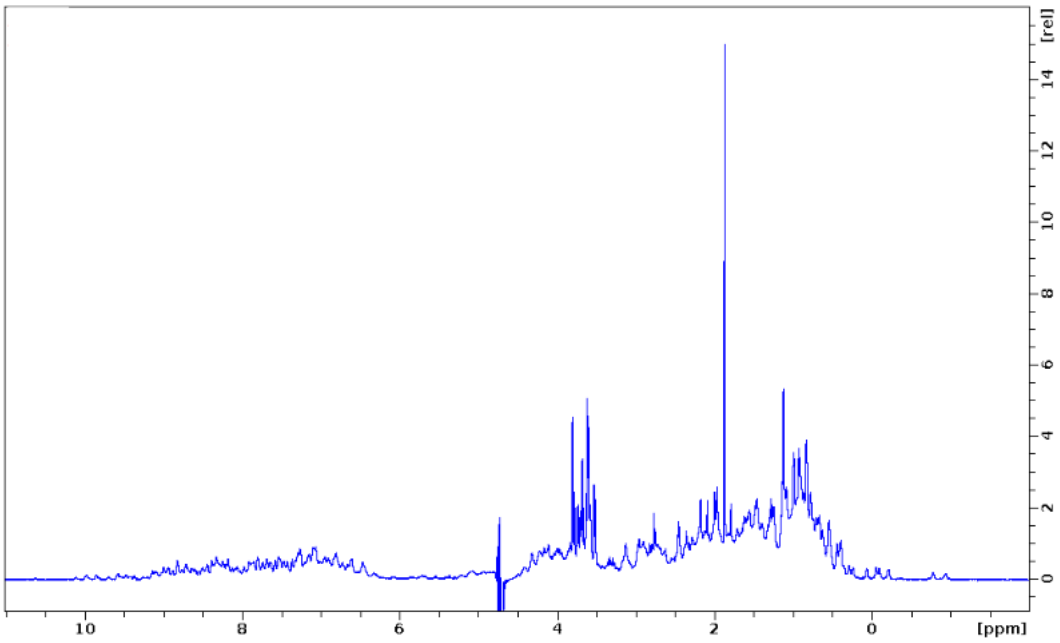



Figure A.11: $^{1}$H NMR spectra of GABARAPL2 which show indications of folded protein.

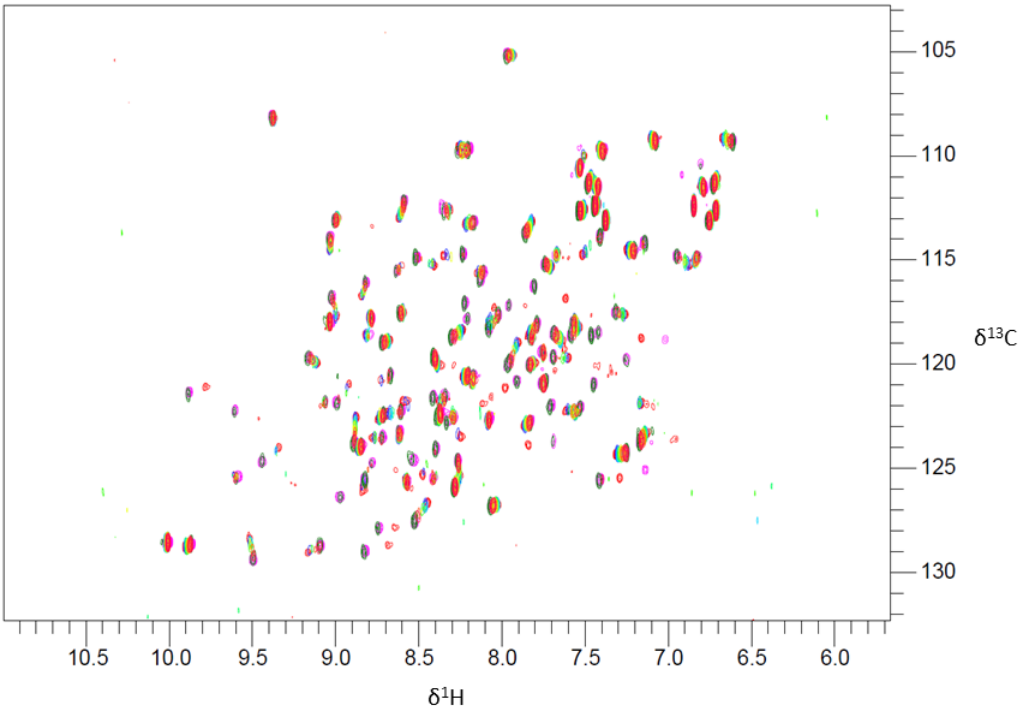



Figure A.12: Spectral overlay of the $^{1}$H-$^{15}$N-HSQC spectra of the titration of the WT p62 AIM peptide into $^{15}$N-GABARAPL2.

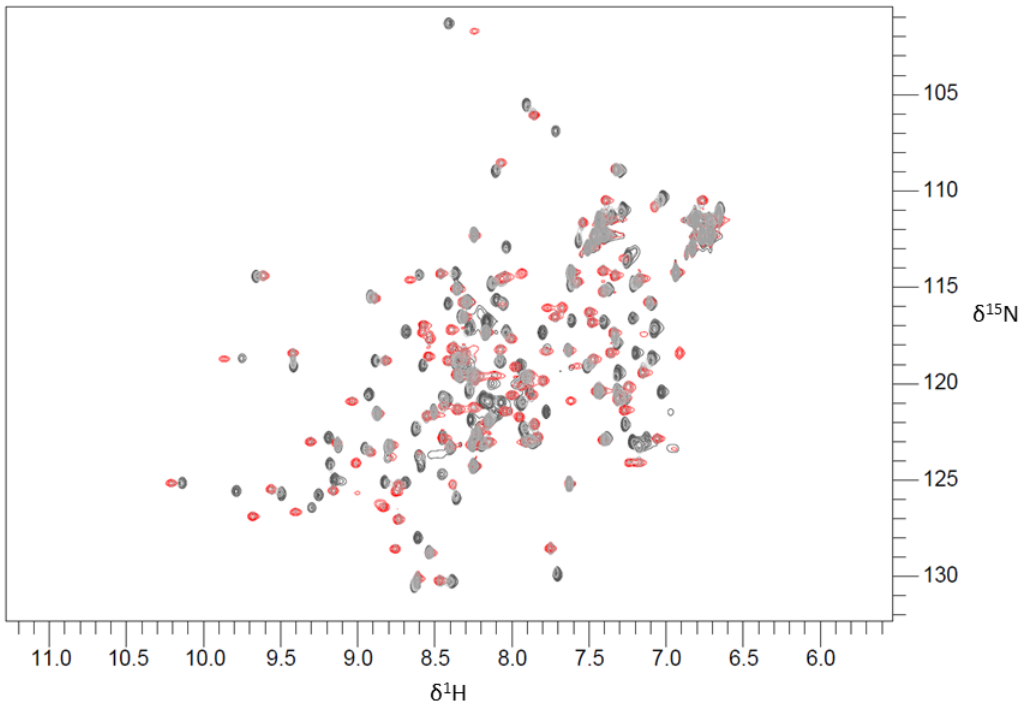



Figure A.13: Spectral overlay of the $^{1}$H-$^{15}$N-HSQC spectra of the titration of the pS342 p62 AIM peptide into $^{15}$N-LC3B.

