INVESTIGATING THE ROLE OF BECN1 CONFORMATIONALLY FLEXIBLE REGIONS AND INVARIANT CYS-X-X-CYS MOTIFS IN AUTOPHAGY

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Shreya Mukhopadhyay

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Department:
Chemistry & Biochemistry

November 2020

Fargo, North Dakota

North Dakota State University
Graduate School

Title

INVESTIGATING THE ROLE OF BECN1 CONFORMATIONALY FLEXIBLE REGIONS AND INVARIANT CYS-X-X-CYS MOTIFS IN AUTOPHAGY

By

Shreya Mukhopadhyay

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Dr. Sangita Sinha
Chair

Dr. Kalpana Katti

Dr. Guodong Liu

Dr. Christopher L. Colbert

Approved:

12/11/2020
Date

Dr. Gregory Cook
Department Chair

ABSTRACT

Autophagy is an essential catabolic cellular homeostasis process conserved in all eukaryotes. BECN1, a key autophagy protein involved in autophagosome nucleation, comprises of a large, poorly conserved, N-terminal intrinsically disordered region (IDR); a flexible helical domain; a coiled-coil domain (CCD) that forms anti-parallel homodimers in the absence of other interacting partners; and a β-α repeated autophagy specific domain (BARAD).

The IDR of higher eukaryotes includes a BCL2 homology 3 domain (BH3D), which undergoes dramatic disorder-to-helix transitions upon binding to anti-apoptotic and anti-autophagic BCL2s. We show that the BH3D is not required for starvation-induced autophagy upregulation suggesting that BCL2-binding to the BH3D does not directly impede a pro-autophagy function of the BH3D, rather it may impact structure, oligomerization, interactions and function of other BECN1 domains.

CCD C-terminal residues, named the overlap helix (OH), pack in two mutually-exclusive states stabilized by the same interface residues: against either the partner helix in a CCD dimer or the BARAD. We show that mutation of these interface residues abrogates starvation-induced autophagy upregulation. Together with our complementary structural studies, this suggests that autophagy-inactive BECN1 adopts conformations preventing the BARAD from membrane-association, with BECN1 heterodimerization with ATG14 or UVRAG disrupting this inhibitory conformation. In the BECN1 homodimer, the OH packs against a nuclear export signal sequence (NES) at the N-terminus of the partner CCD. We show that when released from this interaction, the NES can interact with the complex of the nuclear exporter, Chromosomal Region Maintenance 1 protein and a GTP-bound small G-protein, Ran. This interaction is essential for BECN1 export to the cytoplasm, and for autophagy.

Two invariant CxxC motifs bookend the IDR. We find that both CxxC motifs are required, but the intervening IDR is less important, for starvation-triggered upregulation of autophagy. We demonstrate that BECN1 binds Zn^{2+} in a 1:1 molar ratio. Further, mutation of the invariant cysteines or treatment with reducing agent abrogates Zn^{2+} co-ordination, demonstrating that the invariant CxxC motifs are responsible for binding Zn^{2+}. We use diverse biophysical methods to show that Zn^{2+}-binding impacts the conformation and structural transitions of the BECN1 IDR, thereby playing an important role in regulating autophagy.

ACKNOWLEDGMENTS

I deeply express my gratitude to my advisor Dr. Sangita Sinha for her immense contribution in my doctoral work. I thank her for allowing me to join her lab as a graduate student and pursue my doctoral degree. She has always directed me towards approaching scientific problems objectively and constantly provided me feedback for my work whenever needed. I am grateful to her to give me the opportunity for learning various types of techniques that is needed for performing cut edge research in biochemistry. I thank her for keeping high standards for research that has made me think more critically about my experiments before performing it. This training will serve as a guideline for performing further research in my scientific career and will help me build upon the knowledge and experience I have obtained during the course of my doctoral degree. I am also extremely thankful to her for allowing me to pursue an industry internship that helped me explore the research opportunities in industry, initiate to build industry network and also vet myself in the current job market.

I would also like to thank my committee members Dr. Christopher Colbert, Dr. Guodong Liu and Dr. Kalpana Katti for providing valuable feedback during my committee meetings and graduate seminars. Dr. Colbert is also our collaborator and I am extremely fortunate to receive his guidance and feedback about my work during lab meetings and also on during one-one meetings. He has openly shared his ideas, knowledge, and expertise that I could put to use while planning and executing experiments. I thank Dr. Liu and Dr. Katti for bringing in diverse discussions during every meeting that allowed me to think beyond my current research.

I also extend my heartfelt thanks to my current and past lab mates of both the Sinha lab and Colbert lab. Thanks to Dr. Minfei Su, Dr. Yang Mei for helping me get started with my research work in Sinha lab. I am grateful to Dr. Karen Glover, Dr. Yue Li from Sinha group and

Dr. Jaime Jensen from Colbert group for always being open to scientific discussions and also for teaching me several lab techniques. My other lab members from Sinha lab Srinivasulu Dasanna, Elizabeth Bueno, Samuel Wyatt, Daniel Asa, Robyn Duttenhefner and from Colbert lab Shane Wyborny, Benjamin LeVahn, Beau Jernberg, Tajnin Sultana and Victor Arnold also have an instrumental role in my graduate career. Thank you for your constant support whether it be engaging in scientific discussions or maintaining a joyful and lighthearted spirit in the lab. I think I have been fortunate to get the best labmates one could expect in graduate school.

I also thank our collaborators Dr. Srinivas Chakravarthy and Dr. Jesse Hopkins at Bio-CAT, APS, Argonne for helping during data collection and also data analysis. I thank Dr. Pawel Borowicz at Advanced Imaging & Microscopy lab at NDSU for helping me during experimental design and also for teaching me to use the confocal microscope.

This research was funded by NIH grants RO3 NS090939, R15 GM122035 (S. S.), P20 RR015566 (S S.), R21 AI078198 (S. S.), and R15 GM113227 (C. C.); an NSF grant MCB-1413525 (S. S.). Work performed at Bio-CAT was supported by NIH NIGMS 9P41 GM103622 and use of the Pilatus 3 1M detector funded by NIH NIGMS 1S10OD018090-01.

DEDICATION

This dissertation is dedicated to my parents, sister and brother for their unconditional support, love, and motivation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Å	Angstrom
AFM	Aromatic Finger Mutation
AMBRA1	Autophagy And Beclin 1 Regulator 1
AMP	Adenosine Monophosphate
AMPK	AMP-activated Protein Kinase
ADP	Adenosine Diphosphate
ATP	Adenosine Triphosphate
ATG	Autophagy Related Protein
BARAD	β-α Repeated, Autophagy-Specific Domain
BCL2	B Cell Lymphoma 2
BH3D	BH Homology 3 Domain
BioXTAS	Biological X-ray Total Analysis System
CCD	Coiled-Coil Domain
CD	Circular Dichroism Spectroscopy
ΔG	Change in Gibbs Free Energy
ΔH	Change in Enthalpy
ΔS	Change in Entropy
D_{max}	Maximum Particle Size
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic Acid
EFA	Evolving Factor Analysis
EM	Electron Microscopy
EOM	Ensemble Optimization Method

FBS	Fetal Bovine Serum
FHD	Flexible Helical Doman
FL	Full Length
γHV68	γ Herpes Virus 68
GFP	Green Fluorescent Protein
GST	Glutathione S-Transferase
GTP	Guanosine Triphosphate
HDX-MS	Hydrogen-Deuterium Exchange Mass Spectrometry
His$_6$	Hexa-histidine
I(0)	Scattering Intensity at Zero Angle
I(q)	Scattering Intensity
ICP-MS	Inductively Coupled Mass Spectrometry
IDR	Intrinsically Disordered Region
IPTG	Isopropyl Thio-β-D-Galactoside
ITC	Isothermal Titration Calorimetry
K$_d$	Dissociation Constant
LC3	Light Chain 3
MBP	Maltose Binding Protein
MgOAc	Magnesium Acetate
MoRF	Molecular Recognition Feature
mTOR	Mammalian Target Of Rapamycin Complex 1
MWCO	Molecular Weight Cut-Off
Ni-NTA	Nickel-Nitriloacetic Acid
NEB	New England Biolabs
OH	Overlap Helix

PAGE	Polyacrylamide Gel Electrophoresis
PDB	Protein Data Bank
P(r)	Pairwise-Distance Distribution
PI	Phosphatidylinositol
PI3KC3	Class III Phosphatidylinositol 3 Kinase
PI3P	Phosphatidylinositol 3-Phosphate
R_g	Radius of Gyration
ROS	Reactive Oxygen Species
SAXS	Small Angle X-ray Scattering
SDS	Sodium Dodecyl Sulfate
SEC	Size Exclusion Chromatography
SVD	Singular Value Decomposition
TEV	Tobacco Etch Virus
TFE	2,2,2-Trifluoreothanol
TM	Transmembrane
T_m	Melting Temperature
TORC1	Target Of Rapamycin Complex 1
VMP1	Vacuole Membrane Protein 1
VPS	Vacuolar Protein Sorting
WT	Wild Type
θ	Molar ellipticity

CHAPTER 1: INTRODUCTION

1.1. Introduction to autophagy

Autophagy is a conserved lysosomal degradation pathway that plays a housekeeping role in removing misfolded, aggregated, mutated, or damaged proteins, organelles, and pathogens. So, autophagy is generally thought of as a survival mechanism, although its deregulation is linked to non-apoptotic cell death. There are three different kinds of autophagy: macroautophagy, microautophagy and chaperone mediated autophagy. In macroautophagy, the cellular components are sequestered in bi-layered vesicles called autophagosomes that fuse with lysosomes, enabling the degradation of the sequestered contents, and consequently the recycling of nutrients(Klionsky *et al.*, 2011). In microautophagy, the cytosolic components are directly taken up by the lysosome itself through invagination of the lysosomal membranes. Both macro and micro autophagy are able to engulf large structures through both selective and non-selective mechanisms. In chaperone mediated autophagy (CMA), targeted proteins are translocated across the membranes in complex with chaperone proteins (Hsc70) that are recognized by the lysosomal membrane receptor LAMP-2A (Lysosomal-associated protein 2A), in a complex with chaperone proteins resulting in their unfolding and degradation. Macroautophagy is the most extensively studied of these and is commonly referred to as autophagy. Henceforth, in this dissertation, all mentions of autophagy will refer to macroautophagy. Autophagy is upregulated to recycle nutrients and energy under stress conditions which includes intracellular stress conditions such as accumulation of damaged proteins and organelles and high bioenergetic demands; and extracellular stress conditions such as nutrient starvation, infection, hypoxia, formation of reactive oxygen species (ROS) and DNA damage (Kroemer *et al.*, 2010). Aberrant autophagy has been implicated in numerous diseases like neurodegenerative disorders, muscular diseases, cardiomyopathy, cancer and infectious diseases(Levine & Kroemer, 2008). Autophagy is mediated by several protein-protein interactions(Popelka *et al.*, 2014), that are often dependent on the conformational flexibility of these proteins(Glover *et al.*, 2017).

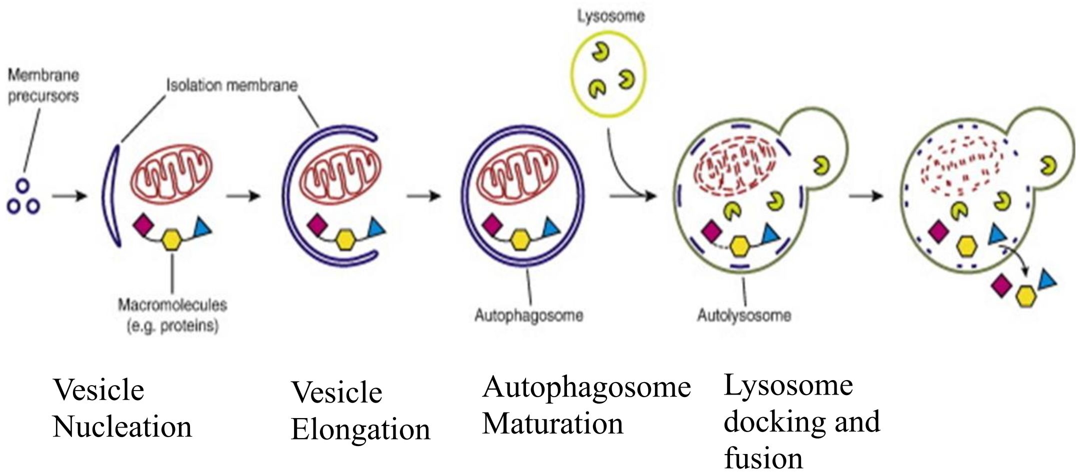


Figure 1.1: General steps of macroautophagy (Zaffagnini & Martens, 2016)

Autophagy involves several proteins that are conserved in eukaryotes. Many of these conserved autophagy-related effectors, called ATG proteins (Levine & Klionsky, 2004, Mizushima, 2010) form multi-protein complexes responsible for different stages of autophagy such as (1) vesicle nucleation, (2) vesicle elongation, (3) autophagosome maturation and (4) lysosome docking and fusion, followed by recycling of nutrients (Kroemer *et al.*, 2010). Autophagy is triggered by several cellular stress conditions, with lack of nutrition or starvation being one of the main inducers of autophagy. Glucose metabolism is the major source of ATP generation, which is the major energy currency in the cells. Therefore, carbon sources such as glucose directly affect the level of ADP/AMP and ATP. A high ADP+AMP to ATP concentration ratio indicates a cellular nutrient-deprivation stress condition. At high concentrations of ADP+AMP, AMP binds to and activates protein kinase (AMPK). Activated AMPK phosphorylates ULK1, which transduces the starvation signal to Vacuolar Protein Sorting 34 (VPS34) complex 1 by phosphorylation. Activated AMPK also activates VPS34 by phosphorylating mammalian (m)TORC1 (mTORC1) and so the ULK1 is no longer deactivated by mTORC1 due to phosphorylation. Another major source of cellular stress is starvation-induced oxidative stress. Nutrient deprivation in cells leads to the malfunction of mitochondrial electron transport. Electrons leak from NADH-Complex Q oxidoreductase and Coenzyme Q-Cytochrome c oxidoreductase, resulting in partial oxygen reduction to superoxide radicals ($O_2^{\bullet-}$). The superoxide radicals may be converted to hydrogen peroxide (H_2O_2) by mitochondrial superoxide dismutase and released from the mitochondrial matrix to cytoplasm. Starvation induced H_2O_2 may also activate AMPK, which may be redox sensitive (Shao *et al.*, 2014).

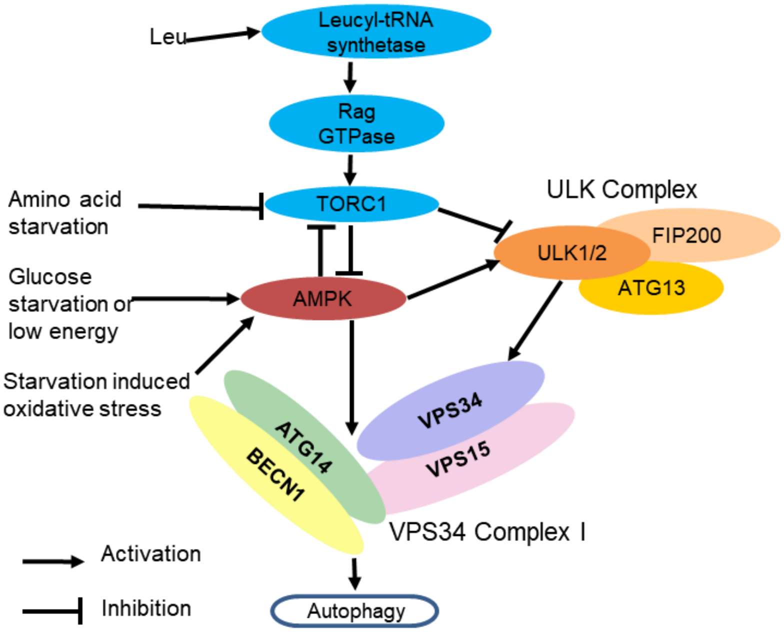


Figure 1.2: Upstream signals regulating starvation induced autophagy

Starvation induces autophagy via a protein signaling complex that includes UNC-51 like autophagy-activating Ser/Thr kinase (ULK)1 or ATG1 signaling complex composed of ULK1/2 kinase, ATG13, and focal adhesion kinase family interacting protein of 200 kD (Mizushima, 2010). Under nutrient rich conditions the mammalian target of rapamycin (mTOR) complex I (mTORC1) inhibits autophagy by ULK1 phosphorylation, whereas under starvation condition, mTORC1 dissociates from the ULK1 complex, which can then trigger autophagosome nucleation and elongation (Figure 1.2) (Sinha *et al.*, 2019).

The protein complex responsible for vesicle nucleation, the vacuolar protein sorting protein 34 (VPS34) Complex I, comprises of four core components: VPS34/PI3KC3, VPS15/P150, VPS30/BECN1/ATG6 and ATG14 (Itakura & Mizushima, 2009). This complex is activated and recruited to the site of phagophore nucleation, where PI3KC3 catalyzes phosphorylation of phosphatidylinositol to generate the important signaling molecule, phosphatidylinositol-3-phosphate. The other components of the complex regulate the activity of PI3KC3. Apart from the previously mentioned four core components of VPS34 Complex I, there is also a fifth subunit, ATG38/Nuclear Receptor Binding Factor 2 (NRBF2) which associates with the active VPS34/PI3KC3 Complex I and makes it 10-fold more active (Young *et al.*, 2016, Young *et al.*, 2019).

The phagophores are elongated to form autophagosomes. The proteins involved in regulating this phase are classified into two ubiquitin-like conjugation systems in which one protein activates the other protein through a cascade of events. The first is the ATG12 conjugation system (Figure 1.3). In this system, first, Atg12 protein is activated by Atg7, an E1-like enzyme. Next, activated Atg12 is transferred to Atg10, an E2-like conjugating enzyme. Finally, Atg12 is covalently conjugated to Atg5, wherein the Atg12-Atg5 then forms a larger conjugation system with Atg16. The other ubiquitin-like system is the ATG8 or LC3 conjugation system (Figure 1.4). In this system, autophagosome formation results in processing of microtubule-associated protein light chain 3 (LC3B). LC3B, expressed in most cell types as a full-length cytosolic protein, is proteolytically cleaved by Atg4 (a protease) to expose the C-terminal glycine to generate LC3B-I. LC3B-I is activated by Atg7, an E1-like enzyme, followed by transfer to the Atg3, an E2-like enzyme. Lastly, Atg8 forms a conjugate with phosphatidylethanolamine to generate LC3B-II.

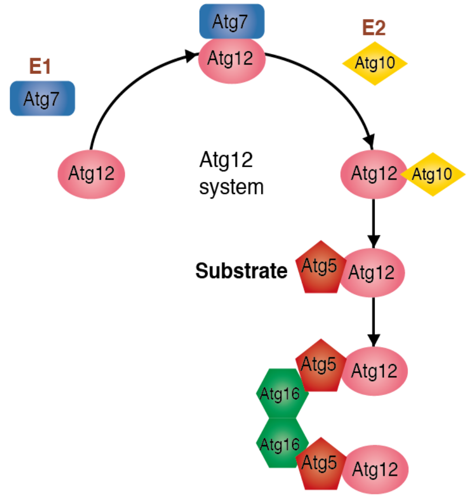


Figure 1.3: Atg12-Atg5-Atg16 oligomer formation (https://www.mblbio.com/bio/g/product/autophagy/pickup/atg8.html)

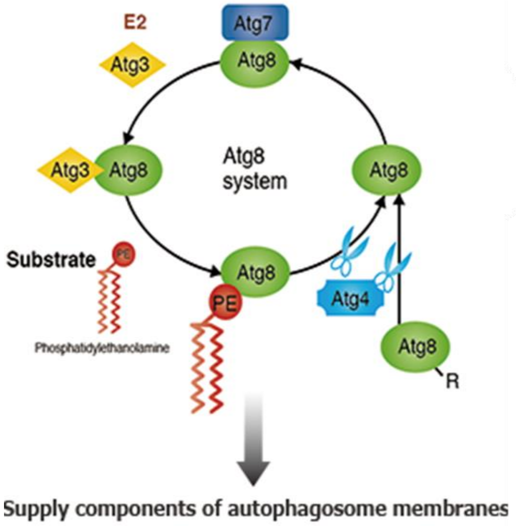


Figure 1.4: Atg8 or LC3 conjugation system (https://www.mblintl.com/products/lc3-at8/)

The last stage of autophagy, autophagosome maturation, results in the autophagosome membrane fusing with the lysosome membrane to form the autolysosome. The sequestered contents are processed further and degraded by lysosomal hydrolases, and finally released into cytosol for recycling. In yeast, this fusion process involves a family of SNARE proteins (Darsow *et al.*, 1997, Sato *et al.*, 1998, Mayer & Wickner, 1997, Burri & Lithgow, 2004). In mammalian cells, monomeric GTPases such as Rab22 and Rab24 are also required for autophagosome maturation. VPS34 Complex II, which has a composition very similar to the VPS34 Complex I, except that ATG14 is replaced by VPS38/UVRAG, also plays a role in autophagosome maturation (Liang *et al.*, 2006). VPS38 is also responsible for vacuolar protein sorting.

VPS34 Complex I and II have a very similar V-shaped structure (Baskaran *et al.*, 2014, Rostislavleva *et al.*, 2015a) (Figure 1.5). One arm of the V-shape is formed by VPS34/PI3KC3 and VPS15/p150 intertwining in an antiparallel fashion. VPS15 consists of a kinase domain, a helical (KINHEAT) domain and a WD40 domain. VPS34 consists of a C2 domain, a helical domain and a kinase domain. The VPS15 kinase domain interacts with the VPS34 activation loop, regulating its activity(Rostislavleva *et al.*, 2015a). Recently it was found that autophagy induction due to starvation or stress involves the activation of the VPS34 kinase domain in VPS34 Complex I by NRBF2, due to removal of the inhibitory contact of VPS15 and thus acts as a positive regulator of VPS34 Complex I (Figure1.6). NRBF2 has an N-terminal Microtubule Interacting and Trafficking domain (MIT) domain and a C-terminal coiled coil domain. One of the two MIT domains in the NRBF2 homodimer (Young *et al.*, 2016) bends the helical solenoid of the VPS15 scaffold, thereby displacing the VPS15 kinase domain of VPS15 away from the VPS34 activation loop (Young *et al.*, 2019), while the other NRBF2 MIT domain stabilizes the VPS34 lipid kinase domain in an active conformation with an unrestricted activation loop facilitating access to the membranes (Young *et al.*, 2019).

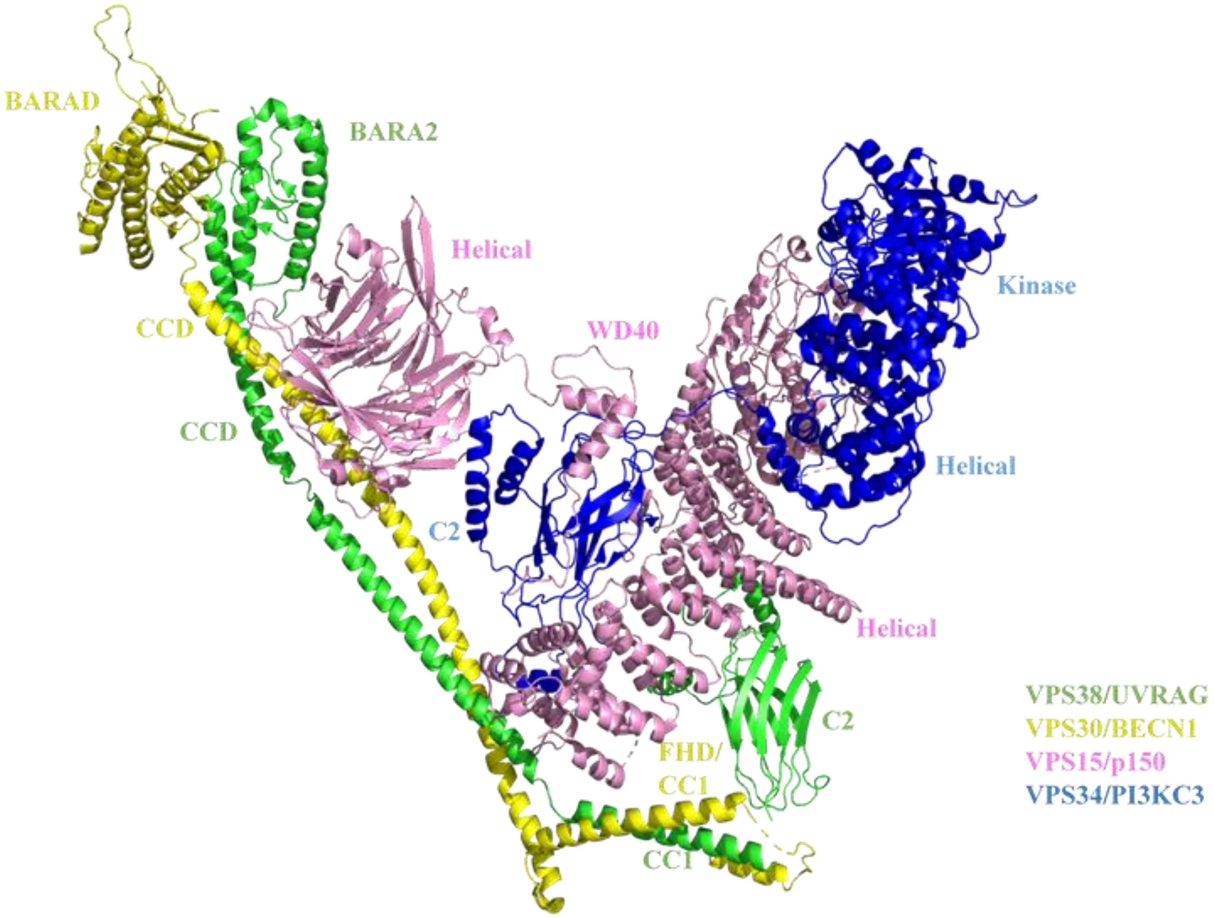


Figure 1.5: Crystal structure of VPS34 Complex II (PDB ID: 5DFZ)

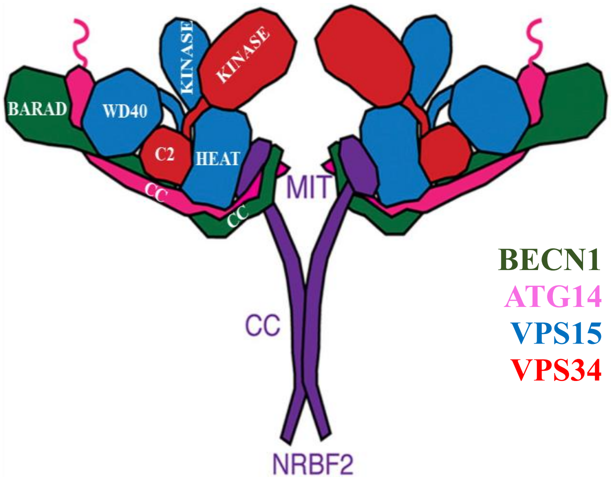


Figure 1.6: Dimeric PI3KC3-C1 containing NRBF2. NRBF2 is localized to the base of PI3KC3-C1 and induces dimerization through its central coiled-coil domain (Young *et al.*, 2016).

1.2. BECN1 in autophagy

BECN1 is a key autophagy protein which is a core component of two mutually exclusive, quaternary class III phosphatidylinositol 3-kinase (PI3KC3) complexes, Complex I and Complex II, which mediate autophagosome nucleation and maturation respectively (Itakura & Mizushima, 2009, Matsunaga *et al.*, 2009). Complex I and Complex II comprise of PI3KC3, p150 (a Ser/Thr kinase regulatory protein), BECN1 and either ATG14 or UVRAG, respectively. BECN1, which is highly conserved across diverse eukaryotes, is a conformationally flexible, multi-domain protein (Mei, Glover, *et al.*, 2016b, Mei, Su, *et al.*, 2016b) (Figure 1.7). Human BECN1 consists of 450 residues and has at least four distinct structural domains (i) a large, poorly conserved, N-terminal intrinsically disordered region (IDR) comprising residues 1-140, which also consists of a BCL2 homology 3 domain (BH3D) comprising residues 105-130; (ii) a flexible helical domain (FHD) comprising residues 141-171 which is disordered at its N-terminus and helical at its C-terminus but becomes completely helical upon binding to appropriate partners; (iii) a coiled-coil domain (CCD) comprising residues 175-265 that forms an anti-parallel homodimer in the absence of other interacting partners (Mei, Glover, *et al.*, 2016b), with the CCD N-terminal region comprising residues 180-190 constituting a conserved Nuclear Export Signal sequence (NES) and the CCD C-terminal region comprising residues 248-265 constituting a conserved conformationally-dynamic region named the Overlap Helix; and (iv) a β-α repeated autophagy specific domain (BARAD) comprising residues 266-450 that is involved in membrane binding (Rostislavleva *et al.*, 2015b).

