

# Structural characterization and interaction studies of ubiquitin-like proteins

## **Dissertation**

zur Erlangung des Doktorgrades  
der Naturwissenschaften

Vorgelegt dem Fachbereich  
Biochemie, Chemie und Pharmazie (FB 14)  
der Goethe-Universität Frankfurt  
in Frankfurt am Main

von  
Alexis Rozenknop

Frankfurt 2011  
D30

vom Fachbereich Biochemie, Chemie und Pharmazie (FB14) der  
Goethe-Universität Frankfurt als Dissertation angenommen

Dekan: Prof. Dr. Dieter Steinhilber

Gutachter: Prof. Dr. Volker Dötsch  
Prof. Dr. Ivan Dikić

Datum der Disputation: . . . 2011

## Eidesstattliche Versicherung

Ich erkläre hiermit an Eides statt, dass ich die vorgelegte Dissertation über „Structural characterization and interaction studies of ubiquitin-like proteins“ selbständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere, dass aus Schriften Entlehnungen, soweit sie in der Dissertation nicht ausdrücklich als solche mit Angabe der betreffenden Schrift bezeichnet sind, nicht stattgefunden haben.

Alexis Rozenknop

Frankfurt, den . . . 2011



To you

Des chercheurs qui cherchent, on en trouve.  
Des chercheurs qui trouvent, on en cherche...

(Researchers who search, we find them.  
Researchers who find, we search them...)

Charles de Gaulle

## Table of Contents

<b>TABLE OF CONTENTS</b> .....	<b>I</b>
<b>LIST OF FIGURES</b> .....	<b>V</b>
<b>LIST OF TABLES</b> .....	<b>VII</b>
<b>ABBREVIATIONS</b> .....	<b>IX</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>XIII</b>
<b>1. ABSTRACT</b> .....	<b>1</b>
<b>2. ZUSAMMENFASSUNG</b> .....	<b>3</b>
<b>3. INTRODUCTION</b> .....	<b>9</b>
3.1. THE UBIQUITIN SYSTEM .....	9
3.2. UBIQUITIN-LIKE PROTEINS .....	10
3.3. AUTOPHAGY .....	11
3.3.1. <i>Autophagosome</i> .....	11
3.3.2. <i>Ubiquitin-like systems in autophagy</i> .....	13
3.3.3. <i>Autophagy effector proteins, the MAP1LC3 protein family</i> .....	13
3.3.4. <i>Autophagy receptor proteins</i> .....	17
3.3.5. <i>Interaction between autophagy effectors and autophagy receptors</i> .....	20
3.3.6. <i>Autophagy and diseases</i> .....	21
3.4. TBK1 AND A PUTATIVE ULD .....	22
3.5. AN UBIQUITIN-TAG SUITABLE FOR BIOPHYSICAL METHODS .....	23
3.6. GOALS OF THIS STUDY .....	24
<b>4. MATERIALS</b> .....	<b>27</b>
4.1. EQUIPMENTS .....	27
4.2. MOLECULAR BIOLOGY TOOLS .....	28
4.2.1. <i>Bacterial strains</i> .....	28
4.2.2. <i>Vectors</i> .....	29
4.2.2.1. pGEX-4T1 .....	29
4.2.2.2. pETM-60 .....	29
4.2.3. <i>Oligonucleotides</i> .....	29
4.3. BUFFERS, EQUIPMENTS AND SOLUTIONS .....	31
4.3.1. <i>Cloning</i> .....	31
4.3.2. <i>Expression media</i> .....	31
4.3.3. <i>Purification buffers</i> .....	33
4.3.4. <i>NMR buffers</i> .....	33
4.3.5. <i>Gel chromatography buffers</i> .....	34
4.4. SOFTWARE .....	34
<b>5. METHODS</b> .....	<b>35</b>
5.1. PRIMARY SEQUENCE ANALYSIS .....	35
5.2. CLONING .....	35
5.2.1. <i>Substitution of GST to NusA</i> .....	35
5.2.2. <i>Substitution of NusA to ubiquitin</i> .....	35
5.2.3. <i>Mutagenesis</i> .....	36
5.3. ISOLATION OF PLASMID-DNA AND MEASUREMENT OF CONCENTRATION .....	36
5.4. TRANSFORMATION.....	36
5.5. SEQUENCING .....	37
5.6. CELL GROWTH AND EXPRESSION .....	37
5.6.1. <i>Analytical expression</i> .....	37
5.6.2. <i>Preparative expression</i> .....	37
5.6.3. <i>Selective labeling</i> .....	37
5.7. PROTEIN ISOLATION AND PURIFICATION .....	38

## Table of Contents

---

5.7.1.	<i>Cell lysis</i> .....	38
5.7.2.	<i>GST purification</i> .....	38
5.7.3.	<i>NiNTA purification</i> .....	38
5.7.4.	<i>Tag cleavage</i> .....	38
5.7.5.	<i>Cation-exchange chromatography</i> .....	39
5.7.6.	<i>Buffer exchange</i> .....	39
5.7.7.	<i>Size-exclusion chromatography</i> .....	39
5.8.	PEPTIDE PREPARATION .....	39
5.9.	CONCENTRATION SAMPLES .....	40
5.10.	SAMPLE STATE .....	40
5.10.1.	<i>Gel chromatographies</i> .....	40
5.10.2.	<i>Concentration determination</i> .....	40
5.11.	CIRCULAR DICHROISM (CD) SPECTROSCOPY .....	40
5.12.	ISOTHERMAL TITRATION CALORIMETRY (ITC) .....	41
5.13.	NUCLEAR MAGNETIC RESONANCE (NMR) .....	42
5.13.1.	<i>Conditions</i> .....	42
5.13.2.	<i>1D NMR</i> .....	43
5.13.3.	<i>2D NMR: [<sup>15</sup>N, <sup>1</sup>H] TROSY-HSQC and [<sup>13</sup>C, <sup>1</sup>H] HSQC</i> .....	43
5.13.4.	<i>Amide/proton exchange experiment</i> .....	44
5.13.5.	<i>Analysis of cell lysate samples from Ub-constructs</i> .....	44
5.13.6.	<i>Resonance assignment</i> .....	44
5.13.7.	<i>Secondary structure determination</i> .....	45
5.13.8.	<i>Tertiary structure calculation</i> .....	45
5.13.8.1.	TBK1_ULD structure determination .....	46
5.13.8.2.	GABARAPL-1/NBR1-LIR structure determination .....	46
5.13.9.	<i>Interaction studies by NMR</i> .....	47
5.13.9.1.	TBK1_ULD project .....	47
5.13.9.2.	Autophagy project .....	47
5.13.9.3.	<i>Analysis of titration data</i> .....	47
<b>6.</b>	<b>RESULTS</b> .....	<b>49</b>
6.1.	TBK1_ULD PROJECT .....	49
6.1.1.	<i>Expression and purification of TBK1_ULD</i> .....	49
6.1.2.	<i>Secondary structure determination</i> .....	52
6.1.2.1.	Optimization of NMR conditions .....	52
6.1.2.2.	Assignment .....	54
6.1.2.3.	Secondary structure calculation .....	54
6.1.3.	<i>Tertiary structure determination</i> .....	55
6.1.4.	<i>Expression and purification of IRF3</i> .....	57
6.1.5.	<i>Interaction studies of TBK1_ULD/IRF3_IAD-SRR</i> .....	59
6.2.	AUTOPHAGY PROJECT .....	61
6.2.1.	<i>Expression and purification of MAP1LC3 proteins</i> .....	61
6.2.2.	<i>Expression and purification of peptides containing LIR motifs</i> .....	63
6.2.3.	<i>Interaction Atg8/LIR</i> .....	64
6.2.3.1.	LC3B vs. LIRs .....	64
6.2.3.1.1.	LC3B vs. p62-LIR .....	64
6.2.3.1.2.	LC3B vs. Nix-LIRs .....	66
6.2.3.1.3.	LC3B vs. NBR1-LIR .....	69
6.2.3.2.	Other MAP1LC3 proteins vs. LIRs .....	70
6.2.3.2.1.	LC3A vs. Nix-LIRs .....	70
6.2.3.2.2.	GABARAPL-1 vs. Nix-LIR_W36 .....	71
6.2.3.3.	Specificity of NBR1-LIR .....	72
6.2.3.3.1.	GABARAPL-1 vs. NBR1-LIR by NMR .....	72
6.2.3.3.2.	GABARAPL-1 vs. NBR1-LIR by ITC .....	75
6.2.4.	<i>Flexibility of Atg8 proteins</i> .....	77
6.2.4.1.	LC3B .....	77
6.2.4.2.	GABARAPL-1 .....	78
6.2.5.	<i>3D structure of GABARAPL-1/NBR1-LIR</i> .....	78
6.3.	UBIQUITIN PROJECT .....	82



---

6.3.1.	<i>Design of ubiquitin tag constructs</i> .....	82
6.3.2.	<i>Expression and purification of pure proteins/peptides</i> .....	83
6.3.3.	<i>Expression and purification of ubiquitin fused proteins/peptides</i> .....	85
6.3.4.	<i>Use of ubiquitin-fused proteins for CD spectroscopy</i> .....	86
6.3.5.	<i>Use of ubiquitin-fused peptides for ITC</i> .....	86
6.3.6.	<i>Use of ubiquitin-fused proteins/peptides for NMR</i> .....	88
6.3.6.1.	Characterization of ubiquitin-fused proteins/peptides.....	88
6.3.6.2.	Use of ubiquitin-fused constructs from cell lysate .....	90
6.3.6.3.	Interaction studies.....	92
6.3.6.4.	Assignment of peptides using ubiquitin-fused constructs .....	94
6.3.7.	<i>Solubility enhancement</i> .....	95
<b>7.</b>	<b>DISCUSSION</b> .....	<b>97</b>
7.1.	EXPRESSION AND PURIFICATION OF PROTEINS FOR BIOPHYSICAL STUDIES .....	97
7.1.1.	<i>TBK1 and IRF3 proteins</i> .....	97
7.1.2.	<i>MAP1LC3 proteins and LIR domains</i> .....	99
7.1.3.	<i>Ubiquitin as solubility tag suitable for biophysical methods</i> .....	101
7.1.3.1.	Ubiquitin-fused constructs for CD spectroscopy.....	102
7.1.3.2.	Ubiquitin-fused constructs for ITC.....	102
7.1.3.3.	Ubiquitin-fused constructs for NMR.....	103
7.2.	TBK1 .....	105
7.2.1.	<i>TBK1_ULD structure</i> .....	105
7.2.2.	<i>TBK1_ULD vs. IRF3_IAD-SRR</i> .....	106
7.2.3.	<i>TBK1_ULD differs from most common ULDs</i> .....	107
7.3.	AUTOPHAGY .....	108
7.3.1.	<i>MAP1LC3 proteins and LIR peptides</i> .....	108
7.3.2.	<i>Interaction studies</i> .....	108
7.3.3.	<i>Differences between LIR motifs</i> .....	109
7.3.3.1.	p62 .....	109
7.3.3.2.	NBR1 .....	109
7.3.3.3.	Nix .....	111
7.3.4.	<i>Differences between MAP1LC3 proteins</i> .....	113
7.3.4.1.	LC3B vs. LC3A.....	113
7.3.4.2.	LC3- vs. GABARAP- proteins .....	114
7.3.5.	<i>Structure of GABARAPL-1/NBR1-LIR</i> .....	115
7.3.6.	<i>MAP1LC3 proteins are not only involved in autophagy</i> .....	117
7.4.	CONCLUSION AND OUTLOOK .....	119
<b>8.</b>	<b>REFERENCES</b> .....	<b>121</b>
<b>9.</b>	<b>APPENDIX</b> .....	<b>131</b>
9.1.	PROTEIN SEQUENCES .....	131
9.2.	PEPTIDE SEQUENCES .....	131
9.3.	ITC RAW DATA .....	132
9.4.	GABARAPL-1/NBR1-LIR INTERMOLECULAR NOES .....	133
9.5.	NMR STRUCTURAL STATISTICS.....	134
<b>10.</b>	<b>LIST OF PUBLICATIONS</b> .....	<b>137</b>
<b>11.</b>	<b>CURRICULUM VITAE</b> .....	<b>139</b>



## List of Figures

Figure 1: The ubiquitin conjugation system .....	10
Figure 2: Autophagosome formation .....	12
Figure 3: The mammalian Atg8 conjugation system .....	16
Figure 4: Structure of Atg8 proteins .....	17
Figure 5: Autophagy receptor proteins and the LIR motif .....	18
Figure 6: The TBK1 system .....	22
Figure 7: Alignment of the primary sequences of ubiquitin and TBK1_ULD .....	23
Figure 8: Properties of pGEX-4T1 with control regions .....	29
Figure 9: Properties of pETM-60 with control regions .....	29
Figure 10: Analytical expression of GST_TBK1_ULD <sub>s</sub> and GST_TBK1_ULD <sub>l</sub> .....	49
Figure 11: Isolation GST_TBK1_ULD <sub>s</sub> and GST_TBK1_ULD <sub>l</sub> .....	50
Figure 12: Purification of cleaved of GST_TBK1_ULD <sub>s</sub> .....	50
Figure 13: Optimization of buffer conditions for cleaved UL D <sub>s</sub> .....	51
Figure 14: <sup>1</sup> H NMR spectra of TBK1_ULD .....	52
Figure 15: [ <sup>15</sup> N, <sup>1</sup> H] TROSY-HSQC of TBK1_ULD .....	52
Figure 16: [ <sup>1</sup> H, <sup>1</sup> H] NOESY spectrum of TBK1_ULD .....	53
Figure 17: Assignment of TBK1_ULD .....	54
Figure 18: Secondary structure of TBK1_ULD .....	55
Figure 19: NMR Structure of TBK1_ULD .....	56
Figure 20: NMR structure of TBK1_ULD and ubiquitin .....	56
Figure 21: Overlay of the NMR structures of TBK1_ULD and ubiquitin .....	57
Figure 22: Isolation of GST_IAD and GST_IAD-SRR .....	58
Figure 23: Isolation and purification of NusA_IAD-SRR .....	59
Figure 24: Interaction of TBK1_ULD with IRF3_IAD-SRR .....	60
Figure 25: Interaction surface of IAD-SRR on TBK1_ULD .....	60
Figure 26: NMR titration of GST_IAD-SRR against TBK1_ULD .....	61
Figure 27: Expression and purification of GST_MAP1LC3 proteins .....	61
Figure 28: Expression and purification of NusA_LC3B .....	62
Figure 29: Expression and isolation of NBR1-LIR constructs .....	63
Figure 30: LC3B interaction with p62-LIR .....	65
Figure 31: Interaction of LC3B with p62-LIR .....	66
Figure 32: LC3B interaction with Nix-LIRs .....	66
Figure 33: Interaction of LC3B with Nix-LIR_W36 and Nix-LIR_W140/144 .....	68
Figure 34: LC3B interaction with NBR1-LIR .....	69
Figure 35: Interaction of LC3A with Nix-LIRs .....	71
Figure 36: Interaction of GABARAPL-1 with Nix-LIR_W36 .....	71
Figure 37: GABARAPL-1 interaction with NBR1-LIR .....	73
Figure 38: NBR1-LIR interaction with GABARAPL-1 .....	74
Figure 39: GABARAPL-1 interaction with NBR1-LIR_Y732W .....	75
Figure 40: Flexibility of LC3B .....	77
Figure 41: GABARAPL-1/NBR1-LIR interaction proved by intermolecular NOEs .....	79
Figure 42: NMR structure of the GABARAPL-1/NBR1-LIR complex .....	80
Figure 43: LIR binding sites of GABARAPL-1 in presence of the NBR1-LIR domain .....	81
Figure 44: Interaction of the acidic residues of NBR1-LIR with positively charged amino acids of GABARAPL-1 .....	81
Figure 45: Flexibility in the interaction of NBR1-LIR with GABARAPL-1 .....	82
Figure 46: Expression of Ub-fused constructs .....	82

<b>Figure 47: Purification of cleaved Ub_NBR1-LIR</b> .....	84
<b>Figure 48: Purification of Ub_NBR1-LIR</b> .....	85
<b>Figure 49: CD spectroscopy with Ub-fused proteins</b> .....	86
<b>Figure 50: Effect of the ubiquitin moeity on the interaction of Ub_p62 with LC3B</b> .....	87
<b>Figure 51: Effect of the ubiquitin moeity on the interaction of Ub-fused peptides with protein</b> .....	88
<b>Figure 52: NMR spectra of LIR domains fused to ubiquitin</b> .....	89
<b>Figure 53: NMR spectra of Ub-fused proteins</b> .....	90
<b>Figure 58: Stability of Ub_TBK1_ULD</b> .....	95
<b>Figure 54: NMR spectra from the cell lysate of Ub-fused proteins</b> .....	91
<b>Figure 55: LC3B interaction with p62 and Ub_p62</b> .....	92
<b>Figure 56: Interaction studies of Ub_NBR1-LIR with MAP1LC3 proteins</b> .....	93
<b>Figure 57: Assignment of Ub-fused peptides</b> .....	94
<b>Figure 59: Crystal structure of IRF3_IAD-SRR</b> .....	99
<b>Figure 60: Melting curves of ABIN1 fused to ubiquitin</b> .....	102
<b>Figure 61: ITC raw data of the interaction of MAP1LC3 proteins with LIR domains</b> .....	132

---

**List of Tables**

<b>Table 1: Amino acid sequence identity among Atg8 proteins.....</b>	<b>15</b>
<b>Table 2: List of bacterial strains used .....</b>	<b>28</b>
<b>Table 3: Typical NMR experiments used .....</b>	<b>43</b>
<b>Table 4: Thermodynamic parameters obtained by ITC for the interaction of LC3B with different LIR domains.....</b>	<b>64</b>
<b>Table 5: Thermodynamic parameters obtained by ITC for the interactions of GABARAPL-1 with NBR1-LIR wild type and mutants at position 1. ....</b>	<b>76</b>
<b>Table 6: Thermodynamic parameters obtained by ITC for the interactions of GABARAPL-1 with NBR1-LIR wild type and mutants at position -1 and -2. ....</b>	<b>76</b>
<b>Table 7: NUS-NMR experiments for the assignment of Ub_NBR1-LIR.....</b>	<b>104</b>
<b>Table 8: List of peptides used. ....</b>	<b>131</b>
<b>Table 9: Structural statistics of the 20 energy-minimized conformers of TBK1_ULD.....</b>	<b>134</b>
<b>Table 10: Structural statistics of the 20 energy-minimized conformers of GABARAPL-1/NBR1-LIR.....</b>	<b>135</b>



## Abbreviations

A	absorbance
Å	angström
AcA	acrylamide
APS	ammonium peroxydisulfate
Atg	autophagy-related protein
ATP	adenosine triphosphate
BMRB	biological magnetic resonance bank
bp	base pair
C	celsius
cal	calorie
CD	circular dichroism
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio] propanesulfonic acid
CL	cell lysate
CSI	chemical shift index
CSP	chemical shift perturbation
CV	column volume
Da	dalton
DNA	desoxyribonucleic acid
dNTPs	desoxyribonucleic triphosphate
DTT	dithithreitol
<i>E. Coli</i>	<i>Escherischia Coli</i>
EDTA	ethylenediaminetetraacetic acid
El	elution
ER	endoplasmic reticulum
EtBr	ethidium bromide
FT	flow-through
g	gramm
GABARAP	$\gamma$ -aminobutyric acid receptor-associated protein
GABARAPL-1	$\gamma$ -aminobutyric acid receptor-associated protein-like 1
GABARAPL-2	$\gamma$ -aminobutyric acid receptor-associated protein-like 2
GATE-16	Golgi-associated ATPase enhancer of 16kDa
GEC1	glandular epithelial cell protein 1
GFB	gel filtration buffer
GST	glutathione S-transferase
h	hour
His <sub>6</sub>	hexa-histidine tag
hp	hydrophobic pocket
HSQC	heteronuclear single quantum coherence
Hz	hertz
IAD-SRR	interferon associated domain-serine rich region
IEC	ion-exchange chromatography
IMAC	immobilized metal ion affinity chromatography
IKK	I $\kappa$ B kinase
IPTG	isopropyl $\beta$ -D-1-thiogalactopyranoside
IRF3	interferon 3
ITC	isothermal titration calorimetry
k	kilo
K	kelvin

## Abbreviations

---

$K_d$	dissociation constant
L	liter
LB medium	lysogeny broth (Luria-Bertani) medium
LC3-	light chain 3 proteins
LC3A	light chain 3 alpha of microtubule-associated protein 1
LC3B	light chain 3 beta of microtubule-associated protein 1
LIR	LC3- interacting region
m	meter
m	milli
M	mega
M	molar ( $M L^{-1}$ )
M9	minimal media
MAP1LC3	light chain 3 of microtubule-associated protein 1 proteins
MBP	maltose binding protein
min	minute
MW	molecular mass
n	nano
NaCl	sodium chloride
NaP	sodium phosphate
NBR1	neighbor of BRCA1 gene 1 protein
NiNTA	nickel-nitrilotriacetic acid
Nix	Nip-like protein x
NMR	nuclear magnetic resonance
NOESY	nuclear overhauser effect spectroscopy
NSF	<i>N</i> -ethylmaleimide sensitive fusion protein
NUS	non uninominal sampling
OD	optical density
OMM	outer mitochondrial membrane
PAGE	polyacrylamide gel electrophoresis
PAS	preautophagosomal structure
PB1	Phox and Bem1p
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
PE	phosphatidylethanolamine
PI3K	phosphoinositol 3-kinase
ppm	parts per million
RMSD	root mean square deviation
rpm	round per minute
RT	room temperature
S	supernatant
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography
SUMO	small ubiquitin-like modifier
TALOS	torsion angle likeliness obtained from shift & sequence similarity
TAE	Tris acetic acid EDTA buffer
TBK1	TANK-binding kinase 1
TCEP	tris(2-carboxyethyl)phosphine
TEMED	tetramethylethylenediamide
TEV protease	tobacco etch virus protease



TGS	Tris Glycine SDS buffer
TOCSY	total correlation spectroscopy
TRIS	Tris(hydroxymethyl)aminomethane
TROSY	transverse relaxation optimized spectroscopy
U	unit
Ub	ubiquitin
UBA	ubiquitin associated domain
UBD	ubiquitin binding domain
UBL	ubiquitin-like protein
UBX	ubiquitin-regulatory X domain
ULD	ubiquitin-like domain
UV	ultraviolet
V	volt
Zn finger	zinc finger
(v/v)	volume per volume
(w/v)	weight per volume
$\Delta G$	Gibbs free energy
$\Delta H$	change in enthalpy
$\Delta S$	change in entropy
$\epsilon$	extinction coefficient
$\mu$	micro



## Acknowledgements

Even though this work was most of the time a though lonely journey, it was also sometimes smoother with the help of some people.

I am first very thankful to Prof. Volker Dötsch and Prof. Ivan Dikic for giving me the opportunity to work in their labs and providing me the facilities and tools to perform successful research in challenging and interesting projects.

I am grateful to Sigrid Oğuzer-Fachinger, Birgit Lipke, Rebecca Pfeiffer and Milena Klumbis for their help in understanding and going through the german university administration.

I would like to thank the whole group at the Institute of Biochemistry II at the Uni Klinikum for their reception and their critical and challenging questions during group meetings in order to pursue my research in the good directions. I especially thank Dr. Fumiyo Ikeda for the possibility to start my PhD with the successful TBK1\_ULD project. Thanks also to Dr. Vladimir Kirkin who worked on the autophagy project and introduced me to this side project, which became the main part of my thesis and to Dr. Ivana Novak, Dr. David McEwan and Philip Wild, who followed up the project and furnished me with different plasmids.

At the Institute for Biophysical Chemistry at the Campus Riedberg, I would like to first thank Natalja Rogova, Juliana Winkler and Birgit Schäfer for their constant work in the shadow but allowing us to have a great running lab. Natalja also helped me with cloning and provided different constructs essential for my work. Thanks to Dr. Vladimir Rogov, who introduced me in the world of protein purification, for his guidance and suggestions for successful research and finally for his idea about the Ub-tag, which was very useful. Many thanks to Dr. Frank Löhr for providing me enough NMR measuring time and maintaining the NMR facility successful. I am glad for the help of Prof. Peter Güntert in structure calculation but also for his constructive corrections for my paper. Of course, I would like to thank all my actual colleagues as well as the former lab members for the nice working atmosphere in the lab and for the scientific discussions during all these years. Sorry for the non-football fans about our long discussion during lunch... Thanks to Gregor Deutsch for his advices and corrections of my different german/english manuscripts and the time spent in the U3 (and recently U8) and to Elisabeth Zielonka and Laura Luh for their help to correct my "Zusammenfassung". A great thanks to Friederike Junge and Susanne Stefer for supporting me in "our office" although I was often yelling in different languages at Word and its formatting.

## Acknowledgements

---

Out of the lab, I would like to thank the french guys in Frankfurt, who have still no clue about what I was doing in the lab but they were always here to maintain a normal life for me outside of the lab. 143 to the Toulouse butterfly and hold on for the ones who still have to finish their PhD, we will meet again soon.

Finally, I would like to thank my parents for their support and letting me perform long, long, long studies.

Last but not least, thanks to Kate for being by my side during all these years and to keep my motivation high also when nothing was working in the lab. We will make it, Chaka!

## 1. Abstract

Ubiquitin is a highly conserved protein involved in several cellular processes like protein degradation, endocytosis, signal transduction and DNA repair. The discovery of ubiquitin-like proteins (UBL) and ubiquitin-like domains (ULD) increases the number of regulation pathways where the property of the ubiquitin-fold (Ub-fold) is profitable.

Autophagy is the catabolic pathway used in cells to deliver cytosolic components and dysfunctional organelles to the lysosome for degradation. Light chain 3 of microtubule-associated protein 1 (MAP1LC3) proteins are UBL involved in the expansion of the autophagosome, which sequesters cytosolic substrates. Additionally, this protein family, including the light chain 3 proteins (LC3-) and the  $\gamma$ -aminobutyric acid receptor-associated proteins (GABARAP-) subfamilies, bind to autophagy receptors linked to polyubiquitinated proteins aggregates. For this project, the three dimensional structure of a protein complex containing the autophagy effector GABARAPL-1 in the presence of the LC3- interacting region (LIR) of the autophagy receptor neighbor of BRAC1 gene 1 protein (NBR1) was determined by Nuclear Magnetic Resonance (NMR). The results confirmed that  $\gamma$ -aminobutyric acid receptor-associated protein-like 1 (GABARAPL-1) belongs to the MAP1LC3 protein family, structurally characterized by an Ub-fold, consisting of a central  $\beta$ -sheet formed by four  $\beta$ -strands and two  $\alpha$ -helices on one side of the  $\beta$ -sheet, preceded N-terminally by two  $\alpha$ -helices, resulting in the formation of two hydrophobic pockets, hp1 and hp2. The autophagy receptor NBR1 interacts with GABARAPL-1 through these pockets with its LIR motif taking an extended beta conformation upon binding, forming an intermolecular  $\beta$ -sheet with the second  $\beta$ -strand of GABARAPL-1. This LIR motif consists of an  $\Theta xx\Gamma$  sequence preceded by negatively charged amino acids, where  $\Theta$  and  $\Gamma$  are represented by any aromatic and hydrophobic residues, respectively. Interaction studies of the LIR domains of different autophagy receptors (p62, Nix and NBR1) with different members of the MAP1LC3 protein family indicated that the presence of a tryptophan in the LIR motif increases the binding affinity. Substitution to other aromatic amino acids or increasing the number of negatively charged residues at the N-terminus of the LIR motif has, however, little effect on the binding affinity due to enthalpy-entropy compensation, suggesting that autophagy effector proteins can interact with a wide variety of different sequences with similar and moderate binding affinities.

Additionally to be present in proteins dealing with protein folding and degradation, ULD were found in proteins involved in the regulation of signal transduction like TANK binding kinase 1 (TBK1), a serine/threonine kinase involved in the induction of the immune response. In this second project, based on the NMR chemical shifts of the domain of TBK1 including amino acids 302 and 383, secondary structure prediction programs (TALOS and CSI) confirmed the presence of an ULD in TBK1 by identifying one  $\alpha$ -helix and four  $\beta$ -strands sequentially aligned as follows,  $\beta\beta\alpha\beta\beta$ . This alignment corresponds to the secondary structure elements of ubiquitin and proved that TBK1\_ULD belongs to the ubiquitin-like protein superfamily. The similarity to ubiquitin is even bigger by the presence in addition of a small  $\beta$ -strand and a short  $\alpha$ -helix, which are observed in ubiquitin as the  $\beta$ 5-strand and a  $3_{10}$ -helix, respectively. The first attempts on the 3D structure determination confirmed the Ub-fold but due to the lack of assignment in TBK1\_ULD, only a preliminary model was determined. Interaction studies of TBK1\_ULD with the interferon associated domain-serine rich region (IAD-SRR) domain of interferon 3 (IRF3) showed that both sides of the molecule seem to be involved. Consequently, the TBK1/IRF3 interaction is more complex than a one to one binding process. Unfortunately, the instability of TBK1\_ULD associated to difficulties in IAD-SRR purification did not allow to further study this interaction more precisely.

Finally, to overcome the difficulty encountered in NMR experiments because of low expression or poor solubility, an expression vector using the intrinsic properties of ubiquitin was designed. Fused to protein and peptide targets, this construct produced proteins and peptides in a larger amount than with traditional expression vectors. Using this construct, labeled peptides were also produced for NMR structural studies with a less cost than the chemical synthesis of pure labeled peptides. The presence of a hexa-histidine tag was useful for the isolation and the purification of the constructs. A TEV cleavage site was included to keep the possibility of releasing the ubiquitin moiety from the expressed protein or peptide. Moreover, the ubiquitin-tag could also still be attached to the protein/peptide of interest when biophysical methods like NMR, Isothermal Titration Calorimetry (ITC) or Circular Dichroism (CD) spectroscopy are applied, providing the same results than for the protein/peptide moiety alone.

## 2. Zusammenfassung

Die Strukturbiologie von Proteinen ist ein wichtiges Forschungsfeld für Wissenschaftler, die die Funktion eines Proteins auf struktureller Sicht erläutern oder die Funktion in Abhängigkeit von der Struktur verstehen wollen. Die Bestimmung der Struktur eines Proteins, sowie die Art der Bindung an seinen Liganden sind bedeutungsvolle Erkenntnisse um biologische Systeme zu verstehen und neue Effektoren zu entwickeln, die von der pharmazeutischen Industrie verwendet werden könnten. In dieser Arbeit wurde kernmagnetische Resonanz (NMR) verwendet, nicht nur, um die Sekundärstruktur einer Proteindomäne, sondern auch die dreidimensionale Struktur eines Proteins mit seinem Liganden zu bestimmen. Gekoppelt mit Isothermer Titrations Kalorimetrie (ITC) wurde NMR auch zur Charakterisierung von Protein-Peptid-Interaktionen ausgeübt.

Ubiquitin ist ein hoch konserviertes Protein, das in verschiedene zelluläre Prozesse wie Proteinabbau, Endocytose, Signaltransduktion und DNA-Reparatur involviert ist. Die Entdeckung der Ubiquitin-ähnlichen Proteine (UBL) und Ubiquitin-ähnlichen Domänen (ULD) erhöht die Anzahl der Regulierungswege, bei denen Erkenntnisse über das Grundgerüst der Ubiquitin-Faltung profitabel wären.

Zellen verwenden Autophagie als Abbauweg, um cytosolische Komponenten sowie dysfunktionale Organellen den Lysosomen zur Degradation zu liefern. MAP1LC3 Proteine sind UBL Proteine, die bei dem Ausbau eines doppelten Membransacks, des Autophagosoms, beteiligt sind. Dieser trennt die Substrate aus dem Cytosol und fusioniert mit dem Lysozym, wobei Hydrolasen die im Autophagosom enthaltenen Proteine entfalten und lysieren. Des Weiteren binden MAP1LC3 Proteine (LC3- und GABARAP- Unterfamilien) auch Autophagie-Rezeptoren in Verbindung mit polyubiquitinierten Proteineaggregaten. Polyubiquitinierung ist eine prinzipielle posttranslationale Signalmodifikation, die für die Markierung von Proteinen für deren Degradation verwendet wird. Mit der Charakterisierung des ersten Autophagie-Rezeptor p62 wurde ein wesentliches peptidisches Motiv für die Bindung an Autophagie Effektoren identifiziert. Diese LC3-Interaktions-Region (LIR) besteht aus einer WxxL Sequenz mit vorangestellten sauren Aminosäureresten. Mit der Entdeckung weiterer Autophagie-Rezeptoren mit unterschiedlichen LIR Motiven, mußte die Definition des LIR Sequenz überdacht werden. Anstelle von Tryptophan und Leucin werden jeweils aromatische und hydrophobe Aminosäuren benötigt, wobei immer noch die Anwesenheit von negativ geladenen Aminosäuren N-terminal zu diesem Motiv unabkömmlich ist. Die MAP1LC3

Proteinfamilie ist strukturell durch eine Ubiquitin-Faltung gekennzeichnet. Diese besteht aus einem zentralen  $\beta$ -Faltblatt aus vier  $\beta$ -Strängen und zwei  $\alpha$ -Helices auf einer Seite des  $\beta$ -Faltblatt und desweiteren aus zwei N-terminalen  $\alpha$ -Helices. Der Ubiquitin Kern, flankiert von den beiden zusätzlichen  $\alpha$ -Helices, führt zur Bildung von zwei hydrophoben Taschen, hp1 und hp2. Vorherige Studien haben gezeigt, dass hp1 und hp2 an der Bindung des LIR Motivs durch die Interaktion mit der aromatischen und der hydrophoben Aminosäure von LIR beteiligt sind.

Für dieses Projekt wurde die dreidimensionale Struktur des Komplexes GABARAPL-1/NBR1-LIR mittels NMR bestimmt. Die Struktur bestätigte, dass GABARAPL-1 die gleichen strukturellen Eigenschaften wie auch die anderen MAP1LC3 Proteine besitzt. Der autophagische Rezeptor NBR1 interagiert mit GABARAPL-1 mittels hp1 und hp2 durch den Tyrosin- und Isoleucin-Reste des LIR Motivs. Desweiteren besteht eine Interaktion zwischen den Lysin in der N-terminalen  $\alpha$ -Helices von GABARAPL-1 und den Glutaminsäuren und die Asparaginsäuren N-terminal des LIR Motivs. Darüber hinaus nimmt das LIR Motiv eine erweiterte  $\beta$ -Konformation bei der Bindung ein und bildet so ein intermolekulares  $\beta$ -Faltblatt mit dem zweiten  $\beta$ -Strang von GABARAPL-1. Interessanterweise ist der aromatische Rest in NBR1-LIR ein Tyrosin, wohingegen sich es in den meisten der LIR Motiven um ein Tryptophan handelt. Bei einem genaueren Blick in die Struktur des GABARAPL-1/NBR1-LIR Komplexes konnte die Seitenkette von Tyrosin unterschiedliche Positionen in der hp1 Tasche einnehmen, die mit dem unterschiedlichen berechneten strukturellen Konformeren übereinstimmen. Diese mögliche Flexibilität des NBR1-LIR Motivs ist im Einvernehmen mit den Studien zu Wechselwirkungen der LIR Domänen von unterschiedlichen Autophagie-Rezeptoren mit den verschiedenen Mitgliedern der MAP1LC3 Protein Familie. Allerdings erhöht die Anwesenheit von Tryptophan im LIR Motiv die Bindungsaffinität zu MAP1LC3 Proteinen, wie auch in p62, das somit ein besserer Interaktion Partner als NBR1 ist. ITC Experimente bestätigen, dass die Mutation von Tyrosin zu Tryptophan in NBR1-LIR zu einer stärkeren Interaktion im Vergleich zu Wildtyp mit GABARAPL-1 führt. Darüber hinaus zeigten die NMR-Titrations von NBR1-LIR Wildtyp und der Mutant unterschiedliche chemische Verschiebungen, langsam für die Tryptophan-Mutant und intermediär (nah zu langsam) für den Wildtyp, was die stärkere Interaktion in Anwesenheit eines Tryptophan anstelle eines Tyrosins in dem LIR Motiv bestätigt. Obwohl die beiden LIR Motive im Protein Nix einen Tryptophan-Rest besitzen, interagiert dieser autophagische Rezeptor noch schwächer als NBR1 mit MAP1LC3 Proteinen. Folglich ist die alleinige Gegenwart eines Tryptophan als aromatischer Rest nicht ausreichend, um eine starke Bindung zu erreichen. Des weiteren sind die Anwesenheit des hydrophoben Rests sowie



negativ geladener Aminosäuren notwendig. Wahrscheinlich wirkt der erweiterte hydrophobe Patch in NBR1-LIR unterstützend, um die Substitution von Tryptophan durch Tyrosin zu kompensieren. Hinsichtlich der Bedeutung von negativ geladenen Resten im LIR Motiv für die Interaktion mit MAP1LC3 Proteinen könnte die Anwesenheit von mehreren Serinen in den verschiedenen Autophagie-Rezeptor Sequenzen vermuten lassen, dass posttranslationale Modifikationen wie Phosphorylierung die Interaktion verbessern könnten. Im Falle von NBR1-LIR hatte die Erhöhung der Anzahl der negativ geladenen Reste am N-Terminus des LIR Motivs durch Enthalpie-Entropie Kompensation wenig Einfluss auf die Bindungsaffinität. Dieser Effekt wurde auch bei der Substitution von Tyrosin zu einer weiteren aromatischen Aminosäure, Phenylalanin, beobachtet. Diese Ergebnisse implizieren, dass die Autophagie Effektor-Proteine mit einer Vielzahl von verschiedenen Sequenzen mit ähnlichen und moderaten Bindungsaffinitäten interagieren können.

Neben ihrer Funktion in Proteinfaltung und Degradation wurden ULDs in regulatorischen Proteinen der Signaltransduktion gefunden, wie auch in TANK-binding kinase 1 (TBK1), einer Serin/Threonin-Kinase, die an der Induktion der Immunantwort beteiligt ist. NMR wurde verwendet um nach zu weisen, dass die Domäne zwischen den Resten 302 und 383 in TBK1 eine Ubiquitin-ähnliche Domäne ist, wie bereits *in silico* vorhergesagt wurde. Aufgrund Löslichkeits- und Stabilitätsprobleme wurde die Strukturbestimmung von TBK1\_ULD anspruchsvoller als zuvor angenommen. Dennoch konnte die Zuordnung des Peptid-Rückgrats von TBK1\_ULD mit Hilfe selektiv markierter Proben zusätzlich zu den einheitlich markierten Proben durchgeführt werden. Die  $C^\alpha$ ,  $C^\beta$  und  $C'$  chemischen Verschiebungen von TBK1\_ULD waren notwendig, um die sekundäre Struktur dieser Domäne mithilfe von TALOS und CSI Software zu berechnen. Beide Programme bestätigen das Vorhandensein einer Ubiquitin-ähnlichen Domäne in TBK1 durch die Ermittlung einer  $\alpha$ -Helix und vier  $\beta$ -Strängen in der Abfolge  $\beta\beta\alpha\beta\beta$ . Diese Organisation entspricht genau der sekundären Strukturelemente von Ubiquitin und beweist, dass TBK1 zur ULD Protein-Superfamilie gehört. Die Ähnlichkeit mit Ubiquitin wurde noch deutlicher durch die Anwesenheit von einem zusätzlichen kleinen  $\beta$ -Strang und einer kurzen  $\alpha$ -Helix, die dem  $\beta$ 5-Strang und der  $3_{10}$ -Helix in Ubiquitin entsprechen. Die ersten Versuche die zur Bestimmung der 3D-Struktur führen sollten, bestätigen zwar die Ub-Faltung, aufgrund der fehlenden Zuordnung in TBK1\_ULD konnte allerdings eine Struktur nur auf der Basis eines Ubiquitin Modells ermittelt werden. Es wurde bereits gezeigt, dass TBK1\_ULD an die Interferon-assoziierte Domän-serin-reiche Region (IAD-SRR) von Interferon-regulatorischer Faktor 3 (IRF3) bindet und anschließend diese folglich durch die

TBK1\_Kinase phosphoryliert wird, um in den Zellkern transportiert zu werden. Während der NMR Titration von IAD-SRR gegen TBK1\_ULD zeigten beide chemischen Verschiebungen Störungen, welche nach neuen Signalen für Aminosäure Reste aussahen, die sich auf beiden Seiten des Moleküls befinden. Dies lässt darauf schließen, dass die TBK1\_ULD/IRF3\_IAD-SRR Interaktion komplexer als ein „eins zu eins“-Bindungsprozess ist. Leider hat die Instabilität von TBK1\_ULD verbunden mit der Schwierigkeit bei der IAD-SRR Reinigung nicht erlaubt, weitere Studien dieser Interaktion genauer durchzuführen.

Die beiden ersten Projekte haben die Schwierigkeit bei der Bestimmung von Proteinstrukturen durch NMR aufgrund der Notwendigkeit der Herstellung einer hohen Menge an löslichem und gleichzeitig reinem Protein gezeigt. Um dieses Problem einer niedrigen Expression und/oder schlechter Löslichkeit zu bewältigen, wurde ein Expressionsvektor auf Basis von Ubiquitin hergestellt. Durch die inhärenten Eigenschaften von Ubiquitin wurde die Expression und Löslichkeit des fusionierten Proteins oder Peptids bereits weitgehend verbessert. Des weiteren war die Anwesenheit eines Hexahistidin-Tag nützlich für die Isolierung und Reinigung der Konstrukte. Eine TEV-Spaltstelle bietet zudem die Möglichkeit der Abtrennung der Ubiquitin-Einheit von dem restlichen Protein oder Peptid. Dieser Expressionsvektor wurde verwendet um Proteine in einer größeren Menge als bei herkömmlichen Expressionsvektoren zu produzieren. Peptide wie NBR1-LIR (18 Aminosäuren) wurden ebenfalls auf dieser Weise mit geringeren Kosten als bei chemischen Synthesen von Peptiden hergestellt, insbesondere, wenn diese für die NMR Strukturuntersuchungen markiert werden müssen. Außerdem könnte der Ubiquitin-tag mit daran gebundenem Protein/Peptid bei biophysikalische Methoden wie NMR, ITC oder Circular dichroismus (CD) Spektroskopie von Interesse sein. Im Fall der NMR verstärkte die Präsenz des Ubiquitin-Einheit die Stabilität und die Löslichkeit von TBK1\_ULD, welches zuvor zur Aggregation und Präzipitation neigte. Diese führten zu Schwierigkeiten bei der Durchführung von Langzeit-NMR-Experimenten. Aufgrund der Verwendung des Ubiquitin-Vektors konnte eine hohe Ausbeute erzielt werden, wodurch NMR-Spektren für die Ubiquitin fusionierten Protein- oder Peptidproben aus Zelllysate aufgenommen werden konnten. Tatsächlich zeigte ein so aufgenommenes NMR Spektrum in Zelllysate das gleiche Signal-Muster und nahe zu eine gleich gute Qualität wie Spektren von gereinigten Protein. Gereinigte, Ub-fusionierte Peptide wurden für NMR-Experimente zur Titration verwendet und zeigten die gleichen Signale, jedoch mit kleinen Unterschieden zu dem synthetisierten Peptid. Diese Unterschiede basieren auf leichten Veränderungen im Puffer. Letztendlich könnte die Produktion von Ub-fusionierten Peptide für Non-Uniform Sampling (NUS) NMR mit einer starken Verminderung der benötigten Spektroskopiezeit verwendet werden. Gekoppelt an das automatisierte

Zuordnungsprotokoll wurden Ub-fusionierten Peptide leichter, schneller und mit dem gleichen Ergebnis von unseren Kooperationspartnern zugeordnet. Wie schon für NMR wurden fusionierte Ub-Peptide auch für ITC Experimente verwendet, um die thermodynamischen Parameter der Interaktion von MAP1LC3 Proteinen mit LIR Peptide zu bestimmen. Die gleichen Werte ( $K_d$ ,  $\Delta H$ ,  $\Delta S$  und  $\Delta G$ ) wurden sowohl für p62 als auch für Ub\_p62 gegen LC3B erhalten. Die Kontrollexperimente mit Ubiquitin, titriert gegen das Zielprotein, zeigten, dass Ubiquitin kaum Einfluss auf die Bindung hat. Folglich wurden die Studien zu Wechselwirkungen von GABARAPL-1 mit NBR1-LIR sowie von NBR1-LIR\_Y732W mit NBR1-LIR fusioniert an Ubiquitin durchgeführt. Hierbei wurden die Unterschiede durch die Substitution von Tyrosin durch Tryptophan durch ITC- und NMR- Experimente untersucht. Schließlich wurden bei der CD-Spektroskopie ebenfalls Ub-Fusionsproteine eingesetzt. Am Beispiel der GABARAPL-1 wurden CD-Kurven von Ub-GABARAPL-1, GABARAPL-1 allein und Ub allein aufgenommen: die erhaltene Kurve für Ub\_GABARAPL-1 ist die Überlagerung der beiden Kurven aus GABARAPL-1 und Ubiquitin allein. Dies weist darauf hin, dass die Anwesenheit von Ubiquitin in dem fusionierten Konstrukt keine weiteren sekundären Strukturelemente als die bereits im Ubiquitin enthaltenen in das Zielprotein einfügt.

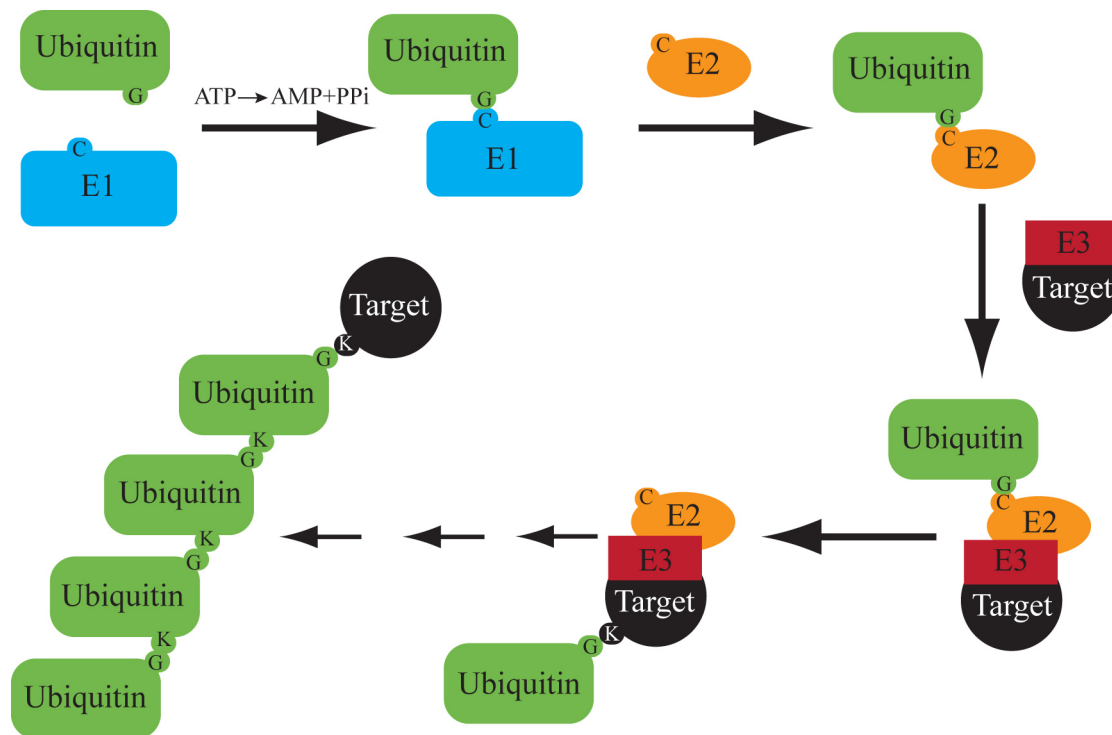


### 3. Introduction

In his *Elementary Treatise on Chemistry*<sup>1</sup> published in 1789, the French chemist Lavoisier declared that nothing is created, nothing is destroyed, everything is only exchanged. Applied to living systems, this definition means that cells are not creating proteins, organelles and other components *ab nihilo* but use already existing nutrients to build them. In the case of protein metabolism, the pool of amino acids could be reused to build new proteins after degradation. Thus, protein recycling pathways are essential for cells to survive, to develop and to adapt to the different stimuli affecting them. Autophagy and the ubiquitin proteasome are two main catabolic pathways used in cells. Whereas the ubiquitin proteasome pathway, which is deeply involved in development and apoptosis by regulating the degradation of short-lived proteins, has been extensively reviewed,<sup>2</sup> the studies on autophagy since the last decade raised a lot of new questions. Surprisingly, these both pathways have more in common than it was first suspected.

#### 3.1. The ubiquitin system

Their successful studies on the ubiquitin proteasome pathway provided the Nobel Prize in Chemistry to Aaron Ciechanover, Avram Hershko and Irwin Rose in 2004.<sup>3-5</sup> Proteins, after being tagged by a chain of several ubiquitin, are brought to the proteasome to be degraded for the regulation of their cell concentrations or because of misfolding. Ubiquitin is a 76 amino acid protein highly conserved from yeast to human and polyubiquitination occurs as a three steps conjugation system.<sup>2; 6; 7</sup> First, ubiquitin is activated under action of ATP by the activating enzyme E1 with the formation of an intermediate ubiquitin adenylate, which binds through its C-terminal carbonyl group to the active site cysteine of E1 through a thioester bond. Second, ubiquitin is transferred to the active site cysteine of an ubiquitin-carrier protein E2 during the conjugation process. Third, ubiquitin is ligated through its C-terminal glycine to the  $\epsilon$ -amino group of a lysine on the target protein via an amide bond under the action of an ubiquitin ligase E3 (**Figure 1**). By reproducing this process, a new ubiquitin molecule could bind to the previous one through a lysine to create a polyubiquitin chain. Usually the polyubiquitin chain is formed by linkage on the lysine at position 48 of each ubiquitin and the tagged protein is then degraded by the proteasome complex. Different linkages on the ubiquitin are also possible leading to other functions like endocytosis, signal transduction and DNA repair.<sup>8-12</sup>



**Figure 1: The ubiquitin conjugation system.** The ubiquitination of target protein is a three-step enzymatic reaction with first an activation of ubiquitin via an E1 enzyme, then a transfer through an E2 enzyme and finally an E3 enzyme makes the ligation of ubiquitin to the target protein.

### 3.2. Ubiquitin-like proteins

As suggested in their names, ubiquitin-like proteins (UBL) share a high similarity to ubiquitin.<sup>13</sup> Although not presenting obligatory a high sequence homology, UBL present the same structure than ubiquitin, the  $\beta$ -grasp fold.<sup>14</sup>

UBL of type I, commonly named UBL, are conjugated to different targets and act as signal messengers controlling many cellular functions like cell proliferation, apoptosis, proteolysis, endocytosis, transcription, DNA repair, signal transduction and autophagy.<sup>12; 15</sup> Several human diseases, including many types of tumors, are linked to the misregulation of UBL or proteins involved in their conjugation.<sup>16</sup>

UBL of type II, also named ubiquitin-like domain (ULD), define proteins possessing a domain sharing the same 3D structure than ubiquitin whereas their primary sequences do not have obligatory a high similarity. At least, the hydrophobic patch around isoleucine at position 44 in ubiquitin is conserved between the different ULD.<sup>17</sup> This domain could be also found along other domains of a protein. In opposition to ubiquitin and UBL, ULD are neither processed nor linked to other proteins to act as a post-translational modifier.<sup>18</sup> The ubiquitin-fold (Ub-fold) found in various proteins has a significant importance because it represents a common docking

site for protein-protein interaction.<sup>19</sup> UBL and ULD are present in diverse proteins involved in the cellular machinery or in the folding and degradation of proteins as well as in proteins regulating enzymatic activity and signal transduction.<sup>18; 20</sup>

### **3.3. Autophagy**

Recent studies showed a possible link between ubiquitination and autophagy.<sup>21</sup> Autophagy is a highly conserved catabolic pathway used in cells to conduct cytosolic components to the lysosome for degradation.<sup>22</sup> The term autophagy describes three types of processes:

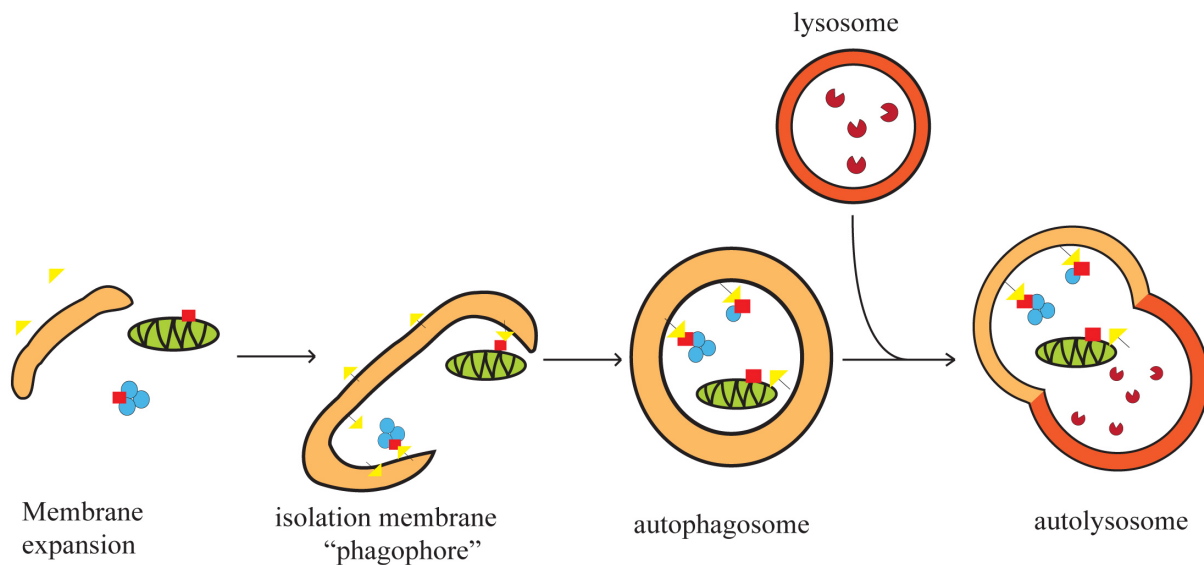
- macroautophagy defined by the formation of a double-membrane sack, the autophagosome, that is capable of fusing with the lysosome,<sup>23</sup>
- microautophagy where the lysosome itself engulf cytosolic components,<sup>24</sup>
- chaperone-mediated autophagy with proteins translocated through the membrane of the lysosomes.<sup>25</sup>

Macroautophagy, being the main autophagy pathway, will be described hereafter under the generic term of autophagy. Autophagy was first considered to occur as a response against cell starvation<sup>26</sup> but this self-digestion process for the turnover in cells is also involved in the removal of damaged organelles,<sup>27</sup> the degradation of long-lived proteins<sup>28</sup> and the containment of infectious agents or the antigen presentation.<sup>29</sup>

#### ***3.3.1. Autophagosome***

To avoid that cytosolic components are randomly engulfed during autophagy, a key step for selective autophagy is the recognition and the isolation of substrates, which entails surrounding of the substrates with the autophagosome that is capable of fusing with the lysosome.<sup>30; 31</sup> Under action of hydrolases, the different components brought from the autophagosomes to the lysosomes are degraded until single elements are again available for the cell needs (**Figure 2**).<sup>32</sup> The origin of the autophagosomes is still subject to controversy. In yeast, already formed vesicles have been localized in the cytoplasm of cell forming the phagophore assembly site (PAS). At this site, a core of machinery proteins has been identified to be coupled to the PAS,<sup>33</sup> which matures from a phagophore to an autophagosome.<sup>34</sup> Atg9 is a transmembrane protein present as well in the Golgi apparatus, in late endosomes as in the PAS and initiates the formation of the phagophore.<sup>35</sup> Atg9 is brought to the PAS by Atg23<sup>36</sup> and Atg27.<sup>37</sup> Atg9 is then released from the PAS by the Atg1 kinase complex associated to Atg2 and Atg18.<sup>36</sup> The expansion of the phagophore is mediated by Atg8 coupled to phosphatidylethanolamine (PE) and

by Atg12 modified with the Atg5/Atg16 complex.<sup>38</sup> Whereas Atg12 is only present in the outer membrane of the phagophore,<sup>39</sup> Atg8 conjugated to PE is found in the inner and the outer membrane.<sup>40</sup>



**Figure 2: Autophagosome formation.** The expansion of membrane in the cytosol results in the formation of a double membrane sack, the autophagosome, which entails its surrounding and fuses with a lysosome where hydrolases degrade its content.