## 10.3 ITC

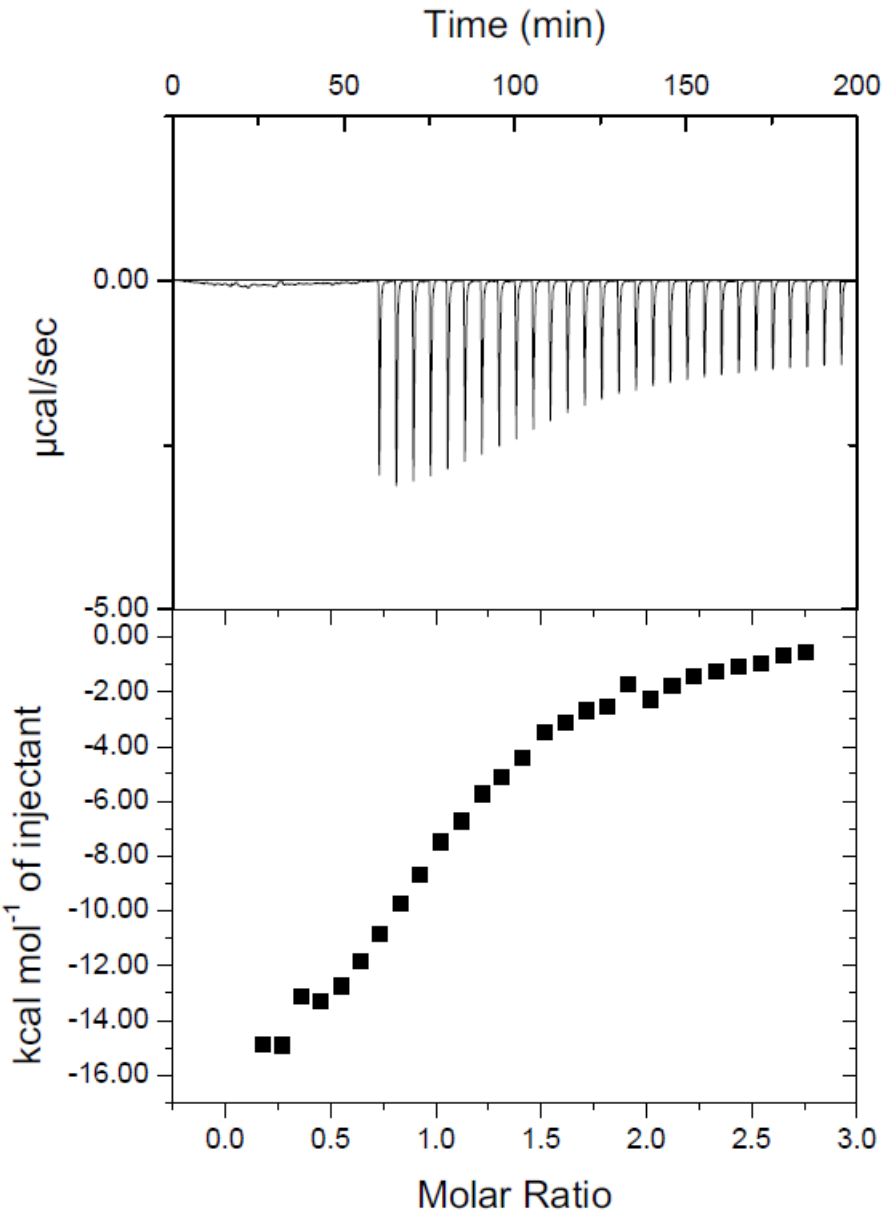



Figure A.14: ITC data for the titration of the WT p62 AIM peptide into LC3A. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

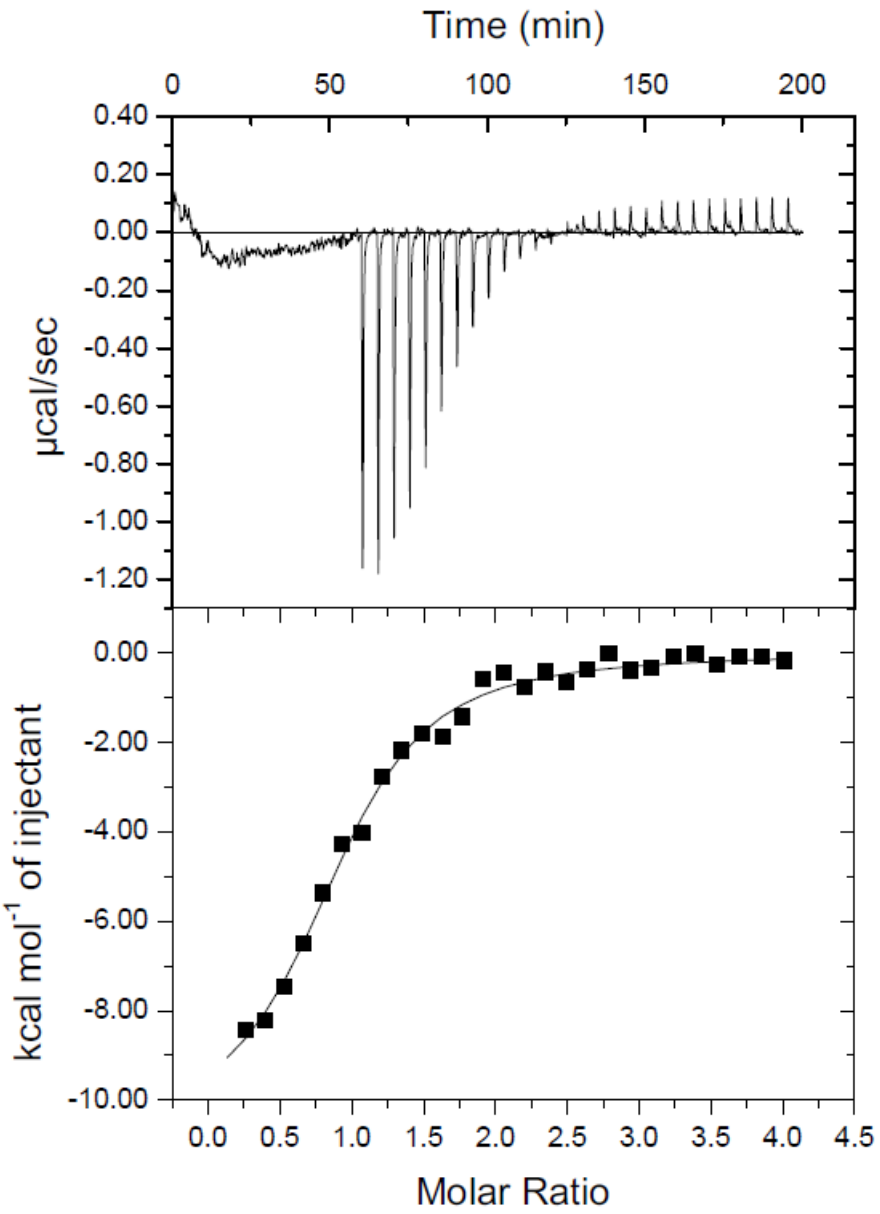



Figure A.15: ITC data for the titration of the WT p62 AIM peptide into LC3B. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