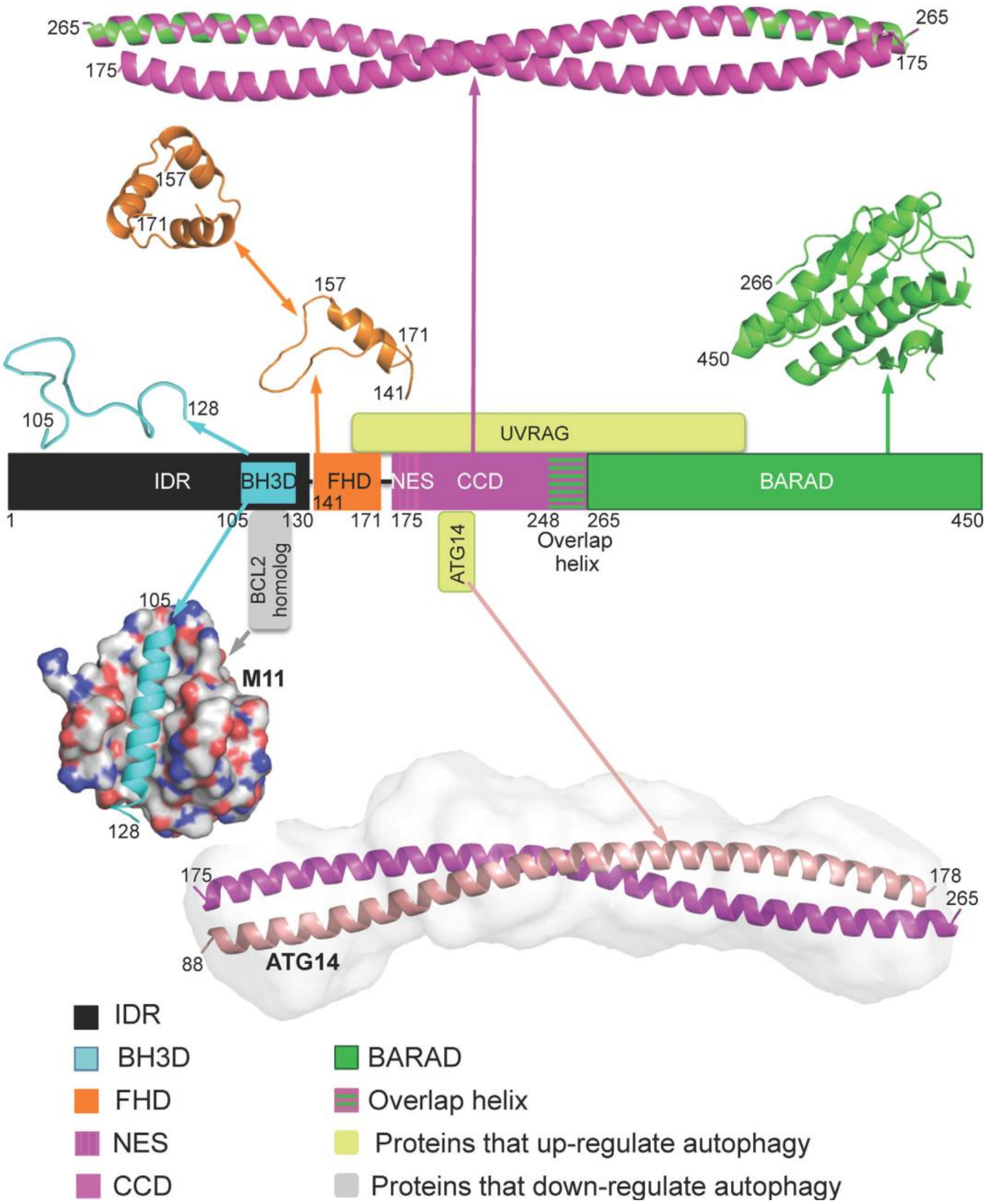


Figure 1.7: Domain architecture of BECN1 and selected interactions. Domains structures are displayed in ribbon above the schematic. Two structurally characterized interactions are displayed below the schematic: (i) The BECN1 BH3D (cyan ribbon) bound to M11 (grey molecular surface) and (ii) The BECN1 CCD (magenta ribbon):ATG14 CCD (salmon ribbon) heterodimer is shown modelled into its SAXS-derived molecular envelope. Yellow-green boxes represent interacting proteins that up-regulate autophagy while grey boxes represent interacting proteins that down-regulate autophagy (Mei, Glover, *et al.*, 2016a). The PDB IDs for FHD, CCD, BARAD, and M11:BECN1 complex structures are: 5EFM, 5HHE, 4DDP, and 3DVU).

1.3. Intrinsically Disordered Regions in autophagy

The flexibility in protein conformation may arise due to presence of intrinsically disordered regions (IDRs) in proteins. IDRs lack the ability to form stable ordered secondary structure like helices and β-strands, and therefore do not form stable, ordered tertiary structures. Approximately 35-50% of eukaryotic proteins are predicted to bear IDRs (van der Lee *et al.*, 2014), indicative of an important role in eukaryotes (van der Lee *et al.*, 2014). IDRs generally have lower sequence complexity and a higher fraction of polar or charged residues rather than hydrophobic residues, which prevents formation of a well-packed hydrophobic core (van der Lee *et al.*, 2014). IDRs also tend to have low sequence conservation (van der Lee *et al.*, 2014). Despite their overall low sequence conservation, IDRs may contain evolutionarily invariant sequence motifs that may be important for the structural and functional regulation of the protein. The sequence and structural flexibility of IDR is thought to enable diverse and multiple interactions with high specificity and reversibility enabling them to regulate cell signaling pathways (Diella *et al.*, 2008, Jones & Thornton, 1996, Wright & Dyson, 1999, Dunker & Obradovic, 2001, Dyson & Wright, 2005). A previous bioinformatics study shows that the orthologs have IDRs in positions similar to those predicted in IDRs in human autophagy proteins, but these are poorly conserved, indicating that these IDR regions may have diverse functions in autophagy (Mei *et al.*, 2013). The ELM server(Kumar *et al.*, 2020) was used to search for Eukaryotic Linear Motifs (ELMs) which are short, evolutionarily plastic, linear sequence motifs experimentally shown to be key for various protein-protein interactions. The program ANCHOR(Mészáros *et al.*, 2009) was used to predict protein sequences flanking or overlapping the IDRs that would be stabilized as secondary structures upon binding to a globular protein partner. In ULK1 and AMBRA1, ELMs were predicted within the Anchors, which suggests that those ELMs might nucleate interactions to stabilize secondary structure. Additionally, in the same study it was also shown that the BH3D (residues 105-130) of the key autophagy inducing protein BECN1 is completely unstructured in the absence of other interactions and undergoes binding-induced helical transition nucleated by an anchor. Therefore, IDRs play an extremely important role in regulating autophagy, either by protein-protein interaction or by binding-associated helical transition (Mei *et al.*, 2013).

1.4. Zinc binding in autophagy

Amongst biologically relevant transition metals, zinc (Zn^{2+}) is the second most important metal found within the cells after iron (Porcheron et al., 2013). Bioinformatics analysis of the human genome suggests that up to 3000 proteins, which corresponds to ~10% of all encoded proteins, bind Zn^{2+} (Andreini et al., 2006). A similar fraction of human protein structures deposited in structural databases bind Zn^{2+} (Kochanczyk et al., 2015). The most abundant class of zinc-binding proteins in humans are proteins containing zinc-fingers, with Cys_4 and Cys_2His_2 being the most common types of coordination environment (Andreini et al., 2006). Zn^{2+}-cysteine complexes are critical mediators of protein structure, catalysts and regulation (Pace & Weerapana, 2014). It has been shown that Zn^{2+} binding is critical for autophagy(Hwang *et al.*, 2010, Liuzzi *et al.*, 2014). When the MCF7 cells are treated with tamoxifen, autophagy is induced and results in the death of these cells(Hwang *et al.*, 2010). Levels of LC3B or LC3-II also increase, and GFP-LC3 accumulates around autophagosomes in cells that gets exposed to tamoxifen, indicating that autophagy is involved in tamoxifen-induced changes. Live cell confocal microscopy assay which uses a fluorescent Zn^{2+} binding dye, FluoZin-3, shows that labile Zn^{2+} ion accumulates in most acidic LC3(+) autophagic vesicles. When Zn^{2+} ions are chelated using N,N,N’,N’-tetrakis (2-pyridylmethyl) ethylenediamine (TPEN), the increase in

12hosphor-ERK and LC3-II levels is blocked, and autophagic vesicle formation and cell death are attenuated. Conversely, treatment of MCF7 cells with $ZnCl_2$, markedly increases tamoxifen-induced Extracellular-Signal Regulated Kinases (ERK) activation, autophagy, and cell death, indicating the importance of Zn^{2+} in these events (Hwang et al., 2010, Zheng et al., 2007). ERK ½ has been shown to regulate both basal and starvation-induced autophagy, via two different pathways. It can activate the Beclin1-PI3K complex by phosphorylating Bcl-2 proteins which down-regulate BECN1-mediated autophagy (Botti et al., 2006), or activate autophagy by promoting disassembly of mTORC1 complex through a non-classical pathway. Zn^{2+} starvation has been implicated in inactivating TORC1, thereby upregulating autophagy (Kawamata *et al.*, 2017). However, TORC1 has no Zn^{2+}-binding site, so this regulation may be indirect. On the other hand, Zn^{2+} downregulates tamoxifen-,H_2O_2-, ethanol-, and dopamine-mediated autophagy (Hung *et al.*, 2013, Lee & Koh, 2010, Hwang *et al.*, 2010, Liuzzi & Yoo, 2013)

1.5. Nuclear Export Signal and cellular homeostasis

Eukaryotic cells have a double membraned nuclear envelope that inhibits free flux of the macromolecules between cytoplasm and nucleus. There are approximately 3000 nuclear pores in the mammalian nuclear envelope which contain embedded Nuclear Pore Complexes (NPCs) (Figure 1.8). The morphological structure of the NPC evaluated previously by EM (Callan & Tomlin, 1950, Gall, 1967) and most recently by Cryo-EM (Akey & Radermacher, 1993, Beck *et al.*, 2004, Stoffler *et al.*, 1999, Yang *et al.*, 1998) reveals that the overall structure of the NPC has an eight fold rotational symmetry, containing three ring-like structures (cytoplasmic ring, inner ring, nucleocytoplasmic ring) which are conserved from yeast to mammals (Miyamoto *et al.*, 2018, Maul, 1971). The NPC acts as a molecular sieve limiting the free passage of macromolecules across the nuclear envelope, while allowing passive diffusion of small molecules such as ions, metabolites, and small proteins of less than ~ 40kDa. Large proteins inside the nucleus bind via a special leucine-rich sequence to nuclear transport factors (called Exportins), in the presence of a small GTPase called Ran which exists in GTP-bound form (RanGTP) inside the nucleus and GDP-bound form in the cytoplasm. This asymmetric distribution of Ran depends on the presence of two regulatory proteins that allows cycling of Ran between the GTP/GDP bound states (Figure 1.9). The Regulator of chromosome condensation 1 (RCC1) protein, which is chiefly associated with nucleosomes through interactions with histones H2A and H2B (Nemergut & Macara, 2000), triggers the conversion of RanGDP to RanGTP, resulting in predominance of RanGTP inside the nucleus. Conversely, Ran GTPase-activating protein 1 (RanGAP1) is localized in the cytoplasm, which stimulates GTPase activity along with the RanGTP-binding protein 1 (RanBP1), which causes GTP hydrolysis, thereby leading to reduced cytoplasmic concentration of RanGTP(Miyamoto *et al.*, 2018). Therefore, this highly compartmentalized distribution of Ran regulatory proteins allows for directional transport of cargo proteins between the nucleus and cytoplasm.

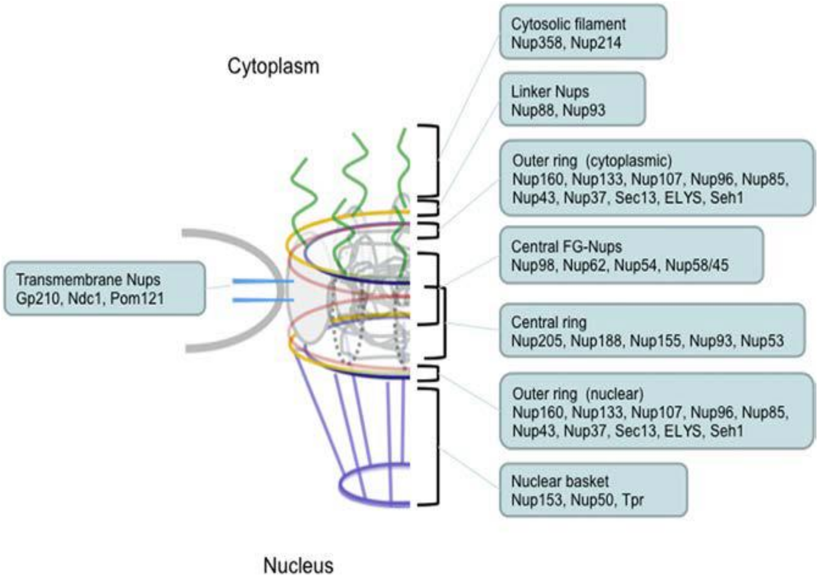


Figure 1.8: Nuclear Porin Complex (NPC) and nucleoporins (Nups) (Miyamoto *et al.*, 2018)

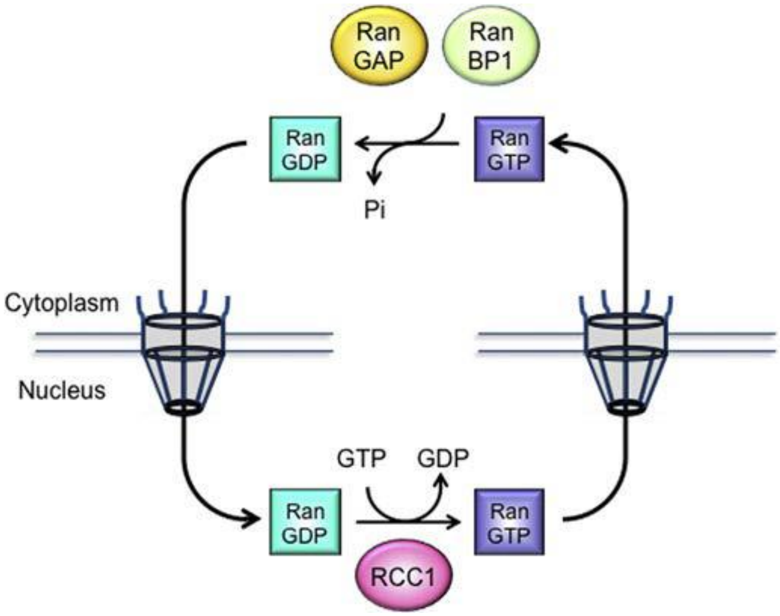


Figure 1.9: Ran recycling between nucleus and cytoplasm(Miyamoto *et al.*, 2018).

1.6. Overview of methods used in this research

1.6.1. Cellular autophagy assays

Autophagy levels are evaluated by monitoring the formation of autophagosomes which can be quantified by the levels of an autophagy protein, LC3, that localize to autophagosomes (Figure 1.9)(Mizushima *et al.*, 2010). Among the 4 different LC3 isoforms (Mizushima *et al.*, 2010), LC3B is most widely accepted as a marker for this fluorescence-based assay. Soon after translation, Atg4 processes the nascent LC3 at its C terminus to form LC3-I, which has a glycines at its C-terminus and is dispersed in the cytoplasm (Mizushima *et al.*, 2010). LC3-I is subsequently conjugated with phosphatidylethanolamine (PE) to form LC3-II (or LC3-PE) (Mizushima *et al.*, 2010) by a ubiquitination-like enzymatic reaction, resulting in association with both the outer and inner membrane of the autophagosome (Figure 1.10). After fusion with the lysosome, LC3 on the outer membrane is cleaved off by Atg4 and LC3 on the inner membrane is degraded by the lysosomal enzymes, resulting in very low LC3 content in the autolysosome. Thus, endogenous LC3 or GFP-LC3 is visualized by fluorescence microscopy either as a diffuse cytoplasmic pool or as punctate structures that primarily represent autophagosomes (Figure 1.11).

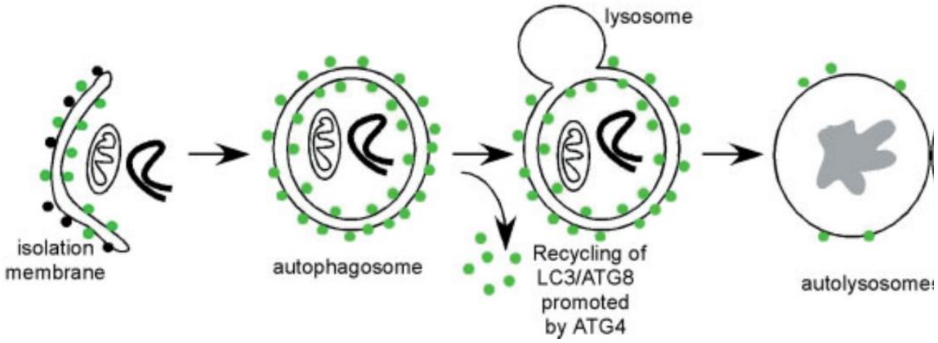


Figure 1.10: The GFP-LC3 (or endogenous LC3) puncta formation assay counts the average number of punctate structures per cell by fluorescence microscopy (Glick *et al.*, 2010).

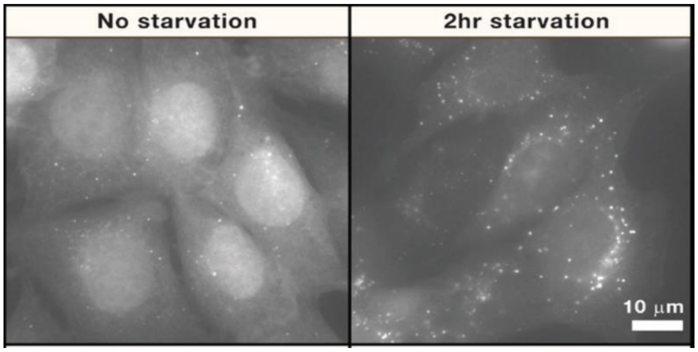


Figure 1.11: Mammalian cells in culture stably expressing GFP-LC3 with (right) or without (left) 2 hours of starvation (depletion of both amino acids and serum). There is not only an increase in GFP-LC3 puncta number, but also a decrease in total GFP-LC3 fluorescent signals during the 2 hours incubation period (Mizushima *et al.*, 2010).

1.6.2. Affinity pull down assays

Affinity pull down assays are qualitative protein-protein interaction assays. When performed using purified proteins, they can be used to confirm direct protein-protein interactions. In this “bait and prey” assay one protein is expressed with a specific affinity tag and allowed to bind its cognate affinity resin. Excess unbound protein is washed off forming the “bait.” (Figure 1.12) A potential binding partner, the prey which is expressed without the affinity tag used for the first protein, is allowed to interact with the affinity resin-bound protein bound to it, followed by several column volume washes to remove any unbound protein. Lastly, bound protein is eluted from the affinity column, and wash and elution fractions analyzed by SDS PAGE. The presence of bands corresponding to both proteins in the elution fractions is indicative of a direct protein-protein interaction, enabling retention of the potential binding partner on the resin. Reverse pull-down assays can be performed if the potential binding partner were also expressed with an affinity tag, different from the first protein.

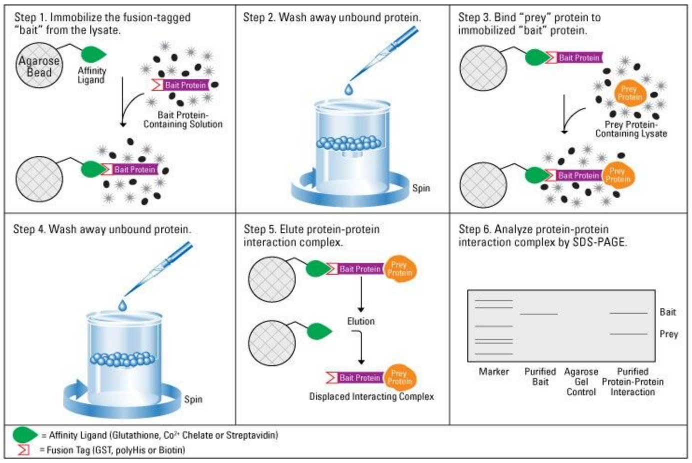


Figure 1.12: General set up of an affinity pulldown experiment (https://www.thermofisher.com/us/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/pull-down-assays.html)

1.6.3. Inductive Coupled Mass Spectrometry (ICP-MS)

ICP-MS is a sensitive technique for detecting metal ions in samples. It is capable of detecting almost all elements in the periodic table and can be applied to solutions, solids, and gases. ICP-MS analysis of proteins requires treating the protein with 33% metal grade HNO_3, to denature, digest and precipitate proteins from the solution, and release any metal ions bound to the protein. The precipitate is pelleted out by centrifugation at high speeds and the supernatant analyzed by ICP-MS. A typical configuration of the ICP-MS (Figure 1.13) includes a nebulizer in which solutions are vaporized and are converted to gases by laser ablation. These gases are then introduced into the argon plasma, which consists of electrons and positively charged Argon ions (Ar^+), at temperature of 7000-10000 K. In the Argon plasma the sample gas is ionized to individual atoms that lose electrons and produce cations. Since Ar^+ is the predominant ion present, the elements get ionized and results in equilibration of the sample ions with Ar^+. Argon has a very high ionization potential (IP), and the advantage of using Ar^+ plasma is that almost every element, except for the He, F, Ne elements, has a lower IP, so through electron transfer and collision reactions all the elements are detected as respective monovalent cations. So, even though an element may be very stable in a multivalent state under normal pressure or temperature, under high temperature and pressure all the elements are identified only as monovalent cations. The cation beam generated then passes through the quadrupole mass analyzer where the ions are separated according to their mass/charge (m/z) ratio (Singh, 2016). All the ions with a particular m/z arrive at the mass detector within a span of 1-10 ns and ions with the same m/z ratio are then counted by the ion counting detectors (Beauchemin, 2017). The concentration of metal ions in samples is determined by correlation of metal isotope counts detected by the instrument to known concentrations in an ICP-MS calibration solution. Thus, ICP-MS analysis can be used to determine the stoichiometry of metal to protein binding.

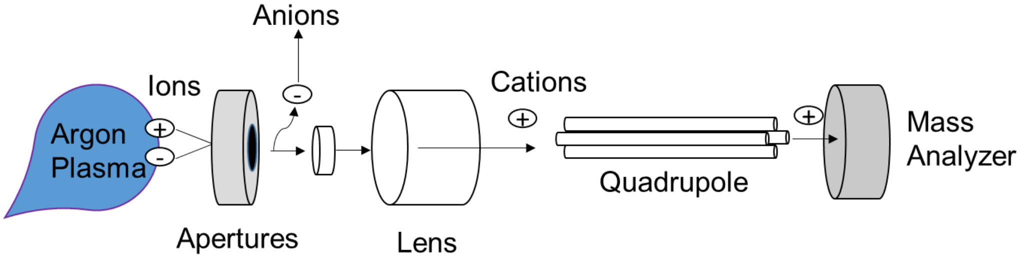


Figure 1.13: Components of inductivity coupled plasma (ICP) mass spectrometer (Singh, 2016)

1.6.4. Circular Dichroism Spectroscopy

Circular dichroism (CD) spectroscopy is a method that measures the difference in the absorption of left-handed circular polarized light and right-handed circular polarized light by a sample. Chiral molecules such as proteins absorb left- and right-handed circularly polarized light differently. A polypeptide chain in protein is primarily made up of secondary amides. Polypeptides form different secondary structures dictated by different backbone conformations that have different arrangements of secondary amides. At far UV wavelengths (i.e. below 250 nm), the geometric relation between the amide groups of the protein backbone determines the characteristic shape of CD spectra. Therefore, each of the different protein secondary structures have a characteristic molar ellipticity across different far UV wavelengths giving rise to CD spectra. A-helices have two negative peaks at 208nm and 222 nm and a positive peak at 193 nm; β-strands have a positive peak at 195 nm and a negative peak at 218 nm; and unfolded proteins or coils have a negative peak at 195 nm and very low ellipticity at wavelengths beyond 210 nm.

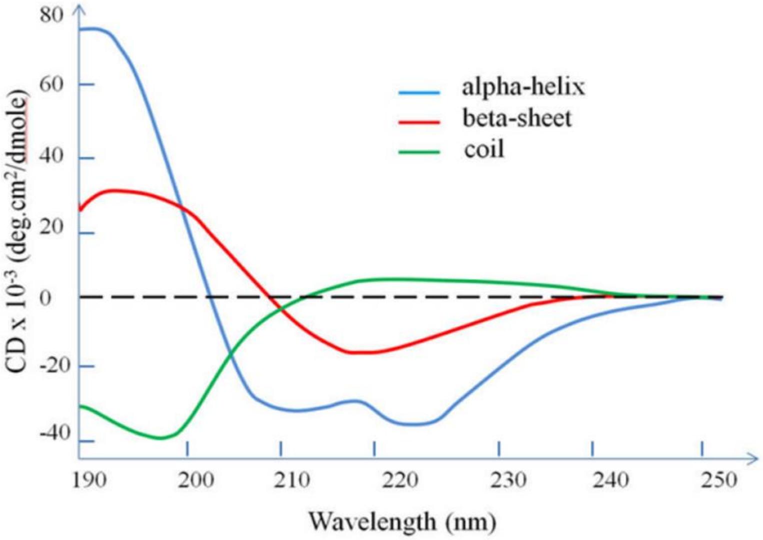


Figure 1.14: CD spectra of polypeptides with representative secondary structures. (Wei *et al.*, 2014)

The CD spectra is typically displayed in molar ellipticity in degrees.cm^2.dmol^{-1} (on y-axis) versus wavelength in nm (on x-axis), where degree of ellipticity is defined as the tangent of the ratio of the minor to major elliptical axis and mean residue molar ellipticity is the molar ellipticity divided by the total no. of residues in a protein (Figure 1.14). The equation $[\theta] = \theta$ x

10^6/ (C x L), is used to calculate the mean residue molar ellipticity where C is the µM concentration of the protein multiplied by the no. of residues and L is path length in mm. The secondary structure content of a protein can be estimated from CD spectra by software such as CD Pro that includes four different programs SELCON3, CDSSTR, CONTIN, CLUSTAL and seven reference sets (Sreerama & Woody, 2000) or using the online Dichroweb software (Whitmore & Wallace, 2008, Lobley *et al.*, 2002, Whitmore & Wallace, 2004). The reference sets used to train the software contain CD spectra of proteins with known secondary structures.

CD can also be used to study secondary-structure transitions in proteins. Compounds such as 2,2,2-trifluroethanol (TFE) can be used to approximate hydrophobic binding interfaces to monitor changes in secondary structure, especially helicity (Figure 1.15). (Luo & Baldwin, 1997, Culik *et al.*, 2014, Glover *et al.*, 2016).

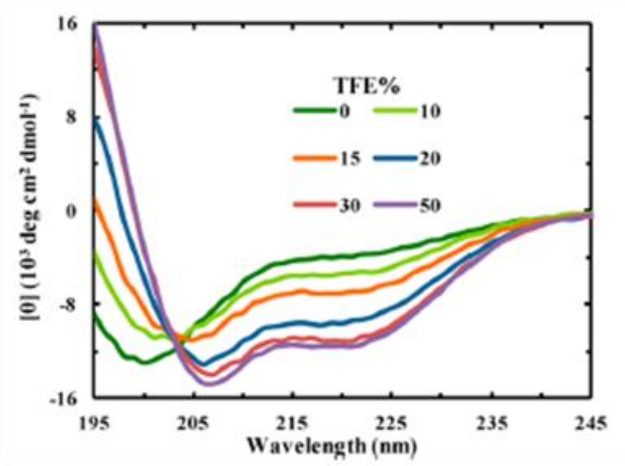


Figure 1.15: Effect of TFE on a natively unstructured protein. (Culik et al., 2014)

1.6.5. Small Angle X-ray Scattering (SAXS)

SAXS can be used to obtain low resolution structural information for proteins in solution, thereby providing useful structural information for flexible proteins that cannot be crystallized.

In a typical SAXS experiment the protein sample is exposed to the X-ray source and the elastic scattering of X-ray beams recorded at very low angles, between 0.1° to 10°. Scattering from the buffer is subtracted from the scattering curve of the protein solution, and the resultant scattering intensity (I) can be represented as a function of the magnitude of the scattering vector, q, where

$$q = \frac{4\pi \sin\theta}{\lambda}$$

2θ is the scattering angle and λ is the wavelength of the incident X-ray beam (Putnam *et al.*, 2007). Structural information about the protein can then be obtained by analysis of the SAXS data. Different parameters are used to verify the quality of the samples used for the data collection and also for obtaining the size and shape information.

A Guinier Plot (Figure 1.16) is a plot of Ln(I) vs q^2. A linear Guinier plot at low q regions, i.e. where $q \times R_g \leq 1.3$, indicates that the sample is a monodispersed in solution. A flat Guinier fit residual indicates the protein is monodispersed in solution, whereas if the fit has a shape of a frown or smile, then that indicates interparticle repulsion and aggregation, respectively (Hopkins *et al.*, 2017). The linear Guinier plot at the low q region enables estimation of the radius of gyration (R_g), scattering intensity at zero angle (I(0)) of diffraction. The slope of the linear fit equals $-R_g^2/3$, and the X-axis intercept equals $e^{(I(0))}$. R_g is the average root-mean square distance from the center of the density of the molecule weighted by the scattering length of the density of the molecule. Despite equivalent molecular mass, a protein with a compact shape such as globular proteins will have a smaller R_g than a protein that is extended or disordered (Kikhney & Svergun, 2013).

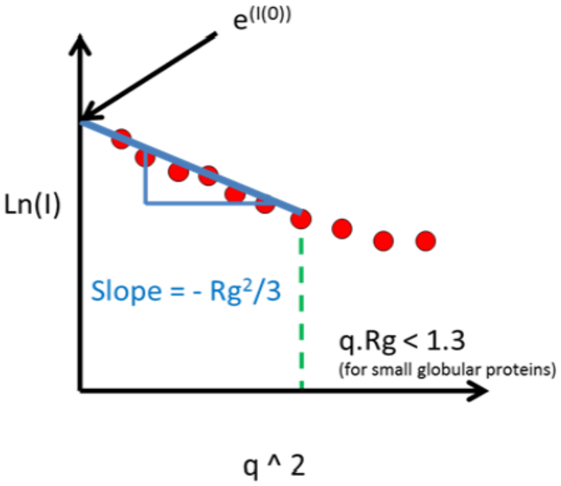


Figure 1.16: A sample Guinier Plot

The P*(r)* distribution plot (Figure 1.17) is a histogram of distances of all possible pairs of distances between atoms in a particle. The scattering pattern I(q) of a particle is related to the Fourier transform of its P*(r)* function by the equation:

$$P(r) = \frac{r}{2\pi^2}\int_0^{\infty} I(q)sin\,(qr)dq$$

where r is all possible intra-particle distances, and I(q) is the intensity as a function of the magnitude of the scattering vector q. D_{max} or maximum particle dimension is the radius of the protein in the longest dimension, which is the value of r when P(r) equals zero (Putnam *et al.*, 2007). The shape of the P*(r)* distribution plot provides information about the shape of the protein. A symmetric bell-shaped curve indicates that the protein is globular and well-folded, whereas an asymmetrical curve with elongated tail indicates the protein is a long rod-shaped protein or a disordered protein (Kikhney & Svergun, 2013).

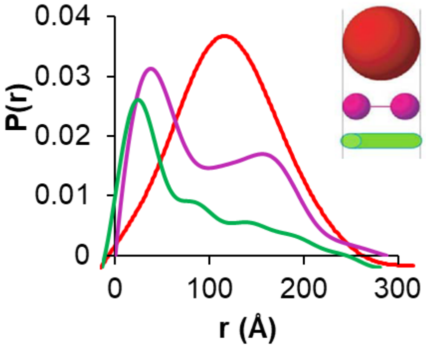


Figure 1.17: A sample P(r) distribution Plot

The dimensionless Kratky plot, which is a plot of I(q) × q^2 vs. q, is a measure of folding state of proteins. The Kratky plot for a well-folded protein is bell-shaped, with the curve converging back to the q axis at the high q region. Conversely, if the protein is unfolded or disordered then the Kratky plot does not converge back to the q axis at the high q region (Figure1.18) (Kikhney & Svergun, 2013).