In mammals, no PAS has been identified but an isolation membrane is induced to grow and to expand under the action of ULK1, homolog of Atg1, coupled to the phosphatidylinositol 3-kinase (PI3K) complex and to mAtg9, homolog of yeast Atg9.<sup>41</sup> The further expansion and closing of the autophagosomes depends of the mammalian homologs of the Atg12/Atg5/Atg16 complex and of Atg8 represented by the Atg16L complex and by the light chain 3 of microtubule-associated protein 1 (MAP1LC3) proteins, respectively.<sup>42</sup>

However, for yeast and mammals, the question about the origin of the autophagosome membrane is still debated between four hypotheses:<sup>43</sup>

- creation of the membrane *de novo* using lipids delivery,
- vesicular transport,
- a cisternal assembly with the fusion of vesicles, in a sleeping state in the cell until induction,
- a membrane remodeling and extension from preexisting structures.

For this last theory, the endoplasmic reticulum (ER) was first supposed to be at the origin of the autophagosome membrane<sup>44</sup> as it was shown that PI3P, deeply involved in the formation of autophagosomes, was also connected to ER.<sup>45</sup> In the meantime, Atg5 and Atg8/MAP1LC3



proteins, key markers of autophagosomes, are present in the membrane of the ER but also of mitochondria<sup>46</sup> and PE, essential for the insertion of Atg8 proteins in membrane, is mainly produced by mitochondria.<sup>47</sup> These two observations develop the theory of a mitochondrial origin of the autophagosomal membrane. The starting point of the autophagosomes is still unclear but it seems that the different structures enunciated previously are involved in their creation but at different time point and under different cell conditions.

### ***3.3.2. Ubiquitin-like systems in autophagy***

Based on genetic studies on *Saccharomyces cerevisiae*, 31 ATG genes have been identified with more than the half involved in autophagosome formation.<sup>48</sup> The Atg proteins already introduced above could be functionally distinguished in five groups:

- the Atg1 kinase complex,
- the phosphatidylinositol 3-kinase complex,
- the integral membrane protein Atg9 and the Atg2-Atg18 complex,
- the Atg12 conjugation system,
- the Atg8 conjugation system.<sup>49</sup>

A major role for the Atg8 conjugation system is its implication in the formation and in the maturation of the autophagosome.<sup>50; 51</sup> The Atg8 conjugation system is highly interesting because Atg8 is an UBL that follows the same conjugation procedure than ubiquitin. Unlike ubiquitin, which has a glycine in C-terminal position, Atg8 exhibits one only after cleavage of the arginine last residue by Atg4.<sup>52</sup> The presence of a glycine as last residue for Atg8 is essential for the following processes in the Atg8 conjugation pathway. Atg8 is first activated by one E1-like protein, Atg7.<sup>53</sup> Atg8 is then transferred to one E2-like protein, Atg3.<sup>52</sup> Finally, Atg8 is conjugated to PE through the free C-terminus of the glycine residue by an amine bond. PE can be later released from Atg8 under the cleaving action of Atg4.<sup>51</sup> No E3-like enzyme has been described so far for the conjugation step but knowing the fact that the Atg12/Atg5/Atg16 complex induces Atg8-PE formation, it is likely that these proteins act like an E3-like enzyme in the Atg8 conjugation system.<sup>54</sup> Conjugated to PE, Atg8 binds to the membrane of the autophagosome, already during its preformation but also all along the expansion process. Atg8 is thus used as an autophagosomal marker.<sup>40</sup>

### ***3.3.3. Autophagy effector proteins, the MAP1LC3 protein family***

In the eukaryote domain, the ATG genes are highly conserved in plants<sup>55</sup> and in mammals.<sup>42; 56</sup> Whereas only one Atg8 protein has been found in yeast, several mammalian

homologs of Atg8 have already been identified, which form the MAP1LC3 protein family consisting of two subgroups, the light chain 3 proteins (LC3-) and the  $\gamma$ -aminobutyric acid receptor-associated proteins (GABARAP-) subfamily. These proteins are also named autophagy effectors.

The first mammalian homolog of Atg8 was discovered in rat and defined as the light chain 3 of the microtubule-associated protein, LC3.<sup>57</sup> Depending on its conjugation state, two denominations are used for LC3. LC3-I corresponds to a native integral form present in the cytoplasm of cells and LC3-II is the PE conjugated form associated to membranes.<sup>58</sup> The identification of three human homologs of the rat LC3 leads to the denomination of the LC3-subfamily constituted by LC3A, LC3B and LC3C.<sup>59</sup> These proteins share a high sequence homology with the rat LC3 protein with 82%, 96% and 55% identity, respectively. Whereas LC3A and LC3C expose a glycine as last residue after their C-terminal proteolytic cleavage, no cleavage is observed for LC3B, which uses a lysine, also located in the C-terminal region, for further conjugation.<sup>59</sup> Human LC3- proteins, after following different post-translational modifications and being conjugated to PE, are also tightly bound to the membrane of the autophagosome.<sup>59</sup>

The GABARAP- subfamily consists of GABARAP, GABARAPL-1 (also known as glandular epithelial cell protein 1 (GEC1)), GABARAPL-2 (previously Golgi-associated ATPase enhancer of 16 kDa (GATE-16)) and GABARAPL-3. GABARAP, localized in all tissues, has 31% identity in the primary sequence with the rat LC3 protein, indicating the possibility of a homologous functions for this protein. Xin et al.<sup>60</sup> identified three paralogs of GABARAP, ubiquitously expressed. As for LC3- proteins, GABARAP- proteins are also converted in a form II conjugated to PE and localized into the membrane of autophagosomes.<sup>61; 62</sup>

The primary sequences of Atg8 and its mammalian homologs LC3A, LC3B, LC3C GABARAP, GABARAPL-1 and GABARAPL-2 were analyzed with ClustalW.<sup>63</sup> Atg8 shares on average 55% and 35% sequence identity with GABARAP- and LC3- subfamily, respectively. Into the subfamilies, the sequence identity is always above 55% but is below 40% compared to each other (**Table 1**). All these proteins have similar size (117 to 147 amino acids).

GABARAP, GABARAPL-2 and LC3B are the most studied proteins in the human MAP1LC3 protein family. Due to their high sequence homology, the same general function is expected for these proteins. However, different specificities for LC3- and GABARAP- proteins are hypothetical depending on time action and space localization.<sup>42; 64</sup> Further studies have shown different functions in the autophagosome formation for LC3- and GABARAP- proteins,

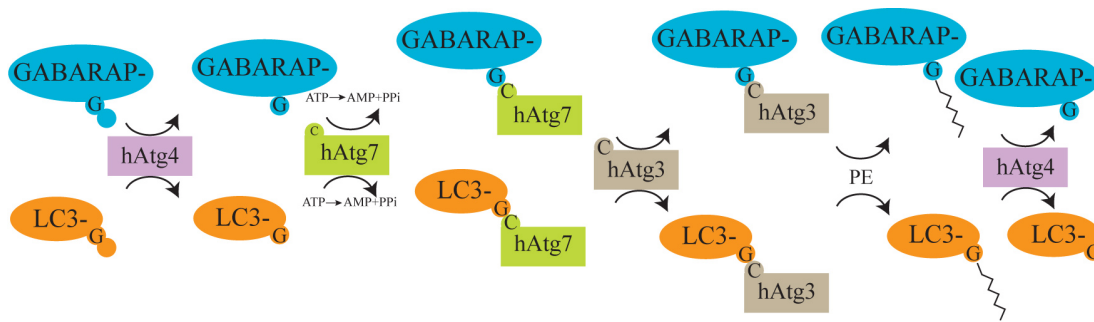
these proteins being involved in the elongation of the phagophore membrane and in the maturation of the autophagosome,<sup>64</sup> respectively.

Sequence 1	Size (aa)	Sequence 2	Size (aa)	Identity Score (%)
Atg8	117	LC3A	121	35
Atg8	117	LC3B	125	34
Atg8	117	LC3C	147	38
Atg8	117	GABARAP	117	54
Atg8	117	GABARAPL-1	117	54
Atg8	117	GABARAPL-2	117	55
Atg8	117	GABARAPL-3	117	52
LC3A	121	LC3B	125	81
LC3A	121	LC3C	117	58
LC3A	121	GABARAP	147	29
LC3A	121	GABARAPL-1	117	33
LC3A	121	GABARAPL-2	117	40
LC3A	121	GABARAPL-3	117	31
LC3B	125	LC3C	147	52
LC3B	125	GABARAP	117	30
LC3B	125	GABARAPL-1	117	31
LC3B	125	GABARAPL-2	117	37
LC3B	125	GABARAPL-3	117	29
LC3C	147	GABARAP	117	37
LC3C	147	GABARAPL-1	117	38
LC3C	147	GABARAPL-2	117	41
LC3C	147	GABARAPL-3	117	36
GABARAP	117	GABARAPL-1	117	86
GABARAP	117	GABARAPL-2	117	57
GABARAP	117	GABARAPL-3	117	82
GABARAPL-1	117	GABARAPL-2	117	60
GABARAPL-1	117	GABARAPL-3	117	93
GABARAPL-2	117	GABARAPL-3	117	57

**Table 1: Amino acid sequence identity among Atg8 proteins.**

The multiple alignment of protein sequences was performed by ClustalW2. All sequences are from human proteins except Atg8 from *Saccharomyces cerevisiae*.

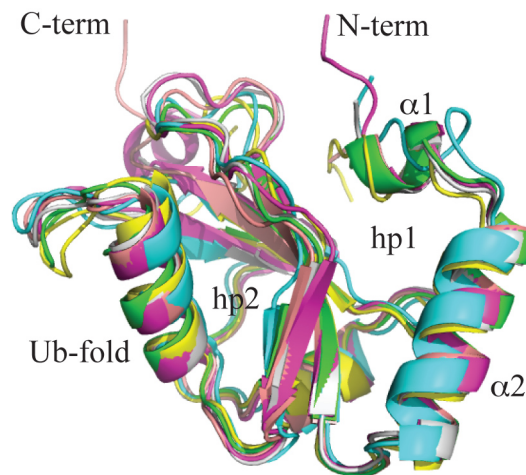
The general mechanism proposed for the modification of the mammalian homologs of Atg8 is to be first processed by the cleavage of the C-terminus and the exposition of a glycine as last residue under the action of autophagins like Apg4b, homolog of yeast Atg4.<sup>65</sup> MAP1LC3 proteins are then activated by the mammalian homolog of Atg7 acting as E1-like protein and transferred to an E2-like protein, the mammalian homolog of Atg3, which conjugates MAP1LC3 proteins to PE.<sup>58; 66; 67</sup> MAP1LC3 proteins could be later cleaved again by the autophagin in order to be released from the membrane and this cycle could be then repeated (**Figure 3**).<sup>68</sup>



**Figure 3: The mammalian Atg8 conjugation system.** GABARAP- and LC3- proteins are first cleaved at the C-terminus by hAtg4 to expose a glycine residue, then activated by hAtg7 (E1-like), transferred to hAtg3 (E2-like) and finally ligated to lipids like PE.

Using X-Ray crystallography as well as nuclear magnetic resonance (NMR), the three dimensional structure of Atg8,<sup>69; 70</sup> LC3,<sup>71; 72</sup> GABARAP<sup>73-75</sup> and GABARAPL-2<sup>76</sup> has already been determined (**Figure 4**). All MAP1LC3 proteins share an Ub-fold motif,<sup>77</sup> which is N-terminally extended by two  $\alpha$ -helices. The Ub-fold consists of a central  $\beta$ -sheet formed by four  $\beta$ -strands and two  $\alpha$ -helices on one side of the  $\beta$ -sheet. The  $\beta$ -strands in the middle are parallel between them but antiparallel to the two other outer  $\beta$ -strands. The special characteristic of MAP1LC3 proteins is the presence of two  $\alpha$ -helices at the N-terminus of the protein, specific of this protein family. Interestingly these helices have different ionic charges along the proteins of the family. Indeed, whereas  $\alpha$ -helix 1 is basic and  $\alpha$ -helix 2 acidic for LC3- proteins,  $\alpha$ -helix 1 is acidic and  $\alpha$ -helix 2 is basic for GABARAP- proteins.<sup>71</sup> Nevertheless, most of the residues presenting ionic charges in these helices are conserved between the different homologs like K08, E14, R16, D/E19 and R24 on LC3B sequence (UniProt accession number Q9GZQ8). These residues are involved in stabilizing the protein structure by keeping these  $\alpha$ -helices close to the Ub-fold through hydrogen bonds between the conserved residues E14 and E36, but also by salt bridges between the conserved residues R16 and D106 and between D19 and K51.<sup>78</sup> One important point found in the case of GABARAP is the presence of two conformations in one crystal structure: a closed state with the  $\alpha$ -helices at the N-terminus along the Ub-core and an open state where  $\alpha$ -helix 1 is rotated by 180°.<sup>74</sup> In the open conformation,  $\alpha$ -helix 1 of GABARAP interacts with the C-terminus of another GABARAP molecule, leading to oligomerization in a head to tail manner. Until now, only GABARAP showed two conformations in X-ray crystallography and NMR depending on the conditions. LC3 and Atg8 showed instead only slight changes in the conformation of their N-terminal regions but after lipidation.<sup>79</sup> The structure of the other Atg8 homologs corresponds to GABARAP closed conformation, which will be mentioned as the common fold of Atg8 proteins hereafter. Thus, the overall structure

consists of a central  $\beta$ -sheet surrounded by two  $\alpha$ -helices resulting in the formation of two hydrophobic pockets, hp1 and hp2. Using the nomenclature of LC3B, hp1 is situated between  $\alpha$ -helix 1,  $\alpha$ -helix 2 and the  $\beta$ -sheet and is formed by the residues D19, I21, P30, I32, K51, L53 and F108 and hp2 is situated between  $\alpha$ -helix 3 and the  $\beta$ -sheet and is formed by the residues F52, V54, P55, V58, L63, I66 and I67.<sup>80</sup> These hydrophobic pockets are formed due to the presence of the  $\alpha$ -helices specific of Atg8 proteins and are constituted by conserved amino acids along proteins of the family.



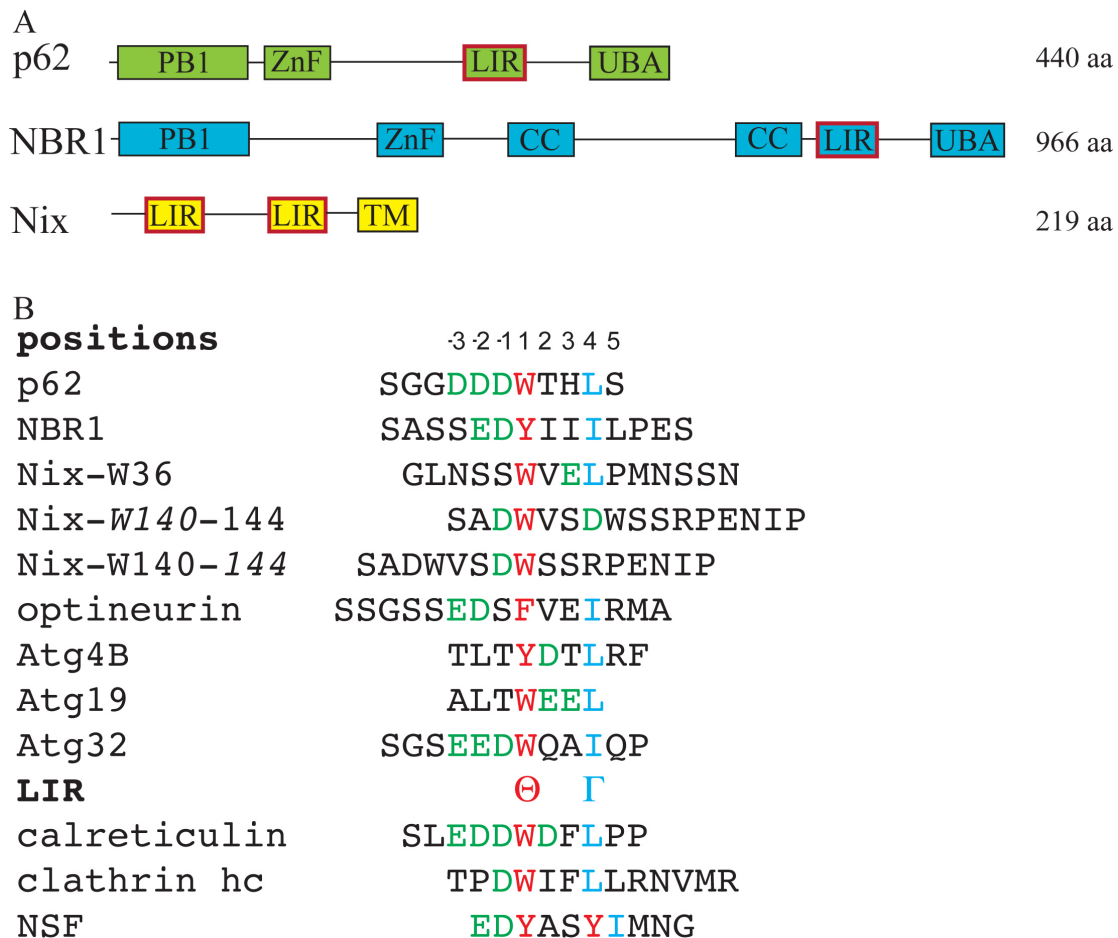
**Figure 4: Structure of Atg8 proteins.** Overlay of the structure of Atg8 (cyan), LC3A (gray), LC3B in presence of p62 (magenta), GABARAP (yellow), GABARAPL-1 (green) and GABARAPL-2 (pink). Each protein possesses an Ub-fold preceded by two N-terminal  $\alpha$ -helices, resulting in the formation of two hydrophobic pockets, hp1 and hp2. The similarity between these structures is reflected by an overall RMSD of 1.5 Å.

### 3.3.4. Autophagy receptor proteins

To regulate the degradation of protein aggregates, autophagy is also involved. Autophagy receptor proteins bind to a polyubiquitin chain linked to a protein, polyubiquitination being a common modification for misfolded proteins.<sup>12; 81</sup> These receptors are also localized into autophagosomes after interacting with autophagy effectors.<sup>82</sup>

The p62 protein is an adaptor protein capable of binding to a broad range of proteins through different domains. p62 is thus involved in several pathways of cell signaling. The protein contains a N-terminal Phox and Bem1p (PB1) domain responsible for its oligomerization and capable also to bind to the atypical protein kinase C (aPKC).<sup>83; 84</sup> Beside a zinc (Zn) finger domain, p62 possesses a C-terminal ubiquitin-associated domain (UBA), which binds to ubiquitin or polyubiquitin signals (**Figure 5**).<sup>85</sup> Thus, one function of p62 is to act as transporter for ubiquitinated cargo. p62 is co-localized with ubiquitin inclusion bodies and found in LC3 positive structures like autophagosomes, linking the ubiquitinated substrates that should be

degraded to the autophagy effector proteins in the membrane of the autophagosomes or, more generally, linking protein aggregation to autophagy.<sup>86</sup>



**Figure 5: Autophagy receptor proteins and the LIR motif.** The domain organization of autophagy receptors is shown in A. Autophagy receptors like p62 or NBR1 bind to polyubiquitinated cargo via their UBA domain whereas Nix is inserted into the membrane of mitochondria through its TM domain. In addition, p62 and NBR1 possess a PB1 domain responsible for their oligomerization but also the polymerization of p62. The polymerization of NBR1 takes place through its coil-coil domain. The LIR sequence, involved in the interaction with autophagy effector proteins, is the common motif between the autophagy receptors. The amino acid sequence of different LIR motifs are represented in B. The LIR sequences present in autophagy receptors and enzymes modifying MAP1LC3 proteins as well as non-autophagy proteins are represented above and below the LIR motif, respectively. Aromatic residues corresponding to position 1 are marked in red, hydrophobic residues at position 4 in blue and negatively charged amino acids in green. All sequences are human except for Atg19 and Atg32, which are from *Saccharomyces cerevisiae*.

Moreover, p62 possesses a DDDWTHLS sequence motif that enables p62 to interact with MAP1LC3 proteins and therefore is called a LC3- interacting region (LIR).<sup>87; 88</sup> Originally, LIR sequences were defined as a WxxL motif. Further studies on proteins, including autophagy receptors, interacting with MAP1LC3 proteins showed that neither the tryptophan nor the leucine residues are strictly conserved. A more general definition of the LIR sequence should be xx⊖xxΓ where ⊖ and Γ are aromatic and hydrophobic residues, respectively, with one or more acidic residues being required as well.<sup>89</sup> Hereafter, the annotation for the position of the residues

in the LIR motif is used as follows: 1 for the aromatic residue  $\Theta$  and 4 for the hydrophobic residue  $\Gamma$ ; -1, -2 and -3 for residues located at the N-terminus of the aromatic residue  $\Theta$ ; 2 and 3 for the residues between the aromatic residue  $\Theta$  and the hydrophobic residue  $\Gamma$ ; 5 for residues located at the C-terminus of the hydrophobic residue  $\Gamma$  (**Figure 5**).

Because knocked-out of p62 only reduces autophagy but does not inhibit it totally,<sup>90; 91</sup> other autophagy receptor proteins are also involved in autophagy.

Neighbor of breast cancer 1 gene 1 (NBR1) was initially characterized as potential gene of the ovarian cancer antigen CA125<sup>92</sup> but was also identified as an autophagy receptor that shares similar features with p62.<sup>93; 94</sup> Although NBR1 is more than twice as big as p62, the protein presents the same profile with the presence of a PB1 domain, a Zn finger, a UBA domain and a LIR domain. Unlike in p62, the PB1 domain of NBR1 does not lead into the oligomerization of NBR1, which takes place through the coil-coil domains situated between the Zn finger and the LIR domain.<sup>95</sup> This PB1 domain is involved in several interaction of whom p62 (**Figure 5**).<sup>83</sup> Like p62, NBR1 can simultaneously bind to the ubiquitin cargo and to MAP1LC3 proteins.<sup>93; 94</sup> Although the deletion of the LIR domain NBR1 as well as failure of autophagy lead to an accumulation of NBR1 in cells, NBR1 knocked-out cells are still presenting an autophagic activity even though reduced in regard to normal cells.<sup>93</sup> A major difference to p62 involves the LIR domain. While most LIR sequences show a tryptophan at position 1 and a leucine at position 4, the LIR domain identified in NBR1 consists instead of a tyrosine and an isoleucine, respectively (**Figure 5**). These residues could be essential for the selectivity of the interaction with LC3 proteins.

Mitophagy is a special aspect of selective autophagy involving mitochondria. Damaged mitochondria for toxicity reasons and “normal” mitochondria depending of the cell turnover are degraded in a similar way than misfolded protein.<sup>96</sup> The characterization of mitochondrial proteins acting as autophagy receptors explains the mechanism of mitophagy.<sup>97-99</sup> In yeast, Atg32 is an outer mitochondrial membrane (OMM) protein overexpressed in damaged mitochondria.<sup>97; 98</sup> Atg32 binds directly to Atg8 present in the PAS<sup>40</sup> or indirectly through Atg11.<sup>100</sup> This protein complex targets mitochondria into autophagosomes for degradation after fusion to lysosome. An interesting characteristic of Atg32 is the presence of a LIR domain represented by a EEDWQAI motif necessary for direct binding to Atg8 (**Figure 5**).<sup>98</sup> In mammals, although there is no homology with Atg32, the NIP3-like protein X (Nix) is also an OMM protein presenting the same function as Atg32<sup>101; 102</sup> but possessing two LIR motifs (**Figure 5**). The first one is situated in the N-terminal region of the protein and its SSWVEL

sequence corresponds to the common LIR motif. Situated close to the BH3-like domain, the ADWVSDWSSR sequence of the second LIR motif consists of the juxtaposition of two LIR motifs. Although the sequence of the second LIR motif in Nix shows differences regarding the common LIR motif, it binds to all MAP1LC3 proteins (**Figure 5**).<sup>99</sup> Both LIR motifs allow the binding of Nix to MAP1LC3 proteins but mutations in the first LIR abolish totally the interaction whereas mutations in the second LIR only decrease this interaction.<sup>99</sup> During erythropoiesis, mitochondria colocalized with MAP1LC3 proteins in autophagosomes via the overexpressed Nix protein. This removal of mitochondria is essential for the development of red blood cells.<sup>99</sup>

### ***3.3.5. Interaction between autophagy effectors and autophagy receptors***

Binding of tryptophan-containing LIR domains to LC3- and GABARAP- proteins is structurally and functionally well characterized.<sup>87; 88</sup> In the structure of p62 in complex with LC3, the  $\Theta$  and  $\Gamma$  residues of p62-LIR interact with the hydrophobic pockets hp1 and hp2 on the surface of the autophagy effector.<sup>87; 103</sup> The first hydrophobic pocket hp1, situated between the Ub-like fold core of MAP1LC3 proteins and the two N-terminal  $\alpha$ -helices, interacts with the aromatic  $\Theta$  residue. The hydrophobic  $\Gamma$  residue binds to the hydrophobic pocket hp2 located on the surface of the Ub-core domain. The LIR motif takes an extended beta conformation upon binding, forming an intermolecular  $\beta$ -sheet with the second  $\beta$ -strand of LC3.

The consensus LIR motif WxxL undergoes different modifications upon the characterization of new binding partners to Atg8 and MAP1LC3 proteins, independently of their functions in autophagy or not. The  $\Theta$  residue is more commonly a tryptophan but identification of NBR1 and Atg4b shows that it could be a tyrosine<sup>93; 104</sup> as well as a phenylalanine in the case of optineurin.<sup>105</sup> The  $\Gamma$  residue is preferentially a leucine but an isoleucine is present in the LIR motif of NBR1 or calreticulin and it has been suggested that valine could also fit at this position.<sup>93; 106</sup> Finally, there is a lot of freedom for the residues at position -3, -2, -1, 2 and 3 but at least one negative charged residue, if not more, is always found at one of these positions (**Figure 5**).<sup>89</sup> No structure of a complex or detailed interaction study of autophagy effector proteins with non-tryptophan LIR domains has been reported so far. The NBR1-LIR (YIILL) has a much more hydrophobic nature in comparison to p62-LIR (WTHL), although they share the same functional features. This increase of hydrophobicity as well as other modifications in the motif could influence the interaction with autophagy effector proteins and might even lead to different or multiple orientations of the peptide sequence in the binding site.

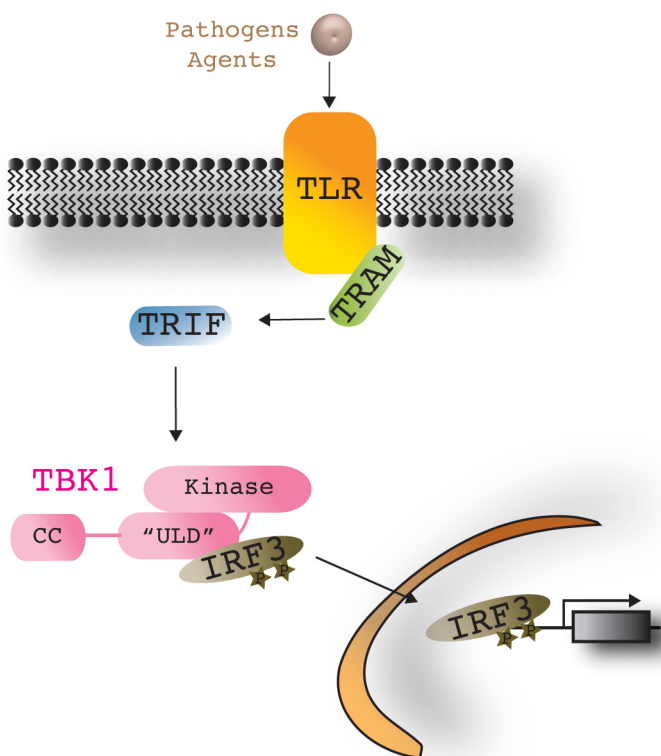


### 3.3.6. *Autophagy and diseases*

By removing damaged organelles or misfolded proteins, presenting antigen and containing infection, autophagy preserves the integrity of the cells by avoiding the accumulation of toxic components.<sup>107</sup> Moreover, autophagy allows cells to survive by adapting their conditions to different stimuli like starving conditions. Regarding the number of proteins involved in the different steps necessary for autophagy, it is easily understandable that dysfunctions like absence of proteins or the incapability of binding to interaction partners have tremendous consequences. Dysregulation of autophagy pathways provides less tolerance to starvation conditions leading to several diseases.<sup>107</sup> Mice studies have shown that organelle aggregation and accumulation of ubiquitinated protein aggregates in different tissues due to defection of autophagy are the first steps in cell malignance before apoptosis, neurodegeneration, tumor formation or cell death.<sup>108</sup> Several human diseases like inflammatory diseases, liver injury, diabetes, neurodegenerative disorders and cancer are caused by mutations in proteins involved in autophagy.<sup>108</sup> The following description of diseases involving proteins studied in this thesis is not exhaustive but shows the severe consequences faced in autophagy dysfunctions. The Paget disease, causing an abnormal and deforming bones growth,<sup>109</sup> is observed in patients presenting mutations in the UBA domain of p62 leading to deficiencies in the clearance of ubiquitinated substrates.<sup>110</sup> Alzheimer's and Parkinson's diseases are characterized by cytoplasmic inclusions of toxic proteins in the brain. p62 colocalizes with these inclusions bodies, being not degraded by autophagy.<sup>111</sup> Due to their functional similarities, p62 and NBR1 are involved in similar diseases. Mallory bodies are inclusions found in liver cells of patients presenting alcoholic liver diseases and are formed due to the presence of ubiquitinated substrates non-cleared by the p62/NBR1 cargo.<sup>93</sup> Failures in mitophagy result in tumorigenesis and abnormal cell death that are observed in Parkinson's disease.<sup>112</sup> Knockout of or mutation in the Nix gene lead to tumor growth whereas the overexpression of Nix protects cells from the expansion of tumor cells.<sup>113-115</sup> Moreover, Nix being directly involved in the erythroid maturation by clearance of mitochondria, anemia is observed in Nix deficient cells.<sup>116</sup> Until now, no disease has been directly correlated with mutations of MAP1LC3 proteins. Due to the high sequence and structure similarity of the different members of the family, it is likely that in case of mutations on one of the proteins, another protein of the family could substitute its function without the cells being affected as it was shown after GABARAP knock-out.<sup>60; 117</sup> Nevertheless, multiple knock-out of MAP1LC3 proteins blocks autophagosome formation. Any disruption in the conjugation of MAP1LC3 proteins to lipids or in the interaction with LIR domains leads to an accumulation of protein aggregates having the same consequences than describe above.

### 3.4. TBK1 and a putative ULD

TANK binding kinase 1 (TBK1) is an I $\kappa$ B kinase related kinase (IKK-related kinase) protein involved in the innate immune response. IKK-related kinases have a crucial role in the innate immune response by regulating the activity of the interferon regulatory factors and the NF- $\kappa$ B transcription factors.<sup>118</sup> Upon infections by bacteria and viruses, Toll-like receptors present on the membrane of macrophages, of fibroblasts as well as of dendritic, B- and T- cells activate TBK1 through the adaptor proteins TRAM and TRIF.<sup>119</sup> Bioinformatical analysis on the amino acid sequence of TBK1 showed that the protein could possess a ULD between residues 305 and 385.<sup>120</sup> Another feature of TBK1 is the presence of a kinase domain responsible for the phosphorylation of the interferon regulatory factor 3 (IRF3). In its inactive conformation, the kinase and the potential ULD of TBK1 are distant to each other and can not interact. In the active form of TBK1, these two domains can interact. After activation, TBK1 recruits IRF3 upon the interaction of the expected ULD and the interferon associated domain (IAD) of IRF3.<sup>120</sup> The kinase domain of TBK1 phosphorylates IRF3 on its C-terminal region rich in serines and threonines (SRR).<sup>121</sup> As an inactive form, IRF3 is located in the cytoplasm as a monomer. Upon phosphorylation, IRF3 dimerizes and translocates then to the nucleus, where IFN genes are induced after interaction on the promoter region and engage the immune response to the bacterial and viral infections (**Figure 6**).<sup>122</sup> The putative ULD domain of TBK1 is thus essential for the activity of the protein.



**Figure 6: The TBK1 system.** The activation of TBK1 is induced by LPS or poly (I:C) and TBK1 interacts then through its ULD domain with the IAD domain of IRF3. Upon its phosphorylation via the kinase domain of TBK1, IRF3 translocates into the nucleus and binds to the promoter region of IFN-inducible genes in order to start the autoimmune response.

The sequence similarities given by the sequence alignment software ClustalW between ubiquitin and TBK1<sub>305-385</sub> is low with only 9% of sequence identity. Nevertheless, the hydrophobic patch around isoleucine at position 44 (leucine-isoleucine-phenylalanine) of ubiquitin is conserved in TBK1, centered on isoleucine at position 353 (leucine-isoleucine-tyrosine) (**Figure 7**). This hydrophobic patch is indispensable for the intramolecular interaction of the ULD and the kinase domain of TBK1 but it is not necessary for the intermolecular interaction of TBK1\_ULD with IRF3.<sup>120</sup>

```

          ↓
TBK1_ULD  TSDVLHRMVIHVFSLQHMHTAHKIYIHSYNTAAVPHLVYKQTKIVSSNQELIYEGRRVLVLELGRLAQHFPKTEENPIFVTS----
Ubiquitin  -----MQIFVKLTGKTIT---LEVEPSDTIENVKAKIQDKGIPPDQORLIFAGKQ--LEDGRTLSDYNIQKESTLHLVLRIRGG
          : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
          : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  
```

**Figure 7: Alignment of the primary sequences of ubiquitin and TBK1\_ULD .** TBK1\_ULD and ubiquitin present a amino acid sequence identity of 9% Identical, conserved and semi-conserved amino acids are represented by (\*), (:), and (.), respectively. The conserved isoleucine at position 44 in ubiquitin and at position 353 in TBK1 is signaled with a bold character and an arrow.

### 3.5. An ubiquitin-tag suitable for biophysical methods

NMR provides information on the structure and the dynamics of a protein,<sup>123</sup> the site(s) of interaction with binding partner<sup>124</sup> and to quantify this interaction.<sup>125</sup> Limiting parameters in the study of proteins by NMR have to be considered like:

- the size: up to 30 kDa to avoid overcrowded spectra and broadening peaks due to fast relaxation,
- the concentration: more than 3 mg of <sup>15</sup>N or <sup>15</sup>N/<sup>13</sup>C isotopic labeled protein for a 15 kDa protein due to the low sensitivity of the technique,
- the purity: to avoid the presence of extra peaks due to contaminants,
- the stability: several days are needed to record one 3D-nuclear Overhauser enhancement spectroscopy (NOESY) experiment of the studied sample.

The first step is to optimize the best conditions of the sample to be suitable for further studies by NMR.<sup>126</sup> First have to be established the expression conditions to produce a sufficient protein yield and the purification methods, which are different for each protein. Once the protein sample is concentrated and does not seem to show any aggregation, the next step is to record a [<sup>15</sup>N, <sup>1</sup>H] transverse relaxation optimized spectroscopy (TROSY) - hetero single quantum correlation (HSQC) experiment.<sup>127</sup> A good dispersion and a small line broadening of the peaks corresponding to each peptide bond of the protein are good signs for a “good behaving protein” for NMR. Further NMR experiments for 3D structure calculation or interactions studies could be performed. To face the different problems occurring in proteins structural biology, diverse expression, solubility and purification tags like Glutathione S-Transferase (GST),<sup>128</sup> NusA,<sup>129</sup>

small ubiquitin-related modifier (SUMO),<sup>130</sup> maltose-binding protein (MBP),<sup>128; 131</sup> hexahistidine (His<sub>6</sub>)<sup>131</sup> tags have been developed to increase the amount of protein produced. Nevertheless, removal of these moieties is often needed for biophysical studies necessitating further purification steps leading to a loss of material. Thus the advantages gained due to favorable expression and solubility is not obvious anymore. Using the high expression and the great solubility properties of ubiquitin, proteins and peptides could take advantage of being fused to ubiquitin for an enhanced expression.<sup>132-134</sup>

### 3.6. Goals of this study

Autophagy is the catabolic pathway used in cells to deliver cytosolic components and dysfunctional organelles to the lysosome for degradation after being engulfed in the autophagosome. Selective autophagy is possible due to the presence of autophagy effectors, also named MAP1LC3 proteins by mammalian, which bind to the membrane of the autophagosome and of autophagy receptors, which interact with polyubiquitinated substrates and with autophagy effectors through a LIR motif. Binding of tryptophan-containing LIR domains to the MAP1LC3 proteins is structurally and functionally well characterized.<sup>87; 88; 94; 103; 104; 135-137</sup> However, only a model of a complex of NSF<sup>138</sup> with GABARAP but no structure of a complex or a detailed interaction study of autophagy effector proteins with non-tryptophan LIR domains from autophagy receptors have been reported. The NBR1-LIR domain (YIIL) has a more hydrophobic nature than the p62-LIR domain (WTHLS). This increased hydrophobicity could influence the interaction with MAP1LC3 effector proteins and might even lead to different or multiple orientations of the peptide in the binding site. To address these questions, the interaction of the NBR1-LIR domain with mammalian Atg8 proteins was studied and compared to the LIR domains of p62 and of Nix, which was identified as a mammalian mitophagy receptor. Mutants of NBR1-LIR with tryptophan or phenylalanine in place of tyrosine were generated in order to understand the importance of position 1 for the binding mode. To answer a similar question, negatively charged amino acids located at the N-terminus of the core LIR domain were additionally introduced. The results show a tighter interaction between Atg8 family proteins and LIRs in the presence of tryptophan as well as the importance of the surrounding residues. The NMR structure of the GABARAPL-1–NBR1-LIR complex was determined and is the first example of the structure of a complex involving a non-tryptophan autophagy receptor–LIR domain. The NMR structure confirmed that NBR1-LIR and GABARAPL-1 form intermolecular  $\beta$ -strands and that NBR1-LIR occupied both hydrophobic pockets of GABARAPL-1.

TBK1 is deeply involved in the regulation of interferon inducible genes for the immune response against bacterial and viral infections. Bioinformatics analysis determined the presence of an ULD in TBK1. Moreover, an intact ULD in TBK1 is essential because this domain was shown to interact with the kinase domain of TBK1 as well as with its substrate IRF3.<sup>120</sup> To confirm the ULD in TBK1, the determination of a structured Ub-fold characterized by the presence of the following secondary structure elements  $\beta\beta\alpha\beta\beta$  has to be done. The same sequential arrangement and the same length as the corresponding structure elements in ubiquitin are a special feature of UBL and ULD. In addition, the structural characterization of TBK1\_ULD interaction with IRF3 and with its kinase domain could better explain the mechanism of TBK1 activation.

The scientist community is always looking forward for the development of new biotechnological tools to produce efficiently soluble proteins and peptides for biochemical studies. Different expression vectors are already available and could provide enough material for biophysical studies. Nevertheless, several purification steps including cleavage of co-expressed tag lead often to a decrease of the final available amount of protein. The use of new expression vector including a modified ubiquitin could increase the production of proteins and peptides, which benefit from the presence of ubiquitin for a better expression in a soluble form. Moreover, due to the relative small size of ubiquitin, release of this expression tag should not be obligatory to use biophysical methods for the studies of proteins and peptides. The features of this new expression tool could provide a great advantage for the researchers in the biochemistry field.



## 4. Materials

### 4.1. Equipments

Äkta purifier FPLC system	GE Healthcare (Germany)
Agarose gel running chamber	Peqlab (Germany)
Autoclave	Tecnomara AG (Switzerland), Gettinge AB (Sweden)
Balances	Sartorius AG (Germany)
Circular Dichroism Spectrometer	Jasco Inc. (USA)
Centrifuges:	
Biofuge RS 28	Haereus Sepatech AG (Germany)
Centrifuge Sorvall RC 28S	Sorvall Instruments (Germany)
Centrifuge Sorvall RC 5B	Sorvall Instruments (Germany)
Refrigerated centrifuge 5417R	Eppendorf (Germany)
Columns:	
GST, Superdex 75 16/60,	GE Healthcare (Germany)
Q-sepharose, SP-sepharose,	
NiNTA, HiTrap Desalting 16/60,	
HiTrap SP FF	
Electrophoresis system:	
electrophoresis apparatus	Bio-Rad GmbH (Germany)
Power supply	Bio-Rad GmbH (Germany)
Freeze Dryer (Sublimator VaCo 10)	Zirbus GmbH (Germany)
French Pressure Cell Press	SLM Instruments (USA)
Incubator	Memmert (Germany)
ITC (VP-ITC)	MicroCal Inc (USA)
Nanodrop 1000	PeqLab (Germany)
NMR spectrometers (599 MHz, 600 MHz, 700 MHz, 800 MHz, 900 MHz)	Bruker Avance (Germany)
PCR cycler	Biometra AG and Eppendorf (Germany)
pH-meter PHM 210	Radiometer (Denmark)
Pipettes and Tips	Eppendorf (Germany)

Rotors (GSA, GS3, ss-34, F28/50)	Sorvall Instruments (Germany)
Shakers:	
Incubating shakers	Infors AG (Germany), New Brunswick Scientific Inc. (USA)
Table shaker	Heidolph Electro GmbH (Germany)
Spectrophotometers	
Cary 3 UV	Varian (Australia)
Hitachi U-1100	RD Analytik (Germany)
Sonifier Labsonic	B. Braun Biotech GmbH (Germany)
Vortex mixer	IKA (Stauffen, Germany)

## 4.2. Molecular biology tools

### 4.2.1. Bacterial strains

Strains	Description	Reference
E. Coli DH5 $\alpha$	[F-,endA1, hsdR17 (rk-mk-), supE44, thi-1, recA1, gyrA (Nalr), relA1, $\Delta$ (lacZYA-argF)U169, $\phi$ 80lacZ $\Delta$ M15]	Hanahan 1983 Woodcock 1989
E. Coli BL21 (DE3)	fhuA2 [lon] ompT gal ( $\lambda$ DE3) [dcm] $\Delta$ hsdS $\lambda$ DE3 = $\lambda$ sBamHIo $\Delta$ EcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 $\Delta$ in5	New England Biolabs
E. Coli NEB T7 Express	fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::miniTn10--Tet <sup>S</sup> )2 [dcm] R(zgb-210::Tn10--Tet <sup>S</sup> ) endA1 $\Delta$ (mcrCmrr)114::IS10	New England Biolabs
E.Coli DL39	ilvE12 tyrB507 aspC13 rph-1 fnr-25 LAM-	LeMaster 1988

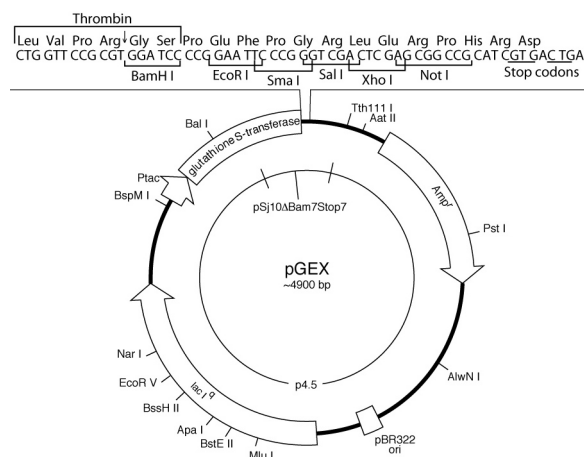
**Table 2: List of bacterial strains used**



## 4.2.2. Vectors

### 4.2.2.1. pGEX-4T1

The pGEX-4T1 vector was purchased at GE Healthcare (Germany).

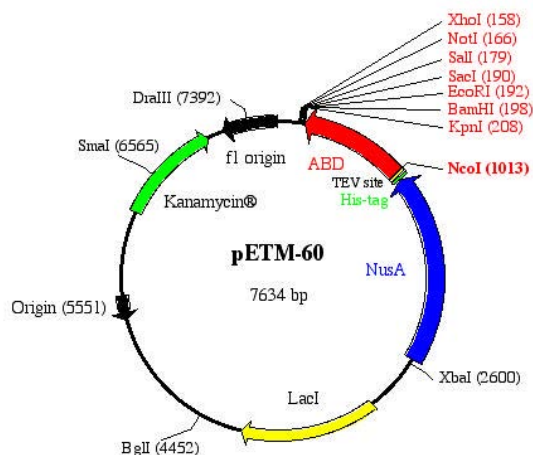


**Figure 8: Properties of pGEX-4T1 with control regions.**

- \*Vector size (bp) 4969;
- \*Glutathione S-transferase gene region: *tac* promoter: -10: 205-211; -35: 183-188; *lac* operator: 217-237; Ribosome binding site for GST: 244; Start codon (ATG) for GST: 258; Coding region for Thrombin cleavage: 918-935;
- \*Multiple Cloning Site: 930-966;
- \*b-lactamase gene region: Promoter: -10: 1330-1335; -35: 1307-1312; Start codon (ATG): 1377; Stop codon (TAA): 2235;
- \**lacIq* gene region: Start codon (GTG): 3318; Stop codon (TGA): 4398;
- \*Plasmid replication region: Site of replication initiation: 2995; Region necessary for replication: 2302-2998;

### 4.2.2.2. pETM-60

The pETM-60 vector was provided by Stier at EMBL (Germany).



**Figure 9: Properties of pETM-60 with control region.**

- \*Vector size (bp): 7634;
- \*NusA gene region: T7 promoter: 2633-2651; T7 terminator: 1-129; NusA coding sequence: 1068-2564 (substituted by Ubiquitin in Ubiquitin fused constructs); His-tag coding sequence: 1050-1067; TEV cleavage site coding sequence: 1020-1040; *lacI* operator: 3038-4120;
- \*Multiple cloning site: 158-208;
- \*Plasmid replication region: pBR322 origin: 5535-6154; fl origin: 7300-7606; KanamycinR: 6260-7075.

## 4.2.3. Oligonucleotides

All oligonucleotides were purchased from Eurofins MWG Operon (Germany).

Cloning primers for TBK1\_ULD\_P3F5

Forward: 5'-GATCTGGTTC CGCGTGGATCCGAGGAAGAAACCAGTGATGTGCTTCACCG-3'

Reverse: 5'-CGGTGAAGCACATCACTGGTTTCTTCCTCGGATCCACGCGGAACCAGATC-3'

## Materials

---

### Cloning primers for TBK1\_ULD\_VL10AD

Forward: 5'-GAATTCACCAAGTGTGATGCGGATCACCGAATG-3'

Reverse: 5'-CATTCGGTGATCCGCATCACTGGTGAATTC-3'

### Cloning primers for TBK1\_ULD\_V14D

Forward: 5'-CTTCACCGAATGGATATCCATGTCTTC-3'

Reverse: 5'-GAAGACATGGATATCCATTCGGTGAAG-3'

### Cloning primers for TBK1\_ULD\_F18D

Forward: 5'-GTAATCCATGTGCGATTCGCTACAACAC-3'

Reverse: 5'-GTGTTGTAGCGAATCGACATGGATTAC-3'

### Cloning primers for TBK1\_ULD\_Y29D

Forward: 5'-ACGGCGCATAAGATTGACATTCACAGC-3'

Reverse: 5'-GCTGTGAATGTCAATCTTATGCGCCGT-3'

### Cloning primers for TBK1\_ULD\_I30D

Forward: 5'-CGCATAAGATTTACGATCACAGCTATAAC-3'

Reverse: 5'-GTTATAGCTGTGATCGTAAATCTTATGCG-3'

### Cloning primers for TBK1\_ULD\_Y33D

Forward: 5'-GATTTACATTCACAGCGATAAACTGCTGC-3'

Reverse: 5'-GCAGCAGTGTTATCGCTGTGAATGTAAATC-3'

### Cloning primers for TBK1\_ULD\_Y44D

Forward: 5'-TTCCATGAACTGGTCGACAAACAAACCAAG-3'

Reverse: 5'-CTTGGTTTGTGTCGACCAGTTCATGGAA-3'

### Cloning primers for TBK1\_ULD\_P91H92

Forward: 5'-GTCACGAGCCTCGAGCGGGAGGAACGTGACTGACTGACGATCTGCC-3'

Reverse: 5'-GGCAGATCGTCAGTCAGTCACGTTCCCTCCCGCTCGAGGCTCGTGAC-3'

### Cloning primers for TBK1\_ULD\_L88R90R93

Forward: 5'-CCTATCTTTGTCACGAGCCGCGAGCAGCCGCATGATGACTGACTGACG-3'

Reverse: 5'-CGTCAGTCAGTCATCATGCGGCTGCTCGCGGCTCGTGACAAAGATAGG-3'

Cloning primers for Ub\_NBR1-LIR\_Y732W

Forward: 5'-GCTCTGCTTCCTCAGAGGATTGGATCATCATCCTGCCGG-3'

Reverse: 5'-CCGGCAGGATGATGATCCAATCCTCTGAGGAAGCAGAGC-3'

Sequencing primers for N-terminal of GST-constructs

Forward: 5'-GGGCTGGCAAGCCACGTTTGGTG-3'

Reverse: 5'-CACCAAACGTGGCTTGCCAGCCC-3'

Sequencing primers for C-terminal of GST-constructs

Forward: 5'-CCGGGAGCTGCATGTGTCAGAGG-3'

Reverse: 5'-CCTCTGACACTTGACAGTCCCGG-3'

Sequencing primers for pETM-60 Upstream

Forward: 5'-ATGCGTCCGGCGTAGA-3'

Reverse: 5'-TCTACGCCGGACGCAT-3'

### 4.3. Buffers, equipments and solutions

All chemicals were purchased from Roth (Germany) except when notified.

All buffers and solutions were prepared with Milli Q water, filtered through 0.2 µm membrane and kept at +4°C if not stated otherwise.

Heat stable glassware as well as temperature resistant solutions were autoclaved for 30 min at 120°C and 2 bar.

#### 4.3.1. Cloning

**Agarose 1%:** 1% (w/v) agarose boiled in 0.5% TBE buffer. Store at RT.

**DNA-sample buffer:** 50% (v/v) glycerin and 0.1% (w/v) bromophenol blue, dissolve in 100 mM Tris-HCl, pH 8.0. Store at RT.

**Ethidiumbromide stock solution:** 10 mg ml<sup>-1</sup> ethidiumbromide stock solution

**Mini and Midi DNA preparation kits** (Macherey-Nagel, Germany)

**Polymerases and polymerases buffer** (New England Biolabs)

**Restriction enzymes** (New England Biolabs)

#### 4.3.2. Expression media

**Amino acid stocks** (Sigma-Aldrich, Germany):

**Antibiotics stock solutions** (stored at -20°C)

**Ampicillin:** 150 mg ml<sup>-1</sup> Na<sup>+</sup>-ampicillinsalt in 50% EtOH

**Kanamycin:** 75 mg ml<sup>-1</sup> kanamycin sulfate in H<sub>2</sub>O

**Glycerol cultures:** Mix 60 µl of glycerin at 50% with 100 µl cell suspension and store at -80°C.

**IPTG stock solution:** 1 M in H<sub>2</sub>O

**LB medium:** 10 g tryptone, 5 g yeast extract, 10 g NaCl, for 1L

**M9 medium:** I: 290 mL H<sub>2</sub>O + 200 mL M9x5 + 10 mL <sup>15</sup>NH<sub>4</sub>Cl, 1 mL MgSO<sub>4</sub>

II: 490 mL H<sub>2</sub>O + 10 mL glucose (or 20 mL <sup>13</sup>C-glucose), 4 mL glycerol at 50%,  
1 mL CaCl<sub>2</sub>, 300 µL thiamine

Mix I and II, add 1.0 mL TRACE-elements-I

**Ammonium chloride (NH<sub>4</sub>Cl or labeled <sup>15</sup>NH<sub>4</sub>Cl):** 1 g in 10 mL H<sub>2</sub>O at pH 7.0.

**M9-salts (Maniatis) (5x):** 44.5 g Na<sub>2</sub>HPO<sub>4</sub>, 2 H<sub>2</sub>O; 15 g KH<sub>2</sub>PO<sub>4</sub>; 2.5 g NaCl. pH 7.2 for 1L.

**Calcium chloride (CaCl<sub>2</sub>):** 1M in H<sub>2</sub>O

**Glucose (<sup>13</sup>C labeled and unlabeled):** 2 g in 10 mL H<sub>2</sub>O

**Magnesium sulfate (MgSO<sub>4</sub>)** (Sigma-Aldrich, Germany): 1M in H<sub>2</sub>O

**Thiamine:** 20 mg ml<sup>-1</sup> stock solution in H<sub>2</sub>O. Store at -20°C

**TRACE-elements-I:** 2.50 g EDTA; 0.25 g FeCl<sub>3</sub>, 6 H<sub>2</sub>O; 0.025 g ZnO; 0.005 g CuCl<sub>2</sub>, 2 H<sub>2</sub>O;  
0.005 g Co(NO<sub>3</sub>)<sub>2</sub>, 6 H<sub>2</sub>O; 0.005 g (NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>, 4 H<sub>2</sub>O. pH 7.0 in 500 mL H<sub>2</sub>O.

**Protease inhibitors cocktail (50x)** (home made, protease inhibitors from Sigma-Aldrich, Germany): 59.88 mg AEBSF; 11.89 mg leupeptin; 7.71 mg bestatin; 4.88 mg aprotinin; 4.46 mg E-64; 1.71 mg pepstatin. In 50% MeOH.