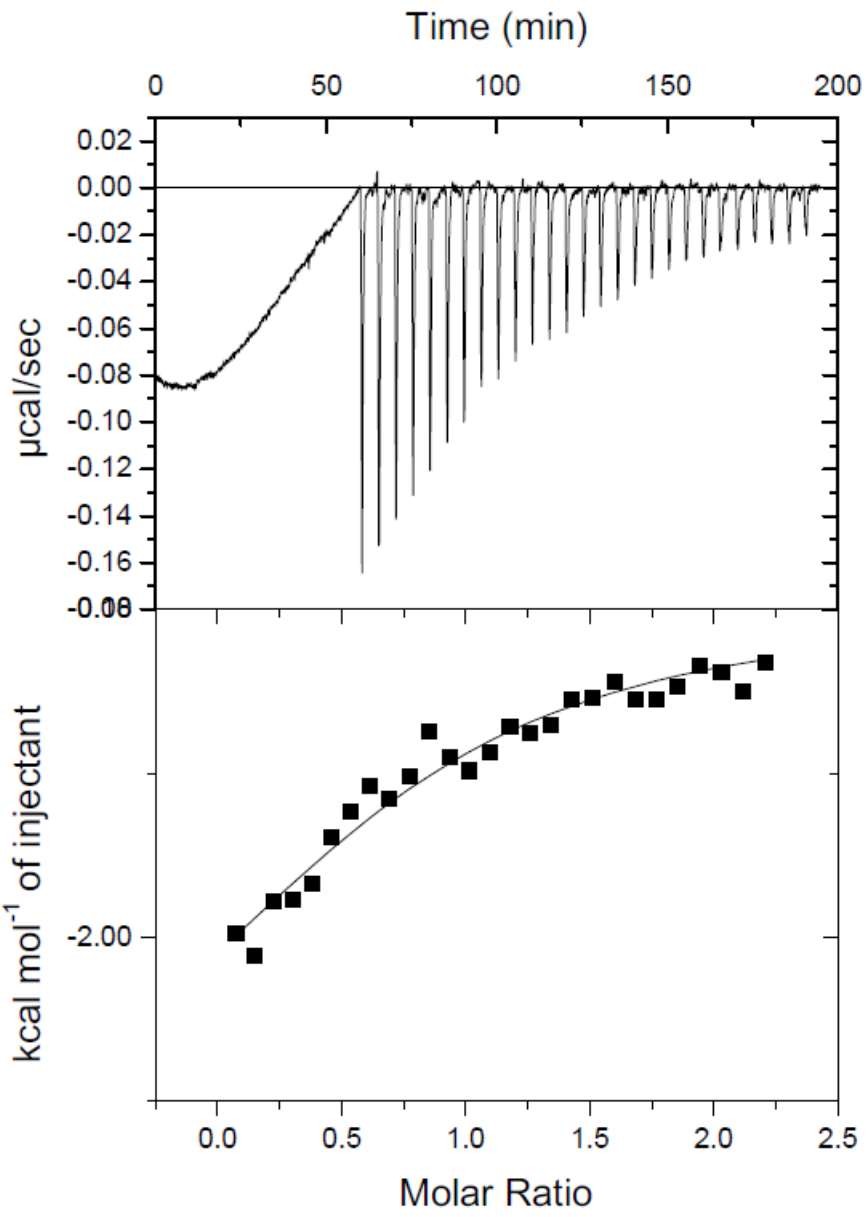



Figure A.16: ITC data for the titration of the WT p62 AIM peptide into LC3C. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

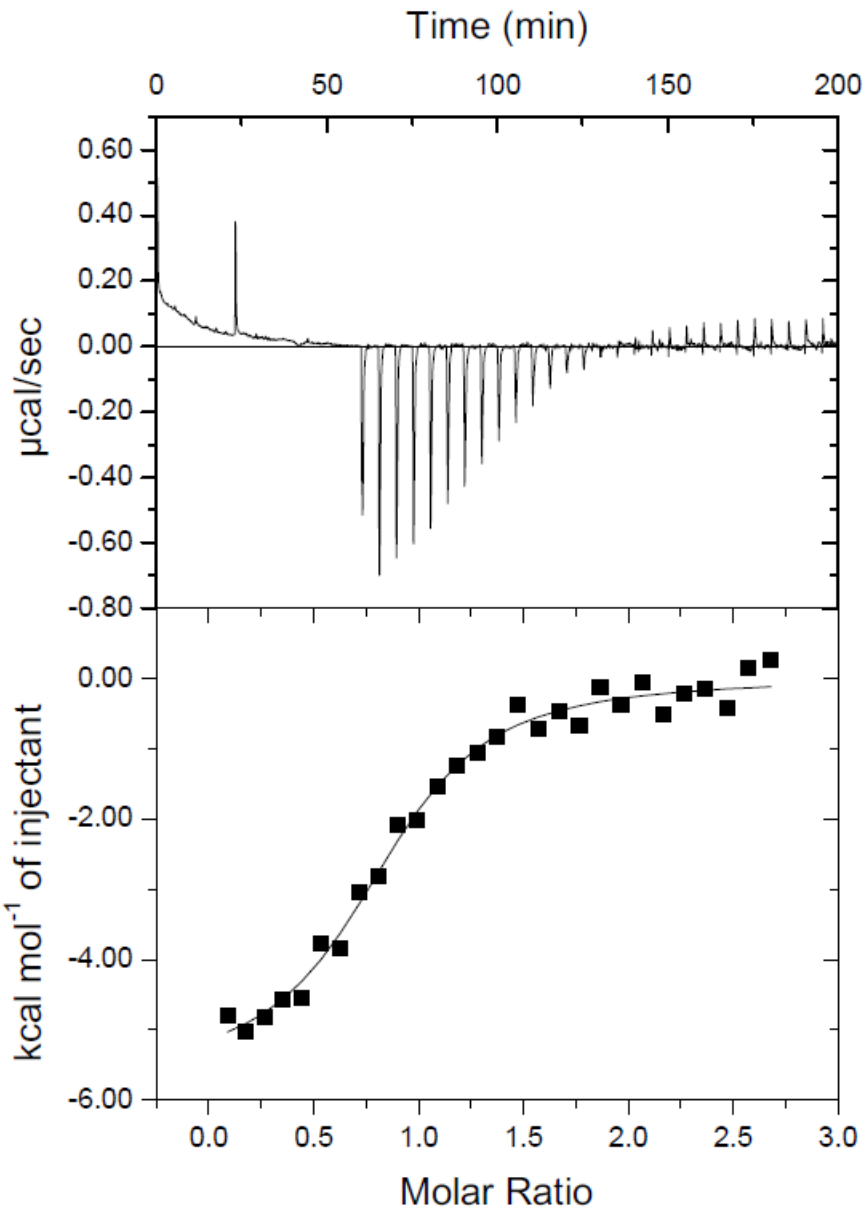



Figure A.17: ITC data for the titration of the WT p62 AIM peptide into GABARAPL1. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

