

Biophysical Studies of Ligand Selectivity for the hATG8 Family of Autophagy Proteins

Kevin Butler, MSci (Hons.)

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

January 2021

Abstract

Selective autophagy is a process by which potentially toxic material such as aggregated proteins, organelles and pathogens are specifically targeted and transported to autophagosomes and subsequent degradation by the lysosome. The process relies on selective autophagy receptor proteins (SARs) recognising cargo and then interacting with the six human autophagy related receptor proteins (hATG8) anchored to the membrane of the autophagosome. The interaction between SARs and hATG8 proteins requires a short amino acid sequence called an LC3 interacting region (LIR). LIRs consist of a core motif $[W/F/Y]_0-X_1-X_2-[L/V/I]_3$ with the aromatic residue often preceded by acidic residues. The LIR interacts with ATG8 proteins at a conserved binding site consisting of two hydrophobic pockets with aromatic and hydrophobic residues sitting in hydrophobic pockets one and two respectively. Neighbour of BRAC1 gene 1 (NBR1), unlike most SARs contains two LIRs and whilst both contain the core LIR motif they vary in the amino acids at the aromatic and hydrophobic position and the composition of amino acids preceding the aromatic residue. Using Biophysical methods the interaction of the two LIRs of NBR1 (Sequences containing 11-12 amino acids) with the hATG8 proteins were studied to gain a better understanding of the effect the different LIR sequence has on binding.

NMR titration experiments with subsequent binding site mapping indicated that whilst both LIRs interact with LC3B WT at the hATG8 canonical binding site there are variations in how the LIR interacts with LC3B. The interaction of NBR1 LIR1 (ASSEDYIILPE) containing more acidic amino acids preceding the aromatic forms more interactions with the face of LC3B than that of NBR1 LIR2 (AQDLLSFELLD).

An LC3B binding site mutant that altered the charge on the face of LC3B WT showed preferential binding for NBR1 LIR2. NMR binding site studies indicated that this change allowed NBR1 LIR2 to fit tighter against the binding surface of LC3B.

Using NMR binding site mapping a phosphomimetic mutant showed that phosphorylation of LC3B at threonine 29 made very little difference to how NBR1 LIR1 interacted with LC3B. However, it made a large difference in the interaction with LIR2 with the aromatic residue of the LIR no longer sitting in HP1.

The LIRs of NBR1 showed a preference to binding to LC3A over the other members of the LC3 subfamily. NMR binding site mapping showed the LIRs of NBR1 bound to LC3A at the canonical binding site. However, the structures generated in HADDOCK (a molecular docking program) for the interaction of LC3A and LIR1 showed the aromatic binding into HP2 rather than HP1. For NBR1 LIR2 Phe563 was seen to sit more across the face of LC3A and changes around the hydrophobic pocket on LC3A when compared to LC3B allowed a more favourable interaction of the LIR with the face of LC3A.

Acknowledgements

I would like to thank Professor Mark Searle for giving me the opportunity to carry out this project and for all the help, support and direction he has offered over the last 9 years. I would not have even thought of starting a PhD without the encouragement of Dr. Adrienne Davis to develop my research skills and who was there to support me when the stresses of doing a job and part time PhD got too much.

Dr. Jed Long deserves special thanks for having the patience of taking someone who had never worked in a microbiology lab and teaching me how to produce proteins (well most of the time). His constant advice and help with problem solving kept the project on track. I would like to thank Dr. Huw Williams for always having time to talk through the project and offer advice. He also provided much needed down time through the suggestion of a coffee break.

I would also like to thank all the members of the Searle, Sultanas and Oldham group past and present whose help and friendship has made working in A20 an enjoyable experience. I would also like to thank all the technical staff in the CBS and Chemistry. Without whom I wouldn't be able to do my job or PhD. The members of the chemistry and university student support teams deserve a lot thanks for advising me on how to proceed during some extremely difficult and stressful times.

Finally I would like to thank my wife Hannah and children for offering lots of support and love to keep me going when it all got a bit fraught. They have always been understanding when I have had to work extra hours and always there to give me much needed relaxation time through "playing cars".

Contents

| | |
|---|----|
| 1. Introduction | 1 |
| 1.1 Autophagy | 1 |
| 1.2 The ATG8 Family | 5 |
| 1.3 Selective Autophagy and the LIR | 8 |
| 1.4 Phosphorylation of LC3B | 15 |
| 1.5 Project Goals | 16 |
| 2. Materials and Methods | 18 |
| 2.1 Chemicals | 18 |
| 2.2 Biological Materials | 19 |
| 2.3 Molecular Biology Techniques | 22 |
| 2.3.1 Sterilisation | 22 |
| 2.3.2 Small-scale Overnight <i>E. coli</i> Cell Cultures | 22 |
| 2.3.3 Agar Plates | 22 |
| 2.3.4 Glycerol Stocks | 22 |
| 2.3.5 Site Directed Mutagenesis | 22 |
| 2.3.6 DNA Digestion and Ligation | 23 |
| 2.3.7 Agarose Gels | 23 |
| 2.3.8 Transformation of Plasmid DNA into <i>E. coli</i> Cells | 24 |
| 2.3.9 Plasmid Purification | 24 |
| 2.3.10 DNA Sequencing | 24 |
| 2.3.11 Protein Over-expression | 25 |
| 2.4 Protein Purification | 26 |
| 2.4.1 Affinity Chromatography of GST-fusion Proteins | 26 |
| 2.4.2 Cationic Column FPLC | 26 |
| 2.4.3 Hi Trap Desalt FPLC | 26 |
| 2.4.4 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-Page) | 27 |
| 2.5 Biophysical Techniques | 28 |
| 2.5.1 Mass Spectrometry | 28 |
| 2.5.2 Nuclear Magnetic Resonance | 29 |
| 3. Producing the ATG8 Proteins and the Mutants of LC3B | 39 |
| 3.1 Mutants of LC3B | 39 |
| 3.2 Modification of the ATG8 Proteins | 41 |

| | |
|--|-----|
| 3.3 Protein Production and Purification | 44 |
| 3.4 Initial Analysis of the hATG8 Proteins..... | 46 |
| 3.4.1 Initial MS of the hATG8 Proteins | 46 |
| 3.4.2 One Dimensional NMR of the hATG8 Proteins..... | 48 |
| 3.5 Conclusion..... | 49 |
| 3.5.1 LC3B Mutants..... | 49 |
| 3.5.2 ATG8 Proteins | 50 |
| 4. Probing Binding Selectivity of hATG8s with the LIRs of NBR1..... | 51 |
| 4.1 Binding of LIRs of NBR1 with the hATG8 Proteins | 52 |
| 4.2 Competition Binding of LIRs of NBR1 to the LC3B Mutants Using ESI-MS | 54 |
| 4.2.1 LC3B Mutants Binding to NBR1 LIR1..... | 55 |
| 4.2.2 LC3B Mutants Binding to NBR1 LIR2..... | 59 |
| 4.3 Competition Binding of LIRs of NBR1 to the ATG8s Using ESI-MS | 61 |
| 4.3.1 LC3s Binding to NBR1 LIR1 | 62 |
| 4.3.2 LC3s Binding to NBR1 LIR2 | 64 |
| 4.3.3 GABARAPs Binding to NBR1 LIR1..... | 65 |
| 4.3.4 GABARAPS Binding to NBR1 LIR2..... | 66 |
| 4.4 Discussion..... | 68 |
| 4.4.1 LC3B Mutants Binding to the LIRs of NBR1..... | 69 |
| 4.4.2 hATG8 Proteins Binding to LIRs of NBR1 | 70 |
| 4.5 Summary | 73 |
| 5. NMR Assignment of WT LC3B and Studies of its Interactions with LIRs of P62 and NBR1 | 74 |
| 5.1 Backbone Assignment of LC3B..... | 75 |
| 5.2 Interaction of WT LC3B with the LIR1 of NBR1..... | 78 |
| 5.3 Interaction of WT LC3B with the LIR2 of NBR1..... | 83 |
| 5.4 Interaction of WT LC3B with the LIR of p62..... | 88 |
| 5.5 Discussion..... | 91 |
| 5.5.1 Titrations with the LIRs | 91 |
| 5.5.2 Comparing the Binding of the LIRs | 94 |
| 5.5.3 Comparison of Titration of LC3B WT with NBR1 LIR1 and p62 LIR..... | 95 |
| 5.5.4 Comparison of Titration of LC3B WT with NBR1 LIR2 and p62 LIR..... | 98 |
| 5.5.5 Comparison of Titration of LC3B WT with NBR1 LIR1 and NBR1 LIR2 | 101 |
| 5.6 Summary | 104 |
| 6. Investigating Selectivity Between the two LIRs of NBR1 | 105 |
| 6.1 Backbone Assignment of LC3B KL and T29D Mutant | 105 |

| | |
|---|-----|
| 6.1.1 LC3B KL Mutant Assignment | 105 |
| 6.1.2 LC3B T29D Mutant Assignment | 108 |
| 6.2 NMR Titrations of LC3B Mutants with the LIRs of NBR1 | 108 |
| 6.2.1 Titrations with KL Mutant | 109 |
| 6.2.2 Titrations with T29D Mutant | 118 |
| 6.3 Discussion..... | 124 |
| 6.3.1 Assigning the KL and T29D Mutant..... | 124 |
| 6.3.2 Titrations of the NBR1 LIRs with the LC3B Mutants | 126 |
| 6.3.3 Comparison for Mutants with LIRs | 129 |
| 6.3.4 Comparison of LC3B KL and LC3B WT Titration with NBR1 LIR1 | 129 |
| 6.3.5 Comparison of LC3B KL and LC3B WT Titration with NBR1 LIR2 | 132 |
| 6.3.6 Comparison of LC3B KL Titration with NBR1 LIR1 and LIR2..... | 135 |
| 6.3.7 Comparison of LC3B T29D and LC3B WT Titration with NBR1 LIR1 | 138 |
| 6.3.8 Comparison of LC3B T29D and LC3B WT Titration with NBR1 LIR2 | 142 |
| 6.3.9 Comparison of LC3B T29D Titration with NBR1 LIR1 and LIR2 | 145 |
| 6.5 Summary | 148 |
| 7. Interaction of LC3A with the LIRs of NBR1 | 149 |
| 7.1 NMR Backbone Assignment of LC3A | 151 |
| 7.2 NMR Titrations of LC3A with the LIRs of NBR1..... | 154 |
| 7.2.1 Titrations with NBR1 LIR1 | 154 |
| 7.2.2 Titrations with NBR1 LIR2 | 159 |
| 7.3 Discussion..... | 163 |
| 7.3.1 Titrations of LC3A with the LIRs of NBR1..... | 163 |
| 7.3.2 Comparison of LC3B and LC3A LIR1 | 165 |
| 7.3.3 Comparison of LC3B and LC3A LIR2 | 169 |
| 7.3.4 Comparison of LIR1 and LIR2 with LC3A..... | 173 |
| 7.4 Summary | 177 |
| 8. Conclusion..... | 178 |
| 8.1 Investigating the Changes in Binding Interaction between the Two LIRs of NBR1 | 178 |
| 8.2 Binding Preferences of the hATG8 Proteins | 181 |
| 8.3 Characterization of the Interaction of NBR1 LIRs with LC3A..... | 183 |
| 8.4 Future Work..... | 186 |
| 8.5 Conclusion..... | 187 |
| Bibliography | 188 |
| Appendix | 198 |

| | |
|--------------------|-----|
| A. Appendix A..... | 198 |
| B. Appendix B..... | 202 |
| C. Appendix C..... | 207 |

List of Figures

All figures and schemes contained within the thesis were generated by the author unless otherwise indicated in the legend.

| | |
|--|----|
| Figure 1.1: mTOR signalling network. | 2 |
| Figure 1.2: Formation of ATG16L complex. | 3 |
| Figure 1.3: Model showing formation of autophagosome and subsequent fusion with the lysosome to form the autolysosome. | 4 |
| Figure 1.4: Schematic showing process of conjugation of phosphatidylethanolamine to ATG8s | 5 |
| Figure 1.5: ATG8 interactions on the phagophore. | 6 |
| Figure 1.6: Crystal structure of ubiquitin (blue) and LC3B (green). The first two α -helices on LC3B are coloured red. | 6 |
| Figure 1.7: Human ATG8 homologues structure comparison shown from the front (A) and top (B). Proteins shown are LC3A (yellow), LC3B (pink), LC3C (grey), GABARAP (green), GABARAPL1 (blue) and GABARAPL2 (magenta). | 7 |
| Figure 1.8: Structure of p62 LIR bound to LC3B. Hydrophobic pocket 1 (HP1) and 2 (HP2) are shown in red and blue respectively. Trp338 and Leu341 are shown in green with the rest of the LIR coloured yellow. | 8 |
| Figure 1.9: WebLogo from Terje Johansen and Trond Lamark paper [64] showing sequence content of 100 LIRs in bits (upper panel) and residue probabilities at each position (lower panel). Hydrophobic pockets are marked by HP1 and HP2. | 11 |
| Figure 1.10: Domain architecture p62 and NBR1 members of the sequestosome-1-like receptor group of selective autophagy receptors. PB1, Phox and Bem1 domain; (pink), ZZ, ZZ-type zinc finger domain (light blue); CC, coiled-coil domain (orange); FW, four tryptophan domain (green); LIR, LC3 interacting region (dark blue); UBA, ubiquitin-associated domain (dark blue); NLS, Nuclear localisation signals (grey); NES, nuclear export signal (grey); KIR, Keap interacting region (grey). Images are not to scale and represent relative positions of domains. SAR size (number of amino acids) is given far right. | 13 |
| Figure 1.11: Structure of NBR1 LIR1 bound to GABARAPL1. LIR is shown in green with Y732 and I735 coloured purple. Amino acids on GABARAPL1 that were shown to be close in the NMR data to Y732 are coloured blue and those close to I735 are coloured red. | 14 |
| Figure 1.12: Phosphorylation sites on LC3A/B. Serine's and threonine's that can be phosphorylated are shown in blue. Where the amino acid present in LC3A is different from LC3B the LC3A sequence is given first (LC3A/LC3B). | 16 |
| Figure 2.1: pGEX-4T-3 Plasmid Map (GE Healthcare Life Sciences)..... | 19 |
| Figure 2.2: Sequences for the hATG8 proteins as supplied in the pETM-30 plasmid | 21 |
| Figure 2.3: pGEX-4T-3 Plasmid Map (Amersham). | 23 |

| | |
|--|----|
| Figure 2.4: Magnetisation transfer for triple resonance backbone assignment NMR experiments. Nuclei coloured in green give signals in the 3D NMR spectrum. | 33 |
| Figure 2.5: Example ¹³ C CA and CB chemical shifts of alanine, glycine, serine and threonine amino acids. Data Shown is from a HNCACB experiment. | 34 |
| Figure 2.6: Examples of exchange regimes of peaks in a ¹ H- ¹⁵ N HSQC whilst carrying out a titration of LC3B with NBR1 LIRs. Peaks are shown exhibiting fast (A), intermediate (B) and slow (C) exchange. Start and end points of titration are indicated..... | 36 |
| Figure 3.1: 3D structure of LC3B showing the location of the KL mutation in relation to the hydrophobic pocket and p62 LIR. | 40 |
| Figure 3.2: 3D structure of LC3B showing location of Thr29 in relation to the hydrophobic pockets and p62 LIR. | 40 |
| Figure 3.3: SDS PAGE for LC3A and LC3B 1L over-expression in LB; lane annotations: M = marker, 1 = LC3A uninduced, 2 = LC3A after induction - total, 3 = LC3A after induction – soluble, 4 = LC3B uninduced, 5 = LC3B after induction - total, 6 = LC3B after induction – soluble. Expression bands for LC3A and LC3B are circled in orange. | 41 |
| Figure 3.4: Agarose Gel showing cut DNA from pETM30 plasmid; lane annotations M = marker, 1 = LC3A, 2 = LC3C, 3 = GABARAP, 4 = GABARAPL1. | 42 |
| Figure 3.5: SDS PAGE gel of hATG8 test growths for LC3B, LC3C, GABARAP and GABARAPL1. Labels for lanes are as follows: M= marker, 1=uninduced, 2=total cell lysate, 3 hours after induction, 3=soluble cell lysate, 3 hours after induction, 4=total cell lysate, 21 hours after induction and 5=soluble cell lysate, 21 hours after induction. Band of interest is indicated with red arrow. | 43 |
| Figure 3.6: Ion exchange chromatogram of LC3B T29D. A ₂₈₀ is shown in blue, and conductivity in red. The protein of interest is indicated by the red arrow. | 45 |
| Figure 3.7: Desalt chromatogram of LC3B KL. A ₂₈₀ is shown in blue, and conductivity in red. | 45 |
| Figure 3.8: SDS PAGE gel monitoring the various steps in production and purification of LC3B T29D. Labels for lanes are as follows: M=marker, 1 = uninduced, 2 = after induction - total, 3 = after induction – soluble, 4 = cleaved protein from column, 5 = supernatant after cell extraction and 6,7,8 = protein containing fractions from ion exchange column..... | 46 |
| Figure 3.9: Example of native MS spectrum of LC3B (A) and denatured (B)..... | 47 |
| Figure 3.10: ¹ H NMR of LC3B. | 49 |
| Figure 4.1: Sequences for LIRs of NBR1 (LIR1 top and LIR2 bottom). Amino acids that bind to the hydrophobic pockets of LC3B are shown in red and residues that directly precede the aromatic residue are shown in green. | 51 |
| Figure 4.2: Mass spectra of LC3A bound to LIRs of NBR1. Spectrum shown are LC3A-LIR1 (A) and LC3A-LIR2 (B). Key indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are. | 53 |
| Figure 4.3: MS competition experiment for mutants of LC3B with LIR1. Sample contained LC3B WT, KL and T29D (¹⁵ N) and NBR1 LIR1 peptide in equal concentrations (10 μM). Key | |

| | |
|--|----|
| indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are..... | 56 |
| Figure 4.4: MS competition experiment for mutants of LC3B with LIR1. Sample contained LC3B WT, KL and T29D (¹⁵ N) and NBR1 LIR1 peptide. Top spectrum shows proteins with 0.5 equivalent of LIR (1:1:1:0.5) and bottom spectrum shows 3 equivalents (1:1:1:3). Key indicates the peaks that belong to each protein and protein-LIR complex. | 57 |
| Figure 4.5: Plot of ratios of bound to unbound protein against increasing equivalents of NBR1 LIR1 peptide. | 58 |
| Figure 4.6: MS competition experiment for mutants of LC3B with LIR2. Sample contained LC3B WT, KL and T29D (¹⁵ N) and NBR1 LIR2 peptide in equal concentrations (10 μM). Key indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are found. | 59 |
| Figure 4.7: MS competition experiment for mutants of LC3B with LIR2. Sample contained LC3B WT, KL and T29D (¹⁵ N) and NBR1 LIR1 peptide. Top spectrum shows proteins with 0.5 equivalent of LIR (1:1:1:0.5) and right bottom spectrum shows 3 equivalents (1:1:1:3). Key indicates the peaks that belong to each protein and protein-LIR complex. | 60 |
| Figure 4.8: Plot of ratios of bound to unbound protein against increasing equivalents of NBR1 LIR2 peptide. | 61 |
| Figure 4.9: MS competition experiment for hATG8 LC3 family NBR1 LIR1 peptide. Sample contained LC3A, LC3B, LC3C and NBR1 LIR1 peptide in equal concentrations (10 μM). Key indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are found..... | 63 |
| Figure 4.10: Bar chart of calculated values for ratio of bound to unbound for LC3 subfamily with LIR1 MS competition experiment. | 63 |
| Figure 4.11: MS competition experiment for hATG8 LC3 family NBR1 LIR2 peptide. Sample contained LC3A, LC3B, LC3C and NBR1 LIR2 peptide in equal concentrations (10 μM). Key indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are found. | 64 |
| Figure 4.12: Bar chart of calculated values for ratio of bound to unbound for LC3 subfamily with LIR2 MS competition experiment. | 65 |
| Figure 4.13: MS competition experiment for hATG8 GABARAP family NBR1 LIR1 peptide. Sample contained GABARAP, GABARAPL1, GABARAPL2 and NBR1 LIR1 peptide in equal concentrations (10 μM). Key indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are found..... | 65 |
| Figure 4.14: Bar chart of calculated values for ratio of bound to unbound for GABARAP subfamily with LIR1 MS competition experiment. | 66 |
| Figure 4.15: MS competition experiment for hATG8 GABARAP family NBR1 LIR2 peptide. Sample contained GABARAP, GABARAPL1, GABARAPL2 and NBR1 LIR2 peptide in equal concentrations (10 μM). Key indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are found..... | 67 |

| | |
|---|----|
| Figure 4.16: Bar chart of calculated values for ratio of bound to unbound for GABARAP subfamily with LIR2 MS competition experiment. | 68 |
| Figure 4.17: Surface representations of LC3B mutants with p62 LIR shown bound. LC3B KL is shown on the left and LC3B T29D on the right. Mutation is shown in red. | 69 |
| Figure 4.18: Sequences of LC3 proteins. Parts of the sequence where all three LC3s have different amino acids are shaded green and where LC3A, LC3B and LC3C are different are shaded pink, yellow and blue respectively. | 70 |
| Figure 4.19: Structure of LC3B WT bound to p62 LIR coloured to show differences in sequences between LC3s. Parts of the sequence where all three LC3s have different amino acids are shaded green and where LC3A, LC3B and LC3C are different are shaded pink, yellow and blue respectively. | 71 |
| Figure 4.20: Sequences of GABARAP proteins. Parts of the sequence where all three GABRAPs have different amino acids are shaded green and where GABARAP, GABARAPL1 and GABARAPL2 are different are shaded pink, yellow and blue respectively. | 71 |
| Figure 4.21: Surface of GABARAPL1 bound to NBR1 LIR1 coloured to show where changes in sequence between the GABARAPs occur. Parts of the sequence where all three GABARAPs have different amino acids are shaded green and where GABARAP, GABARAPL1 and GABARAPL2 are different are shaded pink, yellow and blue respectively. | 72 |
| Figure 5.1: Sequences for LIRs of p62 and NBR1. Amino acids that bind to the hydrophobic pockets of LC3B are shown in red and acidic residues that directly precede the aromatic residue are shown in green. Amino acids that are greyed out were not included in peptide sequence used but are included for reference. | 74 |
| Figure 5.2: Example of linking C α and C β peaks using HNCACB (green and blue peaks) and HN(CO)CACB (pink and orange peaks). | 76 |
| Figure 5.3: LC3B WT ^1H - ^{15}N HSQC backbone assignment. | 77 |
| Figure 5.4: Overlay of a selection of ^1H - ^{15}N HSQC spectrum for titration of LC3B WT and NBR1 LIR1 peptide. Spectra shown are the ratios of LC3B WT to LIR 1:0 (red), 1:0.5 (orange), 1:1 (green) and 1:2 (blue). | 78 |
| Figure 5.5: ^1H - ^{15}N HSQC backbone assignment of NBR1 LIR1 peptide bound to LC3B WT. | 80 |
| Figure 5.6: Graph of CSP values for binding of NBR1 LIR1 peptide to LC3B. The light blue line indicates the mean, the blue line is the mean and one standard deviation and the dark blue line the mean and two standard deviations. | 81 |
| Figure 5.7: CSP mapping for the binding of NBR1 LIR1 peptide to LC3B. Residues are coloured purple if the CSP is greater than the mean, light blue if greater than the mean and one standard deviation and blue if above the mean and two standard deviations. Residues with no significant shift are coloured grey. | 81 |
| Figure 5.8: HADDOCK structure of NBR1 LIR1 peptide bound to LC3B. The peptide is coloured yellow with the active residues Tyr733 and Ile736 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation. P55 on LC3B WT is coloured black for reference. | 82 |

Figure 5.9: Overlay of a selection of ^1H - ^{15}N HSQC spectrum for titration of LC3B WT and NBR1 LIR2 peptide. Spectrum shown are the ratios of LC3B WT to LIR 1:0 (red), 1:0.5 (orange), 1:1 (green) and 1:2 (blue).83

Figure 5.10: Graph of CSP values for binding of NBR1 LIR2 peptide to LC3B. The light blue line indicates the mean, the blue line is the mean and one standard deviation and the dark blue line the mean and two standard deviations.84

Figure 5.11: CSP mapping for the binding of NBR1 LIR2 peptide to LC3B. Residues are coloured purple if the CSP is greater than the mean, light blue if greater than the mean and one standard deviation and blue if above the mean and two standard deviations. Residues with no significant shift are coloured grey.85

Figure 5.12: HADDOCK structure of NBR1 LIR2 peptide bound to LC3B. The peptide is coloured yellow with the active residues Phe556 and Leu566 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation. P55 on LC3B WT is coloured black for reference.86

Figure 5.13: Second highest scoring HADDOCK cluster with structure of reverse NBR1 LIR2 peptide bound to LC3B. The peptide is coloured yellow with the active residues Phe556 and Leu566 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation. P55 on LC3B WT is coloured black for reference.87

Figure 5.14: Overlay of a selection of ^1H - ^{15}N HSQC spectra for titration of LC3B WT and NBR1 LIR2 peptide. Spectra shown are the ratios of LC3B WT to LIR 1:0 (red), 1:0.5 (orange), 1:1 (green) and 1:2 (blue).88

Figure 5.15: Graph of CSP values for binding of p62 LIR peptide to LC3B. The light blue line indicates the mean, the blue line is the mean and one standard deviation and the dark blue line the mean and two standard deviations.89

Figure 5.16: CSP mapping for the binding of p62 LIR peptide to LC3B. Residues are coloured light purple if the CSP is greater than the mean, purple if greater than the mean and one standard deviation and blue if above the mean and two standard deviations. Residues with no significant shift are coloured grey.90

Figure 5.17: HADDOCK structure of p62 LIR peptide bound to LC3B. The peptide is coloured yellow with the active residues Trp338 and Leu341 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation. P55 on LC3B WT is coloured black for reference.90

Figure 5.18: Comparison of p62 LIR binding to LC3B with crystal structure in green[61] and the top scoring HADDOCK structure in yellow.91

| | |
|---|-----|
| Figure 5.19: Overlay of ^1H - ^{15}N HSQC spectra for end points of LC3B WT and LIR titrations. Spectra shown are LC3B WT (blue), LC3B WT and NBR1 LIR1 (magenta), LC3B WT and NBR1 LIR2 (green) and LC3B WT and p62 LIR (orange). | 92 |
| Figure 5.20: HADDOCK structure of NBR1 LIR1 bound to LC3B indicating interactions of the LIR with the face of LC3B. LC3B β -2 is shown in blue and LIR in gold, amide groups are coloured red and oxygens green. Potential hydrogen bonds between β -sheet and LIR are circled (purple). | 93 |
| Figure 5.21: Structure of LC3B showing knock on effect of LIR binding to the hydrophobic pockets to the secondary structure. Residues that have significant shifts are coloured in red. | 94 |
| Figure 5.22: Plot of differences between CSPs for titrations LC3B WT with NBR1 LIR1 and p62 LIR. Significant differences are shown in blue (mean + one standard deviation) and most significant differences is shown in dark blue (mean + two standard deviations). | 96 |
| Figure 5.23: A: Graphical representation of CSPs that are significantly different between the titration with p62 (blue) and NBR1 LIR1 (orange). B: Overlay of HADDOCK structure for the binding of p62 LIR (yellow) and NBR1 LIR1 (pink) on the surface of LC3B. Significant shifts are coloured from light blue and most significant shifts are coloured dark blue. | 97 |
| Figure 5.24: Plot of differences between CSPs for titrations LC3B WT with NBR1 LIR2 and p62 LIR. Value for significant differences (mean and one significant difference) is shown in blue and most significant differences (mean and two significant difference) is shown in dark blue. | 99 |
| Figure 5.25: A: Graphical representation of CSPs that are significantly different between the titration with p62 (blue) and NBR1 LIR2 (orange). B: Overlay of HADDOCK structure for the binding of p62 LIR (yellow) and NBR1 LIR2 (green) on the surface of LC3B. Significant shifts are coloured light blue and the most significant shifts are coloured dark blue. | 100 |
| Figure 5.26: Plot of differences between CSPs for titrations LC3B WT with NBR1 LIR1 and p62 LIR. Value for significant differences (mean and one significant difference) is shown in blue and most significant differences (mean and two significant difference) is shown in dark blue. | 101 |
| Figure 5.27: A: Graphical representation of CSPs that are significantly different between the titration with NBR1 LIR1 (blue) and NBR1 LIR2 (orange). B: Overlay of HADDOCK structure for the binding of NBR1 LIR1 (pink) and NBR1 LIR2 (green) on the surface of LC3B. Significant shifts are coloured light blue with most significant shifts coloured dark blue. | 103 |
| Figure 6.1: Overlay of LC3B WT (blue) and LC3B KL (pink) ^1H - ^{15}N HSQC. | 106 |
| Figure 6.2: ^1H - ^{15}N HSQC backbone amide assignment of LC3B KL. | 107 |
| Figure 6.3: Overlay of ^1H - ^{15}N HSQC spectrum of LC3B WT (red) and LC3B T29D (blue). | 108 |
| Figure 6.4: Overlay of a selection of ^1H - ^{15}N HSQC NMR spectra for the titration of LC3B KL with NBR1 LIR1 (A) and NBR1 LIR2 (B). Spectra shown are LC3B KL with zero (red), 0.5 (green), 1 (blue) and 2 (purple) equivalents of NBR1 LIR peptide. | 110 |
| Figure 6.5: Graph of CSP values for binding of NBR1 LIR1 (A) and LIR2 (B) peptide to LC3B KL. The light blue line indicates the mean, the blue line is the mean and one standard deviation and the purple line the mean and two standard deviations. | 111 |

Figure 6.6: CSP mapping for the binding of NBR1 LIR1 (A) and LIR2 (B) peptide to LC3B KL mutant. Residues are coloured light purple if the CSP is greater than the mean, purple if greater than the mean and one standard deviation and blue if above the mean and two standard deviations. Residues with no significant shift are coloured grey.....113

Figure 6.7: Highest scoring HADDOCK structure of NBR1 LIR1 peptide binding to LC3B KL. The peptide is coloured yellow with the active residues Tyr732 and Ile735 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.114

Figure 6.8: Highest scoring HADDOCK structure of NBR1 LIR2 peptide binding to LC3B KL. The peptide is coloured yellow with the active residues Phe563 and Leu566 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.115

Figure 6.9: Highest scoring HADDOCK structure from second cluster of NBR1 LIR2 peptide binding to LC3B KL. The peptide is coloured yellow with the active residues Phe563 and Leu566 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.116

Figure 6.10: Overlay of a selection of ^1H - ^{15}N HSQC NMR spectra for the titration of LC3B T29D with NBR1 LIR1 (A) and LIR2 (B) peptide. Spectra shown are LC3B T29D with zero (red), 0.5 (green), 1 (blue) and 2 (purple) equivalents of NBR1 LIR1 peptide.....119

Figure 6.11: Graph of CSP values for binding of NBR1 LIR1 (A) and LIR2 (B) peptide to LC3B T29D. The light blue line indicates the mean, the blue line is the mean and one standard deviation and the purple line the mean and two standard deviations.120

Figure 6.12: CSP mapping for the binding of NBR1 LIR1 (A) and LIR2 (B) peptide to LC3B T29D mutant. Residues are coloured light purple if the CSP is greater than the mean, purple if greater than the mean and one standard deviation and blue if above the mean and two standard deviations. Residues with no significant shift are coloured grey.....121

Figure 6.13: Highest scoring HADDOCK structure of NBR1 LIR1 peptide binding to LC3B T29D. The peptide is coloured yellow with the active residues Tyr732 and Ile735 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.122

Figure 6.14: Highest scoring HADDOCK structure of NBR1 LIR2 peptide binding to LC3B T29D. The peptide is coloured yellow with the active residues Phe563 and Leu566 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.123

Figure 6.15: Graph of CSPs calculated between LC3B WT and KL ¹H-¹⁵N HSQC spectrum. Lines represent significance cut offs and the values are mean + one standard deviation (light blue) and mean + two standard deviations (dark blue).....124

Figure 6.16: Significant CSP values mapped onto surface of LC3B. KL mutation shown in purple, significant shift shown in blue and most significant in red.125

Figure 6.17: Changes in the HSQC between LC3B WT and LC3B T29D mapped onto the surface of LC3B. Mutation site shown in red.....126

Figure 6.18: Significant CSPs for titrations of LC3B with NBR1 LIR1 titrations mapped on to LC3B WT structure. Those coloured green appear as most significant amino acids in all titrations. Those coloured yellow are significant or most significant in all titrations. Those coloured purple appear in at least one titration.127

Figure 6.19: Plot of differences between CSPs for titrations LC3B WT and LC3B KL with NBR1 LIR1. Value for significant differences is shown in blue and most significant differences is shown in dark blue.....129

Figure 6.20: A: Graphical representation of CSPs that are significantly different between the titration with LC3B WT and NBR1 LIR1 (blue) and LC3B KL and NBR1 LIR1 (orange). B: Surface of LC3B WT with significant differences between titration of LC3B WT and LC3B KL with LIR1 mapped. Significant and most significant differences are coloured light blue and blue respectively. HADDOCK structure for LIR1 position in binding site are shown for LC3B WT (green) and LC3B KL (red).131

Figure 6.21: Plot of differences between CSPs for titrations LC3B WT and LC3B KL with NBR1 LIR2. Value for significant differences is shown in blue and most significant differences is shown in dark blue.....132

Figure 6.22: A: Graphical representation of CSPs that are significantly different between the titration with LC3B WT and NBR1 LIR2 (blue) and LC3B KL and NBR1 LIR2 (orange). B: Surface of LC3B WT with significant differences between titration of LC3B WT and LC3B KL with LIR2 mapped. Significant and most significant differences are coloured light blue and blue respectively. HADDOCK structure for LIR2 position in binding site are shown for LC3B WT (green) and LC3B KL (red).134

Figure 6.23: Plot of differences between CSPs for titrations NBR1 LIR1 and LIR2 with LC3B KL. Value for significant differences is shown in blue and most significant differences is shown in dark blue.136

Figure 6.24: A: Graphical representation of CSPs that are significantly different between the titration with LC3B KL and NBR1 LIR1 (blue) and NBR1 LIR2 (orange). B: Surface of LC3B KL with significant differences between titration of LC3B KL with NBR1 LIR1 and LIR2. Significant and most significant differences are coloured light blue and blue respectively. HADDOCK structure for LIR position in binding site are shown for LIR1 (green) and LIR2 (red).138

Figure 6.25: Plot of differences between CSPs for titrations LC3B WT and LC3B T29D with NBR1 LIR1. Value for significant differences is shown in blue and most significant differences is shown in dark blue.139

Figure 6.26: A: Graphical representation of CSPs that are significantly different between the titration with LC3B WT and NBR1 LIR1 (blue) and LC3B T29D and NBR1 LIR1 (orange). B:

Surface of LC3B WT with significant differences between titration of LC3B WT and LC3B T29D with LIR1 mapped. Significant and most significant differences are coloured purple and blue respectively. HADDOCK structure for LIR1 position in binding site are shown for LC3B WT (green) and LC3B T29D (red).141

Figure 6.27: Plot of differences between CSPs for titrations LC3B WT and LC3B T29D with NBR1 LIR2. Value for significant differences is shown in blue and most significant differences is shown in dark blue.142

Figure 6.28: A: Graphical representation of CSPs that are significantly different between the titration with LC3B WT and NBR1 LIR2 (blue) and LC3B T29D and NBR1 LIR2 (orange). B: Surface of LC3B WT with significant differences between titration of LC3B WT and LC3B T29D with LIR2 mapped. Significant and most significant differences are coloured light blue and blue respectively. HADDOCK structure for LIR2 position in binding site are shown for LC3B WT (green) and LC3B T29D (red). The top image shows the front face of LC3B WT with the bottom image being side on.....144

Figure 6.29: Plot of differences between CSPs for titrations NBR1 LIR1 and LIR2 with LC3B T29D. Value for significant differences is shown in blue and most significant differences is shown in dark blue.....145

Figure 6.30: A: Graphical representation of CSPs that are significantly different between the titration with LC3B T29D and NBR1 LIR1 (blue) and NBR1 LIR2 (orange). B: Surface of LC3B T29D with significant differences between titration of LC3B T29D with NBR1 LIR1 and LIR2. Significant and most significant differences are coloured light blue and blue respectively. HADDOCK structure for LIR position in binding site are shown for LIR1 (green) and LIR2 (red). The top image shows the front face of LC3B T29D with the bottom image being side on...147

Figure 7.1: Sequences of LC3A and LC3B. Differences in the sequences are highlighted with green denoting a change to a similar type of amino acid and blue denoting a change to an amino acid of a different type.149

Figure 7.2: Overlay of LC3A and LC3B Structure. LC3A is shown in red (PDB structure 3WAL) and LC3B is shown in green (PDB structure 3VTU).....150

Figure 7.3: Differences between LC3A and LC3B sequences mapped onto the structure of LC3B. Differences in the sequences are highlighted with green denoting a change to a similar type of amino acid and blue denoting a change to an amino acid of a different type.150

Figure 7.4: Example of linking C_{α} and C_{β} peaks using HNCACB (blue and red peaks) and HN(CO)CACB (pink and green peaks).....151

Figure 7.5: ^1H - ^{15}N HSQC assignment of LC3A WT.....153

Figure 7.6: A selection ^1H - ^{15}N HSQC's from the titration of NBR1 LIR1 Peptide with LC3A WT. Where red is 0 equivalents of peptide and blue is 2 equivalents. Orange, pink and green are 0.3, 0.5 and 0.8 equivalents of peptide respectively.....154

Figure 7.7: Plot of CSPs for titration of LC3A WT with peptide representing LIR1 of NBR1. The light blue line is the mean value, the blue is the mean plus one standard deviation and the dark blue line is the mean and two standard deviations.155

Figure 7.8: CSPs for NMR titration of LC3A with peptide representing NBR1 LIR1 mapped to structure of LC3A. Amino acids with CSPs with a value greater than the mean, mean and

| | |
|--|-----|
| standard deviation and mean and two standard deviations are coloured light blue, purple and dark blue respectively..... | 155 |
| Figure 7.9: Highest scoring HADDOCK structure of NBR1 LIR1 peptide bound to LC3A. The peptide is coloured green with the active residues Tyr733 and Ile736 coloured yellow. The surface of LC3A is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation. | 156 |
| Figure 7.10: Alternative HADDOCK structure of NBR1 LIR1 peptide bound to LC3A. The peptide is coloured green with the active residues Tyr733 and Ile736 coloured yellow. The surface of LC3A is coloured to indicate the significant shifts seen in the bound form. Residues light blue had CSPs greater than the mean and one standard deviation and those coloured blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation..... | 157 |
| Figure 7.11: A selection ^1H - ^{15}N HSQC's from the titration of NBR1 LIR2 Peptide with LC3A WT. Where red is 0 equivalents of peptide and blue is 2 equivalents. Orange, pink and green are 0.3, 0.5 and 0.8 equivalents of peptide respectively..... | 160 |
| Figure 7.12: Plot of CSPs for titration of LC3A WT with peptide representing LIR2 of NBR1. The light blue line is the mean value, the blue is the mean plus one standard deviation and the dark blue line is the mean and two standard deviations. | 161 |
| Figure 7.13: CSPs for NMR titration of LC3A with peptide representing NBR1 LIR2 mapped to structure of LC3A. Amino acids with CSPs with a value greater than the mean, mean and standard deviation and mean and two standard deviations are coloured light blue, purple and dark blue respectively..... | 161 |
| Figure 7.14: Highest scoring HADDOCK structure of NBR1 LIR2 peptide bound to LC3A. The peptide is coloured green with the active residues Phe563 and Leu566 coloured yellow. The surface of LC3A is coloured to indicate the significant shifts seen in the bound form. Residues light blue had CSPs greater than the mean and one standard deviation and those coloured blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation..... | 162 |
| Figure 7.15: Surface of LC3A WT with amino acids that are un-assigned in the ^1H - ^{15}N HSQC shown in purple. | 163 |
| Figure 7.16: ^1H - ^{15}N HSQC of LC3A WT showing start (blue) and endpoints for titration with NBR1 LIR1 (purple) and LIR2 (red) peptides. | 164 |
| Figure 7.17: Significant CSPs mapped onto the protein surface for the titration of LC3A (left) and LC3B (right) with NBR1 LIR1 peptide. Significant shifts shown in light blue and most significant shifts shown in dark blue..... | 165 |
| Figure 7.18: Plot of difference in CSPs between the titration of NBR1 LIR1 peptide with LC3A and LC3B. The blue is the mean plus one standard deviation and the dark blue line is the mean and two standard deviations. | 166 |
| Figure 7.19: A: Graphical representation of CSPs that are significantly different between the titration with LC3A and NBR1 LIR1 (blue) and LC3B and NBR1 LIR1 (orange). B: Amino acids with significant CSP differences between titration of NBR1 LIR1 peptide with LC3A and LC3B | |

| | |
|--|-----|
| mapped onto the surface of LC3B. NBR1 LIR1 peptide is shown as predicted from LC3B with LIR1 titration data. Significant shifts are coloured light blue and most significant shifts dark blue. | 168 |
| Figure 7.20: Significant CSPs mapped onto the protein surface for the titration of LC3A (left) and LC3B (right) with NBR1 LIR2 peptide. Significant shifts shown in light blue and most significant shifts shown in dark blue..... | 169 |
| Figure 7.21: Plot of difference in CSPs between the titration of NBR1 LIR2 peptide with LC3A and LC3B. The blue is the mean plus one standard deviation and the dark blue line is the mean and two standard deviations. | 170 |
| Figure 7.22: A: Graphical representation of CSPs that are significantly different between the titration with LC3A and NBR1 LIR2 (blue) and LC3B and NBR1 LIR2 (orange). B: Significant differences in CSP between LC3A-LIR2 and LC3B-LIR2 titrations mapped onto the surface of LC3B. HADDOCK structure of NBR1 LIR2 peptide shown bound to LC3A (green) and LC3B (pink). Amino acids with significant differences in CSPs are coloured light blue and most significant differences are coloured dark blue. | 172 |
| Figure 7.23: Significant CSPs mapped onto the protein surface for the titration of NBR1 LIR1 (left) and LIR2 (right) with LC3A WT. Significant shifts shown in light blue and most significant shifts shown in dark blue..... | 173 |
| Figure 7.24: Plot of difference in CSPs between the titration of LC3A WT with NBR1 LIR1 and LIR2 peptide. The blue line is the mean plus one standard deviation and the dark blue line is the mean and two standard deviations..... | 174 |
| Figure 7.25: A: Graphical representation of CSPs that are significantly different between the titration with LC3A and NBR1 LIR1 (blue) and NBR1 LIR2 (orange). B: Surface of LC3A WT with significant differences in CSP between the NMR titrations of LIR1 and LIR2 with LC3A. Significant shifts are coloured light blue and most significant shifts dark blue. NBR1 LIR2 HADDOCK structure for titration is shown in green. | 176 |
| Figure 8.1: HADDOCK structures for NBR1 LIR1 (A) and LIR2 (B) when bound to LC3B WT (green), LC3B KL (blue) and LC3B T29D (pink). | 179 |
| Figure 8.2: Comparison of position of LIR2 when binding to LC3B WT (green) and LC3B KL (blue). | 180 |
| Figure 8.3: Comparison of position of LIR2 when binding to LC3B WT (green) and LC3B T29D (blue). | 181 |
| Figure 8.4: Significant CSPs mapped onto the protein surface for the titration of LC3A (left) and LC3B (right) with NBR1 LIR1 (A) and LIR2 (B). Significant shifts shown in light blue and most significant shifts shown in dark blue. | 183 |
| Figure 8.5: HADDOCK structure of NBR1 LIR1 peptide bound to LC3A. The peptide is coloured green with the active residues Tyr733 and Ile736 coloured yellow. The surface of LC3A is coloured to indicate the significant shifts seen in the bound form. Residues light blue had CSPs greater than the mean and one standard deviation and those coloured blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation. | 184 |

Figure 8.6: HADDOCK structure of NBR1 LIR2 peptide bound to LC3A. The peptide is coloured green with the active residues Phe563 and Leu566 coloured yellow. The surface of LC3A is coloured to indicate the significant shifts seen in the bound form. Residues light blue had CSPs greater than the mean and one standard deviation and those coloured blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.185

List of Tables

| | |
|---|-----|
| Table 1.1: Sequence similarity between the human ATG8 proteins. Figure given as a percentage. | 7 |
| Table 1.2: Mammalian selective autophagy receptors and their LIR sequences. Amino acids critical for interaction with ATG8 family proteins are shown in red. | 10 |
| Table 1.3: Selection of mammalian autophagy proteins with LIRs. | 11 |
| Table 2.1: Buffer and stock solution composition. | 18 |
| Table 2.2: Peptide sequences for LIRs of p62 and NBR1 used in this work. | 21 |
| Table 2.3: PCR conditions. | 23 |
| Table 2.4: contents of minimal media. Made up in 1 L of MilliQ water. | 25 |
| Table 2.5: Values for settings on SYNAPT MS. | 28 |
| Table 2.6: Values for settings on Bruker Impact II MS. | 28 |
| Table 2.7: Points used for a standard titrations of ATG8 protein with LIR peptide. | 29 |
| Table 2.8: Values for settings for triple resonance NMR experiments. | 32 |
| Table 3.1: Primers used to introduce <i>XhoI</i> and <i>BamHI</i> cut sites into hATG8 containing pETM30 plasmid. | 42 |
| Table 3.2: Masses of hATG8 proteins from MS data. | 48 |
| Table 4.1: Predicted values for free LC3A and LC3A NBR1 LIR complex compared to peaks in ESI-MS spectrum. | 54 |
| Table 4.2: Predicted values for peaks in MS spectrum for competition experiment of LC3B mutants with LIR1 of NBR1. | 54 |
| Table 4.3: Predicted values for MS peaks for ¹⁵ N labelled LC3B T29D. | 55 |
| Table 4.4: Predicted MS peaks for hATG8 protein MS competition peaks. | 62 |
| Table 5.1: Amino acids with significant CSPs and HADDOCK results for titrations of LC3B WT with LIR peptides. Those highlighted green appear as a most significant CSP in all titrations. | 95 |
| Table 5.2: List of significant and most significant differences between the titrations of LC3B WT with NBR1 LIR1 and p62 LIR. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration. | 96 |
| Table 5.3: List of significant and most significant differences between the titrations of LC3B WT with NBR1 LIR2 and p62 LIR. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration. | 99 |
| Table 5.4: List of significant and most significant differences between the titrations of LC3B WT with NBR1 LIR1 and NBR1 LIR2. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration. | 102 |
| Table 6.1: Parameters and results of HADDOCK runs carried out to investigate reverse configuration LIR for LC3B KL titration with NBR1 LIR2 peptide. | 117 |

| | |
|---|-----|
| Table 6.2: Amino acids with significant CSPs for titrations of LC3B mutants with NBR1 LIR peptides. Also included are LC3B significant amino acids for reference. Those highlighted green appear as a most significant CSP in all titrations. Those highlighted yellow appear as any kind of significant shift in all titrations..... | 127 |
| Table 6.3: HADDOCK run results for titrations of LIR peptides with LC3B mutants..... | 128 |
| Table 6.4: List of significant and most significant differences between the titrations of LC3B WT and LC3B KL with LIR1. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration..... | 130 |
| Table 6.5: List of significant and most significant differences between the titrations of LC3B WT and LC3B KL with LIR2. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration..... | 133 |
| Table 6.6: List of significant and most significant differences between the titrations of LC3B LIR1 and LIR2. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration..... | 137 |
| Table 6.7: List of significant and most significant differences between the titrations of LC3B WT and LC3B T29D with LIR1. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration..... | 139 |
| Table 6.8: List of significant and most significant differences between the titrations of LC3B WT and LC3B T29D with LIR2. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration..... | 143 |
| Table 6.9: List of significant and most significant differences between the titrations of LC3B T29D with LIR1 and LIR2. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration..... | 146 |
| Table 7.1: Summary of Haddock runs carried out to probe interaction of NBR1 LIR1 peptide with LC3A. | 158 |
| Table 7.2: Amino acids with significant differences in CSP between the titrations of NBR1 LIR1 peptide with LC3A and LC3B. | 167 |
| Table 7.3: Amino acids with significant differences in CSP between the titrations of NBR1 LIR2 peptide with LC3A and LC3B. | 170 |
| Table 7.4: Amino acids with significant differences in CSP between the titrations of LC3A WT with NBR1 LIR1 and LIR2..... | 175 |
| Table 8.1: Ratios of bound to unbound protein in MS competition experiments for LC3B WT, KL and T29D at a 1:1:1:1 ratio with the relevant LIR. | 178 |
| Table 8.2: Ratios of bound to unbound protein for MS competition experiments for LC3 and GABARAP subfamilies. Proteins were in a 1:1:1:1 ratio with the relevant LIR..... | 182 |

Abbreviations

| | |
|--------------------|---|
| aa | Amino acids |
| AIM | ATG8 interacting motif |
| ALS-FTLD | Amyotrophic lateral sclerosis and frontotemporal lobar degeneration |
| AMPK | Activated protein kinase |
| APS | Ammonium persulfate |
| ATG | Autophagy protein |
| CC | Coiled-coil domain |
| CSP | Chemical shift perturbation |
| ddH ₂ O | Double-distilled water |
| DNA | Deoxyribonucleic acid |
| <i>E.coli</i> | <i>Escherichia coli</i> |
| EDTA | Ethylenediaminetetraacetic acid |
| ER | Endoplasmic reticulum |
| ESI | Electrospray Ionization |
| FIP200 | Family-interacting protein of 200 kDa |
| FPLC | Fast Protein Liquid Chromatography |
| FW | Four tryptophan domain |
| FYCO1 | FYVE and coiled-coil domain-containing protein |
| GABARAP | Gamma-aminobutyric acid receptor-associated protein |
| GIM | GABARAP interaction motif |
| GST | Glutathione S-Transferase |
| hATG8 | Human ATG8 proteins |
| HOPS | Homotypic fusion and protein sorting |
| HP1 | Hydrophobic pocket 1 |
| HP2 | Hydrophobic pocket 2 |
| hsc70 | Heat shock-cognate protein of 70 kDa |
| HSQC | Heteronuclear Single Quantum Coherence |
| IPTG | Isopropyl- β -D-thiogalactoside |
| KIR | Keap interacting region |
| KL | LC3B K51A, L53A mutant |
| KO | Knock out |
| LB | Luria-Bertani broth, (luria broth, Lysogeny broth) |
| LC3 | Microtubule-associated protein 1 Light chain 3 |
| LIR | LC3 Interacting Region |
| M9 | Minimal medium M9 |
| MEF | Mouse-embryonic fibroblasts |
| MS | Mass spectrometry |
| mTOR | Mammalian Target of Rapamycin |
| NBR1 | Next to BRCA1 gene 1 protein |
| NES | Nuclear export signal |
| NLS | Nuclear localisation signals |
| NMR | Nuclear Magnetic Resonance |
| NSF | N-ethylmaleimide-sensitive fusion |
| OD ₅₉₅ | Optical density at 595 nm |
| PB1 | Phox and Bem1 domain |
| PCR | Polymerase chain reaction |

| | |
|----------|--|
| PDB | Protein Data Bank |
| PE | Phosphatidylethanolamine |
| pI | Isoelectric point |
| PLEKHM1 | Pleckstrin homology domain containing protein family member 1 |
| ppm | Parts Per Million |
| PTM | Posttranslational modification |
| qTOF | Quadrupole time of flight |
| RMSD | Root mean square deviation |
| rpm | Revolutions Per Minute |
| SAR | Selective autophagy receptor protein |
| SDS-PAGE | Sodium dodecyl sulphate polyacrylamide gel electrophoresis |
| SNAREs | Soluble <i>N</i> -ethylmaleimide-sensitive fusion attachment protein receptors |
| SSB | Shifted sine bell |
| T29D | LC3B T29D phosphomimetic mutant |
| TAE | Tris-acetate-EDTA |
| TEV | Tobacco Etch virus |
| TMED | Tetramethylethylenediamine |
| TOF | Time of Flight |
| UBA | Ubiquitin-associated domain |
| ULK1 | Unc51-like kinase |
| UV | Ultra violet |
| WT | Wild Type |
| 1D | One-dimensional |
| 2D | Two-dimensional |
| 3D | Three-dimensional |

Amino Acid Abbreviations

| | | |
|---------------|-----|---|
| Alanine | Ala | A |
| Arginine | Arg | R |
| Asparagine | Asn | N |
| Aspartic Acid | Asp | D |
| Cysteine | Cys | C |
| Glutamic Acid | Glu | E |
| Glutamine | Gln | Q |
| 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 |

Chapter 1

1. Introduction

1.1 Autophagy

Autophagy has been a widely studied cellular process since Christian de Duve [1] first discovered the lysosome in 1955. In 1963 he employed the term autophagy [2] (literally meaning self-eating in Greek), for the process of regions of the cytoplasm being sequestered in a membrane and delivered to the lysosome for degradation. The importance of the research in the area was recognised in 2016 with the award of the Nobel Prize in Physiology or Medicine to Yoshinori Ohsumi, for his early identification of the autophagy machinery [3]

Autophagy is a catabolic process responsible for the basal turnover of cellular components within the cytoplasm. In response to certain stress stimuli, autophagy is rapidly upregulated to maintain cellular homeostasis. The most well-known stress response that autophagy is linked to is starvation, where recycling the contents in the cytoplasm replenishes cellular precursors in response to nutrient depletion and removes potentially toxic materials [4]. Autophagy also undertakes an important role in protein turnover and organelle quality control, ensuring appropriate numbers of macromolecules are present in the cytoplasm, and disposing of damaged or dysfunctional organelles [5]. As well as cellular homeostasis, autophagy is a response to cellular stress caused by xenobiotics, oxidants, infectious agents, hypoxia, and chemicals that perturb endoplasmic reticulum (ER) function [6].

There are three types of autophagy. These are microautophagy, chaperone-mediated autophagy and macroautophagy.

Chaperone-mediated autophagy targets individual proteins and transports them to the lysosome for degradation. Proteins that are to be degraded contain a pentapeptide motif, which is recognised by heat shock-cognate protein of 70 kDa (hsc70). Once bound to the chaperone the substrate is targeted to the lysosome for degradation. The selectivity of the process allows the removal of damaged or abnormal proteins and sub units of multi complex proteins, without disturbing the rest of the contents of the cytoplasm [7, 8].

Microautophagy involves the lysosome directly enveloping a portion of the cytoplasm degrading proteins and organelles. This can either be done by lysosomal protrusion, or lysosomal or endosomal invagination. Lysosomal protrusion is where the lysosome either elongates and wraps the cytoplasm, or extends an arm like protrusion to envelope an area of cytoplasm. Lysosomal and endosomal invagination involves part of the lysosome, or endosome folding back on itself to envelop the material to be degraded. Microautophagy can also selectively degrade targeted proteins, by enveloping areas of the cytoplasm that contain proteins that have been transported by hsc70 and NBR1, which are normally associated with chaperone-mediated autophagy and macroautophagy respectively [9, 10].

Finally macroautophagy (which will be referred to as autophagy for the following work) is where an isolation membrane sequesters portions of the cytoplasm containing proteins and

organelles to form an autophagosome. The autophagosome is then transported to the lysosome with which it fuses and the contents are then degraded. Autophagy is a highly conserved process with many mammalian homologues of yeast Atg genes having been identified and characterised [11, 12].

One of the major regulatory components of autophagy is the protein kinase Tor (target of Rapamycin). In Mammals mTor complex 1 (mTORC1) phosphorylates ULK1 (unc51-like kinase), sequestering ULK1 in a complex with ATG13 and FIP200 (family-interacting protein of 200 kDa). During nutrient deprivation, AMPK (activated protein kinase) inhibits mTor reducing ULK1 phosphorylation [13]. The subsequent AMPK-mediated phosphorylation of ULK1, at a different residue then leads to AMPK translocating to the site of autophagosome formation, and the subsequent initiation of autophagy [14, 15] (Figure 1.1).

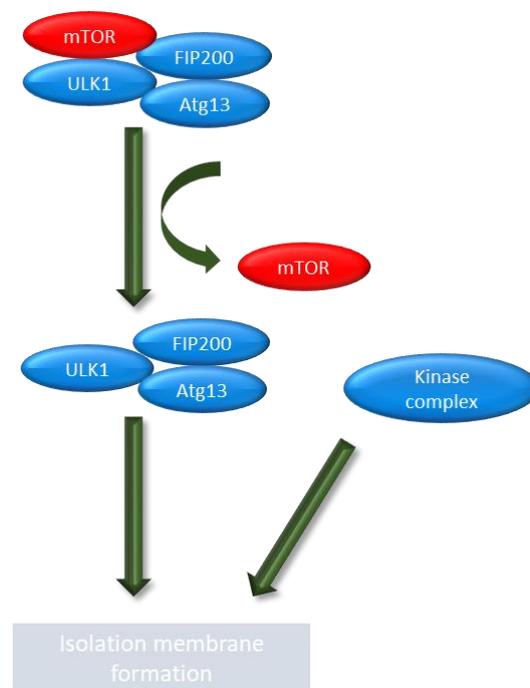


Figure 1.1: mTOR signalling network.

mTOR (in red) phosphorylates ULK1 sequestering it in complex with ATG13 and FIP200.

Phosphorylation of ULK1 leads to translocation to the site of autophagosome formation.

Once autophagy is initiated the initial phagophore is formed. The location of vesicle nucleation is still not fully understood, and numerous sites have been proposed for the formation of the autophagosome membrane, including the Endoplasmic reticulum (ER), mitochondria, Golgi and the plasma membrane as well as it being synthesized *de novo* [16, 17].

Once formed the double membraned phagophore expansion is then mediated by the Atg16L complex localising to the phagophore membrane and directing the site at which lipidation occurs [16]. The ATG16L complex is formed when ATG12 is covalently conjugated to ATG5, through an isopeptide bond through a series of ubiquitination-like reactions, with

ATG7 acting as an E1 like enzyme and ATG10 as an E2 like enzyme. The ATG12-ATG5 conjugate then associates with ATG16L and homo-oligomerises (Figure 1.2).

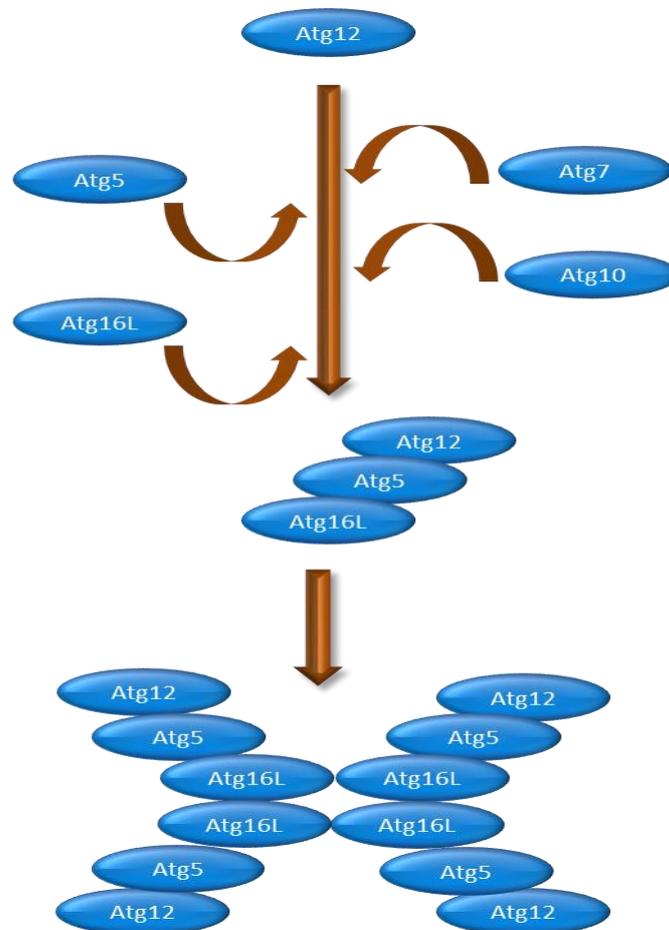


Figure 1.2: Formation of ATG16L complex.
ATG5 is conjugated to ATG12 and the ATG16L. This unit then homo-oligomerises to form the ATG16L complex.

ATG16L complex recruits members of the ATG8 family of proteins (MAP1LC3A (LC3A), MAP1LC3B (LC3B), MAP1LC3C (LC3C), GABARAP, GABARAPL1 and GABARAPL2), which have been conjugated to phosphatidylethanolamine (PE). This then facilitates expansion of the bi-lipid layer of the phagophore. The phagophore is then closed by membrane scission [18], enclosing the sequestered material within the mature autophagosome.

The autophagosome is then transported to the lysosome using the dynein machinery on microtubules [19], through an interaction with FYVE and coiled-coil domain-containing protein (FYCO1), facilitating kinesin-dependant plus-end movement of phagosomes [20].

Once at the lysosome the autophagosome and lysosome fuse to form an autolysosome. The fusion process has been attributed to a number of protein complexes. One process is through Rab7 and small GTPases interacting with LC3s on the surface of the autophagosome, stimulating autophagosome-lysosome fusion [21]. Another set of molecules involved are soluble N-ethylmaleimide-sensitive fusion (NSF) attachment protein receptors (SNAREs) which promote fusion [22]. There is also pleckstrin homology domain

containing protein family member 1 (PLEKHM1), which interacts with ATG8s on the autophagosome, and homotypic fusion and protein sorting (HOPS) tethering complex and mediating autophagosome fusion [23]. Fusion with the lysosome forms the autolysosome, and cargo is degraded by lysosomal hydrolytic enzymes (Figure 1.3).

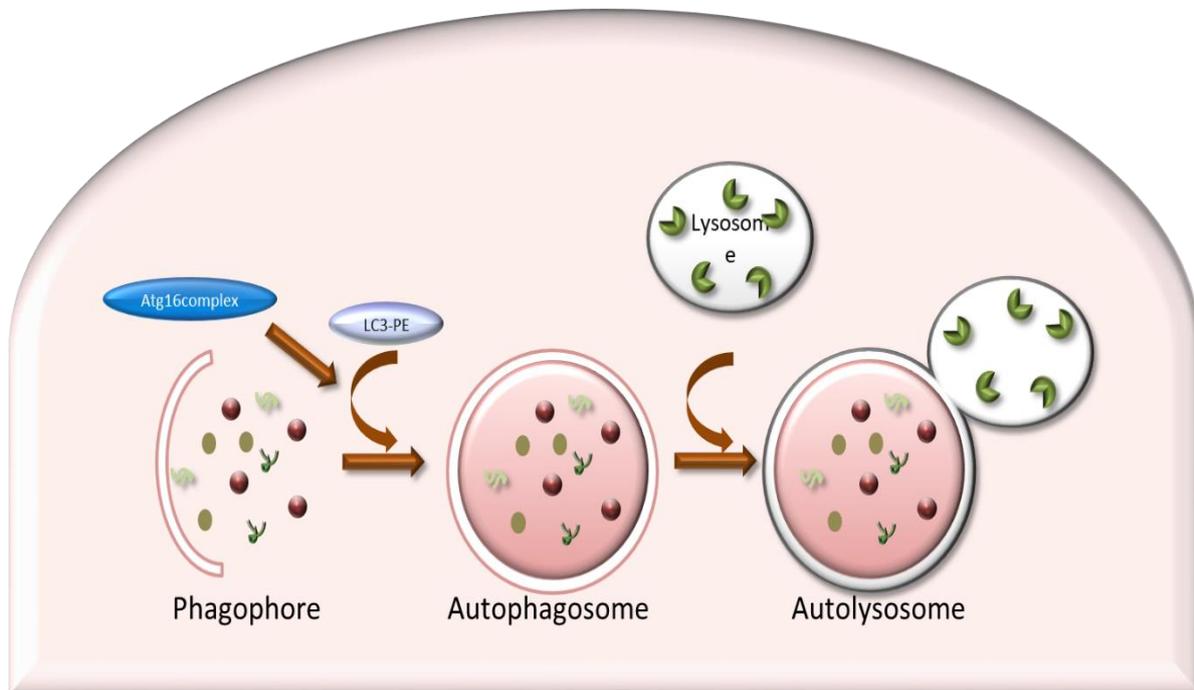


Figure 1.3: Model showing formation of autophagosome and subsequent fusion with the lysosome to form the autolysosome.

As well as the non-selective degradation of bulk material from the cytoplasm, autophagy is also a selective process by which specific materials are transported to the autophagosome for degradation [24-26]. Examples of these materials are aggregated proteins (aggrephagy)[27], mitochondria (mitophagy)[28], peroxisomes (pexophagy)[29] and microbes and viruses (xenophagy)[30]. This process is regulated by interactions between the ATG8 family of proteins, and selective autophagy receptors (SARs), and will be covered in greater details later in this chapter.

Autophagy has been linked with a number of neurodegenerative diseases. Studies in mice have shown that removing important genes involved in autophagy leads to neurodegenerative disease [31, 32]. It has also been shown that p62, a key cargo receptor protein for selective autophagy, appears at reduced levels in Alzheimer's disease patients [33, 34]. Autophagic dysregulation has also been shown in sufferers of Amyotrophic Lateral Sclerosis Frontotemporal lobar degeneration (ALS-FTLD) [35, 36].

1.2 The ATG8 Family

In yeast a single protein Atg8 is essential for phagophore expansion [37]. Whilst yeast has only Atg8 to fill this role, almost all other eukaryotes express at least two ATG8 homologues. These ATG8 homologues are split between two subfamilies, the microtubule-associated protein 1 light chain 3 (LC3s) and γ -aminobutyric acid receptor-associated protein (GABARAPs) [38, 39]. In humans there are six ATG8 (hATG8) homologue proteins with LC3A, LC3B and LC3C in the LC3 subfamily, and GABARAP, GABARAPL1 and GABARAPL2 (also known as GATE-16) in the GABARAP subfamily.

The ATG8 family of proteins are expressed in a full length cytosolic form, which is processed and conjugated to phosphatidylethanolamine (PE), before anchoring into the inner and outer membranes of the phagophore [40-42]. First the ATG8s are cleaved by the cysteine protease ATG4, exposing a C-terminal glycine. ATG7, an E1-like enzyme forms a thioester bond with the carboxyl group of the exposed glycine. The ATG8 glycine residue is then transferred to ATG3 an E2-like enzyme, releasing ATG7. The glycine of the ATG8 is then conjugated to PE localised to the autophagosome, with the ATG12-ATG5-ATG16 complex acting as an E3-like enzyme [43] releasing ATG3. A summary of this process is given in Figure 1.4. The conjugation of PE to the ATG8 proteins is a reversible process, with ATG4 able to act as a deconjugating enzyme [44].

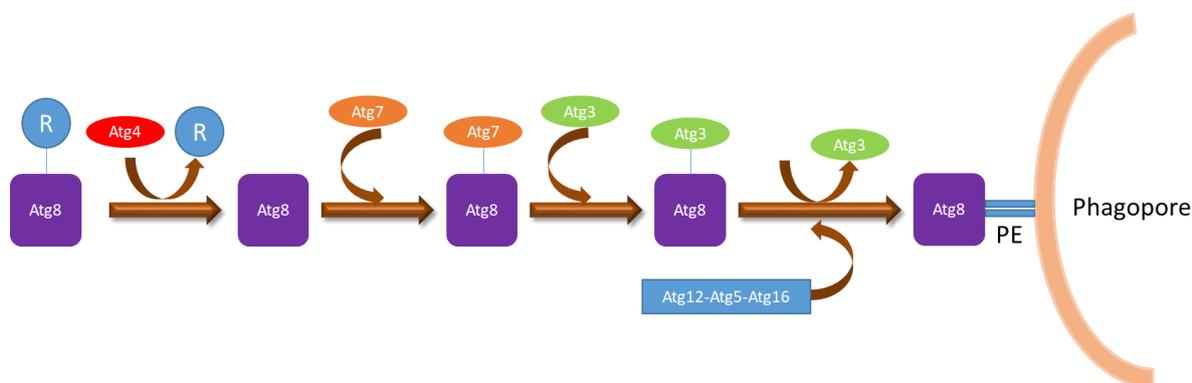


Figure 1.4: Schematic showing process of conjugation of phosphatidylethanolamine to ATG8s

Once the ATG8s are lipidated and anchored on the surface of the phagophore they can then act as adaptors, recruiting other proteins to the inner and outer surfaces of the growing phagophore. The recruited proteins interact with the ATG8 family protein through a LC3 interacting region (LIR), which binds to the face of ATG8 proteins via two hydrophobic pockets [45]. The LIR more recently has also been called an autophagy interacting motif (AIM), GABARAP interacting motif (GIM) and LC3 interacting motif (LIM). However, as the literature still predominately uses LIR then for the purpose of this work the sequence will be referred to as an LIR. Core autophagy components that interact with the ATG8s through a LIR include those involved in ATG8 lipidation, ATG16L [16] and ATG4B [46], in phagophore production like the ULK1-FIP200-ATG13 complex [47] and the ATG-ATG complex [48], SAR proteins like p62 [49] and NBR1 [50] docking cargo to the phagophore, and proteins involved in autophagosome transport and lysosome fusion like FYCO1 [51] and PLEKHM1 [52] (Figure 1.5).

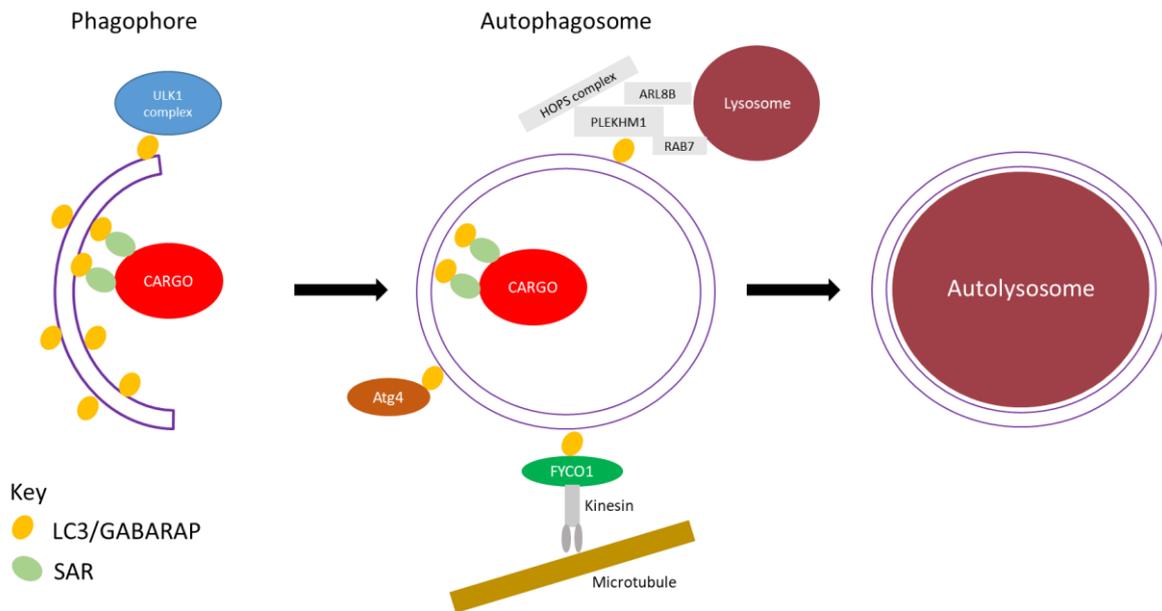


Figure 1.5:ATG8 interactions on the phagophore.

The ATG8 proteins show a structural similarity to ubiquitin (Figure 1.6), but with the hATG8 proteins having two N-terminal alpha helices before the ubiquitin like core [53]. In the human ATG8 proteins, the LC3 family have alpha helices that are basic and for the GABARAPs they are acidic [54]. The core of the hATG8 proteins contains a four-stranded beta-sheet with the two central stands running parallel (β -1 and β -4) and the two outer strands antiparallel (β 2 and β -3). Between the beta stands lie two alpha helices, with α -3 lying between β -2 and β -3 and α -4 lying between β -3 and β -4 [55] (Figure 1.6).

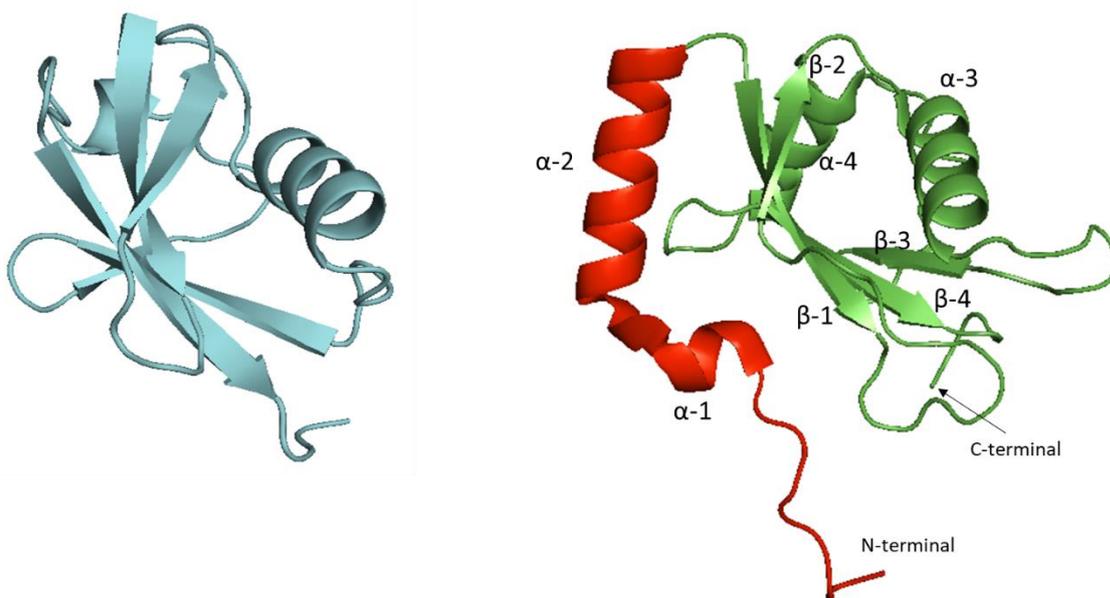


Figure 1.6: Crystal structure of ubiquitin (blue) and LC3B (green). The first two α -helices on LC3B are coloured red.

The six human ATG8 homologues have a high level of structural alignment (Figure 1.7). However, the amount of sequence similarity (Table 1.1) has large variations not only between the LC3s and GABRAPs, but also between proteins within the subfamilies. There is very low sequence similarity between the LC3s and GABARAPS, and whilst LC3A and LC3B and GABARAP and GABARAPL1 have very similar sequences (83% and 87% respectively), LC3C and GABARAPL2 have a much lower sequence similarity with other members of the subfamily (Table 1.1).

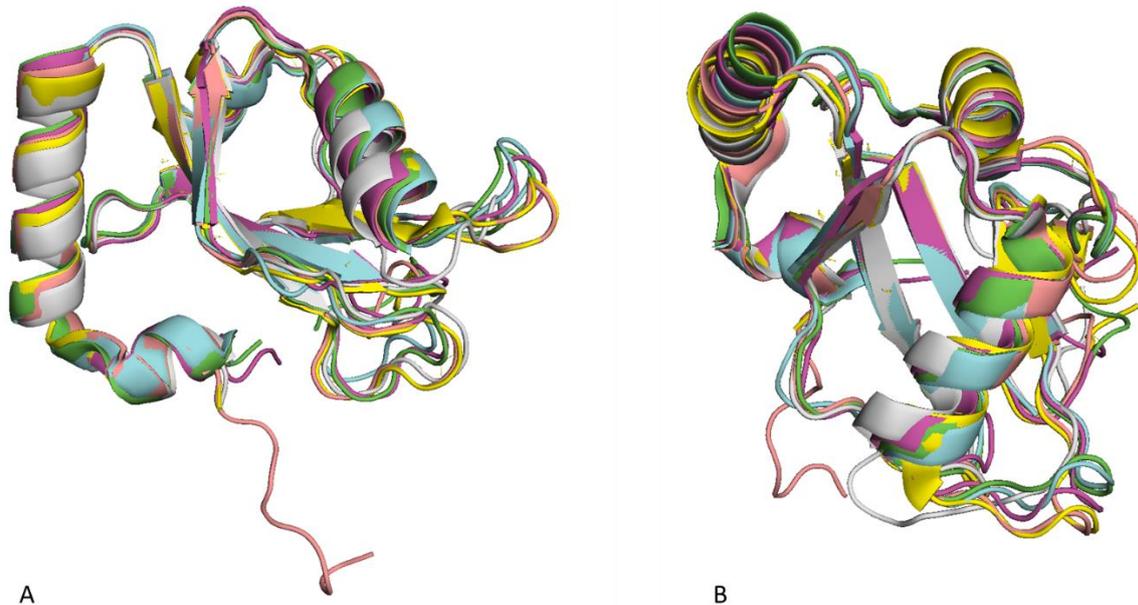


Figure 1.7: Human ATG8 homologues structure comparison shown from the front (A) and top (B). Proteins shown are LC3A (yellow), LC3B (pink), LC3C (grey), GABARAP (green), GABARAPL1 (blue) and GABARAPL2 (magenta).

| | LC3A | LC3B | LC3C | GABARAP | GABARAPL1 | GABARAPL2 |
|-----------|------|------|------|---------|-----------|-----------|
| LC3A | | 83 | 59 | 31 | 34 | 42 |
| LC3B | 83 | | 55 | 31 | 33 | 40 |
| LC3C | 59 | 55 | | 39 | 40 | 43 |
| GABARAP | 31 | 31 | 39 | | 87 | 58 |
| GABARAPL1 | 34 | 33 | 40 | 87 | | 61 |
| GABARAPL2 | 42 | 40 | 43 | 58 | 61 | |

Table 1.1: Sequence similarity between the human ATG8 proteins. Figure given as a percentage.

The existence of six human ATG8 homologues raises the question as to why there are so many when yeast only has one. Redundancy is one reason for the number of ATG8s however, having six proteins to act as redundancy for a single process suggests that the six hATG8 proteins have functional specialization.

A study showed that whilst both LC3 and GABARAP subfamilies are required for efficient autophagy, and both sub families are involved in the formation of the initial phagophore, that LC3s are involved in the elongation of the phagophore membrane and GABARAPs are essential for later stages of autophagosome maturation [56].

However, in another set of experiments where all six hATG8 family homologues were knocked out in human Hela cells it was found that this did not stop autophagy [57]. This did lead to slower initial rate of autophagosome formation and a large effect on the autophagosome maturation which combined to make the whole process inefficient. The same study showed the ATG8 proteins are essential for efficient autophagosome-lysosome fusion. Another KO study carried out with the six hATG8 homologues [58] underlined the importance that the GABARAPs in selective autophagy, as when the sub family was selectively removed the LC3s could not support all steps of selective autophagy. It has also been shown that bulk sequestration of cytoplasm is independent of LC3s, and requires GABARAPs [59].

1.3 Selective Autophagy and the LIR

Selective autophagy is mediated by selective autophagy receptors (SARs) that interact with the ATG8 family proteins on the inner membrane of the phagophore, through an LIR motif, thus ensuring encapsulation of the cargo within the autophagosome. The first human SAR discovered and studied was p62/SQSTM1 and its interaction with LC3B, facilitating the removal of protein aggregates [49, 60].

Investigations into the structure of the LC3B-p62 interaction [61, 62] showed that Trp338 and Leu341 of the LIR bind with two hydrophobic pockets on the surface of LC3B (Figure 1.8).

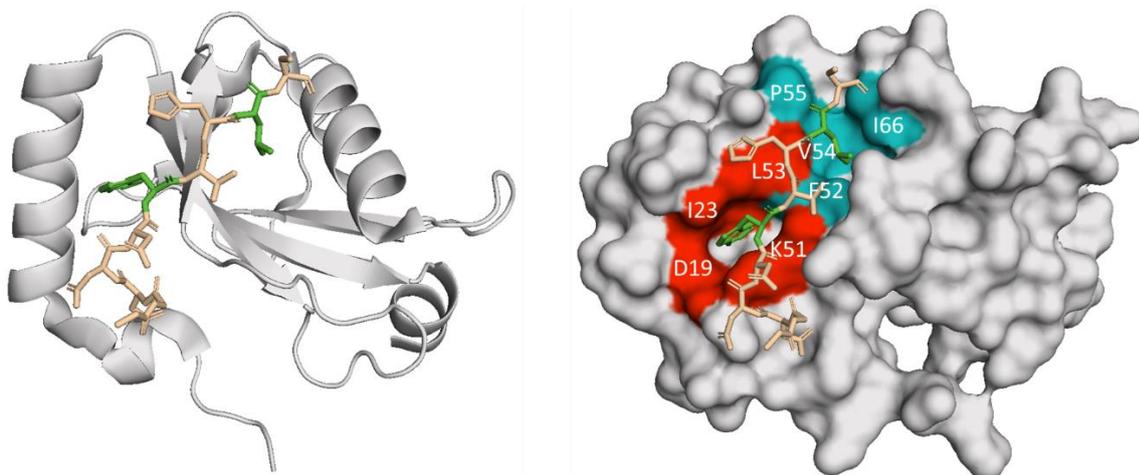


Figure 1.8: Structure of p62 LIR bound to LC3B. Hydrophobic pocket 1 (HP1) and 2 (HP2) are shown in red and blue respectively. Trp338 and Leu341 are shown in green with the rest of the LIR coloured yellow.

Trp338 binds deeply in hydrophobic pocket 1 (HP1), formed by α -helix 2 packing against the central β -sheet containing the amino acids D19, I23, K51 and L53. Leu341 of the LIR then sits in hydrophobic pocket 2 (HP2), formed by α -helix 3 packed against the β -sheet and containing the amino acids F52, V54, P55 and I66. As well as the interaction with the hydrophobic pockets, the interaction was also found to be stabilised by electrostatic interactions with three aspartic acid residues that preceded Trp338 with the face of LC3B. Although these are not essential for p62 to bind with LC3B (Figure 1.8).

At the same time the interaction of Atg8 in yeast with Atg19 was shown to be through an LIR that contained a tryptophan and Leucine, that bound with two hydrophobic pockets on Atg8 [62], and so the LIR motif was termed to be W-x-x-L (with x being any amino acid).

Since the discovery of p62 a large number of other selective autophagy receptors have been discovered transporting numerous different cargos to the autophagosome [63]. These cargos as well as protein aggregates (aggrephagy) include mitochondria (mitophagy), peroxisome (pexophagy), lysosomes (lysophagy), Secretory granules (zymophagy), ER (ERphagy), Ferritin (ferritinophagy), glycogen (glycophagy), bacteria (xenophagy), Viral capsids (virophagy), ribosomes (ribophagy), midbody rings (midbody autophagy) and nuclear lamina (nuclear lamina autophagy). A summary of a number of SARs, the cargo they transport and their LIRs is given in Table 1.2.

From the list of SARs given it can be seen that the vast majority interact with the ATG8 proteins through an aromatic and hydrophobic acid, with the notable exceptions of NDP52 and TAX1BP1. Whilst a number of the SARs have a single role, there are a number that interact with various different cargos.

As well as the SARs there are a number of proteins that are members of the core autophagy machinery that interact with the ATG8 family of proteins through an LIR, allowing the proteins to anchor on the convex side of the autophagosome membrane and regulate autophagosome formation, transport and fusion. These include members of the ULK1 complex, ULK1, ATG13 and FIP200 [47, 64, 65], component of the lipidation system ATG4B [66] and proteins involved with transport and fusion FYCO1 [51] and PLEKHM1 [52]. A summary of some of the core autophagy proteins and their LIR sequence is given in Table 1.3.

| Protein | Pathway | LIR Sequence | LIR Position |
|------------|---|---------------------------|--------------|
| p62 | Aggrephagy [49, 60] Mitophagy [67, 68] Pexophagy [69] Zymophagy [70] Xenophagy [30] Virophagy [71] Midbody autophagy [72, 73] | SGGDDD W THLSS | 332-343 |
| NBR1 | Aggrephagy [50] Pexophagy [74] Midbody autophagy [75] | ASSED Y IILPE | 727-738 |
| | | AQDLL S FELLD | 557-567 |
| Optineurin | Aggrephagy [76] Xenophagy [77] Mitophagy [67, 68, 78] | GSEDS F VEIRM | 173-183 |
| | | | |
| NDP52 | Xenophagy [79, 80] Mitophagy [81] | PENEED I LVVTT | 127-138 |
| | | | |
| TAX1BP1 | Mitophagy [82] | DEGNSD M LVVTT | 134-145 |
| BNIP3 | Mitophagy [83] | ESLQGS W VELHF | 12-23 |
| FUNDC1 | Mitophagy [84] | ESDDDS Y EVLDL | 12-23 |
| AMBRA1 | Mitophagy [85] | SGVEYY W DQL | 1012-1023 |
| Bcl2L13 | Mitophagy [86] | SLGPES W QQIAMDPE | 270-284 |
| FKBP8 | Mitophagy [87] | VPPLED F LVLDGVEDA | 18-33 |
| PHB2 | Mitophagy [88] | QELPS M YQRLGLD | 115-127 |
| NLRX1 | Mitophagy [89] | IRTEEE F QLLHIF | 457-469 |
| FAM134B | ERphagy [90] | DD F EELLD | 453-459 |
| Protein | Pathway | LIR Sequence | LIR Position |
| SEC62 | ERphagy [91] | SGNGND F EMITKEE | 357-370 |
| RTN3L | ERphagy [92] | DDR F TLLTA | 202-210 |
| | | PTE Y SKVEG | 214-222 |
| | | ESP F EVIID | 245-253 |
| | | ILT W DLVPQ | 339-347 |
| | | SKN F EELVS | 552-560 |
| | | QRS Y DILER | 787-795 |
| CCPG1 | ERphagy [93] | SDSSCG W TVISHE | 8-20 |
| ATL3 | ERphagy [94] | FTE Y GR LAMD | 189-198 |
| | | HCE F KQLALD | 357-396 |
| TEX264 | ERphagy [95] | GSS F EELDLK | 270-279 |
| Stbd1 | Glycophagy [96] | HEE W EMV | 200-206 |
| NUFIP1 | Ribophagy [97] | WM F WAMLPPP | 37-46 |
| NIX | Mitophagy [98] | AGLNSS W VELPM | 30-41 |

Table 1.2: Mammalian selective autophagy receptors and their LIR sequences. Amino acids critical for interaction with ATG8 family proteins are shown in red.

the ATG8 proteins (for example R10, R11, K49 and K50 on LC3B) [47, 65, 99, 100]. As well as acidic residues stabilising the interaction, it has been shown that phosphorylation of threonine or serine residues that precede the X₀ can also stabilise the interaction [101, 102]. LIRs usually bind as an extended beta-sheet to the binding site on the ATG8 proteins [45, 62].

Mutations within the LIRs have been shown to have large effects on the binding affinity of the LIR to ATG8 proteins, which can then play a part in neurodegenerative disorders like amyotrophic lateral sclerosis and frontotemporal lobar degeneration (ALS-FTLD) [103].

With so many SARs and core autophagy containing proteins possessing LIRs that bind to the ATG8 proteins, it raises the possibility that the reason for six ATG8 proteins is due to certain SARs having preferred binding partners within the hATG8 family, and the LIR sequence drives this specificity. Whilst most LIR containing proteins will bind to all members of the ATG8 family, there are a number of examples in the literature of LIR containing proteins binding selectively to the GABARAP family [47, 104-106] and currently a smaller number binding to the LC3 family [51, 87].

A recent study looked at a range of LIR containing proteins to derive a consensus GABARAP interaction motif (GIM) [107]. This study used a peptide-based assay to screen 30 LIR sequences against the ATG8 family proteins, to ascertain if a LIR showed a preference for the LC3 or GABARAP sub family. They found that 12 of the LIRs showed selectivity towards GABARAP over the LC3 sub family, and only one LIR (FUNDC1) showed a preference to LC3 over GABARAP. Analysis of these sequences lead them to propose a GIM of [W/F]-[V/I]-X₂-V. However, they were unable to identify a set of LIRs that showed a clear preference for the LC3 sub family. Further investigation of the binding affinity's of PLEKHM1-LIR peptide binding to the hATG8 proteins, showed K_d values that were around 10 times lower for the interaction with the GABARAP sub family (0.55 μM – 0.93 μM) than the LC3 subfamily (4.22 μM – 6.33 μM).

It has also been shown the C-terminus directly after the LIR motif can confer binding specificity to the GABARAP sub family [65]. The study showed that for proteins ULK1, ATG13 and PCM1, that X₅₋₁₀ of the LIR motifs formed hydrogen bonds and hydrophobic interactions with residues on the surface of GABARAP (L55/F62/L63). However, in LC3A/B the expanded hydrophobic surface is not present, and is partially charged (V58/K65/I68), and so the extended C-terminal interaction does not happen.

The most studied group of mammalian SARs are the soluble sequestosome-1-like receptors (SLRs). This group includes the SAR protein p62, NBR1, NDP52, TAX1BP1 and OPTN. All of the SLRs contain a dimerization or multimerization domain, an LIR binding domain and ubiquitin-binding domain. The cargo targeted by SLRs, as shown in Table 1.3, is diverse with just p62 being involved in degradation of protein aggregates, organelles and bacteria. The cargos are commonly ubiquitinated, which then allows the SLR to interact with it through the ubiquitin binding domain [24]. Through the LIR, the SLR interacts with the ATG8 proteins on the internal surface of the phagophore delivering the cargo for degradation. Polymerization of p62 through its PB1 domain is essential for selective autophagy and the formation of p62

bodies [49, 60, 108], which facilitates co-aggregation with the cargo and a tighter interaction between p62, the labelled cargo and the ATG8 protein present at the phagophore [109].

Neighbour of BRCA1 gene 1 (NBR1) is a member of the SLR group and has a similar domain structure to p62, containing an N-terminal PB1 domain followed by a ZZ type zinc finger and a C-terminal UBA (Figure 1.10).

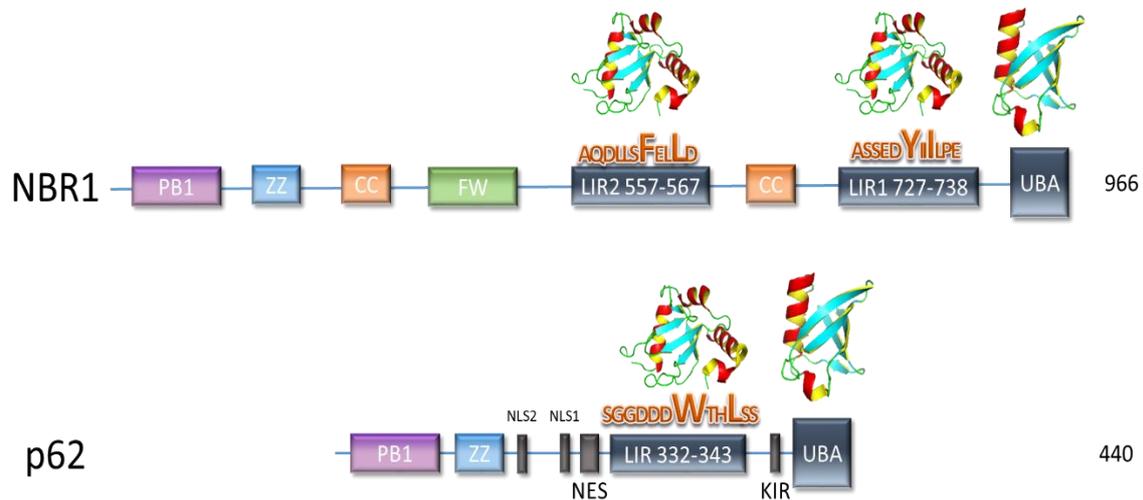


Figure 1.10: Domain architecture p62 and NBR1 members of the sequestosome-1-like receptor group of selective autophagy receptors. PB1, Phox and Bem1 domain; (pink), ZZ, ZZ-type zinc finger domain (light blue); CC, coiled-coil domain (orange); FW, four tryptophan domain (green); LIR, LC3 interacting region (dark blue); UBA, ubiquitin-associated domain (dark blue); NLS, Nuclear localisation signals (grey); NES, nuclear export signal (grey); KIR, Keap interacting region (grey). Images are not to scale and represent relative positions of domains. SAR size (number of amino acids) is given far right.

Both NBR1 and p62 target aggregated proteins and peroxisomes for degradation by selective autophagy, and have been shown to collaborate when targeting these cargos [50, 74]. NBR1 was found to promote pexophagy without the presence of p62, but the presence of p62 increased the efficiency of the process. Defects in aggrephagy leads to the accumulation of ubiquitin-positive protein inclusions which can lead to severe liver injury [110] and neurodegeneration [33]. An inability to maintain peroxisome numbers is linked to various neurodegenerative and developmental disorders such as X-linked adrenoleukodystrophy, and Krabbe disease [111].

NBR1 is roughly twice the size of p62 and is found at lower levels than p62 in the cell. NBR1 undergoes homo-oligomerization through a coiled-coil domain [50]. This is required for efficient peroxisomes degradation but it is unclear if required for aggrephagy. NBR1 contains a PB1 domain like p62, but unlike p62 is not used to polymerize but instead to interact with p62. NBR1 also contains a novel, evolutionary conserved four tryptophan (FW) domain (also known as the NBR1 box) [112]. NBR1 like p62 contains a UBA domain for binding to ubiquitin, but unlike p62 the isolated NBR1 UBA domain binds very well to both K48 and K63 linked di-ubiquitin [50]. Unlike p62 NBR1 contains two LIRs. The first LIR (LIR1)

has the sequence ASSEYIIILPE (amino acids 727-738 on NBR1) and the second LIR (LIR2) AQDLLSFELLD (amino acids 557-567 on NBR1) [50]. Both contain the core LIR Motif with an aromatic at the X₀ position, with LIR1 having a tyrosine and LIR2 containing a phenylalanine, and both have a hydrophobic residue in the X₃ position, with LIR1 and LIR2 having isoleucine and leucine at these positions respectively. LIR1 follows the majority of LIR sequence containing acidic residues at positions X₋₁ and X₋₂, whereas LIR2 contains no acidic residues until X₋₄. Neither sequence follows the motif for a GIM with both lacking the valine at X₃. The interaction of NBR1 LIR1 with the ATG8 family has been studied and a structure of the interaction with GABARAPL1 was produced[113] (Figure 1.11).

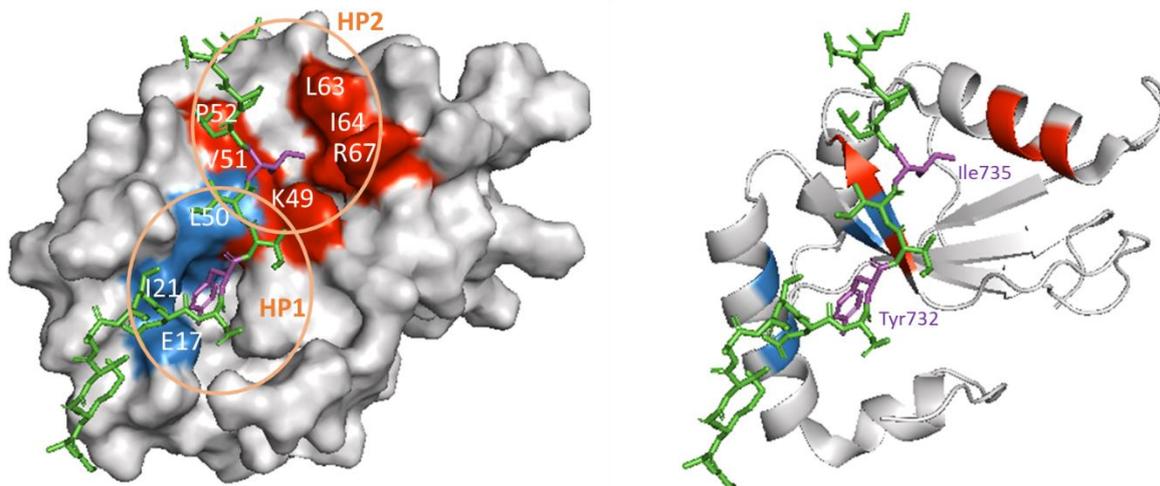


Figure 1.11: Structure of NBR1 LIR1 bound to GABARAPL1. LIR is shown in green with Y732 and I735 coloured purple. Amino acids on GABARAPL1 that were shown to be close in the NMR data to Y732 are coloured blue and those close to I735 are coloured red.

NBR1 LIR1 is shown binding to the canonical ATG8 binding site, with Tyr732 sitting in HP1 and Ile735 in HP2. Tyr732 on NBR1 binds deeply in the pocket, and makes close contact with E17, I21, P30 and L50 on GABARAPL1. Ile735 in HP2 then has close contact with Y49, L50, V51, P52, L63, I64 and R67. The acidic residues Glu730 and Asp731 then have electrostatic interactions with the positively charged amino acids on α -helix 2 (K16, K20 and K24).

As well as the structure of NBR1 LIR1 with GABARAPL1 the study also investigated the effect of mutating the LIR at various positions, and the effect this had on binding with LC3B. Using ITC they found that when Tyr732 was mutated to a tryptophan that the K_d went from 3 μ M to 0.4 μ M. However, mutation to a phenylalanine made very little difference. Adding extra acidic amino acids by mutating positions X₋₃ and X₋₄ from serine to Glutamic acid made very little difference to the K_d values.

Whilst the interaction of NBR1 LIR1 with the ATG8 proteins has been studied, the interaction with NBR1 LIR2 has not, and in fact in the literature the LIR was only located to a region of 100 amino acids [50]. Work within the Searle group then narrowed this down to the sequence given above and some initial interaction studies were carried out [114]. However, this initial analysis did not fully explore why NBR1 contains two LIRs, and the NMR titration data for the interaction of the two LIRs with LC3B, whilst indicating the LIRs bound

to the canonical binding site, the data was not analysed sufficiently to show the details of the interaction and the differences in interaction between the two LIRs.

NBR1 is not alone in proteins that contain two LIRs with ATL3 and RTN3L containing multiple LIRs. ATL3 contains two LIRs and it was found that both LIRs were required for the interaction, as if either of the aromatic residues in the LIR is mutated then the interaction with GABARAP was weakened. For RTN3L there are six LIRs and it was found that all six LIRs had to be deleted for the interaction with LC3B and GABARAPL1 to be interrupted. For NBR1 studies of the binding affinity of the two LIRs it was found that both LIRs bound to LC3B, GABRAP and GABARAPL2. However, in HEK293 cells it was found that LIR1 was by far the stronger binder with GABARAPL2 than LIR2 [50]. The reasons as to why NBR1 has two LIRs is not understood. The fact that the two LIRs are different in the amino acids at X₀ and X₃ and in the acidic residues present suggest that they have different preferred binding partners within the ATG8 family, and hence perform different roles within the SAR and autophagy.

1.4 Phosphorylation of LC3B

Posttranslational modification (PTM) of proteins is an important process in controlling the functionality of proteins, regulating activity, localization and interaction with other cellular materials. PTMs are used to regulate processes that covalently add functional groups to proteins, those that cleave proteins and those involved in degradation of entire proteins. These modifications include phosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, acetylation, lipidation and proteolysis. Proteins can be modified multiple times, and PTMs can be either reversible or irreversible [115]. PTMs within autophagy play an important role regulating ATG protein function with phosphorylation, ubiquitination, acetylation and lipidation playing important roles in regulating ATG proteins [116-118]. Within selective autophagy, lipidation is required to recruit the ATG8 proteins to the phagophore, ubiquitination labels aggregated proteins for degradation and phosphorylation of the SAR Optineurin, is required to stabilize the interaction with ATG8 proteins.

Phosphorylation, a PTM where protein kinases attach a phosphate group to the OH group on serine, threonine and tyrosine has been shown to be an important PTM when regulating the interactions of the ATG8 proteins. Phosphorylation of S12 on LC3A reduced its recruitment and participation in autophagy [119]. Phosphorylation of LC3B by protein kinase C significantly inhibited starvation induced processing of LC3B attenuating the addition of PE to the glycine exposed truncated form of LC3B. The phosphorylation sites were mapped to T6 and T29 on mouse LC3B [120]. Phosphorylation of the ATG8 proteins has been shown to have an important role in selective autophagy and the clearance of a specific cargo [121]. This study showed that the phosphorylation of LC3B at T50 by Sterile-20 kinases STK3 and STK4 was critical for fusion between autophagosomes and lysosome. As well as this important roll they also showed that mouse-embryonic fibroblasts (MEFs) deficient in STK3/STK4 were unable to efficiently clear *Streptococci* which would be a normal target of xenophagy. However, reconstitution of the deficient cells with a LC3B phosphomimetic mutant (T50E) allowed for the bacteria to be cleared from the cells. Further studies of the phosphorylation at T50 on LC3B has also been shown to inhibit binding to p62, NBR1 and FYCO1 [122].

All three phosphorylation sites found on human LC3A/B (on LC3A S6 is P6) are sited around the two hydrophobic pockets, with S/T12 (LC3A/LC3B) present on N-terminal α -helices that sit alongside the central β -sheet close to HP1, S/T29 (LC3A/LC3B) near HP2 and T50 located close to HP1 (Figure 1.12).

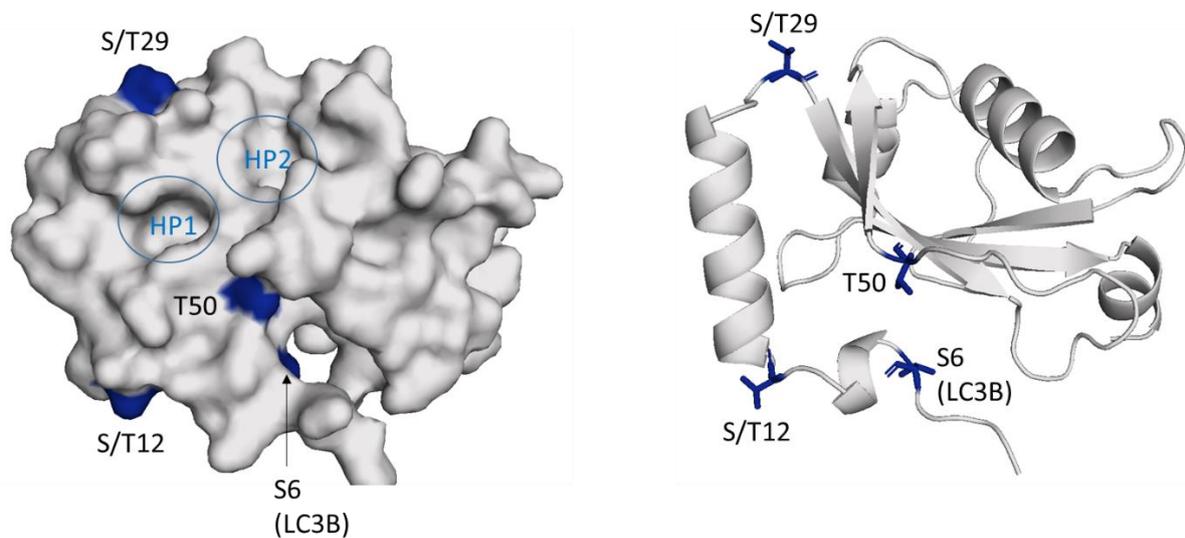


Figure 1.12: Phosphorylation sites on LC3A/B. Serine's and threonine's that can be phosphorylated are shown in blue. Where the amino acid present in LC3A is different from LC3B the LC3A sequence is given first (LC3A/LC3B).

The position of the phosphorylation sites leads to the possibility that phosphorylation of LC3A/B at these positions, could alter the interaction with the SAR LIRs with the hydrophobic pockets. This is of particular interest for S/T29 and T50 due to their close proximity to HP2 and HP1 respectively. S/T29 sits at the top of α -helix 2, and close to HP2 and where the LIR sits. Phosphorylation at this position could therefore change the interaction of the LIR with LC3A/B. T50 then sits close to HP1, and where the amino acids preceding the aromatic of the LIR interact with the face of LC3A/B. Hence a change in charge at this position through phosphorylation could potentially alter the interaction between LC3A/B and LIRs.