**Selective labeling (for 1L of expression in minimal media, only half of the amount used when the amino acid is labeled):**

- |                         |                         |
|-------------------------|-------------------------|
| - alanine: 500 mg       | - leucine: 230 mg       |
| - arginine: 400 mg      | - lysine: 420 mg        |
| - aspartic acid: 400 mg | - methionine: 250 mg    |
| - asparagine: 650 mg    | - phenylalanine: 130 mg |
| - cysteine: 50 mg       | - proline: 100 mg       |
| - glutamic acid: 650 mg | - serine: 2100 mg       |
| - glutamine: 400 mg     | - threonine: 230 mg     |
| - glycine: 550 mg       | - tryptophan: 56 mg     |
| - histidine: 100 mg     | - tyrosine: 170 mg      |
| - isoleucine: 230 mg    | - valine: 230 mg        |

**SOC medium:** 20.0 g tryptone; 5.0 g yeast extract; 0.5 g NaCl; 950 mL H<sub>2</sub>O. Shake it till totally dissolved. Add 10 mL of a 250 mM KCl solution (1.86 g KCl in 100 mL H<sub>2</sub>O). Adjust pH to 7.0. Adjust total volume to 1 L. Autoclave. Add 5 mL of a sterile 2 M MgCl<sub>2</sub> solution (19 g MgCl<sub>2</sub> in 100 mL H<sub>2</sub>O). Add 20 mL of a sterile 1 M glucose solution (18 g glucose in 100 mL H<sub>2</sub>O, filtered).

#### 4.3.3. Purification buffers

**Ammonium acetate buffer:** 10 mM ammonium acetate at pH 7.3

**DNase I from bovine pancreas (Sigma-Aldrich, Germany):** 5 000 units in 20 mM Tris and 100 mM MgSO<sub>4</sub> (for 2 L of expression)

**DTT:** 500 mM stock solution

**GST purification buffer:**

- Loading buffer with 1xPBS at pH 7.5
- Elution buffer with 20 mM GSH in 1xPBS at pH 8.0

**IEC purification buffer:**

- Loading buffer with 25 mM Na<sub>2</sub>HPO<sub>4</sub>, 2 H<sub>2</sub>O and 25 mM NaCl
- Elution buffer with 25 mM Na<sub>2</sub>HPO<sub>4</sub>, 2 H<sub>2</sub>O and 1 M NaCl

pH adjusted depending on proteins pI.

**Lysis buffer:** 50 mM Tris-HCl, 250 mM NaCl, 5% glycerol in H<sub>2</sub>O at pH 7.5

**Nickel-NTA purification buffer:**

- Loading buffer with 250 mM NaCl, 50 mM Tris-HCl, 1% glycerol, 10 mM imidazole at pH 7.7
- Elution buffer with 250 mM NaCl, 50 mM Tris-HCl, 1% glycerol, 400 mM imidazole at pH 7.7

**PBS (10x):** 1.37 M NaCl, 0.03 M KCl, 80 mM Na<sub>2</sub>HPO<sub>4</sub> and 15 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.5. at RT.

**Sodium azide (NaN<sub>3</sub>)** (Fluka Sigma-Aldrich, Germany): 0.3% (w/v) stock solution in H<sub>2</sub>O.

#### 4.3.4. NMR buffers

**NMR buffer I:** 25 mM Na<sub>2</sub>HPO<sub>4</sub>, 2 H<sub>2</sub>O, 100 mM NaCl, in H<sub>2</sub>O at pH 7.3

**NMR buffer II:** 100 mM Na<sub>2</sub>HPO<sub>4</sub>, 2 H<sub>2</sub>O, 70 mM NaCl, 1.4 mM KCl and 0.9 mM KH<sub>2</sub>PO<sub>4</sub> at pH 7.3

**NMR buffer III:** 100 mM Na<sub>2</sub>HPO<sub>4</sub>, 2H<sub>2</sub>O, 30 mM NaCl, 5 mM KCl at pH 7.3

**ULD buffer:** 50 mM Tris in H<sub>2</sub>O at pH 7.2

**LC3- buffer:** 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 2 H<sub>2</sub>O, 100 mM NaCl at pH 7.0

#### 4.3.5. Gel chromatography buffers

**10xTGS:** 30 g Tris, 144 g Glycerin, 10 g SDS at pH 8.6 for 1L

**Coomassie Staining solution:** 50% (v/v) ethanol (96%), 10% acetic acid (100%), 0.1% G250 coomassie brilliant blue, completed with water

**Destaining solution:** 40% (v/v) MeOH (96%), 10% (v/v) acetic acid (100%), completed with water

#### **Polyacrilamide gels (for 2 gels):**

- **Running gel 12%:** 2.5 mL H<sub>2</sub>O, 1.0 mL 1% SDS, 2.5 mL buffer A, 4.0 mL AcA-30, 50 μL APS 10%, 5 μL TEMED Elution buffer with 250 mM NaCl, 50 mM Tris-HCl, 1% glycerol, 400 mM imidazole at pH 7.7
- **Running gel 16.5%:** 1.5 mL H<sub>2</sub>O, 1.0 mL 1% SDS, 2.5 mL buffer A, 5.0 mL AcA-30, 50 μL APS 10%, 5 μL TEMED
- **Running gel 18%:** 0.5 mL H<sub>2</sub>O, 1.0 mL 1% SDS, 2.5 mL buffer A, 6.0 mL AcA-30, 50 μL APS 10%, 5 μL TEMED
- **Stacking gel 4%:** 2.5 mL H<sub>2</sub>O, 0.5 mL 1% SDS, 1.3 mL buffer B, 0.7 mL AcA-30, 25 μL APS 10%, 5 μL TEMED
- **Buffer A:** 1.5 M Tris at pH 8.8
- **Buffer B:** 0.5 M Tris-HCl at pH 6.8

**Protein sample buffer:** 150 mM Tris-HCl, 12% (w/v) SDS, 30% (v/v) glycerin, 6% (v/v) β-mercaptoethanol, 0.05% coomassie blue at pH 7.0

#### **Tricine gels:**

- **Anode buffer:** 100 mM Tris, 22.5 mM HCl (37%) at pH 8.9
- **Cathode buffer:** 100 mM Tris, 100mM Tricine, 0.1% (v/v) SDS at pH 8.25

#### 4.4. Software

Adobe package

Microsoft Office

Nanodrop ND-1000 v.3.7.0

Origin 6.1

Sparky 3.115

Topspin

Unicorn 5.11

Spectra Manager

## 5. Methods

### 5.1. Primary sequence analysis

ClustalW2 was used to align primary sequences of the different proteins and to calculate the percentage of homology between these sequences.

Physico-chemical parameters (amino acid and atomic compositions, isoelectric point, extinction coefficient...) based on the protein sequences were determined by the ProtParam software from the ExPASy Proteomics Server.

### 5.2. Cloning

Initial GST-fused constructs (TBK1\_ULD, GABARAPL-1, GABARAPL-2, LC3A, LC3B) were provided from the Dikic group.

#### 5.2.1. *Substitution of GST to NusA*

After Polymerase Chain Reaction (PCR), pGEX\_TBK1\_ULD was digested with the restriction endonucleases NcoI and BamHI. The gene product of interest and the expression vector were then separated by agarose gel electrophoresis and extracted from the gel by using the QIAquick gel extraction kit. The ligation of the TBK1\_ULD PCR product with the open pETM-60 vector was performed overnight at 4°C according to the manufacturer's instructions using the T4-DNA-ligase from NEB. 10 µl of the ligation product were transformed the next day into *E. coli* DH5α cells. After the cell growth in 3 ml LB of overnight culture, the plasmid DNA was then purified and the insertion of TBK1\_ULD PCR product into pETM-60 vector backbone was analyzed by restriction digest. Positive clones were submitted for sequencing.

#### 5.2.2. *Substitution of NusA to ubiquitin*

PCR fragments encoded human LC3 modifiers (LC3A, LC3B and GABARAPL-1) were cloned into pETM-60 vector where the NusA leader was substituted with modified ubiquitin between NcoI and BamHI sites. For the peptides (p62-LIR, Nix-LIR\_W36, Nix-LIR\_W140/144 and NBR1-LIR), entire DNA fragments were ordered at MWG Bioscience and cloned at the same sites.

### 5.2.3. *Mutagenesis*

To introduce single mutation on plasmid DNA, PCR based site-directed mutagenesis was used as follows:

- initialization for 60 seconds at 95°C
  - denaturation for 30 seconds at 95°C
  - annealing for 30 seconds at 60°C cycle
  - extension for 16 minutes at 68°C
  - end at 4°C
- } x 18 times

with a mixture composed of:

- 10 ng plasmid DNA
- 20 pmol 5' primer oligonucleotide
- 20 pmol 3' primer oligonucleotide
- 2.5 mmol dNTPs
- 2.5 U polymerase
- 5 µL 10 times Polymerase buffer
- 4 mM MgSO<sub>4</sub> (samples were also prepared in parallel without MgSO<sub>4</sub>)
- complete to 50 µL with sterile water

Once the PCR achieved, 1 µL dpNI were added to each probe and let at 1 hour for 37°C to cleave methylated DNA. Amplified DNA were analyzed on a 1% agarose gel.

### 5.3. Isolation of plasmid-DNA and measurement of concentration

Positive cloning samples were purified on analytical scale with 3 mL overnight culture using the NucleoPlasmid PC Kit („Mini“) and on preparative scale with 100 mL overnight culture using the NucleoSpin Kit („Midi“) following manufacturer protocols (Macherey-Nagel, Düren). DNA pellets were resuspended in H<sub>2</sub>O. The measurement of plasmid-DNA concentration was performed by measuring the absorbance at 260 nm using the Nano-Drop Spectrometer.

### 5.4. Transformation

Transformation of competent bacterial cells to take up plasmid DNA was based on the heat shot protocol.<sup>139</sup> Frozen competent cells were removed from the -80°C freezer and put on ice until melting. 50 ng DNA were added to the cells and kept on ice for 30 minutes then placed for



30 seconds at 42°C for heat-shock and again on ice for 2 minutes. Then 450 µL SOC medium were added and the tube placed at 37°C in the shaker. After 1 hour growth, 150 µL of the transformed cells were spread on an agar plate with the appropriate antibiotic and incubate at 37°C overnight.

Plasmids encoded GST constructs and NusA or Ub-fused proteins/peptides were respectively transformed into *E.Coli* DH5α and *E.Coli* NEB T7 Express (equivalent to BL21) strains.

## 5.5. Sequencing

After analytical expression (cf. below), putative “positive” clones were sent to SeqLab (Göttingen) for sequencing using 600 to 900 ng of plasmid, 20 pmol of primer and completed with sterile water for a final volume of 7 µL.

## 5.6. Cell growth and expression

### 5.6.1. Analytical expression

Single colonies were scratched from agar plate and resuspended in 3 mL LB medium with 1 mM antibiotics. After reaching an optical density at 600 nm ( $OD_{600}$ )  $\approx$  1, cells were induced with different concentration of isopropyl-β -D-thiogalactopyranosid (IPTG), 1mM, 0.5 mM, 0.25 mM and 0.1 mM at 37°C for 3 hours or at 20°C for 16 h.

### 5.6.2. Preparative expression

Preculture of transformed cells were grown in 50 mL LB overnight with the appropriate antibiotics. For DNA plasmid presenting ampicillin resistance, the preculture cells were spinned down the next morning, the supernatant was poured off and the cells resuspended in 1L medium. In case of kanamycin, cells were then directly inoculate into 1L medium.

Bacteria were grown in LB medium for expression of unlabeled proteins/peptides or in M9 medium for expression of <sup>15</sup>N-labeled (1g/L <sup>15</sup>NH<sub>4</sub>Cl) and <sup>13</sup>C/<sup>15</sup>N-labeled (1g/L <sup>15</sup>NH<sub>4</sub>Cl and 2g/L <sup>13</sup>C-glucose) proteins/peptides. After  $OD_{600} \approx 1$ , IPTG at final concentration 1mM was added to induce gene expression. All cells were expressed for 3 hours at 37°C, except TBK1\_ULD cells, which were expressed for 16h at 20°C.

### 5.6.3. Selective labeling

Selective labeled samples were produced only for TBK1\_ULD. A sample selectively <sup>15</sup>N labeled for the lysine residues was produced by expression in BL21 bacteria. Samples selectively

<sup>15</sup>N labeled for the phenylalanine, tyrosine, leucine or isoleucine residues were expressed in the auxotrophic strain DL39 following standard protocols. The amount of amino acids used for one liter of expression in minimal media is listed in the material section.

## **5.7. Protein isolation and purification**

### **5.7.1. Cell lysis**

Resuspended pellets from cell cultures were either lysed by French Press under a 1200 psi pressure or using sonication with 50 pulses during 30 seconds with 1 minute rest, this cycle being repeated 8 times.

The cell lysate was harvested at 18 000 rpm for 30 min at 4°C to remove cell debris. Resulting pellets were discarded whereas the supernatant was loaded on the appropriate affinity chromatography column for purification after after being filtrated through a 0.2 µm filter.

### **5.7.2. GST purification**

The supernatant of the cell lysate was loaded with an external pump on a glutathione sepharose 4 Fast Flow (GE Healthcare) column. The column was washed with 4 column volumes (CV) of 1xPBS buffer, pH 7.5 and GST-fused proteins were eluted with 1xPBS, 20 mM reduced glutathione, pH 7.9. Loading, washing and eluting steps were performed at a flow-rate of 2 mL min<sup>-1</sup>.

### **5.7.3. NiNTA purification**

The supernatant of the cell lysate was loaded onto the Ni-NTA column with an external pump. The column was previously equilibrated with the loading buffer (50 mM Tris, 250 mM NaCl, 1% glycerol and 10 mM imidazole, pH 7.7). After washing with 3 CV of loading buffer, the NusA constructs as well as the Ub-fused protein/peptides constructs containing a hexahistidine tag were eluted with a linear gradient (10-400 mM) of imidazole with a flow rate of 2 mL min<sup>-1</sup>.

### **5.7.4. Tag cleavage**

GST-fusion proteins were cleaved with thrombin (1 unit used for 100 µg proteins) (GE Healthcare) directly in GST elution buffer overnight at 16°C.

NusA and ubiquitin were cleaved by tobacco etch virus (TEV) protease (produced in house) after adding 2 mM EDTA and 10 mM β-mercaptoethanol to NiNTA elution buffer overnight at 16°C.

### 5.7.5. *Cation-exchange chromatography*

A 20 mL sp-sepharose column was used with a flow rate of 2 mL min<sup>-1</sup> for the purification of cleaved Ub-constructs and a 5 mL HiTrap sp ff column to repurify samples after NMR and ITC experiments. The column was equilibrated with 3 CV of 25 mM sodium phosphate, 25 mM NaCl at pH 6.3. The sample was applied with an external loop and unbound proteins were removed by washing the column for 10 CV with 2 ml min<sup>-1</sup>. The bound proteins were eluted with a linear gradient from 25 mM to 1M NaCl at 2 ml min<sup>-1</sup>.

### 5.7.6. *Buffer exchange*

To exchange buffer conditions, samples with a volume below 5 mL were applied to HiPrep 26/10 Desalting (GE Healthcare) column with a 4 mL min<sup>-1</sup> flow rate. For samples with a bigger volume, buffer exchange was achieved using 3 to 8 kDa cut-off dialysis membranes, stirred solutions of 500 times the sample volume at 4°C and two buffer changes within 24 h.

### 5.7.7. *Size-exclusion chromatography*

In order to purify proteins by size-exclusion chromatography, a Superdex 75 16/60 column was used. 5 mL of filtered solution through 0.2 µm filter were injected on the equilibrated column in final buffer conditions through an external loop with a 1 mL min<sup>-1</sup> flow rate. 3 mL elutions fraction collection were collected.

All affinity chromatography techniques were used on ÄKTA™ systems at 4°C.

Elution fractions were collected following protein absorbance at 280 nm and their purity was checked by sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

## 5.8. Peptide preparation

p62-LIR and NBR1-LIR peptides were purchased from Genecust (Luxembourg). Nix-LIR\_W36 and Nix-LIR\_W140/144 were provided by AK Schwalbe. NBR1-LIR mutants (NBR1-LIR\_Y732F, NBR1-LIR\_S729E, NBR1-LIR\_S728,729E) were purchased from Peptide Speciality Laboratories GmbH (Heidelberg, Germany).

The solid-phase synthesized lyophilized peptides were resuspended in 50 mM sodium phosphate, 100 mM NaCl at pH 7.0 to make stock solutions according to their molecular size.

Ub-fused peptides samples for structural determination by NMR (unlabeled, <sup>15</sup>N, <sup>13</sup>C and <sup>15</sup>N/<sup>13</sup>C labeled) were cleaved with TEV protease after NiNTA elution and purified by size-exclusion chromatography in 10 mM ammonium acetate at pH 7.3. The resulting elution

fractions were combined in a glass vial and frozen in liquid nitrogen. Once the refrigerating unit reached  $-40^{\circ}\text{C}$  and a pressure of 0.1 mbar, the sample was attached to the freezer dryer unit and let overnight under vacuum. After lyophilization, peptides were resuspended in 50 mM sodium phosphate, 100 mM NaCl at pH 7.0.

### **5.9. Concentration samples**

Centrifuge concentrators from Millipore with a cut-off of 3 or 10 kDa were used to concentrate protein solutions at 3 000 rpm until reaching the final wished concentration/volume.

### **5.10. Sample state**

#### ***5.10.1. Gel chromatographies***

To identify the expression, the isolation, the purification and the stability efficiency of the target proteins, SDS-PAGE were performed. Depending on the size of the proteins of interest, different gel compositions were used (12%, 16.5% or 18% polyacrylamide or tricine gels).

Running gel was first poured and then sealed with water on the top. After polymerization, water was removed and replaced stacking gel. A comb with 10 or 15 slots were inserted to load samples. Protein samples diluted 1:5 in SDS buffer were heated up for 5 minutes at  $95^{\circ}\text{C}$ . Unstained protein marker from Fermentas (St. Leon-Rot) was used as reference. SDS-PAGE run for 15 minutes under 100 V and then 50 minutes under 200 V with 1xTGS buffer pH. Tricine gels run first for 20 minutes under 70 V followed with 45 minutes under 150 V cathode buffer was placed in the inside chamber and anode buffer outside. Gels were stained in Coomassie solution for 30 minutes and, after wash under water, destained in destaining solution for also 30 minutes.

#### ***5.10.2. Concentration determination***

Based on their extinction coefficient, pure protein concentration was calculated from the absorption at 280 nm using NanoDrop or Cary UV following the Beer-Lamber law:

$$A = \varepsilon * l * C$$

where A is the absorbance at 280 nM,  $\varepsilon$  the extinction coefficient of the protein, l the path length of the cuvette and C the concentration of the sample.

### **5.11. Circular Dichroism (CD) spectroscopy**

CD spectroscopy uses the interaction of polarised light with biological molecules like proteins, based on their chiral properties. Due to their structural asymmetry, chiral molecules

exhibit circular dichroism, meaning a differential absorbance of left-hand circularly polarised and right-hand circularly polarised light.

CD spectroscopy of Ub\_GABARAPL-1, GABARAPL-1 and ubiquitin proteins was done in a 10 mM sodium phosphate buffer at pH 7.0 with a Jasco J-810 spectropolarimeter (Jasco Labortechnik, Gross-Umstadt, Germany) using a quartz cuvette with a path length of 0.1 cm. Assays were carried out at standard sensitivity with a band width of 1 nm and a response of 1 s. The data pitch was 0.1 nm and the scanning rate 50 nm per minute. The CD spectra were recorded from 300 to 190 nm. The presented data are the average of 20 scans. Melting curves for the same proteins in the same conditions were recorded monitoring ellipticity at 222 nm from 4°C to 100°C with the rate of 1°C per minute. The temperature was controlled using a digitized water bath integrated with the instrument. The spectra were signal-averaged over at least three scans, baseline corrected by subtracting a buffer spectrum and smoothed by the mean value averaging.

### **5.12. Isothermal Titration Calorimetry (ITC)**

ITC measures the gain or loss of heat appearing during reactants mixture. Applied to an interacting system, the amount of heat released or absorbed is directly connected the amount of binding in this system. Reaction occurs in a sample cell containing the titrand and where a spinning syringe injects regularly precise amount of titrant. Compared to the reference cell maintained at a constant temperature, a feedback regulation as electrical power is applied to the sample cell to respond to the thermal changes due to the interaction and to reach back the equilibrium. At saturation, the heat signal decreases and only background heat dilution is observed, which need to be subtracted from the heat interaction.

ITC experiments were performed at 25°C using a VP-ITC calorimeter (MicroCal Inc., Northampton, MA, USA) and analyzed with the ITC-Origin software (MicroCal Inc.) based on the assumption of one-site binding reactions.

For LC3A and LC3B experiments, 400 µM p62-LIR was titrated into 15 µM LC3B in 26 steps (10 µL per injection), 1.6 mM Nix-LIR\_W36 into 80 µM LC3B, 2.0 mM Nix-LIR\_W140/144 into 100 µM LC3B and 450 µM Ub\_NBR1-LIR into 35 µM LC3B in 16 steps (15 µL per injection).

For GABARAPL-1 experiments, 350 µM Ub\_NBR1-LIR, Ub\_NBR1-LIR\_Y211W, NBR1-LIR\_Y732F, NBR1-LIR\_S729E and NBR1-LIR\_S728,729E were titrated into 25 µM GABARAPL-1 in 26 steps (10 µL per injection).

To measure the heat contribution due to the titrant alone after saturation of the interaction, heat dilution effect of the different peptides were analyzed by titrating the titrant against the buffer. These data were then subtracted from the interaction data to only have the heat due to interaction involved in the thermodynamics data calculation.

### **5.13. Nuclear Magnetic Resonance (NMR)**

Detailed description of the NMR principle is exhaustively available in the literature. In short, NMR is based on the magnetic properties of the nuclei of each atom. Under a high magnetic field, the nuclei have spinning charges, which are all oriented in the same direction and precess about the magnetic field at the Larmour frequency. By applying a radio-frequency pulse, a magnetic resonance occurs when the spectrometer frequency matches with the nuclear precessing frequency and produces thus a transverse magnetization for each spin. This transverse magnetization precesses differently for each spin due to the presence of the surrounding spins, which create additional local fields. The detection of the resulting precession provides information of individual nuclei affected differently by their local environment, reflected by a chemical shift. A chemical shift is the difference between a resonance frequency and that of a reference substance. Thus, the total dispersion of the chemical shifts of a molecule is a specific picture corresponding to the different atoms (in 1D NMR) or to the relation between atoms (in 2D or 3D NMR) of the molecule.<sup>140</sup>

#### ***5.13.1. Conditions***

All NMR experiments were performed on Bruker Avance spectrometers operating at proton frequencies between 600 and 900 MHz. All experiments were done at 298K except when mentioned. The resulting spectra were processed by Topspin (Bruker) and analyzed by Sparky (UCSF) programs.

The list of the different NMR experiments used during this work as well as their common signification are listed in **Table 3**.

NMR experiments	Purposes	
	Correlation	Common Use
$[^{15}\text{N}, ^1\text{H}]$ HSQC-TROSY	$^{15}\text{N}$ and attached $^1\text{H}$	Protein backbone fingerprint
$[^{13}\text{C}, ^1\text{H}]$ HSQC	$^{13}\text{C}$ and attached $^1\text{H}$	Protein backbone fingerprint
$[^1\text{H}, ^1\text{H}]$ TOCSY	$^1\text{H}$ and attached $^1\text{H}$	Backbone and side assignment
$[^{15}\text{N}, ^1\text{H}]$ TOCSY-TROSY	NH with $^1\text{H}$ along side chain	$^1\text{H}$ side chain assignment
$[^1\text{H}, ^1\text{H}]$ NOESY	$^1\text{H}$ - $^1\text{H}$ distance	NOE constraints
$^{15}\text{N}$ -separated NOESY-HSQC	NH and H using NOE	$^1\text{H}$ side chain assignment
$^{15}\text{N}$ -separated TOCSY-HSQC	$^{15}\text{N}$ and attached $^1\text{H}$	NHHN resonance pair assignment
$^{13}\text{C}$ -separated NOESY-HSQC	$^{13}\text{C}$ and attached $^1\text{H}$	CHHC resonance pair assignment
$[^{13}\text{C}, ^{15}\text{N}]$ HNCO	NH with CO of preceding residue	Carbonyl assignment
$[^{13}\text{C}, ^{15}\text{N}]$ HN(CA)CO	NH with CO of own and preceding residue	Carbonyl assignment
$[^{13}\text{C}, ^{15}\text{N}]$ HNCA	NH with $\text{C}^\alpha$ of own and preceding residue	$\text{C}^\alpha$ assignment
$[^{13}\text{C}, ^{15}\text{N}]$ HN(CO)CA	NH with $\text{C}^\alpha$ of preceding residue	$\text{C}^\alpha$ assignment
$[^{13}\text{C}, ^{15}\text{N}]$ HNCACB	NH with $\text{C}^\alpha$ and $\text{C}^\beta$ of own and preceding residue	$\text{C}^\alpha$ and $\text{C}^\beta$ assignment
$[^{13}\text{C}, ^{15}\text{N}]$ (H)CC(CO)NH-TOCSY	NH with C of preceding residue	$\text{C}^\alpha$ , $\text{C}^\beta$ and $\text{C}^{\text{side chain}}$ assignment
$[^{13}\text{C}, ^{15}\text{N}]$ TROSY-H(CCCO)NH-TOCSY	NH with H of preceding residue	$\text{H}^\alpha$ , $\text{H}^\beta$ and $\text{H}^{\text{side chain}}$ assignment
$^{13}\text{C}$ H(C)CH-TOCSY	CH with attached H	$\text{H}^\alpha$ , $\text{H}^\beta$ and $\text{H}^{\text{side chain}}$ assignment

**Table 3: Typical NMR experiments used**

### 5.13.2. 1D NMR

Analysis of the 1D  $^1\text{H}$  NMR spectrum of a protein gives information of the general state of folding of the protein depending of the dispersion of the protein signals but the spectrum is crowded with too many peaks in order to identify each residues. A simple  $^1\text{H}$  NMR was recorded for each experiment to first check the well-fold of the protein and then to set-up parameters for 2D experiments.

### 5.13.3. 2D NMR: $[^{15}\text{N}, ^1\text{H}]$ TROSY-HSQC and $[^{13}\text{C}, ^1\text{H}]$ HSQC

Using 2D NMR like  $[^{15}\text{N}, ^1\text{H}]$  TROSY-HSQC,<sup>127</sup> the signals are dispersed between the two frequency axes and each cross peak corresponds to an amide bond of a residue of the protein, providing a fingerprint specific to each protein. Because the resonance of each cross peak depends of its surrounding, modifications of the protein due to conformation states, internal mobility, denaturation, environment changes (buffer, pH, temperature...) and presence/absence of a ligand are reflected in the different NMR spectra. Performing titration experiments by adding gradually a non labeled ligand against a  $^{15}\text{N}$  labeled protein and then observing the

chemical shift perturbations (CSP) of the cross peaks provide information about the affected amino acids; the strongest CSP are observed for the peaks, which are the most modified by the new conditions, meaning the residues involved in the interaction. By plotting the CSP between the reference state (protein alone) and each titration point until the final state (protein saturated with ligand) against the protein concentration used, a titration curve is obtained from which the dissociation constant ( $K_d$ ) could be calculated.

#### ***5.13.4. Amide/proton exchange experiment***

A [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC reference spectrum of LC3B was first recorded in NMR buffer conditions, 25 mM sodium phosphate, 100 mM NaCl at pH 7.3. The sample was then lyophilized and later resolubilized in  $\text{D}_2\text{O}$ . After a waiting time of 10 minutes, a new TROSY spectrum was recorded, then the next day and after several weeks, using the same acquisition parameters.

#### ***5.13.5. Analysis of cell lysate samples from Ub-constructs***

After 3 hours expression at  $37^\circ\text{C}$ , the expressed cells were harvested for 10 minutes at 6 000 rpm. The pellet of the Ub-fused proteins was resuspended in 100 mM sodium phosphate and 2 mM TCEP at pH 7.2. The cells were then lysed by sonication with 5 times 4 pulses. After undergoing  $\text{N}_2$  flow for 1 minute, the sample was centrifuged for 5 minutes at 14 000 rpm. The supernatant was finally filtered through a  $0.2\ \mu\text{M}$  filter before recording NMR spectra. 3 ml of the cell cultures were enough to visualize all resonances but 15 ml were necessary for 3D experiments.

#### ***5.13.6. Resonance assignment***

Nevertheless, more information are needed to understand to which residue corresponds each cross peak. 3D NMR experiments allow to characterize the interaction between the spin systems of the protein like for example, the amide bond from a residue to the  $\text{C}^\alpha$  and the  $\text{C}^\beta$  (for a HNCACB experiment) or to the CO (for HNCO) of the same residue and from the previous residue of the primary sequence. Performing a sequential assignment, each cross peak corresponding to the backbone atoms of the protein could be then identified. To complete the information about the protein, the same type of experiments has to be performed in order to assign the side chains of the amino acids constituting the protein of interest. Added to total correlation spectroscopy (TOCSY) experiments where the cross peaks are observed upon bond relation up to three covalent bonds and others available experiments, an assignment of all



protons of the protein is possible. Assignment of the protein is not enough for structure determination where each atoms of the protein has to be located in the space. Nuclear Overhauser effect spectroscopy (NOESY) experiments are the widely used NMR experiments to determine the structure of protein. The nuclear Overhauser effect (NOE) is based on the magnetization transfer between two nuclear spins via cross-relaxation, the two nuclear spins being close enough in space. NOESY experiments provide information on intramolecular distances illustrated by cross peaks between two protons which are distanced by less than 5 Å in space. This proton-proton interaction could be directly related to the protein structure, the NOESY intensity signal being inversely related to the proton-proton distance. Additionnaly in the study of a complex between two proteins, intermolecular NOEs are an important parameter that could also be observed, illustrating the close interaction between these proteins.

For the assignment of TBK1\_ULD, HNCA, HNCACB,  $^{15}\text{N}$ -separated 3D-NOESY,  $^{15}\text{N}$ -separated 3D-TOCSY, as well as homonuclear 2D-TOCSY and 2D-NOESY spectra were recorded as well as [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC for the selective labeled samples.

HN, NH,  $\text{C}^\alpha$  and  $\text{C}^\beta$  resonances for LC3B and GABARAPL-1 proteins were obtained from individual HNCACB experiments with minor guidance of reported resonances assignment from BMRB entries 5958 and 5058 for LC3B and GABARAPL-1, respectively. Backbone and side chain resonances for GABARAPL-1 in complex with NBR1-LIR peptide were assigned with the standard set of the 3D experiments: HNCA, HNCACB, HNCOC, HN(CA)CO, (H)CC(CO)NH-TOCSY, TROSY-H(CCCO)NH-TOCSY. The assignment was proven and completed with  $^{15}\text{N}$ -edited and  $^{13}\text{C}$ -edited (for aliphatic and aromatic regions) NOESY experiments. For the resonance assignment of NBR1-LIR in complex with GABARAPL-1, the same set of NMR experiments was used with additional 3D H(C)CH-TOCSY experiment.

#### ***5.13.7. Secondary structure determination***

Based on the analysis of the  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts, CSI<sup>141</sup> and TALOS<sup>142</sup> programs identified secondary structure elements of TBK1\_ULD. Torsion angles constraints determined by TALOS based on chemical shift perturbations were added to CYANA for 3D structure calculation.

#### ***5.13.8. Tertiary structure calculation***

A first step of the calculation of protein 3D structure is the assignment on all NOEs signals and that is automated by the CANDID module of the CYANA software. Based on distance restraints obtained from the NOEs and torsion angle restraints provided by TALOS, a first cycle

of structure calculation is started by CYANA, followed by new cycles where new restraints acquired from the previous cycle are added until unique assignment for each  $^1\text{H}$ - $^1\text{H}$  pair and elimination of the non-assigned. An ensemble of 20 conformers presenting the smallest target function, the best agreement between the obtained structure and the different restraints used, is selected and submitted to refinement based on simulated annealing, an alternance of high and low temperature phases to obtain the best energy minimization. The final precision of the structure is determined by the root-mean-square deviation (RMSD) showing the deviation of the mean structure to the average structure.

#### **5.13.8.1. TBK1\_ULD structure determination**

CANDID module of CYANA 1.0.5 program<sup>143</sup> was implemented to calculate the structure of TBK1\_ULD and to assign NOE signals in 3D  $^{15}\text{N}$ - and  $^1\text{H}$ -NOESY spectra. 32 hydrogen bonds observed in Ubiquitin structure<sup>144</sup> were transposed to TBK1\_ULD sequence and were added to the initial calculations. In total, 1781 signals (866 from [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TBK1\_ULD and 915 from [ $^1\text{H}$ ,  $^1\text{H}$ ] TBK1\_ULD) together with 91 torsion angles restraints (aco) predicted by TALOS<sup>142</sup> and 26 upper limits restraints (upl) describing  $\alpha$ -helical hydrogen bonds (as consensus of TALOS and CSI<sup>141</sup> programs) were used in the initial CYANA run. 702 meaningful distance restraints, 91 torsion angle restraints and 58 upper distance limits restraints for hydrogens bonds provided by the initial calculations were used to calculate the final 20 (from 100) CYANA conformers with CYANA 3.0 (Table in Appendix 9.5). The final 20 (from 100) CYANA structures have a backbone RMSD 0.65 Å within a structured part (residues 14 to 88); 80 % NOE signals were assigned.

#### **5.13.8.2. GABARAPL-1/NBR1-LIR structure determination**

CANDID module of CYANA 1.0.5 program<sup>143</sup> was implemented to calculate structure of GABARAPL-1/NBR1-LIR complex and to assign NOE signals in 3D  $^{15}\text{N}$ - and  $^{13}\text{C}$ - NOESY spectra. In total, 4732 signals (4140 from [ $^{13}\text{C}$ ,  $^{15}\text{N}$ ] labeled GABARAPL-1 and 592 from [ $^{13}\text{C}$ ,  $^{15}\text{N}$ ] labeled NBR1-LIR) together with 203 torsion angles constraints from TALOS<sup>142</sup> and 46 upper limits constraints for  $\alpha$ -helical hydrogen bonds (identified as consensus of the TALOS and CSI<sup>141</sup> programs) were used in the initial CYANA calculations. According to CYANA algorithm, the complex was simulated by connection of the GABARAPL-1 and NBR1-LIR polypeptides with imaginary linker of 82 L5 pseudogroups. 1364 meaningful distance restraints, 203 torsions angles restraints and 84 upper distance limits restraints for hydrogens bonds provided by the initial calculations were used to calculate the final 20 (from 100) CYANA

conformers with CYANA 3.0 (Table in Appendix 9.5). These 20 final conformers with the lowest target function values were subjected to restrained energy minimization in explicit solvent against the AMBER force field<sup>145</sup> using the program OPALp.<sup>146</sup> The final 20 (from 100) CYANA structures have backbone RMSD 0.65 Å within a structured part (GABARAPL-1 residues 12-114 and NBR1-LIR residues 729-737); near 85 % of the NOE signals were assigned.

### **5.13.9. Interaction studies by NMR**

#### **5.13.9.1. TBK1\_ULD project**

The titration of IAD-SRR against <sup>15</sup>N\_TBK1\_ULD was done by co-concentrating both proteins, purified by size-exclusion chromatography, until reaching a molar ratio of 1/2:1, 1:1 and 2:1. For the titration of GST\_IAD-SRR against <sup>15</sup>N\_TBK1\_ULD, proteins were co-concentrated until molar ratio 1:1 and 2:1.

[<sup>15</sup>N, <sup>1</sup>H] TROSY-HSQC spectra were recorded at each step.

#### **5.13.9.2. Autophagy project**

Titration of p62-LIR, Nix-LIR\_W36 and Nix-LIR\_W140/144 against <sup>15</sup>N-labeled proteins (~100µM) was performed at 25°C by gradually adding stock solutions of unlabeled peptides until reaching a molar ratio 1:10 (protein to peptide).

Unlabeled Ub\_NBR1-LIR was added on an eight step titration to a 120 µM <sup>15</sup>N-labeled LC3B solution until reaching five molar excess. <sup>15</sup>N-labeled GABARAPL-1 was concentrated to 100 µM. Titration experiments were performed by adding gradually unlabeled Ub\_NBR1-LIR and Ub\_NBR1-LIR\_Y211W to obtain a final 1:3 molar ratio. [<sup>15</sup>N, <sup>1</sup>H] TROSY-HSQC spectra were recorded at each step.

Titration experiments were performed by adding gradually unlabeled GABARAPL-1 to <sup>15</sup>N-labeled NBR1-LIR and <sup>13</sup>C-labeled NBR1-LIR (1 mM concentration) until reaching 1:1 molar ratio. After each addition of GABARAPL-1, [<sup>15</sup>N, <sup>1</sup>H] TROSY-HSQC or [<sup>13</sup>C, <sup>1</sup>H] HSQC spectra were recorded for <sup>15</sup>N\_NBR1-LIR and <sup>13</sup>C\_NBR1-LIR, respectively.

#### **5.13.9.3. Analysis of titration data**

For each individual amid group, the chemical shift differences, Δδ, at each titration step were calculated using the following formula:<sup>147</sup>

$$\Delta\delta = \sqrt{\Delta\delta_{HN}^2 + (\Delta\delta_{NH} / 6.5)^2}$$

with  $\Delta\delta_{HN}$  and  $\Delta\delta_{NH}$  representing the chemical shift differences for the amide proton and the amide nitrogen of one residue, respectively.

For the interactions where a fast exchange mode was observed and a “one binding site” mode was assumed,  $K_d$  was calculated using the least-square method. The following equation for the  $K_d$  calculations was used:<sup>148; 149</sup>

$$\Delta\delta = \frac{\Delta\delta_{\max}}{2C} (C + C * MR + K_d) - \sqrt{(C * MR + C + K_d)^2 - 4 * MR * C^2}$$

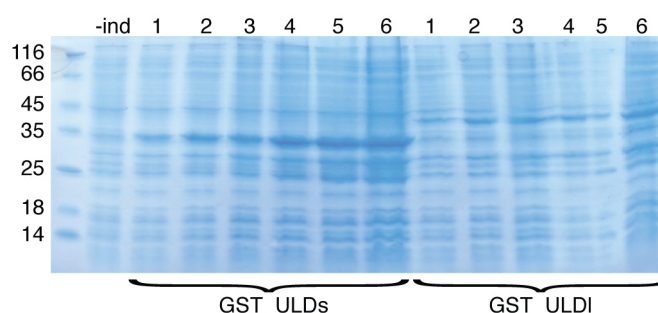
with  $\Delta\delta$  the current chemical shift difference value (Y-axis value),  $\Delta\delta_{\max}$  the maximal chemical shift difference upon the titration, C the protein concentration in the NMR tube,  $K_d$  the dissociation constant and MR the molar ratio where the  $\Delta\delta$  was observed (X-axis value). The global fitting over all individual titration curves was done using a global minimization algorithm (MatLab 6.5 software package).

## 6. Results

### 6.1. TBK1\_ULD project

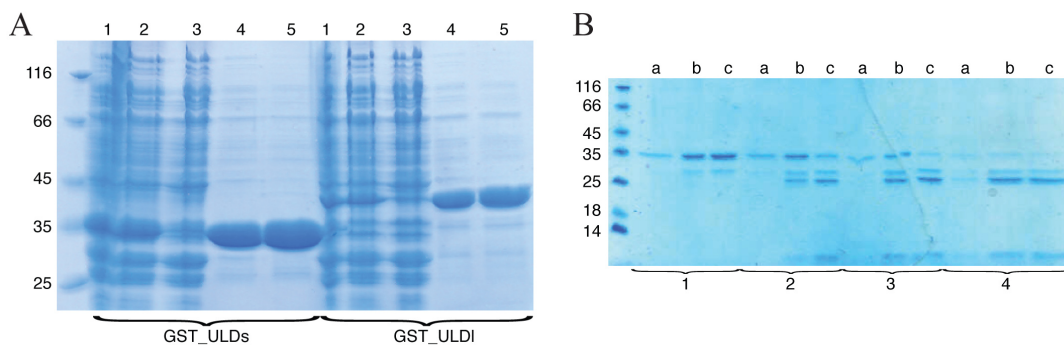
#### 6.1.1. Expression and purification of TBK1\_ULD

To confirm by NMR the presence of the ubiquitin motif in TBK1 predicted *in silico*, two TBK1 constructs were provided by Dr. Fumiyo Ikeda to check which one could be the most preferable for NMR experiments. The first construct TBK1\_ULD<sub>s</sub> (residues 302-383) contains only the predicted ULD domain whereas the second construct TBK1\_ULD<sub>l</sub> (residues 282-403) is larger. The TBK1\_ULD constructs were first analytically expressed to screen the best conditions of expression. The first expression try showed that both TBK1\_ULD constructs expressed as soluble proteins at 37°C but with a low expression yield. The expression at 25°C overnight increased the amount of GST\_TBK1\_ULD produced. Therefore, this protocol was used for protein expression in a preparative scale (**Figure 10**). The expression of the protein of interest was induced by addition of IPTG. Different concentrations were used for the induction at 25°C (200 μM, 500 μM and 1 mM final concentration). The best expression yield was obtained for 1 mM IPTG.



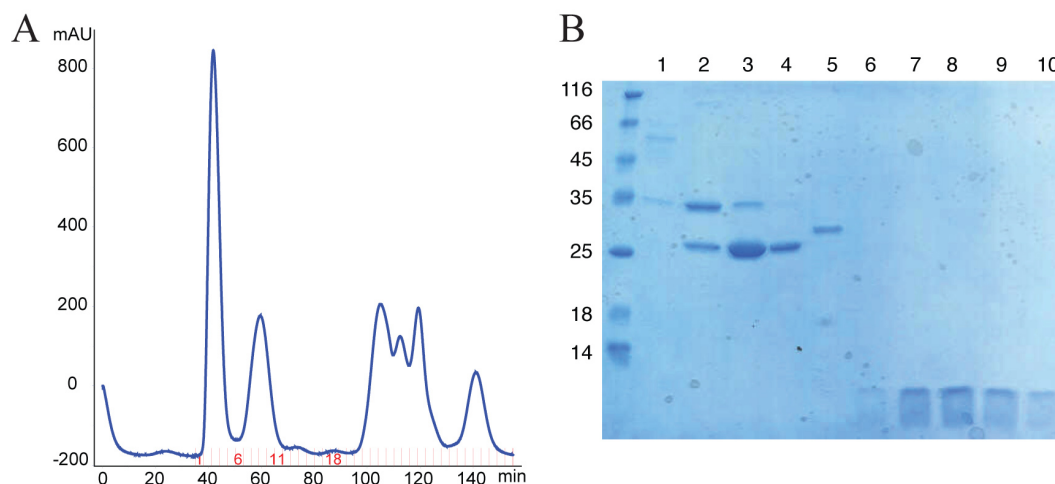
**Figure 10: Analytical expression of GST\_TBK1\_ULD<sub>s</sub> and GST\_TBK1\_ULDI.** The results of the analytical expression was analyzed by SDS-PAGE. The first line represents the non-induced GST\_TBK1\_ULD<sub>s</sub> after reaching  $OD_{600} \approx 1.0$ . For each construct, a sample was used after 1 to 6 hours (lines 1 to 6) induction with 1 mM IPTG at 25°C.

After the cell lysis, the TBK1\_ULD constructs presenting a GST tag were purified by affinity chromatography on a glutathione sepharose column (**Figure 11A**). The collected elution fraction was cleaved by thrombin protease. A cleavage screening was performed to adapt the amount of thrombin needed, the temperature, the duration of cleavage and the buffer conditions. Even by applying the best found conditions (3 units enzyme for 100 μg protein in GST elution buffer at pH 7.7 overnight at 18°C), the cleavage rate was not really efficient with the remaining presence of non-cleaved product in solution. More thrombin protease was needed than the advice of the manufacturers (**Figure 11B**).



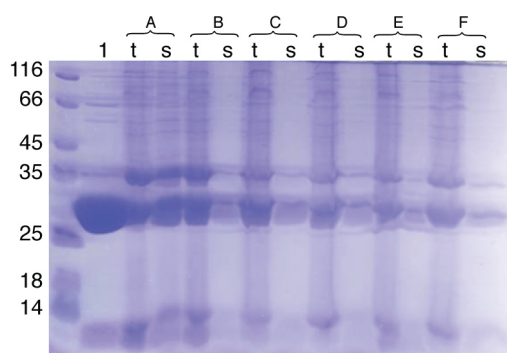
**Figure 11: Isolation and purification GST\_TBK1\_ULD<sub>s</sub> and GST\_TBK1\_ULDI.** The isolation of both constructs on a glutathione sepharose column is represented in A and the their analytical cleavage in B. In A, line 1 represents the cell lysate fraction, the line 2 the flow-through, the line 3 the washing step and the lines 4 and 5 the elution fractions. In B, the bracket 1 represents the control without any thrombin, the bracket 2 with thrombin for 1h, in the bracket 3 for 3 hours and in the bracket 4 overnight. Thrombin cleavage was performed under different pH conditions: pH = 6 in lines a, pH = 6.8 in lines b and pH = 7.7 in lines c.

The cleaved TBK1\_ULD<sub>s</sub> was separated from the GST tag using size-exclusion chromatography. Monitoring the absorbance at 280 nm on a Superdex S75 column, one small monodispersed peak eluted at approximately 90 mL, corresponding to a molecular mass of around 10 kDa following the calibration measurements, which matches with the size of TBK1\_ULD<sub>s</sub>. A broad peak eluting close to the dead volume contained uncleaved GST\_TBK1\_ULD<sub>s</sub> as well as cleaved GST (**Figure 12A**).



**Figure 12: Purification of GST\_TBK1\_ULD<sub>s</sub>.** The elution profile of GST\_ULD<sub>s</sub> purification by size exclusion chromatography with a Superdex S75 after thrombin cleavage is represented in A. The void volume elutes around 45 mL and corresponds to elution fractions 1 to 5. The cleaved GST is a dimer in solution with a molecular size of around 56 kDa eluting at 60 mL. The peak 18 corresponds to an elution around 90 mL, which is correlated to a molecular size of 10 kDa. The SDS-PAGE profile of the elution peaks in A is illustrated in B. The line 1 represents the first elution fraction corresponding to the void volume. The lines 2 to 5 and the lines 6 to 10 represent the elution fractions corresponding to the peaks eluting from 45 to 60 mL and around 80 mL, respectively.

SDS-PAGE confirmed that each peak eluting as the expected molecular mass corresponds to GST\_TBK1\_ULD<sub>s</sub>, GST or TBK1\_ULD<sub>s</sub>. The examination of the fraction elutions containing TBK1\_ULD<sub>s</sub> did not reveal the presence of any contaminants (**Figure 12B**). Pure samples were then concentrated to a volume of 300  $\mu$ L for NMR measurements. During the concentration steps, some precipitation was already observed using 1.5x gel filtration buffer (GFB) II buffer at pH 8.6. TBK1\_ULD<sub>s</sub> was concentrated in different buffer compositions and with different additives (arginine/glutamic acid, NaCl, MgCl<sub>2</sub>, CHAPS, glycerol, pH and phosphate) to reduce the appearance of aggregates in the samples. By comparing the amount of soluble TBK1\_ULD<sub>s</sub> present to the total amount of protein signal in a SDS-PAGE, the best conditions potentially suitable for NMR for TBK1\_ULD<sub>s</sub> were found to be 50 mM Tris, 50 mM arginine/glutamic acid at pH 7.2. The presence of other additives did not show any improvements in the final concentration of the protein (**Figure 13**).



**Figure 13: Optimization of buffer conditions for ULDS.** The line 1 is the reference conditions in 1.5x GFB II buffer at pH 8.6. In A, the pellet was resuspended with 50 mM Tris and 0.1% SDS, in B with 50 mM Tris, in C with 50 mM Tris and 50 mM arginine/glutamic acid, in D with 50 mM Tris and 500 mM NaCl, in E with 50 mM Tris and 10 mM MgSO<sub>4</sub> and in F with PBS buffer at pH 7.5. t represents the total fraction and the s the soluble fraction.

In parallel, mutations have been introduced in TBK1\_ULD<sub>s</sub>. None of these new constructs where residues from N- and C- terminal region were substituted (PF3,5EE; VL10,11AD; V14D; F18D; PH91,92EE; LRR88,90,93RQD) as well as residues close to the hydrophobic patch (Y29D; I30D; Y33D) improved the protein solubility. Precipitations after cleavage were even worse than for the wild type, some mutants did not let the GST-tag to be cleaved by the thrombin protease and some of the constructs did not even express.

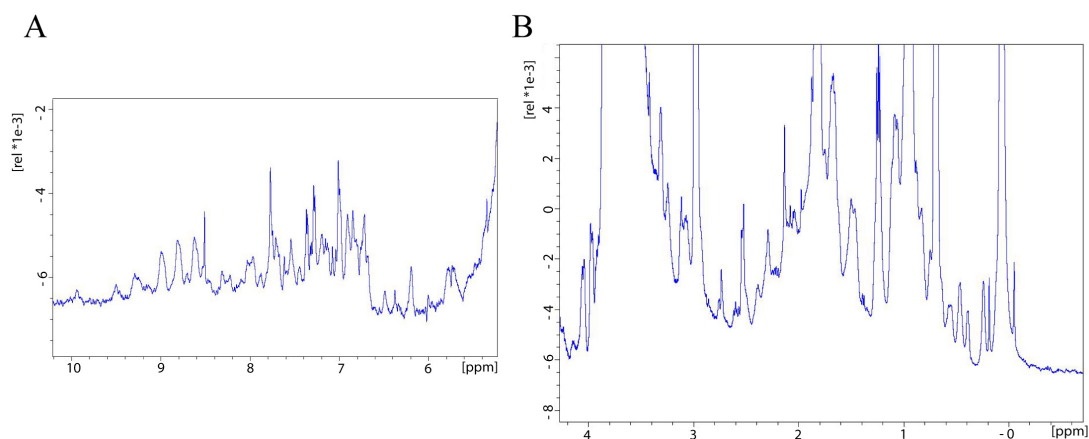
TBK1\_ULD<sub>s</sub> was also expressed under NusA expression vector but, even if an increase of the expression level was observed, NusA\_TBK1\_ULD was present almost exclusively in inclusion bodies. Extraction and resolubilisation tries did not allow to isolate and purify the target protein (results not shown).

Even though less than 1 mg of pure TBK1\_ULD<sub>s</sub> under GST expression per liter of expression were obtained, the yield provided enough material to perform NMR experiments.

## 6.1.2. Secondary structure determination

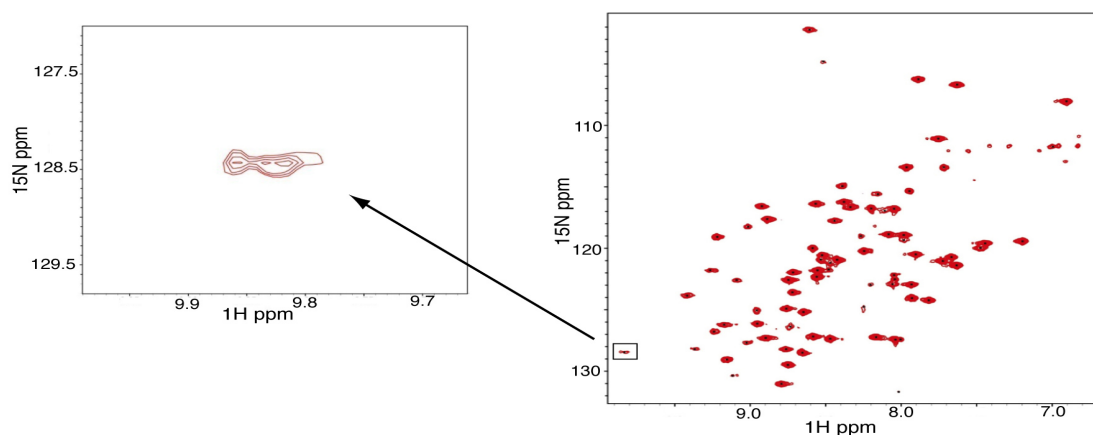
### 6.1.2.1. Optimization of NMR conditions

After the first expression and purification tries, a proton NMR ( $^1\text{H}$  NMR) spectrum of TBK1\_ULD was recorded in 1.5x GFB II buffer at pH 8.6 at 298K. Peaks in the low and high field regions were sharp and showed a good chemical shift dispersion indicating for a well-folded protein (**Figure 14**).



**Figure 14:**  $^1\text{H}$  NMR spectra of TBK1\_ULD. The good dispersion from the protons signals, both for amino (A) and methyl groups (B) of TBK1\_ULD, is an indication for a well-folded protein.

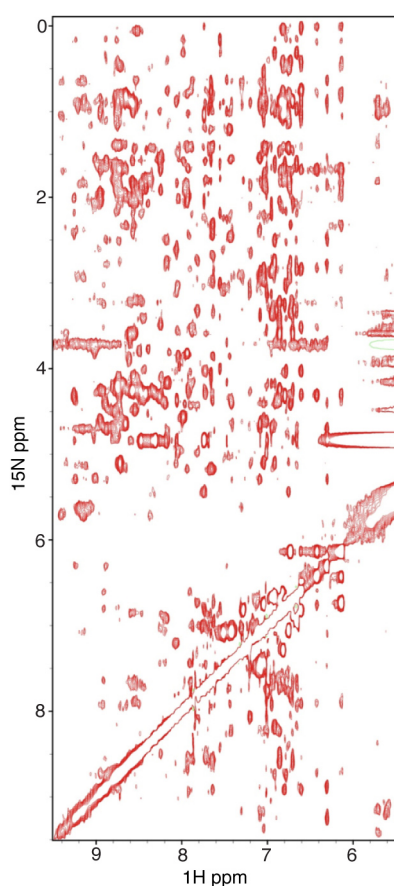
A first [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC was recorded under the same conditions. 76 well-dispersed peaks on an expected total of 94 were observed but with line broadening, several centers and different line-shape (**Figure 15**). Even though the protein is folded in these conditions, the shape of the peaks was not ideal to perform further NMR experiments. According to the screening conditions of TBK1\_ULD purification, the amount, the shape and the dispersion of TBK1\_ULD resonances in [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra were monitored under these different conditions.



**Figure 15:** [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC of TBK1\_ULD. The first NMR spectrum recorded for TBK1\_ULD revealed a nice peak dispersion illustrating a well-folded protein but the presence of broad peaks with several centers (F84 in the zoomed box) reveals some dynamics in TBK1\_ULD.



By adding salt (up to 400 mM NaCl) in the sample, ten more signals were then observable and some of the peaks presenting several centers disappeared but the peaks were still broad. Adding progressively CHAPS to the reference spectrum of TBK1\_ULD, although some peaks became sharper and a diminution of double peaks could be noticed, no real improvement was detectable (results not shown).