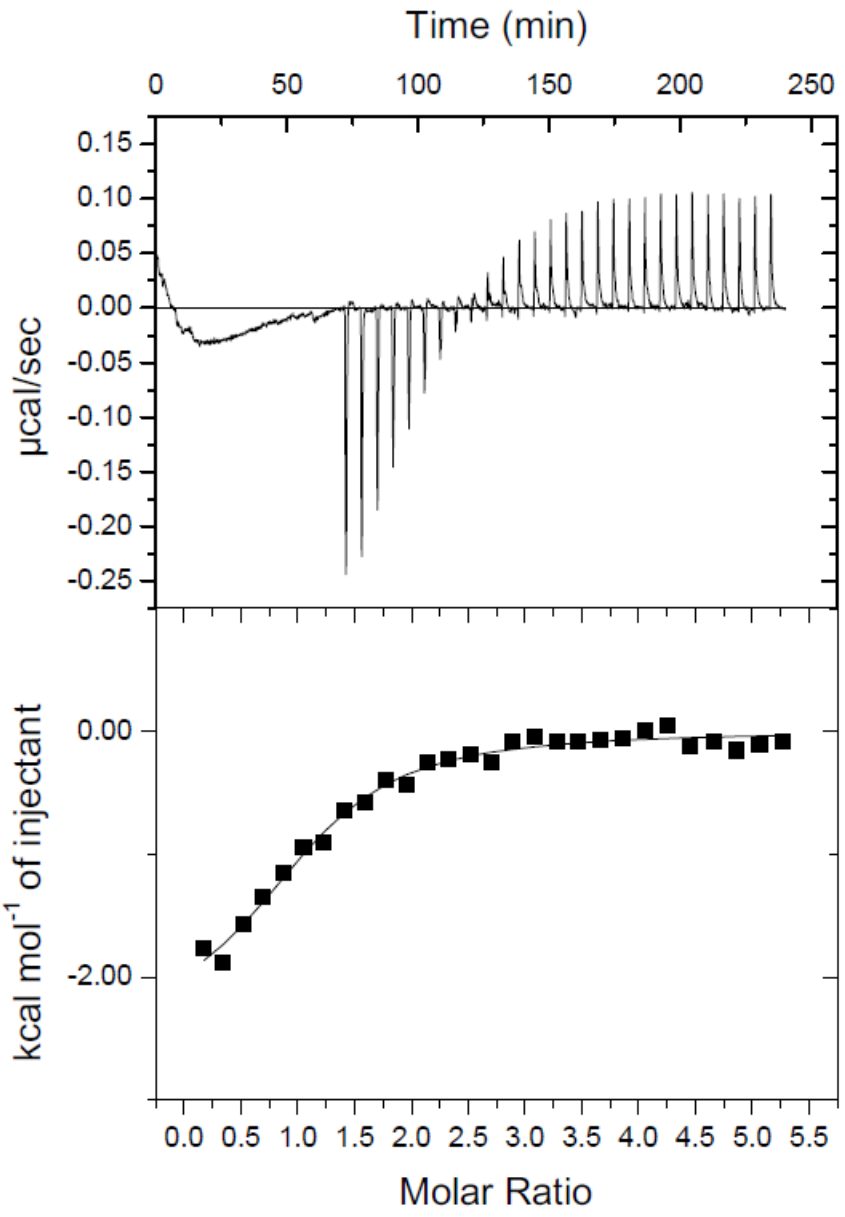



Figure A.18: ITC data for the titration of the WT p62 AIM peptide into GABARAPL2. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

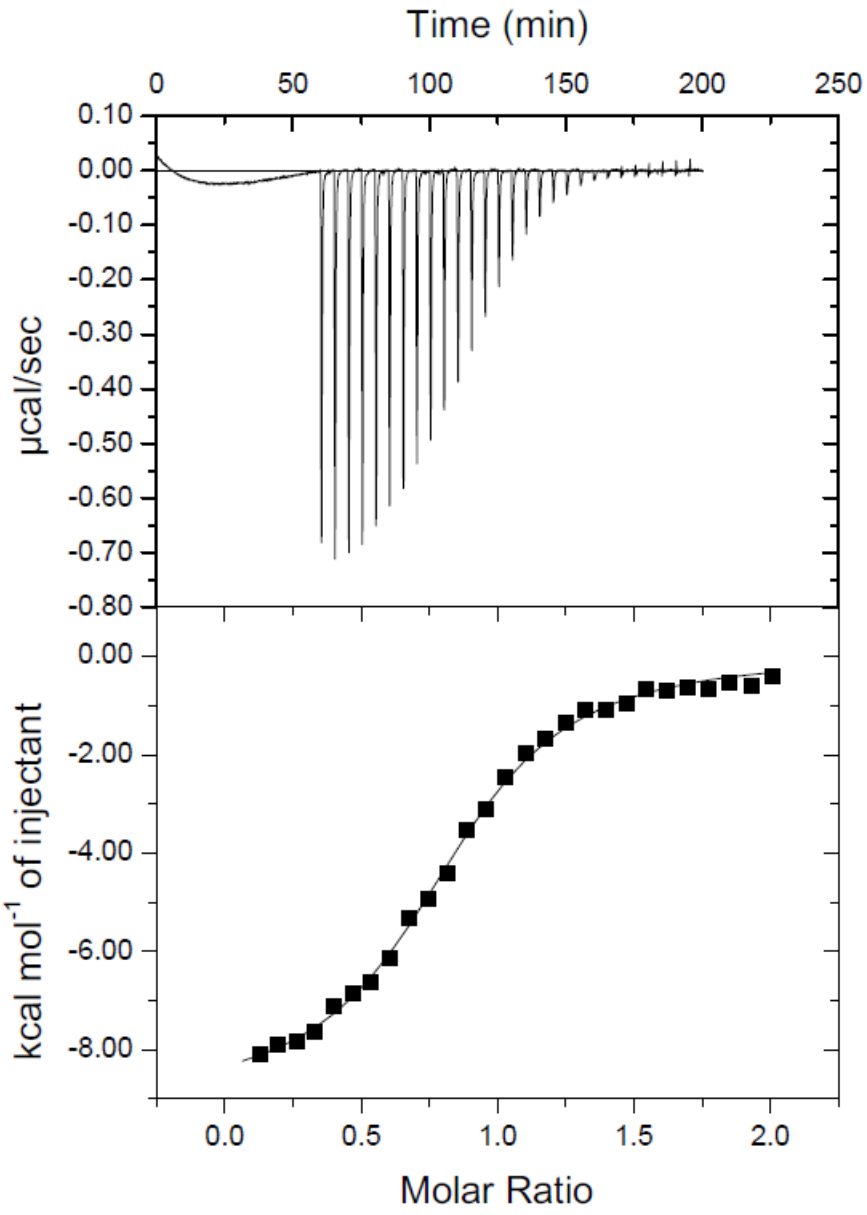



Figure A.19: ITC data for the titration of the pS342 p62 AIM peptide into LC3A. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

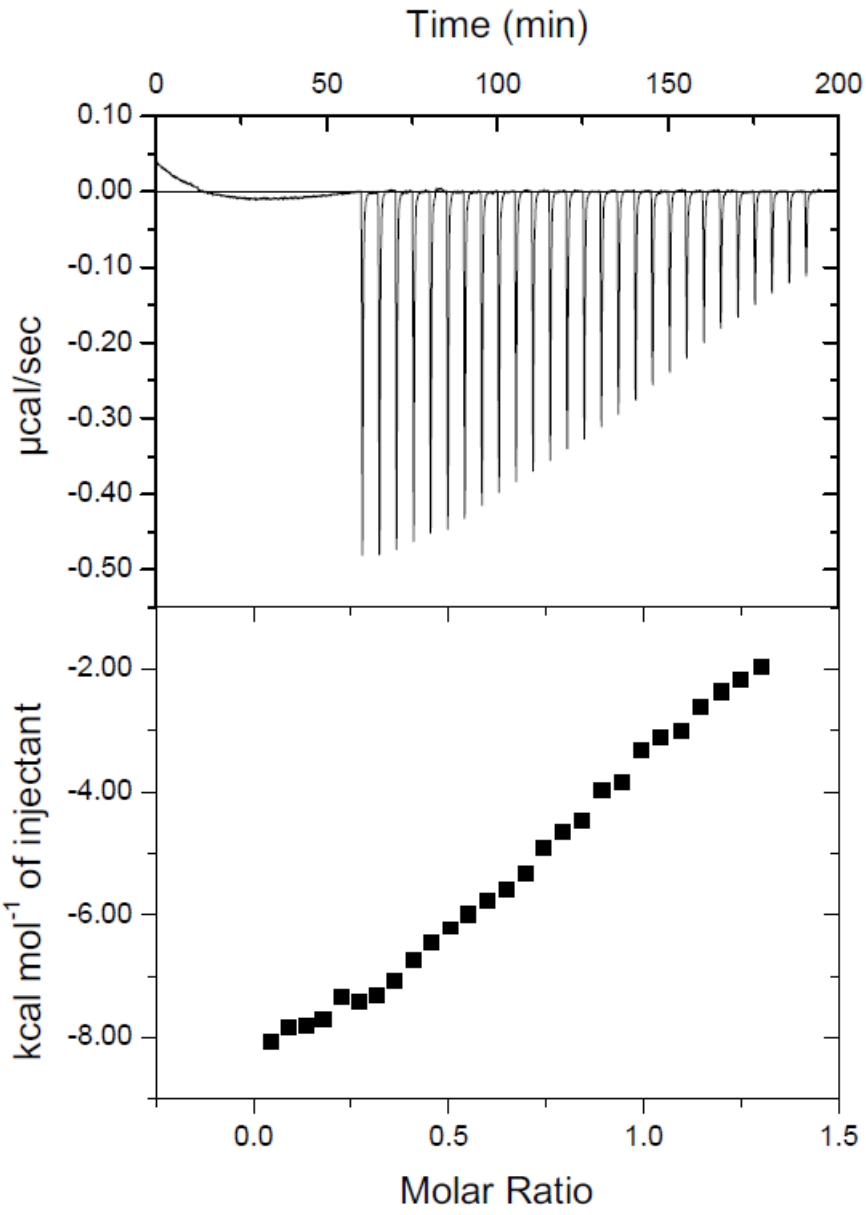



Figure A.20: ITC data for the titration of the pS342 p62 AIM peptide into LC3B. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