1.5 Project Goals

NBR1 plays an important role in clearing aggregated material and peroxisomes from the cell by localising them to the phagosome and subsequent degradation at the lysosome. Whilst p62, which has a similar domain structure to NBR1, has been well studied, NBR1 has not. The process of selective autophagy, by which the cargo is degraded, is reliant on the interaction of SARs like NBR1 with members of the ATG8 family of proteins through an LIR. NBR1 unlike other members of the SLR group of proteins has two LIRs. The interaction of LIR1 of NBR1 with GABARAP1 has been studied but LIR2 seems to have been all but forgotten in the literature, and the interaction of NBR1 LIR2 with the ATG8 proteins has not been investigated.

The two LIRs of NBR1 both bind to the ATG8 family but have differences at critical positions in the LIR, with the sequence varying between LIRs at the aromatic and hydrophobic amino

acid, and with LIR1 containing acidic amino acids preceding the aromatic position, but LIR2 does not. This raises the question of whether there are two LIRs for redundancy or if the LIRs have preferred binding partners amongst the six ATG8 proteins, and hence allowing NBR1 to fulfil a specific role.

This project will use a combination of mass spectrometry and NMR to study the interactions of the human ATG8 family of proteins with the two LIRs of NBR1. Initially the interaction of the two LIRs of NBR1 with LC3B will be studied. This work will present a complete backbone amide NMR assignment of LC3B and NMR titration data of the two LIRs with LC3B to look at the detail of the interaction between the two LIRs of NBR1 and LC3B. This will be done by analysing the chemical shift perturbations from NMR titrations to focus on the most significant changes. This selection of amino acids will be used in docking software HADDOCK webserver[123] to calculate and model the likely structure of the LIRs binding to LC3B. This data will then be compared to similar data generated for the interaction of LC3B with the LIR of p62.

The interaction of the two LIRs with LC3B was investigated further using a number of LC3B mutants to understand the difference in binding interactions. A K51A, L53A LC3B mutant (LC3B KL) was used as a binding site mutant which altered the surface around HP1. This change to HP1 and alteration in charge allowed the study of the difference in aromatic residue, and the preceding N-terminal amino acids in the two amino acids. A T29D LC3B phosphomimetic was then used to see what effect phosphorylation has on the interaction of the LIRs with LC3B, and if it affects the LIRs differently. MS competition binding experiments would initially be used to identify if the LIRs preferentially bind to one of the LC3B mutants (WT, KL and T29D). Any selectivity shown in the MS competition experiments would then be investigated further using NMR titrations of the KL and T29D mutants with the LIRs and HADDOCK webserver.

Finally MS competition experiments would be employed to probe if the two LIRs of NBR1 bind preferentially to any of the human ATG8 proteins. NMR titrations and HADDOCK webserver would then be used to further investigate any binding preference discovered.

Chapter 2

2. Materials and Methods

2.1 Chemicals

All chemicals were purchased from Sigma-Aldrich, Melford, Fisher, Enzyme Research Lab or Promega unless otherwise stated. A list of the composition of common buffers and stock solutions is given in Table 2.1

| Buffer | Contents |
|---|---|
| Tris-acetate-EDTA (TAE) 1x | 25 mM acetic acid, 1 mM EDTA, pH 4 (NaOH) |
| Lysis Buffer (1x) | 10 mM Tris, 150 mM NaCl, pH 7.5 |
| Cleavage Buffer (1x) | 20 mM Tris, 150 mM NaCl, 2.5 mM CaCl ₂ , pH 8.4 (HCl) |
| Gel Filtration Buffer | 30 mM Potassium phosphate, 100 mM NaCl, pH 7.0 |
| Desalt Buffer | 25 mM Ammonium acetate |
| Sodium dodecyl sulphate (SDS) Running Buffer (1x) | 25 mM Tris, 250 mM Glycine, and 0.1% SDS (w/v), pH 8.3 |
| 2x SDS Loading Dye | 100 mM Tris, 4% SDS (w/v), 20% Glycerol (v/v), 0.2% Bromophenol Blue (w/v), and 200 mM DTT, pH 6.8 |
| Coomassie Blue Stain | 0.25% Coomassie Brilliant Blue (w/v), 45% Methanol (v/v), and 10% Acetic acid (v/v) |
| NMR Buffer | 50 mM Potassium phosphate, 50 mM NaCl, 0.04% NaN ₃ , 5% D ₂ O (v/v), pH 7.0 |
| Ion Exchange buffer A | 5 mM Potassium phosphate, pH 7.0 |
| Ion Exchange buffer B | 5 mM Potassium phosphate, 2 M NaCl, pH 7.0 |
| 15% polyacrylamide solution | 375 mM Tris, 15% acrylamide/methylene bisacrylamide (37.5:1 ratio) solution (v/v), 0.1% SDS (w/v), pH 8.8 |
| Stacking buffer | 125 mM Tris, 4% acrylamide/methylene bisacrylamide (37.5:1 ratio) solution (v/v), 0.1% SDS (w/v), pH 6.8 |

Table 2.1: Buffer and stock solution composition.

2.2 Biological Materials

Glycerol stocks of *E. coli* cell lines: XL1 Blue and C41 (DE3) were provided by Dr. Jed Long (Searle group, School of Chemistry, University of Nottingham). Dr Jed Long also provided MAP1LC3B wild type and mutants in the pGEX-4T-3 plasmid (Figure 2.1). The LC3B wild type and mutant sequences were inserted between the EcoRI and Sall cloning sites. Finally Prof. Robert Layfield (University of Nottingham Medical School, QMC) provided MAP1LC3A, MAP1LC3B, MAP1LC3C, GABARAP, GABARAPL1, and GABARAPL2 in pETM-30 plasmid. Sequences for the ATG8 proteins are shown in Figure 2.2.

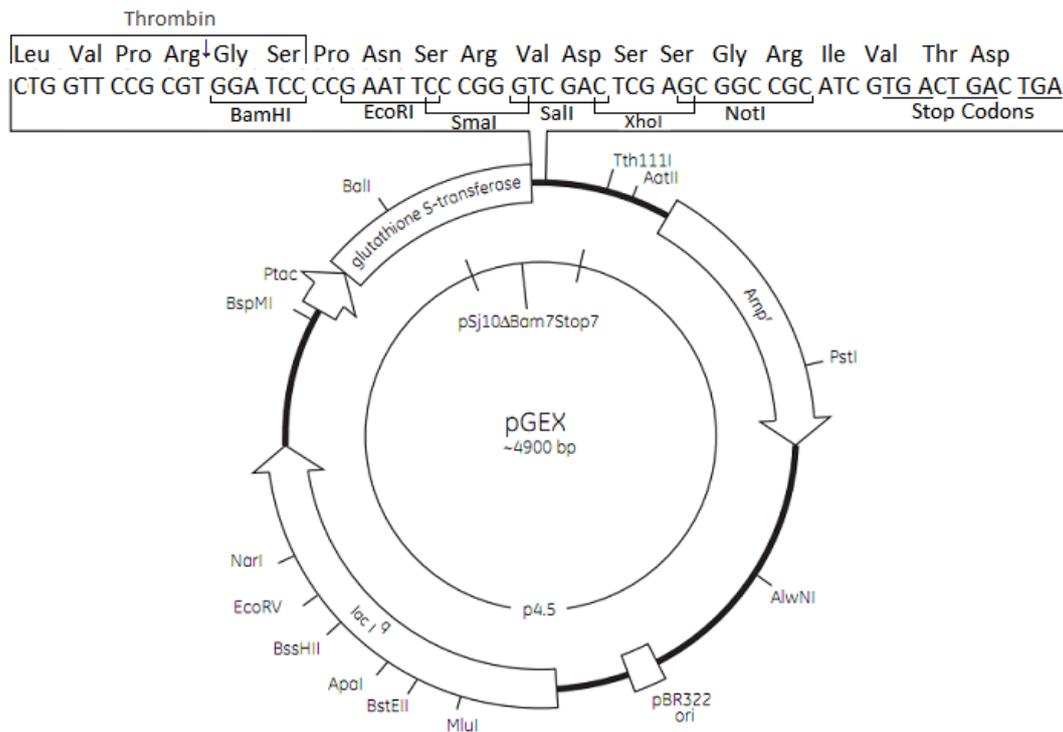


Figure 2.1: pGEX-4T-3 Plasmid Map (GE Healthcare Life Sciences)

LC3A:

```

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

```

LC3B:

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

LC3C:

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

GABARAP:

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

GABARAPL1:

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

GABARAPL2:

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

Figure 2.2: Sequences for the hATG8 proteins as supplied in the pETM-30 plasmid

Peptides for the LIRs of p62 and NBR1 were purchased from Peptide Synthetics. The peptide sequences used are listed in Table 2.2

| LIR | Peptide Sequence |
|-----------|------------------|
| p62 | SGGDDDWTHLS |
| NBR1 LIR1 | ASSEDYIIILPE |
| NBR1 LIR2 | AQDLLSFELLD |

Table 2.2: Peptide sequences for LIRs of p62 and NBR1 used in this work.

2.3 Molecular Biology Techniques

2.3.1 Sterilisation

All materials used that were not already provided sterile were sterilised by autoclaving (121°C, 15 min) using a Rodwell Ensign Autoclave. Materials that were heat sensitive were sterilised by filtration using a 0.22 µm filter.

2.3.2 Small-scale Overnight *E. coli* Cell Cultures

To carry out protein over-expression and plasmid DNA production small scale overnight cultures were used. 10-20 mL of sterilised Lysogeny broth (LB) media (25 g/L) was poured into 30 mL universals to which was added ampicillin (100 µg/mL) when using *E. coli* containing pGEX-4T-1/3 plasmids or tetracycline (50 µg/mL) when using *E. coli* containing pETM-30 plasmid. To this was added either a single colony from an agar plate or a scraping from a glycerol stock depending on the work being carried out. The culture was then incubated with shaking (37°C, 180rpm) overnight.

2.3.3 Agar Plates

Single colonies of bacteria were grown by putting 0.5 mL of overnight *E. coli* culture onto LB agar plates. LB agar was suspended in milliQ water (40 g/L), autoclaved (Chapter 2.3.1) and left to cool. Plates were produced as required by heating LB agar in a microwave until fully melted and the molten agar was allowed to cool to 42°C, before the appropriate antibiotic was added. The molten agar was poured into disposable Petri dishes and allowed to set. Plates that were not required straight away were then stored at 4°C.

2.3.4 Glycerol Stocks

Glycerol stocks were produced for storage of cells for later use. A single colony was picked from an agar plate and grown overnight (Chapter 2.3.2) and stored in glycerol (15%, v/v) at -80°C.

2.3.5 Site Directed Mutagenesis

Site directed mutagenesis was used to modify the DNA of the hATG8 proteins in the pETM-30 plasmid to introduce the correct restriction sites (*Bam*HI and *XHO*I) to allow the transfer into pGEX-4T-1 plasmid (Figure 2.3). Samples were made up of forward and reverse primer (2.5 µL, 10 µM), plasmid (1 µL, 100-200 ng/µL), dNTP (1 µL), Q5 reaction buffer (5X, 10 µl) and made up to 49.5 µL with nuclease free water. The polymerase chain reaction (PCR) conditions specified in Table 2.3 were then used.

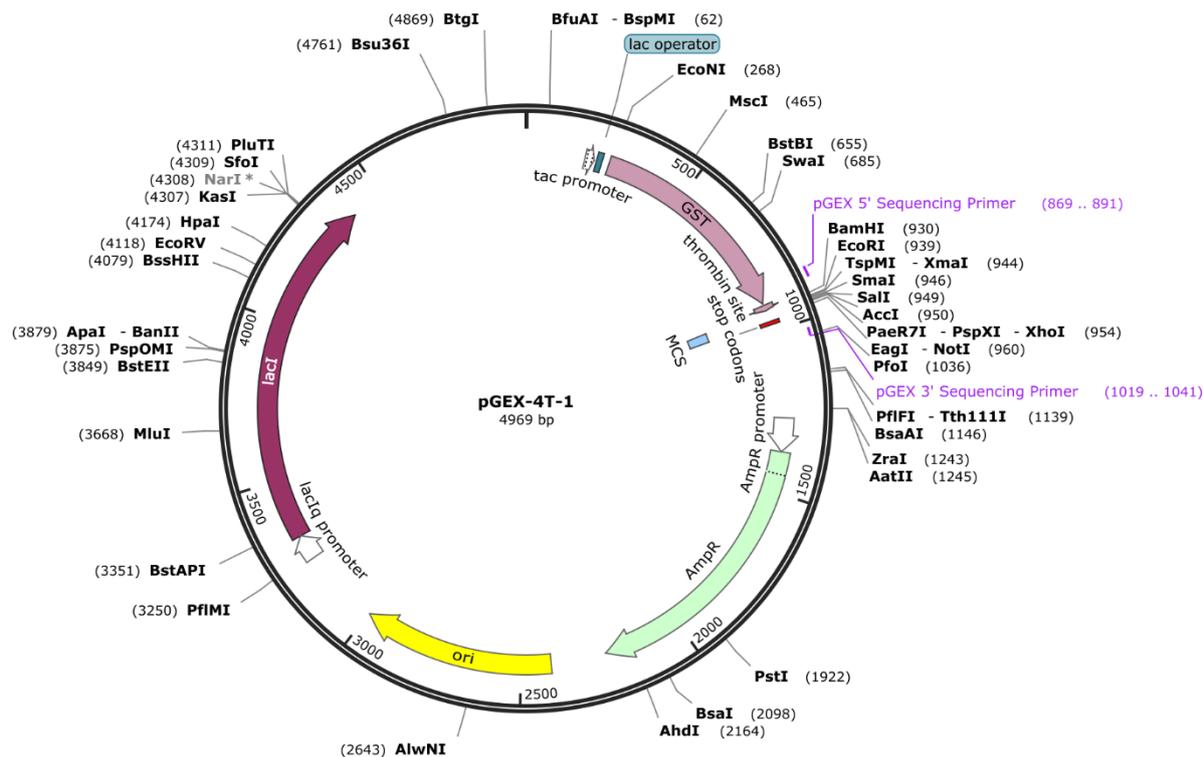


Figure 2.3: pGEX-4T-3 Plasmid Map (Amersham).

| Step | Temperature/ °C | Time/seconds |
|----------------------|-----------------|--------------|
| Initial Denaturation | 98 | 30 |
| 35 cycles | 98 | 10 |
| | 72 | 30 |
| | 72 | 30 |
| Final Extension | 72 | 120 |
| Hold | 4 | |

Table 2.3: PCR conditions.

2.3.6 DNA Digestion and Ligation

Enzyme digest was carried by mixing PCR product (43 μ L, 10-20 ng/ μ L), cutsmart buffer (5 μ L) and *XhoI* and *BamHI* (1 μ L each). These samples were then incubated at 37°C for one hour. The resulting cut DNA was extracted using gel extraction and the cut DNA (3 μ L) was added to pre-cut pGEX-4T-1 plasmid (1 μ L), ligase 10x buffer (2 μ L), nuclease water (13 μ L) and finally ligase (1 μ L). The solution was incubated for 20 minutes at room temperature.

2.3.7 Agarose Gels

Agarose gels were used to confirm the presence of DNA in PCR product and to purify material from enzyme digests. 1% gels were produced by dissolving agarose (0.50 g) in TAE (50 mL) buffer by heating in the microwave. The solution was then allowed to cool to roughly 40°C at which point ethidium bromide (5 μ L) was added. The resulting molten gel

was then carefully poured into a pre-assembled gel kit, the well comb was put in place and then the gel was left to cool and solidify.

Once the gel was set it was submerged in TAE buffer in a gel tank and PCR product in 6x DNA loading dye (NEB) was loaded. If the gel was being used to check the presence of DNA after PCR then 5 μ L of sample was used and if for the purpose of enzyme digest purification the whole sample was loaded. A 2-log ladder (NEB) was also loaded and the gel was run at a constant voltage of 70 V until sufficient separation had occurred. DNA bands were observed and photographed using a SynGene G:Box gel imager in transilluminator mode.

2.3.8 Transformation of Plasmid DNA into *E. coli* Cells

Cultures of the required *E. coli* strain were prepared as in Chapter 2.3.2 but instead of being grown overnight were grown (37°C, 180 rpm) until they reached an OD₅₉₅ of 0.3 AU. The cells were then harvested using a centrifuge (1000 *g*, 10 min, 4°C) and the supernatant discarded. The resulting pellet was then re-suspended in ice cold sterile filtered CaCl₂ (4 mL, 50 mM) and left on ice for at least an hour. The sample was then returned to the centrifuge (1000 *g*, 10 min, 4°C), the supernatant discarded and the resulting pellet re-suspended in ice cold sterile filtered CaCl₂ (800 μ L, 50 mM) to produce competent cells.

For each transformation 100 μ L aliquot of the competent cells was taken and 1 μ L of plasmid added. The sample was then left for 30 minutes on ice to allow the plasmid and *E. coli* to mix. The sample was then heat shocked (42°C, 40 secs) and placed back on ice for 5 minutes, followed by the addition of 400 μ L of LB and then placed in the incubator (37°C, 180 rpm, 1 hr). The sample was then plated onto ampicillin agar plates and incubated overnight at 37°C.

2.3.9 Plasmid Purification

Plasmid DNA was purified from a 10 mL overnight culture using either a QIAGEN QIAprep Spin Miniprep Kit or a Sigma GenElute Plasmid MiniPrep Kit, following the manufacturer's instructions.

2.3.10 DNA Sequencing

DNA sequencing was carried out by the sequencing facility, University of Nottingham Medical School, QMC. The sequencer used was a3130xl ABI PRISM Genetic Analyser and for the pGEX-4T-1 plasmid pGEX-reverse primer (CCG GGA GCT GCA TGT GTC AGA GG) was used.

2.3.11 Protein Over-expression

Protein over-expression of unlabelled protein

1 litre of LB in a 2 litre baffled flask was pre-warmed (37°C, 180 rpm, 30 min). To which was added ampicillin (1 mL, 100 µg/mL) and 2 overnight cultures (20 mL) of the desired protein (Chapter 2.3.2). The flask was then incubated with shaking (37°C, 180 rpm) until an OD₅₉₅ between 0.7 and 1 AU was achieved. To the flask was then added Isopropyl-β-D-thiogalactoside (IPTG) (0.5 mL, 1M) and ampicillin (0.5 mL, 100 µg/mL). The flask was then incubated overnight (30°C, 180 rpm) and the OD₅₉₅ checked to ensure it had grown. The cell suspension was then centrifuged (3000 g, 10 min, 4°C) and the resulting pellet collected and stored at -80°C.

Protein over-expression of ¹³C and ¹⁵N labelled protein

To 1 litre of sterile minimal media (Table 2.4) in a 2 litre baffled flask was added ampicillin (1 mL, 100 µg/mL), MEM vitamins (1 mL) and trace elements (1 mL). Depending on the type of labelling to be carried out varying amounts of ammonium chloride and glucose were dissolved in MilliQ water (20 mL) and added via sterile filtration. For ¹⁵N labelling then ¹⁵N labelled ammonium chloride (1 g) and glucose (4 g) were used. When producing ¹⁵N, ¹³C labelled protein then ¹⁵N labelled ammonium chloride (1 g) and ¹³C labelled glucose (2 g) were used. The minimal media was then pre-warmed to 37°C. 3 overnights of LB (15 mL) of the desired protein (Chapter 2.3.2) then had the cells harvested using a centrifuge (1000 g, 10 min, 4°C). The resulting pellets were then re-suspended in minimal media from the flask (15 mL) and the minimal media and the universals containing the re-suspended pellets place in an incubator (37°C, 180 rpm) for 30 minutes. The re-suspended pellets we then poured into the minimal media flask and then incubated with shaking (180 rpm, 37°C) until an OD₅₉₅ between 0.7 and 1 AU was achieved. To the flask was then added IPTG (0.5 mL, 1M) and ampicillin (0.5 mL, 100 µg/mL). The flask was then incubated overnight (180 rpm 30°C) and the OD₅₉₅ checked to ensure growth had occurred. The cell suspension was then centrifuged (3000 g, 10 minutes, 4°C) and the resulting pellet collected and stored at -80°C.

| Component | Amount/g |
|-------------------------------------|----------|
| Disodium hydrogen orthophosphate | 6 |
| Potassium dihydrogen orthophosphate | 3 |
| Sodium chloride | 0.5 |
| Magnesium sulphate | 0.3 |
| Calcium chloride dihydrate | 0.015 |

Table 2.4: contents of minimal media. Made up in 1 L of MilliQ water.

2.4 Protein Purification

2.4.1 Affinity Chromatography of GST-fusion Proteins

The protein pellet from 1 litre growth stored in the -80°C freezer was thawed in a room temperature water bath and then suspended and transferred to a sonication flask using 5mL of lysis buffer (Table 2.1) with the addition of DNase 1 (1 mg/ml) and protease inhibitor (1:1000). The suspension was then sonicated at 10 microns for 10 cycles of 30s of sonication followed by 30 s of rest. The resulting homogeneous suspension was then transferred to a centrifuge tube with 4ml of lysis buffer and centrifuged (30000 *g*, 30 min, 4°C) and the supernatant poured into a clean centrifuge tube and centrifuged again (35000 *g*, 30 min, 4°C) to give the clarified supernatant.

An empty P10 column was washed through with water and lysis buffer. A 1 mL aliquot of Glutathione Sepharose 4B beads was then added which formed a layer at the bottom of the column. This was then washed with water (10 mL) followed by lysis buffer (10 mL) to remove excess ethanol and equilibrate the beads. The clarified supernatant was then applied to the column and placed on a 360° rotator at 4°C for 1 hour to allow the GST tagged protein to bind to the beads. The column was then drained of liquid and washed with lysis buffer (2 x 10 mL) followed by cleavage buffer (10 mL). Thrombin (10 uL 5 units) was then made up to 1ml using cleavage buffer (Table 2.1) and added to the column. The column was then rotated (4°C, 14 hrs) to allow thrombin cleavage of the GST tag.

The protein was then eluted from the column. The column and the end caps were then washed carefully with 1.5 ml of cleavage buffer to make a total of 2.5 ml of sample and EDTA (40 uL, 0.5M) was added. The sample was then applied immediately to either a gel filtration column or cationic column depending on the purification process to be used.

2.4.2 Cationic Column FPLC

A 5 mL HiTrap™ SP HP cation exchange column (GE Healthcare) attached to an AKTA prime system (GE Healthcare) was used for protein purification. The column was pre-equilibrated with ion exchange buffer A (Table 2.1) and the protein sample made up to 50 mL with MilliQ water was loaded via a 50 mL Superloop™ (GE Healthcare). The protein was then eluted in 10 mL fractions using a flow rate of 2 mL/min with a 0-15 % gradient of ion exchange buffer B (Table 2.1) over 300 mL. By monitoring UV response protein containing fractions were isolated and lyophilised.

2.4.3 Hi Trap Desalt FPLC

If the protein to be desalted had initially been purified by gel filtration then the 1 mL of protein sample was loaded into a syringe with 1.5 ml of MilliQ used to wash the spin concentrator and added to the protein solution resulting in 2.5 mL of protein sample. If the protein had been purified using a cationic column then the lyophilised protein sample was dissolved in 2.5 mL of MilliQ water. The 2.5 mL of protein sample was loaded onto a HiTrap

Desalt column (GE Healthcare) on an AKTA prime system. The Desalt column consisted of 5x5ml columns and had been pre-equilibrated in ammonium acetate buffer (25 μ M) and washed sodium chloride (2 M). The sample was then run through at 0.5 ml/min with fraction sizes of 10 ml. The UV out was monitored to see when the protein was eluted and the conductivity checked to make sure the salt levels were low. The protein fractions were then frozen at -80°C and freeze dried. Once dry the sample was then dissolved in water, frozen again and then lyophilised. This water wash step was repeated twice to remove any traces of ammonium acetate.

2.4.4 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-Page)

SDS-Page allows the separation of a mixture of proteins into distinct bands through allowing denatured and reduced proteins to migrate through the polyacrylamide gel. As the charge of the proteins is dependent on size when an electric field is applied the proteins are separated.

The 15% Polyacrylamide resolving gels were poured using a casting system that was set up as per the manufacturer's instructions. Polyacrylamide gel solution (7 ml of stock 15%, Table 2.1) was mixed with freshly prepared ammonium persulfate (10 μ l, 100 mg/mL, APS) and tetramethylethylenediamine (10 μ l, TEMED). The gel was then immediately poured into the gel kit and levelled off with water. Once polymerisation had finished stacking gel was made by combining stacking buffer (5 mL, Table 2.1), APS (30 μ l) and TEMED (10 μ l). This was then poured on-top of the set gel and the lane comb inserted.

Gel samples were prepared in different ways depending on the sample by taking the required amount of sample and adding the required amount of 2x loading buffer (Table 2.1). Protein pellets were suspended in lysis buffer (300 μ l) and sonicated. Soluble and insoluble fractions were taken of the induced protein (10 μ l) and 2X loading buffer added (20 μ l). For GST-affinity samples flow through (10 μ l) was combined with 2X loading buffer (20 μ l). Bead samples were mixed with 2X loading buffer (50 μ l). For samples from purification (gel filtration and desalt) sample (10 μ l) was mixed with 2X loading buffer (20 μ l). All gel samples after the addition of loading buffer were heated at 95°C for 5 minutes.

Samples were loaded onto the gel along with Sigma Low Molecular weight markers 6.5 kDa to 66 kDa (for comparison purposes). The gel was run at 180 V in a cell containing 1x SDS (Table 2.1) until the dye had reached the end of the gel and then stained with coomassie blue. The gel was then rocked for 30 minutes, before being washed with distilled water and then covered with destain. The sample was then rocked with destain until it was judged that amount of dye remaining was enough to distinguish the bands from the background. A digital photo was then taken.

2.5 Biophysical Techniques

2.5.1 Mass Spectrometry

Mass spectrometry (MS) was used to determine the mass of protein samples, to check the formation of protein-peptide complexes and binding stoichiometry and to carry out binding competition experiments.

MS experiments were carried out on either a SYNAPT High Definition Mass Spectrometry System (Waters) using electrospray ionization (ESI) with a quadrupole time of flight (qTOF) mass analyser or a Bruker Impact II mass spectrometer (Bruker Daltonics) using ESI and a qTOF mass analyser. Values for settings used on the SYNAPT and Bruker Impact II are given in Table 2.5 and Table 2.6 respectively.

| Setting | Value |
|-------------------|---------------------------|
| Ion Mode | positive |
| Capillary voltage | 2.5 kV |
| Cone voltage | 40 V |
| Trap CE | 8 V |
| Transfer CE | 5 V |
| Backing pressure | ~3.8 mbar |
| Trap pressure | 2.1×10^{-2} mbar |
| Acquisition range | 500 to 4000 m/z |

Table 2.5: Values for settings on SYNAPT MS.

| Setting | Value |
|--------------|-----------------|
| Ion Mode | positive |
| Cone voltage | 2.5 kV |
| Mass range | 500 to 3000 m/z |
| Temperature | 150°C |

Table 2.6: Values for settings on Bruker Impact II MS.

2.5.1.1 Data Acquisition

Samples were made up of varying concentrations of lyophilised protein and peptide in ammonium acetate (25 mM) and were loaded into a syringe. These were then injected into the MS by direct infusion using a 100 μ L syringe on a syringe pump set at a rate of 180-250 μ L/min depending on the sample concentration.

2.5.1.2 Data Processing

Data processing was carried out using mestrenova (Mestrelab Research). For protein competition experiments the sum of the intensities of the peaks for the bound and unbound protein were calculated for each protein present. The ratio of the bound to unbound was then calculated which then allowed comparison of binding affinities between the proteins.

2.5.2 Nuclear Magnetic Resonance

Nuclear Magnetic resonance (NMR) experiments were carried out on a Bruker AV(III)800 (Bruker) spectrometer with a QCI cryoprobe. Samples were loaded into the spectrometer and allowed to equilibrate to 298 K before standard setup of tuning and matching, lock, shimming and calibration of 90° ¹H transmitter pulse. Standard Bruker pulse sequences with water suppression were used for all experiments.

2.5.2.1 Sample Preparation

Samples for standard analysis were prepared by dissolving the required amount of lyophilised protein in NMR buffer (600 µL, Table 2.1). The resulting sample was then centrifuged (13,000 *g*, 10 min) and carefully pipetted into an NMR tube to ensure a homogenous sample.

For NMR titrations two samples were prepared, the unbound and the bound. The required amount of protein was dissolved in NMR buffer (600 µL, Table 2.1) to form a stock. For the unbound sample half of the stock was diluted with NMR buffer (300 µL). For the bound sample, 2 equivalents of LIR peptide was dissolve in NMR buffer (300 µL; Table 2.1) and 300 µL of the stock solution then added. These samples would then constitute the start and end point of the titration. Once these had been run portions of the protein sample were then removed and replaced with the 2:1 stock solution to give different ratios of protein to peptide for points in the titration. An example of the points and concentrations of a standard titration with 250 µM of protein is given in Table 2.7.

| Point | Protein Concentration/µM | LIR peptide concentration/ µM | Equivalents of LIR peptide to protein |
|-------|--------------------------|-------------------------------|---------------------------------------|
| 0 | 250 | 0 | 0 |
| 1 | 250 | 25 | 0.1 |
| 2 | 250 | 50 | 0.2 |
| 3 | 250 | 75 | 0.3 |
| 4 | 250 | 100 | 0.4 |
| 5 | 250 | 125 | 0.5 |
| 6 | 250 | 150 | 0.6 |
| 7 | 250 | 175 | 0.7 |
| 8 | 250 | 200 | 0.8 |
| 9 | 250 | 225 | 0.9 |
| 10 | 250 | 250 | 1 |
| 11 | 250 | 500 | 2 |

Table 2.7: Points used for a standard titrations of ATG8 protein with LIR peptide.

2.5.2.2 Data Acquisition

Proton

^1H NMR experiments were carried out with 64 scans, using 32K points and a relaxation delay of one second. The transmitter frequency was set to be the frequency of water in the sample and the experiment was run with a spectral width of 15 ppm. Water suppression was carried out using excitation sculpting.

The pulse program used was zgesgp, which used a 2 ms 180 degree shaped pulse (Squa100.1000) and gradients (SMSQ10.100) set to 31% and 11% for GPZ1 and GPZ2 respectively. Power levels were then set based upon the 90 degree pulse calculated in the initial setup experiment.

Heteronuclear Single Quantum Coherence (HSQC)

^1H - ^{15}N HSQCs were acquired with varying numbers of scans depending on the concentration of the sample. The transmitter frequency was set to the water peak and 2k points with a spectral width of 16 ppm were used in F2. The number of points in F1 was varied from 64 to 256 points depending on the quality of data required and the sweep width and transmitter frequency were set according to the protein being run. Water suppression was carried out using a watergate sequence.

HSQC data was obtained using the standard Bruker pulse program hsqccfpf3gpplhwg. A proton-nitrogen coupling constant (CNST4) of 90 Hz was used and a relaxation delay of 1 s was set. The shape pulse used was a 1 ms squa100.1000 and the gradients were set to 50%, 80% and 30% for GPZ3, GPZ4 and GPZ5 respectively. For ^{15}N decoupling garp4 was used with PCPD3 set to 269.43 μs . All powers were then calculated from the 90 degree pulse obtained earlier.

Triple Resonance Backbone Assignment

To allow the assignment of the backbone amides in the ^1H - ^{15}N HSQC a series of 3D experiments were used. These were carried out on ^{13}C and ^{15}N isotopically enriched protein samples. The experiments carried out were HNCO, HN(CA)CO, HNCA, HNCACB, HN(CO)CA and HN(CO)CACB.

The number of scans used was based upon the concentration of the sample and the spectral widths and transmitter frequencies were set in F1 and F2 to those used to obtain a ^1H - ^{15}N HSQC for the sample. The number of points in F1 and F2 were set 2k and 32K respectively. Values for the settings for the various 3D experiments are given in Table 2.8.

HNCO and HN(CA)CO

| Experiment | HNCO | HN(CA)CO |
|---|--|--|
| Pulse program | hncogpwg3d | hcaconhgp3d |
| ¹³ C Transmitter frequency | 176 ppm | 176 ppm |
| ¹³ C spectral width | 22 ppm | 22 ppm |
| ¹³ C points used | 64 | 64 |
| ¹⁵ N points used | 40 | 40 |
| ¹ H 90 degree shaped pulse | 1 ms, Squa100.1000 | 1 ms, Squa100.1000 |
| ¹ H decoupling | 60 μs, waltz65 | 60 μs, waltz65 |
| CO offset | NA | 176 ppm |
| C-alpha offset | 54 ppm | 54 ppm |
| ¹³ C 90 degree shaped pulse | 240 μs SP2 = Q5.1000 SP8 = Q5tr.100 | 240 μs SP2 = Q5.1000 SP8 = Q5tr.100 |
| ¹³ C 180 degree shaped pulse | 192 μs, Q3.1000 | 192 μs, Q3.1000 |
| ¹⁵ N decoupling | 269.43 μs, garp4 | 269.43 μs, garp4 |
| Gradients | SMSQ10.100 GPZ1 = 50% GPZ2 = 40% GPZ3 = 60% GPZ4 = 30% | SMSQ10.100 GPZ1 = 50% GPZ2 = 40% GPZ3 = 60% GPZ4 = 30% |

HNCA and HN(CO)CA

| Experiment | HNCA | HN(CO)CA |
|---|--|--|
| Pulse program | hncagpwg3d | hncocagpwg3d |
| ¹³ C Transmitter frequency | 54 ppm | 54 ppm |
| ¹³ C spectral width | 32 ppm | 32 ppm |
| ¹³ C points used | 64 | 64 |
| ¹⁵ N points used | 40 | 40 |
| ¹ H 90 degree shaped pulse | 1 ms, Squa100.1000 | 1 ms, Squa100.1000 |
| ¹ H decoupling | 60 μs, waltz65 | 60 μs, waltz65 |
| CO offset | 176 ppm | 176 ppm |
| C-alpha offset | NA | 54 ppm |
| ¹³ C 90 degree shaped pulse | 240 μs SP2 = Q5.1000 SP8 = Q5tr.100 | 240 μs SP2 = Q5.1000 SP8 = Q5tr.100 |
| ¹³ C 180 degree shaped pulse | 192 μs, Q3.1000 | 192 μs, Q3.1000 |
| ¹⁵ N decoupling | 269.43 μs, garp4 | 269.43 μs, garp4 |
| Gradients | SMSQ10.100 GPZ1 = 50% GPZ2 = 40% GPZ3 = 60% GPZ4 = 30% | SMSQ10.100 GPZ1 = 50% GPZ2 = 40% GPZ3 = 60% GPZ4 = 30% |

HNCACB and HN(CO)CACB

| Experiment | HNCACB | HN(CO)CACB |
|---|--|--|
| Pulse program | hncacbgpwg3d | hncocacbgpwg3d |
| ¹³ C Transmitter frequency | 39 ppm | 39 ppm |
| ¹³ C spectral width | 75 ppm | 75 ppm |
| ¹³ C points used | 64 | 64 |
| ¹⁵ N points used | 40 | 40 |
| ¹ H 90 degree shaped pulse | 1 ms, Squa100.1000 | 1 ms, Squa100.1000 |
| ¹ H decoupling | 60 μs, waltz65 | 60 μs, waltz65 |
| CO offset | 176 | 176 ppm |
| C-alpha offset | NA | 54 ppm |
| C-aliphatic offset | NA | 39 ppm |
| ¹³ C 90 degree shaped pulse | 240 μs SP2 = Q5.1000 SP8 = Q5tr.100 | 240 μs SP2 = Q5.1000 SP8 = Q5tr.100 |
| ¹³ C 180 degree shaped pulse | 192 μs, Q3.1000 | 192 μs, Q3.1000 |
| ¹⁵ N decoupling | 269.43 μs, garp4 | 269.43 μs, garp4 |
| Gradients | SMSQ10.100 GPZ1 = 50% GPZ2 = 40% GPZ3 = 60% GPZ4 = 30% | SMSQ10.100 GPZ1 = 50% GPZ2 = 40% GPZ3 = 60% GPZ4 = 30% |

Table 2.8: Values for settings for triple resonance NMR experiments.

2.5.2.3 Data Processing

Raw data was processed using topspin 3.5 or topspin 4. All data was zero filled by a factor of 2. For 1D proton data an exponential function was used with a 2 Hz line broadening and for 2D and 3D data a shifted sine bell (SSB) window function was applied. Forward linear prediction was used in the ¹⁵N and ¹³C dimensions for 2D and 3D data sets.

2.5.2.4 Assigning backbone amides in the ¹H-¹⁵N HSQC spectrum using CcpNmr Analysis

The backbone amides in the ¹H-¹⁵N HSQC were assigned by first linking the amide NH peaks in the HSQC to carbon peaks for the amino acid in the 3D data [124]. Then by linking the carbonyl, C_α and C_β peaks in the forward (HN(CA)CO, HNCA and HNCACB) and reverse (HNCO, HN(CO)CA and HN(CO)CACB) it is possible to work out which sets of peaks belong to amino acids that are next to each other within the protein [125-129].

The linking of the amino acids required all 3 sets of complimentary 3D NMR experiments to maximise the chance of being able to find the matching carbon peaks in the forward and backward spectra. How the magnetisation is transferred along the amino acid backbone of the protein is shown in Figure 2.4.

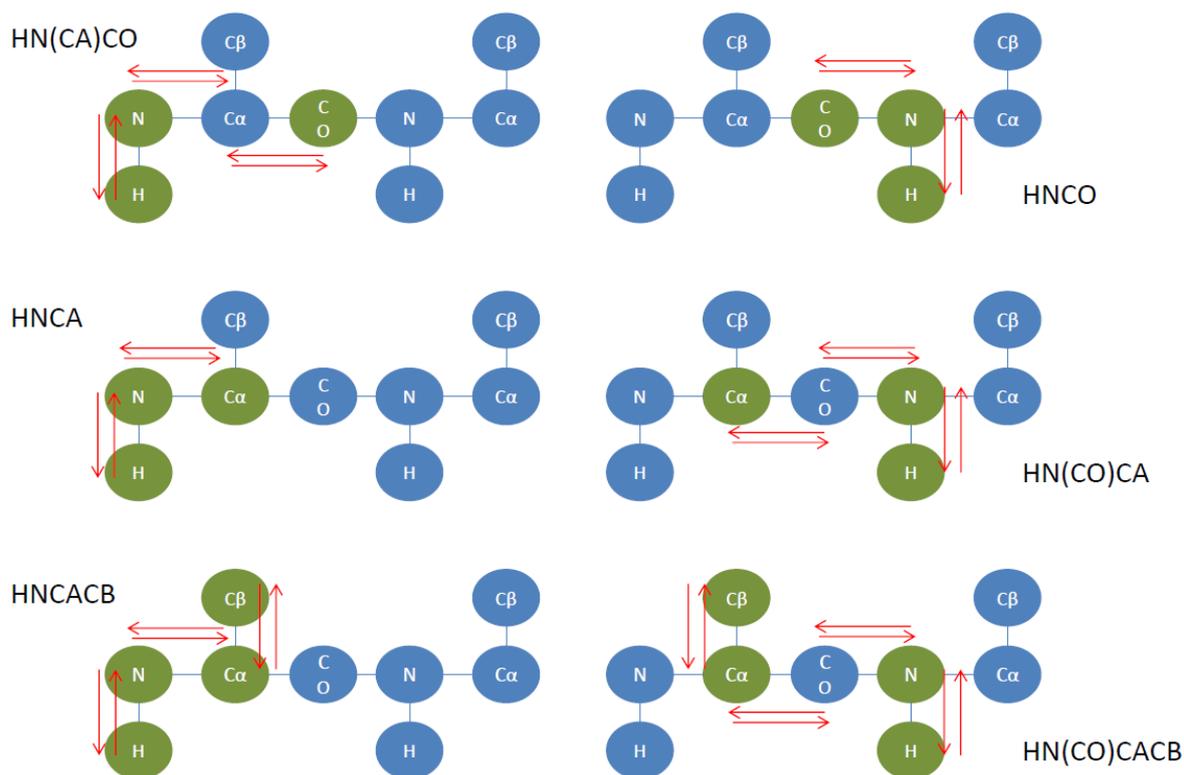


Figure 2.4: Magnetisation transfer for triple resonance backbone assignment NMR experiments. Nuclei coloured in green give signals in the 3D NMR spectrum.

However, the linking of peaks can be challenging when peaks are missing due to low signal to noise or protein dynamics. Whilst the CACB experiments gave the most useful information, the sensitivity was much lower than that of the carbonyl and CA experiments, and CB peaks can be missing. This could sometimes be improved by running longer experiments, or CACB experiments that were optimised to show the CB peak.

To enable the linking of peaks CcpNmr analysis [130] was used. Once the spectra had been imported CcpNMR facilitates the linking of peaks to spin systems, by first picking the peaks in the ^1H - ^{15}N HSQC spectrum and assigning them to a spin system. By then selecting each spin system it is possible to navigate to the same proton and nitrogen frequencies in the 3D spectra and link peaks for carbonyls, carbon alpha and carbon beta to the spin system. Each spin system then has the carbonyl, carbon alpha and carbon beta for the current amino acid (i) and the preceding amino acid ($i-1$) associated with it. Using the software it is then possible to link spin systems. On selecting a spin system the i peaks for the amino acid are displayed and then the software presents a series of matches for the carbon frequencies which are the $i-1$ for other spin systems (Figure 2.4). By then looking at this selection it is possible to select the peaks that best match and then link spin systems.

However, this is not always as simple as it appears as sometimes the software does not pick up matches and so looking through spectrum at a certain frequency is required. Also not all peaks appear in the spectrum and so it is not always possible to link all spins systems and so gaps appear in the sequence.

Once links had been made between peaks in the NMR data it was then possible to identify the type of amino acid the spin system was. All amino acids have a range of chemical shifts over which the CA and CB is expected to appear and by using certain amino acids that have very indicative chemical shifts it is then possible to get a foothold on the sequence. Alanine, glycine and serine and threonine all have chemical shifts that are very distinctive and so relatively easy to spot in the CACB 3D spectrum and hence know the type of amino acids the spin system corresponds to. Alanine contains a methyl CB and so the peaks appears at a much lower ppm value than the rest of the amino acids (most of the time). Glycine only has one peak which as a CA and a CH₂ has a different phase to the other CA (the CACB experiments are multiplicity edited) and appears at a frequency in-between the CA and CB. Finally serine and threonine contain a CB that is downfield from the CA. Some examples of these ppm shifts are for these 3 amino acids are given in Figure 2.5.

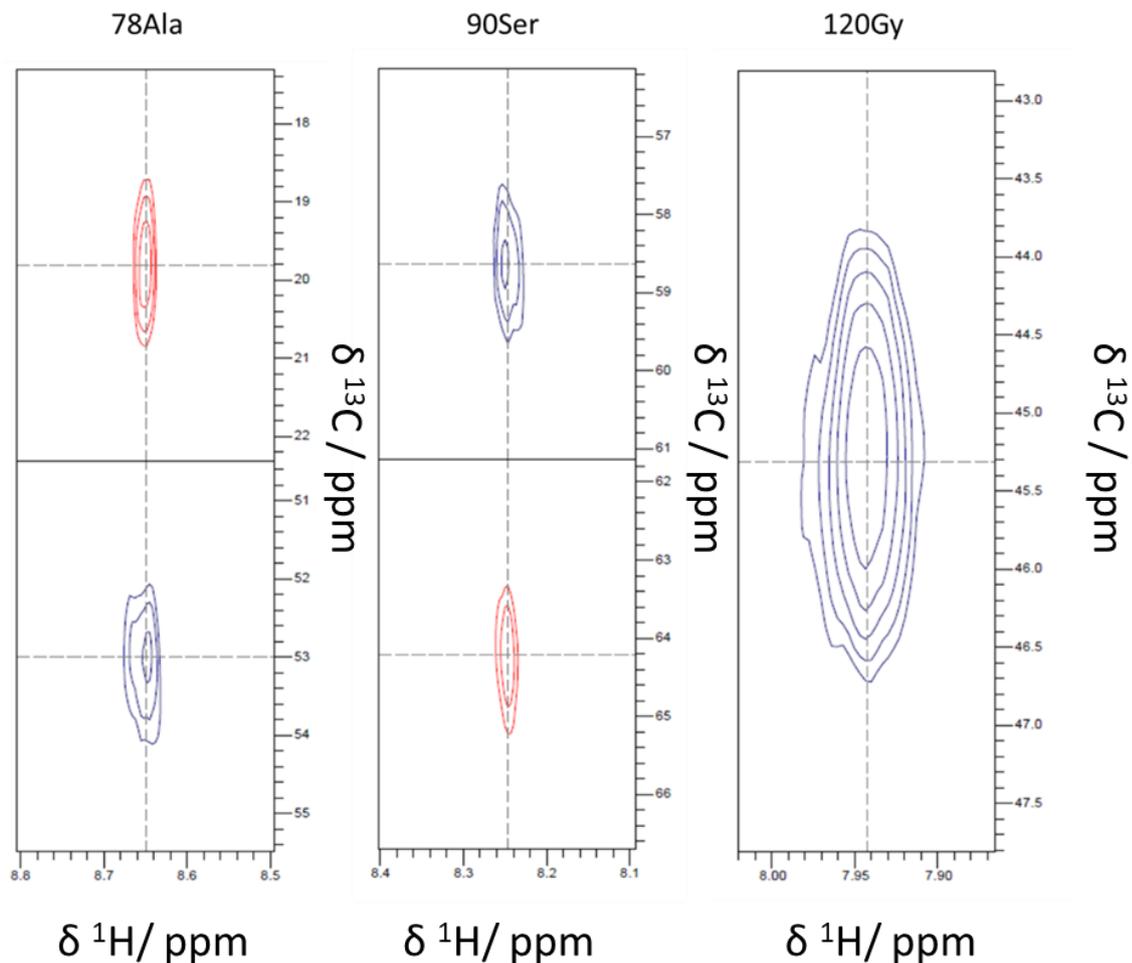


Figure 2.5: Example ^{13}C CA and CB chemical shifts of alanine, glycine, serine and threonine amino acids. Data Shown is from a HNCACB experiment.

With these amino acids identified in the linked peaks it is then possible to compare these partially identified amino acid sequences to the sequence of LC3B and see where they match. By then also looking at the chemical shifts for the unidentified amino acids and eliminating types of amino acid they can't be it is then possible to assign the peaks to the sequence of LC3B.

2.5.2.5 Analysis of titration data

To analyse the data produced by titrating the LC3 proteins with the various LIRs (Chapter 2.5.2.1) the individual ^1H - ^{15}N HSQC spectrum for the points in the titration were imported into CcpNmr analysis. The amide backbone assignment for the protein was then copied to the zero point. Using the copy resonance facility in the software it was then possible to copy the assignment to the end point by following the peak in the ^1H - ^{15}N HSQC spectrum as the peptide concentration was increased. However, this was not always as easy as following a peak as the ligand was titrated with the protein depending on the rate of exchange.

The exchange regime depends on the rate of the ligand binding and dissociating from the protein and for a ligand binding reversibly to a single site the dissociation constant is equal to

$$K_d = \frac{[P][L]}{[PL]}$$

Equation 2.1

Where [P] and [L] are the concentration of free protein and ligand respectively and [PL] is the concentration of protein-ligand complex. The dissociation constant k_d can then be thought of as the concentration of ligand to saturate half the binding sites. The forward and back rates are then $[P][L]k_{on}$ and $[PL]k_{off}$ respectively which at equilibrium when the rates are equal gives the following equation for the dissociation constant

$$K_d = \frac{k_{off}}{k_{on}}$$

Equation 2.2

So depending on the rate the protein is dissociating leads to different exchange regimes in the ^1H - ^{15}N HSQC as the ligand is titrated and hence how the peaks behave as the ligand concentration increases.

The 3 exchange regimes the peaks in the ^1H - ^{15}N HSQC can undergo are fast, intermediate and slow. If k_{off} is much greater than the difference in Hz between the chemical shifts of free and bound protein then the signals will move smoothly from their position in the free spectrum to the bound as the titration is carried out. This is fast exchange.

When the exchange rates are similar to the chemical shift difference, the signals broaden and shift at the same time. This is intermediate exchange, which leads to peaks undergoing a mixture of the characteristics of slow and fast exchange. The peak may move smoothly for the first few point in the titration until it broadens out and then moves to the position of the bound protein. This broadening of the peak often means that the peak is lost in the noise, and hence following the peak in the titration can be very difficult.

Finally when k_{off} is slower than the difference in chemical shifts the unbound signal gradually decreases in intensity, whilst the bound peak appears and increases in intensity (the

intensity of the peak is relative to the amount of the species present). This is slow exchange. This leads to seeing two peaks in the titration HSQC spectrum, one for the free protein and another for the bound. As the titration progresses the intensity of the free protein peak decreases as the bound peak increases until the end point is reached and only the bound peak remains. Peaks undergoing slow exchange can be extremely difficult to assign as if there is a large change in the chemical shift for the peak of the bound protein then following the peak in the titration is not possible.

Some examples of peaks undergoing the different exchange regimes are shown in Figure 2.6.

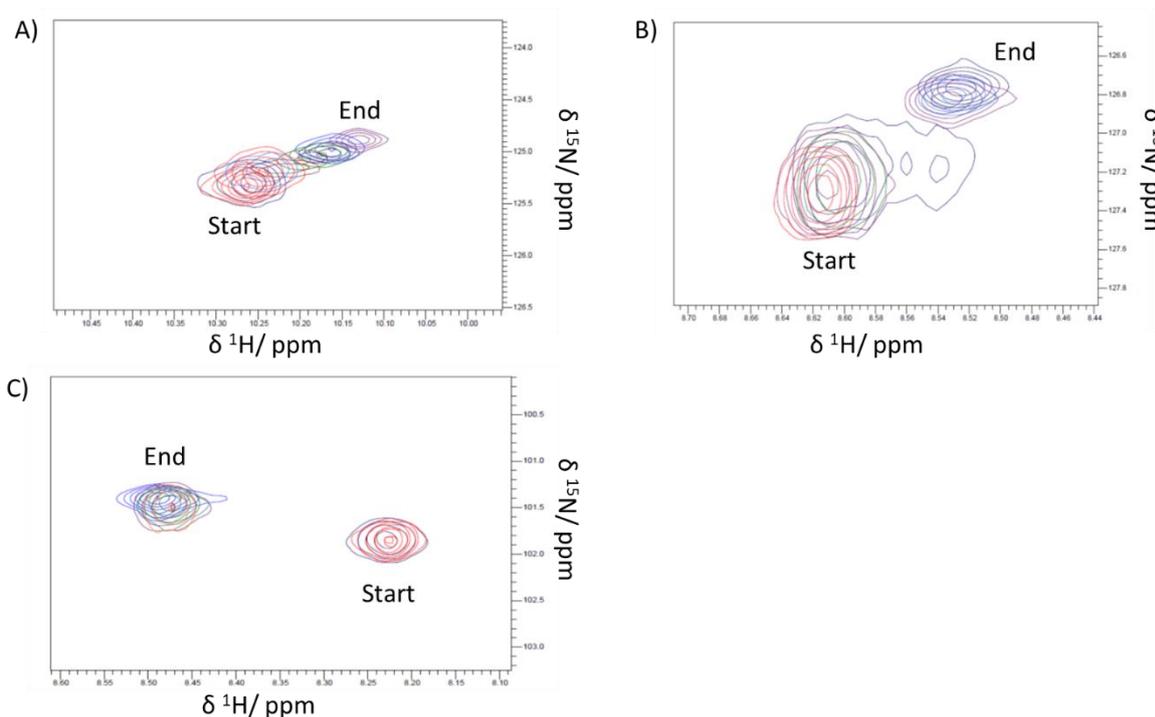


Figure 2.6: Examples of exchange regimes of peaks in a ^1H - ^{15}N HSQC whilst carrying out a titration of LC3B with NBR1 LIRs. Peaks are shown exhibiting fast (A), intermediate (B) and slow (C) exchange. Start and end points of titration are indicated.

2.5.2.6 Determination of Chemical Shift Perturbations

With the start and end points HSQCs assigned the next step was to calculate the chemical shift perturbation (CSP) for each peak so as to give a value for the magnitude of the change each amide undergoes on binding [131].

$$CSP = \sqrt{\frac{1}{2}[\delta_H^2 + (1.4 \cdot \delta_N^2)]}$$

Equation 2.3

CSPs were calculated using Equation 2.3 where δ_H and δ_N were the difference in chemical shift between the free and bound peaks for the proton and nitrogen dimension respectively.

With the CSPs calculated it was then important to analyse the data carefully so as to determine which changes were significant and hence important for defining the binding site. It is important to use a defined way to calculate a cut off for significance so as to not unduly influence the location of the binding site to a desired result by setting arbitrary value.

$$Cutoff = \mu + x \cdot \sigma$$

Equation 2.4

The cut off value for significance was calculated using Equation 2.4 where μ is the mean of the CSPs for all the peaks in the HSQC and σ the standard deviation. The value x was then set depending on the level of significance to be achieved. For the purpose of visualising the binding site on the protein surface x was set to 0. For then focusing in on residues that played the most significant part in the binding values of 1 and 2 were used.

2.5.2.7 Generating hATG8 – LIR structure with HADDOCK

Structures of the LIR peptides binding to the hATG8 proteins were generated with HADDOCK webserver [123]. HADDOCK uses standard energetics and shape complementarity for structure calculation of molecular complexes. It can also incorporate experimental data as restraints to guide the docking of ligands onto proteins. HADDOCK uses ambiguous interaction restraints (AIRs) to translate raw data such as NMR CSPs into distance restraints. These AIRs are separated into two categories: active and passive. Those that are active are considered to be directly involved in the interaction and passive residues are considered to be surrounding surface residues. When carrying out the simulation the active residues are restrained to be part of the docking interface or else incur a scoring penalty. Passive residues are of less importance and if missing from the binding interface there is no penalty.

When using HADDOCK for this project the first step was to provide structures for the hATG8 and LIR in PDB format. For the hATG8 protein complex the active residues were defined to be those amino acids on the front face of the hATG8 that had the largest peak shifts in the ^1H - ^{15}N HSQC (larger than the mean and two standard deviations) or those that could not be assigned in the titration end point. The passive residues were then set as the other significant shifts (larger than the mean and one standard deviation). For the LIR peptide the active residues were set to be the aromatic and hydrophobic amino acids in the LIR and no passive residues were selected.

Once the HADDOCK run has started the first step will then generate 1000 models through rigid body energy minimisation. At this point the AIRs are included in the energy function and have a large effect on the complexes generated. These are then refined to the best 200 based on energetics when flexibility is introduced to the interface followed by water.

These remaining structures are then grouped into clusters based on similarity and assigned a score which is the average score of the top 4 structures in each cluster. This score reflects the van der Waals energy, intermolecular electrostatic energy, desolvation energy and the AIR energy. Using these scores it is then possible to select the top scoring cluster and interpret the structure generated.

HADDOCK is an extremely useful tool for generating possible structures for the protein interactions. However, care needs to be taken when viewing the results to remember, that whilst the structures generated are potentially how the proteins dock that this might not be the case. Due to HADDOCK being very dependent on the AIRs used, it is possible to heavily influence the initial set of structures generated. If a very strict set of AIRs is used there is then very narrow sampling on the subsequent energy refinement process, which will lead to very similar structures. Whereas a sparser set of restraints can give very different solutions.

For the purpose of this project HADDCOK will be used as a tool to visualise how the LIRs are potentially docking onto the hATG8s which will help interpret the CSPs observed in the NMR titrations.

Chapter 3

3. Producing the ATG8 Proteins and the Mutants of LC3B

To enable the study of the interaction of human ATG8 (hATG8) proteins with various LIRs, a number of different proteins were produced both as unlabelled and as ^{13}C and ^{15}N isotopically enriched proteins. The production of the various proteins was split into two separate pieces of work, the first being the production of LC3B, and a number of mutants of LC3B to probe the interaction with the LIRs of NBR1. The second part was changing the plasmid containing the hATG8 proteins and then expressing and purifying these for study against the LIRs of NBR1

3.1 Mutants of LC3B

Previous work within the Searle group had produced the full length LC3B WT (125aa) and a number of mutants of this full length LC3B. Whilst this full length protein is not the biologically active form of LC3, which is a truncated (120 aa) and conjugated to a lipid as part of the membrane of the autophagosome, it was felt this would be the most appropriate form to use. Carrying out the MS and NMR experiments planned on the membrane bound form would be difficult if not impossible, and it was hoped that by using the 125 amino acid version, that the extra amino acids at the C-terminal would go some way to replicating the extension seen of the truncated version conjugated to the lipid. This form is also the one that is widely studied in the literature.

The mutants produced were a binding site mutant LC3B KL (K51A, L53A) and a phosphorylation mimic mutant LC3B T29D. The LC3B KL binding site mutant contained a double mutation with both Lys51 and Leu53 changed to alanine. Both these residues sit on the front face of LC3B near HP1 and so their mutation should have a large effect on the binding interaction with the LIRs. As these residues are on the surface of the LC3B, the effect on the 3D structure should be minimal. These mutations are on the central beta sheet of LC3B surrounding the first hydrophobic pocket into which the LIR's aromatic residue sits (Figure 3.1). This modification of HP1 would alter the binding of the LIRs, investigating the importance of the aromatic residue, and allow further study of the effect the amino acids preceding the aromatic residue have when binding to the hATG8s.

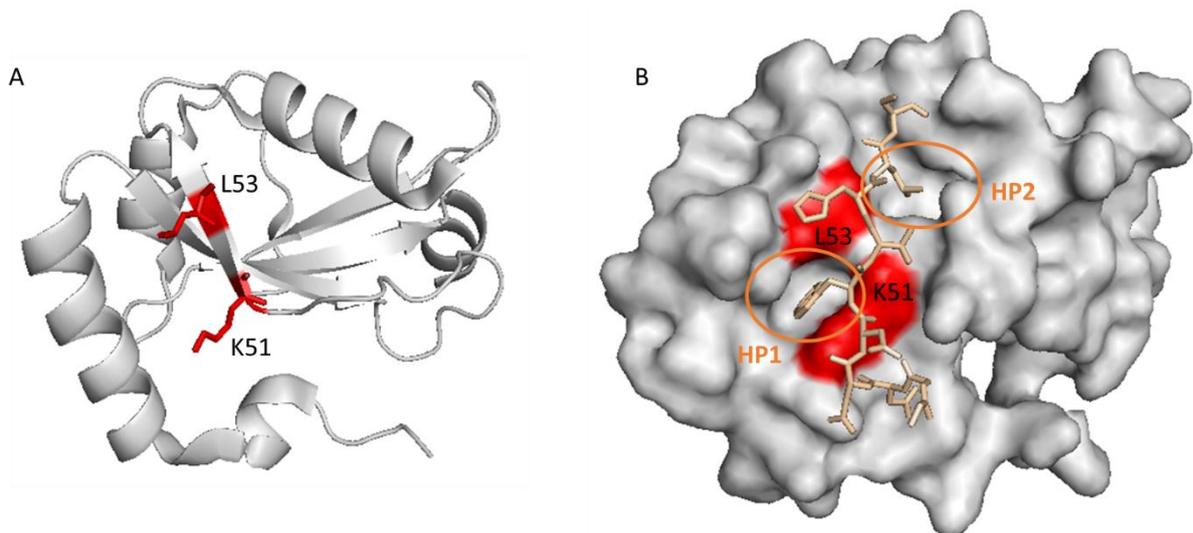


Figure 3.1: 3D structure of LC3B showing the location of the KL mutation in relation to the hydrophobic pocket and p62 LIR.

The LC3B T29D phosphomimetic mutant would act as another binding site mutant. There are a number of phosphorylation sites on LC3B that have previously been discovered at Thr6, Ser12 and Thr29 [120, 121]. Of these sites Thr29 was chosen due to it being located at the top of the alpha helix that sits next to the central beta sheet that forms HP1 and also due to its proximity to HP2 (Figure 3.2). A common method to examine phosphorylation when using *E. coli* overexpressed protein, is to mimic the effect by mutating the residue to aspartic acid [132, 133]. LC3B T29D does not completely replicate phosphorylation at a site due to the differences in size and number of oxygen's. However, it does provide an accessible experimental tool to gain insight into the possible mediation of LIR interaction by phosphorylation.

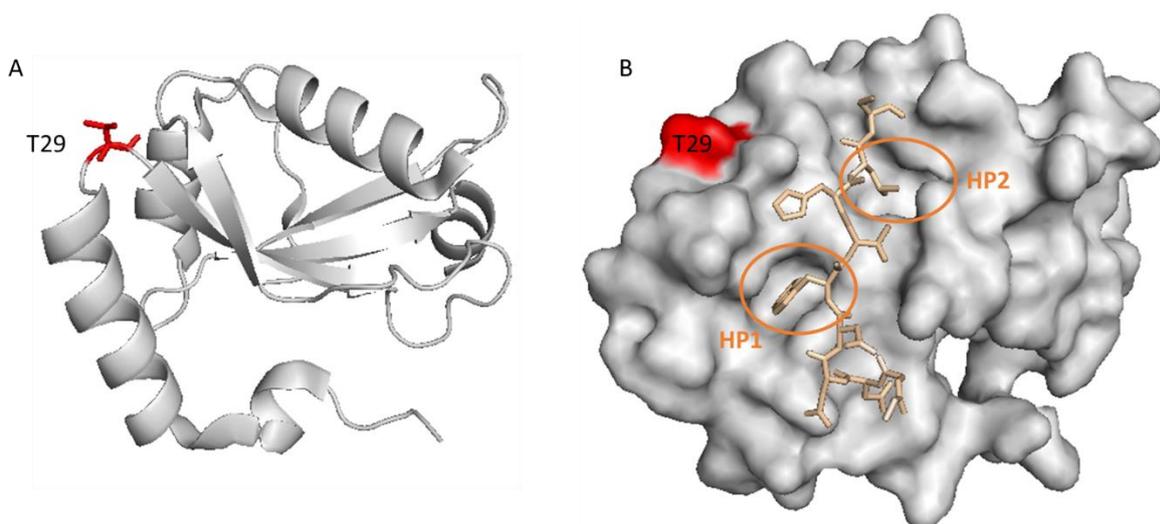


Figure 3.2: 3D structure of LC3B showing location of Thr29 in relation to the hydrophobic pockets and p62 LIR.

The WT LC3B and its two mutants produced would contain an extra 5 amino acids (GSPNS) at the N-terminus due to the linker present for the GST-tag. However, it was felt that this was a minimal extension and should have no effect on the binding due to its location.

3.2 Modification of the ATG8 Proteins

Plasmid containing the 6 hATG8 protein gene was supplied by Prof. Robert Layfield (University of Nottingham Medical School, QMC). Upon transforming the plasmid into C41 cells (Chapter 2.3.8) initial large scale protein over-expression (Chapter 2.3.11) showed expression bands for LC3A and LC3B (Figure 3.3). However, during cleavage of the GST with Tobacco Etch virus (TEV) the protein became insoluble and impossible to purify further.

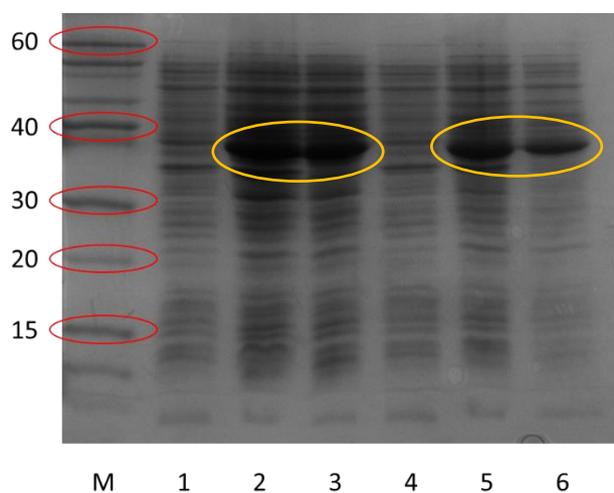


Figure 3.3: SDS PAGE for LC3A and LC3B 1L over-expression in LB; lane annotations: M = marker, 1 = LC3A uninduced, 2 = LC3A after induction - total, 3 = LC3A after induction – soluble, 4 = LC3B uninduced, 5 = LC3B after induction - total, 6 = LC3B after induction – soluble. Expression bands for LC3A and LC3B are circled in orange.

As LC3B had successfully been produced using the pGEX4T-1 plasmid all the hATG8 proteins were cloned onto the pGEX4T-1 vector which would then allow purification of the protein through removing the GST tag with a thrombin cleavage site.

To enable this transformation between plasmids the correct restriction sites would have to be added to the sequence in the pETM30 plasmid to allow the protein to be introduced into the correct part of the pGEX4T-1 for the thrombin cleavage site (Figure 2.1). The sites decided upon were the *XhoI* and *BamHI* as this would then leave a minimal extension of glycine and serine on the N-terminus of the hATG8 which should then not have any effect on the binding studies to be carried out.

To introduce the cut sites primers were designed (Table 3.1) and the hATG8 containing pETM30 plasmids subjected to site directed mutagenesis (Chapter 2.3.5).

| Protein | Cut Site | Primer |
|-----------|---------------|-----------------------------------|
| LC3A | <i>Bam</i> HI | CGCGGATCCATGCCCTCAGACCG |
| | <i>Xho</i> I | GCGATCCGCTCGAGTCAGAAGCCGAAGGTTTC |
| LC3B | <i>Bam</i> HI | CGCGGATCCATGCCGTCGGAGAAGACC |
| | <i>Xho</i> I | GCGATCCGCTCGAGTTACTGACAATTTTCATC |
| LC3C | <i>Bam</i> HI | CGCGGATCCATGCCGCCTCCACAGAAAATC |
| | <i>Xho</i> I | CGCTAGGGCTCGAGTTAGAGAGGATTGCAGGG |
| GABARAP | <i>Bam</i> HI | CGCGGATCCATGAAGTTCGTGTACAAAG |
| | <i>Xho</i> I | GCGATCCGCTCGAGTCACAGACCGTAGACAC |
| GABARAPL1 | <i>Bam</i> HI | CGCGGATCCATGAAGTTCAGTACAAG |
| | <i>Xho</i> I | GCGATCCGCTCGAGTCATTTCCCATAGACACTC |
| GABARAPL2 | <i>Bam</i> HI | CGCGGATCCATGAAGTGGATGTTCAAG |
| | <i>Xho</i> I | GCGATCCGCTCGAGTCAGAAGCCAAAAGTGTC |

Table 3.1: Primers used to introduce *Xho*I and *Bam*HI cut sites into hATG8 containing pETM30 plasmid.

The resulting PCR processed plasmid was then cut using *Xho*I and *Bam*HI restriction enzymes. The cut DNA was then run on an agarose gel (Chapter 2.3.7) to confirm the presence of cut DNA and to separate it from the remaining plasmid DNA (Figure 3.4).

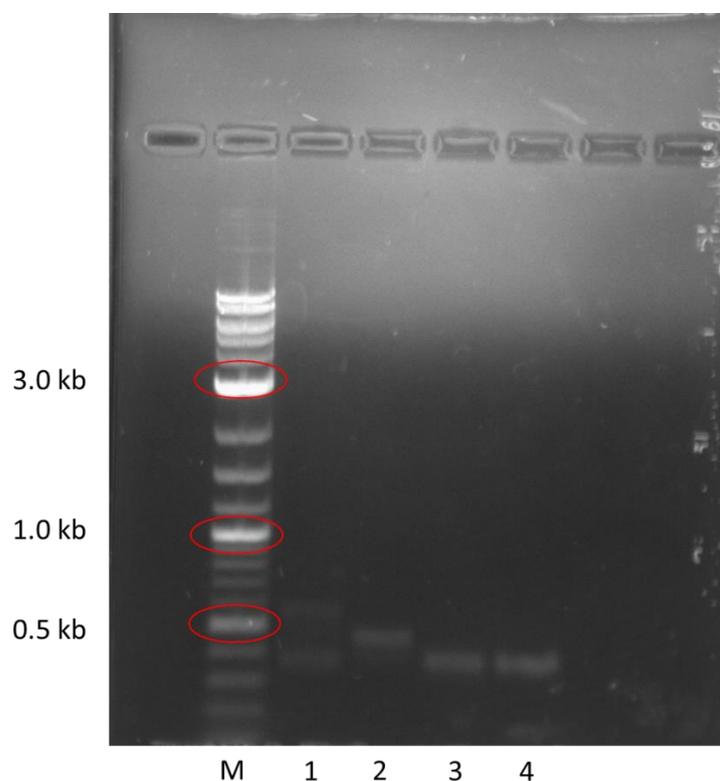


Figure 3.4: Agarose Gel showing cut DNA from pETM30 plasmid; lane annotations M = marker, 1 = LC3A, 2 = LC3C, 3 = GABARAP, 4 = GABARAPL1.

The bands for the cut DNA were the cut out of the agarose gel and purified using a gel extraction kit (Qiagen, QIAquick Gel Extraction Kit)). The extracted DNA then underwent ligation with a stock of pre-cut pGEX-4T-1 plasmid (Chapter 2.3.6) and the resulting plasmid was transformed into XL1 blues (Chapter 2.3.8). Colonies from the agar plates were then grown and plasmid DNA extracted (Chapter 2.3.9) and submitted for DNA sequencing (Chapter 2.3.10). Unfortunately on analysis of the DNA sequencing it was found that only plasmid was present and so the pre-cut plasmid had closed under ligation conditions without inserting the DNA for the hATG8 proteins. The digestion and ligation step was then repeated with the extracted cut DNA. However, this time uncut pGEX-4T-1 plasmid was produced fresh using the same *XhoI* and *BamHI* enzyme digest as the PCR product. However, again after work up and sequencing only closed plasmid without the DNA for the hATG8s was produced. At this point due to time constraints Dr Jed Long from the Searle group optimised the conditions for the enzyme digest and ligation of the cut hATG8 DNA with the pGEX-4T-1 plasmid and produced plasmid for all the hATG8s. These were then transformed into C41 DE3 *E. coli* cells (Chapter 2.3.8).

Initial test growths for the hATG8 proteins were carried out using 20 mL of LB in universals and after induction were incubated at 30°C. These showed expression bands for the proteins (Figure 3.5) and so large scale protein overexpression could proceed.

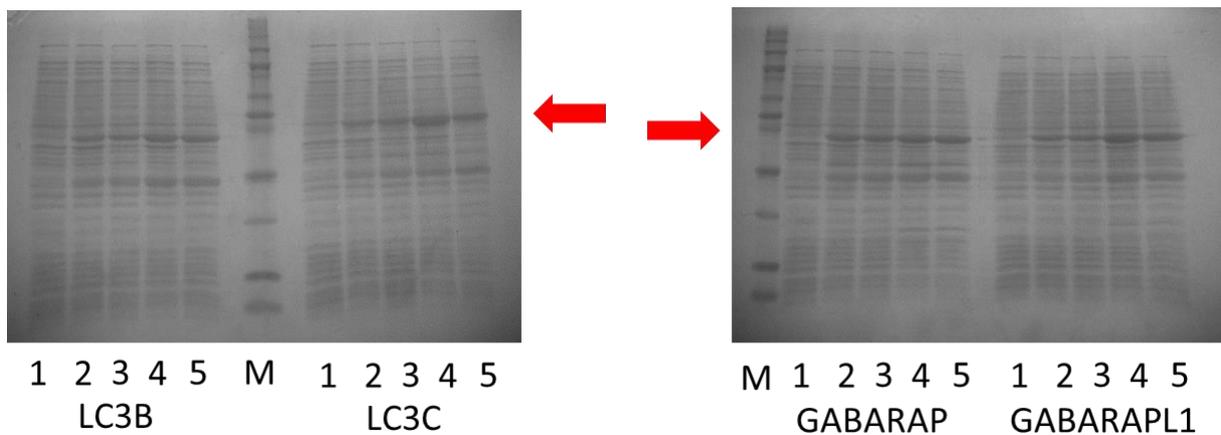


Figure 3.5: SDS PAGE gel of hATG8 test growths for LC3B, LC3C, GABARAP and GABARAPL1. Labels for lanes are as follows: M= marker, 1=uninduced, 2=total cell lysate, 3 hours after induction, 3=soluble cell lysate, 3 hours after induction, 4=total cell lysate, 21 hours after induction and 5=soluble cell lysate, 21 hours after induction. Band of interest is indicated with red arrow.

3.3 Protein Production and Purification

LC3B mutant and hATG8 proteins were produced by first putting on samples of LB containing the desired construct and leaving them to grow overnight (Chapter 2.3.2) and subsequently transferred to fresh media and grown on (Chapter 2.3.11). Initial growths carried out in minimal media with only ammonium chloride, glucose and ampicillin gave poor yields of protein. However, after trying to vary a number of conditions (temperature of incubation, amounts used), it was found that the addition of vitamins and trace elements improved the yields of protein produced to amounts that would be expected for expression in minimal media.

There were also issues with the large scale growths not growing and hence not producing any protein. The growths would reach an optical density 0.5-0.7 AU which was the range that had been used with previous proteins grown in the group to trigger induction. However, on induction the *E. coli* would fail to grow any further and no protein would be produced. It was found that by increasing the density of *E. coli* before induction from the usual 0.5-0.7 AU to 0.7 – 1.0 AU that the *E. coli* would continue to grow overnight and produce protein. The cells in the growths were harvested and stored for future purification.

Pure protein was then obtained by first separating the GST-tagged protein from the *E. coli* cells using affinity chromatography followed by thrombin cleavage of the GST tag leading to the protein being collected (Chapter 2.4.1). This soluble protein with some impurities was then further cleaned.

Whilst previous work within the Searle group had used fast protein liquid chromatography (FPLC) with gel filtration to purify the LC3B mutants this method had still left impurities within the sample. To improve the protein purification process it was decided to use ion exchange chromatography instead. Firstly the isoelectric point (pI) for the proteins was calculated so as to determine what type of column would be best suited for the purification and the pH of the buffer to be used. The pIs of all the hATG8 proteins were calculated (using ExPASy ProtParam[134]) and found to be greater than 8 for all proteins. Thus if the buffer used had a pH of 7 then the protein would be positively charged and a cationic column could be used. A HiTrap™ SP column was selected which use sulfopropyl groups as the negative stationary phase. The protein was loaded and a salt gradient could be run which as the positive charge in the buffer increased would overload the cationic column until the protein was released. Using the method in Chapter 2.4.4 pure protein was collected. An example trace for the ion exchange chromatography is shown in Figure 3.6

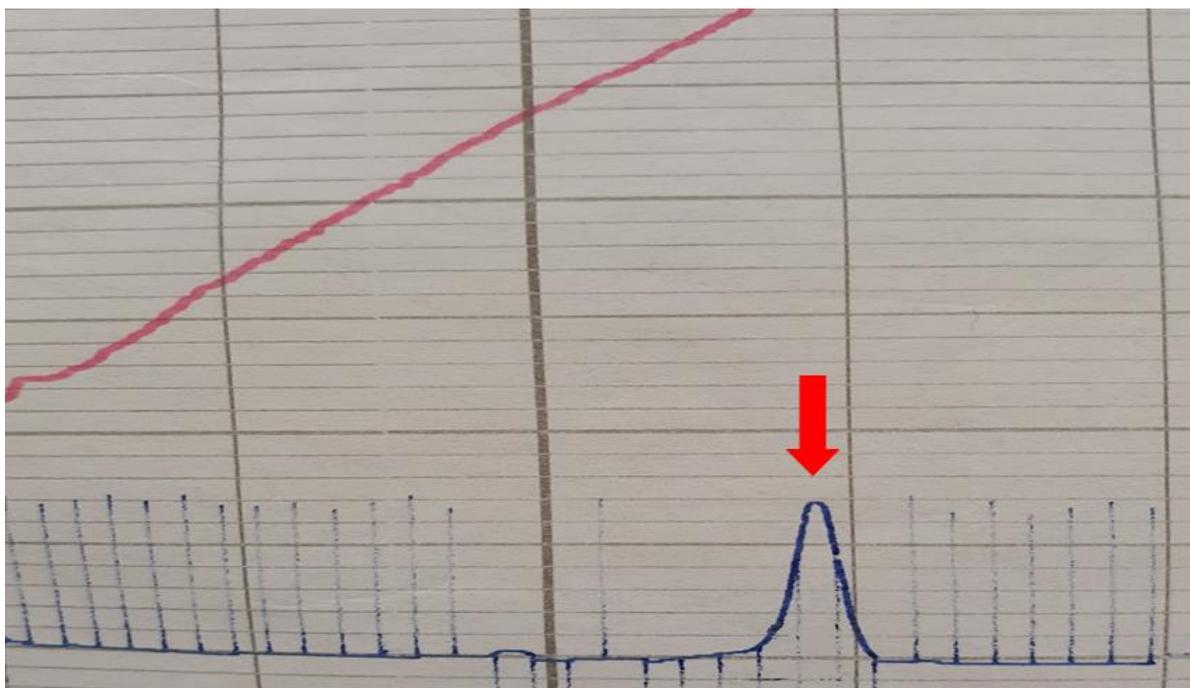


Figure 3.6: Ion exchange chromatogram of IC3B T29D. A₂₈₀ is shown in blue, and conductivity in red. The protein of interest is indicated by the red arrow.

The protein containing fractions were isolated and then lyophilised. The resulting samples contained large amounts of salt due to the nature of the ion exchange purification and would therefore not be suitable for NMR and MS analysis. To remove the excess salt FPLC was again used with a desalt column attached (Chapter 2.4.4). An example trace of the desalt process is given in Figure 3.7.

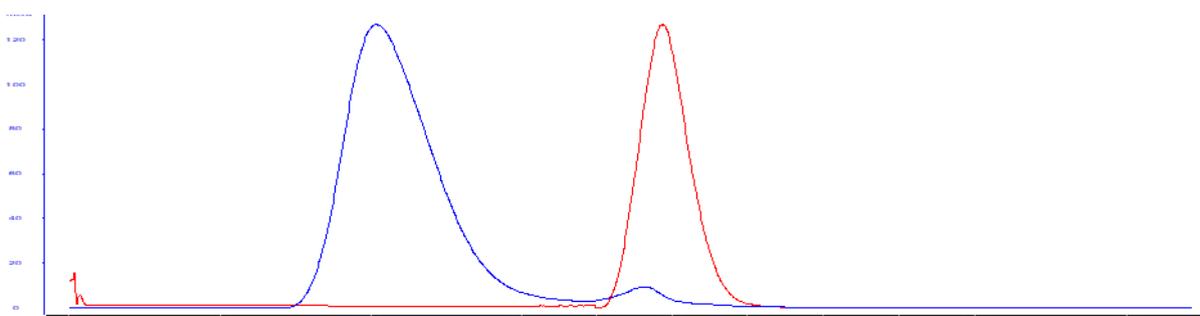


Figure 3.7: Desalt chromatogram of LC3B KL. A₂₈₀ is shown in blue, and conductivity in red.

This yielded pure lyophilised protein in varying quantities. Whilst non-isotopically labelled protein samples of LC3A, LC3B and the LC3B mutants tended to produce between 10 to 15 mg of protein the yields for the other hATG8s proteins was considerably less. LC3C would produce approximately 5 mg of protein per litre of LB media and the GABARAPS only 1-2 mg. This was enough for the experiments that were to be carried out initially and so although there was some discrepancy between the yields of hATG8 protein it was decided not to look at further optimisation due to time constraints. However, for any future work it would be recommended to look closer at these yield discrepancies. For isotopically labelled samples of LC3A and LC3B between 5 – 6mg of protein was produced.

SDS Page Gels were run (Chapter 2.4.4) to check new purifications and when there were problems with the purification. An example of this is shown for LC3B T29D in Figure 3.8

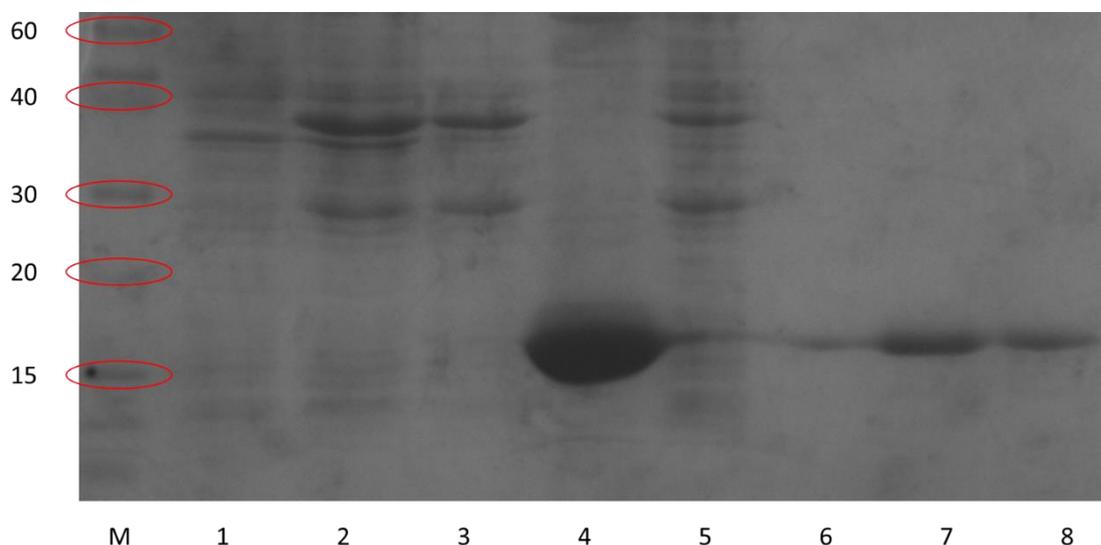


Figure 3.8: SDS PAGE gel monitoring the various steps in production and purification of LC3B T29D. Labels for lanes are as follows: M=marker, 1 = uninduced, 2 = after induction - total, 3 = after induction – soluble, 4 = cleaved protein from column, 5 = supernatant after cell extraction and 6,7,8 = protein containing fractions from ion exchange column.

3.4 Initial Analysis of the hATG8 Proteins

With the hATG8 proteins successfully produced it was important to do some initial analysis of the proteins to confirm the samples were the correct mass, did not contain large impurities and that the protein was folded and had not become denatured during the purification process. MS and NMR were used to perform this initial analysis.

3.4.1 Initial MS of the hATG8 Proteins

Mass spectrometry is a widely used tool for accurate mass determination and characterisation of proteins. The use of Electrospray ionisation (ESI) a “soft” ionization technique allows the production of intact folded protein ions from large biomolecules in solution [135, 136]. For the initial analysis of the hATG8 proteins ESI-MS would be able to confirm an accurate mass for the native protein and determine whether the protein was in its native state or denatured. If the protein was in its native state (folded) there would only be a few peaks present in the mass spectrum representing charge states of the protein. However, if the protein was denatured/unfolded there would be considerably more peaks as greater areas of the protein would be accessible and hence more charge states would be available. An example of native and denatured protein is given in Figure 3.9. It would also give some information on whether there were any major impurities present within the sample.

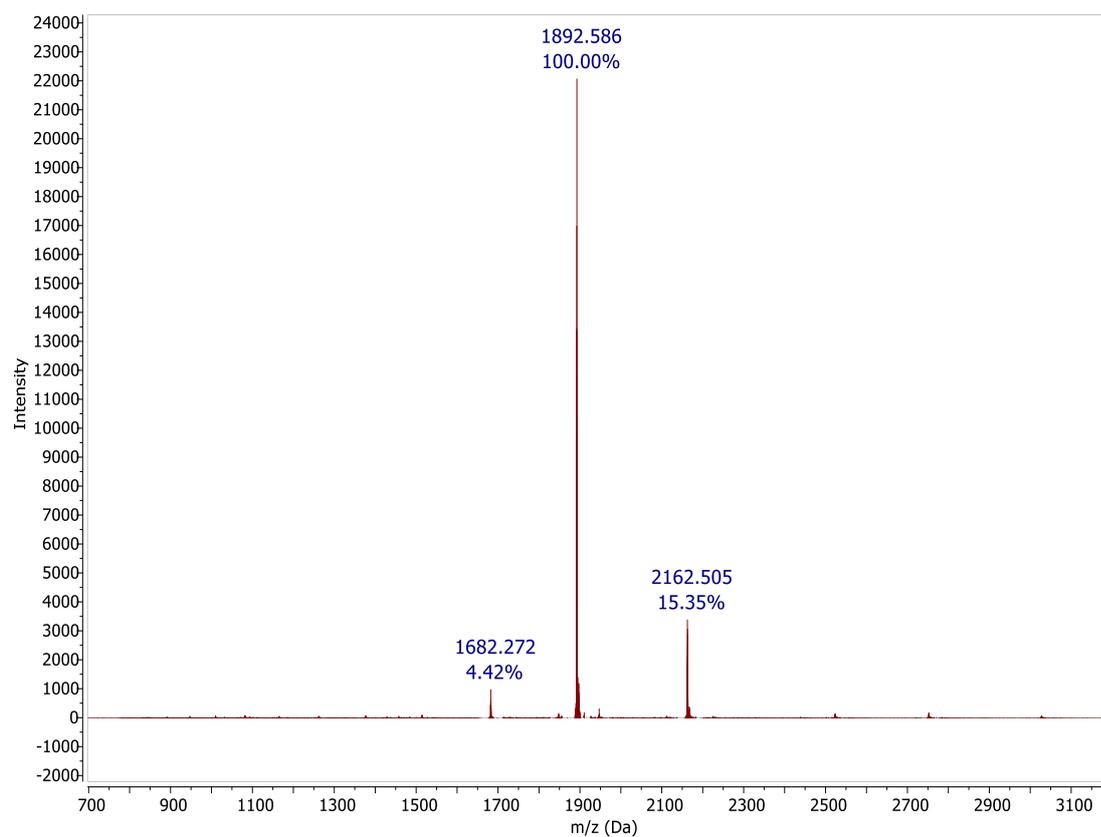
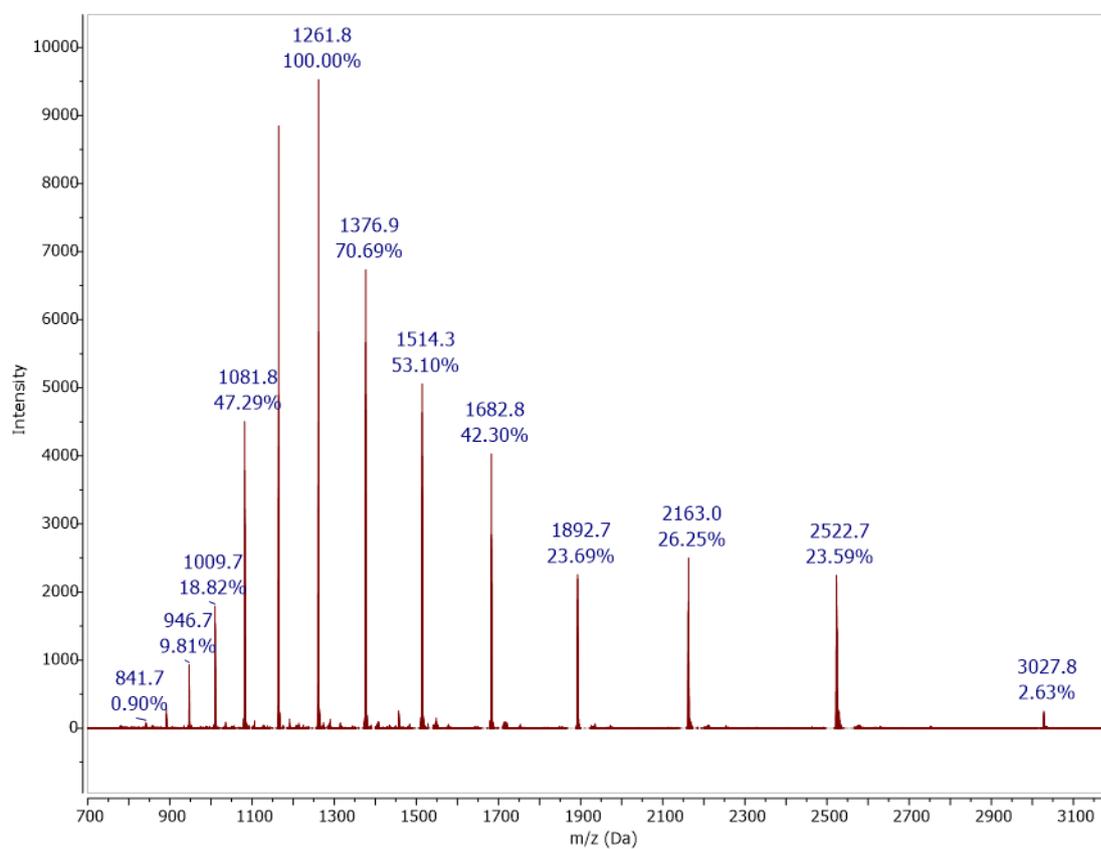
A**B**

Figure 3.9: Example of native MS spectrum of LC3B (A) and denatured (B).

The MS spectra of the hATG8 proteins and mutants showed a number of peaks for the protein corresponding to the different ionisation states of the protein. In general for the hATG8 proteins the +9, +8, +7 and +6 states were most prevalent. The exception to this was LC3C which has an extended unstructured N- and C-terminal and contained more peaks in the MS spectrum due to the increased number of charge states available. The masses obtained for the hATG8 and LC3B mutant proteins are given in Table 3.2. MS spectrum for the other proteins produced are given in Appendix B.

| Protein | Mass of protein | Mass from MS | Native Charge States |
|-----------------------------|-----------------|--------------|--|
| LC3B WT (initial construct) | 15131.4 | 15130.37 | +8 (1892.4) +7 (2162.4) +6 (2523.8) |
| LC3B KL | 15032.2 | 15031.18 | +8 (1879.9) +7 (2148.3) +6 (2506.2) |
| LC3B T29D | 15329 | 15327.99 | +8 (1917.0) +7 (2190.7) +6 (2555.7) |
| LC3A | 14416.6 | 14418.18 | +8 (1803.0) +7 (2060.4) |
| LC3B | 14832.2 | 14834.43 | +8 (1855.3) +7 (2120.0) |
| LC3C | 16995.7 | 16994.02 | +11 (1546.0) +10 (1700.6) +9 (1889.5) +8 (2125.4) +7 (2428.7) +6 (2833.1) |
| GABARAP | 14062.2 | 14062.15 | +8 (1759.2) +7 (2010.4) +6 (2344.8) |
| GABARAP-L1 | 14188.2 | 14189.81 | +8 (1774.5) +7 (2028.2) |
| GABARAP-L2 | 13811 | 13812.72 | +8 (1727.6) +7 (1974.5) |

Table 3.2: Masses of hATG8 proteins from MS data.

3.4.2 One Dimensional NMR of the hATG8 Proteins

Although the 1D proton NMR of proteins are too complex for it to be possible to assign peaks to particular protons in the protein they can be a useful diagnostic on the state of the protein being analysed. If the protein is folded then the dispersion of peaks in the spectrum will be much greater than that of an unfolded protein. This is due to the large effect secondary structure has in creating a much larger number of environments that the protons can reside in. As well as giving a good indication as to whether a protein is folded a ^1H

spectrum indicates if there are any large impurities in the sample (as long as they contain protons) and whether the concentration you think you have is sensible or if in fact there is a lot less protein in the sample. Finally by looking at the resolution of the spectrum it is possible to infer some information about the dynamics of the system.

^1H proton NMR were acquired as stated in Chapter 2.5.2.2 for the LC3A, LC3B, LC3C and GABARAP. Due to the low yield of protein for GABARAPL1 and GABARAPL2 and the requirements for running MS experiments ^1H spectra were not acquired for these proteins. An example ^1H spectrum is shown in Figure 3.10 with the regions of interest highlighted. Spectra for the other proteins are in Appendix A.

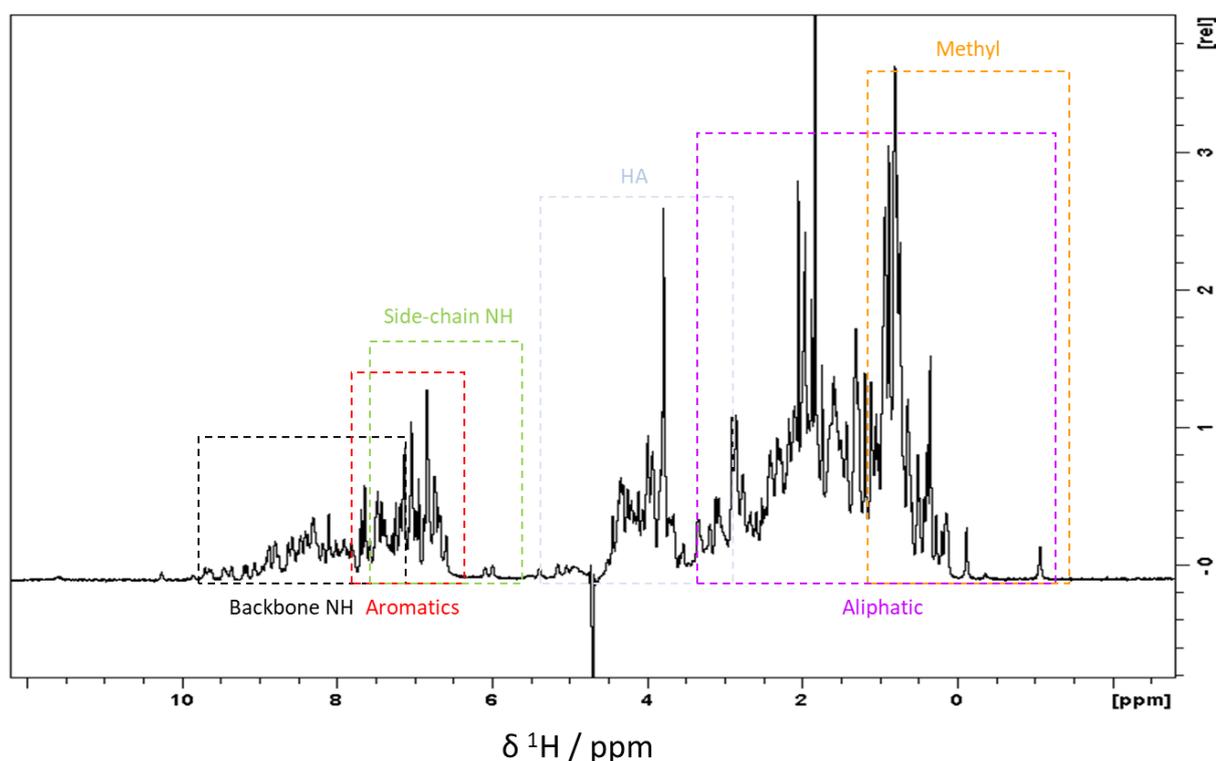


Figure 3.10: ^1H NMR of LC3B.

3.5 Conclusion

3.5.1 LC3B Mutants

The three proteins LC3B proteins, LC3B WT (initial construct), LC3B KL and LC3B T29D all expressed well under standard conditions and after purification produced good yields of pure protein. Subsequent analysis with NMR and MS confirmed that the proteins were folded and the correct mass.

3.5.2 ATG8 Proteins

Initial work to produce protein using constructs provided in pETM30 plasmid appeared to be initially successful, with over expression being seen in the SDS page gels. However, purification of the proteins was not possible and so the plasmid used was changed. On transforming the ATG8 proteins into pGEX4T-1 plasmid over expression and purification of the protein was then possible. The protein was produced with reasonable yields, and MS experiments confirmed the mass and purity of the protein samples. NMR experiments then confirmed that the protein was folded (except for GABARAPL1 and GABARAPL2 where it was not possible to obtain a spectrum). In conclusion the proteins required for the interaction studies were produced in large enough quantities for the proposed MS and NMR experiments.

Chapter 4

4. Probing Binding Selectivity of hATG8s with the LIRs of NBR1

This chapter investigates the binding preference of the two LIRs of NBR1 with LC3B mutants and hATG8 proteins using mass spectrometry (MS). It will initially confirm the binding of the two LIRs of NBR1 to the LC3B mutants and hATG8 proteins. With binding confirmed MS competition experiments will then be used to identify which protein if any, binds preferentially to the NBR1 LIRs.

NBR1 contains two LIR sequences which is unique amongst the cargo adaptor proteins. The LIRs are dissimilar in the amino acids that make up the classic recognition motif. They have different aromatic (X_0) and hydrophobic (X_3) residues with LIR1 containing tyrosine and isoleucine in these positions, and LIR2 has phenylalanine and leucine. A large number of known LIRs have acidic residues preceding the aromatic residue which is true for LIR1. However, whilst LIR2 does contain an acidic residue at position X_{-4} (Figure 4.1), this is not directly preceding the aromatic, where the acidic residues traditionally sit. Due to the difference in sequences it is likely that the two LIRs have different preferred binding partners, and are used for different roles within selective autophagy.

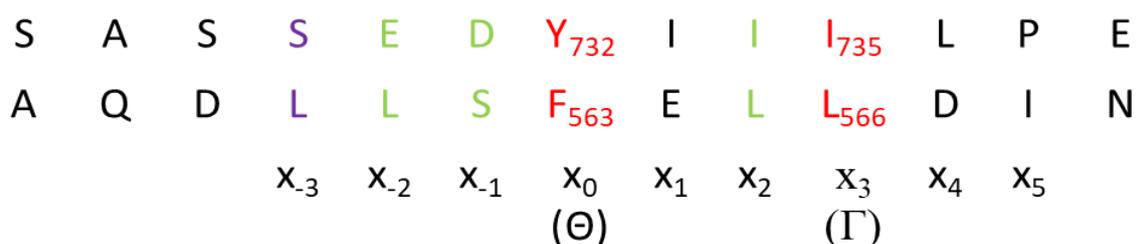


Figure 4.1: Sequences for LIRs of NBR1 (LIR1 top and LIR2 bottom). Amino acids that bind to the hydrophobic pockets of LC3B are shown in red and residues that directly precede the aromatic residue are shown in green.

There are two sub families of hATG8 protein, the LC3s and GABARAPs, and whilst they share structural features, and the LIR binding site is the same across the hATG8 family, there is low sequence alignment between the two subfamilies. Even within the LC3 and GABARAP sub families there are some large differences. LC3A has a high sequence alignment with LC3B and GABARAP with GABARAPL1 (80-90%). However, LC3C and GABARAPL2 are quite different from the rest of the sub family, with about 60% sequence alignment. So with the differences between the 6 hATG8 proteins, it seems plausible that the LIRs of NBR1 could bind preferentially to one hATG8 over another.

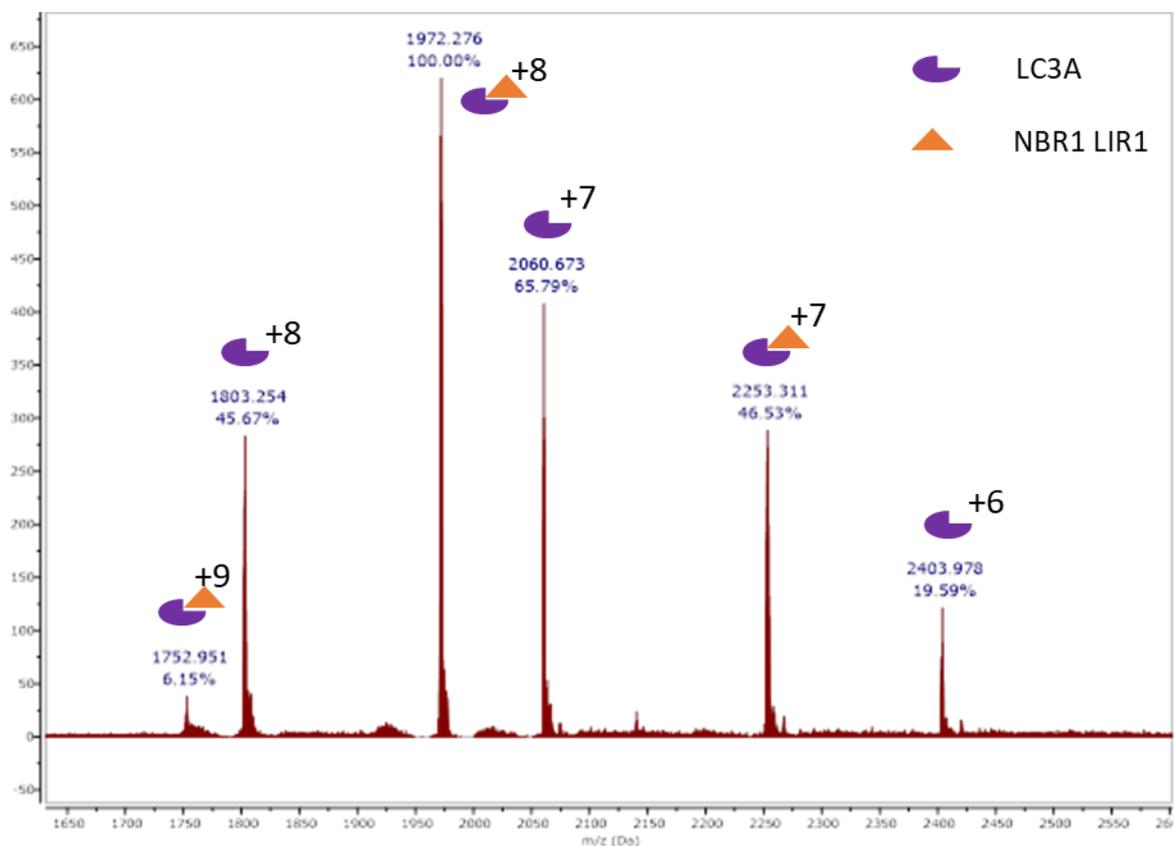
The KL and T29D mutant of LC3B, which change the binding surface and mimic phosphorylation respectively, could also preferentially bind to one of the LIRs, which would shed more light on how the sequence of the LIRs drives selectivity. To investigate if the LIRs showed any binding preferences between the LC3B mutants and the 6 hATG8 proteins a series of MS competition experiments were carried out.

4.1 Binding of LIRs of NBR1 with the hATG8 Proteins

To check that the hATG8 proteins produced (Chapter 3) did in fact bind to the peptide sequences for the LIRs of NBR1, MS was used to see if when protein and peptide were mixed, the hATG8-LIR complex could be observed. ESI MS has been shown to be a useful method for studying protein ligand-interactions, with the non-disruptive nature of the ESI source allowing noncovalent interactions to be transferred intact from solution to gas phase [137, 138].

Example spectra of LC3A and NBR1 LIR1 and LIR2 peptide at 1:1 ratio are shown in Figure 4.2.

A



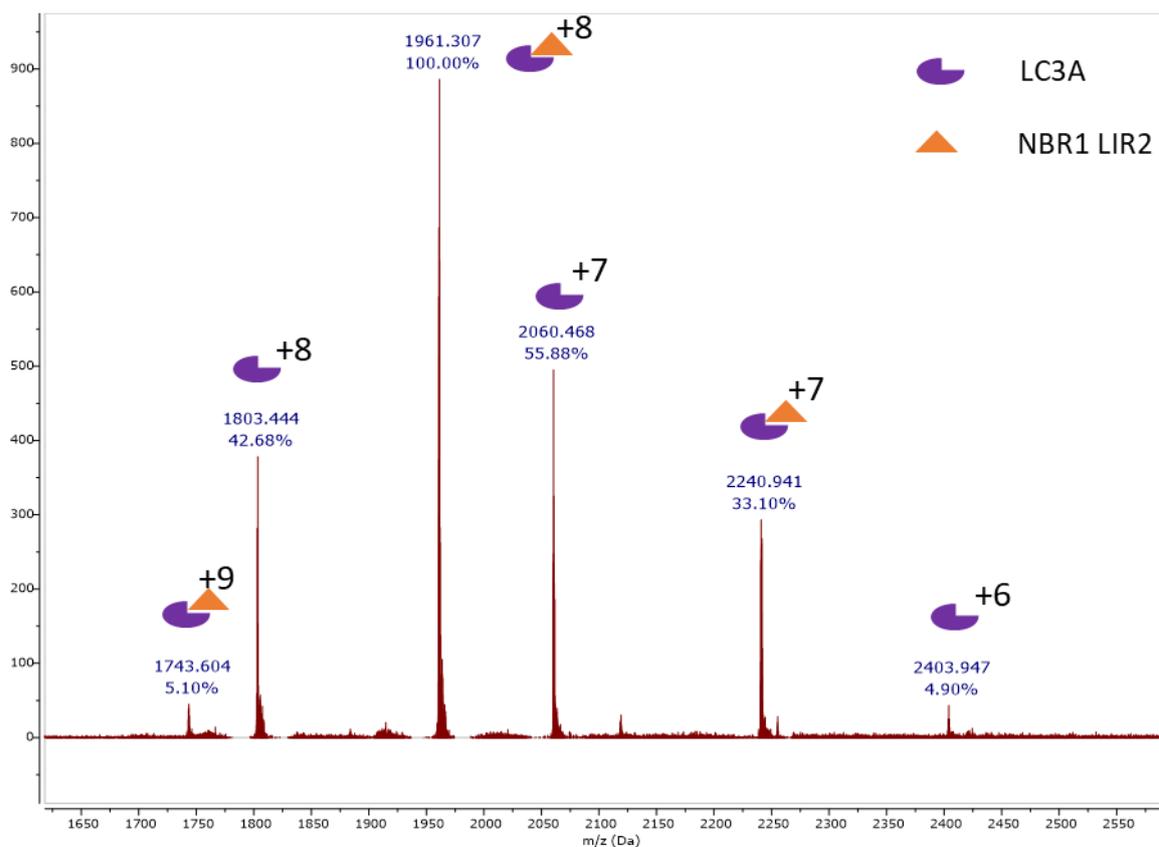
B

Figure 4.2: Mass spectra of LC3A bound to LIRs of NBR1. Spectrum shown are LC3A-LIR1 (A) and LC3A-LIR2 (B). Key indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are.