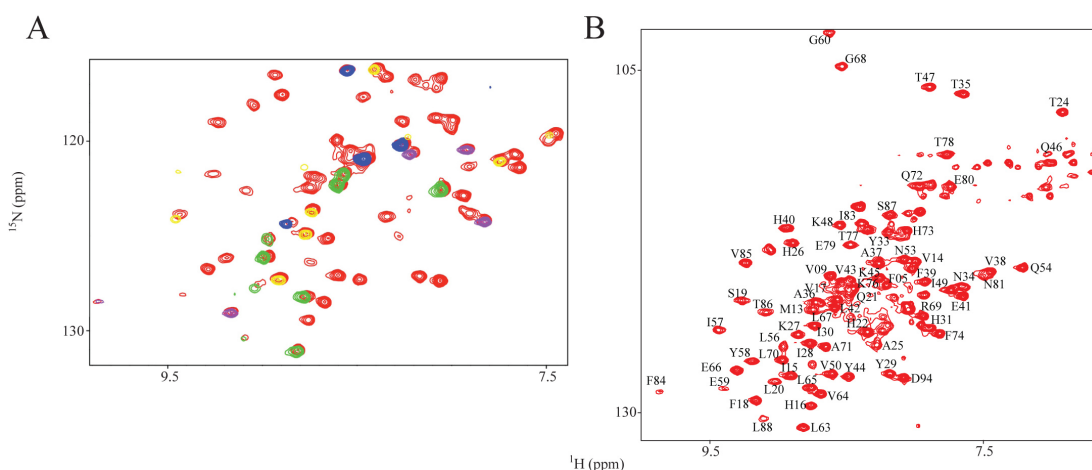
**Figure 16:** [ $^1\text{H}$ ,  $^1\text{H}$ ] NOESY spectrum of TBK1\_ULD. The presence of several NOESY cross-peaks confirmed that TBK1\_ULD in these conditions was a well-folded and globular protein.

After several attempts, the best conditions for TBK1\_ULD were established as 50 mM Tris, 50 mM arginine/glutamic acid at pH 7.2. The [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum of the isolated domain (amino acids 302 to 383) showed a chemical shift dispersion that is indicative of a well-folded and globular domain and the rich NOESY spectrum confirmed the folding of the protein (**Figure 16**). Unfortunately, no good “structural NMR” samples with at least a concentration of 0.5 mM and stable for more than 3 days could be produced. Due to these conditions, the secondary structure of the potential ULD domain of TBK1 has been later determined only with calculations based on the assignment of the chemical shift values of TBK1<sub>302-383</sub>.

### 6.1.2.2. Assignment

For the sequential assignment, HNCA, HNCACB,  $^{15}\text{N}$ -separated 3D-NOESY,  $^{15}\text{N}$ -separated 3D-TOCSY as well as homonuclear 2D-TOCSY and 2D-NOESY spectra were measured on Bruker Avance spectrometers operating at proton frequencies between 500 and 800 MHz. Using this standard approach, good peak connectivities were observed for sequential assignment but only around 40% of the resonances could be assigned at this point.

To complete the assignment, selective labeling was used. Different samples with labeled lysine, leucine, isoleucine or phenylalanine were prepared. For each selectively labeled amino acid, the expected number of peaks corresponding to the number of residues of this type in TBK1\_ULD was present in the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum except for one leucine. These cross-peaks overlaid with the cross-peaks present in the fully labeled sample (**Figure 17A**). With the help of selective labeling, most of the backbone assignment (90%) was achieved with this method (**Figure 17B**).

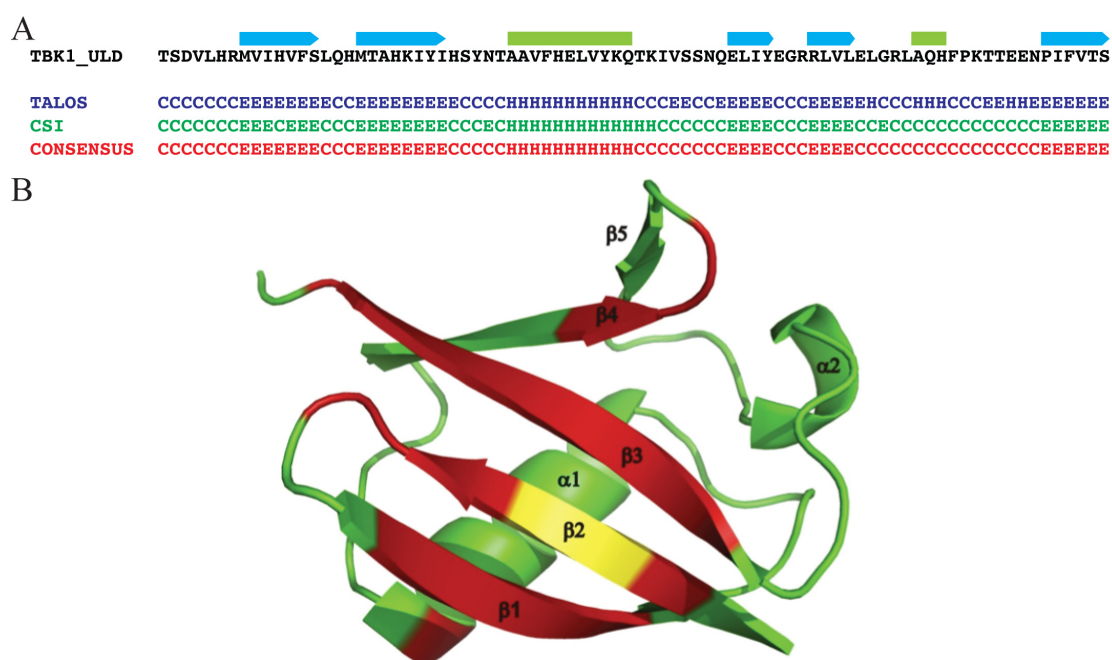


**Figure 17: Assignment of TBK1\_ULD.** The overlay of the representative region of [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of TBK1\_ULD using selective labeling with lysine (blue), leucine (green), isoleucine (yellow) and phenylalanine (purple) is shown in A. The representation of [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum of TBK1\_ULD with the full assignment of the amide resonances of TBK1\_ULD is represented in B.

### 6.1.2.3. Secondary structure calculation

Analysis of the secondary structure elements with the programs TALOS<sup>142</sup> and CSI,<sup>141</sup> based on the  $^{13}\text{C}^\alpha$ ,  $^{13}\text{C}^\beta$ ,  $^{13}\text{CO}$  and  $^1\text{H}^\alpha$  chemical shifts, identified one long  $\alpha$ -helix and five  $\beta$ -strands. As a consensus between the two predictions, the  $\alpha$ -helix contains the residues A332 to Q342, the first  $\beta$ -strand the residues M309 to S315, the second  $\beta$ -strand the residues M319 to I326, the third  $\beta$ -strand the residues E351 to Y354, the fourth  $\beta$ -strand the residues R358 to R361 and finally the fifth  $\beta$ -strand the residues P378 to L384. In addition to these main structural features, the chemical shifts indicated, only for TALOS prediction, a short helical segment between amino

acids A367 and H369 that corresponds to the  $3_{10}$ -helix of ubiquitin (**Figure 18A**). The secondary structure elements of TBK1\_ULD have the same sequential arrangement ( $\beta\beta\alpha\beta\beta$ ) and the same length as the corresponding secondary structure elements in ubiquitin proving that TBK1 belongs to the ULD superfamily. Despite these structural similarities between the ULD and ubiquitin, differences exist in their dynamic behavior. Interestingly, the whole  $\beta$ -sheet and in particular the two  $\beta$ -strands in TBK1\_ULD corresponding to  $\beta$ -strand 1 and  $\beta$ -strand 3 in ubiquitin exhibit broad resonances with several maxima, which is characteristic for conformational exchange processes. Being plotted on the ubiquitin three-dimensional structure, these residues form a well-defined surface (**Figure 18B**).

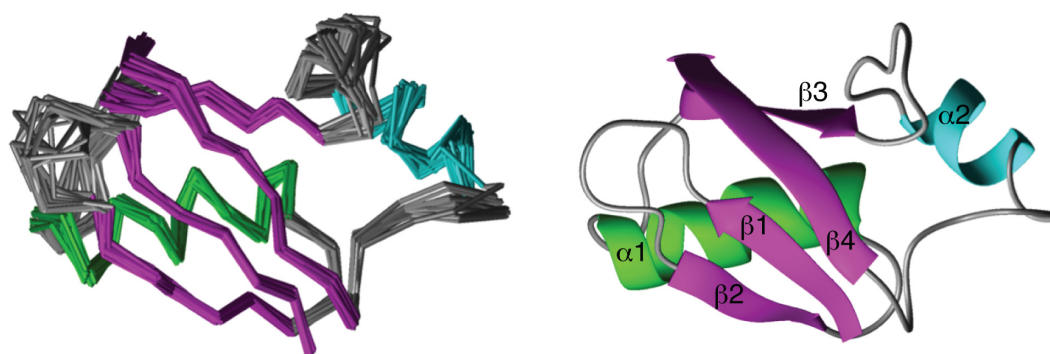


**Figure 18: Secondary structure of TBK1\_ULD.** The alignment of the secondary structure prediction from TALOS and CSI softwares (A) confirmed the ULD fold in TBK1 with the presence of the secondary structure motif  $\beta\beta\alpha\beta\beta$  followed by one short  $\alpha$ -helix and one  $\beta$ -strand. The residues of TBK1\_ULD presenting broad resonances in NMR spectra are plotted on the structure of ubiquitin depending of their intensity (peaks with several maxima in red, normal peak in green).

### 6.1.3. Tertiary structure determination

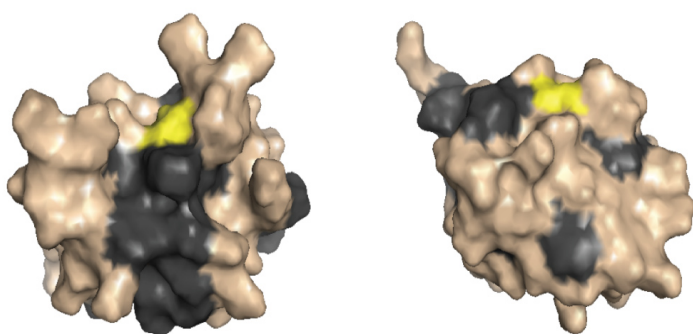
The low solubility and the tendency to aggregate (highest reachable concentration around 0.2 mM with only a stability of approximately 2 days) for all tested fragments did not allow to perform a complete assignment nor to calculate a high-quality 3D structure for TBK1\_ULD. Nevertheless, using proton assignments from the 3D  $^{15}\text{N}$ -NOESY and the 2D  $^1\text{H}$ -NOESY spectra, some side chains resonances were also assigned with the help of the topology similarity between ubiquitin and TBK1\_ULD. It was thus possible to calculate a preliminary structural

model of TBK1\_ULD with the use of the hydrogen bonds observed in the X-ray structure of ubiquitin (PDB: 1UBQ)<sup>144</sup> transposed to the sequence of TBK1\_ULD.



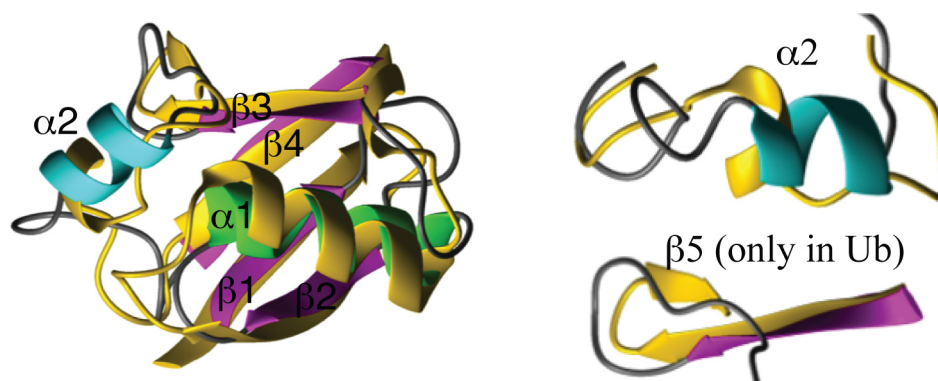
**Figure 19: NMR Structure of TBK1\_ULD.** Left: overlay of the backbone atoms of the 20 energy-refined conformers of TBK1\_ULD. Loops are shown in gray,  $\alpha$ -helices in green and cyan and  $\beta$ -strands in magenta. Right: the mean structure of the complex is presented as a ribbon diagram with annotation of the secondary structure elements.

The 3D structure of TBK1\_ULD confirmed the secondary structure elements predicted by CSI and TALOS: ULD of TBK1 adopts an ubiquitin like ( $\beta$ -grasp) fold. In the  $\beta\beta\alpha\beta\beta$  core, the four  $\beta$ -strands form a mixed  $\beta$ -sheet ordered  $\beta 2$ - $\beta 1$ - $\beta 5$ - $\beta 3$  where  $\beta 1$  (M309 to S315) and  $\beta 5$  (N377 to S383) are parallel but anti-parallel to  $\beta 2$  (M319 to I326) and  $\beta 3$  (E351 to Y354). The first  $\alpha$ -helix contains the residues from A332 to Q342 and connects  $\beta 2$  and  $\beta 3$  to form a right-handed  $\beta$ - $\alpha$ - $\beta$  unit. The second  $\alpha$ -helix, longer than the one predicted by TALOS and CSI, starting from L363 and finishing with H369 is present outside of the ubiquitin core (**Figure 19** and structural statistics in Appendix 9.5). Finally, around I353 is concentrated several residues with hydrophobic side groups forming a large hydrophobic patch on the surface of TBK1\_ULD (**Figure 20**).



**Figure 20: NMR structure of TBK1\_ULD and ubiquitin.** The surface representation shows TBK1\_ULD (left) and ubiquitin in (right). Hydrophobic residues (leucine, isoleucine, valine, phenylalanine and tyrosine) are in black, other residues in wheat. I353 in TBK1\_ULD and I44 in ubiquitin, center of the hydrophobic patch in both proteins are colored in yellow.

The superposition of TBK1\_ULD structure with the structure of ubiquitin showed strong similarities with a global RMSD of 3 Å, reduced to 1.5 Å when only the structured region ( $\alpha$ -helices and  $\beta$ -strands) are taken in account. Some differences are still observed like the fourth  $\beta$ -strand predicted (R358 to R361), which is not structured like it is in the case of ubiquitin (**Figure 21**).



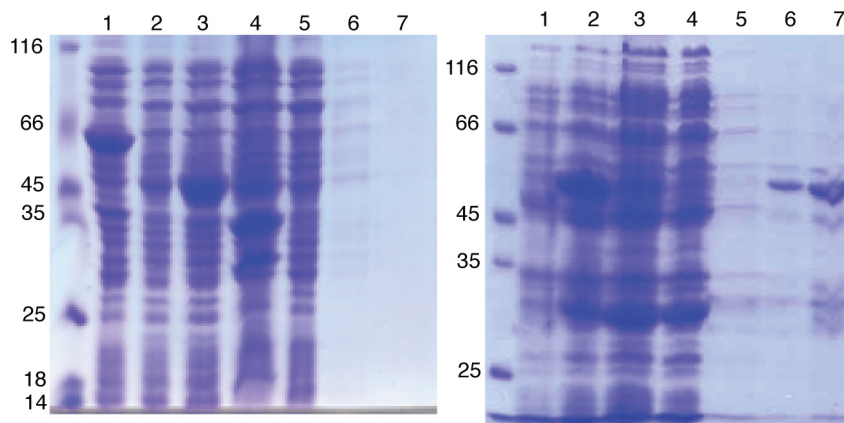
**Figure 21: Overlay of the NMR structures of TBK1\_ULD and ubiquitin.** TBK1\_ULD is represented in the same colors than in Fig. 18, ubiquitin is in gold. Left: Overlay of the full structures. Right: Zoom on  $\beta$ -strand 4 of TBK1\_ULD, equivalent to  $\beta$ -strand 5 in ubiquitin in (top) and on  $\alpha$ -helix 2 in (bottom).

#### 6.1.4. Expression and purification of IRF3

Three different constructs under the GST expression system were provided by Dr. Fumiyo Ikeda to confirm by NMR the interaction of the interferon regulatory factor association domain (IAD) of interferon regulatory factor 3 (IRF3) with TBK1\_ULD by NMR. The first construct, GST\_IAD, contained strictly the IAD domain (residues 190 to 384). The second construct, GST\_IAD-SRR, is an extension of IAD until the C-terminal of IRF3 (residues 190 to 427) including the serine rich region (SRR). The last construct, GST\_IRF3, is the full IRF3 (residues 1 to 427). First attempts were performed on GST\_IAD, which showed nice expression after 3 hours induction at 37°C with 1 mM IPTG. A fast degradation was already observed by SDS-PAGE with the appearance of multiple smaller bands in the cell lysate fraction in comparison to the expression fraction. No protein were detected in the elution fraction from the GST purification (**Figure 22**).

The second tries were done with the GST\_IAD-SRR construct. Here, the expression yield was still good even though lower than for GST\_IAD under the same conditions. In this case, it was possible to detect GST\_IAD-SRR in the elution fractions (**Figure 22**). To remove the GST moiety, analytical cleavage of GST\_IAD-SRR with thrombin protease was performed. The cleavage rate was very low even by increasing amounts of thrombin protease, changing the temperature or increasing the time of incubation. Moreover, because of the similar size of GST and IAD-SRR ( $\approx 26$  kDa), cleaved IAD-SRR was not distinguishable from the GST moiety by

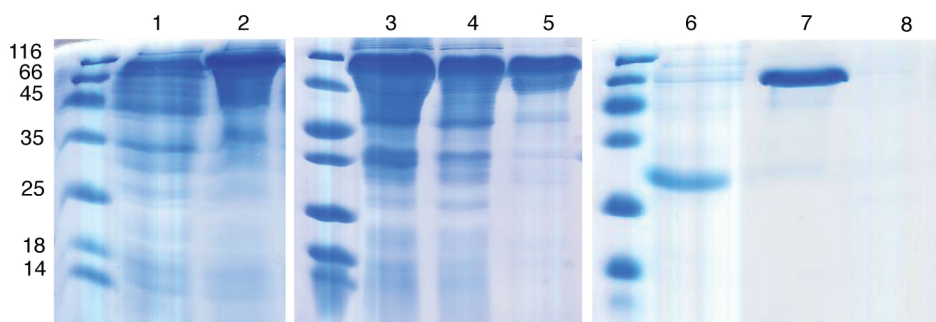
SDS-PAGE. GST is a dimer in solution so size-exclusion chromatography should separate the cleaved IAD-SRR from the GST moiety was loaded. Gel filtration results were not clear but the elution fractions that should correspond to pure IAD-SRR were collected and concentrated for NMR titration experiments (results not shown).



**Figure 22: Isolation of GST\_IAD and GST\_IAD-SRR.** For the isolation of GST\_IAD on a glutathione sepharose column (left), the line 2 represents the non-induced GST\_IAD, the line 3 the final expressed protein, the line 4 the cell lysate, the line 5 the flow-through and the lines 6 and 7 the elution fractions (The line 1 represents tetra ubiquitin, not related to GST\_IAD purification). For the isolation of GST\_IAD-SRR (right), the line 1 represents the non-induced GST\_IAD-SRR, the line 2 the final expressed protein, the line 3 the cell lysate, the line 4 the flow-through, line 5 the washing step and the lines 6 and 7 the elution fractions.

To overcome the problem of separating GST from IAD-SRR, a NusA expression vector was used. The expression yield was sufficient at 37°C for 3 hours with 1 mM IPTG but isolation on Ni-NTA column showed that half of NusA\_IAD-SRR did not bind to the nickel matrix and was found in the flow-through. Cleavage with TEV protease of both elution and flow-through fractions from the Ni-NTA purification provided IAD-SRR products. After purification by size-exclusion chromatography, IAD-SRR was found in the void volume and no protein was found at the elution time expected for a protein of around 25 kDa (**Figure 23**). The attempts to concentrate anyway the eluted IAD-SRR failed with the fast appearance of precipitation.

Because of the unsuccessful purification with the two previous constructs, GST\_IRF3 was also used even though the size of IRF3 could be problematic for NMR titration experiments. Using the same expression protocol as for GST\_IAD-SRR, GST\_IRF3 expression yield was lower and some proteolytic cleavage was already observable during the first purification steps (results not shown). These results were not favorable to continue with this construct in addition to the size of IRF3, which could be a disadvantage for further studies by NMR.

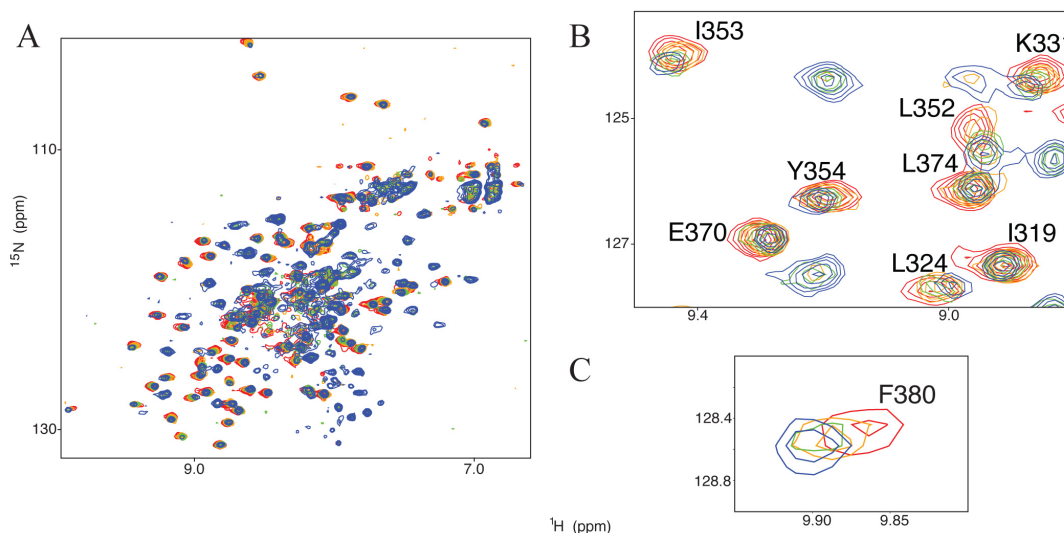


**Figure 23: Isolation and purification of NusA\_IAD-SRR.** NusA\_IAD-SRR was isolated on a Ni-NTA column. Line 1 represents the non-induced NusA\_IAD-SRR, the line 2 the final expressed protein, the line 3 the cell lysate, the line 4 the flow-through, and the line 5 the elution fractions. After cleavage by TEV protease, the protein was purified by size-exclusion chromatography. The line 6 represents the void volume, the line 7 the fraction eluting at 42 mL (75 kDa) and the line 8 the fraction eluting at 80 mL (25 kDa).

Finally, in order to get some complementary information on TBK1\_ULD interaction with IRF3, an uncleaved GST\_IAD-SRR sample was purified for further experiments. Even though the purity was not total, the sample was used anyway for primary information.

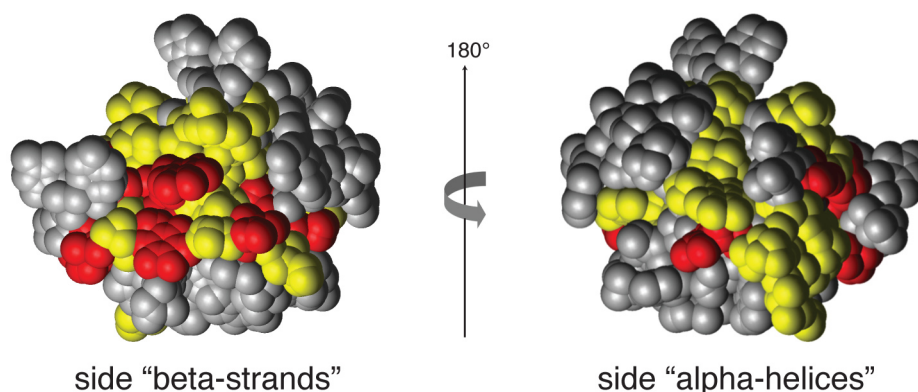
#### 6.1.5. Interaction studies of TBK1\_ULD/IRF3\_IAD-SRR

First NMR titration experiments were performed using uniformly  $^{15}\text{N}$  labeled TBK1\_ULD and unlabeled IAD-SRR. Due to purification issues and fast degradation of IAD-SRR, it was only possible to study this interaction with few titration steps. After addition of IAD-SRR to  $^{15}\text{N}$ \_TBK1\_ULD at molar ratio of 1:1/2, 1:1 and 1:2 (TBK1\_ULD:IAD-SRR), two main effects were observed upon the titration. First effect, several residues were strongly affected by chemical shift perturbations (CSP) indicating intermolecular interactions in a fast exchange mode. This effect was well observed for residues L352 and Y354 that are present in the hydrophobic patch of TBK1\_ULD but also for residues L352 and F380. Second effect, the appearance of new peaks in the spectra caused by the possible formation of a tight complex between the two proteins. No experiments for the assignment of TBK1\_ULD in presence of IAD-SRR were done, so it was impossible to know to which residues correspond the new resonances. Nevertheless, resonances corresponding to residues K331 and I353 showed a decreased of the intensity. In the meantime, new peaks appeared but not close to the peaks disappearing during titration. The number of appearing peaks being higher than the disappearing ones, it seems that several TBK1\_ULD residues are present in solution in two states, free and bound to IAD-SRR (**Figure 24**).



**Figure 24: Interaction of TBK1\_ULD with IRF3\_IAD-SRR.** The overlay of representative regions of [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra from TBK1\_ULD in presence of IRF3\_IAD-SRR with different molar ratio, 1:0 (red), 1:1/2 (yellow), 1:1 (green) and 1:2 (blue) is shown. While in (A) the full spectrum is illustrated, (B) and (C) represent different regions where CSP as well as disappearance/reappearance of peaks are observed.

Being plotted on the TBK1\_ULD molecule, those perturbations appeared on both side of the molecule. Even though the residues affected the strongest are more preferentially situated on the TBK1\_ULD side where the  $\beta$ -strands are localized, several residues from the  $\alpha$ -helices present also strong CSP upon titration with IAD-SRR of IRF3 (**Figure 25**). The TBK1\_ULD/IRF3 interaction seems therefore to be a more complex process than a 1 to 1 binding model.

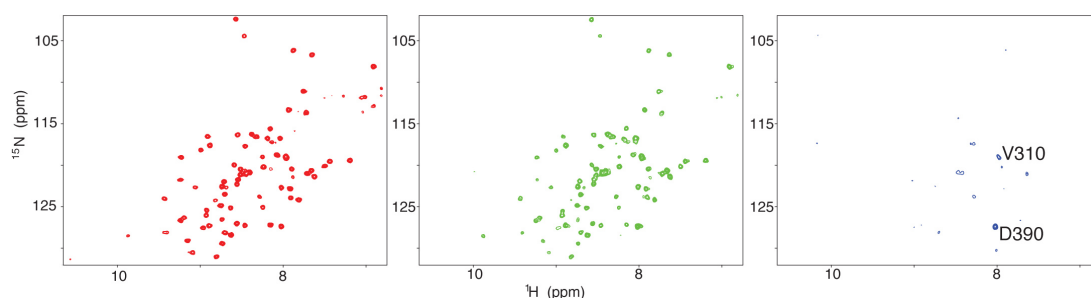


**Figure 25: Interaction surface of IAD-SRR on TBK1\_ULD.** The residues of TBK1\_ULD are represented under different colors depending on the strength of their perturbation due to the presence of IAD-SRR. In red are represented the residues presenting two-states or strong chemical shift perturbations, in yellow residues with small perturbations and in grey residues not affected by the presence of IAD-SRR.

To monitor this interaction, a second experiment was performed by titrating uncleaved GST\_IAD-SRR against TBK1\_ULD in order to overcome the issue of IAD-SRR purification. Almost every peak disappeared or decreased significantly during the titration due to the slow



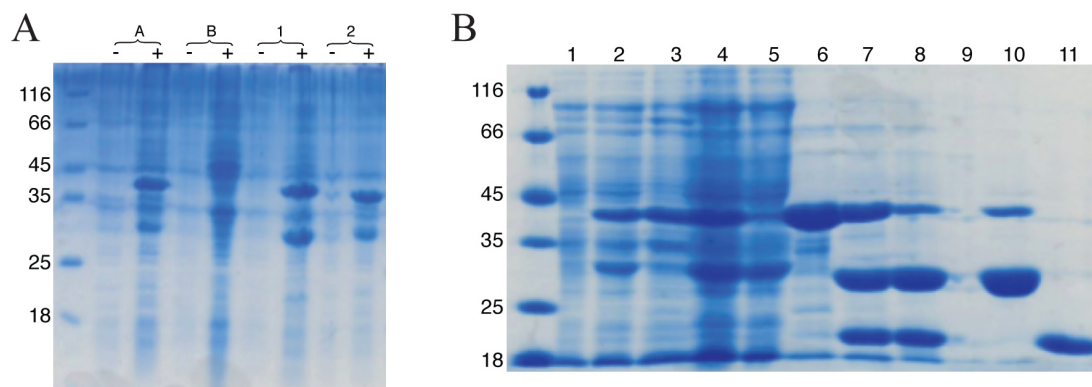
tumbling effect of the big complex formed by GST\_IAD-SRR/TBK1\_ULD. V310 and D390, present at the N- and C-terminal regions of TBK1\_ULD, were the only visible residues at the end of the titration, because they could have a faster tumbling effect than the rest of the complex (Figure 26).



**Figure 26: NMR titration of GST\_IAD-SRR against TBK1\_ULD.** The reference spectrum of TBK1\_ULD (red) shows a nice peak dispersion for all residues of the protein. In presence of GST\_IAD-SRR at a molar ratio 1:1 (green), most of the resonance peaks are still present but with a decrease of intensity coupled to small chemical shift perturbations. At a molar 1:2, most of the peaks disappeared. Only V310 and D390 are still visible.

## 6.2. Autophagy project

### 6.2.1. Expression and purification of MAP1LC3 proteins

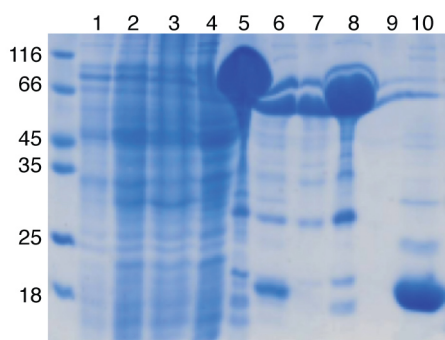


**Figure 27: Expression and purification of GST\_MAP1LC3 proteins.** The expression of the different MAP1LC3 proteins fused to GST are represented in (A). The (-) and (+) symbols represent the non-induced and the final expression state for each protein, respectively. The lines A, B, 1 and 2 correspond to LC3A, LC3B, GABARAPL-1 and GABARAPL-2 proteins, respectively. The expression and the purification profiles of GST\_GABARAPL-1 are shown in (B). The line 1 represents the non induced protein, the line 2 the final expression, the line 3 the pellet of the cell lysate, the line 4 the cell lysate, the line 5 the flow-through of the glutathione sepharose column, the line 6 the elution fraction, the line 7 the elution fraction with thrombin protease after 16 hours and the line 8 after 64 hours. After 64 hours cleavage action of the thrombin protease, the elution fraction was purified by size-exclusion chromatography on a Superdex 75 column. The line 9 represents the elution corresponding to the void volume, the line 10 the elution fraction after 55 mL and the line 11 the elution fraction after 75 mL.

The different sequences provided by Dr. Vladimir Kirkin and Dr. David McEwan were cloned into pGEX-4T1 expression vectors with a deletion of the C-terminus including the last

glycine residue to avoid any possible conjugation (sequences listed in Appendix 9.1). All constructs (GST-GABARAPL-1, GST-GABARAPL-2, GST-LC3A and GST-LC3B) could be expressed nicely upon screening conditions where the best temperature, the IPTG concentration and the duration of expression were tested. The ideal conditions found were 3 hours of expression at 37°C with 1 mM IPTG (**Figure 27A**). GST-fusion proteins were isolated by glutathione affinity chromatography, cleaved by thrombin protease and purified through size-exclusion chromatography following the same protocols as described for the TBK1\_ULD project. Even though the cleavage was more successful than for the previous project, a large amount of uncleaved protein was still present in solution. The use of more thrombin protease for proteins expressed on a large scale was not the best economic solution (**Figure 27B**).

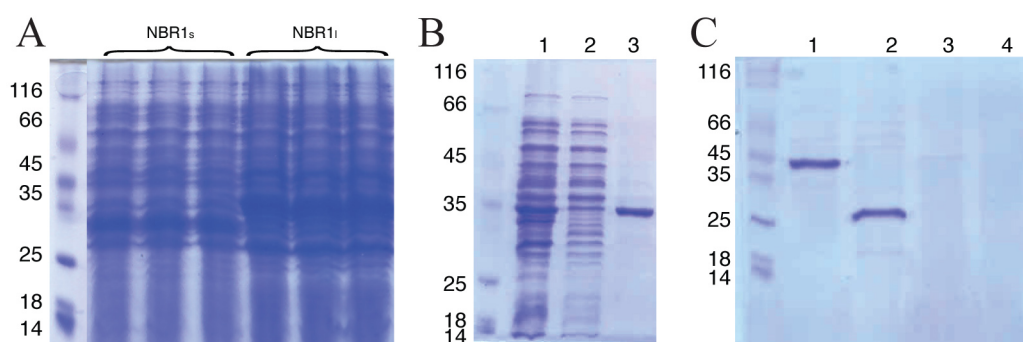
DNA fragments coding for the different MAP1LC3 proteins were cloned into the expression vector pETM-60. This vector presents the NusA protein to enhance the expression of the His<sub>6</sub>-tagged recombinant proteins fused to the C-terminus of NusA through a TEV protease cleavage site recognition sequence. The analytical expression of these new constructs presented higher expression yield than for the GST constructs under the same conditions, 3 hours at 37°C with 1 mM IPTG (**Figure 28**). After expression, cells were harvested and the resuspended pellet was lysed. The cell lysate was then loaded onto a Ni-NTA affinity column where the NusA constructs bind through the His<sub>6</sub>-tag prior to the TEV cleavage site. The flow-through did not contain any of the proteins of interest. The elution fractions were collected and cleaved by TEV protease using standard protocols. Cleavage rate with TEV protease was maximal and pure proteins in a high yield were more easily obtained than in case of GST based constructs. MAP1LC3 proteins were then separated from the NusA tag and the uncleaved protein by ion-exchange chromatography followed by size-exclusion chromatography (**Figure 28**). Each MAP1LC3 protein was eluted monodispersly around 80 mL on a Superdex S75 column corresponding to a molecular size of approximately 14 kDa, the average molecular size of these proteins. MAP1LC3 proteins were also obtained from Ub-constructs with the procedure described in chapter 6.3.2.



**Figure 28: Expression and purification of NusA\_LC3B.** The line 1 represents the non induced protein, the line 2 the final expression, the line 3 the cell lysate, the line 4 the flow-through of the NiNTA purification, the line 5 the elution fraction, the line 6 the elution fraction after action of TEV protease, the lines 7, 8 and 9 the flow-through of the IEC after action of TEV protease and the line 10 the elution of the IEC.

Based on the protocols used by different groups to determine the structures of some Atg8 and MAP1LC3 proteins by NMR, a 25 mM sodium phosphate and 100 mM NaCl buffer solution at pH 7.3 was first used. The purity of the size-exclusion chromatography elution fractions was checked upon SDS-PAGE and samples could be concentrated for NMR and ITC applications until 1 mM (40 mg for 1L expression of protein fused to NusA) without noticing any precipitation. The first NMR samples showed nicely shaped resonance peaks in a [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum and a similar dispersion pattern compared to the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra published in the biological magnetic resonance bank (BMRB) database for the corresponding proteins. By performing some slight adjustments in the buffer conditions (50 mM sodium phosphate and 100 mM NaCl at pH 7.0), the broadening of the peaks was reduced. These new conditions were chosen as a reference buffer for NMR and ITC.

### 6.2.2. Expression and purification of peptides containing LIR motifs



**Figure 29: Expression and isolation of NBR1-LIR constructs.** The analytical expression of NBR1-LIR is shown in (A). Both constructs were expressed under a GST expression vector for 3 hours at 37°C after induction with 1 mM IPTG. The isolation profile of GST\_NBR1<sub>l</sub> is shown in (B). The line 1 represents the cell lysate, the line 2 the flow-through and the line 3 the elution fraction of GST\_NBR1<sub>l</sub> on a glutathione sepharose column. The cleavage of GST\_NBR1<sub>l</sub> is shown in (C). The line 1 represents the elution fraction of GST\_NBR1<sub>l</sub> on a glutathione sepharose column and the line 2 this elution fraction after cleavage with thrombin protease.

Plasmids coding for peptides containing LIR motifs were cloned under pGEX-4T1 expression vectors. Dr. Vladimir Kirkin provided generously constructs containing the coding sequence of NBR1-LIR protein with different lengths varying from 46 to 276 amino acids (NBR1<sub>short</sub> (residues 718-763), NBR1<sub>long</sub> (residues 718-802), NBR1<sub>874</sub> (residues 691-874) and NBR1<sub>966</sub> (residues 691-966)). The two shorter versions of NBR1, GST\_NBR1<sub>short</sub> and GST\_NBR1<sub>long</sub>, expressed nicely in a soluble form as well at 37°C as at 22°C (**Figure 29**). After purification by glutathione affinity chromatography, first attempts showed a good cleavage rate by thrombin protease but only the GST moiety was observable and no peptide. Because of their small sizes, NBR1<sub>short</sub> and NBR1<sub>long</sub> peptides were not seen on SDS-PAGE (**Figure 29**) even with a

SDS-PAGE containing 18% polyacrylamide. Even passed through size-exclusion chromatography, no reasonable peaks were detected at 280 nm. The bigger constructs, NBR1<sub>874</sub> and NBR1<sub>966</sub>, expressed also nicely under the same conditions but the cleavage was not total and proteins were not easily detectable (results not shown). Nevertheless, the uncleaved constructs were very stable in solution. GST\_NBR1<sub>874</sub> and GST\_NBR1<sub>966</sub> being potentially too big for further NMR interaction studies, these constructs were purified for Dr. Vladimir Kirkin for his first interaction tries using Surface Plasmon Resonance.

In the meantime, the project using ubiquitin as an expression vector was initiated. Using peptides containing LIR motif were a good target for a proof of concept. Thus NBR1-LIR wild type and NBR1-LIR\_Y732W peptides produced from Ub constructs were also used for structural and interaction studies. These expression and purification protocols are precisely described in chapter 6.3.3.

Finally for the last titration experiments, mutated NBR1-LIR peptides (NBR1-LIR\_Y732F, NBR1-LIR\_S729E and NBR1-LIR\_S728,729E) were used as chemically synthesized peptides for time reason (peptidic sequences are listed in the Appendix 9.2).

### 6.2.3. Interaction Atg8/LIR

#### 6.2.3.1. LC3B vs. LIRs

##### 6.2.3.1.1. LC3B vs. p62-LIR

LC3B vs LIR	$\Delta H$ kcal mol <sup>-1</sup>	$\Delta S$ cal mol <sup>-1</sup> K <sup>-1</sup>	$\Delta G$ kcal mol <sup>-1</sup>	$K_d$ μM
p62	-10.5	-8.7	-8.0	1.5
Nix_W36	-2.7	+9.4	-5.5	91
Nix_W140/144	-2.8	+5.0	-4.3	670
NBR1	-4.4	+10.7	-7.6	2.9

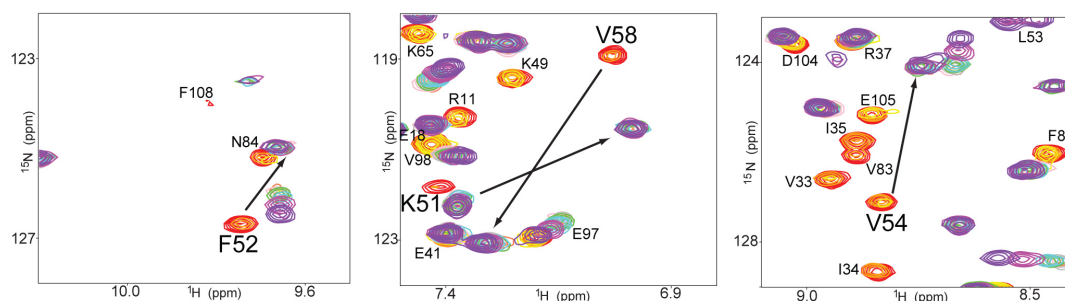
**Table 4: Thermodynamic parameters obtained by ITC for the interaction of LC3B with different LIR domains.**

All experiments were performed at 25°C.  $\Delta H$ ,  $\Delta S$  and  $K_d$  values were measured with the assumption of a one site model. Statistical errors for all calculations were quite small (below 3%) and the uncertainty of the sample concentrations was approximately 5% (or more in case of peptides containing no tryptophan or tyrosine residues). Therefore, a total error of 10% for all parameters could be expected.

The interaction of LC3B with p62-LIR was already well characterized on a structural level in the literature showing a strong binding<sup>87; 103</sup> and was consequently used in this project as reference for the interactions of MAP1LC3 proteins with the LIR motifs of autophagy receptors.

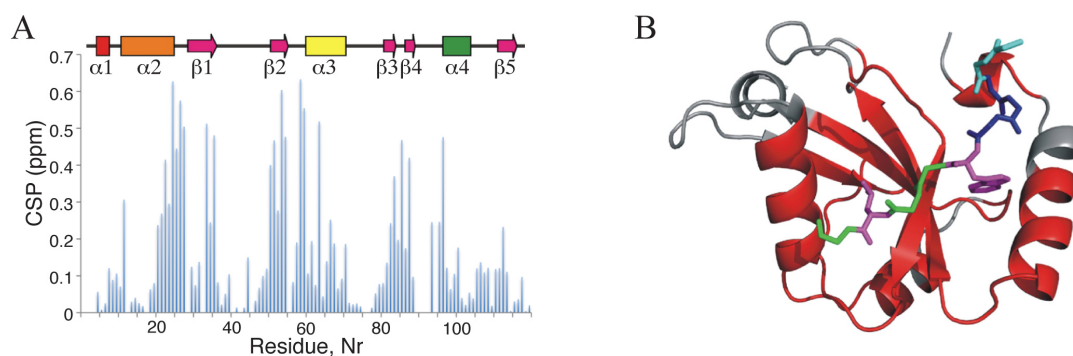
ITC experiments revealed a strong binding affinity for the LC3B/p62-LIR interaction with a  $K_d$  of 1.5  $\mu\text{M}$ . Whereas the strong negative enthalpy ( $\Delta H = -10.5 \text{ kcal mol}^{-1}$ ) plays in favor of binding, the contribution of negative entropy ( $\Delta S = -8.7 \text{ cal mol}^{-1} \text{ K}^{-1}$ ) is unfavorable, resulting in a free energy  $\Delta G$  of  $-7.9 \text{ kcal mol}^{-1}$  (**Table 4** and raw data in **9.3**).

By NMR, p62-LIR titration against LC3B revealed a typical strong interaction in a slow exchange mode, where the two peaks corresponding to the “free” and “bound” states could be observed for each HN resonance of LC3B residues affected by the presence of p62-LIR. The intensity of each peak is proportional to the population of free LC3B and LC3B bound to the p62-LIR peptide in solution (**Figure 30**).



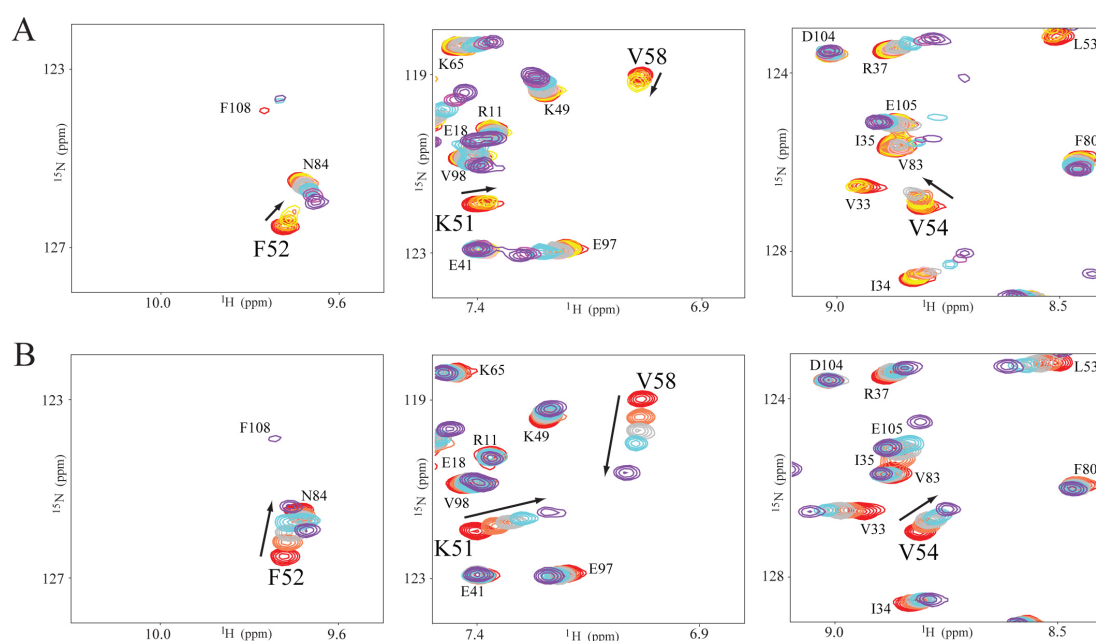
**Figure 30: LC3B interaction with p62-LIR.** The overlay of representative regions (left for F52, middle for K51 and V58 and right for V54) in  $[\text{}^{15}\text{N}, \text{}^1\text{H}]$  TROSY-HSQC spectra of LC3B in presence of p62-LIR is represented. Contours are presented in different colors corresponding to individual titration points (red as reference with LC3B alone and in presence of p62-LIR at different molar ratio: orange 1:1/8, yellow at 1:1/4, coral at 1:1/2, pink at 1:3/4, gray at 1:1, cyan at 1:2, magenta at 1:5 and purple at 1:10).

Although almost every peak is affected in this two-state transition, the CSP are more important (up to 0.6 ppm) concerning residues involved in the formation of hp1 for L53, of hp2 for V58 and L63 or close to these pockets for R24, Q26, H27, V33, F52, V54 and N59 (**Figure 31A**). Because of this two-state transition for almost all resonances, no dissociation constant could be calculated based on the chemical shift dispersion. Nevertheless, due to the slow exchange mode characterized during this interaction, a dissociation constant in the low micromolar range is expected. Being mapped on the LC3B structure, the resonances with meaningful CSP form not a localized surface but the whole volume of the molecule, indicating specific differences in structural or mobility properties between free LC3B and LC3B in complex with p62-LIR (**Figure 31B**).



**Figure 31: Interaction of LC3B with p62-LIR.** The CSP of LC3B in presence of p62-LIR are shown in (A) and are plotted on the structure of LC3B (PDB: 2K6Q) in (B). The residues presenting a double-state transition are colored in red and the residues, which did not show CSP or very small ones are in grey.

### 6.2.3.1.2. LC3B vs. Nix-LIRs



**Figure 32: LC3B interaction with Nix-LIRs.** The overlay of representative regions (left for F52, middle for K51 and V58 and right for V54) in  $^{15}\text{N}$ ,  $^1\text{H}$  TROSY-HSQC spectra of LC3B in presence of different Nix-LIR\_W36 (A) and Nix-LIR\_W140/144 (B) is represented. Contours are presented in different colors corresponding to individual titration points (red as reference with LC3B alone and in presence of LIRs at different molar ratio: orange 1:1/8, yellow at 1:1/4, coral at 1:1/2, pink at 1:3/4, gray at 1:1, cyan 1:2, magenta at 1:5 and purple at 1:10). For Nix-LIR\_W140/144 against LC3B, only the experiments representing the molar ratio 1:0, 1:1/2, 1:1, 1:2 and 1:10 were performed and for NBR1-LIR, the last titration point is at molar ratio 1:5.

In Nix, two potential LIR domains have been identified, around W36 and around W140/144. Both Nix-LIR domains have a significantly weaker interaction with LC3B than p62-LIR ( $K_d$  of  $91\ \mu\text{M}$  for Nix-LIR\_W36 and around  $700\ \mu\text{M}$  for Nix-LIR\_W140/144). This weaker interaction is also reflected by a less negative binding enthalpy ( $\Delta H = -2.7\ \text{kcal mol}^{-1}$  for Nix-LIR\_W36 and  $\Delta H = -2.8\ \text{kcal mol}^{-1}$  for Nix-LIR\_W140/144) but the positive entropy ( $\Delta S = 9.4\ \text{cal mol}^{-1}\cdot\text{K}^{-1}$  for Nix-LIR\_W36 and  $\Delta S = 5.0\ \text{cal mol}^{-1}\cdot\text{K}^{-1}$  for Nix-LIR\_W140/144) contributes to the binding.

The free energy for both interactions is similar ( $\Delta G = -5.5 \text{ kcal mol}^{-1}$  for Nix-LIR\_W36 and  $\Delta G = -4.3 \text{ kcal mol}^{-1}$  for Nix-LIR\_W140/144) but less negative than for LC3B/p62 (**Table 4** and raw data in Appendix 9.3).

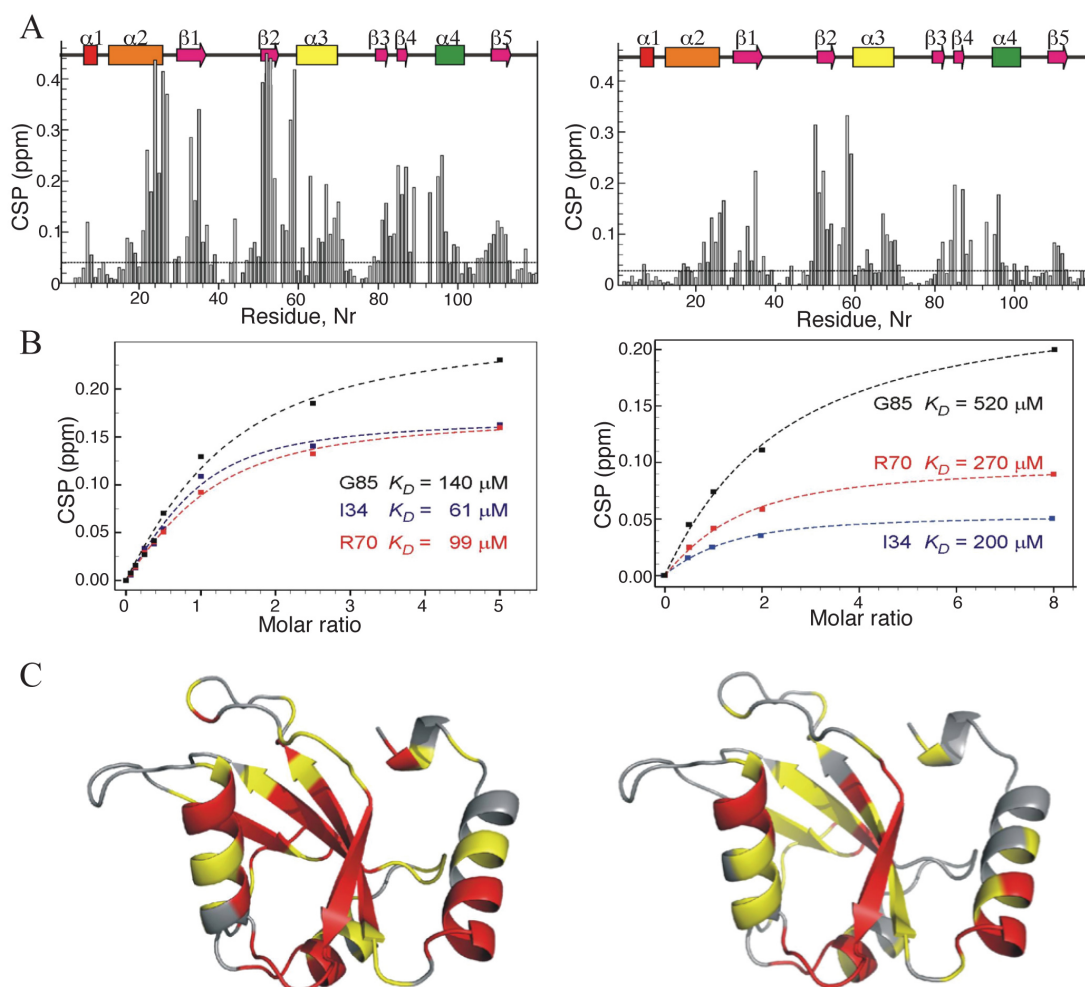
The NMR titration of LC3B with the Nix-LIR\_W140/144 peptide showed an interaction in a fast exchange mode. For each titration step, sharp peaks were observed for each residue resonance, shifted from its previous position with almost no decrease of the peak intensity until reaching almost a complete saturation. The overall CSP in this case are smaller, indicating a weaker binding of this peptide to LC3B. The residues from LC3B that are the most affected are the ones present in  $\alpha$ -helix 2 (R24, Q26 and H27), in  $\beta$ -strand 1 (I35), in  $\beta$ -strand 2 (K51, F52, L53 and V54) and in  $\alpha$ -helix 3 (I67) (**Figure 32A**).

The NMR titration of LC3B with Nix-LIR\_W36 peptide represented a fast (close to intermediate) exchange mode. Here the affected peaks are shifted with a decrease of the peak intensity at the initial stages of the titration and a following increase near the complete saturation. The residues showing the strongest CSP are localized in the structured region of LC3B like I34 and R37 from  $\beta$ -strand 1, R69 and R70 from  $\alpha$ -helix 3, L82 and G85 from  $\beta$ -strand 3 or I95 and E97 from  $\alpha$ -helix 4. For some of the peaks, after decrease of the peak intensity, no reappearance of the peaks was noticed in the neighborhood even for a ten times molar excess of the peptide in molar ratio. This effect was seen for the residues belonging to  $\alpha$ -helix 2 (R24, E25, Q26, H27 and T29), to  $\beta$ -strand 1 (V33 and I35), to  $\beta$ -strand 2 (T50, K51, F52, L53 and V54) and to  $\alpha$ -helix 3 (L63, I64, I66, I67 and L71), which were also the residues the most affected in the titration of Nix-LIR\_W140/144 against LC3B (**Figure 32B**). The values of the CSP are in between the ones for Nix-LIR\_W140/144 and p62-LIR titration, however, the affected residues are generally the same.