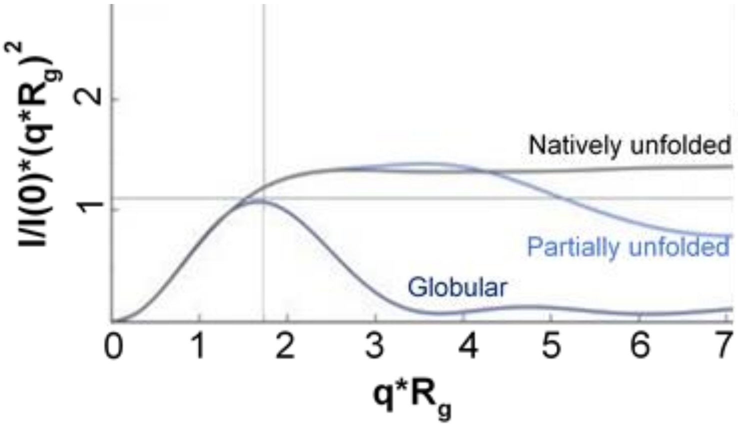


Figure 1.18: A sample Dimensionless Kratky Plot (Durand et al., 2010)

Ensemble of Optimization Method (EOM) is a method in which a genetic algorithm is used to generate a large pool of all possible conformation based on the given protein sequence and the structural information about the polypeptide chain (Bernado *et al.*, 2007). A genetic algorithm repeats several times to select an ensemble of conformers from the pool that best fit the experimental SAXS data (Tria *et al.*, 2015). The discrepancy between the theoretical scattering calculated for each conformation and the experimental scattering data is assessed by a χ^2 test to evaluate the quality of fit for each conformer. Thus, EOM allows disordered regions in a protein or a disordered protein to be feasibly modelled. The average R_g and $D_{\max}$ distributions of the pool, as well as the selected conformer ensemble, are determined and compared (Figure 1.19).

The conformational flexibility of the selected conformer ensemble relative to the pool is further described by two metrics R_{flex} and R_σ (Tria *et al.*, 2015). R_{flex} is a metric used to evaluate the flexibility of the distribution of conformers, and R_σ is a metric to evaluate the variance of the distributions of conformers (Tria *et al.*, 2015).

$R_{\text{flex}} = -H_b(S)$, where H_b is the entropy as a function of probability density functions

$S = (X,P)$, where $P = (p_1,\ldots,p_n)$ is the probability ascribed to the interval $X = (x_1,\ldots,x_n)$ such that the characteristics of the selected ensemble are compared to those displayed by the pool, which allows us to assess the flexibility of the system. The value of $Hb(S)$ can range from (-1 to 0), where a value of -1 indicates extreme (theoretical) flexibility, which is expected to be close to the $Hb(S)$ calculated for the pool and a value of 0 represents extreme (theoretical) rigidity. So the value of R_{flex} could be between 0 % to 100% (Figure 1.20).

$R_\sigma = \sigma_S / \sigma_P$, where σ_S and σ_P are the standard deviations of the R_g distributions of the selected ensemble and of the pool respectively.

For a completely rigid system R_{flex} is 0 and for completely flexible system R_{flex} is to 100%. R_σ approaches 1 for fully flexible systems and decreases with decreasing flexibility. When R_{flex} of the selected conformer ensemble is smaller than that of the randomly generated pool distribution, R_σ is usually less than 1 (Tria *et al.*, 2015). Conversely, when R_{flex} of the selected conformer ensemble tends to value greater than the pool, R_σ is greater than 1. For cases where R_{flex} of selected conformer ensemble is significantly smaller than pool, but $R_\sigma > 1$, further evaluation of the data is required.(Tria *et al.*, 2015) If the frequency distributions shows a narrower frequency distribution of R_g and $D_{\max}$ than that of the randomly generated pool, this indicates that conformers in the experimentally selected ensemble are more compact as those in the larger pool (Figure 1.19 A,B). But if the frequency distributions of the selected conformer ensemble and the pool match well, it indicates that conformers in the experimentally selected conformer ensemble are as flexible as those in the larger pool generated computationally (Figure 1.19 D,E).

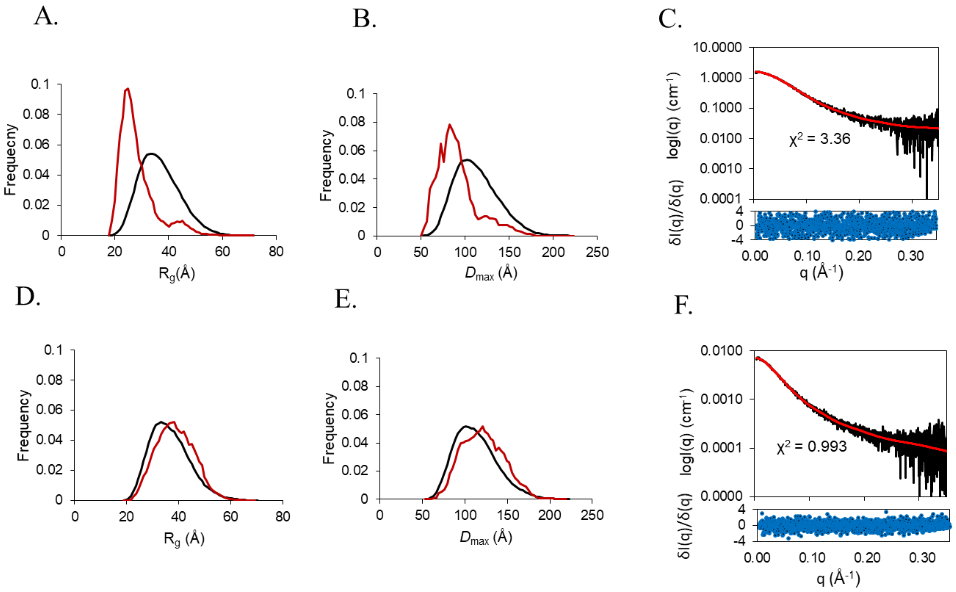


Figure 1.19: A sample EOM generated Rg and Dmax distribution plot. A, D) R_g distribution for the EOM-generated conformer pool (black) and the selected conformer ensemble (red). B,E) D_{max} distribution for the pool (black) and selected conformer ensemble (red). C,F) Top plot: Experimental scattering profile (black) with the fit of the selected conformer ensemble (red); Bottom plot: the normalized fit residual (blue).

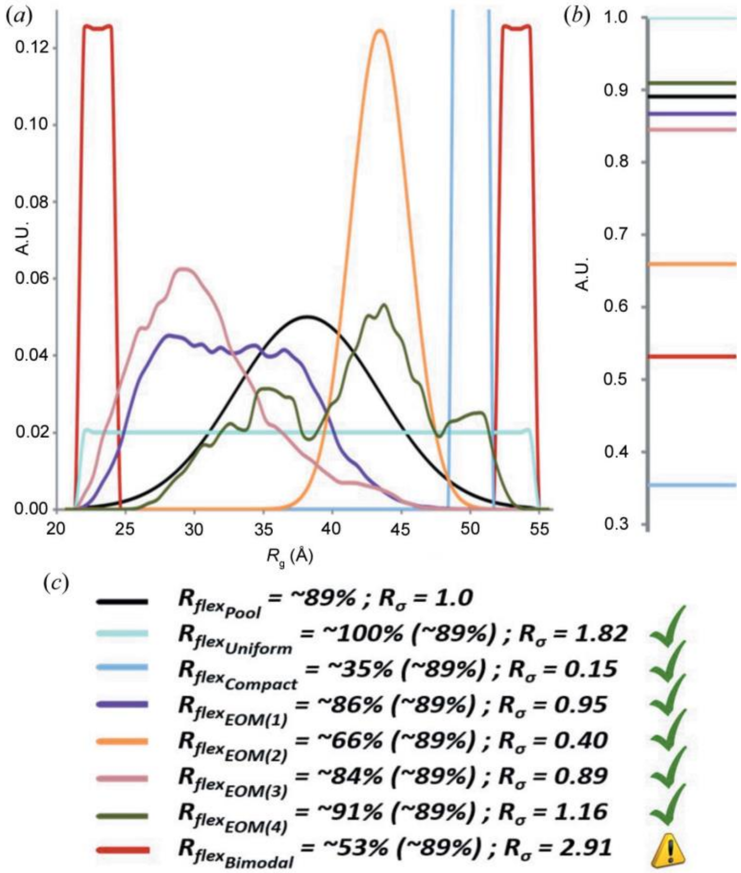


Figure 1.20: Qualitative characterization of particle flexibility from various characteristic Rg distributions. (a) Pool (black), which represents randomly generated conformers; in terms of R_g distributions; uniform (cyan), compact (light blue), bimodal (red) (b) *Hb(S)* values computed from the distributions in (a). (c) Combination of R_{flex} values for all the distributions (and compared to the threshold of randomness computed from the pool, in brackets, 89%) with the associated R_σ values. The last example (red curve) indicates a potentially inconsistent result. (Tria *et al.*, 2015)

Singular Value Decomposition (SVD) analysis enables the SAXS scattering intensity to be used for identification of species that are structurally significantly different (Figure 1.21 A), yet cannot be separated by SEC. These significant species may indicate different conformations or l different proteins in a protein complex or different oligomeric states of the same protein. A species is considered significant if it has an autocorrelation value roughly > 0.6 to 0.7. Once significant species or components are identified, Evolving Factor Analysis (EFA) in the BioXTAS RAW software (Hopkins *et al.*, 2017) is used to generate individual scattering curves which can then be used to separately analyze SAXS parameters such as the R_g, $D_{\max}$, P(r) distribution and Kratky plot, as a result of removal of influence of one component to the other. Reliability of the EFA analysis (Figure 1.21 B,C) depends on the Chi-square versus the scattering intensity frame being ~ 1.2 to 1.3, and concentration versus scattering intensity frame showing unique curves for each component. Both Chi-square and the concentration plot can be manipulated using user defined intensity frames to obtain a reliable EFA analysis.

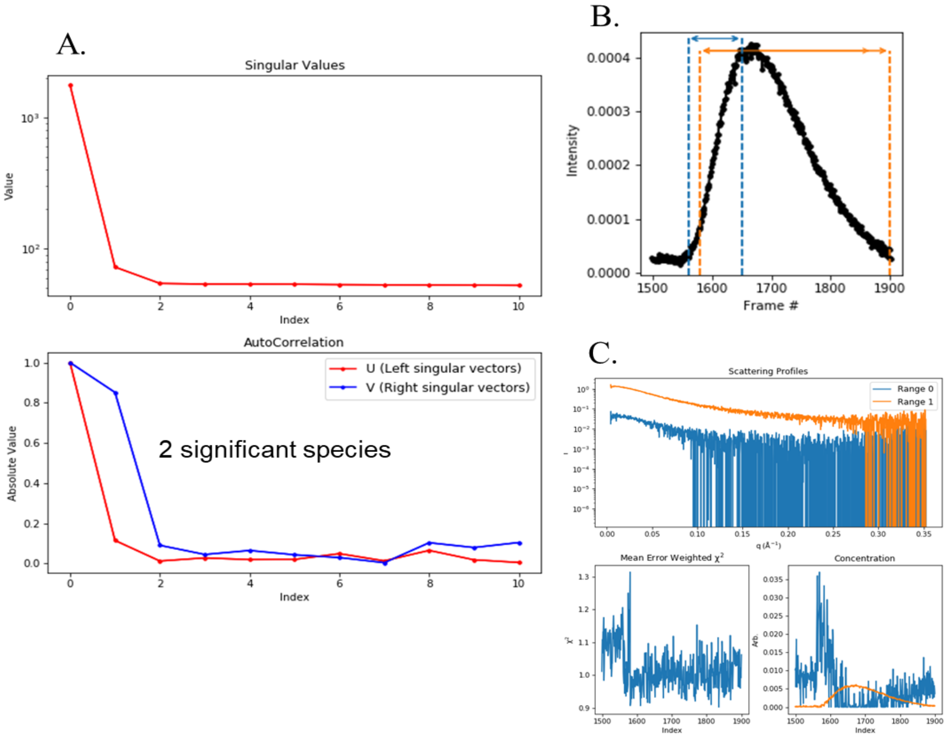


Figure 1.21: Singular Value Decomposition and Evolving Factor Analysis. A. Singular Value Decomposition Analysis and Autocorrelation versus scattering frame index, B. Ranges selected for each component across the scattering profile from Evolving Factor Analysis, C. Individual scattering profile and each of the χ^2 fit and concentration plots from Evolving Factor Analysis

1.6.6. Multi Angle Light Scattering (MALS)

Multi Angle Light Scattering (MALS) can be used to determine absolute molecular mass of flexible, extended proteins as well as different oligomeric states and conformations of proteins in solution. In this technique, the molecular mass determination is independent of SEC elution time and is solely dependent on the size of the macromolecule in solution. When a beam of light passes through a protein solution, it can diffract light at multiple angles.

Light can be defined as an oscillating wave of electric and magnetic fields. When light interacts with proteins, an oscillating electric field partially separates positive and negative charges in the particle, creating an oscillating dipole. This oscillating dipole can then re-radiate light predominantly in the plane perpendicular to the plane of polarization of light. The amount of charge separation depends on the polarizability of the protein and we can describe polarizability with the specific refractive index increment dn/dc (n is the refractive index of the solution and c is the molecular concentration of the protein) (Wyatt, 1993). So, if in a protein solution there are homodimers of a protein that will scatter twice as much as light compared to its corresponding monomer. Thus, this method is able determine the oligomeric state of a protein in solution.

The diagram (Figure 1.22) shows only one detector that can measure the intensity of scattered light at a single angle. In a MALS equipment there multiple such detectors that are arranged around the sample which measures the intensity of scattered light each at different angles. The Wyatt DAWN HELIOS II detector can be used to simultaneously measure light scattering data at 18 different angles.

The theoretical concept behind MALS can be explained by the Zimm plot:

$$\frac{K * c}{R(\theta, c)} = \frac{1}{M_w P(\theta)} + 2A_2 c$$

In the above equation,

- $R(\theta, c)$ is the excess Ryleigh ratio of the solution as a function of the scattering angle θ and solute concentration c. It is directly proportional to the intensity of the scattered light in excess of the light scattered by the pure solvent.

- M_w is average molecular mass
- A_2 is the second virial coefficient in the virial expansion of the osmotic pressure
- K^* is a constant $4\Pi^2(dn/dc)^2 n_0^2/N_a \lambda_0^4$ (n_0 = refractive index of the solvent, N_a = Avogadro's number, λ_0 = vacuum wavelength of laser light)
- $P(\theta)$ is the angular dependence of the scattered light

So, from the Zimm plot it is clear that the combined average molecular weight calculation from several different scattering angle gives a more accurate and precise estimation of the molecular weight.

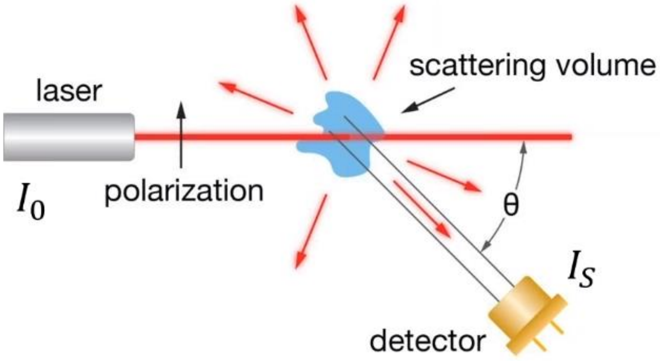


Figure 1.22: General principle for Multi Angle Light Scattering. Incident light with an intensity at 0° (I_0) upon interaction with a particle scatters light at different angles and the detector shown here detects the light scattering at a single angle (θ), with a scattered light intensity I(s)

1.6.7. Isothermal Titration Calorimetry

Isothermal Titration Calorimetry (ITC) is a biophysical technique used to quantify thermodynamic parameters of a molecular interaction. As chemical interactions are often associated with a heat exchange with the environment, this makes ITC a convenient method to quantify thermodynamics if the interaction. Further favorable features of ITC are that it is a label-free technique. It does not require immobilization of ligand or protein, and it uses relatively small amounts of ligands or proteins for a reliable estimation of thermodynamic parameters of interaction (Luis *et al.*, 2011).

Currently, there are two ITC machines that are widely used: the VP-ITC or iTC-200 from Microcal and Low Volume Nano-ITC from TA instruments. The latter was used in this study. These instruments can determine dissociation constants in a wide range from 10^{-9} to 10^{-4} M (Luis *et al.*, 2011).This instrument has a reference cell and a sample cell, made of highly efficient, thermally-conducting and chemically-inert material, and sensitive thermopile circuits used to detect temperature differences between the two cells. One binding partner is placed in the sample cell while the other is titrated into the cell at a much higher concentration. As the molecules interact in the cell, heat is either released or absorbed, but the sample and reference cells are maintained at equal temperatures by decreasing or increasing respectively the heating power supplied to the sample cell (Figure 1.23 A).

The heat change represents the change in enthalpy of the reaction, and can be calculated by integrating the heater power over time (secs) (Freyer & Lewis, 2008). The heat is discharged or consumed all along the reaction, with more heat being consumed (negative signal) or released (positive signal) at the beginning of the reaction (Figure 1.23 B). During the data analysis, the area under the peak is integrated and plotted versus the molar ratio of the molecule in the syringe to the molecule in the cell (Figure 1.23 B). The titration curve is then fitted to a suitable binding model to extract the enthalpy of interaction (ΔH), the association constant (K_a) and stoichiometry (n) (Figure 1.23 B). ΔH is represented by the distance between the two asymptomatic lines of the binding curve corresponding to the minimal and maximal heat formation (Figure 1.23 B). The slope at the inflection point of the binding curve reflects the K_a. The molar ratio at the inflection point of the binding curve determines the reaction stoichiometry. The change in the Gibbs free energy (ΔG) and the change in the entropy (ΔS) of the reaction can be calculated from ΔH, K_a and reaction temperature (Núñez *et al.*, 2012).

The C-value, defined as the product of the K_a times the molar concentration of the molecule in the sample cell times stoichiometry, should be ~ 10-100 (Wiseman *et al.*, 1989) for the ITC experiment. The C-value determines the binding parameters, which is important for obtaining a sigmoidal curve. Therefore, simulations are important for optimizing the ITC experiment and to achieve a balance between detectable heats and a sigmoidal binding curve. However, a good C-value may be difficult to achieve due to the intrinsic properties of some binding reactions such as an inability to reach to a sufficiently high protein concentration.

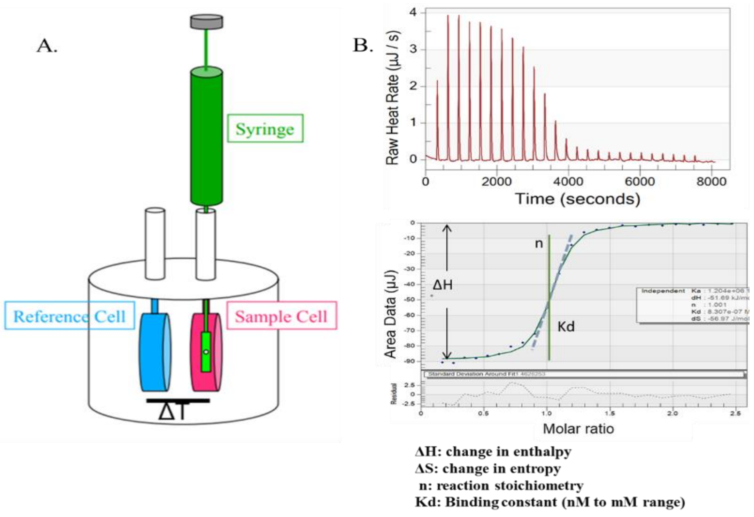


Figure 1.23: Basic Principle of ITC. A) General set up of an ITC equipment, B) An example ITC raw data and thermodynamic parameters measured

1.6.8. Hydrogen Deuterium Exchange Mass Spectrometry

Hydrogen deuterium exchange mass spectrometry (HDX-MS) is a powerful technique that can provide insightful links between protein structure, conformational dynamics, and function (Figure 1.24). The hydrogen atoms of O-H, N-H, and S-H groups freely exchange with the hydrogens of the aqueous environment. Disordered regions and surface exposed regions rapidly exchange with the surrounding water. Stably hydrogen bonded or residues sequestered from solvent by protein folding exchange at a slower rate. Therefore, if the protein is placed in

D_2O instead of H_2O, the exposed hydrogen atoms of O-H, N-H, and S-H are exchanged with deuterium. This results in an increase in the mass of the deuterium-exchanged protein molecules. While deuterium exchange does occur on side chains, they are typically more dynamic and less likely to remain unexchanged. Therefore, HDX-MS measures the changes in mass associated with the rate of exchange of the amide hydrogens of the protein backbone. The N-H → N-D rate of exchange depends on accessibility, which may be restricted due to the formation of intermolecular hydrogen bonds N-H···O=C in structured regions of a protein or because of being buried within the protein, thereby indicative of the folded state of the protein and its dynamics. The protection factor $P = k_{ch}/k_{HDX}$, where k_{ch} is the N-H → N-D conversion rate constant of N-H from disordered, solvent exposed residues; and k_{HDX} is N-H → N-D conversion rate constant of N-H from protected residues. P may exceed 10^6 for well-folded regions of the protein.

HDX-MS is also used to study protein conformational changes (Figure 1.24) and protein-protein interactions (Figure 1.25). Conformational fluctuations may cause some protected N-H groups to undergo N-H → N-D conversion at a measurable rate. However, exposed protein regions that become buried in a protein-protein interface become more protected in the presence of a binding partner. This will result in less deuteration compared to the deuteration levels in the absence of a binding partner.

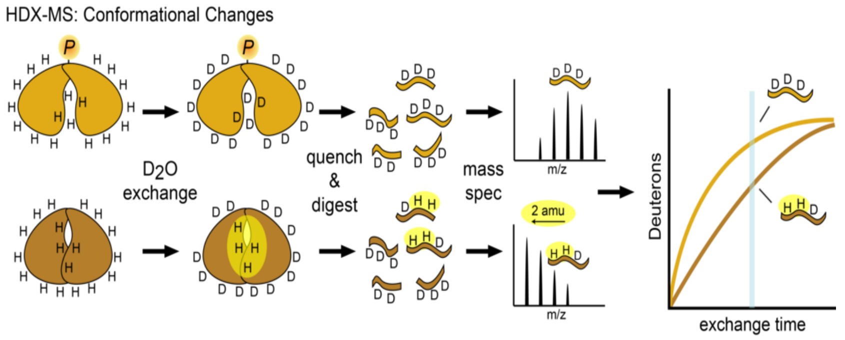


Figure 1.24: Schematic of analysis of protein conformational changes by HDX-MS (https://www.bbmb.iastate.edu/underbakke-lab/hydrogen-deuterium-exchange-mass-spectrometry)

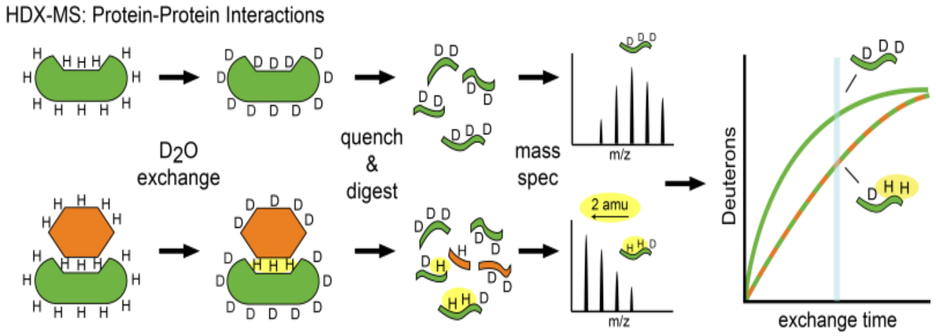


Figure 1.25: Schematic of protein-protein interaction analysis by HDX-MS (https://www.bbmb.iastate.edu/underbakke-lab/hydrogen-deuterium-exchange-mass-spectrometry)

In typical HDX-MS experiments, deuteration is performed under basic conditions (pD 8-10) at specific time points (such as 0s, 30s, 60s, 120s, 300s etc.) at around 25°C. The exchange is then quenched by rapid acidification, using a low pH solution of 3 M guanidine, 0.1% formic acid, pH 2.66 to prevent back-exchange (Konermann *et al.*, 2011). The deuteration is further prevented from back-exchange by handling the acidified samples at low temperatures (0°C – 4°C). This reduces the kinetic energy of the molecules in solution, reducing the rate of diffusion and leading to fewer collisions between the protein and hydrogen ions in solution. The quenched protein samples are then injected into a refrigerated Liquid Chromatography Mass Spectrometry system and digested in-line using an immobilized pepsin column while maintaining low pH and temperature. Pepsin is ideally suited for this since this stomach aspartic acid protease is maximally active at pH 1.5 – 2.5. The resulting peptides are captured on a reverse phase chromatography column; are separated by analytical liquid chromatography, using a 7-85% acetonitrile gradient in 0.1% formic acid over a predefined time; and analyzed by Electrospray Ionization Mass Spectrometry.

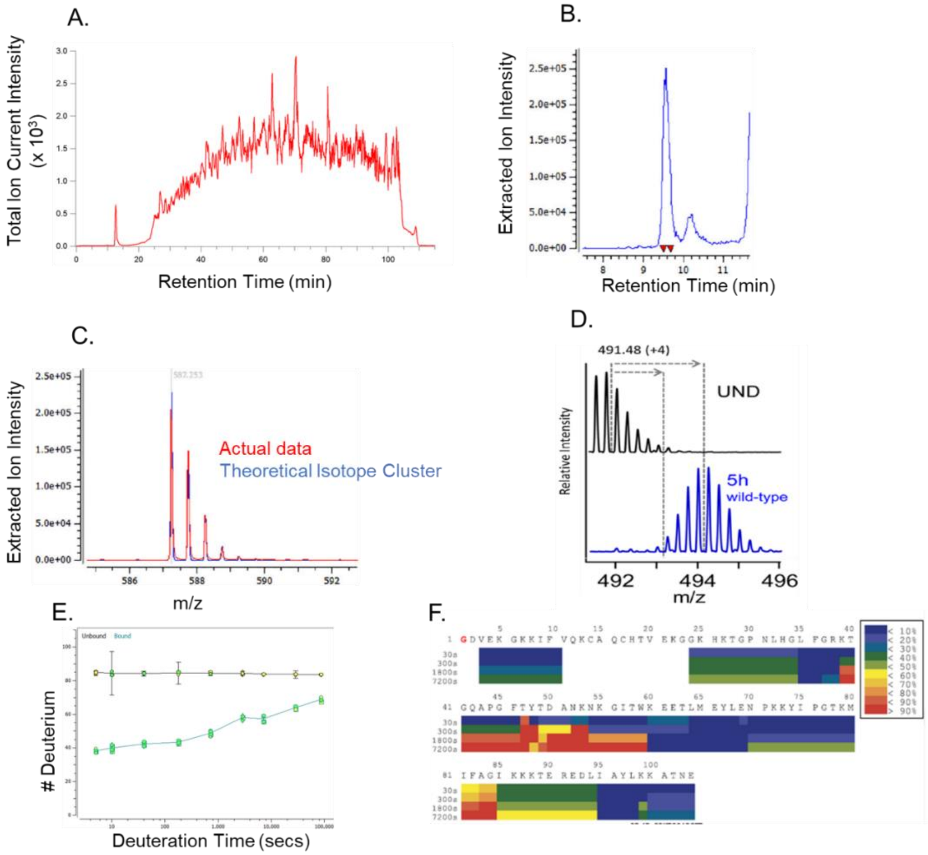


Figure 1.26: HDX-MS data analysis of an example protein. A. Total Ion Chromatogram. B. Extracted Ion Chromatogram of a specific peptide with Extracted ion intensity (on y-axis) and retention time in minutes(on x-axis), C. Match of the Actual (red) and Theoretical (blue) Isotope Clusters of a specific peptide over mass to charge (m/z) of the particular peptide, D. Deuterium Uptake Plot for deuterium (bound), E. Deuterium uptake (y-axis) over time (secs), F. Heat map of the entire protein, colors from blue to red represent different deuterium exchange rates as illustrated by the legend. (massspec.com/hdexaminer, (Campobasso & Huddler, 2015))

After the digested peptides are separated based on m/z ratio by LC- MS/MS, a total ion chromatogram is obtained (Figure 1.26 A). The total ion chromatogram (Figure 1.26 A) depicts the elution profile of digested peptides as detected by the mass spectrometer. Specific peptides in the total ion chromatogram can be extracted to generate the extracted ion chromatogram (Figure 1.26 B). Software such as the Protein Lynx Global Server 3.0 (PLGS) (Waters Corporation, Milford, MA) is then used for peptide mapping. Initially, a known protein sequence information is provided by the user to the PLGS software, along with other sample information such as posttranslational modifications, mutagenesis and the protease used for protein digestion during the experiment. The theoretical peptides thus generated are mapped such that the size of the observed peptides is matched to the theoretical peptides to identify the best fit between experimentally observed peptides and theoretically observed peptides. The experimentally observed peptides that show a good match in the retention time and mass/charge ratio (m/z) between the theoretically expected and experimentally observed peptides are further mapped to the entire sequence of the protein.

After the peptide map is generated from PLGS, the list of theoretical peptides that showed best fit with the experimentally observed peptides are further used to generate a peptide pool that can be analyzed for deuteration. This analysis is done in several steps. Initially the HDExaminer software is provided with the FASTA sequence for the protein, followed by creating a peptide pool from the peptide map obtained PLGS, then the LC-MS/MS data files for the undeuterated condition and deuterated conditions at various time points are loaded into HDExaminer. After the LC-MS/MS data gets imported, the software performs an iterative process for obtaining a good match between the theoretical isotope clusters and experimental isotope clusters for a given peptide (Figure 1.26 C). This data curation is done such that exactly identical peptides are analyzed for the various time points. Thus, data curation can be done by adjusting the retention time of the extracted ion chromatogram of a peptide. After the data curation is finished, a list of peptides based on the match between a theoretical isotope cluster and experimental isotope cluster; where a high confidence peptide indicates an excellent match, medium confidence peptides indicate a good match, and the low confidence peptides show little to no match. The high and medium confidence peptides can then be analyzed for deuteration uptake using a deuteration uptake plot. In a deuteration uptake plot, the isotopic distribution pattern for a peptide of interest shifts to higher mass number as incubation time increases due to increased deuterium exchange with hydrogens over time by comparison of the spectra (Figure 1.26 D, top versus bottom panel).

The number of deuterons (D) incorporated by each peptide at each D_2O time point (D(t)) is calculated using the following equation:

$$D(t) = N\frac{m(t) - m_0}{m_{100} - m_0}$$

where, N is the number of exchangeable amide hydrogens; *m(t)* is the centroid mass of the peptide at time, t; m_{100} and m_0 are the theoretical mass of the peptide with 100% and 0% deuteration.