The LC3A-LIR MS spectra contain numerous peaks that represent the multiple charge states of the free and bound protein. To better understand the spectrum, which peaks are for which species and the charge states, the predicted peak positions were calculated and then compared to the spectrum (Table 4.1)

| Species | Peak | Predicted | Charge state |
|-------------|---------|-----------|--------------|
| LC3A + LIR1 | 1752.91 | 1752.79 | +9 |
| LC3A | 1803.25 | 1803.08 | +8 |
| LC3A + LIR1 | 1972.28 | 1971.76 | +8 |
| LC3A | 2060.67 | 2060.51 | +7 |
| LC3A + LIR1 | 2253.31 | 2253.30 | +7 |
| LC3A | 2403.98 | 2403.77 | +6 |

| Species | Peak | Predicted | Charge state |
|-------------|---------|-----------|--------------|
| LC3A + LIR2 | 1743.60 | 1743.22 | +9 |
| LC3A | 1803.44 | 1803.08 | +8 |
| LC3A + LIR2 | 1961.31 | 1961.00 | +8 |
| LC3A | 2060.47 | 2060.51 | +7 |
| LC3A + LIR2 | 2240.94 | 2241.00 | +7 |
| LC3A | 2403.95 | 2403.77 | +6 |

Table 4.1: Predicted values for free LC3A and LC3A NBR1 LIR complex compared to peaks in ESI-MS spectrum.

From Table 4.1 it can be seen that the peaks in the MS spectrum closely match the predicted values for both free and bound LC3A (within the error of the experiment). This confirmed the presence of the LC3A-LIR complex and hence that the protein and LIR peptide bound.

These experiments were repeated for all the mutants of LC3B (WT, KL and T29D) and the 6 hATG8s. For all the proteins, when mixed with the LIRs of NBR1, the MS spectrum showed that the protein-LIR complex was present. The spectrum and a table of peaks and various charge states observed are shown in Appendix C.

4.2 Competition Binding of LIRs of NBR1 to the LC3B Mutants Using ESI-MS

With binding of the mutants of LC3B to the NBR1 LIR1 peptides confirmed, MS competition experiments were carried out. The MS competition experiments would involve mixing all 3 LC3B proteins (WT, KL and T29D) in equal concentrations with the LIR NBR1 peptide. The Protein-LIR mixtures would then be run on the MS, and the ratio of bound peaks to free protein analysed, to determine if there was a preference for the LIR to bind to a particular LC3B mutant.

Due to the large number of species present in the MS competition experiments it was important to check if there would be any overlap of peaks that would prevent full analysis of the data. The predicted position of the peaks in the MS spectrum for the three LC3B mutants (both free and bound to LIR1) were calculated (Table 4.2) and checked for potential problems.

| Charge state | LC3B WT | LC3B KL | LC3B T29D | LC3B WT + LIR1 | LC3B KL + LIR1 | LC3B T29D + LIR1 |
|--------------|----------|----------|-----------|----------------|----------------|------------------|
| +6 | 2522.735 | 2506.205 | 2525.073 | 2747.651 | 2731.121 | 2749.99 |
| +7 | 2162.487 | 2148.318 | 2164.491 | 2355.272 | 2341.104 | 2357.277 |
| +8 | 1892.301 | 1879.903 | 1894.055 | 2060.988 | 2048.591 | 2062.742 |
| +9 | 1682.157 | 1671.136 | 1683.715 | 1832.101 | 1821.081 | 1833.66 |

Table 4.2: Predicted values for peaks in MS spectrum for competition experiment of LC3B mutants with LIR1 of NBR1.

The values indicate that the peaks of LC3B WT and LC3B T29D (both free protein and bound to LIR1) would be extremely close which would cause problems fully analysing the MS data.

To avoid the peak overlap it was decided to use ^{15}N isotopically labelled LC3B T29D. The calculated values for the peaks for the ^{15}N labelled LC3B T29D are shown in Table 4.3. ^{15}N isotopically labelled proteins are commonly used in NMR binding studies, and do not change the way the protein interacts with binding partners, and so the use of ^{15}N labelled LC3B T29D for the MS competition experiments should not have any effect on the interaction with the NBR1 LIR peptides.

| Charge state | ^{15}N LC3B T29D | ^{15}N LC3B T29D + LIR1 |
|--------------|---------------------------|----------------------------------|
| +6 | 2555.673 | 2780.59 |
| +7 | 2190.72 | 2383.505 |
| +8 | 1917.005 | 2085.692 |
| +9 | 1704.115 | 1854.06 |

Table 4.3: Predicted values for MS peaks for ^{15}N labelled LC3B T29D.

These predicted peak values in Table 4.3 for ^{15}N labelled LC3B T29D indicate that there would be no overlap with the peaks for the LC3B mutants, and so the MS competition experiments would be carried out with unlabelled LC3B WT, LC3B KL and ^{15}N labelled LC3B T29D.

4.2.1 LC3B Mutants Binding to NBR1 LIR1

A sample of the three LC3B mutant proteins and the peptide representing the LIR1 of NBR1 in a 1:1:1:1 ratio was made up and then run on the mass spectrometer. This resulted in a spectrum containing multiple peaks for the different charge states of the free protein and the protein-LIR complex shown in Figure 4.3.

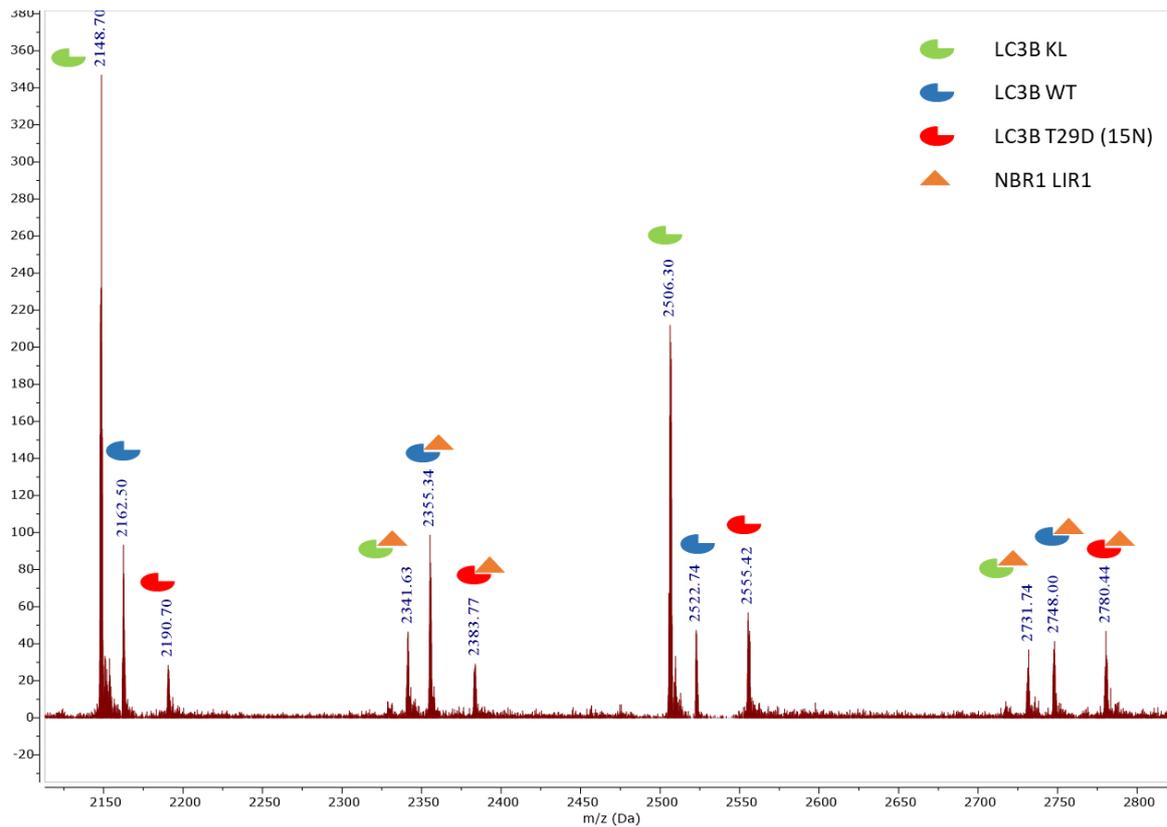


Figure 4.3: MS competition experiment for mutants of LC3B with LIR1. Sample contained LC3B WT, KL and T29D (^{15}N) and NBR1 LIR1 peptide in equal concentrations (10 μM). Key indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are.

Due to the number of species present in the mixture the MS spectra can initially seem difficult to interpret. However, by calculating the values for the peaks expected it makes it easier to assign the peaks in the actual spectrum to the protein or protein-LIR complexes. Initial analysis of this data suggested that the LC3B-LIR complex was present for all 3 mutants. When comparing the amount of free protein versus bound it was clear that there was a much larger amount of free KL protein present when compared to WT and T29D.

With the initial LC3B Mutant and LIR1 competition experiment carried out proving that it was possible to get usable data, a series of samples with varying concentrations of LIR1 varying from 0.5 to 3 equivalents of LIR1 to LC3B protein were produced. These samples were then run on the mass spectrometer and a series of spectra with varying concentrations of LIR1 were produced. Figure 4.4 shows MS spectra at 0.5 and 3 equivalents of LIR1 to LC3B mutant proteins.

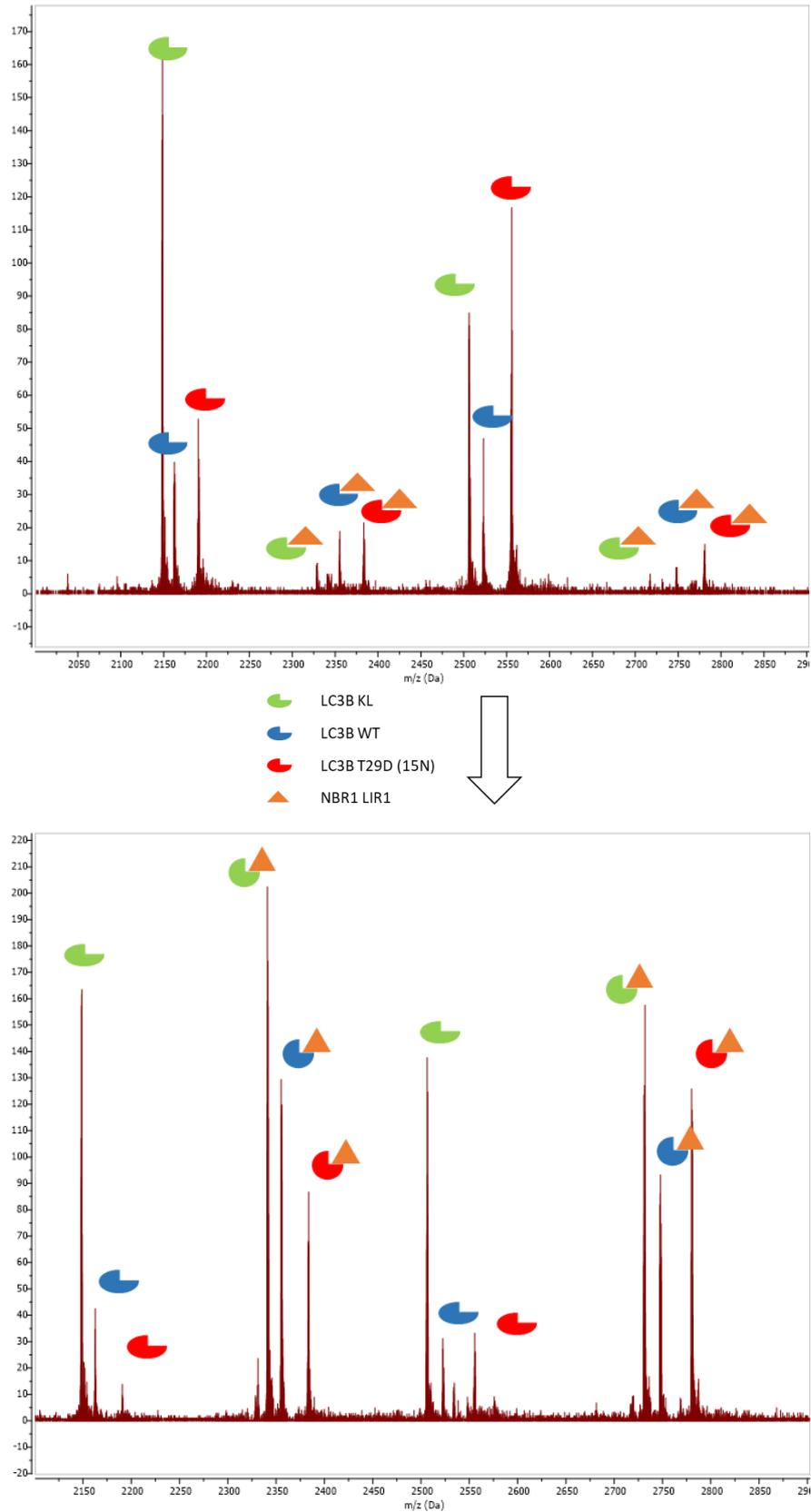


Figure 4.4: MS competition experiment for mutants of LC3B with LIR1. Sample contained LC3B WT, KL and T29D (¹⁵N) and NBR1 LIR1 peptide. Top spectrum shows proteins with 0.5 equivalent of LIR (1:1:1:0.5) and bottom spectrum shows 3 equivalents (1:1:1:3). Key indicates the peaks that belong to each protein and protein-LIR complex.

These spectra were analysed to determine which peaks corresponded to which species in the MS spectra. The obvious change was the increase in intensity of the peaks for the LC3B-LIR complex. However, to fully understand the changes in the spectrum, and to see if LIR1 showed a preference for a particular LC3B mutant, the ratio of bound to unbound protein were determined. To do this the intensity for all the unbound peaks for this particular protein were summed and the same done for the LIR1-protein complex peaks for the protein. The bound to unbound protein ratio could then be calculated to determine the preference for the LIR1 at that particular equivalence. These ratios were then calculated for each equivalence of LIR1 and plotted to see the trend as the concentration of LIR was increased (Figure 4.5).

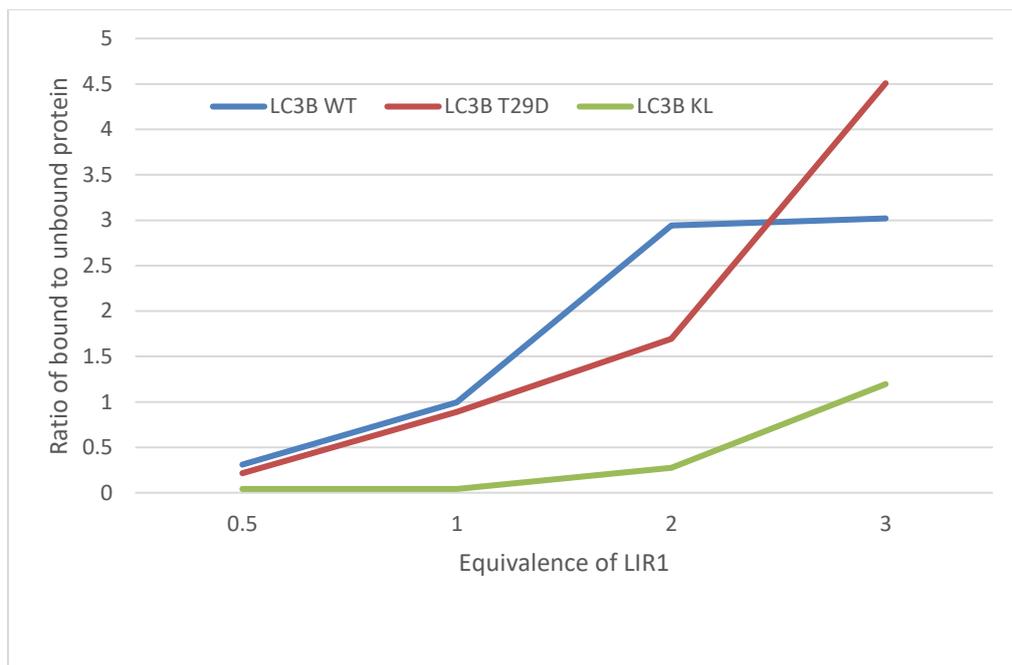


Figure 4.5: Plot of ratios of bound to unbound protein against increasing equivalents of NBR1 LIR1 peptide.

From Figure 4.5 it is clear that LIR1 shows a preference for LC3B WT and LC3B T29D over that of LC3B KL. As LC3B KL was designed to be a binding site knockout this is what would be expected, as the interaction of NBR1 LIR1 with LC3B KL would not be as strong as LC3B. However, the preference for LIR between LC3B WT and LC3B T29D is less clear cut. Although potentially at higher concentrations of LIR1 there is preference for LC3B T29D over LC3B WT. However, with the points used with this set of data it would be difficult to say this for definite without higher concentrations of LIR1 confirming the trend.

4.2.2 LC3B Mutants Binding to NBR1 LIR2

The same process as used for NBR1 LIR1 was repeated for NBR1 LIR2. The initial experiment was carried out with unlabelled LC3B WT, LC3B KL and ^{15}N labelled LC3B T29D with 1 equivalent of the NBR1 LIR2 peptide. The resulting sample was then run on the ESI-MS and the resulting spectrum analysed (Figure 4.6).

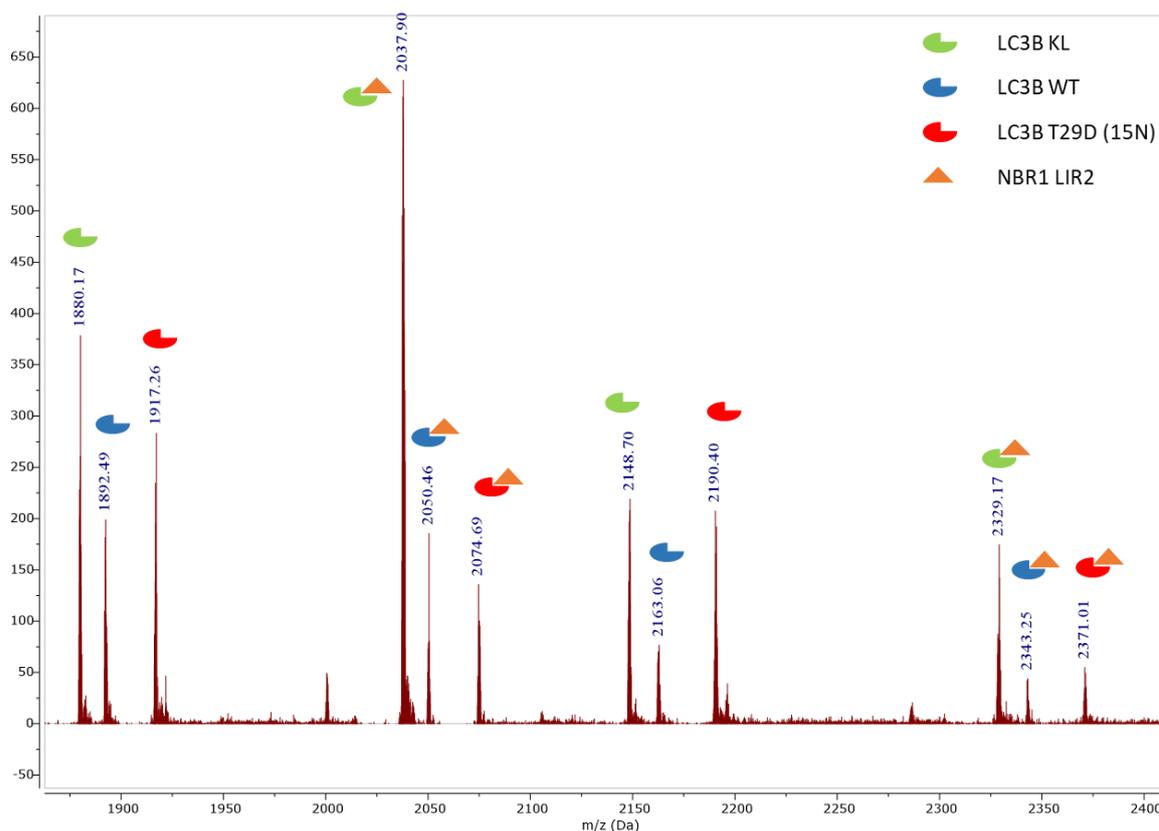


Figure 4.6: MS competition experiment for mutants of LC3B with LIR2. Sample contained LC3B WT, KL and T29D (^{15}N) and NBR1 LIR2 peptide in equal concentrations (10 μM). Key indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are found.

As before peaks for multiple charge states are present for all three of the LC3B mutants as free protein and in complex with the NBR1 LIR2 peptide. However, on initial analysis the intensity of the peaks of the LC3B KL in complex with the LIR2 peptide are much greater than those of the peaks for the WT and T29D in complex with LIR2.

To further investigate this initial observation a series of MS experiments with the three LC3B mutants mixed with different equivalences of NBR1 LIR2 peptide (from 0.5 to 3 equivalents) were run. The data was analysed and the start and endpoint are shown in Figure 4.7.

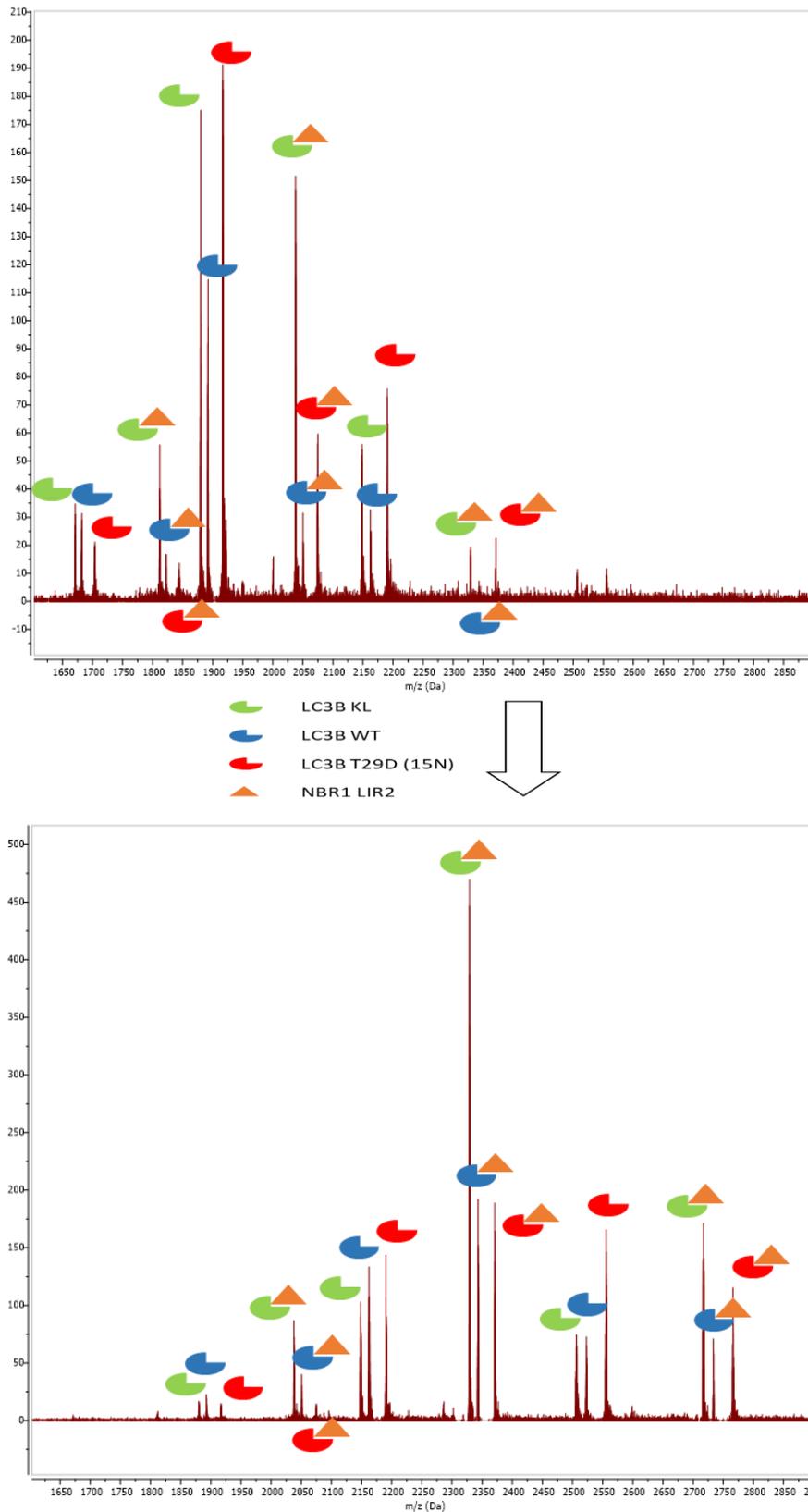


Figure 4.7: MS competition experiment for mutants of LC3B with LIR2. Sample contained LC3B WT, KL and T29D (¹⁵N) and NBR1 LIR1 peptide. Top spectrum shows proteins with 0.5 equivalent of LIR (1:1:1:0.5) and right bottom spectrum shows 3 equivalents (1:1:1:3). Key indicates the peaks that belong to each protein and protein-LIR complex.

The data again shows that the intensity for the peaks of the LC3B mutants in complex with the NBR1 LIR2 peptide as the concentration of peptide is increased. The ratio of bound protein to unbound protein was then calculated for the data and plotted (Figure 4.8).

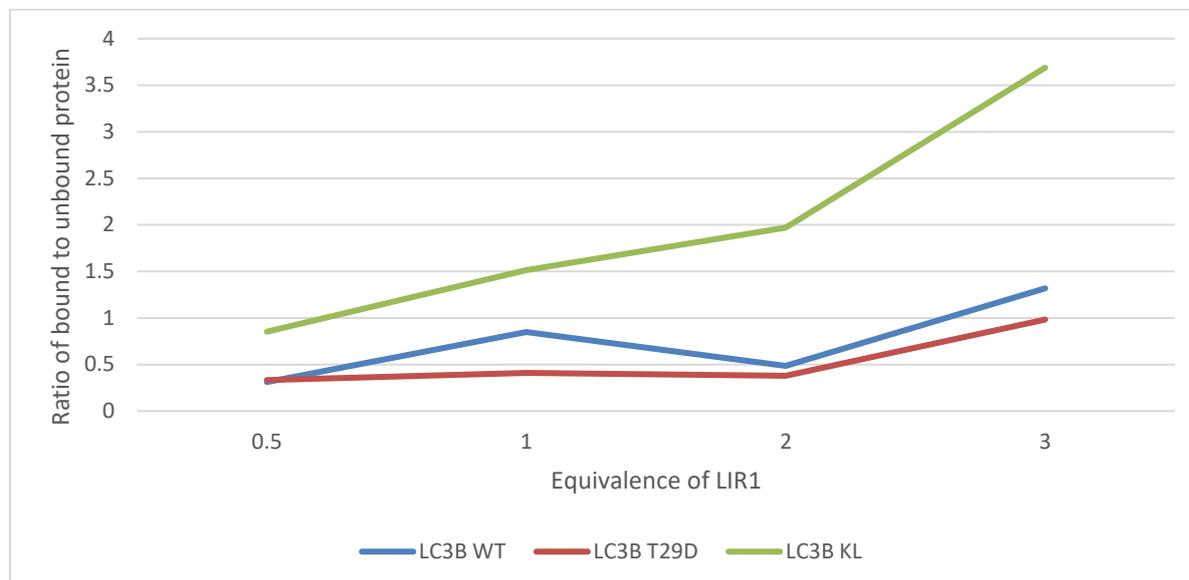


Figure 4.8: Plot of ratios of bound to unbound protein against increasing equivalents of NBR1 LIR2 peptide.

The plot in Figure 4.8 confirms the initial observation that the NBR1 LIR2 peptide shows a greater preference for the LC3B KL mutant. This is extremely interesting as the LC3B KL mutant is supposed to be a binding site knockout mutant and so should be a weaker binder. This suggests that something else, other than the main binding site, is playing a role in making the interaction of LC3B KL with the peptide of NBR1 LIR2 preferred over that of LC3B WT and T29D.

4.3 Competition Binding of LIRs of NBR1 to the ATG8s Using ESI-MS

With the binding of the hATG8s to the two peptides representing the LIRs of NBR1 confirmed (Chapter 4.1), it was then possible to use MS competition experiments to identify if there is selective binding of the NBR1 LIR peptides for a particular hATG8 protein. However, due to the number of proteins present it was not practical to mix all six together and add the peptide. Firstly the spectra would be extremely difficult to interpret due to the number of peaks present. The spectra were already complicated when just mixing the three mutants of LC3B (Chapter 4.2), and so doubling the number of peaks would make analysis extremely challenging. This is shown in Table 4.4 where the calculated m/z for common charge states is given for the free proteins and the protein-peptide complexes.

| Charge State | Species | | | | | |
|--------------|------------------|------------------|------------------|-----------------|-------------------|-------------------|
| | LC3A | LC3B | LC3C | GAB | GABL1 | GABL2 |
| +6 | 2403.77 | 2473.03 | 2833.62 | 2344.70 | 2365.70 | 2302.83 |
| +7 | 2060.51 | 2119.89 | 2428.96 | 2009.89 | 2027.89 | 1974.00 |
| +8 | 1803.08 | 1855.03 | 2125.46 | 1758.78 | 1774.53 | 1727.38 |
| +9 | 1602.84 | 1649.02 | 1889.41 | 1563.47 | 1577.47 | 1535.56 |
| | | | | | | |
| | LC3A-LIR1 | LC3B-LIR1 | LC3C-LIR1 | GAB-LIR1 | GABL1-LIR1 | GABL2-LIR1 |
| +6 | 2628.68 | 2697.95 | 3058.53 | 2569.62 | 2590.62 | 2527.75 |
| +7 | 2253.30 | 2312.67 | 2621.74 | 2202.67 | 2220.67 | 2166.79 |
| +8 | 1971.76 | 2023.71 | 2294.15 | 1927.46 | 1943.21 | 1896.06 |
| +9 | 1752.79 | 1798.97 | 2039.36 | 1713.41 | 1727.41 | 1685.50 |
| | | | | | | |
| | LC3A-LIR2 | LC3B-LIR2 | LC3C-LIR2 | GAB-LIR2 | GABL1-LIR2 | GABL2-LIR2 |
| +6 | 2614.33 | 2683.60 | 3044.18 | 2555.27 | 2576.27 | 2513.40 |
| +7 | 2241.00 | 2300.37 | 2609.44 | 2190.37 | 2208.37 | 2154.49 |
| +8 | 1961.00 | 2012.95 | 2283.39 | 1916.70 | 1932.45 | 1885.30 |
| +9 | 1743.22 | 1789.40 | 2029.79 | 1703.84 | 1717.84 | 1675.93 |

Table 4.4: Predicted MS peaks for hATG8 protein MS competition peaks.

Hence it was decided to simplify the spectra by separating the six hATG8 proteins into the two subfamilies, and so run two sets of MS competition experiment, one with the LC3 proteins and the other with the GABARAP proteins. Whilst this could potentially still leave some peaks that were close in m/z, depending on the charge states present, with the spectra having fewer peaks the resolution of the MS would be enough to differentiate between the peaks.

4.3.1 LC3s Binding to NBR1 LIR1

LC3A, LC3B and LC3C were mixed in equal concentration (10 μ M) and one equivalent of NBR1 LIR1 peptide was added. This sample was then run on the ESI-MS and the spectrum collected. As expected this gave a spectrum containing six sets of peaks corresponding to the three LC3s, both as free protein and in complex with the NBR1 LIR1 peptide. The spectrum is shown in Figure 4.9 with the peaks labelled to show the species they belong to.

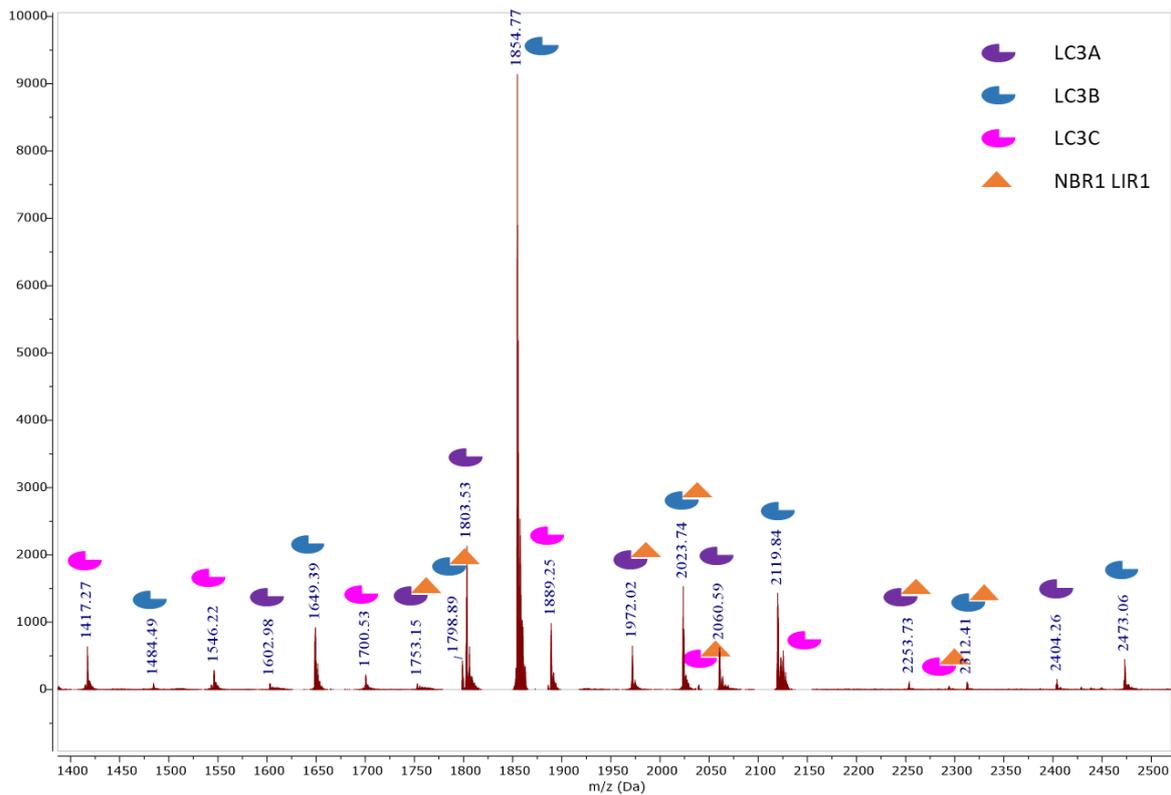


Figure 4.9: MS competition experiment for hATG8 LC3 family NBR1 LIR1 peptide. Sample contained LC3A, LC3B, LC3C and NBR1 LIR1 peptide in equal concentrations (10 μ M). Key indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are found.

On initial analysis, the first thing that stands out is that there is almost no LC3C-LIR1 complex, with the majority of LIR1 binding to LC3A and LC3B. If the ratio of bound to unbound protein is then calculated it indicates that the peptide of NBR1 LIR1 shows a selectivity for LC3A, with almost double the amount of bound to unbound compared to LC3B (Figure 4.10). There is then only a small amount of LC3C-LIR1 complex present pointing to NBR1 LIR1 binding preferentially to LC3A over the other LC3 proteins.

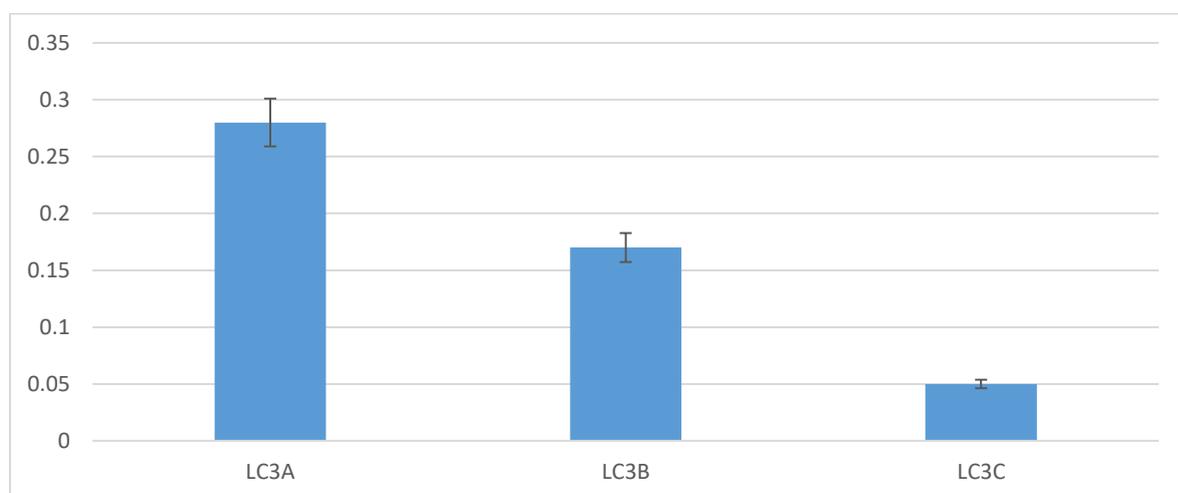


Figure 4.10: Bar chart of calculated values for ratio of bound to unbound for LC3 subfamily with LIR1 MS competition experiment.

4.3.2 LC3s Binding to NBR1 LIR2

As with LIR1 these competition experiments were carried out by mixing the 3 LC3s in equal amounts with 1 equivalent of the NBR1 LIR2 peptide which was then run on the ESI-MS. The resulting spectrum is shown in Figure 4.11 with the peaks labelled to indicate the protein/protein-complex the peak corresponds to.

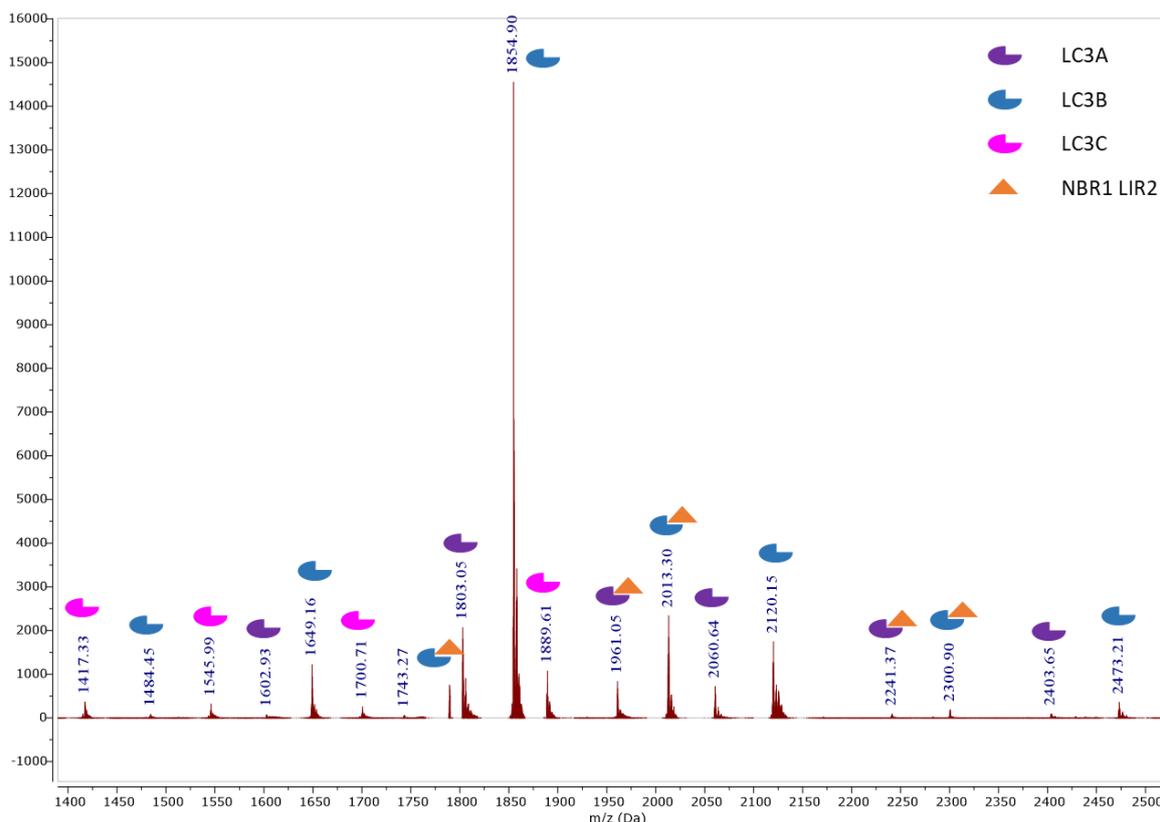


Figure 4.11: MS competition experiment for hATG8 LC3 family NBR1 LIR2 peptide. Sample contained LC3A, LC3B, LC3C and NBR1 LIR2 peptide in equal concentrations (10 μ M). Key indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are found.

When this spectrum was analysed it was again apparent that the majority of protein-ligand complex peaks present were for LC3A-LIR2 and LC3B-LIR2. In fact, with the LIR2 MS competition experiment it was not possible to find any peaks corresponding to the LC3C-LIR2 complex. When the data was analysed the NBR1 LIR2 peptide seemed to show a strong selectivity for LC3A, with almost double the amount of bound to unbound protein present compared to LC3B (Figure 4.12).

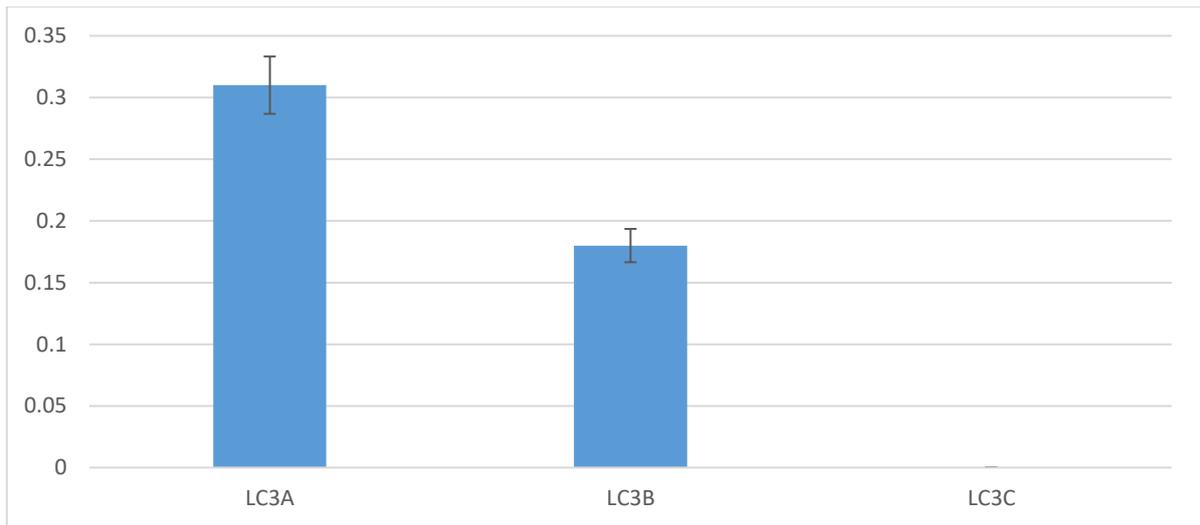


Figure 4.12: Bar chart of calculated values for ratio of bound to unbound for LC3 subfamily with LIR2 MS competition experiment.

4.3.3 GABARAPs Binding to NBR1 LIR1

As with the LC3 family MS competition experiments, all 3 GABARAP proteins were mixed in equal amounts with 1 equivalent of NBR1 LIR1 peptide. The spectrum obtained using ESI-MS is shown in Figure 4.13.

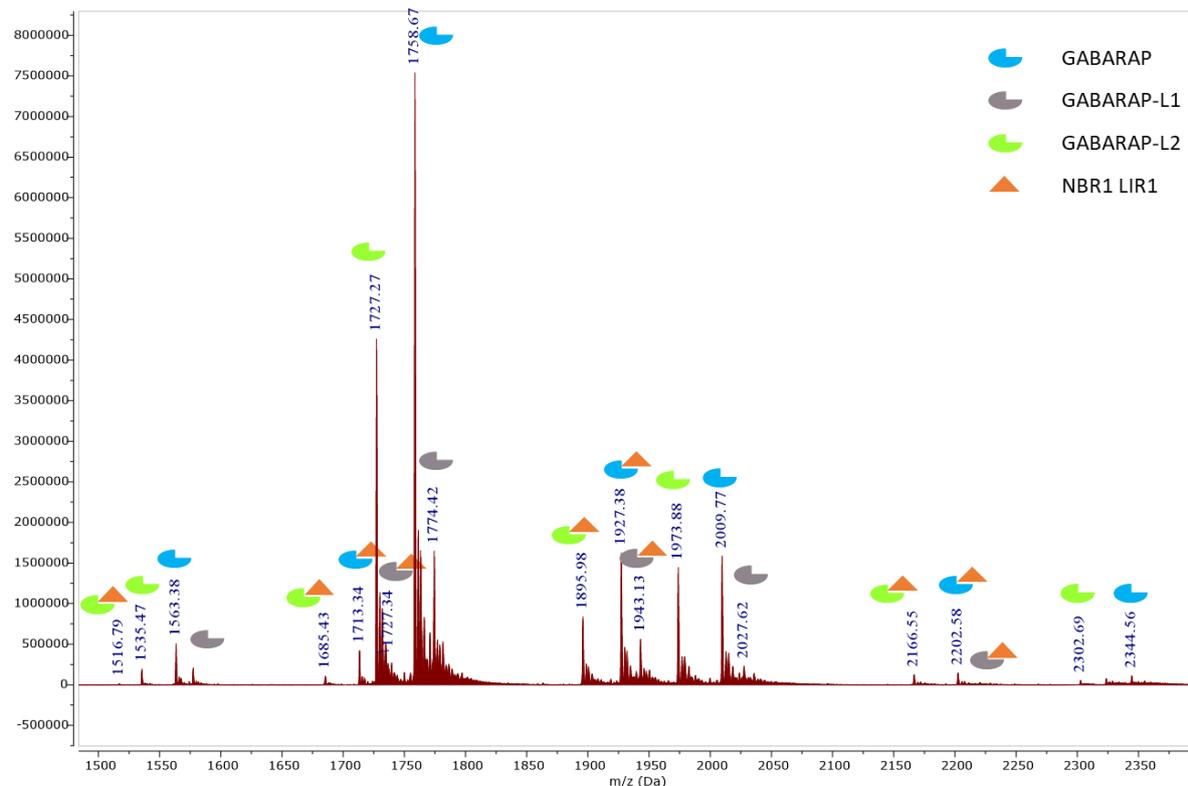


Figure 4.13: MS competition experiment for hATG8 GABARAP family NBR1 LIR1 peptide. Sample contained GABARAP, GABARAPL1, GABARAPL2 and NBR1 LIR1 peptide in equal concentrations (10 μM). Key indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are found.

As with previous MS competition experiments, peaks were present for bound and unbound protein. On comparison of the intensity of bound to unbound MS peaks for the charge states of each species, it was found that GABARAP and GABARAPL2 had a very similar affinity to the NBR1 LIR1 peptide. However, the ratio of bound to unbound protein of GABARAPL1 was double that of GABARAP and GABARAPL2 (Figure 4.14).

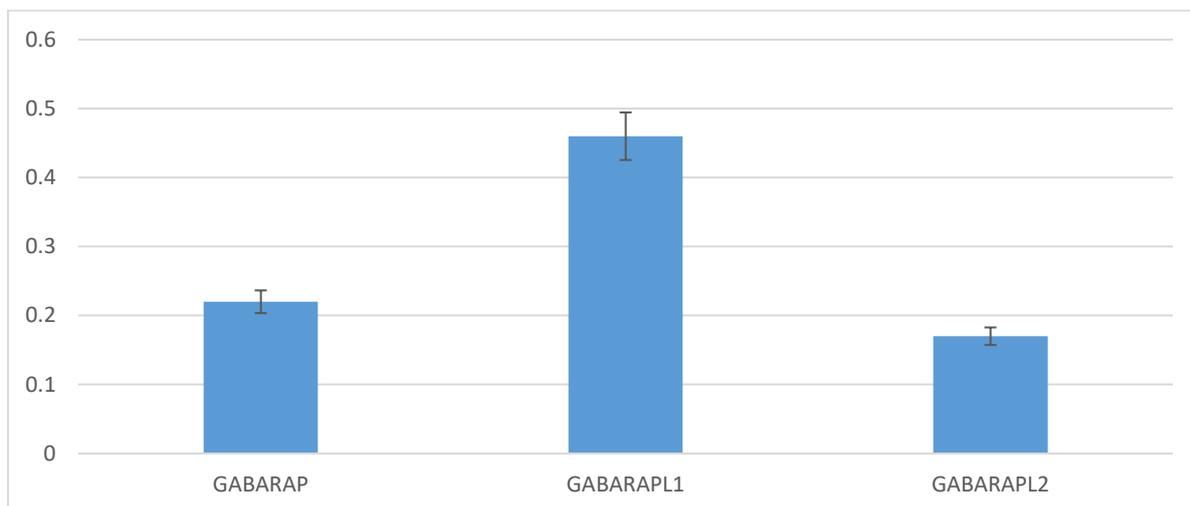


Figure 4.14: Bar chart of calculated values for ratio of bound to unbound for GABARAP subfamily with LIR1 MS competition experiment.

4.3.4 GABARAPS Binding to NBR1 LIR2

Finally, the 3 proteins of the GABARAP family were mixed in equal amounts with one equivalent of the NBR1 LIR2 peptide. The resulting ESI-MS spectrum is shown in Figure 4.15.

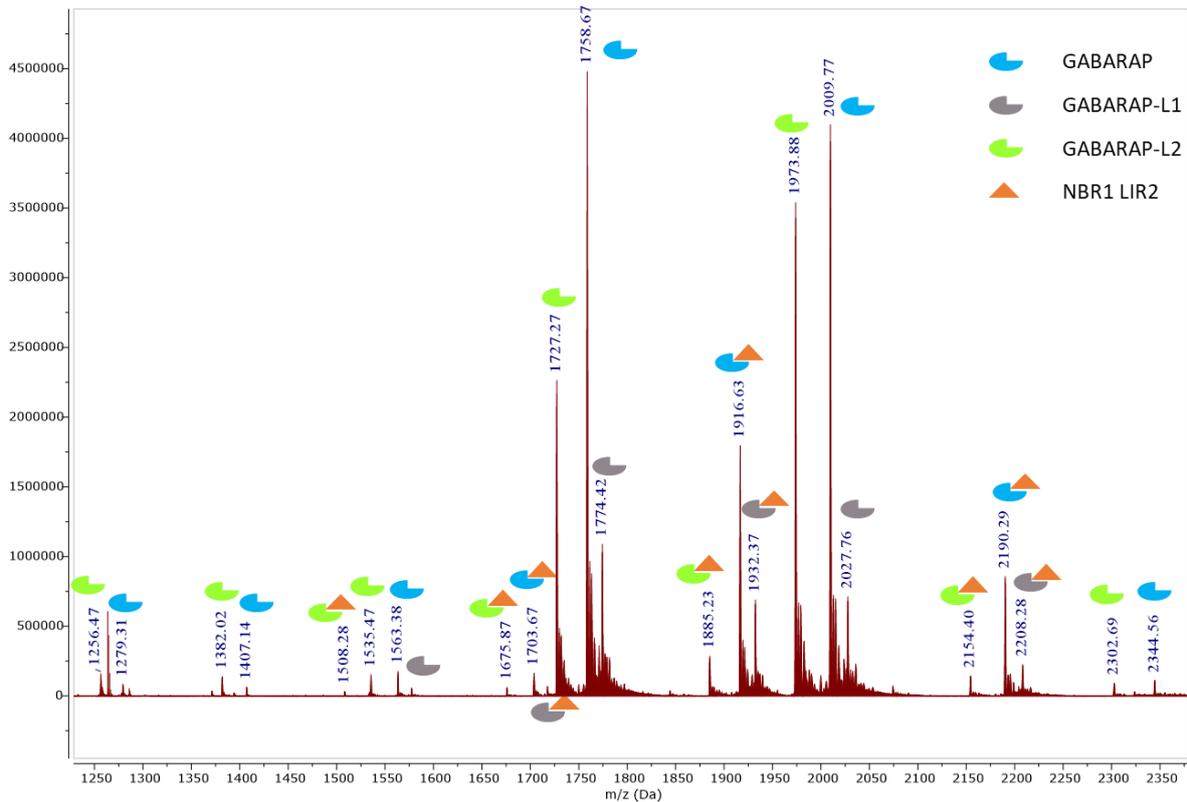


Figure 4.15: MS competition experiment for hATG8 GABARAP family NBR1 LIR2 peptide. Sample contained GABARAP, GABARAPL1, GABARAPL2 and NBR1 LIR2 peptide in equal concentrations (10 μ M). Key indicates the peaks that belong to each protein and protein-LIR complex. Spectrum expanded to show region where the main protein and complex peaks are found.

The MS spectrum was analysed to calculate the “bound to unbound” ratios for each GABARAP protein as with the previous MS competition experiments. This showed that there was a selectivity for the NBR1 LIR2 peptide to bind to GABARAPL1, with almost double the ratio of bound to unbound protein when compared to GABARAP (similar to LIR1). However, the GABARAP-LIR complex was much more favoured than the GABARAPL2-LIR complex, where the ratio of GABARAP complex was over 3 times more prevalent than the complex with GABARAPL2 (Figure 4.16).

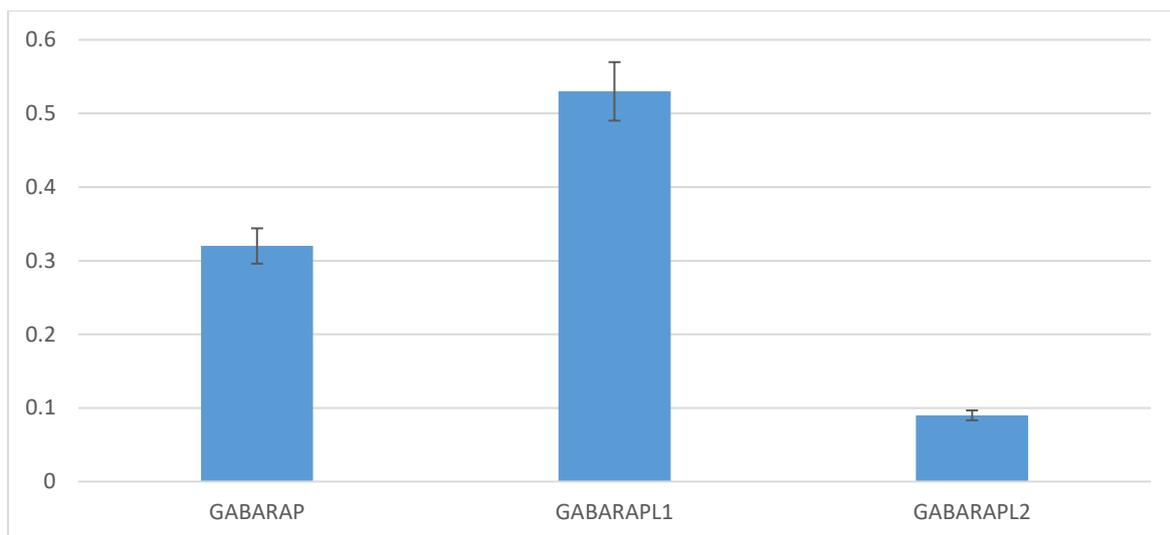


Figure 4.16: Bar chart of calculated values for ratio of bound to unbound for GABARAP subfamily with LIR2 MS competition experiment.

This is different from the LC3 family, where the ranking of binding preference between the 3 proteins was shown for both NBR1 LIRs, this time there was a clear difference in binding preference for GABARAPL2 between the LIRs.

Unfortunately due to issues with samples it was not possible to carry out the MS competition experiments whilst varying the number of equivalents of NBR1 LIR2 as in Chapter 4.2. However, these initial experiments gave an indication of which protein the two LIRs of NBR1 preferentially bind to.

4.4 Discussion

The MS experiments carried out in this chapter were used to probe the interaction of the LC3B mutants and hATG8 proteins with the NBR1 LIR peptides. The initial experiments confirmed that all the proteins that had been produced (LC3B mutants and hATG8s), that all bound to the two peptides used to represent the LIRs of NBR1. Whilst it had previously been shown that this was the case for LC3B, it was important to confirm the same interaction was seen for all of the proteins produced, before further work was carried out.

With the binding interactions confirmed between the LIR peptides and the proteins, the next step was to carry out MS competition experiments, to give some information of the binding preferences of the LIRs of NBR1 with the various proteins. These experiments showed some clear evidence of LIRs showing a particular selectivity for certain proteins, which when further analysed can shed more light on the binding action of LIRs with the hATG8 proteins.

4.4.1 LC3B Mutants Binding to the LIRs of NBR1

The two LC3B mutants were produced to probe the binding interaction of the LIRs of NBR1, and to help understand why there were two LIRs on NBR1. As can be seen from Figure 4.17 the LC3B KL mutant was designed to disrupt the binding around the two hydrophobic pockets. The LC3B T29D mutant was then used to see if phosphorylation on the large alpha helix, which sits against the central beta sheet to form the hydrophobic pockets, would alter the binding of the LIRs.

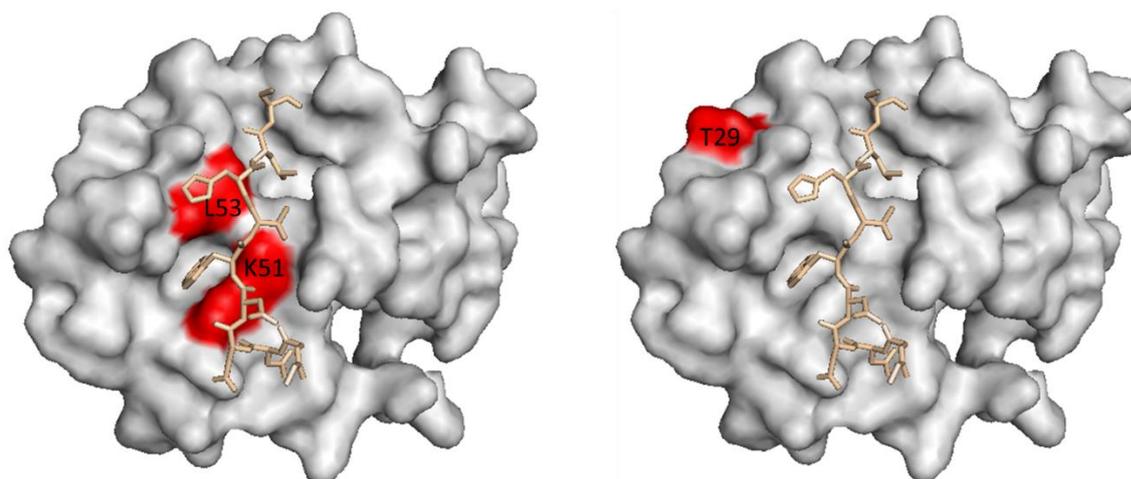


Figure 4.17: Surface representations of LC3B mutants with p62 LIR shown bound. LC3B KL is shown on the left and LC3B T29D on the right. Mutation is shown in red.

The MS completion experiment for LIR1 with the LC3B mutants showed that the KL (Lys51/Leu53) mutant had the desired effect in disrupting the interaction with LIR1. It was by far the least preferred of binding partners, with very little LC3B KL-LIR1 complex present, until higher concentrations of LIR1 were reached. Given where the mutation is on LC3B (Figure 4.17) around the hydrophobic pocket, it is possible to see that the reduced binding affinity could be explained in several ways. Firstly as K51 and L53 surround a large amount of the hydrophobic pocket any change would alter how the aromatic on the LIR would fit in the pocket, potentially making the interaction of HP1 with Tyr732 less favourable. The mutation of a lysine for alanine at position 51 would also make this area less basic, and hence destabilise any interaction with the acidic residues (Asp731 and Glu730) that precede Tyr732 in the LIR sequence.

The T29D mutant on the other hand shows little difference in affinity from LC3B WT, and it is only at higher concentrations of LIR1 that there is potentially a higher affinity of the LIR1 for T29D over WT. However, this is a single point and to confirm whether this was an anomaly or the start point in a trend would require the competition experiment to be run to higher equivalents of LIR1.

Whilst the KL mutant had the desired effect of disrupting the interaction with LIR1 the opposite seems true with LIR2. The MS competition experiments of the mutants of LC3B with LIR2 in fact show that the KL mutant is the preferred binding partner with LIR2. Whilst the less basic face of LC3B KL destabilise the interaction with NBR1 LIR1 it has the opposite effect with LIR2, which contains no acidic amino acids preceding Phe563. Hence LIR2 binds

selectively to the KL mutant over LC3B WT and T29D with their more basic faces. This also points to the KL mutant not actually altering HP1 in such a way as to significantly decrease the interaction of the aromatic residue on the LIR with the pocket, as if this was the case and the aromatic residue was significantly impeded in binding with the hydrophobic pocket, you would not expect it to have such a high affinity for LIR2.

Whilst the competition experiments are interesting and shed some light on selectivity between the two LIRs of NBR1, they do not give any information on how the LIR is binding to the proteins. To do this NMR titration experiments will be carried out in the next chapters to probe the LIR-LC3B interactions.

4.4.2 hATG8 Proteins Binding to LIRs of NBR1

The competition experiments for the hATG8 family proteins showed that the LIRs of NBR1 had a degree of selectivity for particular proteins. For the LC3 sub family LIR1 bound preferentially to LC3A, with almost double the ratio of complex to free protein of LC3B. The other interesting result was that there was almost no LC3C and LIR1 complex. This result was then repeated for LIR2 which preferentially bound to LC3A roughly twice as much as LC3B, and then there was no LC3C-LIR2 complex. Whilst it was not possible to run these experiments at increased concentrations of LIR to observe trends, it does show that both LIRs have a slight selectivity for LC3A, and that the interaction with LC3A and LC3B is considerably more favoured than that with LC3C.

As has been mentioned previously the LC3s are structurally similar and there is high sequence alignment between LC3A and LC3B at 82.5%. However there is much less sequence alignment of LC3C with LC3A (59.2%) and LC3B (55%). Looking at the sequences of the three LC3 proteins (Figure 4.18) there are 10 amino acids which are different in all three LC3s, three unique to LC3A, eight which are unique to LC3B and 36 which are unique to LC3C. We can see that for LC3A and LC3B the differences are all mainly located at the N-terminus. However, the changes in sequence between LC3C and LC3A and LC3B are spread throughout the sequence.

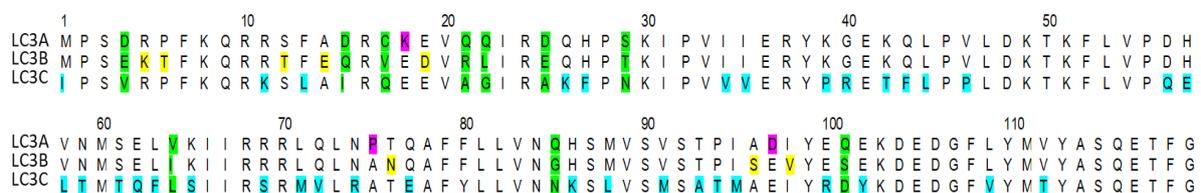


Figure 4.18: Sequences of LC3 proteins. Parts of the sequence where all three LC3s have different amino acids are shaded green and where LC3A, LC3B and LC3C are different are shaded pink, yellow and blue respectively.

To visualise where on the LC3s these changes in amino acid are, and to see how they could affect the interaction with the LIRs, the differences were mapped to the surface of LC3B bound to the p62 LIR (Figure 4.19)

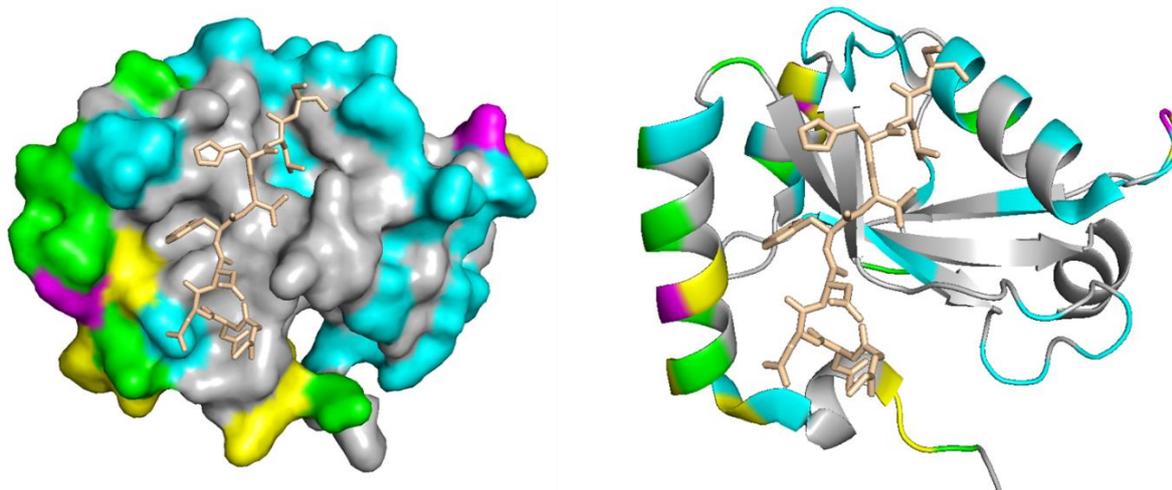


Figure 4.19: Structure of LC3B WT bound to p62 LIR coloured to show differences in sequences between LC3s. Parts of the sequence where all three LC3s have different amino acids are shaded green and where LC3A, LC3B and LC3C are different are shaded pink, yellow and blue respectively.

From these mapped changes in sequence we can see the main differences in sequence on LC3A and LC3B appear on the first two alpha helices, which pack against the β -sheet that forms HP1, and hence these changes could lead to the preferential binding for LC3A by the two LIRs. For LC3C the changes in sequence are spread all over the protein, but a number are located at the top of the binding surface around HP2. LC3C has been shown to bind to non-standard LIR motifs [80], which do not have an aromatic residue (ILVV), and so rely on a valine binding into HP2. This lack of aromatic residue is then compensated for by other hydrophobic contacts with LC3C. These changes on LC3C, whilst enabling the binding of an atypical LIR, does then alter how LC3C will interact with standard LIRs. So whilst the LIRs of NBR1 will bind to LC3C, when in competition with LC3A and LC3B there is very little binding to LC3C.

For the GABARAP sub family both NBR1 LIR1 and LIR2 bind preferentially to GABARAPL1. Then for LIR1, GABARAP and GABARAPL2 have similar affinities, whereas for LIR2 GABARAP is then bound preferentially over GABARAPL2. As with the LC3 family, the structures of the three GABARAP are extremely similar and GABARAP and GABARAPL2 have high sequence alignment at 87.9%. However, the sequence alignment of GABARAPL2 with GABARAP and GABARAPL1 at 57.8% and 61.2% respectively are considerably less. The sequences for the three GABARAPs are shown in Figure 4.20 with changes in sequence coloured.

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--------|---|---|----|---|----|---|----|---|----|---|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 10 | | 20 | | 30 | | 40 | | 50 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| GAB | M | K | F | Y | K | E | H | P | F | E | R | R | S | E | G | E | K | I | R | K | K | Y | P | D | R | V | P | V | I | V | E | K | A | P | K | A | R | I | G | D | L | D | K | K | Y | L | V | P | S | D | | | | |
| GAB-L1 | M | K | F | C | Y | K | E | D | H | P | F | E | R | R | S | E | G | E | K | I | R | K | K | Y | P | D | R | V | P | V | I | V | E | K | A | P | K | A | R | V | G | D | L | D | K | R | K | Y | L | V | P | S | D | |
| GAB-L2 | M | K | W | M | F | K | E | D | H | S | L | E | H | R | C | V | E | S | A | K | I | R | A | K | Y | P | D | R | V | P | V | I | V | E | K | V | S | G | S | Q | I | V | D | I | D | K | R | K | Y | L | V | P | S | D |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--------|----|---|----|---|----|---|----|---|-----|---|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 60 | | 70 | | 80 | | 90 | | 100 | | 110 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| GAB | L | T | V | G | Q | F | Y | F | L | I | R | K | R | I | H | L | R | A | E | D | A | L | F | F | F | V | N | N | V | I | P | P | T | S | A | T | M | G | Q | L | Y | P | H | H | E | E | D | F | F | L | Y | I | A | Y | S | D | E | S | V | Y | G | |
| GAB-L1 | L | T | V | G | Q | F | Y | F | L | I | R | K | R | I | H | L | R | P | E | D | A | L | F | F | F | V | N | N | T | I | P | P | T | S | A | T | M | G | Q | L | Y | E | D | N | H | E | E | D | F | F | L | Y | V | A | Y | S | D | E | S | V | Y | G |
| GAB-L2 | I | T | V | A | Q | F | W | I | I | R | K | R | I | Q | L | P | S | E | K | A | I | F | L | F | V | D | K | T | V | P | Q | S | L | T | M | G | Q | L | Y | E | K | E | K | D | E | D | G | F | L | Y | V | A | Y | S | G | E | N | T | F | G | | |

Figure 4.20: Sequences of GABARAP proteins. Parts of the sequence where all three GABRAPs have different amino acids are shaded green and where GABARAP, GABARAPL1 and GABARAPL2 are different are shaded pink, yellow and blue respectively.

There are eight amino acids that are different for each GABARAP, three that are unique for GABARAP, one is unique for GABARAPL1 and 34 unique for GABARAPL2. As with LC3A and LC3B, the majority of changes in sequence for GABARAP and GABARAPL1 appear at the N-terminus (although not as many), with the variations in sequence for GABARAPL2 being spread throughout the protein. To better understand how these changes might affect the binding with the LIRs, the structure of GABARAPL1 bound to NBR1 LIR1 was coloured where the sequences were different between the three proteins (Figure 4.21).

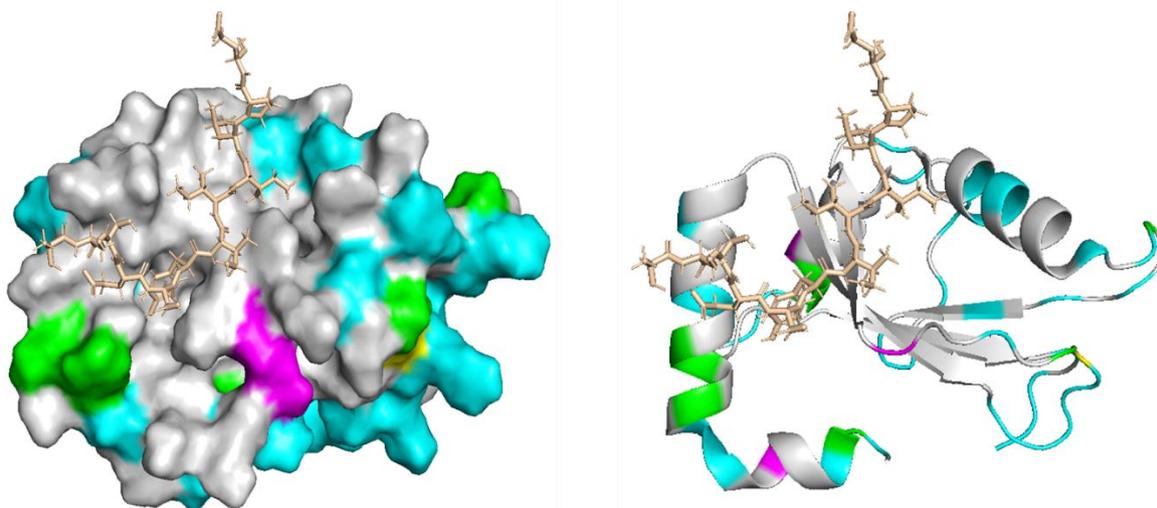


Figure 4.21: Surface of GABARAPL1 bound to NBR1 LIR1 coloured to show where changes in sequence between the GABARAPs occur. Parts of the sequence where all three GABARAPs have different amino acids are shaded green and where GABARAP, GABARAPL1 and GABARAPL2 are different are shaded pink, yellow and blue respectively.

Having visualised where the differences in the GABARAP subfamily are in relation to where the LIR binds, it is possible that the selectivity could be due to the changes in amino acid at positions 13 and 16, which would potentially interact with the amino acids preceding the aromatic on the LIR. The amino acid on the LIR that is closest to this region on the GABARAPs is X₋₂, which for LIR1 is Glu730, and for LIR2 this is Leu561. These are different types of amino acids (acidic vs hydrophobic), and so would interact with the surfaces of the GABARAPs in different ways. GABARAPL2 also has a number of changes in sequence around HP2, which could alter how the hydrophobic residues Ile735 (LIR1) and Leu566 (LIR2) binds, and lead to the interaction being lower affinity than with the other GABARAPs. The small difference in size between the LIR containing a Leucine or Isoleucine could then account for the change in selectivity between the LIRs. However, why GABARAPL1 is preferentially bound for both sets of competition experiments is not clear.

4.5 Summary

This chapter has described the MS experiments used to investigate the binding of the NBR1 LIRs with the LC3B mutants and ATG8 family proteins. In summary the following observations were made

- NBR1 LIR peptides bind to LC3B mutants and ATG8 proteins
- MS competition experiments with the LC3B mutants with LIR1 show that LIR1 binds preferentially to LC3B WT and T29D. The KL mutant with its modified hydrophobic pocket and less basic face is the least preferred of the binding partners
- Competition experiments with the LC3B mutants and LIR2 show the opposite result to LIR1 with LC3B KL now being the preferred binding partner. This is potentially driven by the less basic face of LC3B KL being a better binding partner as LIR2 contains no acidic residues preceding the aromatic in the LIR sequence. Potentially, the K51A and L53A mutations provide additional hydrophobic contacts and reduce steric interactions.
- LIR1 and LIR2 bind with some selectivity to LC3A over LC3B with almost no LC3C-LIR complex detected.
- For the GABARAP sub family NBR1 LIR1 binds preferentially to GABARAPL1 with GABARAP and GABARAPL2 showing similar affinity.
- NBR1 LIR2 binds preferentially to GABARAPL1 with almost double the amount of complex than with GABARAP. GABARAP-LIR complex is then almost 3 times more prevalent than for GABARAPL2.

Chapter 5

5. NMR Assignment of WT LC3B and Studies of its Interactions with LIRs of P62 and NBR1

This chapter will investigate the interactions of LC3B WT with the two LIRs of NBR1 using NMR titrations, and then compare them with the interaction of LC3B WT with the p62 LIR.

The interaction of wild type LC3B and the LIR of p62 has been well studied previously. So comparing this with the interaction of LC3B with the two LIRs of NBR1, was considered a good starting point for understanding the interactions of the LIRs.

First the backbone amides in the ^1H - ^{15}N HSQC spectrum of wild type LC3B will be assigned. This assignment will then be used to track the peak perturbations in the NMR titrations of the three LIRs. This data will then be analysed to focus on the amino acids that experience the most significant changes in LC3B WT when binding with the LIRs. These results will then be used to generate structures of the LIRs binding with LC3B WT using HADDOCK webserver (Chapter 2.5.2.7). Finally the titration data and HADDOCK structures will be compared between the three LIRs, to look at the effect the differences in sequences of the LIRs have upon the binding of LC3B WT.

The LIR of p62 contains the canonical LIR motif with a WxxL motif, and whilst the LIRs of NBR1 follow the sequence of an aromatic and hydrophobic amino acid separated by two others, there are significant differences (Figure 5.1).

| | | | | | | | | | | | | | |
|-----------|---|---|-----------------|-----------------|-----------------|-----------------|------------------|----------------|----------------|------------------|----------------|----------------|---|
| P62 | S | G | G | D | D | D | W ₃₃₈ | T | H | L ₃₄₁ | S | S | K |
| NBR1 LIR1 | S | A | S | S | E | D | Y ₇₃₂ | I | I | I ₇₃₅ | L | P | E |
| NBR1 LIR2 | A | Q | D | L | L | S | F ₅₆₃ | E | L | L ₅₆₆ | D | I | N |
| | | | x ₋₄ | x ₋₃ | x ₋₂ | x ₋₁ | Θ | x ₁ | x ₂ | Γ | x ₃ | x ₄ | |

Figure 5.1: Sequences for LIRs of p62 and NBR1. Amino acids that bind to the hydrophobic pockets of LC3B are shown in red and acidic residues that directly precede the aromatic residue are shown in green. Amino acids that are greyed out were not included in peptide sequence used but are included for reference.

Whilst all 3 LIR sequences have aromatic residues that bind to the first hydrophobic pocket on LC3B, they all use a different aromatic residue, with p62 containing a tryptophan, and NBR1 LIR1 and LIR2 containing a tyrosine and phenylalanine respectively. This difference in aromatic residue at the Θ position will alter how the LIR interacts with LC3B WT, due to steric and electronic interactions. Previous work has shown that the binding is strongest when a tryptophan is present, and the p62 LIR has been shown to be a stronger binder than that of the NBR1 LIR1, which supports this data [113]. There is also some variation of the hydrophobic residue, with two LIRs (p62 and LIR2) containing leucine and NBR1 LIR1

containing isoleucine. This hydrophobic residue binds into the second hydrophobic pocket on LC3B, and the small difference between isoleucine and leucine might have a small steric effect. Finally the LIR of p62 contains three acid residues, and NBR1 LIR1 contains two, in the positions X₋₃ to X₋₁ preceding the aromatic residue. NBR1 LIR2 does contain one acidic residue, but this is further away from the aromatic residue in position X₋₄. However, NBR1 LIR1 contains serines in positions X₋₃ and X₋₄, and LIR2 contains one at position X₋₁, which although not classed as an acidic residue, still contains a negative charge on the oxygen of the amino acid side chain, which could interact with the positively charged face of LC3B. Previous studies have shown that LIRs with acidic residues have a slightly stronger binding interaction with LC3B, and the sequence analysis of known LIRs does show that there is a preference for acidic residues in positions X₋₁ to X₋₃, but after aspartic and glutamic acid, serine is then the most common amino acid in these positions. The serines at these positions can potentially be phosphorylated, altering the binding affinity of the LIR with LC3B.

To better understand how all these differences affect the interaction of LIRs with LC3B, ¹H-¹⁵N HSQC NMR titrations would be used. Whilst other methods can confirm there is an interaction and give values for binding coefficients, NMR titrations will also give information on which amino acids on LC3B are involved with the interaction, and allow mapping of the binding site.

5.1 Backbone Assignment of LC3B

To facilitate the NMR titrations of LC3B with the 3 LIRs, an assignment of the backbone amides in the ¹H-¹⁵N HSQC for LC3B would be required. There is a published assignment of a construct of LC3B for the first 120 amino acids (aa) [139]. However, as the NMR titrations would be carried out with the full length LC3B WT (125aa), an assignment would be carried out on this form of LC3B so as to make sure the assignment of the backbone amides in the ¹H-¹⁵N HSQC for the 125 amino acid LC3B is correct, and provide a firm basis for further titration studies.

To assign LC3B, a series of 2D and 3D heteronuclear NMR experiments were collected on a sample of LC3B WT uniformly labelled with ¹⁵N and ¹³C. The NMR experiments used were ¹H-¹⁵N HSQC, HNC0, HN(CA)CO, HNCA, HN(CO)CA, HNCACB and HN(CO)CACB (Chapter 2.5.2.2). This set of 3D experiments allowed assignment of the backbone amides in the ¹H-¹⁵N HSQC using the methods in Chapter 2.5.2.4. An example linking of the C_α and C_β peaks for the LC3B WT sample is given in Figure 5.2.

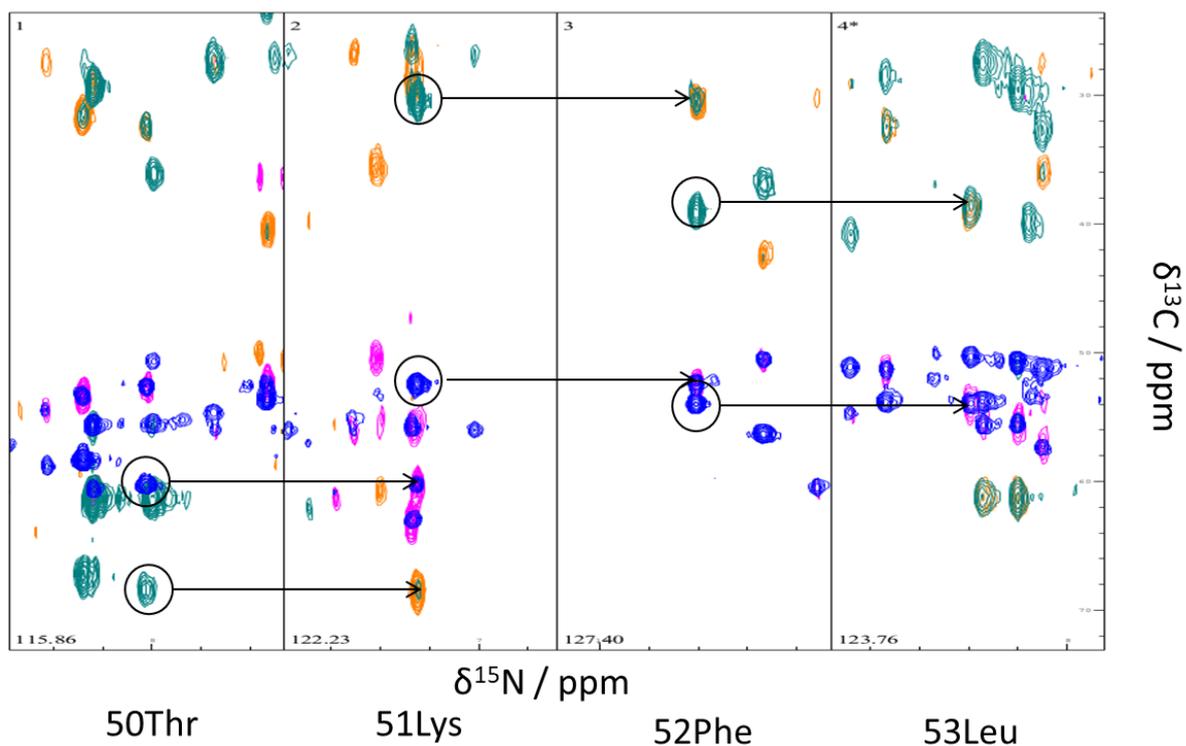


Figure 5.2: Example of linking C_{α} and C_{β} peaks using HNCACB (green and blue peaks) and HN(CO)CACB (pink and orange peaks).

Following this process of linking peaks and using characteristic chemical shifts to assign spin systems to a type of amino acid, it was possible to fully assign the ^1H - ^{15}N HSQC of LC3B (Figure 5.3)

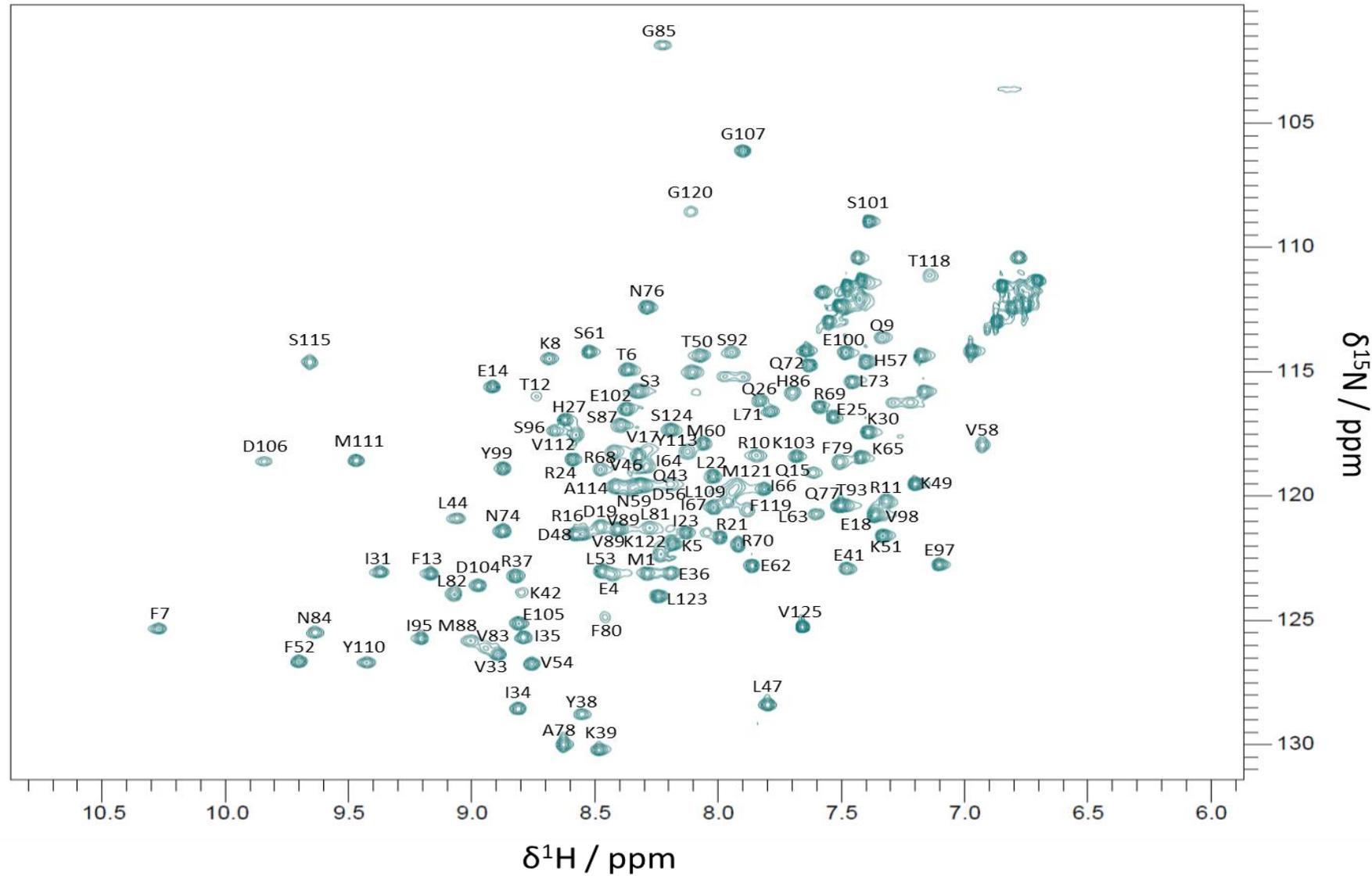


Figure 5.3: LC3B WT ^1H - ^{15}N HSQC backbone assignment.

5.2 Interaction of WT LC3B with the LIR1 of NBR1

With the backbone amide peaks in the ^1H - ^{15}N HSQC for LC3B assigned, it was then possible to evaluate that data for titrations of LC3B WT with NBR1 LIR1 that had been carried out previously. The same peptide representing NBR1 LIR1, used in the MS experiments, had been used for the NMR titrations (residues 727-738 in NBR1, ASSEDIILLPE). This was titrated against LC3B until a ratio of 2:1 peptide to protein was reached, acquiring ^1H - ^{15}N HSQC at each point (Chapter 2.5.2.1). An overlay of the start and endpoint of the titration of LC3B WT with NBR1 LIR1 is given in Figure 5.4

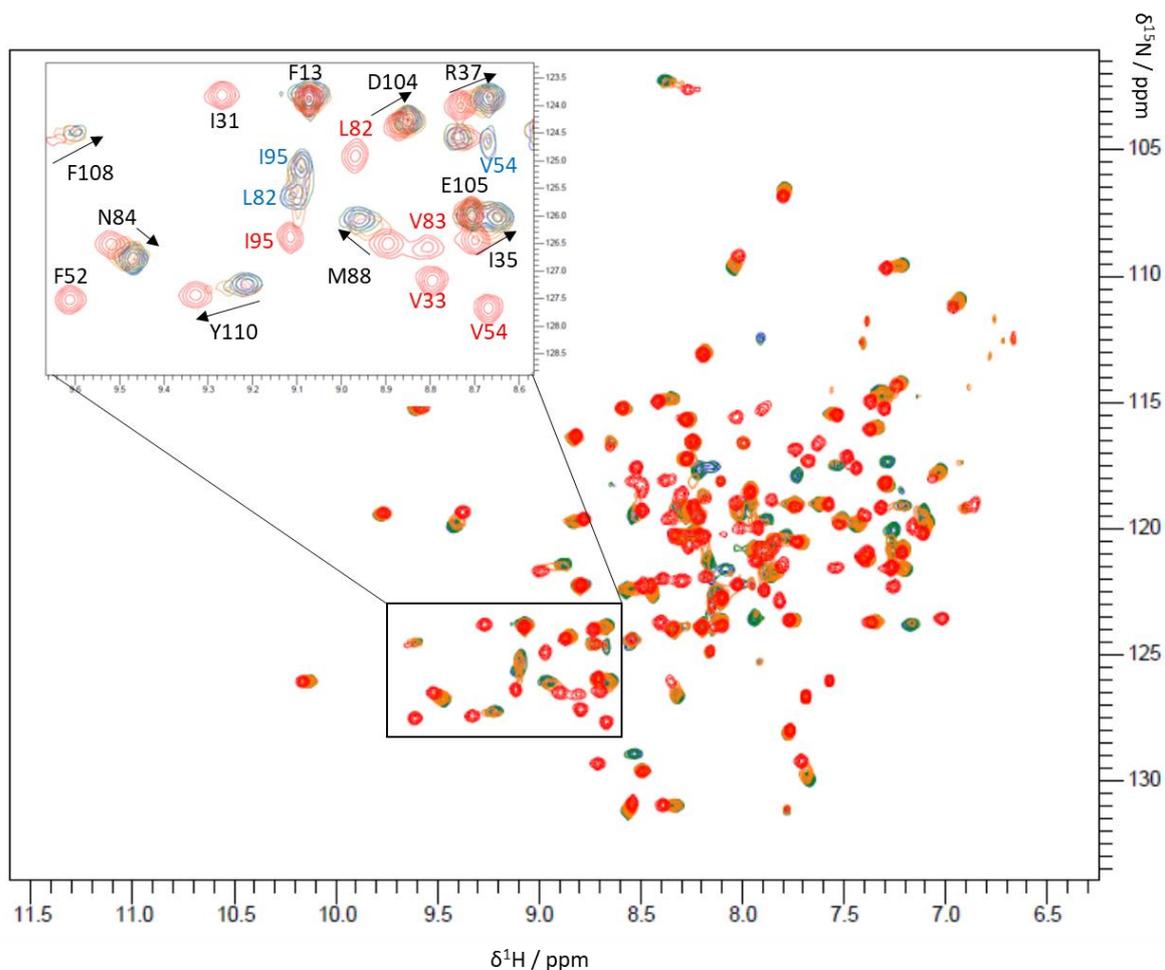


Figure 5.4: Overlay of a selection of ^1H - ^{15}N HSQC spectrum for titration of LC3B WT and NBR1 LIR1 peptide. Spectra shown are the ratios of LC3B WT to LIR 1:0 (red), 1:0.5 (orange), 1:1 (green) and 1:2 (blue).

As the LIR peptide was titrated with the LC3B, peaks within the ^1H - ^{15}N HSQC began to move until by the end of the titration a large number of peaks had moved to a new position in the spectrum. Some of these shifts were relatively small, but some peaks had disappeared and reappeared somewhere entirely different in the spectrum. How the peaks moved depends on the exchange regime the particular amino acid is undergoing and a description of these exchange regimes is given in Chapter 2.5.25.

Understanding the nature of the binding interaction between LC3B and NBR1 LIR1 relies on following the peaks in the titration, and if the peaks undergo intermediate or slow exchange this can be extremely difficult. The titration of LC3B with NBR1 LIR1 had a large number of peaks (>20) in slow and intermediate exchange, where it was not possible to follow the peaks in the titration. To get the clearest picture possible of the binding of the NBR1 LIR1 peptide to LC3B, a backbone amide assignment on a ^{13}C , ^{15}N labelled sample of LC3B with unlabelled peptide was conducted.

As with the unbound LC3B a ^{15}N and ^{13}C enriched LC3B WT protein sample was produced, mixed with 2 equivalents of NBR1 LIR1 peptide and a series of 3D NMR experiments was run. The same set of HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB and HN(CO)CACB experiments were run, and the carbonyl, CA and CB peaks linked to the NH peak in the ^1H - ^{15}N HSQC. These peaks were then linked using the method in Chapter 2.5.2.4.

It was then possible to almost completely assign the backbone amides in the ^1H - ^{15}N HSQC for NBR1 LIR1 peptide bound to LC3B, with all but 3 amino acids assigned (R24, H27 and V91, Figure 5.5)

With the ^1H - ^{15}N HSQC of LC3B bound to NBR1 LIR1 assigned it was then possible to know the end point for the peaks. The chemical shift perturbations were then calculated using the equation given in Chapter 2.5.2.4 (equation 2.1). As a large number of peaks had moved in the ^1H - ^{15}N HSQC, it was important to analyse the CSPs so as to focus on those amino acids that underwent a significant change when binding to the NBR1 LIR1 peptide. The significance was analysed using the method in Chapter 2.5.2.4, and these cut offs and the CSP values calculated were plotted (Figure 5.6).

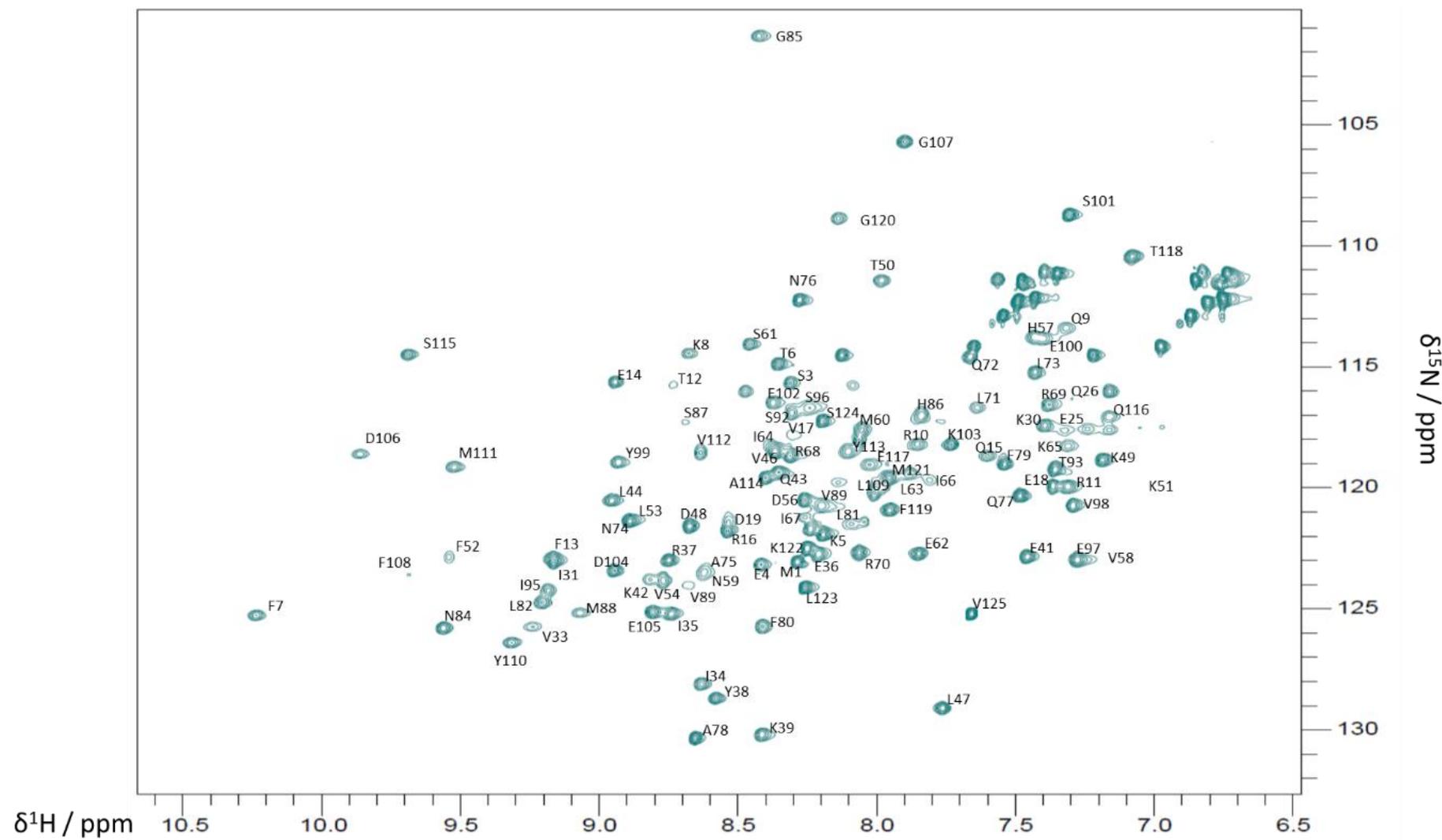


Figure 5.5: ^1H - ^{15}N HSQC backbone assignment of NRB1 LIR1 peptide bound to LC3B WT.

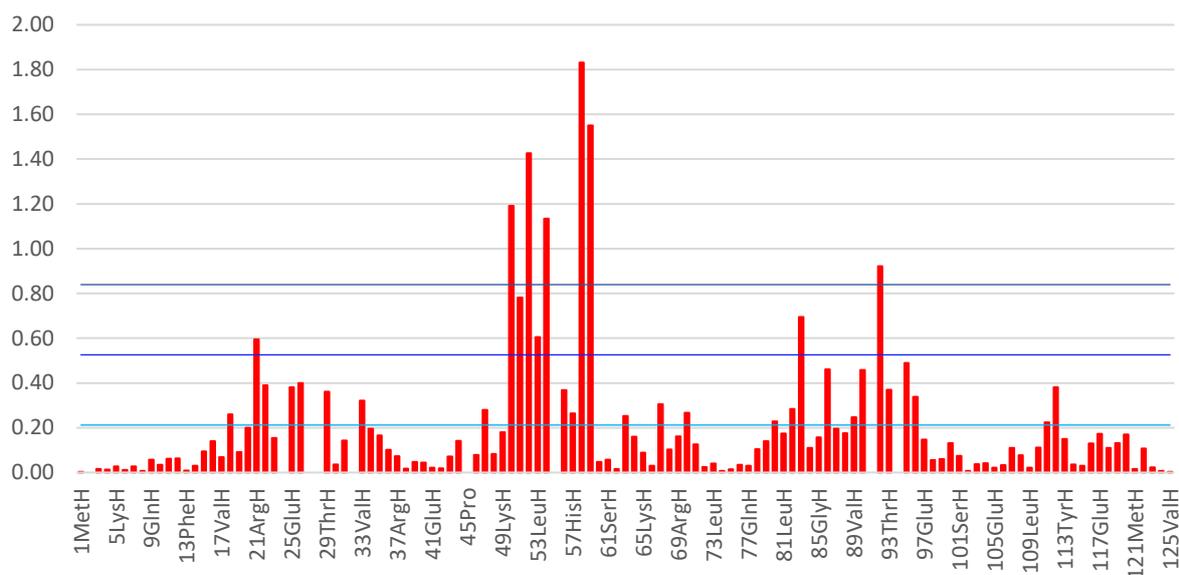


Figure 5.6: Graph of CSP values for binding of NBR1 LIR1 peptide to LC3B. The light blue line indicates the mean, the blue line is the mean and one standard deviation and the dark blue line the mean and two standard deviations.

In total 13 peaks had significant change to the chemical shift (>0.53 ppm), and of those 9 had shifts greater than 0.84 ppm (or could not be assigned in the bound data), indicating substantial change to the structure of the protein around that residue.

The residues that had significant perturbations and those that had moved more than the mean were then mapped to the surface of LC3B, to identify the structural perturbations from the binding of the NBR1 LIR1 peptide to LC3B (Figure 5.7).

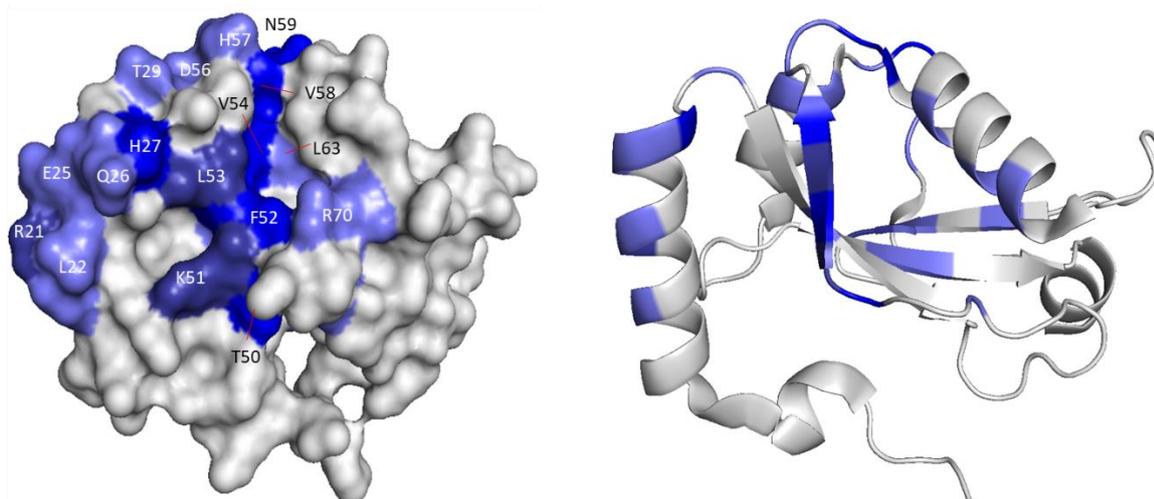


Figure 5.7: CSP mapping for the binding of NBR1 LIR1 peptide to LC3B. Residues are coloured purple if the CSP is greater than the mean, light blue if greater than the mean and one standard deviation and blue if above the mean and two standard deviations. Residues with no significant shift are coloured grey.

The large perturbations of the residues on the N-terminal α -helix (α 2), and those on the β -sheet of the ubiquitin like domain, are indicative of the NBR1 LIR1 peptide interacting with the two hydrophobic pockets of LC3B. H27, K51, F52 and L53 all have significant CSPs, and these amino acids surround HP1, indicating that Tyr732 on LIR1 is interacting here. There are then significant CSPs seen on V54 and V58, which are located around HP2, and is indicative of Ile735 interacting here.

To gain a better understanding of how the NBR1 LIR1 peptide binds to LC3B, the HADDOCK webserver was utilised following the method in Chapter 2.5.2.7. The HADDOCK run generated 170 structures distributed in 10 clusters. The top scoring cluster contained 66 structures with a HADDOCK score of -91.6 ± 4.4 . The structure produced by Haddock, with the significant amino acids coloured, is shown below (Figure 5.8). The next highest scoring cluster only scored -62.1 ± 0.9 , and contained only 32 structures, and so was not investigated further due to the lower score and fewer structures.