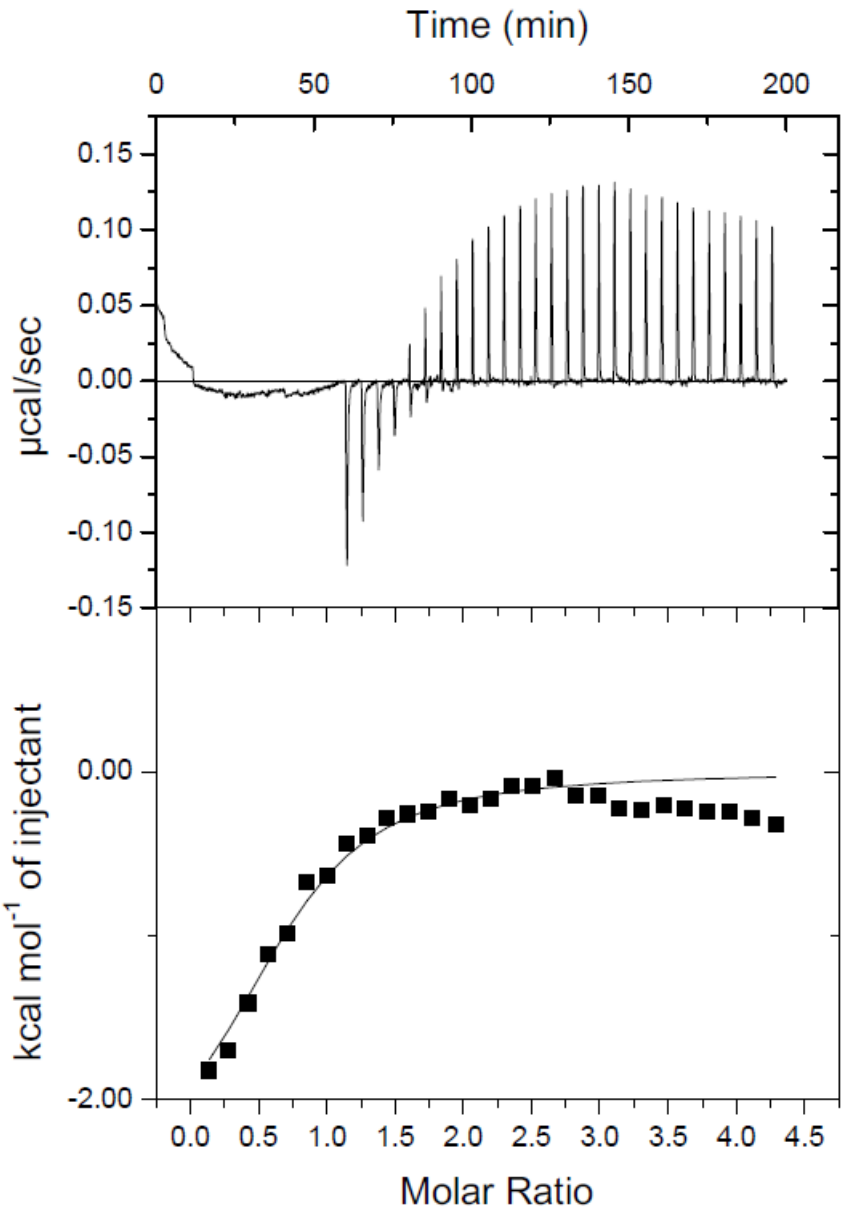



Figure A.21: ITC data for the titration of the pS342 p62 AIM peptide into LC3C. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

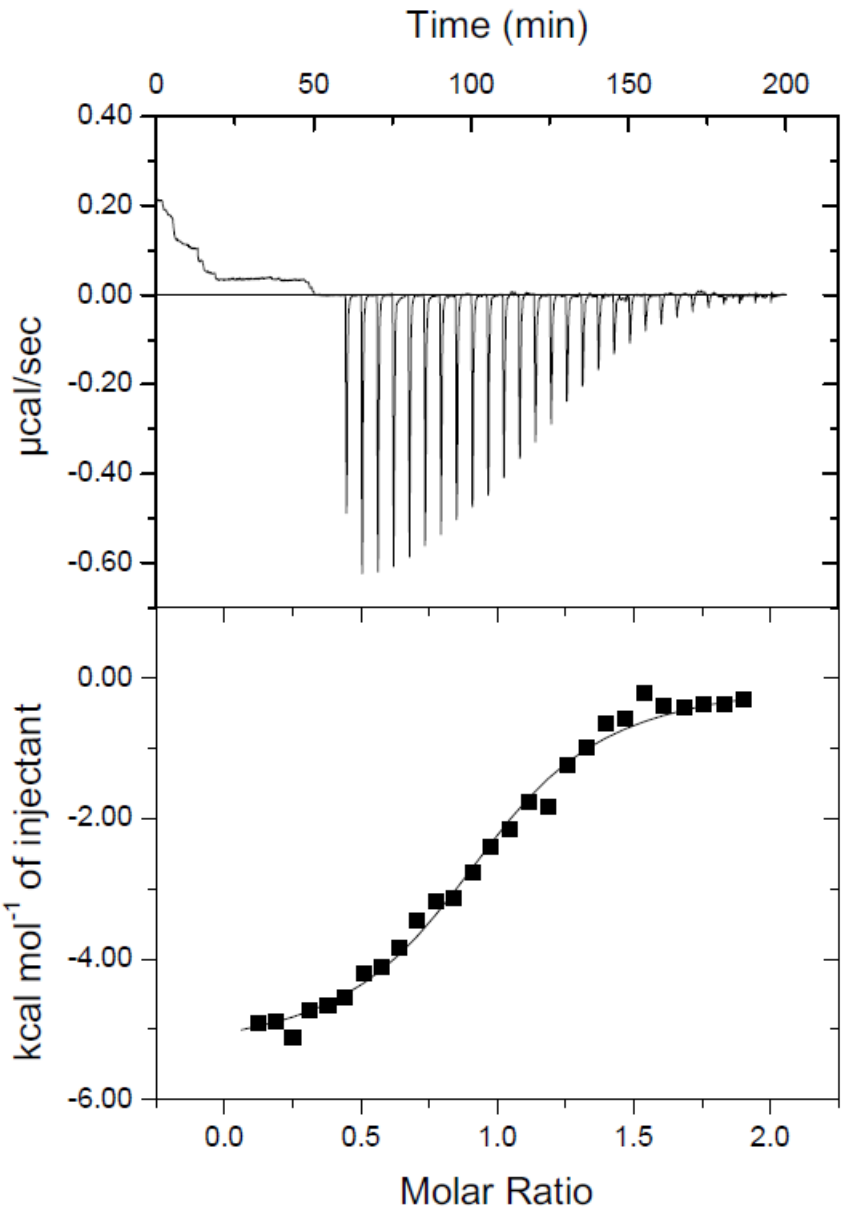



Figure A.22: ITC data for the titration of the pS342 p62 AIM peptide into GABARAPL1. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