Deuterium uptake plots are generated by plotting D(t) against time (Figure 1.26 E). After we have obtained the deuteration uptake plots with low error (obtained from the standard deviation of repeated trials for each deuteration time point) for each time point, a heat map is generated (Figure 1.26 F). The heat map represents the computed deuteration level at each residue of the protein, which is the best least-square match to the individual peptide measurements (Figure 1.26 F). Thus, when the heat map shows high deuteration, it indicates a region that was solvent exposed under the experimental conditions compared to regions with low deuteration, which are buried because of a compact conformational state or interaction partner preventing deuterium uptake.

1.7. Specific aims of this research

BECN1 plays a very important role in regulating autophagy as it is a protein interaction hub. Therefore, it is important for us to understand BECN1 structure and the structural basis of its interactions with other proteins, and how it regulates autophagy. The goal of this dissertation has been to delineate the structural and functional importance of the intrinsically disordered region in BECN1 and selected mechanisms by which other ordered regions in BECN1 regulate autophagy.

The specific aims of this research are:

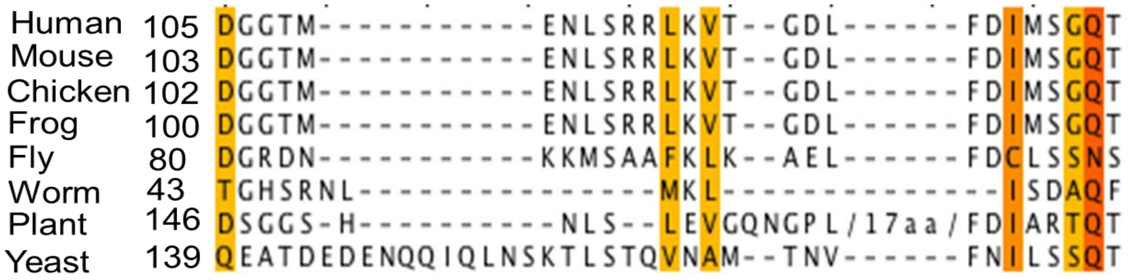
- To investigate the importance of the BECN1 non-conserved, unstructured BH3 domain and the conserved, ordered overlap helix in regulating autophagy. This study involved the quantitative analysis of the number of autophagosomes per cell under nutrient-rich and starvation-induced autophagy, using light microscopy.
- To delineate the minimal region of BECN1 required for interaction with CRM1. This was accomplished by performing affinity pull downs of purified CRM1, RanGTP and BECN1 NES containing constructs.
- To delineate the role of Zn^{2+} binding by the invariant C-x-x-C motifs in modulation of BECN1 IDR structure and function in autophagy. Purified FL BECN1 and the BECN1 IDR were assessed for metal binding using ICP-MS. The structure of the BECN1 IDR was assessed using CD spectroscopy to evaluate secondary structure content and potential binding-induced helicity, SEC-SAXS and HDX-MS to study conformational dynamics, and SEC-MALS for accurate molecular mass estimation. Lastly, the functional importance of BECN1 CxxC motifs and the BECN1 IDR was assessed using mammalian autophagy assays.

CHAPTER 2: IMPORTANCE OF CONSERVED, ORDERED OVERLAP HELIX AND NON-CONSERVED, UNSTRUCTURED BH3 DOMAIN OF BECN1 IN REGULATING AUTOPHAGY

2.1. Introduction

BECN1, Beclin 1 or VPS30 upregulates PI3KC3 activity in response to diverse signals. The exact mechanisms by which BECN1 performs this function is unknown, although BECN1 appears to be a major interaction hub for autophagy that interacts with at least twenty different cellular proteins. As described in Chapter 1, starvation is a common signal that induces autophagy. While the mechanism(s) by which BECN1 binds to so many diverse partners is/are unknown, conformational flexibility often facilitates multivalent protein interactions, allowing one protein to interact with various other binding partners to carry out diverse functions. Thus, structural, and biophysical studies investigating the conformational flexibility in BECN1 are important for understanding how eukaryotic organisms respond to and survive different stressors.

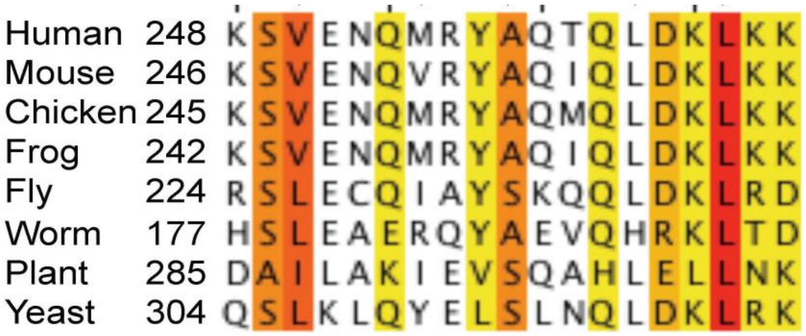
The BECN1 N-terminal region contains a BCL2 homology 3 domain or BH3D (human BECN1 residues 105-130), which is disordered but becomes helical upon binding to BCL2 proteins (Figure 2.1). BH3Ds were originally discovered in the context of the apoptosis regulators, as anti-apoptotic BCL2s bind to BH3Ds within pro-apoptotic Bcl-2 family members and down-regulate apoptosis. But several studies suggest that BH3Ds also enable down-regulation of autophagy (Sinha & Levine, 2008). Anti-apoptotic BCL2 proteins such as the γHV68 anti-apoptotic BCL2 protein, M11, bind to the BECN1 BH3D, to down-regulate autophagy. While the mechanism by which M11 and other BCL2s bind to the BECN1 BH3D has been extensively studied, the molecular mechanism by which this interaction inhibits autophagy is unknown. Further, while the BECN1 BH3D binds various autophagy regulators, to date, the role of the BECN1 BH3D in BECN1-mediated autophagy has not been investigated.



Human	105	DGGTM-----------ENLSRRLKVT--GDL------FDIMSGQT
Mouse	103	DGGTM-----------ENLSRRLKVT--GDL------FDIMSGQT
Chicken	102	DGGTM-----------ENLSRRLKVT--GDL------FDIMSGQT
Frog	100	DGGTM-----------ENLSRRLKVT--GDL------FDIMSGQT
Fly	80	DGRDN-----------KKMSAAFKLK--AEL------FDCLSSNS
Worm	43	TGHSRNL---------------MKL--------------ISDAQF
Plant	146	DSGGS-H----------NLS--LEVGQNGPL/17aa/FDIARTQT
Yeast	139	QEATDEDENQQIQLNSKTLSTQVNAM--TNV------FNILSSQT

Figure 2.1: Sequence alignment of BH3D in BECN1 orthologs. Increasing background color intensity corresponds to increasing residue conservation

Additionally, a highly-conserved BECN1 helical region comprising residues 248-265 which we named the ‘Overlap Helix’ (OH) (Figure 2.2) has been found to crystallize in two differently-packed states: packed against either the partner helix in a CCD homodimer or the BARAD (Figure 2.3) (Li *et al.*, 2012b, Mei, Su, *et al.*, 2016b) . Due to extensive steric conflicts between the BARAD and the partner helix of CCD, the OH cannot pack in both conformations simultaneously (Glover *et al.*, 2017). Calculation of buried surface areas show that packing of the BECN1 OH with the BARAD buries a larger surface area compared to packing with the partner helix within the CCD (Glover *et al.*, 2017). The five OH residues that are critical for this packing are Val^{250}, Met^{254}, Ala^{257}, Leu^{261} and Leu^{264} (Figure 2.3). The BECN1 CCD-BARAD was found to form a weaker homodimer than the BECN1 CCD, with a binding affinity comparable to a BECN1 CCD ‘TETRAD’ mutant where four critical OH residues (Val^{250}, Met^{254}, Leu^{261} and Leu^{264}) were mutated to Ala. This suggests that the OH contributes significantly to homodimerization of the CCD, and in the CCD-BARAD, packs preferentially against the BARAD. Further, the tetrad mutation in the CCD-BARAD further weakens CCD-BARAD homodimerization, which indicates that BECN1 OH transiently contributes to homodimerization in this construct as well (Glover *et al.*, 2017). As the OH interface residues are so important for regulating BECN1 conformation, we investigated the importance of these residues in regulating autophagy.



```
Human   248 KSVENQMRYAQTQLDKLKK
Mouse   246 KSVENQVRYAQIQLDKLKK
Chicken 245 KSVENQMRYAQMQLDKLKK
Frog    242 KSVENQMRYAQIQLDKLKK
Fly     224 RSLECQIAYSKQQLDKLRD
Worm    177 HSLEAERQYAEVQHRKLTD
Plant   285 DAILAKIEVSQAHLELLNK
Yeast   304 QSLKLQYELSLNQLDKLRK
```

Figure 2.2: Sequence alignment of the overlap helix in BECN1 orthologs. Increasing background color intensity corresponds to increasing residue conservation: with red corresponding to invariant residues of the Overlap helix(Glover et al., 2017)

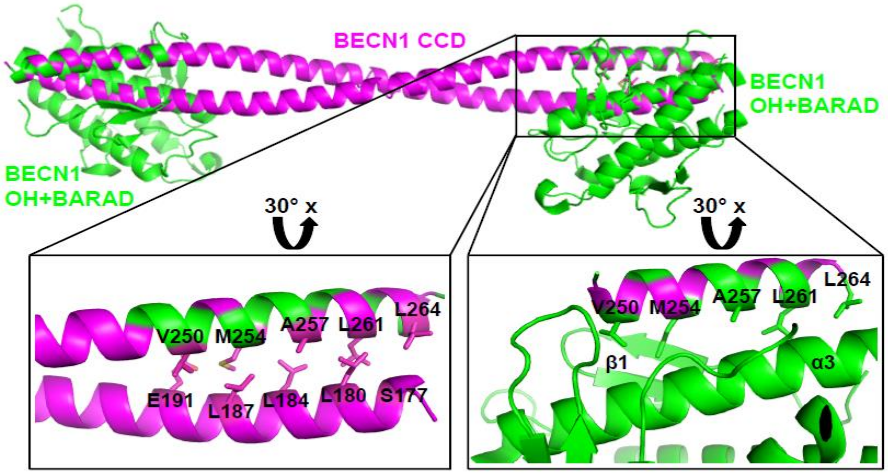


Figure 2.3: The OH has two different packing states. The BECN1 CCD (magenta) and BARAD (green) are shown in ribbon. The OHs in the CCD and BARAD structures are superimposed. Boxes indicate regions that are rotated 30° about the x-axis and enlarged to show the mutually exclusive packing arrangement of interacting OH side chains displayed as sticks. The first β-sheet and third helix of the BARAD, against which the OH side chains pack, are labeled β1 and α3, respectively.(Glover et al., 2017)

2.2. Materials and methods

2.2.1. Plasmid preparation

A pcDNA3.1 mammalian expression vector encoding full-length, human FLAG- BECN1 WT (WT) was subjected to site-directed mutagenesis to generate OH TETRAD mutant (V250A, M254A, L261A, and L264A). The FLAG-BECN1 ΔBH3D construct was prepared from the WT construct by deleting residues 105 to 130 by site-directed mutagenesis. The BECN1 OH TETRAD mutant construct was prepared by Dr. Karen Glover and the BECN1 ΔBH3D mutant construct was prepared by Dr. Yue Li.

2.2.2. Optimizing protein over-expression and autophagy assay

We investigated how deletion of the BECN1 BH3D or OH impacts cellular autophagy levels mediated by exogenously expressed BECN1, by quantifying and comparing the impact of either ΔBH3D or TETRAD mutants to WT. As BECN1 is known to be required for autophagosome nucleation, we evaluated cellular autophagy by monitoring and comparing levels of puncta labeled with GFP-tagged LC3, an autophagosome-specific marker, in cells grown in either nutrient-rich or starvation medium. We used human breast adenocarcinoma MCF7 cells because they lack detectable endogenous expression of BECN1, resulting in very low basal levels of autophagy (Liang *et al.*, 1999, Pattingre *et al.*, 2005, Sinha, Colbert, Becker, Wei & Levine, 2008) even in starvation conditions unless BECN1 is ectopically expressed. This allows the effect of BECN1 mutants to be assayed in the absence of endogenous BECN1. So, before performing autophagy assay we first ensured comparable exogenous expression of BECN1 WT and its mutants in starvation and nutrient-rich conditions. Cellular autophagy levels were quantified by monitoring cellular localization of GFP-tagged LC3 protein (Kabeya *et al.*, 2000). Each chamber of an 8-well culture slide (Nunc^{TM} Lab-Tek^{TM} Chambered Coverglass) was seeded with 1 ×10^5 MCF7 cells and cultured overnight in DMEM (Gibco) with 10% fetal bovine serum (Gibco) until 80–90% confluent. The cells were co-transfected with 200 ng of GFP-LC3 and 300 ng of WT or 300 ng of TETRAD or 150 ng of ΔBH3D mutant expression plasmids, using Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's instructions. After transfection and incubation at 37 °C for 24 h, the cells were cultured in either rich (DMEM, 10% FBS, 2 × essential amino acids, and 2 × nonessential amino acids) or starvation (Earle's balanced salt solution) medium for 4 h. The cells were counterstained with DAPI to visualize nuclei and facilitate total cell counts, fixed to slides with 4% (v/v) paraformaldehyde, and then stored in 70% (v/v) glycerol. Cells were washed with PBS between the counterstaining, fixation, and storage stages. GFP-LC3-positive puncta were observed under a Zeiss LSM700 Confocal Laser Scanning Microscope (CLSM) (Zeiss, Thornwood, NY), using excitation at 488 nm and emission set at 590 nm using Plan-Apochromat 40x/1.4 oil immersion lens (Figure 2.5). The GFP-LC3 labelled puncta was quantified by counting a minimum of 50 cells for duplicate samples per condition using Imaris software (Bitplan AG, Zurich, Switzerland) in three independent experiments. The significance of alterations in autophagy levels was determined by a two-tailed, heteroscedastic Student's *t* test, wherein $p \leq 0.05$ is considered significant. Similar expression levels of WT, ΔBH3D or TETRAD mutants were verified by Western blotting using commercial mouse monoclonal anti-FLAG M2-peroxidase (1:1000) antibody (Sigma). Actin levels in MCF7 cell lysates, detected using mouse anti-β-actin antibody (1:1000) (EMD-Millipore) served as a loading control. In the expression assay for the TETRAD mutant, the anti-β-actin treated membrane was further treated with Goat-Anti-Mouse IgG-HRP (1:2000) (Abcam) antibody. The detection of both anti-FLAG and secondary antibody-treated anti-β-actin were done using Thermo Scientific™ Pierce™ ECL 2 Western Blotting Substrate and scanned by

Storm 860 imager (GE Healthcare). In the expression assay for ΔBH3D, membranes were treated with an Irdye 800CW goat anti-mouse secondary antibody and scanned using the 800nm channel of Odyssey® CLx Imager (LI-COR).

2.3. Results

2.3.1. Autophagy assay results for the BECN1 ΔBH3D mutant

Transient expression of BECN1 in MCF7 cells does not increase autophagy levels in nutrient-rich conditions ($p = 0.775$ for WT expression *versus* no expression; Figure 2.4 B,C) but leads to a marked increase in autophagy upon starvation ($p = 0.0102$ for WT expression in starvation condition *versus* nutrient-rich condition) (Figure 2.5 B,C). We find that deletion of the BH3D does not impact autophagy levels in nutrient-rich media ($p = 0.245$ for ΔBH3D *versus* WT), nor does it prevent increased autophagy in starvation media ($p = 0.048$ for ΔBH3D *versus* WT) (Figure 2.5 B,C), indicating that the BECN1 BH3D is not essential for upregulating starvation-induced autophagy. Therefore, binding of BCL2 homologs to the BECN1 BH3D does not by itself explain how BCL2 homologs down-regulate autophagy.

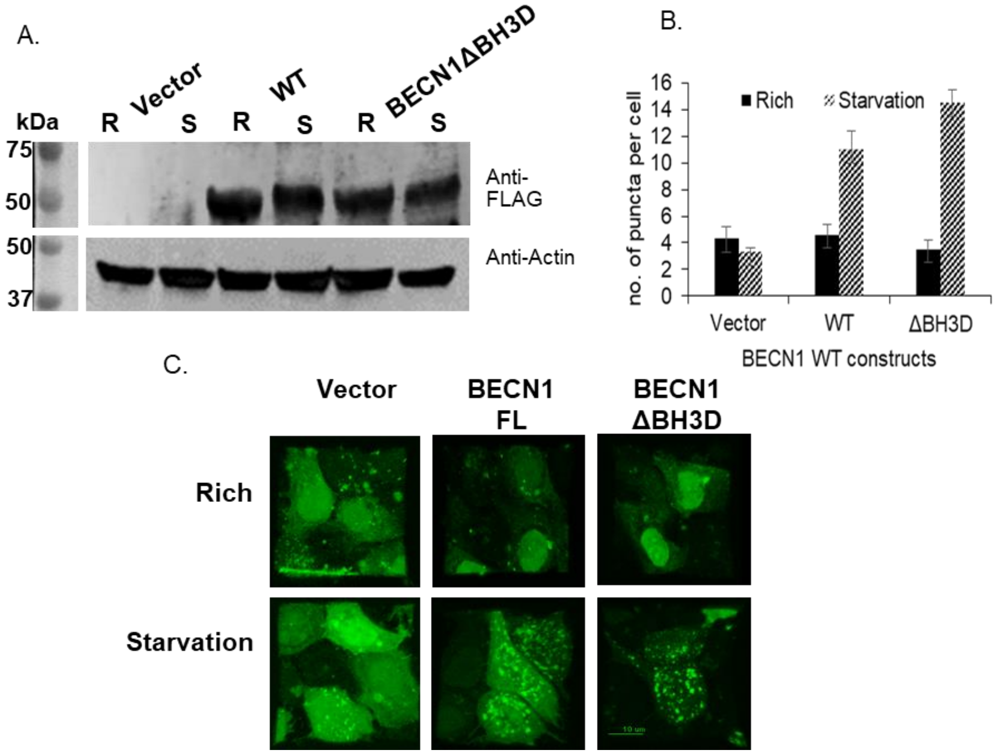


Figure 2.4: Effect of the BECN1 BH3D mutation on autophagy. A) Western blot of MCF7 cell extracts showing comparable expression levels of WT and mutant FLAG-BECN1 in nutrient rich (R) and starvation (S) conditions, with Actin as a loading control. B) Light microscopy quantification of discrete GFP-LC3 puncta per cell in GFP-positive MCF7 cells co-transfected with GFP-LC3 and WT or mutant FLAG-BECN1 as indicated below the x-axis. Bars represent the number of puncta per cell. Error bars represent standard deviation. C) Representative images of GFP-LC3 (green) fluorescence in cells grown in starvation or nutrient rich media and transfected with mutant FLAG-BECN1 as indicated.

2.3.2. Autophagy assay results for the BECN1 OH TETRAD mutant

We investigated the impact of destabilization of the OH on cellular autophagy levels mediated by exogenously expressed full-length BECN1, by quantifying and comparing the impact of either WT or TETRAD (V250A, M254A, L261A, and L264A) mutant of BECN1. Since BECN1 is known to be required for autophagosome nucleation, we evaluated cellular autophagy by monitoring and comparing levels of puncta labeled with GFP-tagged LC3, an autophagosome-specific marker, in cells grown in either nutrient-rich or starvation media. Expression of WT and TETRAD was comparable in starvation and nutrient-rich conditions

(Figure 2.5 A). Transient expression of BECN1 in MCF7 cells did not increase autophagy levels in nutrient-rich conditions (p = 0.951 for WT expression versus no expression; Figure 2.5 B, C); but led to a marked increase in autophagy upon starvation (p = 0.0047 for starved versus nutrient-rich cells; Figure 2.6 B, C). We find that the TETRAD does not impact autophagy levels in nutrient rich media (p = 0.836 for TETRAD versus WT); consistent with the lack of impact of BECN1 expression on basal autophagy levels in nutrient-rich conditions. Strikingly however, the TETRAD abrogates the starvation-induced up-regulation of autophagy (p = 0.004 for TETRAD versus WT), indicating that these residues are essential for up-regulating starvation-induced autophagy. This is consistent with the established importance of other conserved regions of BECN1 in starvation-triggered autophagy, rather than basal autophagy levels (Mei, Ramanathan, *et al.*, 2016, Mei, Su, *et al.*, 2016b).

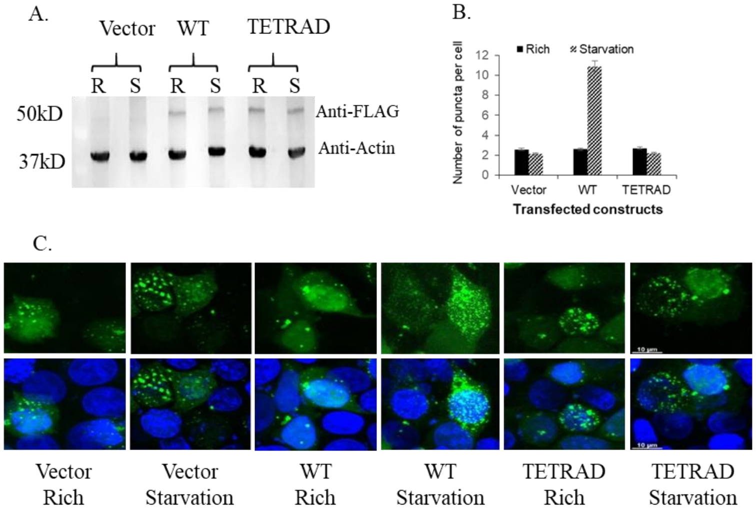


Figure 2.5: Effect of the OH Tetrad mutation on autophagy. A) Western blot of MCF7 cell extracts showing comparable expression levels of WT and TETRAD mutant FLAG-BECN1 in nutrient rich (R) and starvation (S) conditions, with Actin as a loading control. B) Light microscopy quantification of discrete GFP-LC3 puncta per cell in GFP-positive MCF7 cells co-transfected with GFP-LC3 and WT or TETRAD as indicated below the x-axis. Bars represent the number of puncta per cell. Error bars represent standard deviation. C) Representative images of GFP-LC3 (green) fluorescence and DAPI (blue) staining in cells grown in starvation or nutrient rich media and transfected with mutant FLAG-BECN1 as indicated.(Glover *et al.*, 2017)

2.4. Discussions and conclusions

The cellular autophagy assays indicate that the deletion of the BH3D does not impact the starvation-induced upregulation of autophagy by BECN1, i.e. the BH3D does not play a direct role in the execution of autophagy. This indicates that binding of the anti-apoptotic BCL2 proteins to the BECN1 BH3D downregulates autophagy by a mechanism other than directly impeding a pro-autophagy function of the BH3D. Additional unpublished data from our lab indicates that binding of the γHV68 BCL2, M11, to a BECN1 BH3D-FHD-CCD fragment improves self-association affinity from a K_d of 16.4 µM to 4.0 µM. BECN1 homodimers represent the autophagy-inactive state of BECN1 as in this state BECN1 cannot interact with other autophagy proteins that bind via the CCD such as ATG14 or UVRAG. Therefore, it appears that the BCL2-binding to the BECN1 BH3D stabilizes autophagy-inactive oligomeric states and prevents interactions of other BECN1 domains that are required for BECN1 function in autophagy. Notably, other BECN1 binding partners such as VMP1 that bind via the BH3D are known to upregulate autophagy. Thus, while the BH3D is not directly required for autophagy, BCL2-binding to the BH3D likely prevents binding of non-essential partners such as VMP1 that up-regulate autophagy, and also appears to impact structure, oligomerization, interactions, and function of other BECN1 domains.

The BECN1 OH packs in two mutually exclusive states, involving the same set of interface residues. Our cellular autophagy assays indicate that mutation of these BECN1 OH interface residues (the OH TETRAD mutation) does not impact basal autophagy, but abrogates the starvation-induced up-regulation of autophagy (Glover *et al.*, 2017). In this study we also showed that in full-length BECN1, the OH packs preferentially against the BARAD, rather than the CCD, weakening the autophagy-inactive homodimer. Further, a previous SAXS-constrained model showed that OH also contributes critical binding determinants to the parallel CCD heterodimer with ATG14 (Mei, Su, *et al.*, 2016b), as well as to the parallel CCD heterodimer with UVRAG (Rostislavleva *et al.*, 2015a), required to form VPS34 complexes essential for autophagy. Together, these studies suggest a mechanism when homodimerization via the BECN1 CCD results in BECN1 conformations that prevent the BARAD aromatic finger from associating with membrane, whereas BECN1 heterodimerization with ATG14 and UVRAG disrupts this inhibitory conformation.

CHAPTER 3: PROVING THAT CRM1 INTERACTS WITH BECN1

3.1. Introduction

BECN1 contains a Nuclear Export Signal (NES) sequence, ^{180}LQMELKELALE^{190}, located at the N-terminus of the CCD. An early study had used confocal microscopy to show that BECN1 colocalized to intracytoplasmic organelles and nuclei in COS7 monkey kidney and MCF7 human breast carcinoma cells (Liang *et al.*, 2001) (Figure 3.1). Chromosomal Region Maintenance 1 (CRM1) is a major, highly conserved nuclear exporter. Notably, when either CRM1-dependent nucleocytoplasmic transport is inhibited by leptomycin B, or the BECN1 NES sequence is mutated (L184A,L187A), BECN1 localizes predominantly to the nucleus (Liang *et al.*, 2001) (Figure 3.1). The BECN1 NES mutant also abrogated both basal and starvation induced autophagy in MCF7 cells (Liang *et al.*, 2001).

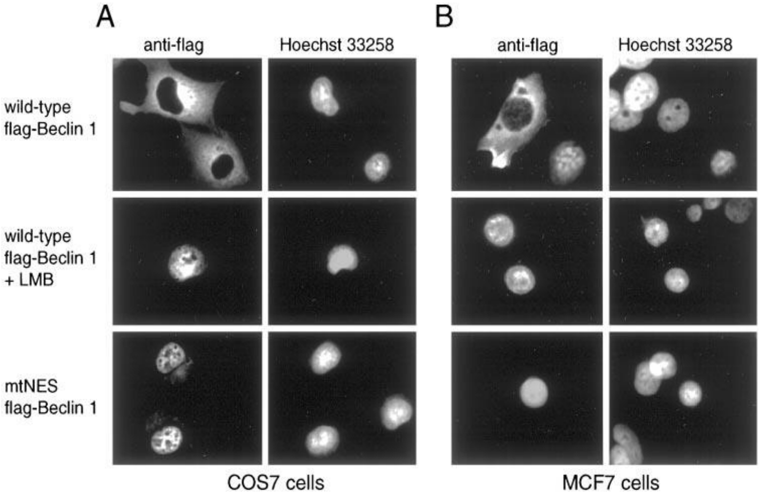


Figure 3.1: Leptomycin B treatment and mutation of the Beclin 1 NES alters the subcellular localization of Beclin 1 in COS7 and MCF7 cells. COS7 cells (A) and MCF7 cells (B) transfected with wild-type flag-beclin 1 (top and middle rows of each panel) in the absence (top rows of each panel) or presence of leptomycin B (middle rows of each panel) or transfected with a NES mutant (mtNES flag-Beclin 1) (lower rows of each panel) (Liang et al., 2001).

The function of BECN1 in the nucleus is poorly understood. In a recent study, BECN1 was proposed to promote DNA damage repair(Xu *et al.*, 2017). Knock down of BECN1 expression in HeLa cells exposed to ionizing radiation, significantly reduces the rate of repair of the double stranded breaks by both homologous recombination and nonhomologous end-joining pathway (Xu *et al.*, 2017). This reduction in double-stranded break repair may lead to an accumulation of mutations that result in aging, genetic disease such as cancers or neurodegenerative diseases (Bohgaki *et al.*, 2010). BECN1 has also been shown to localize to the nucleus during post-natal development in mice(Xu *et al.*, 2017). During embryonic development of neonatal mice BECN1 is found in cytoplasm; but by the time the mice are 15 days old, about 50% of BECN1 is found in the nucleus; and by 20 days the majority of BECN1 is in the nucleus(Xu *et al.*, 2017). BECN1 residues 1-50 and 254-278 are needed for the nuclear localization of BECN1(Xu *et al.*, 2017).

As described in Chapter 2, BECN1 residues 250-264 constitute an overlap helix (OH) that packs in two mutually exclusive states: either against the N-terminal region of the partner helix of the CCD (Li *et al.*, 2012b, Mei, Su, *et al.*, 2016a) or against the C-terminal BARAD (Huang *et al.*, 2012). Recently, we have shown (Glover *et al.*, 2017) that BECN1 OH preferentially packs against the BARAD, which would release the N-terminal region of the BECN1 CCD containing the NES sequence for other interactions, such as that with CRM1. However, evidence of direct protein-protein interaction between CRM1 and either BECN1 or the BECN1 NES is currently unavailable.

CRM1 is known to export around 1000 different proteins (Kırlı *et al.*, 2015) from either yeast, human or frog cell nuclei. It exports proteins from the nucleus to the cytoplasm, across the nuclear membrane, through the Nuclear Pore Complexes (NPCs) (Beck & Hurt, 2017). Inside nuclei, CRM1 binds cargo molecules cooperatively with the GTP-bound small G-protein Ran (RanGTP), and the heterotrimeric RanGTP-exportin-cargo complex is transported to the cytoplasm, where GTP is hydrolyzed, triggering release of cargo and RanGDP. The free CRM1 can then re-enter nuclei to export the next cargo.

CRM1 is structurally well-characterized. It is a ring-shaped protein comprising of 20 Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A), and yeast kinase TOR1 (HEAT) repeats (Figure 3.2). The CRM1 ring is not planar (Figure 3.3), as HEAT repeats 2-7 (H2-H7; most of H1 is disordered) twist away from the plane of the ring with an ~45° rotation (Dong, Biswas & Chook, 2009). Residues 1031-1057 comprise a 27-residue C-terminal helix (C-helix or HEAT-repeat helix 21B) that lies across the plane of the ring and connects helix H20B to the B helices of the H8-H12.