Similar to the LC3B/p62-LIR interaction, the CSP induced by both Nix-LIR\_W36 and Nix-LIR\_W140/144 peptides envelop the whole LC3B molecule. LC3B residues participating in the formation of the intermolecular  $\beta$ -structure in the LC3B/LIR complex (K51, L53) and directly in contact with the two hydrophobic moieties of the LIR motif (I23, I34, K51, F52 and L53 for W340; V33, I35, L63, I67 and R70 for L343) are strongly affected. Additionally, residues forming the barrier between the two hydrophobic pockets (those interacting with T341 and H342 in p62-LIR) show strong CSP (H27, K49, R69 and R70). However, there are some residues shown to be in very close contacts with the hydrophobic motif of LIR but presenting rather small or not meaningful CSP values (D19, V20, I66 and F108). From the other hand, the biggest CSP values were observed for residues, which are not in direct contact with the LIR

## Results

moiety. Certain structural elements of LC3B out of hp1 and hp2, like L22, R24, E25 and Q26 in the  $\alpha$ 2-helix; V58 and N59 in the loop between strand  $\beta$ 2 and  $\alpha$ 3; residues 83-87 ( $\beta$ 3- $\beta$ 4 strands and loop in-between); residues 90-96 (beginning of  $\alpha$ 4 and loop prior it), can rather adopt a new conformation for a more efficient interaction with the LIR motif (**Figure 33A**).



**Figure 33: Interaction of LC3B with Nix-LIR\_W36 and Nix-LIR\_W140/144.** The CSP of LC3B in presence of Nix-LIRs are shown in (A). Individual  $K_d$ , which have been calculated from three residues selected to represent areas in proximity to hp1 (I34), hp2 (R70) and distant from both (G85) are represented in (B). The CSP are plotted on the structure of LC3B (PDB: 2K6Q) in (C). The residues are represented in different colors following the strength of the CSP: below 0.04 ppm in grey, between 0.04 and 0.08 ppm in yellow and above 0.08 ppm in red. The left pictures concern the interaction of LC3B with Nix-LIR\_W36 and the right ones the interaction of LC3B with Nix-LIR\_W140/144.

An average dissociation constant of the interaction was calculated for the residues showing prominent deviation of their chemical shifts during the NMR titration experiments according to an one-site binding model. The CSP of three different LC3B residues localized either in one hydrophobic pocket (I34 belonging to hp1) or close to one (R70 involved in  $\alpha$ -helix 3 and close to hp2) and out of it (G85 in the loop between  $\beta$ -strand 3 and  $\alpha$ -helix 4) were used as an average representant of the different LC3B resonances.  $K_d$  values calculated from the CSP of LC3B upon

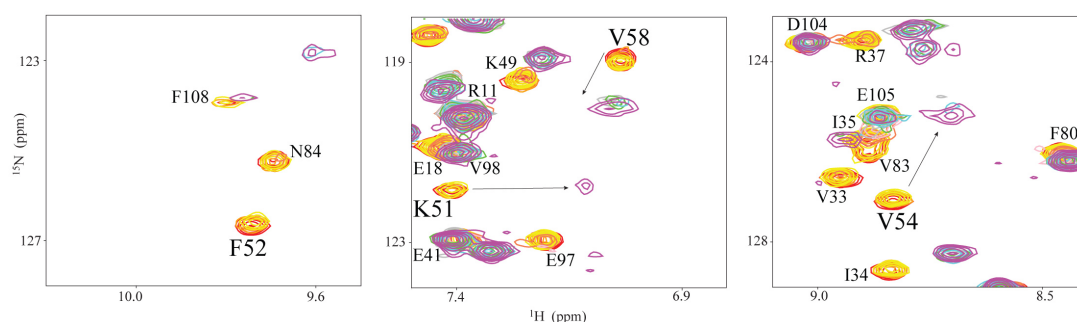


titrations with Nix-LIR\_W36 (100  $\mu\text{M}$ ) and Nix-LIR\_W140/144 peptides (375  $\mu\text{M}$ ) are consistent to the ITC-derived values (**Figure 33B**).

The LC3B residues affected during these interactions with Nix-LIRs as well as with p62-LIR were plotted on the structure of LC3B/p62 (PDB 2K6Q)<sup>87</sup> following the intensity of the CSP (**Figure 33C**). The structure mapping of LC3B in presence of p62-LIR, Nix-LIR\_W36 or Nix-LIR\_W140/144 revealed many similarities in the residues and the regions affected by the interaction of these peptides with LC3B but showed differences in the strength of the perturbation.

### 6.2.3.1.3. LC3B vs. NBR1-LIR

Monitoring the LC3B/NBR1-LIR interaction by ITC, a dissociation constant  $K_d$  of 2.9  $\mu\text{M}$  with a negative enthalpy ( $\Delta H = -4.4 \text{ kcal mol}^{-1}$ ) and a positive entropy ( $\Delta S = +10.7 \text{ cal mol}^{-1} \text{ K}^{-1}$ ) were calculated resulting to a free energy  $\Delta G$  of  $-7.6 \text{ kcal mol}^{-1}$  close to the one for LC3/p62 interaction (**Table 4** and raw data in Appendix 9.3).



**Figure 34: LC3B interaction with NBR1-LIR.** The overlay of representative regions (left for F52, middle for K51 and V58 and right for V54) in [<sup>15</sup>N, <sup>1</sup>H] TROSY-HSQC spectra of LC3B in presence of NBR1-LIR is represented. Contours are presented in different colors corresponding to individual titration points (red as reference with LC3B alone and in presence of LIRs at different molar ratio: orange at 1:1/8, yellow at 1:1/4, coral at 1:1/2, pink at 1:3/4, gray at 1:1, cyan at 1:2, magenta at 1:5 and purple at 1:10).

The  $K_d$  obtained by ITC showed that NBR1-LIR binds stronger than both Nix-LIR to LC3B but still weaker than p62-LIR. NMR titration experiments of NBR1-LIR against LC3B also showed this tendency. In one hand, a binding pattern similar to p62-LIR interaction with a slow exchange mode was observed where the free and bound states are present for some residues (Q15 and L22 from  $\alpha$ -helix 2, I31 and I34 from  $\beta$ -strand 1, K49 from  $\beta$ -strand 2, E62 and R69 from  $\alpha$ -helix 3, Y99 and S101 from  $\alpha$ -helix 4 and L44, L47 or V58 from the loop regions). In the other hand, a faster exchange mode like for Nix-LIR\_W36 was observed, leading to the disappearance of the majority of the resonances upon titration (E25 from  $\alpha$ -helix 2, E36 and

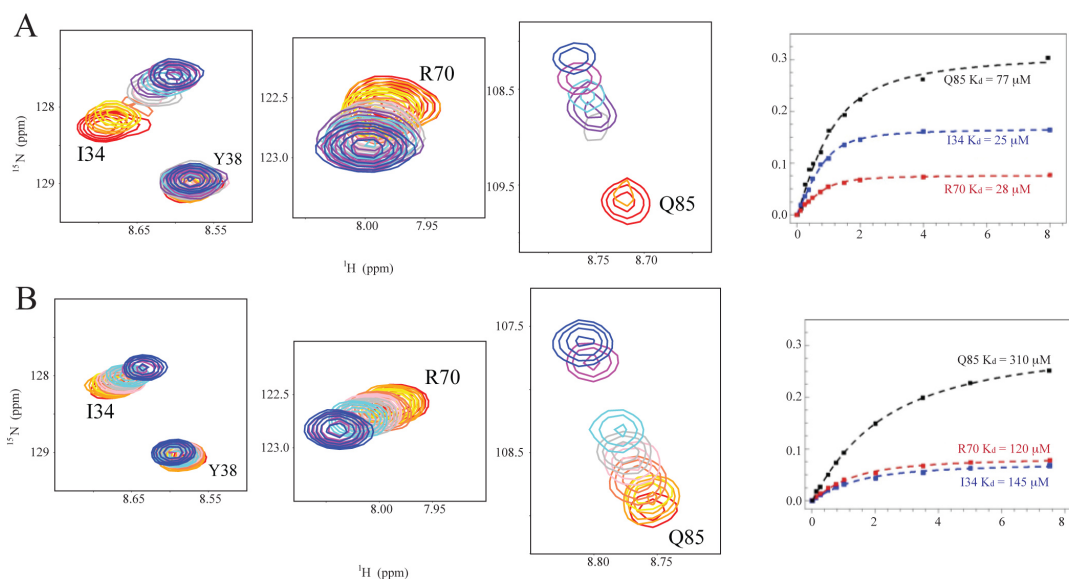
M37 from  $\beta$ -strand 1, D48, F52 and D56 from  $\beta$ -strand 2, R69 and L71 from  $\alpha$ -helix 3, V83 from  $\beta$ -strand 3, I95, S96 and E100 from  $\alpha$ -helix 4 and K39, H57, N59, A78, N84, H86, S97 and M88 from the loop region) whereas these peaks were still present at the same molar ratio for p62-LIR (**Figure 34**). The co-presence of these two situations brings the idea that LC3B interacts strongly with NBR1 but in an intermediate (close to slow) exchange mode. Moreover, these examples showed again that not only one region but also the whole of LC3B molecule is perturbed by the presence of a LIR domain, here from NBR1.

### 6.2.3.2. Other MAP1LC3 proteins vs. LIRs

#### 6.2.3.2.1. LC3A vs. Nix-LIRs

LC3A interacts with both Nix-LIRs with dissociation constants of 28  $\mu$ M and 130  $\mu$ M for Nix-LIR\_W36 and Nix-LIR\_W140/144, respectively. LC3A interaction with both Nix-LIRs is entropy driven with a high positive entropy ( $\Delta S = 17.9 \text{ cal mol}^{-1} \text{ K}^{-1}$  for Nix-LIR\_W36 and  $\Delta S = 15.0 \text{ cal mol}^{-1} \text{ K}^{-1}$  for Nix-LIR\_W140/144) and a reduced contribution of enthalpy ( $\Delta H = -1.0 \text{ kcal mol}^{-1}$  for Nix-LIR\_W36 and  $\Delta H = -0.8 \text{ kcal mol}^{-1}$  for Nix-LIR\_W140/144). As for LC3B, the free energy of the interaction is similar for both Nix peptides ( $\Delta G = -6.3 \text{ kcal mol}^{-1}$  for Nix-LIR\_W36 and  $\Delta G = -5.3 \text{ kcal mol}^{-1}$  for Nix-LIR\_W140/144) and less negative than for p62 and NBR1 peptides (raw data in Appendix 9.3).

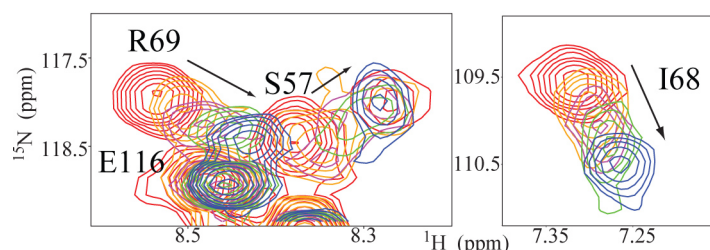
By NMR titration, a fast exchange mode for LC3A against Nix-LIR\_W140/144 and an intermediate (close to fast exchange) mode for LC3A against Nix-LIR\_W36 were observed as for LC3B. In the presence of Nix-LIR\_W36, LC3A peaks are shifted up to 0.6 ppm. R24 from  $\alpha$ -helix 2, T50 from  $\beta$ -strand 2, V58 from the hp2 and N59 and N74, both at the limit of  $\alpha$ -helix 3, are the residues the most affected by presence of Nix-LIR\_W36. Peaks corresponding to the residues K51, F52 and L53, which are at the barrier of hp1 and hp2, disappeared upon titration (**Figure 35A**). Concerning Nix-LIR\_W140/144 titration, the most affected peaks are T50, F52, V54 from  $\beta$ -strand 2 and part of hp2, the residues N59 and Q85 and the C-terminus of  $\beta$ -strand 4 with a maximum CSP up to 0.35 ppm. Only L53 disappeared after addition of Nix-LIR\_W140/144 (**Figure 35B**). Using the same corresponding residues than for LC3B (I34, R70 and Q85), dissociation constants of the interaction of LC3A with Nix-LIRs were calculated:  $K_d$  of 43  $\mu$ M for LC3A/Nix-LIR\_W36 and of 192  $\mu$ M for LC3A/Nix-LIR\_W140/144 (**Figure 35**).



**Figure 35: Interaction of LC3A with Nix-LIRs.** The overlay of representative regions (left for I34, middle for R70 and right for Q85) in [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of LC3A in presence of Nix-LIR\_W36 (A) and Nix-LIR\_W140/144 (B), associated to the individual dissociation constant calculated for these residues is represented. Contours are presented in different colors corresponding to individual titration points (red as reference with LC3A alone and in presence of Nix-LIR at different molar ratio: orange at 1:1/8, yellow at 1:1/4, coral at 1:1/2, pink at 1:3/4, gray at 1:1, purple at 1:1.5, cyan at 1:2, magenta at 1:5 and blue at 1:10). Three residues were selected to represent areas in proximity to hp1 (I34), hp2 (R70) and distant to both (Q85) in order to determine  $K_d$  of the interaction.

#### 6.2.3.2.2. *GABARAPL-1* vs. *Nix-LIR\_W36*

In this case, only few titrations points were done, starting from molar ratio 1:1 to a five-times molar excess of Nix-LIR\_W36. After adding Nix-LIR\_W36, most of the residues located in hp1, hp2 or in close neighborhood presented the strongest CSP (I68, R69 and I72) or disappeared (M32 to V35, K50 to S57, F64, L67 and F108). The residues situated far away from these hydrophobic pockets showed smaller CSP without modification in the peak intensities upon titration. This observation supports an intermediate (close to fast) exchange mode for the interaction of *GABARAPL-1* with Nix-LIR\_W36 (**Figure 36**).



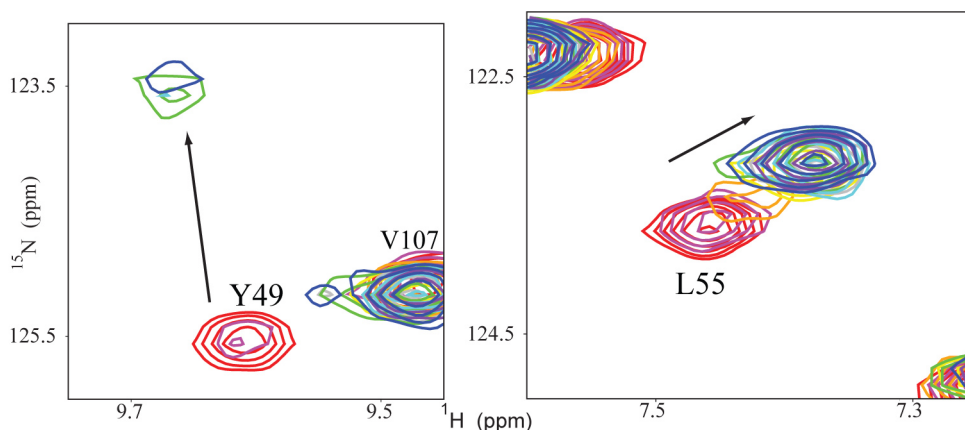
**Figure 36: Interaction of *GABARAPL-1* with *Nix-LIR\_W36*.** The overlay of representative regions (left for R69 and right for I68) in [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of *GABARAPL-1* in presence of Nix-LIR\_W36 is represented. Contours are presented in different colors corresponding to individual titration points (red as reference with *GABARAPL-1* alone and in presence of Nix-LIR\_W36 at different molar ratio: orange at 1:1, magenta at 1:2, green at 1:3 and blue at 1:5).

### 6.2.3.3. *Specificity of NBR1-LIR*

#### 6.2.3.3.1. *GABARAPL-1 vs. NBR1-LIR by NMR*

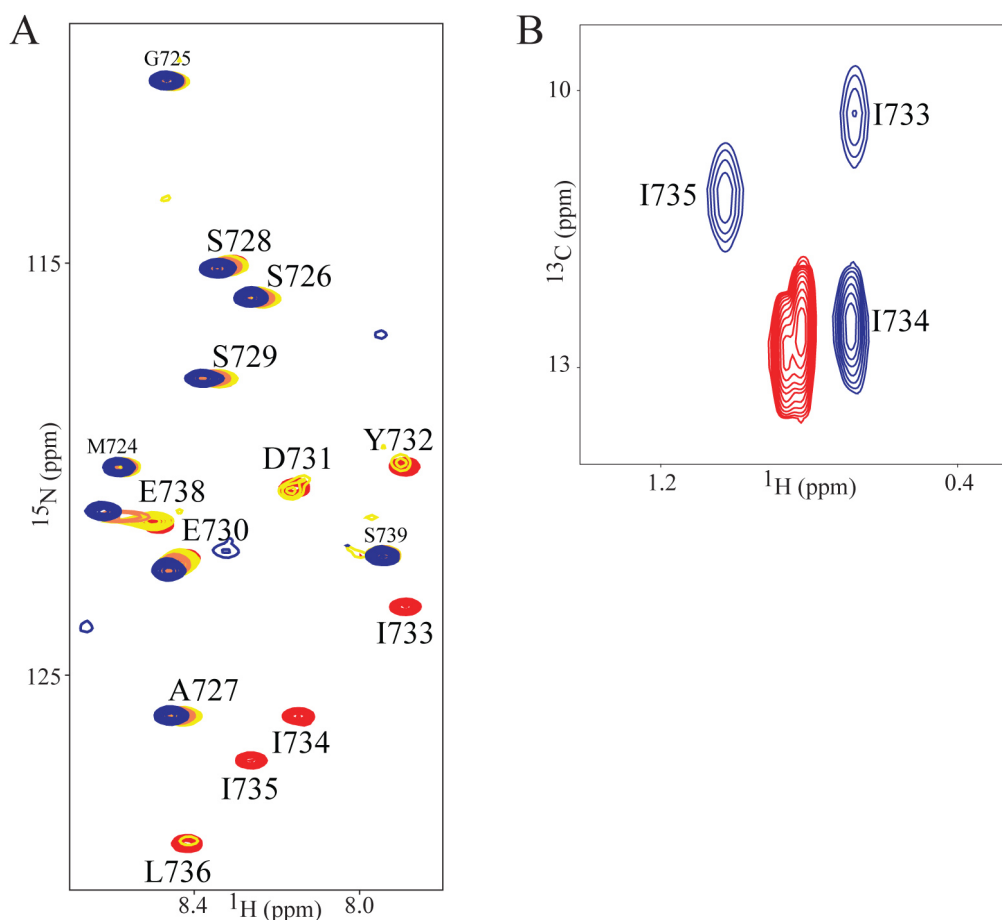
The NMR spectra of a complex of LC3B with the NBR1-LIR domain revealed significant line-broadening of both peptide and protein resonances, making a detailed structural NMR analysis impossible. Therefore, investigations were done to see if the interaction of NBR1-LIR with other MAP1LC3 proteins would result in the formation of a protein complex and to see then which member of the protein family provide the most favorable spectral characteristics. While all of the investigated family members (LC3A, LC3B, GABARAPL-1 and GABARAPL-2) exhibited broad lines due to exchange-broadening, the spectra of NBR1-LIR in complex with GABARAPL-1 showed the most favorable spectral characteristics to further investigate this interaction by NMR.

The previously described studies of LC3B interaction with three different LIR domains showed three different kinetic exchange modes observable by NMR, representing three different interaction strengths. Interestingly in the case of GABARAPL-1 interaction with NBR1-LIR, almost all residues were perturbed but in the three different ways: i) CSP for residues interacting weakly in a fast exchange mode concerning residues not directly involved in the interaction with NBR1-LIR like in the loop regions or in  $\alpha$ -helix 4 (Q97, L98, Y99 or E100); ii) presence of two peaks corresponding to the free and bound states for the residues showing the strongest interaction in a slow exchange mode represented by the residues situated in  $\alpha$ -helix 3 closing the hp2 (L67, I68, R70 or R71); iii) interaction in an intermediate exchange mode with a decrease of the peak intensity then an increase at an other position and finally reaching a full peak intensity approaching to the saturation. A majority of the GABARAPL-1 residues and more precisely the ones situated in the hp1 (L54) and hp2 (Y53 and L59), in  $\alpha$ -helix 2 (E16, K24 and K28) or in the close neighborhood (V37 or T60) are from the type iii) in presence of NBR1-LIR, meaning that these residues are the ones the most probably involved in the interaction. The saturation of the interaction being only approached during this experiment, the intensity of the peaks corresponding to the bound form of GABARAPL-1 is often weaker than the one corresponding to the free form. The presence of a broad peak overlapping free and bound state peaks (Q63, F83, A93) or the presence of a peak between these two states (G22, L59, Y65, I72) indicates also an intermediate (close to fast) kinetic exchange between these two states leading to a weaker interaction than for LC3B. Some peaks were also not observable anymore or difficult to assign like residues I25 and K52 from hp1 and V55 and F64 from hp2 after adding NBR1-LIR, confirming this intermediate exchange (**Figure 37**).



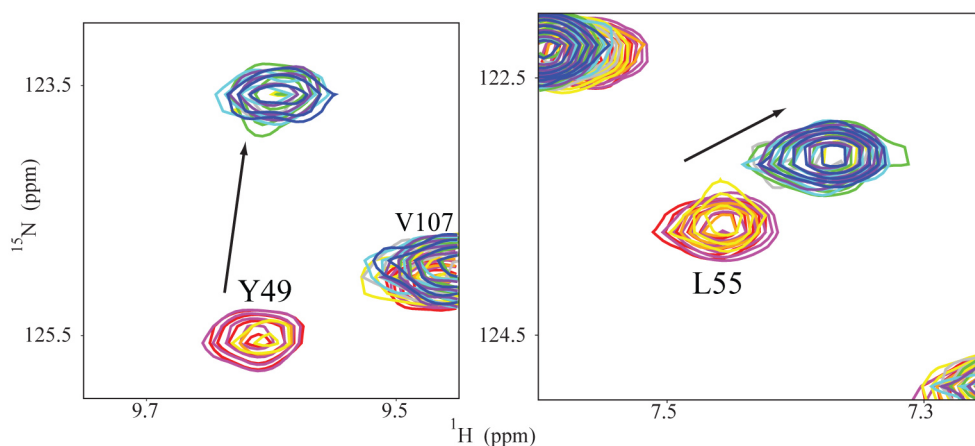
**Figure 37: GABARAPL-1 interaction with NBR1-LIR.** The overlay of representative regions (left for Y49 and right for L55) of [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of GABARAPL-1 in presence of NBR1-LIR is represented. Contours are presented in different colors corresponding to individual titration points (red as reference with GABARAPL-1 alone and in presence of LIR domains at different molar ratio: magenta at 1:1/8, orange at 1:1/4, yellow at 1:1/2, gray at 1:3/4, green at 1:1, purple at 1:1.5, brown at 1:2 and blue at 1:3).

Detailed structural analysis of the GABARAPL-1/NBR1-LIR complex also required the observation of NMR signals of NBR1-LIR in the presence of GABARAPL-1. In titration experiments of  $^{15}\text{N}$ -labeled NBR1-LIR with GABARAPL-1, the amide resonances corresponding to amino acids D731, Y732, I733, I734, I735 and L736 disappeared, all belonging to the core of the LIR motif. At the same time, the resonances of the amino acids surrounding NBR1-LIR motif did not disappear but showed only more or less strong CSP depending of their distance to the core motif (**Figure 38A**). In the presence of a large excess of GABARAPL-1 (molar ratio 1:10), some of the peaks reappeared but the sample was too unstable for further NMR experiments due to the appearance of precipitation (results not shown). However, by performing titration experiments with  $^{13}\text{C}$ -labeled NBR1-LIR, CH signals for all residues of the LIR domain were observable at lower molar ratios. Focusing on the spectral region corresponding to the  $\delta$ -methyl groups of isoleucine, three overlapping peaks were present in the reference spectrum of the free peptide, which shifted significantly upon titration to three well distinguishable resonance positions (**Figure 38B**). Based on the good quality of these spectra, multi-dimensional NOESY experiments, where intra- and intermolecular NOEs were observed, were recorded for the structure determination of the complex.



**Figure 38: NBR1-LIR interaction with GABARAPL-1.** The overlay of the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of NBR1-LIR in presence of GABARAPL-1 at 1:0 (red), 1:1/4 (yellow), 1:1/2 (coral) and 1:1 (blue) molar ratio in (A) and of the representative region of isoleucine CH $\delta$  in [ $^{13}\text{C}$ ,  $^1\text{H}$ ] HSQC for NBR1-LIR (red as reference) and in presence of GABARAPL-1 (blue at molar ratio 1:1) in (B) is represented.

To investigate the role of the aromatic residue in position 1, a mutant form of NBR1-LIR was created in which the tyrosine residue was mutated to tryptophan, the aromatic amino acid the most observed at this position in LIR domains. Upon NMR titration of NBR1-LIR\_Y732W against GABARAPL-1, almost every residue was also perturbed but in majority presenting only two states, free and bound, like residues L50 and Y49, L55, L63 and I64 constituting hp1 and hp2, respectively, suggesting a tighter interaction with NBR1-LIR mutant. Moreover, whereas some GABARAPL-1 peaks did not reappear after addition of NBR1-LIR even close to saturation, this effect was not seen in the case of NBR1-LIR\_Y732W, confirming the formation of a tighter complex for GABARAPL-1 with the tryptophan mutant form of NBR1-LIR than in case of the wild type. NMR titration experiments of the  $^{15}\text{N}$ -labeled GABARAPL-1 both with NBR1-LIR and the NBR1-LIR\_Y732W mutant showed that the exchange-broadening of the tryptophan mutant was reduced compared to the wild type peptide, indicating stronger interaction in a slower exchange mode between the bound and the free form (**Figure 39**).



**Figure 39: GABARAPL-1 interaction with NBR1-LIR\_Y732W.** The overlay of representative regions (left for Y49 and right for L55) of  $^{15}\text{N}$ ,  $^1\text{H}$  TROSY-HSQC spectra of GABARAPL-1 in presence of NBR1-LIR\_Y732W is represented. Contours are presented in different colors corresponding to individual titration points (red as reference with GABARAPL-1 alone and in presence of NBR1-LIR\_Y732W at different molar ratio: magenta at 1:1/8, orange at 1:1/4, yellow at 1:1/2, gray at 1:3/4, green at 1:1, purple at 1:1.5, brown at 1:2 and blue at 1:3).

#### 6.2.3.3.2. GABARAPL-1 vs. NBR1-LIR by ITC

To understand thermodynamically the involvement of different residues in the LIR motif of NBR1 in the interaction with GABARAPL-1, ITC experiments were performed in which NBR1-LIR and several of its mutants were titrated against GABARAPL-1. Titration of NBR1-LIR against GABARAPL-1 showed a strong binding ( $K_d = 1.3 \mu\text{M}$ ) equivalently driven by the contributions of the enthalpy and entropy of binding ( $\Delta H = -4.1 \text{ kcal mol}^{-1}$  and  $\Delta S = 11.6 \text{ cal mol}^{-1} \text{ K}^{-1}$ ) leading to a free energy  $\Delta G$  of  $-7.6 \text{ kcal mol}^{-1}$  equivalent of the interaction of LC3B with NBR1-LIR (**Table 5** and raw data in Appendix 9.3).

The measurement of the binding affinity of GABARAPL-1 to NBR1-LIR\_Y732W by ITC showed a stronger binding indicated by a reduction of the dissociation constant to  $0.4 \mu\text{M}$ . The tighter binding of NBR1-LIR\_Y732W is also reflected in a more negative binding enthalpy contribution ( $\Delta H = -5.7 \text{ kcal mol}^{-1}$ ). As in presence of the wild type, the binding entropy for GABARAPL-1/NBR1-LIR\_Y732W interaction is large and positive ( $\Delta S = 10.3 \text{ cal mol}^{-1} \text{ K}^{-1}$ ) but the free energy is slightly more negative ( $\Delta G = -8.8 \text{ kcal mol}^{-1}$ ) (**Table 5** and raw data in Appendix 9.3).

To further investigate the importance of the nature of the aromatic residue of LIR in position 1 in the interaction, the tyrosine residue in NBR1-LIR was mutated to phenylalanine and NBR1-LIR\_Y732F was titrated against GABARAPL-1. A strong interaction ( $K_d = 2.9 \mu\text{M}$ ) with a very low binding entropy ( $\Delta S = 2.6 \text{ cal mol}^{-1} \text{ K}^{-1}$ ) was revealed suggesting a tighter interaction that was reflected also by a more negative binding enthalpy ( $\Delta H = -6.8 \text{ kcal mol}^{-1}$ ) relative to

that of the wild type peptide but with the same free energy ( $\Delta G = -7.6 \text{ kcal mol}^{-1}$ ) than for NBR1-LIR wild type (**Table 5** and raw data in Appendix **9.3**).

GABARAPL-1 vs NBR1-LIR	$\Delta H$ kcal mol <sup>-1</sup>	$\Delta S$ cal mol <sup>-1</sup> K <sup>-1</sup>	$\Delta G$ kcal mol <sup>-1</sup>	$K_d$ μM
NBR1	-4.1	+11.6	-7.6	3
NBR1_Y732W	-5.7	+10.3	-8.8	0.4
NBR1_Y732F	-6.8	+2.6	-7.6	2.9

**Table 5: Thermodynamic parameters obtained by ITC for the interactions of GABARAPL-1 with NBR1-LIR wild type and mutants at position 1.**

All experiments were performed at 25°C.  $\Delta H$ ,  $\Delta S$  and  $K_d$  values were measured with the assumption of a one site model. Although statistical errors for all calculations were quite small (below 3%), the uncertainty of the sample concentrations was approximately 5% (or more in case of peptides containing no tryptophan or tyrosine residues). Therefore, a total error of 10% for all parameters could be expected.

In order to investigate if the addition of negatively charged residues preceding the tyrosine in NBR1-LIR could increase the affinity to GABARAPL-1, glutamate mutants were created. The mutation S729E resulted in a slightly increased affinity to GABARAPL-1 characterized by a more negative binding enthalpy ( $\Delta H = -4.8 \text{ kcal mol}^{-1}$ ) and a decrease of the binding entropy ( $\Delta S = 9.5 \text{ cal mol}^{-1} \text{ K}^{-1}$ ). These effects were further enhanced with the double mutant NBR1-LIR\_S728,729E ( $\Delta H = -5.5 \text{ kcal mol}^{-1}$  and  $\Delta S = 7.2 \text{ cal mol}^{-1} \text{ K}^{-1}$ ). In both cases, however, the overall effect was however rather small owing to a remarkable compensation of enthalpic and entropic effects, resulting in very similar dissociation constant and free energy values that are not significantly different from the wild type peptide ( $K_d$  of 2.7 and 2.4 μM and  $\Delta G$  of -7.6 and -7.7 kcal mol<sup>-1</sup>, for the single and double mutant, respectively) (**Table 6** and raw data in Appendix **9.3**).

GABARAPL-1 vs NBR1-LIR	$\Delta H$ kcal mol <sup>-1</sup>	$\Delta S$ cal mol <sup>-1</sup> K <sup>-1</sup>	$\Delta G$ kcal mol <sup>-1</sup>	$K_d$ μM
NBR1	-4.1	+11.6	-7.6	3
NBR1_S729E	-4.8	+9.5	-7.6	2.7
NBR1_S728,729E	-5.5	+7.2	-7.7	2.4

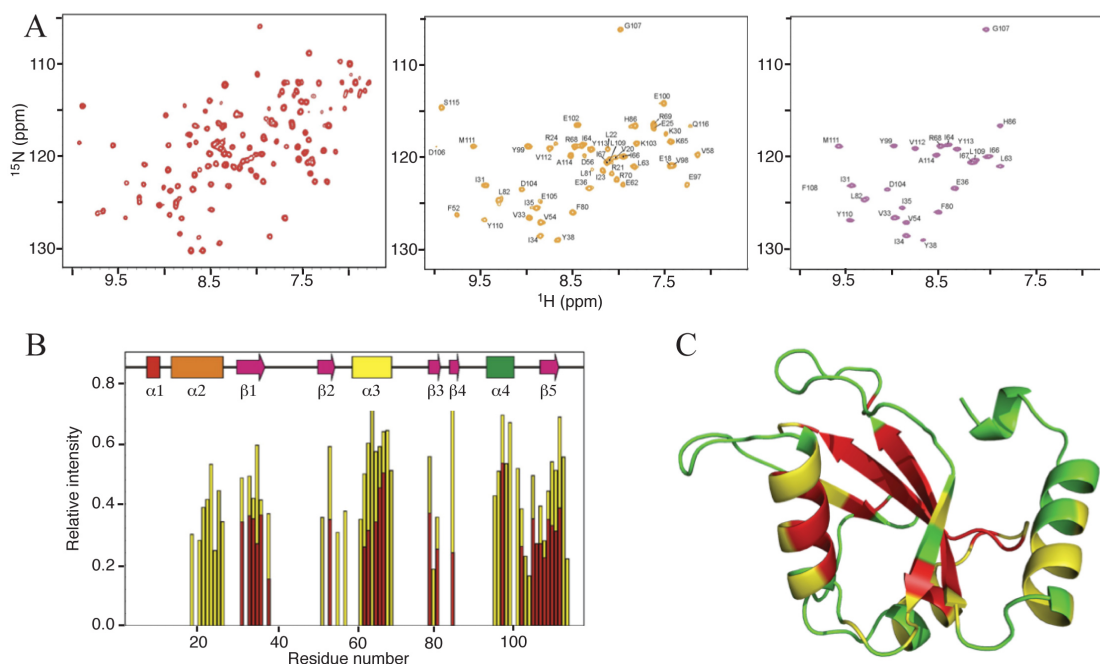
**Table 6: Thermodynamic parameters obtained by ITC for the interactions of GABARAPL-1 with NBR1-LIR wild type and mutants at position -1 and -2.**

All experiments were performed at 25°C.  $\Delta H$ ,  $\Delta S$  and  $K_d$  values were measured with the assumption of a one site model. Although statistical errors for all calculations were quite small (below 3%), the uncertainty of the sample concentrations was approximately 5% (or more in case of peptides containing no tryptophan or tyrosine residues). Therefore, a total error of 10% for all parameters could be expected.



## 6.2.4. Flexibility of Atg8 proteins

### 6.2.4.1. LC3B



**Figure 40: Flexibility of LC3B.** Amide proton exchange experiments of LC3B are shown in (A). The left spectrum is the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC reference spectrum for LC3B. The middle and the right spectra show the replacing of amide protons of LC3B by deuterons after placing the lyophilized sample into  $\text{D}_2\text{O}$  for 10 minutes and 1 hour, respectively. The relative intensity (decrease of intensity regarding to the intensity of reference peaks) of LC3B residues after amide protons exchange is shown in (B) and is plotted on the structure of LC3B in (C). The residues disappearing before 10 minutes, before 1 hour and still present after 1 hour are represented in green, yellow and red, respectively.

To monitor the flexibility of LC3B, amide/proton exchange experiments were performed by NMR. After measuring a [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC reference spectrum for LC3B, the sample was lyophilized and then resuspended into  $\text{D}_2\text{O}$ . After 10 minutes, a new [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum for the sample was recorded and some amide protons were already completely replaced with deuterons and only 51 residues peaks were present against 96 for the reference sample. All resonances of amino acids belonging to the first  $\alpha$ -helix, to the N-terminus of the second  $\alpha$ -helix, to the loop between  $\beta$ -strands 1 and 2, to the N-terminus of  $\beta$ -strand 2, to the loop between  $\alpha$ -helix 3 and  $\beta$ -strand 3 and to the loop between  $\beta$ -strand 3 and  $\alpha$ -helix 3 were the first to disappear on the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum. Over the next hour, remaining of the amide protons present in  $\alpha$ -helix 2, in  $\beta$ -strand 2, in  $\alpha$ -helix 4 and at the C-terminus of LC3B exchanged completely and were not observable anymore in the next [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum (**Figure 40A**). However, some amide protons (24 residues peaks) participating in the central antiparallel sheet of LC3B were still present in the TROSY spectrum after some months like for residues belonging to  $\beta$ -strand 1, to the central part of  $\alpha$ -helix 3, to  $\beta$ -strand 3, to a part

of the loop between  $\alpha$ -helix 4 and  $\beta$ -strand 4 and to  $\beta$ -strand 4 (**Figure 40B**). Plotted on the surface of LC3B, the fast exchange appears for amide groups exposed to the solvent whereas those buried in the core of LC3B should present a slower exchange rate (**Figure 40C**).

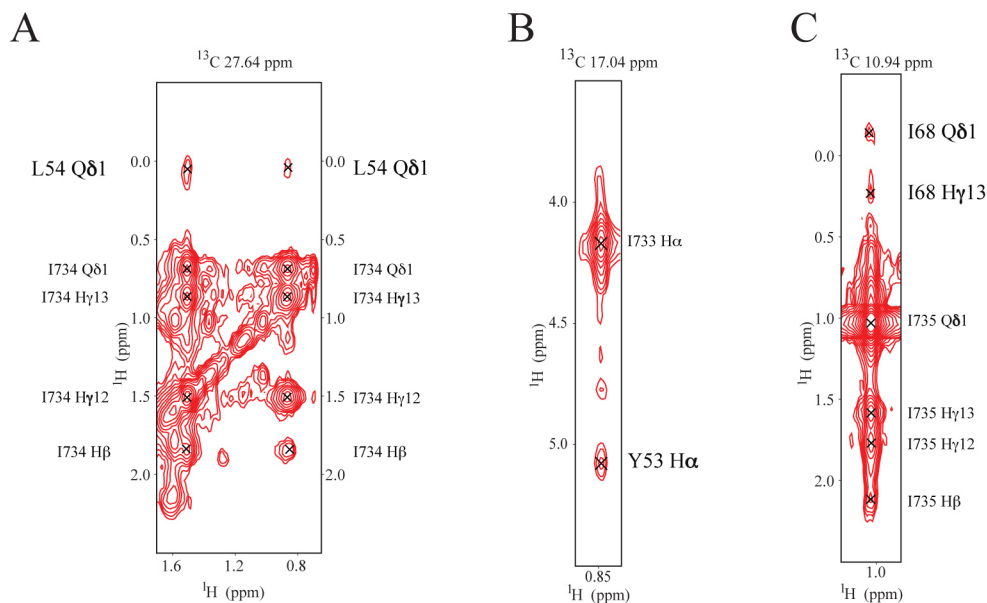
#### **6.2.4.2. GABARAPL-1**

Residues Q8 to D12 constituting  $\alpha$ -helix 1 but also F15, R18 and I25 from  $\alpha$ -helix 2 were hardly seen in each recorded [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum of GABARAPL-1. Although the absence of these residues could be due to proteolytic cleavage, the NMR signals corresponding to these residues were present in [ $^{13}\text{C}$ ,  $^1\text{H}$ ] HSQC and  $^{13}\text{C}$ -edited NOESY. The flexibility of the two N-terminal  $\alpha$ -helices of GABARAPL-1, occurring in a slow process regarding to NMR time-scale, could probably explain the disappearance of the amide proton peaks during the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC experiment. Similarly, broad or double peaks were observed for residues, R18 and E23 as well as K19 and K24, which belongs to  $\alpha$ -helix 2. This observation confirmed that the N-terminal region of GABARAPL-1 is subjected to some dynamics (results not shown).

#### **6.2.5. 3D structure of GABARAPL-1/NBR1-LIR**

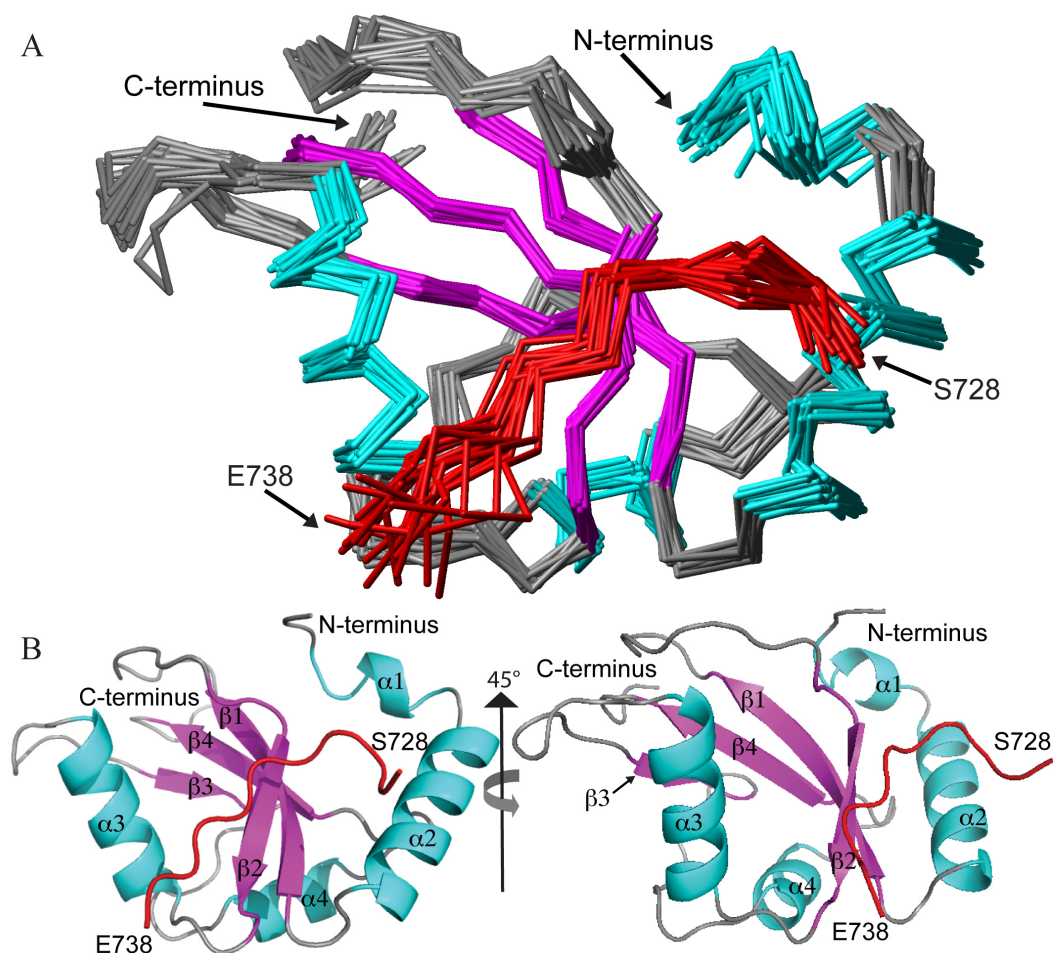
NMR titration experiments showed that a  $\sim 1.5$ - $2.0$  molar excess of one interaction partner was required to saturate the binding site. To obtain structural constraints for both partners, NOESY spectra of complexes in which either GABARAPL-1 or the NBR1-LIR domain was  $^{13}\text{C}/^{15}\text{N}$  labeled in the presence of a 1.5 molar excess of the unlabeled interaction partner were recorded. Under these conditions, almost all (backbone and side chain) resonances in the spectra were assigned (except for the N-terminal part of GABARAPL-1 near and along the small  $\alpha$ -helix 1). Due to dynamic contributions and protein concentration limitations (aggregation at  $\sim 1.5$  mM concentration) intermolecular 3D J-resolved NOESY spectra did not provide reliable intermolecular resonances. However, in the 3D NOESY spectra of NBR1-LIR in presence of GABARAPL-1, and vice versa, eight NOE cross-peaks could be unambiguously manually assigned as intermolecular NOEs but two of them could not be manually assigned for which intermolecular NOEs. Using CYANA for peak picking and automatic assignment, a total of 69 intermolecular NOEs were unambiguously assigned (**Figure 41**). The different intermolecular NOEs are listed in the Appendix 9.4. Based on the backbone and the side-chains assignment of GABARAPL-1 and NBR1-LIR and the selection of all NOE signals in 3D  $^{15}\text{N}$  and  $^{13}\text{C}$  NOESY spectra, the CANDID module of CYANA<sup>143</sup> assigned these NOESY cross peaks. With the

additional torsion angles and upper distance limit restraints provided by the TALOS<sup>142</sup> and CSI<sup>141</sup> programs, 20 final conformers were calculated with CYANA.



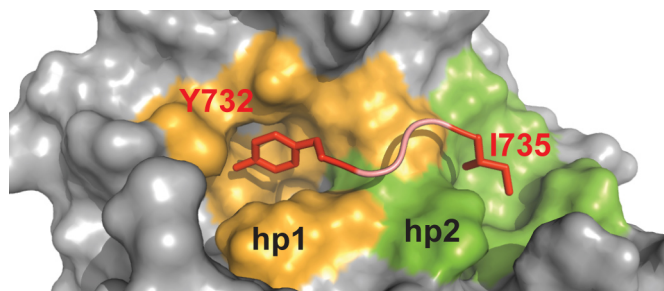
**Figure 41: GABARAPL-1/NBR1-LIR interaction proved by intermolecular NOEs.** A partial strip corresponding to I734 H $\gamma$ 12 and I734 H $\gamma$ 13 (of NBR1) is shown in (A). In addition to intramolecular NOEs to I734 CH $_3$  $\delta$ 1, I734 H $\gamma$ 12, I734 H $\gamma$ 13 and I734 H $\beta$ , an intermolecular NOE at 0.045 ppm corresponding to L50 CH $_3$  $\delta$ 1 could be unambiguously manually assigned. In (B), the strip of I733 CH $_3$  $\gamma$ 2 shows a peak at 5.078 ppm, which could be unambiguously manually assigned as an intermolecular NOE corresponding to Y49 H $\alpha$ . NOE cross peaks for I735 CH $_3$  $\delta$ 1 are shown in (C) with the resonances at -0.141 ppm and 0.232 ppm assigned to intermolecular NOEs corresponding to I64 CH $_3$  $\delta$ 1 and V51 CH $_3$  $\gamma$ 1, respectively.

As expected, the three dimensional structure of GABARAPL-1 in the presence of the NBR1-LIR domain is similar to the structure of other MAP1LC3 proteins. The typical ubiquitin core  $\beta\beta\alpha\beta\alpha\beta$  is preceded by two  $\alpha$ -helices 1 and 2. These two N-terminal  $\alpha$ -helices ( $\alpha$ 1, Q4 to D8;  $\alpha$ 2, F11 to K24) are the result of the position of a proline (P10) in the middle of a longer  $\alpha$ -helix by creating a kink. The  $\beta$ -sheet domain is composed of four  $\beta$ -strands, two being parallel ( $\beta$ 1, R28 to K35;  $\beta$ 4, L105 to S110) and the other two are attached antiparallel ( $\beta$ 2, K48 to P52;  $\beta$ 3, F77 to V80). The two N-terminal  $\alpha$ -helices close hp1 formed by the convex side of the  $\beta$ -sheet. The other hydrophobic pocket, hp2, including the concave side is covered by  $\alpha$ -helices  $\alpha$ 3 (V57 to R67) and  $\alpha$ 4 (M91 to D97) (**Figure 42** and structural statistics in Appendix 9.5). The hydrophobic pocket hp1 is formed by the residues E17, I21, P30, I32, K48, L54 and F99 while Y49, V51, P52, L55, F60, L63 and I64 are constituting the hydrophobic pocket hp2. The structure of GABARAPL-1 in complex with the NBR1-LIR domain shows only small changes relative to the non-complexed crystal structure (PDB entry 2R2Q reference) with a backbone RMSD of 1.78 Å over the structured part (Q4 to S110).



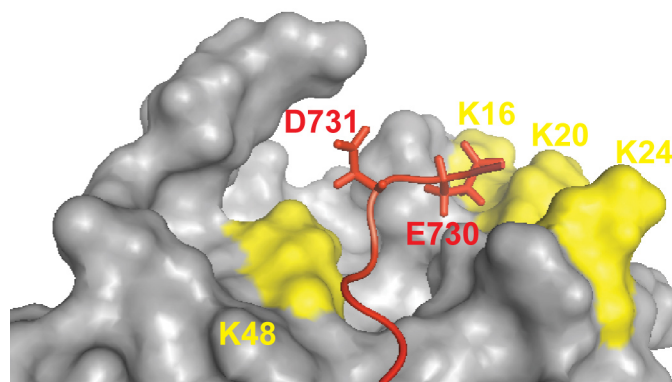
**Figure 42: NMR structure of the GABARAPL-1/NBR1-LIR complex.** The overlay of the backbone atoms of the 20 energy-refined conformers of GABARAPL-1 in complex with NBR1-LIR is shown in (A). Loops are shown in gray,  $\alpha$ -helices in cyan,  $\beta$ -strands in magenta and NBR1-LIR is colored in red. Residues 1 to 7 of GABARAPL-1 are disordered and were therefore excluded. Only the structured region of NBR1 (S728 to E738) is shown. The mean structure of the complex is presented as a ribbon diagram in two orientations (the view on the right corresponds to a  $45^\circ$  rotation) in (B).

In the complex, the LIR motif of the NBR1 peptide adopts an extended conformation and adds a  $\beta$ -strand to the central  $\beta$ -sheet of GABARAPL-1. The side chain of Y732 ( $\ominus$  NBR1-LIR aromatic residue, position 1) binds deep into hp1 and makes close contact with E17, I21, P30, K48, L50 and F104 with distances less than  $4 \text{ \AA}$  (**Figure 43**). hp2 interacts with the side chain of I735 ( $\Gamma$  NBR1-LIR hydrophobic residue, position 4), with short distances to amino acids Y49, V51, P52 and L55 on one side of the pocket and L63, I64 and R67 on the other side (**Figure 43**).



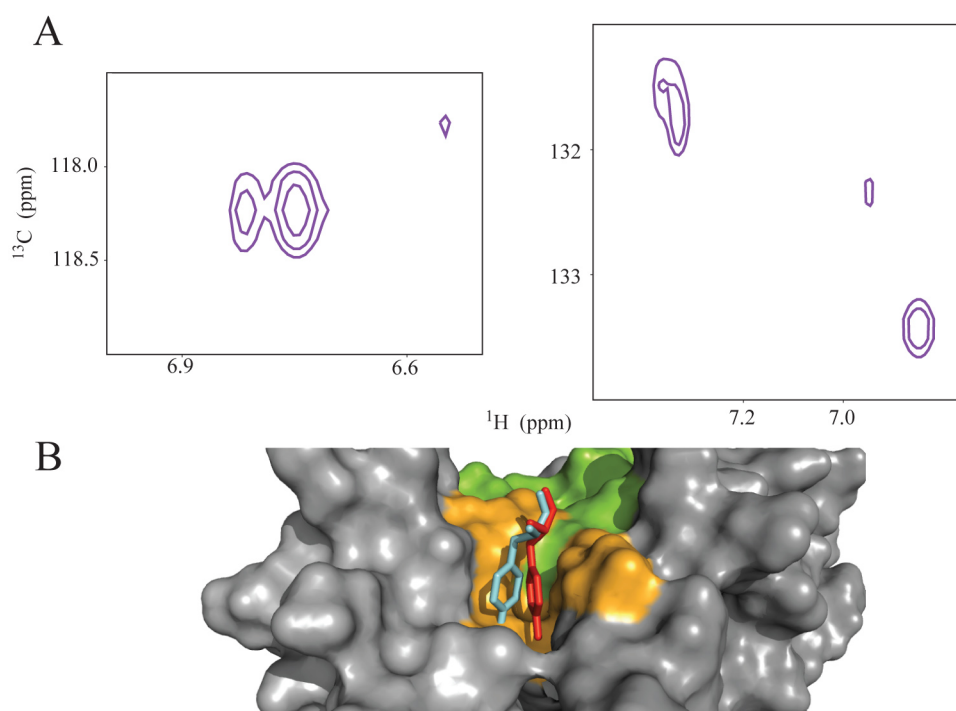
**Figure 43: LIR binding sites of GABARAPL-1 in presence of the NBR1-LIR domain.** The position of Y732 and I735 (red) within NBR1-LIR and their interaction with the hp1 (orange) and hp2 (green) on GABARAPL-1 (grey) are shown.

The other two hydrophobic amino acids, I733 and I734 at position 2 and 3, are in contact with the surface of GABARAPL-1 (along  $\beta$ -strand 2). L736, the last hydrophobic residue at position 5, is in close proximity to hp2 but its side chain is oriented in a different direction and makes contact instead with V51 and P52. Finally the two acidic amino acids E730 and D731 at positions -2 and -1 interact with the positively charged  $\alpha$ -helix 2 (closer residues are E17, K20 and I21) (**Figure 44**).



**Figure 44: Interaction of the acidic residues of NBR1-LIR with positively charged amino acids of GABARAPL-1.** E730 and D731, located at the N-terminus of Y732 in NBR1-LIR, make close contacts with positively charged residues of GABARAPL-1, especially with residues in  $\alpha$ -helix 2 (K16, K20 and K24).

Careful analysis of the aromatic region of the NMR spectra revealed that the  $\epsilon$ -protons of Y732 show two different peaks in the complex with GABARAPL-1, probably representing a major and a minor conformation (**Figure 45A**). Unfortunately, the quality of the sample was not sufficient to observe specific NOEs for both conformations but only for the more intense resonance. However, during the structure calculation two different populations of the peptide with different positions of the tyrosine side chain in the hydrophobic pocket were consistent with the restraints, suggesting that potentially more than one conformation of the tyrosine side chain might exist (**Figure 45B**).

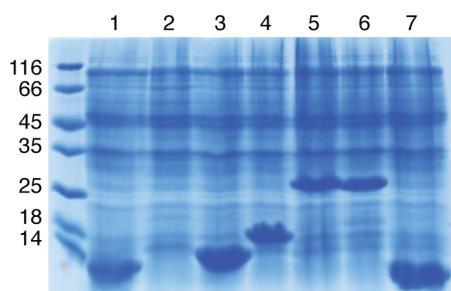


**Figure 45: Flexibility in the interaction of NBR1-LIR with GABARAPL-1.** The regions from a [ $^{13}\text{C}$ ,  $^1\text{H}$ ] HSQC spectrum showing different signals corresponding to the tyrosine resonance positions  $\text{CH}\epsilon$  (left) and  $\text{CH}\delta$  (right) of the NBR1-LIR in presence of GABARAPL-1 at a molar ratio 1:8 are illustrated in (A). The potential positions of the tyrosine residue of NBR1-LIR in the hp1 (orange) of GABARAPL-1 are represented in (B). Both conformers are compatible with all NMR restraints and were observed in the structure calculation of the GABARAPL-1/NBR1-LIR complex.

### 6.3. Ubiquitin project

#### 6.3.1. Design of ubiquitin tag constructs

Cloned into the pETM-60 vector, the target proteins were fused to the C-terminus of ubiquitin presenting a TEV protease cleavage site preceded by a  $\text{His}_6$ -tag. The first Ub-fused protein constructs involving TBK1\_ULD showed already degradation products after only few hours of expression (results not shown). To overcome the degradation issue, the ubiquitin used in the next construct was slightly modified with the two last glycines substituted to S75 and A76. Compared to the constructs containing ubiquitin wild type, the modifications within ubiquitin increased the internal stability without any cleavage products appearing upon expression and the solubility of the expressed fusion constructs (**Figure 46**).