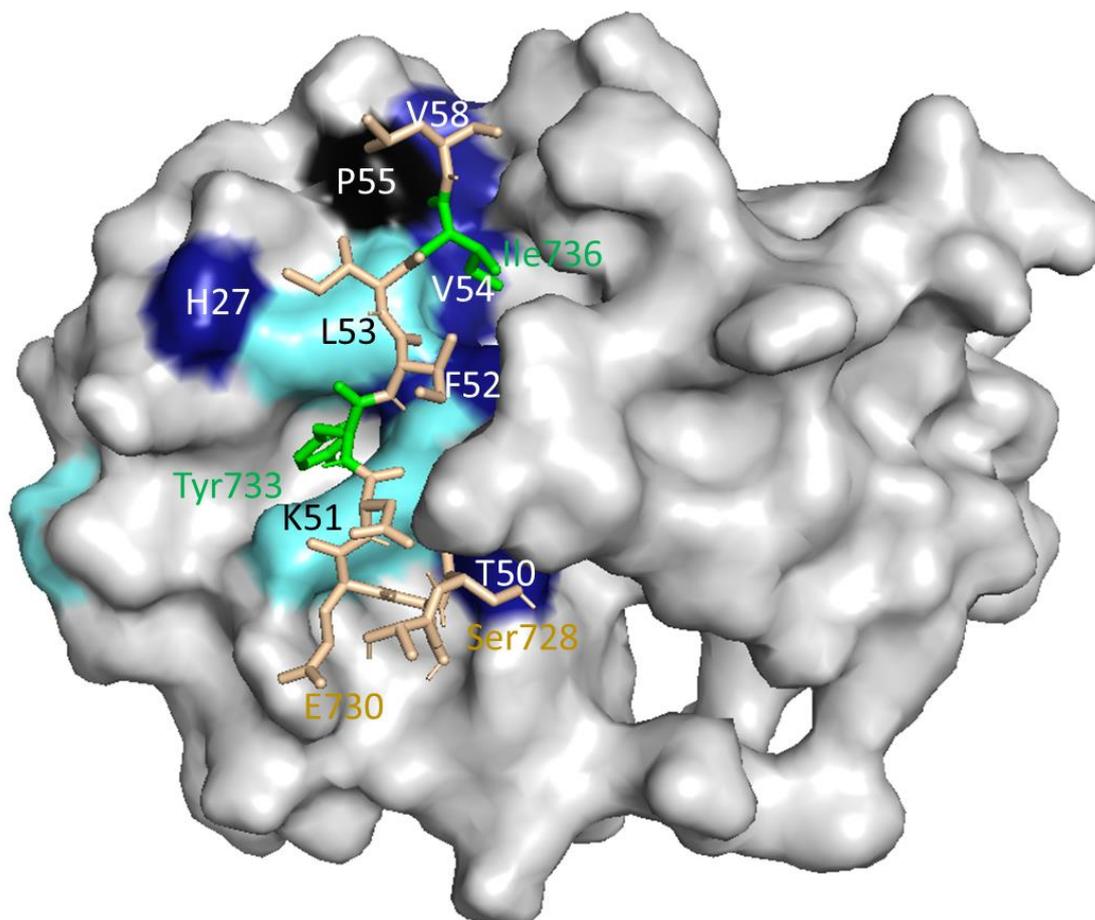


Figure 5.8: HADDOCK structure of NBR1 LIR1 peptide bound to LC3B. The peptide is coloured yellow with the active residues Tyr733 and Ile736 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation. P55 on LC3B WT is coloured black for reference.

This structure produced by HADDOCK shows a similar mode of binding to previous LIRs studied. With K51, F52 and L53 forming a hydrophobic pocket occupied by Tyr732, and Ile735 fitting into a second hydrophobic pocket containing the highly perturbed residues F52, V54 and V58. Acidic residues E730 and D731 on LIR1 then sit close to R11 and K49 on LC3B, respectively, with a likely interaction between the acidic residues on the LIR with the basic face of LC3B. Whilst the CSPs for R11 and K49 are not significant for the NMR titration, as CSPs are just an indication of how the environment for the amide of the amino acid changes, it does not rule out that there is not an interaction. There is also a large CSP on T50, which when reviewing the HADDOCK structure indicates an interaction with Ser728 on LIR1 (probably a hydrogen bond).

5.3 Interaction of WT LC3B with the LIR2 of NBR1

Having studied the interaction of the NBR1 LIR1 peptide with LC3B using NMR titrations, the same process was repeated for LIR2 with the titration data that had been previously acquired. For LIR2 of NBR1 a peptide with the sequence AQDLLSFELLD was used to represent residues 557-567 in NBR1. This had been titrated against LC3B until a ratio of 2:1 peptide to protein was reached acquiring ^1H - ^{15}N HSQC at each point (Figure 5.9).

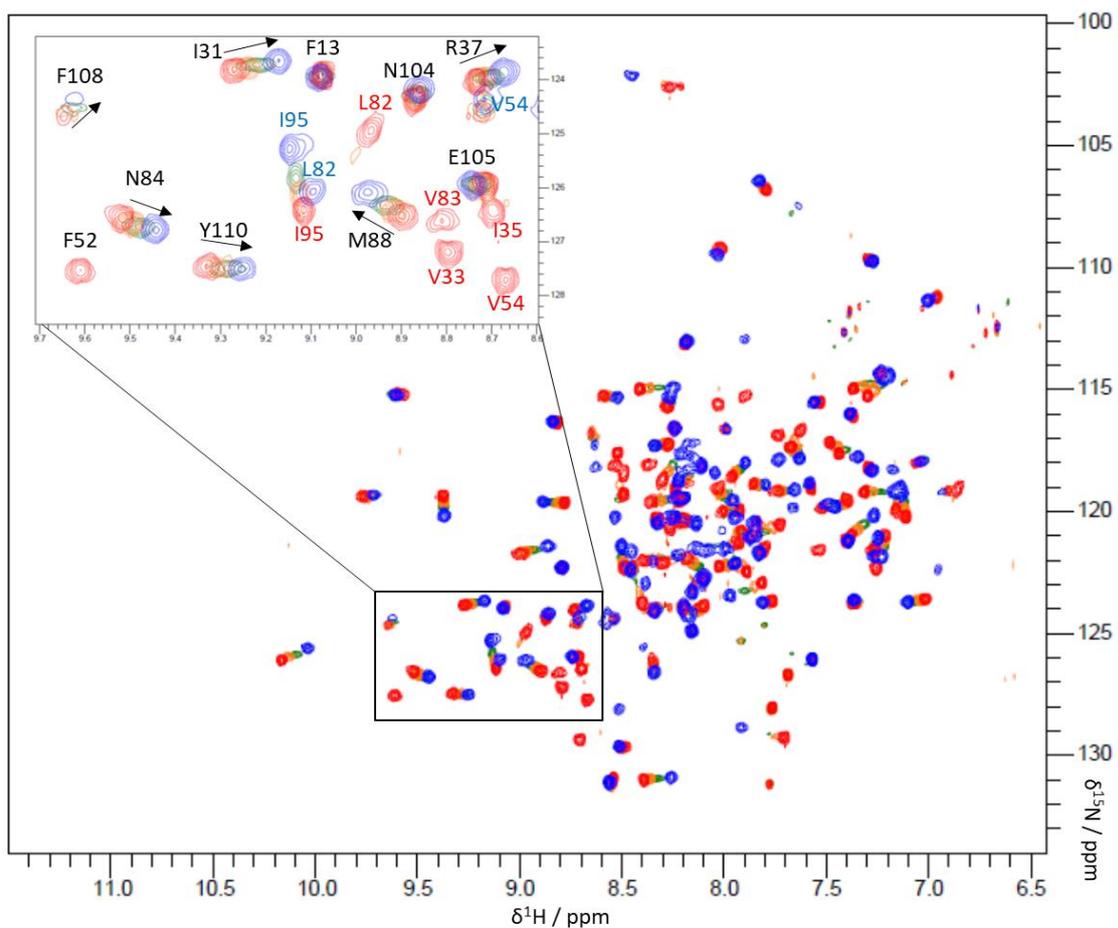


Figure 5.9: Overlay of a selection of ^1H - ^{15}N HSQC spectrum for titration of LC3B WT and NBR1 LIR2 peptide. Spectrum shown are the ratios of LC3B WT to LIR 1:0 (red), 1:0.5 (orange), 1:1 (green) and 1:2 (blue).

As for the titration of the NBR1 LIR1 peptide, the majority of peaks in the ^1H - ^{15}N HSQC spectrum moved, and as before there was a mixture of slow, medium and fast exchange. However, this time it was possible using the already assigned LC3B and NBR1 LIR1 peptide spectrum, to aid assignment of the slow and medium exchanging peaks in the NBR1 LIR2 peptide spectrum. Although not all the peaks in the LIR2 spectrum moved in exactly the same way, there was enough similarity to allow assignment by comparison with the LIR1 data.

Using this method all but 3 peaks were assigned (R24, H27 and V91). CSPs were then calculated for each peak and plotted (Figure 5.10).

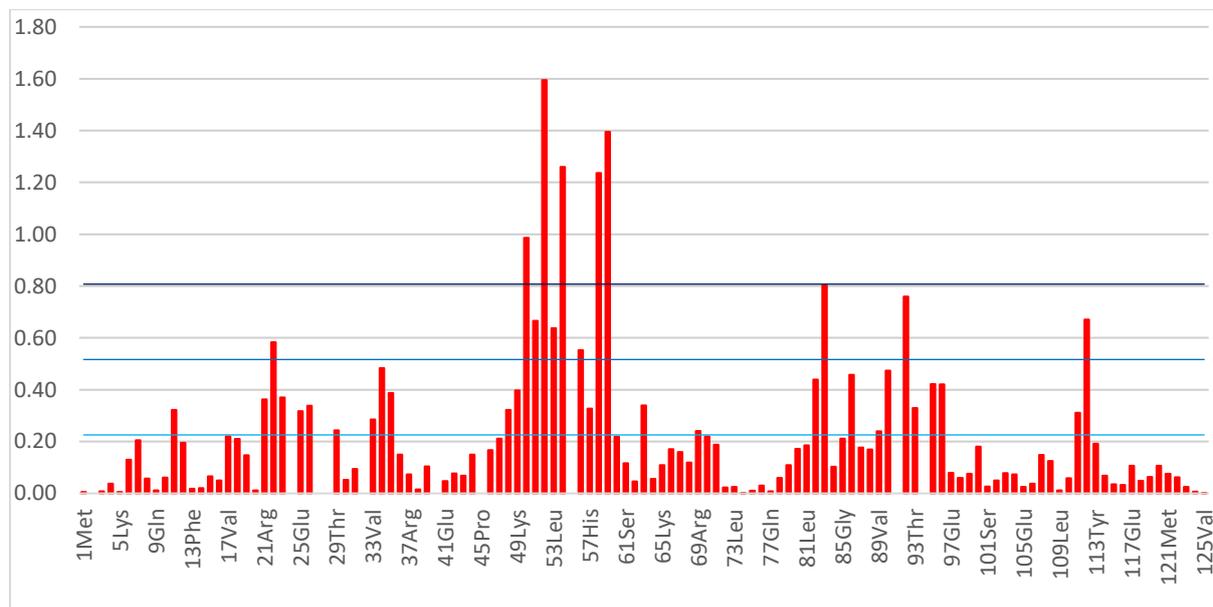


Figure 5.10: Graph of CSP values for binding of NBR1 LIR2 peptide to LC3B. The light blue line indicates the mean, the blue line is the mean and one standard deviation and the dark blue line the mean and two standard deviations.

The same method as used for NBR1 LIR1 peptide for determining which peaks had moved significantly was used (Chapter 2.5.2.6). This gave 17 peaks that had significant changes in the ^1H - ^{15}N HSQC spectrum (> 0.52 ppm), and 8 of these were above the higher significance limit of 0.84 ppm (or could not be followed in the titration). This data was then used to colour the surface of LC3B to better visualise the binding site (Figure 5.11).

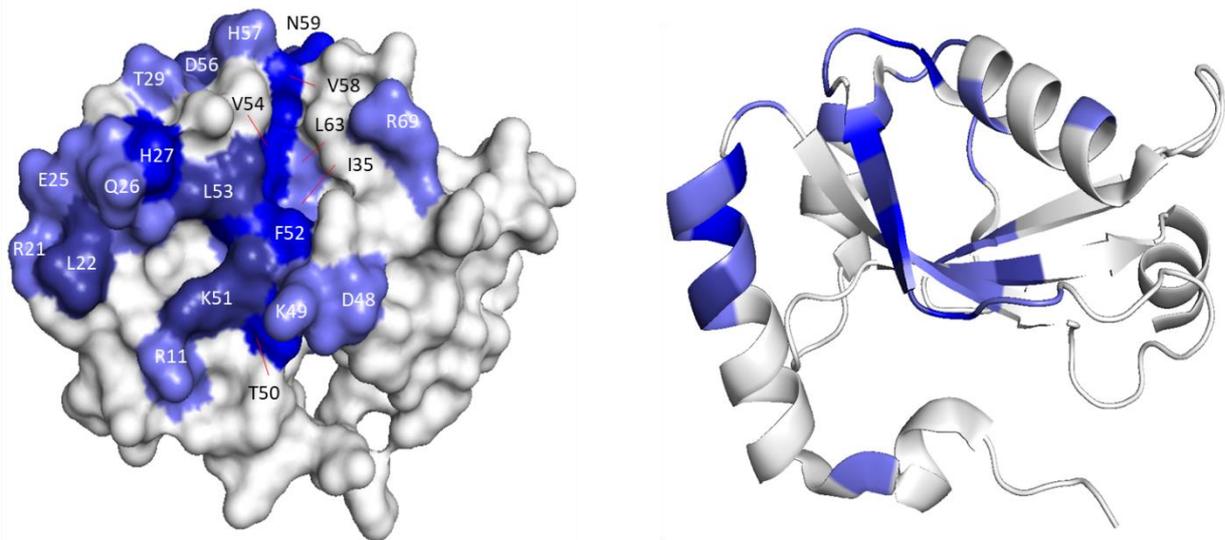


Figure 5.11: CSP mapping for the binding of NBR1 LIR2 peptide to LC3B. Residues are coloured purple if the CSP is greater than the mean, light blue if greater than the mean and one standard deviation and blue if above the mean and two standard deviations. Residues with no significant shift are coloured grey.

The CSPs were then used to act as restraints in HADDOCK webserver, which generated 167 structures in 11 clusters. The top scoring cluster contained 49 structures, with the top structure HADDOCK scoring -102.5 ± 2.9 (Figure 5.12).

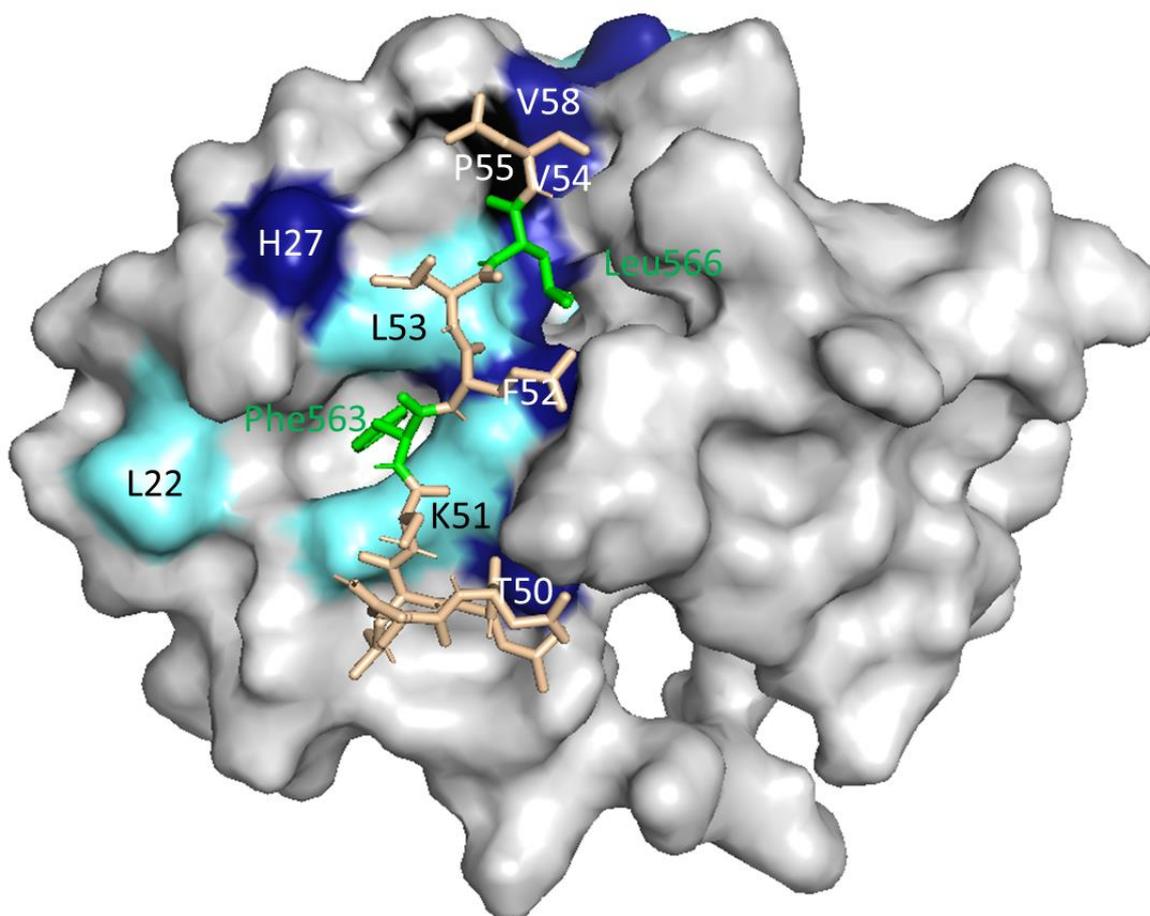


Figure 5.12: HADDOCK structure of NBR1 LIR2 peptide bound to LC3B. The peptide is coloured yellow with the active residues Phe556 and Leu566 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation. P55 on LC3B WT is coloured black for reference.

This structure shows again a similar binding mode to previously studied LIRs. Phe563 sits in the first hydrophobic pocket, with large perturbations on K51 and L53 that form part of the pocket. Leu 566 on LIR2 then sits in the second hydrophobic, with large shift changes on F52, V54 and V58. There is also a significant shift at T50 where Leu560 on LIR2 is potentially forming a hydrogen bond. Asp559 also sits close to K49 with the acidic residue interacting with the basic face of LC3B.

However, the second highest scoring cluster contained 47 structures, with a high score of -94.6 ± 1.3 . This was close enough in number of structures and score to the highest scoring structure to warrant further investigation. The top scoring structure for this cluster is shown in Figure 5.13.

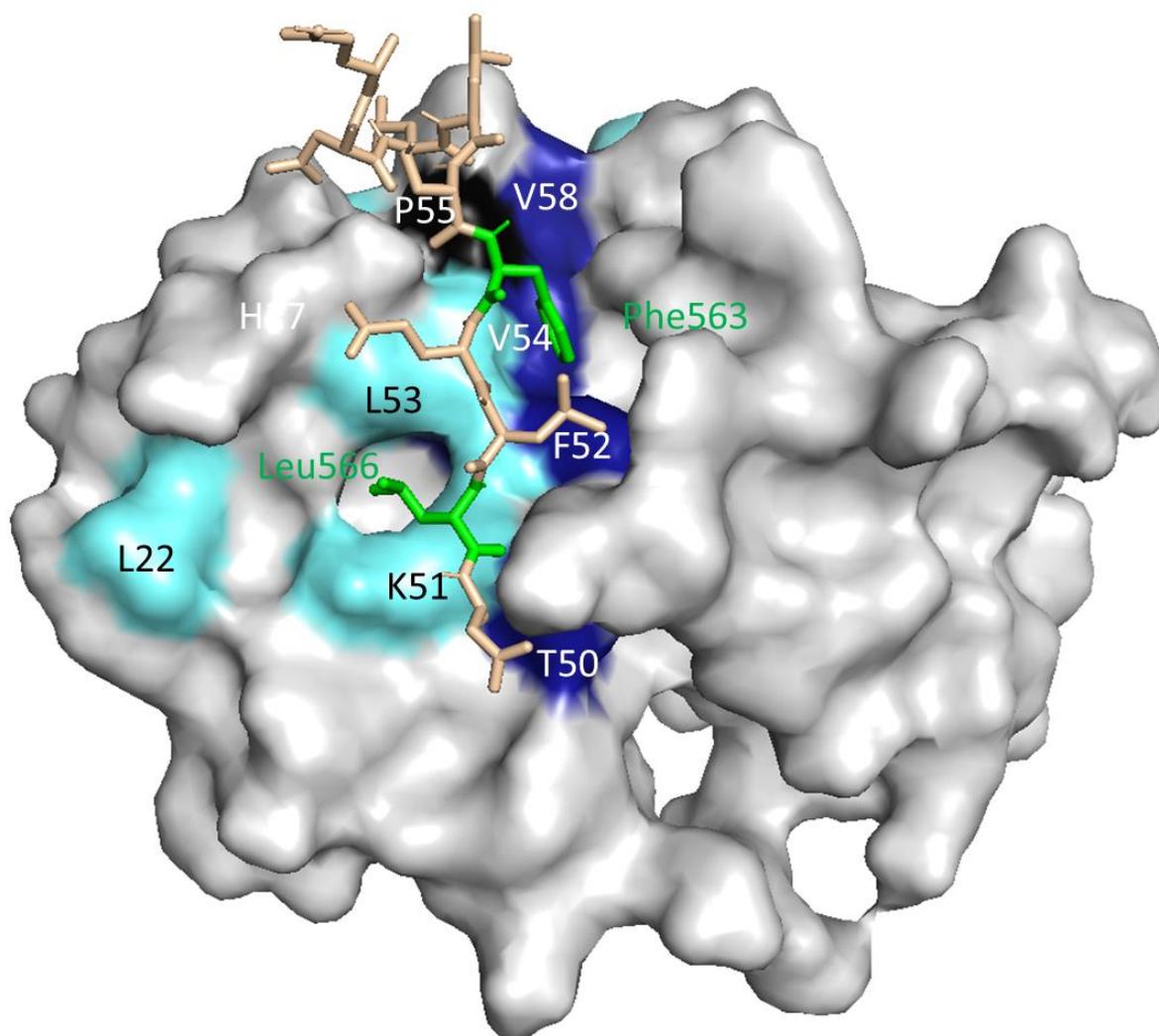


Figure 5.13: Second highest scoring HADDOCK cluster with structure of reverse NBR1 LIR2 peptide bound to LC3B. The peptide is coloured yellow with the active residues Phe556 and Leu566 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation. P55 on LC3B WT is coloured black for reference.

This structure shows the LIR in a “reverse” orientation, with Leu566 bound into HP1 and Phe563 bound in HP2. This structure whilst unlikely to be correct due to the CSP sites mapping out a typical LIR-hATG8 interaction, does suggest that the sequence of LIR2 leads to the possibility of this “reverse” orientation arising. This is likely due to there being no acidic residues directly before Phe563 to interact with the basic face of LC3B, and instead the unfavourable interaction of Leu561 with the basic face of LC3B WT, leads to the generation of the relatively high scoring “reverse” HADDOCK structure.

5.4 Interaction of WT LC3B with the LIR of p62

The LIR of p62 was one of the first LIRs that was studied interacting with LC3B, and contains the canonical WxxL binding site that has been used when describing other LIRs that have been discovered. So if the titration with p62 gave a similar interaction to that of the LIRs of NBR1 it would be confirmation of the methods used. Also due to the large number of changes over the whole of LC3B on binding, it was felt prudent to check that this was seen on a known 1:1 binding peptide, so as to check the mode of binding for the LIRs of NBR1. Previous work with p62 had used longer peptides to study the binding, and so it was decided to use a shorter peptide to model the binding interaction, so as to be directly comparable to the peptides used to mimic the LIRs of NBR1, and rule out any differences being down to interactions arising from using an extended peptide. The sequence of the peptide used was SGGDDDWTHLS (representing residues 334 to 344 in p62), which at 11 amino acids long was of similar length to those used for the LIRs of NBR1 (12 and 11). Previous work carried out had shown that the difference in the length of p62 peptide made no difference to the CSPs observed [103].

As before the peptide was titrated against LC3B until a ratio of 2:1 peptide to protein was reached, acquiring ^1H - ^{15}N HSQC at each point (Figure 5.14).

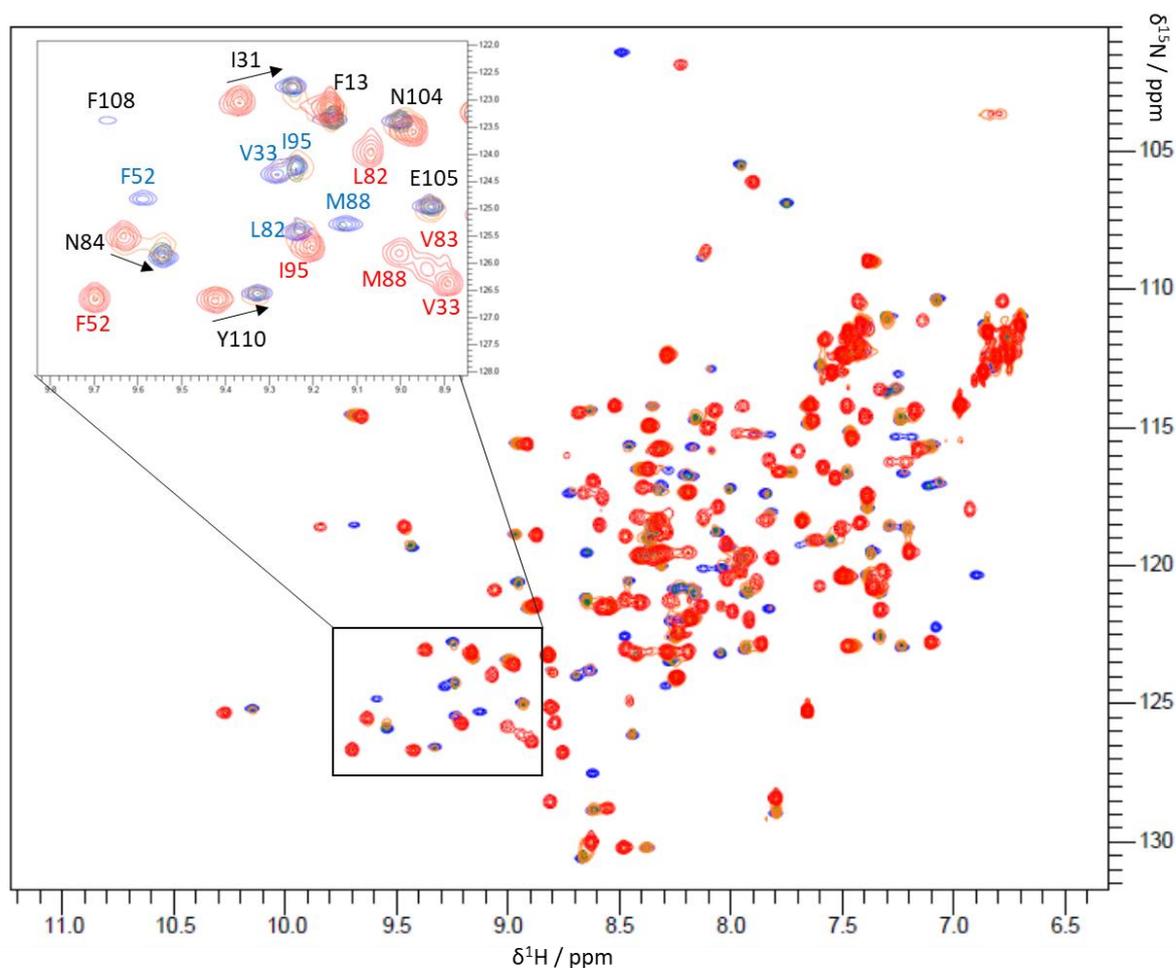


Figure 5.14: Overlay of a selection of ^1H - ^{15}N HSQC spectra for titration of LC3B WT and NBR1 LIR2 peptide. Spectra shown are the ratios of LC3B WT to LIR 1:0 (red), 1:0.5 (orange), 1:1 (green) and 1:2 (blue).

As in the ^1H - ^{15}N HSQC for the NBR1 LIR peptides most of the peaks moved. However, it was possible from following peaks, and from previous work carried out to assign all but one amino acid in the bound spectrum (V91). CSP values were then calculated for each amino acid and plotted (Figure 5.15).

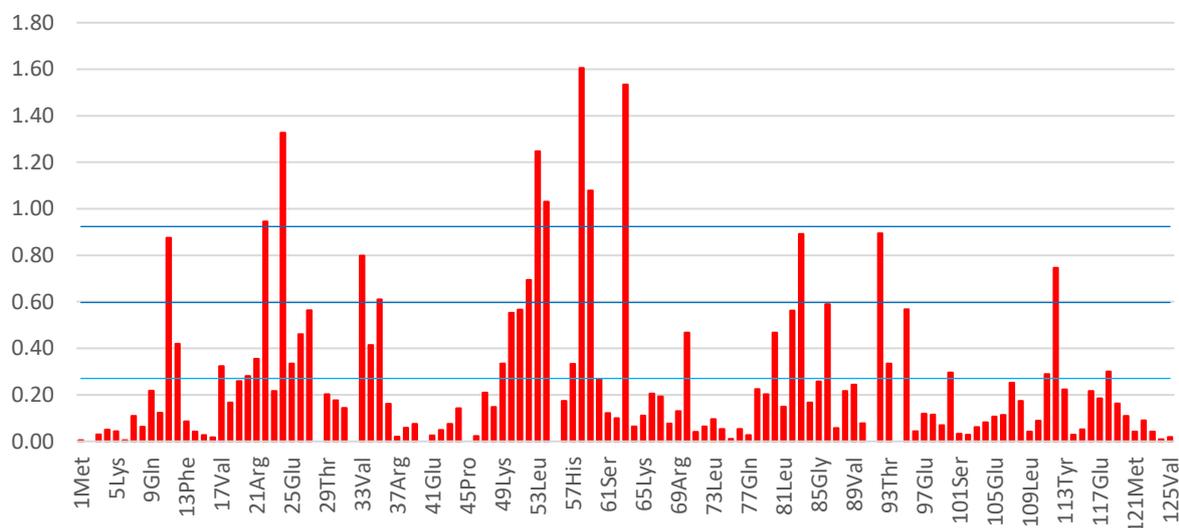


Figure 5.15: Graph of CSP values for binding of p62 LIR peptide to LC3B. The light blue line indicates the mean, the blue line is the mean and one standard deviation and the dark blue line the mean and two standard deviations.

These CSPs were then analysed to select those amino acids that underwent the most significant changes using the method in Chapter 2.5.2.6. In total there were 15 significant perturbations (>0.6 ppm), and of those eight were >0.92 ppm, and so play a major part in the binding site on LC3B when bound to the LIR of p62. The amino acids with significant CSPs were then mapped onto the surface of LC3B to identify the binding site of the shorter p62 LIR peptide (Figure 5.16).

To confirm the method used with the HADDOCK webserver was valid the CSP data was used as previously to generate a structure for the p62 LIR bond to LC3B. The top structure generated had a HADDOCK score of -119.7 ± 2.7 (Figure 5.17).

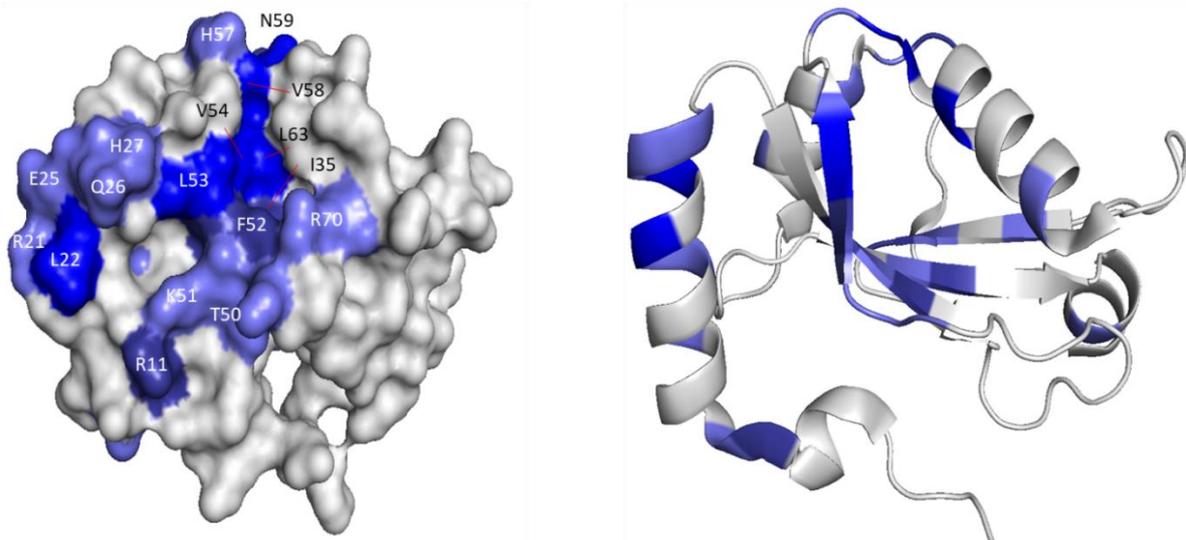


Figure 5.16: CSP mapping for the binding of p62 LIR peptide to LC3B. Residues are coloured light purple if the CSP is greater than the mean, purple if greater than the mean and one standard deviation and blue if above the mean and two standard deviations. Residues with no significant shift are coloured grey.

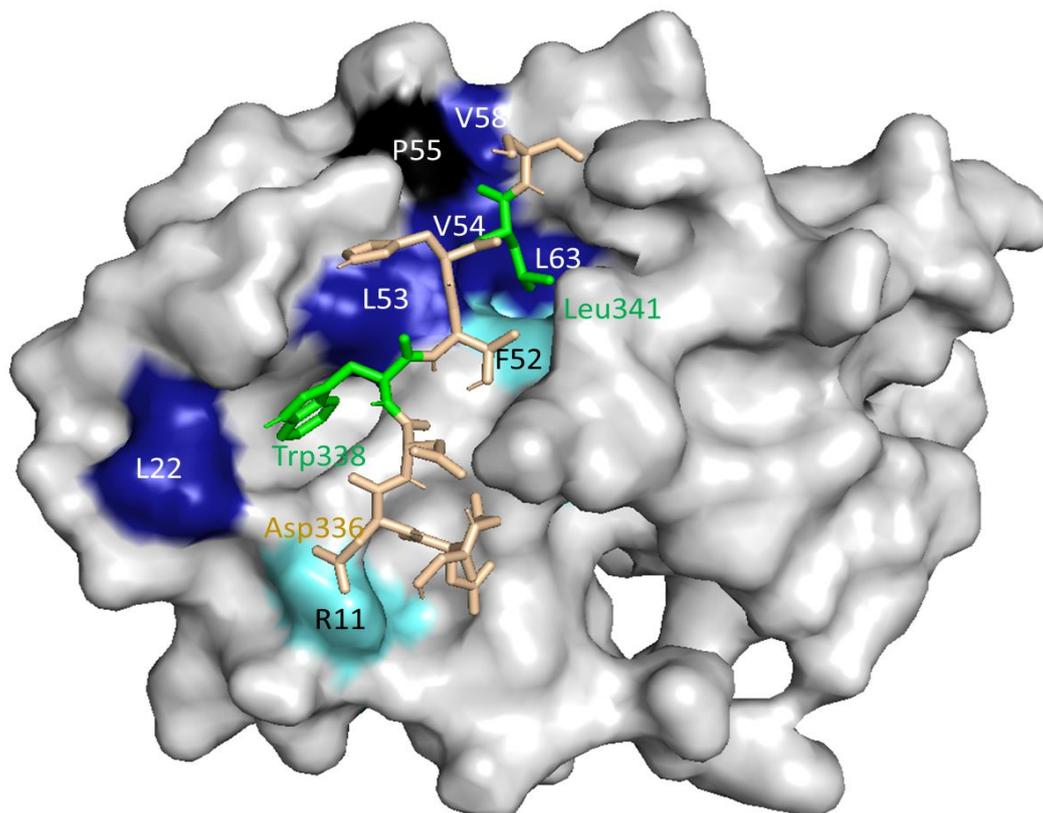


Figure 5.17: HADDOCK structure of p62 LIR peptide bound to LC3B. The peptide is coloured yellow with the active residues Trp338 and Leu341 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation. P55 on LC3B WT is coloured black for reference.

The structure produced by HADDOCK was as expected for the binding of p62 LIR peptide with Trp338 and Leu341 sitting in the two hydrophobic pockets. When compared to the published data for the binding of the p62 LIR to LC3B, the position of the peptides on the face of LC3B were very similar (Figure 5.18). There were some differences in the position of Trp338 and the preceding amino acids. However, the RMSD calculated for the difference between two structures was 1.094, which confirms that there is a good alignment between the two positions of LIR [140, 141]. As the published structure is of a crystal complex, some variation is to be expected due to the difference in conditions between how the two sets of data were collected (solid vs solution).

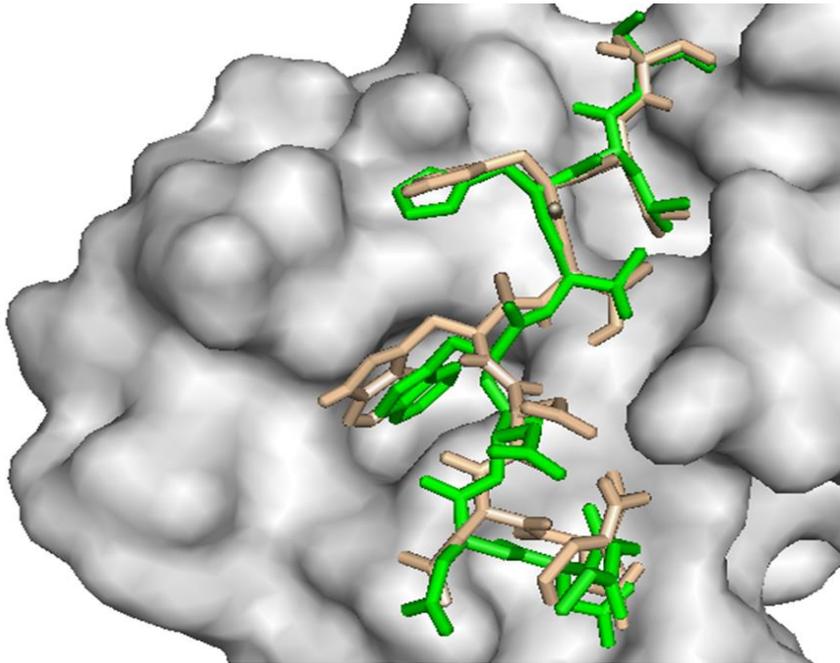


Figure 5.18: Comparison of p62 LIR binding to LC3B with crystal structure in green[61] and the top scoring HADDOCK structure in yellow.

5.5 Discussion

The data acquired in this chapter has allowed for a better understanding of the mode of binding for the peptides representing the two LIRs of NBR1 and their comparison with the well-studied interaction of LC3B with the p62 LIR peptide.

5.5.1 Titrations with the LIRs

The titrations with LC3B WT and the three LIRs all showed global changes in the ^1H - ^{15}N HSQC with large numbers of peaks shifting (Figure 5.19). The majority of the peaks that shifted during the titration underwent slow and intermediate exchange, making assignment of the end point difficult. Through the assignment of the LC3B WT bound to LIR1 ^1H - ^{15}N HSQC spectrum, and using data from a previous LC3B WT p62 LIR titration, it was possible to assign the majority of peaks in the endpoint spectrum.

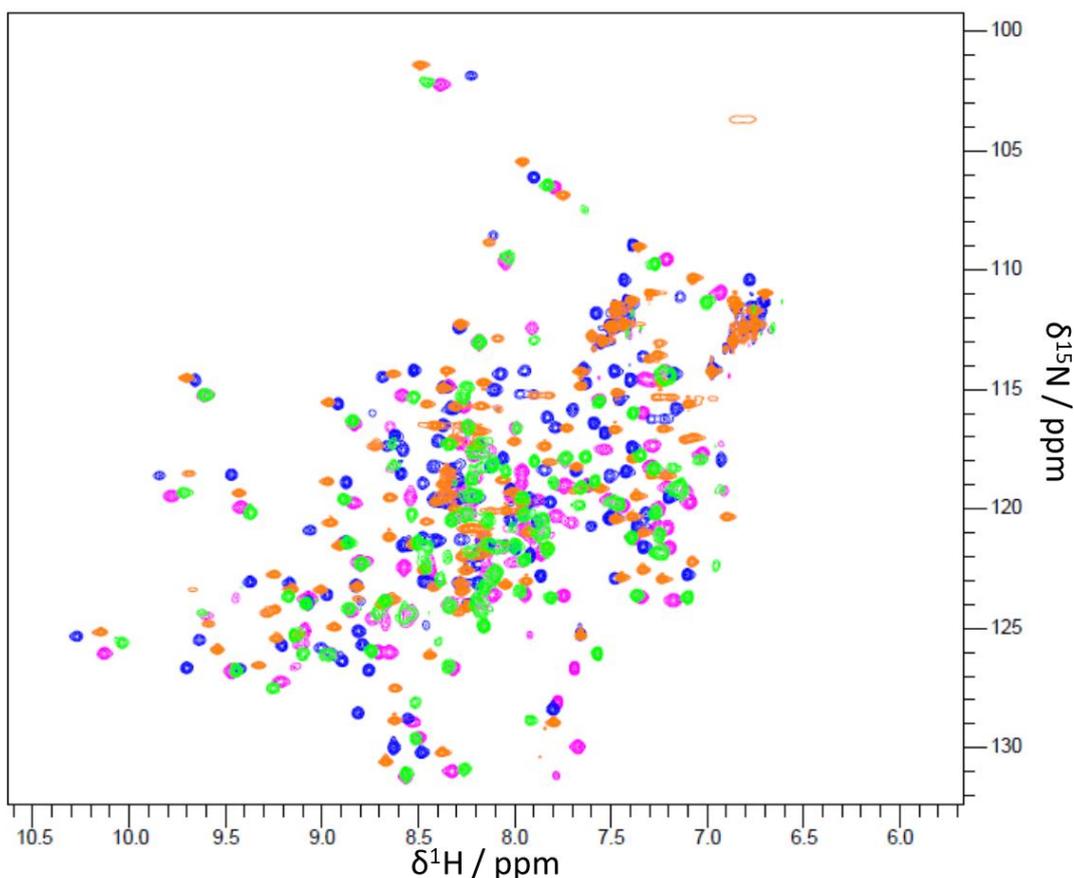


Figure 5.19: Overlay of ^1H - ^{15}N HSQC spectra for end points of LC3B WT and LIR titrations. Spectra shown are LC3B WT (blue), LC3B WT and NBR1 LIR1 (magenta), LC3B WT and NBR1 LIR2 (green) and LC3B WT and p62 LIR (orange).

By analysing the changes in ^1H - ^{15}N HSQC spectrum it was possible determine which were the most significant shifts. All three titrations had a similar number of significant shifts, with the titration with NBR1 LIR1 having 13, NBR1 LIR2 having 17 and p62 having 15. By plotting these on to the surface of LC3B it was clear that all 3 LIRs had very similar interaction patches. The majority of amino acids with significant CSPs were found around the two hydrophobic pockets into which the LIR binds. All three titrations had significant perturbations on R24 and L53 HP1, where the aromatic amino acid on the LIR binds. Around HP2 all three LIRs had significant shifts on F52, V54 and V58. As well as these common significant shifts, there were also a number of other significant shifts around these pockets that varied slightly depending on the LIR binding to the hydrophobic pocket. A list of these significant shifts is shown in Table 5.1.

When binding to the face of LC3B as well as the interactions of the aromatic and hydrophobic residues in HP1 and HP2, there are a series of extra hydrogen bonds formed between the backbones of LC3B and the LIR, with the LIR acting like an additional β -strand to the central beta sheet (Figure 5.20). The CSPs of the amides are acutely sensitive to hydrogen bonding, and hence the long range effects that are observed on LC3B when binding to the LIRs.

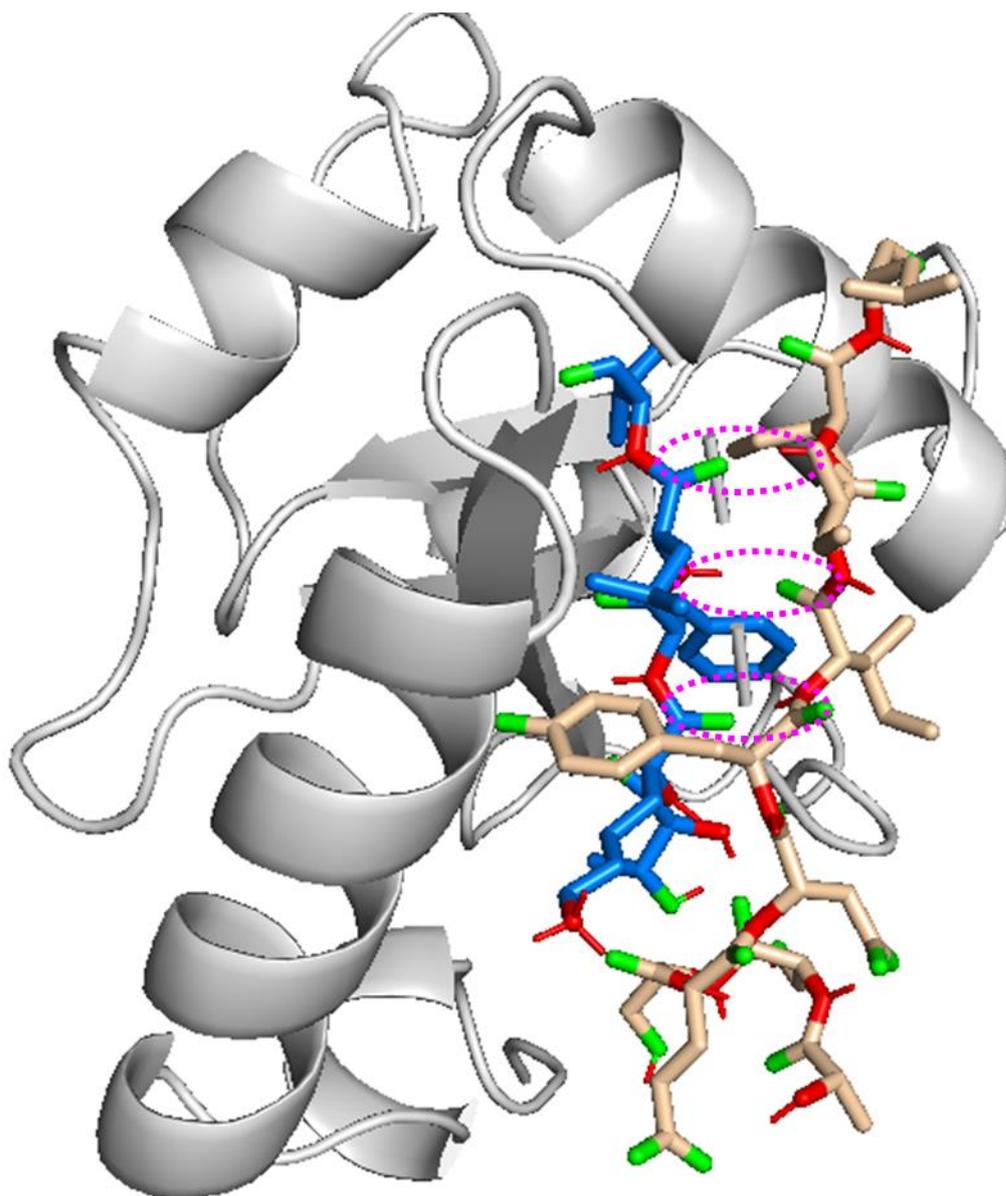


Figure 5.20: HADDOCK structure of NBR1 LIR1 bound to LC3B indicating interactions of the LIR with the face of LC3B. LC3B β -2 is shown in blue and LIR in gold, amide groups are coloured red and oxygens green. Potential hydrogen bonds between β -sheet and LIR are circled (purple).

As well as the significant shifts seen on the front face of LC3B associated with the binding of the LIR, there are also significant perturbations in the core of LC3B and on the rear face. On closer inspection these significant shifts can be separated into two distinct groups that are related to the LIR binding into the two separate hydrophobic pockets.

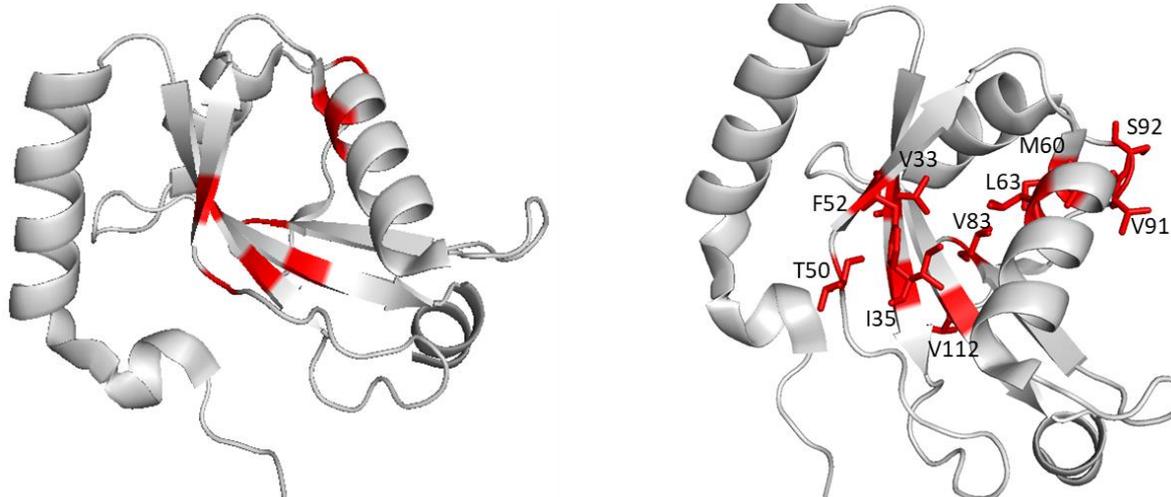


Figure 5.21: Structure of LC3B showing knock on effect of LIR binding to the hydrophobic pockets to the secondary structure. Residues that have significant shifts are coloured in red.

When plotted onto LC3B (Figure 5.21) the significant shifts are localised in two areas, the first traveling through the central β -sheet, and the second at the top of α -helix 3. The amino acids F52, T50 and I33 form part of the first hydrophobic pocket, and so as the aromatic residue binds to the pocket this causes movement within the secondary structure, which is then passed through the core of LC3B and to the rear face. A similar re-organisation happens as hydrophobic residue of the LIR binds into the second hydrophobic pocket, which contains Leu63. This then causes movement throughout the top of the alpha helix. This structural re-organisation could also optimise the interaction with bound ligand. The long range perturbations seen on LC3B suggest some degree of induced fit and global tightening to optimise the interaction, all of which will contribute to the binding free energy.

5.5.2 Comparing the Binding of the LIRs

Using the HADDOCK server to dock the peptides representing the LIRs onto LC3B gave structures showing where the peptides were binding to LC3B. It is important not to put too much weight behind structures generated by a simplified docking procedure. However, they do allow the visualisation of the interactions of the different LIRs with LC3B.

All 3 of the HADDOCK structures showed the LIR peptide binding to LC3B at the known binding site, with the aromatic residue sitting in HP1, and either the leucine or isoleucine hydrophobic residue of the LIR binding into HP2.

However, there were differences in how the aromatic residue binds to the first hydrophobic pocket reflecting the difference in the interaction with tryptophan, phenylalanine and tyrosine. There was a difference in how the hydrophobic residue sat in the second hydrophobic pockets and how the other amino acids of the peptide sit on the face of LC3B. In Table 5.1 is a summary of the amino acids with significant amino acids for the three LC3B WT titrations and the HADDOCK results.

| Titration | WT-NBR1LIR1 | WT-NBR1LIR2 | WT-p62LIR |
|---------------------------------------|---|---|--|
| Amino acids with most significant CSP | R24, H27, T50, F52, V54, V58, N59, V91, S92 | R24, H27, T50, F52, V54, V58, N59, V91 | L22, R24, L53, V54, V58, N59, L63, V91 |
| Amino acids with significant CSP | R21, K51, L53, V83 | L22, K51, L53, D56, N59, V83, S92, V112 | R11, V33, I35, F52, V83, S92, V112 |
| HADDOCK run results | 170 structures 10 clusters | 167 structures 11 clusters | 164 structures 9 cluster |
| Top scoring HADDOCK cluster | 66 structures Top score -91.6±4.4 | 49 structures Top score -102.7±2.9 | 79 structures Top score -119.7±2.7 |

Table 5.1: Amino acids with significant CSPs and HADDOCK results for titrations of LC3B WT with LIR peptides. Those highlighted green appear as a most significant CSP in all titrations.

All of the HADDOCK runs produced a similar number of structures, with the interaction of the p62 LIR with LC3B WT having the highest scoring structure. However, the scores for the top structures for the two LIRs were similar to each other, and still close to that of p62. NBR1 LIR2 also generated an alternative structure that had a very similar score to the top scoring cluster (-94.5±1.3/ 47 structures), that showed the LIR in a “reverse” orientation. Looking at the similarity in the significant shifts between the three LIRs it seems unlikely that the “reverse” orientation is what is occurring when LIR binds to LC3B WT. However, the HADDOCK score suggest that the sequence present in LIR2, makes the “reverse” orientation feasible, which may be of significance for non-standard LIR sequences, of which a number have been identified.

5.5.3 Comparison of Titration of LC3B WT with NBR1 LIR1 and p62 LIR

To probe the differences between the interaction of the LIR of p62 and LIR1 of NBR1 the CSPs were subtracted from each other to identify any ligand specific differences (as with the CSPs from the titration data). As the CSP reflects how the protein “adapts” on the LIR binding, then significant changes in these values between the two interactions would flag up differences in how the LIR binds to LC3B. These differences in CSP and the significance cut offs were then plotted (Figure 5.22).

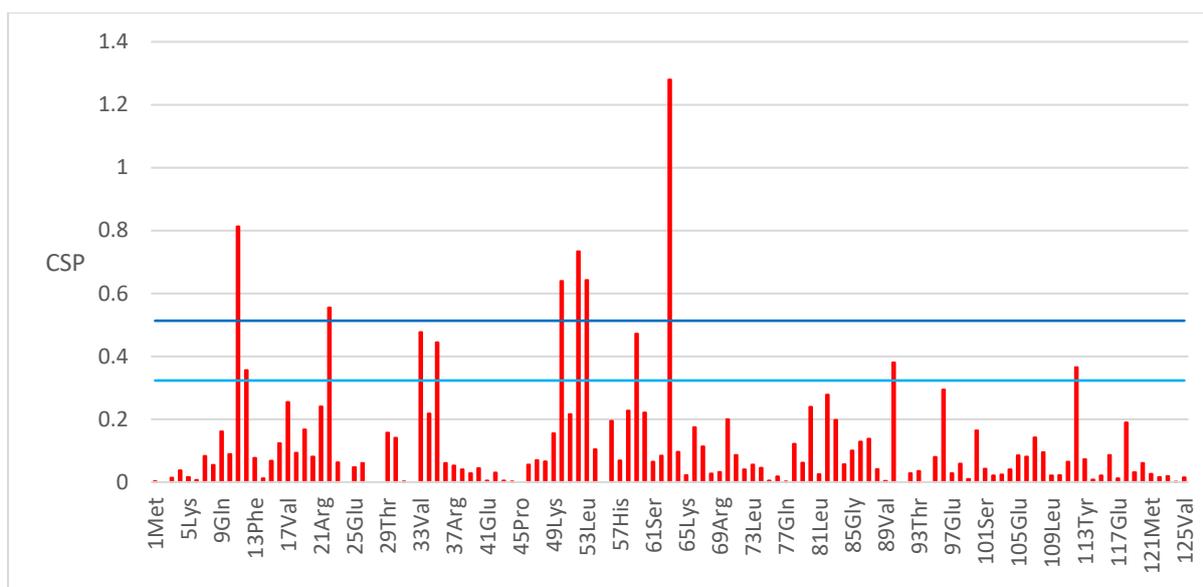


Figure 5.22: Plot of differences between CSPs for titrations LC3B WT with NBR1 LIR1 and p62 LIR. Significant differences are shown in blue (mean + one standard deviation) and most significant differences is shown in dark blue (mean + two standard deviations).

In total there were 13 CSPs that were significantly different between the two titrations, and 7 of these were classed as the most significant (Table 5.2), with differences in CSP of up to 1.2ppm observed, which are shown graphically and plotted onto the surface of LC3B in Figure 5.23.

Significant Differences in CSP

| Amino Acid | P62 LIR CSP | NBR1 LIR1 CSP |
|------------|-------------|---------------|
| T12 | 0.42 | 0.06 |
| V33 | 0.80 | 0.32 |
| I35 | 0.61 | 0.17 |
| N59 | 1.08 | 1.55 |
| S90 | 0.08 | 0.46 |
| V112 | 0.75 | 0.38 |

Most Significant Differences in CSP

| Amino Acid | P62 LIR CSP | NBR1 LIR1 CSP |
|------------|-------------|---------------|
| R11 | 0.88 | 0.06 |
| L22 | 0.94 | 0.39 |
| H27 | 0.56 | CNF |
| T50 | 0.55 | 1.19 |
| F52 | 0.69 | 1.43 |
| L53 | 1.24 | 0.61 |
| L63 | 1.53 | 0.25 |

Table 5.2: List of significant and most significant differences between the titrations of LC3B WT with NBR1 LIR1 and p62 LIR. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration.

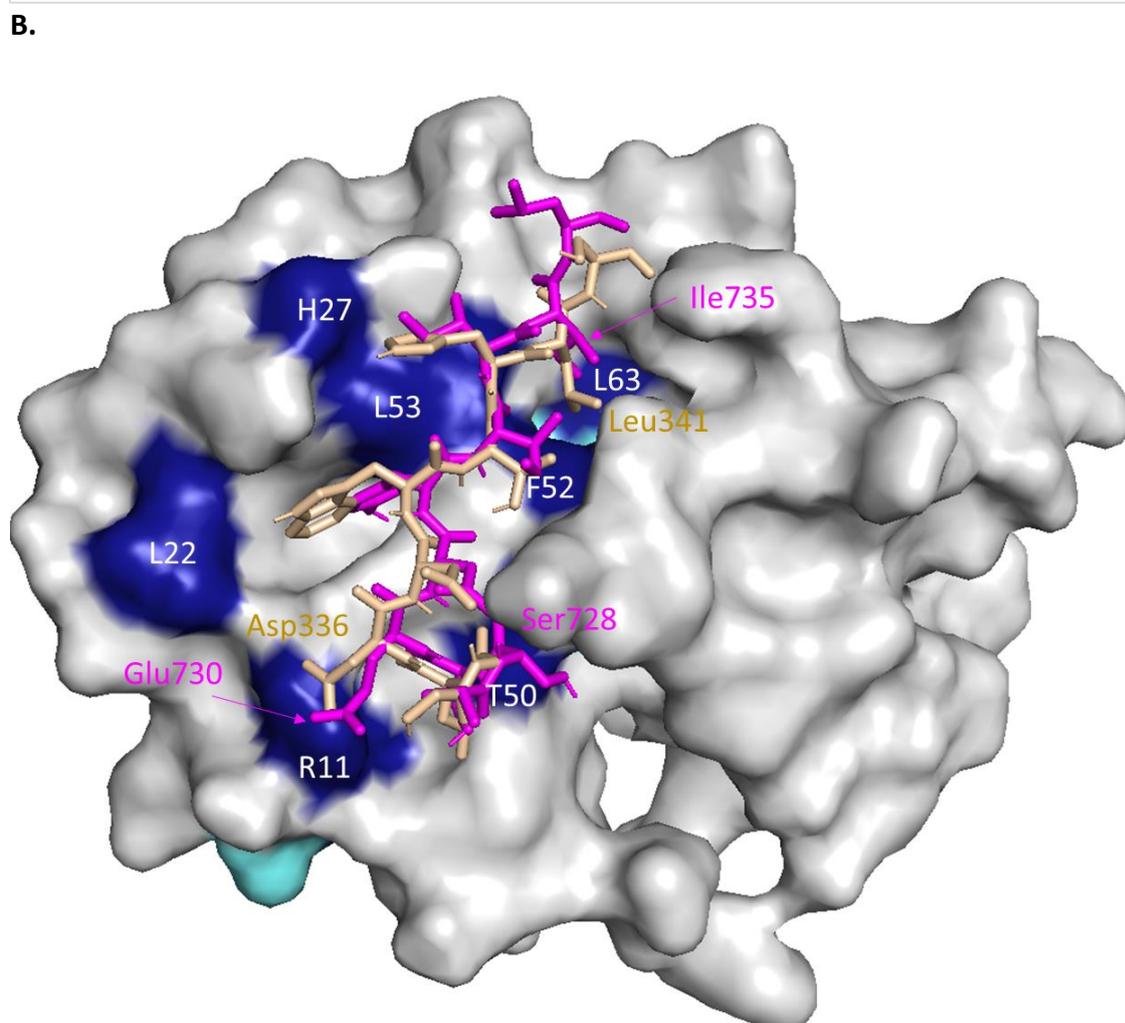
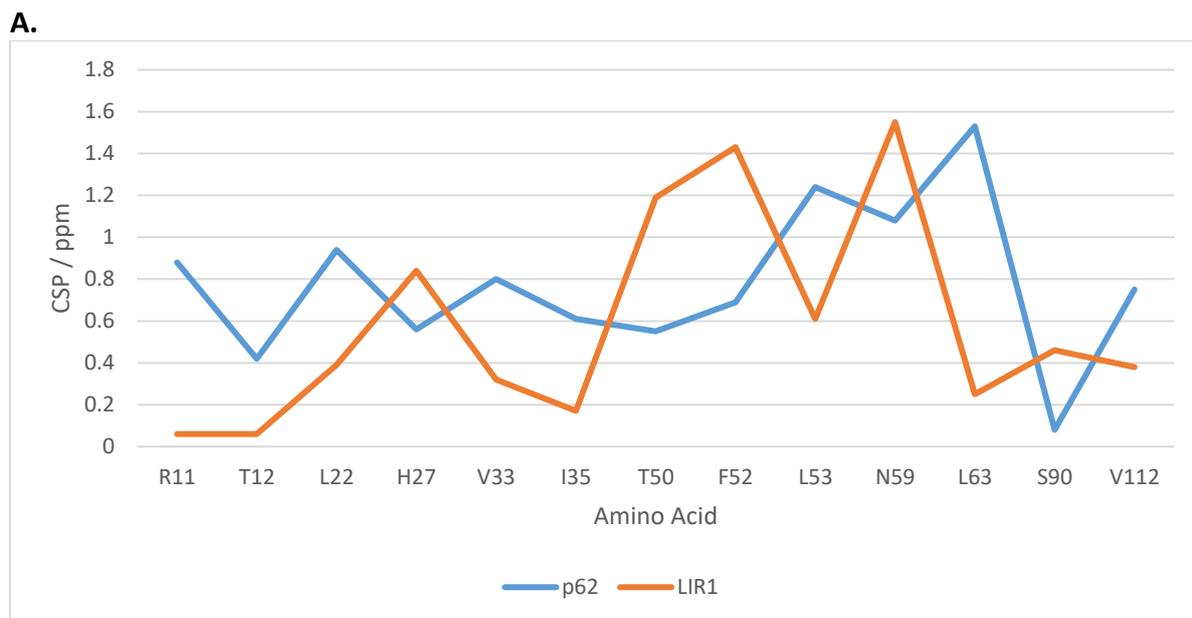


Figure 5.23: A: Graphical representation of CSPs that are significantly different between the titration with p62 (blue) and NBR1 LIR1 (orange). B: Overlay of HADDOCK structure for the binding of p62 LIR (yellow) and NBR1 LIR1 (pink) on the surface of LC3B. Significant shifts are coloured from light blue and most significant shifts are coloured dark blue.

The largest changes were seen on amino acids R11, L22, V33, T50, F52, L53 and L63. The binding with the p62 peptide gave a very large perturbation for R11 on LC3B, whereas for the NBR1 LIR1 peptide there was hardly any change. From the HADDOCK model we can see that the amino acids of the peptide interacting with R11, are Asp336 on p62 and Glu730 on NBR1. From looking at the structure this difference can be attributed to the position of the peptide giving the shorter chain of Asp336 a better orientation to interact with R11, whilst Glu730 on NBR1 LIR1 is pointing away from LC3B WT.

There is a larger perturbation for the NBR1 LIR1 on LC3B WT for T50 where Ser728 is likely to be forming a side chain hydrogen bond, whereas p62 has Gly334 at this position, which would be unable to interact with T50 in the same way.

The differences around the first hydrophobic pocket show that for L22 and L53 there is a larger perturbation for p62 than for NBR1 LIR1, whereas the change in H27 and F52 is larger for NBR1 LIR1 than the LIR of p62. This suggests that Trp338 of p62 is sitting more centrally between the α -helix 2 and the central beta sheet, whereas Tyr733 on NBR1 sits close to the central β -sheet. The CSP on V33 is also bigger for p62, suggesting Trp338 sits deeper into the pocket. The change in aromatic residue binding in HP1 would also cause larger CSP changes for Trp338 due to larger ring current effects from a tryptophan, leading to large changes in environment for amino acids that sit close to it in HP1.

Finally the CSP for L63 when binding with p62 was considerably bigger than that of the binding with NBR1 LIR1 (> 1 ppm), suggesting that Leu341 sits much further into the second hydrophobic pocket than Ile735 on NBR1 LIR1. This could be due to the difference in the flexibility of side chain fitting into the hydrophobic pocket. Isoleucine is a β -branched residue and so has restricted rotation about the C_{α} - C_{β} bond, whereas leucine is more extended and so can fit into the pocket more easily.

In the literature [113] the interactions of p62 and NBR1 LIR1 with LC3B have been studied, and it was shown that p62 binds with a slightly stronger binding affinity for LC3B WT ($K_d = 1.5 \mu\text{M}$) than NBR1 LIR1 ($K_d = 2.9 \mu\text{M}$). This seems to be in agreement with the titration data, with the p62 having a different binding interaction with Trp338, causing larger perturbations in HP1, and Leu341 being able to sit further into HP2 due to it being slightly less bulky than Ile735 on NBR1 LIR1. There is also the extra interaction p62 Asp336 is able to make with R11 on LC3B, although this is potentially offset by Ser728 on NBR1 LIR1 interacting with T50.

5.5.4 Comparison of Titration of LC3B WT with NBR1 LIR2 and p62 LIR

As for NBR1 LIR1, the difference between the CSPs of p62 and NBR1 LIR2 were calculated. These values were then plotted with the significance cut offs (Figure 5.24).

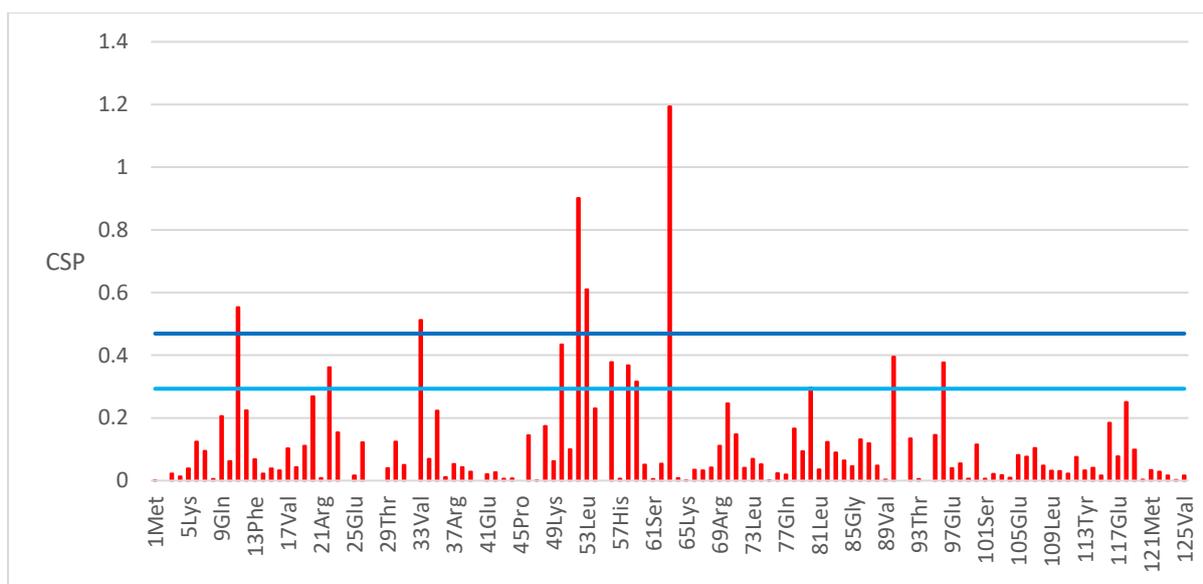


Figure 5.24: Plot of differences between CSPs for titrations LC3B WT with NBR1 LIR2 and p62 LIR. Value for significant differences (mean and one significant difference) is shown in blue and most significant differences (mean and two significant difference) is shown in dark blue.

In total there were 14 significant differences in CSPs between the two titrations of NBR1 LIR2 and p62 LIR with LC3B WT, and of these 6 were classed as most significant. The amino acids that had significant differences between titrations and the CSP values are given in Table 5.3, and shown graphically and used to colour the surface of LC3B in Figure 5.25

Significant differences in CSP

| Amino Acid | P62 LIR CSP | NBR1 LIR2 CSP |
|------------|-------------|---------------|
| L22 | 0.94 | 0.58 |
| T50 | 0.55 | 0.99 |
| D56 | 0.17 | 0.55 |
| V58 | 1.60 | 1.24 |
| N59 | 1.08 | 1.39 |
| F80 | 0.47 | 0.17 |
| S90 | 0.07 | 0.47 |
| S96 | 0.04 | 0.42 |

Most Significant Differences in CSP

| Amino Acid | P62 LIR CSP | NBR1 LIR2 CSP |
|------------|-------------|---------------|
| R11 | 0.88 | 0.32 |
| H27 | 0.56 | CNF |
| V33 | 0.80 | 0.28 |
| F52 | 0.69 | 1.60 |
| L53 | 1.25 | 0.64 |
| L63 | 1.53 | 0.34 |

Table 5.3: List of significant and most significant differences between the titrations of LC3B WT with NBR1 LIR2 and p62 LIR. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration.

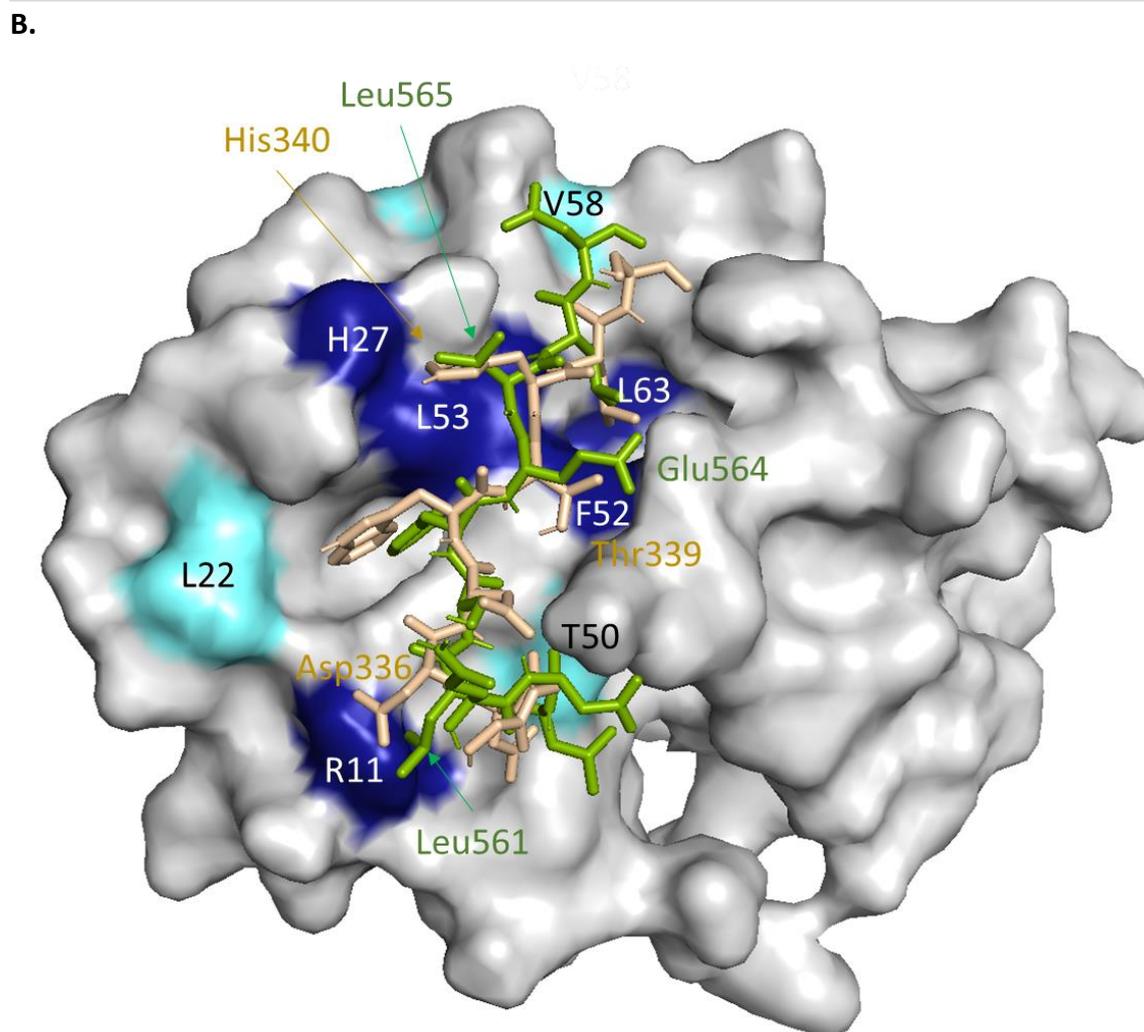
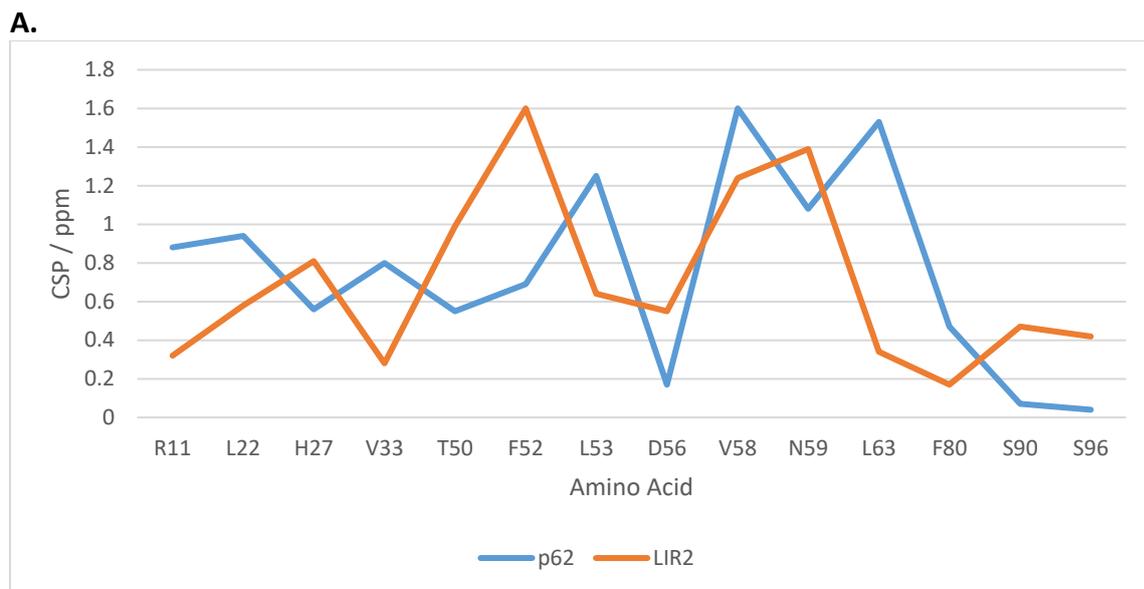


Figure 5.25: A: Graphical representation of CSPs that are significantly different between the titration with p62 (blue) and NBR1 LIR2 (orange). B: Overlay of HADDOCK structure for the binding of p62 LIR (yellow) and NBR1 LIR2 (green) on the surface of LC3B. Significant shifts are coloured light blue and the most significant shifts are coloured dark blue.

Again there were a number of large changes around the binding site and surrounding surface. The differences in CSPs were seen at R11, V33, F52, L53 and L63. There were still significant differences in the CSPs for L22 and T50. However, these were not as large as the differences between p62 and NBR1 LIR1. As seen with NBR1 LIR1 There is large CSP for the interaction of p62 with LC3B at R11, and hardly any with NBR1 LIR2. For NBR1 LIR2 the HADDOCK model shows R11 in close contact with Leu561, whereas for the p62 LIR it is Asp336. This then explains why the CSP for R11 is so much smaller for the peptide of NBR1 LIR2, as there is not the electrostatic interaction that is present for p62. The CSP for L22 and L53 for p62 is still larger than that of Phe563 on NBR1 LIR2, suggesting it sits closer to the central β -sheet than Trp338 on p62. The CSP for V33 is also bigger for p62, suggesting that Phe563 on NBR1 LIR2 does not fit as deeply into the hydrophobic pocket as W338 on p62. The CSP for F52 is considerably bigger for NBR1 LIR2 than p62 (~ 0.9 ppm), showing that Glu564 of NBR2 is sitting closer than Thr339 of p62. Again the CSP for p62 is considerably bigger than that for NBR1 LIR2, most likely down to the extra ring current effect present due to the p62 LIR containing a tryptophan and NBR1 LIR2 having a phenylalanine.

Work carried out in the Searle group [114] has indicated that NBR1 LIR2 is a weaker binder than NBR1 LIR1, and so subsequently would be a weaker binder than p62. This is confirmed by the difference seen between the titration data of the LIR of p62 and NBR1 LIR2, with Trp338 and Leu341 sitting more deeply in the hydrophobic pockets, and the electrostatic interactions between R11 and Asp336 on p62 stabilising the binding.

5.5.5 Comparison of Titration of LC3B WT with NBR1 LIR1 and NBR1 LIR2

Finally the CSPs of the titration of the two LIR's of NBR1 were compared. The significant differences between the CSPs were calculated and plotted with the cut off points (Figure 5.26).

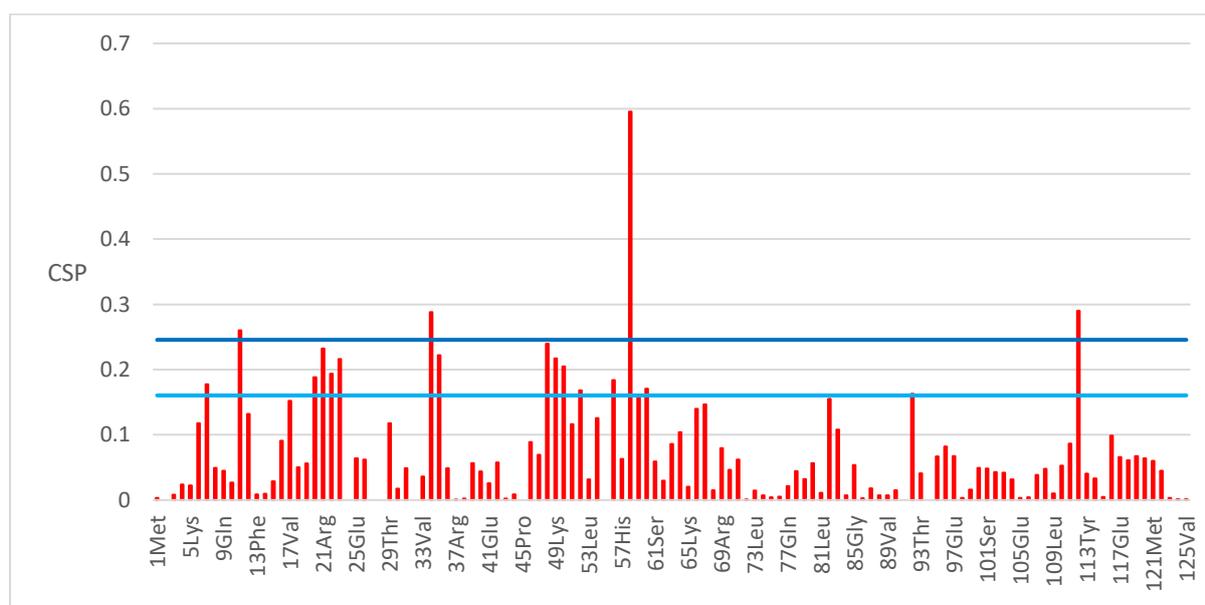


Figure 5.26: Plot of differences between CSPs for titrations LC3B WT with NBR1 LIR1 and p62 LIR. Value for significant differences (mean and one significant difference) is shown in blue and most significant differences (mean and two significant difference) is shown in dark blue.

In total there were 17 amino acids that had a significant difference between the CSP for the titration of NBR1 LIR1 and LIR2 with LC3B WT. However, only four of these were classed as most significant. A list of the amino acids and the CSP for each titration is given in Table 5.4, and shown graphically and mapped onto the surface of LC3B with HADDOCK structure for the LIR binding overlaid (Figure 5.27).

Significant Differences in CSP

| Amino Acid | NBR1 LIR1 CSP | NBR1 LIR2 CSP |
|------------|---------------|---------------|
| F7 | 0.03 | 0.20 |
| V20 | 0.20 | 0.01 |
| R21 | 0.59 | 0.36 |
| L22 | 0.59 | 0.36 |
| I23 | 0.15 | 0.37 |
| I35 | 0.17 | 0.39 |
| D48 | 0.08 | 0.32 |
| K49 | 0.18 | 0.40 |
| T50 | 1.19 | 0.99 |
| F52 | 1.43 | 1.60 |
| D56 | 0.37 | 0.55 |
| M60 | 0.05 | 0.22 |
| S92 | 0.92 | 0.76 |

Most Significant Differences in CSP

| Amino Acid | NBR1 LIR1 CSP | NBR1 LIR2 CSP |
|------------|---------------|---------------|
| R11 | 0.06 | 0.32 |
| I34 | 0.20 | 0.48 |
| V58 | 1.83 | 1.24 |
| V112 | 0.38 | 0.67 |

Table 5.4: List of significant and most significant differences between the titrations of LC3B WT with NBR1 LIR1 and NBR1 LIR2. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration.

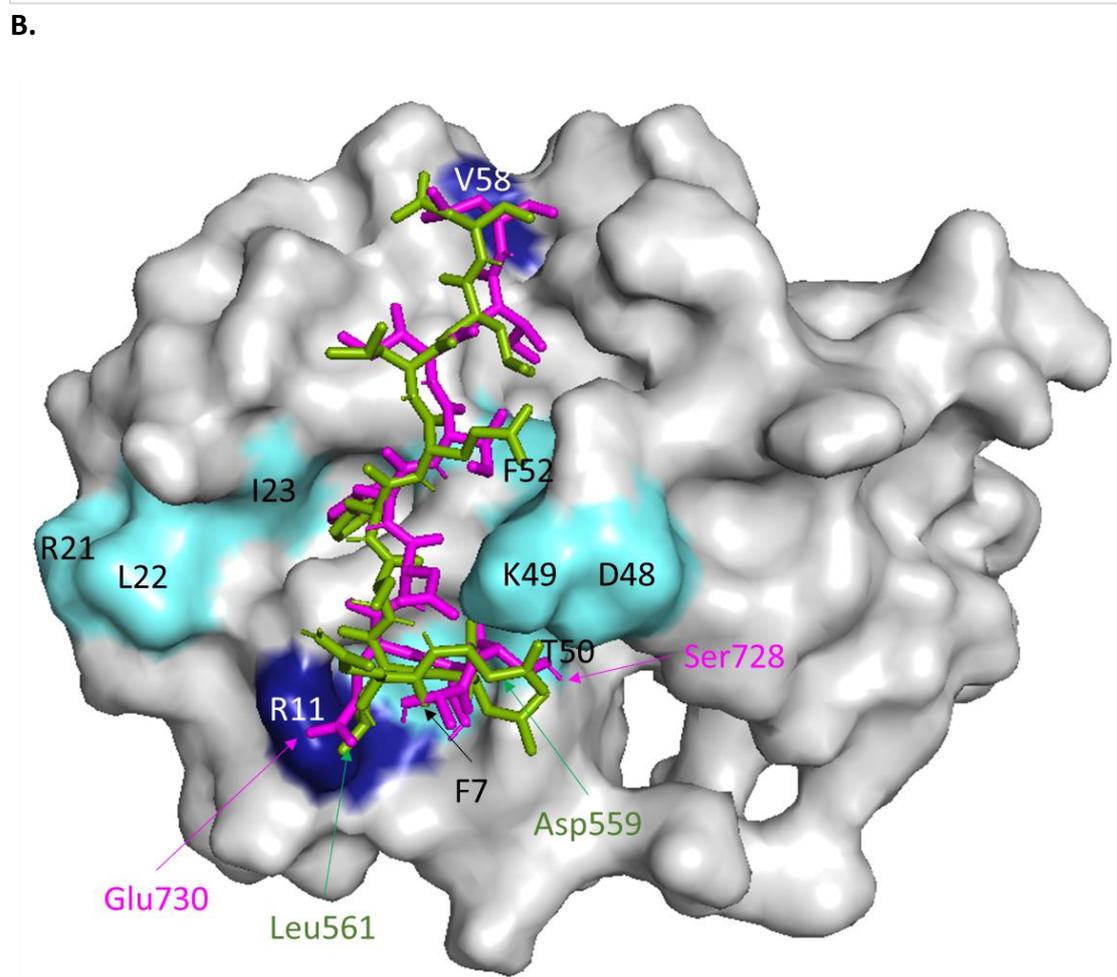
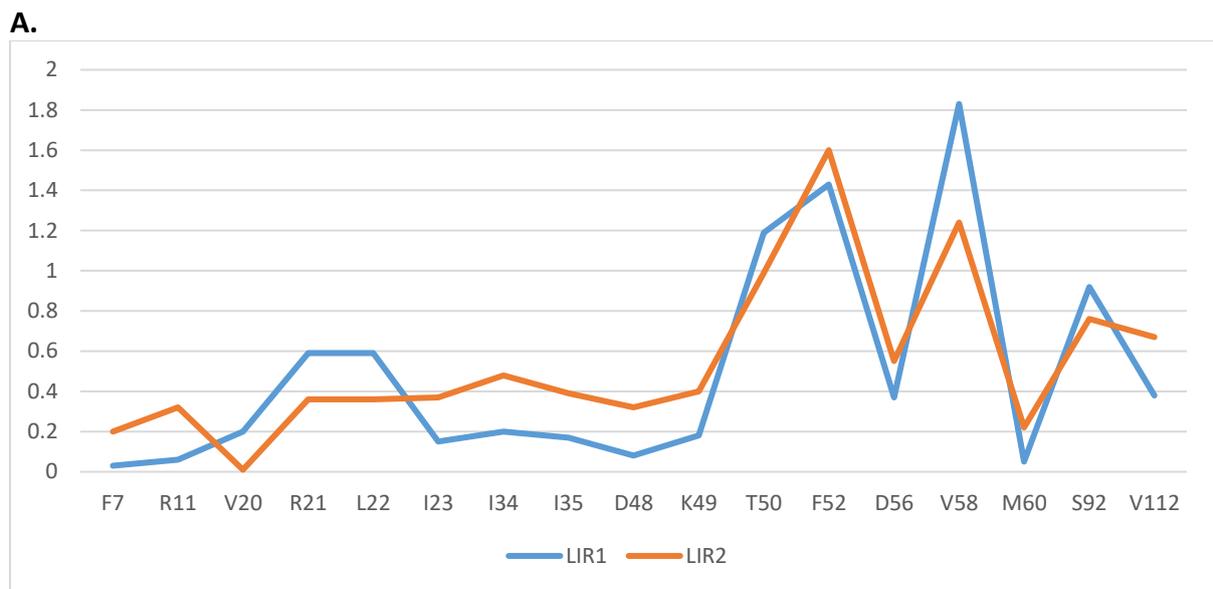


Figure 5.27: A: Graphical representation of CSPs that are significantly different between the titration with NBR1 LIR1 (blue) and NBR1 LIR2 (orange). B: Overlay of HADDOCK structure for the binding of NBR1 LIR1 (pink) and NBR1 LIR2 (green) on the surface of LC3B. Significant shifts are coloured light blue with most significant shifts coloured dark blue.

When comparing the CSPs of the two LIRs there were only a few residues that were of the largest significance. These were R11, I34 and V58. The rest of the significant differences that were of a lower value were R21, L22, I23 on alpha helix 2, and D48, K49, T50 and F52. There is a big difference for R11 potentially due to LC3B having to reorganise more due to the unfavourable interaction with Leu561. The CSPs for D48 and K49 are larger for LIR2 than LIR1, indicating that Asp559 is interacting with K49 which is then keeping the amide backbone away from the face of LC3B. The CSPs around the second hydrophobic pocket are very similar, suggesting that Ile736 of LIR and Leu566 of LIR2 sit in that pocket in a similar fashion.

As mentioned earlier previous work in the Searle group has indicated that the binding affinity of NBR1 LIR1 is stronger than that of LIR2. This would seem to mainly be facilitated by the amino acids preceding the aromatic residue in the LIR sequence. This is backed up by published results that show that there was no change in the binding affinities of LIR1 with the Y733F LIR1 peptide mutant [113].

5.6 Summary

This chapter has looked at the interaction of LC3B with the LIR of p62 and the two LIRs of NBR1 using NMR titrations to understand the details of the binding. The key results were

- Full assignment of the amide backbone in the ^1H - ^{15}N HSQC for LC3B and LC3B when bound to NBR1 LIR1 peptide.
- Titrations confirmed the mode of binding of the two LIRs of NBR1 with LC3B were similar to that of the LIR of p62.
- HADDOCK docking structures were produced using restraints generated from the CSPs.
- Using the HADDOCK structures and looking in detail at the differences in CSPs for the binding of the LIRs with LC3B it was possible to gain greater understanding of the relative binding affinities of the LIRs.
- Having a tryptophan at the aromatic residue position of the LIR increases the binding affinity but the preceding amino acids to this play a key part to the binding action, with interactions between R11, K49 and T50 on LC3B and Asp336 (p62), Asp559 (NBR1 LIR2) and Ser728 (NBR1 LIR1) respectively being observed.
- The amino acids preceding Tyr732 on LIR1 have more favourable interactions with the face of LC3B than those preceding Phe563 on LIR2 contributing to LIR1 binding more favourably than LIR2 with LC3B WT.

Chapter 6

6. Investigating Selectivity Between the two LIRs of NBR1

This chapter aims to investigate further the difference between the two LIRs of NBR1 when binding to LC3B. This will be done by utilising two mutants of LC3B which alter the binding sites, and then using NMR titrations to map the binding surface of the interaction. The data will then be used with HADDOCK to generate structures of the LIR binding to the LC3B mutants. Finally results from these titrations and HADDOCK structures will be compared with the data obtained in Chapter 5 for LC3B WT.

Whilst both LIRs contain the standard LIR binding motif of an aromatic and hydrophobic residue to interact with the two hydrophobic pockets on LC3B, they do differ in the makeup of amino acids that precede the aromatic residue. NBR1 LIR1 follows the canonical consensus sequence of an LIR, where the aromatic residue is preceded by acidic residues, whereas the sequence of NBR1 LIR2 contains mainly hydrophobic amino acids.

The LC3B K51A, L53A mutant (KL mutant) had shown some interesting results in the MS competition experiments, where the NBR1 LIR2 peptide had bound preferentially to the LC3B KL mutant, rather than LC3B WT (chapter 4). This can be attributed to the change in charge on the face of LC3B by mutating a lysine to an alanine, which removes the effect of the acidic residues preceding the aromatic acid on NBR1 LIR1 interacting with the face of LC3B, and hence changing the preferred binding partner (for NBR1 LIR1 LC3B WT was the preferred binding partner over the KL mutant).

The phosphomimetic mutant (T29D mutant) showed similar affinity to LC3B WT for LIR1. However, for LIR2 the binding affinity for the T29D mutant was slightly less than for the wild type protein.

6.1 Backbone Assignment of LC3B KL and T29D Mutant

As with the titrations of the NBR1 LIRs with LC3B WT the backbone amide peaks in the ^1H - ^{15}N HSQC assignment would be required to carry out the titration experiments. It was hoped that as the changes being made to LC3B were relatively small and on the surface, it would be possible to transfer the assignment for the wild type directly to the KL and T29D mutant.

6.1.1 LC3B KL Mutant Assignment

On running a ^1H - ^{15}N HSQC of a ^{15}N enriched sample of the KL mutant it was clear that there was considerable movement for a number of peaks (Figure 6.1).

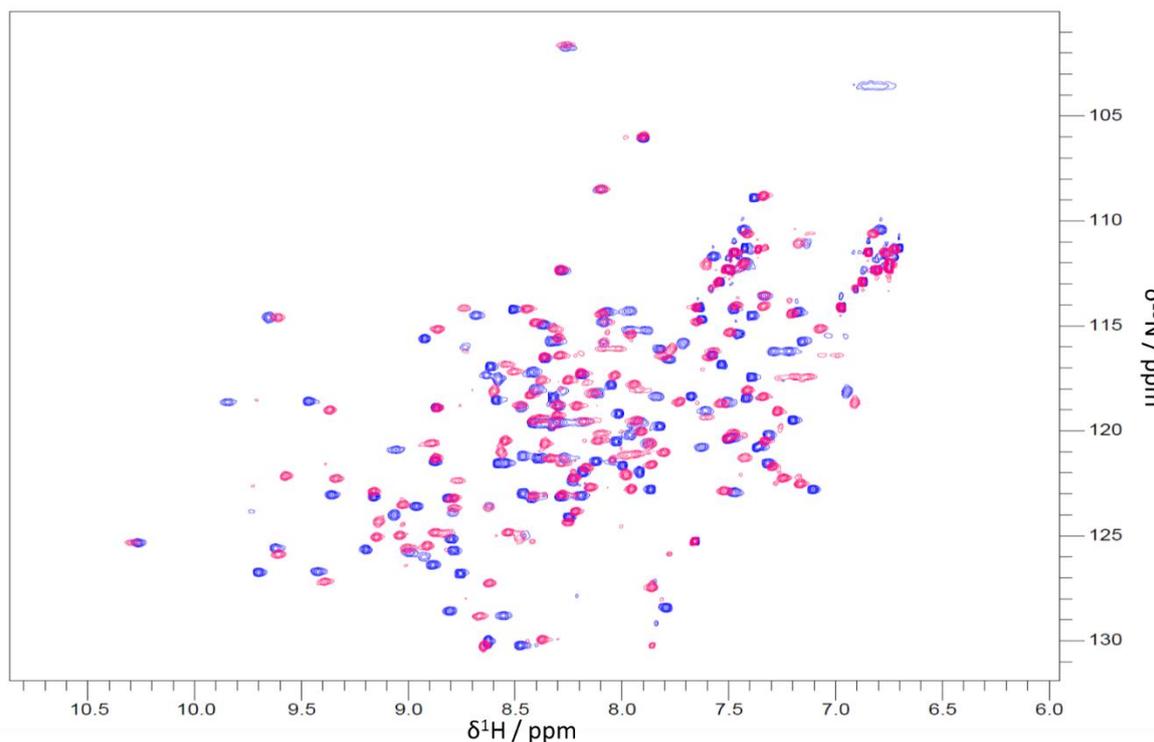


Figure 6.1: Overlay of LC3B WT (blue) and LC3B KL (pink) ^1H - ^{15}N HSQC.

Whilst it was possible for some peaks to transfer the assignment, for a large number in the congested regions of the spectra it was not possible. It was thus decided to carry out a full backbone assignment of the KL mutant so as to assign those peaks in congested regions, and to confirm the assignment of peaks where it was possible to directly transfer the assignment from WT to KL spectra.

As with the ^1H - ^{15}N HSQC backbone amide assignment a ^{15}N and ^{13}C enriched LC3B KL protein sample was produced (Chapter 2.3.11) and then a series of 3D NMR experiments were collected. As before, the experiments used were ^1H - ^{15}N HSQC, HNC(O), HN(CA)CO, HNCA, HN(CO)CA, HNCACB and HN(CO)CACB (Chapter 2.5.2.2). The ^1H - ^{15}N HSQC backbone amide assignment was then carried out using the method in Chapter 2.5.2.4.

Using this method it was possible to assign all but S90, T29 and V91 (97% complete), which would be more than enough to carry out the titrations with the LIRs of NBR1 (Figure 6.2).

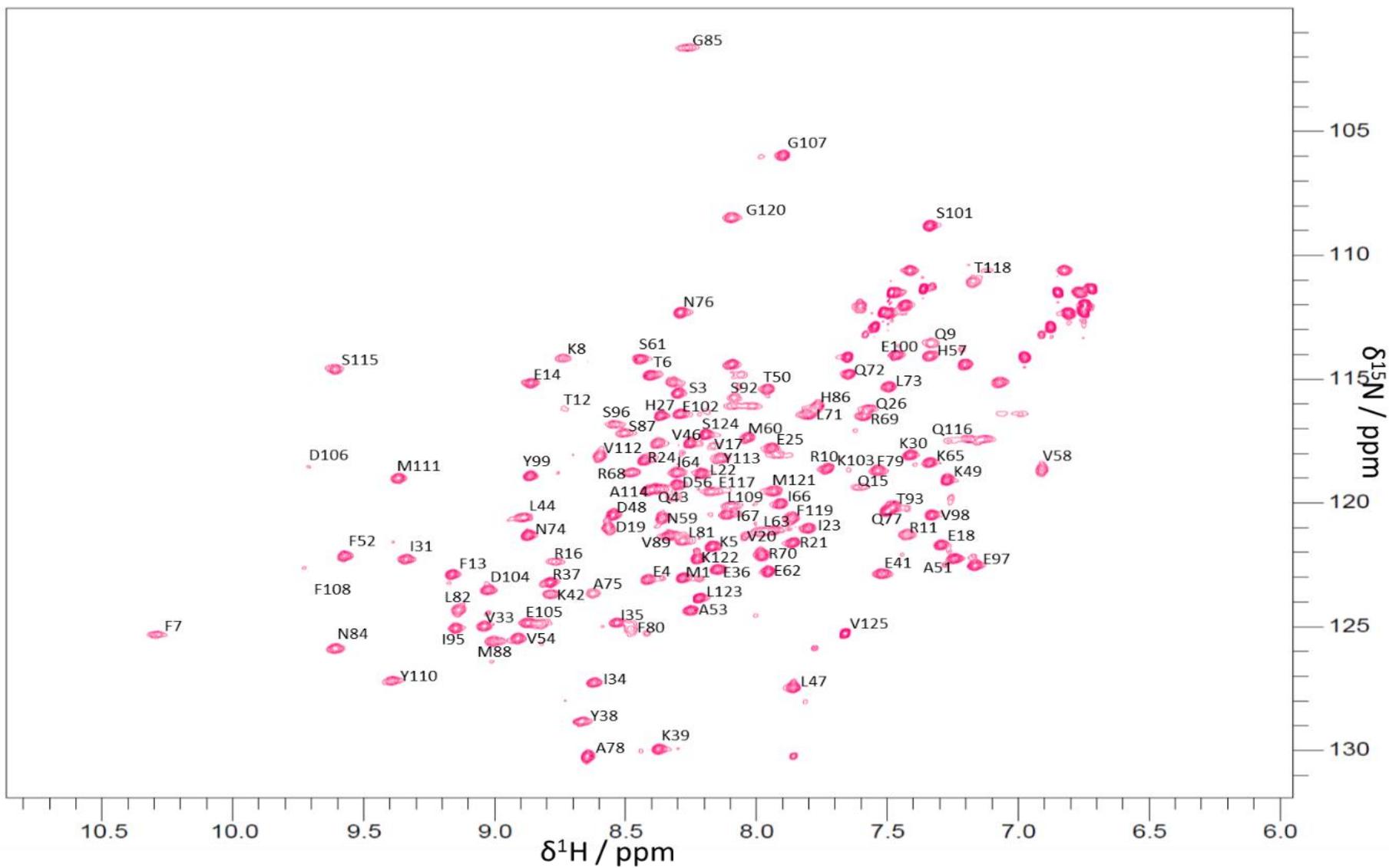


Figure 6.2: ^1H - ^{15}N HSQC backbone amide assignment of LC3B KL.

6.1.2 LC3B T29D Mutant Assignment

The LC3B T29D mutant, unlike the LC3B KL (and not unsurprisingly), showed very few changes in the ^1H - ^{15}N HSQC spectrum compared to LC3B (Figure 6.3), and so it was possible to transfer almost all of the assignment directly from the WT data to the T29D HSQC spectrum.

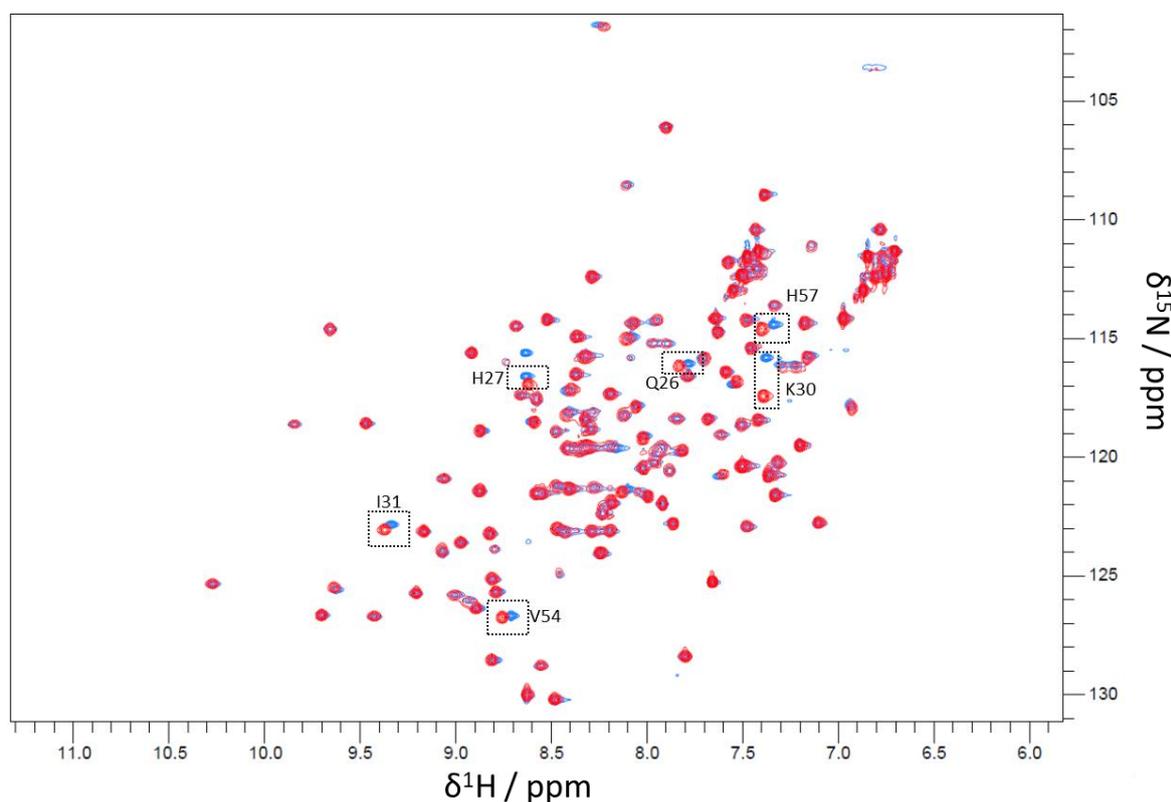


Figure 6.3: Overlay of ^1H - ^{15}N HSQC spectrum of LC3B WT (red) and LC3B T29D (blue).

There were a few peaks that had shifted in the HSQC spectrum (labelled in Figure 6.3), but apart from D29 these were easy to re-assign. The peak for D29 was in a totally new position from T29 in the WT spectrum, and so could not be assigned. This would not cause a problem though with the titrations as all other peaks were assigned, and so there would be enough information from the other peaks in the titration to understand the interaction with the LIRs.