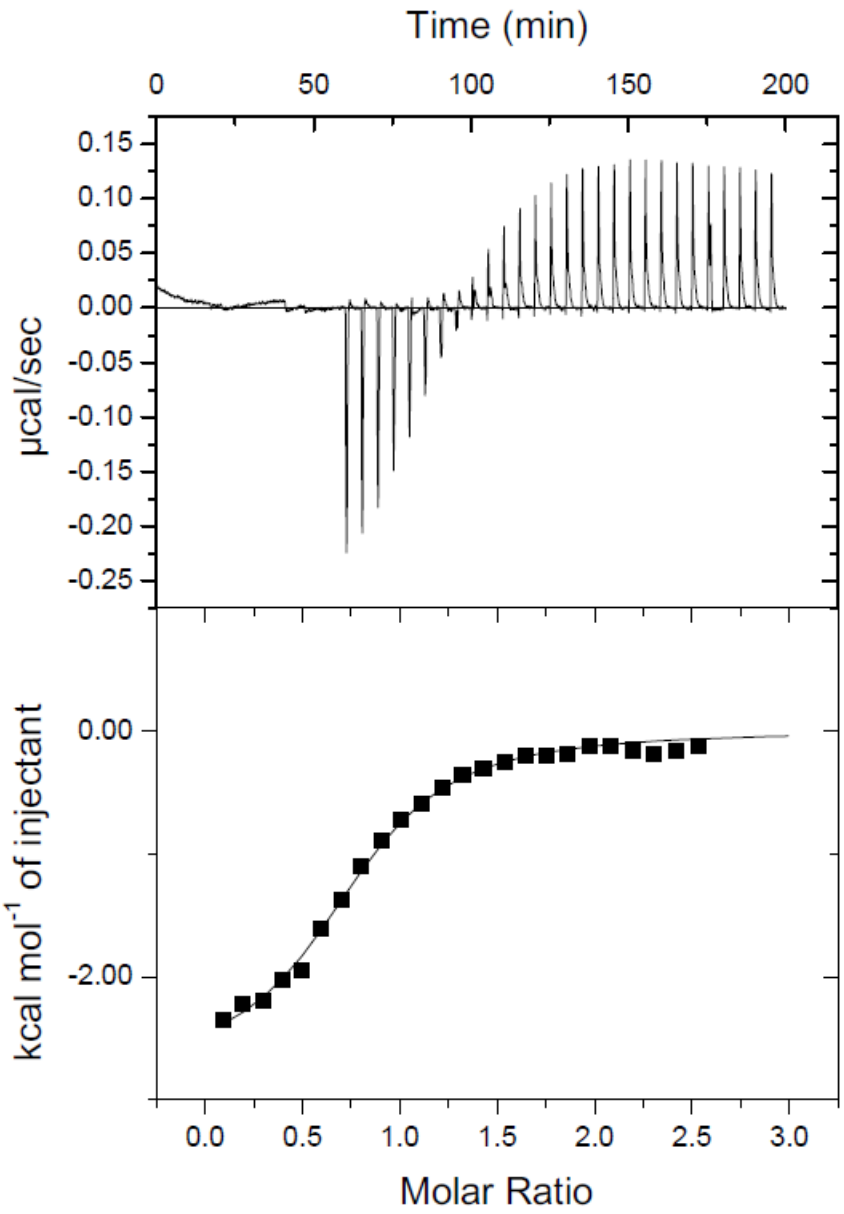



Figure A.23: ITC data for the titration of the pS342 p62 AIM peptide into GABARAPL2. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

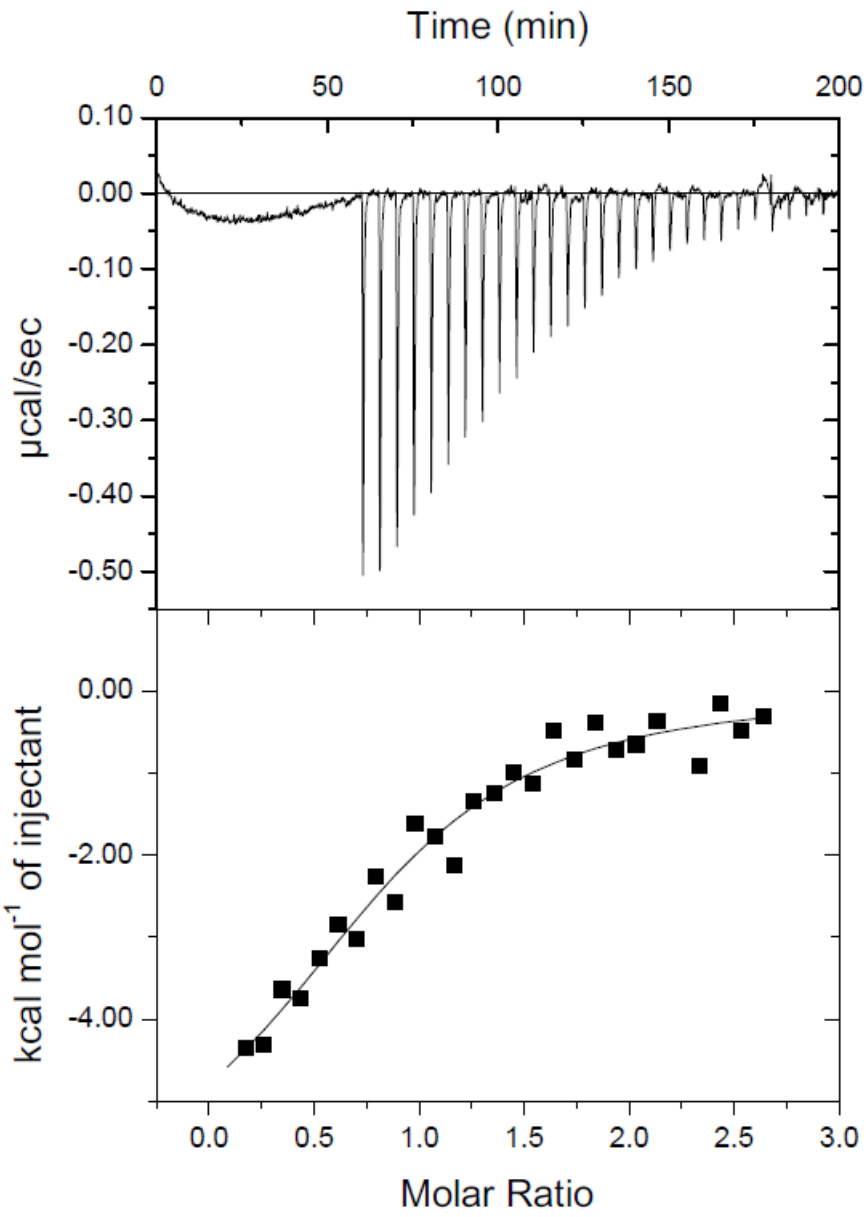



Figure A.24: ITC data for the titration of the L341V p62 AIM peptide into LC3A. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