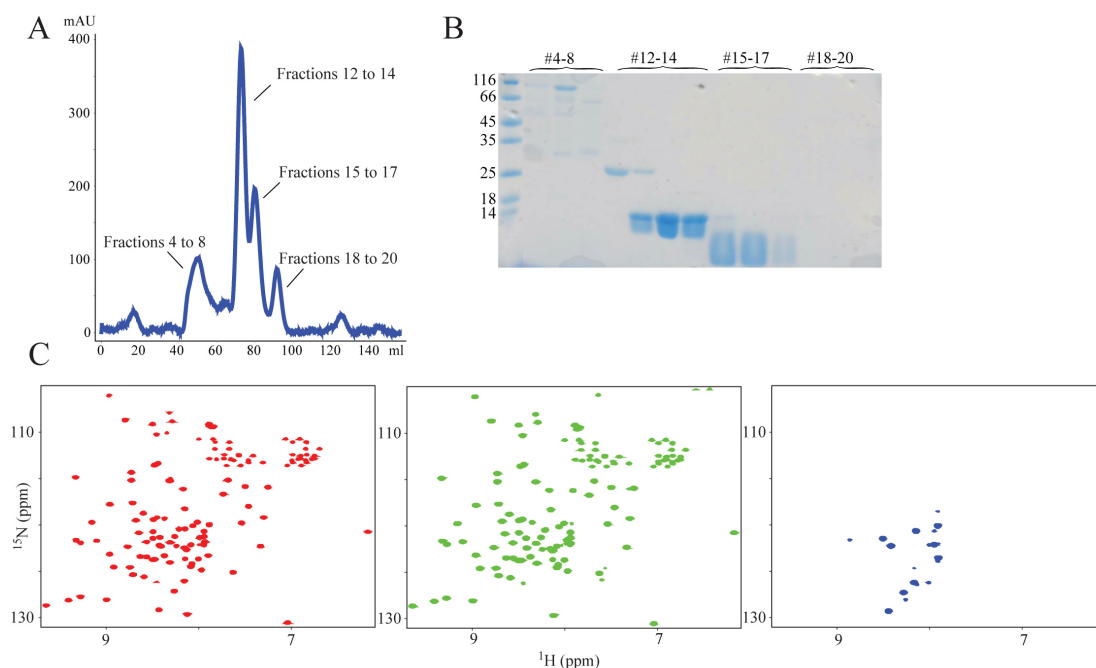
**Figure 46: Expression of Ub-fused constructs.** The expressed proteins from 1 to 7 are optineurin-LIR, TBK1\_ULD, p62-LIR, NBR1-LIR\_long, LC3A, LC3B and Nix-LIR\_W140/144, respectively. Each final expression was obtained after 3 hours of expression at  $37^\circ\text{C}$  with 1 mM IPTG.

### 6.3.2. *Expression and purification of pure proteins/peptides*

First experiments showed that mammalian Atg8 proteins used for the “autophagy project” (chapter 6.2.1) could be expressed in a very large amount ( $100 \text{ mg mL}^{-1}$ ) in less than 3 hours expression at  $37^\circ\text{C}$  with  $1 \text{ mM}$  IPTG without any degradation products. Screening the expression conditions by changing the amount of IPTG, the temperature, the degree and the time of expression did not improve significantly the expression yield (**Figure 46**).

Proteins and peptides expressed with the designed ubiquitin expression tag were purified by immobilized metal ion affinity chromatography (IMAC) using a Ni-NTA column due to the presence of the His<sub>6</sub>-tag. After the cell lysis using a French Press and after the removal of the cell debris by centrifugation, the cell lysate was loaded onto a Ni-NTA column. Flow-through, washing and elution fractions were analyzed by SDS-PAGE and showed that ubiquitin fused proteins were nicely isolated. The presence of a TEV cleavage site between the C-terminus of ubiquitin and the target protein gave the possibility to produce pure protein by removing the ubiquitin moiety. The collected elution fractions were cleaved by TEV protease using standard protocols. The cleavage rate was very efficient and no uncleaved product was detectable by SDS-PAGE (results not shown).

Because the molecular size of the target protein was often close to the one of ubiquitin, size-exclusion chromatography did not always separate both proteins efficiently. A second purification step by ion exchange chromatography with an sp-sepharose column was successfully used for Ub\_GABARAPL-1 but, for Ub\_LC3B, the target protein was still present in the flow-through where only ubiquitin should have been found. Performing a new Ni-NTA purification of the cleaved elution fraction after dialysis appeared to separate successfully the target protein in the flow-through from ubiquitin, which bound to the matrix because of the remaining His<sub>6</sub>-tag. Then a last purification step using size-exclusion chromatography was done to remove the remaining contaminants and obtain a pure final sample. Moreover, the target protein was eluted at the retention volume corresponding to its expected molecular size and was subjected to buffer exchange at the same time (results not shown). LC3A, LC3B and GABARAPL-1 were easily purified from Ub-constructs, providing an average expression yield of  $30 \text{ mg L}^{-1}$ . With this amount, the proteins could be further used for NMR and ITC experiments.



**Figure 47: Purification of cleaved Ub\_NBR1-LIR.** After the action of TEV protease, cleaved Ub\_NBR1-LIR was purified by size-exclusion chromatography. The elution profile is illustrated in (A) and the SDS-PAGE with the corresponding collected fractions is shown in (B). The [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra corresponding to fractions #12-14, #15-17 and #18-20 are represented in (C) from the left to the right, respectively.

The methods used for protein purification could not exactly be applied for peptides owing their small size. Indeed concentration steps had to be avoided because of the risk of peptide loss through the concentrating membrane, which have a 3 kDa cut-off when most of the peptide used in this study have a lower molecular size. Thus, slight modifications in the purification protocol were applied to successfully purify peptides from the Ub-construct. The first attempts on peptides purification expressed from an Ub-fused construct were performed with Ub\_NBR1-LIR. After TEV cleavage of the Ni-NTA elution product, size-exclusion chromatography in ammonium chloride buffer was performed to separate the peptide from ubiquitin. Unfortunately, elution peaks of ubiquitin and target peptides were not always perfectly separated (**Figure 47A**). Nevertheless, after checking the fraction purity with SDS-PAGE, some elution fractions presenting absorbance at 280 nm were free of ubiquitin (**Figure 47B**). These fractions were then lyophilized and resuspended in the final NMR buffer (50 mM sodium phosphate, 100 mM NaCl at pH 7.0). Because of the small size of the NBR1-LIR peptide, no band corresponding to the protein was observed by SDS-PAGE but the UV absorbance at 280 nm, which could be related to the presence of the tyrosine residue in NBR1-LIR, confirmed the purification of the target peptide. The first [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum showed a nice dispersion of all expected peaks, 12 peaks corresponding to the LIR motif plus 3 peaks from cloning artifacts. P737 as well as the two first residues of the construct belonging to the cloning

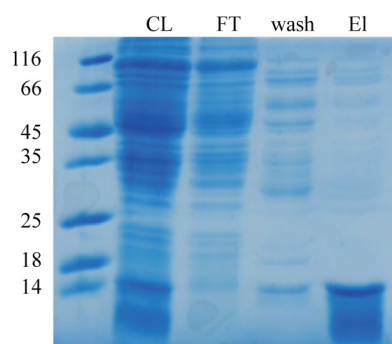


site were not seen in the spectrum. The presence of ubiquitin in the background was very subtle and did not interfere with NBR1-LIR resonances (**Figure 47C**). The [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum of NBR1-LIR was a proof that pure labeled NBR1-LIR was correctly isolated. Thus the same protocol was used to purify the unlabeled NBR1-LIR necessary as interaction partner of GABARAPL-1 for the structure determination of the protein complex. Because no direct biochemical detections of the unlabeled NBR1-LIR was possible, it was assumed that the UV detection of the protein absorbance associated with no SDS-PAGE signal was a proof of the good purification of unlabeled NBR1-LIR. At the end, 10 to 20 mg of pure peptide was obtained per liter of expression.

### 6.3.3. Expression and purification of ubiquitin fused proteins/peptides

As shown above, the purification of peptides expressed with the designed ubiquitin expression tag was successfully achieved. This method for the production of pure peptides needs a lot of steps causing the loss of a lot of materials during these procedures. To overcome the peptide lost issue, the expression and the purification of peptide fused to ubiquitin and not cleaved from it was done.

The expression and the purification protocols were the same as the one previously described providing the same amount of ubiquitin-fused protein/peptide. In some cases, the Ni-NTA elution fraction was quite pure and buffer exchange was sufficient to have enough of pure materials to perform further experiments. To increase the sample purity, size-exclusion chromatography was used in addition. The Ub-fused peptide eluted monodispersly at the expected volume, no cleaved ubiquitin was detected and no precipitations were observable during concentration (**Figure 48**). A final average amount of 100 to 150 mg of ubiquitin-fused protein/peptide per liter of expression was obtained for the different constructs used.

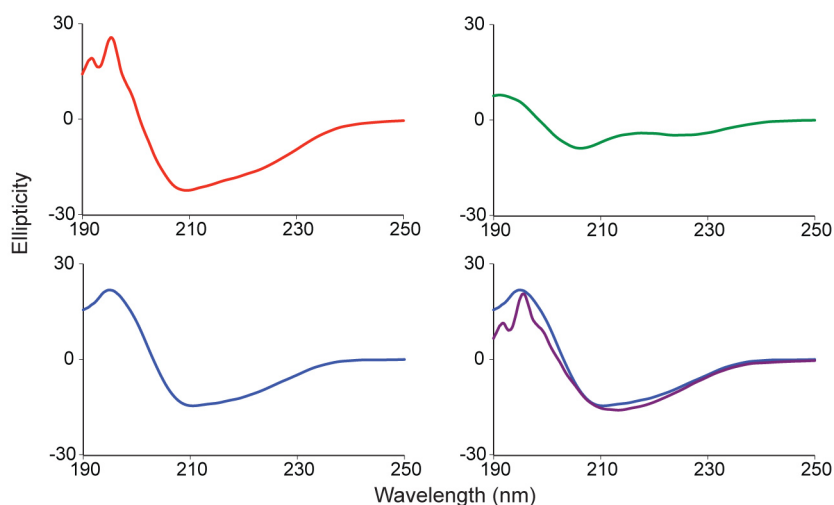


**Figure 48: Purification of Ub\_NBR1-LIR.** The purification was performed on a NiNTA column. CL represents the cell lysate, FT the flow-through, wash the washing step and EI the elution fraction. The smear observed in the elution fraction is due to the presence of imidazole in the sample.

Ub\_NBR1-LIR and Ub\_NBR1-LIR\_Y732W were produced under these conditions and used for NMR and ITC experiments. These protocols were also effective for the expression and the purification of bigger proteins fused to ubiquitin, like GABARAPL-1 with a molecular size of 15 kDa.

#### 6.3.4. Use of ubiquitin-fused proteins for CD spectroscopy

To verify if the presence of ubiquitin interferes with the structure of the protein of interest, CD spectroscopy was used to provide structural information about the Ub-fused protein, the ubiquitin moiety and the target protein.



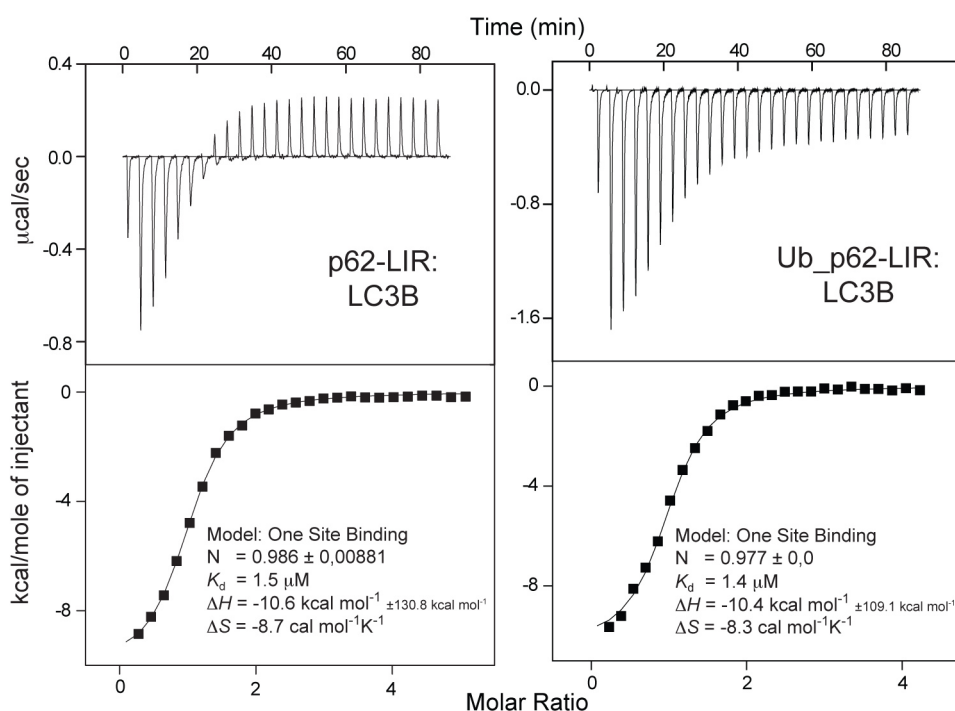
**Figure 49: CD spectroscopy with Ub-fused proteins.** The intact Ub\_GABARAPL-1 is represented in red, ubiquitin in green and GABARAPL-1 in blue. The curve resulting from the subtraction of the values obtained from ubiquitin alone to the ones of uncleaved Ub\_GABARAPL-1 is in purple. This curve overlays with the curve of GABARAPL-1 alone.

CD spectra were recorded for Ub\_GABARAPL-1 as well as for ubiquitin and GABARAPL-1 alone. The CD spectrum of Ub\_GABARAPL-1 fits with the CD spectrum resulting from the addition of ubiquitin and GABARAPL-1 spectra. By subtracting the data obtained during the CD spectrum of ubiquitin from the one of Ub\_GABARAPL-1, a similar CD spectrum than the one issued from GABARAPL-1 alone was observed. No structural elements appeared or disappeared due to the fusion of ubiquitin to GABARAPL-1 (**Figure 49**).

#### 6.3.5. Use of ubiquitin-fused peptides for ITC

In the “autophagy project”, the binding affinity of LC3B to p62-LIR using a chemically synthesized peptide was well studied by ITC. To prove that the ubiquitin moiety does not interfere with the interaction and in consequence that a Ub-fused peptide could be used instead of a peptide alone, the ITC titration of p62-LIR against LC3B was reperformed in the same

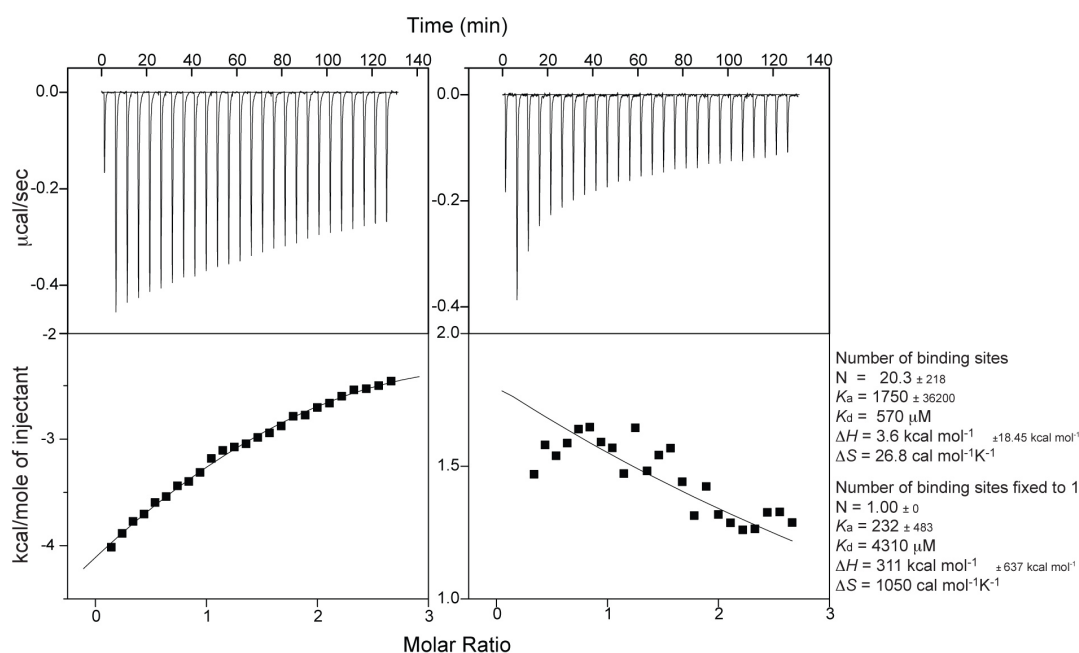
buffer conditions as with the free peptide but using p62-LIR fused to ubiquitin this time. The dissociation constants obtained for both titration experiments were almost identical with a  $K_d$  of 1.5  $\mu\text{M}$  and 1.4  $\mu\text{M}$  for LC3B/p62-LIR and LC3B/Ub\_p62-LIR interactions, respectively. Looking to the thermodynamical data, the binding enthalpy and entropy obtained for LC3B interaction with Ub\_p62-LIR ( $\Delta H = -10.4 \text{ kcal mol}^{-1}$  and  $\Delta S = -8.3 \text{ cal mol}^{-1} \text{ K}^{-1}$ ) were very similar than those obtained with the synthesized p62-LIR ( $\Delta H = -10.5 \text{ kcal mol}^{-1}$  and  $\Delta S = -8.7 \text{ cal mol}^{-1} \text{ K}^{-1}$ ) (**Figure 50**). Regarding to these data, ubiquitin did not interfere in the interaction of LC3B with p62-LIR and Ub\_p62-LIR could be used for interaction studies by ITC instead of synthetic p62-LIR peptide.



**Figure 50: Effect of the ubiquitin moiety on the interaction of Ub-p62 with LC3B.** The raw data and the thermodynamic values of the titration of p62-LIR against LC3B (left) and of Ub\_p62-LIR against LC3B are shown. All experiments were performed at 25°C in the same conditions as the GABARAPL-1 interaction studies with NBR1-LIR wild type and mutants.

To confirm the non-influence of ubiquitin in the ITC data obtained with Ub-fused proteins, a control titration experiment of ubiquitin against GABARAPL-1 was done (**Figure 51**). In this case, some unspecific binding effects were observed with a dissociation constant in the hundred micromolar range with a number of binding site fixed to one and a dissociation constant in the millimolar range without a fixed number of binding sites. These dissociation constant values are insignificant regarding to the strong binding of GABARAPL-1 with Ub\_NBR1-LIR and Ub\_NBR1-LIR\_Y732W. Indeed, Ub-fused NBR1-LIR constructs were used in chapter 6.2.3.3.2 to study the interaction of GABARAPL-1 with NBR1-LIR and NBR1-LIR\_Y732W. The

dissociation constants as well as the thermodynamic data were different between NBR1-LIR wild type and the mutant form. The only difference between both constructs is the substitution of the tyrosine by a tryptophan. This effect on the interaction by this single mutation was measured with the peptide linked to the ubiquitin leading sequence. The data obtained during ubiquitin titration were also used as a baseline for GABARAPL-1/Ub\_NBR1-LIR interaction to subtract the unspecific interaction from the total interaction.



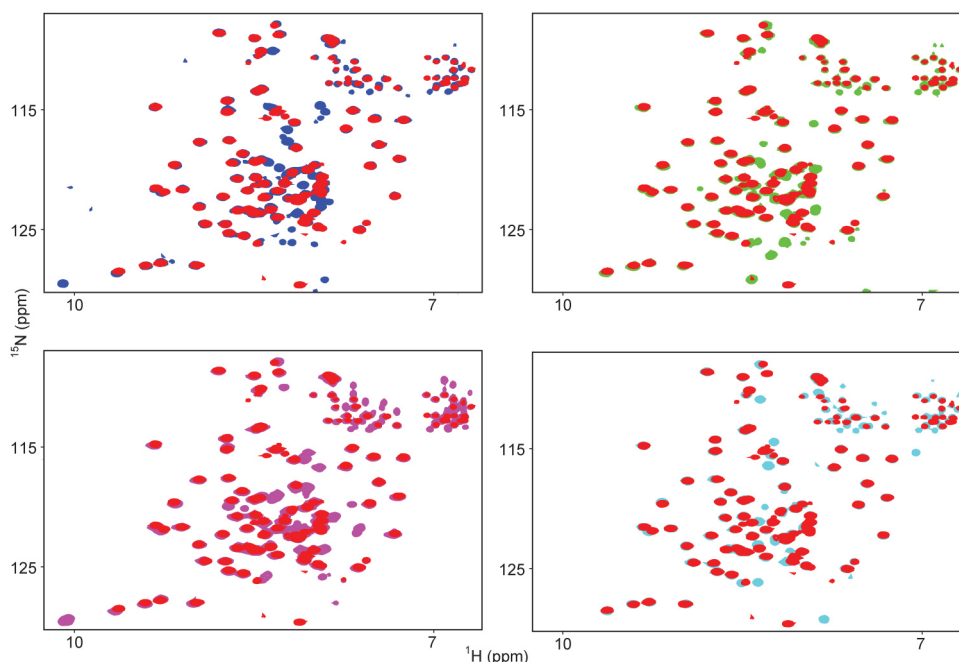
**Figure 51: Effect of the ubiquitin moiety on the interaction of Ub-fused peptides with protein.** The raw data and the thermodynamic values for the ubiquitin control experiment are shown. The titration of ubiquitin against buffer (dilution heat) on the left side and of ubiquitin against GABARAPL-1 on the right side are represented.

### 6.3.6. Use of ubiquitin-fused proteins/peptides for NMR

#### 6.3.6.1. Characterization of ubiquitin-fused proteins/peptides

To prove the benefit of Ub-fused constructs for NMR, the small hydrophobic peptides corresponding to the LIR domains of p62, NBR1, Nix-LIR\_W140/144 and of a new autophagy receptor, optineurin,<sup>105</sup> were first used. [<sup>15</sup>N, <sup>1</sup>H] TROSY-HSQC spectra were recorded for each of the precited peptides fused to ubiquitin. A number of extra resonances corresponding to the number of amino acids present in the peptidic sequence were observed in the [<sup>15</sup>N, <sup>1</sup>H] TROSY-HSQC spectra in addition to the resonances corresponding to ubiquitin (**Figure 52**). Thus for each peptide fused to ubiquitin, a singular NMR pattern was present as it will have been without the presence of ubiquitin. It should also be mentionned that because of its high flexibility/mobility, the resonances of the residues in the linker region containing the His<sub>6</sub>-tag

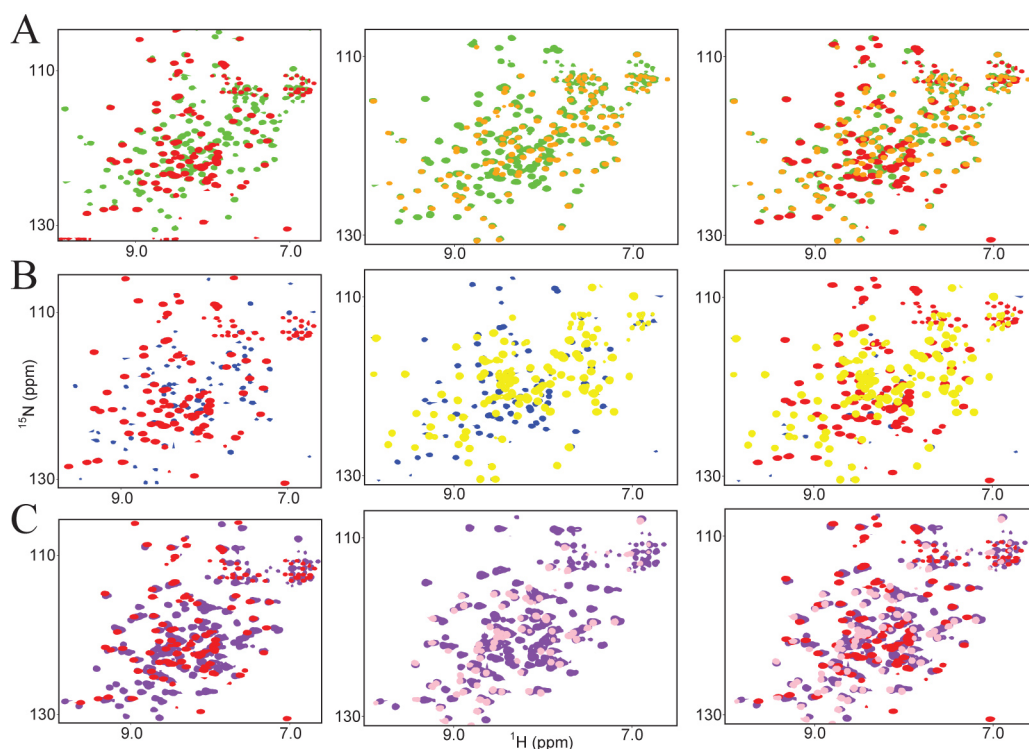
between ubiquitin and the protein of interest are not seen in the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra.



**Figure 52: NMR spectra of LIR domains fused to ubiquitin.** The overlay of [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of Ub\_p62-LIR (blue), Ub\_NBR1-LIR (green), Ub\_Nix-LIR\_W140/144 (magenta) and Ub\_optineurin (cyan) with ubiquitin (red) is represented. The resulting extra peaks observed in each spectrum reflect the specificity of each LIR motif.

What was true for small peptides was also applicable for proteins. LC3A was used as an initial example. First, a [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum of a purified LC3A fused to ubiquitin was recorded. Compared to the spectra of previously purified LC3A samples expressed as a NusA fusion construct and of ubiquitin from a previous work, the spectra are similar: Ub\_LC3A is the superposition of the previous [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of ubiquitin and LC3A. Except some overlays, the peaks belonging to ubiquitin or to LC3A were easily distinguished in the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum of Ub\_LC3A. The presence of ubiquitin in the sample did not interfere with LC3A because no chemical shift perturbations appeared regarding to the reference [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of ubiquitin and LC3A alone. The peaks had the same intensity and the same shape for the protein alone as well as for the Ub-fused LC3A. The size of the full construct was not an issue for this NMR experiment. Some slight differences were anyhow observed but only caused by the small differences in buffer compositions and pH. The same results as for Ub\_LC3A were observed for Ub-fused constructs used with other proteins like LC3B (although the purification protocol lead to a large amount of lost protein) or for TBK1\_ULD. The [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of Ub-fused proteins showed the same

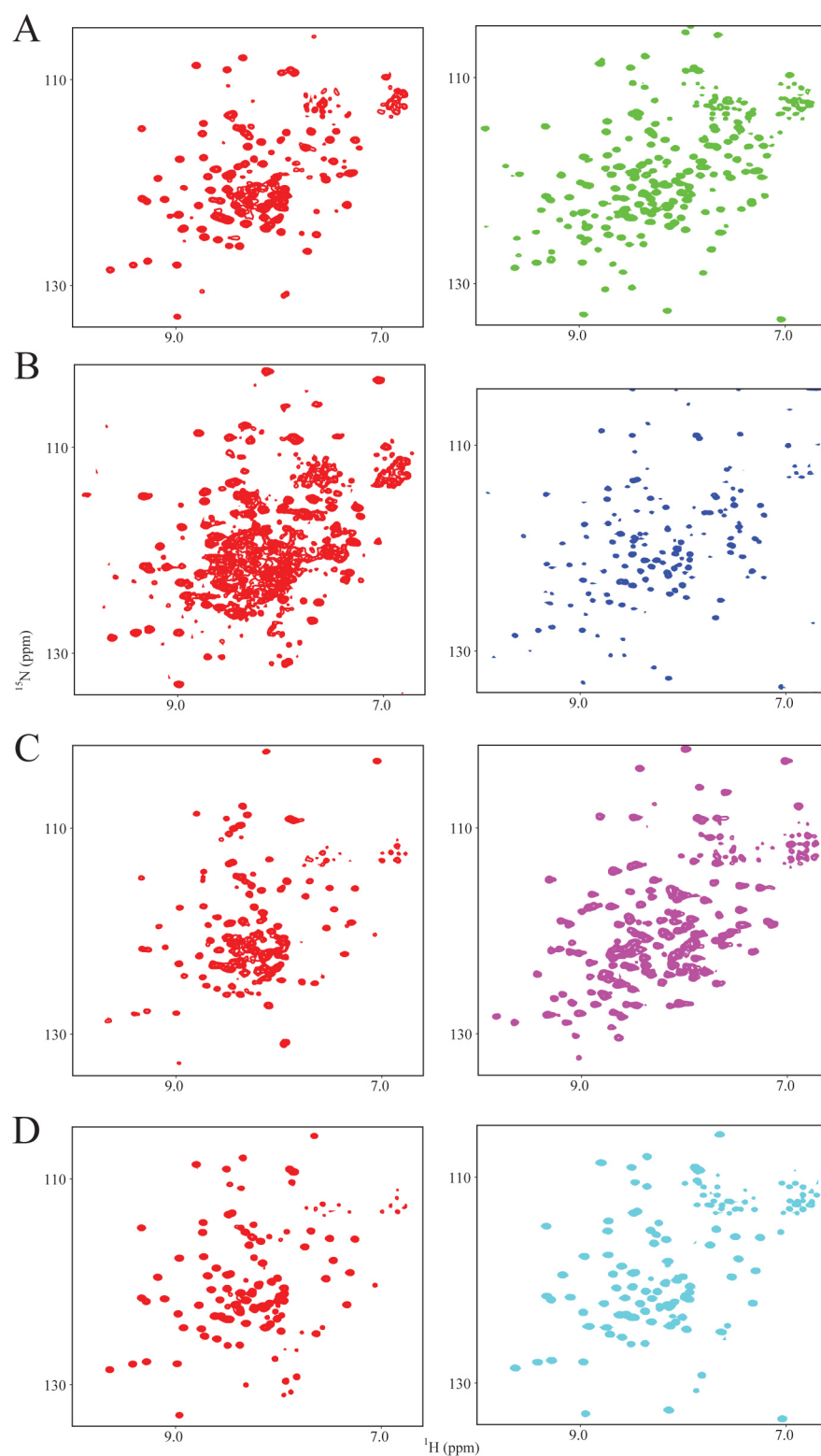
pattern that their own reference [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra with extra resonances belonging to ubiquitin only (**Figure 53**).



**Figure 53: NMR spectra of Ub-fused proteins.** The overlay of [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of Ub-fused protein with ubiquitin (left), with the pure protein (middle) and with both (right) is shown. In every spectrum, ubiquitin is represented in red. In (A), Ub\_LC3A is in green and LC3A alone in orange. In (B) Ub\_LC3B is in blue and LC3B alone in yellow. In (C), Ub\_TBK1\_ULD is in purple and TBK1\_ULD alone in pink. Except some small CSP caused by the buffer conditions, the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of the protein fused to ubiquitin is totally covered by the spectra of ubiquitin and of the concerned protein alone.

#### 6.3.6.2. Use of ubiquitin-fused constructs from cell lysate

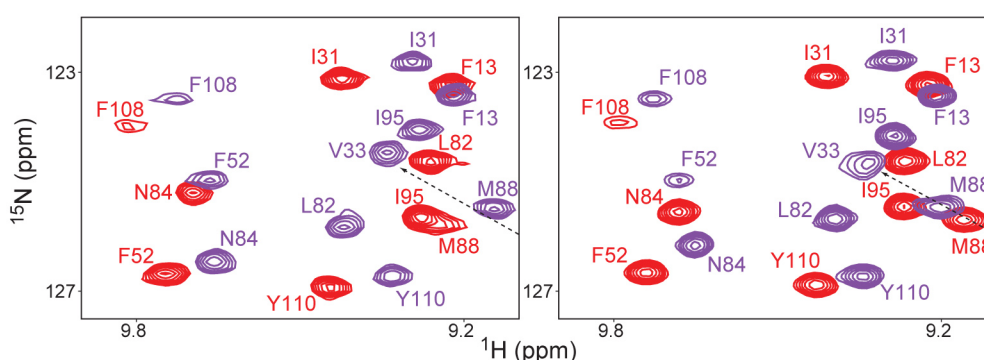
After the expression and the harvest of only 100 mL of growing cells containing the DNA encoding for Ub-fused proteins, a [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum was recorded for the supernatant of the cell pellet resuspended in the NMR buffer after sonication and centrifugation. The [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra obtained for the different Ub-constructs prepared in less than one hour showed a nice peak dispersion, which corresponds to the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of the same proteins obtained after few days of purification. Naturally, some background was observed for all spectra but, in most of the cases, with less intensity than the peaks corresponding to the target proteins. The same results were also obtained by using only 3 mL of protein culture. Surprisingly, Ub\_LC3B provided spectra with less peak intensity and with more background than for Ub\_LC3A whereas the expression yield observed by SDS-PAGE was similar (**Figure 54**).



**Figure 54: NMR spectra from the cell lysate of Ub-fused proteins.** The overlay of [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of Ub-fused proteins from cell lysate (left) and of Ub-fused proteins after purification (right) is shown. LC3A, LC3B, TBK1\_ULD and optineurin, all fused to ubiquitin, are represented in (A), (B), (C) and (D), respectively. A similar peak distribution between the cell lysate and the purified Ub-fused proteins is observed but with less background and sharper peaks for the purified proteins (except for Ub\_LC3B with the loss of material during the purification protocol).

In the case of Ub\_p62-LIR cell lysate, several short [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra have been recorded over 48 hours to monitor the appearance of degradation over time. Only a few resonances corresponding to degradation products were present after this time. Most of the resonances were still present in the last [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum proving that the sample was stable in these conditions for more than 24 hours (results not shown). Adjustments of the buffer conditions with more protease inhibitor cocktail and more reducing agents were done on the sample preparation to protect it from degradation. Thus an enhanced stability of the protein was observable in the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra.

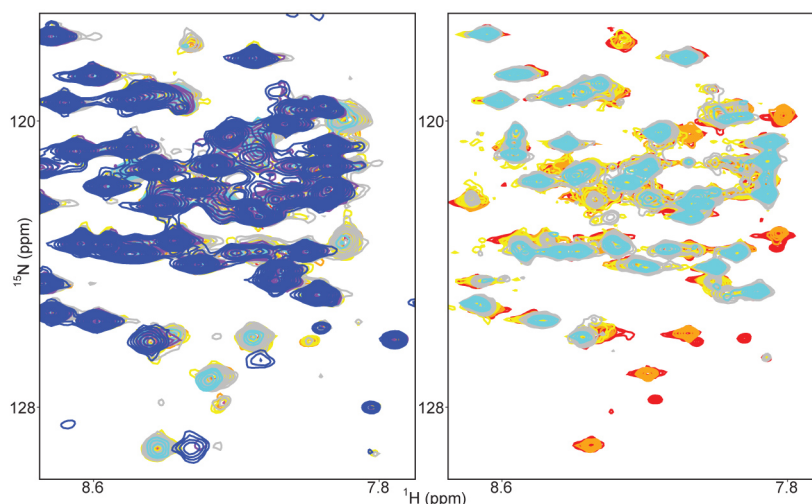
### 6.3.6.3. Interaction studies



**Figure 55: LC3B interaction with p62 and Ub\_p62.** The NMR titrations of p62 (left) and Ub\_p62 (right) against LC3B (protein alone in red and in presence of LIR in purple) show a similar pattern, reflecting that ubiquitin does not interfere in the interaction. The small differences observable are caused by slight buffer differences.

Another possible NMR application for a protein fused to ubiquitin is to study protein-protein interactions. In chapter 6.2.3.1.1, the interaction of LC3B with chemically synthesized p62-LIR was characterized. This NMR titration experiment was repeated but with a Ub-fused peptide this time. The same titration pattern than previously was observed with the disappearance of the cross-peaks corresponding to the free state and the appearance of the cross-peaks corresponding to the bound state upon titration as a proof for a strong interaction in a slow exchange mode. By a closer look on the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra, some very small chemical shift differences could be noticed between both [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra (**Figure 55**). A probable reason for these changes is the persisting presence of some chemical agents in the synthetic peptides, which could affect the general conditions of the sample.





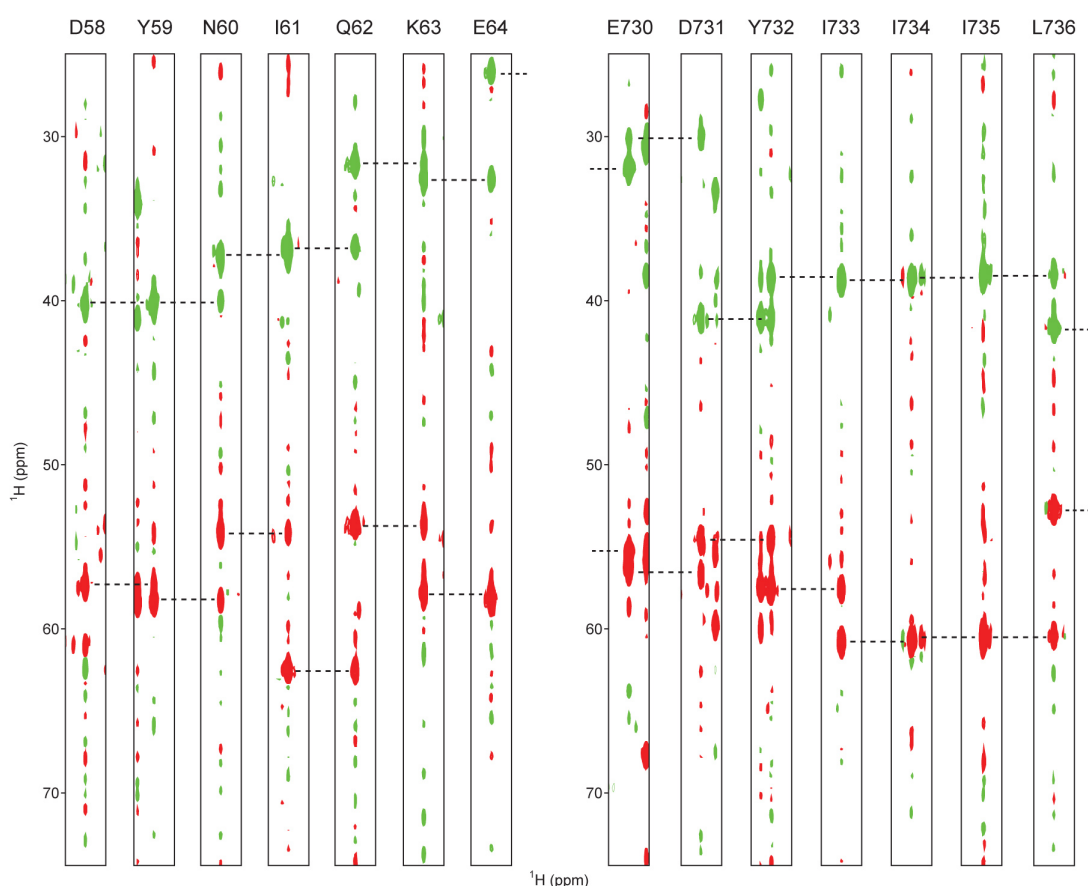
**Figure 56: Interaction studies of Ub\_NBR1-LIR with MAP1LC3 proteins.** The overlay of [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra of Ub-NBR1-LIR in presence of LC3B (left) or GABARAPL-1 (right) is represented. Contours in different colors correspond to individual titration points (red as reference, magenta in presence of MAP1LC3 proteins at molar ratio 1:1/8 (only for LC3B), orange at 1:1/4, yellow at 1:1/2, gray at 1:1, cyan at 1:2, purple at 1:4 (only for LC3B) and blue at 1:8 (only for LC3B)).

With the idea of first monitoring the effect of the interaction of MAP1LC3 proteins with the LIR domains of autophagy receptors to be able to then structurally determine this interaction, NMR interactions studies were performed by titrating unlabeled LC3B and GABARAPL-1 against labeled Ub\_NBR1-LIR. The interest was to observe the interaction from the LIR domain side. Thus a labeled LIR peptide was needed and to overcome the cost of chemically synthesized labeled peptide, a Ub-fused NBR1-LIR construct was used. The [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC reference spectrum of Ub\_NBR1-LIR corresponded to ubiquitin and 18 extra peaks as the number of NBR1-LIR residues. At the first titration steps, resonances corresponding to the residues from NBR1-LIR were shifted when the ones from ubiquitin were not. Pursuing the titration, most of the resonances corresponding to the NBR1-LIR moiety disappeared (**Figure 56**) like the resonances of GABARAPL-1 disappeared during the NMR titration of unlabeled NBR1-LIR peptide against labeled GABARAPL-1. Only the resonances corresponding to ubiquitin were then observed. What was first suspected to be a degradation of the peptide moiety was finally understood as the effect of the intermediate (close to slow) exchange mode of LC3B/NBR1-LIR as it was explained in more details in the chapters **6.2.3.1.3** and **6.2.3.3.1**. Approaching the saturation of the interaction, some peaks reappeared partially but the sample was to unstable as revealed by the appearance of precipitation in the NMR tube and signals of degradation products in the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectra. Interestingly, the resonances corresponding to NBR1-LIR disappeared already at a molar ratio of 1:1/2 in presence of GABARAPL-1 but only at 1:4 in presence of LC3B, indicating that GABARAPL-1 interacts

stronger than LC3B with NBR1-LIR. Because the same effect was observed for the interaction studies involving the LIR motif as peptide, these results corroborate the fact that ubiquitin does not interfere with the interaction in both cases. Here, performing a [ $^{13}\text{C}$ ,  $^1\text{H}$ ] HSQC could not have helped because the CH signals from ubiquitin are too many to be distinguished from the CH signals specific of NBR1-LIR.

As it was described above with ITC, GABARAPL-1 interaction with NBR1-LIR wild type and the Y732W mutant was also studied by NMR using peptides fused to ubiquitin as already described in chapter 6.2.3.3.1.

#### 6.3.6.4. Assignment of peptides using ubiquitin-fused constructs

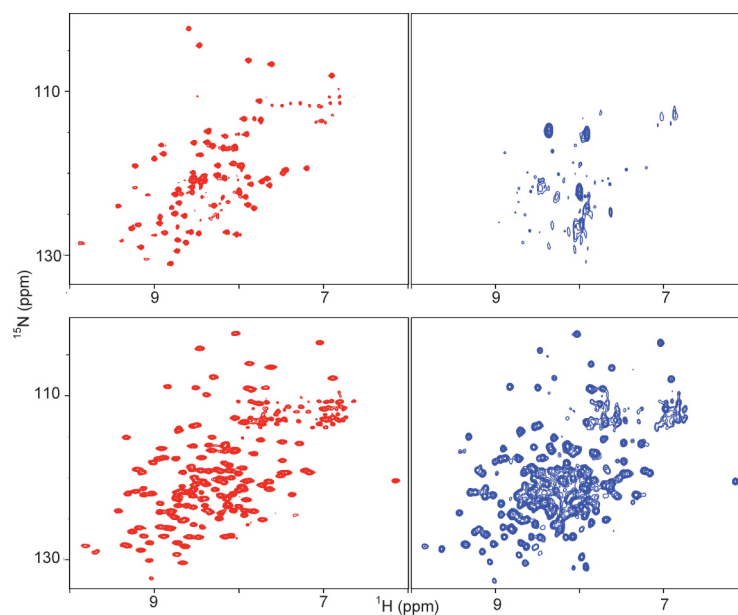


**Figure 57: Assignment of Ub-fused peptides.** The different strips represent the HNCACB correlation peaks for the stretch between D58 and E64 in the ubiquitin portion (left) and for the stretch between E730 and L736 in the NBR1-LIR portion (right) of Ub\_NBR1-LIR. The large presence of artifacts in the background of the strips is explained by the use of NUS-NMR to assign NBR1-LIR.

In comparison to existing ubiquitin assignments, the majority of the ubiquitin resonances from the Ub-fused peptides could be assigned with good presumption. Using traditional NMR experiments but also Non-Uniform Sampling (NUS) NMR, the backbone resonances of p62-LIR as well as of NBR1-LIR fused to ubiquitin were assigned and the assignment of

ubiquitin was also confirmed. In both cases, the resonances corresponding to the residues present in the loop containing the His<sub>6</sub>-tag and the TEV cleavage site between ubiquitin and the peptide were not totally assigned due to the fast flexibility of this region (**Figure 57**). Assignment of peptide side chains was started but further NMR experiments are still needed to obtain a total assignment of p62-LIR and NBR1-LIR moieties.

### 6.3.7. Solubility enhancement



**Figure 58: Stability of Ub\_TBK1\_ULD.** The overlay of [<sup>15</sup>N, <sup>1</sup>H] TROSY-HSQC spectra of TBK1\_ULD (top) and Ub-TBK1\_ULD (bottom) NMR spectra is shown. The initial NMR spectra of the TBK1\_ULD forms pure or fused to ubiquitin at ~300 μM protein concentration are represented in the left plots. The same samples were recorded after being stored 24 hours at 25°C and are shown in the right plots. The strong tendency of TBK1\_ULD to aggregate is decreased under fusion to ubiquitin.

In a purified form, TBK1\_ULD had shown a strong tendency to aggregate, already starting at a concentration of 200 μM. Moreover, TBK1\_ULD samples degraded really fast and after few days several resonances had already disappeared in [<sup>15</sup>N, <sup>1</sup>H] TROSY-HSQC spectra. Fused to ubiquitin, it was possible to concentrate the sample above 500 μM and the Ub\_TBK1\_ULD sample showed good stability after more than 24 hours at room temperature (**Figure 58**). No degradation products were observed in the different [<sup>15</sup>N, <sup>1</sup>H] TROSY-HSQC spectra recorded and no precipitation was observed in the NMR tube after few days unlike the purified TBK1\_ULD expressed with the GST expression vector.



## 7. Discussion

### 7.1. Expression and purification of proteins for biophysical studies

The function of a protein is, in most cases, directly related to its structure. It is commonly admitted that proteins sharing the same structure share also the same function and *vice-versa*. Protein interaction partners are very important in biological systems in order to activate or to inhibit the protein activity. Knowing the structure of a protein allows to assume structurally which molecules could be able to interact with this protein. Thus, the structure determination of proteins raises a high interest in the scientific community in order to get more fundamental knowledge about protein functions but also in the pharmaceutical industry, which is looking for the best interacting compounds as therapeutic products by structure-based drug design.

The Protein Data Bank (PDB: [www.pdb.org](http://www.pdb.org)) compiles structures of proteins and nucleic acids. More than 85% of the submitted structures were determined by X-Ray crystallography. X-ray crystallography is a powerful tool for the determination of the three dimensional structure of proteins and protein complexes of a high molecular mass but not all proteins provide crystals and not all crystals are well-resolved by X-Ray crystallography. Moreover, because the physiological conditions of the protein are sometimes not suitable to obtain a crystal, substitutive conditions are used and could lead to structural artifacts. Finally, the biggest limitation of X-Ray crystallography is the impossibility to study the dynamic of a protein alone or in a complex.

The remaining structures present in the PDB were determined by NMR. Compared to X-Ray crystallography, the big advantage of NMR is the possibility to study the motion of a whole protein or only its sub-domains. Moreover, different NMR experiments are available providing different information about solvent exchange, kinetics and thermodynamics of protein interactions as well as 3D structure determination. The disadvantage of NMR is the necessity to work with a relatively low molecular mass (below 30 kDa) because of the complexity to analyze a such complex and because of the limitations of the technique concerning molecules presenting a low tumbling rate leading to broad overlapped peaks. Moreover, due to the poor sensitivity of the method, a high amount of material is necessary to perform NMR experiments. Expression, isolation and purification of peptides and proteins are the crucial initial steps to be able to further study their structural organization.

#### 7.1.1. *TBK1 and IRF3 proteins*

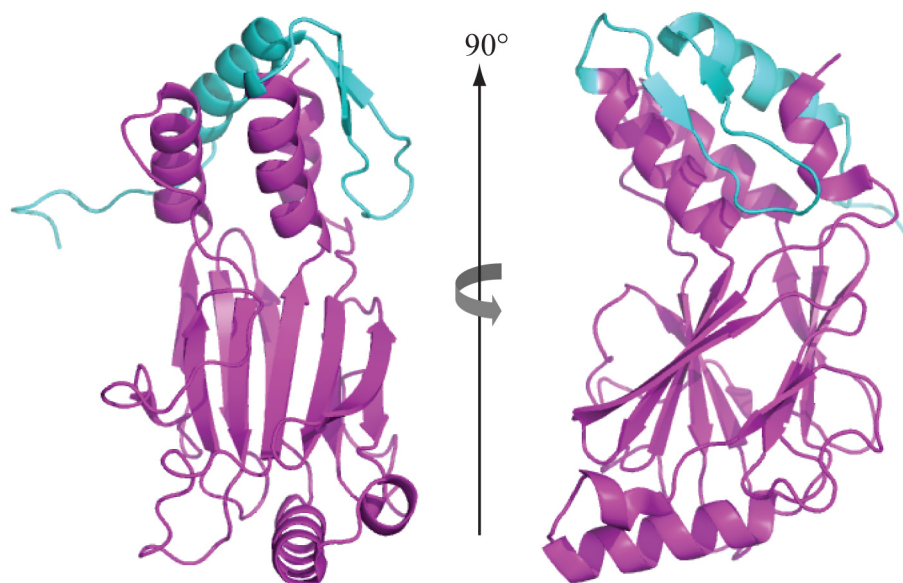
For the “TBK1\_ULD project”, the presence of an ULD in TBK1 had to be structurally confirmed. Other research laboratories tried to solve the structure of TBK1\_ULD using X-Ray

crystallography but the crystals they obtained did not diffract enough to be able to determine any structures.

Two TBK1 constructs containing the putative ULD were cloned into the pGEX-4T1 expression vector provided by Dr Fumiyo Ikeda to check which one could be the most preferable for NMR experiments. The first sequence (residues 302 to 383) contains only the predicted ULD whereas the second sequence consists of the ULD and the C-terminus of TBK1 (residues 282 to 403). The expression of the protein was observed in an analytical expression screening. The advantage of a longer sequence is its probable higher expression yield due to a better stability but no expression improvement being noticed, the shorter sequence was preferentially chosen because of the prevision of an easier NMR assignment by reducing the number of atoms to assign. The expression and the first purification steps on a preparative scale supplied only the minimal amount of material required for such an experiment. Removal of GST by thrombin protease was not always sufficient and more enzyme than expected were needed, rising up the experiments costs. Moreover, the solubility of TBK1\_ULD released from GST moiety was low with some precipitation observed after overnight cleavage of the elution sample of TBK1\_ULD. Nevertheless, pure TBK1\_ULD sample concentrated up to 300  $\mu$ M were produced and used for NMR studies. Increasing the solubility of the TBK1<sub>302-383</sub> protein sample was still a need to complete the NMR structural studies. Several attempts were done to improve the solubility by substituting residues at the N- and C- termini of TBK1\_ULD. Unfortunately, these different mutations did not help and had even sometimes a worse effect on the protein solubility. Other attempts to increase the expression yield were also performed but the substitution of the GST expression vector to the NusA expression vector did not help neither. Even though the expression yield was similar or better than for GST\_TBK1\_ULD, most of the protein was present in inclusion bodies and resolubilization processes did not work out. The use of an other expression vector was more promising. Fused to ubiquitin, the expression yield of TBK1\_ULD was not significantly higher but the sample could be higher concentrated and was stable for a longer time without showing any precipitation. The same Ub\_TBK1\_ULD sample was thus used for several experiments when a new TBK1\_ULD sample had to be prepared for each new NMR experiment.

Ikeda et al.<sup>120</sup> showed by GST pull-down assays that the IAD domain of IRF3, a binding partner of TBK1, was enough to bind to the ULD domain of TBK1. Different constructs corresponding to IRF3 full-length, the IAD domain alone and the IAD followed by the SRR domain were provided. Unfortunately, purification of IAD failed out, which could be explained by the structure of IRF3. The SRR domain folds over the IAD domain and probably stabilizes the whole protein (**Figure 59**).<sup>150</sup> The expression and the purification of IAD-SRR were more

successful and the protein was then used for preliminary interaction studies by NMR. Due to the beneficial effect of the ubiquitin expression tag, the expression of IAD-SRR fused to ubiquitin should be tried in the future and could probably lead to a better expression of the protein and further interactions studies could be performed.



**Figure 59: Crystal structure of IRF3\_IAD-SRR.** The structure of the IAD-SRR domain of IRF3 is presented as a ribbon diagram in two orientations (the view on the right corresponds to a 90° rotation). IAD and SRR domains are represented in magenta and cyan, respectively. The SRR domain interacts with the IAD domain to build a compact structure. (picture adapted from Qin et al.<sup>150</sup>)

### 7.1.2. MAP1LC3 proteins and LIR domains

The different protein and peptide sequences provided by Dr. Vladimir Kirkin and Dr. David McEwan were cloned into pGEX-4T1 expression vectors. Although the constructs expressed well, DNA fragments coding for the different MAP1LC3 proteins and the LIR domains were cloned into the expression vector pETM-60, presenting the NusA protein to enhance the expression, a hexahistidine tag for purification and a TEV protease recognition sequence. The NusA expression tag had a great impact in the final expression yield of these proteins and the cleavage rate with the TEV protease was fast and efficient. Pure protein in large amount was more easily obtained than in the case of GST based constructs. As a matter of proof for the ubiquitin-tag project, these proteins were also expressed as ubiquitin-fused proteins and lead to an even larger expression yield compared to the same proteins fused to GST and NusA expression vectors.

To understand the role of the different residues present in the LIR domain of autophagy receptors, the production of small peptides up to 20 amino acid length was needed. First plasmids containing the coding sequence of the NBR1-LIR protein with different lengths (46 to

276 amino acids) expressed well as GST-fusion proteins but, after removal of the GST moiety, no protein was found in solution. Using the different techniques available in house (SDS-PAGE or size-exclusion chromatography), no protein was detected, probably due to protein degradation after cleavage of the expression tag. Ordering synthetic peptides was then the best option at this time. p62-LIR, Nix-LIR\_W36, Nix-LIR\_W140/144 as well as NBR1-LIR wild type and some mutants were thus used for ITC and NMR titration. Because the LIR peptides were also needed for structure determination by NMR, labeled peptides would have to be produced what would have been a huge cost if ordered as the non-labeled ones were. The use of peptides fused to ubiquitin appeared to be a great method to produce labeled peptide at a lower cost.

Unfortunately, good protein concentration does not necessarily lead to a good NMR spectrum. A first set of experiments was often needed to optimize the buffer conditions in terms of buffer system, salt concentration and pH. The conditions were optimized until obtaining a [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY-HSQC spectrum presenting nice resolved, globular and well-dispersed resonances. The chemical shift dispersion is indicative of a well-folded and of a globular domain but several centers and broad peaks are synonyms of dynamics, which could complicate the assignment of the protein.

Some peaks were sometimes not seen like the N-terminal part of the different MAP1LC3 proteins because of conformational exchange. Generally, additives could increase the solubility of the protein like an arginine/glutamic acid mixture or detergents like CHAPS stabilize hydrophobic surface with reducing conformational exchange leading to sharper peaks with better shape and diminution of double peaks. Although the change of the conditions did not improve the behavior for MAP1LC3 proteins, addition of arginine/glutamic acid showed improvement in the NMR spectra of TBK1\_ULD.