6.2 NMR Titrations of LC3B Mutants with the LIRs of NBR1

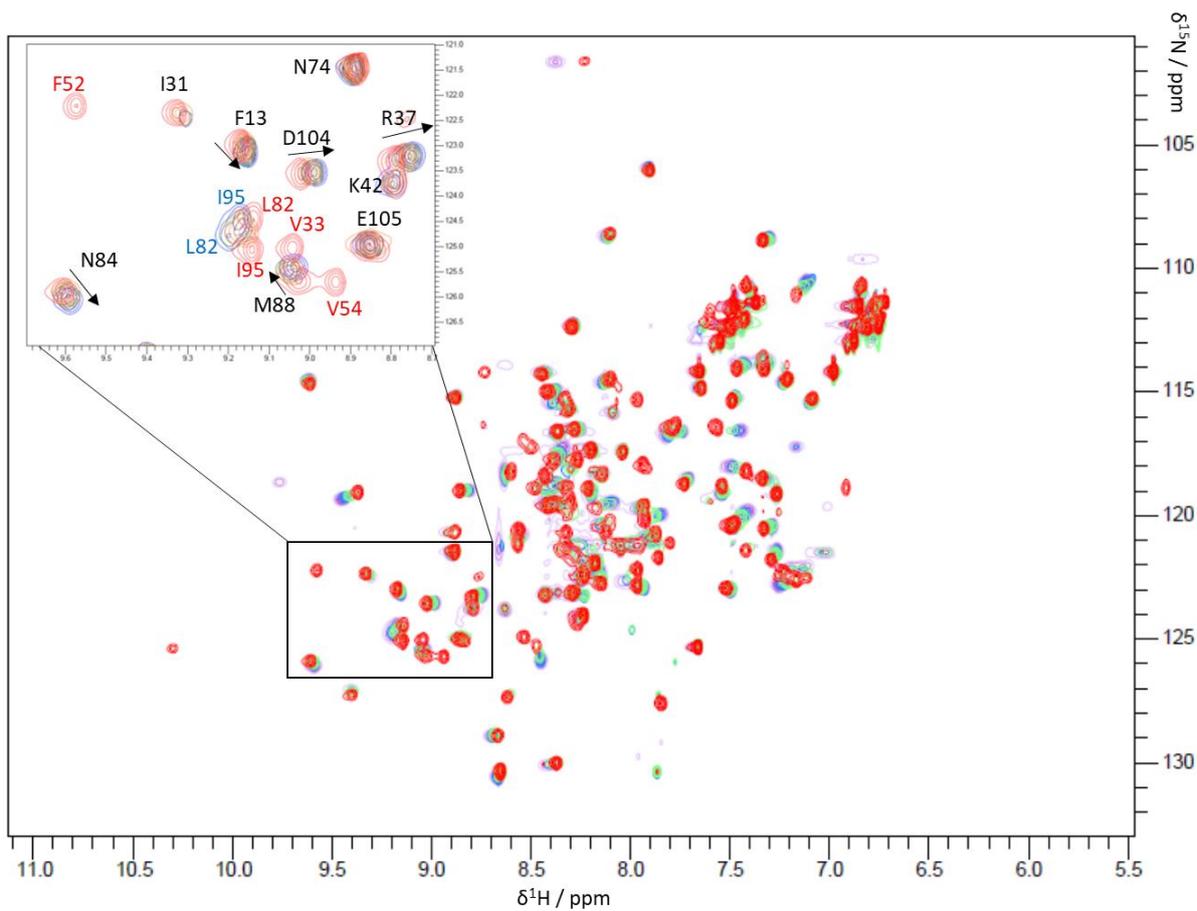
With the backbone amides in the ^1H - ^{15}N HSQC assigned for the majority of amino acids in the LC3B KL and T29D mutant, it was then possible to carry out the NMR titrations with the LIRs of NBR1. As previously these titrations were carried out using peptides representing the LIRs of NBR1, with the sequences ASSEDIILPE (residues 727-738 in NBR1), and AQDLLSFELLD (residues 557-567 in NBR1), used for LIR1 and LIR2 respectively. Samples were prepared using the method in Chapter 2.5.2.1, and a ^1H - ^{15}N HSQC run for each point in the titration series. Titration data was then analysed using the method in Chapter 2.5.2.5. Finally

CSPs for the titrations were generated and analysed using the method in Chapter 2.5.2.6, and HADDOCK structures for the binding interaction generated using the method in Chapter 2.5.2.7.

6.2.1 Titrations with KL Mutant

The NMR titration of the NBR1 LIRs with the LC3B KL mutant showed wide spread movement of peaks in the ^1H - ^{15}N HSQC between the start and end points (Figure 6.4).

A.



B.

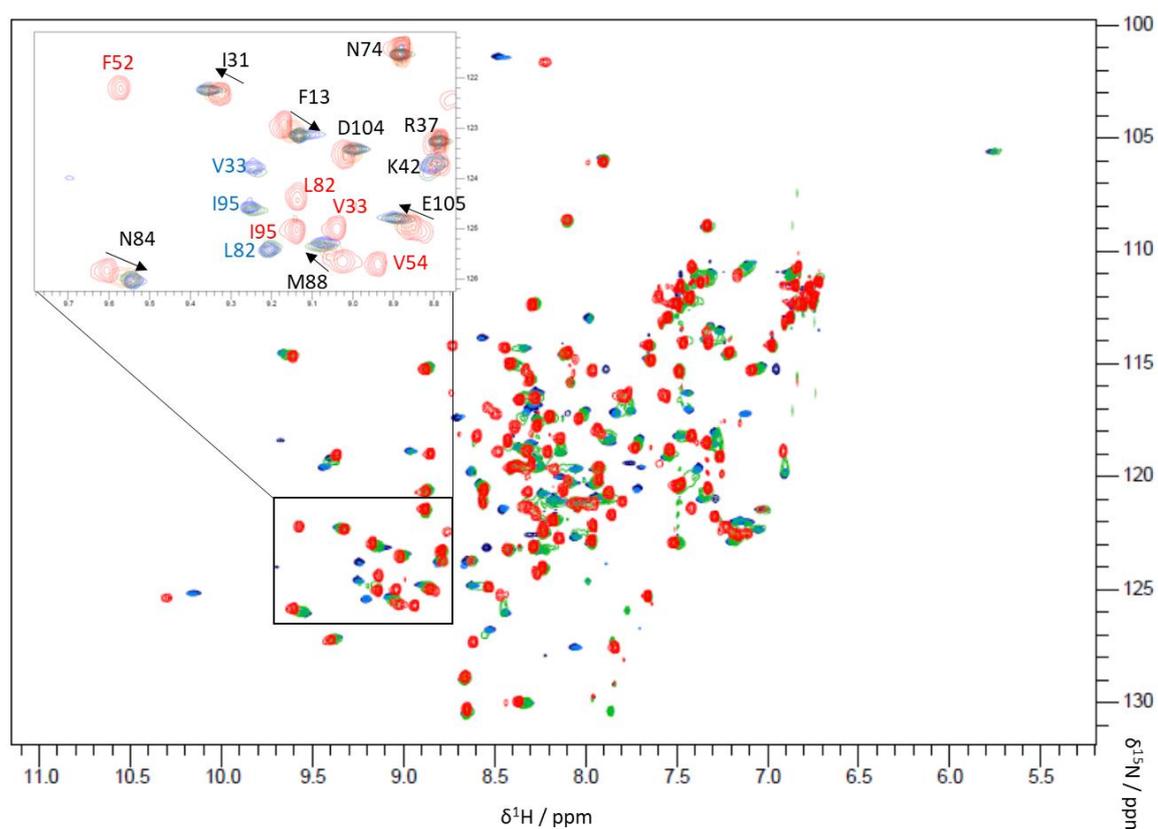


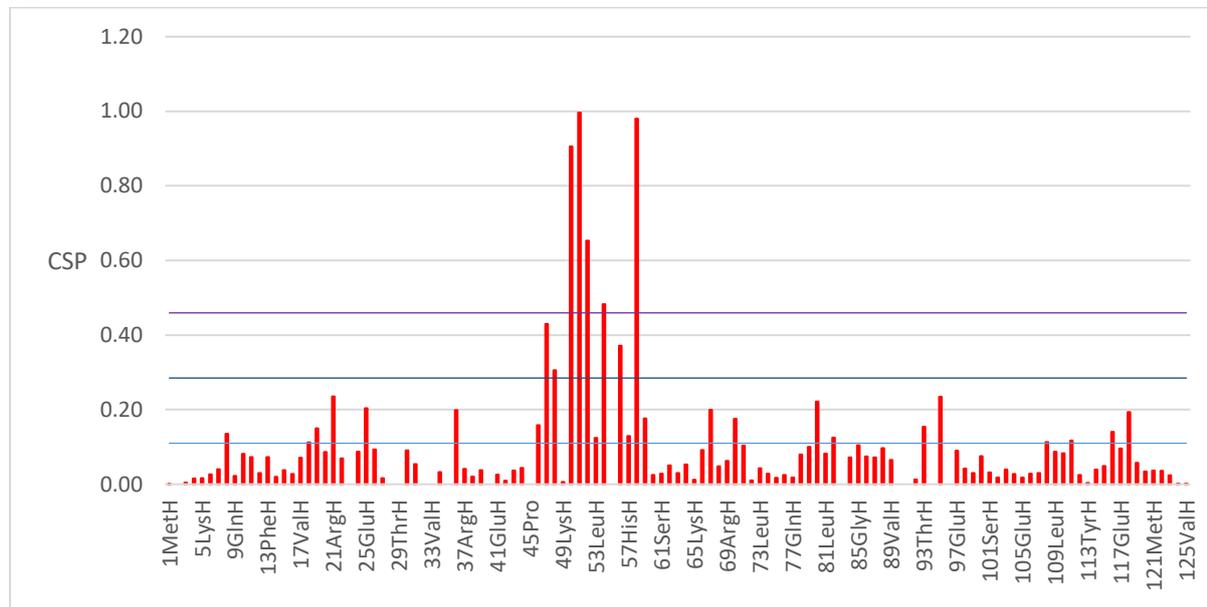
Figure 6.4: Overlay of a selection of ^1H - ^{15}N HSQC NMR spectra for the titration of LC3B KL with NBR1 LIR1 (A) and NBR1 LIR2 (B). Spectra shown are LC3B KL with zero (red), 0.5 (green), 1 (blue) and 2 (purple) equivalents of NBR1 LIR peptide.

As with the titrations with LC3B WT the peaks were in a mixture of fast, intermediate and slow exchange, with the majority of the peaks that had moved undergoing intermediate and slow exchange. Where possible the assignment was copied from the start to the end point by following the peaks in the titration. However, this was not possible for peaks in slow and intermediate exchange. To assign these peaks in slow and intermediate exchange, the assignment of the endpoint for the previous titrations with LC3B WT were used to assist in deciding which peaks belonged to which amino acid. For the titration of KL with LIR2, there were slightly fewer peaks in slow and intermediate exchange making it possible to transfer more of the assignments from the start to end point.

By doing this it was possible to assign 111 of the 119 possible (93%) for the LC3B KL with LIR1 titration, and LC3B KL with 115 of the 119 (97%) for the LIR2 titration. For the LIR1 titration, the peaks that it was not possible to assign were I23, T29, V33, I35, V83, S90, V91 and S96, and for LIR2 it was R24, T29, L53 and S90. Generally for the peaks where it was not possible to assign the endpoint it was due to the peak being in a congested area of the ^1H - ^{15}N HSQC spectrum, and so it was not possible to confidently identify a peak to transfer the assignment to. This was less of a problem for the LIR2 titration due to the peaks moving less. However, the majority of the peaks which were not possible to assign would be those expected to be around the binding site (and neighbouring amino acids), and so a large change in environment, and hence shift in peak position would not be unexpected.

The CSP values for both titrations were calculated and plotted with the significance cut offs (Figure 6.5).

A.



B.

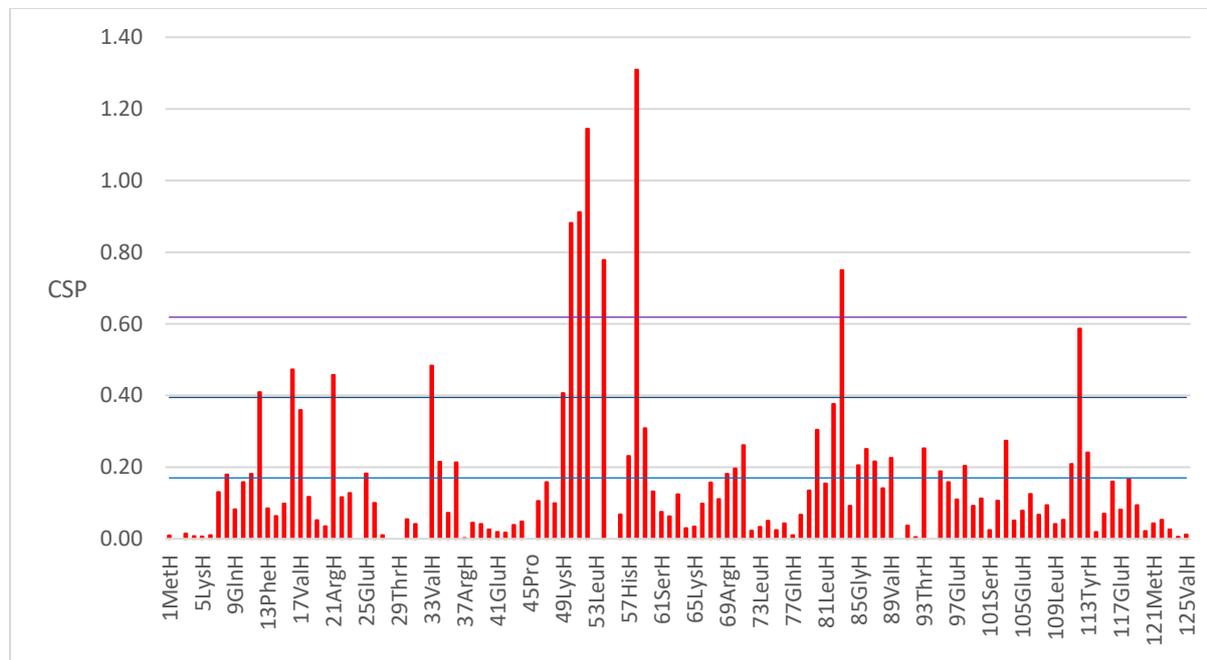


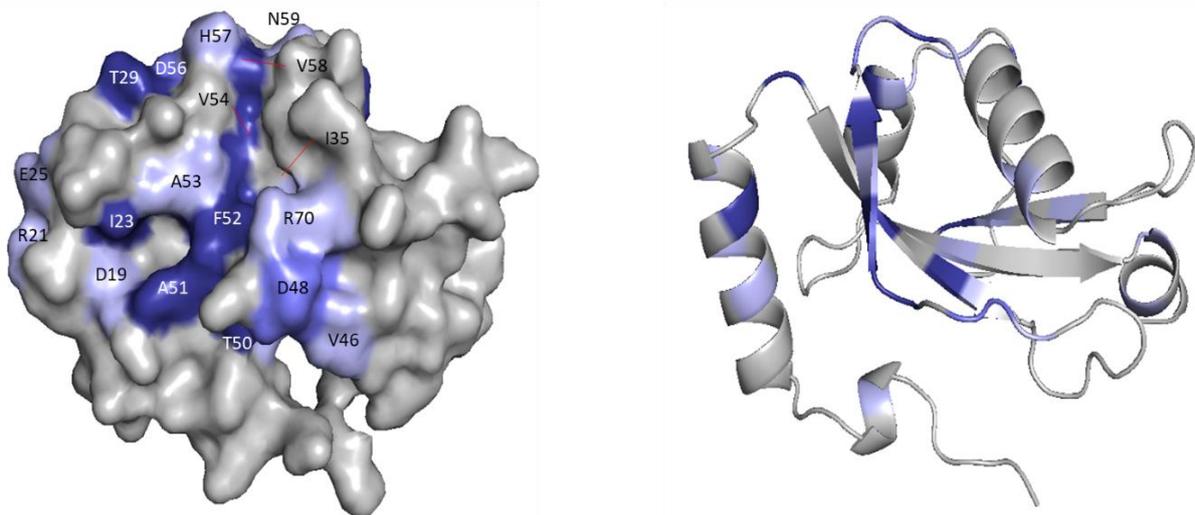
Figure 6.5: Graph of CSP values for binding of NBR1 LIR1 (A) and LIR2 (B) peptide to LC3B KL. The light blue line indicates the mean, the blue line is the mean and one standard deviation and the purple line the mean and two standard deviations.

As observed for LC3B WT there are a number of large perturbations for amino acids that form the hydrophobic pockets on the face of LC3B. However, the CSPs for these most significant amino acids are larger for the titration with LIR2, which in general has larger CSPs

across the board, which is reflected in the values for the significance cut offs for both titrations. For the titration with LIR1 there are 8 amino acids that are classed as significant, with T50, A51, F52, V54 and V58 classed as being most significant and L47, D48, D56 being classed as significant. For the titration with LIR2 titration there are 11 amino acids that are classed as significant, with T50, A51, F52, V54, V58 and V83 classed as being most significant and T12, R16, R21, V33, K49 and V112 being classed as significant. This analysis shows the titration with LIR1 has fewer significant than that of LIR2, and all of them are localised on the β -sheet of LC3B. However, taking into account the peaks that could not be followed in the titrations, and assuming that the peaks are significantly involved in the interaction of LC3B KL with LIR, makes the titrations more comparable with 16 and 15 amino acids that are playing a significant part in the interaction for LIR1 and LIR2, respectively.

To get a fuller understanding of the areas on LC3B KL that show the largest CSPs, and play a part in the interaction with the NBR1 LIR peptides, the CSPs were mapped onto the surface of LC3B KL (Figure 6.6)

A.



B.

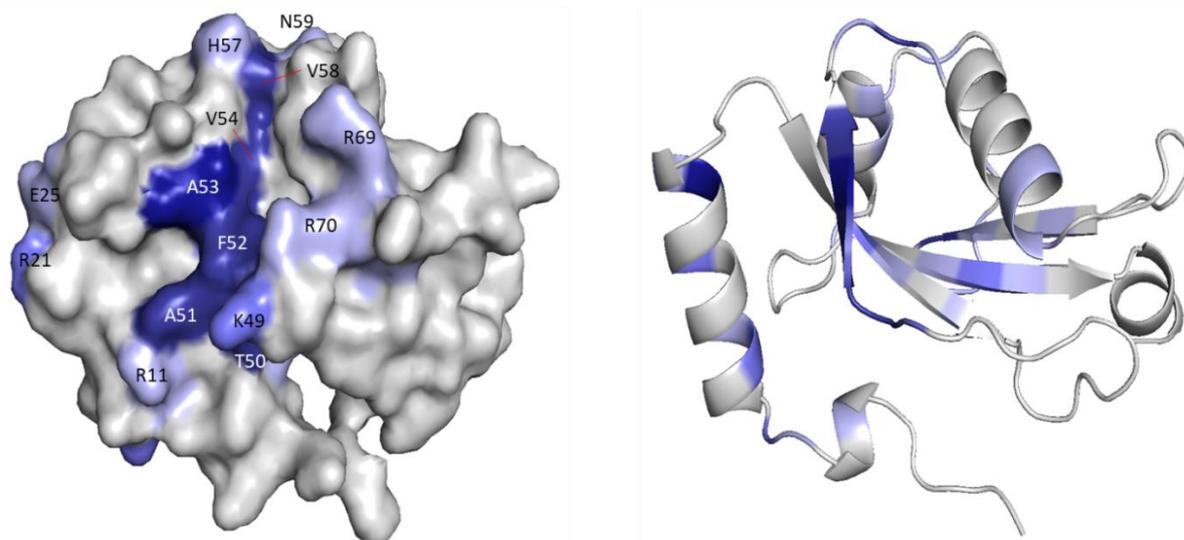


Figure 6.6: CSP mapping for the binding of NBR1 LIR1 (A) and LIR2 (B) peptide to LC3B KL mutant. Residues are coloured light purple if the CSP is greater than the mean, purple if greater than the mean and one standard deviation and blue if above the mean and two standard deviations. Residues with no significant shift are coloured grey.

The two CSP maps show similar areas of interaction for the LIR with LC3B KL. For LIR1 there are significant CSPs around the two hydrophobic pockets, with I23, A51 and F52 around HP1 where Tyr732 on the LIR peptide binds, and Val54 and Val58 round HP2 where Ile735 on the peptide binds. However, there are fewer significant shifts seen, especially on the α -helix that packs against the β -sheet when compared to the LC3B WT titration. There are also very few significant shifts in the hydrophobic pockets, suggesting that the tyrosine and Isoleucine on the peptide are not going as far into the hydrophobic pocket. This suggests that there is less change, and a potentially weaker binding interaction.

For LIR2 there are large perturbations around HP1, with A51, F52 and A53 undergoing significant changes. There are then also large shifts seen around HP2, with F52, A53, V54 and V58 showing significant changes. However, for LIR2 there are no significant changes seen on the far side of HP2. For both titrations there are the same large changes seen through the body of LC3B KL, where changes around the binding site cause changes in amino acids packed behind, as discussed above.

For the titration of NBR1 LIR1 with LC3B KL the HADDOCK run generated 184 structures in 3 clusters. The top scoring cluster contained 139 structures with a HADDOCK score of -96.4 ± 0.8 . The top scoring structure for this first cluster is given in Figure 6.7. The next highest scoring cluster only contained 39 structures with a HADDOCK score of -72.1 ± 2.2 , and so the structures in the first cluster were a much better fit than those in the second cluster.

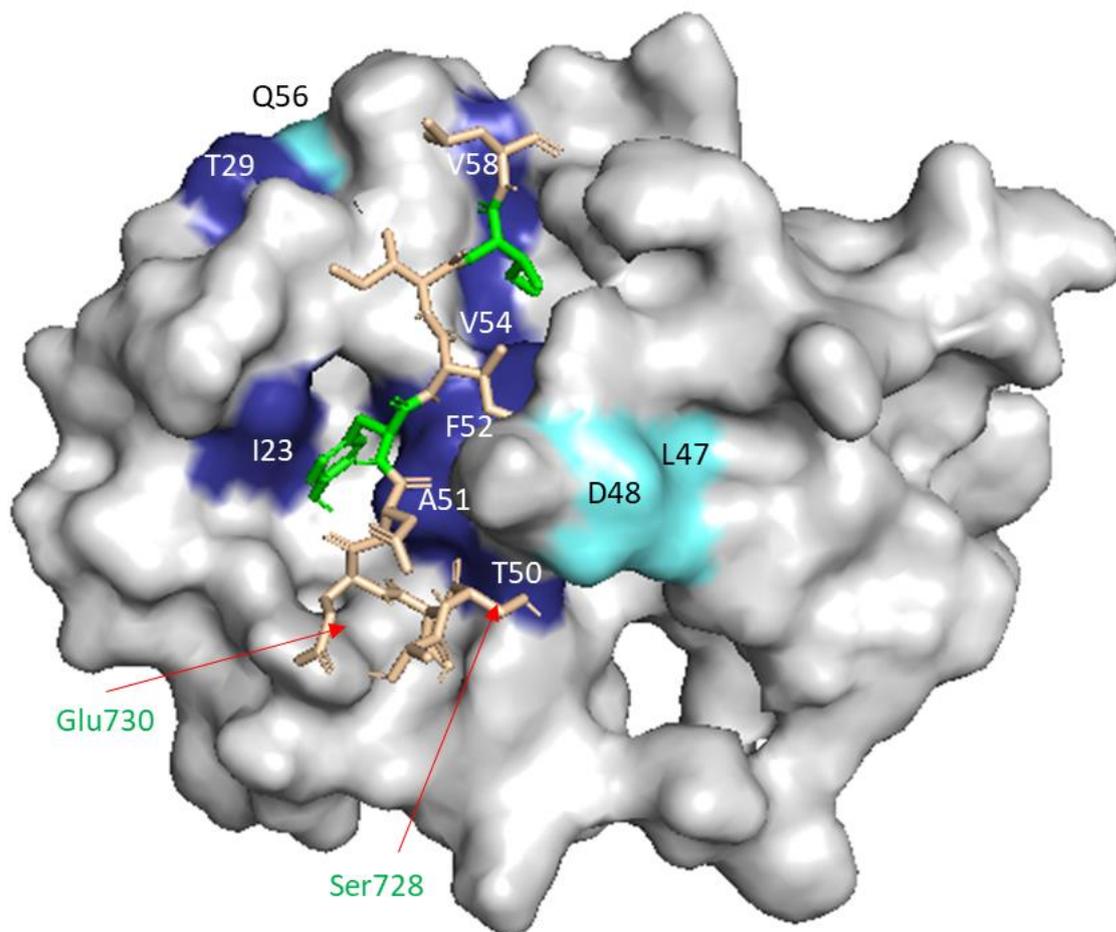


Figure 6.7: Highest scoring HADDOCK structure of NBR1 LIR1 peptide binding to LC3B KL. The peptide is coloured yellow with the active residues Tyr732 and Ile735 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.

The HADDOCK structure in Figure 6.7 shows the NBR1 LIR1 peptide sitting in standard LIR binding orientation, with Tyr732 and Ile735 sitting in HP1 and HP2 respectively. Tyr732 seems to be sitting fairly low in HP1, which is reflected in the lack of a significant shift on A51 on the KL mutant. Ser728 is then close to T50 on LC3B KL, where it is likely to be forming a hydrogen bond. Glu730 is sat above R11 where it can form an electrostatic interaction.

For LIR2 the HADDOCK run produced 191 structures grouped into 4 clusters. The top scoring cluster contained 76 structures with a HADDOCK score of -107.0 ± 6 . The top structure in this cluster is shown in Figure 6.8

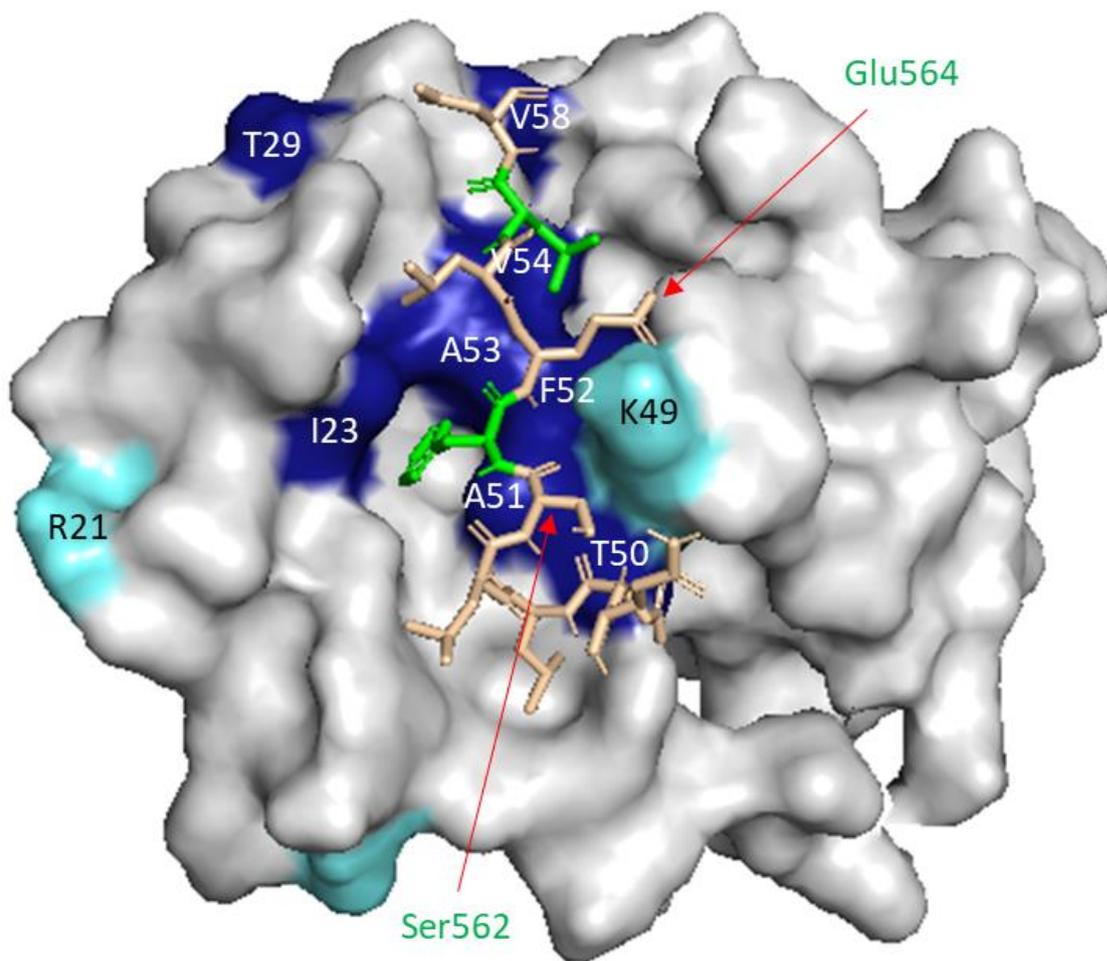


Figure 6.8: Highest scoring HADDOCK structure of NBR1 LIR2 peptide binding to LC3B KL. The peptide is coloured yellow with the active residues Phe563 and Leu566 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.

This structure shows the LIR2 binding to the front face of the KL mutant at the canonical binding site, with Phe563 sitting in HP1 surrounded by the amino acids with significant CSPs A52, F52 and A53. Leu566 then sits in HP2. There are also possible interactions between Ser562 and T50 and Glu564 and K49.

However, the HADDOCK run produced another cluster that contained 102 structures and a HADDOCK score of -91.2 ± 3.8 . This was a larger cluster size than the cluster showing the LIR in the standard configuration, and its HADDOCK score was close enough to the highest scoring cluster to warrant further study. The top scoring structure in the cluster is shown below (Figure 6.9).

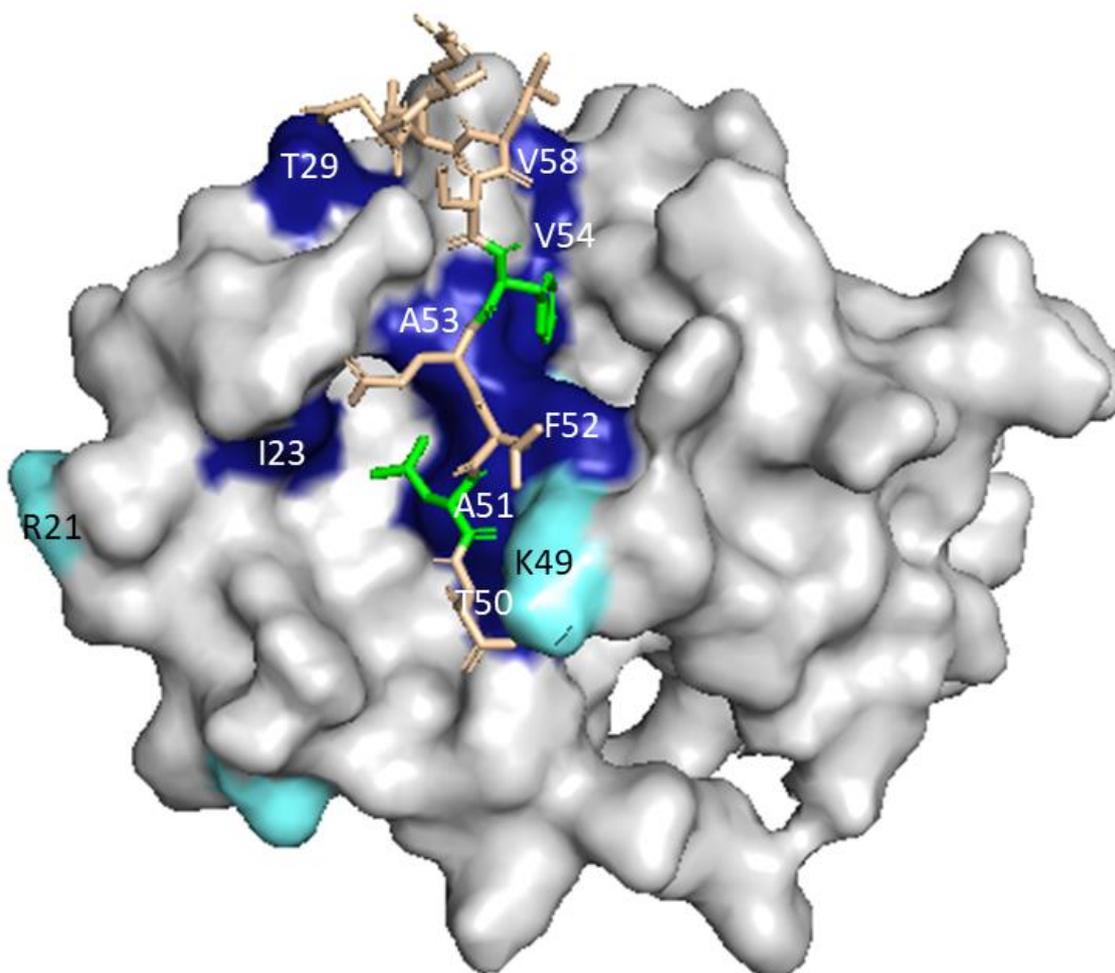


Figure 6.9: Highest scoring HADDOCK structure from second cluster of NBR1 LIR2 peptide binding to LC3B KL. The peptide is coloured yellow with the active residues Phe563 and Leu566 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.

This structure shows the LIR in a reverse configuration, with Leu566 on the LIR binding with HP1 and Phe563 binding in HP2. This goes counter to the titration data, which gave the expected set of amino acids that showed significant perturbations in the NMR data, but does agree with the result found for the LC3B WT LIR2 titration.

To further probe the cause of this reverse configuration being produced with such a high (and close to the standard orientation) HADDOCK score, a series of HADDOCK runs were carried out. These included a run where the WT CSPs were applied to the KL mutant structure, where the KL CSPs were applied to the WT structure, where only a set of core CSPs were used and finally a set of “select” amino acids that sit central on the binding patch of LC3B KL were used as active residues. Details of these runs and their results are given in Table 6.1.

| Run | Titration Data | WT CSPs with KL | KL CSPs with WT | Core CSPs | Select CSPs |
|---|---------------------|---------------------|---------------------|---|---------------------------|
| CSPs used | KL + LIR2 titration | WT + LIR2 titration | KL + LIR2 titration | Active = 50,51,52,53,54,58 Passive = 33,49 | 51,52,53,54 (active only) |
| Structure Used | LC3B KL | LC3B KL | LC3B WT | LC3B KL | LC3B KL |
| Top scoring cluster (number of structures) | -107.0±6 (76/191) | -101.9±4.8 (73/168) | -103.0±2.2 (96/187) | -99.7±0.5 (67/184) | -103.3±2.6 (103/185) |
| Top scoring LIR orientation | Standard | Standard | Standard | Standard | Standard |
| Next highest scoring cluster (number of structures) | -91.2±3.8 (102/191) | -96.4±2.1 (63/168) | -91.9±1.7 (85/187) | -93.2±1.9 (95/184) | -91.8±1.0 (27/185) |
| Next highest scoring LIR orientation | Reverse | Reverse | Reverse | Reverse | Reverse |

Table 6.1: Parameters and results of HADDOCK runs carried out to investigate reverse configuration LIR for LC3B KL titration with NBR1 LIR2 peptide.

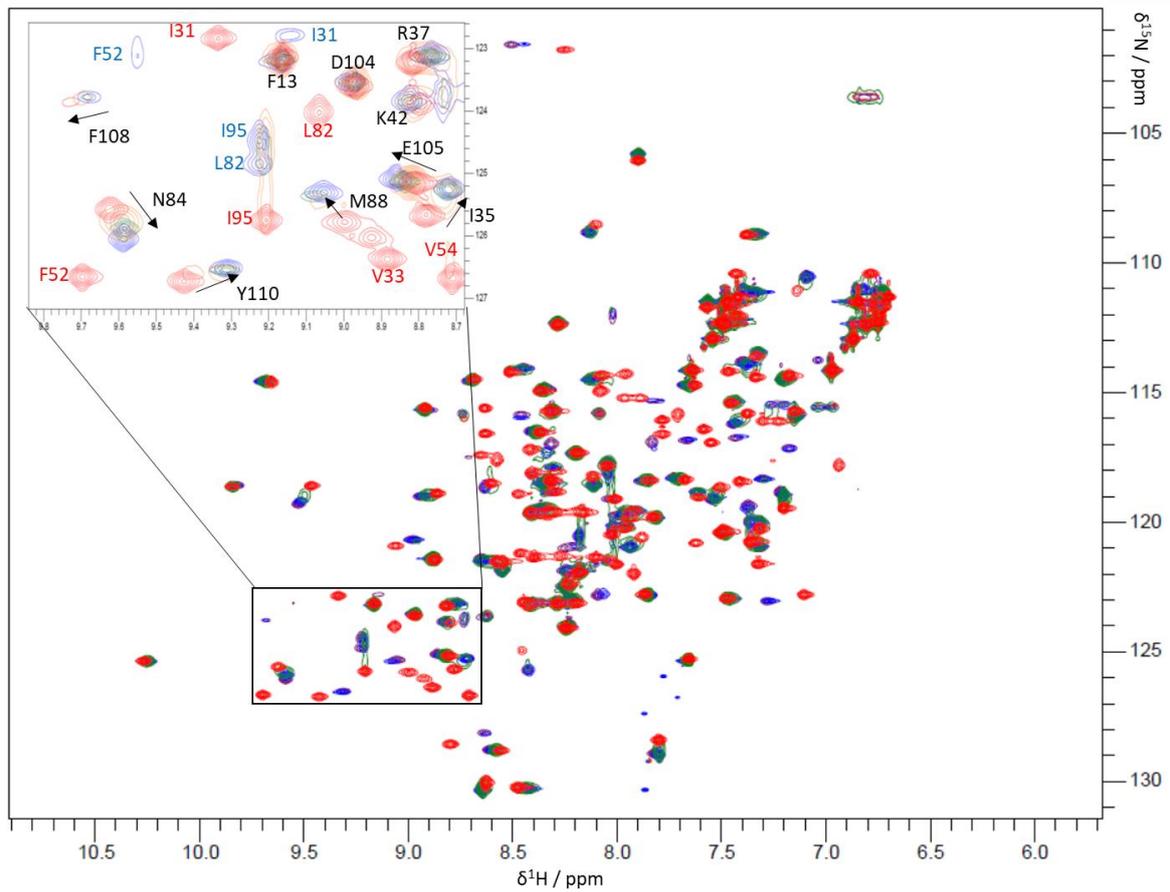
The run results show that for all the HADDOCK runs carried out the standard LIR configuration is present as the top scoring set of structures, and that the reverse LIR configuration is then the second set of highest scoring structures. The highest scoring structure, and the biggest difference between the standard and reverse LIR configuration are seen for the actual KL mutant with NBR1 LIR2 run data. The lowest score for the standard run configuration is for the set of core CSPs and this also has the smallest difference between standard and reverse configuration. However, there are not big differences in HADDOCK score between all the runs.

This suggests that the result is not due to an issue with the CSPs, but with the change in structure of the KL mutant from the WT, which combined with the LIR2 peptide, makes the LIR interaction with LC3B KL in a reverse configuration relatively high scoring in HADDOCK.

6.2.2 Titrations with T29D Mutant

As with previous titrations run with wild type LC3B and the KL mutant, a large number of peaks had moved over the course of the titration for the LIRs of NBR1 with LC3B T29D (Figure 6.10).

A.



B.

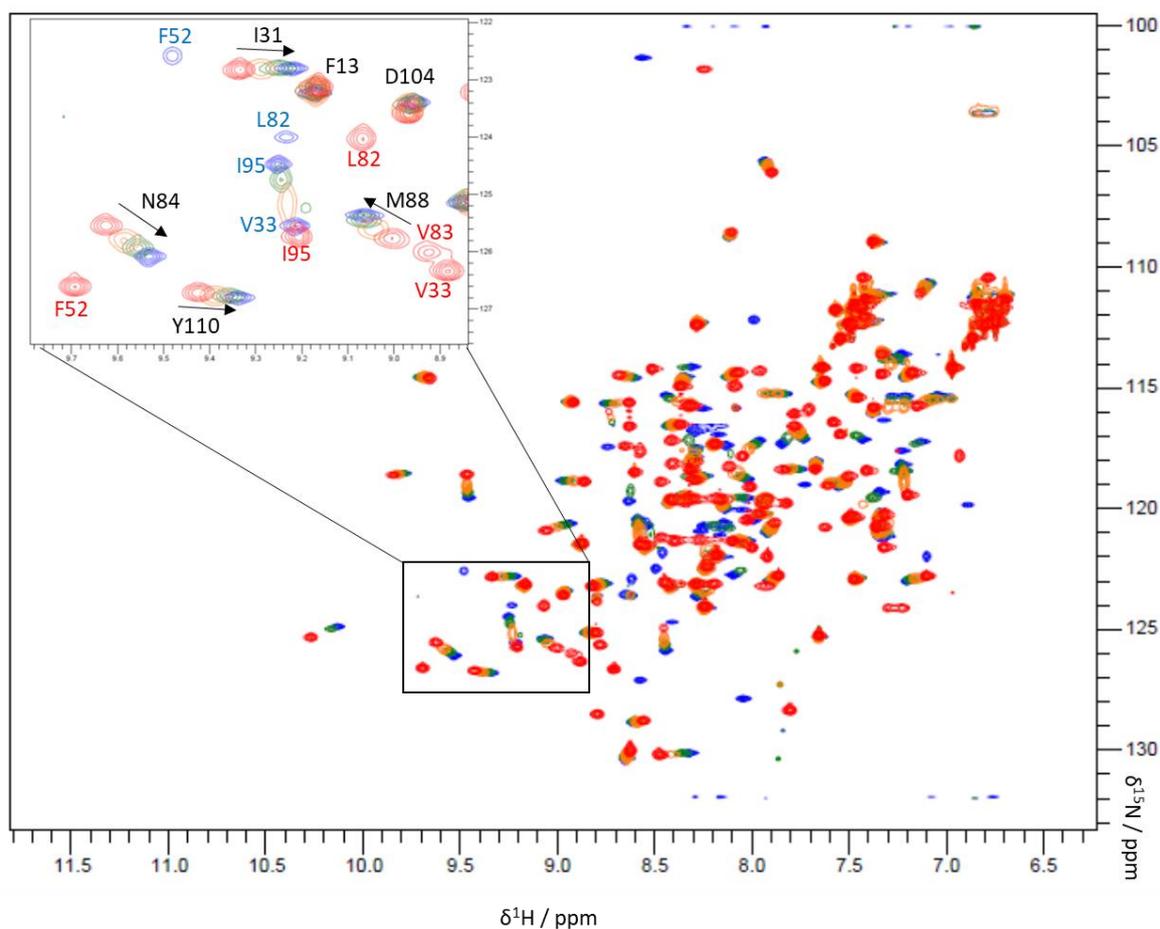
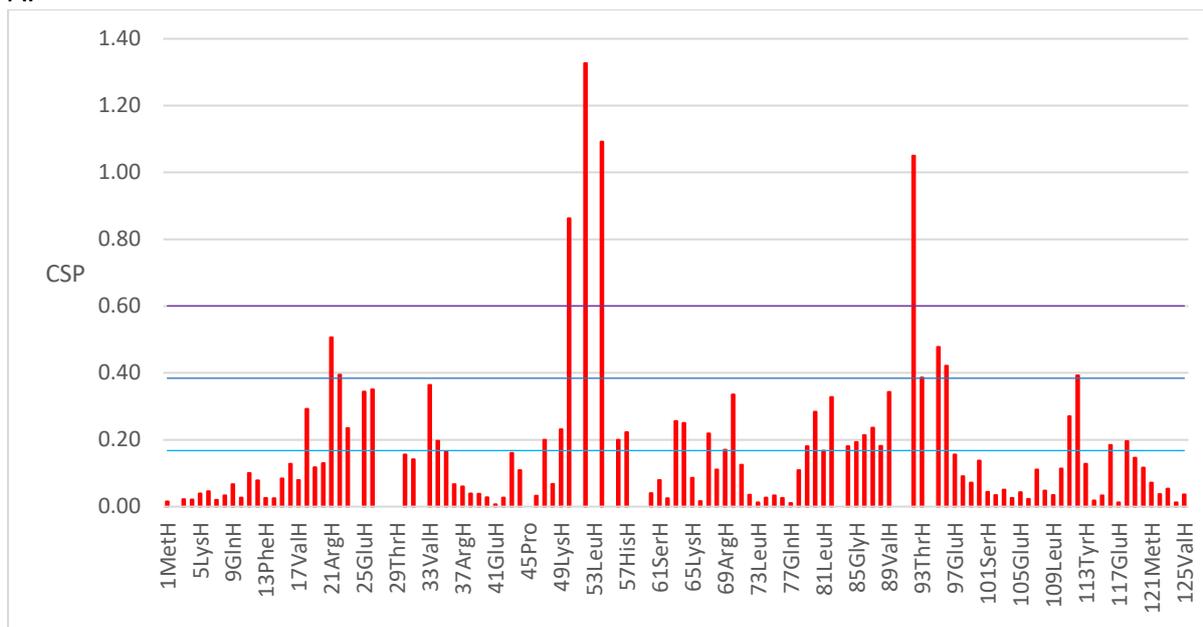


Figure 6.10: Overlay of a selection of ^1H - ^{15}N HSQC NMR spectra for the titration of LC3B T29D with NBR1 LIR1 (A) and LIR2 (B) peptide. Spectra shown are LC3B T29D with zero (red), 0.5 (green), 1 (blue) and 2 (purple) equivalents of NBR1 LIR1 peptide.

The peaks that had moved in the ^1H - ^{15}N HSQC spectrum of the T29D mutant during the titration with the two NBR1 LIR1 titrations did so under a mixture of exchange regimes, with peaks showing fast, intermediate and slow exchanges. As previously, there were a large number of peaks that underwent slow and intermediate exchange, leading to problems with assigning peaks in the spectrum. Where possible the endpoint ^1H - ^{15}N HSQC was assigned by following the peaks as the amount of LIR increased. However, this was not always possible, and the data from previous titrations was then used to help assign those peaks it was not possible to follow in the titration.

This allowed 110 out of 119 (92%) to be assigned for LIR1, and 111 of 119 (93%) for LIR2. For the LIR1 titration the peaks that could not be assigned were R24, H27, K51, L53, V58, Q59, V83, S90 and V91, and for LIR2 they were V20, L22, R24, H27, L63, I66, V91 and S96. Although these peaks could not be followed in the titrations, it is likely that they are significantly involved in the titration. The calculated CSPs for the two LIR titrations with T29D with the significance cut offs were then plotted (Figure 6.11).

A.



B.

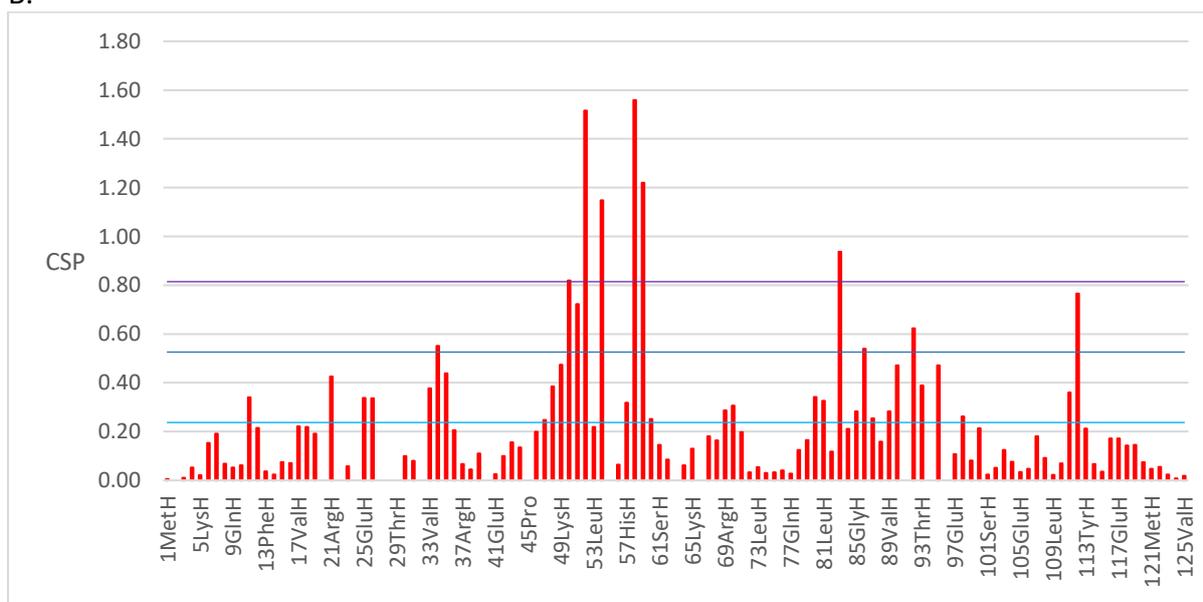


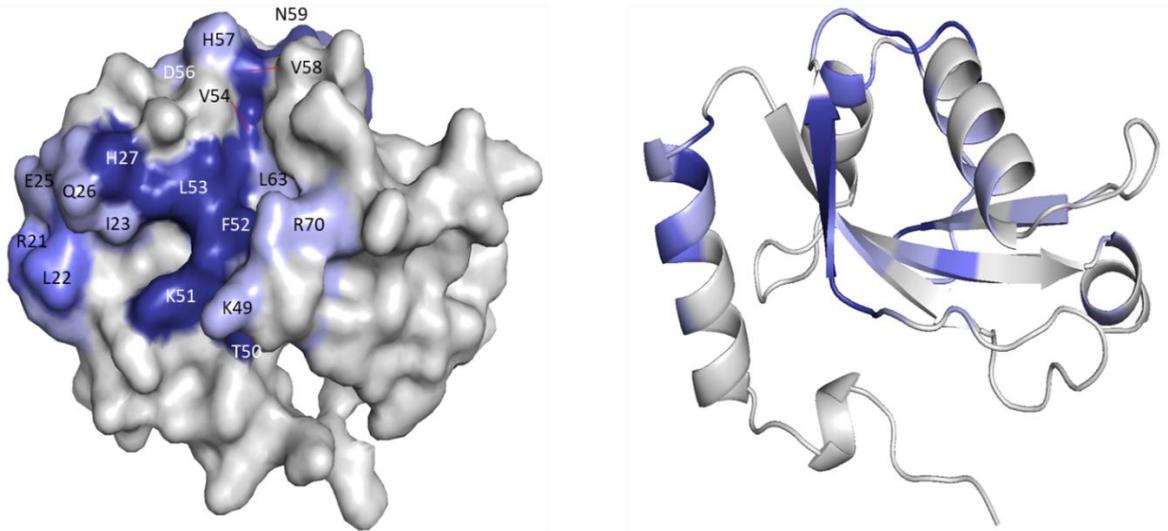
Figure 6.11: Graph of CSP values for binding of NBR1 LIR1 (A) and LIR2 (B) peptide to LC3B T29D. The light blue line indicates the mean, the blue line is the mean and one standard deviation and the purple line the mean and two standard deviations.

The CSP plots showed a number of amino acids with significant values around the canonical binding site. For the titration with LIR1 there are 10 amino acids that are classed as significant, with T50, F52, V54 and S92 classed as being most significant, and R21, L22, T93, I95, S96, V112 being classed as significant. For the titration with LIR2 there are 11 amino acids that are classed as significant, with T50, F52, V54, V58, N59 and V83 classed as being most significant, and I34, K51, H86, S92 and V112 being classed as significant. Both titrations have a similar number of significant amino acids, with LIR2 having slightly larger CSP values. Taking into account the peaks that could not be followed in the titrations, and assuming that

the peaks are significantly involved in the interaction of LC3B KL, means that both titrations of LIR with T29D have 19 amino acids that have significant changes when binding.

Using the amino acids with significant CSPs to colour LC3B T29D allowed the mapping of the binding site for the two LIRs of NBR1 (Figure 6.12).

A.



B.

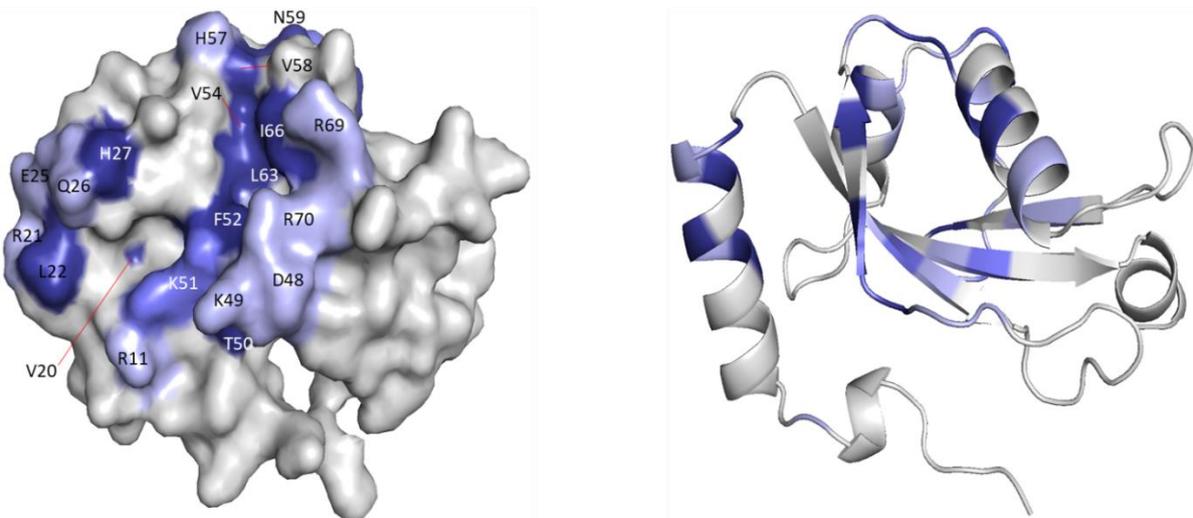


Figure 6.12: CSP mapping for the binding of NBR1 LIR1 (A) and LIR2 (B) peptide to LC3B T29D mutant. Residues are coloured light purple if the CSP is greater than the mean, purple if greater than the mean and one standard deviation and blue if above the mean and two standard deviations. Residues with no significant shift are coloured grey.

The mapped binding site for the interaction of NBR1 LIR1 and T29D follows that of previous titrations, with large changes seen around HP1 on K51, F52 and L53, and around HP2 on V54, V58 and L63. However, for LIR2, whilst seeing large perturbations on F52, V54, V58, L63 and I66 around HP2, there are fewer amino acids with significant changes around HP1. K51 is not at the highest level of significance like in previous titrations, and L53 is not classed as significant at all. To better understand these results the HADDOCK structures were analysed.

For NBR1 LIR1 the HADDOCK run produced 189 structures in 3 clusters. The top scoring cluster had a HADDOCK score of -95.5 ± 4.1 and contained 176 structures. The next best scoring cluster only contained seven structures, and scored -52.8 ± 2.6 , suggesting with high confidence that the top scoring structure was the best fit for the data from the HADDOCK simulation. The top scoring HADDOCK structure is shown in Figure 6.13

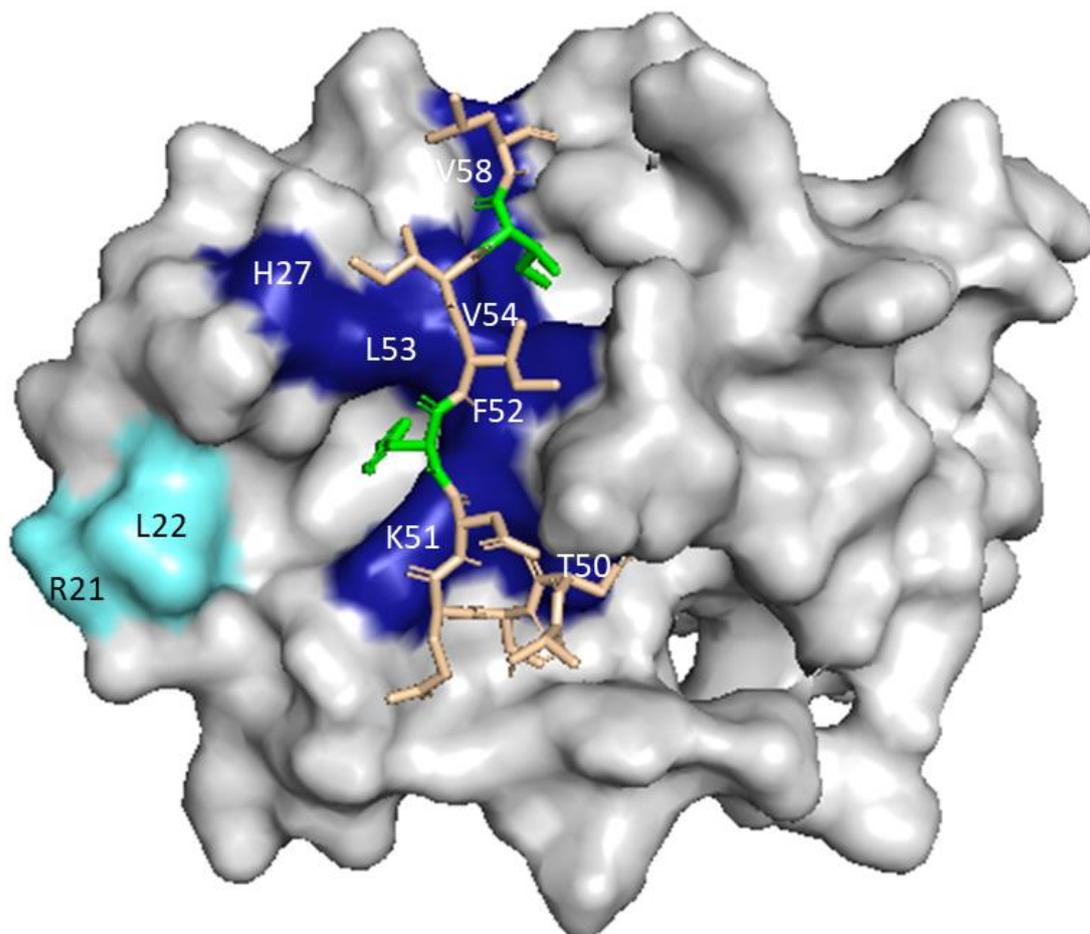


Figure 6.13: Highest scoring HADDOCK structure of NBR1 LIR1 peptide binding to LC3B T29D. The peptide is coloured yellow with the active residues Tyr732 and Ile735 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.

From this HADDOCK structure we can see that that NBR1 LIR1 is binding with LC3B T29D at the canonical LC3B binding site. Tyr732 on the LIR is inserted into HP1 and Ile735 is sitting in HP2. Tyr732 sits in the centre of HP1 pointing straight down. Ser728 then sits next to T50 on T29D, potentially forming a hydrogen bond. Glu730 and Asp731 on LIR2 then point towards R11 and K49, respectively, forming a potential electrostatic interaction. Whilst the CSPs for T29D on R11 and K49 were not classed as significant, they were still large (being above the mean) so the interaction is still possible, but the environment of R11 and K49 was not changed such to warrant a significant CSP.

For LIR2 the HADDOCK run created 193 structures in only two clusters. The top scoring cluster had a HADDOCK score of -98.2 ± 0.6 and contained 133 structures. The top structure for this cluster is shown in Figure 6.14.

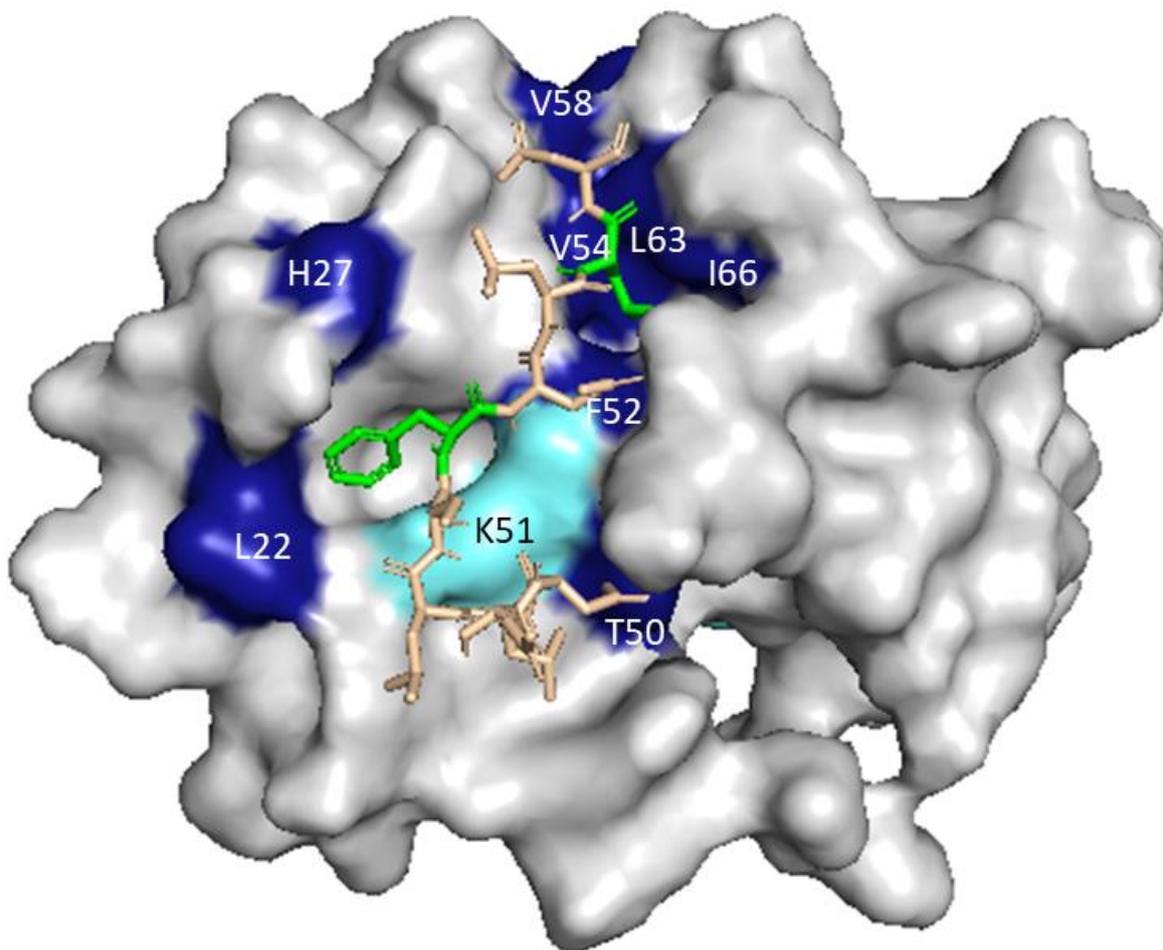


Figure 6.14: Highest scoring HADDOCK structure of NBR1 LIR2 peptide binding to LC3B T29D. The peptide is coloured yellow with the active residues Phe563 and Leu566 coloured green. The surface of LC3B is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured dark blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.

From this HADDOCK structure we can see that the NBR1 LIR2 peptide is binding to LC3B T29D at the canonical binding site. Phe563 is interacting with HP1, and Leu566 is sat in HP2. However, rather than sitting in HP1, Phe563 is sitting across the pocket. This configuration accounts for the fact there are large perturbations seen on L22 and H27, and much smaller changes on K51 and L53. With Phe563 sat so far across the face of LC3B T29D it prevents Glu564 interacting with K49, and Ser562 interacting with T50, as has been seen previously with LIR2. Asp559 can also not interact with K49, and instead is located close to T50.

The second highest scoring HADDOCK structure did score -89.8 ± 3.5 , which whilst being relatively high scoring only contained half the structures of cluster one, and the structure shown was not the “reverse” LIR configuration seen with LC3B WT and KL, but instead it

showed just Phe563 in HP2. From the titration CSPs this would not account for the perturbations seen around HP1 and so will not be explored further.

6.3 Discussion

6.3.1 Assigning the KL and T29D Mutant

Both mutants of LC3B had differences in the ^1H - ^{15}N HSQC when compared with the LC3B WT spectrum. This is not unexpected due to altering the structure of LC3B. However, as only two amino acids were being changed for the KL mutant, and one for the T29D mutant, the differences would not be expected to be large. Whilst this was true for the T29D mutant (only a few peaks in the HSQC had shifted), this was not the case for the KL mutant.

For the KL mutant, as mentioned earlier in the chapter, the majority of peaks had moved in the ^1H - ^{15}N HSQC when compared with the wild type spectrum. However, the majority had not moved by much, and a lot of the difficulty of transferring the assignment had been when there were clusters of peaks and not knowing which one was which.

To check how significant the peak movement was, and to check that for the majority of peaks it was just due to some structural readjustment, CSPs were calculated for the peaks between the wild type and KL mutant ^1H - ^{15}N HSQC and the significance cut offs were calculated (Chapter 2.5.2.6). The graph of the CSPs with the significance lines is shown in Figure 6.15.

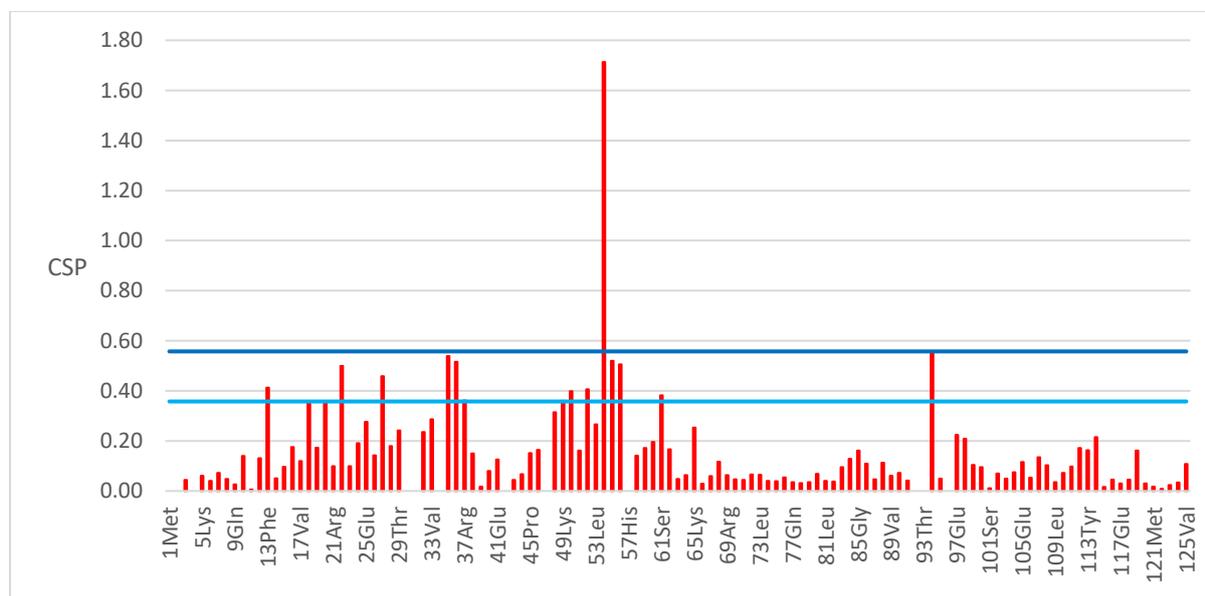


Figure 6.15: Graph of CSPs calculated between LC3B WT and KL ^1H - ^{15}N HSQC spectrum. Lines represent significance cut offs and the values are mean + one standard deviation (light blue) and mean + two standard deviations (dark blue).

The most significant peak to shift in the ^1H - ^{15}N HSQC between the wild type and KL mutant spectra is that for Phe52. This is not a surprise as it is located between the two mutations on the face of LC3B, and this would be expected to have a large change in environment. The rest of the changes were less significant and are mainly located around the mutations. To

understand these changes better these significant changes were mapped onto a surface of LC3B (Figure 6.16)

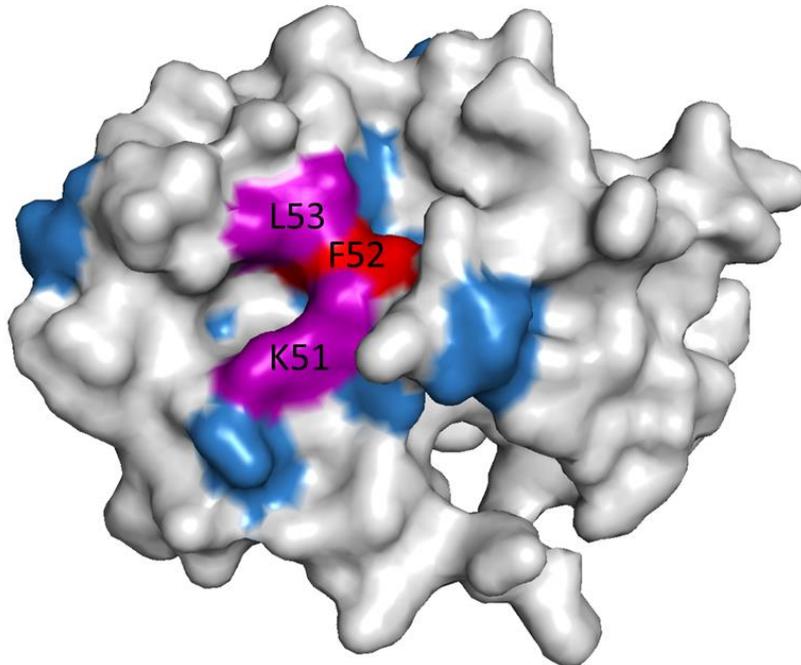


Figure 6.16: Significant CSP values mapped onto surface of LC3B. KL mutation shown in purple, significant shift shown in blue and most significant in red.

Looking at the surface in Figure 6.16 the majority of the significant shifts occur around the mutations. However, there are a small number that appear at the top of the α -helix packed against the binding site (Val20 and Glu25). However, these are quite close to the mutations internally, which explains the relatively large shift in the ^1H - ^{15}N HSQC.

For LC3B T29D there were 6 peaks in the ^1H - ^{15}N HSQC that had shifted when compared to the WT spectrum. When mapped onto the surface of LC3B T29D (Figure 6.17), we can see that the main changes in the HSQC between wild type LC3B and the T29D mutant are next to the mutation. However, there are some peak shifts seen on V54 and H57, which are near to the mutation in the 3D structure of LC3B. These two amino acids border HP2, and in previous titrations have had significant CSPs.

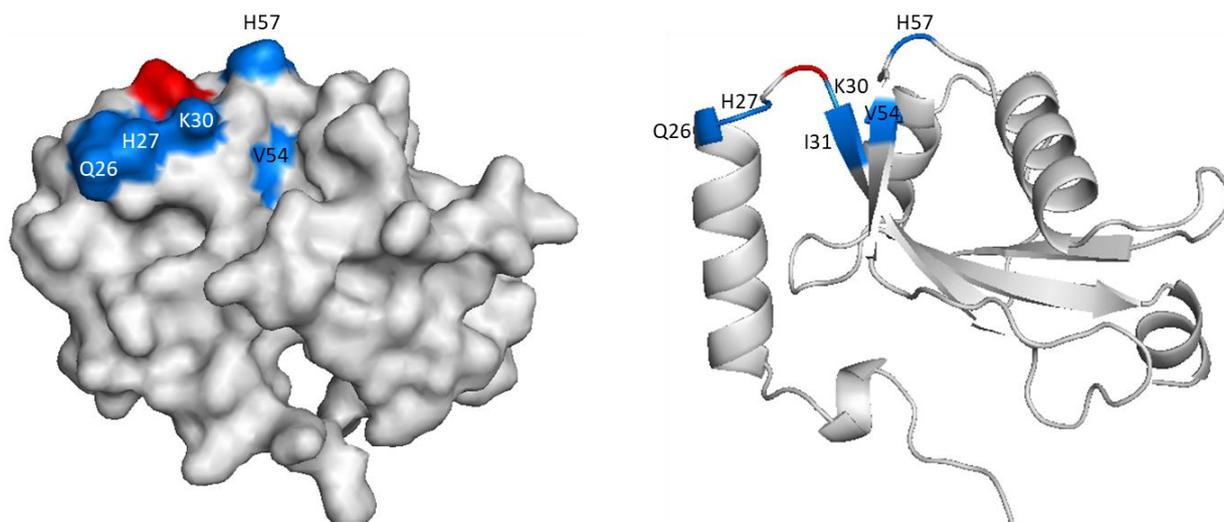


Figure 6.17: Changes in the HSQC between LC3B WT and LC3B T29D mapped onto the surface of LC3B. Mutation site shown in red.

6.3.2 Titrations of the NBR1 LIRs with the LC3B Mutants

As with the previous titrations of LC3B WT with the LIR peptides, the peaks in the ^1H - ^{15}N HSQC underwent widespread changes between the start and endpoints of the titration with the LC3B mutants.

These global perturbations across the whole of LC3B mutants were similar to those seen with the wild type, and required further analysis of the CSP data to focus on those amino acids playing a large part in the interaction with the LIRs.

With the most significant amino acids selected, all the titrations of the LC3B mutants with LIRs showed that the most significant changes were seen around HP1, where Tyr732 (LIR1) and Phe563 (LIR2) interact, and HP2 where Ile735 (LIR1) and Leu566 (LIR2) interact. To further analyse similarities in the binding site on the LC3B wild type and its mutants, the amino acids with significant CSPs are shown in Table 6.2.

| Titration | KL-LIR1 | KL-LIR2 | T29D-LIR1 | T29D-LIR2 | WT-LIR1 | WT-LIR2 |
|--|--|---|---|--|---|--|
| Amino acids with most significant CSP | I23, V33, I35, T50, A51, F52, V54, V58, V83, S90, V91, S96 | T50, A51, F52, A53, V54, V58, V83, S90, | R24, H27, T50, K51, F52, L53, V54, V58, N59, V83, S90, V91, S92 | V20, L22, R24, H27, T50, F52, V54, V58, N59, L63, I66, V83, V91, S96 | R24, H27, T50, F52, V54, V58, N59, V91, S92 | R24, H27, T50, F52, V54, V58, N59, V91 |
| Amino acids with significant CSP | L47, D48, D56 | T12, R16, R21, V33, K49, V112 | R21, L22, T93, I95, S96, V112 | I34, K51, H86, S92, V112 | R21, K51, L53, V83 | L22, K51, L53, D56, V58, N59, V83, S92, V112 |

Table 6.2: Amino acids with significant CSPs for titrations of LC3B mutants with NBR1 LIR peptides. Also included are LC3B significant amino acids for reference. Those highlighted green appear as a most significant CSP in all titrations. Those highlighted yellow appear as any kind of significant shift in all titrations.

Looking at the list of significant amino acids we can see that there are four amino acids that appear in every titration with the NBR1 LIR peptides. These are T50, F52, V54 and V58, which form a central core on the front face of LC3B, forming the sides of hydrophobic pocket one and two. As well as these four amino acids that have most significant CSPs in all the titrations, there is also K51 (A51 in KL mutant) and V83 that are in all six titrations as either a significant, or most significant CSP. K/A51 sits around HP1, and V83 although not on the surface of LC3B, sits on the beta sheet behind HP1. To visualise the amino acids position on LC3B we can map the significant shifts seen in the titrations onto the structure of LC3B (Figure 6.18).

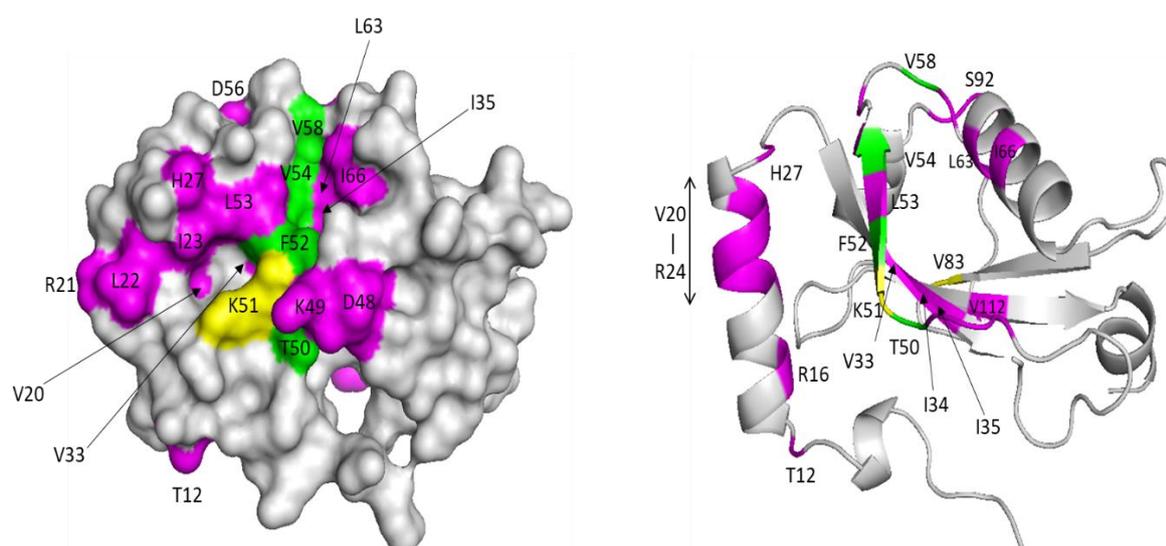


Figure 6.18: Significant CSPs for titrations of LC3B with NBR1 LIR1 titrations mapped on to LC3B WT structure. Those coloured green appear as most significant amino acids in all titrations. Those coloured yellow are significant or most significant in all titrations. Those coloured purple appear in at least one titration.

From Figure 6.18 we can see that the areas experiencing largest perturbations on LC3B WT and the mutants are all localised in the same areas. As well as these six amino acids that appear in all the titrations there are also areas of large perturbation that appear in all the titrations. There are significant shifts seen on the alpha helix (V20, R21, L22, I23, R24 and H27) sitting alongside the central beta sheet which forms HP1. These are also large perturbations seen in most of the titrations on the top of LC3B, next to HP1 pockets (S90, V91, S92, T93 and S96). There are then other amino acids which have significant CSPs, that appear in some of the titrations, that sit inside or around the hydrophobic pockets (V33, I35, L63 and I35), and depend upon how the LIR interacts with the hydrophobic pockets. In general though the amino acids that are experiencing the most significant perturbations are located around the canonical binding site, and thus the LIRs of NBR1 are binding in a very similar fashion to LC3B WT and the mutants. The variation in amino acids that are playing a significant part in the interaction, and the size of the CSP is then dependant on the precise position of the LIR on the face of LC3B.

Using the list of amino acids with significant CSPs it was then possible to dock the LIRs on the LC3B mutants using HADDOCK. All four HADDOCK runs carried out produced top scoring clusters that contained structure showing the LIR binding to the canonical binding site on LC3B, with the aromatic and hydrophobic residue on the LIR binding to hydrophobic pocket 1 and 2 respectively. A summary of the run results is given in Table 6.3.

| Titration | KL + LIR1 | KL + LIR2 | T29D + LIR1 | T29D + LIR2 |
|--------------------------------------|--|---|--|---|
| Structures created (clusters) | 184(3) | 186(5) | 188(3) | 193(2) |
| Top scoring run | -96.4±0.8 Cluster 1 139 structures | -107.9±3.2 Cluster 2 60 structures | -95.5±4.1 Cluster 1 176 structures | -98.2±0.6 Cluster1 133 structures |
| Top scoring LIR orientation | Standard | Standard | Standard | Standard |
| Other run of interest | | Cluster 1 -93.9±2.7 110 structures Reverse LIR orientation | | |

Table 6.3: HADDOCK run results for titrations of LIR peptides with LC3B mutants.

For LC3B WT and KL titrations with LIR2, as well as the top scoring cluster, which showed the LIR in the standard configuration with the aromatic and hydrophobic amino acids in hydrophobic pocket 1 and 2 respectively, there was also a cluster with a HADDOCK score around 10 less, showing the LIR in reverse orientation, with the hydrophobic residue on the LIR in HP1 and the aromatic in HP2. Whilst looking at the mapping of the significant CSPs, and the pattern around HP1, suggests that this structure is not what is happening when the LIR is binding to LC3B, it is an interesting result for the LIRs, as it suggests the changes in

sequence between LIR1 and LIR2 then make the docking of LIR2 in the reverse orientation a comparably high scoring cluster in HADDOCK.

6.3.3 Comparison for Mutants with LIRs

To better understand how the mutations on LC3B KL and T29D effect the interaction with the LIRs, and to further understand how the difference in sequences between LIR1 and LIR2 have an impact on binding selectivity, the titrations were compared with those of LC3B WT. For all the titrations large numbers of peaks moved in the ^1H - ^{15}N HSQC. The analysis of data gave sets of amino acids that had undergone significant changes, which were located around the canonical binding site for all the interactions of the LIRs with the LC3B mutants. To better understand where the differences were, the difference in CSPs between the titrations were calculated, and then those which were significant were selected (Chapter 2.5.2.6).

6.3.4 Comparison of LC3B KL and LC3B WT Titration with NBR1 LIR1

For the titration of LC3B KL with NBR1 LIR1 The difference between CSPs was calculated and plotted with the significance cut offs (Figure 6.19).

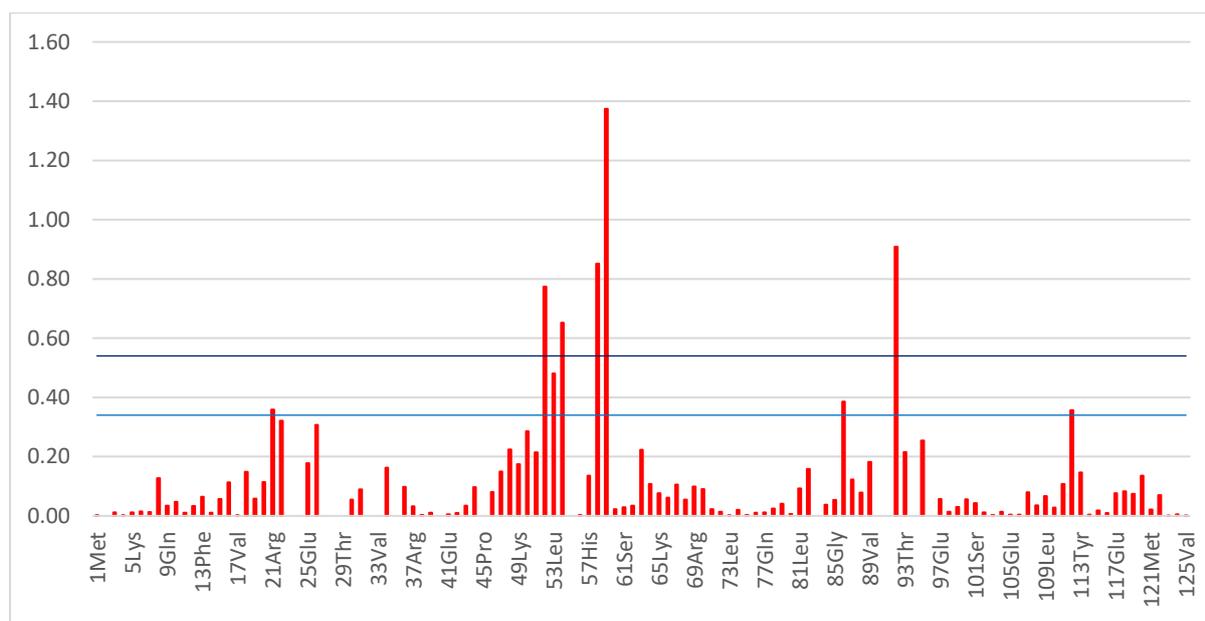


Figure 6.19: Plot of differences between CSPs for titrations LC3B WT and LC3B KL with NBR1 LIR1. Value for significant differences is shown in blue and most significant differences is shown in dark blue.

From this plot we can see that although the majority of CSPs have small differences between the titrations, that there are 16 that see significant change between the two titrations. Of the 16 amino acids that had significant differences between titrations, six are significant and 12 are in the most significant grouping (Table 6.4).

Significant Differences in CSP

| Amino Acid | LC3B WT CSP | LC3B KL CSP |
|------------|-------------|-------------|
| R21 | 0.59 | 0.24 |
| L/A53 | 0.61 | 0.12 |
| H86 | 0.46 | 0.07 |
| V112 | 0.38 | 0.02 |

Most Significant Differences in CSP

| Amino Acid | LC3B WT CSP | LC3B KL CSP |
|------------|-------------|-------------|
| I23 | 0.15 | CNF |
| R24 | CNF | 0.09 |
| H27 | CNF | 0.02 |
| V33 | 0.32 | CNF |
| I35 | 0.17 | CNF |
| F52 | 1.43 | 0.65 |
| V54 | 1.13 | 0.48 |
| V58 | 1.83 | 0.98 |
| N59 | 1.55 | 0.18 |
| V83 | 0.69 | CNF |
| S90 | 0.46 | CNF |
| S96 | 0.34 | CNF |

Table 6.4: List of significant and most significant differences between the titrations of LC3B WT and LC3B KL with LIR1. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration.

From the tabulated data the majority of significant differences that are around the binding site have a larger CSP for the titration of LC3B WT with NBR1 LIR1 than the KL mutant. The exception to this is for I23, V33 and I35, where there are small CSPs in the wild type titration, but it was not possible to follow the peaks in the KL mutant titration, suggesting the amino acid is playing significant part in the interaction with the LIR. These significant differences were then mapped onto the surface of LC3B with the HADDOCK structure for the LIRs to visualise where on the protein the major differences in interaction with LIR1 were occurring (Figure 6.20).

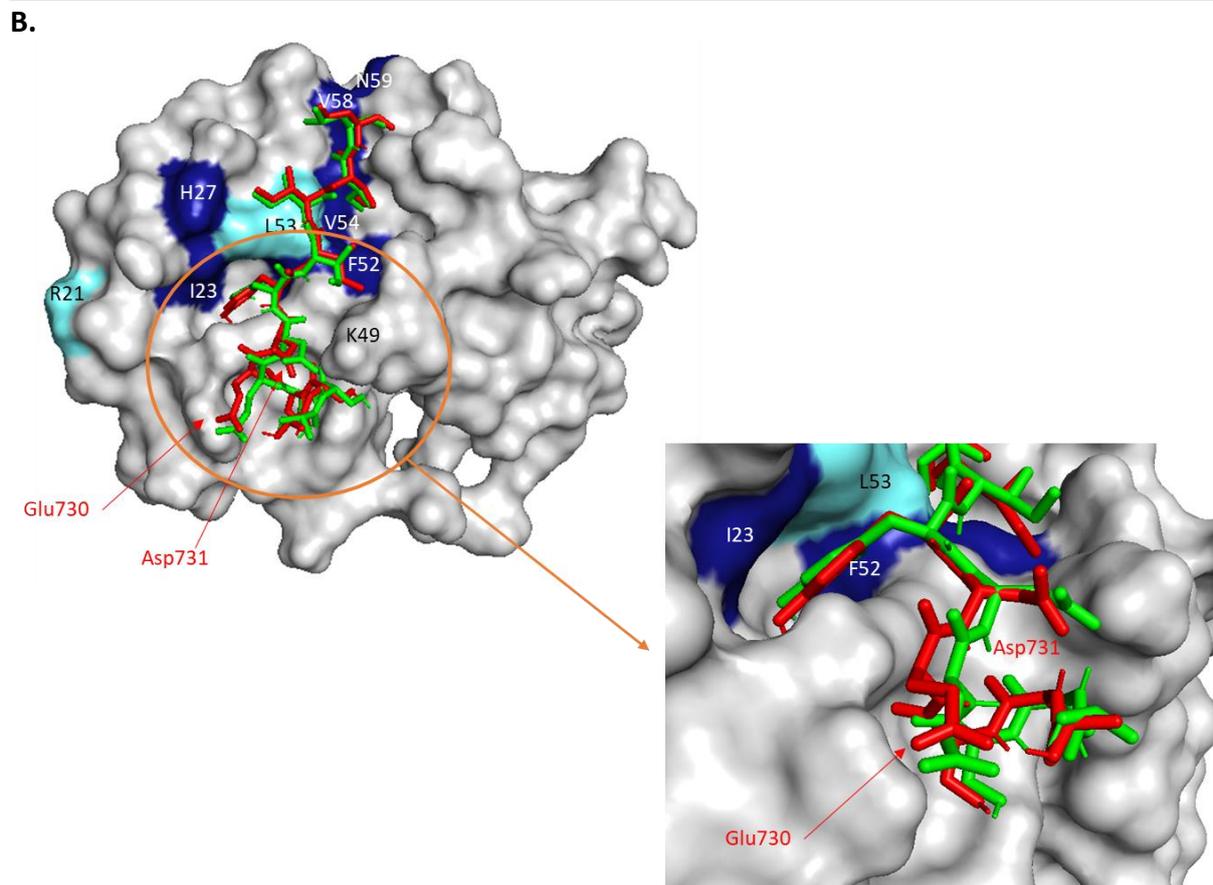
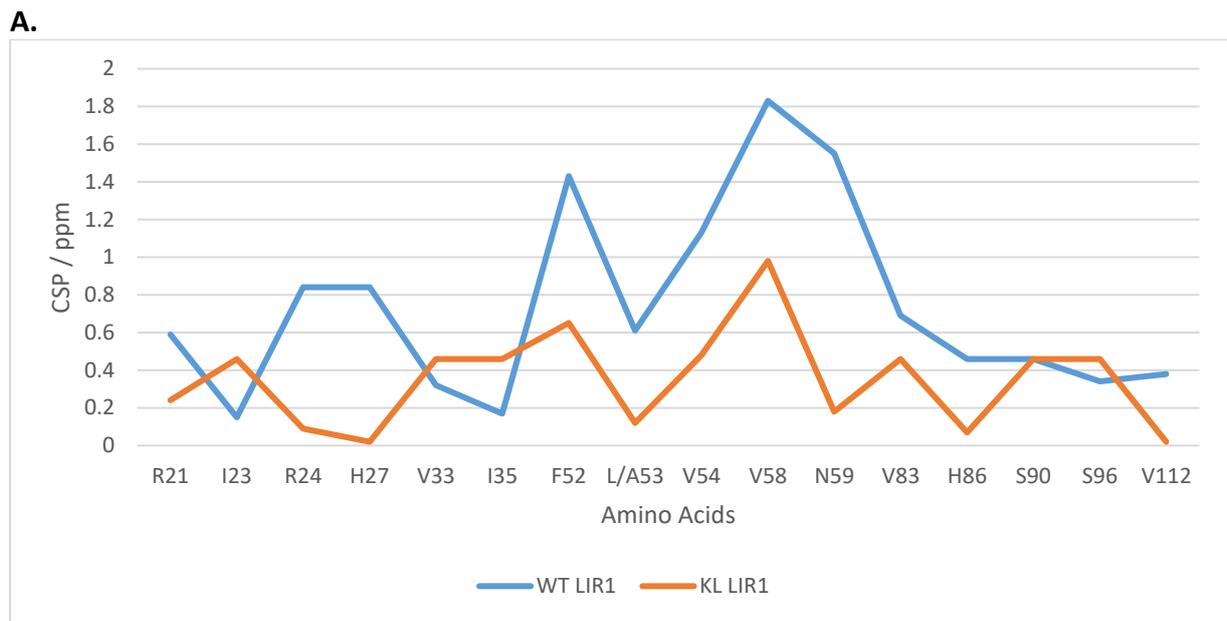


Figure 6.20: A: Graphical representation of CSPs that are significantly different between the titration with LC3B WT and NBR1 LIR1 (blue) and LC3B KL and NBR1 LIR1 (orange). B: Surface of LC3B WT with significant differences between titration of LC3B WT and LC3B KL with LIR1 mapped. Significant and most significant differences are coloured light blue and blue respectively. HADDOCK structure for LIR1 position in binding site are shown for LC3B WT (green) and LC3B KL (red).

The HADDOCK structure for LIR1 binding to LC3B shown for the wild type and KL mutant titration are very similar, with the main differences seen on Tyr732 and the preceding amino acids. Tyr732 sits slightly further to the left side of HP1, and hence the difference in perturbation seen at I23 between the two titrations. Glu730 is also moved up on the face of LC3B KL. The knock on effect of the change in position on Tyr732 and Glu730 seems to have been to pull Asp731 away from K49, and lessen the interaction. These slight differences are enough to alter the interaction, which leads to larger CSPs for the interaction of LIR1 with LC3B WT, implying larger changes to LC3B WT on binding with LIR1. This tied with the MS competition data from chapter 4, confirms that KL mutant has weakened the interaction of LIR1 with LC3B KL.

6.3.5 Comparison of LC3B KL and LC3B WT Titration with NBR1 LIR2

Whilst the titration of LC3B KL with NBR1 LIR2 show the LIR interacting at the canonical binding site, and have similar significant CSPs to the wild type titration, there are noticeable differences. Whilst the wild type titration has CSPs for L22 and H27 that are classed as most significant, they are not for the KL mutant titration. A51 and A53 on the KL mutant titration (K51 and L53 on LC3B WT) have CSPs that are categorised as most significant, whereas they are only significant in the wild type titration. The Differences in CSPs were plotted with the significance cut offs (Figure 6.21).

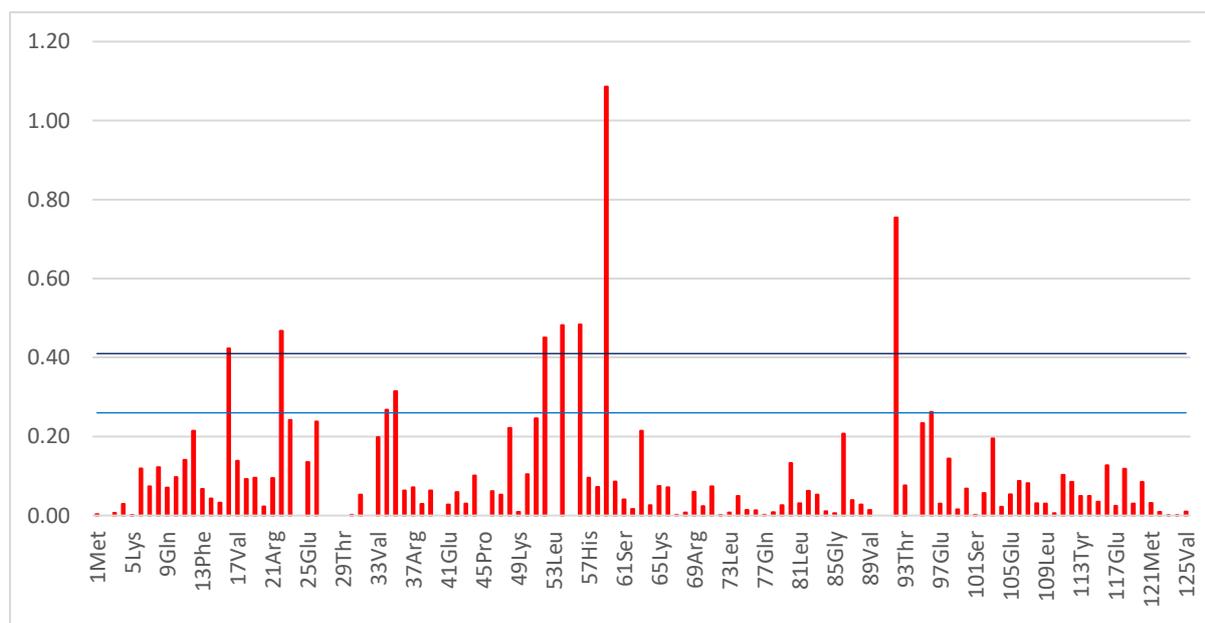


Figure 6.21: Plot of differences between CSPs for titrations LC3B WT and LC3B KL with NBR1 LIR2. Value for significant differences is shown in blue and most significant differences is shown in dark blue.

From comparing the two titrations it was found that there were a total of 14 amino acids that had differences between the two titrations above the significance threshold, with 11 amino acids being classed as most significant (Table 6.5).

Significant Differences in CSP

| Amino Acid | LC3B WT CSP | LC3B KL CSP |
|------------|-------------|-------------|
| I34 | 0.48 | 0.22 |
| I35 | 0.39 | 0.07 |
| S96 | 0.42 | 0.16 |

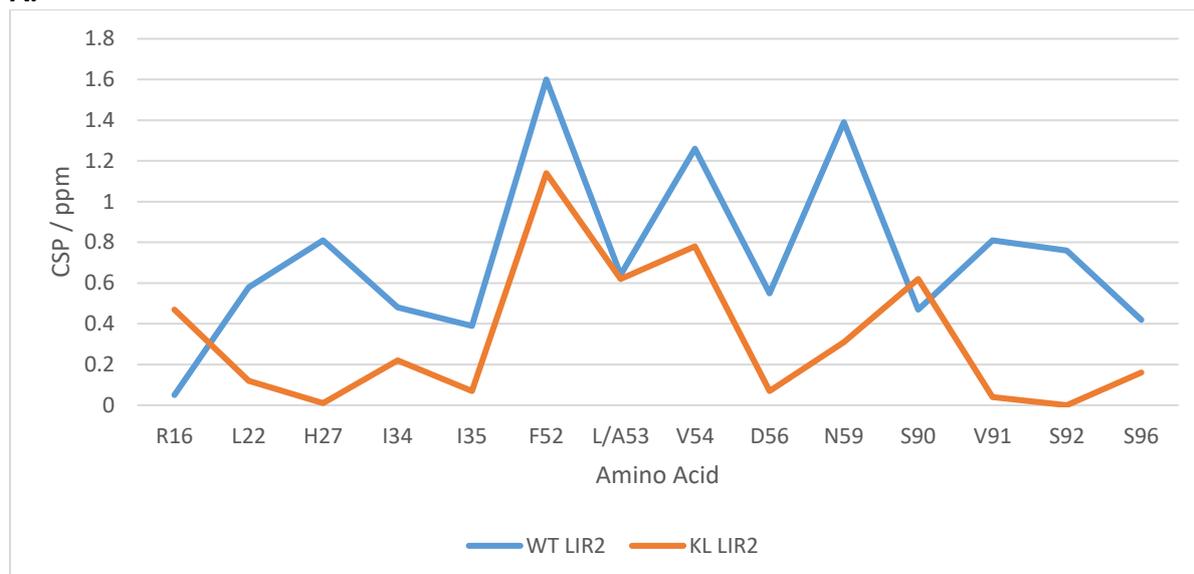
Most Significant Differences in CSP

| Amino Acid | LC3B WT CSP | LC3B KL CSP |
|------------|-------------|-------------|
| R16 | 0.05 | 0.47 |
| L22 | 0.58 | 0.12 |
| H27 | CNF | 0.01 |
| F52 | 1.60 | 1.14 |
| L/A53 | 0.64 | CNF |
| V54 | 1.26 | 0.78 |
| D56 | 0.55 | 0.07 |
| N59 | 1.39 | 0.31 |
| S90 | 0.47 | CNF |
| V91 | CNF | 0.04 |
| S92 | 0.76 | 0 |

Table 6.5: List of significant and most significant differences between the titrations of LC3B WT and LC3B KL with LIR2. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration.

Looking at the CSPs for the two titrations for these significant differences (Table 6.5) it can be seen that those that are for amino acids that form the binding site all fall in the most significant category. To visualise where the differences are precisely the amino acids were coloured on the surface of LC3B, and the HADDOCK structure for the LIR position for the two titrations shown (Figure 6.22).

A.



B.

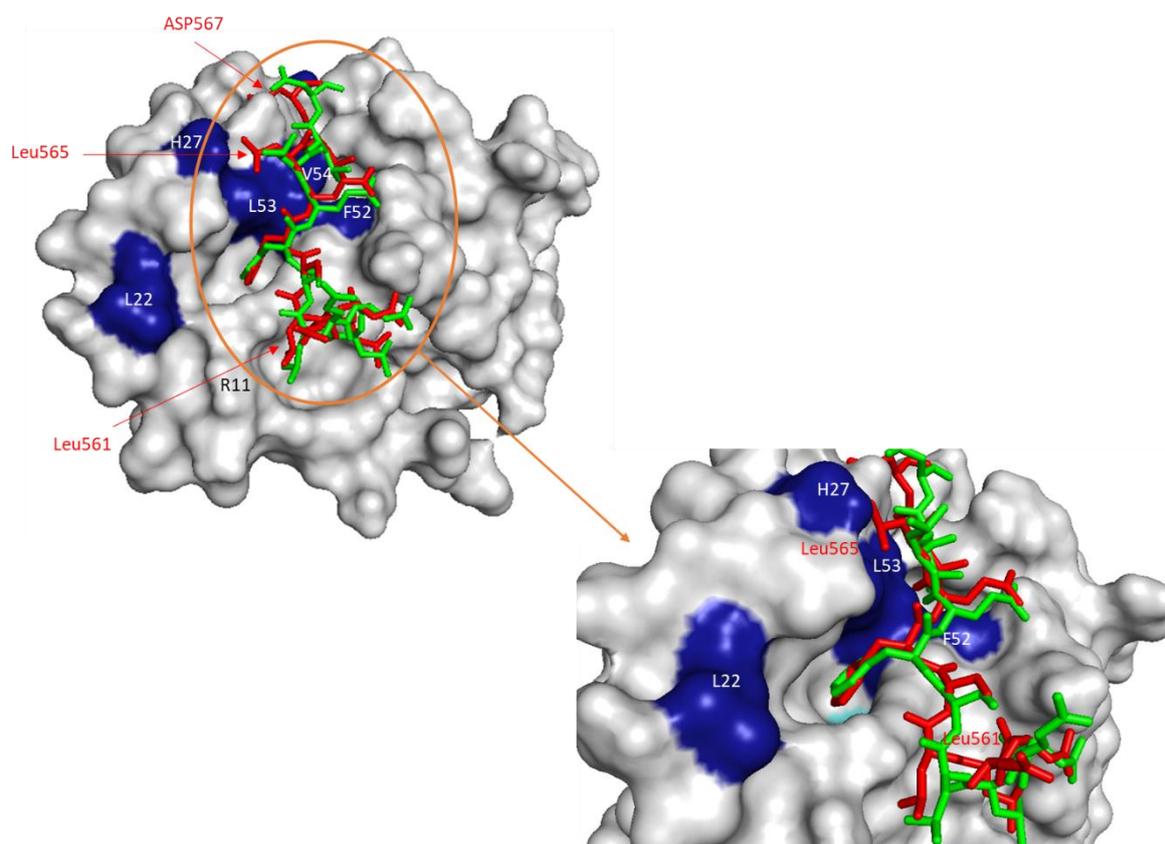


Figure 6.22: A: Graphical representation of CSPs that are significantly different between the titration with LC3B WT and NBR1 LIR2 (blue) and LC3B KL and NBR1 LIR2 (orange). B: Surface of LC3B WT with significant differences between titration of LC3B WT and LC3B KL with LIR2 mapped. Significant and most significant differences are coloured light blue and blue respectively. HADDOCK structure for LIR2 position in binding site are shown for LC3B WT (green) and LC3B KL (red).

Comparing the two HADDOCK structure for the LIRs and the significant CSP do show more differences between the WT and KL data. Phe563 is in a very similar position, apart from being angled slightly back, when interacting with the KL mutant and hence the difference seen on F52, which whilst having a significant difference between the two titrations, was still classed as one of the most significant CSPs in both titrations. L53 also shows a significant difference, and from the HADDOCK structures, Leu561 on LIR sits closer to the face of the KL mutant, whereas for wild type it points more towards R11 (another significant difference between the two titrations). The position of Leu566 in HP2 is then very similar between the wild type and KL mutant, with it being slightly tighter against LC3B KL than with LC3B WT. Finally Asp567 sits closer to the face of LC3B KL than LC3B WT whilst being in the same orientation.

These differences all combine to allow LIR2 to be tighter to the face of LC3B KL. This then supports the observation from the MS competition experiments in chapter four, that showed that LIR2 bound preferentially to the KL mutant over LC3B WT. From the NMR titrations it is possible to deduce that with lysine 51 mutated to an alanine, making the face of LC3B KL less basic, and also altering the environment of R11. Leu561 is then positioned between A51 and R11, a more favourable interaction than it was with the more basic wild type LC3B. However, in general there are only small differences in the CSP between the titrations. These small differences between the interactions of the LIRs then leads to a large difference in binding preference for LIR2 with the LC3B mutants. Whilst the work here was done with a mutant that does not appear naturally, it does support the idea that different ATG8 proteins, whilst having high sequence alignment, can have small changes around the binding site which then have a large impact on selectivity of LIR binding partner.

6.3.6 Comparison of LC3B KL Titration with NBR1 LIR1 and LIR2

With the titrations of NBR1 LIR1 and LIR2 with LC3B KL compared with the wild type titrations, the two titrations of the LIRs with KL mutant were compared. As with all the previous titrations there are a very similar set of amino acids with significant amino acids, and so to get a true reflection of where the differences are between the interactions of the two LIRs with LC3B KL, the difference in CSPs were calculated and plotted (Figure 6.23).