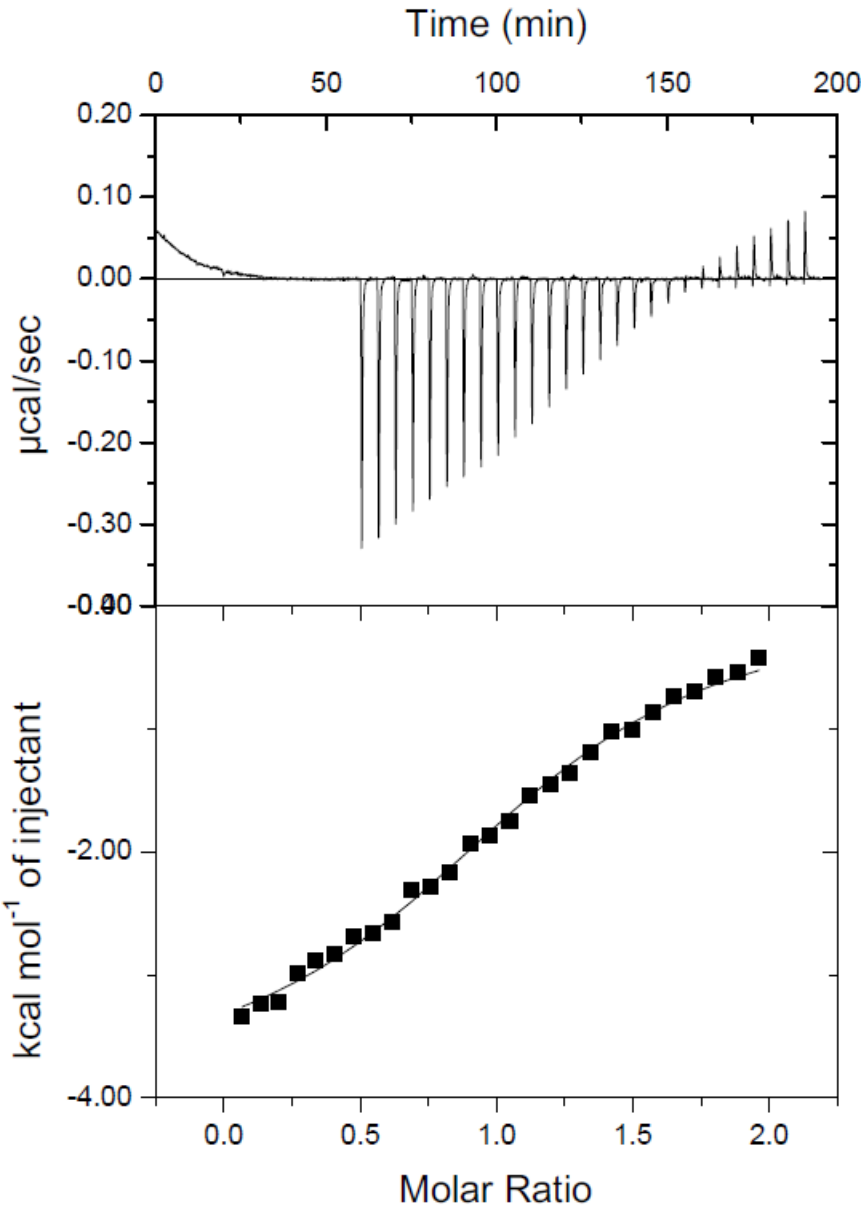



Figure A.25: ITC data for the titration of the L341V p62 AIM peptide into LC3B. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

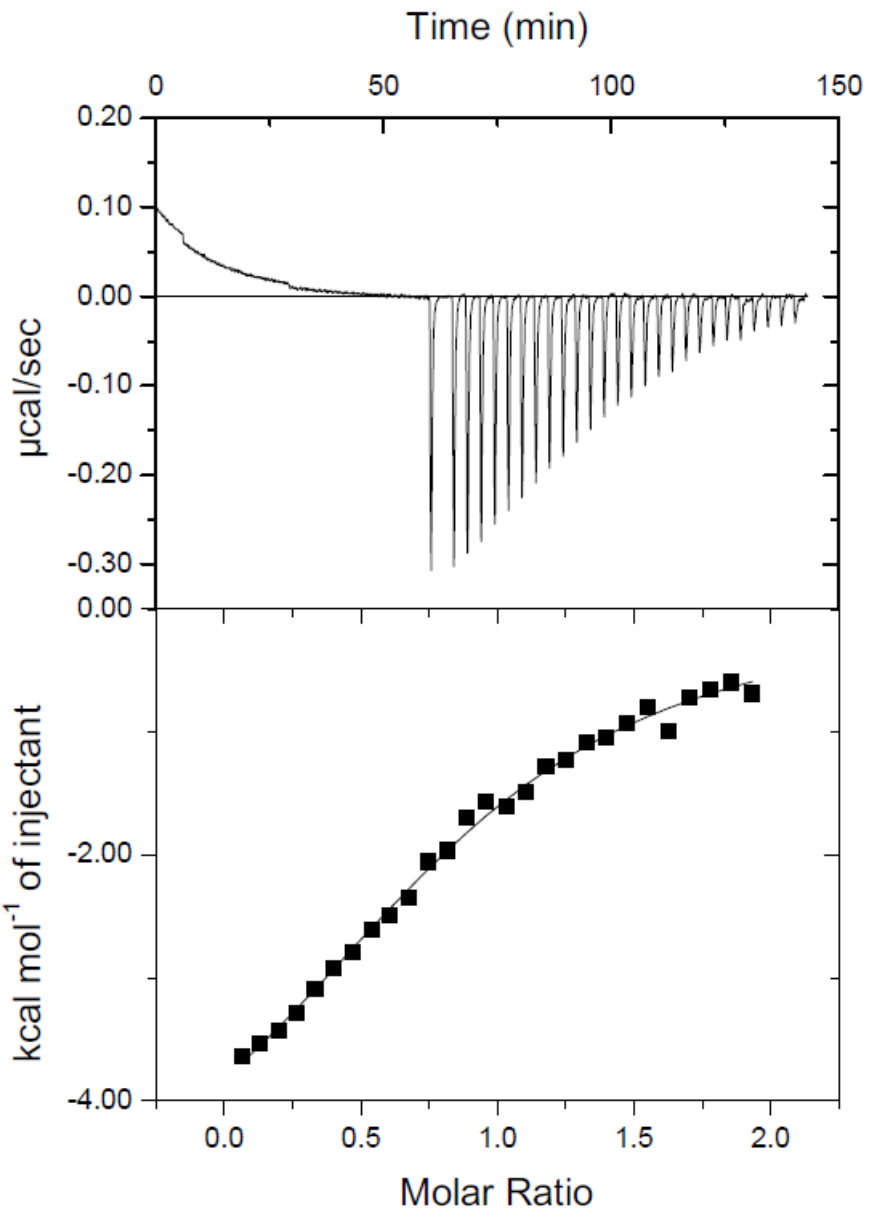



Figure A.26: ITC data for the titration of the L341V p62 AIM peptide into LC3C. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