Thus, all the proteins required for the “TBK1\_ULD” and the “Autophagy” projects were expressed and purified. Different expression vectors and different buffers had to be used to get to these results. Good NMR spectra were recorded for all these proteins, meaning that the proteins were well-folded under the purification conditions. This had made possible to study further these proteins with biophysical methods. NMR was successfully used during this thesis as technique to determine the secondary structure of the putative ULD domain of TBK1 when X-Ray crystallography failed. NMR was also used to structurally characterize the GABARAPL-1/NBR1-LIR complex. Finally, the description of the interaction of mammalian Atg8 proteins with different LIR domains was done by NMR and ITC.



### 7.1.3. Ubiquitin as solubility tag suitable for biophysical methods

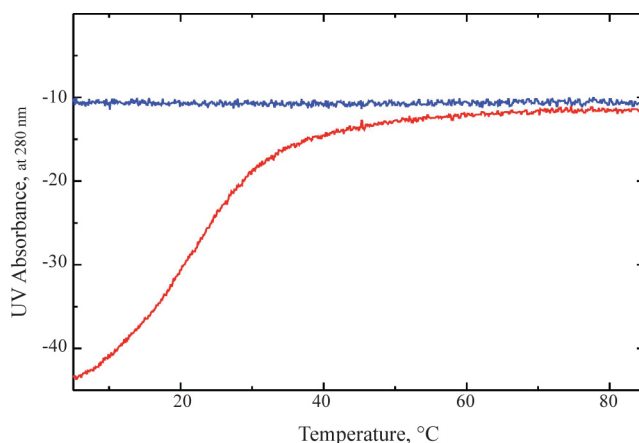
Most biophysical applications require to remove the expression tag and to operate with highly purified proteins. The possibility to lose the benefits of a high yield protein within fusion constructs increases with each purification step as it was showed for TBK1\_ULD, which precipitated after the removal of the GST moiety. It was even worth concerning short peptides like for NBR1-LIR, which could not be easily treated with the most common methods provided in biochemistry laboratories. Chemical laboratories could easily synthesize peptides for reasonable price but it becomes very expensive if they are labeled ( $^{15}\text{N}$ ,  $^{13}\text{C}$  or both) and companies even failed to synthesize some of the peptides.

Ubiquitin, presenting a fast and high expression yield as well as a high protein stability, possesses the ideal qualities to be used as an expression tag. With the addition of slight modifications within ubiquitin, the internal stability and solubility of the expressed fusion constructs was increased.<sup>151</sup> First experiments showed that proteins (LC3A and LC3B) as well as peptides (Nix-LIR\_W36 and Nix-LIR\_W140/144) are expressed in a very large amount (100 mg mL<sup>-1</sup>) without any degradation products. Ubiquitin could be successfully removed from the fused constructs and, after few purification steps, pure proteins and peptides could be produced even with isotopic labeling. With the establishment and the optimization of purification protocols, the high expression yield obtained with Ub-fused peptides provided large amount of pure peptides useful for biophysical studies. Pure NBR1-LIR was produced this way in order to determine the structure of GABARAPL-1 in presence of NBR1-LIR by NMR.

The main idea behind ubiquitin behaving as an expression tag was to use the full construct without any cleavage for biophysical methods like NMR, ITC or CD spectroscopy. With ubiquitin being thermodynamically very stable, its effect during NMR, CD spectroscopy and ITC experiments should be always the same. Recording blank experiments with ubiquitin gave information of its background contribution during the experiments with Ub-fused protein/peptide. Due to its small size (8 kDa against more than 50 kDa for the dimeric GST and the monomeric NusA) and its stable fold, ubiquitin should not affect NMR spectra like a bigger tag would do. Indeed, big proteins have a slow tumbling rate leading to fast relaxation and thus peak broadening. The presence of GST\_IAD-SRR in the titration against TBK1\_ULD lead to the disappearance of all resonances in the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY\_HSQC spectra for example. To show that ubiquitin fulfills these conditions, the fact that the target protein possesses the same features within the fusion construct as well as being purified was proven.

### 7.1.3.1. Ubiquitin-fused constructs for CD spectroscopy

CD spectroscopy was performed to verify if the structural information provided for the Ub-fused constructs was actually corresponding to the sum of the structural characteristics of ubiquitin and of the target protein alone. Using GABARAPL-1 as example for an ubiquitin-fused protein, the CD spectrum of Ub\_GABARAPL-1 was perfectly the addition of ubiquitin and GABARAPL-1 spectra. No additional structural elements appeared or disappeared due to the fusion of GABARAPL-1 to ubiquitin. Another application possible is to check the thermal stability by CD spectroscopy. Unfortunately the data provided by the melting curves of Ub\_GABARAPL-1 were not useful for a clear statement but, in theory, the subtraction of the melting curve of ubiquitin from the one of the Ub-fused protein should give information on the thermal stability of the protein. Experiments done by Dr. Vladimir Rogov with Ub\_ABIN1 were more positive concerning the possibility to study thermal stability of protein based on the results obtained with the Ub-fused protein (**Figure 60**).<sup>151</sup> Point mutations on a protein are often performed to prove the functional importance of specific residues of a protein and can also be useful to enhance the expression and solubility of the protein for further structural studies. The Ub-fused construct could be also used as a fast method to determine by CD spectroscopy if the mutations have any effect on the protein folding after short protein expression and without big purification steps.



**Figure 60: Melting curves of ABIN1 fused to ubiquitin.** The only observed thermal transition in the Ub\_ABIN1 melting curve (red) is due to ABIN1. The small size of ubiquitin and its thermal stability (ubiquitin melting curve in blue) at biologically relevant conditions are crucial advantages of this system in comparison to the NusA- or GST-tags in CD studies.

### 7.1.3.2. Ubiquitin-fused constructs for ITC

ITC furnishes important information on the interaction between two molecules like the dissociation constant and the thermodynamic values of the interaction. As it was previously explained, production of sufficient materials to perform the experiment could be a limiting

factor. If the protein/peptide is not an ubiquitin binding partner, an Ub-fused protein/peptide is totally suitable to study interactions using ITC. Dissociation constants, binding enthalpy and entropy obtained for LC3B interaction with Ub\_p62-LIR and with synthesized p62-LIR in the same experimental conditions were very similar, proving that the ubiquitin moiety did not interfere in the interaction. The observed slight differences could be explained by an error due to the uncertainty in the sample concentration. Within their fusion to ubiquitin, the effect of single mutations in the interaction of NBR1-LIR with GABARAPL-1 was successfully studied.<sup>152</sup>

### ***7.1.3.3. Ubiquitin-fused constructs for NMR***

Due to its stability and its small size, ubiquitin is well characterized by NMR and frequently used as a model for new NMR experiments.<sup>153</sup> The use of different MAP1LC3 proteins fused to ubiquitin has shown that the peak dispersion in NMR experiments was good enough to distinguish resonances belonging to the protein target from the ubiquitin moiety based on published ubiquitin assignment (BMRB accession number 4768). To avoid the loss of protein through several purification steps, NMR experiments for proteins issued directly from the cell lysate were performed. The resulting [<sup>15</sup>N, <sup>1</sup>H] TROSY\_HSQC spectra were very similar compared to the ones obtained from pure proteins. Naturally, because of the absence purification steps, more background was present in the [<sup>15</sup>N, <sup>1</sup>H] TROSY\_HSQC spectra. Interestingly, LC3A and LC3B presented different behaviors in cell lysate when they were fused to ubiquitin. While LC3A had relatively no background, the [<sup>15</sup>N, <sup>1</sup>H] TROSY\_HSQC spectra of LC3B were more crowded. In the case of Ub\_LC3B, the loss of quality resolution in the spectrum could be due to the formation of a large protein complex with remaining endogenous proteins, which is hardly observable in [<sup>15</sup>N, <sup>1</sup>H] TROSY\_HSQC spectra. Using protease inhibitors and sodium azide, Ub-fused proteins or peptides isolated from the cell lysate could be used for 24 hours measurements without any problems. For measurements needing longer time (approximately a week), further purification steps should be added to avoid any degradation.

Structure determination of proteins is often laborious and bioinformaticians try to automatize this work. First tries using “traditional assignment methods” with Ub-fused peptides (Ub\_p62-LIR and Ub\_NBR1-LIR) were already performed and most of the backbone resonances could be assigned. These constructs were in the meantime transferred to the group of Prof. Peter Güntert (Goethe University Frankfurt), which used computational methods to study biomolecular systems. Applying NUS-NMR methods, which only need 10 to 25% of the total spectroscopy time of traditional experiments, in combination with automated assignment protocols (FLYA<sup>154</sup> and AUTOASSIGN<sup>155</sup>), the resonances of Ub\_p62-LIR and Ub\_NBR1-LIR

backbone and side chains were assigned. Both methods provided the same results but the speed of execution plays in favor of NUS-NMR (Table 7).<sup>156</sup> Further experiments using Ub-fused proteins with higher molecular weight than p62-LIR and NBR1-LIR should be performed to confirm the beneficial effect of Ub-fusion coupled to automated NMR methods for assignment and structure determination. Thus, this approach allows using Ub-fused targets of any length independently from rates of resonance overlapping.

Experiments	Time in hour	Gain of Time
hncoca	1.2	15 %
ihnca	2.4	15 %
hnco	1.0	10 %
hncaco	8.0	20 %
hncocacb	7.5	20 %
ihncacb	9.4	25 %
hbhacbcaonh	11.9	20 %
hbhacbcanh	11.9	20 %
ihncaco	10.0	25 %
ctocsynh	14.3	18 %
htocsynh	14.3	18 %
ihbhacbcanh	14.9	25 %
ictocsynh	15.9	20 %
ihtocsynh	15.9	20 %

**Table 7: NUS-NMR experiments for the assignment of Ub\_NBR1-LIR.**

NUS-NMR was used to assign more than 80 % of the backbone and side chain resonances of Ub\_NBR1-LIR. Only 20 % of the total spectroscopy time needed with traditional NMR experiments was used.

With Ub-fused constructs, titration experiments were also performed showing the same pattern than using pure peptides cleaved from their expression tag. As it was shown for ITC experiments, interaction studies of LC3B with p62-LIR by NMR showed the same results using synthesized peptide and Ub-fused peptide with the resonances corresponding to LC3B affected in a similar manner. Some differences were also seen but most probably due to buffer effect like remaining of purification solvent for the synthesized peptide. To understand the effect of the aromatic residue in the LIR domain, unlabeled NBR1-LIR and NBR1-LIR\_Y732W were used as ubiquitin fused peptides to titrate GABARAPL-1 by NMR. Titration of both Ub\_NBR1-LIR and Ub\_NBR1-LIR\_Y732W against GABARAPL-1 showed different patterns with different binding modes. Because the titration of ubiquitin against GABARAPL-1 did not show any interaction, these differences were not caused by the ubiquitin part but by the only difference between the both sequences, the substitution of tyrosine to tryptophan. The result of the titration was in

adequation with the ITC results showing a stronger interaction for the tryptophan mutant form of NBR1-LIR.

Thus, labeled and unlabeled proteins or peptides fused to ubiquitin are suitable for NMR studies. Respecting the NMR limiting factor of an overall molecular mass lower than 30 kDa, a nice peak dispersion was seen and the extra resonances were easily recognizable as belonging to the protein/peptide moiety. If ubiquitin is not interfering in the interaction, ubiquitin-fused constructs could be used as well for interaction studies, especially when short peptides are needed.

## 7.2. TBK1

### 7.2.1. TBK1\_ULD structure

To confirm the presence of an ULD in TBK1 as predicted *in silico* by Dr. Kay Hofmann, the structural characteristics of this domain were studied by NMR. NMR spectroscopy of TBK1\_ULD was really challenging because the protein was poorly soluble and tended to form aggregates. The different attempts to improve the quality of TBK1\_ULD samples were unsuccessful to determine the three dimensional structure of the protein. Amino acid mutagenesis or use of additives did not allow to reach higher protein concentrations needed to perform 3D NOESY experiments with a good quality necessary for the full assignment of TBK1\_ULD, prior to any structure calculation. To determine the secondary structure of the potential ULD domain of TBK1 under these conditions, calculations were only based on the assignment of the chemical shift values of TBK1<sub>302-385</sub>. Using selective labeling, the backbone of TBK1\_ULD was assigned and the secondary structure elements were identified by NMR spectroscopy. The secondary structure elements of TBK1\_ULD have the same sequential arrangement ( $\beta\beta\alpha\beta\beta$ ) and the same length as the corresponding secondary structure elements in ubiquitin. TBK1 belongs thus to the UBL superfamily, especially the type II of ULD, defined by the presence of an ULD as an element of a bigger structure. Using the known ubiquitin hydrogen bonds and by transposing them onto the TBK1\_ULD structure, first calculations for the three dimensional structure determination of TBK1\_ULD confirmed also the Ub-fold. Nevertheless, the full backbone and side chain assignment of TBK1\_ULD has still to be achieved to determine the real structure. Improvement in the quality of TBK1\_ULD NMR sample is required first. Initial expression and purification assays with TBK1\_ULD fused to ubiquitin were promising and further improvements in the method could help to provide enough pure TBK1\_ULD for structural studies.

Despite these structural similarities between the ULD of TBK1 and ubiquitin, differences exist in their dynamic behavior, reflected by broader signals for TBK1\_ULD as for ubiquitin. This difference observed on the structured part of TBK1\_ULD, especially on the  $\beta$ -sheet, could be important for the binding of the protein, which depends on the Ub-fold. This difference is enhanced by the presence of an enlarged hydrophobic surface of the ULD in TBK1, concentrated around the conserved hydrophobic patch on the  $\beta$ -grasp fold (352-354 residues LIY in TBK1\_ULD and 43-45 residues LIF in ubiquitin), while the rest of ubiquitin and TBK1\_ULD amino acids have a low sequence similarity. This conserved motif is important because it is known to play a role in the interaction with binding partners for ubiquitin and other UBL. This bigger hydrophobic surface could be one of the reason why the ULD domain of TBK1 precipitated quickly and was not easily purified as it could be for ubiquitin.

### **7.2.2. *TBK1\_ULD* vs. *IRF3\_IAD-SRR***

Ikeda et al.<sup>120</sup> showed that TBK1\_ULD was interacting with the IAD domain of IRF3 using GST pull-down assays. The structure determination of the TBK1\_ULD/IRF3\_IAD complex should help to understand how TBK1 recruits its substrate through the ULD for later phosphorylation via kinase domain of TBK1. ULD binds to IAD alone but the crystal structure of IRF3 showed that the core of the protein is a  $\beta$ -sandwich, which is covered up on one side by the association of the structural elements of the IAD and the SRR domains stabilizing the hydrophobic surface provided by  $\alpha$ -helices of IAD.<sup>122; 150; 157</sup> That could explain why the IAD domain was predicted to be alone very unstable and why only IAD-SRR could be isolated. The difficulties to obtain the structure of TBK1\_ULD added to the instability of IAD-SRR complicated these studies. Nevertheless, the preliminary interaction studies showed that the TBK1-ULD/IRF3 interaction is a more complex process than a one to one binding mode with the whole TBK1\_ULD molecule affected by the presence of IAD-SRR.

The main characteristic of TBK1\_ULD is its direct function as a regulator of the kinase domain of TBK1 but also to bind to other IKK-related kinases. No characterization of TBK1\_ULD interaction with its kinase domain has been determined so far. However, the presence of the hydrophobic patch in TBK1\_ULD seems to be essential for IRF3 to be phosphorylated. Mutations on this hydrophobic motif reduce the phosphorylation activity of TBK1. IRF3 does not dimerize, does not become active and does not transfer to the nucleus, as it appears with TBK1 wild type. In contrast, mutation of the hydrophobic patch does not affect the binding of TBK1\_ULD to IRF3. TBK1\_ULD plays a crucial role as a protein-protein interaction

domain in the regulation of the function of TBK1 by recruiting substrates to be phosphorylated in one hand and by controlling the activity of TBK1 kinase domain in the other hand.<sup>120</sup> This double function of TBK1\_ULD by interacting with its kinase domain as well as with IRF3 could explain why TBK1\_ULD seems to be affected on both sides in presence of IAD-SRR. Thus, a model of the interaction and of the activity regulation between TBK1 and IRF3 could be explained as follows: first IRF3 binds weakly to the ULD of TBK1 via its IAD domain; this fleeting binding brings the kinase domain of TBK1 close to the SRR domain of IRF3, which is then phosphorylated; finally this post-translational modification has for consequence that IRF3 dimerizes via its IAD domain.<sup>120</sup>

Complementary studies should be done for a better understanding of these interactions by characterizing the binding affinity and mapping the interaction on the surface of TBK1\_ULD with IRF3 but also with its own kinase domain.

### ***7.2.3. TBK1\_ULD differs from most common ULDs***

TBK1 belongs to the type II of UBL proteins where the ULD is “only” an ubiquitin-like region in the sequence of the protein. In contrast, type I UBL are post-translational modifiers conjugated in a similar way than ubiquitin. UBL are defined by their Ub-fold but some structural differences are still observed between them, which are categorized in three sub-families of ULD: RAD23 (and TBK1\_ULD), very similar structurally to ubiquitin; ubiquitin-regulatory X domain (UBX) with an expanded surface loop; PB1, which is part of a higher order structure. Although TBK1\_ULD structure is close to ubiquitin (RMSD over  $\alpha$ -helix 1 and  $\beta$ -strands 1 to 4 is 1.5 Å and the global RMSD is 3 Å), their primary amino acid sequences differ a lot and TBK1\_ULD presents a different function as well as other interaction partners. Other proteins presenting a  $\beta$ -grasp superfold have a broad range of functions, from a role in protein folding and degradation to the regulation of signal transduction and enzymatic activity. For proteins presenting an Ub-fold, it could be optimistic to think that the substitution of ULD by ubiquitin would have no effect on the protein function. It is not the case for TBK1 since ubiquitin does not bind to IRF3 as TBK1\_ULD. This is different for each ULD. While ubiquitin could substitute the function of RAD23\_ULD, the function of IKK $\beta$  is abolished in the presence of ubiquitin instead of IKK $\beta$ \_ULD. Interestingly, proteins possessing an ULD bind also to ubiquitin via ubiquitin-binding domains (UBD) but not for TBK1. All together, the unique structural features of TBK1\_ULD could explain the functional specificity of this ubiquitin homolog.

### **7.3. Autophagy**

The aim of this project was to characterize the interactions responsible for selective autophagy between the LIR domains of autophagy receptors with different mammalian homologs of the yeast autophagy effector Atg8.

The three-dimensional structure determination of the GABARAPL-1/NBR1-LIR complex added insights of the interaction at a structural point of view.

#### ***7.3.1. MAP1LC3 proteins and LIR peptides***

The first issue when studying proteins is of course to be able to have enough of pure protein. Even though the first plasmids provided could express nicely MAP1LC3 protein as GST fusions, the purification steps were challenging and did not offer enough materials. Replacement of GST as expression tag, first by NusA then by ubiquitin, could cover this problem and the obtained amount of MAP1LC3 proteins was sufficient for the studies. In parallel, peptides containing the LIR motifs of the different autophagy receptors had to be isolated. The first tries using different constructs of NBR1 as GST fusion produced some proteins but the purification was challenging or even unsuccessful. In a first time, ordering synthetic peptides was a solution to perform titration experiments. Nevertheless this solution would have been expensive when labeled peptides were needed for the structure determination studies. Expression of Ub-fused peptides and their use as free peptide or peptide still fused to ubiquitin was a successful solution to perform NMR and ITC studies.

#### ***7.3.2. Interaction studies***

The specificity of autophagy effector/receptor interactions is essential for selective degradation by autophagy. This interaction was shown to take place between the hydrophobic pockets of MAP1LC3 proteins and the LIR domain of autophagy receptors. The constant discovery of new autophagy receptors brings along a diverse composition of LIR motifs, which could lead to a big difference in their binding to MAP1LC3 proteins. Based on their studies on the structural basis of LC3B/p62 and Atg8/Atg19 interactions, Noda et al.<sup>87</sup> determined in autophagy receptor proteins a WxxL motif crucial for these interactions. The LIR motif consists of the following sequence xx $\Theta$ xx $\Gamma$  where  $\Theta$  and  $\Gamma$  are aromatic and hydrophobic residues, respectively, and x should represent at least one acidic residue.

Whereas ITC provided very interesting quantitative information about the thermodynamics of the interaction of MAP1LC3 proteins with the different LIR motifs, NMR brings the interaction studies at a structural level and shows the involvement of the residues in the interaction.



### 7.3.3. Differences between LIR motifs

#### 7.3.3.1. p62

Identified in parallel by different groups,<sup>88</sup> p62 is the prototypical LIR containing protein with three aspartates followed by a tryptophan as aromatic residue and a leucine as hydrophobic one. LC3B has been the most studied mammalian Atg8 family members and its structure in presence of p62 has been determined.<sup>88; 89</sup> Because the affinity of p62 to LC3B seems to be high enough to get the structure of complex by both X-ray crystallography and NMR methods, LC3B as mammalian Atg8 model and its interaction with p62 as LIR model were chosen as reference to be able to compare the binding of different LIR motifs to MAP1LC3 proteins. The NMR and ITC experiments confirmed the strong interaction of LC3B with p62 with a dissociation constant in the low micromolar range. The extremely negative value for the enthalpy plays in favor of the binding by increasing the binding surface. This effect is counterbalanced by a decrease of the mobility of both protein and peptide upon binding explained by the negative entropy value reducing the overall binding affinity.

p62 was the first autophagy receptor discovered and its interaction with MAP1LC3 proteins is the strongest, regarding to the LIR motif of other autophagy receptors. Autophagy has a crucial role in cell surviving and lack of p62 lead to autophagy deficiency. Thus, the tightly binding of p62 to MAP1LC3 proteins anchored into the autophagosome membrane make of p62 the essential autophagy receptor needed for a good functioning of selective autophagy.

#### 7.3.3.2. NBR1

For NBR1, it was shown that a YIII peptidic sequence was crucial for its interaction with all members of the MAP1LC3 protein family.<sup>93</sup> Interestingly in the LIR motif here, the aromatic residue is a tyrosine, the hydrophobic residue an isoleucine and only two acidic residues, glutamate and aspartate, are present before the aromatic residue. Finally, the presence of isoleucines at position 2 and 3 as well as a leucine at 5, provides a strong hydrophobic patch at the C-terminus of NBR1-LIR. To establish the importance of the aromatic residue at position 1, the binding behavior of the non-tryptophan NBR1-LIR domain was compared with the LIR domains of p62 and Nix using NMR and ITC.

Compared to p62, the interaction of LC3B with NBR1 presented a less favorable binding enthalpy, which is partially compensated by a positive entropy contribution. These differences in binding between NBR1- and the p62- LIR domains are likely caused by the more favorable hydrophobic character of the tryptophan present in p62-LIR and a reduced fitting to the hydrophobic pocket on LC3B due to the additional hydroxyl group of the tyrosine in NBR1-LIR.

The binding of the side chains of the NBR1-LIR domain to the surface and the hydrophobic pockets on LC3B is less optimal than for p62. To further study the specificity of the LIR motif of NBR1, GABARAPL-1 was chosen as MAP1LC3 protein. Already illustrated by a lower dissociation constant than in case of LC3B/NBR1-LIR interaction, the entropy and enthalpy values of GABARAPL-1/NBR1-LIR interaction showed that less hydrophobic interactions are made and that the binding complex is not as tight.

To understand the influence of the tyrosine as aromatic residue in a LIR motif for the binding of NBR1 to MAP1LC3 proteins, two NBR1-LIR mutants were created: NBR1-LIR\_Y732W and NBR1-LIR\_Y732F by substituting the tyrosine by a tryptophan and a phenylalanine, respectively. GABARAPL-1 was chosen as representant of MAP1LC3 proteins in this study because it was in parallel the best candidate for NMR structural studies. Both NMR and ITC titration experiments indicated that the presence of a tyrosine residue in position 1 in the NBR1-LIR domain instead of a tryptophan residue in most LIR motifs results in an increased flexibility in the interaction with MAP1LC3 proteins. The less positive entropy in ITC for the tryptophan mutant suggests that this residue is making tighter contacts to GABARAPL-1 whereas there is more flexibility in the presence of the tyrosine. These data confirmed the NMR results suggesting a stronger interaction with NBR1-LIR\_Y732W and a faster exchange mode for NBR1-LIR wild type. Phenylalanine instead of tyrosine lead also to a more negative binding enthalpy but at the same time to a less positive binding entropy. These two opposite effects lead to a similar dissociation constant for NBR1-LIR wild type and for the phenylalanine mutant. Nevertheless, NBR1-LIR\_Y732W showed a stronger interaction to GABARAPL-1 than NBR1-LIR\_Y732F and NBR1-LIR wild type. By replacing the tyrosine with a phenylalanine, the binding surface in the hydrophobic pocket increased by the deletion of the hydroxyl group. At the same time, increasing the number of hydrophobic bonds did not help so much to get a stronger interaction because GABARAPL-1 could not adapt its hydrophobic pocket 1 for a better fitting for this aromatic residue like it is for the tyrosine residue. Presence of a tryptophan at position 1 of the LIR domains seems therefore to be essential for a strong interaction with MAP1LC3 proteins and its substitution by an other aromatic residue reduces equally the binding effect either with a tyrosine or a phenylalanine.

The structure of the GABARAPL-1/NBR1-LIR complex showed that the residues preceding the N-terminal part of the LIR motif were close to the two first  $\alpha$ -helices of GABARAPL-1. Several positive charged residues being present in this region and negative charged residues being characteristic of the LIR motif, ionic bonds could stabilize the interaction. To investigate whether a higher number of negatively charged amino acids in the LIR motif could compensate

for the substitution of tryptophan with tyrosine, the stretch of negatively charged amino acids directly N-terminal to Y732 in NBR1 was extended by mutating two serines to glutamates resulting to two other constructs, NBR1-LIR\_S729E and NBR1-LIR\_S728,729E. Introducing additional negatively charged amino acids also resulted in a more negative enthalpy but a less positive entropy relative to the wild type peptide. These opposite effects in the interaction are illustrated by a dissociation constant very similar to the one for NBR1-LIR wild type. Although an increase in the binding surface due to the addition of negative charged residue is noticed, these amino acids did not increase the binding affinity.

NBR1 and p62 share the same role in autophagy with the selective sequestration of ubiquitinated substrates into the autophagosome. Still, p62 shows a stronger interaction to MAP1LC3 proteins than NBR1. Moreover, p62 seems, for the moment, to bind to more ubiquitinated targets than NBR1 but only because these studies have not been performed for NBR1. Thus, even though NBR1 does bind to MAP1LC3 proteins as strong as p62, NBR1 is an essential auxiliary autophagy receptor. Both autophagy receptors co-localize in ubiquitinated bodies and interact together to act as ubiquitin cargo receptors.

#### 7.3.3.3. *Nix*

In the case of the Nix protein, the two identified LIR motifs have both a tryptophan residue at position 1 but the interaction with MAP1LC3 proteins is even weaker than for NBR1. Whereas Nix-LIR\_W36 has also a leucine as hydrophobic residue at position 4, no second hydrophobic residue is present at the correct position after W140 nor W144 but an aspartate and an arginine, respectively. There is no N-terminal strong negatively charged block present before W36 and only one aspartate before each tryptophan in Nix-LIR\_W140/144.<sup>99</sup>

For Nix, a weaker interaction with MAP1LC3 proteins is illustrated by a dissociation constant in the hundred micromolar range and a less negative enthalpy as well as a low positive entropy compared to p62. The interaction between LC3B and Nix-LIR\_W36 or Nix-LIR\_W140/144 is less saturated with only a few residues in the peptides that could find correct partners on LC3B surface to form proper polar and/or nonpolar contacts. The absence of negatively charged residues N-terminally to the aromatic residue reduces the ionic interactions with the N-terminal residues of MAP1LC3 proteins. While Nix-LIR\_W140/144 contains one acidic residue at position -1 (compared to three for p62 and two for NBR1), most importantly it does not have any hydrophobic residues at position 4. Consequently, the data showed the weakest binding for Nix-LIR\_W140/144. In the interaction with LC3B, ITC data showed that Nix-LIR\_W140/144 was the weakest binder followed by Nix-LIR\_W36, p62 being the strongest one. These

observations were confirmed by NMR with different binding patterns for each interaction partner: slow exchange mode correlated to a strong interaction for p62 and fast exchange mode illustrating a weak interaction for Nix-LIR\_W140/144. In between, an intermediate close to fast exchange mode was observed for Nix-LIR\_W36. Using LC3A instead of LC3B as MAP1LC3 protein, the CSP were more important in presence of Nix-LIR\_W36 than of Nix-LIR\_W140/144. While Nix-LIR\_W140/144 titration against LC3A and LC3B showed a typical fast exchange mode, a fast (close to intermediate) exchange mode was observed in presence of Nix-LIR\_W36. In parallel, Nix-LIR\_W36 titration against GABARAPL-1 by NMR confirmed the tendency of Nix-LIR\_W36 to interact in a fast-intermediate exchange mode with MAP1LC3 proteins. To reassign LC3B residues in presence of the LIR domains, no assignment using NMR experiments was done but only by following the chemical shift perturbations. So it is possible that some of the peaks corresponding to the bound form reappeared approaching to saturation but so far away from the free form that it could not be assign like this. Nevertheless, the number of new peaks appearing being lower than the disappearing one for Nix-LIR\_W36, it confirmed that NH resonances are going into a intermediate (fast) kinetic exchange.

While the comparison of p62 and NBR1 interaction with MAP1LC3 proteins showed that the presence of a tryptophan as aromatic residue is more preferential, a tryptophan alone is not sufficient to assess a strong interaction profile to a LIR motif. Thus, in addition to the presence of the aromatic residue buried in hp1, the hydrophobic interactions of the LIR motif with hp2 are also essential and have to be complemented with electrostatic interactions mediated by negatively charged amino acids. Nix binds less preferentially to MAP1LC3 proteins than NBR1 and p62. While p62 and NBR1 have similarity in their domain organization, Nix differs from them. Especially, Nix presents a transmembran domain to be anchored in the outer mitochondrial membran and, most important, does not bind to ubiquitin. For the moment, no clear evidence of p62 or NBR1 binding to ubiquitin chains conjugated to the membran of mitochondria has been shown. Until data proving their role in the selective degradation of mitochondria by autophagy, Nix supplies this function as mitochondrial autophagy receptor.

Taken together, these results demonstrate the importance of the aromatic residue in the LIR motif but also the involvement of the residues in the neighborhood for the interaction with MAP1LC3 proteins. The replacement of tryptophan at position 1 with other aromatic residues leads to a weaker interaction. The presence of a tyrosine as aromatic residue reduces the overall binding affinity but due to its flexibility in the receiving hydrophobic pocket, the interaction surface could adapt itself to still allow the binding. Other substitutions like increasing the

number of negatively charged residues at the N-terminus of the aromatic residue of the LIR domain, however, result in a remarkable enthalpy-entropy compensation allowing different sequences to interact with an overall similar binding affinity.

The characterization of a new autophagy receptor, optineurin, showed that the phosphorylation of the serines preceding the LIR domain increases its interaction with LC3A and LC3B.<sup>105</sup> The substitution of serines by glutamates is often seen as a mimetic of phosphorylation. In the case of NBR1, the phosphorylation of the LIR motif does not have such an obvious effect in the binding to MAP1LC3 proteins like for optineurin. Interestingly, instead of acidic residues, both Nix-LIR motifs present serines, which could be phosphorylated and then carry a negative charge and finally could interact stronger to MAP1LC3 proteins. *In vivo* studies as well as the identification of kinases will be necessary for a clear statement if phosphorylation of NBR1 and/or Nix could increase their binding properties to MAP1LC3 proteins.

p62, NBR1 and Nix have a similar function as autophagy receptors, which tether entire organelles or individual proteins to the autophagosomal membrane.<sup>82</sup> Nevertheless, differences in their LIR motifs and in their binding affinity to MAP1LC3 proteins suggest some functional divergences. Further studies of the role of posttranslational modifications that might regulate the interaction between autophagy receptors and effectors will be of high interest.

#### **7.3.4. Differences between MAP1LC3 proteins**

The different MAP1LC3 proteins present different exchange mode of interactions with the different LIR motifs suggesting that differences in protein constitution have also to be taken in account to understand the specificity of the interaction.

In yeast, Atg8 is the only protein involved in autophagosome formation whereas seven homologues have been identified so far in mammalian and could be divided into two sub-families: LC3- (LC3A, LC3B and LC3C) and GABARAP- (GABARAP, GABARAPL-1, GABARAPL-2 and GABARAPL-3) proteins.

##### **7.3.4.1. LC3B vs. LC3A**

The interaction of LC3B with autophagy receptors being established as reference, the next interest was to compare their binding to other MAP1LC3 proteins.

Like for LC3B, NMR interaction studies of LC3A with Nix-LIRs showed that Nix-LIR\_W36 binds stronger than Nix-LIR\_W140/144, confirming the GST pull-down assays performed by Novak et al.<sup>99</sup> For LC3A and LC3B, corresponding residues situated in the hydrophobic pockets 1 and 2 or in the closest neighborhood, meaning in direct contact with the LIR motif, were the

most affected residues upon titration but the CSP were higher in presence of Nix-LIRs for LC3A than for LC3B. The ITC titration confirmed that LC3A was a better interaction partner to both Nix-LIRs than LC3B and that Nix-LIR\_W36 binds stronger to LC3- proteins than Nix-LIR\_W140/144. Interestingly, the free energy of the system is the same for LC3A and LC3B with both Nix-LIR\_W36 and Nix-LIR\_W140/144 ( $\Delta G \approx -6 \text{ kcal mol}^{-1}$ ) but, while the enthalpy and entropy have opposite effects in LC3B/Nix-LIRs interaction, LC3A/Nix-LIRs is entropy driven. The higher entropy values for LC3A interaction with Nix-LIRs could reflect the possibility for more hydrophobic bonds than for LC3B even though LC3A and LC3B have a very similar hydrophobic pattern. However, NMR studies of both proteins showed already differences in their behavior, especially in the flexibility of the N-terminal part. Further studies focused on the structural insights of the interaction will be necessary to be able to explain the important difference between LC3A and LC3B in their interaction with Nix-LIRs.

#### **7.3.4.2. LC3- vs. GABARAP- proteins**

Although the high sequence and structure similarity between MAP1LC3 proteins, handling these proteins was challenging. For GABARAP- proteins, GABARAPL-1 and GABARAPL-2 presented a different behavior in the first purification steps and GABARAPL-1 seemed to be more preferential for further studies due to a better solubility and stability of the protein. Concerning LC3- proteins, the NMR spectra of LC3A and LC3B fused to ubiquitin and isolated from the cell lysate showed different patterns. Whereas the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY\_HSQC spectrum of LC3A was similar to the one of the purified protein, the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY\_HSQC spectrum of LC3B looked like one of a bigger protein or of a protein involved in a huge complex. Intriguingly, during the different NMR experiments involving MAP1LC3 proteins, several broad or even double peaks were observed as well in [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY\_HSQC spectra as in HNCA, HNCACB or NOESY spectra. This fact can be interpreted as a certain degree of flexibility for these proteins. Moreover, during assignment attempts for GABARAPL-1 but also for LC3B or for LC3A, difficulties were encountered concerning the N-terminal part of these proteins, especially around the two  $\alpha$ -helices. Depending of the conditions, the [ $^{15}\text{N}$ ,  $^1\text{H}$ ] TROSY\_HSQC spectra of LC3A presented some differences that could be interpreted as a switch between conformations. The presence of broad or double peaks as well as the disappearance of some peaks is probably provoked by proton amide exchange of these exposed residues with the buffer due to local dynamics created by the flexibility at the N-terminus of GABARAPL-1 like it was previously reported for GABARAP by Coyle et al.<sup>74</sup>

The fact that MAP1LC3 proteins share a similar structure although they differ in their tissue localizations raises the question about different functions for LC3- and GABARAP- proteins in autophagy. Weidberg et al.<sup>64</sup> showed already that all MAP1LC3 proteins are indispensable for autophagosome formation but acting at different time points. Moreover, each MAP1LC3 protein does not bind identically to the different autophagy receptors. The main structural characteristic of MAP1LC3 proteins is the presence of two N-terminal  $\alpha$ -helices. Interestingly, the amino acid composition of these helices differs between MAP1LC3 proteins. Whereas  $\alpha$ -helix 1 and  $\alpha$ -helix 2 in LC3- proteins are basic and neutral, respectively, they present the opposite ionic character in GABARAP- proteins. Added to the knowledge earned by amino acid mutations in NBR1-LIR, it is likely that the initial sequence of LIR motif in autophagy receptors and the potential posttranslational modifications correlated to the electrostatic surface of the N-terminal  $\alpha$ -helices of autophagy effectors could determine the specificity of the interaction and could explain their particular functions. Studies of the effects of mutagenesis of the N-terminal region of MAP1LC3 proteins on the interaction with LIR motifs is still needed to be able to understand the importance of each residue and to explain the specificity of each autophagy effector and receptor on the interacting point of view.

### **7.3.5. Structure of GABARAPL-1/NBR1-LIR**

One of the goals of this study was to further characterize the interaction of the NBR1-LIR domain with autophagy effector proteins by detailed NMR investigations on a structural point of view in order to better understand the involvement of the tyrosine residue in the LIR motif of NBR1 compared to the standard tryptophan in the majority of other LIR.

In this thesis, the first NMR complex structure of a mammalian Atg8 homolog, GABARAPL-1, in presence of a non-tryptophan receptor LIR domain, NBR1 has been solved. Three dimensional structures (mainly solved by X-ray crystallography) of different MAP1LC3 proteins have already been solved and show a high structural similarity with an average backbone RMSD of 1.5 Å. In addition, the binding mode of peptides derived from different autophagy receptors is also very similar compared to the structure of the GABARAPL-1/NBR1-LIR complex (RMSD 1.85 Å with the NMR resolved structure of LC3B/p62 (PDB: 2K6Q)<sup>87</sup> over analogous secondary structure elements; RMSD 1.59 Å with the X-Ray crystallography of GABARAP/calreticulin (PDB: 3DOW)<sup>135</sup> over the structured part E12-V114). The structure of GABARAPL-1 in presence of NBR1-LIR presents the typical Ub-fold preceded by two N-terminal  $\alpha$ -helices, forming the two hydrophobic pockets

characteristic for MAP1LC3 proteins. Nevertheless, some differences are still observable. Even though the LIRs of NBR1 and p62 share the same general motif, their amino acid composition differs, reflected especially by a different occupation of the hydrophobic pocket for the aromatic residue at position 1. The structure of the GABARAPL-1/NBR1-LIR complex shows that the side chain of NBR1-LIR Y732 is flexible and can adopt different positions in the hydrophobic pocket 1 of GABARAPL-1 whereas p62 shows less mobility with LC3B. In the structure determination of GABARAP by Coyle et al.,<sup>74</sup> two conformations of the protein, open and closed, were identified. For all other structures of MAP1LC3 proteins, alone or in a complex, only the close conformation was observed. This is also the case for the GABARAPL-1/NBR1-LIR complex but the difficulty of assignment in the first two  $\alpha$ -helices provoked by some flexibility in the N-terminal region could be engaged due to this conformational exchange. Nevertheless, the presence of different resonances for Y732 of NBR1-LIR associated to the possibility for this same residue to take different positions in the hp1 of GABARAPL-1, all in accord with the NMR calculation of the structure of the GABARAPL-1/NBR1-LIR complex, plays also in favor of different possible conformations for the protein. Using different conditions in the analysis of GABARAPL-1 alone and in complex could maybe allow to better distinguish these two conformations for GABARAPL-1. The same studies on other MAP1LC3 proteins will be necessary to be able to structurally differentiate proteins from the same protein family.

It was already shown that the type of amino acid at position 4 does not dramatically affect the interaction as long as the residue at this position provides enough hydrophobic surface.<sup>136</sup> In the NBR1-LIR domain this position is occupied by isoleucine and its side chain in the GABARAPL-1/NBR1-LIR complex structure is oriented in hp2 in the same manner as leucine in the LC3B/p62 complex.<sup>87; 103</sup> The NBR1-LIR domain contains also more hydrophobic residues than p62 at positions 2, 3 and 5 (isoleucine, isoleucine and leucine for threonine, histidine and serine). Such a hydrophobic track might lead to multiple binding modes or even competition of the individual amino acids for the hydrophobic pockets. The broad NMR resonances of amino acids in the core of the NBR1-LIR domain in complex with GABARAPL-1 show that conformational averaging indeed occurs, but at the same time did not allow to characterize the different potential states in more details. While the reappearance of NBR1 resonances at high molar ratios argues that a significant contribution to the observed line broadening is due to intermediate exchange of the peptide between the bound and free state, the observation of multiple resonances for Y732 suggests that, at least for some residues, multiple conformations in the bound state might occur. However, no multiple resonances or conformations during structure calculations have been observed for the isoleucine residues. The



increased hydrophobic interaction throughout the entire NBR1-LIR domain might, however, compensate the tight interaction lost by the replacement of the typical tryptophan with a tyrosine.

### ***7.3.6. MAP1LC3 proteins are not only involved in autophagy***

The conserved amino acid sequences for each MAP1LC3 protein, from yeast to mammals, can be a sign of crucial function for this protein family. However, no special phenotype has been observed so far for GABARAP knockout mice. The high sequence and the structure homology between MAP1LC3 proteins could probably allow one protein to substitute the function of a homolog if necessary. Nevertheless, differences in the binding to LIR motifs coupled with differences in flexibility and in the potential oligomerization state lead to the hypothesis that each MAP1LC3 protein has more specific function in addition to selective autophagy.

Interestingly, none of the MAP1LC3 proteins was discovered from its role in autophagy but essentially in intracellular protein trafficking by enhancing vesicle fusion.

LC3- proteins were first identified as proteins associated to microtubules.<sup>57</sup> Present in neurons, these proteins bind to tubulin and regulate the microtubule binding activity of MAP1A and MAP1B. Whereas LC3B and LC3C were ubiquitously present in all tissues, LC3A is absent from peripheral blood leukocytes. Their expression level depend of their localization with a constant weaker expression for LC3C.<sup>59</sup>

GABARAP- proteins were first denominated after the discovery of GABARAP, a protein associated to GABA<sub>A</sub> receptors. First studies showed that GABARAP could bind to the  $\gamma 2$ -subunit of GABA<sub>A</sub> receptors as well as to tubulin at the same time, participating in targeting and clustering of GABA<sub>A</sub> receptors.<sup>158</sup> Vergnier-Magnin et al.<sup>117</sup> characterized GEC1 as a protein related to GABARAP due to their similarity with 87% of sequence identity and was consequently renamed GABARAPL-1. Expressed in all tissues and binding to tubulin, GABARAPL-1 acts also as a linker between membrane receptors and microtubules. GABARAPL-2 also known as GATE-16 was described as a soluble transporter that could bind to NSF and Golgi membrane proteins leading to the modulation of intra-Golgi transport and being involved in membrane fusion.<sup>67</sup>

Studies about the interaction partners of GABARAP are well characterized.<sup>137</sup> Although this work focused on the interaction of MAP1LC3 proteins with autophagy receptors, it should be noticed that mammalian Atg8 proteins also interact with modifying Atg8 protein enzymes but also with non-autophagy proteins, showing a wide role for MAP1LC3 proteins.

The structure of the GABARAPL-1/NBR1-LIR complex determined in this work is analog to the binding mode observed in the structure of GABARAP with calreticulin,<sup>135</sup> LC3 with p62,<sup>87</sup>

<sup>103</sup> LC3 with Atg4B<sup>104</sup> or Atg8 with Atg19<sup>87</sup> and with Atg3<sup>159</sup> and suggested for the interaction of LC3 with Atg32,<sup>89</sup> of GABARAP with  $\gamma$ 2 subunit of the GABA<sub>A</sub> receptor,<sup>74</sup> NSF,<sup>138</sup> clathrin Heavy Chain<sup>80</sup> or Nix<sup>137</sup> and of GABARAPL-1 with  $\kappa$ -opioid receptor.<sup>160</sup> The involvement of the two hydrophobic pockets of the mammalian Atg8 effector proteins binding to an aromatic residue and to a hydrophobic residue in hp1 and hp2, respectively, is characteristic of the MAP1LC3 proteins. Additionally, the presence of an extended  $\beta$  conformation in the interaction partner resulting in an intermolecular parallel  $\beta$ -sheet with  $\beta$ -strand 2 of the Ub-fold is a known feature in protein-protein interaction involving ubiquitin-like proteins.<sup>161; 162</sup>

Except tubulin using the N-terminal region,<sup>163</sup> most of the interaction partners of MAP1LC3 proteins binds to the same surface including the Ub-fold and especially the two characteristic hydrophobic pockets. The same binding surface is involved even though the binding partners have a variety of functions from autophagy to the transport and clustering of diverse proteins and vesicles. Although these interaction partners have different functions, they all have the presence of an aromatic residue associated to a hydrophobic residue and negatively charged residues in the binding site in common. Docking experiments demonstrated that the indole ring of tryptophan is positioned between the side chains of K48 and L50 of GABARAP while the carboxyl group points out of the hydrophobic pocket.<sup>80</sup> Because of the high structure homology of GABARAP to other MAP1LC3 proteins and the conservation of the residues involved in hp1, it is extremely probable that the specificity of a ligand possessing an aromatic residue showed for GABARAP is also true for other LC3- and GABARAP- proteins. However, the discrepancy between different LIR motifs suggests that not a specific amino acid sequence is obligatory but hydrophobic interactions lead by aromatic residues are the main driving force in the binding mode of MAP1LC3 proteins with the interaction partners.

High conserved primary sequences, a wide tissue expression and a broad range of interaction partners for mammalian Atg8 proteins suggest a role as adaptor proteins that bind simultaneously to multivalent receptor proteins on one hand and to membranes after association to lipids on the other. In a general and synthetic description, MAP1LC3 proteins are thus involved in membrane trafficking events but some questions are still remaining. Is each MAP1LC3 protein specifically essential for selective autophagy? Does the role of each MAP1LC3 protein depend on different factors like the tissue localization or the cell conditions? Elazar and co-workers showed already functional differences between LC3- and GABARAP- proteins in the autophagosome formation.<sup>64</sup> Is it also the same for autophagy receptor recognition? Are NBR1 and p62 always involved in parallel to recognize ubiquitinated

substrates? The last decade has seen a lot of research performed, which led to a better understanding of the autophagy pathway but some questions are still remaining for the next generations of PhD students.

#### **7.4. Conclusion and Outlook**

Ubiquitin is a highly conserved protein. Considering the direct importance of ubiquitin in different cellular processes, the discovery of ubiquitin-like proteins involved in a great variety of biological pathways enhances the importance of their characteristic Ub-fold. Depending on the structure elements associated to the Ub-fold, interaction partners bind differently. Nevertheless, the same epitope on ubiquitin-like domains is used along one protein family to functionally differentiate proteins in between the ubiquitin-like superfamily.

In this work, two kind of ubiquitin-like proteins have been studied and both of them showed differences in their binding surface. Whereas the ubiquitin-like fold is preceded by two  $\alpha$ -helices to form essential hydrophobic pockets on MAP1LC3 proteins for the binding of autophagy receptors, these hydrophobic pockets are not present in TBK1\_ULD. The surface of the Ub-fold on TBK1\_ULD is differently used depending of binding to IRF3 or its own kinase domain. During this work, the structure of two proteins has been solved, the ULD domain of TBK1 and the complex issued from the interaction of GABARAPL-1 with NBR1-LIR. Preliminary interaction studies of TBK1\_ULD with IRF3\_IAD-SRR have been started and the interaction of the LIR motif of autophagy receptors with MAP1LC3 proteins has been characterized. Concerning the “TBK1\_ULD project”, in addition to a real characterization of the interaction with IAD-SRR, studies on the kinase domain of TBK1 will provide a better understanding of the immune response mechanism involving IKK proteins. For the “autophagy project”, characterization of posttranslational modifications on autophagy receptors and effectors could explain the specificity of the different proteins.

Due to the solubility and stability provided by the structure of ubiquitin, the properties of the Ub-fold provide a great interest for protein engineering. Fused to proteins or peptides, ubiquitin was used as expression and solubility enhancer, which did not obligatory need to be removed for protein studies using biophysical methods. In this thesis, the use of Ub-fused constructs has been shown to be successful to produce pure labeled peptide for structure determination by NMR and to perform interaction studies by NMR and ITC. In order to answer to the remaining questions related to the two previous projects, a further use of Ub-fused constructs will be a method of proof for the benefits of ubiquitin in the expression of proteins and peptides but also for the use of biophysical methods without the constraints of difficult purification steps.



## 8. References

1. A. Lavoisier. (1789). *Traité Élémentaire de Chimie*, Cuchet.
2. A. Hershko and A. Ciechanover, The ubiquitin system, *Annu. Rev. Biochem.* **67** (1998), pp. 425-479.
3. A. Ciechanover, Intracellular protein degradation: from a vague idea, through the lysosome and the ubiquitin-proteasome system, and onto human diseases and drug targeting (Nobel lecture), *Angew Chem Int Ed Engl* **44** (2005), pp. 5944-5967.
4. A. Hershko, The ubiquitin system for protein degradation and some of its roles in the control of the cell-division cycle (Nobel lecture), *Angew Chem Int Ed Engl* **44** (2005), pp. 5932-5943.
5. I. Rose, Ubiquitin at Fox Chase (Nobel lecture), *Angew Chem Int Ed Engl* **44** (2005), pp. 5926-5931.
6. A. Hershko, H. Heller, S. Elias and A. Ciechanover, Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown, *J. Biol. Chem.* **258** (1983), pp. 8206-8214.
7. A. Hershko, Lessons from the discovery of the ubiquitin system, *Trends Biochem. Sci.* **21** (1996), pp. 445-449.
8. F. Ikeda and I. Dikic, Atypical ubiquitin chains: new molecular signals. 'Protein Modifications: Beyond the Usual Suspects' review series, *EMBO Rep.* **9** (2008), pp. 536-542.
9. D. Finley, A. Ciechanover and A. Varshavsky, Thermolability of ubiquitin-activating enzyme from the mammalian cell cycle mutant ts85, *Cell* **37** (1984), pp. 43-55.
10. T. Sommer and S. Jentsch, A protein translocation defect linked to ubiquitin conjugation at the endoplasmic reticulum, *Nature* **365** (1993), pp. 176-179.
11. J. Spence, R. R. Gali, G. Dittmar, F. Sherman, M. Karin and D. Finley, Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain, *Cell* **102** (2000), pp. 67-76.
12. A. M. Weissman, Themes and variations on ubiquitylation, *Nat. Rev. Mol. Cell Biol.* **2** (2001), pp. 169-178.
13. M. Hochstrasser, Origin and function of ubiquitin-like proteins, *Nature* **458** (2009), pp. 422-429.
14. A. M. Burroughs, S. Balaji, L. M. Iyer and L. Aravind, Small but versatile: the extraordinary functional and structural diversity of the beta-grasp fold, *Biol Direct* **2** (2007), pp. 1-28.
15. V. Kirkin and I. Dikic, Role of ubiquitin- and Ubl-binding proteins in cell signaling, *Curr. Opin. Cell Biol.* **19** (2007), pp. 199-205.
16. D. Hoeller, C.-M. Hecker and I. Dikic, Ubiquitin and ubiquitin-like proteins in cancer pathogenesis, *Nat. Rev. Cancer* **6** (2006), pp. 776-788.
17. A. Buchberger, From UBA to UBX: new words in the ubiquitin vocabulary, *Trends Cell Biol.* **12** (2002), pp. 216-221.
18. C. Grabbe and I. Dikic, Functional roles of ubiquitin-like domain (ULD) and ubiquitin-binding domain (UBD) containing proteins, *Chem. Rev.* **109** (2009), pp. 1481-1494.
19. L. Hicke, H. L. Schubert and C. P. Hill, Ubiquitin-binding domains, *Nat. Rev. Mol. Cell Biol.* **6** (2005), pp. 610-621.
20. R. Hartmann-Petersen and C. Gordon, Integral UBL domain proteins: a family of proteasome interacting proteins, *Semin. Cell Dev. Biol.* **15** (2004), pp. 247-259.

## References

---

21. V. I. Korolchuk, F. M. Menzies and D. C. Rubinsztein, Mechanisms of cross-talk between the ubiquitin-proteasome and autophagy-lysosome systems, *FEBS Lett.* **584** (2010), pp. 1393-1398.
22. D. J. Klionsky and Y. Ohsumi, Vacuolar import of proteins and organelles from the cytoplasm, *Annu. Rev. Cell Dev. Biol.* **15** (1999), pp. 1-32.
23. D. J. Klionsky, Autophagy: from phenomenology to molecular understanding in less than a decade, *Nat. Rev. Mol. Cell Biol.* **8** (2007), pp. 931-937.
24. A. Uttenweiler and A. Mayer, Microautophagy in the yeast *Saccharomyces cerevisiae*, *Methods Mol. Biol.* **445** (2008), pp. 245-259.
25. A. Massey, R. Kiffin and A. M. Cuervo, Pathophysiology of chaperone-mediated autophagy, *Int. J. Biochem. Cell Biol.* **36** (2004), pp. 2420-2434.
26. F. Reggiori and D. J. Klionsky, Autophagy in the eukaryotic cell, *Eukaryot. Cell* **1** (2002), pp. 11-21.
27. J. C. Farre, R. Krick, S. Subramani and M. Thumm, Turnover of organelles by autophagy in yeast, *Curr. Opin. Cell Biol.* **21** (2009), pp. 522-530.
28. B. Levine and D. J. Klionsky, Development by self-digestion: molecular mechanisms and biological functions of autophagy, *Dev. Cell* **6** (2004), pp. 463-477.
29. I. Nakagawa, A. Amano, N. Mizushima, A. Yamamoto, H. Yamaguchi, T. Kamimoto, A. Nara, J. Funao, M. Nakata, K. Tsuda, S. Hamada and T. Yoshimori, Autophagy defends cells against invading group A *Streptococcus*, *Science* **306** (2004), pp. 1037-1040.
30. M. Baba, K. Takeshige, N. Baba and Y. Ohsumi, Ultrastructural analysis of the autophagic process in yeast: detection of autophagosomes and their characterization, *J. Cell Biol.* **124** (1994), pp. 903-913.
31. Z. Xie and D. J. Klionsky, Autophagosome formation: core machinery and adaptations, *Nat. Cell Biol.* **9** (2007), pp. 1102-1109.
32. W. A. Dunn, Jr., Autophagy and related mechanisms of lysosome-mediated protein degradation, *Trends Cell Biol.* **4** (1994), pp. 139-143.
33. K. Suzuki, T. Kirisako, Y. Kamada, N. Mizushima, T. Noda and Y. Ohsumi, The pre-autophagosomal structure organized by concerted functions of APG genes is essential for autophagosome formation, *EMBO J.* **20** (2001), pp. 5971-5981.
34. J. Kim, W. P. Huang, P. E. Stromhaug and D. J. Klionsky, Convergence of multiple autophagy and cytoplasm to vacuole targeting components to a perivacuolar membrane compartment prior to de novo vesicle formation, *J. Biol. Chem.* **277** (2002), pp. 763-773.
35. F. Reggiori and D. J. Klionsky, Atg9 sorting from mitochondria is impaired in early secretion and VFT-complex mutants in *Saccharomyces cerevisiae*, *J. Cell Sci.* **119** (2006), pp. 2903-2911.
36. F. Reggiori, K. A. Tucker, P. E. Stromhaug and D. J. Klionsky, The Atg1-Atg13 complex regulates Atg9 and Atg23 retrieval transport from the pre-autophagosomal structure, *Dev. Cell* **6** (2004), pp. 79-90.
37. W. L. Yen, J. E. Legakis, U. Nair and D. J. Klionsky, Atg27 is required for autophagy-dependent cycling of Atg9, *Mol. Biol. Cell* **18** (2007), pp. 581-593.
38. N. Fujita, T. Itoh, H. Omori, M. Fukuda, T. Noda and T. Yoshimori, The Atg16L complex specifies the site of LC3 lipidation for membrane biogenesis in autophagy, *Mol. Biol. Cell* **19** (2008), pp. 2092-2100.
39. N. Mizushima, A. Yamamoto, M. Hatano, Y. Kobayashi, Y. Kabeya, K. Suzuki, T. Tokuhiya, Y. Ohsumi and T. Yoshimori, Dissection of autophagosome formation using Apg5-deficient mouse embryonic stem cells, *J. Cell Biol.* **152** (2001), pp. 657-668.
40. T. Kirisako, M. Baba, N. Ishihara, K. Miyazawa, M. Ohsumi, T. Yoshimori, T. Noda and Y. Ohsumi, Formation process of autophagosome is traced with Apg8/Aut7p in yeast, *J. Cell Biol.* **147** (1999), pp. 435-446.