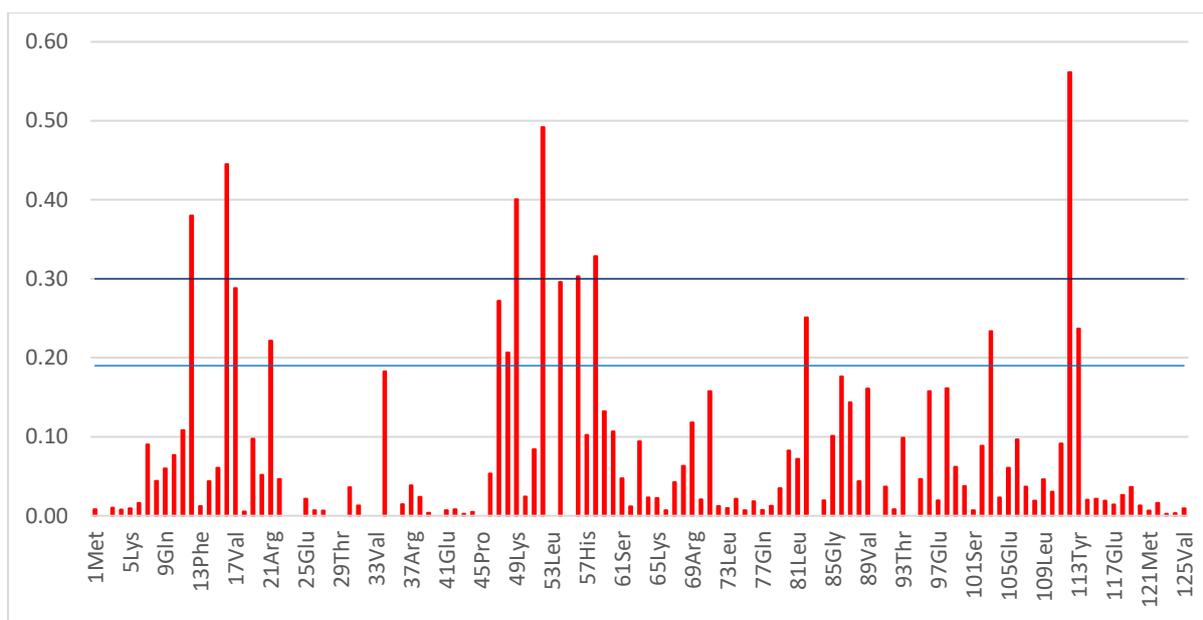


Figure 6.23: Plot of differences between CSPs for titrations NBR1 LIR1 and LIR2 with LC3B KL. Value for significant differences is shown in blue and most significant differences is shown in dark blue.

From this analysis there are 21 amino acids that have significant differences between the titration of which 12 are classed as most significant. These amino acids and the CSPs for each titration are given in Table 6.6.

Significant Differences in CSP

| Amino Acid | LIR1 | LIR2 |
|------------|------|------|
| V17 | 0.07 | 0.36 |
| R21 | 0.24 | 0.46 |
| L47 | 0.43 | 0.16 |
| D48 | 0.31 | 0.10 |
| V54 | 0.48 | 0.78 |
| D56 | 0.37 | 0.07 |
| L82 | 0.13 | 0.38 |
| K103 | 0.04 | 0.27 |
| Y113 | 0 | 0.24 |

Most Significant Differences in CSP

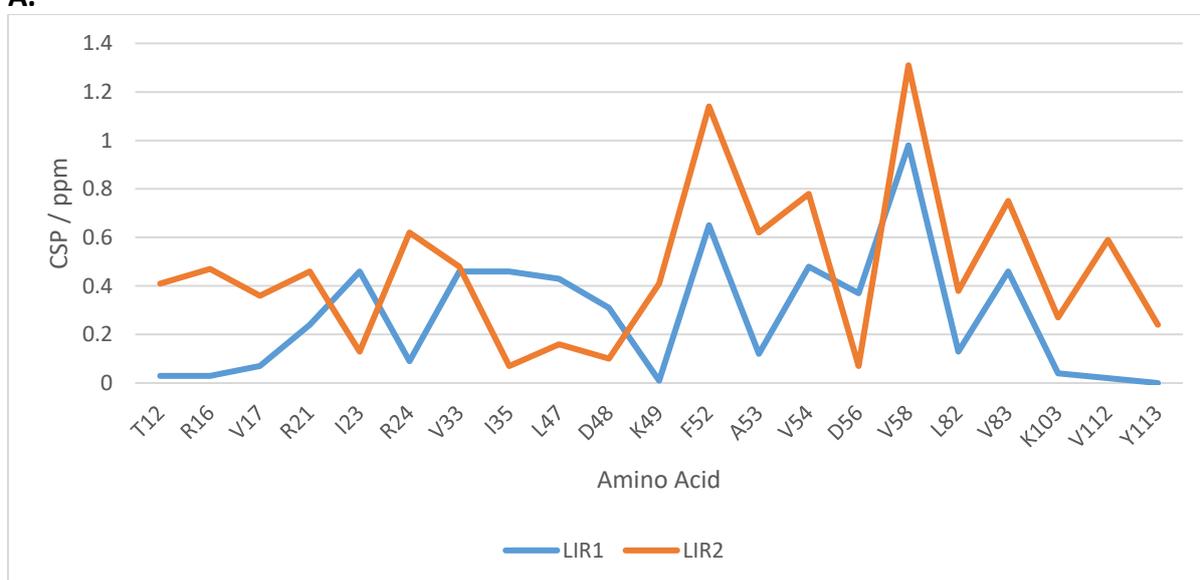
| Amino Acid | LIR1 | LIR2 |
|------------|------|------|
| T12 | 0.03 | 0.41 |
| R16 | 0.03 | 0.47 |
| I23 | CNF | 0.13 |
| R24 | 0.09 | CNF |
| V33 | CNF | 0.48 |
| I35 | CNF | 0.07 |
| K49 | 0.01 | 0.41 |
| F52 | 0.65 | 1.14 |
| A53 | 0.12 | CNF |

| | | |
|------|------|------|
| V58 | 0.98 | 1.31 |
| V83 | CNF | 0.75 |
| V112 | 0.02 | 0.59 |

Table 6.6: List of significant and most significant differences between the titrations of LC3B LIR1 and LIR2. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration.

This is a large number of amino acids that show significant differences which appear to be spread over LC3B KL. To gain better insight into the areas where the LIR interactions differ the amino acids were used to colour the surface of LC3B KL (Figure 6.24)

A.



B.

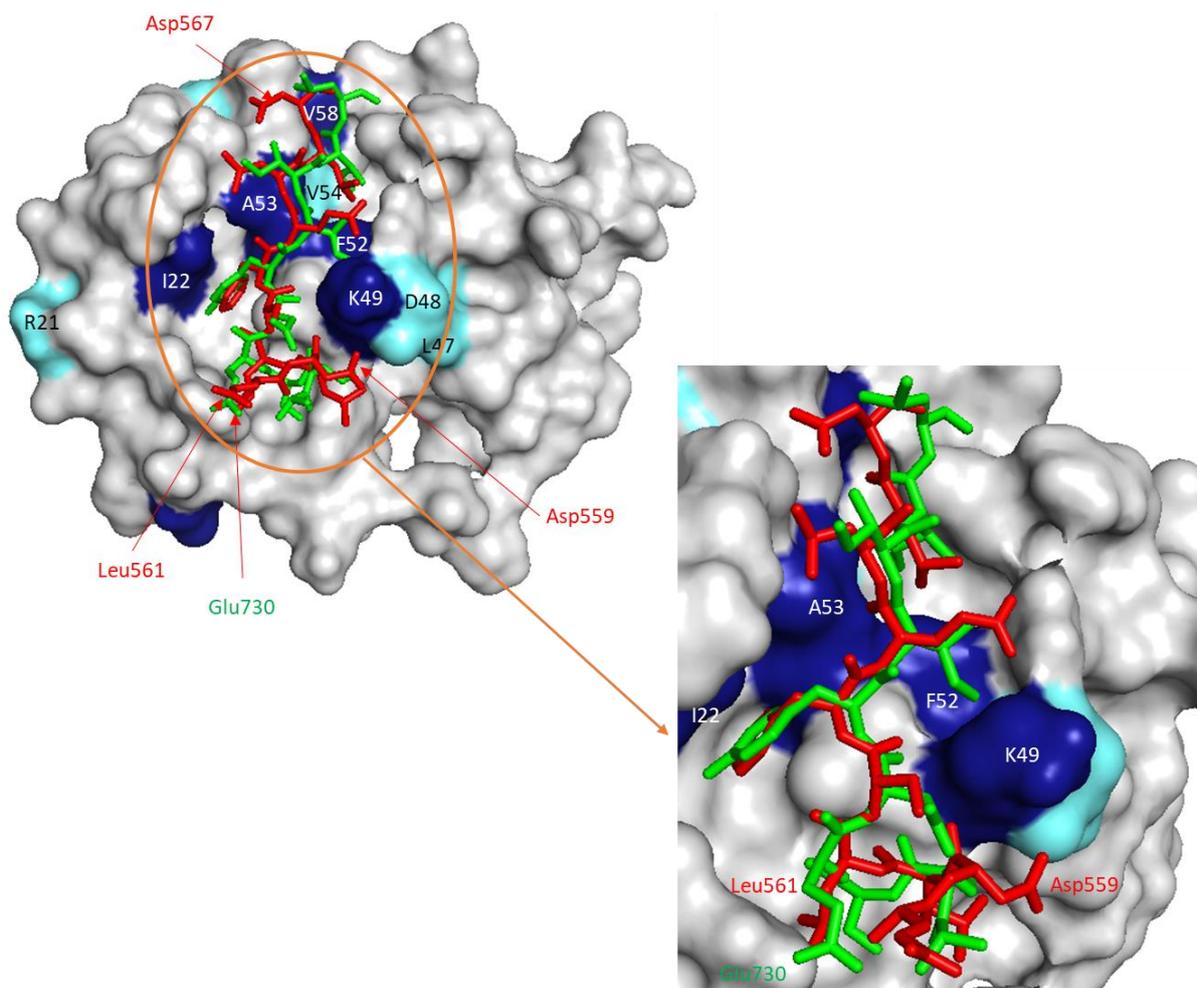


Figure 6.24: A: Graphical representation of CSPs that are significantly different between the titration with LC3B KL and NBR1 LIR1 (blue) and NBR1 LIR2 (orange). B: Surface of LC3B KL with significant differences between titration of LC3B KL with NBR1 LIR1 and LIR2. Significant and most significant differences are coloured light blue and blue respectively. HADDOCK structure for LIR position in binding site are shown for LIR1 (green) and LIR2 (red).

Comparing the HADDOCK structures of LIR1 and LIR2 interacting with LC3B KL, we can see that Phe563 on LIR2 sits more centrally in HP1 than Tyr732 on LIR1. The amino acids that follow the aromatic residues are in similar positions, but with those on LIR2 pulled further across the face of LC3B KL. This seems to be driven by an interaction between D56 on KL and Asp567 on LIR2. The end result is that LIR2 sits closer to the face of the KL mutant. For the amino acids preceding the amino acids on the LIRs, interactions with Asp559 on LIR2 and K49 and Glu730 on LIR1 with R11 can be observed. As mentioned previously Leu561 on LIR2 is able to sit closer to the face of the KL mutant.

6.3.7 Comparison of LC3B T29D and LC3B WT Titration with NBR1 LIR1

To see how the phosphomimetic mutant LC3B T29D alters the interaction of LC3B with LIR1 the wild type and T29D data for the titrations with LIR were compared. The difference in CSPs were calculated, and the standard significance test applied to the values. These values

were then plotted and the thresholds for level of significance marked on the plot (Figure 6.25).

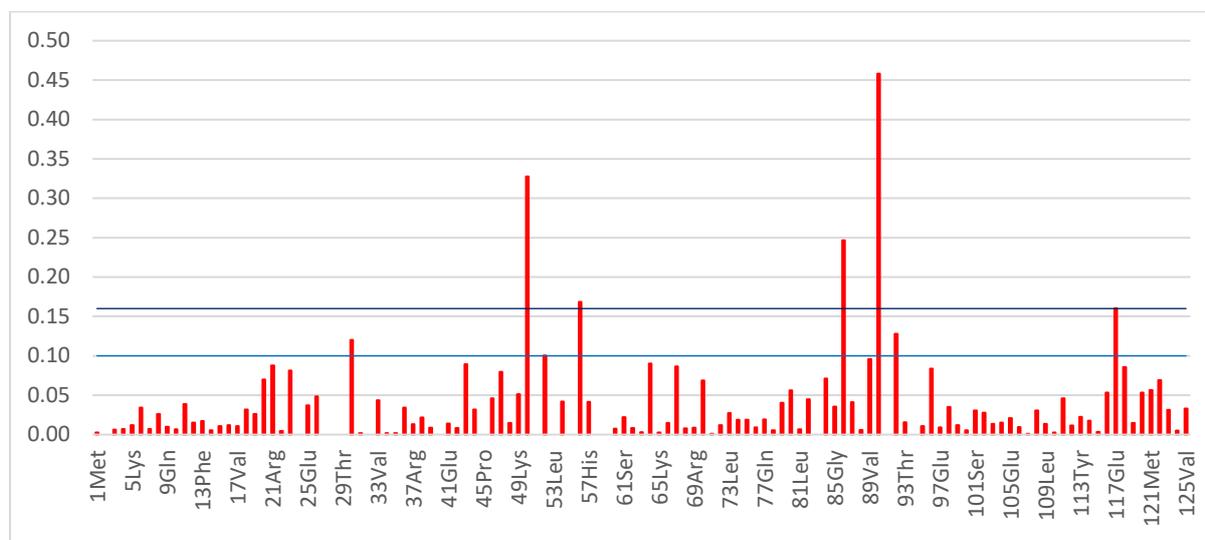


Figure 6.25: Plot of differences between CSPs for titrations LC3B WT and LC3B T29D with NBR1 LIR1. Value for significant differences is shown in blue and most significant differences is shown in dark blue.

This analysis shows that there are 12 amino acids that have a difference in CSPs between the two titrations above the significance threshold (including those that could not be followed in either titration) with 9 of those being in the most significant category. These amino acids and the CSPs for each titration are given in Table 6.7.

Significant Differences in CSP

| Amino Acid | LC3B WT CSP | LC3B T29D CSP |
|------------|-------------|---------------|
| K30 | 0.04 | 0.16 |
| S92 | 0.92 | 1.05 |
| E117 | 0.17 | 0.01 |

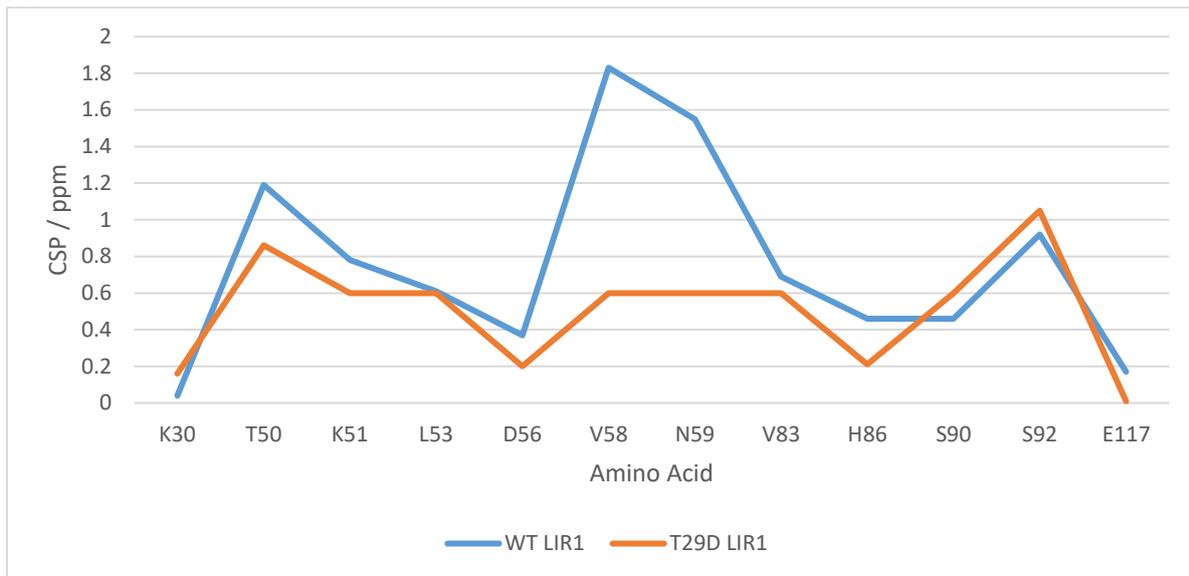
Most Significant Differences in CSP

| Amino Acid | LC3B WT CSP | LC3B T29D CSP |
|------------|-------------|---------------|
| T50 | 1.19 | 0.86 |
| K51 | 0.78 | CNF |
| L53 | 0.61 | CNF |
| D56 | 0.37 | 0.20 |
| V58 | 1.83 | CNF |
| N59 | 1.55 | CNF |
| V83 | 0.69 | CNF |
| H86 | 0.46 | 0.21 |
| S90 | 0.46 | CNF |

Table 6.7: List of significant and most significant differences between the titrations of LC3B WT and LC3B T29D with LIR1. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration.

From this table we can see that the number of amino acids that have been categorised as most significant is perhaps misleading, as there are a number of peaks in the T29D titration that could not be followed. For V58 and N59 the peaks could not be followed in the T29D titration, but also have a large value in the wild type titration, indicating that the amino acids underwent significant changes in both titrations. If these amino acids are ruled out then there are four amino acids, K30, T50, K51 and L53 on the surface of LC3B that have significant differences (Figure 6.26)

A.



B.

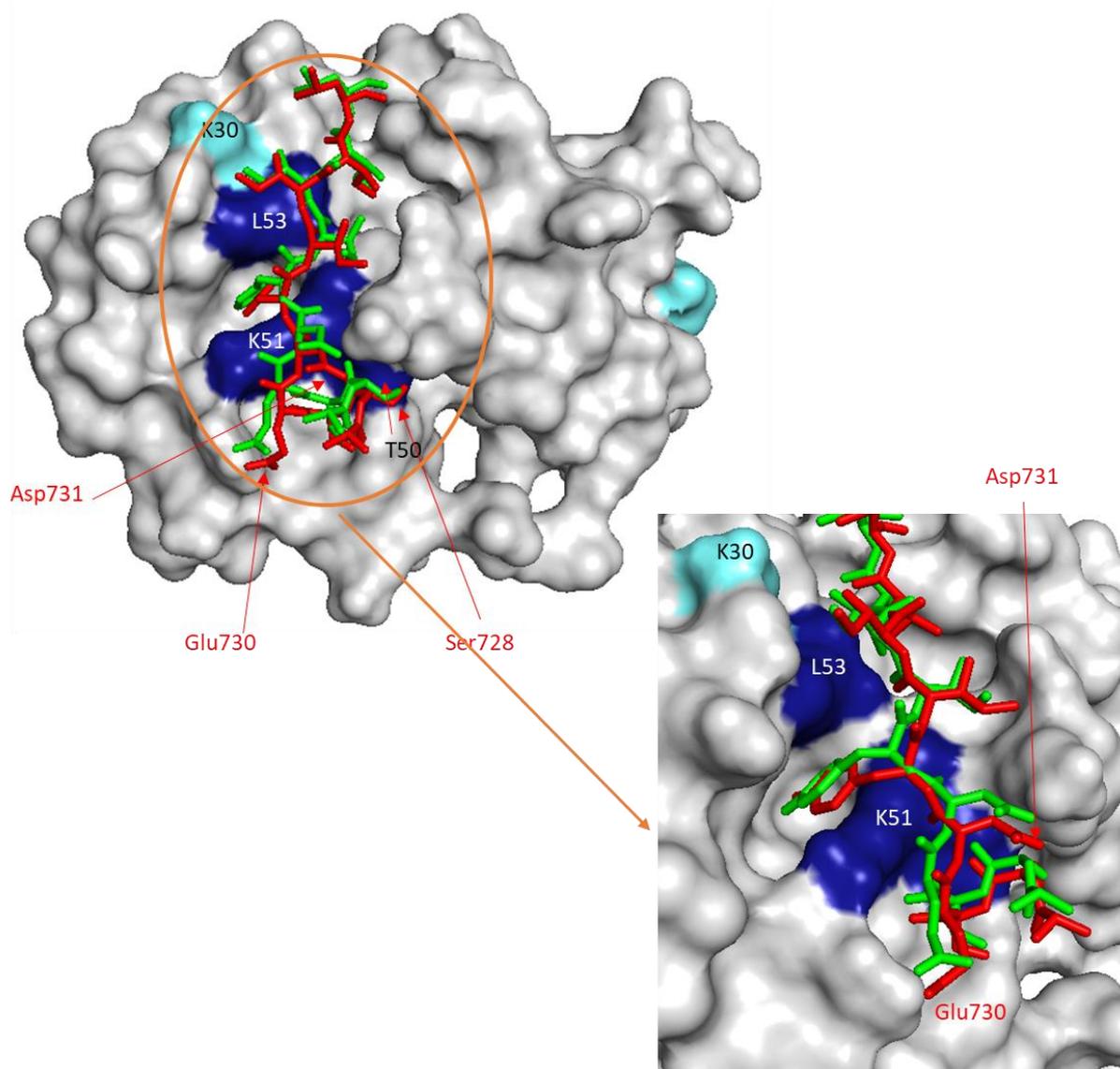


Figure 6.26: A: Graphical representation of CSPs that are significantly different between the titration with LC3B WT and NBR1 LIR1 (blue) and LC3B T29D and NBR1 LIR1 (orange). B: Surface of LC3B WT with significant differences between titration of LC3B WT and LC3B T29D with LIR1 mapped. Significant and most significant differences are coloured purple and blue respectively. HADDOCK structure for LIR1 position in binding site are shown for LC3B WT (green) and LC3B T29D (red).

Looking at where the significant differences between the two titrations occur, in conjunction with the HADDOCK structures for the position of LIR1 on LC3B WT and LC3B T29D, the main variations are around HP1, and how Tyr732 inserts in this pocket and the preceding amino acids on the LIR. T50 shows a larger CSP for the WT titration, with the position of LIR1 closer to the centre of the face of LC3B. Tyr732 then sits slightly closer to K51 on the T29D structure. Glu730 sits in a slightly different orientation close to R11, and so does Asp731 with K49. Ser728 then sits slightly closer to T50 for LIR2. However, the position of LIR1 for the rest of the peptide varies very little between LC3B WT and LC3B T29D. LIR1 does appear to be slightly closer to the face of LC3B WT than for T29D but only marginally.

The T29D mutation appears to have very little effect on how LIR1 binds to LC3B, and this is backed up by the MS competition results in chapter 4, that show there is a slight preference for LC3B WT over T29D.

6.3.8 Comparison of LC3B T29D and LC3B WT Titration with NBR1 LIR2

Whilst the standard set of key amino acids were present in both the titration of LC3B WT and LC3B T29D with LIR2, there were almost double the number classed as most significant in the T29D titration (14 for T29D and 8 for WT). The values for the differences in CSP and the significance cut off were plotted (Figure 6.27).

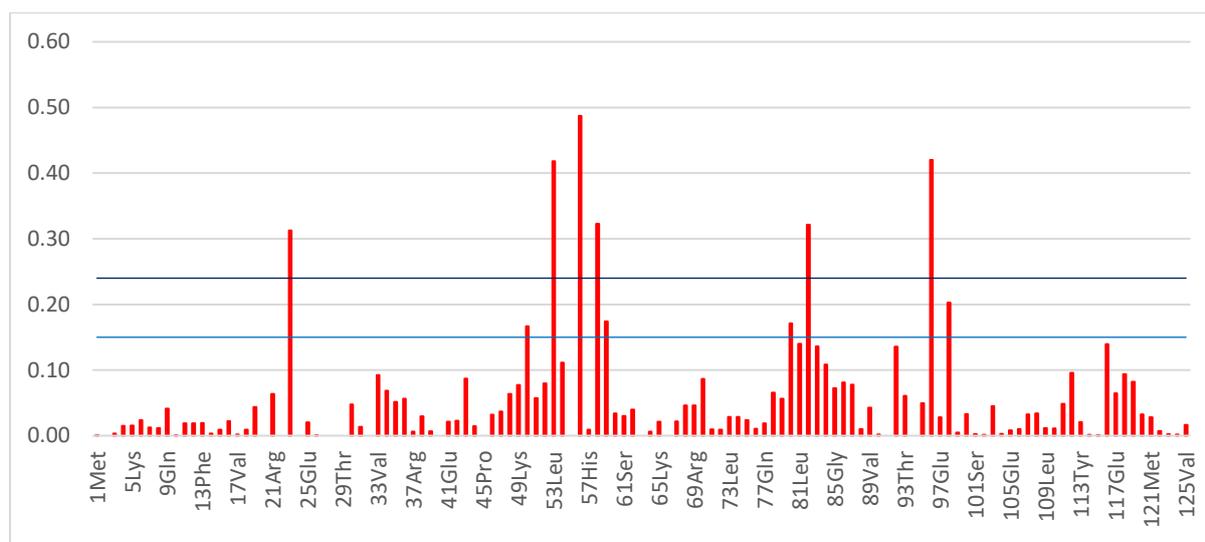


Figure 6.27: Plot of differences between CSPs for titrations LC3B WT and LC3B T29D with NBR1 LIR2. Value for significant differences is shown in blue and most significant differences is shown in dark blue.

In total there were 14 amino acids that had significant differences in CSP between the titration LC3B WT and LC3B T29D with NBR1 LIR2, and of those ten were classed as being most significant. The amino acids and the values for the CSPs for the two titrations are given in Table 6.8.

Significant Differences in CSP

| Amino Acid | LC3B WT CSP | LC3B T29D CSP |
|------------|-------------|---------------|
| T50 | 0.99 | 0.82 |
| N59 | 1.39 | 1.22 |
| F80 | 0.17 | 0.34 |
| V98 | 0.06 | 0.26 |

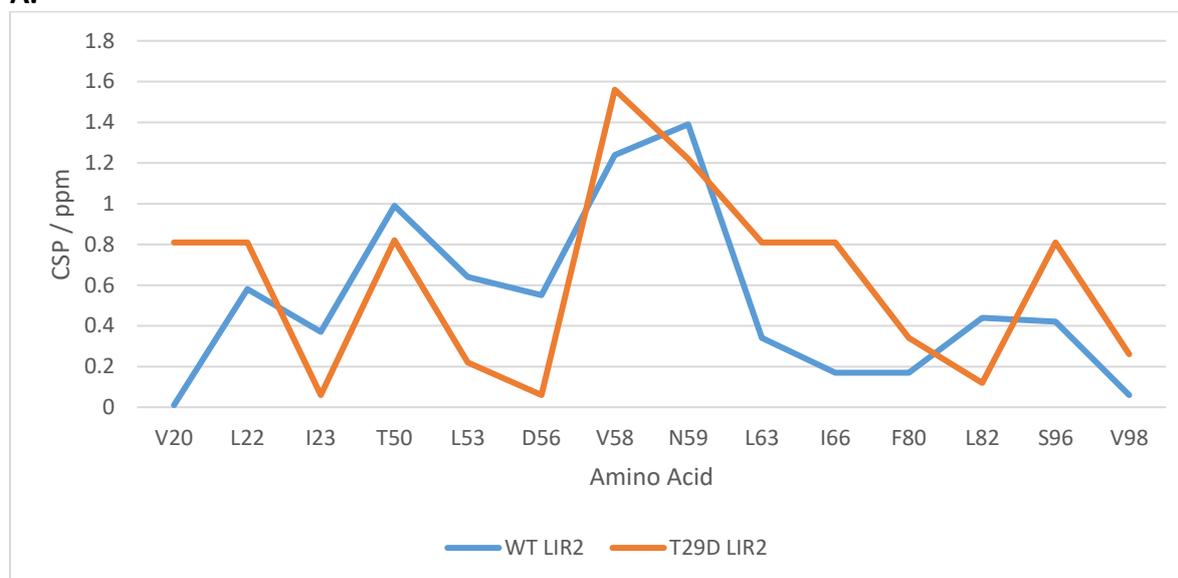
Most Significant Differences in CSP

| Amino Acid | LC3B WT CSP | LC3B T29D CSP |
|------------|-------------|---------------|
| V20 | 0.01 | CNF |
| L22 | 0.58 | CNF |
| I23 | 0.37 | 0.06 |
| L53 | 0.64 | 0.22 |
| D56 | 0.55 | 0.06 |
| V58 | 1.24 | 1.56 |
| L63 | 0.34 | CNF |
| I66 | 0.17 | CNF |
| L82 | 0.44 | 0.12 |
| S96 | 0.42 | CNF |

Table 6.8: List of significant and most significant differences between the titrations of LC3B WT and LC3B T29D with LIR2. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration.

To understand where these differences are on LC3B and to better understand how the interaction of LIR2 varies between the two LC3Bs the significant amino acids were mapped onto LC3B along with the HADDOCK structures for the position of LIR2 calculated from the data from the two titrations (**Error! Reference source not found.**).

A.



B.

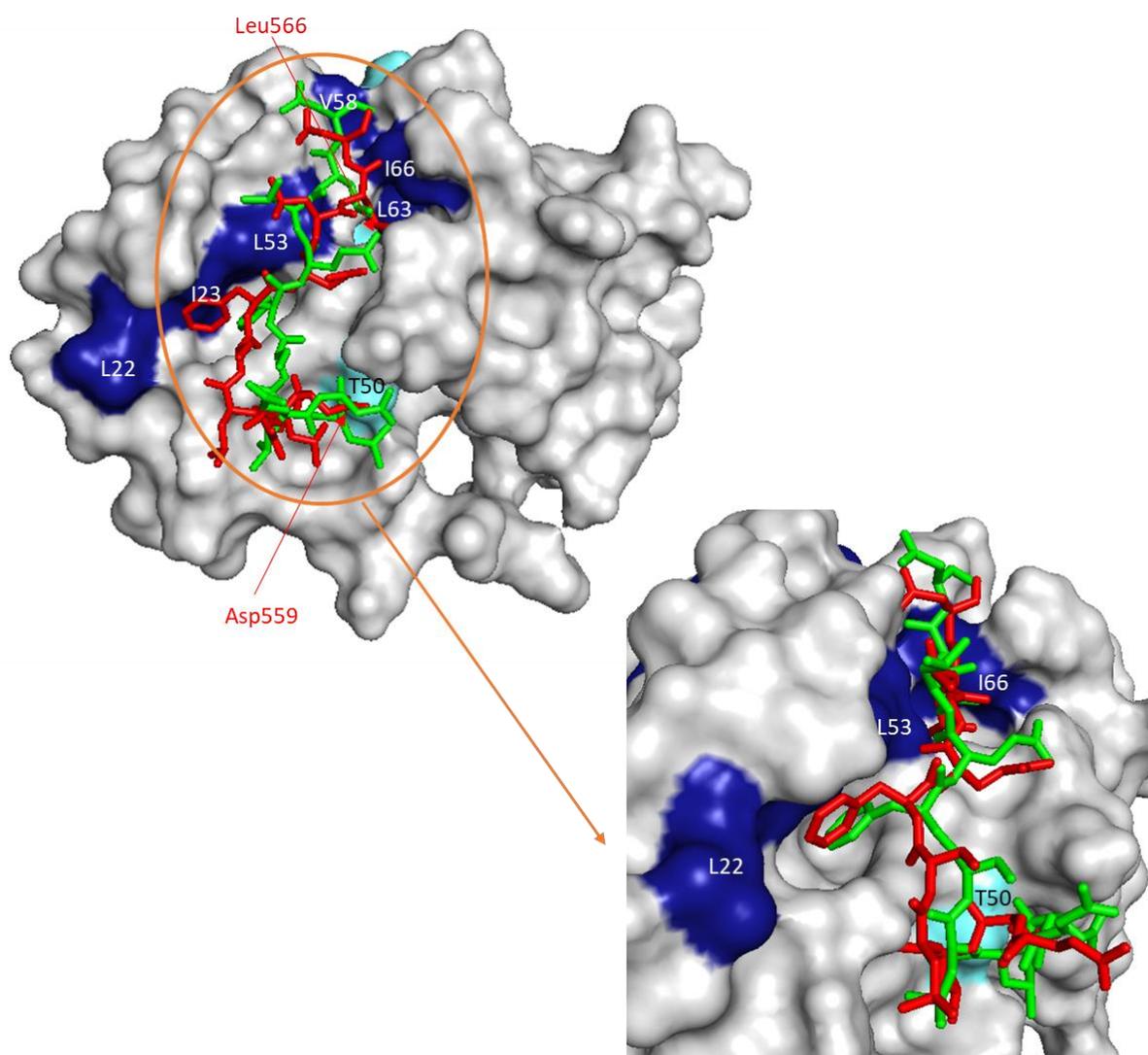


Figure 6.28: A: Graphical representation of CSPs that are significantly different between the titration with LC3B WT and NBR1 LIR2 (blue) and LC3B T29D and NBR1 LIR2 (orange). B: Surface of LC3B WT with significant differences between titration of LC3B WT and LC3B T29D with LIR2 mapped. Significant and most significant differences are coloured light blue and blue respectively. HADDOCK structure for LIR2 position in binding site are shown for LC3B WT (green) and LC3B T29D (red). The top image shows the front face of LC3B WT with the bottom image being side on.

From the HADDOCK structures for the LIR positions, we can see the major change is that Phe563 for the T29D data is sitting across the face of HP1, whereas for the WT titration it sits centrally in the pocket. This has the knock on effect of dragging the amino acids preceding the aromatic across the face of LC3B T29D. The result of this is that Asp559 is not as close to K49 on LC3B T29D as it is for LC3B WT, and so the electrostatic interaction would be lessened. As well as the difference around hydrophobic pocket one, Leu566 also sits deeper in HP2 with V58, L63 and I66 showing larger perturbations in the T29D titration than that of the wild type.

However, it is not clear exactly how the T29D mutation alters the interaction with NBR1 LIR2. One option is that the T29D mutant alters HP2 which allows Leu566 to fit deeper in the hydrophobic pocket, thus pulling the LIR further across the face of T29D preventing Ph563 sitting properly in HP1. Another option is that the mutation alters the alpha helix that forms HP1, which alters the position of Phe563 in the hydrophobic pocket. There is also the change in the charge around HP2, which could alter how Asp567 interacts.

The binding interaction of NBR1 LIR2 with LC3B T29D from the titration data and resulting HADDOCK structure has changed from the LC3B WT, and yet from the MS competition experiments in chapter 4, LIR2 has similar affinity for both LC3B WT and T29D. However, this may not reflect the true preference between the two LC3s, as LIR2 showed such a strong preference for the KL mutant that the difference between LC3B WT and T29D may have been lost in the noise. If the ratios of complex to free protein of LC3B WT and T29D are compared at the end point of the titration, then the value is a third higher for the wild type, implying LIR2 binds to LC3B WT preferentially over T29D. However, to further analyse this and confirm the result, a separate competition experiment would need to be carried out, with just LC3B WT and LC3B T29D present with LIR2.

6.3.9 Comparison of LC3B T29D Titration with NBR1 LIR1 and LIR2

Finally the titration data of the T29D mutant with the two LIRs of NBR1 were compared to analyse the effect the different sequences the LIRs have on the binding interaction. The difference in CSP between the two titrations were calculated and then analysed for significance. The values for the differences in CSP between titrations and the thresholds for the significance levels were then plotted (Figure 6.29).

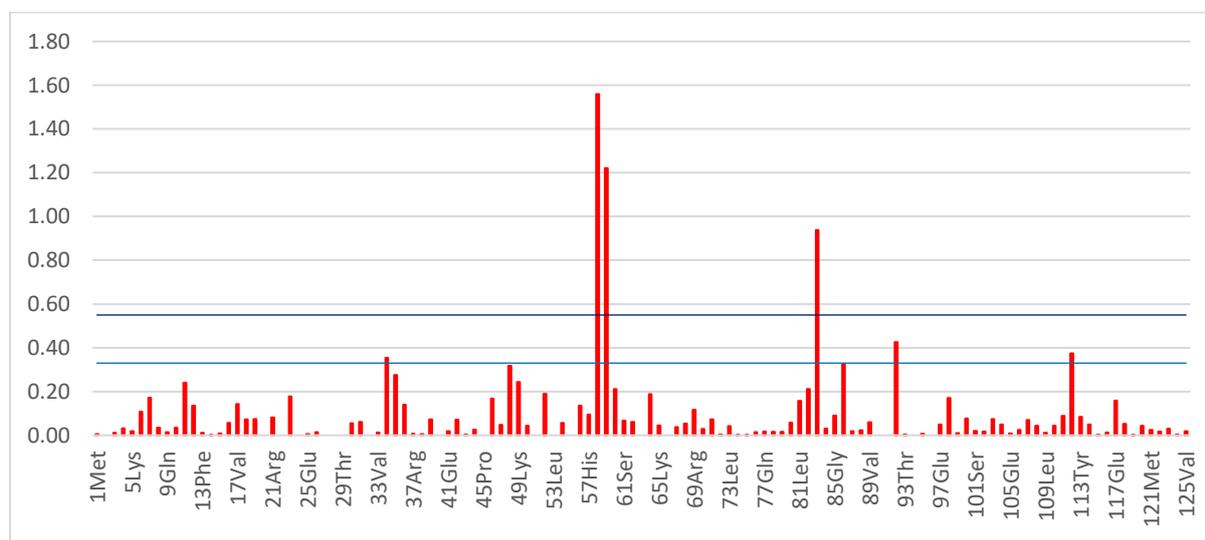


Figure 6.29: Plot of differences between CSPs for titrations NBR1 LIR1 and LIR2 with LC3B T29D. Value for significant differences is shown in blue and most significant differences is shown in dark blue.

In total there are 13 amino acids with significant differences in CSP, and of those 11 were classed as most significant. The amino acids and the values for the CSPs for the titrations of LC3B T29D with NBR1 LIR1 and LIR2 are given in Table 6.9.

Significant Differences in CSP

| Amino Acid | LIR1 | LIR2 |
|------------|------|------|
| I34 | 0.20 | 0.55 |
| V112 | 0.39 | 0.77 |

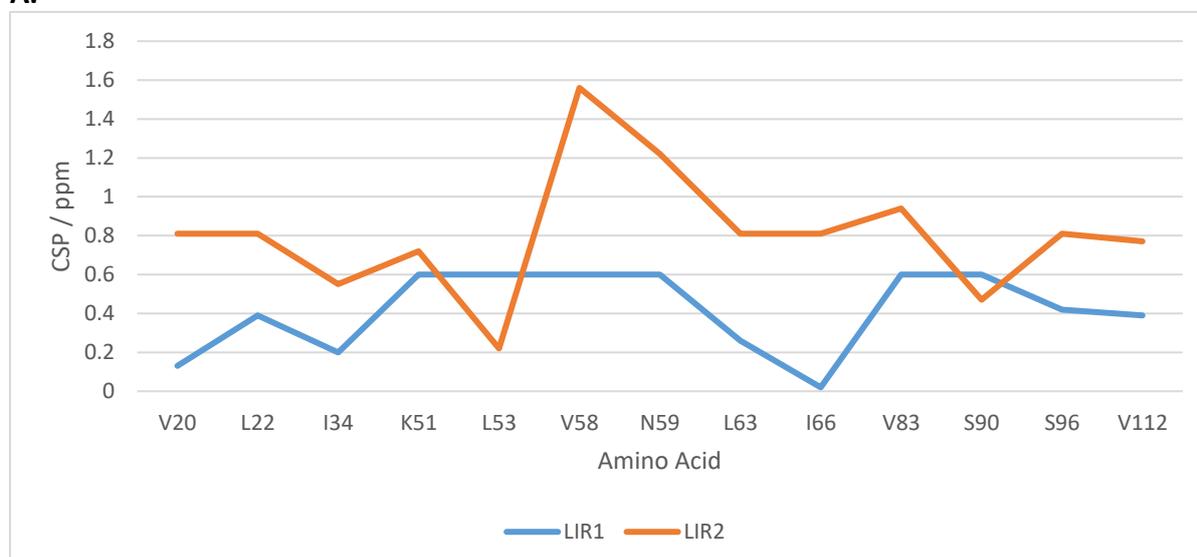
Most Significant Differences in CSP

| Amino Acid | LIR1 | LIR2 |
|------------|------|------|
| V20 | 0.13 | CNF |
| L22 | 0.39 | CNF |
| K51 | CNF | 0.72 |
| L53 | CNF | 0.22 |
| V58 | CNF | 1.56 |
| N59 | CNF | 1.22 |
| L63 | 0.26 | CNF |
| I66 | 0.02 | CNF |
| V83 | CNF | 0.94 |
| S90 | CNF | 0.47 |
| S96 | 0.42 | CNF |

Table 6.9: List of significant and most significant differences between the titrations of LC3B T29D with LIR1 and LIR2. CSP values for each titration are given with CNF denoting peaks that could not be followed in the titration.

Whilst in the titration for T29D with LIR1 the peaks for V58 and N59 could not be followed, they also had a large CSP in the T29D titration with LIR2, and so they will be removed from the set of significant amino acids. To understand where on LC3B T29D that these differences occur the amino acids that had significant differences were mapped onto the surface of LC3B T29D with the HADDOCK structures for the position of LIR1 and LIR2 shown (Figure 6.30).

A.



B.

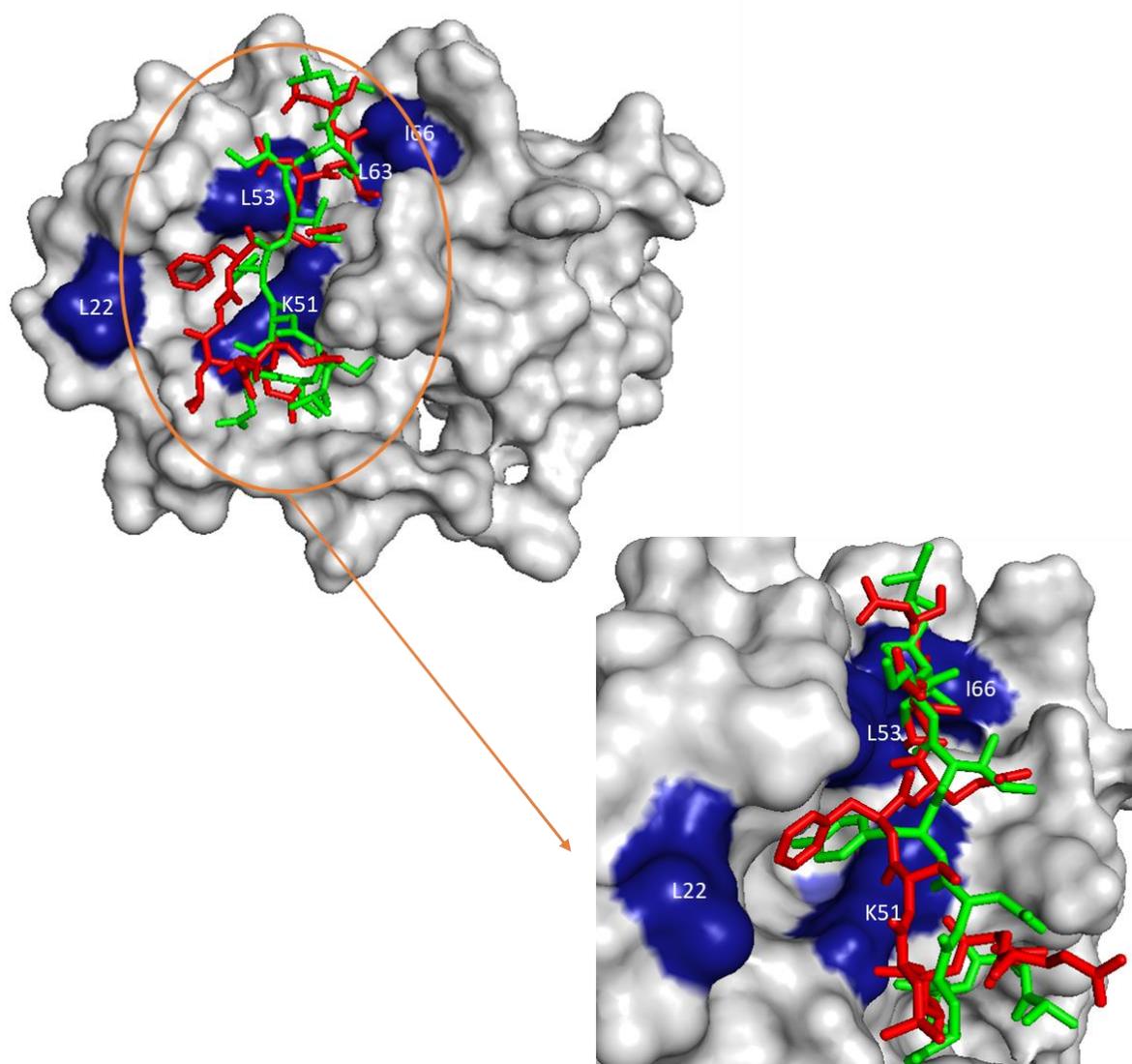


Figure 6.30: A: Graphical representation of CSPs that are significantly different between the titration with LC3B T29D and NBR1 LIR1 (blue) and NBR1 LIR2 (orange). B: Surface of LC3B T29D with significant differences between titration of LC3B T29D with NBR1 LIR1 and LIR2. Significant and most significant differences are coloured light blue and blue respectively. HADDOCK structure for LIR position in binding site are shown for LIR1 (green) and LIR2 (red). The top image shows the front face of LC3B T29D with the bottom image being side on.

The differences are located around hydrophobic pocket 1 and 2. For LIR2 the peak of L22 could not be followed in the titration, whereas there is a relatively small CSP for the LIR1 titration. This ties in with the HADDOCK structure showing Phe563 on LIR2 sitting across the face of HP1. This is then a different position of Tyr732 on LIR1 which sits more centrally in HP1, and hence the larger CSPs seen on K51 and L53 seen for the titration with LIR1. Leu566 sits much further into the HP2 than Ile735 on LIR1, and hence the large CSPs seen on L63 and I66 on LC3B T29D when titrated with LIR2.

The phosphorylation mimic T29D suggests that the phosphorylation at T29, has a large effect on the mode of binding of NBR1 LIR2, but little change in how LIR1 binds. However,

the data is not clear as to whether this is due to the T29D mutant altering HP2 and the interaction with Leu566 which shifts Phe563, or the opposite and it alters HP1 changing the interaction with Phe563, which has the knock on effect of shifting Leu566 or a bit of both. As the interaction of the aromatic with HP1 is the main driving force of the LIR binding to the LC3B this is the most likely, but some structural studies would be required to clarify this (crystal or NMR structure).

6.5 Summary

This chapter has dealt with the NMR titrations of NBR1 LIR1 and LIR2 with the LC3B mutants KL and T29D. Below is a summary of the key points from the chapter

- An assigned ^1H - ^{15}N HSQC for T29D and KL mutant was produced. For the T29D mutant this was a case of just transferring the assignment from the wild type spectrum. However, for KL this was not possible and so 3D data was acquired and the ^1H - ^{15}N HSQC assigned.
- Titrations for both mutants with NBR1 LIR1 and LIR2 were carried out and HADDOCK structures were generated showing the LIRs binding to the canonical binding site on LC3B.
- As well as this “standard” orientation of LIR for the LIR2 titration data HADDOCK produced a relatively high scoring “reverse” orientation with both LC3B WT, KL and T29D with the aromatic in the LIR in HP2 and the hydrophobic residue in HP1. Whilst it is unlikely that this is how the LIR is actually binding it shows that the acidic residues missing from LIR2 and the change in aromatic then makes a HADDOCK structure with the reverse configuration possible.
- The CSPs from the interaction of LIR1 with LC3B KL are on the whole less than that for the wild type titration. The HADDOCK structures show that LIR1 is closer to the face of LC3B WT than LC3B KL. The mutation of lysine 51 to an alanine has reduced how basic the face of LC3B KL is and so the interaction with LIR1 is less stabilised by the preceding acidic residues than that for the wild type.
- For The titration LC3B KL with NBR1 LIR2 has a very similar data to the titration with LC3B WT. However, changing the charge of the face of LC3B allows L560 on LIR2 to sit closer to the face of LC3B.
- The interaction of the T29D mutant with LIR1 has changed very little from the interaction with LC3B WT indicating that phosphorylation at this site would not have a major effect on the interaction with NBR1 LIR1, at least at the structural level.
- The interaction of NBR1 LIR2 with LC3B T29D shows Phe563 sitting across the face of HP1 and Leu566 sitting tight into HP2. This interaction is different from how LIR2 sits on LC3B WT and how LIR1 interacts with T29D. Hence the T29D mutation can alter the binding of LC3B with LIRs but it is based on the sequence of LIR interacting with LC3B.

Chapter 7

7. Interaction of LC3A with the LIRs of NBR1

This chapter aims to investigate further the preference of the two LIRs of NBR1 to bind to LC3A over LC3B discovered in Chapter 4. This will be examined through NMR titration experiments, which will allow the binding site of the LIRs on LC3A to be mapped. The amino acids that are shown to play a significant part in the interaction will then be used with the HADDOCK webserver to produce a structure of the LIRs bound to LC3A. The CSPs and HADDOCK structure of the interaction of the NBR1 LIRs with LC3A will then be compared to those of LC3B, to better understand the preferential binding observed.

From the MS competition experiments carried out in Chapter 4, it was interesting to note that for both LIRs of NBR1, there was a preference to bind to LC3A over LC3B, with roughly double the amount of LC3A-LIR to LC3B-LIR complex present.

LC3A has the most similar structure to LC3B of all the ATG8 proteins with there being 82.5% sequence alignment (Figure 7.1).

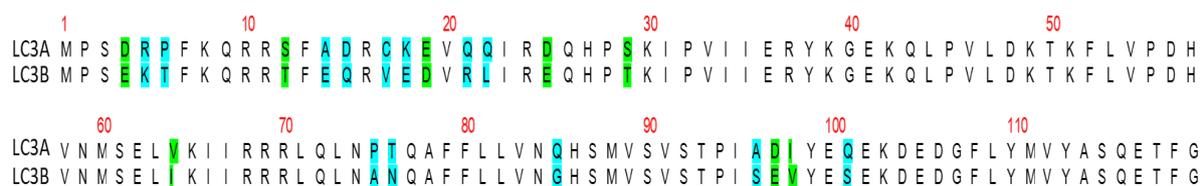


Figure 7.1: Sequences of LC3A and LC3B. Differences in the sequences are highlighted with green denoting a change to a similar type of amino acid and blue denoting a change to an amino acid of a different type.

Of the 21 differences in sequence 8 are a change to an amino acid of the same type (serine for threonine, aspartate for glutamate, asparagine for glutamine etc.). Overlaying the structures of LC3A and LC3B show that the two proteins structures are very similar (Figure 7.2).

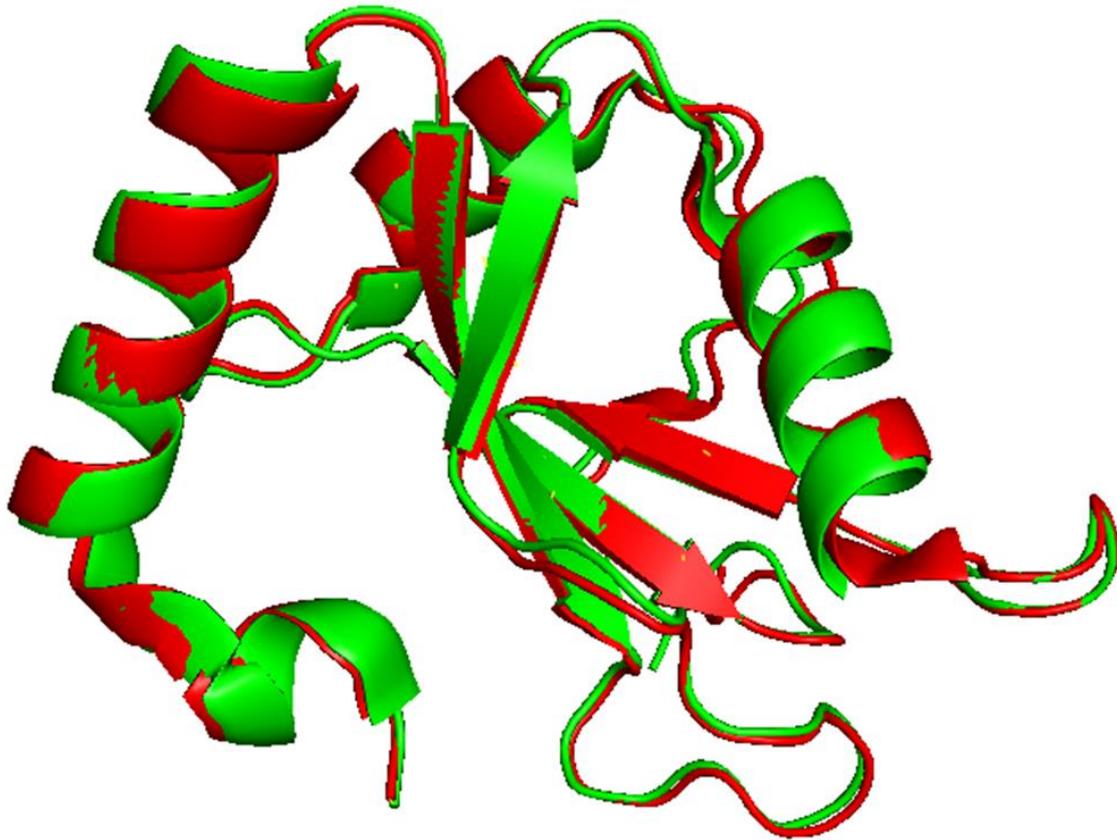


Figure 7.2: Overlay of LC3A and LC3B Structure. LC3A is shown in red (PDB structure 3WAL) and LC3B is shown in green (PDB structure 3VTU).

When mapping the changes in sequence on to the structure of LC3B (Figure 7.3), it can be seen that the majority of the differences appear on the first two α -helices that pack against the central beta sheet.

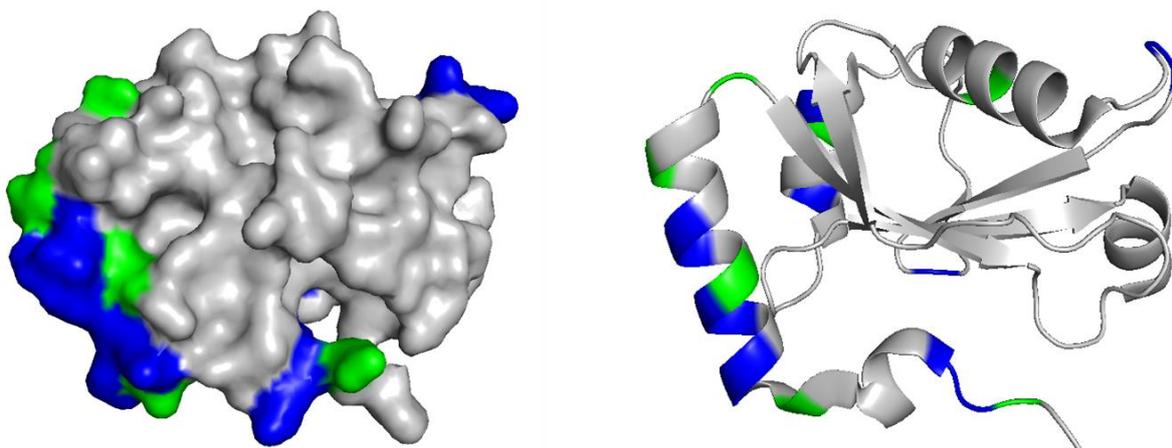


Figure 7.3: Differences between LC3A and LC3B sequences mapped onto the structure of LC3B. Differences in the sequences are highlighted with green denoting a change to a similar type of amino acid and blue denoting a change to an amino acid of a different type.

As the first hydrophobic pocket, into which the aromatic residue of the LIR fits, is formed by the second alpha helix packing against the beta sheet, any changes to the amino acids that are involved could change the interaction with the LIR.

7.1 NMR Backbone Assignment of LC3A

Before the titrations could be carried out an assignment of the backbone amides in the ^1H - ^{15}N HSQC for LC3A was required. A uniformly ^{15}N and ^{13}C labelled LC3A protein sample was produced (Chapter 2.3.11), and a ^1H - ^{15}N HSQC and the standard set of forward and reverse 3D experiments (HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB and HN(CO)CACB) for assignment were acquired. The peaks in the ^1H - ^{15}N HSQC were then assigned using the standard method (Chapter 2.5.2.4), and an example of the linking of peaks in the HNCACB and HN(CO)CACB is given below (Figure 7.4).

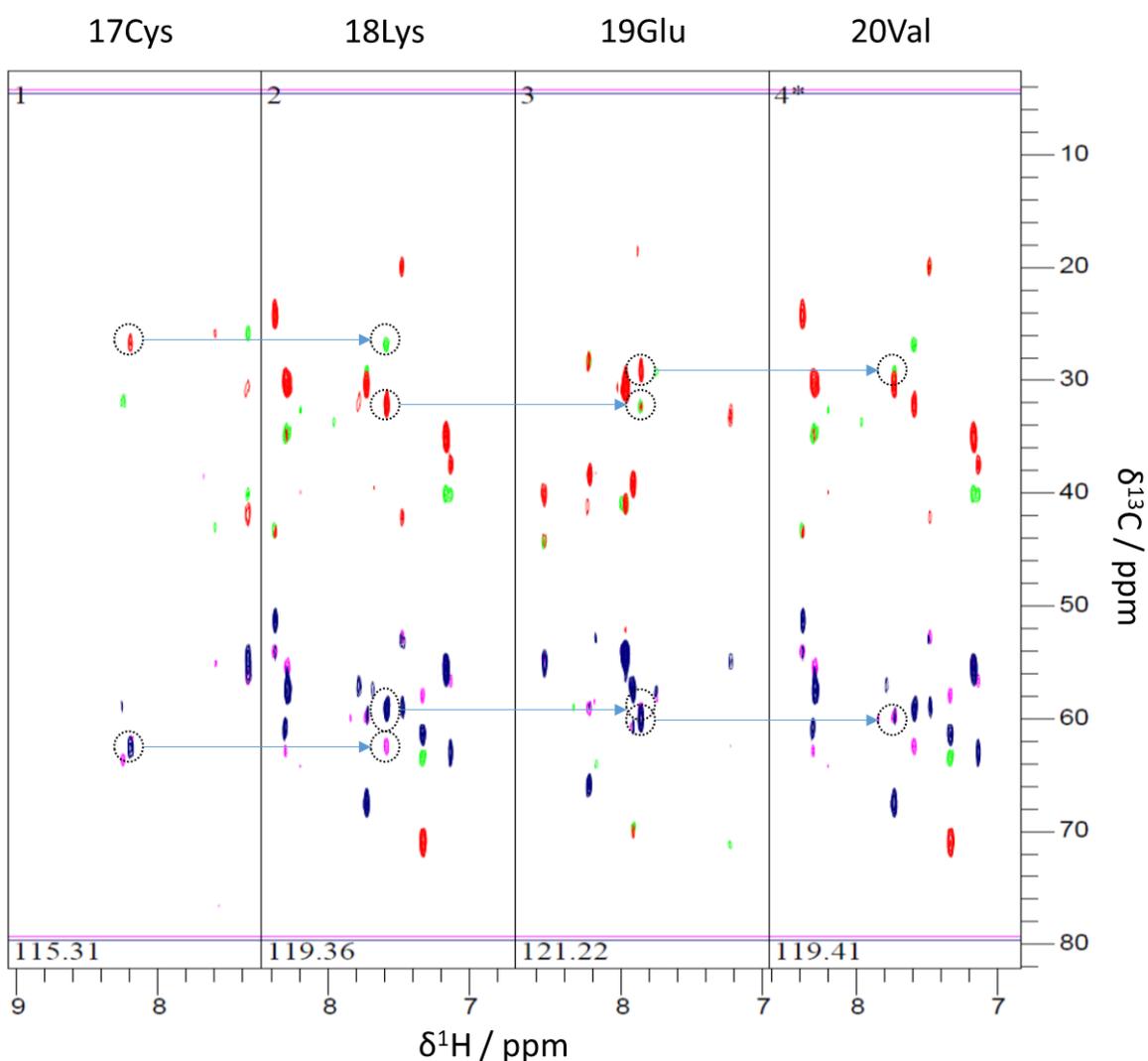


Figure 7.4: Example of linking C_α and C_β peaks using HNCACB (blue and red peaks) and HN(CO)CACB (pink and green peaks).

By following the assignment procedure it was possible to assign 101 out the 113 non-proline amino acids present in LC3A WT (Figure 7.5). This 90% assignment of the ^1H - ^{15}N HSQC would be complete enough to carry out NMR titrations with the LIRs of NBR1, especially as there were no large blocks missing. The missing amino acids were throughout the protein (S3, F7, K8, S12, G40, V83, N84, Q85, V89, V91, I95 and L109), and so would not prevent observation of the interaction through the canonical binding patch.

7.2 NMR Titrations of LC3A with the LIRs of NBR1

7.2.1 Titrations with NBR1 LIR1

As per the previous NMR titrations, a peptide with the sequence of ASSEDIYIILPE was used to represent LIR1 of NBR1, which was then titrated with full length (1-121) ^{15}N enriched LC3A WT. LIR1 peptide was titrated into the protein sample and a ^1H - ^{15}N HSQC acquired until a ratio of 2:1 of peptide to protein was reached (Figure 7.6).

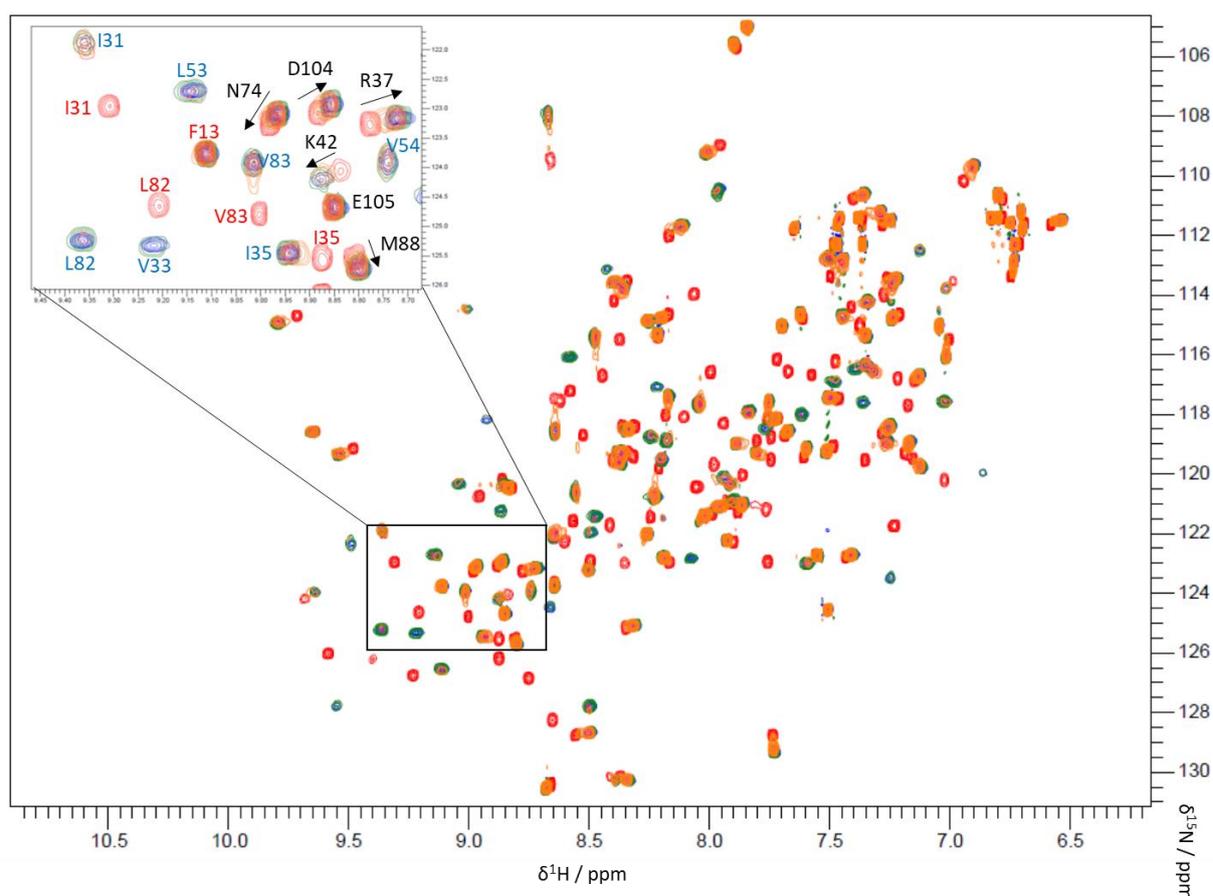


Figure 7.6: A selection ^1H - ^{15}N HSQC's from the titration of NBR1 LIR1 Peptide with LC3A WT. Where red is 0 equivalents of peptide and blue is 2 equivalents. Orange, pink and green are 0.3, 0.5 and 0.8 equivalents of peptide respectively.

As with the titrations of LC3B a large number of peaks on the LC3A ^1H - ^{15}N HSQC moved. The majority of the peaks that moved did so under fast and intermediate exchange regimes, and so it was possible to follow the majority peaks in the titration. Taking the positions of the peaks in the ^1H - ^{15}N HSQC for 0 and two equivalents of LIR (start and finish of the titration), it was then possible to calculate CSPs for the amino acids (Figure 7.7), with some large CSP effects of up to 1.4 ppm observed.

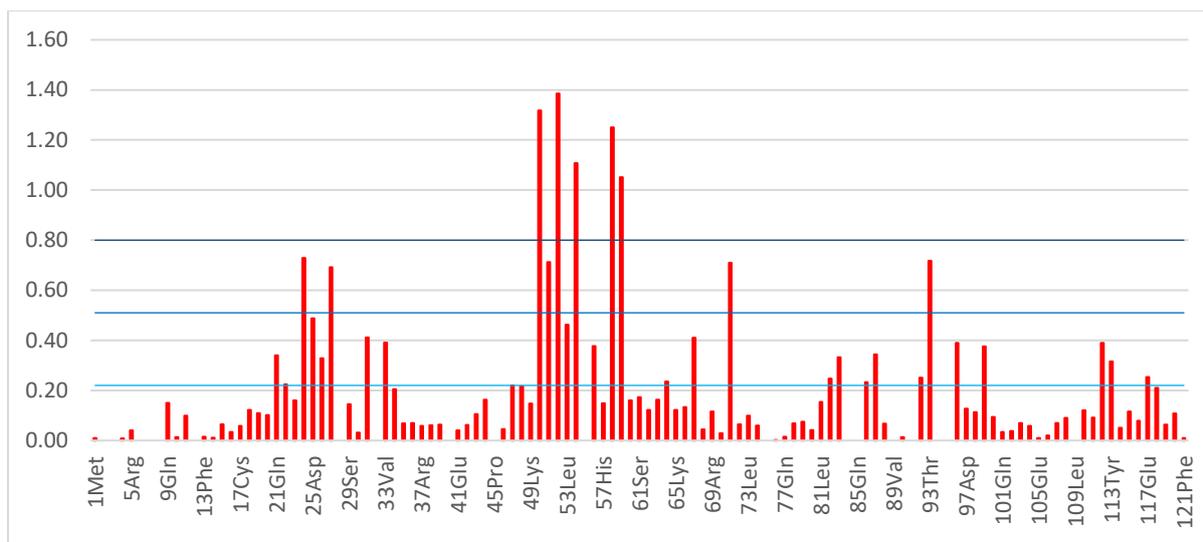


Figure 7.7: Plot of CSPs for titration of LC3A WT with peptide representing LIR1 of NBR1. The light blue line is the mean value, the blue is the mean plus one standard deviation and the dark blue line is the mean and two standard deviations.

To focus on those amino acids that were most involved in LIR1 binding to LC3A the peaks which had the most significant CSPs were selected (Chapter 2.5.2.6). These significant amino acids were then used to colour the structure of LC3A to show the areas of largest perturbation, and allow visualisation of the binding site (Figure 7.8).

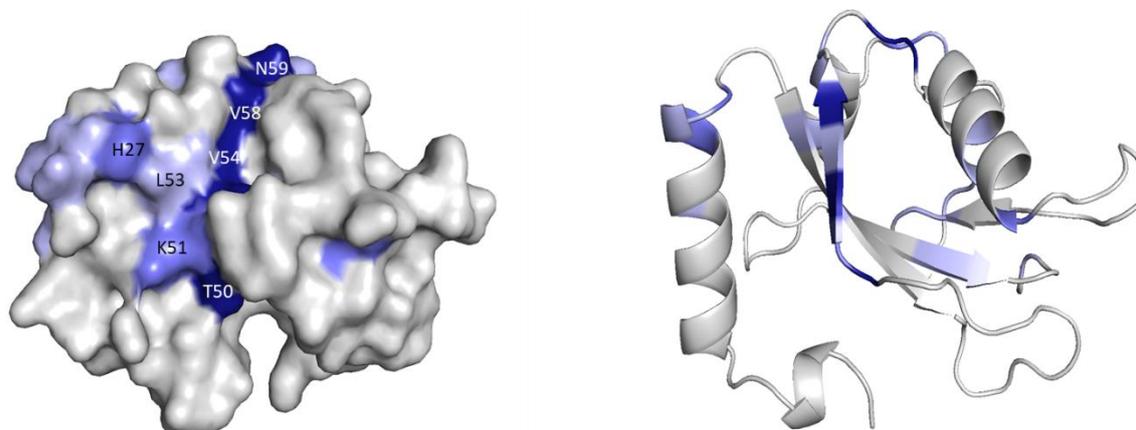


Figure 7.8: CSPs for NMR titration of LC3A with peptide representing NBR1 LIR1 mapped to structure of LC3A. Amino acids with CSPs with a value greater than the mean, mean and standard deviation and mean and two standard deviations are coloured light blue, purple and dark blue respectively.

The most significant perturbations on LC3A are seen by amino acids T50, V54, V58 and N59, which run up the centre of the front face of LC3A, between the two hydrophobic pockets. There are also significant shifts on H27 and K51 that surround HP1. This seems to indicate that LIR1 of NBR1 bind to the face of LC3A in a similar way to that of LC3B, with the aromatic Tyr733 and Ile736 sitting in the hydrophobic pockets on the front face of LC3A.

To further investigate how LIR1 of NBR1 interacts with LC3A HADDOCK server was used to dock LIR1 onto LC3A (Chapter 2.5.2.7). The HADDOCK run produced 169 structures in 9

clusters with the top scoring cluster containing 76 of these structures. However, the HADDOCK score for the top structure was only -72.1 ± 5.1 , which was not particularly high, and when the structure was analysed it did not show the expected structure for the LIR1 peptide binding to LC3A (Figure 7.9).

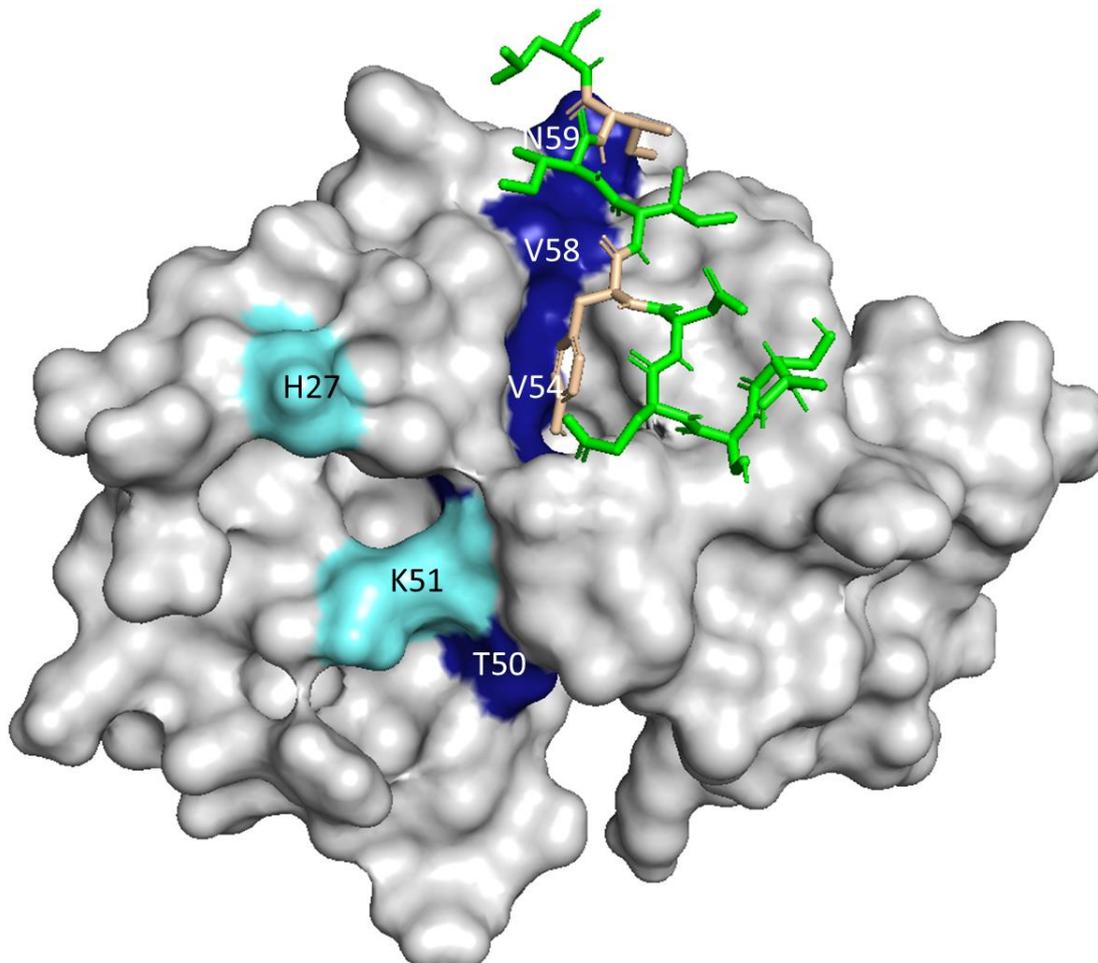


Figure 7.9: Highest scoring HADDOCK structure of NBR1 LIR1 peptide bound to LC3A. The peptide is coloured green with the active residues Tyr733 and Ile736 coloured yellow. The surface of LC3A is coloured to indicate the significant shifts seen in the bound form. Residues coloured light blue had CSPs greater than the mean and one standard deviation and those coloured blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.

Instead of the expected structure where Tyr733 binds to HP1 and Ile736 binds to HP2, we see a structure where Tyr733 sits in HP2, and the rest of the peptide lies across the top of LC3A covering V58 and N59. On analysing the structures in the other Haddock clusters, one cluster did produce a structure more similar to the canonical structure (Figure 7.10).

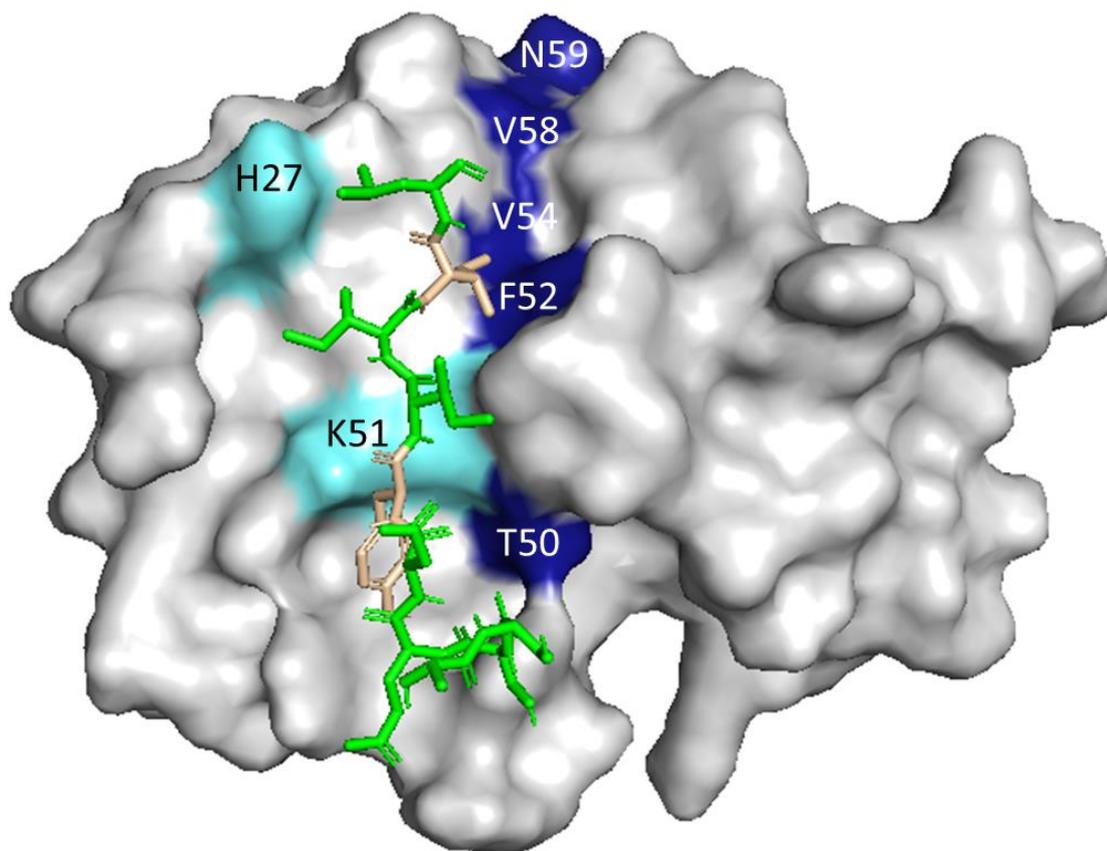


Figure 7.10: Alternative HADDOCK structure of NBR1 LIR1 peptide bound to LC3A. The peptide is coloured green with the active residues Tyr733 and Ile736 coloured yellow. The surface of LC3A is coloured to indicate the significant shifts seen in the bound form. Residues light blue had CSPs greater than the mean and one standard deviation and those coloured blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.

However, the HADDOCK score for this structure was only -42.8 ± 4.7 and the Tyr733 is not actually sitting in HP1 but on the surface at the bottom of LC3A. To further probe why HADDOCK was producing these structures rather than the canonical LIR binding interaction, a number of HADDOCK runs were carried out where the active and passive residues were varied. A summary of the residues used and the results of these runs can be found in Table 7.1.

| Run | 1 | 2 | 3 | 4 | 5 |
|---|----------------------------------|----------------------------------|----------------------------------|-------------------------------|----------------------------------|
| LC3 Structure | LC3A | LC3A | LC3A | LC3B | LC3A |
| NMR titration CSPs used | LC3A + LIR1 | LC3A + LIR1 | LC3B + LIR1 | LC3A + LIR1 | LC3A + LIR2 |
| Active residue | Mean + 2 x standard deviation | Mean + 1 x standard deviation | Mean + 2 x standard deviation | Mean + 2 x standard deviation | Mean + 2 x standard deviation |
| Passive residues | Mean + 1 X standard deviation | Mean | Mean + 1 X standard deviation | Mean + 1 X standard deviation | Mean + 1 X standard deviation |
| Top structure HADDOCK Score and number of structures | -72.1±5.1 76/169 | -62.9±5.9 30/168 | -63.9±2.0 57/156 | -97.1±2.8 63/179 | -69.9±1.9 93/155 |
| Location of LIR1 on top score | Aromatic in hydrophobic pocket 2 | Aromatic in hydrophobic pocket 2 | Aromatic in hydrophobic pocket 2 | Expected canonical structure | Aromatic in hydrophobic pocket 2 |
| Canonical structure present | No | Yes -57.3±3.1 21/168 | No | Yes | No |

Table 7.1: Summary of Haddock runs carried out to probe interaction of NBR1 LIR1 peptide with LC3A.

For HADDOCK run 2 a greater number of residues were used for the active residues (any CSP with a value greater than the mean and one standard deviation) and passive residues (CSP value greater than the mean). This would then more clearly define the regions of interaction. However, as with run 1, the top scoring HADDOCK structure was with Tyr738 in HP2. With this set of active and passive residues there was a cluster of structures that had the LIR peptide binding in the standard LIR-LC3 fashion, and with a similar HADDOCK score to the top scoring aromatic in HP2 structure. However, the HADDOCK scores were relatively low. Run 3 used the CSPs from the titration of LC3B WT with NBR1 LIR1 to see if it was an issue with the CSPs used or with the protein structure. Again the highest scoring cluster contained structures where Tyr738 interacts with HP2. This implies that the change in sequence between LC3A and LC3B is the main cause in HADDOCK producing a structure of LIR1 bound to LC3A that is different from the normal binding site.

To confirm that the difference in sequence between LC3A and LC3B is the main influence on the alternative binding structure produced by HADDOCK, run 4 used the active and passive residues from the LC3A WT with NBR1 LIR1 titration, but with the structure of LC3B. This run gave the top scoring cluster as a series of structures with standard interaction between LC3s and an LIR (aromatic in HP1 and hydrophobic amino acid in HP2). The HADDOCK score for the top structure was also comparable to those for the actual HADDOCK run of LC3B WT with NBR1 LIR1. This gave further confirmation that the difference in sequence between

LC3A and LC3B is the major contributing factor to the difference in HADDOCK binding structures.

Finally for completeness run 5 uses the active and passive residues for the titration of LC3A with NBR1 LIR2 (results to be discussed in the following section). This gives a similar result to run1 and again suggests the CSPs used to select the active and passive residues are not a major contributing factor to the structure produced.

Whilst the HADDOCK structures produced do not necessarily mean that this is how NBR1 LIR1 binds with LC3A, and looking at the CSPs mapped on the surface of LC3A (figure 7.9) it is likely that LIR1 does bind with LC3A in the standard way. It does however, show that there are some interesting differences in the structures of LC3A and LC3B that will influence the binding interactions with LIRs.

7.2.2 Titrations with NBR1 LIR2

For the NMR titrations of LC3A WT with LIR2 of NBR1, the same peptide that had been used with the mutants of LC3B was used, with the sequence AQDLLSFELLD. This was titrated into a sample of ^{15}N enriched LC3A WT until a concentration of two equivalents of peptide to protein was reached. At each point in the titration a ^1H - ^{15}N HSQC was acquired which would be used to follow the peak movements during the titration (Chapter 2.5.2.3).

On overlaying the spectrum it was noticed that as with the previous titration with LIR1 that a large number of peaks had moved (Figure 7.11).

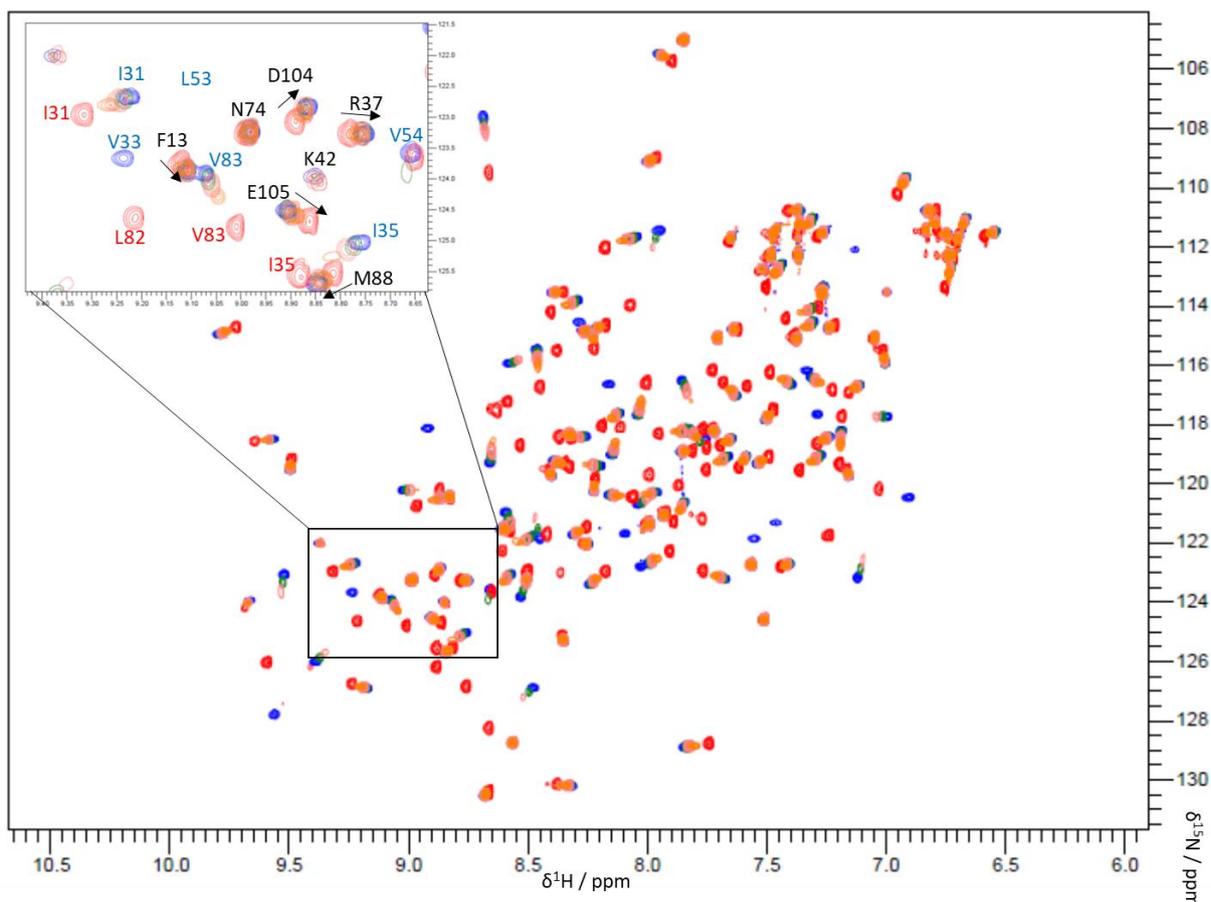


Figure 7.11: A selection ^1H - ^{15}N HSQC's from the titration of NBR1 LIR2 Peptide with LC3A WT. Where red is 0 equivalents of peptide and blue is 2 equivalents. Orange, pink and green are 0.3, 0.5 and 0.8 equivalents of peptide respectively.

For the titration of NBR1 LIR1 with LC3A WT the majority of the peaks had undergone intermediate exchange. However, for the titration with the LIR2 peptide, the vast majority of peaks underwent fast exchange, with only a few undergoing intermediate exchange. This made it possible to fully copy the assignment of LC3A WT with no LIR peptide to the end point of the titration, with two equivalents of peptide present. With the peaks in the end point ^1H - ^{15}N HSQC assigned, it was then possible to calculate CSPs for the amino acids and select those amino acids that had peaks with significant CSPs (Chapter 2.5.2.6). These CSPs and the significance cut offs were then plotted (Figure 7.12).

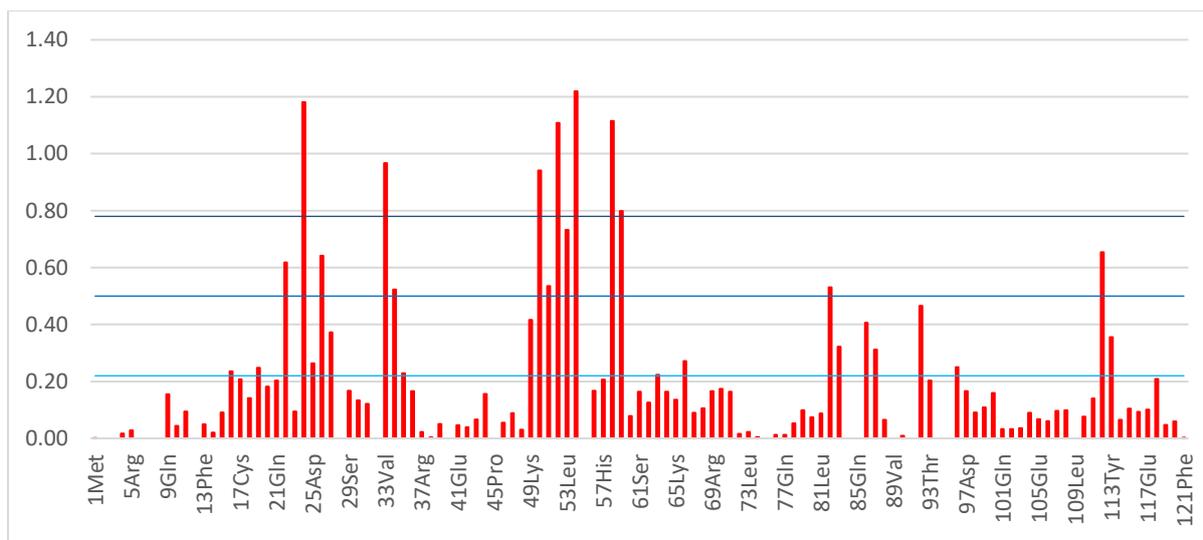


Figure 7.12: Plot of CSPs for titration of LC3A WT with peptide representing LIR2 of NBR1. The light blue line is the mean value, the blue is the mean plus one standard deviation and the dark blue line is the mean and two standard deviations.

With the amino acids that had the most significant changes selected, the structure of LC3A WT was then coloured to indicate where these were, which would hopefully then provide the location of the binding site (Figure 7.13).

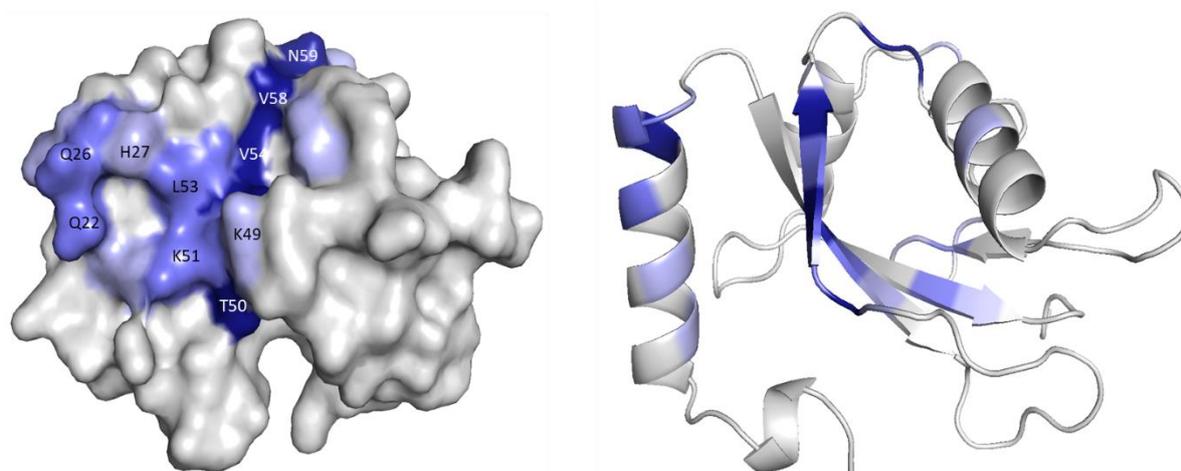


Figure 7.13: CSPs for NMR titration of LC3A with peptide representing NBR1 LIR2 mapped to structure of LC3A. Amino acids with CSPs with a value greater than the mean, mean and standard deviation and mean and two standard deviations are coloured light blue, purple and dark blue respectively.

Looking at the areas of large perturbation there are two regions where significant shifts are seen. The first sits around HP1 where Q22, Q26, L53 and K51 form a pocket into which the Phe563 on the LIR2 peptide can sit. There is then a band of amino acids with significant perturbations on V54, V58 and N59 that sit alongside the second hydrophobic pocket, where you would expect Leu566 on the LIR2 peptide to bind.

This data seems to fit the standard perturbations seen on the face of an LC3 when binding to an LIR. However, to get a better idea of how the peptide for NBR1 LIR2 binds to LC3A, HADDOCK webserver was used (Chapter 2.5.2.7). This HADDOCK run produced a total of 174 structures. The top scoring cluster had a HADDOCK score of -83.2 ± 5.7 and contained 60 structures. On analysing the top structure within this cluster the NBR1 LIR2 peptide is bound to the face of LC3A in the standard configuration (Figure 7.14).

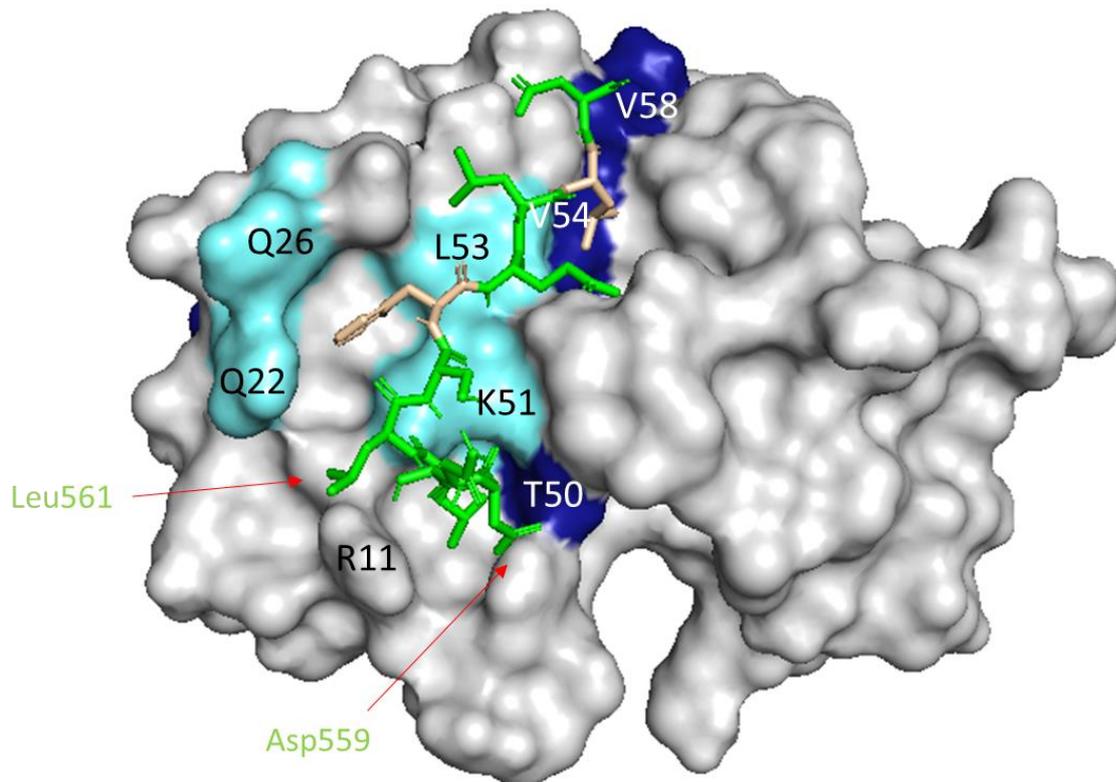


Figure 7.14: Highest scoring HADDOCK structure of NBR1 LIR2 peptide bound to LC3A. The peptide is coloured green with the active residues Phe563 and Leu566 coloured yellow. The surface of LC3A is coloured to indicate the significant shifts seen in the bound form. Residues light blue had CSPs greater than the mean and one standard deviation and those coloured blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.

This HADDOCK structure shows Phe563 on the LIR2 peptide binding into HP1, formed by Q22, Q26, K51 and L53. Leu566 then binds into the hydrophobic pocket surrounded by the active residue F52, V54 and V58. Asp559 on LIR2 then sits in close proximity to T50, potentially forming a hydrogen bond. Interestingly Leu561 also sits in a gap between Q22 and R11 on LC3A away from the surface.

7.3 Discussion

Using the standard set of 2D and 3D NMR experiments it was possible to assign 101 out of the 113 non-proline amino acids in the LC3A WT ^1H - ^{15}N HSQC, which corresponds to 90% of all amino acids. This assignment would be complete enough to carry out the titrations of LC3A with the NBR1 LIR1 peptides. Most importantly the amino acids that were not assigned were mostly in the core or rear of LC3A (Figure 7.15).

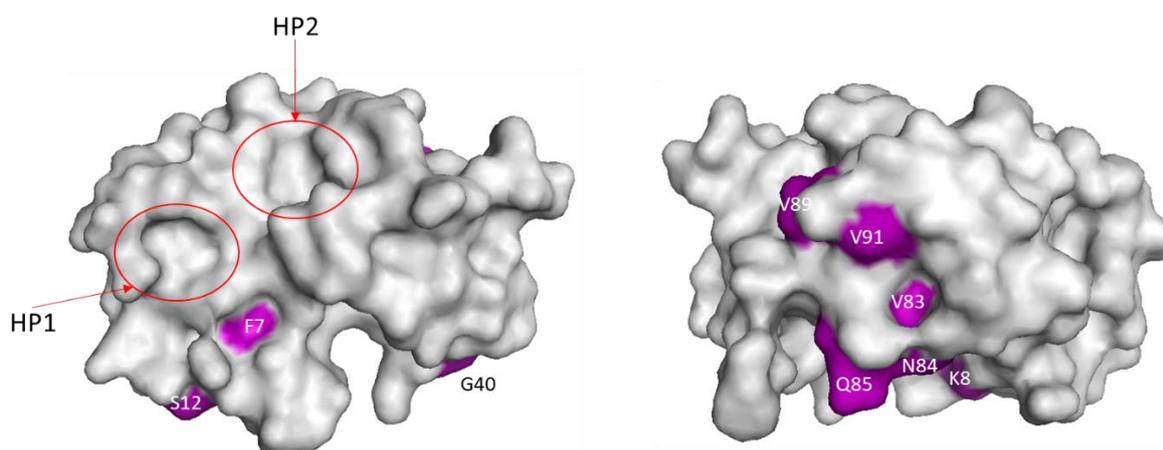


Figure 7.15: Surface of LC3A WT with amino acids that are un-assigned in the ^1H - ^{15}N HSQC shown in purple.

Hence, as the binding of the LIR peptides was anticipated to happen at the canonical binding site for the LC3 family, the critical amino acids for this interaction were assigned in the ^1H - ^{15}N HSQC, and so it would be possible to follow these in the titrations and compare with the interaction with LC3B.

7.3.1 Titrations of LC3A with the LIRs of NBR1

The ^1H - ^{15}N HSQC spectrum of LC3A WT showed a number of peaks moving on titration with the two LIRs of NBR1 indicating an interaction of the LIRs with the two LIR peptides. As with the previous titrations carried out between the LIR peptides and LC3B WT (Chapter 5), a large number of peaks moved for both titrations of LC3A WT (Figure 7.16).

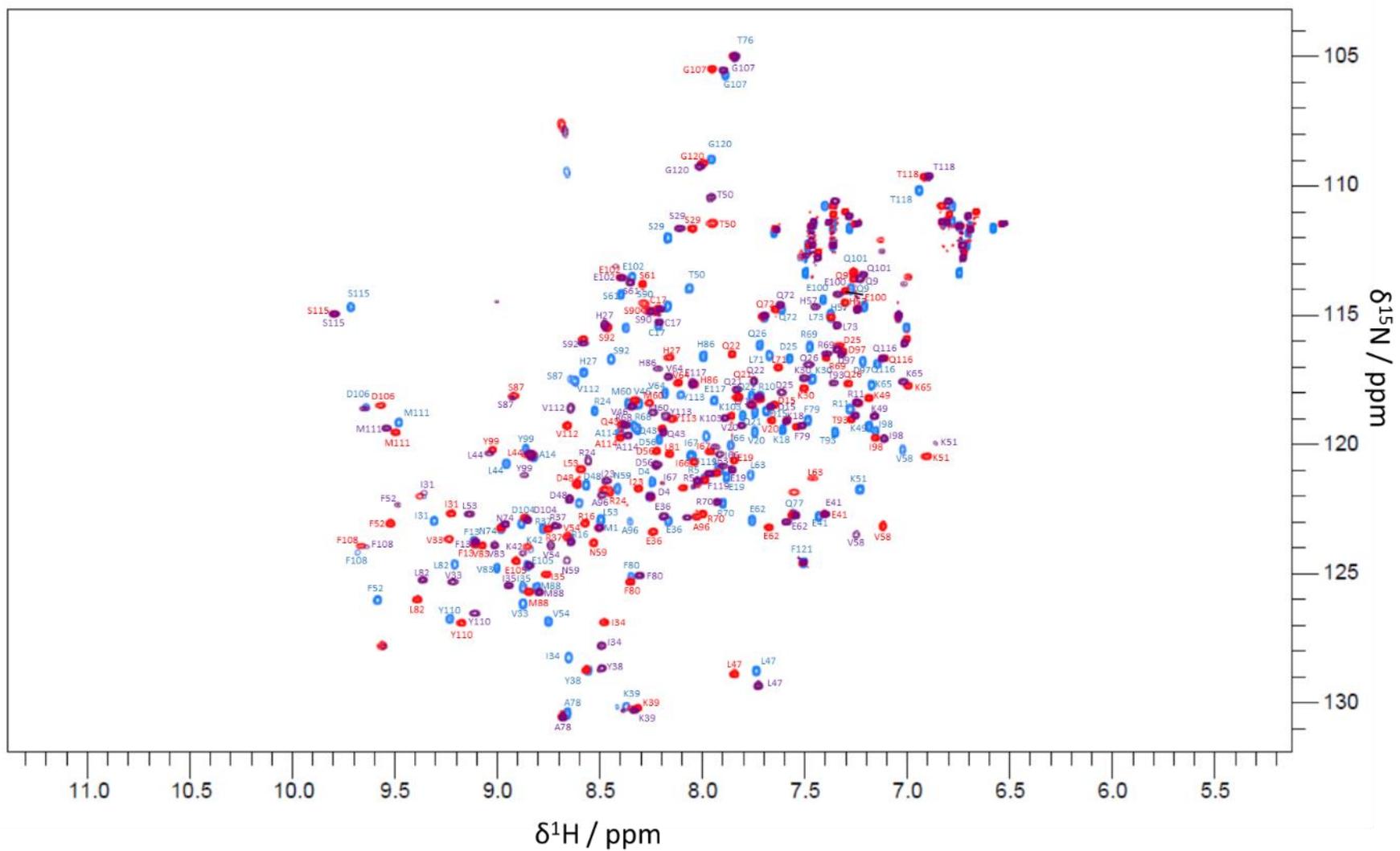


Figure 7.16: ^1H - ^{15}N HSQC of LC3A WT showing start (blue) and endpoints for titration with NBR1 LIR1 (purple) and LIR2 (red) peptides.

However, whilst the titrations were similar to those carried out with LC3B, in the wide scale peak movements in the ^1H - ^{15}N HSQC spectrum, the exchange regimes that the majority of peaks experienced were different. Whilst for the titrations of the NBR1 LIR peptides with LC3B the majority of peaks underwent slow and intermediate exchange, for the titrations with LC3A there was far less slow exchange, and in the case of the titration of LIR2 with LC3B almost all peaks underwent fast exchange. This indicates that the rate of dissociation (k_{off}) of the LIR peptides is slower for the interaction with LC3B than it is for that of LC3A, with the interaction of LC3A with LIR2 having the fastest rate of dissociation.

As a large number of peaks had moved in the HSQC for the titrations of the LIR peptides with LC3A the CSPs were analysed to focus on those that played the most significant part in the binding interaction. This analysis picked a set of significant peaks that indicated the main interaction was on the front face of LC3A at two sites. This seemed to indicate that the binding would be similar to that of the LIRs with LC3B, with the aromatic and hydrophobic residue on the LIR interacting with two hydrophobic pockets on the front face. However, for the titration of LC3A WT with NBR1 LIR1 when these significant amino acids were used in HADDOCK the top predicted structure did not show the LIR binding in the expected fashion to LC3A. Instead Tyr733 on LIR1 was interacting with HP2.

For the titration of LC3A WT with the NBR1 LIR2 peptide the top scoring HADDOCK structure was with the LIR interacting with the front face of LC3A in the expected orientation, with Phe738 sat in HP1 and Leu566 sat on top of HP2.

7.3.2 Comparison of LC3B and LC3A LIR1

To better understand why HADDOCK produced the structure it did for the interaction of LC3A with LIR1 the amino acids that had undergone significant shifts were compared with those of LC3B. At a glance the set of amino acids that had the most significant shifts and mapped area of interaction on the face of LC3A and LC3B appear to be very similar (Figure 7.17).