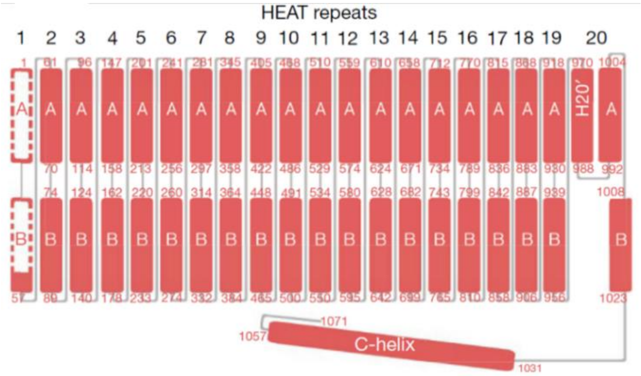


Figure 3.2: HEAT repeat organization of CRM1. Most of H1 is disordered and not modelled in the structure (Dong, Biswas & Chook, 2009)

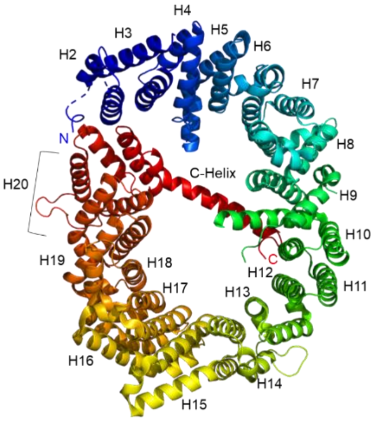
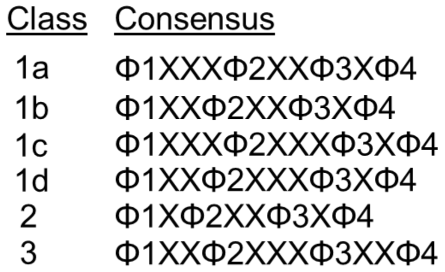


Figure 3.3: Crystal structure of Human CRM1(residues 1-1071). The CRM1 polypeptide chain is colored in rainbow with HEAT repeats 2-20 and the C-terminal helix.

CRM1 recognizes short linear nuclear export signal (NES) sequences in cargo proteins which comprise of 4-5 critical hydrophobic (φ) residues with characteristic spacing (Wen *et al.*, 1994, Fischer *et al.*, 2015, Dong, Biswas, Süel, *et al.*, 2009, Güttler *et al.*, 2010, Monecke *et al.*, 2013). CRM1 binds to 6 different classes of NES sequences called 1a, 1b, 1c, 1d, 2, and 3, (Figure 3.4) each of which have a characteristic spacing of hydrophobic amino acids (Fung *et al.*, 2015). The BECN1 NES, ^{180}**L**QME**L**KE**L**A**L**E^{190}, is a class 1a NES.



Class	Consensus
1a	Φ1XXXΦ2XXΦ3XΦ4
1b	Φ1XXΦ2XXΦ3XΦ4
1c	Φ1XXXΦ2XXXΦ3XΦ4
1d	Φ1XXΦ2XXXΦ3XΦ4
2	Φ1XΦ2XXΦ3XΦ4
3	Φ1XXΦ2XXXΦ3XXΦ4

Figure 3.4: Six different classes of the Nuclear Export Signal (NES) consensus patterns (Φ is Leu, Val, Ile, Phe or Met; X is any amino acid)(Fung *et al.*, 2015).

CRM1 interacts with RanGTP and NES-bearing cargo protein in a cooperative fashion. A molecular dynamic simulation suggested a model for CRM1-RanGTP-cargo export complex assembly and disassembly (Figure 3.5). This complex is assembled in the nucleus and gets disassembled in the cytoplasm. As suggested by EM data (Monecke *et al.*, 2013), CRM1 in its resting state exists predominantly in an extended conformation, with the C-terminal helix 21B occupying the concave surface where RanGTP binds. Another important structural aspect of the resting CRM1 conformation is the flipped-back conformation of the acidic loop packed on the back side of the NES cleft, against the surface of helices 10B, 11B and 12B. Inside the nucleus during the assembly of CRM1-RanGTP-cargo complex, either cargo or RanGTP can bind first to CRM1, which leads to a conformational change in CRM1, that facilitates binding of the second protein (Monecke *et al.*, 2013). When RanGTP binds first to CRM1, the C-terminal helix 21B becomes oriented parallel to helix 21A. This process of release and reorientation of helix 21B is likely an energetically unfavorable process, which leads to CRM1: RanGTP binding with low micromolar affinity. On the other hand, if cargo binds first to CRM1 then the acidic loop gets released from a flipped-back position, allowing the cargo NES sequence to interact with the hydrophobic NES binding cleft on CRM1 formed by helices 10B, 11B and 12B (Monecke *et al.*, 2013).This release and reorientation of the acidic loop of CRM1 is likely also an energetically unfavorable process that allows CRM1: cargo binding with a low micromolar affinity. But when either RanGTP or cargo binds, this leads to a conformational change in CRM1 that makes CRM1 a less extended molecule and the terminal HEAT repeats interact with each other (Monecke *et al.*, 2013). This conformational change allows CRM1 to bind to the third component in the complex with a much tighter, low nanomolar affinity(Monecke *et al.*, 2013). The CRM1-RanGTP-cargo complex thus formed is exported through the NPC and reaches the cytoplasm. Inside the cytoplasm cargo gets disassembled from CRM1-RanGTP complex due to hydrolysis of RanGTP to RanGDP, after which CRM1 and RanGDP gets transported back to the nucleus for subsequent rounds of nuclear transport.

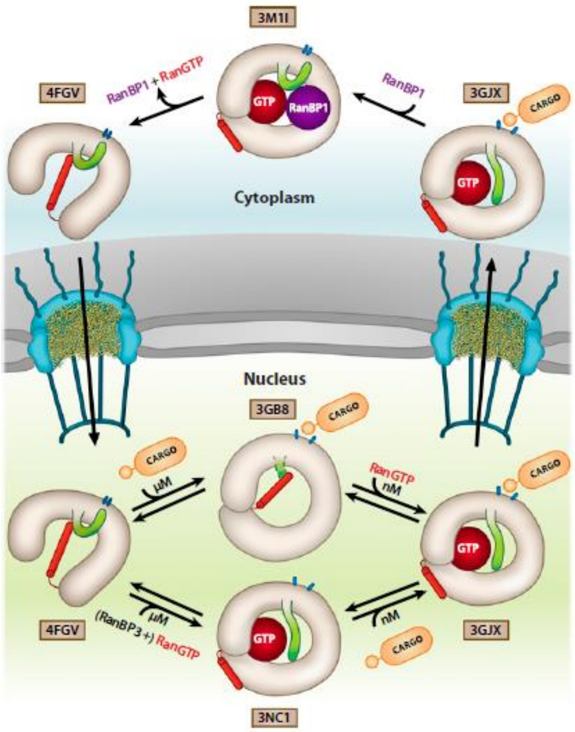


Figure 3.5: Model for cooperative CRM1 export complex assembly and disassembly showing conformational variability and important structural features in different states of the transport cycle. Different CRM1 conformations shown are colored gray with the acidic loop highlighted in green. The CRM1 C-terminal helix is shown in red, and the NES binding cleft is represented in blue ovals. The PDB ID codes of crystal structures corresponding to each state are indicated. (Monecke et al., 2013)

3.2. Materials and methods

3.2.1. Plasmid preparation

Plasmids for these experiments were obtained from Dr. Yuh Min Chook (Dong, Biswas, Süel, et al., 2009, Fung et al., 2015). These included human CRM1 construct pGX4T3 GST-CRM1 (1-1071). GSP1, a yeast homolog of RAN, shares 88% sequence identity with human RAN. Therefore, as a pET21d GSP1(1-179)Q71L-His_6 expression was readily available from Dr. Chook, it was substituted for human RAN in the experiments reported in this chapter. pMBP-Parallel-1 BECN1 (175-265), pMBP-Parallel-1 BECN1 (175-450) and pMBP-Parallel-1 BECN1(1-450) with aromatic finger mutation (F359D, F360D,W361D), constructs were obtained from Dr. Karen Glover and Dr. Yue Li respectively.

Table 3.1: List of protein constructs used for this research

Label	**Description**	**Molecular mass (kDa)**	**Purpose**
GST-CRM1	Human CRM1 (residues 1-1071), GST fusion protein	148.4 kDa	GST pull down assay
CRM1	Human CRM1 (residues 1-1071)	123.4 kDa	Ni-NTA pull down assay
GSP1(1-179)Q71L-His$_6$	Yeast Ran, a C-terminal His-tagged fusion protein	23.7 kDa	Ni-NTA pull down assay
MBP-BECN1 CCD	BECN 1 residues 175 to 265 with MBP tag	54.1 kDa	GST pull down assay
BECN1 CCD	BECN 1 residues 175 to 265 without MBP tag	11.1 kDa	Ni-NTA pull down assay
MBP-BECN1 CCD BARAD^{AFM}	BECN 1 residues 175 to 450, with the aromatic fingers (F359D, F360D,W361D) mutated with MBP tag	75.6 kDa	GST pull down assay
BECN1 CCD BARAD^{AFM}	BECN 1 residues 175 to 450, with the aromatic finger (F359D, F360D,W361D) mutation	32.3 kDa	GST pull down assay
BECN1 ^{AFM}	BECN 1 full length with the aromatic finger (F359D, F360D,W361D) mutation	52.0 kDa	Ni-NTA pull down assay

3.2.2. Expression and protein purifications

BL21(DE3) cells (NEB) were transformed with the pGEX-4T3 GST-CRM1(1-1071) plasmid and used for expression as previously reported (Dong, Biswas, Süel, *et al.*, 2009). The frozen cell pellets from 1L cells were allowed to thaw by adding 25 ml cold buffer (50 mM Tris, pH 7.5, 200 mM sodium chloride, 1 mM EDTA, 2 mM DTT, 20% glycerol) per liter of cells, and stirring at 4°C on a magnetic stir plate. The thawed cells were then sonicated 4 times for 1 minute at 2 minutes intervals. The lysate was clarified at 20,000 g for 30 mins. The supernatant was loaded onto a 5 ml Glutathione Sepharose 4B resin column that has been pre-equilibrated with 20 column volume of the buffer mentioned above, followed by elution with 20mM Glutathione. The eluted protein was further purified to ~95% homogeneity using SEC by use of Superdex 200 10/300 column (GE Healthcare). For preparation of untagged CRM1, the GST tag was cleaved by addition of GST-TEV protease in 1:10 molar ratio and overnight at 4°C. Subsequently, the column was washed with the buffer above and wash fractions were collected. Lastly, CRM1 was purified to ~95% homogeneity by SEC, using a Superdex 200 10/300 column in 50 mM Tris, pH 7.5, 200 mM sodium chloride, 1 mM EDTA, 2 mM DTT, 10% (v/v) glycerol buffer.

BL21(DE3) cells (NEB) were transformed with pET21d GSP1(1-179)Q71L-His$_6$, and expression was performed according to a previously published protocol (Fung *et al.*, 2015). Frozen cells pellets from 1L cells were allowed to thaw by adding 25 ml cold Buffer A (50 mM HEPES pH 7.4, 200 mM NaCl, 2 mM MgOAc, 10% (v/v) glycerol, 25mM Imidazole, 2mM DTT) in the presence of 1 protease inhibitor cocktail tablet per liter cells, and stirring at 4°C on a magnetic stir plate. The thawed cells were then sonicated 4 times for 1 minute at 2-minute intervals. The lysate was clarified at 20,000 g for 30 mins. The supernatant was loaded onto 2 ml Ni-NTA resin column that pre-equilibrated with 20 column volumes of Buffer A. The loaded column was washed with 50 ml Buffer A, followed by elution with Buffer B 50 mM HEPES pH 7.4, 200 mM NaCl, 2 mM MgOAc, 10% (v/v) glycerol, 250 mM Imidazole, 2 mM DTT. Prior to GTP loading, the Ni-NTA eluate was concentrated to 34 mg/ml (1430 µM). The concentrated protein was first treated with 50 mM (molar excess) of EDTA for 30 min on ice, followed by incubation with 50 mM (molar excess) each of GTP (ACROS Organics™) and MgOAc for 30 min at room temperature. The GSP1-GTP was loaded onto a Mono S (5/50 GL) column, and separated using low and high salt buffers (50 mM HEPES pH 7.4, 2 mM MgOAc, 2 mM DTT, 50 mM HEPES pH 7.4, 1M NaCl, 2 mM MgOAc, 2 mM DTT) respectively. Lastly, the cation-exchange purified protein was purified to ~95% homogeneity by SEC using 20 mM HEPES pH 7.4, 110 mM KOAc, 2 mM MgOAc, 10% (v/v) glycerol, 2 mM DTT, on a Superdex 75 increase column.

MBP-BECN1 FL^{AFM} (residues 1-450), MBP-BECN1 CCD-BARAD^{AFM} (residues 175 to 450) and MBP-BECN1 CCD (residues 175-265) were expressed and purified as previously reported (Glover *et al.*, 2017). The MBP-BECN1 FL^{AFM} and MBP-BECN1CCD-BARAD^{AFM} were expressed in ArcticExpress(DE3) cells (Agilent), whereas MBP-BECN1 CCD was expressed in BL21(DE3)pLysS cells grown in LB media containing 100µg/ml of ampicillin. ArcticExpress(DE3) cells expressing MBP-BECN1 FL^{AFM} and MBP-BECN1CCD-BARAD^{AFM} were allowed to grow until the OD$_{600}$ was ~0.6 at 30°C, and then the temperature was reduced to 13°C, prior to inducing with 0.5 mM IPTG when OD$_{600}$ was between 0.8-1.2, and allowing overnight expression. Similarly, BL21(DE3)pLysS cells expressing MBP-BECN1 CCD were allowed to grow until the OD$_{600}$ was ~0.8 at 37°C, prior to equilibrating the temperature to 20°C and inducing with 0.5 mM IPTG for overnight expression. Cells were then harvested by centrifugation at 4000g for 15 mins, and stored at -80°C until used for purification.

Frozen cell pellets were allowed to thaw by adding 25 ml cold Buffer A (25mM Tris pH 7.5, 300 mM NaCl, 10% (v/v) glycerol, 1 mM EDTA, 2 mM DTT) for 1L cells, in presence of one protease inhibitor cocktail tablet per liter of cells while stirring at 4°C on a magnetic stir plate. The thawed cells were then sonicated for 1 minute for 4 times in 2 minutes intervals. The lysate was clarified at 20,000 g for 30 mins. The supernatant from 6 L cells was loaded onto a 10 ml Amylose resin column pre equilibrated with 20 column volumes of Buffer A. After the lysate was loaded, the column was washed by 100 ml Buffer A followed by elution by use of Buffer B (25 mM Tris pH 7.5, 300 mM NaCl, 2 mM DTT). The affinity-purified protein was further purified to ~95% homogeneity by SEC in Buffer C (50 mM Tris pH 8, 300 mM NaCl, 2 mM DTT) using a Superdex 200 10/300 column (Pharmacia Biotech – GE Life Sciences, Pittsburgh, US).

I thank Dr. Karen Glover for generously sharing purified untagged BECN1CCD, untagged BECN1 CCD-BARAD^{AFM} and Dr. Yue Li for untagged BECN1 ^{AFM} proteins for carrying out the pull-down assays.

3.2.3. Pull down assays

Ni-NTA and GST affinity pull down assays were performed to assess binding of CRM1 to BECN1 NES-containing constructs including (1) BECN1 CCD, (2) BECN1 CCD-BARAD^{AFM} and (3) BECN1^{AFM}.

Lysate from 1L cells containing unpurified GST-CRM1 was first allowed to pass through 5 ml Glutathione Sepharose 4B resin (GE Lifesciences), followed by at least 5 column volume (CV) washes with the wash buffer (50mM Tris pH 8, 300 mM NaCl, 2 mM DTT). Next the column was loaded with either 8 mg of purified BECN1 CCD or 2 mg of purified BECN1 CCD-BARAD^{AFM} proteins. The column was washed with 20 column volume wash buffer. The flow-through, first (W_1) and last (W_L) wash fractions were collected for running on SDS PAGE, followed by elution of the column to check for CRM1:BECN1 interaction.

A previous study suggested that binding of RanGTP to CRM1 facilitates binding of NES containing proteins (Monecke *et al.*, 2013). This was tested by Ni-NTA affinity pull downs using an yeast homolog of human Ran GSP(1-179)Q71L-His$_6$-GTP. In this experiment 1:3:10 molar ratio of CRM1: GSP(1-179)Q71L-His$_6$-GTP: BECN1 CCD/CCD BARAD or 1:3:5 molar ratio of CRM1: GSP(1-179) Q71L-His$_6$-GTP: BECN1 FL was used. 60 µM of purified GSP(1-179)Q71L-His$_6$-GTP was loaded on 25µl Ni-NTA resin followed by 5 CV washes with wash buffer (20mM HEPES pH 7.4, 110mM KOAc, 2mM MgOAc, 10%(v/v) glycerol, 2mM DTT) by centrifuging resin at 10000 g for 3 minutes. 20µM of CRM1 was allowed to bind to GSP(1-179)Q71L-His$_6$-GTP bound to the Ni-NTA resin by incubating on ice for 30 mins at 4°C.

Unbound CRM1 was washed off from the column, then either 100 µM MBP-BECN1CCD or MBP-BECN1 CCD BARAD, or 200 µM BECN1^{AFM}, were added to the CRM1: GSP(1-179)Q71L-His$_6$-GTP complex bound to the Ni-NTA resin and incubated for 30 min to 1 hour . Unbound BECN1 NES was washed off the column using 10 CV wash buffer, then eluted using elution buffer (20 mM HEPES pH 7.4, 2 mM MgOAc, 250 mM Imidazole, 10% (v/v) glycerol, 2 mM DTT).

3.3. Results

3.3.1. Purification

The net yield of GST-CRM1 is 1 mg/L of BL21(DE3) cells used for protein expression. The theoretical molecular weight of GST-CRM1 calculated from the amino acid sequence is 148.4 kDa. The apparent molecular weight of GST-CRM1 calculated from the elution volume of (11.7 ml peak, on a Superdex 200 10/300 column) the major peak (Figure 3.6) is 210 kDa. The apparent mass of GST-CRM1 is likely larger than expected because the C-terminal helix (residues 1031-1057) results in an extended conformation.

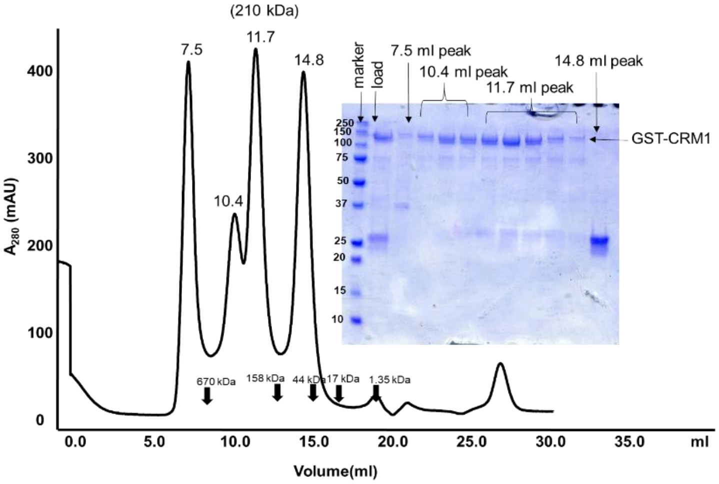


Figure 3.6: Size Exclusion chromatogram and the corresponding SDS-PAGE of GST-CRM1. The elution positions of SEC standards are indicated by arrows.

The net yield of CRM1 is 0.5 mg/L of BL21(DE3) cells used for protein expression. The theoretical molecular weight of CRM1 calculated from its amino acid sequence is 123.4 kDa. The apparent molecular weight of untagged CRM1 calculated from the elution volume (13.3 ml, on Superdex 200 10/300 column) of the major peak (Figure 3.7) is 135 kDa. Like GST-CRM1, the size of untagged CRM1 is also larger than expected likely due to the C-terminal helix forming an extended conformation.

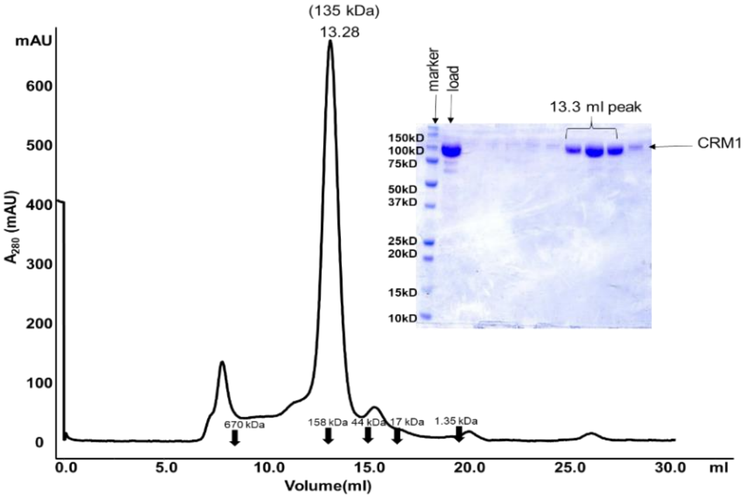


Figure 3.7: Size Exclusion chromatogram and the corresponding SDS-PAGE of CRM1. The elution positions of SEC standards are indicated by arrows.

The net yield of GSP1(1-179)Q71L-His$_6$ is 4.4 mg/L of BL21(DE3) cells used for protein expression. The theoretical molecular weight of GSP1(1-179)Q71L-His$_6$ calculated from the amino acid sequence is 23.7 kDa. The apparent molecular weight of GSP1(1-179)Q71L-His$_6$ calculated from the elution volume (Figure 3.8) of (12.1 ml on a Superdex 75 increase 10/300 column) 31.2 kDa.

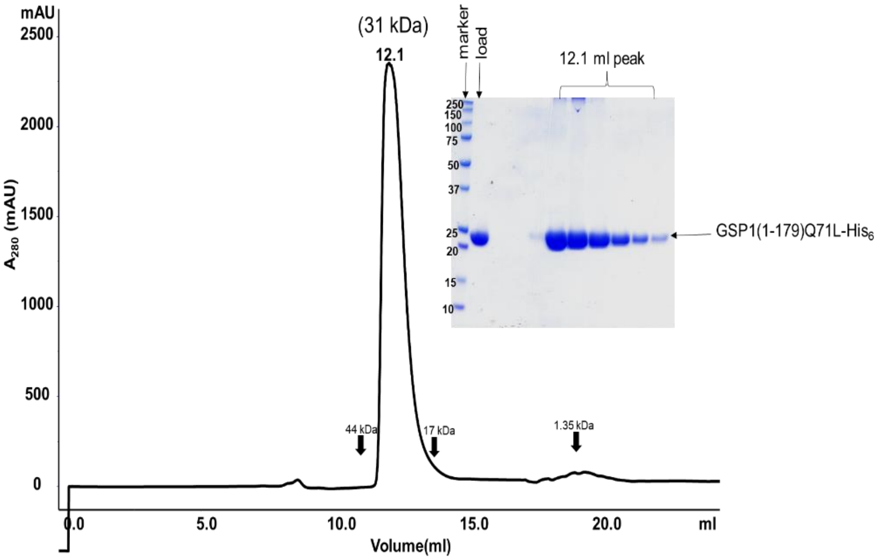


Figure 3.8: Size Exclusion chromatogram and the corresponding SDS-PAGE of GSP1(1-179)Q71L-His$_6$. The elution positions of SEC standards are indicated by arrows.

The net yield of the MBP-BECN1 CCD is 28 mg/L of BL21(DE3) cells used for protein expression. The theoretical molecular weight of MBP-BECN1CCD dimer as per the amino acid sequence is 52 kDa. The apparent molecular weight of MBP-BECN1CCD calculated from the elution volume of the peak (12.4 ml, on a Superdex 200 10/300 column) (Figure 3.9) is 252 kDa. The size of MBP-BECN1 CCD is 4.7 times larger than expected likely because it forms an antiparallel homodimer with a rod-shaped conformation, allowing the protein to have an extended conformation.

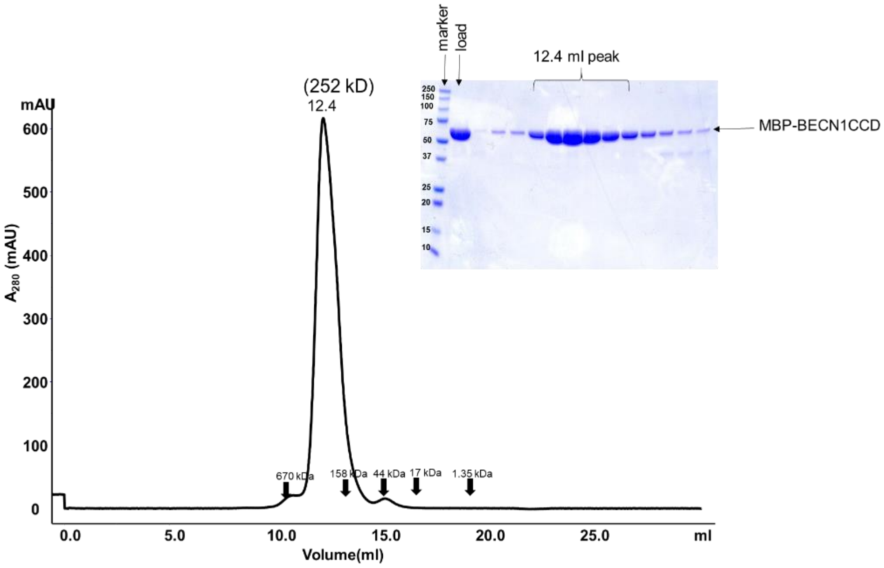


Figure 3.9: Size Exclusion chromatogram and the corresponding SDS-PAGE of MBP-BECN1 CCD. The elution positions of SEC standards are indicated by arrows.

The net yield of the purified MBP-BECN1 CCD-BARAD^{AFM} is 4.2 mg/L of ArcticExpress(DE3) cells used for protein expression. The theoretical molecular weight of MBP-BECN1 CCD BARAD^{AFM} dimer as per the amino acid sequence is 151 kDa. The apparent molecular weight of MBP-BECN1 CCD-BARAD^{AFM} calculated from the elution volume of the peak (Figure 3.10) (12.5 ml, on Superdex 200 10/300 column) is 305 kDa. Like the BECN1 CCD, MBP-BECN1 CCD BARAD^{AFM} runs larger in size than expected as BECN1 CCD BARAD^{AFM} forms an antiparallel homodimer with a rod-shaped, extended conformation.

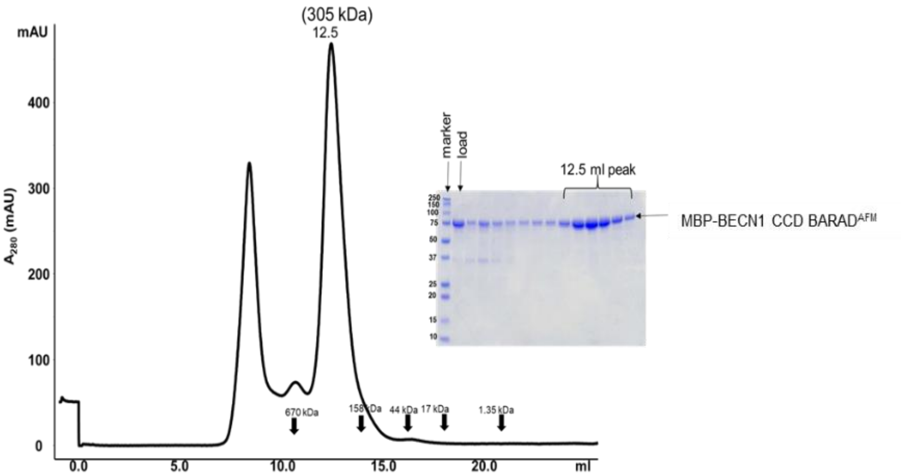


Figure 3.10: Size Exclusion chromatogram and the corresponding SDS-PAGE of MBP-BECN1 CCD-BARAD^{AFM}. The elution positions of SEC standards are indicated by arrows.

3.3.2. Pull down assay results

As expected the GST-CRM1 binds to the glutathione sepharose resin. However, neither the BECN1 CCD nor the BECN1 CCD BARAD appeared to bind to CRM1, as it was present in the flow through and wash fractions, rather than the eluate (Figure 3.11).

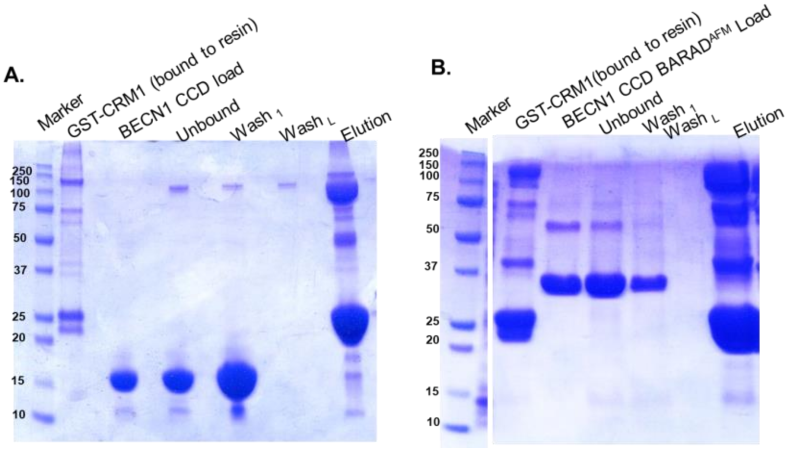
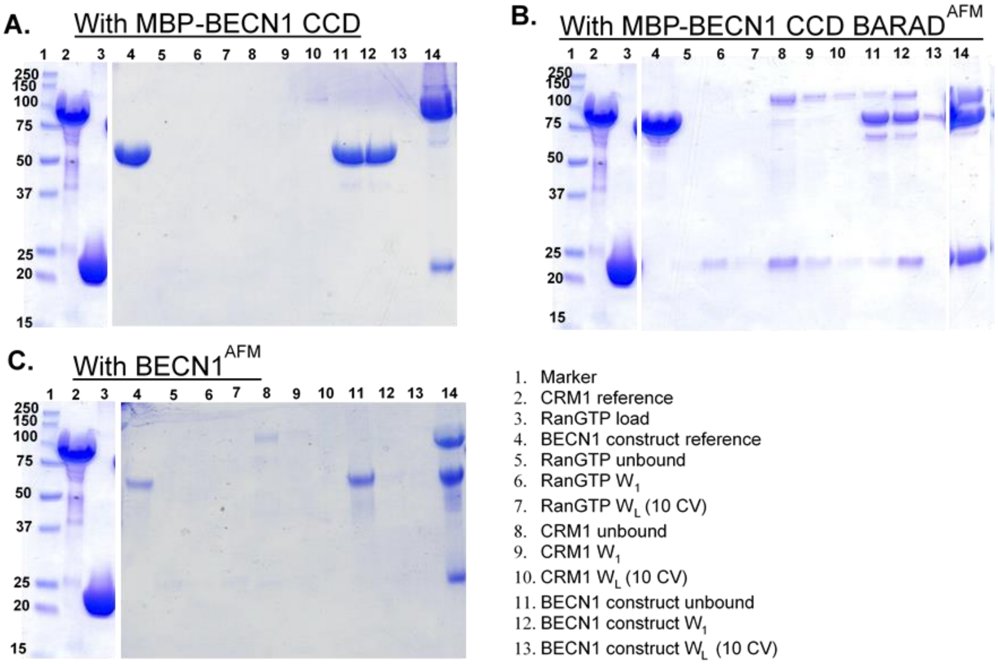


Figure 3.11: GST Pull down assay. Pure GST-CRM1 is loaded on to glutathione sepharose resin and passed through either A) BECN1 CCD or B) BECN1 CCD BARAD^{AFM}

Purified GTP loaded GSP(1-179, Q71L) was allowed to bind to Ni-NTA resin, followed by addition of CRM1 and either MBP-BECN1 CCD or MBP-BECN1 CCD BARAD^{AFM} or BECN1^{AFM} . MBP-BECN1 CCD BARAD^{AFM} and BECN1^{AFM} were found to interact with

CRM1, as it was present in the eluate from Ni-NTA resin (Figure 3.12 B, C). Surprisingly, MBP-BECN1 CCD did not appear to bind to CRM1 as it was present in the flow through and wash fractions, rather than the eluate (Figure 3.12 A)



1. Marker
2. CRM1 reference
3. RanGTP load
4. BECN1 construct reference
5. RanGTP unbound
6. RanGTP W_1
7. RanGTP W_L (10 CV)
8. CRM1 unbound
9. CRM1 W_1
10. CRM1 W_L (10 CV)
11. BECN1 construct unbound
12. BECN1 construct W_1
13. BECN1 construct W_L (10 CV)
14. Elution

Figure 3.12: Ni-NTA Pull down assay. Pure GSP(1-179)Q71L-His$_6$ was loaded in Ni-NTA resin, followed by purified CRM1, then either, A) BECN1 CCD or B) BECN1 CCD BARAD^{AFM} ,C) BECN1^{AFM}

3.4. Discussions and conclusions

Neither of two NES-containing BECN1 constructs, the BECN1 CCD or the CCD-BARAD^{AFM} were found to interact with GST-CRM1 by performing GST pull downs, which is indicative of the fact that the either the BECN1 NES was unable to interact with CRM1 or the interaction is too unstable or weak to determine by pull down assay. However, in the presence of GSP(1-179)Q71L-His$_6$-GTP, CRM1 does bind MBP-BECN1CCD-BARAD^{AFM} , BECN1^{AFM}, which indicates that RanGTP is critical for the CRM1:BECN1 interaction (Monecke *et al.*, 2013).