41. A. R. Young, E. Y. Chan, X. W. Hu, R. Kochl, S. G. Crawshaw, S. High, D. W. Hailey, J. Lippincott-Schwartz and S. A. Tooze, Starvation and ULK1-dependent cycling of mammalian Atg9 between the TGN and endosomes, *J. Cell Sci.* **119** (2006), pp. 3888-3900.
42. I. Tanida, T. Ueno and E. Kominami, LC3 conjugation system in mammalian autophagy, *Int. J. Biochem. Cell Biol.* **36** (2004), pp. 2503-2518.
43. A. Longatti and S. A. Tooze, Vesicular trafficking and autophagosome formation, *Cell Death Differ.* **16** (2009), pp. 956-965.
44. W. A. Dunn, Jr., Studies on the mechanisms of autophagy: maturation of the autophagic vacuole, *J. Cell Biol.* **110** (1990), pp. 1935-1945.
45. E. L. Axe, S. A. Walker, M. Manifava, P. Chandra, H. L. Roderick, A. Habermann, G. Griffiths and N. T. Ktistakis, Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum, *J. Cell Biol.* **182** (2008), pp. 685-701.
46. D. W. Hailey, A. S. Rambold, P. Satpute-Krishnan, K. Mitra, R. Sougrat, P. K. Kim and J. Lippincott-Schwartz, Mitochondria Supply Membranes for Autophagosome Biogenesis during Starvation, *Cell* **141** (2010), pp. 656-667.
47. J. E. Vance and D. E. Vance, Phospholipid biosynthesis in mammalian cells, *Biochem. Cell Biol.* **82** (2004), pp. 113-128.
48. K. Suzuki and Y. Ohsumi, Molecular machinery of autophagosome formation in yeast, *Saccharomyces cerevisiae*, *FEBS Lett.* **581** (2007), pp. 2156-2161.
49. T. Yamamoto, K. Fujita, S. Asari, A. Chiba, Y. Kataba, K. Ohsumi, N. Ohmura, Y. Iida, C. Ijichi, S. Iwayama, N. Fukuchi and M. Shoji, Synthesis and evaluation of isoxazole derivatives as lysophosphatidic acid (LPA) antagonists, *Bioorg. Med. Chem. Lett.* **17** (2007), pp. 3736-3740.
50. N. Mizushima, T. Noda, T. Yoshimori, Y. Tanaka, T. Ishii, M. D. George, D. J. Klionsky, M. Ohsumi and Y. Ohsumi, A protein conjugation system essential for autophagy, *Nature* **395** (1998), pp. 395-398.
51. Y. Ichimura, T. Kirisako, T. Takao, Y. Satomi, Y. Shimonishi, N. Ishihara, N. Mizushima, I. Tanida, E. Kominami, M. Ohsumi, T. Noda and Y. Ohsumi, A ubiquitin-like system mediates protein lipidation, *Nature* **408** (2000), pp. 488-492.
52. T. Kirisako, Y. Ichimura, H. Okada, Y. Kabeya, N. Mizushima, T. Yoshimori, M. Ohsumi, T. Takao, T. Noda and Y. Ohsumi, 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** (2000), pp. 263-276.
53. I. Tanida, N. Mizushima, M. Kiyooka, M. Ohsumi, T. Ueno, Y. Ohsumi and E. Kominami, Apg7p/Cvt2p: A novel protein-activating enzyme essential for autophagy, *Mol. Biol. Cell* **10** (1999), pp. 1367-1379.
54. T. Hanada, N. N. Noda, Y. Satomi, Y. Ichimura, Y. Fujioka, T. Takao, F. Inagaki and Y. Ohsumi, The Atg12-Atg5 conjugate has a novel E3-like activity for protein lipidation in autophagy, *J. Biol. Chem.* **282** (2007), pp. 37298-37302.
55. K. Yoshimoto, Y. Takano and Y. Sakai, Autophagy in plants and phytopathogens, *FEBS Lett.* **584** (2010), pp. 1350-1358.
56. N. Mizushima, H. Sugita, T. Yoshimori and Y. Ohsumi, A new protein conjugation system in human. The counterpart of the yeast Apg12p conjugation system essential for autophagy, *J. Biol. Chem.* **273** (1998), pp. 33889-33892.
57. S. S. Mann and J. A. Hammarback, Molecular characterization of light chain 3. A microtubule binding subunit of MAP1A and MAP1B, *J. Biol. Chem.* **269** (1994), pp. 11492-11497.

58. Y. Kabeya, N. Mizushima, T. Ueno, A. Yamamoto, T. Kirisako, T. Noda, E. Kominami, Y. Ohsumi and T. Yoshimori, LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing, *EMBO J.* **19** (2000), pp. 5720-5728.
59. H. He, Y. Dang, F. Dai, Z. Guo, J. Wu, X. She, Y. Pei, Y. Chen, W. Ling, C. Wu, S. Zhao, J. O. Liu and L. Yu, Post-translational modifications of three members of the human MAP1LC3 family and detection of a novel type of modification for MAP1LC3B, *J. Biol. Chem.* **278** (2003), pp. 29278-29287.
60. Y. Xin, L. Yu, Z. Chen, L. Zheng, Q. Fu, J. Jiang, P. Zhang, R. Gong and S. Zhao, Cloning, expression patterns, and chromosome localization of three human and two mouse homologues of GABA(A) receptor-associated protein, *Genomics* **74** (2001), pp. 408-413.
61. Y. Kabeya, N. Mizushima, A. Yamamoto, S. Oshitani-Okamoto, Y. Ohsumi and T. Yoshimori, LC3, GABARAP and GATE16 localize to autophagosomal membrane depending on form-II formation, *J. Cell Sci.* **117** (2004), pp. 2805-2812.
62. F. Z. Chakrama, S. Seguin-Py, J. N. Le Grand, A. Fraichard, R. Delage-Mourroux, G. Despouy, V. Perez, M. Jouvenot and M. Boyer-Guittaut, GABARAPL1 (GEC1) associates with autophagic vesicles, *Autophagy* **6** (2010), pp. 495-505.
63. M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson and D. G. Higgins, Clustal W and Clustal X version 2.0, *Bioinformatics* **23** (2007), pp. 2947-2948.
64. H. Weidberg, E. Shvets, T. Shpilka, F. Shimron, V. Shinder and Z. Elazar, LC3 and GATE-16/GABARAP subfamilies are both essential yet act differently in autophagosome biogenesis, *EMBO J.* **29** (2010), pp. 1792-1802.
65. J. Hemelaar, V. S. Lelyveld, B. M. Kessler and H. L. Ploegh, A single protease, Apg4B, is specific for the autophagy-related ubiquitin-like proteins GATE-16, MAP1-LC3, GABARAP, and Apg8L, *J. Biol. Chem.* **278** (2003), pp. 51841-51850.
66. I. Tanida, M. Komatsu, T. Ueno and E. Kominami, GATE-16 and GABARAP are authentic modifiers mediated by Apg7 and Apg3, *Biochem. Biophys. Res. Commun.* **300** (2003), pp. 637-644.
67. Y. Sagiv, A. Legesse-Miller, A. Porat and Z. Elazar, GATE-16, a membrane transport modulator, interacts with NSF and the Golgi v-SNARE GOS-28, *EMBO J.* **19** (2000), pp. 1494-1504.
68. I. Tanida, Y. S. Sou, J. Ezaki, N. Minematsu-Ikeguchi, T. Ueno and E. Kominami, HsAtg4B/HsApg4B/autophagin-1 cleaves the carboxyl termini of three human Atg8 homologues and delipidates microtubule-associated protein light chain 3- and GABAA receptor-associated protein-phospholipid conjugates, *J. Biol. Chem.* **279** (2004), pp. 36268-36276.
69. H. Kumeta, M. Watanabe, H. Nakatogawa, M. Yamaguchi, K. Ogura, W. Adachi, Y. Fujioka, N. N. Noda, Y. Ohsumi and F. Inagaki, The NMR structure of the autophagy-related protein Atg8, *J. Biomol. NMR* **47** (2010), pp. 237-241.
70. M. Schwarten, M. Stoldt, J. Mohrluder and D. Willbold, Solution structure of Atg8 reveals conformational polymorphism of the N-terminal domain, *Biochem. Biophys. Res. Commun.* **395** (2010), pp. 426-431.
71. K. Sugawara, N. N. Suzuki, Y. Fujioka, N. Mizushima, Y. Ohsumi and F. Inagaki, The crystal structure of microtubule-associated protein light chain 3, a mammalian homologue of *Saccharomyces cerevisiae* Atg8, *Genes Cells* **9** (2004), pp. 611-618.
72. T. Kouno, M. Mizuguchi, I. Tanida, T. Ueno, T. Kanematsu, Y. Mori, H. Shinoda, M. Hirata, E. Kominami and K. Kawano, Solution structure of microtubule-associated



- protein light chain 3 and identification of its functional subdomains, *J. Biol. Chem.* **280** (2005), pp. 24610-24617.
73. V. N. Bavro, M. Sola, A. Bracher, M. Kneussel, H. Betz and W. Weissenhorn, Crystal structure of the GABA(A)-receptor-associated protein, GABARAP, *EMBO Rep.* **3** (2002), pp. 183-189.
74. J. E. Coyle, S. Qamar, K. R. Rajashankar and D. B. Nikolov, Structure of GABARAP in two conformations: implications for GABA(A) receptor localization and tubulin binding, *Neuron* **33** (2002), pp. 63-74.
75. D. Knight, R. Harris, M. S. McAlister, J. P. Phelan, S. Geddes, S. J. Moss, P. C. Driscoll and N. H. Keep, The X-ray crystal structure and putative ligand-derived peptide binding properties of gamma-aminobutyric acid receptor type A receptor-associated protein, *J. Biol. Chem.* **277** (2002), pp. 5556-5561.
76. Y. Paz, Z. Elazar and D. Fass, Structure of GATE-16, membrane transport modulator and mammalian ortholog of autophagocytosis factor Aut7p, *J. Biol. Chem.* **275** (2000), pp. 25445-25450.
77. C. A. Orengo, D. T. Jones and J. M. Thornton, Protein superfamilies and domain superfolds, *Nature* **372** (1994), pp. 631-634.
78. N. N. Noda, Y. Ohsumi and F. Inagaki, ATG systems from the protein structural point of view, *Chem. Rev.* **109** (2009), pp. 1587-1598.
79. Y. Ichimura, Y. Imamura, K. Emoto, M. Umeda, T. Noda and Y. Ohsumi, In vivo and in vitro reconstitution of Atg8 conjugation essential for autophagy, *J. Biol. Chem.* **279** (2004), pp. 40584-40592.
80. Y. Thielmann, J. Mohrlüder, B. W. Koenig, T. Stangler, R. Hartmann, K. Becker, H.-D. Höltje and D. Willbold, An indole-binding site is a major determinant of the ligand specificity of the GABA type A receptor-associated protein GABARAP, *Chembiochem* **9** (2008), pp. 1767-1775.
81. K. Haglund and I. Dikic, Ubiquitylation and cell signaling, *The EMBO Journal* **24** (2005), pp. 3353-3359.
82. T. Lamark, V. Kirkin, I. Dikic and T. Johansen, NBR1 and p62 as cargo receptors for selective autophagy of ubiquitinated targets, *Cell Cycle* **8** (2009), pp. 1986-1990.
83. T. Lamark, M. Perander, H. Outzen, K. Kristiansen, A. Overvatn, E. Michaelsen, G. Bjorkoy and T. Johansen, Interaction codes within the family of mammalian Phox and Bem1p domain-containing proteins, *J. Biol. Chem.* **278** (2003), pp. 34568-34581.
84. M. I. Wilson, D. J. Gill, O. Perisic, M. T. Quinn and R. L. Williams, PB1 domain-mediated heterodimerization in NADPH oxidase and signaling complexes of atypical protein kinase C with Par6 and p62, *Mol. Cell* **12** (2003), pp. 39-50.
85. R. K. Vadlamudi, I. Joung, J. L. Strominger and J. Shin, p62, a phosphotyrosine-independent ligand of the SH2 domain of p56lck, belongs to a new class of ubiquitin-binding proteins, *J. Biol. Chem.* **271** (1996), pp. 20235-20237.
86. G. Bjorkoy, T. Lamark, A. Brech, H. Outzen, M. Perander, A. Overvatn, H. Stenmark and T. Johansen, p62/SQSTM1 forms protein aggregates degraded by autophagy and has a protective effect on huntingtin-induced cell death, *J. Cell Biol.* **171** (2005), pp. 603-614.
87. N. N. Noda, H. Kumeta, H. Nakatogawa, K. Satoo, W. Adachi, J. Ishii, Y. Fujioka, Y. Ohsumi and F. Inagaki, Structural basis of target recognition by Atg8/LC3 during selective autophagy, *Genes Cells* **13** (2008), pp. 1211-1218.
88. S. Pankiv, T. H. Clausen, T. Lamark, A. Brech, J.-A. Bruun, H. Outzen, A. Overvatn, G. Bjorkoy and T. Johansen, p62/SQSTM1 binds directly to Atg8/LC3 to facilitate degradation of ubiquitinated protein aggregates by autophagy, *J. Biol. Chem.* **282** (2007), pp. 24131-24145.

89. N. N. Noda, Y. Ohsumi and F. Inagaki, Atg8-family interacting motif crucial for selective autophagy, *FEBS Lett.* **584** (2010), pp. 1379-1385.
90. M. Komatsu, S. Waguri, M. Koike, Y. S. Sou, T. Ueno, T. Hara, N. Mizushima, J. Iwata, J. Ezaki, S. Murata, J. Hamazaki, Y. Nishito, S. Iemura, T. Natsume, T. Yanagawa, J. Uwayama, E. Warabi, H. Yoshida, T. Ishii, A. Kobayashi, M. Yamamoto, Z. Yue, Y. Uchiyama, E. Kominami and K. Tanaka, Homeostatic levels of p62 control cytoplasmic inclusion body formation in autophagy-deficient mice, *Cell* **131** (2007), pp. 1149-1163.
91. M. W. Wooten, T. Geetha, J. R. Babu, M. L. Seibenhener, J. Peng, N. Cox, M. T. Diaz-Meco and J. Moscat, Essential role of sequestosome 1/p62 in regulating accumulation of Lys63-ubiquitinated proteins, *J. Biol. Chem.* **283** (2008), pp. 6783-6789.
92. I. G. Campbell, H. M. Nicolai, W. D. Foulkes, G. Senger, G. W. Stamp, G. Allan, C. Boyer, K. Jones, R. C. Bast, Jr. and E. Solomon, A novel gene encoding a B-box protein within the BRCA1 region at 17q21.1, *Hum. Mol. Genet.* **3** (1994), pp. 589-594.
93. V. Kirkin, T. Lamark, Y.-S. Sou, G. Bjorkoy, J. L. Nunn, J.-A. Bruun, E. Shvets, D. G. McEwan, T. H. Clausen, P. Wild, I. Bilusic, J.-P. Theurillat, A. Overvatn, T. Ishii, Z. Elazar, M. Komatsu, I. Dikic and T. Johansen, A Role for NBR1 in Autophagosomal Degradation of Ubiquitinated Substrates, *Mol. Cell* **33** (2009), pp. 505-516.
94. S. Waters, K. Marchbank, E. Solomon, C. Whitehouse and M. Gautel, Interactions with LC3 and polyubiquitin chains link nbr1 to autophagic protein turnover, *FEBS Lett.* **583** (2009), pp. 1846-1852.
95. S. Müller, I. Kursula, P. Zou and M. Wilmanns, Crystal structure of the PB1 domain of NBR1, *FEBS Lett.* **580** (2006), pp. 341-344.
96. I. Kim, S. Rodriguez-Enriquez and J. J. Lemasters, Selective degradation of mitochondria by mitophagy, *Arch. Biochem. Biophys.* **462** (2007), pp. 245-253.
97. T. Kanki, K. Wang, Y. Cao, M. Baba and D. J. Klionsky, Atg32 is a mitochondrial protein that confers selectivity during mitophagy, *Dev. Cell* **17** (2009), pp. 98-109.
98. K. Okamoto, N. Kondo-Okamoto and Y. Ohsumi, Mitochondria-anchored receptor Atg32 mediates degradation of mitochondria via selective autophagy, *Dev. Cell* **17** (2009), pp. 87-97.
99. I. Novak, 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, Nix is a selective autophagy receptor for mitochondrial clearance, *EMBO Rep.* **11** (2009), pp. 45-51.
100. T. Kanki and D. J. Klionsky, Mitophagy in yeast occurs through a selective mechanism, *J. Biol. Chem.* **283** (2008), pp. 32386-32393.
101. R. L. Schweers, J. Zhang, M. S. Randall, M. R. Loyd, W. Li, F. C. Dorsey, M. Kundu, J. T. Opferman, J. L. Cleveland, J. L. Miller and P. A. Ney, NIX is required for programmed mitochondrial clearance during reticulocyte maturation, *Proc. Natl Acad. Sci. USA* **104** (2007), pp. 19500-19505.
102. W.-X. Ding, H.-M. Ni, M. Li, Y. Liao, X. Chen, D. B. Stolz, G. W. Dorn II and X.-M. Yin, Nix is critical to two distinct phases of mitophagy: reactive oxygen species (ROS)-mediated autophagy induction and Parkin-ubiquitin-p62-mediated mitochondria priming, *J. Biol. Chem.* **285** (2010), pp. 27879-27890.
103. Y. Ichimura, T. Kumanomidou, Y.-S. Sou, T. Mizushima, J. Ezaki, T. Ueno, E. Kominami, T. Yamane, K. Tanaka and M. Komatsu, Structural basis for sorting mechanism of p62 in selective autophagy, *J. Biol. Chem.* **283** (2008), pp. 22847-22857.
104. K. Satoo, N. N. Noda, H. Kumeta, Y. Fujioka, N. Mizushima, Y. Ohsumi and F. Inagaki, The structure of Atg4B-LC3 complex reveals the mechanism of LC3 processing and delipidation during autophagy, *EMBO J.* **28** (2009), pp. 1341-1350.

105. P. Wild, H. Farhan, D. G. McEwan, S. Wagner, V. V. Rogov, N. R. Brady, B. Richter, J. Korac, O. Waidmann, C. Choudhary, V. Dotsch, D. Bumann and I. Dikic, Phosphorylation of the Autophagy Receptor Optineurin Restricts Salmonella Growth, *Science* **333** (2011), pp. 228-233.
106. J. Mohrluder, T. Stangler, Y. Hoffmann, K. Wiesehan, A. Mataruga and D. Willbold, Identification of calreticulin as a ligand of GABARAP by phage display screening of a peptide library, *FEBS J.* **274** (2007), pp. 5543-5555.
107. N. Mizushima, B. Levine, A. M. Cuervo and D. J. Klionsky, Autophagy fights disease through cellular self-digestion, *Nature* **451** (2008), pp. 1069-1075.
108. B. Levine and G. Kroemer, Autophagy in the pathogenesis of disease, *Cell* **132** (2008), pp. 27-42.
109. S. H. Ralston, Pathogenesis of Paget's disease of bone, *Bone* **43** (2008), pp. 819-825.
110. G. J. Lucas, A. Daroszewska and S. H. Ralston, Contribution of genetic factors to the pathogenesis of Paget's disease of bone and related disorders, *J. Bone Miner. Res.* **21** (2006), pp. 31-37.
111. D. C. Rubinsztein, J. E. Gestwicki, L. O. Murphy and D. J. Klionsky, Potential therapeutic applications of autophagy, *Nat Rev Drug Discov* **6** (2007), pp. 304-312.
112. R. J. Youle and D. P. Narendra, Mechanisms of mitophagy, *Nat. Rev. Mol. Cell Biol.* **12** (2011), pp. 9-14.
113. P. Fei, W. Wang, S. H. Kim, S. Wang, T. F. Burns, J. K. Sax, M. Buzzai, D. T. Dicker, W. G. McKenna, E. J. Bernhard and W. S. El-Deiry, Bnip3L is induced by p53 under hypoxia, and its knockdown promotes tumor growth, *Cancer Cell* **6** (2004), pp. 597-609.
114. M. Matsushima, T. Fujiwara, E. Takahashi, T. Minaguchi, Y. Eguchi, Y. Tsujimoto, K. Suzumori and Y. Nakamura, Isolation, mapping, and functional analysis of a novel human cDNA (BNIP3L) encoding a protein homologous to human NIP3, *Genes Chromosomes Cancer* **21** (1998), pp. 230-235.
115. S. Daido, T. Kanzawa, A. Yamamoto, H. Takeuchi, Y. Kondo and S. Kondo, Pivotal role of the cell death factor BNIP3 in ceramide-induced autophagic cell death in malignant glioma cells, *Cancer Res.* **64** (2004), pp. 4286-4293.
116. A. Diwan, A. G. Koesters, A. M. Odley, S. Pushkaran, C. P. Baines, B. T. Spike, D. Daria, A. G. Jegga, H. Geiger, B. J. Aronow, J. D. Molkenin, K. F. Macleod, T. A. Kalfa and G. W. Dorn, 2nd, Unrestrained erythroblast development in Nix<sup>-/-</sup> mice reveals a mechanism for apoptotic modulation of erythropoiesis, *Proc. Natl Acad. Sci. USA* **104** (2007), pp. 6794-6799.
117. S. Vernier-Magnin, S. Muller, M. Sallot, J. Radom, J. F. Musard, P. Adami, P. Dulieu, J. P. Remy-Martin, M. Jouvenot and A. Fraichard, A novel early estrogen-regulated gene *gec1* encodes a protein related to GABARAP, *Biochem. Biophys. Res. Commun.* **284** (2001), pp. 118-125.
118. J. F. Clement, S. Meloche and M. J. Servant, The IKK-related kinases: from innate immunity to oncogenesis, *Cell Res.* **18** (2008), pp. 889-899.
119. K. A. Fitzgerald, D. C. Rowe, B. J. Barnes, D. R. Caffrey, A. Visintin, E. Latz, B. Monks, P. M. Pitha and D. T. Golenbock, LPS-TLR4 signaling to IRF-3/7 and NF-kappaB involves the toll adapters TRAM and TRIF, *J. Exp. Med.* **198** (2003), pp. 1043-1055.
120. F. Ikeda, C. M. Hecker, A. Rozenknop, R. D. Nordmeier, V. Rogov, K. Hofmann, S. Akira, V. Dötsch and I. Dikic, Involvement of the ubiquitin-like domain of TBK1/IKK-i kinases in regulation of IFN-inducible genes, *The EMBO Journal* **26** (2007), pp. 3451-3462.
121. J.-S. Gatot, R. Gioia, T.-L. Chau, F. Patrascu, M. Warnier, P. Close, J.-P. Chapelle, E. Muraille, K. Brown, U. Siebenlist, J. Piette, E. Dejardin and A. Chariot,

- Lipopolysaccharide-mediated interferon regulatory factor activation involves TBK1-IKKeppsi-dependent Lys(63)-linked polyubiquitination and phosphorylation of TANK/I-TRAF, *J. Biol. Chem.* **282** (2007), pp. 31131-31146.
122. R. Lin, Y. Mamane and J. Hiscott, Structural and functional analysis of interferon regulatory factor 3: localization of the transactivation and autoinhibitory domains, *Mol. Cell. Biol.* **19** (1999), pp. 2465-2474.
123. K. Wüthrich, Protein structure determination in solution by NMR spectroscopy, *J. Biol. Chem.* **265** (1990), pp. 22059-22062.
124. T. Carlomagno, Ligand-target interactions: what can we learn from NMR?, *Annu. Rev. Biophys. Biomol. Struct.* **34** (2005), pp. 245-266.
125. L. Fielding, NMR methods for the determination of protein-ligand dissociation constants, *Progress in Nuclear Magnetic Resonance Spectroscopy* **51** (2007), pp. 219-242.
126. K. Wuthrich, NMR studies of structure and function of biological macromolecules (Nobel Lecture), *J. Biomol. NMR* **27** (2003), pp. 13-39.
127. K. Pervushin, R. Riek, G. Wider and K. Wüthrich, Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, *Proc. Natl Acad. Sci. USA* **94** (1997), pp. 12366-12371.
128. C. Wang, A. F. Castro, D. M. Wilkes and G. A. Altenberg, Expression and purification of the first nucleotide-binding domain and linker region of human multidrug resistance gene product: comparison of fusions to glutathione S-transferase, thioredoxin and maltose-binding protein, *Biochem. J.* **338** (1999), pp. 77-81.
129. V. De Marco, G. Stier, S. Blandin and A. de Marco, The solubility and stability of recombinant proteins are increased by their fusion to NusA, *Biochem. Biophys. Res. Commun.* **322** (2004), pp. 766-771.
130. J. G. Marblestone, S. C. Edavettal, Y. Lim, P. Lim, X. Zuo and T. R. Butt, Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO, *Protein Sci.* **15** (2006), pp. 182-189.
131. K. D. Pryor and B. Leiting, High-level expression of soluble protein in Escherichia coli using a His6-tag and maltose-binding-protein double-affinity fusion system, *Protein Expr. Purif.* **10** (1997), pp. 309-319.
132. A. M. Catanzariti, T. A. Soboleva, D. A. Jans, P. G. Board and R. T. Baker, An efficient system for high-level expression and easy purification of authentic recombinant proteins, *Protein Sci.* **13** (2004), pp. 1331-1339.
133. T. R. Butt, S. Jonnalagadda, B. P. Monia, E. J. Sternberg, J. A. Marsh, J. M. Stadel, D. J. Ecker and S. T. Crooke, Ubiquitin fusion augments the yield of cloned gene products in Escherichia coli, *Proc. Natl Acad. Sci. USA* **86** (1989), pp. 2540-2544.
134. Y. K. Chae, H. Lee and W. Lee, Ubiquitin Fusion System for Recombinant Peptide Expression and Purification: Application to the Cytoplasmic Domain of Syndecan-4, *Bull. Korean Chem. Soc.* **28** (2007), pp. 1549-1552.
135. Y. Thielmann, O. H. Weiergraber, J. Mohrluder and D. Willbold, Structural framework of the GABARAP-calreticulin interface--implications for substrate binding to endoplasmic reticulum chaperones, *FEBS J.* **276** (2009), pp. 1140-1152.
136. O. H. Weiergraber, T. Stangler, Y. Thielmann, J. Mohrluder, K. Wiesehan and D. Willbold, Ligand binding mode of GABAA receptor-associated protein, *J. Mol. Biol.* **381** (2008), pp. 1320-1331.
137. J. Mohrluder, M. Schwarten and D. Willbold, Structure and potential function of gamma-aminobutyrate type A receptor-associated protein, *FEBS J.* **276** (2009), pp. 4989-5005.

138. Y. Thielmann, O. H. Weiergräber, P. Ma, M. Schwarten, J. Mohrlüder and D. Willbold, Comparative modeling of human NSF reveals a possible binding mode of GABARAP and GATE-16, *Proteins* **77** (2009), pp. 637-646.
139. D. Hanahan, Studies on transformation of *Escherichia coli* with plasmids, *J. Mol. Biol.* **166** (1983), pp. 557-580.
140. J. Keeler. (2005). *Understanding NMR Spectroscopy*. Wiley edit.
141. D. S. Wishart and B. D. Sykes, The <sup>13</sup>C chemical-shift index: a simple method for the identification of protein secondary structure using <sup>13</sup>C chemical-shift data, *J. Biomol. NMR* **4** (1994), pp. 171-180.
142. G. Cornilescu, F. Delaglio and A. Bax, Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J. Biomol. NMR* **13** (1999), pp. 289-302.
143. T. Herrmann, P. Guntert and K. Wuthrich, Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA, *J. Mol. Biol.* **319** (2002), pp. 209-227.
144. S. Vijay-Kumar, C. E. Bugg and W. J. Cook, Structure of Ubiquitin Refined at 1.8 Å Resolution, *J. Mol. Biol.* **194** (1987), pp. 531-544.
145. J. W. Ponder and D. A. Case, Force fields for protein simulations, *Adv. Protein Chem.* **66** (2003), pp. 27-85.
146. R. Koradi, M. Billeter and P. Guntert, Point-centered domain decomposition for parallel molecular dynamics simulation, *Computer Physics Communications* **124** (2000), pp. 139-147.
147. F. A. Mulder, D. Schipper, R. Bott and R. Boelens, Altered flexibility in the substrate-binding site of related native and engineered high-alkaline *Bacillus subtilis*ins, *J. Mol. Biol.* **292** (1999), pp. 111-123.
148. S. Kim, D. N. Cullis, L. A. Feig and J. D. Baleja, Solution structure of the Repl1 EH domain and characterization of its binding to NPF target sequences, *Biochemistry* **40** (2001), pp. 6776-6785.
149. A. M. Eldridge, H. S. Kang, E. Johnson, R. Gunsalus and F. W. Dahlquist, Effect of phosphorylation on the interdomain interaction of the response regulator, NarL, *Biochemistry* **41** (2002), pp. 15173-15180.
150. B. Y. Qin, C. Liu, S. S. Lam, H. Srinath, R. Delston, J. J. Correia, R. Derynck and K. Lin. (2003). Crystal structure of IRF-3 reveals mechanism of autoinhibition and virus-induced phosphoactivation. In *Nat. Struct. Biol.*, Vol. 10, pp. 913-921.
151. V. V. Rogov, A. Rozenknop, N. I. Rogova, F. Löhr, V. Zharavin, S. Tikole, P. Güntert, I. Dikic and V. Dötsch, NMR-compatible expression tag provides new perspectives in the study of proteins and their interactions, *in preparation* (2011), pp.
152. A. Rozenknop, V. V. Rogov, N. I. Rogova, F. Löhr, P. Güntert, I. Dikic and V. Dötsch, Characterization of the interaction of GABARAPL-1 with the LIR motif of NBR1, *J. Mol. Biol.* **410** (2011), pp. 477-487.
153. R. Harris and P. C. Driscoll, The Ubiquitin NMR Resource, *Modern NMR Spectroscopy in Education* **969** (2007), pp. 114-126.
154. B. Lopez-Mendez and P. Guntert, Automated protein structure determination from NMR spectra, *J. Am. Chem. Soc.* **128** (2006), pp. 13112-13122.
155. D. E. Zimmerman, C. A. Kulikowski, Y. Huang, W. Feng, M. Tashiro, S. Shimotakahara, C. Chien, R. Powers and G. T. Montelione, Automated analysis of protein NMR assignments using methods from artificial intelligence, *J. Mol. Biol.* **269** (1997), pp. 592-610.
156. S. Tikole, V. Zharavin, V. Rogov, A. Rozenknop, N. Rogova, F. Löhr, I. Dikic, V. Dötsch and P. Güntert, Optimization of Non-Uniform Sampling and Peak Identification

- for the Automated Chemical Shift Assignment and Interaction Studies of Proteins, *in preparation* (2011), pp.
157. K. Takahashi, N. N. Suzuki, M. Horiuchi, M. Mori, W. Suhara, Y. Okabe, Y. Fukuhara, H. Terasawa, S. Akira, T. Fujita and F. Inagaki, X-ray crystal structure of IRF-3 and its functional implications, *Nat. Struct. Biol.* **10** (2003), pp. 922-927.
  158. H. Wang, F. K. Bedford, N. J. Brandon, S. J. Moss and R. W. Olsen, GABA(A)-receptor-associated protein links GABA(A) receptors and the cytoskeleton, *Nature* **397** (1999), pp. 69-72.
  159. M. Yamaguchi, N. N. Noda, H. Nakatogawa, H. Kumeta, Y. Ohsumi and F. Inagaki, Autophagy-related Protein 8 (Atg8) Family Interacting Motif in Atg3 Mediates the Atg3-Atg8 Interaction and Is Crucial for the Cytoplasm-to-Vacuole Targeting Pathway, *J. Biol. Chem.* **285** (2010), pp. 29599-29607.
  160. Y. Chen, C. Chen, E. Kotsikorou, D. L. Lynch, P. H. Reggio and L. Y. Liu-Chen, GEC1-kappa opioid receptor binding involves hydrophobic interactions: GEC1 has chaperone-like effect, *J. Biol. Chem.* **284** (2009), pp. 1673-1685.
  161. L. Huang, F. Hofer, G. S. Martin and S. H. Kim, Structural basis for the interaction of Ras with RalGDS, *Nat. Struct. Biol.* **5** (1998), pp. 422-426.
  162. C. E. Stebbins, W. G. Kaelin, Jr. and N. P. Pavletich, Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function, *Science* **284** (1999), pp. 455-461.
  163. H. Wang and R. W. Olsen, Binding of the GABA(A) receptor-associated protein (GABARAP) to microtubules and microfilaments suggests involvement of the cytoskeleton in GABARAP-GABA(A) receptor interaction, *J. Neurochem.* **75** (2000), pp. 644-655.

## 9. Appendix

### 9.1. Protein sequences

>LC3A (after cleavage of NusA moiety)

GAMGRPFKQRRSFADRCKEVQQIRDQHPSKIPVIIERYKGEKQLPVLDKTKFLVPDHSVNMSELVKIIRRRRLQLNP  
QAFFLLVNQHSMVSVSTPIADIYEQEKDEDEGFLYMVYASQETF

>LC3B (after cleavage of NusA moiety)

GAMGKTFKQRRTFEQRVEDVRLIREQHPTKIPVIIERYKGEKQLPVLDKTKFLVPDHSVNMSELIKIIRRRRLQLNANQ  
QAFFLLVNGHSMVSVSTPISEVYESEKDEDEGFLYMVYASQETF

>GABARAPL-1 (after cleavage of Ub moiety)

GAMFKFQYKEDHPFEYRKKEGKIRKKYPDRVPVIVEKAPKARVPDLDRKYLVPDITVGFYFLIRKRIHLRPED  
ALFFFVNNTIPPTSATMGQLYEDNHEEDYFLYVAYSDES

>GABARAPL-2 (after cleavage of GST moiety)

GSPEFKWMFKEDHSLEHRCVESAKIRAKYPDRVPVIVEKVSQSQIVDIDKRYLVPDITVAQFMWIIRKRIQLPSE  
KAIFLFVDKTVPQSSLTMGQLYEKEKDEDEGFLYVAYS

>Ubiquitin (after cleavage of fused protein or peptide)

MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQELIFAGKQLEDGRTLSDYNIQKESTLHLVLE  
QLESASGSGHHHHHSAGENLYFQ

>TBK1\_ULD (after cleavage of GST moiety)

GSPEFTSDVLRHVMIVHVSLOHMTAKIYIHSYNTAAVFHELIVYKQTKIVSSNQELIYEGRRLLVLELGR  
LAQHFPKTEENPIFVTSLERPHRD

NB: Amino acids from cloning artefacts and site-directed mutagenesis are underlined.

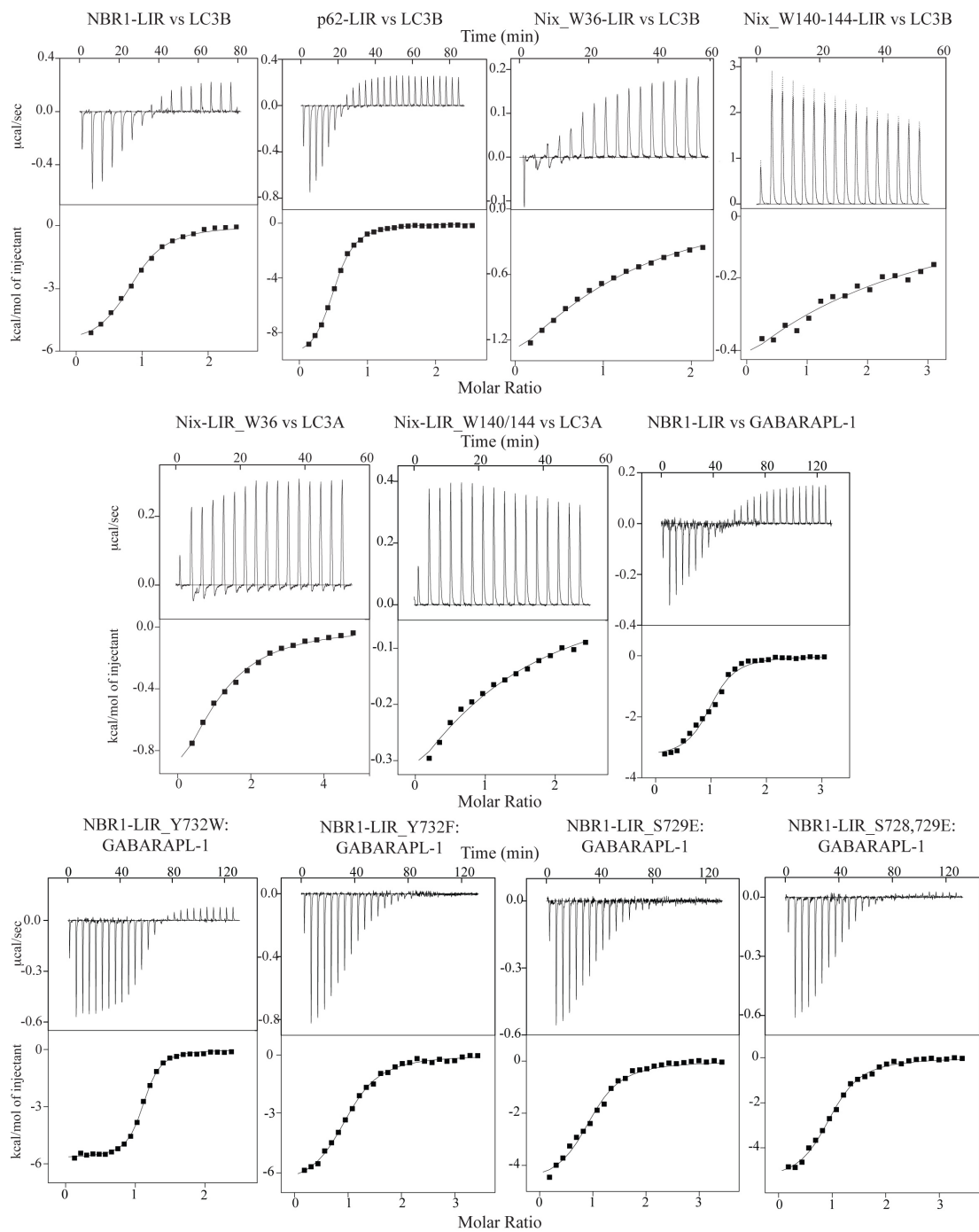
### 9.2. Peptide sequences

LIR motif	Synthesized peptides	Ub-fused peptides
p62	<b>RPEEQMESDNCSSGGDDDWTHLS</b>	<b>GAMGDDDWTHLSS</b>
NBR1		<b>GAMGSASSEDIILPES</b>
NBR1-LIR_Y732W		<b>GAMGSASSEDDWIIILPES</b>
NBR1-LIR_Y732F	<b>GAMGSASSEDFIIILPES</b>	
NBR1-LIR_S729E	<b>GAMGSASEEDYIIILPES</b>	
NBR1-LIR_S728,729E	<b>GAMGSAEEEDYIIILPES</b>	
Nix-W36	<b>GLNSSWVELPMSSN</b>	
Nix-W140/144	<b>SADWVDWSSRPENIP</b>	

**Table 8: List of peptides used.**

The different LIR sequences used in this work are represented in bold characters with the aromatic residue at position 1 shown in red, the hydrophobic residue at position 4 in blue and the negatively charged residues in green. Any additional residues resulting from cloning artifacts are shown in italic characters.

## 9.3. ITC raw data



**Figure 61: ITC raw data of the interaction of MAP1LC3 proteins with LIR domains.** All experiments were performed at 25°C in 50 mM sodium phosphate, 100 mM NaCl at pH 7.0.



## 9.4. GABARAPL-1/NBR1-LIR intermolecular NOEs

24	LYS	HB3	730	GLU	HB2	5.28	53	TYR	QD	735	ILE	QG2	7.46
24	LYS	HB3	730	GLU	HB3	5.60	53	TYR	QD	735	ILE	QD1	7.70
25	ILE	QG2	730	GLU	HN	7.10	53	TYR	QE	733	ILE	QG2	6.88
25	ILE	QG2	730	GLU	QG	7.80	53	TYR	QE	733	ILE	QG1	8.60
25	ILE	HG12	730	GLU	HB2	5.50	53	TYR	QE	733	ILE	QD1	8.80
25	ILE	HG12	730	GLU	HB3	5.50	53	TYR	QE	735	ILE	QG2	7.25
25	ILE	QD1	729	SER	HB2	4.82	53	TYR	QE	735	ILE	HG12	6.88
25	ILE	QD1	729	SER	HB3	6.33	53	TYR	QE	735	ILE	HG13	7.70
25	ILE	QD1	730	GLU	HN	6.30	53	TYR	QE	735	ILE	QD1	8.01
25	ILE	QD1	730	GLU	HB2	6.60	54	LEU	HN	734	ILE	HA	5.05
25	ILE	QD1	730	GLU	HB3	6.60	54	LEU	HN	734	ILE	QD1	6.60
25	ILE	QD1	730	GLU	QG	6.80	54	LEU	HA	734	ILE	QD1	7.10
25	ILE	QD1	732	TYR	QD	7.65	54	LEU	HB2	732	TYR	HB3	4.94
28	LYS	HB2	729	SER	HB2	4.56	54	LEU	HB2	734	ILE	QG2	5.57
28	LYS	HB3	729	SER	HB2	4.63	54	LEU	HB3	732	TYR	QD	6.89
28	LYS	HD2	729	SER	HN	5.60	54	LEU	HB3	734	ILE	HA	4.59
29	TYR	QD	729	SER	HB2	6.93	54	LEU	HB3	734	ILE	HG12	4.85
29	TYR	QD	734	ILE	HG13	6.90	54	LEU	HB3	734	ILE	QD1	5.35
29	TYR	QE	729	SER	HB2	7.08	54	LEU	HG	732	TYR	QD	7.23
29	TYR	QE	734	ILE	HG13	6.60	54	LEU	QD1	732	TYR	HB3	5.54
29	TYR	QE	734	ILE	QD1	6.95	54	LEU	QD1	732	TYR	QD	8.57
34	PRO	HB3	732	TYR	QE	6.76	54	LEU	QD1	732	TYR	QE	8.62
36	ILE	QG1	732	TYR	QE	7.25	54	LEU	QD1	734	ILE	HG12	6.60
52	LYS	HB2	733	ILE	QD1	6.18	54	LEU	QD1	734	ILE	HG13	6.60
52	LYS	HB3	732	TYR	QD	7.06	54	LEU	QD1	734	ILE	QD1	6.56
52	LYS	HB3	732	TYR	QE	6.89	55	VAL	HN	736	LEU	QD2	7.10
53	TYR	HN	732	TYR	QD	7.43	55	VAL	QG1	736	LEU	QD2	7.60
53	TYR	HA	733	ILE	QG2	6.47	56	PRO	HB2	736	LEU	QD2	6.60
53	TYR	HA	735	ILE	HB	5.80	56	PRO	HB3	736	LEU	HB2	5.50
53	TYR	HB3	733	ILE	QG2	6.59	56	PRO	HB3	736	LEU	HB3	5.43
53	TYR	HB3	735	ILE	QD1	6.90	56	PRO	HB3	736	LEU	QD2	5.80
53	TYR	QD	732	TYR	QE	8.90	56	PRO	QD	736	LEU	QD2	6.60
53	TYR	QD	733	ILE	QG2	7.39	68	ILE	HG13	735	ILE	QD1	7.10
53	TYR	QD	733	ILE	QD1	8.80	68	ILE	QD1	735	ILE	QD1	7.70
53	TYR	QD	735	ILE	HB	7.09							

## 9.5. NMR structural statistics

<i>Input restraint statistics</i>	
Total number of meaningful distance restraints	760 (58) <sup>a</sup>
Intraresidual (i = j)	136
Sequential ( i - j =1)	206
Medium range (1 <  i - j  < 4)	203 (26)
Long range ( i - j  > 4)	215 (32)
Torsion angle restraints	91
<i>Restraint violations in final ensemble (20 conformers)</i>	
Distance restraint violations	114
Van der Waals restraints violations	47
Angle restraints violations	26
<i>RMS deviations from experimental restraints</i>	
Distance restraints (Å)	n/a <sup>b</sup>
Angle restraints (deg)	n/a <sup>b</sup>
<i>RMS deviations from idealized covalent geometry</i>	
Bond lengths (Å)	0.014
Bond angles (deg)	1.6
<i>PROCHECK Ramachandran plot analysis (%)</i>	
Residues in most favoured regions	60.8
Residues in additionally allowed regions	21.4
Residues in generously allowed regions	13.0
Residues in disallowed regions	4.9
<i>Structural precision<sup>c</sup>, RMSD (Å) to mean structure</i>	
Backbone atoms N, C <sup>α</sup> , C <sup>β</sup>	0.65 ± 0.11
All heavy atoms	1.44 ± 0.13
<sup>a</sup> The number of included H-bond restraints is indicated in parentheses.	
<sup>b</sup> No energy minimization was performed	
<sup>c</sup> Values for the structured part (residues 14-88)	

Table 9: Structural statistics of the 20 energy-minimized conformers of TBK1\_ULD

<i>Input restraint statistics</i>	
Total number of meaningful distance restraints	1448 (84) <sup>a</sup>
Intraresidual (i = j)	207
Sequential ( i – j =1)	436
Medium range (1< i – j < 4)	295 (46) <sup>a</sup>
Long range ( i – j  > 4)	510 (38) <sup>a</sup>
Intermolecular	69 (8) <sup>b</sup>
Torsion angle restraints	203
<i>Restraint violations in final ensemble (20 conformers)</i>	
Distance restraint violations	
Number > 0.1 Å	0
Maximal violations (Å)	0.09
RMS deviations from experimental restraints	
Distance restraints (Å)	0.0081 ± 0.0004
Angle restraints (deg)	0.55 ± 0.04
<i>RMS deviations from idealized covalent geometry</i>	
Bond lengths (Å)	0.0144 ± 0.0001
Bond angles (deg)	1.730 ± 0.024
<i>PROCHECK Ramachandran plot analysis (%)</i>	
Residues in most favoured regions	85.3
Residues in additionally allowed regions	13.3
Residues in generously allowed regions	1.1
Residues in disallowed regions	0.3
<i>Structural precision<sup>c</sup>, RMSD (Å) to mean structure</i>	
Backbone atoms N, C <sup>α</sup> , C'	0.65 ± 0.11
All heavy atoms	1.21 ± 0.10
<sup>a</sup> The number of included H-bond restraints is indicated in parentheses.	
<sup>b</sup> The number of unambiguously manually assigned intermolecular NOEs is indicated in parentheses.	
<sup>c</sup> Values for the structured part (GABARAPL-1 residues 12-114, NBR1 residues 729-737).	

**Table 10: Structural statistics of the 20 energy-minimized conformers of GABARAPL-1/NBR1-LIR**



**10. List of publications**

- Ikeda F, Hecker CM, **Rozenknop A**, Nordmeier RD, Rogov V, Hofmann K, Akira S, Dötsch V, Dikic I. *Involvement of the ubiquitin-like domain of TBK1/IKK-i kinases in regulation of IFN-inducible genes*. EMBO J (2007), **26**:3451-3462
- Novak I, Kirkin V, McEwan DG, Zhang J, Wild P, **Rozenknop A**, Rogov V, Löhr F, Popovic D., Occhipinti A, Reichert A., Terzic J, Dötsch V, Ney P, Dikic I. *Nix is a selective autophagy receptor for mitochondrial clearance*. EMBO Rep (2010); **11**:45-51
- **Rozenknop A**, Rogov V, Rogova N, Löhr F, Güntert P, Dikic I, Dötsch V. *Characterization of the interaction of GABARAPL-1 with the LIR motif of NBR1*. JMB (2011); **410**:477-487
- Rogov V, **Rozenknop A**, Rogova N, Löhr F, Zharavin V, Tikole S, Güntert P, Dikic I, Dötsch V. *NMR-compatible expression tag provides new perspectives in the study of proteins and their interactions* (in preparation)



## 11. Curriculum Vitae

### Alexis ROZENKNOP

*Date of birth:* January 9<sup>th</sup>, 1981 at Montréal (Canada)

*Nationality:* French



### Education

---

01/2007-08/2011: **Graduation** (Goethe University, Frankfurt Am Main) under the supervision of Prof. Dötsch and Prof. Dikic

**“Structural characterization and interaction studies of ubiquitin-like proteins”**

09/2001-09/2006: **Academic studies of biology** (Toulouse University, France)

Biochemistry, Biotechnologies

*Degree:* Master Biochemistry

*Master thesis:* **“Protein-Ligand Interaction by STD-NMR”**

09/1996-06/1999: **High School Studies**

**Baccalauréat scientifique** (general certificate of education) at Aix-en-Provence, France

### Professional experience

---

01/2007-08/2011: **Scientific co-worker** in the Institute of Biophysical Chemistry, research group of Prof. Dötsch, Goethe University Frankfurt

- Protein expression and purification
- Protein-protein interaction studies by NMR and ITC

02/2006-09/2006: **Internship** in the Structural Biology group at Sanofi-Aventis, Frankfurt

- Protein-ligand interaction studies by STD-NMR

10/2004-01/2005: **Internship** in chemistry laboratory (LSPMIB) at Toulouse University

- Biophysical and chemical studies of antibiotics derivatives

01/2007-12/2010: Organisation and supervision of students practicals

- Basics on Proteins and Molecular Biology tools
- PubMed

10/2004-01/2006: Tutor for young students in Mathematics, French and English

10/2004-06/2005: Coach of softball team of Toulouse University

Sommerjobs (vintage, handling ...)

## Skills and Expertise

---

### *Languages*

- French (native)
- English (fluent: Advanced Level in English Language Skills Certificate at the English Language Institute of the University of Delaware, USA)
- German (fluent)

### *Computer*

- Windows and Mac OS
- MS Office
- Graphic programs: Adobe Illustrator and Photoshop
- Scientific programs: Microcal Origin, Endnote, Pymol, Topsin, Sparky, Unicorn for Äkta systems

## Scientific achievements

---

### *Methods knowledge*

- Molecular and microbiological methods (cloning, transformation, cell growth, protein expression in *E. coli* cells)
- Protein isolation and purification
- Biochemistry and biophysical methods (SDS-PAGE, CD spectroscopy, UV/Vis techniques, NMR and ITC)
- Project and research management
- Scientific information search

### *Further qualifications*

Graduate Program of the Cluster of Excellence Frankfurt (GRACE) workshops

- Negotiation skills (Center for Science & Research Management)
- Presentation skills
- Project Management (Boston Consulting Group)
- Time and self-management

### *Scientific publications*

- Ikeda F, Hecker CM, **Rozenknop A**, Nordmeier RD, Rogov V, Hofmann K, Akira S, Dötsch V, Dikic I. *Involvement of the ubiquitin-like domain of TBK1/IKK-i kinases in regulation of IFN-inducible genes*. EMBO J (2007), **26**:3451-3462
- Novak I, Kirkin V, McEwan DG, Zhang J, Wild P, **Rozenknop A**, et al. *Nix is a selective autophagy receptor for mitochondrial clearance*. EMBO Rep (2010); **11**:45-51
- **Rozenknop A**, Rogov V, Rogova N, Löhr F, Güntert P, Dikic I, Dötsch V. *Characterization of the interaction of GABARAPL-1 with the LIR motif of NBRI*. JMB (2011); **410**:477-487
- Rogov V, **Rozenknop A**, Rogova N, Löhr F, Zharavin V, Tikole S, Güntert P, Dikic I, Dötsch V. *NMR-compatible expression tag provides new perspectives in the study of proteins and their interactions* (in preparation)



*Talks and Poster presentations on internal and external conferences*

- Structural and functional analysis of the ubiquitin-like domain of TBK1 kinases involved in regulation of IFN-inducible genes, *International Conference on Molecular Perspectives on Protein-Protein Interactions* Dubrovnik, Croatia (2008) (poster presentation)
- Structural and functional analysis of the ubiquitin-like domain of TBK1 kinases involved in regulation of IFN-inducible genes, *Annual Meeting of EU-NMR and NMR-LIFE*, Frankfurt, Germany (2008) (poster presentation)
- Physiological and Structural Studies of Ubiquitin Like Domain, *2<sup>nd</sup> Biomolecular Magnetic Resonance (BMRZ) PhD/Postdoc Workshop*, Frankfurt, Germany (2008) (talk)
- Structural characterization & Interaction Studies of Ubiquitin Like Proteins involved in Autophagy, *4<sup>th</sup> (BMRZ) PhD/Postdoc Workshop*, Frankfurt, Germany (2010) (talk)
- Characterization of the interaction of MAP1LC3 proteins with the LIR motif of autophagy receptors, *2<sup>nd</sup> PhD Symposium Research in Molecular and Cell Biological Sciences*, Frankfurt, Germany (2011) (talk)

*ENDOCYTE fellowship (Marie Curie Research Training Networks)*

- Kick-off meeting & 1<sup>st</sup> Training Workshop: "Growth factor signaling & Cytoskeletal control" (May 2007 in Split, Croatia)
- 2<sup>nd</sup> Training Workshop: "Image-based high-throughput screening methods and tools for pathway modeling" (May 2008 in Heidelberg, Germany)
- Mid Term Review Meeting & 3<sup>rd</sup> Training Workshop: "Life And Death in the Nervous System" (September 2008 in Israel)

**Miscellaneous**

---

*Interests*

- Sports (football, rugby, baseball, jogging)
- Travels
- Movies
- Reading

*Others*

- Driver license