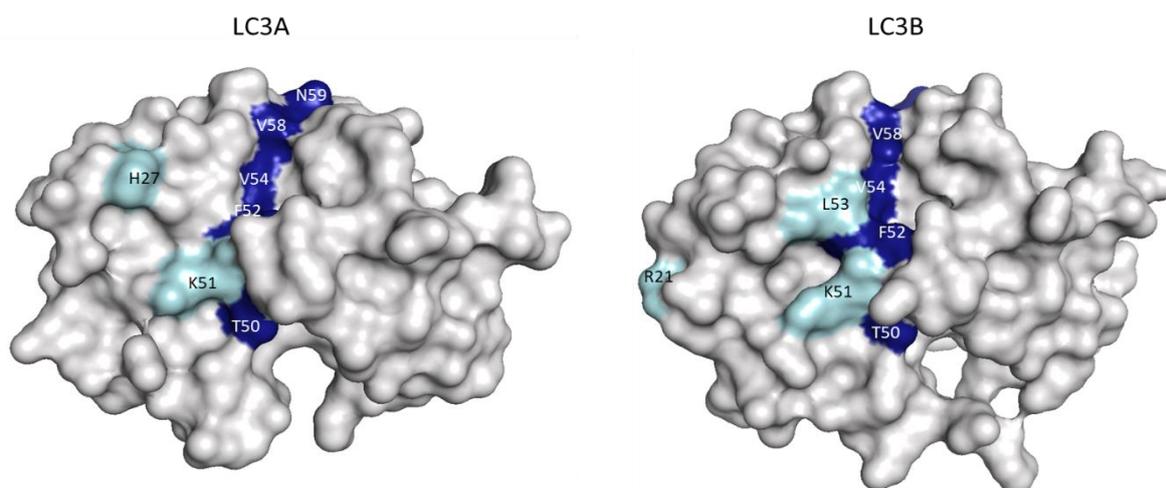


Figure 7.17: Significant CSPs mapped onto the protein surface for the titration of LC3A (left) and LC3B (right) with NBR1 LIR1 peptide. Significant shifts shown in light blue and most significant shifts shown in dark blue.

The most noticeable difference between the significant shifts is that L53 is missing for LC3A and for LC3B H27 is not significant.

To perform a more quantitative analysis of where the differences are between the two titrations with LIR1 the difference in CSPs for LC3A and LC3B were calculated and plotted (Figure 7.18)

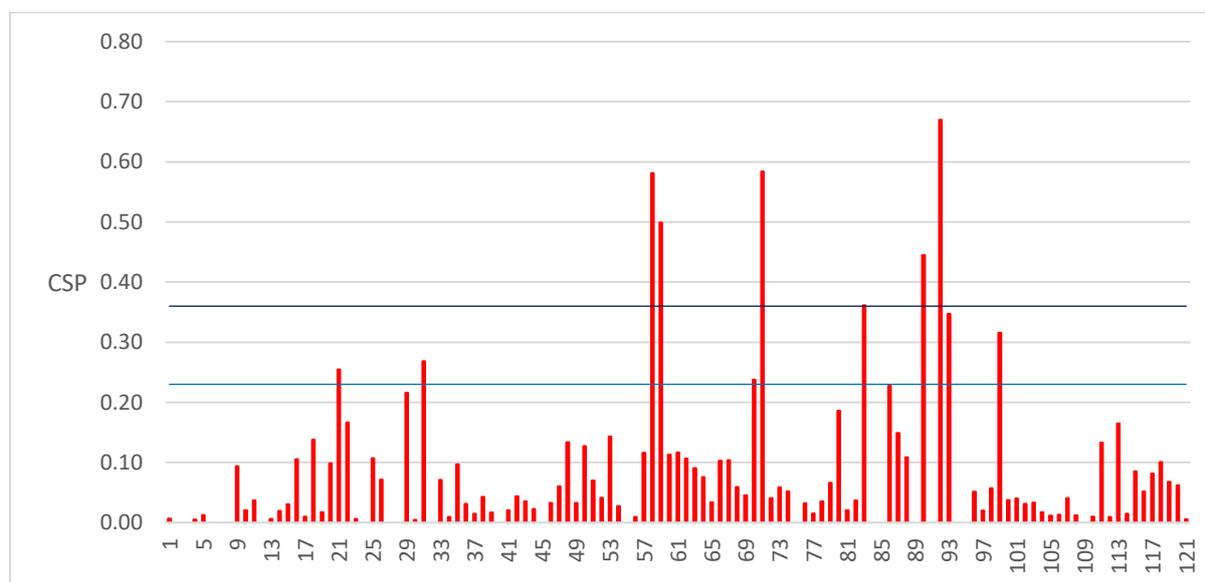


Figure 7.18: Plot of difference in CSPs between the titration of NBR1 LIR1 peptide with LC3A and LC3B. The blue is the mean plus one standard deviation and the dark blue line is the mean and two standard deviations.

The plot of the differences in CSPs between the two titrations suggest that most amino acids have a different value for the CSP. However, a lot of these appear to be relatively small. So to focus in on the main difference the most significant differences were selected (Chapter 2.5.2.6). A table of these significant amino acids is given below (Table 7.2).

Significant

| Sequence No. | LC3A amino acid | LC3B amino acid | LC3A CSP | LC3B CSP |
|--------------|-----------------|-----------------|----------|----------|
| 21 | Gln | Arg | 0.34 | 0.59 |
| 31 | Ile | Ile | 0.41 | 0.14 |
| 70 | Arg | Arg | 0.03 | 0.27 |
| 93 | Thr | Thr | 0.72 | 0.37 |
| 99 | Tyr | Tyr | 0.38 | 0.06 |

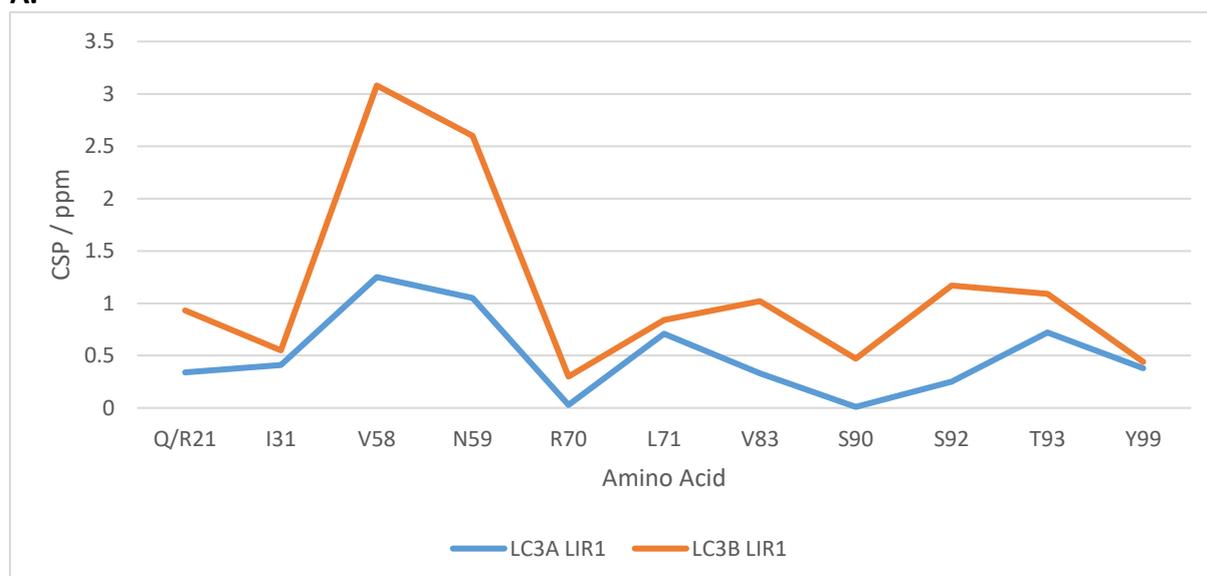
Most significant

| Sequence No. | LC3A amino acid | LC3B amino acid | LC3A CSP | LC3B CSP |
|--------------|-----------------|-----------------|----------|----------|
| 58 | Val | Val | 1.25 | 1.83 |
| 59 | Asn | Asn | 1.05 | 1.55 |
| 71 | Leu | Leu | 0.71 | 0.13 |
| 83 | Val | Val | 0.33 | 0.69 |
| 90 | Ser | Ser | 0.01 | 0.46 |
| 92 | Ser | Ser | 0.25 | 0.92 |

Table 7.2: Amino acids with significant differences in CSP between the titrations of NBR1 LIR1 peptide with LC3A and LC3B.

In total there are 11 amino acids that showed significant differences in CSP between the titration of LC3A and LIR1 and LC3B and LIR1. Of these only one is for an amino acid that is not the same between both LC3s and that is Q21 on LC3A and R21 on LC3B. This is a change in type of amino acid and so the difference in CSP could potentially be due to the change in amino acid. These significant differences were then plotted onto the surface of LC3B WT to visualise the areas in which the CSPs were different (Figure 7.19).

A.



B.

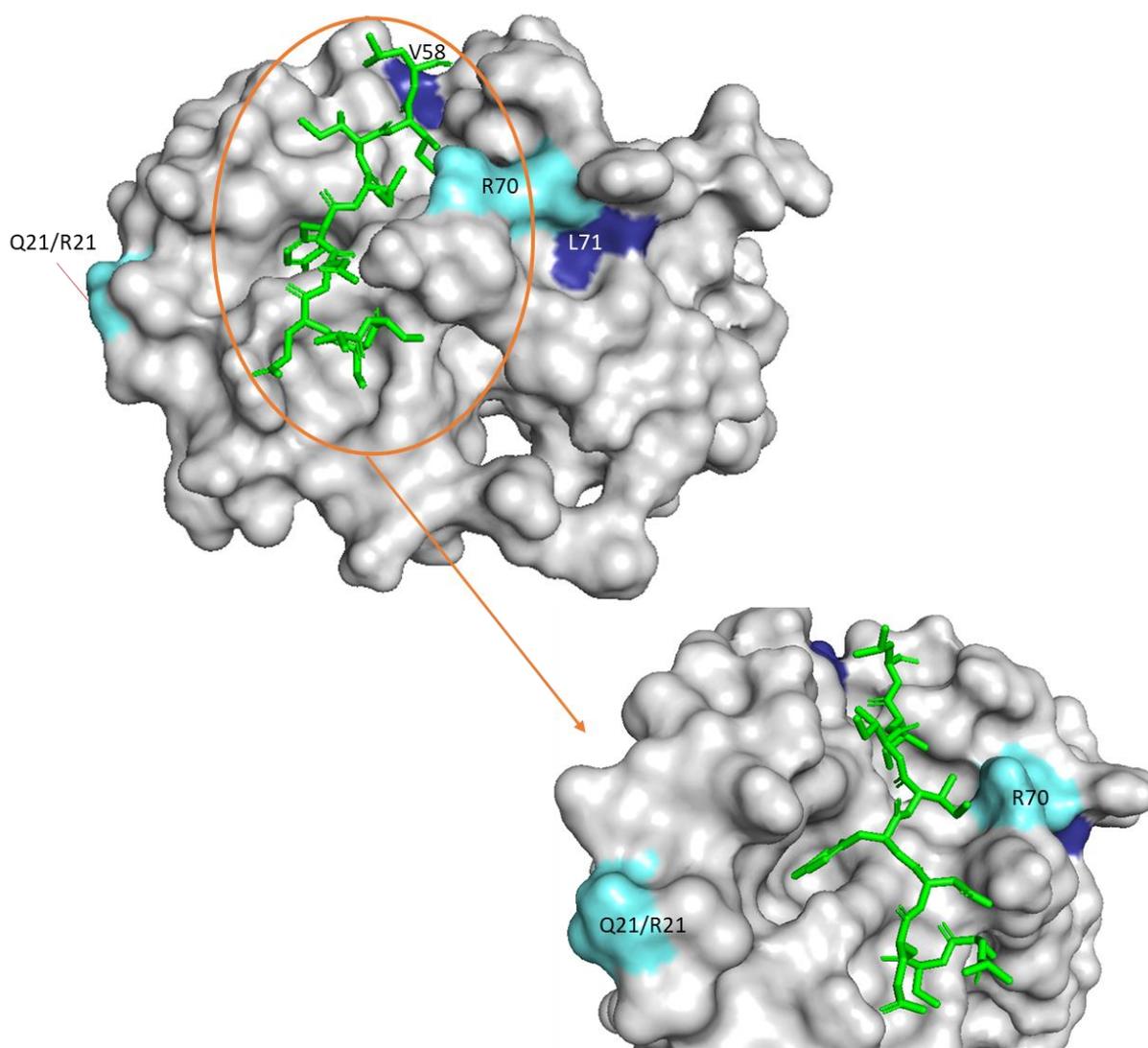


Figure 7.19: A: Graphical representation of CSPs that are significantly different between the titration with LC3A and NBR1 LIR1 (blue) and LC3B and NBR1 LIR1 (orange). B: Amino acids with significant CSP differences between titration of NBR1 LIR1 peptide with LC3A and LC3B mapped onto the surface of LC3B. NBR1 LIR1 peptide is shown as predicted from LC3B with LIR1 titration data. Significant shifts are coloured light blue and most significant shifts dark blue.

From mapping the significant difference onto the face of LC3B it can be seen that the major changes are seen around HP2, with V58, R70 and L71 having significant difference. However, apart from L71 the perturbations seen from the titrations of LC3B with LIR1 are larger than those for LC3A. There are then no major differences around HP1. So from the CSP data alone there seems to be no reason why HADDOCK would generate the structure it has. This seems to be confirmed by the HADDOCK runs carried out earlier in the chapter, where the CSPs for the titration of LC3A with LIR1 were used for a HADDOCK run with LC3B and LIR1, which generated the expected structure for the interaction.

It is thus likely that the structure predicted by HADDOCK is not correct and that NBR1 LIR1 binds to LC3A at the canonical binding site. However, there is a difference in the structure of LC3A that is forcing HADDOCK to not place Tyr733 in HP1.

Revisiting the differences between the sequences LC3A and LC3B at the start of the chapter (Figure 7.1 and 7.3) we can see that the majority of changes in the sequence between LC3A and LC3B appear on the first and second α -helix, which pack against the central β -sheet to form HP1. As the interaction of the aromatic residue with HP1 is the driving force for the binding of LIRs with the hATG8s, a modification to HP1 could have quite a large effect. This could then lead to the HADDOCK run generating the structure it has. Further structural studies (crystal/NMR structure) would be required to understand exactly how LIR1 binds to LC3A and understand the preference seen in Chapter 4.

7.3.3 Comparison of LC3B and LC3A LIR2

To better understand the selectivity seen in chapter 4, with almost double the amount of LC3A-LIR2 complex than LC3B-LIR2 complex present, the two NMR titrations of LC3A-LIR2 and LC3B-LIR2 would be compared.

As with NBR1 LIR1, at first glance the sets of active residues for the two titrations with LIR2 appear to be very similar, with the amino acids that show the most significant shifts located around the two hydrophobic pockets on the front face of LC3A and LC3B (Figure 7.20).

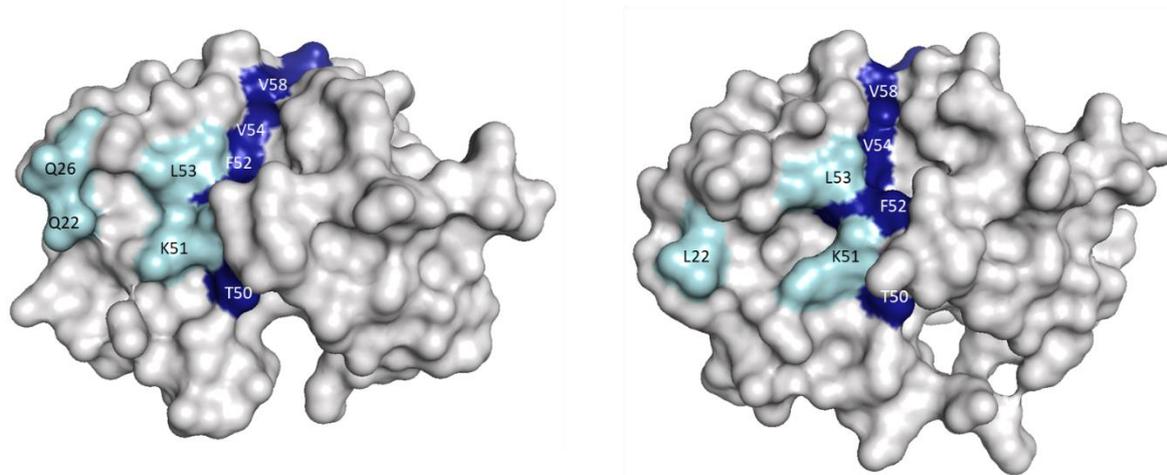


Figure 7.20: Significant CSPs mapped onto the protein surface for the titration of LC3A (left) and LC3B (right) with NBR1 LIR2 peptide. Significant shifts shown in light blue and most significant shifts shown in dark blue.

To further compare the interaction of LC3A with LIR2 and LC3B with LIR2 the difference in CSPs were calculated and plotted (Figure 7.21).

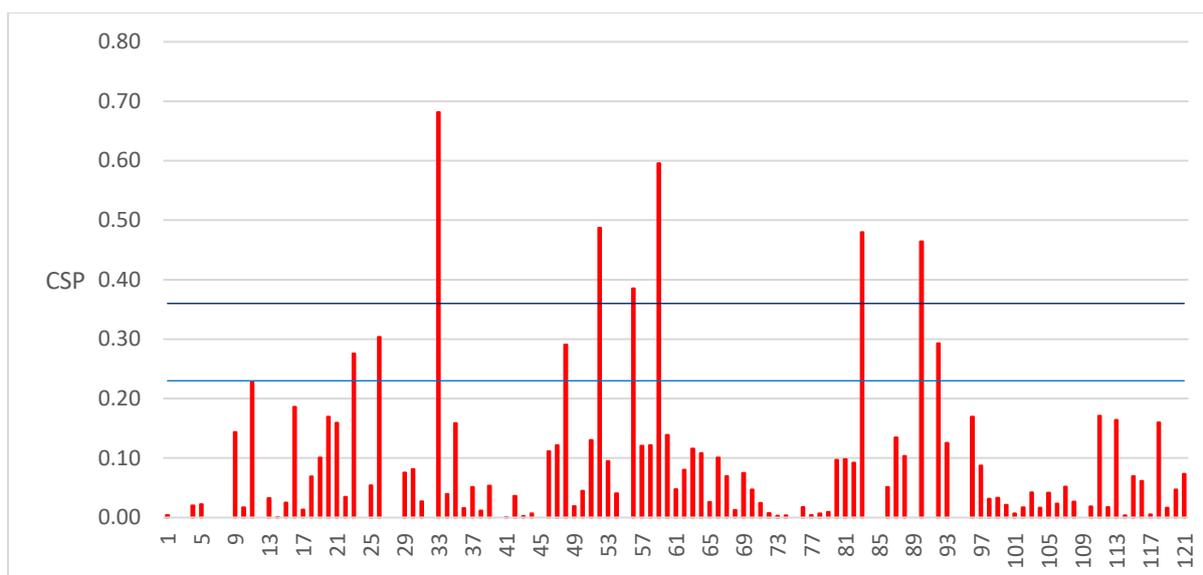


Figure 7.21: Plot of difference in CSPs between the titration of NBR1 LIR2 peptide with LC3A and LC3B. The blue is the mean plus one standard deviation and the dark blue line is the mean and two standard deviations.

From these calculated differences between the LC3A-LIR2 and LC3B-LIR2 titrations it can be seen that the majority of amino acids have a CSP that is different between the two titrations. The peaks that had the most significant differences were then selected (Chapter 2.5.2.6).

Using these significance cut off values there were 11 amino acids that had a CSP that was significantly different between the two titrations of LIR2 with LC3A and LC3B (Table 7.3).

Significant

| Sequence No. | LC3A amino acid | LC3B amino acid | LC3A CSP | LC3B CSP |
|--------------|-----------------|-----------------|----------|----------|
| 11 | Arg | Arg | 0.09 | 0.32 |
| 23 | Ile | Ile | 0.09 | 0.37 |
| 26 | Gln | Gln | 0.64 | 0.34 |
| 48 | Asp | Asp | 0.03 | 0.32 |
| 92 | Ser | Ser | 0.47 | 0.76 |

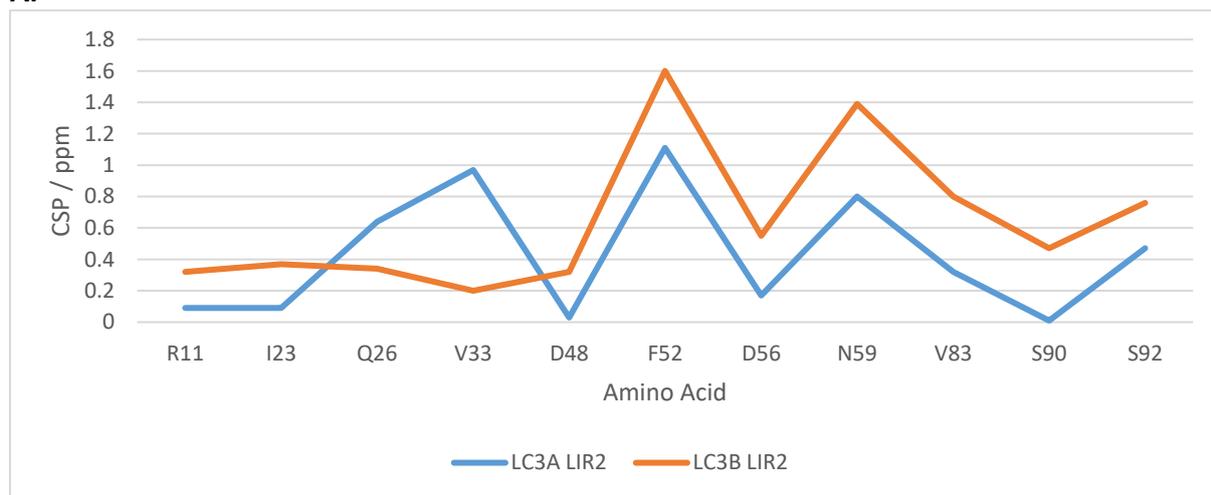
Most significant

| Sequence No. | LC3A amino acid | LC3B amino acid | LC3A CSP | LC3B CSP |
|--------------|-----------------|-----------------|----------|----------|
| 33 | Val | Val | 0.97 | 0.2 |
| 52 | Phe | Phe | 1.11 | 1.60 |
| 56 | Asp | Asp | 0.17 | 0.55 |
| 59 | Asn | Asn | 0.80 | 1.39 |
| 83 | Val | Val | 0.32 | 0.80 |
| 90 | Ser | Ser | 0.01 | 0.47 |

Table 7.3: Amino acids with significant differences in CSP between the titrations of NBR1 LIR2 peptide with LC3A and LC3B.

All of the significant differences between the LC3A and LC3B titration are for amino acids that are the same in both sequences. For the majority of the 11 amino acids the CSP is larger for the titration of LC3B with LIR2 than it is for LC3A and LIR2. The two exceptions to this are D48 and V33, for which the CSP is larger in the LC3A-LIR2 titration. To visualise where these amino acids appear on the LC3s the significant values were mapped onto a surface of LC3B, and the two HADDOCK structures for the positions of the LIR2 peptide were shown to see how they differ (Figure 7.22).

A.



B.

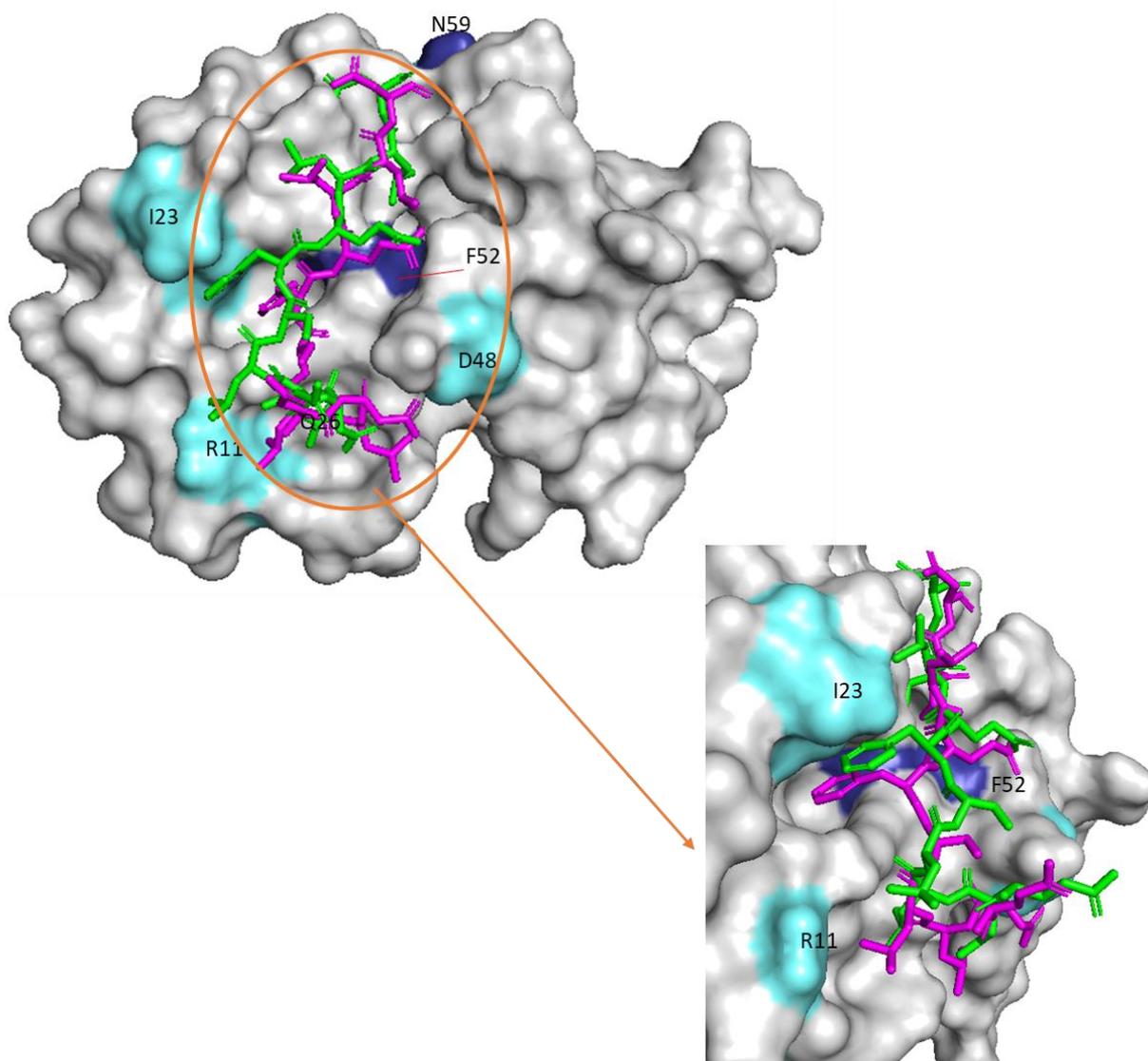


Figure 7.22: A: Graphical representation of CSPs that are significantly different between the titration with LC3A and NBR1 LIR2 (blue) and LC3B and NBR1 LIR2 (orange). B: Significant differences in CSP between LC3A-LIR2 and LC3B-LIR2 titrations mapped onto the surface of LC3B. HADDOCK structure of NBR1 LIR2 peptide shown bound to LC3A (green) and LC3B (pink). Amino acids with significant differences in CSPs are coloured light blue and most significant differences are coloured dark blue.

From these values on the HADDOCK structure it can be seen that Phe563 on LIR2 has shifted considerably for the titration with LC3A when compared with LC3B. This is accounted for by the position of HP1 on LC3A being shifted higher up on LC3A than on LC3B. This then has the effect of pulling L561 further up the face of LC3A and changing the CSP seen on R11 between LC3A and LC3B.

For HP2 there are no calculated significant differences between the CSPs of the two titrations, and looking at the HADDOCK structures calculated they are in similar positions, with Leu566 on the LC3B LIR2 peptide only slightly rotated round from the LC3A peptide.

This change in position for the aromatic residue has the knock on effect of moving the preceding amino acids to a different position on the face of LC3A, which could lead to the hydrophobic amino acids in LIR2 being in a more favourable environment. This in turn could explain why LC3A was a more favourable binding partner for LIR2 than LC3B.

7.3.4 Comparison of LIR1 and LIR2 with LC3A

Having compared the two titrations of the NBR1 LIRs against LC3A with the titrations against LC3B, finally the titration of LIR1 and LIR2 with LC3A would be compared. Looking at the two endpoints for the titrations in the ^1H - ^{15}N HSQC a large number of the peaks move in a similar way. Then comparing the two sets of active residues we can see that the amino acids that have significant CSPs are very similar between the two titrations (Figure 7.23).

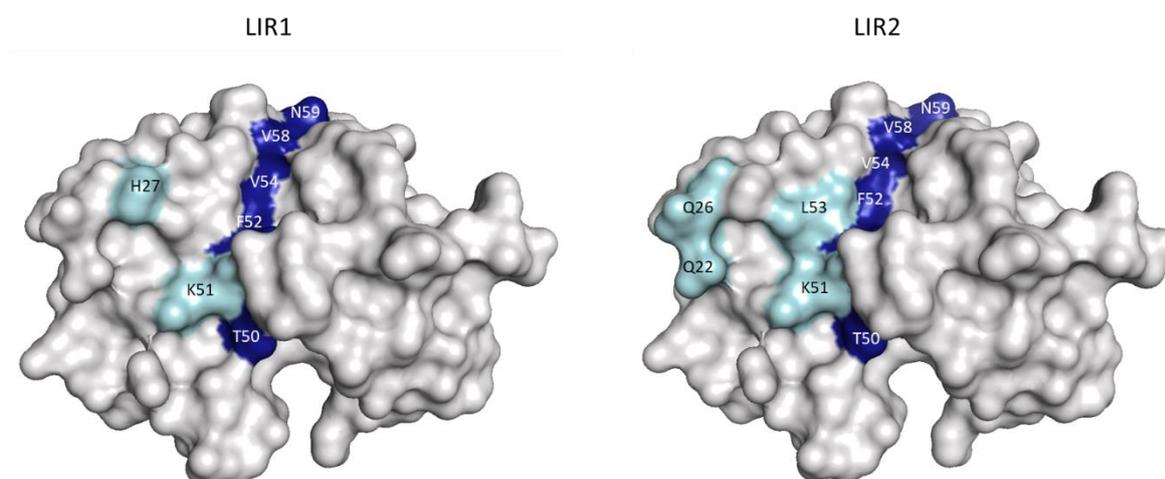


Figure 7.23: Significant CSPs mapped onto the protein surface for the titration of NBR1 LIR1 (left) and LIR2 (right) with LC3A WT. Significant shifts shown in light blue and most significant shifts shown in dark blue.

The main difference appear to be around HP1, where the titration with LIR2 shows significant shifts at Q26, Q22, K51 and L53, whereas the titration with LIR1 only has significant shifts with H27 and K51. From the predicted HADDOCK structure for the titration with LIR2, Phe563 on the LIR sits across the top of the HP1, which potentially explains the CSPs seen on Q26 and Q22. For LIR1 the structures predicted do not even show Tyr733 sitting in the pocket.

To understand the differences between the two titrations and potentially what is happening with LIR1-LC3A interaction the two sets of CSPs for the titrations were compared. The difference between the two CSPs was calculated and plotted (Figure 7.24)

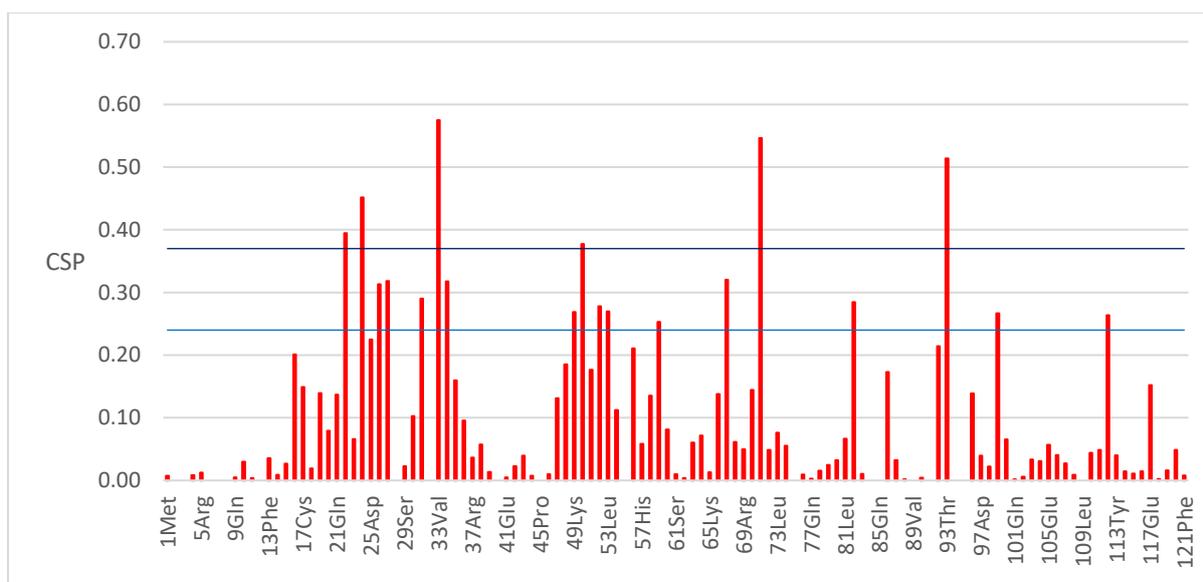


Figure 7.24: Plot of difference in CSPs between the titration of LC3A WT with NBR1 LIR1 and LIR2 peptide. The blue line is the mean plus one standard deviation and the dark blue line is the mean and two standard deviations.

From the differences in CSPs between the two titrations it can be seen that the majority have CSPs that vary between the interaction of LIR1 and LIR2 with LC3A. As previously the data was further analysed to pick out those differences that are significant (Chapter 2.5.2.6). With this analysis carried out it was found that there were 18 amino acids that had significant differences in CSP between the titrations, and of those six selected as being most significant (Table 7.4).

Significant

| LC3A WT amino acid | LIR1 CSP | LIR2 CSP |
|--------------------|----------|----------|
| 26 Gln | 0.33 | 0.64 |
| 27 His | 0.69 | 0.37 |
| 31 Ile | 0.41 | 0.12 |
| 34 Ile | 0.20 | 0.52 |
| 49 Lys | 0.15 | 0.42 |
| 52 Phe | 1.39 | 1.11 |
| 53 Leu | 0.46 | 0.73 |
| 59 Asn | 1.05 | 0.80 |
| 67 Ile | 0.41 | 0.09 |
| 82 Leu | 0.25 | 0.53 |
| 99 Tyr | 0.38 | 0.11 |
| 112 Val | 0.39 | 0.65 |

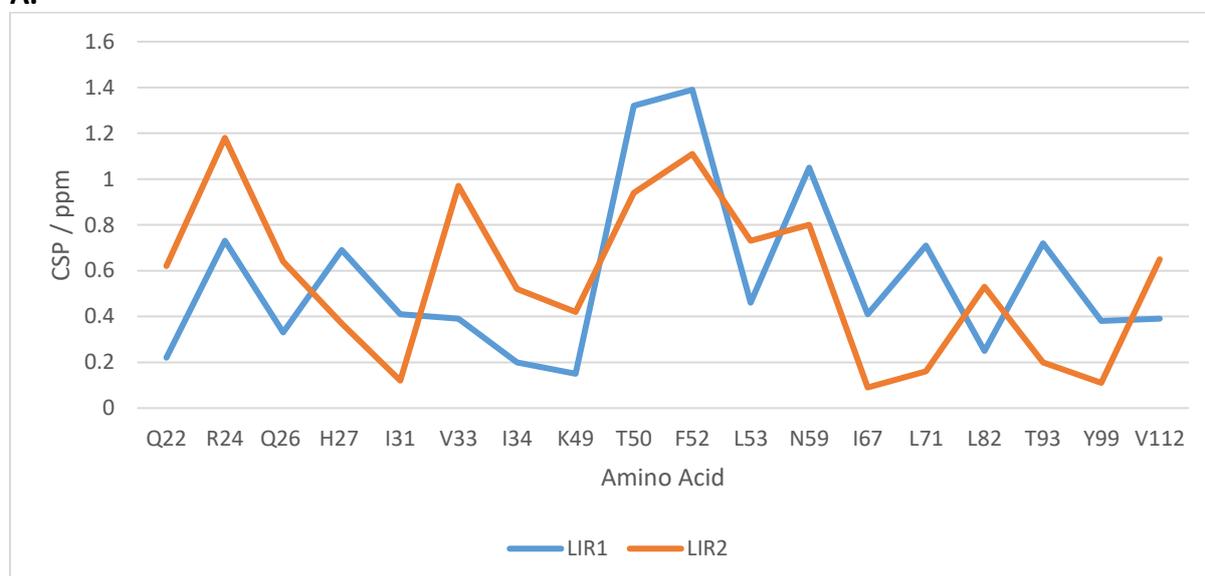
Most Significant

| LC3A WT amino acid | LIR1 CSP | LIR2 CSP |
|--------------------|----------|----------|
| 22 Gln | 0.22 | 0.62 |
| 24 Arg | 0.73 | 1.18 |
| 33 Val | 0.39 | 0.97 |
| 50 Thr | 1.32 | 0.94 |
| 71 Leu | 0.71 | 0.16 |
| 93 Thr | 0.72 | 0.20 |

Table 7.4: Amino acids with significant differences in CSP between the titrations of LC3A WT with NBR1 LIR1 and LIR2.

The amino acids with these significant differences in CSP were then mapped onto the surface of LC3A (Figure 7.25). The HADDOCK structure for the position of LIR2 when interacting with LC3A is also shown for reference.

A.



B.

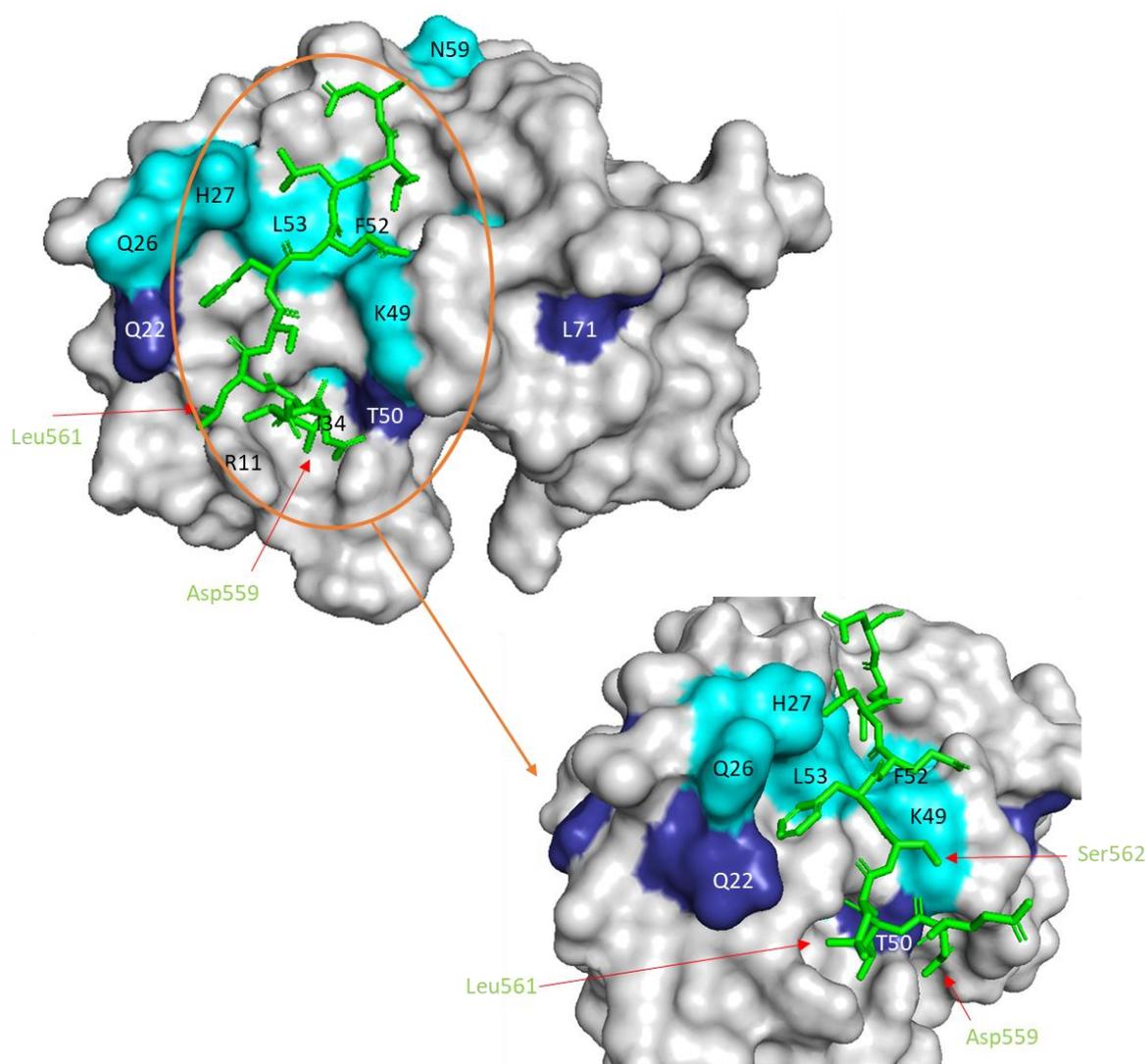


Figure 7.25: A: Graphical representation of CSPs that are significantly different between the titration with LC3A and NBR1 LIR1 (blue) and NBR1 LIR2 (orange). B: Surface of LC3A WT with significant differences in CSP between the NMR titrations of LIR1 and LIR2 with LC3A. Significant shifts are coloured light blue and most significant shifts dark blue. NBR1 LIR2 HADDOCK structure for titration is shown in green.

This analysis reinforces the initial observation that the biggest differences in CSP between the two titration are located around HP1. The CSPs for LIR2 on Q22, R24 and Q26 are all larger than those observed with LIR1, which reinforces the idea that Phe563 sits further to one side of the hydrophobic pocket. LIR1 shows larger CSP for H27, T50 and F52 suggesting the LIR sits close to the centre of LC3A, and that Tyr733 on the LIR sits on the right hand side of the pocket. It is also of interest that Leu561 sits in a gap between Q22 and R11, thus avoiding some of the unfavourable interactions seen with the LC3B interaction with LIR2. There is also an increased CSP for LIR2 on K49, with Glu564 potentially undergoing an electrostatic interaction.

As well as the large number of differences around HP2, I67 and L71 show a significant difference on CSP with larger perturbations for the titration of LIR1. I67 sits inside HP2 and is next to L71 in the 3D structure of LC3A. This implies that Ile736 on LIR1 is fitting deeper into the hydrophobic pocket than Leu566 on LIR2.

All of these observations again help reinforce that the HADDOCK structure for LC3A WT binding with NBR1 LIR1 is not correct and that the LIR is binding at the canonical binding site. It also suggests that LIR1 is sitting closer to the centre of LC3A WT and Ile736 sits relatively far into HP2 when compared with Leu566 on LIR2. It is then possible that the interactions of the Ser728, Glu730 and Asp731 with the face of LC3A (like seen on LC3B), pull the LIR1 closer to the central position, thus making the interaction of Tyr732 in the HP1 less favourable and hence the problems with the HADDOCK structure. However, for LIR2 with fewer interactions with the face of LC3A the LIR is then able to fit better into HP1, which also allows Leu561 to sit in a more favourable position on LC3A than it would on LC3B and for Glu564 to form a favourable interaction with K49. However, to confirm this would require further structural studies.

7.4 Summary

- The ^1H - ^{15}N HSQC of LC3A WT was almost completely assigned with 90% of non-proline residues assigned.
- NMR titrations of LC3A with NBR1 LIR1 and LIR2 show perturbations over the whole of the ^1H - ^{15}N HSQC. However, whereas LIR1 shows a mixture of slow, intermediate and fast exchange the peaks for LIR2 are predominately undergoing fast exchange.
- Analysis of the LC3A WT and NBR1 LIR1 titration data give CSPs that suggest that LIR1 binds to LC3A at the standard LC3 binding site. However, the HADDOCK structure shows the LIR binding with the Tyr733 in hydrophobic pocket.
- Analysis of the LC3A WT and NBR1 LIR2 titration data give CSPs that suggest that LIR2 binds to LC3A at the standard LC3 binding site. The HADDOCK structure then confirms this with Phe563 binding in HP1 and Leu566 in HP2.
- Comparison of the titration of LC3A and LC3B titrations with LIR1 show there are very few significant differences in CSPs for amino acids at the two hydrophobic pockets suggesting that LIR1 does bind in the standard way. However, changes in the sequence of LC3A around the hydrophobic pocket may be influencing how the LIR fits and introducing selectivity over which LIRs bind to LC3A.
- Comparison of the titration of LC3A and LC3B show again that the CSPs are roughly similar but that there are larger perturbations on the far side of HP1 for the titration with LC3A suggesting Phe563 on the LIR sits more over the face of the hydrophobic pocket than when binding to LC3B.
- Comparison of the data for the two LIRs when binding with LC3A show that the majority of differences in CSP are around HP1. LIR2 has larger perturbations on the far side of the pocket whereas LIR1 has larger perturbations on the right hand side of the pocket. This suggest LIR1 sits tighter against the core of LC3A whereas LIR2 sits more to the left. These changes in position of LIR are driven by interactions of other amino acids that precede and come just after the aromatic residue interacting with the face of LC3A.

Chapter 8

8. Conclusion

The work presented within this thesis has furthered the understanding of the interaction the LIRs of NBR1 with the hATG8 family of proteins. Using mass spectrometry and NMR the binding of the LIRs of NBR1 with LC3B have been investigated and how the change in sequence in LIR alters the interaction. Selectivity within the family of hATG8 of proteins was then investigated using MS and the binding of the NBR1 LIRs with LC3A characterised using NMR.

8.1 Investigating the Changes in Binding Interaction between the Two LIRs of NBR1

Two mutants of LC3B WT were produced to study the interaction of the NBR1 LIRs with LC3B. LC3B KL was a binding site mutant altering HP1 and changing the charge on the face of LC3B, and T29D was then a phosphomimetic mutant. MS experiments showed that the LIRs of NBR1 bound to LC3B WT and its mutants. MS competition experiments were used to discover if there was a preferred binding partner for the two LIRs of NBR1 with LC3B WT and its mutants. This data showed that whilst for LIR1 the preferred binding partner was the wild type LC3B, that for LIR2 the preferred binding partner was unexpectedly the K51A/L53A (KL) mutant (Table 8.1). The data for T29D showed a slight preference for WT over T29D for LIR1. However, there was a preference of LC3B WT over T29D for LIR2

| LC3B protein | LIR1 bound to unbound ratio | LIR2 bound to unbound ratio |
|--------------|-----------------------------|-----------------------------|
| WT | 0.99 | 0.84 |
| KL | 0.04 | 1.51 |
| T29D | 0.89 | 0.41 |

Table 8.1: Ratios of bound to unbound protein in MS competition experiments for LC3B WT, KL and T29D at a 1:1:1:1 ratio with the relevant LIR.

The binding preference from the MS experiments warranted further structural studies using NMR. The binding interactions of the LIRs with LC3B WT and its mutants were characterised using chemical shift mapping and the structures for the LIRs binding with LC3B and the mutants were calculated using HADDOCK webserver.

The chemical shift mapping indicated that the LIRs of NBR1 had very similar binding patches to each other and to the LIR of p62. The HADDOCK structures then confirmed that the LIRs bound to LC3B WT, KL and T29D at the canonical binding site, but with variations in position of the LIR (Figure 8.1).

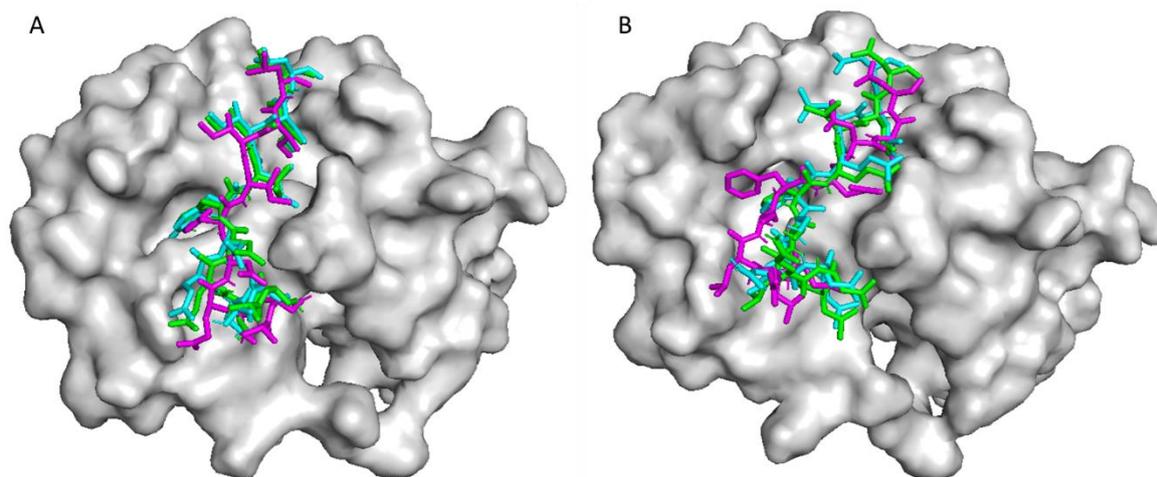


Figure 8.1: HADDOCK structures for NBR1 LIR1 (A) and LIR2 (B) when bound to LC3B WT (green), LC3B KL (blue) and LC3B T29D (pink).

The interactions of the two LIRs of NBR1 with LC3B WT showed that whilst the position of the LIRs were very similar, with Tyr732 on LIR1 and Phe563 on LIR2, sitting in a very similar position in HP1, and Ile735 (LIR1) and Leu566 (LIR2) in HP2. The amino acids preceding the aromatic showed differences in how they interacted with the face of LC3B. NBR1 LIR1 has a number of acidic residues that precede Tyr732 that interact with the face of LC3B WT, with Glu730 and Asp731 interacting with R11 and K59 on LC3B WT respectively. There is also then an interaction between Ser728 on LIR1 and T50 on LC3B WT. However, LIR2 lacking acidic residues directly before Phe563 cannot form these interactions with R11 and K49 on LC3B. However, Asp559 which is 4 positions from the aromatic appears to be forming an interaction with K49. Previous work has noted that acidic residues up to three places away from the aromatic can help stabilize the interaction of the LIR with hATG8s. However, this result suggests that amino acids further away can also help stabilise the binding of LIR to hATG8. This interaction does not help to offset Leu561 on LIR2, a hydrophobic residue, then sitting near R11, and K51 on LC3B causing an unfavourable interaction, and hence why binding affinities calculated previously indicate that LIR1 is a stronger binder than LIR2. It would be interesting to produce crystal structures of the two LIRs binding to LC3B WT to confirm the structures produced by HADDOCK.

The interaction of LIR1 with the KL mutant shows that the differences in the position of the LIR all occur with Tyr732 and the amino acids preceding (Figure 8.2). The mutation of K51 to alanine has the knock on effect of altering the environment of R11. This has the effect of pulling Glu730 further across the face of LC3B KL, and hence moving Asp731 away from K49 lessening the interaction. This then combined with Tyr732 sitting less centrally in HP1 can explain why the KL mutant is a less favourable binding partner for NBR1 LIR1 than LC3B WT. For LIR2 this preference is reversed, with the KL mutant being the more favoured binding partner. With K51 mutated to alanine and the environment of R11 changing this leads to the charge distribution on the face of LC3B changing. This then makes the position of Leu561 less unfavourable than with the wild type LC3B and leads to the KL mutant being a more favoured binding partner. This again points to small changes in sequence leading to changes in binding preference.

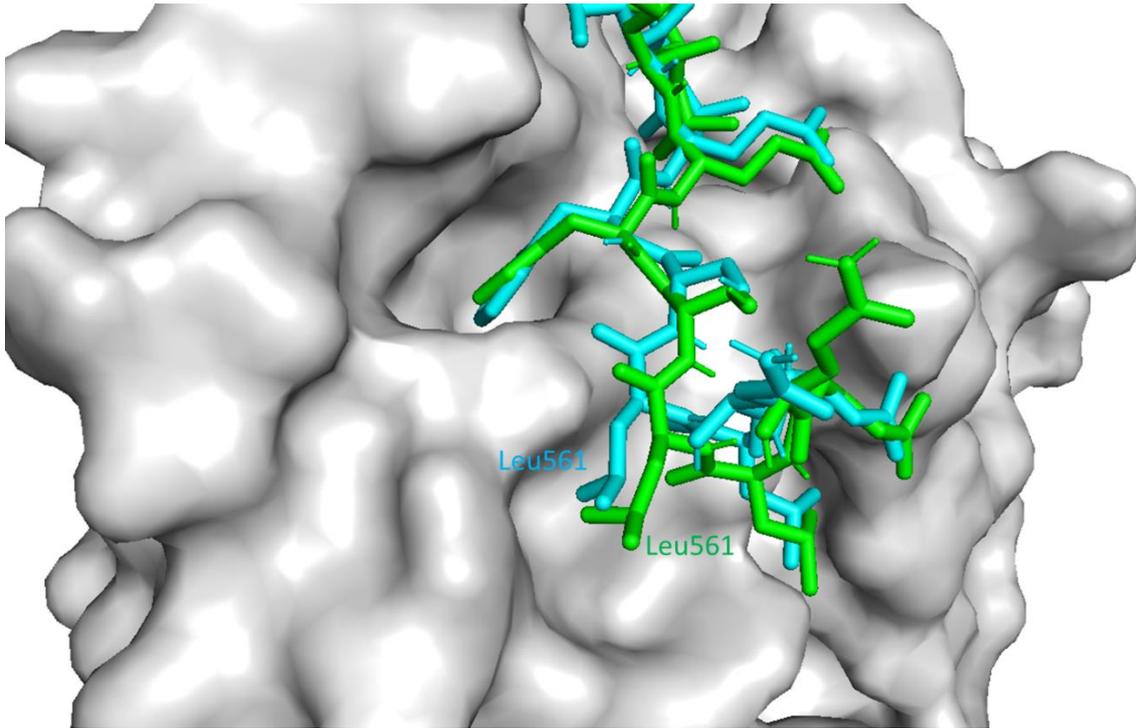


Figure 8.2: Comparison of position of LIR2 when binding to LC3B WT (green) and LC3B KL (blue).

The interactions for the phosphomimetic mutant T29D with LIR1 indicate that the main difference in binding patch when compared to LC3B WT are located around HP1. This leads to Tyr732 sitting closer to K51 on LC3B T29D. This in turn moves the position of the N-terminal amino acids. However, this does not alter the position such that it would eliminate any of the interactions of the N-terminal amino acids with the face of LC3B that are seen with the wild type. With the two structures only varying slightly, and the MS data showing a slight preference for binding with LC3B WT, it is not clear if the phosphorylation at T29 will have any biological significance with the interaction of NBR1 LIR1.

However, the structure for the binding of NBR1 LIR2 binding to T29D shows a very different picture, with the position of LIR2 when interacting with T29D being very different to that of wild type LC3B. Phe563 on LIR2 when binding to LC3B T29D does not sit in the pocket and instead lies across the face (Figure 8.3).

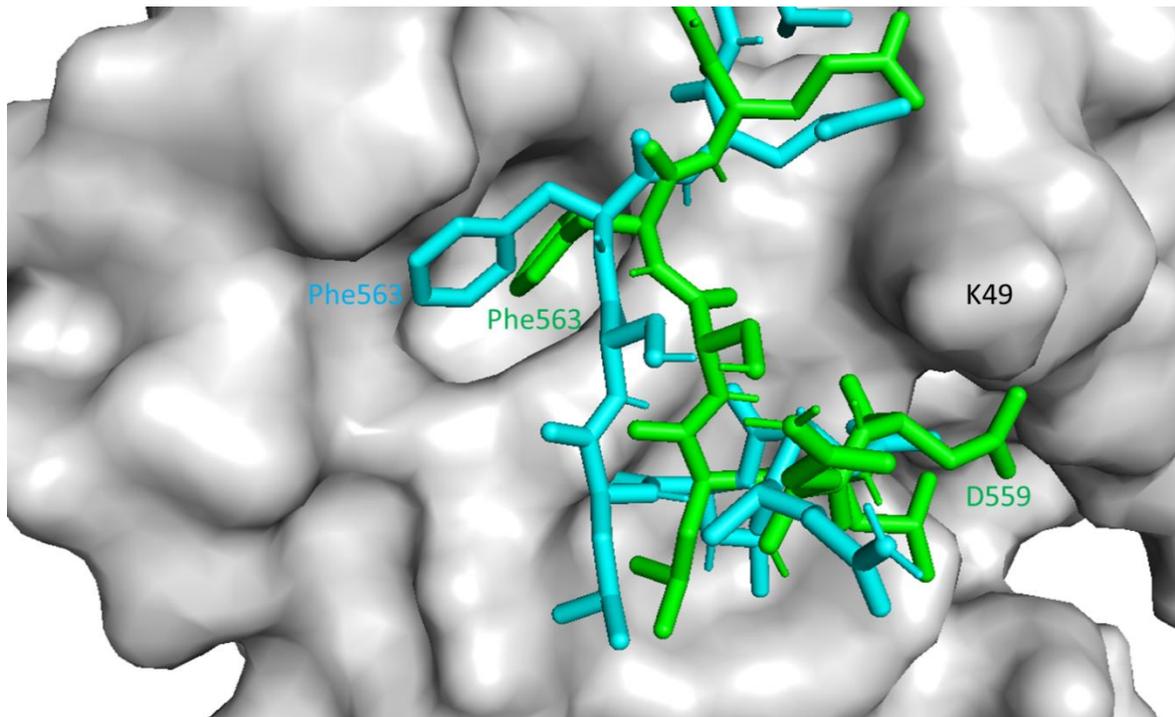


Figure 8.3: Comparison of position of LIR2 when binding to LC3B WT (green) and LC3B T29D (blue).

This in turn moves the whole LIR across the face of LC3B T29D. Combined with the MS competition experiments suggesting that LC3B T29 is a less preferred binding partner than wild type LC3B, suggests that phosphorylation at T29 on LC3B would lead to the inhibition of NBR1 LIR2 with LC3B.

Where LIR1 interacts more strongly with LC3B WT than LIR2 either altering the binding site or phosphorylation of LC3B alters the binding affinity. This points to NBR1 LIR2 interacting with different binding partners than LIR1 and being more susceptible to post translational modifications.

8.2 Binding Preferences of the hATG8 Proteins

MS experiments confirmed that NBR1 LIR1 and LIR2 bound to all six of the hATG8s. MS competition experiments were carried out to investigate if any of the hATG8 family members bound preferentially to the LIRs of NBR1. The results showed that for both NBR1 LIR1 and LIR2 that LC3A was the preferred binding partner of the LC3 subfamily, and for the GABARAPS it was GABARAPL1 (Table 8.2)

| hATG8 protein | LIR1 bound to unbound ratio | LIR2 bound to unbound ratio |
|---------------|-----------------------------|-----------------------------|
| LC3A | 0.28 | 0.31 |
| LC3B | 0.17 | 0.18 |
| LC3C | 0.05 | 0.00 |
| GABARAP | 0.22 | 0.32 |
| GABARAPL1 | 0.46 | 0.53 |
| GABARAPL2 | 0.17 | 0.09 |

Table 8.2: Ratios of bound to unbound protein for MS competition experiments for LC3 and GABARAP subfamilies. Proteins were in a 1:1:1:1 ratio with the relevant LIR.

The ranking of the rest of the LC3 subfamily showed the next preference was for LC3B, with LC3C last. LC3C whilst shown to bind to the two LIRs of NBR1, when in competition with LC3A and LC3B formed very little or undetectable amounts of complex. Whilst LC3A and LC3B have very similar sequences the sequence of LC3C has a number of differences located around the second hydrophobic pocket. This has been shown to lead to LC3C binding to specific types of LIRs [80], and hence when in competition with LC3A and LC3B which bind to the regular LIR motif, that the NBR1 LIRs contain, it is the least preferred binding partner. Due to the similarity in sequences between LC3A and LC3B, and with the clear preference for binding to LC3A, this warranted further structural analysis with NMR.

For the GABARAP subfamily GABARAPL1 was the preferred binder for both LIRs, with GABARAP next and GABARAPL2 the least preferred binding partner. Like in the LC3 subfamily GABARAP and GABARAPL1 have a high sequence alignment with GABARAPL2 less so. As with LC3C, GABARAPL2 has a large number of the changes of sequence around HP2 and whilst no preferred binding partner has been discovered like for LC3C, it is likely there is one, and hence why the standard LIR motif of the NBR1 LIRs binds less preferentially to GABARAPL2.

ITC experiments have been carried out that show LC3B and GABARAPL1 have very similar K_d values [113]. This would then imply that LC3A should be the preferred binding partner of the NBR1 LIRs. However, the work carried out defining the GABARAP interaction motif (GIM) identified NBR1 LIR1 as a GABARAP interacting LIR [107]. These results only used LC3B and GABARAP in their interaction profile, and so it might still be that LC3A has a stronger preference for the NBR1 LIRs. To get a better idea of selectivity in the hATG8 family, competition experiments comparing the LC3s and GABARAPs could be carried out, with the competition experiment containing LC3A and GABARAPL1 being particularly interesting, as these were the preferred binders for both LIRs of NBR1 for the LC3 and GABARAP subfamily's. However, to do this the LIR peptide used would have to be redesigned to take into account that the GABARAPs have hydrophobic interactions with the C-terminal amino acids in the LIR, and the LIR used is too short to include these residues. To get a clearer idea of binding preferences between the NBR1 LIRs and the hATG8 proteins it would be useful to carry out ITC experiments to derive thermodynamic properties and K_d of the interactions.

8.3 Characterization of the Interaction of NBR1 LIRs with LC3A

LC3A has an 82.5% sequence alignment with LC3B with the majority of the differences in sequence occurring at the N-terminal. However, the MS Competition experiments carried out showed that both LIRs of NBR1 bound preferentially to it over the other members of the LC3 subfamily.

NMR titration experiments were used to map the binding site using chemical shift perturbations. These experiments defined a bind patch for both LIRs than indicated a similar binding site to that of LC3B (Figure 8.4).

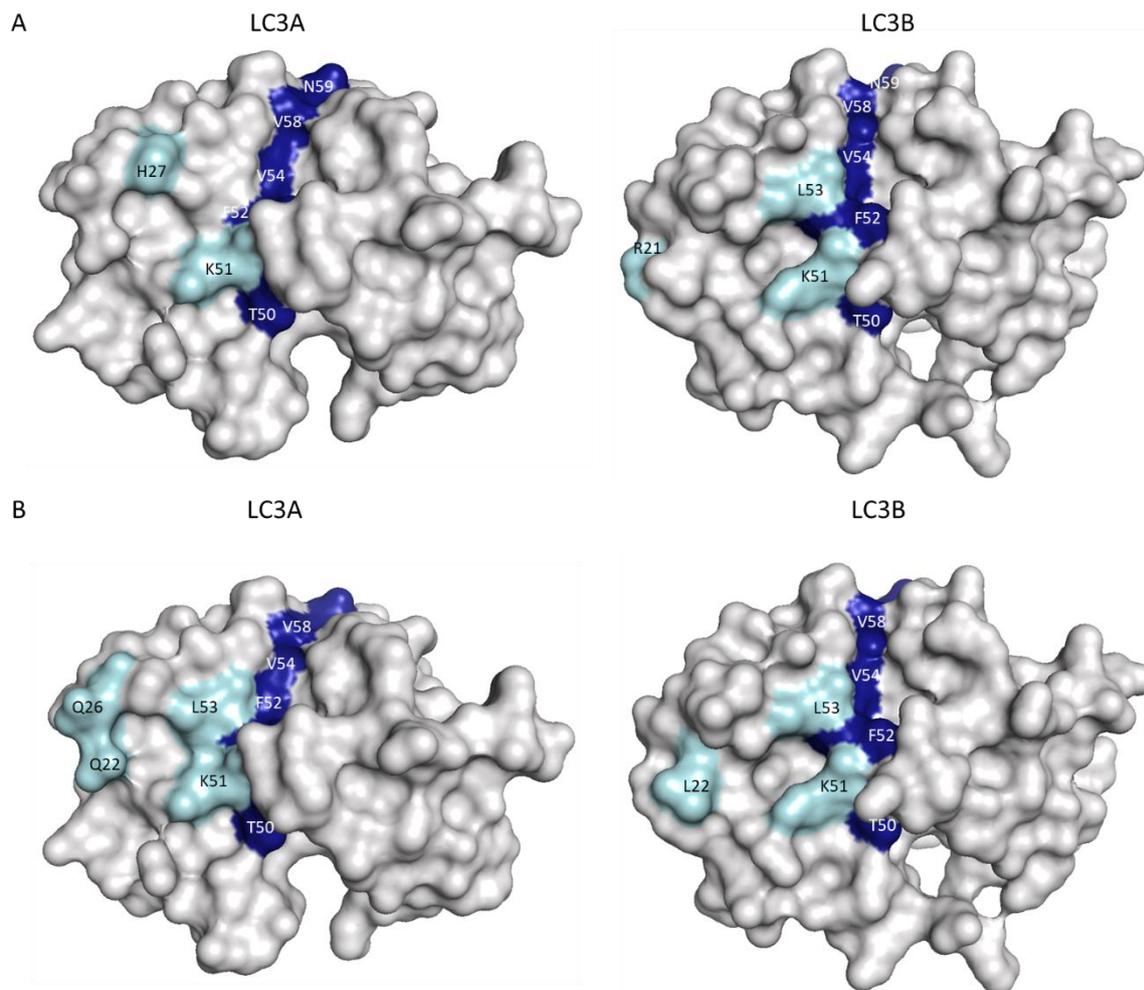


Figure 8.4: Significant CSPs mapped onto the protein surface for the titration of LC3A (left) and LC3B (right) with NBR1 LIR1 (A) and LIR2 (B). Significant shifts shown in light blue and most significant shifts shown in dark blue.

However, when the HADDOCK structure from the CSPs were generated for NBR1 LIR1 the position of the LIR was not as expected. The structure generated showed Tyr732 of LIR1 binding into HP2 with Ile735 sitting across the face of LC3A (Figure 8.5).

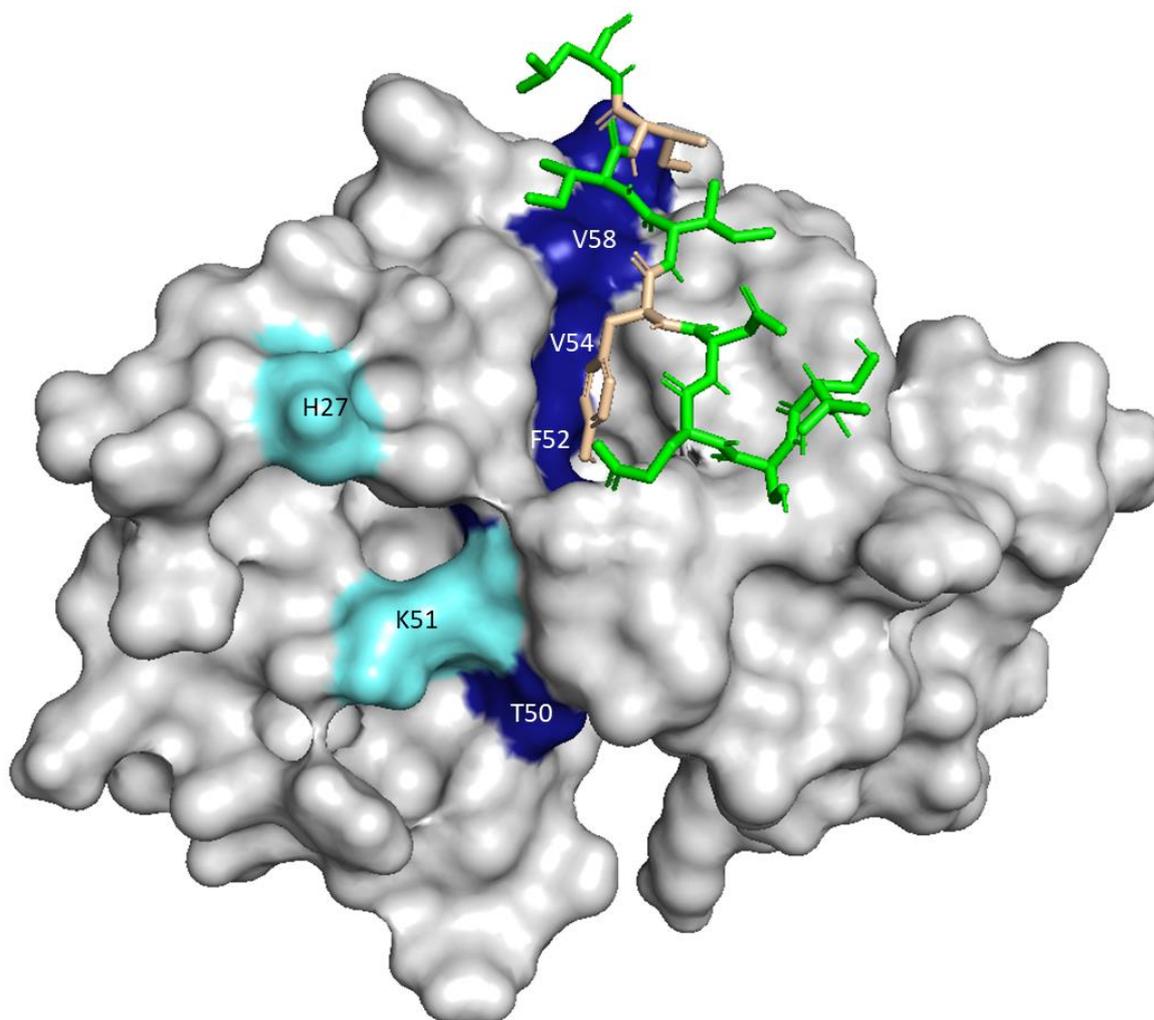


Figure 8.5: HADDOCK structure of NBR1 LIR1 peptide bound to LC3A. The peptide is coloured green with the active residues Tyr733 and Ile736 coloured yellow. The surface of LC3A is coloured to indicate the significant shifts seen in the bound form. Residues light blue had CSPs greater than the mean and one standard deviation and those coloured blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.

When analysed, the differences in CSPs between the interaction of NBR1 LIR1 with LC3A and LC3B showed the significant changes to be around HP2. This suggests that the aromatic on LIR1 is interacting with HP1 on LC3A, despite the HADDOCK structure indicating otherwise. This was confirmed further when HADDOCK generated the structure of LIR1 binding to the canonical binding site on LC3B when using the LC3A CSPs.

For NBR1 LIR2 the HADDOCK structure generated showed LIR binding in the standard orientation with Phe563 in HP1 and Leu566 in HP2 (Figure 8.6).

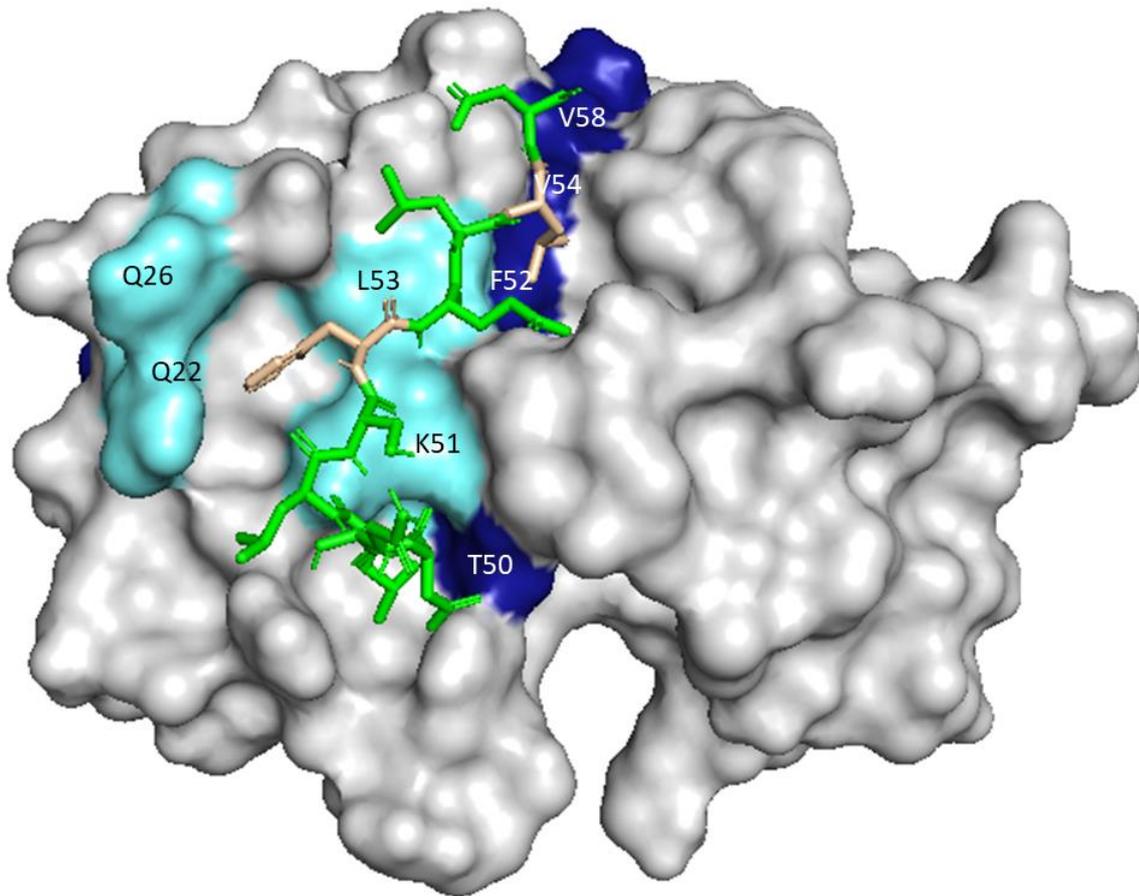


Figure 8.6: HADDOCK structure of NBR1 LIR2 peptide bound to LC3A. The peptide is coloured green with the active residues Phe563 and Leu566 coloured yellow. The surface of LC3A is coloured to indicate the significant shifts seen in the bound form. Residues light blue had CSPs greater than the mean and one standard deviation and those coloured blue had CSPs greater than the mean and two standard deviations. Residues coloured grey had no significant perturbation.

Due to the changes in HP1 on LC3A when compared to LC3B, this allows Glu564 to sit close to K49 where an electrostatic interaction could help stabilise the binding. However, the amino acids that precede Phe563 sit away from the face of LC3A, and do not appear to form any further interactions with LC3A. This then enables Leu561 to point away from the face of LC3A, preventing the unfavourable interaction of a hydrophobic residue with the charged residues R11 and K51 on the face of LC3A.

The differences between LC3A and LC3B are mainly located on the N-terminal alpha helices that pack against the central beta sheet to form HP1. These changes are enough that the HADDOCK software now no longer generates a structure for the binding of LIR1 with LC3A. However, this is not the same for LIR2 which binds to the canonical binding site. It seems unlikely that the small change in aromatic from tyrosine on LIR1 to phenylalanine would produce such a change. Therefore it points to the acidic amino acids preceding the aromatic to be playing a part in the interaction. For the interaction of LIR1 with LC3B Ser728, Glu730 and Asp731 appear to interact with residues on the face of LC3B, and help stabilise the binding of the LIR. However, for LC3A these may have a negative effect with the LIR sitting

centrally and not allowing Tyr732 to bind effectively in HP1. For LIR2 on the other hand, which has fewer interactions with the face of LC3A/B, can sit with Phe563 orientated more favourably in HP1. Unfortunately without a structure to work with it is difficult to postulate why the interaction of LIR1 with LC3A has produced the HADDOCK structure it has, as the area of LC3A where those amino acids sit is the same as LC3B. To better understand what is happening future work would look at producing a structure of NBR1 LIR1 binding to LC3A, either through NMR or a crystal structure.

The data presented shows that despite the binding patches being similar, that the LIRs interact in different ways with LC3B and LC3A. With the growing evidence that the selective autophagy receptors are linked to specific hATG8 family members [52, 56, 142], and hATG8 members having specific functions [56], then the specificity of an LIR is key to a particular pathway. The data collected shows that small changes in sequence in the LIR, whilst maintaining the canonical motif are enough to make the LIRs more specific to certain hATG8s. Whilst the difference in binding affinities is relatively small between a particular LIRs [113], the avidity effect of having large complexes containing multiple SARs binding to a number of hATG8s on the phagophore can amplify this small difference.

It would be interesting to investigate the structure of the LIRs binding to LC3C and GABARAPS so as to get the complete picture of the interactions of an LIR with hATG8 family. However, this is a considerable amount of work and was not possible within the time frame of this work.

8.4 Future Work

Whilst the work carried out in this project has been able to study in-depth the binding of the NBR1 LIRs to the hATG8 proteins, there are some extra experiments that could be carried out to further understand the interaction. The work so far has used predominantly NMR and MS, and so it would be useful to carry out some extra biophysical and biochemical techniques to further characterize the interaction.

To start with isothermal calorimetry experiments (ITC) could be employed to calculate the K_d value for the interaction of the two LIRs of NBR1 with the hATG8s. This would then determine the order of binding preference for the hATG8 proteins with the NBR1 LIRs which would then compliment the MS data acquired in this project. To then probe the observed binding preferences for the NBR1 LIRs for LC3A (LC3 subfamily) and GABARAPL1 (GABARAP subfamily) *in vitro*, cell work could be used. Using flag-tagged hATG8s, and V5-tagged NBR1, and mutants with the LIRs knocked out using CRISPR gene editing (allowing NBR1 to be at physiological concentrations), it would be possible to determine which hATG8s the LIRs preferentially bound to. To complement the Cell work a series of pulldowns could then be employed to further probe binding preference.

NMR could then be further employed to determine the 3D structures of LC3B and LC3A with LIR1 and LIR2 of NBR1. With these structures it would then be possible to firstly confirm the orientation of the LIRs on LC3A/B. They would then be able to shed further light on the fine details of the binding interaction which would complement the NMR titrations carried out and HADDOCK structures generated in this work.

To further study how changes in the LIR sequence effect the binding a series of different peptides could be used with individual amino acids changed to see how this effects the interaction with the hATG8 proteins. Then using the NMR titration experiments and HADDOCK structures it would be possible to visualise the details of these changes. Potential changes would be alter the aromatic and preceding amino acids to mimic the other LIR to see if the interaction could become more like that of the LIR mimicked. This could then be complimented with peptide screens against all six hATG8s using biotinylated peptides then incubated with his-tagged hATG8 proteins.

The effect phosphorylation of T29 on LC3B has on the binding with the LIRS could be further explored using both the cell and pulldown experiments. The T29D mutant and T29A could be used as a constitutive and null respectively. By then using WT and LIR knock out versions of NBR1 the difference seen in the NMR experiments could be investigated.

8.5 Conclusion

In conclusion this thesis has shown that whilst the binding of NBR1 LIR1 to LC3B is stabilised by the amino acids preceding Tyr732 it shows very little variation in interaction when LC3B undergoes modification like phosphorylation. However, for NBR1 LIR2 whilst not interacting as strongly as LIR1 with LC3B, modifications around the binding site and phosphorylation lead to changes in binding affinity. This leads to the idea that within selective autophagy LIR1 is used primarily and LIR2 is then regulated by post translational modifications of the hATG8 family of proteins.

For the hATG8 family of proteins LC3A was shown to be the preferred binding partner. Structural studies of the interaction however, were less clear cut and whilst the CSP data suggests LIRs binds to the canonical binding site on LC3A. Modelling the structure LIR binding to LC3A revealed the LIR1 binding to HP2 with Tyr732 Whereas LIR2 bound to the canonical binding site. As the changes in sequence between LC3A and LC3B are around HP1 this points to this structure being generated by the interaction of the aromatic acid and preceding amino acids with LC3A. This is yet again another sign that the amino acids at the N-terminal of LIRs play a significant part in the binding of LIRs to the hATG8 family of proteins.

Bibliography

1. De Duve, C., B.C. Pressman, R. Gianetto, R. Wattiaux, and F. Appelmans. (1955) Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue. *Biochem. J.* **60**, 604-617
2. Klionsky, D.J. (2008) Autophagy revisited: A conversation with Christian de Duve. *Autophagy.* **4**, 740-743
3. Tsukada, M. and Y. Ohsumi. (1993) Isolation and characterization of autophagy-defective mutants of *Saccharomyces cerevisiae*. *FEBS Letters.* **333**, 169-174
4. Mizushima, N. and M. Komatsu. (2011) Autophagy: Renovation of Cells and Tissues. *Cell.* **147**, 728-741
5. Ravikumar, B., S. Sarkar, J.E. Davies, M. Futter, M. Garcia-Arencibia, Z.W. Green-Thompson, M. Jimenez-Sanchez, V.I. Korolchuk, M. Lichtenberg, S. Luo, D.C.O. Massey, F.M. Menzies, K. Moreau, U. Narayanan, M. Renna, F.H. Siddiqi, B.R. Underwood, A.R. Winslow, and D.C. Rubinsztein. (2010) Regulation of Mammalian Autophagy in Physiology and Pathophysiology. *Physiol. Rev.* **90**, 1383-1435
6. Kroemer, G., G. Mariño, and B. Levine. (2010) Autophagy and the integrated stress response. *Mol Cell.* **40**, 280-293
7. Massey, A., R. Kiffin, and A.M. Cuervo. (2004) Pathophysiology of chaperone-mediated autophagy. *Int. J. Biochem. Cell Biol.* **36**, 2420-2434
8. Cuervo, A.M. and E. Wong. (2014) Chaperone-mediated autophagy: roles in disease and aging. *Cell Research.* **24**, 92-104
9. Galluzzi, L., E.H. Baehrecke, A. Ballabio, P. Boya, J.M. Bravo-San Pedro, F. Cecconi, A.M. Choi, C.T. Chu, P. Codogno, M.I. Colombo, A.M. Cuervo, J. Debnath, V. Deretic, I. Dikic, E.-L. Eskelinen, G.M. Fimia, S. Fulda, D.A. Gewirtz, D.R. Green, M. Hansen, J.W. Harper, M. Jäättelä, T. Johansen, G. Juhasz, A.C. Kimmelman, C. Kraft, N.T. Ktistakis, S. Kumar, B. Levine, C. Lopez-Otin, F. Madeo, S. Martens, J. Martinez, A. Melendez, N. Mizushima, C. Münz, L.O. Murphy, J.M. Penninger, M. Piacentini, F. Reggiori, D.C. Rubinsztein, K.M. Ryan, L. Santambrogio, L. Scorrano, A.K. Simon, H.-U. Simon, A. Simonsen, N. Tavernarakis, S.A. Tooze, T. Yoshimori, J. Yuan, Z. Yue, Q. Zhong, and G. Kroemer. (2017) Molecular definitions of autophagy and related processes. *The EMBO Journal.* **36**, 1811-1836
10. Oku, M. and Y. Sakai. (2018) Three Distinct Types of Microautophagy Based on Membrane Dynamics and Molecular Machineries. *BioEssays.* **40**, 1800008-1800014
11. Klionsky, D.J. and S.D. Emr. (2000) Autophagy as a Regulated Pathway of Cellular Degradation. *Science.* **290**, 1717-1721
12. Behrends, C., M.E. Sowa, S.P. Gygi, and J.W. Harper. (2010) Network organization of the human autophagy system. *Nature.* **466**, 68-76
13. Kim, J., M. Kundu, B. Viollet, and K.L. Guan. (2011) AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1. *Nature cell biology.* **13**, 132-141
14. Hamasaki, M. and T. Yoshimori. (2010) Where do they come from? Insights into autophagosome formation. *FEBS Letters.* **584**, 1296-1301
15. Melia, T.J., A.H. Lystad, and A. Simonsen. (2020) Autophagosome biogenesis: From membrane growth to closure. *J. Cell Biol.* **219**,
16. Fujita, N., T. Itoh, H. Omori, M. Fukuda, T. Noda, and T. Yoshimori. (2008) The Atg16L Complex Specifies the Site of LC3 Lipidation for Membrane Biogenesis in Autophagy. *Mol. Biol. Cell.* **19**, 2092-2100

17. Lamb, C.A., T. Yoshimori, and S.A. Tooze. (2013) The autophagosome: origins unknown, biogenesis complex. *Nat. Rev. Mol. Cell Biol.* **14**, 759-774
18. Knorr, R.L., R. Lipowsky, and R. Dimova. (2015) Autophagosome closure requires membrane scission. *Autophagy*. **11**, 2134-2137
19. Kimura, S., T. Noda, and T. Yoshimori. (2008) Dynein-dependent Movement of Autophagosomes Mediates Efficient Encounters with Lysosomes. *Cell Struct. Funct.* **33**, 109-122
20. Pankiv, S., E.A. Alemu, A. Brech, J.-A. Bruun, T. Lamark, A. Øvervatn, G. Bjørkøy, and T. Johansen. (2010) FYCO1 is a Rab7 effector that binds to LC3 and PI3P to mediate microtubule plus end-directed vesicle transport. *J. Cell Biol.* **188**, 253-269
21. Bento, C.F., C. Puri, K. Moreau, and D.C. Rubinsztein. (2013) The role of membrane-trafficking small GTPases in the regulation of autophagy. *J. Cell Sci.* **126**, 1059-1069
22. Moreau, K., M. Renna, and D.C. Rubinsztein. (2013) Connections between SNAREs and autophagy. *Trends Biochem. Sci.* **38**, 57-63
23. Jiang, P., T. Nishimura, Y. Sakamaki, E. Itakura, T. Hatta, T. Natsume, and N. Mizushima. (2014) The HOPS complex mediates autophagosome-lysosome fusion through interaction with syntaxin 17. *Mol. Biol. Cell.* **25**, 1327-1337
24. Johansen, T. and T. Lamark. (2011) Selective autophagy mediated by autophagic adapter proteins. *Autophagy*. **7**, 279-296
25. Kirkin, V., D.G. McEwan, I. Novak, and I. Dikic. (2009) A Role for Ubiquitin in Selective Autophagy. *Mol Cell.* **34**, 259-269
26. Rogov, V., V. Dötsch, T. Johansen, and V. Kirkin. (2014) Interactions between Autophagy Receptors and Ubiquitin-like Proteins Form the Molecular Basis for Selective Autophagy. *Mol Cell.* **53**, 167-178
27. Hyttinen, J.M.T., M. Amadio, J. Viiri, A. Pascale, A. Salminen, and K. Kaarniranta. (2014) Clearance of misfolded and aggregated proteins by aggrephagy and implications for aggregation diseases. *Ageing Res. Rev.* **18**, 16-28
28. Montava-Garriga, L. and I.G. Ganley. (2020) Outstanding Questions in Mitophagy: What We Do and Do Not Know. *J. Mol. Biol.* **432**, 206-230
29. Li, J. and W. Wang. (2021) Mechanisms and Functions of Pexophagy in Mammalian Cells. *Cells*. **10**, 1094
30. Zheng, Y.T., S. Shahnazari, A. Brech, T. Lamark, T. Johansen, and J.H. Brumell. (2009) The Adaptor Protein p62/SQSTM1 Targets Invading Bacteria to the Autophagy Pathway. *J. Immunol.* **183**, 5909-5916
31. Komatsu, M., S. Waguri, T. Chiba, S. Murata, J.-i. Iwata, I. Tanida, T. Ueno, M. Koike, Y. Uchiyama, E. Kominami, and K. Tanaka. (2006) Loss of autophagy in the central nervous system causes neurodegeneration in mice. *Nature*. **441**, 880-884
32. Hara, T., K. Nakamura, M. Matsui, A. Yamamoto, Y. Nakahara, R. Suzuki-Migishima, M. Yokoyama, K. Mishima, I. Saito, H. Okano, and N. Mizushima. (2006) Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. *Nature*. **441**, 885-889
33. Du, Y., M.C. Wooten, and M.W. Wooten. (2009) Oxidative damage to the promoter region of SQSTM1/p62 is common to neurodegenerative disease. *Neurobiol. Dis.* **35**, 302-310
34. Babu, J.R., T. Geetha, and M.W. Wooten. (2005) Sequestosome1/p62 shuttles polyubiquitinated tau for proteasomal degradation. *J. Neurochem.* **94**, 192-203

35. Sasaki, S. (2011) Autophagy in Spinal Cord Motor Neurons in Sporadic Amyotrophic Lateral Sclerosis. *Journal of Neuropathology & Experimental Neurology*. **70**, 349-359
36. Chua, J.P., S.J. Barmada, C.H. De, and E. Kabashi. (2021) Autophagy and ALS: mechanistic insights and therapeutic implications. *Autophagy*. 1-29
37. Nakatogawa, H., K. Suzuki, Y. Kamada, and Y. Ohsumi. (2009) Dynamics and diversity in autophagy mechanisms: lessons from yeast. *Nat. Rev. Mol. Cell Biol.* **10**, 458-467
38. Shpilka, T., H. Weidberg, S. Pietrokovski, and Z. Elazar. (2011) Atg8: an autophagy-related ubiquitin-like protein family. *Genome Biol.* **12**, 226-237
39. Jatana, N., D.B. Ascher, D.E.V. Pires, R.S. Gokhale, and L. Thukral. (2020) Human LC3 and GABARAP subfamily members achieve functional specificity via specific structural modulations. *Autophagy*. **16**, 239-255
40. Mizushima, N., T. Noda, T. Yoshimori, Y. Tanaka, T. Ishii, M.D. George, D.J. Klionsky, M. Ohsumi, and Y. Ohsumi. (1998) A protein conjugation system essential for autophagy. *Nature*. **395**, 395-398
41. Ichimura, Y., T. Kirisako, T. Takao, Y. Satomi, Y. Shimonishi, N. Ishihara, N. Mizushima, I. Tanida, E. Kominami, M. Ohsumi, T. Noda, and Y. Ohsumi. (2000) A ubiquitin-like system mediates protein lipidation. *Nature*. **408**, 488-492
42. Mizushima, N., T. Yoshimori, and Y. Ohsumi. (2011) The Role of Atg Proteins in Autophagosome Formation. *Annu. Rev. Cell Dev. Biol.* **27**, 107-132
43. Hanada, T., N.N. Noda, Y. Satomi, Y. Ichimura, Y. Fujioka, T. Takao, F. Inagaki, and Y. Ohsumi. (2007) The Atg12-Atg5 Conjugate Has a Novel E3-like Activity for Protein Lipidation in Autophagy. *J. Biol. Chem.* **282**, 37298-37302
44. Kirisako, T., Y. Ichimura, H. Okada, Y. Kabeya, N. Mizushima, T. Yoshimori, M. Ohsumi, T. Takao, T. Noda, and Y. Ohsumi. (2000) The reversible modification regulates the membrane-binding state of Apg8/Aut7 essential for autophagy and the cytoplasm to vacuole targeting pathway. *J Cell Biol.* **151**, 263-276
45. Birgisdottir, Å.B., T. Lamark, and T. Johansen. (2013) The LIR motif – crucial for selective autophagy. *J. Cell Sci.* **126**, 3237-3247
46. Skytte Rasmussen, M., S. Mouilleron, B. Kumar Shrestha, M. Wirth, R. Lee, K. Bowitz Larsen, Y. Abudu Princely, N. O'Reilly, E. Sjøttem, S.A. Tooze, T. Lamark, and T. Johansen. (2017) ATG4B contains a C-terminal LIR motif important for binding and efficient cleavage of mammalian orthologs of yeast Atg8. *Autophagy*. **13**, 834-853
47. Alemu, E.A., T. Lamark, K.M. Torgersen, A.B. Birgisdottir, K.B. Larsen, A. Jain, H. Olsvik, A. Øvervatn, V. Kirkin, and T. Johansen. (2012) ATG8 Family Proteins Act as Scaffolds for Assembly of the ULK Complex: Sequence Requirements for LC3-Interacting Region (LIR) Motifs. *J. Biol. Chem.* **287**, 39275-39290
48. Kaufmann, A., V. Beier, Henri G. Franquelim, and T. Wollert. (2014) Molecular Mechanism of Autophagic Membrane-Scaffold Assembly and Disassembly. *Cell*. **156**, 469-481
49. Bjørkøy, G., T. Lamark, A. Brech, H. Outzen, M. Perander, A. Overvatn, H. Stenmark, and T. Johansen. (2005) p62/SQSTM1 forms protein aggregates degraded by autophagy and has a protective effect on huntingtin-induced cell death. *J Cell Biol.* **171**, 603-614
50. Kirkin, V., T. Lamark, Y.-S. Sou, G. Bjørkøy, J.L. Nunn, J.-A. Bruun, E. Shvets, D.G. McEwan, T.H. Clausen, P. Wild, I. Bilusic, J.-P. Theurillat, A. Øvervatn, T. Ishii, Z. Elazar, M. Komatsu, I. Dikic, and T. Johansen. (2009) A Role for NBR1 in Autophagosomal Degradation of Ubiquitinated Substrates. *Mol Cell*. **33**, 505-516

51. Olsvik, H.L., T. Lamark, K. Takagi, K.B. Larsen, G. Evjen, A. Øvervatn, T. Mizushima, and T. Johansen. (2015) 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**, 29361-29374
52. McEwan, David G., D. Popovic, A. Gubas, S. Terawaki, H. Suzuki, D. Stadel, Fraser P. Coxon, D. Miranda de Stegmann, S. Bhogaraju, K. Maddi, A. Kirchof, E. Gatti, Miep H. Helfrich, S. Wakatsuki, C. Behrends, P. Pierre, and I. Dikic. (2015) PLEKHM1 Regulates Autophagosome-Lysosome Fusion through HOPS Complex and LC3/GABARAP Proteins. *Mol Cell.* **57**, 39-54
53. Shpilka, T., N. Mizushima, and Z. Elazar. (2012) Ubiquitin-like proteins and autophagy at a glance. *J. Cell Sci.* **125**, 2343-2348
54. Sugawara, K., N.N. Suzuki, Y. Fujioka, N. Mizushima, Y. Ohsumi, and F. Inagaki. (2004) The crystal structure of microtubule-associated protein light chain 3, a mammalian homologue of *Saccharomyces cerevisiae* Atg8. *Genes to Cells.* **9**, 611-618
55. Noda, N.N., Y. Ohsumi, and F. Inagaki. (2009) ATG Systems from the Protein Structural Point of View. *Chem. Rev.* **109**, 1587-1598
56. Weidberg, H., E. Shvets, T. Shpilka, F. Shimron, V. Shinder, and Z. Elazar. (2010) LC3 and GATE-16/GABARAP subfamilies are both essential yet act differently in autophagosome biogenesis. *The EMBO Journal.* **29**, 1792-1802
57. Nguyen, T.N., B.S. Padman, J. Usher, V. Oorschot, G. Ramm, and M. Lazarou. (2016) Atg8 family LC3/GABARAP proteins are crucial for autophagosome-lysosome fusion but not autophagosome formation during PINK1/Parkin mitophagy and starvation. *J. Cell Biol.* **215**, 857-874
58. Vaites, L.P., J.A. Paulo, E.L. Huttlin, and J.W. Harper. (2018) Systematic Analysis of Human Cells Lacking ATG8 Proteins Uncovers Roles for GABARAPs and the CCZ1/MON1 Regulator C18orf8/RMC1 in Macroautophagic and Selective Autophagic Flux. *Mol. Cell. Biol.* **38**, e00392-00317
59. Szalai, P., L.K. Hagen, F. Sætre, M. Luhr, M. Sponheim, A. Øverbye, I.G. Mills, P.O. Seglen, and N. Engedal. (2015) Autophagic bulk sequestration of cytosolic cargo is independent of LC3, but requires GABARAPs. *Experimental Cell Research.* **333**, 21-38
60. Pankiv, S., T.H. Clausen, T. Lamark, A. Brech, J.-A. Bruun, H. Outzen, A. Øvervatn, G. Bjørkøy, and T. Johansen. (2007) p62/SQSTM1 Binds Directly to Atg8/LC3 to Facilitate Degradation of Ubiquitinated Protein Aggregates by Autophagy. *J. Biol. Chem.* **282**, 24131-24145
61. Ichimura, Y., T. Kumanomidou, Y.-s. Sou, T. Mizushima, J. Ezaki, T. Ueno, E. Kominami, T. Yamane, K. Tanaka, and M. Komatsu. (2008) Structural Basis for Sorting Mechanism of p62 in Selective Autophagy. *J. Biol. Chem.* **283**, 22847-22857
62. Noda, N.N., H. Kumeta, H. Nakatogawa, K. Satoo, W. Adachi, J. Ishii, Y. Fujioka, Y. Ohsumi, and F. Inagaki. (2008) Structural basis of target recognition by Atg8/LC3 during selective autophagy. *Genes to Cells.* **13**, 1211-1218
63. Gatica, D., V. Lahiri, and D.J. Klionsky. (2018) Cargo recognition and degradation by selective autophagy. *Nat. Cell Biol.* **20**, 233-242
64. Johansen, T. and T. Lamark. (2019) Selective Autophagy: ATG8 Family Proteins, LIR Motifs and Cargo Receptors. *J. Mol. Biol.* **432**,
65. Wirth, M., W. Zhang, M. Razi, L. Nyoni, D. Joshi, N. O'Reilly, T. Johansen, S.A. Tooze, and S. Mouilleron. (2019) Molecular determinants regulating selective binding of autophagy adapters and receptors to ATG8 proteins. *Nat. Commun.* **10**, 2055

66. Satoo, K., N.N. Noda, H. Kumeta, Y. Fujioka, N. Mizushima, Y. Ohsumi, and F. Inagaki. (2009) The structure of Atg4B-LC3 complex reveals the mechanism of LC3 processing and delipidation during autophagy. *The EMBO journal*. **28**, 1341-1350
67. Heo, J.-M., A. Ordureau, Joao A. Paulo, J. Rinehart, and J.W. Harper. (2015) The PINK1-PARKIN Mitochondrial Ubiquitylation Pathway Drives a Program of OPTN/NDP52 Recruitment and TBK1 Activation to Promote Mitophagy. *Mol Cell*. **60**, 7-20
68. Lazarou, M., D.A. Sliter, L.A. Kane, S.A. Sarraf, C. Wang, J.L. Burman, D.P. Sideris, A.I. Fogel, and R.J. Youle. (2015) The ubiquitin kinase PINK1 recruits autophagy receptors to induce mitophagy. *Nature*. **524**, 309-314
69. Kim, P.K., D.W. Hailey, R.T. Mullen, and J. Lippincott-Schwartz. (2008) Ubiquitin signals autophagic degradation of cytosolic proteins and peroxisomes. *Proc. Natl. Acad. Sci*. **105**, 20567-20574
70. Grasso, D., A. Ropolo, A. Lo Ré, V. Boggio, M.I. Molejón, J.L. Iovanna, C.D. Gonzalez, R. Urrutia, and M.I. Vaccaro. (2011) Zymophagy, a novel selective autophagy pathway mediated by VMP1-USP9x-p62, prevents pancreatic cell death. *J Biol Chem*. **286**, 8308-8324
71. Orvedahl, A., S. MacPherson, R. Sumpter, Jr., Z. Tallóczy, Z. Zou, and B. Levine. (2010) Autophagy protects against Sindbis virus infection of the central nervous system. *Cell Host Microbe*. **7**, 115-127
72. Isakson, P., A.H. Lystad, K. Breen, G. Koster, H. Stenmark, and A. Simonsen. (2013) TRAF6 mediates ubiquitination of KIF23/MKLP1 and is required for midbody ring degradation by selective autophagy. *Autophagy*. **9**, 1955-1964
73. Pohl, C. and S. Jentsch. (2009) Midbody ring disposal by autophagy is a post-abscission event of cytokinesis. *Nat. Cell Biol*. **11**, 65-70
74. Deosaran, E., K.B. Larsen, R. Hua, G. Sargent, Y. Wang, S. Kim, T. Lamark, M. Jauregui, K. Law, J. Lippincott-Schwartz, A. Brech, T. Johansen, and P.K. Kim. (2013) NBR1 acts as an autophagy receptor for peroxisomes. *J. Cell Sci*. **126**, 939-952
75. Kuo, T.-C., C.-T. Chen, D. Baron, T.T. Onder, S. Loewer, S. Almeida, C.M. Weismann, P. Xu, J.-M. Houghton, F.-B. Gao, G.Q. Daley, and S. Doxsey. (2011) Midbody accumulation through evasion of autophagy contributes to cellular reprogramming and tumorigenicity. *Nat. Cell Biol*. **13**, 1214-1223
76. Korac, J., V. Schaeffer, I. Kovacevic, A.M. Clement, B. Jungblut, C. Behl, J. Terzic, and I. Dikic. (2013) Ubiquitin-independent function of optineurin in autophagic clearance of protein aggregates. *J. Cell Sci*. **126**, 580-592
77. Wild, P., H. Farhan, D.G. McEwan, S. Wagner, V.V. Rogov, N.R. Brady, B. Richter, J. Korac, O. Waidmann, C. Choudhary, V. Dötsch, D. Bumann, and I. Dikic. (2011) Phosphorylation of the Autophagy Receptor Optineurin Restricts Salmonella Growth. *Science*. **333**, 228-233
78. Wong, Y.C. and E.L.F. Holzbaur. (2014) Optineurin is an autophagy receptor for damaged mitochondria in parkin-mediated mitophagy that is disrupted by an ALS-linked mutation. *Proc. Natl. Acad. Sci*. **111**, E4439-E4448
79. Thurston, T.L.M., G. Ryzhakov, S. Bloor, N. von Muhlinen, and F. Randow. (2009) The TBK1 adaptor and autophagy receptor NDP52 restricts the proliferation of ubiquitin-coated bacteria. *Nat. Immunol*. **10**, 1215-1221
80. von Muhlinen, N., M. Akutsu, B.J. Ravenhill, Á. Foeglein, S. Bloor, T.J. Rutherford, S.M.V. Freund, D. Komander, and F. Randow. (2012) LC3C, bound selectively by a