Strikingly the MBP-BECN1 CCD did not show any interaction with CRM1. Within the anti-parallel BECN1 CCD homodimer, the NES packs with the overlap helix(OH), suggesting that this packing impedes interaction with CRM1. Notably, our previous biophysical studies on the BECN1 CCD and overlap helix(OH) (Glover *et al.*, 2017) showed that self-association of BECN1 CCD is ~1.5 times stronger than a BECN1 CCD$^{\mathrm{TETRAD}}$, wherein four hydrophobic OH interface residues are mutated to alanine, and reduces helicity by ~9.5 residues, with an equivalent increase in coil content. This indicates that packing of the OH with the NES stabilizes the anti-parallel BECN1 CCD homodimer. However, in the same study we also showed that the BECN1 OH preferentially packs against the BARAD in FL BECN1 (Glover *et al.*, 2017), suggesting that in FL BECN1, the NES is accessible to CRM1.

CHAPTER 4: INVESTIGATING THE ROLE OF BECN1 INVARIANT CYSTEINES IN AUTOPHAGY

4.1. Introduction

Many autophagy proteins, including BECN1, have intrinsically disordered regions (IDRs) indicative of an important role for IDRs in this pathway(Mei *et al.*, 2013). IDRs do not form stable, ordered secondary structures like helices and β-strands, or tertiary structure. IDRs usually have lower sequence complexity and a higher fraction of polar or charged residues rather than hydrophobic residues, which prevents formation of a well-packed hydrophobic core(van der Lee *et al.*, 2014). IDRs also have low sequence conservation(van der Lee *et al.*, 2014), yet may contain evolutionarily invariant sequence motifs important for protein conformation or function. The sequence and structural flexibility of IDRs is thought to enable diverse and multiple interactions with high specificity and reversibility enabling them to regulate cell signaling pathways(Diella *et al.*, 2008, Jones & Thornton, 1996, Wright & Dyson, 1999, Dunker & Obradovic, 2001, Dyson & Wright, 2005). Many IDRs fold upon binding to binding partners or ligands(Habchi *et al.*, 2014).

The BECN1 IDR comprises residues 1-140(Mei, Glover, *et al.*, 2016b). Like typical IDRs, the human BECN1 IDR lacks a well-packed hydrophobic core and is rich in disorder promoting residues, containing 51% polar and charged residues, 6% glycine and 6% proline(Mei, Glover, *et al.*, 2016b). In higher eukaryotes, the BECN1 IDR includes a BCL2 homology 3 domain (BH3D) comprised in humans of residues 105-130 (Mei, Glover, *et al.*, 2016b), which undergo a dramatic binding associated disorder-to-helix transition upon binding to BCL2 homologs (Sinha, Colbert, Becker, Wei & Levin, 2008, Sinha & Levine, 2008, Oberstein *et al.*, 2007, Ku *et al.*, 2008, Su *et al.*, 2014). Further, it has been proposed that the BECN1 IDR may include additional α-MORFs besides the BH3D that are responsible for binding to other protein partners(Glover *et al.*, 2016). While all BECN1 homologs bear an IDR at their N-terminus, the BECN1 IDR sequence is very poorly conserved(Mei, Glover, *et al.*, 2016b), except for two CxxC motifs (^{18}CxxC^{21} and ^{137}CxxC^{140} in humans), which are invariant across all eukaryotes. The second of these CxxC motifs delineates the end of the IDR, as it is followed by highly conserved, largely structured regions of BECN1. Despite being invariant, the role of the BECN1 CxxC motifs has not yet been investigated to date.

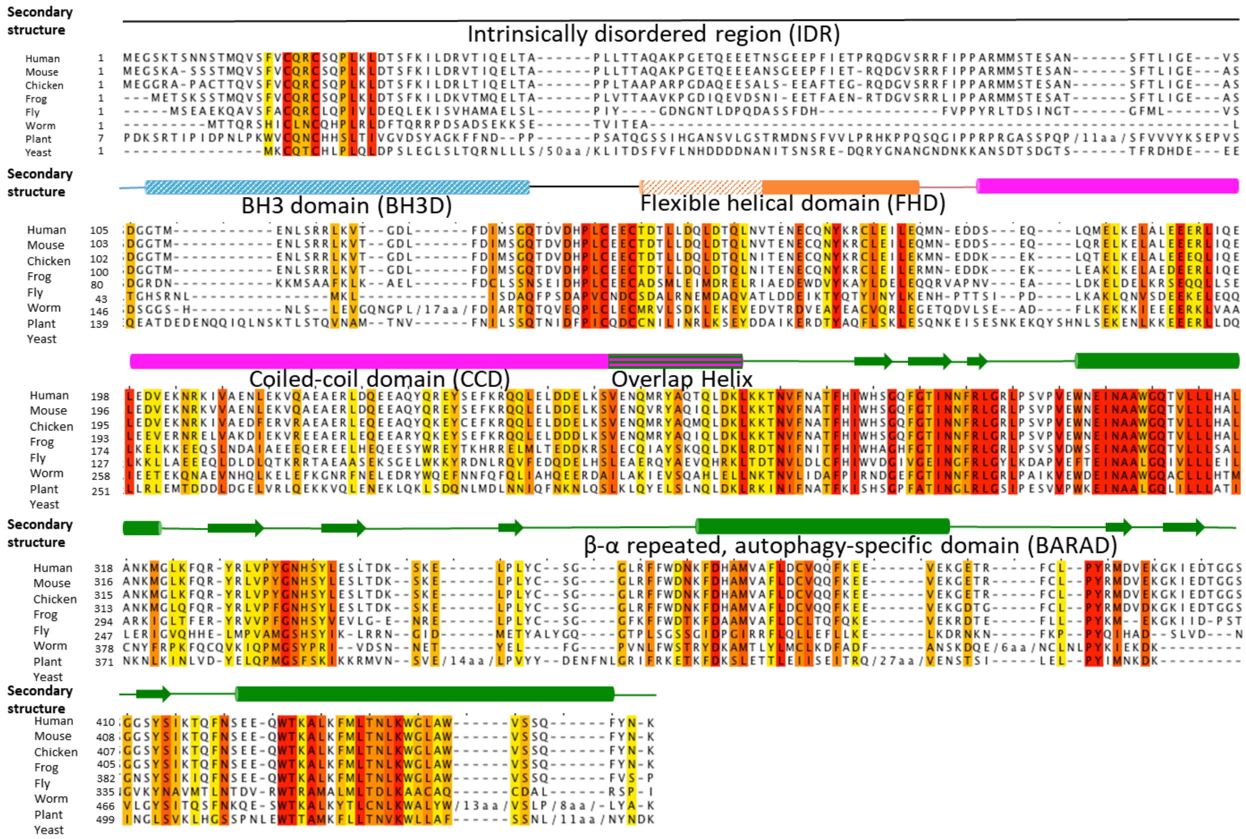


Figure 4.1: Sequence alignment of full length BECN1 orthologs. Increasing background color intensity from yellow to red indicates increased sequence conservation with red corresponding to the invariant residues. Experimentally determined secondary structures are shown above the sequence alignment, with helices, strands and coil represented as cylinders, arrows, and lines. The secondary structures are color-coded by domain boundaries, with diagonal hashes indicating parts of domains known to undergo disorder-to-order transitions upon binding and horizontal hashes indicating regions that form mutually exclusive helices in different BECN1 states.

CxxC motifs are known to perform essential functions, including binding of various metal ions. Bioinformatics analyses of the human genome suggests that Zn^{2+} is bound by up to 3000 proteins, which corresponds to ~10% of all encoded proteins (Andreini *et al.*, 2006). Consistent with that, a similar fraction of Zn^{2+}-binding proteins are found in human protein structures deposited in structural databases (Kochanczyk *et al.*, 2015). The most abundant class of zinc-binding proteins in humans are proteins containing zinc-fingers, with Cys_4 and Cys_2His_2 being the most common types of coordination environment (Andreini *et al.*, 2006). Zn^{2+}-cysteine complexes are critical mediators of protein structure, catalysts and regulation (Pace & Weerapana, 2014).

Therefore, we used Inductively Coupled Plasma Mass Spectrometry (ICP-MS) to determine which metals, if any, are bound by the CxxC motifs of BECN1. We investigated the impact of these CxxC motif on the BECN1 IDR conformation using size-exclusion chromatography in tandem with small angle X-ray scattering (SEC-SAXS) and on BECN1 IDR secondary structure using circular dichroism (CD) spectroscopy. We used Multi Angle Light Scattering (MALS) to estimate the molecular mass of BECN1 IDR. We used Isothermal Titration Calorimetry (ITC) to assess if the CxxC motifs play a role in BECN1 IDR dimerization and in interaction of the BECN1 IDR interaction with a viral BCL2 homolog, M11. Lastly, used cellular assays evaluating the change in the number of autophagosomes per cell to assess role of the BECN1 IDR and the CxxC motifs in regulating cellular autophagy levels.

4.2. Materials and methods

4.2.1. Plasmid preparation

A human BECN1 IDR construct, analogous to the construct previously used to investigate the structure of the BECN1 IDR(Lee et al., 2016) was created by cloning human

BECN1 residues 1-150 between *BamHI* and *HindIII* restriction enzyme sites of the pET DUET1 expression vector, to enable bacterial expression as fusion protein, wherein a hexahistidine (His_6) purification tag followed by a short linker sequence (MGSSHHHHHHSQDP) (Lee et al., 2016) preceded the IDR. A tyrosine was added at the C-terminus of this construct to facilitate quantification of protein concentration by UV spectroscopy. This His_6-BECN1(1-150)Y construct was used for site-directed mutagenesis to create CxxC motif double mutants, His_6-BECN1(1-150)Y^{CysDM1} (C18S+C21S) or His_6-BECN1(1-150)Y^{CysDM2} (C137S+C140S); and the tetrad mutant, His_6-BECN1(1-150)Y^{CysTM}.

In addition to the His_6-tagged BECN1 IDR constructs, human full length (FL), wild-type (WT) BECN1 was cloned between the *BamHI* and *NotI* restriction enzyme sites of the bacterial expression vector pMBP parallel 1(Sheffield *et al.*, 1999). The MBP-tag with the TEV protease site and a linker comprising residues GAMDP precede BECN1. All FL BECN1 constructs used in this research included the aromatic finger mutations(AFM), F359D, F360D and W361D to improve protein solubility(Glover *et al.*, 2017) by Quick Change II site directed mutagenesis (Agilent Technologies). Site directed mutagenesis of the MBP-BECN1 construct was used to create CxxC motif double mutants, MBP-$BECN1^{CysDM1}$ (C18S+C21S) or MBP-$BECN1^{CysDM2}$ (C137S+C140S) and the tetrad mutant, MBP-$BECN1^{CysTM}$ (C18S+C21S+C137S+C140S). All the full-length constructs were prepared by Dr. Yue Li and are described in her dissertation. In addition to the FL constructs, we also used a MBP-BECN1 (Δ31-123) construct that did not have the AFM, as it was prepared before it was determined that AFM was required for maintaining solubility of BARAD containing proteins.

During the course of this research several BECN1 IDR truncations were prepared which allowed us to determine the minimal region of IDR that could be expressed stably and purified in quantities sufficient for biophysical studies. This included an MBP-BECN1(1-140) construct created by Dr. Sangita Sinha, by cloning human BECN1 residues 1-140 between *BamHI* and *HindIII* restriction enzyme sites of the pMBP parallel 1 expression vector to enable expression of MBP-BECN1(1-140). A Strep-SUMO tagged BECN1 IDR (1-143) construct was prepared by cloning human BECN1 residues 1-143 between *BamHI* and *HindIII* restriction enzyme sites of the pET28a Strep-SUMO expression vector, in which the linker between Strep tag II and SUMO was shortened to "SSAS" to allow minimal flexibility. The corresponding cysteine tetrad mutant Strep-SUMO-BECN1(1-143)$^{\mathrm{CysTM}}$ was also prepared as mentioned above. Lastly, a human BECN1 IDR construct, analogous to the construct previously used to investigate the structure of the BECN1 IDR (Lee *et al.*, 2016) was created by cloning human BECN1 residues 1-150 between *BamHI* and *HindIII* restriction enzyme sites of the pMBP parallel 1 expression vector to enable expression of MBP-BECN1(1-150). A tyrosine was added at the C-terminus of this construct to facilitate quantification of protein concentration by UV spectroscopy. Various CxxC mutants were produced by site-directed mutagenesis. Constructs created were: MBP-BECN1(1-150)Y; the CxxC double mutants, MBP-BECN1(1-150)Y$^{\mathrm{CysDM1}}$, MBP-BECN1(1-150)Y$^{\mathrm{CysDM2}}$; and the CxxC tetrad mutant (MBP-BECN1(1-150)Y$^{\mathrm{CysTM}}$).

Lastly, for cellular studies, a pcDNA3.1 mammalian vector encoding human WT, FL BECN1 (Sinha, Colbert, Becker, Wei & Levin, 2008) was mutated by site directed mutagenesis to generate CxxC double mutants FL$^{\mathrm{DM1}}$ or FL$^{\mathrm{DM2}}$; the tetrad mutant FL$^{\mathrm{TM}}$ as well as a FL$^{(\Delta 31\text{-}123)}$ mutant.

Table 4.1: List of proteins used in this project

Serial no.	**Label**	**Description**	**Theoretical Monomer Molecular Mass (kDa)**	**Purpose**
1	His$_6$-BECN1(1-150)Y	BECN1 residues (1-150, C-terminal tyrosine) H$_6$ fusion protein	18.3	ICP-MS, SAXS, MALS, CD
2	His$_6$-BECN1(1-150)Y^{CysTM}	BECN1 residues (1-150, C18S, C21S, C137S C140S, C-terminal tyrosine) H$_6$ fusion protein	18.3	ICP-MS, SAXS, MALS, CD
3	His$_6$-BECN1(1-150)Y^{CysDM1}	BECN1 residues (1-150, C-terminal tyrosine, C18S, C21S) H$_6$ fusion protein	18.3	ICP-MS
4	His$_6$-BECN1(1-150)Y^{CysDM2}	BECN1 residues (1-150, C-terminal tyrosine, C137S, C140S) H$_6$ fusion protein	18.3	ICP-MS
5	MBP-BECN1	BECN1 full length (1-450, F359D, F360D,W361D) MBP fusion protein	93.3	ICP-MS
6	MBP-BECN1^{CysTM}	BECN1 full length (1-450, C18S C21S, C137SC140S, F359D, F360D,W361D) MBP fusion protein	93.3	ICP-MS
7	MBP-BECN1^{CysDM1}	BECN1 full length (1-450, C18S C21S, F359D, F360D,W361D) MBP fusion protein	93.3	ICP-MS
8	MBP-BECN1^{CysDM2}	BECN1 full length (1-450, C137S C140S, F359D, F360D,W361D) MBP fusion protein	93.3	ICP-MS
9	MBP-BECN1(1-150)Y	BECN1 residues (1-150, C-terminal tyrosine) fusion protein	60.4	ICP-MS, ITC
10	MBP-BECN1(1-150)Y^{CysTM}	BECN1 residues (1-150, C18S, C21S, C137S C140S, C-terminal tyrosine) MBP fusion protein	60.4	ICP-MS, ITC
11	MBP-BECN1(1-150)Y^{CysDM1}	BECN1 residues (1-150, C-terminal tyrosine, C18S, C21S) MBP fusion protein	60.4	ICP-MS
12	MBP-BECN1(1-150)Y^{CysDM2}	BECN1 residues (1-150, C-terminal tyrosine, C137S, C140S) MBP fusion protein	60.4	ICP-MS
13	BECN1(1-150)Y	BECN1 residues (1-150, C-terminal tyrosine)	60.4	ICP-MS, SAXS, MALS, CD
14	BECN1(1-150)Y^{CysTM}	BECN1 residues (1-150, C18S, C21S, C137S C140S, C-terminal tyrosine)	60.4	ICP-MS, SAXS, MALS, CD
15	MBP-BECN1(Δ31-123)	BECN1 residues 31 to 123 deleted from the full-length construct, MBP fusion protein	83.2	ICP-MS
16	MBP-BECN1(1-140)	BECN1 residues 1-140, fusion protein, MBP fusion protein	59.3	Attempted purification
17	Strep-SUMO-BECN1(1-143)	BECN1 residues (1-143), Strep-SUMO fusion protein	28.8	Attempted purification
18	Strep-SUMO-BECN1(1-143)^{CysTM}	BECN1 residues (1-143, C18S, C21S, C137S C140S), Strep-SUMO fusion protein	28.8	Attempted purification

4.2.2. Protein expression and purification

Bacterial expression of each construct was performed by transforming E. coli Arctic Express cells with one of the expression vectors described above, and growing the transformed cells in LB medium with 100µg/ml of ampicillin at 30°C. At each stage of purification, protein quality was evaluated by SDS-PAGE stained with Coomassie Blue. In each case, the final purified protein was estimated to be >90% pure by Coomassie Blue-stained SDS PAGE. All His_6-BECN1(1-150)Y cultures were grown to an A_{600} of ~0.8 before equilibrating the temperature to 13°C, and addition of 0.5 mM isopropyl thio-β-D-galactoside (IPTG) to allow overnight protein expression. Cells were harvested by centrifugation at 4000 g at 4°C and stored at -80°C until lysis before purification. Frozen cell pellets for each His_6-BECN1(1-150)Y protein were resuspended in 30 ml lysis buffer (20 mM Tris pH 8, 150 mM NaCl, 25 mM Imidazole, 0.1mM $Zn(Ac)_2$ and 5 mM βME) in presence of 2 protease inhibitor cocktail tablet/L cells. The cells were lysed by sonication and the cell lysate was clarified by centrifugation at 20000 g at 4°C prior to purification. Unlike the MBP-tagged BECN1(1-150)Y constructs, EDTA was not used during purification since EDTA may chelate out Ni^{2+} from the Ni-NTA affinity resin. So, it was critical to perform the experiment quicker than the MBP-tagged constructs, and so only 3L cells were purified each time instead of the 6L purifications performed for MBP-BECN1(1-150)Y constructs, thus minimizing protein degradation during the affinity stage. The proteins were purified from clarified crude lysate using Ni-NTA affinity chromatography, by washing with 20 CV of lysis buffer and 10 CV of Wash buffer (Wash buffer : 20 mM Tris pH 8, 150 mM NaCl and 5 mM βME) prior to elution (Elution buffer: Wash buffer + 250 mM Imidazole). Each of the eluted proteins was further purified by anion exchange chromatography using a 5/50 GL MonoQ column (GE Healthcare) with Mono Q buffer A comprising of 20 mM Tris pH 7.5, 5 mM βME and Mono Q buffer B: Buffer A + 1 M NaCl. Lastly, each protein was purified to homogeneity by size exclusion chromatography (SEC), using a 10/300 Superdex 75 Increase column (GE Healthcare) in a buffer comprising of 20 mM Tris pH 8, 150 mM NaCl). Each of the pure proteins was concentrated using a 3,000 Da MWCO Amicon Ultra- 0.5 ml centrifugal concentrator (EMD Millipore, Billerica, MA) and stored at -80°C.

All FL BECN1 expressions were done by Dr. Yue Li and detailed expression and purification protocols are described in her dissertation. The purification protocols for MBP-BECN1 constructs (as described in Dr. Yue Li's dissertation) were also followed for MBP-BECN1(1-150)Y constructs. The only difference in purifications between MBP-BECN1 constructs and MBP-BECN1(1-150)Y constructs is, the FL constructs were not concentrated before loading on SEC, to prevent any aggregation. Cell pellets for each MBP-tagged BECN1 protein were resuspended in 30 ml Lysis Buffer (50 mM Tris pH 8.0, 300 mM NaCl, 1 mM EDTA, 2 mM DTT) in presence of 2 protease inhibitor cocktail tablet/L cells and loaded onto an amylose affinity column. The amylose affinity column was then washed with 20-column volume of Lysis Buffer, followed with a 10-column volume wash with Wash Buffer 1 (50 mM Tris pH 8.0, 300 mM NaCl), 5-column volume wash with Wash Buffer 2 (50 mM Tris pH 8.0, 300 mM NaCl, 0.1mM $Zn(Ac)_2$) and finally another 10-column volume wash by Wash Buffer 1 before elution with Elution Buffer (50 mM Tris pH 8.0, 300 mM NaCl, 2mM DTT, 25 mM Maltose). The EDTA in the lysis buffer and the initial wash was essential to prevent protein degradation, therefore, Zn^{2+} that may have been stripped out during purification was replaced in the later washes. SEC using a Superdex 200 column (GE Healthcare), in a buffer comprised of 50 mM Tris pH 8, 300 mM NaCl, was used to purify proteins to homogeneity. The fusion protein was concentrated in a 30,000 Da MWCO Amicon Ultra- 0.5 ml centrifugal concentrator (EMD Millipore, Billercia, MA).

MBP-BECN1(1-150)Y cultures were grown to an A_{600} of ~0.8, before equilibrating the temperature to 13°C, followed by addition of 0.5 mM IPTG to induce protein expression overnight. A purification protocol similar to that for FL proteins was used to purify untagged BECN1(1-150)Y or BECN1(1-150)Y^{CysTM} constructs, except that the washes were followed by another wash with a DDT-containing buffer (50 mM Tris pH 8.0, 300 mM NaCl, 2 mM DTT) followed by overnight TEV protease cleavage. The cleaved protein was washed off using 50 mM Tris pH 8.0, 300 mM NaCl, then GST-TEV-His_6 removed by incubating the protein for 30 mins at 4°C by binding to GST resin, and any MBP contaminant removed by binding over fresh amylose resin. The protein that did not bind to the amylose resin was concentrated in a 3,000 Da MWCO Amicon Ultra- 15 ml centrifugal concentrator (EMD Millipore, Billercia, MA), then purified further using Superdex 200 10/300 GL tandem Superdex 75 10/300 GL. The SEC elutions were further cleaned up by incubating for 30 mins at 4°C, followed by passing the fractions through a fresh GST resin to remove any excess GST, GST-TEV-His_6 and through a fresh Amylose resin to remove any excess MBP. So, the SEC fractions containing BECN1(1-150)Y or BECN1(1-150)Y^{CysTM} were pooled and purified to >90% homogeneity. The final product obtained was run on Superdex 75 increase 10/300 column and the corresponding SDS PAGE gel was run to verify for the removal of contaminants.

MBP-BECN1(Δ31-123) cultures were grown to an A_{600} of ~0.6, while all other MBP-BECN1(1-150)Y cultures were grown to an A_{600} of ~0.8, before equilibrating the temperature to 13°C, followed by addition of 0.5 mM IPTG to induce protein expression overnight.

MBP-BECN1(Δ31-123) and MBP-BECN1(1-140) were purified without any exogenous zinc supplementation. The frozen cell pellets for each of these above constructs were resuspended in 50 ml lysis buffer (25 mM Tris pH 8, 150 mM NaCl, 2 mM βME) in presence of 1 protease inhibitor cocktail tablet/L cells. The resuspended cells were further lysed by emulsification. The cell lysate was clarified by centrifugation at 20,000 g at 4°C prior to purification. The supernatant was loaded on to a 10 ml amylose resin followed by washing with 10 CV of the lysis buffer before collecting eluate with elution buffer (Lysis buffer + 20 mM Maltose). Each of these two constructs were further purified to homogeneity using Superdex 200 10/300 (GE Healthcare) using the same lysis buffer and then pooled and concentrated in a 30,000 Da MWCO Amicon Ultra- 0.5 ml centrifugal concentrator (EMD Millipore, Billercia, MA).

In addition to the MBP tagged constructs, Strep-SUMO BECN1(1-143) constructs were grown in presence of 15µg/ml of Kanamycin. The cultures were grown to an A_{600} of ~0.8, before equilibrating the temperature to 13°C, followed by addition of 0.5 mM IPTG to induce protein expression overnight. Cells were harvested by centrifugation at 4000 g at 4°C and stored at -80°C until purification. Frozen cells were thawed and lysed by emulsification, and the cell lysate is clarified by centrifugation at 20,000 g at 4°C prior to purification.

The frozen cell pellets for each of the two constructs above were resuspended in 50 ml lysis buffer (25 mM Tris pH 8, 150 mM NaCl, 2 mM βME) in presence of 1 protease inhibitor cocktail tablet/L cells. The resuspended cells were further lysed by emulsification and the cell lysate. The cell lysate was clarified by centrifugation at 20,000 g at 4°C prior to purification. The supernatant and then loaded on to a 10 ml Strep-Tactin resin (IBA Lifesciences) followed by washing with 10 CV of the lysis buffer before eluting with elution buffer (Lysis buffer + 2.5 mM desthiobiotin) (IBA Lifesciences). Each protein was further purified to homogeneity using Superdex 200 10/300 (GE Healthcare) using the same lysis buffer and then pooled and concentrated in a 30,000 Da MWCO Amicon Ultra- 0.5 ml centrifugal concentrator (EMD Millipore, Billercia, MA).

4.2.3. Element analyses

Quantitative ICP-MS, performed at the Redox Biology Center, University of Nebraska Lincoln, was used to identify and quantify concentrations of elements found in different purified BECN1 protein samples. FL BECN1 samples tested included MBP-BECN1, the double mutants MBP-BECN1^{CysDM1}, MBP-BECN1 ^{CysDM2} and MBP-BECN1^{CysTM}. Additionally, minimal IDR constructs were tested, which includes His$_6$-BECN1(1-150)Y, MBP-BECN1(1-150)Y, the double mutants His$_6$-BECN1(1-150)Y^{CysDM1} and His$_6$-BECN1(1-150)Y^{CysDM2}, MBP-BECN1(1-150)Y^{CysDM1} and MBP-BECN1(1-150)Y^{CysDM2} or tetrad mutant His$_6$-BECN1(1-150)Y^{CysTM} or MBP-BECN1(1-150)Y^{CysTM}. MBP-BECN1(Δ31-123) was also tested. For each of these proteins samples, the concentrations of different elements were buffer subtracted using the corresponding buffers used for SEC, which includes 50 mM Tris pH 8, 300 mM NaCl for MBP-tagged full length and minimal IDR constructs, 20mM Tris pH 8, 150 mM NaCl for all the His$_6$-tagged constructs and 25 mM Tris pH 8, 150 mM NaCl, 2 mM βME for MBP-BECN1(Δ31-123) constructs. Additionally, Zn^{2+} chelation sensitivity of the BECN1 CxxC motifs was tested by treating 12 µM MBP-BECN1 with 240 µM TPEN (N,N,N′,N′-tetrakis(2-pyridinylmethyl)-1,2-ethanediamine), a high-affinity chelating agent specific for transition metals. Likewise, the sensitivity of the BECN1 CxxC motifs to reducing agents was tested by treating 12 µM MBP-BECN1 with either 50 mM DTT or 2.5 mM TCEP (tris(2-carboxyethyl)phosphine), a highly stable reducing agent. The cumulative sensitivity of the CxxC motifs to 240 µM TPEN + 2.5 mM TCEP or 240 µM TPEN + 50 mM DTT was also tested. After addition of TPEN or / and reducing agent, each protein sample was vigorously stirred for 2 hrs at 4°C. These samples were then dialyzed overnight against 50 mM Tris pH 8, 300 mM NaCl buffer for TPEN-treated samples, or against 50 mM Tris pH 8, 300 mM NaCl, 50 mM DTT buffer for TCEP, DTT, TPEN + DTT, TPEN + TCEP, samples prior to ICP-MS evaluation. For each of these treated samples, the concentrations of different elements were buffer subtracted using the corresponding dialysis buffers. The protein samples in solution were digested by metal grade 33% v/v HNO_3 which caused precipitation of the protein and subsequent release of all the attached metal ions to the supernatant, which was then quantitatively screened by ICP-MS for 20 elements: Li, B, Na, K, S, P, Mg, Ca, Mn, Fe, Ni, Co, Cu, Zn, As,V,Cr, Se, Mo and Cd, with Ga used as an internal standard.

4.2.4. CD Spectroscopy

CD spectra were recorded by continuous scanning from 190 nm to 260 nm at 4°C and 20°C using a Jasco J-815 spectrophotometer equipped with a Peltier thermoelectric temperature control. 300 µL of 4-7 µM protein in CD buffer comprising 5 mM sodium phosphate monobasic, 5 mM sodium phosphate dibasic, pH 7.4, 100 mM $(NH_4)_2SO_4$, was placed in a quartz cell of 0.1 cm path length). The His$_6$-BECN1(1-150)Y or His$_6$-BECN1(1-150)Y ^{CysTM} and BECN1(1-150)Y or BECN1(1-150)Y^{CysTM} protein fragments were premixed with 2,2,2-trifluroethanol (TFE) at different v/v ratios, 0%, 15%, 25% and 40%, and incubated on ice for 30 minutes. Secondary structure content was estimated using CONTIN & CDSSTR programs with Reference set SDP48 from the CD Pro software (Sreerama *et al.*, 2000, Sreerama & Woody, 2000, Sreerama *et al.*, 2001).