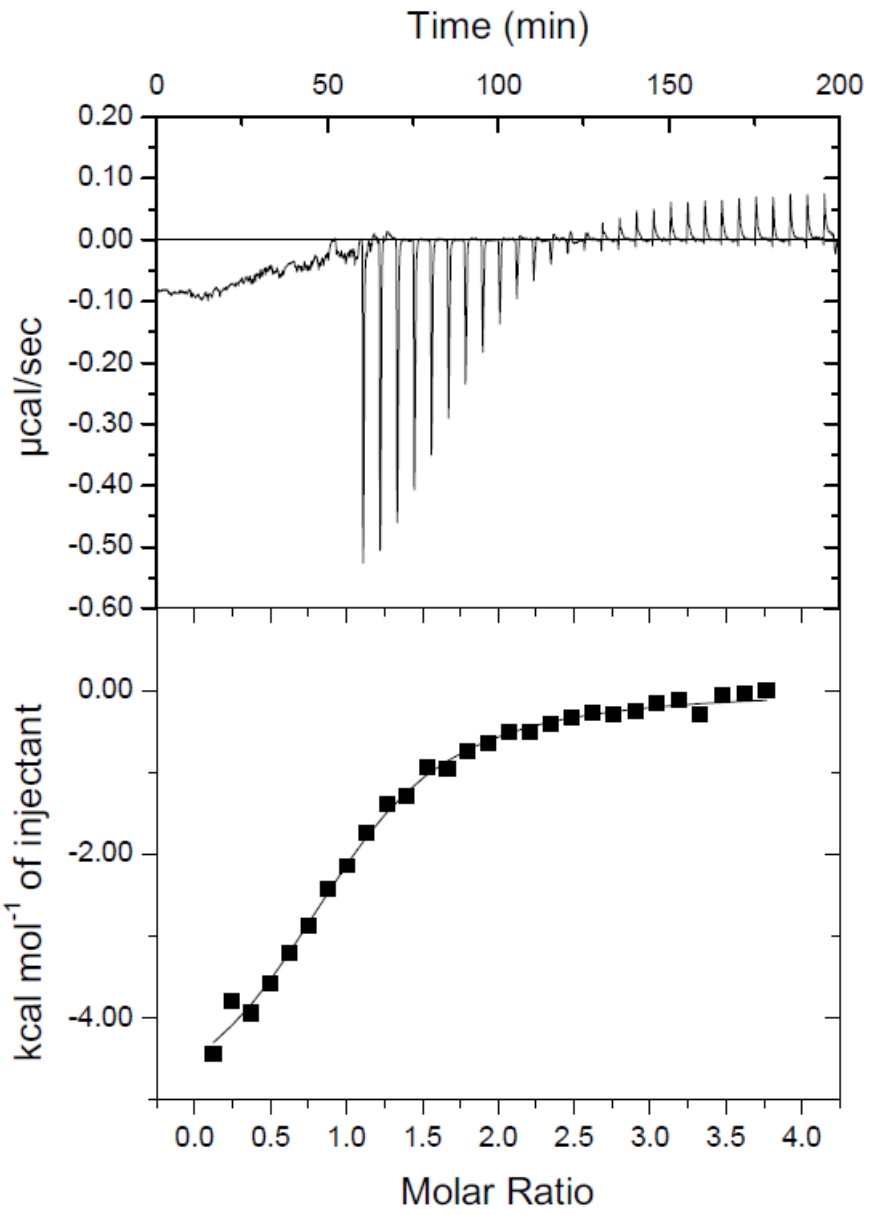



Figure A.27: ITC data for the titration of the L341V p62 AIM peptide into GABARAPL1. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

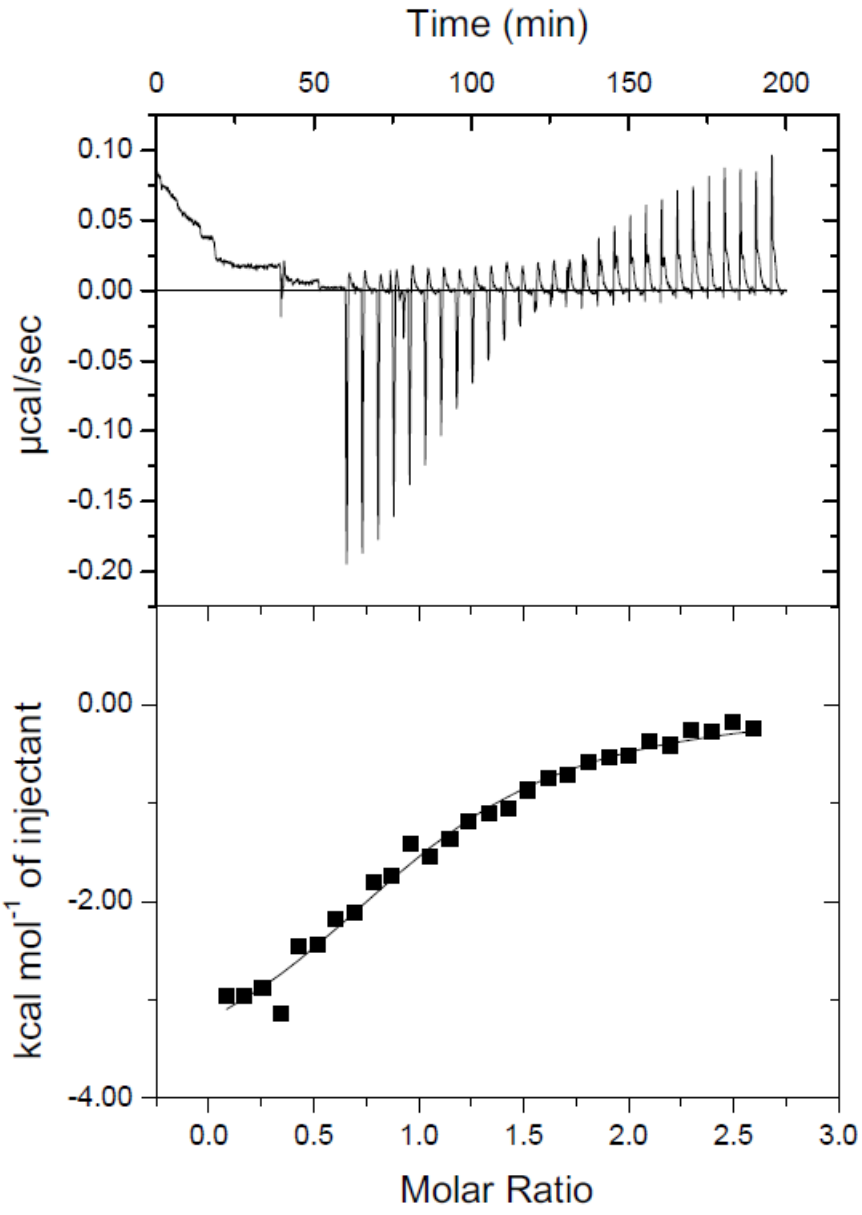



Figure A.28: ITC data for the titration of the L341V p62 AIM peptide into GABARAPL2. The applied power raw data is shown in the top panel showing peaks for each peptide injection. A plotted binding curve from the integration of these peaks and a line showing the fit of this data is shown in the bottom panel.

## 10.4 Professional Internship for PhD Students

**Note to examiners**:

This statement is included as an appendix to the thesis in order that the thesis accurately captures the PhD training experienced by the candidate as a BBSRC

Doctoral Training Partnership student. The Professional Internship for PhD Students is a compulsory 3-month placement which must be undertaken by DTP students. It is usually centred on a specific project and must not be related to the PhD project. This reflective statement is designed to capture the skills development which has taken place during the student's placement and the impact on their career plans it has had.

## PIPS Reflective Statement

My three-month placement was carried out at Sygnature Discovery which is a contract research organisation involved in work throughout the process of drug discovery. They have a broad range of in house expertise split into bioscience, chemistry, computational and DMPK departments. During my time at Sygnature I worked on a project within the bioscience department which attempted to expand their repertoire of biophysical techniques. I had not previously encountered this particular technique but this work fit well within my broader experience of biophysical characterisation of protein interactions. As such this made the topic accessible but was still challenging in a variety of ways which were useful to my development.

Firstly, it was particularly useful to observe the differences between my PhD lab and Sygnature. Many of these differences were small, due to differences in equipment or set up, but adapting to these differences is a key skill which I have advanced.

One of the key goals I hoped to achieve through this placement was to improve my verbal communication skills; presentations in particular. I had a great opportunity to practice this skill throughout my placement as I was presenting my work at least once a week. This also highlighted the skill needed in tailoring a presentation to a particular audience which involves the balance of broad descriptions of the work and specific scientific detail.

One additional aspect of this desire to improve my communication is in the informal discussion of research; a key skill for networking at conferences and elsewhere. I felt this improved by discussions with colleagues throughout my placement. Additionally, there were a number of issues throughout the project which required that I approach people and accurately describe the issues I was facing and subsequently discuss potential solutions.

This project produced significantly larger quantities of data than the types of experiments I typically utilise and this presented me with a number of novel challenges. I had to improve my organisational abilities to deal with large volumes of data by developing more standardised ways of analysing the data. This involved use of the

software package Origin which I had not previously used as well as developing my understanding further for the capabilities within GraphPad. Development of data fitting methodologies was also required which involved derivation of equations. Finally, rigorous statistical analysis was needed for much of this work. I also hoped to gain further insight into the industrial drug discovery pipeline.

By attending a variety of talks and meetings, I was able to expand my knowledge into how a contract research organisation like Sygnature fits into this pipeline. My time there produced insight into how arrangements are initially made with clients about what work is required and how this is planned. I was able to witness how this research then developed in the lab and how this is communicated back to the clients to produce useful research outcomes.

As well as these professional development outcomes, my experiments were able to produce useful data for Sygnature which allowed them to determine when it would be appropriate to use these new methodologies. This data has subsequently been included in a poster and presentations to advertise their biophysics capabilites.