- noncanonical LIR motif in NDP52, is required for antibacterial autophagy. *Mol Cell.* **48**, 329-342
81. Padman, B.S., T.N. Nguyen, L. Uoselis, M. Skulsuppaisarn, L.K. Nguyen, and M. Lazarou. (2019) LC3/GABARAPs drive ubiquitin-independent recruitment of Optineurin and NDP52 to amplify mitophagy. *Nat. Commun.* **10**, 408
 82. Tumbarello, D.A., P.T. Manna, M. Allen, M. Bycroft, S.D. Arden, J. Kendrick-Jones, and F. Buss. (2015) The Autophagy Receptor TAX1BP1 and the Molecular Motor Myosin VI Are Required for Clearance of Salmonella Typhimurium by Autophagy. *PLOS Pathogens.* **11**, e1005174
 83. Hanna, R.A., M.N. Quinsay, A.M. Orogo, K. Giang, S. Rikka, and Å.B. Gustafsson. (2012) Microtubule-associated Protein 1 Light Chain 3 (LC3) Interacts with Bnip3 Protein to Selectively Remove Endoplasmic Reticulum and Mitochondria via Autophagy. *J. Biol. Chem.* **287**, 19094-19104
 84. Liu, L., D. Feng, G. Chen, M. Chen, Q. Zheng, P. Song, Q. Ma, C. Zhu, R. Wang, W. Qi, L. Huang, P. Xue, B. Li, X. Wang, H. Jin, J. Wang, F. Yang, P. Liu, Y. Zhu, S. Sui, and Q. Chen. (2012) Mitochondrial outer-membrane protein FUNDC1 mediates hypoxia-induced mitophagy in mammalian cells. *Nat. Cell Biol.* **14**, 177-185
 85. Strappazon, F., F. Nazio, M. Corrado, V. Cianfanelli, A. Romagnoli, G.M. Fimia, S. Campello, R. Nardacci, M. Piacentini, M. Campanella, and F. Cecconi. (2015) AMBRA1 is able to induce mitophagy via LC3 binding, regardless of PARKIN and p62/SQSTM1. *Cell Death Differ.* **22**, 419-432
 86. Murakawa, T., O. Yamaguchi, A. Hashimoto, S. Hikoso, T. Takeda, T. Oka, H. Yasui, H. Ueda, Y. Akazawa, H. Nakayama, M. Taneike, T. Misaka, S. Omiya, A.M. Shah, A. Yamamoto, K. Nishida, Y. Ohsumi, K. Okamoto, Y. Sakata, and K. Otsu. (2015) Bcl-2-like protein 13 is a mammalian Atg32 homologue that mediates mitophagy and mitochondrial fragmentation. *Nat. Commun.* **6**, 7527-7541
 87. Bhujabal, Z., Å.B. Birgisdottir, E. Sjøttem, H.B. Brenne, A. Øvervatn, S. Habisov, V. Kirkin, T. Lamark, and T. Johansen. (2017) FKBP8 recruits LC3A to mediate Parkin-independent mitophagy. *EMBO reports.* **18**, 947-961
 88. Wei, Y., W.-C. Chiang, R. Sumpter, Jr., P. Mishra, and B. Levine. (2017) Prohibitin 2 Is an Inner Mitochondrial Membrane Mitophagy Receptor. *Cell.* **168**, 224-238
 89. Zhang, Y., Y. Yao, X. Qiu, G. Wang, Z. Hu, S. Chen, Z. Wu, N. Yuan, H. Gao, J. Wang, H. Song, S.E. Girardin, and Y. Qian. (2019) Listeria hijacks host mitophagy through a novel mitophagy receptor to evade killing. *Nat. Immunol.* **20**, 433-446
 90. Khaminets, A., T. Heinrich, M. Mari, P. Grumati, A.K. Huebner, M. Akutsu, L. Liebmann, A. Stolz, S. Nietzsche, N. Koch, M. Mauthe, I. Katona, B. Qualmann, J. Weis, F. Reggiori, I. Kurth, C.A. Hübner, and I. Dikic. (2015) Regulation of endoplasmic reticulum turnover by selective autophagy. *Nature.* **522**, 354-358
 91. Fumagalli, F., J. Noack, T.J. Bergmann, E. Cebollero, G.B. Pisoni, E. Fasana, I. Fregno, C. Galli, M. Loi, T. Soldà, R. D'Antuono, A. Raimondi, M. Jung, A. Melnyk, S. Schorr, A. Schreiber, L. Simonelli, L. Varani, C. Wilson-Zbinden, O. Zerbe, K. Hofmann, M. Peter, M. Quadroni, R. Zimmermann, and M. Molinari. (2016) Translocon component Sec62 acts in endoplasmic reticulum turnover during stress recovery. *Nat. Cell Biol.* **18**, 1173-1184
 92. Grumati, P., G. Morozzi, S. Hölper, M. Mari, M.-L.I.E. Harwardt, R. Yan, S. Müller, F. Reggiori, M. Heilemann, and I. Dikic. (2017) Full length RTN3 regulates turnover of tubular endoplasmic reticulum via selective autophagy. *eLife.* **6**, e25555

93. Lahiri, V. and D.J. Klionsky. (2018) CCPG1 is a noncanonical autophagy cargo receptor essential for reticulophagy and pancreatic ER proteostasis. *Autophagy*. **14**, 1107-1109
94. Chen, Q., Y. Xiao, P. Chai, P. Zheng, J. Teng, and J. Chen. (2019) AT13 Is a Tubular ER-Phagy Receptor for GABARAP-Mediated Selective Autophagy. *Curr. Biol.* **29**, 846-855.e846
95. An, H., A. Ordureau, J.A. Paulo, C.J. Shoemaker, V. Denic, and J.W. Harper. (2019) TEX264 Is an Endoplasmic Reticulum-Resident ATG8-Interacting Protein Critical for ER Remodeling during Nutrient Stress. *Mol Cell*. **74**, 891-908.e810
96. Jiang, S., C.D. Wells, and P.J. Roach. (2011) Starch-binding domain-containing protein 1 (Stbd1) and glycogen metabolism: Identification of the Atg8 family interacting motif (AIM) in Stbd1 required for interaction with GABARAP1. *Biochem. Biophys. Res. Commun.* **413**, 420-425
97. Wyant, G.A., M. Abu-Remaileh, E.M. Frenkel, N.N. Laqtom, V. Dharamdasani, C.A. Lewis, S.H. Chan, I. Heinze, A. Ori, and D.M. Sabatini. (2018) NUFIP1 is a ribosome receptor for starvation-induced ribophagy. *Science*. **360**, 751-758
98. Novak, I., V. Kirkin, D.G. McEwan, J. Zhang, P. Wild, A. Rozenknop, V. Rogov, F. Löhr, D. Popovic, A. Occhipinti, A.S. Reichert, J. Terzic, V. Dötsch, P.A. Ney, and I. Dikic. (2010) Nix is a selective autophagy receptor for mitochondrial clearance. *EMBO reports*. **11**, 45-51
99. Noda, N.N., Y. Ohsumi, and F. Inagaki. (2010) Atg8-family interacting motif crucial for selective autophagy. *FEBS Letters*. **584**, 1379-1385
100. Shvets, E., E. Fass, R. Scherz-Shouval, and Z. Elazar. (2008) The N-terminus and Phe52 residue of LC3 recruit p62/SQSTM1 into autophagosomes. *J. Cell Sci.* **121**, 2685-2695
101. Rogov, Vladimir V., H. Suzuki, E. Fiskin, P. Wild, A. Kniss, A. Rozenknop, R. Kato, M. Kawasaki, David G. McEwan, F. Löhr, P. Güntert, I. Dikic, S. Wakatsuki, and V. Dötsch. (2013) Structural basis for phosphorylation-triggered autophagic clearance of Salmonella. *Biochem. J.* **454**, 459-466
102. Wirth, M., S. Mouilleron, W. Zhang, E. Sjøttem, Y. Princely Abudu, A. Jain, H. Lauritz Olsvik, J.-A. Bruun, M. Razi, H.B.J. Jefferies, R. Lee, D. Joshi, N. O'Reilly, T. Johansen, and S.A. Tooze. (2021) Phosphorylation of the LIR Domain of SCOC Modulates ATG8 Binding Affinity and Specificity. *J. Mol. Biol.* **433**, 166987
103. Goode, A., K. Butler, J. Long, J. Cavey, D. Scott, B. Shaw, J. Sollenberger, C. Gell, T. Johansen, N.J. Oldham, M.S. Searle, and R. Layfield. (2016) Defective recognition of LC3B by mutant SQSTM1/p62 implicates impairment of autophagy as a pathogenic mechanism in ALS-FTLD. *Autophagy*. **12**, 1094-1104
104. Birgisdottir, Å.B., S. Mouilleron, Z. Bhujabal, M. Wirth, E. Sjøttem, G. Evjen, W. Zhang, R. Lee, N. O'Reilly, S.A. Tooze, T. Lamark, and T. Johansen. (2019) Members of the autophagy class III phosphatidylinositol 3-kinase complex I interact with GABARAP and GABARAP1 via LIR motifs. *Autophagy*. **15**, 1333-1355
105. Genau, Heide M., J. Huber, F. Baschieri, M. Akutsu, V. Dötsch, H. Farhan, V. Rogov, and C. Behrends. (2015) CUL3-KBTBD6/KBTBD7; Ubiquitin Ligase Cooperates with GABARAP Proteins to Spatially Restrict TIAM1-RAC1 Signaling. *Mol Cell*. **57**, 995-1010
106. Lystad, A.H., Y. Ichimura, K. Takagi, Y. Yang, S. Pankiv, Y. Kanegae, S. Kageyama, M. Suzuki, I. Saito, T. Mizushima, M. Komatsu, and A. Simonsen. (2014) Structural determinants in GABARAP required for the selective binding and recruitment of ALFY to LC3B-positive structures. *EMBO reports*. **15**, 557-565

107. Rogov, V.V., A. Stolz, A.C. Ravichandran, D.O. Rios-Szwed, H. Suzuki, A. Kniss, F. Löhr, S. Wakatsuki, V. Dötsch, I. Dikic, R.C. Dobson, and D.G. McEwan. (2017) Structural and functional analysis of the GABARAP interaction motif (GIM). *EMBO reports*. **18**, 1382-1396
108. Jakobi, A.J., S.T. Huber, S.A. Mortensen, S.W. Schultz, A. Palara, T. Kuhm, B.K. Shrestha, T. Lamark, W.J.H. Hagen, M. Wilmanns, T. Johansen, A. Brech, and C. Sachse. (2020) Structural basis of p62/SQSTM1 helical filaments and their role in cellular cargo uptake. *Nat. Commun.* **11**, 440
109. Lamark, T., S. Svenning, and T. Johansen. (2017) Regulation of selective autophagy: the p62/SQSTM1 paradigm. *Essays Biochem.* **61**, 609-624
110. Komatsu, M., S. Waguri, T. Ueno, J. Iwata, S. Murata, I. Tanida, J. Ezaki, N. Mizushima, Y. Ohsumi, Y. Uchiyama, E. Kominami, K. Tanaka, and T. Chiba (2005) Impairment of starvation-induced and constitutive autophagy in Atg7-deficient mice. *J. Cell Biol.* **169**, 425-434
111. Singh, I., A.K. Singh, and M.A. Contreras. (2009) Peroxisomal dysfunction in inflammatory childhood white matter disorders: an unexpected contributor to neuropathology. *J Child Neurol.* **24**, 1147-1157
112. Kraft, C., M. Peter, and K. Hofmann. (2010) Selective autophagy: ubiquitin-mediated recognition and beyond. *Nat. Cell Biol.* **12**, 836-841
113. Rozenknop, A., V.V. Rogov, N.Y. Rogova, F. Löhr, P. Güntert, I. Dikic, and V. Dötsch. (2011) Characterization of the Interaction of GABARAPL-1 with the LIR Motif of NBR1. *J. Mol. Biol.* **410**, 477-487
114. Sollenberger, J. (2011) MSci(Hons) research project report - Biophysical studies of the molecular interactions of autophagy marker LC3B, Univ. Nottingham.
115. Deribe, Y.L., T. Pawson, and I. Dikic. (2010) Post-translational modifications in signal integration. *Nat. Struct. Mol. Biol.* **17**, 666-672
116. Xie, Y., R. Kang, X. Sun, M. Zhong, J. Huang, D.J. Klionsky, and D. Tang. (2015) Posttranslational modification of autophagy-related proteins in macroautophagy. *Autophagy.* **11**, 28-45
117. Gubas, A. and I. Dikic. (2021) A guide to the regulation of selective autophagy receptors. *The FEBS Journal.* **n/a**,
118. Vainshtein, A. and P. Grumati. (2020) Selective Autophagy by Close Encounters of the Ubiquitin Kind. *Cells.* **9**, 2349
119. Cherra, S.J., 3rd, S.M. Kulich, G. Uechi, M. Balasubramani, J. Mountzouris, B.W. Day, and C.T. Chu. (2010) Regulation of the autophagy protein LC3 by phosphorylation. *J Cell Biol.* **190**, 533-539
120. Jiang, H., D. Cheng, W. Liu, J. Peng, and J. Feng. (2010) Protein kinase C inhibits autophagy and phosphorylates LC3. *Biochem. Biophys. Res. Commun.* **395**, 471-476
121. Wilkinson, D.S., J.S. Jariwala, E. Anderson, K. Mitra, J. Meisenhelder, J.T. Chang, T. Ideker, T. Hunter, V. Nizet, A. Dillin, and M. Hansen. (2015) Phosphorylation of LC3 by the Hippo kinases STK3/STK4 is essential for autophagy. *Mol Cell.* **57**, 55-68
122. Shrestha, B.K., M. Skytte Rasmussen, Y.P. Abudu, J.-A. Bruun, K.B. Larsen, E.A. Alemu, E. Sjøttem, T. Lamark, and T. Johansen. (2020) NIMA-related kinase 9-mediated phosphorylation of the microtubule-associated LC3B protein at Thr-50 suppresses selective autophagy of p62/sequestosome 1. *J. Biol. Chem.* **295**, 1240-1260
123. van Zundert, G.C.P., J.P.G.L.M. Rodrigues, M. Trellet, C. Schmitz, P.L. Kastiris, E. Karaca, A.S.J. Melquiond, M. van Dijk, S.J. de Vries, and A.M.J.J. Bonvin. (2016) The

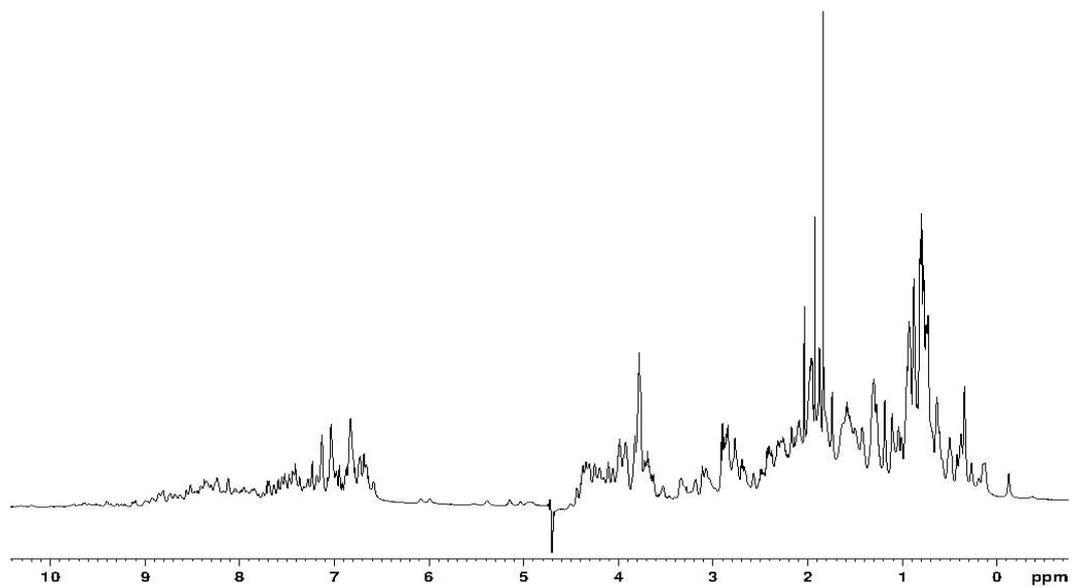
- HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **428**, 720-725
124. Cavanagh, J., W.J. Fairbrother, A.G. Palmer, M. Rance, and N.J. Skelton. (2007) Chapter 10 -Sequential assignment, structure determination and other applications *Protein NMR Spectroscopy (Second Edition)*. 781-817
 125. Grzesiek, S. and A. Bax. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.* **114**, 6291-6293
 126. Kay, L.E., M. Ikura, R. Tschudin, and A. Bax. (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance (1969)*. **89**, 496-514
 127. Clubb, R.T., V. Thanabal, and G. Wagner. (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intrareidue ¹HN, ¹⁵N, and ¹³C chemical shifts in ¹⁵N ¹³C-labelled proteins. *Journal of Magnetic Resonance (1969)*. **97**, 213-217
 128. Bax, A. and M. Ikura. (1991) An efficient 3D NMR technique for correlating the proton and ¹⁵N backbone amide resonances with the α -carbon of the preceding residue in uniformly ¹⁵N/¹³C enriched proteins. *J. Biomol. NMR.* **1**, 99-104
 129. Grzesiek, S. and A. Bax. (1992) An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *Journal of Magnetic Resonance (1969)*. **99**, 201-207
 130. Skinner, S.P., R.H. Fogh, W. Boucher, T.J. Ragan, L.G. Mureddu, and G.W. Vuister. (2016) CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J. Biomol. NMR.* **66**, 111-124
 131. Williamson, M.P. (2013) Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1-16
 132. Potter, L.R. and T. Hunter. (1998) Phosphorylation of the Kinase Homology Domain Is Essential for Activation of the A-Type Natriuretic Peptide Receptor. *Mol. Cell. Biol.* **18**, 2164-2172
 133. Hao, M., A.M. Lowy, M. Kapoor, A. Deffie, G. Liu, and G. Lozano. (1996) Mutation of phosphoserine 389 affects p53 function in vivo. *J Biol Chem.* **271**, 29380-29385
 134. Gasteiger, E., C. Hoogland, A. Gattiker, S.e. Duvaud, M.R. Wilkins, R.D. Appel, and A. Bairoch. (2005) Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook*. 571-607
 135. Fenn, J., M. Mann, C. Meng, S. Wong, and C. Whitehouse. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science.* **246**, 64-71
 136. Banerjee, S. and S. Mazumdar. (2012) Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *Int. J. Anal. Chem.* **2012**, 282574-282614
 137. Katta, V. and B.T. Chait. (1991) Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry. *J. Am. Chem. Soc.* **113**, 8534-8535
 138. Kitova, E.N., A. El-Hawiet, P.D. Schnier, and J.S. Klassen. (2012) Reliable Determinations of Protein–Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet? *J. Am. Soc. Mass. Spectrom.* **23**, 431-441

139. Kouno, T., M. Mizuguchi, I. Tanida, T. Ueno, E. Kominami, and K. Kawano. (2004) Letter to the Editor: ^1H , ^{13}C , and ^{15}N Resonance Assignments of Human Microtubule-associated Protein Light Chain-3. *J. Biomol. NMR.* **29**, 415-416
140. Bordogna, A., A. Pandini, and L. Bonati. (2011) Predicting the accuracy of protein-ligand docking on homology models. *J. Comput. Chem.* **32**, 81-98
141. Kurkcuoglu, Z., P.I. Koukos, N. Citro, M.E. Trellet, J. Rodrigues, I.S. Moreira, J. Roel-Touris, A.S.J. Melquiond, C. Geng, J. Schaarschmidt, L.C. Xue, A. Vangone, and A. Bonvin. (2018) Performance of HADDOCK and a simple contact-based protein-ligand binding affinity predictor in the D3R Grand Challenge 2. *J. Comput. Aided Mol. Des.* **32**, 175-185
142. Joachim, J. and S.A. Tooze. (2016) GABARAP activates ULK1 and traffics from the centrosome dependent on Golgi partners WAC and GOLGA2/GM130. *Autophagy.* **12**, 892-893

Appendix

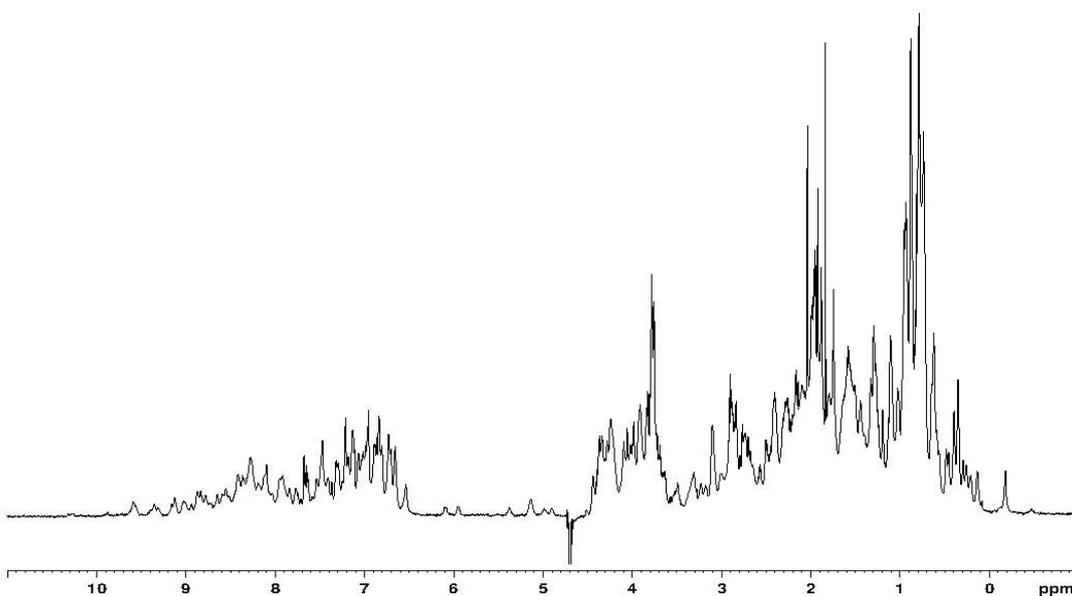
A. Appendix A

A.1. LC3B WT (original construct)



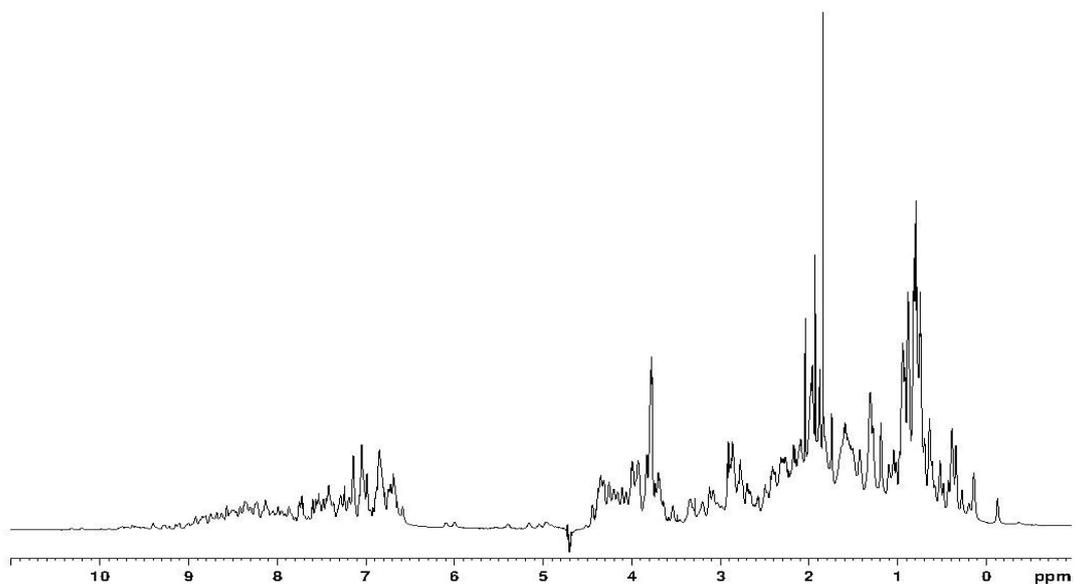
¹H NMR Spectrum of LC3B WT (0.3mM, pH7.0), recorded at 298K on Bruker Avance(III) 800MHz

A.2. LC3B KL



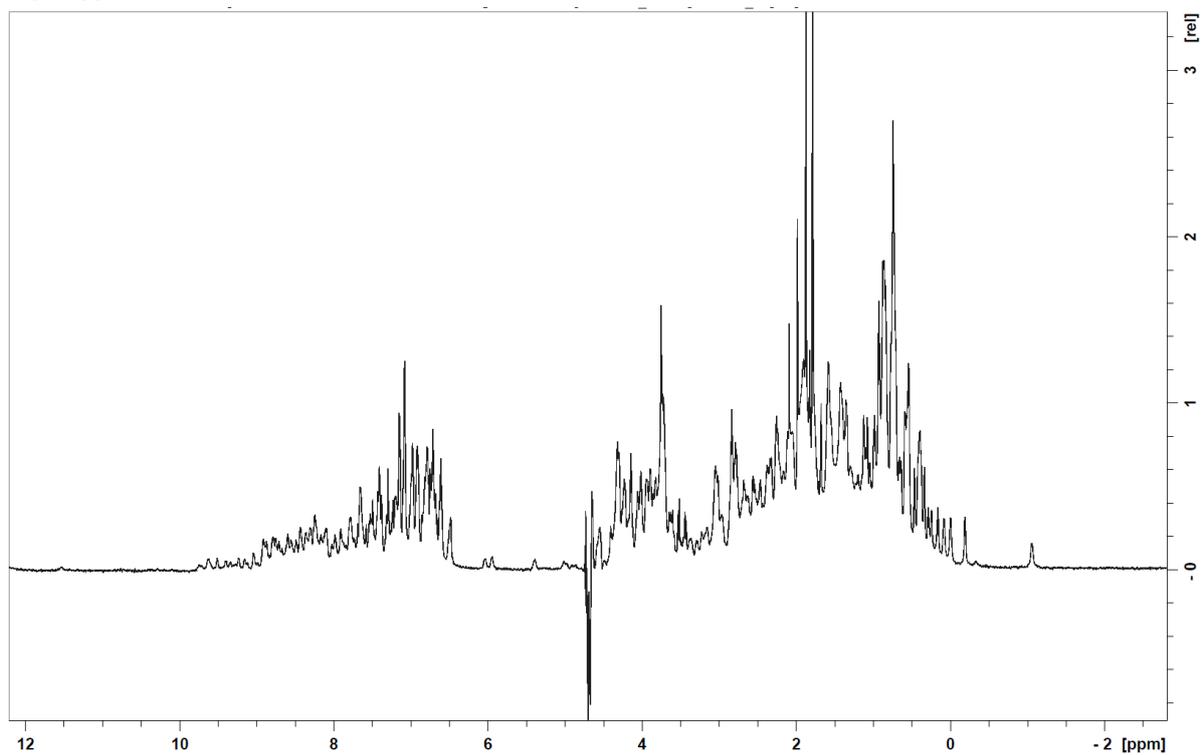
¹H NMR Spectrum of LC3B KL (0.3mM, pH7.0), recorded at 298K on Bruker Avance(III) 800MHz

A.3. LC3B T29D



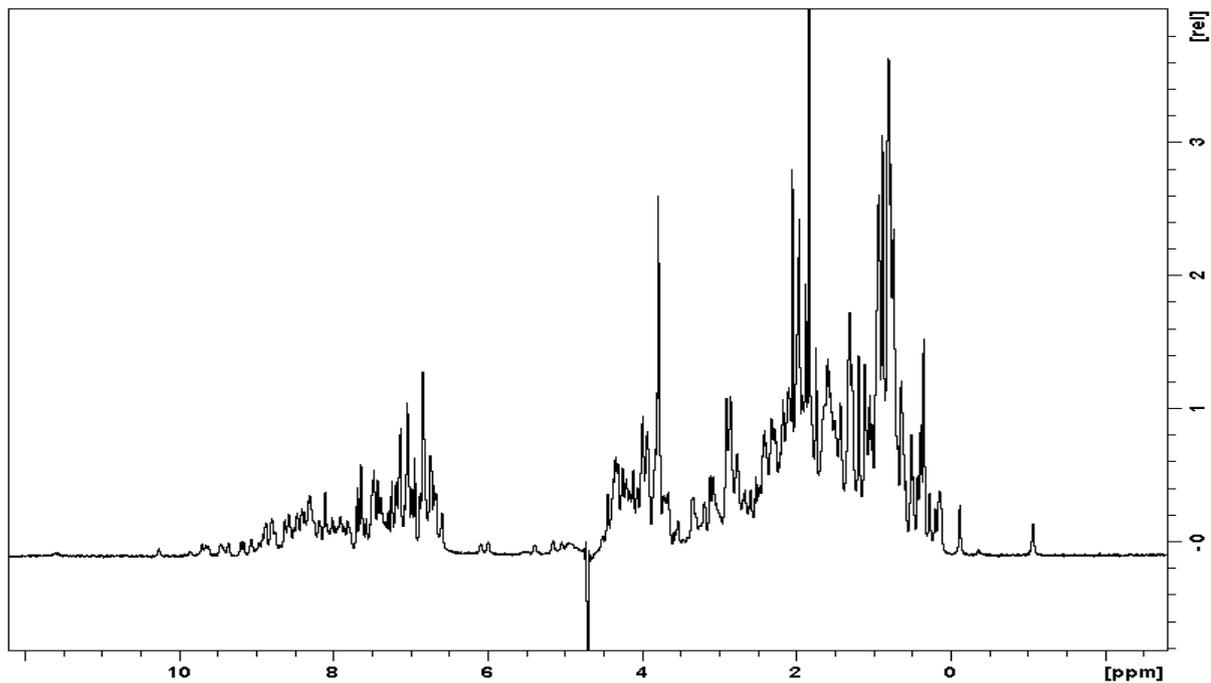
^1H NMR Spectrum of LC3B T29D (0.3mM, pH7.0), recorded at 298K on Bruker Avance(III) 800MHz

A.5. LC3A



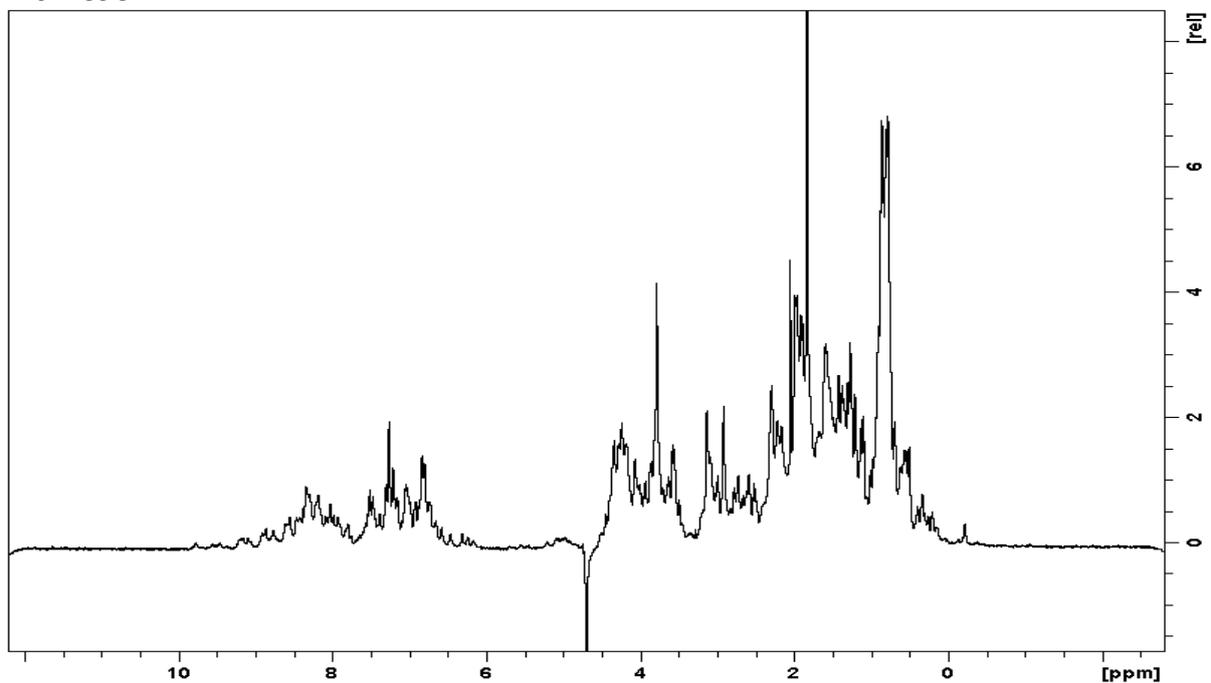
^1H NMR Spectrum of LC3A WT (0.25mM, pH7.0), recorded at 298K on Bruker Avance(III) 800MHz

A.5. LC3B WT (new construct)



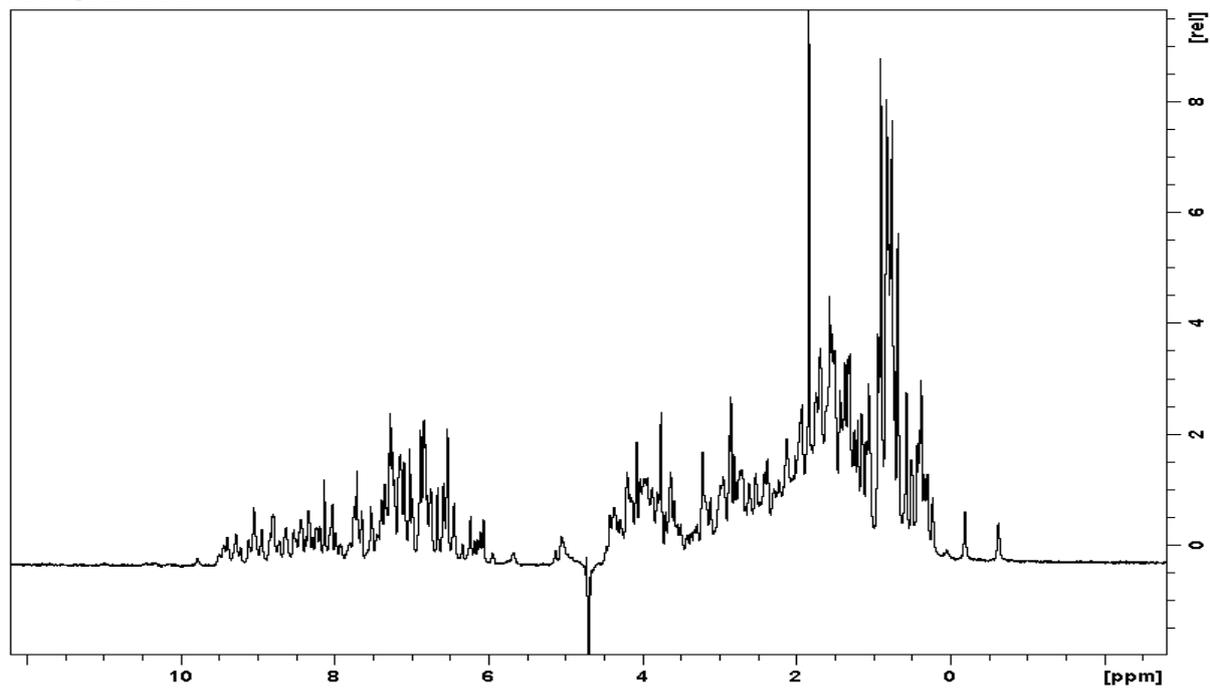
¹H NMR Spectrum of LC3B WT from new construct (0.3mM, pH7.0), recorded at 298K on Bruker Avance(III) 800MHz

A.6. LC3C



¹H NMR Spectrum of LC3C WT (0.3mM, pH7.0), recorded at 298K on Bruker Avance(III) 800MHz

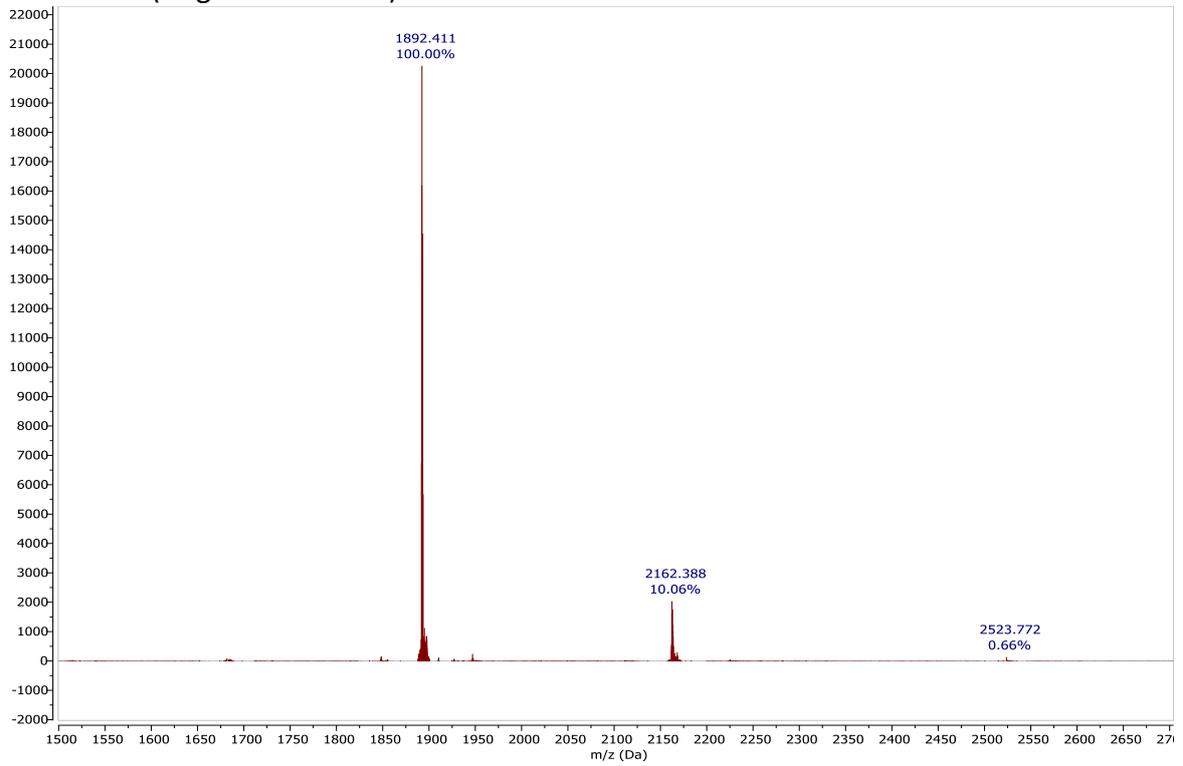
A.7. GABARAP



¹H NMR Spectrum of GABARAP WT (0.3mM, pH7.0), recorded at 298K on Bruker Avance(III) 800MHz

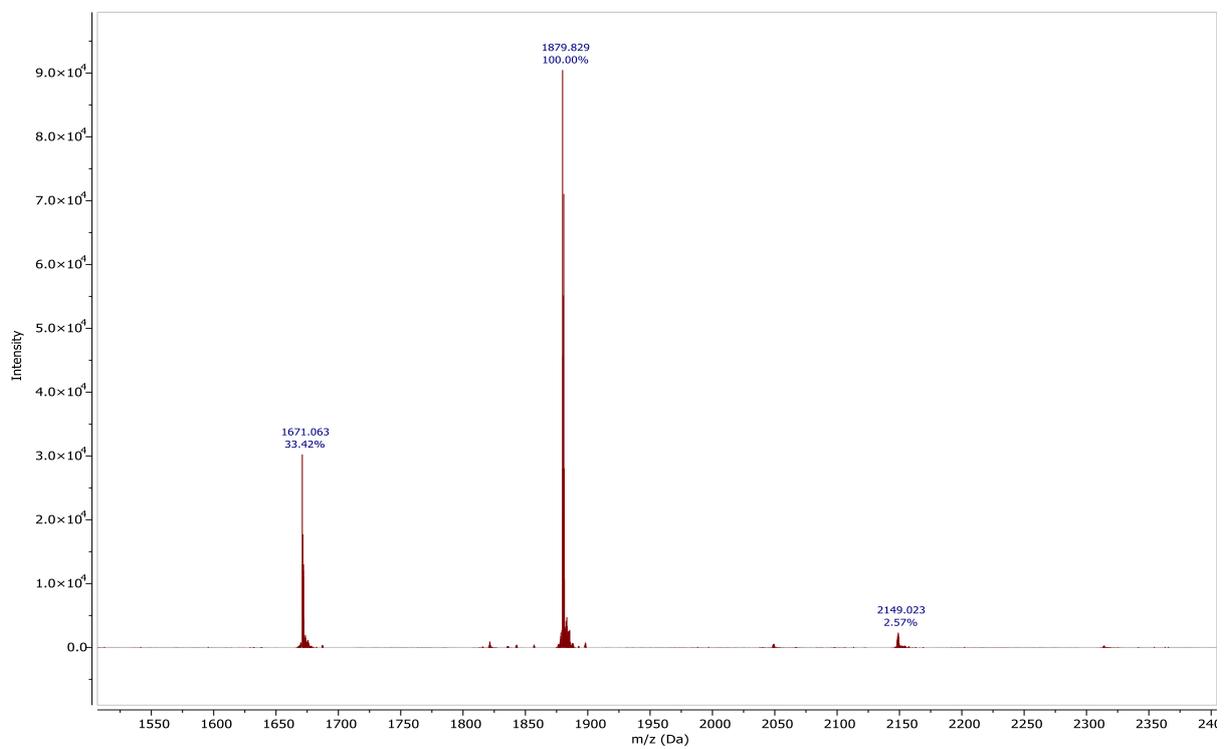
B. Appendix B

B.1 LC3B WT (original construct)



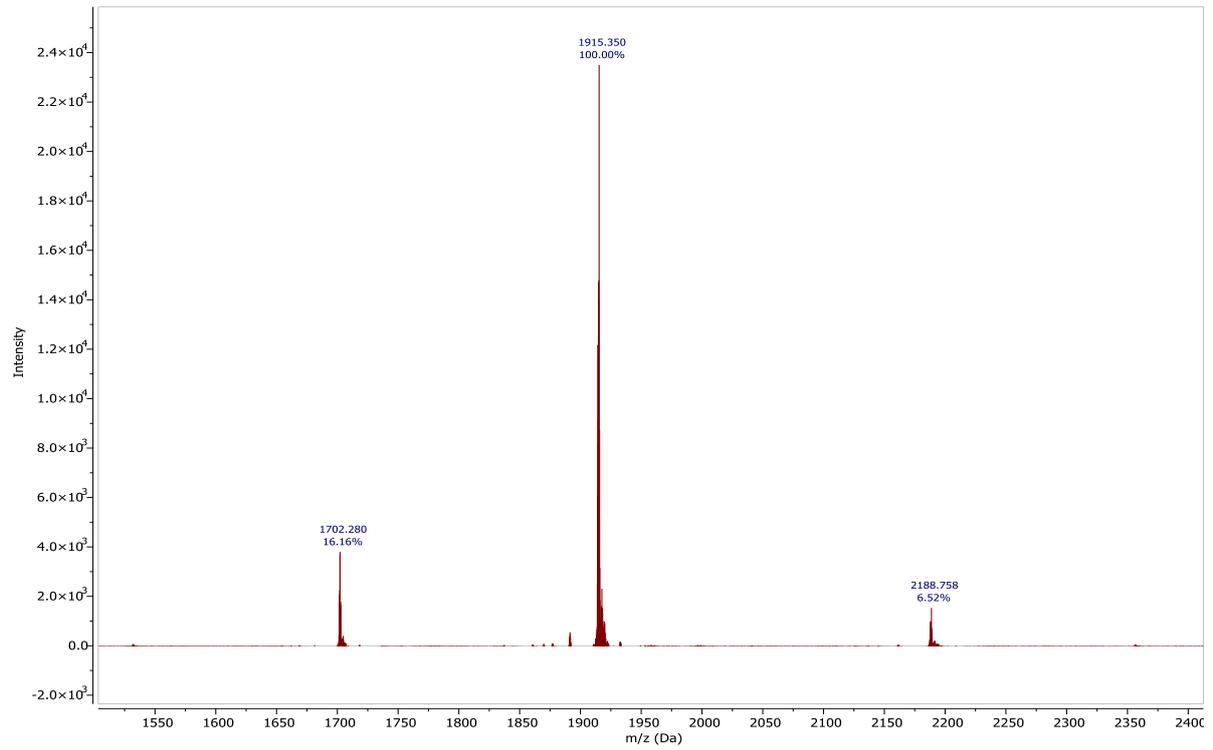
ESI-TOF MS spectrum of LC3B WT (5 μM)

B.2 LC3B KL



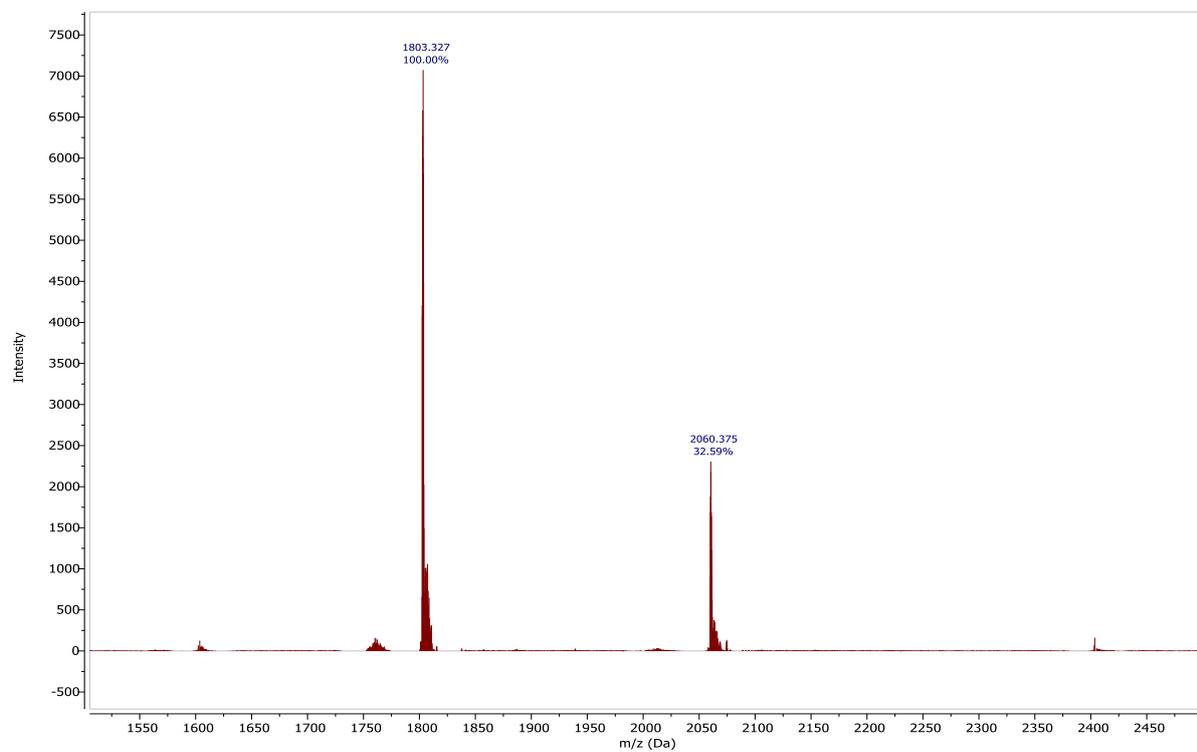
ESI-TOF MS spectrum of LC3B KL (5 μM)

B.3 LC3B T29D



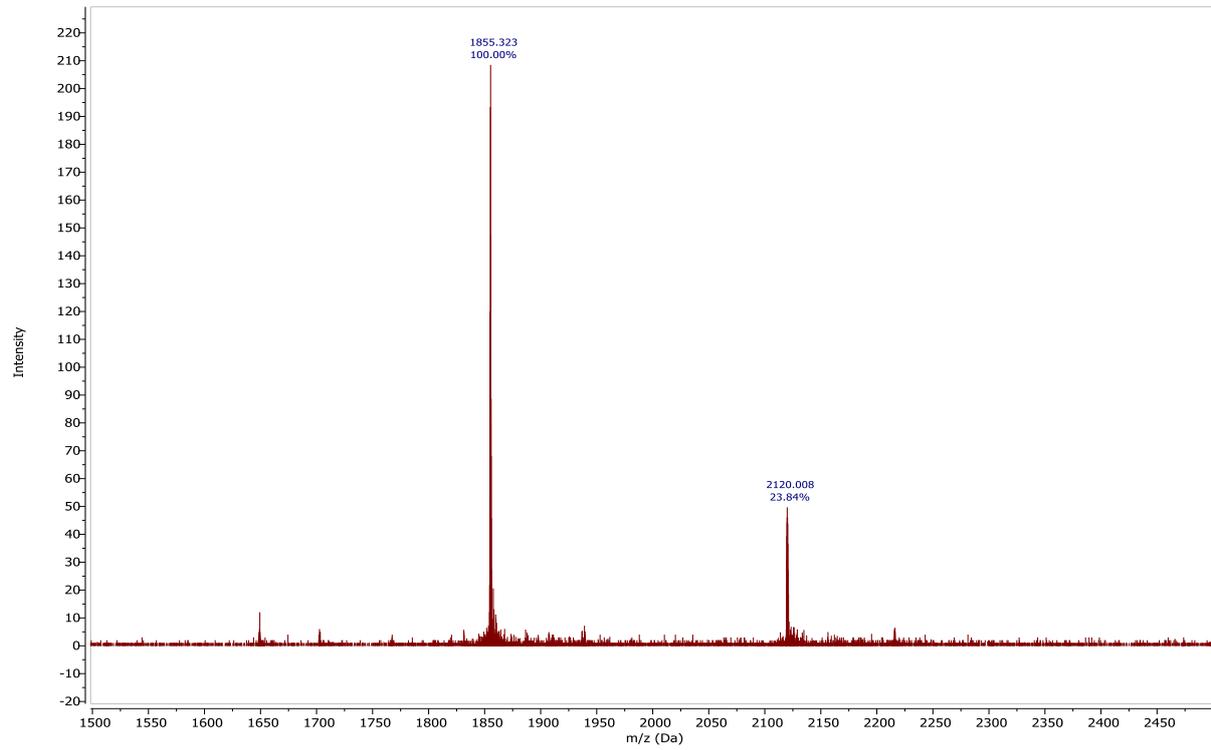
ESI-TOF MS spectrum of LC3B T29D (5 μM)

B.4 LC3A



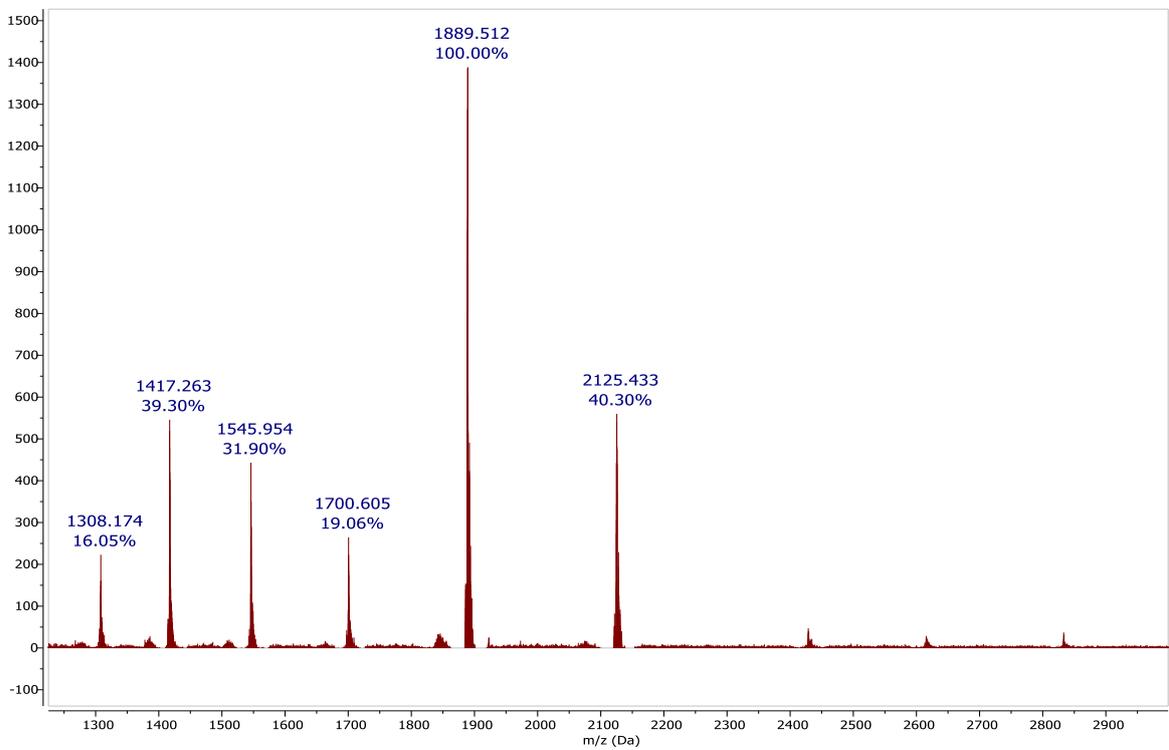
ESI-TOF MS spectrum of LC3A WT (5 μM)

B.5 LC3B (new construct)



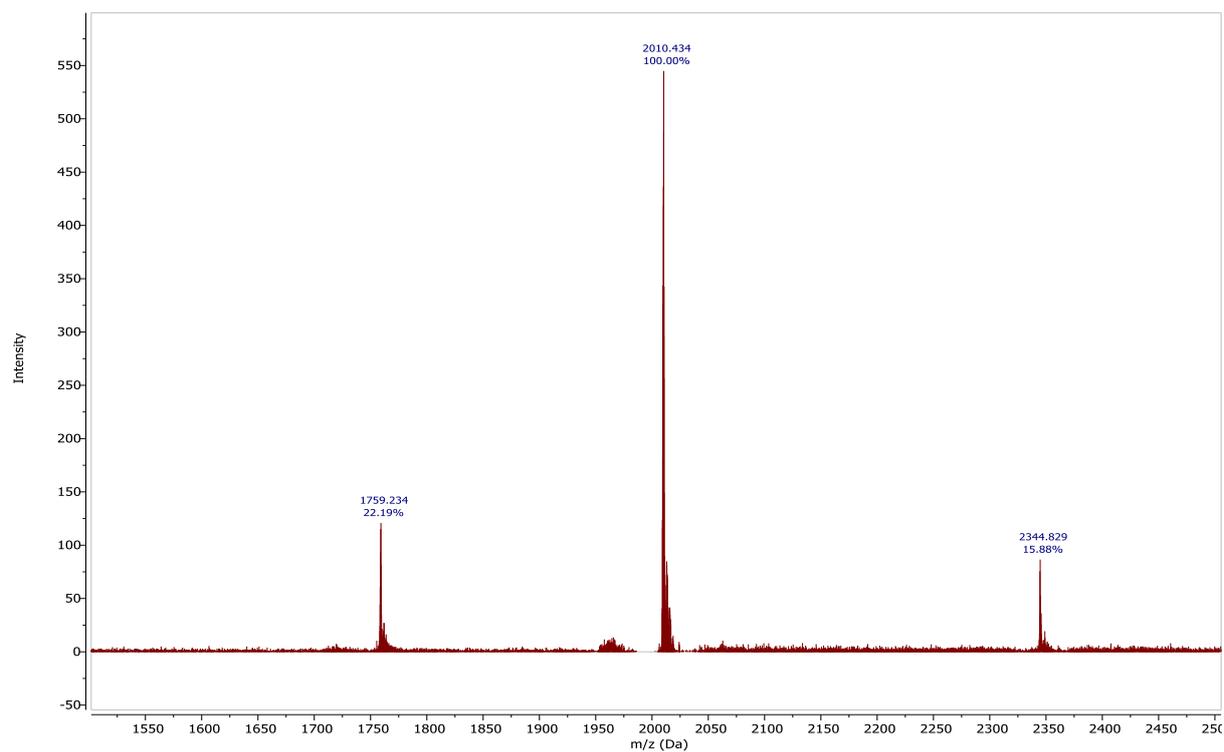
ESI-TOF MS spectrum of LC3A WT (5 μM)

B.6 LC3C



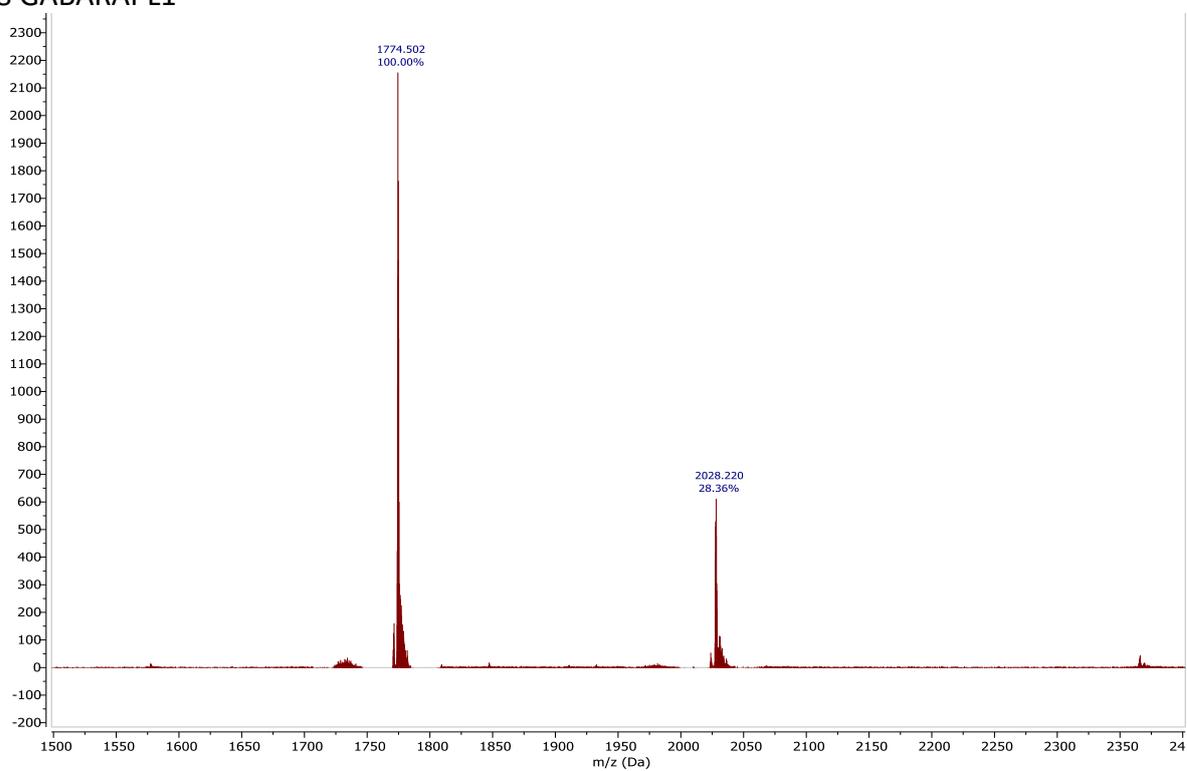
ESI-TOF MS spectrum of LC3C WT (5 μM)

B.7 GABARAP



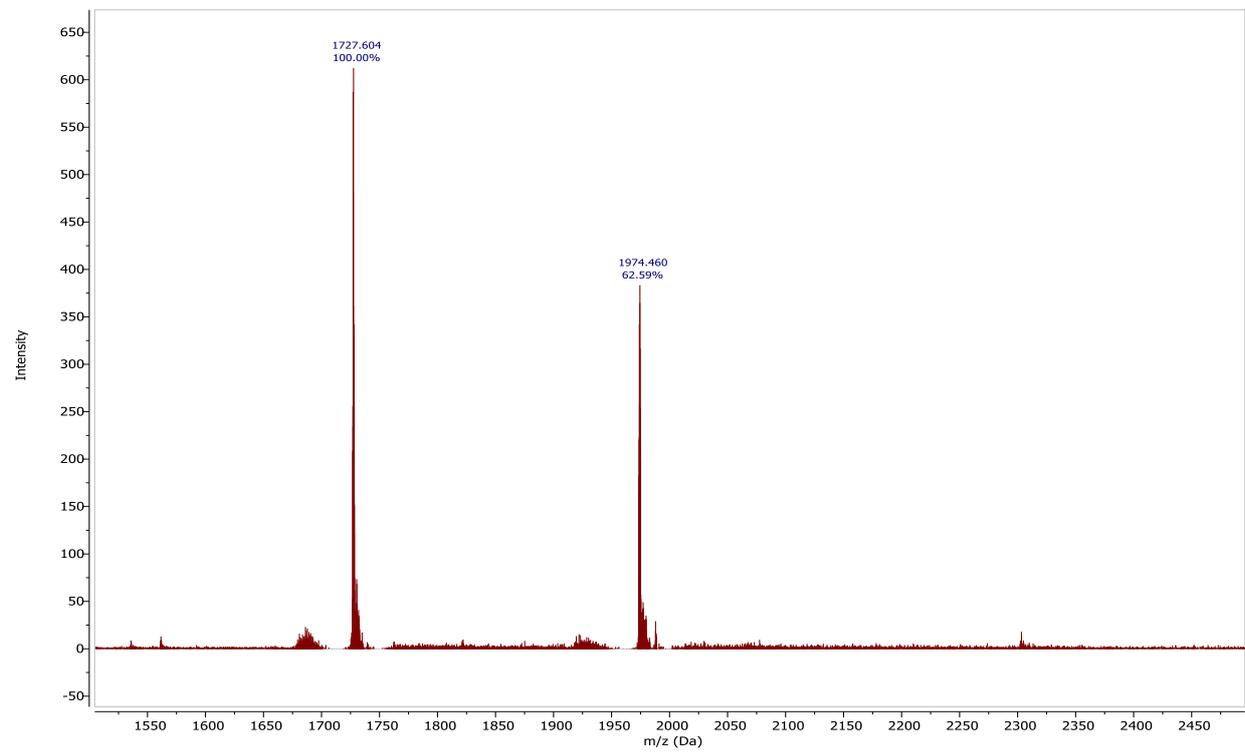
ESI-TOF MS spectrum of GABARAP WT (5 μM)

B.8 GABARAPL1



ESI-TOF MS spectrum of GABARAPL1 (5 μM)

B.9 GABARAPL2

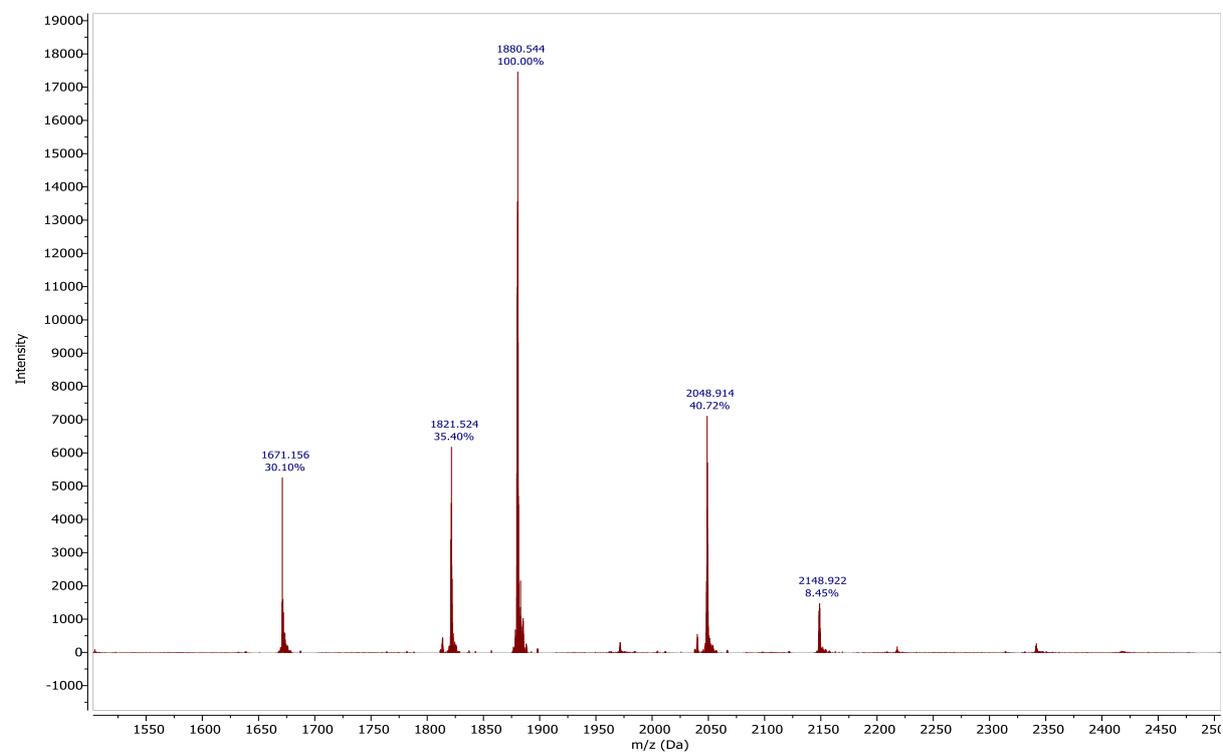


ESI-TOF MS spectrum of GABARAPL2 (5 μM)

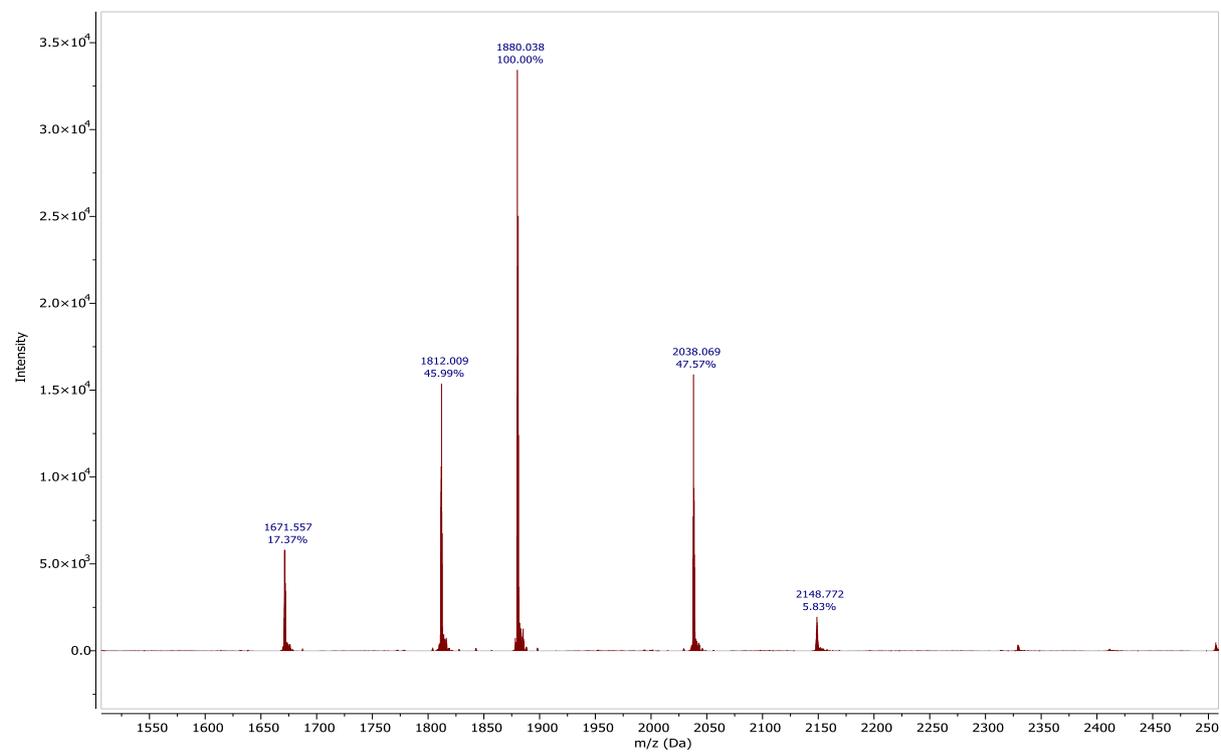
C. Appendix C

C.1 LC3B KL

KL LIR1



KL LIR2

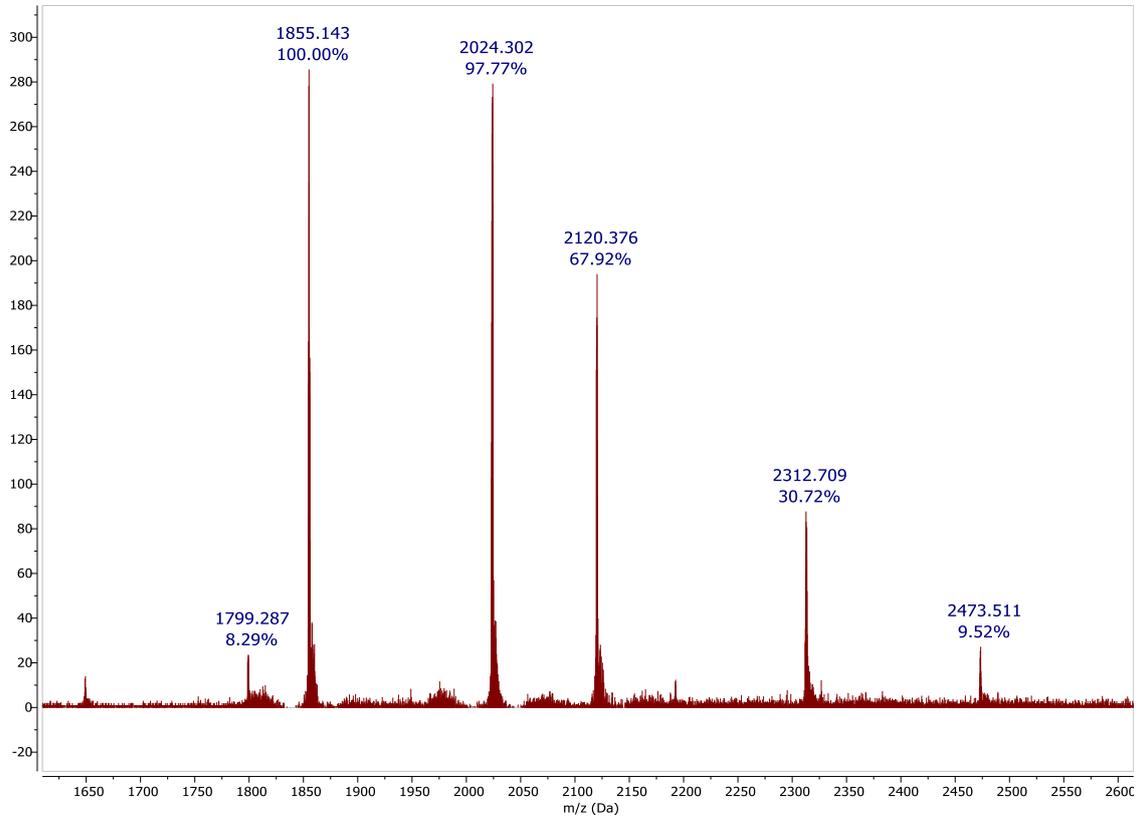


| Peak | Species | Charge state | Predicted |
|---------|----------------|--------------|-----------|
| 1671.16 | LC3B KL | +9 | 1671.14 |
| 1821.52 | LC3B KL + LIR1 | +9 | 1821.08 |
| 1880.54 | LC3B KL | +7 | 1879.90 |
| 2048.91 | LC3B KL + LIR1 | +7 | 2048.59 |
| 2148.92 | LC3B KL | +6 | 2148.32 |
| | | | |
| 1671.56 | LC3B KL | +9 | 1671.14 |
| 1812.01 | LC3B KL + LIR2 | +9 | 1811.52 |
| 1880.04 | LC3B KL | +8 | 1879.90 |
| 2038.07 | LC3B KL + LIR2 | +8 | 2037.83 |
| 2148.77 | LC3B KL | +7 | 2148.32 |

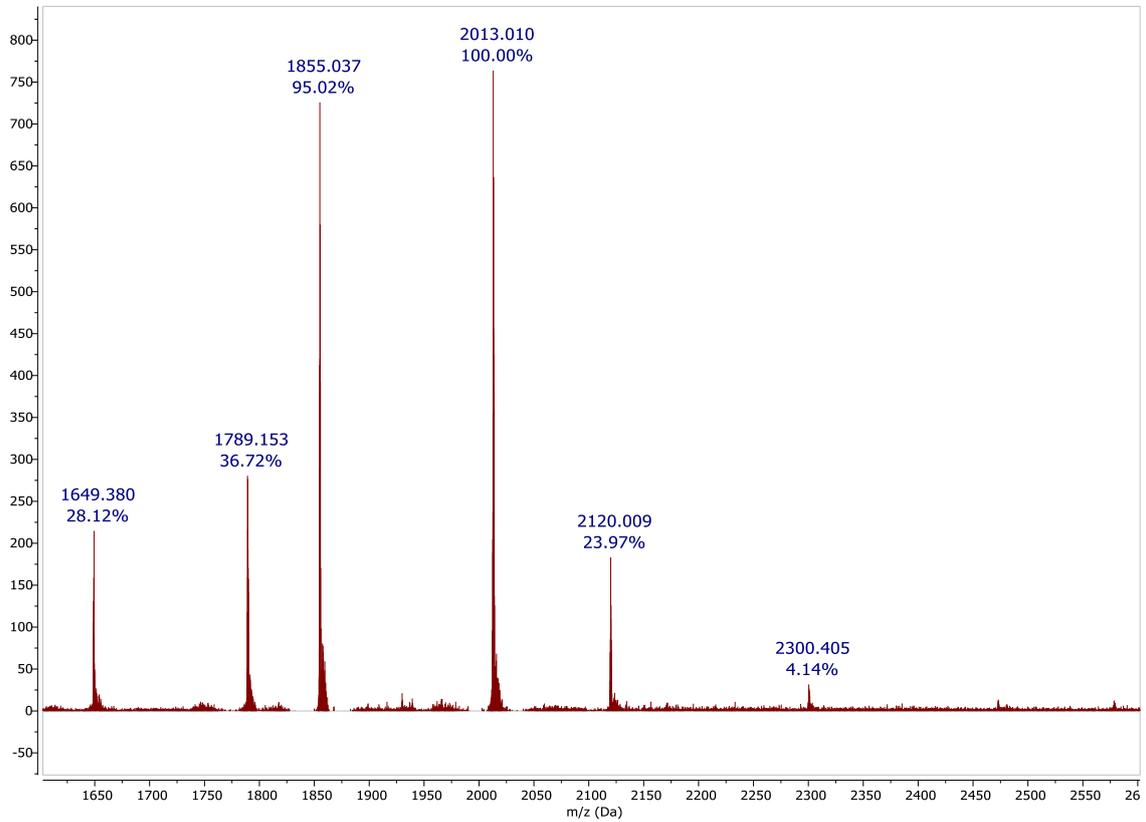
MS spectra for LC3B KL binding with LIRs of NBR1 and table of charge states observed.

C.2 LC3B WT (new construct)

LC3B-LIR1



LC3B-LIR2

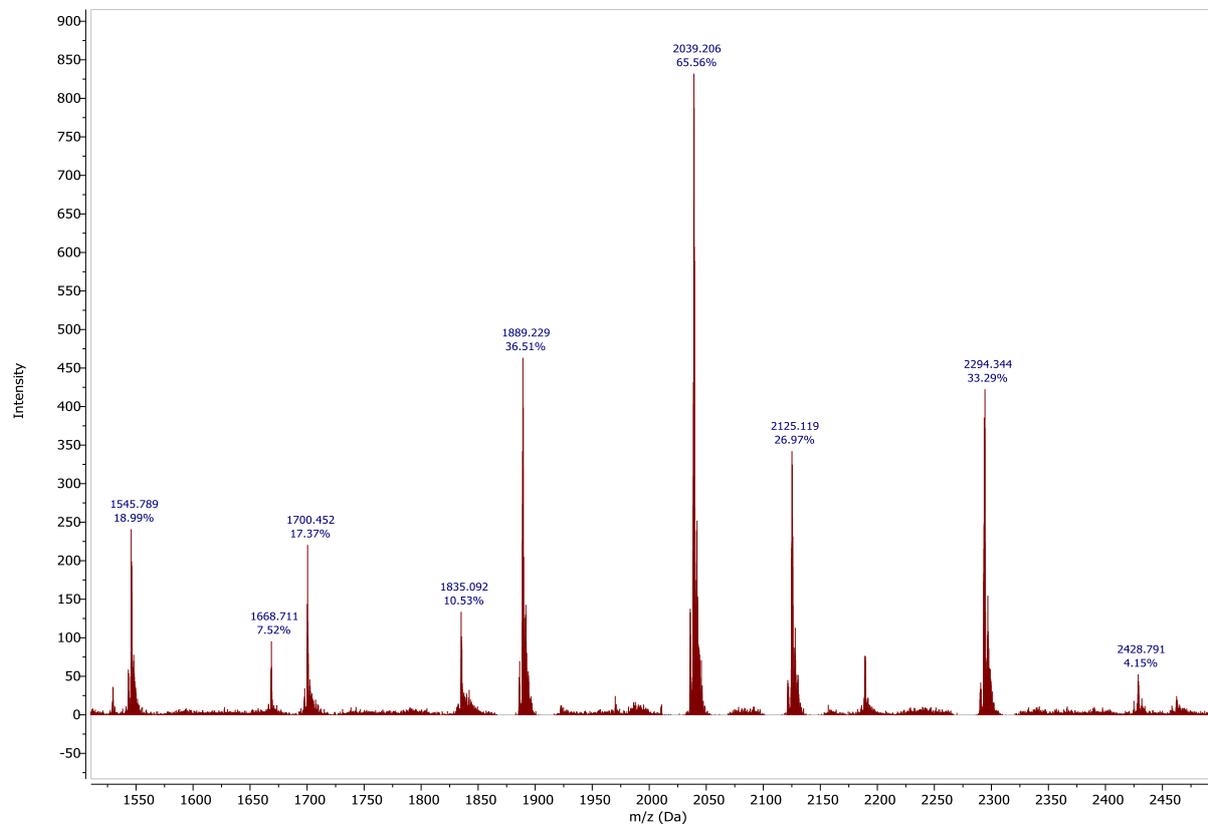


| Peak | Species | Charge state | Predicted |
|---------|-------------|--------------|-----------|
| 1799.29 | LC3B + LIR1 | +9 | 1798.97 |
| 1855.14 | LC3B | +8 | 1855.03 |
| 2024.30 | LC3B + LIR1 | +8 | 2023.71 |
| 2120.38 | LC3B | +7 | 2119.89 |
| 2312.71 | LC3B + LIR1 | +7 | 2312.67 |
| 2473.51 | LC3B | +6 | 2473.03 |
| | | | |
| 1649.38 | LC3B | +9 | 1649.02 |
| 1789.15 | LC3B + LIR2 | +9 | 1789.40 |
| 1855.04 | LC3B | +8 | 1855.03 |
| 2013.01 | LC3B + LIR2 | +8 | 2012.95 |
| 2120.01 | LC3B | +7 | 2119.89 |
| 2300.41 | LC3B + LIR2 | +7 | 2300.37 |

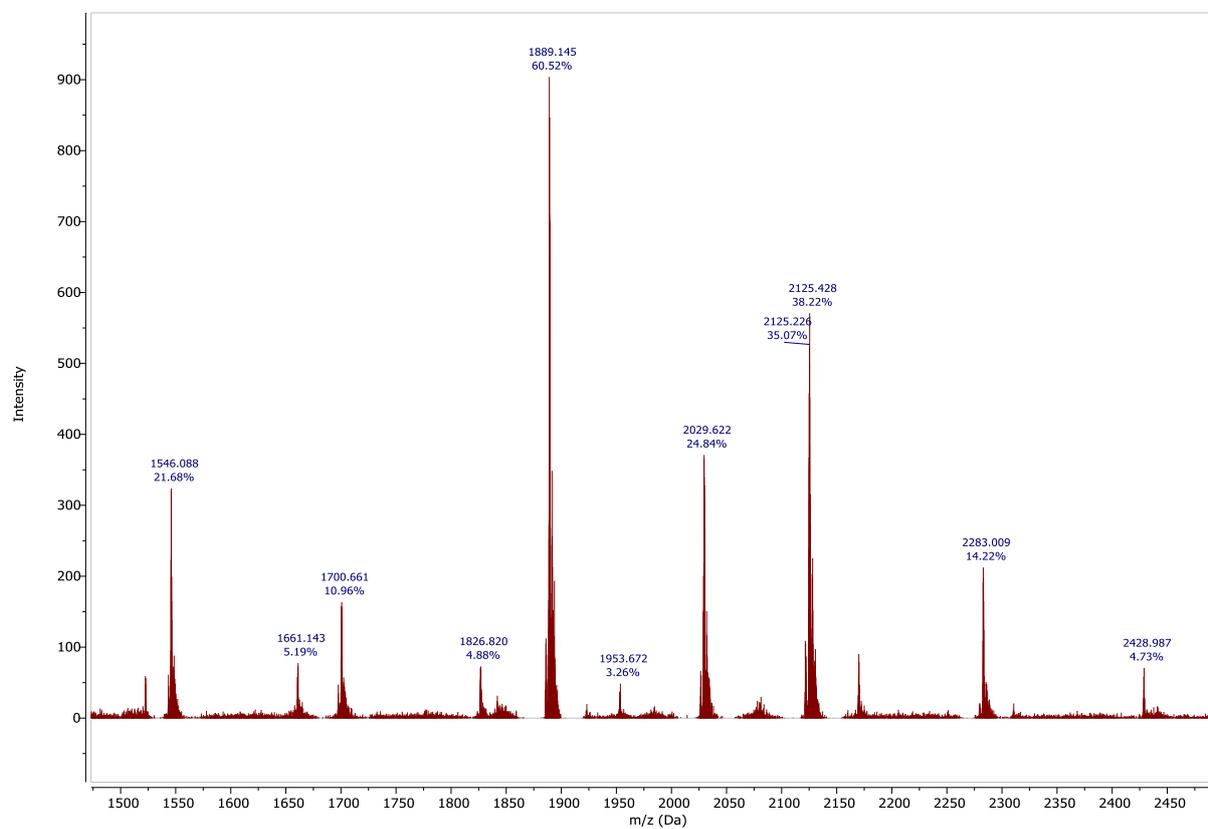
MS spectra for LC3B WT binding with LIRs of NBR1 and table of charge states observed.

C.3 LC3C

LC3C-LIR1



LC3C-LIR2

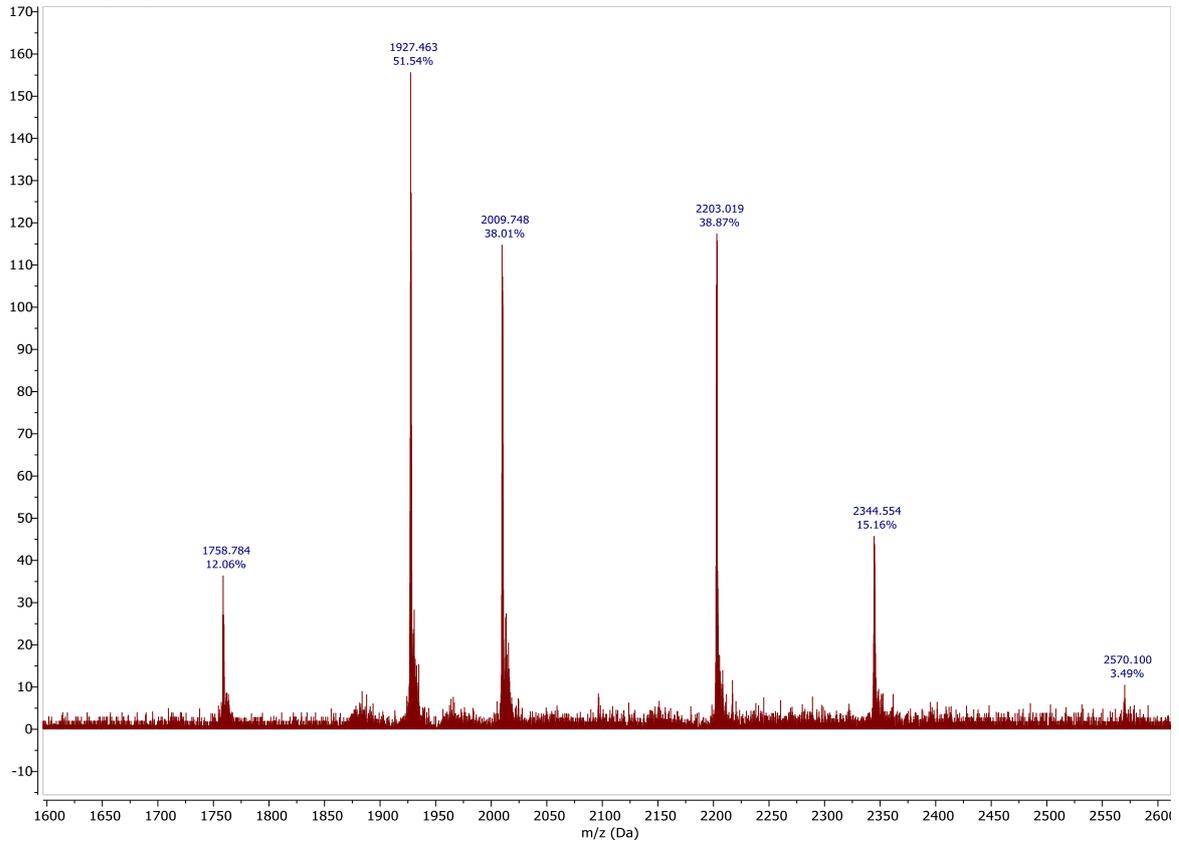


| Peak | Species | Charge state | Predicted |
|---------|-------------|--------------|-----------|
| 1545.79 | LC3C | +11 | 1546.06 |
| 1668.71 | LC3C + LIR1 | +11 | 1668.75 |
| 1700.45 | LC3C | +10 | 1700.57 |
| 1835.09 | LC3C + LIR1 | +10 | 1835.52 |
| 1889.23 | LC3C | +9 | 1889.41 |
| 2039.21 | LC3C + LIR1 | +9 | 2039.36 |
| 2125.12 | LC3C | +8 | 2125.46 |
| 2294.34 | LC3C + LIR1 | +8 | 2294.15 |
| 2428.79 | LC3C | +7 | 2428.96 |
| | | | |
| 1546.09 | LC3C | +11 | 1546.06 |
| 1661.14 | LC3C + LIR2 | +11 | 1660.92 |
| 1700.66 | LC3C | +10 | 1700.57 |
| 1826.82 | LC3C + LIR2 | +10 | 1826.91 |
| 1889.15 | LC3C | +9 | 1889.41 |
| 2029.62 | LC3C + LIR2 | +9 | 2029.79 |
| 2125.43 | LC3C | +8 | 2125.46 |
| 2283.01 | LC3C + LIR2 | +8 | 2283.39 |
| 2428.99 | LC3C | +7 | 2428.96 |

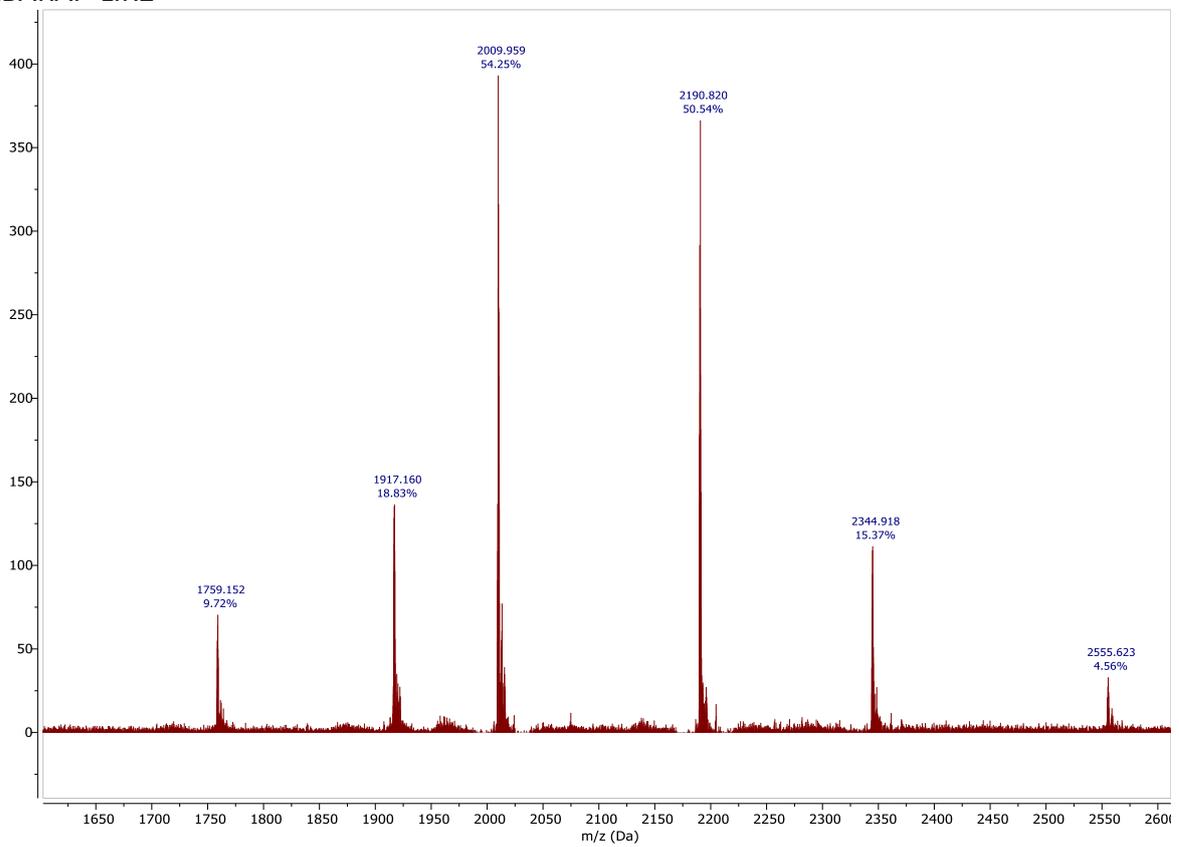
MS spectra for LC3C WT binding with LIRs of NBR1 and table of charge states observed.

C.4 GABARAP

GABARAP-LIR1



GABARAP-LIR2

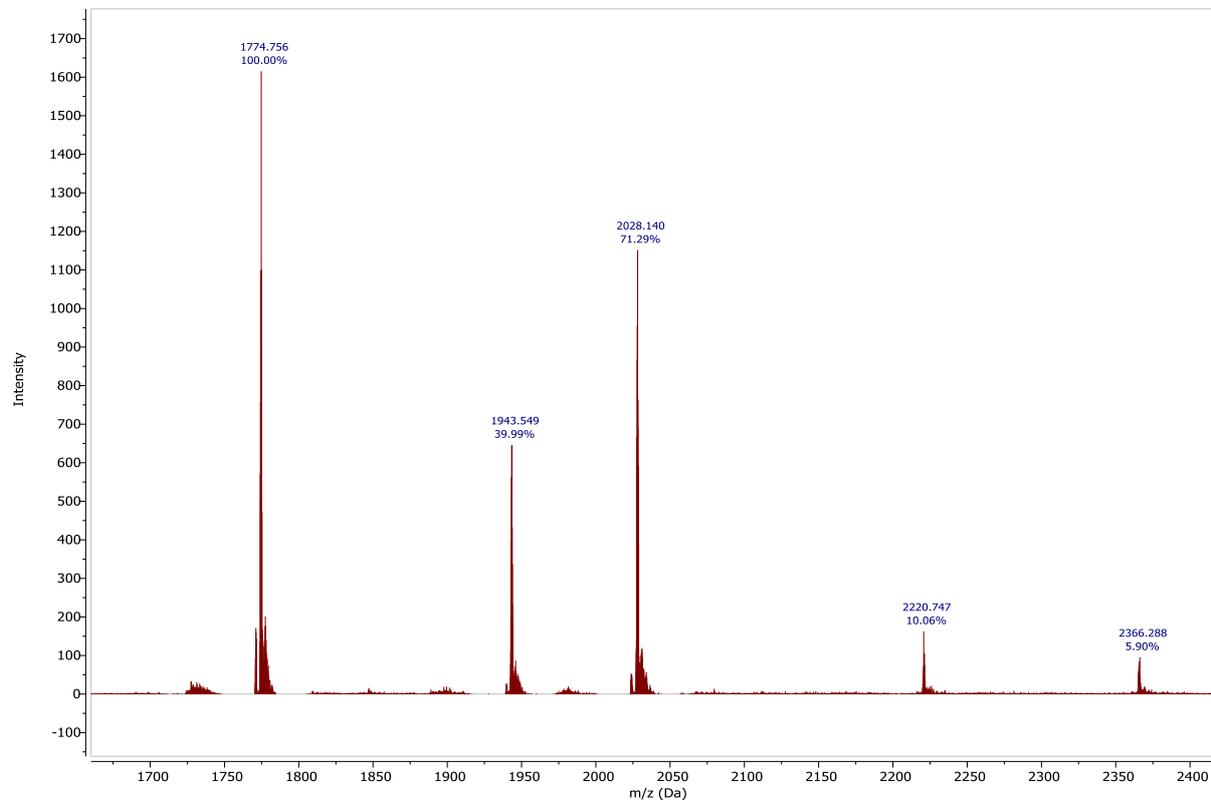


| Peak | Species | Charge state | Predicted |
|---------|----------------|--------------|-----------|
| 1758.78 | GABARAP | +8 | 1758.78 |
| 1927.46 | GABARAP + LIR1 | +8 | 1927.46 |
| 2009.75 | GABARAP | +7 | 2009.86 |
| 2203.02 | GABARAP + LIR1 | +7 | 2202.67 |
| 2344.55 | GABARAP | +6 | 2344.70 |
| 2570.10 | GABARAP + LIR1 | +6 | 2569.62 |
| | | | |
| 1759.15 | GABARAP | +8 | 1758.78 |
| 1917.16 | GABARAP + LIR2 | +8 | 1916.70 |
| 2009.96 | GABARAP | +7 | 2009.86 |
| 2190.82 | GABARAP + LIR2 | +7 | 2190.37 |
| 2344.55 | GABARAP | +6 | 2344.70 |
| 2555.62 | GABARAP + LIR2 | +6 | 2555.27 |

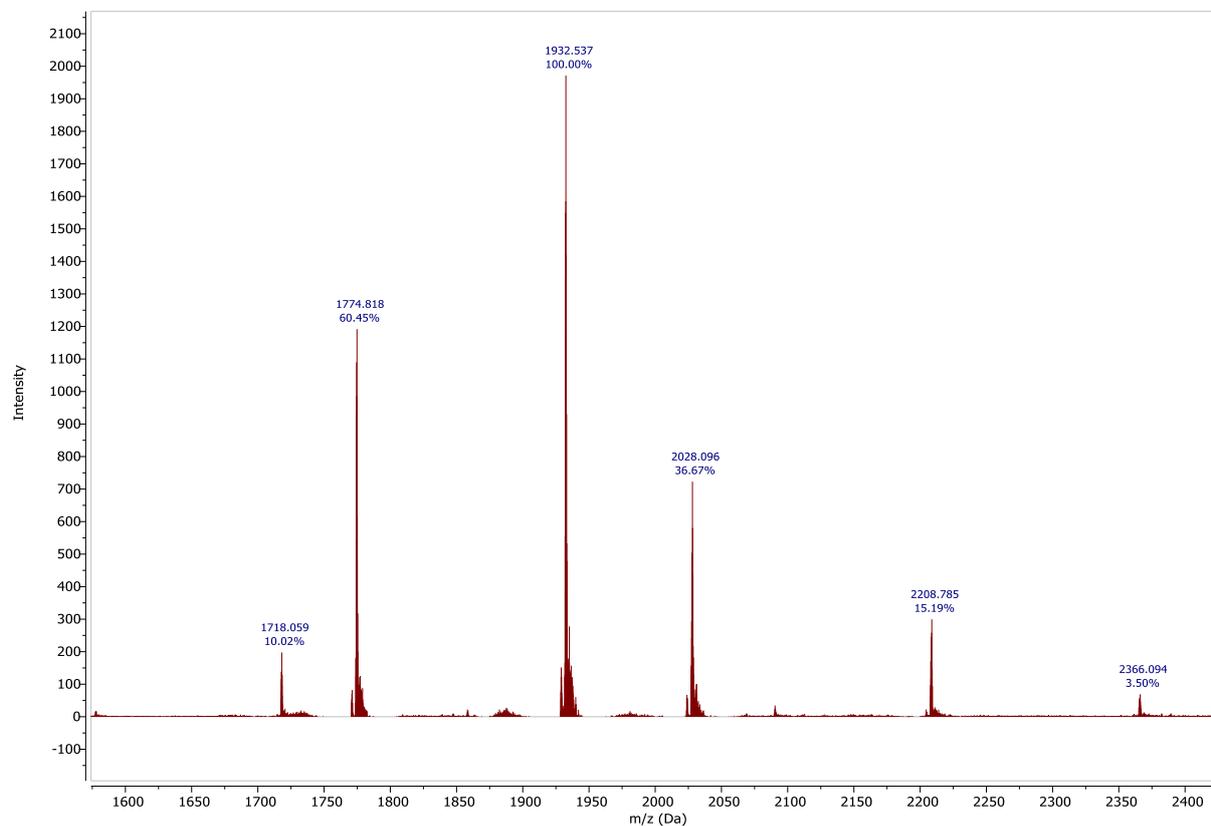
MS spectra for GABARAP WT binding with LIRs of NBR1 and table of charge states observed.

C.5 GABARAPL1

GABARAPL1-LIR1



GABARAPL2-LIR2

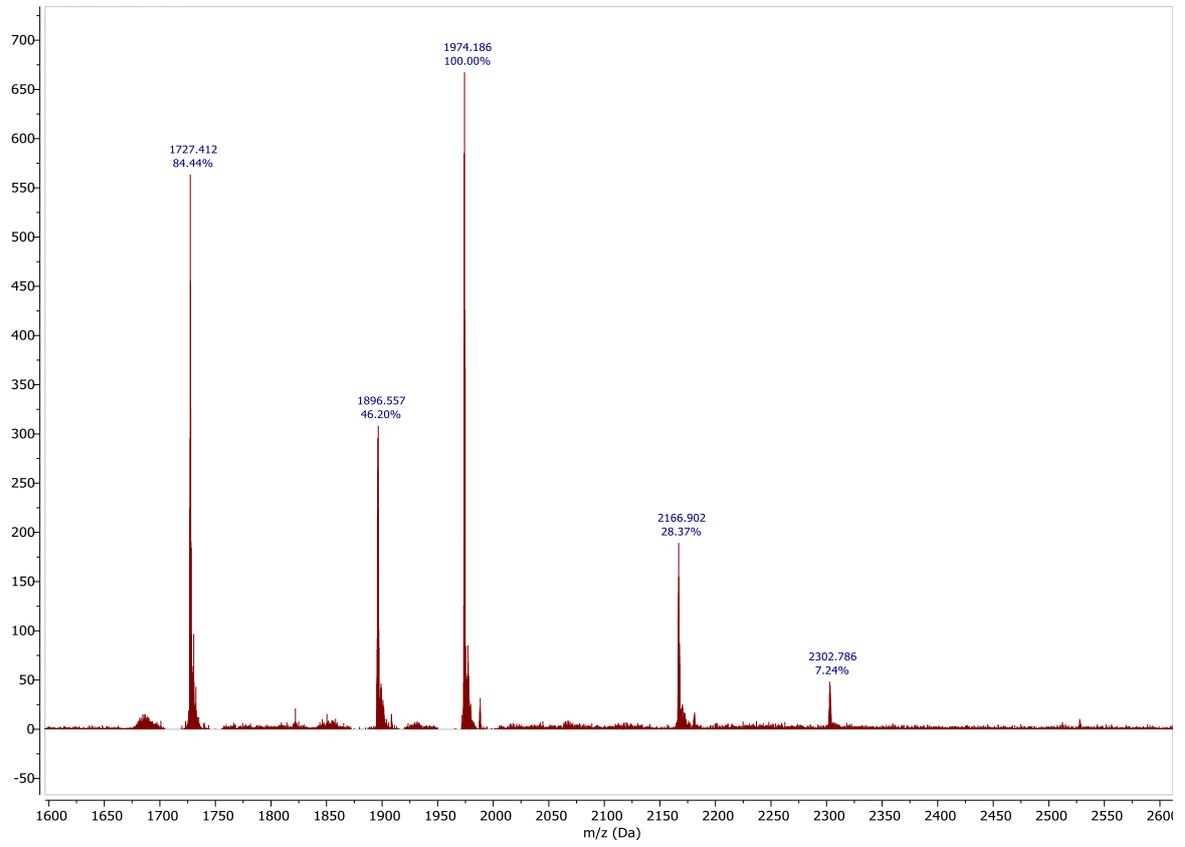


| Peak | Species | Charge state | Predicted |
|---------|------------------|--------------|-----------|
| 1774.76 | GABARAPL1 | +8 | 1774.53 |
| 1943.55 | GABARAPL1 + LIR1 | +8 | 1943.21 |
| 2028.14 | GABARAPL1 | +7 | 2027.89 |
| 2020.75 | GABARAPL1 + LIR1 | +7 | 2020.67 |
| 2366.29 | GABARAPL1 | +6 | 2365.70 |
| | | | |
| 1718.06 | GABARAPL1 + LIR2 | +9 | 1717.84 |
| 1774.82 | GABARAPL1 | +8 | 1774.53 |
| 1932.54 | GABARAPL1 + LIR2 | +8 | 1932.45 |
| 2028.10 | GABARAPL1 | +7 | 2027.89 |
| 2208.79 | GABARAPL1 + LIR2 | +7 | 2208.37 |
| 2366.09 | GABARAPL1 | +6 | 2365.70 |
| | | | |

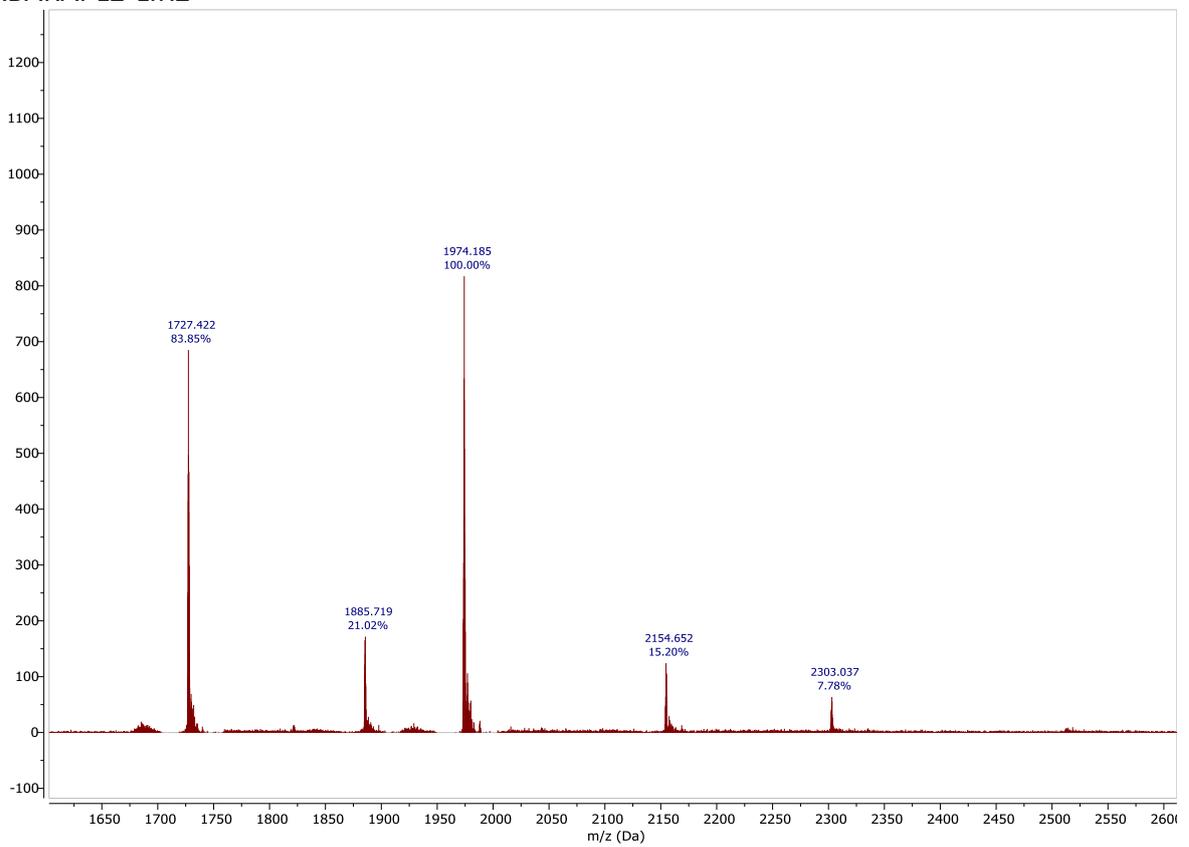
MS spectra for GABARAPL1 WT binding with LIRs of NBR1 and table of charge states observed

C.6 GABARAPL2

GABARAPL2-LIR1



GABARAPL2-LIR2



| Peak | Species | Charge state | Predicted |
|---------|------------------|--------------|-----------|
| 1727.41 | GABARAPL2 | +8 | 1727.38 |
| 1896.56 | GABARAPL2 + LIR1 | +8 | 1896.06 |
| 1974.19 | GABARAPL2 | +7 | 1974.00 |
| 2166.90 | GABARAPL2 + LIR1 | +7 | 2166.79 |
| 2302.79 | GABARAPL2 | +6 | 2302.83 |
| | | | |
| 1727.42 | GABARAPL2 | +8 | 1727.38 |
| 1885.72 | GABARAPL2 + LIR1 | +8 | 1885.30 |
| 1974.19 | GABARAPL2 | +7 | 1974.00 |
| 2154.65 | GABARAPL2 + LIR1 | +7 | 2154.49 |
| 2303.04 | GABARAPL2 | +6 | 2302.83 |
| | | | |

MS spectra for GABARAPL2 WT binding with LIRs of NBR1 and table of charge states observed.