4.2.5. SEC–MALS-SAXS data collection and analyses

The shape, size and oligomerization of purified His$_6$-BECN1(1-150)Y or His$_6$-BECN1(1-150)Y^{CysTM} and BECN1(1-150)Y or BECN1(1-150)Y^{CysTM} was investigated using Multi Angle Light Scattering (MALS) and Small Angle X-ray Scattering (SEC-SAXS) in-line with Size Exclusion Chromatography at the BioCAT beamline at Sector 18 of the Advanced Photon

Source (Argonne, IL). BECN1(1-150)Y, His_6-BECN1(1-150)Y and His_6-BECN1(1-150)Y^{CysTM} were loaded at a concentration of 5 mg ml^{-1} on a 10/300 GL Superdex 75 increase column at a flow rate of 0.6 ml min^{-1} , BECN1(1-150)Y^{CysTM} construct was loaded at a same concentration and at 0.4 ml min^{-1} , using an Infinity II 1260 HPLC (Agilent) system. The SEC eluate was passed through a UV monitor followed by DAWN HELIOS II (17 angle MALS plus one DLS channel) and Optilab T-rEX (refractive index) detectors (Wyatt), and finally into the SAXS flow cell, a 1.0 mm ID quartz capillary with 50 µm walls using coflow sample geometry (Kirby *et al.*, 2016) to prevent radiation damage. The MALS data was analyzed using the Astra 7 software (Wyatt Technology). Data processing included initial baseline subtraction, corrected for scattering to the 90° light scattering (LS) detector (i.e. LS source 11 of the 17-angle detector). The peaks obtained using the LS, UV at 280 nm and dRI signals were aligned and the peaks defined. The MALS data were normalized using a BSA standard run on the same column using the same SEC-MALS buffer as the sample protein to enable accurate calculation of molecular mass.

Each sample was exposed to an X-ray beam of size 160 (h) x 75 (v) µm2, wavelength 1.03 Å (12 keV energy) for 0.5 s/frame, with a periodicity of 1 s/frame, and the SAXS data were recorded on a Pilatus3 X 1M detector (Dectris) set at a sample to detector distance of ~3.7 m. The range of momentum transfer (q) was accessed from ~0.004 to 0.36 Å-1. Commercially available glassy carbon plates were used for absolute intensity calibration standard for SAXS data collection (NIST SRM 3600). X-ray radiation damage was monitored by automated frame-by-frame comparison of relevant regions using CORMAP (Franke et al., 2015), which is implemented in BioXTAS RAW (version 1.6.4). The partial specific volume $\bar{\upsilon}$ (cm^3 g^{-1}) and particle contrast from sequence (10^{-10} cm^{-2}) was calculated using an online tool, ModULes For The Analysis Of Contrast Variation Data (MULCh) (Whitten et al., 2008).

Scattering data were normalized to the transmitted X-ray beam intensity and scattering from buffer was subtracted using the BioXTAS RAW software (version 1.6.4) (Hopkins *et al.*, 2017) prior to further analysis. SEC did not fully separate out the different species in solution. Therefore, Singular Value Decomposition (SVD) was used to identify the number of distinct conformational species in each elution, and Evolving Factor Analysis (EFA) (Meisburger *et al.*, 2016) as implemented in BioXTAS RAW was used to deconvolute the overlapping data into scattering profiles for each individual species. This removed the influence of each species on the other to ensure that the reported SAXS profiles and calculated results are for each independent species. Guinier fits were used to calculate the radius of gyration (R_g) and scattering at zero angle I(0), and dimensionless Kratky plots(Durand *et al.*, 2010) were used to evaluate disorder within the samples. GNOM (Svergun, 1992) was used to calculate the maximum particle size (D_{max}) and the indirect Fourier transform to obtain the pair-distance distribution function, P(r)(Glatter, 1977). The P(r) function was also used to calculate the R_g and I(0). The molecular mass was calculated using the Volume of correlation (V_c) method (Rambo & Tainer, 2013) as implemented in RAW.

For each BECN1 IDR sequence, Ensemble of Optimization 2.0 (EOM)(Bernado *et al.*, 2007, Tria *et al.*, 2015, Bernado & Svergun, 2012) in the ATSAS(Franke *et al.*, 2017) software suite (version 2.8.0) was used to generate a pool of 10,000 random models representing all possible conformers in solution, assuming P1 symmetry and random coil like protein chains. A genetic algorithm was then run 100 times to select an ensemble of conformers from the pool that best fit the experimental SAXS data. Insight into the shape and size of the proteins in solution was obtained by comparing the R_g and D_{max} distribution of each pool of randomly generated set of conformers with the corresponding selected ensemble. The conformational flexibility of the selected ensemble relative to the pool is further described by two metrics R_{flex} and R_{σ}(Tria *et al.*, 2015). R_{flex} is a metric to evaluate the flexibility of the distribution of conformers, whereas R_{σ} is a metric to evaluate the variance of the distributions of conformers (Tria *et al.*, 2015). R_{flex} ranges from 0 to 100% for completely rigid to completely flexible systems. R_{σ} approaches 1 for fully flexible systems and decreases with decreasing flexibility. When R_{flex} of the selected ensemble is smaller than that of the pool distribution, R_{σ} is usually less than 1 (Tria *et al.*, 2015). Conversely, when R_{flex} of the ensemble tends to values greater than the pool, R_{σ} is greater than 1. For the cases where R_{flex} of ensemble is significantly smaller than pool, but $R_{\sigma} > 1$, further evaluation of the data is required (Tria *et al.*, 2015).

Table 4.2: SEC-SAXS sample details for His$_6$-BECN1(1-150)Y constructs

	His$_6$-BECN1(1-150)Y	**His$_6$-BECN1(1-150)Y^{CysTM}**
Organism	*Escherichia coli*	*Escherichia coli*
Source	Agilent (230192)	Agilent (230192)
Description	Uniprot ID: Q14457, BECN1 residues 1-150 with a tyrosine at the C-terminus to facilitate UV absorption at 280 nm	Uniprot ID: Q14457, BECN1 residues 1-150 with C18S, C21S and C137S, C140S and tyrosine at the C-terminus to facilitate UV absorption at 280 nm
Extinction coefficient ε (280 nm , in M^{-1} cm^{-1})	1490	1490
Partial specific volume $\bar{\upsilon}$ (cm^3 g^{-1})	0.722	0.722
Particle contrast from sequence and solvent constituents, ($\rho_{protein}$ - $\rho_{solvent}$; 10^{-10} cm^{-2})	3.12 (12.53 – 9.40)	3.13 (12.53 – 9.40)
Molecular mass *M* from chemical composition (Da)	18343	18279
SEC-SAXS column	Superdex 75 increase 10/300 GL column	
Injection volume (µl)	250 µl	250 µl
Flow rate (ml min^{-1})	0.6	0.6
Concentration (range/values) measured and method	5 mg/ml (Bradford)	5 mg/ml (Bradford)
Solvent composition and source	20 mM Tris pH 8, 150 mM NaCl	

Table 4.3: SEC-SAXS sample details for BECN1(1-150)Y constructs

	BECN1(1-150)Y	**BECN1(1-150)Y$^{\text{CysTM}}$**
Organism	*Escherichia coli*	*Escherichia coli*
Source	Agilent (230192)	Agilent (230192)
Description	Uniprot ID: Q14457, GAMDP-BECN1 residues 1-150 with a tyrosine at the C-terminus to facilitate UV absorption at 280 nm	Uniprot ID: Q14457, GAMDP-BECN1 residues 1-150 with C18S, C21S and C137S, C140S and tyrosine at the C-terminus to facilitate UV absorption at 280 nm
Extinction coefficient ε (280 nm , in M^{-1} cm^{-1})	1490	1490
Partial specific volume $\bar{\upsilon}$ (cm^3 g^{-1})	0.725	0.725
Particle contrast from sequence and solvent constituents, (ρ_{protein} - ρ_{solvent}; 10^{-10} cm^{-2})	3.01 (12.50 – 9.40)	3.01 (12.50 – 9.40)
Molecular mass *M* from chemical composition (Da)	17174	17109
SEC-SAXS column	Superdex 75 increase 10/300 GL column	
injection volume (µl)	250 µl	250 µl
flow rate (ml min^{-1})	0.6	0.4
Concentration (range/values) measured and method	5 mg/ml (Bradford)	5 mg/ml (Bradford)
Solvent composition and source	20mM Tris pH 8, 300mM NaCl	

Table 4.4: SEC-SAXS data collection parameters

Instrument /data processing	BioCAT beamline (18-ID) (Advanced Photon Source, Argonne, IL) with Pilatus3 X 1M (Dectris) detector
Wavelength (Å)	1.0332
Beam size (µm^2)	160 (h) x 75 (v)
Beam geometry (size, sample-to-detector distance) (m)	3.694
q-measurement range (Å^{-1})	~0.004 to 0.36
Absolute scaling method	Glassy Carbon, NIST SRM 3600
Basis for normalization to constant counts	To transmitted intensity, by beamstop counter
Method for monitoring radiation damage, X-ray dose where relevant	Automated frame-by-frame comparison of relevant regions using CORMAP (Franke *et al.*, 2015) implemented in BioXTAS RAW
Exposure time, number of exposures	0.5 s/frame, with a periodicity of 1 s/frame
Sample configuration including path length and flow rate where relevant	SEC-SAXS. Separation by size using ÄKTA Pure with a Superdex 75 Increase 10/300 GL column. SAXS data measured in a 1.0 mm ID quartz capillary with a sheath-flow cell(Kirby *et al.*, 2016) and effective path length 0.542 mm.
Sample temperature (ºC)	22

Table 4.5: Software employed for SAXS data reduction, analysis, and interpretation

SAXS data reduction	BioXTAS RAW 1.6.4
Calculation of ε from sequence	ProtParam tool- Expasy
Calculation of $\Delta\bar{\rho}$ and $\bar{\upsilon}$ values from chemical composition	MULCh (Whitten *et al.*, 2008)

4.2.6. ITC Experiments to evaluate BECN1 IDR self -dissociation

ITC experiments were performed at 20°C using a Low volume Nano ITC (TA Instruments). For self-dissociation experiments, MBP-BECN1(1-150)Y and MBP-BECN1(1-150)Y^{CysTM} were each dialyzed against the ITC buffer comprising 20 mM Tris pH 8 and 150 mM NaCl. MBP-tagged proteins were used to initially as it was easier to obtain enough amounts of MBP-tagged proteins with similar molar concentrations, used for SAXS and MALS analyses. 258 µM (15 mg/ml) MBP-BECN1(1-150)Y or 243 µM (14.6 mg/ml) MBP-BECN1(1-150)Y^{CysTM} was titrated into 196 µl of ITC buffer in the cell, with a stir rate of 250 rpm. A blank profile obtained by titrating MBP into buffer were subtracted from the experimental profiles. To confirm that the presence of the MBP tag did not impact the results, the self-dissociation ITC experiments were repeated using either 116 µM (2 mg/ml) or 355 µM (6 mg/ml) MBP-cleaved BECN1(1-150)Y; and a blank profile obtained by titrating buffer into buffer was subtracted from the experimental profiles.

4.2.7. ITC Experiments to Evaluate M11 interaction with the BECN1 IDR

To investigate whether the CxxC motifs impact the interaction of M11 with BECN1, M11, MBP-BECN1(1-150)Y, and MBP-BECN1(1-150)Y^{CysTM} were co-dialyzed in separate dialysis cassettes against the ITC buffer comprising 50 mM Tris pH 8, 150 mM NaCl. All ITC experiments were performed by titrating between 122-134 µM M11 in 25 injections of 2 µL each into 196 µl of either 34.6-35.2 µM MBP-BECN1(1-150)Y or 26.6-30.2 µM MBP-BECN1(1-150)Y^{CysTM} in the cell. Blank profiles obtained by titrating M11 into buffer were subtracted from the experimental profiles. All experiments were repeated in triplicate. Data were analyzed using either NanoAnalyze (TA instruments) program with an independent, single-site model, or NIT-PIC (Scheuermann & Brautigam, 2015) for data integration, followed by data processing with

SEDPHAT (Zhao *et al.*, 2015, Brautigam *et al.*, 2016, Zhao & Schuck, 2015) and plotting of isotherms in GUSSI (Brautigam *et al.*, 2016). The processing included data refinement considering the local incompetent fraction as a function of the concentration compensation factor (Houtman *et al.*, 2007).

4.2.8. Hydrogen/Deuterium Exchange Mass Spectrometry (HDX-MS)

To better understand the impact of Zn^{2+}-binding on the conformation of BECN1 IDR we performed HDX-MS for His_6-BECN1(1-150)Y and His_6-BECN1(1-150)Y^{CysTM}. Individual proteins were purified by SEC in (20 mM Tris pH 8, 150 mM NaCl) and frozen in liquid nitrogen and sent to Biomolecular & Proteomics Mass Spectrometry Facility (BPMSF) at UCSD for HDX-MS data collection. Each sample was diluted 15-fold into either a sample buffer comprising 20mM Tris pH 8, 150 mM NaCl, or into a deuterated version of this buffer (20 mM Tris pD 8, 150 mM NaCl), prepared by drying the non-deuterated buffer in heat and resuspending in D_2O. The stock protein concentration of each sample was 5 µM. For His_6-BECN1(1-150)Y, 2.5 µM zinc acetate was added to the non-deuterated and deuterated buffer, to make sure that the protein is 100% saturated with Zn^{2+}, as per ICP-MS. For each deuteration time point such as 0, 30, 60, 120, and 300 seconds, 4 µL of the protein stock was equilibrated to 25 °C for 5 min and then mixed with 56 µL D_2O buffer (20 mM Tris pD 8, 150 mM NaCl), i.e. the stock protein was diluted 15 times. The exchange was quenched with an equal volume of quench solution (3 M guanidine, 0.1% formic acid, pH 2.66).

50 µL of the quenched sample was injected into the sample loop, followed by digestion on an in-line pepsin column (immobilized pepsin, Pierce, Inc.) at 15°C. We performed Hydrogen/deuterium exchange mass spectrometry (HDX-MS) using a Waters Synapt G2Si equipped with a nanoACQUITY UPLC system with HDX technology and a LEAP

autosampler. The resulting peptides were captured on a BEH C18 Vanguard pre-column, separated by analytical chromatography (Acquity UPLC BEH C18, 1.7 µM, 1.0 X 50 mm, Waters Corporation) using a 7-85% acetonitrile gradient in 0.1% formic acid over 7.5 min, and electrospray into the Waters SYNAPT G2Si quadrupole time-of-flight mass spectrometer. The mass spectrometer was set to collect data in the Mobility, ESI+ mode; mass acquisition range of 200–2,000 (m/z); scan time 0.4 sec. Continuous lock mass correction was accomplished with the infusion of leu-enkephalin (m/z = 556.277) every 30 sec (mass accuracy of 1 ppm for calibration standard). For peptide identification, the mass spectrometer was set to collect data in MS^E, ESI+ mode instead.

The peptides were identified from triplicate MS^E analyses of 5 µM His$_6$-BECN1(1-150)Y and His$_6$-BECN1(1-150)Y^{CysTM}, and data were analyzed using PLGS 3.0 (Waters Corporation). Peptide masses were identified using a minimum number of 250 ion counts for low energy peptides and 50 ion counts for their fragment ions. The peptides identified in PLGS were then analyzed in HDExaminer (Sierra Analytics). The relative deuterium uptake for each peptide was calculated by comparing the centroids of the mass envelopes of the deuterated samples vs. the undeuterated controls following previously published methods (Wales *et al.*, 2008). The peptides were then analyzed based on including signal to noise ratio and how well the theoretical isotope cluster matches the actual data and that decides the confidence level of the peptide. In general, a score of 0.9 or higher represents a high confidence match, while a score of 0.8 or higher represents a medium confidence match, and a score below 0.8 is considered a low confidence match. Therefore, data processing involves data curation to improve peptide scores so that most peptides match sequences with either high- or medium-confidence. The low confidence coverage peptides are not considered for any downstream calculations. Finally, D_2O uptake by the peptides was extracted and visualized in the uptake charts and heat maps.

4.2.9. Cellular autophagy assays

Autophagy levels were evaluated by monitoring cellular localization of GFP-tagged LC3. $3x10^5$ MCF7 cells were seeded in each chamber of an 8-well culture slide (Nunc^{TM} Lab-Tek^{TM} Chambered Coverglass) and allowed to grow in DMEM (Gibco) media with 10 % fetal bovine serum (Gibco) until ~80-90 % confluency was reached. Cells were co-transfected using Opti-MEM (Gibco) and Lipofectamine 2000 (Invitrogen, Carslbad, CA) as per the manufacturer's instructions, with 200 ng of GFP-LC3 expression plasmid and varying amounts of BECN1 WT and mutant plasmids optimized to ensure comparable protein expression, i.e. 300 ng of either FL, FL^{DM1}, FL^{DM2}, FL^{TM} or 9.4 ng FL$^{(\Delta31\text{-}123)}$. WT and mutant FLAG-BECN1 expression levels were verified by western blot using commercial mouse monoclonal Anti-FLAG M2 peroxidase antibody (Sigma). Actin levels were monitored as a loading control, using mouse Anti-Actin antibody clone C4 (EMD-Millipore). Western blots were developed using GE Storm 865 imaging system (GE Life Sciences).

24 hours after transfection, cells were cultured for an additional 4 hours in either rich (DMEM, 10 % FBS, and 2X amino acid mixture) or starvation (EBSS) medium. Cells were then washed with cold PBS and fixed with 4% paraformaldehyde (w/v in PBS), followed by counterstaining of the cells with DAPI to visualize nuclei and facilitate total cellular counts. Slides were stored in 70% v/v glycerol and were imaged using a Zeiss LSM 700 Confocal Laser Scanning Microscope (CLSM) (Zeiss, Thornwood, NY) with a Plan-Apochromat 40x/1.4 oil immersion lens, using excitation at 488 nm and emission set at 590 nm. GFP-LC3 labeled puncta was quantified by counting a minimum of 50 cells for duplicate samples per condition using

Imaris (8.2.0) (Bitplane, South Windsor, CT) in three independent experimental repeats. The significance of differences in autophagy levels was evaluated using a two-tailed heteroscedastic student's *t* test, wherein $p \leq 0.05$ is considered significant.

4.3. Results

4.3.1. Protein purification

4.3.1.1. Purification of WT and CxxC mutant His_6-BECN1(1-150)Y

The SEC chromatogram for His_6-BECN1(1-150)Y shows a single peak at 10.8 ml with two shoulders on the leading edge of the peak, one between 9.5 ml and 10.1 ml, and a smaller shoulder within the void-volume at 8.6 ml, preceding the major peak and shoulder (Figure 4.2). The SDS-PAGE results indicate that the major peak, including the leading edge shoulders, all contain His_6-BECN1(1-150)Y. The void-volume shoulder likely contained aggregated protein and was not considered further. The molecular masses estimated from the SEC profile are 64.1 kDa for the peak, and between 89.3 kDa to 150.7 kDa for the shoulder closest to the peak; however the theoretical molecular mass of the His_6-BECN1(1-150)Y is 18.3 kDa. Thus, the estimated molecular mass of His_6-BECN1(1-150)Y is significantly larger than expected. This likely indicates His_6-BECN1(1-150)Y is a highly disordered protein, resulting in an extended conformation, although an oligomeric species cannot be ruled out. Since the leading edge of the protein elutes as a larger species, it may represent a larger oligomer or a more extended protein conformation, relative to the protein eluting in the peak. Furthermore, as the leading-edge shoulder is only 0.5 ml behind the void volume shoulder, it is likely that protein eluting in the leading-edge shoulder may include protein aggregates or large oligomers, despite careful selection of protein fractions. Of the total protein eluted under the leading shoulder and peak, the shoulder constitutes 18 % (0.1 mg/L), and the peak constitutes 82 % (0.7 mg/L). This suggests that His$_6$-BECN1(1-150)Y primarily exists in a more compact conformation eluting under the peak, rather than the oligomer or extended conformation in the leading-edge shoulder of the peak.

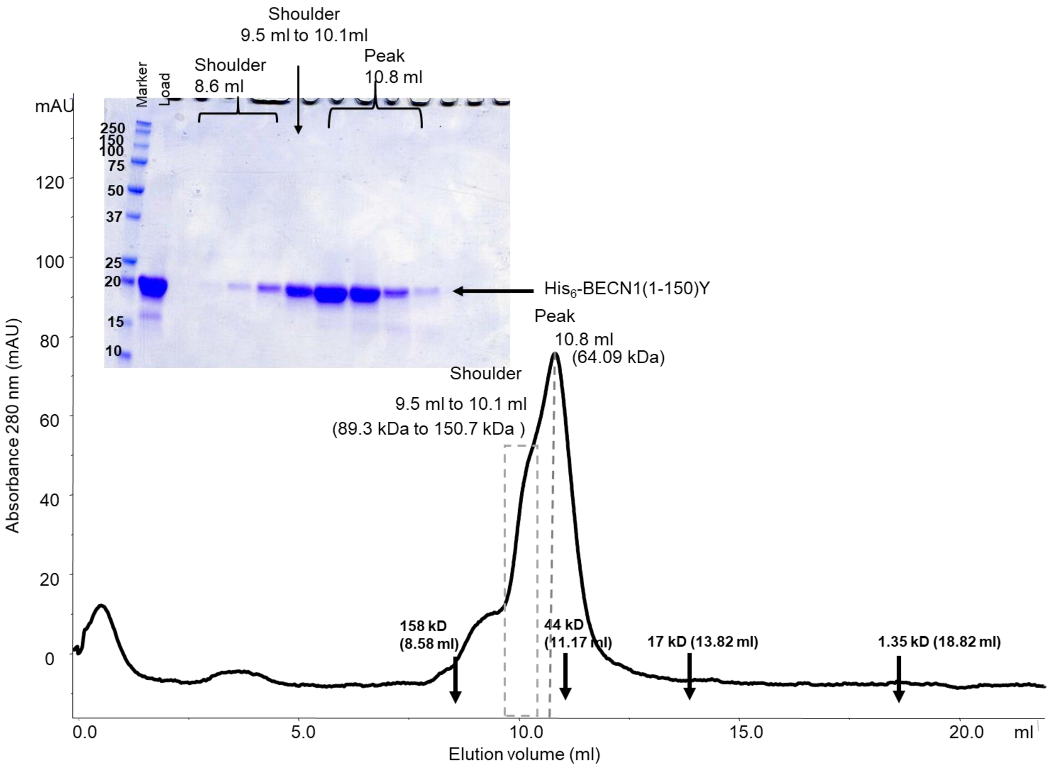


Figure 4.2: Size exclusion chromatogram and the corresponding SDS-PAGE of His$_6$-BECN1(1-150)Y. The elution positions of SEC standards are indicated by arrows. The dashed line indicates the peak and the box represents the leading edge shoulder region.

Unlike His$_6$-BECN1(1-150)Y, the His$_6$-BECN1(1-150)Y^{CysTM} SEC chromatogram from a Superdex 75 increase 10/300 column (Figure 4.3) has a void-volume peak at 8.6 ml. The SDS PAGE result indicates that the void-volume peak contains a higher molecular mass contaminant protein while the major peak contains only His$_6$-BECN1(1-150)Y^{CysTM}. The net yield of the purified His$_6$-BECN1(1-150)Y^{CysTM} protein obtained for the major peak is 1 mg/L of bacterial culture. The molecular mass of His$_6$-BECN1(1-150)Y^{CysTM} estimated from this SEC profile is 86 kDa, however the theoretical molecular mass is 18.3 kDa. The estimated molecular mass of His$_6$-BECN1(1-150)Y^{CysTM} is significantly larger than expected. This likely indicates His$_6$-BECN1(1-

150)Y$^{\mathrm{CysTM}}$ is a highly disordered protein, resulting in an extended conformation, although an oligomeric species cannot be ruled out. Notably, a comparison of the SEC elution profiles of His$_6$-BECN1(1-150)Y (Figure 4.2) and His$_6$-BECN1(1-150)Y$^{\mathrm{CysTM}}$ (Figure 4.3) shows that the SEC elution of His$_6$-BECN1(1-150)Y$^{\mathrm{CysTM}}$ is left shifted, suggesting that in the absence of CxxC motifs, the BECN1 IDR has a more extended conformation.

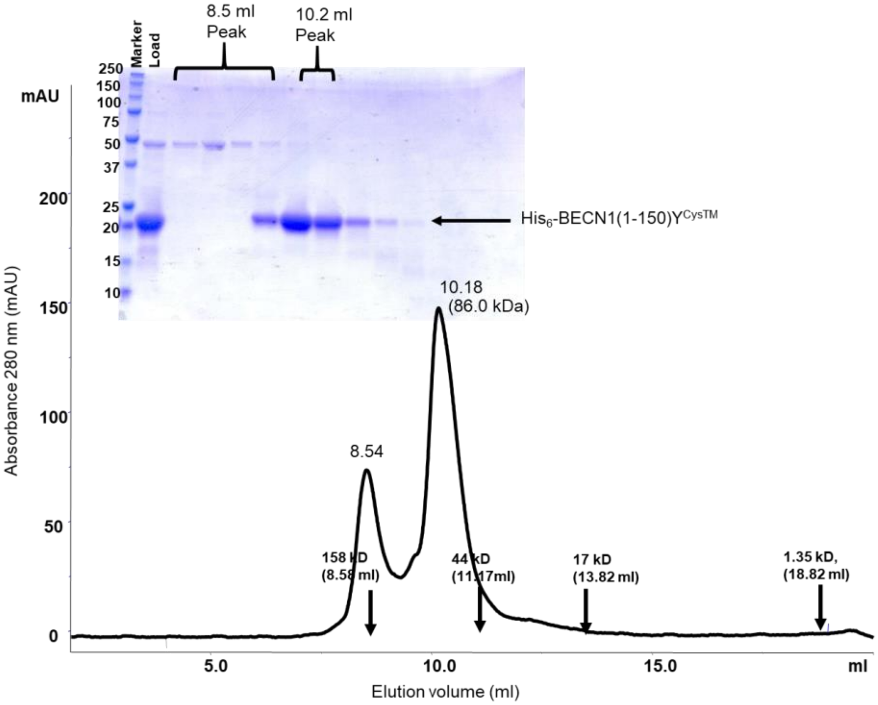


Figure 4.3: Size exclusion chromatogram and the corresponding SDS-PAGE of His$_6$-BECN1(1-150)Y$^{\mathrm{CysTM}}$. The elution positions of SEC standards are indicated by arrows.

The SEC chromatogram from a Superdex 75 increase 10/300 column for His$_6$-BECN1(1-150)Y$^{\mathrm{CysDM1}}$ shows a void volume peak and a well-separated major peak (Figure 4.4) at 10.36 ml. SDS PAGE result indicates that the void-volume peak contains a higher molecular mass contaminant protein and the major peak contains only His$_6$-BECN1(1-150)Y$^{\mathrm{CysDM1}}$. The net yield of the purified His$_6$-BECN1(1-150)Y$^{\mathrm{CysDM1}}$ protein from the major peak is 0.23 mg/L of bacterial culture. The estimated molecular mass from the major peak elution volume is 79 kDa, significantly larger than the theoretical molecular mass of 18.3 kDa. This likely indicates His$_6$-

BECN1(1-150)Y^{CysDM1} is a highly disordered protein, resulting in an extended conformation, although an oligomeric species cannot be ruled out. A comparison of the SEC elution profiles shows that the elution peak of His_6-BECN1(1-150)Y^{CysDM1} is left shifted relative to His_6-BECN1(1-150)Y (Figure 4.2), although not as much as His_6-BECN1(1-150)Y^{CysTM} (Figure 4.3), suggesting that mutation of one of the CxxC motifs, results in an extended conformation, but not as much as the tetrad mutation.

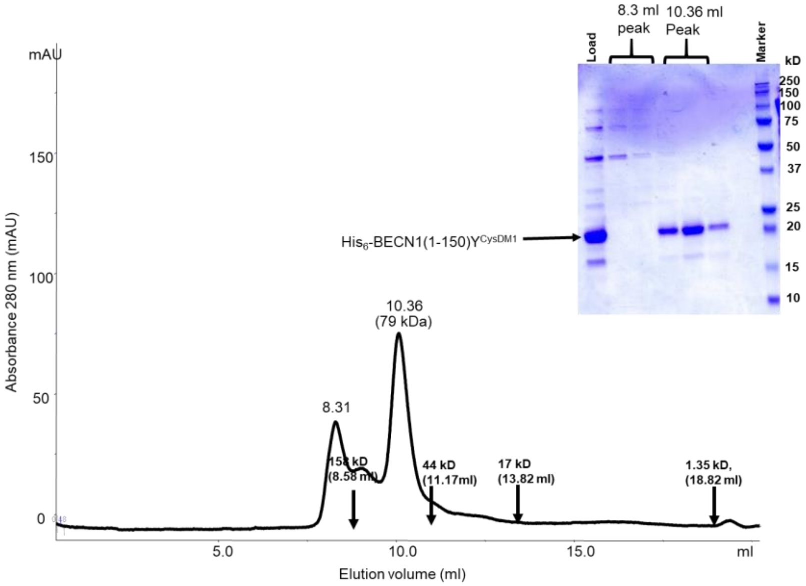


Figure 4.4: Size exclusion chromatogram and the corresponding SDS-PAGE of His_6-BECN1(1-150)Y^{CysDM1}. The elution positions of SEC standards are indicated by arrows**.**

The Superdex 75 increase SEC chromatogram for His_6-BECN1(1-150)Y^{CysDM2} shows a void volume peak at 8.6 ml and a single, well-separated major peak eluting at 10.1 ml and another peak at 16.2 ml. SDS PAGE result indicates that void-volume peak contains a higher molecular mass contaminant protein and the major peak (Figure 4.5) contains His_6-BECN1(1-150)Y^{CysDM2} protein and the last peak likely contains a low concentration degradation peak. The void volume peak and the degradation peak were not analyzed further. The net yield of the purified His_6-BECN1(1-150)Y^{CysDM2} protein obtained from the major peak is 0.82 mg/L of bacterial culture. The molecular mass estimated from the major peak elution volume is 87 kDa, which is significantly larger than the theoretical molecular mass of His$_6$-BECN1(1-150)Y$^{\mathrm{CysDM2}}$ which is 18.3 kDa. This likely indicates His$_6$-BECN1(1-150)Y$^{\mathrm{CysDM2}}$ is a highly disordered protein, resulting in an extended conformation, although an oligomeric species cannot be ruled out. A comparison of the SEC elution profiles shows that the SEC elution peak of His$_6$-BECN1(1-150)Y$^{\mathrm{CysDM2}}$ is left shifted relative to that His$_6$-BECN1(1-150)Y (Figure 4.2), although not compared to as His$_6$-BECN1(1-150)Y$^{\mathrm{CysTM}}$, suggesting that mutation of one of the CxxC motifs, results in an extended conformation, but not as much as the tetrad mutation.

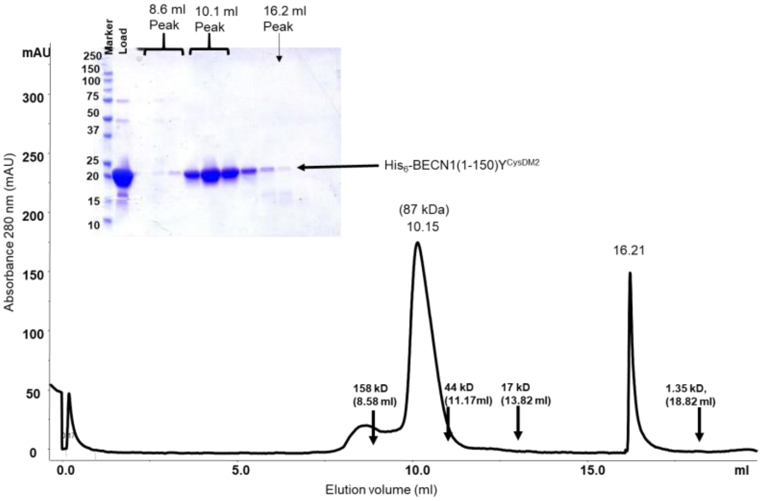


Figure 4.5: Size exclusion chromatogram and the corresponding SDS-PAGE of His$_6$-BECN1(1-150)Y$^{\mathrm{CysDM2}}$. The elution positions of SEC standards are indicated by arrows.

Although we were able to purify the WT and mutant His$_6$- BECN1(1-150)Y proteins to >90% purity in quantities sufficient for various biophysical and structural studies, ICP-MS analyses of these proteins (see Section 4.3.2.1) indicated inconsistent metal binding possibly due to reduction of the cysteine motifs during affinity and Mono Q chromatography, and consequent loss of bound metal. As this inconsistent metal-binding was not seen when the MBP-tag was used for purification, MBP-tagged proteins were used for all analyses for publication.

4.3.1.2. Purification of WT and mutant MBP-BECN1(1-150)Y

The SEC chromatogram from a Superdex 200 10/300 column for MBP-BECN1(1-150)Y (Figure 4.6) shows three distinct peaks: a void volume peak at 8.2 ml; a major peak at 14.3 ml with a shoulder on the leading edge between 11.5 ml to 13.5 ml, extending into the void volume at 10 ml; and also a smaller, overlapping peak at 16.0 ml. The void-volume peak probably corresponds to aggregated protein. SDS-PAGE indicates that both the leading edge shoulder and peak contain pure MBP-BECN1(1-150)Y, while the 16.0 ml peak contains degraded protein as well as MBP-BECN1(1-150)Y. Protein contained in the 16.0-ml peak was not analyzed further. The molecular mass estimated from the elution volumes of the shoulder and the major peak are 111 kDa to 382 kDa, and 73 kDa respectively; significantly larger than the theoretical molecular mass of the MBP-BECN1(1-150)Y is 60.4 kDa. This suggests that MBP-BECN1(1-150)Y is partially disordered, and has an extended conformation, although an oligomeric species cannot be ruled out. Since the leading edge of the protein elutes as a larger species, it may represent a larger oligomer or a more extended protein conformation, relative to the protein eluting in the peak. The net yield of the purified MBP-BECN1(1-150)Y protein obtained for the shoulder region and the major peak are 1.3 mg/L and 3 mg/L of bacterial culture respectively. This suggests that MBP-BECN1(1-150)Y primarily exists in the more compact conformation eluting under the peak, rather than the oligomer or extended conformation in the leading-edge shoulder.

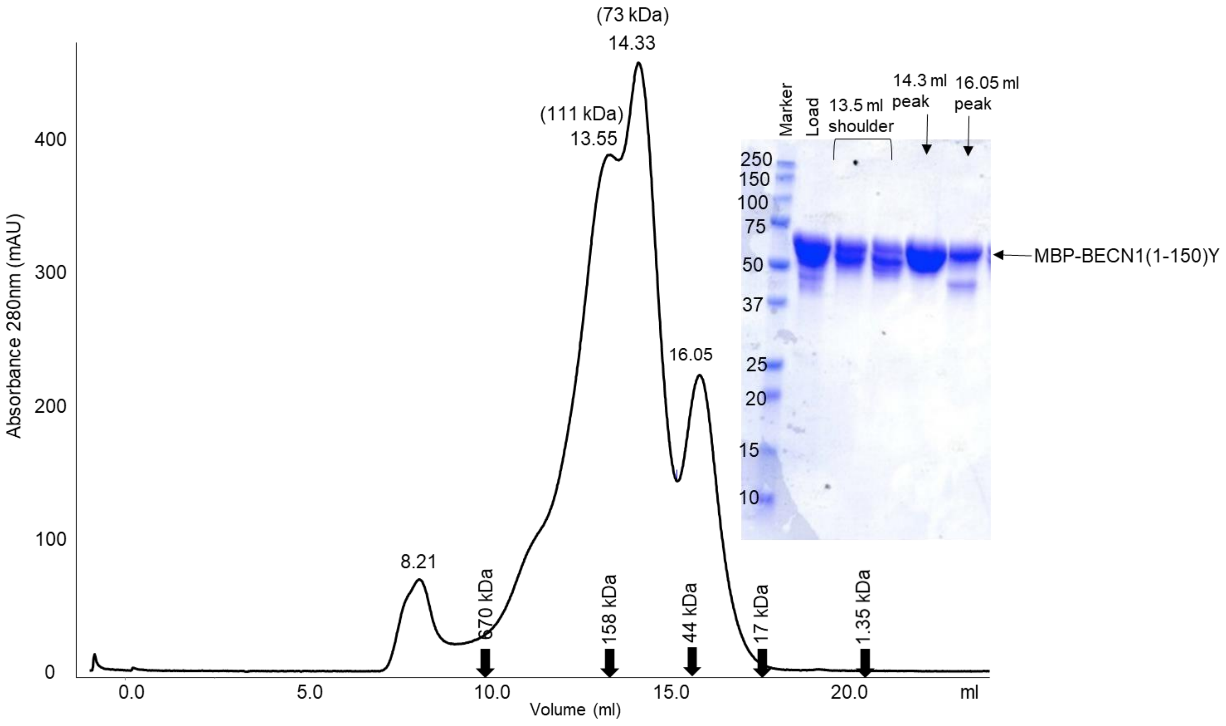


Figure 4.6: Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-BECN1(1-150)Y. The elution positions of SEC standards are indicated by arrows.

The SEC chromatogram from a Superdex 200 10/300 column for MBP-BECN1(1-150)Y^{CysTM} (Figure 4.8) shows four distinct peaks, a void volume peak at 8.3 ml, a small peak at 12.1 ml, a major peak at 13.9 ml bearing a shoulder at the trailing edge, followed by an overlapping third peak at 15.93 mL. The void-volume peak probably corresponds to aggregated protein. SDS-PAGE indicates that both the major peak and the trailing edge shoulder contain pure MBP-BECN1(1-150)Y^{CysTM}, while the protein contained in the 15.9 ml peak contains degraded protein as well as MBP-BECN1(1-150)Y^{CysTM}. The protein in the 15.9 ml peak was not analyzed further. The small peak at 12.1 ml likely contain MBP-BECN1(1-150)Y^{CysTM} contaminated with aggregated protein from the void volume, and was also not analyzed further. The molecular mass estimated from the elution volume of the major peak is 93 kDa; however the theoretical molecular mass of the MBP-BECN1(1-150)Y^{CysTM} is 60.4 kDa. This significantly larger than expected estimation of molecular mass may be because MBP-BECN1(1-150)Y^{CysTM}

is partially disordered, allowing the protein to have an extended conformation, although an oligomeric species cannot be ruled out. The net yield of the purified MBP-BECN1(1-150)Y^{CysTM} protein obtained in the major peak of the SEC chromatogram (Figure 4.8) is 0.55 mg/L of bacterial culture. A comparison of the SEC elution profiles shows that the major elution peak of MBP-BECN1(1-150)Y^{CysTM} is left shifted relative to of MBP-BECN1(1-150)Y (Figure 4.6), suggesting that in the absence of the CxxC motifs, BECN1 IDR has a more extended conformation.

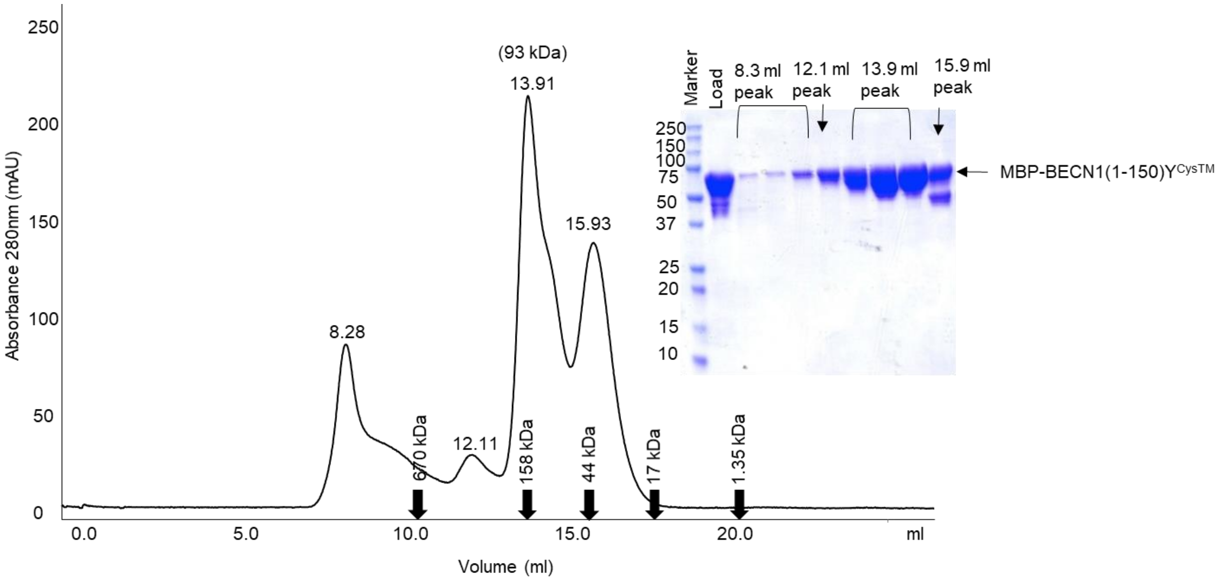


Figure 4.7: Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-BECN1(1-150)Y^{CysTM}. The elution positions of SEC standards are indicated by arrows**.**

The SEC chromatogram from a Superdex 200 10/300 column for MBP-BECN1(1-150)Y^{CysDM1} shows three peaks (Figure 4.8) , a void-volume peak at 8.4 ml, a major peak at 13.9 ml including a trailing-edge shoulder at 14.5 ml, followed by an overlapping third peak at 15.9 ml. SDS PAGE indicates that the void-volume peak, major peak and shoulder all contain MBP-BECN1(1-150)Y^{CysDM1}, while the 15.9 mL peak contains degraded protein. The protein in the 15.9 ml peak was not analyzed further. The void-volume peak probably corresponds to aggregated protein and was also not analyzed further. The molecular mass estimated from the major peak elution volume is 114 kDa; however the theoretical molecular mass of the MBP-BECN1(1-150)Y$^{\text{CysDM1}}$ is 60.4 kDa. The significantly larger than expected estimation of the molecular mass is likely because MBP-BECN1(1-150)Y$^{\text{CysDM1}}$ is a partially disordered protein, allowing the protein to have an extended conformation, although an oligomeric species cannot be ruled out. The net yield of the purified MBP-BECN1(1-150)Y$^{\text{CysDM1}}$ protein is 0.93 mg/L of bacterial culture. A comparison of the SEC elution profiles shows that the major elution peak of MBP-BECN1(1-150)Y$^{\text{CysDM1}}$ is left shifted relative to of MBP-BECN1(1-150)Y (Figure 4.6), but not compared to MBP-BECN1(1-150)Y$^{\text{CysTM}}$ (Figure 4.7), suggesting that mutation of one of the CxxC motifs, results in an extended conformation, but not as much as the tetrad mutation.

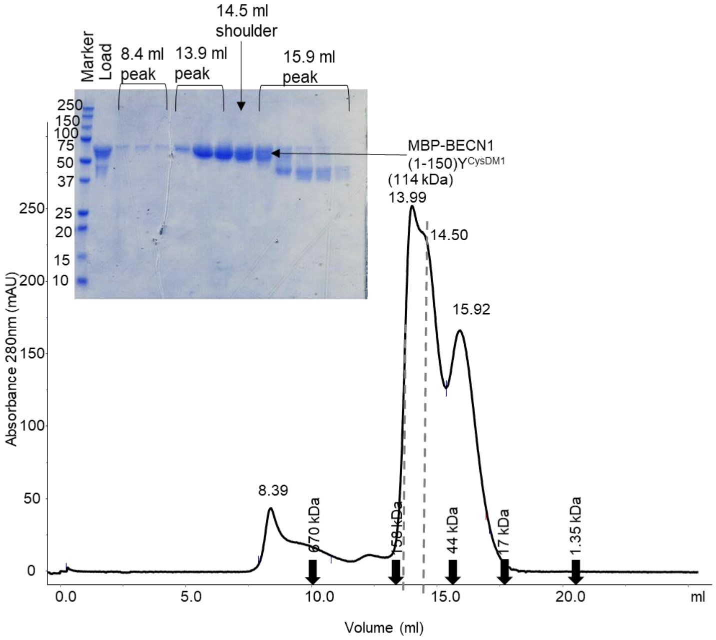


Figure 4.8: Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-BECN1(1-150)Y$^{\text{CysDM1}}$. The elution positions of SEC standards are indicated by arrows**.** The dashed lines indicate the peak.

The SEC chromatogram from a Superdex 200 10/300 column for MBP-BECN1(1-150)Y$^{\text{CysDM2}}$ (Figure 4.9) shows three distinct peaks, a void-volume peak at 8.2 ml, a major peak at 12.67 ml, and a small peak at 16.1ml. SDS PAGE indicates that the void-volume peak and major peak both contain MBP-BECN1(1-150)Y$^{\text{CysDM2}}$, while the 16.1 mL peak contains degraded protein. The void-volume peak probably corresponds to aggregated protein. The protein in the 8.2 ml and 16.1 ml peaks was not analyzed further. The molecular mass estimated from the elution volume of the major peak is 194 kDa, which is significantly larger than expected, likely because MBP-BECN1(1-150)Y$^{\text{CysDM2}}$ is a partially disordered protein, allowing the protein to have an extended conformation, although an oligomeric species cannot be ruled out. The net yield of the purified MBP-BECN1(1-150)Y$^{\text{CysDM2}}$ protein obtained from the major peak is 3.6 mg/L of bacterial culture. A comparison of the SEC elution profiles shows that the elution peak of MBP-BECN1(1-150)Y$^{\text{CysDM2}}$ (Figure 4.9) is most left shifted relative to MBP-BECN1(1-150)Y (Figure 4.6) , MBP-BECN1(1-150)Y$^{\text{CysTM}}$ (Figure 4.7), and MBP-BECN1(1-150)Y$^{\text{CysDM1}}$ (Figure 4.8), suggesting that absence of the second CxxC motif (C137, C140), results in the most extended conformation of the BECN1 IDR.

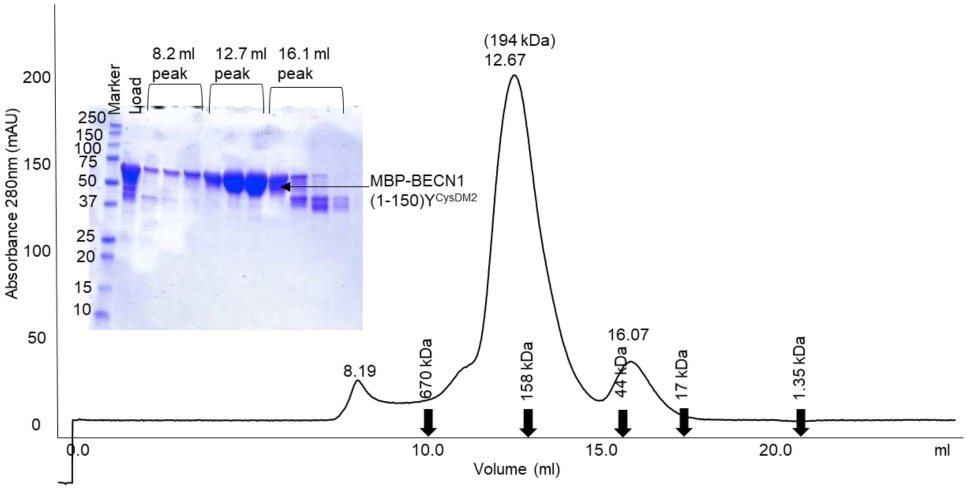


Figure 4.9: Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-BECN1(1-150)Y$^{\text{CysDM2}}$. The elution positions of SEC standards are indicated by arrows**.**

4.3.1.3. Purification of WT and mutant BECN1(1-150)Y

The BECN1(1-150)Y SEC profile from a Superdex 75 increase 10/300 column (Figure. 4.10) has a major peak at 10.7 ml with a shoulder on the leading-edge of the peak between 8.0 ml to 9.9 ml that extends into the void volume and a small shoulder on the trailing edge of the peak at 13.0 ml, which had no measurable amount of protein and was not analyzed further. The SDS-PAGE indicates that both the leading edge shoulder and peak contain BECN1(1-150)Y. The molecular mass estimated from the peak is 67.2 kDa, and between 96.8 to 242 kDa for the leading edge shoulder; however the theoretical molecular mass of the BECN1(1-150)Y is 17.2 kDa. Since the leading edge shoulder is partly in the void volume, it is likely that it contains aggregates or large oligomers of BECN1(1-150)Y. The estimated molecular mass of the major peak is also significantly larger than expected for BECN1(1-150)Y. This likely indicates a highly disordered protein, resulting in an extended conformation, although an oligomeric species cannot be ruled out. The net yield of the purified protein, including both the leading-edge shoulder and the major peak is 0.3 mg/L of bacterial culture, of which 83% (0.25 mg/L) and 17% (0.05 mg/L) of the total protein is eluted in the major peak fractions and in the leading-edge shoulder fractions respectively. This suggests that BECN1(1-150)Y primarily exists in a more compact conformation eluting under the peak, rather than the oligomeric or extended conformation in the leading-edge shoulder of the peak.

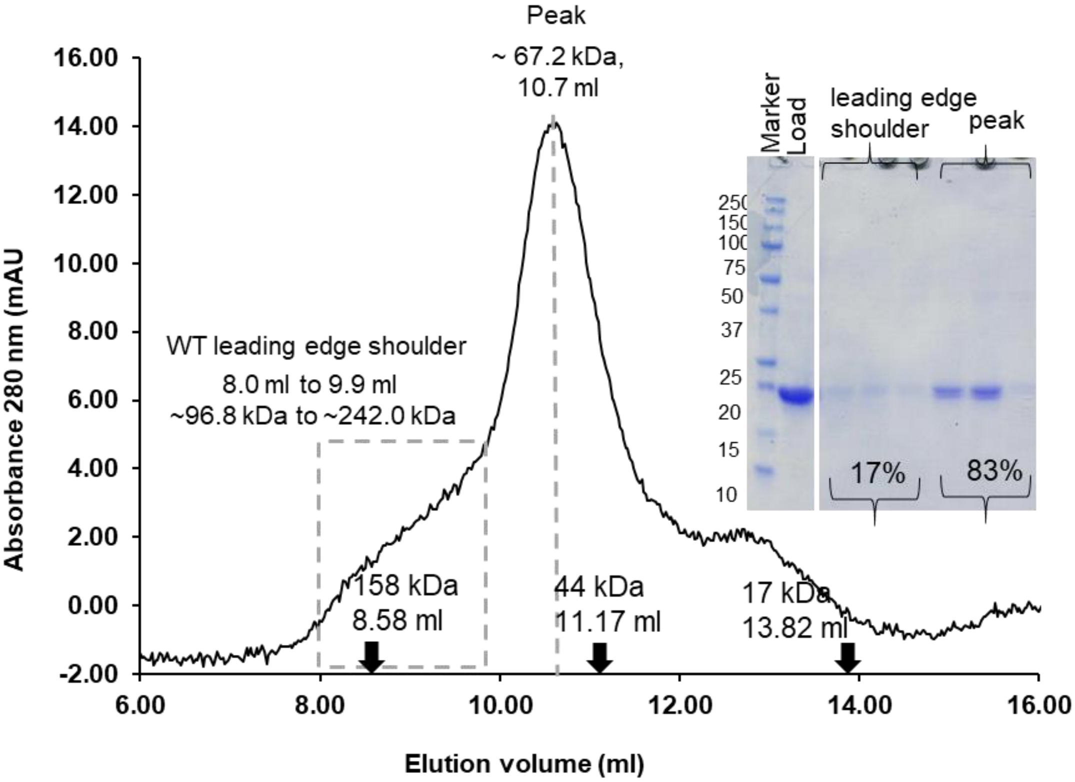


Figure 4.10: Size exclusion chromatogram and the corresponding SDS-PAGE of BECN1(1-150)Y. The elution positions of SEC standards are indicated by arrows**.** The dashed line indicates the peak and boxed region indicates the leading-edge shoulder region

The BECN1(1-150)Y$^{\text{CysTM}}$ SEC profile (Figure 4.10) displays a single major peak at 9.9 ml on a Superdex 75 increase 10/300 column, with a small trailing-edge shoulder. SDS-PAGE indicates the presence of BECN1(1-150)Y$^{\text{CysTM}}$ in the major peak fractions and no detectable protein in the trailing shoulder. The net yield of the purified BECN1(1-150)Y$^{\text{CysTM}}$ obtained for the major peak is 0.27 mg/L of bacterial culture. The molecular mass estimated from the elution volume of the major peak is 95.9 kDa; however the theoretical molecular mass of BECN1(1-150)Y$^{\text{CysTM}}$ is 17.1 kDa. The estimated molecular mass of BECN1(1-150)Y$^{\text{CysTM}}$ is significantly larger than expected. This likely indicates BECN1(1-150)Y$^{\text{CysTM}}$ is highly disordered, resulting in an extended conformation, although an oligomeric species cannot be ruled out. A comparison of the SEC elution profiles of BECN1(1-150)Y (Figure 4.10) and BECN1(1-150)Y$^{\text{CysTM}}$ (Figure

4.11) shows that the SEC elution of BECN1(1-150)Y^{CysTM} is left shifted, suggesting that in the absence of the CxxC motifs, BECN1 IDR shows increased conformational flexibility.

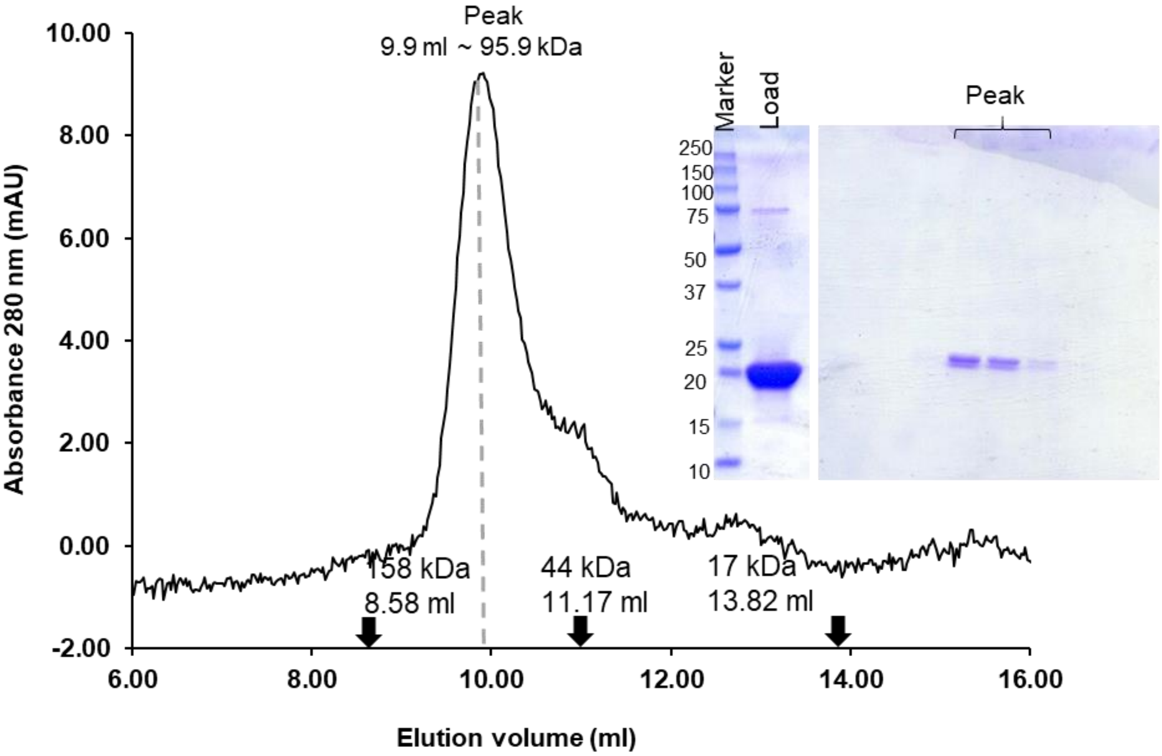


Figure 4.11: Size exclusion chromatogram and the corresponding SDS-PAGE of BECN1(1-150)Y^{CysTM}. The elution positions of SEC standards are indicated by arrows**.** The dashed line indicates the peak

4.3.1.4. Purification of MBP-BECN1(Δ31-123):

The SEC chromatogram from a Superdex 200 10/300 column for MBP-BECN1(Δ31-123) (Figure 4.12) shows two peaks: a void-volume peak at 8.4 ml and a major peak at 10.9 ml. SDS PAGE indicates that both peaks contain MBP-BECN1(Δ31-123). The void-volume peak likely represents aggregated MBP-BECN1(Δ31-123), and is larger than in the preceding purifications due to the absence of the aromatic finger mutation (F359D, F360D,W361D) in the BECN1 BARAD which is known to increase solubility of BARAD-containing BECN1 proteins (Glover *et al.*, 2017). The net yield of the purified MBP-BECN1(Δ31-123) protein obtained for the major peak is 3.4 mg/L of bacterial culture. The molecular mass estimated from the elution volume of the peak is 334 kDa, which is substantially larger than expected as the theoretical molecular mass of MBP-BECN1(Δ31-123) is 83.2 kDa. This observed molecular mass is larger than expected likely because MBP-BECN1(Δ31-123) forms an antiparallel homodimer with an elongated, rod shaped conformation.

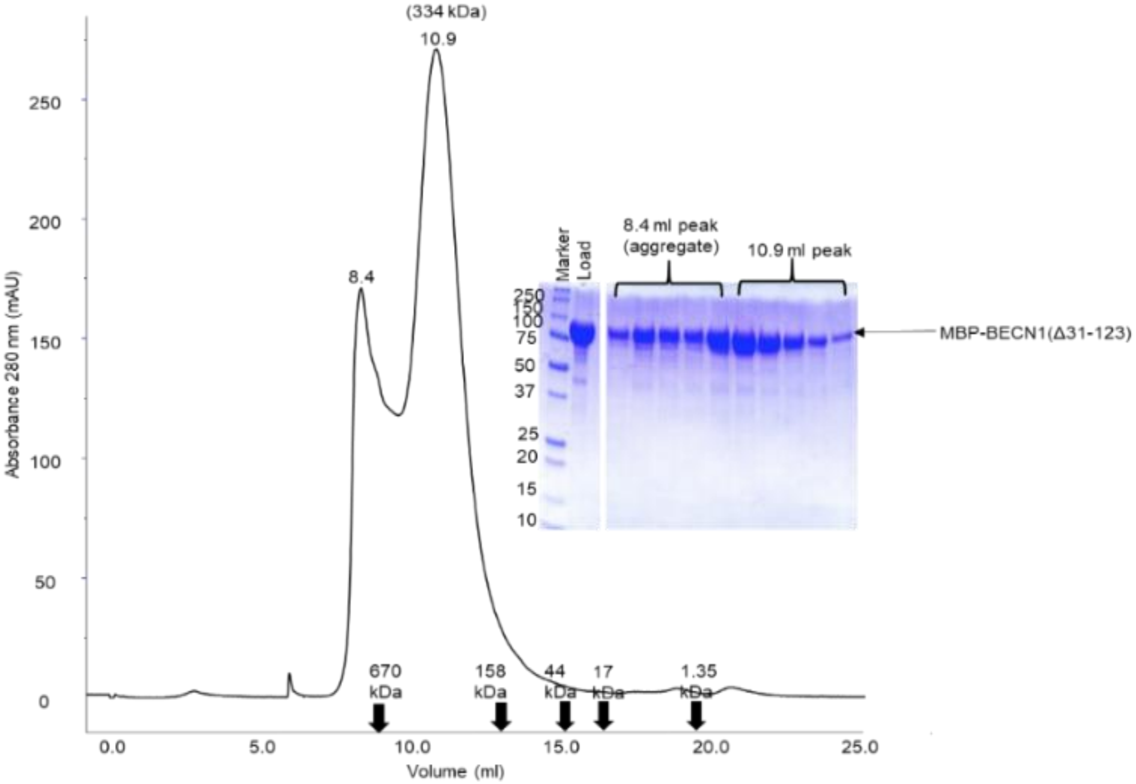


Figure 4.12: Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-BECN1(Δ31-123). The elution positions of SEC standards are indicated by arrows**.**

4.3.1.5. Purification of MBP-BECN1(1-140) constructs:

The SEC chromatogram from a Superdex 200 10/300 column for MBP-BECN1(1-140) (Figure 4.13) shows a major peak at 13.7 ml and an overlapping peak on the leading edge with an elution volume of 12.2 ml. SDS PAGE indicates that both peaks contain MBP-BECN1(1-140). The net yield of the purified MBP-BECN1(1-140) protein obtained from the major peak is 3.1 mg/L of bacterial culture. The estimated molecular mass of MBP-BECN1(1-140) obtained for the major peak is 72 kDa; however, the theoretical molecular mass of MBP-BECN1(1-140) is 59.3 kDa. The molecular mass estimated from the elution volume is significantly larger than expected, likely because it is a partially disordered protein, allowing the protein to have an extended conformation.

This purification was done successfully only once and there were significant problems in reproducing similar quality protein after this purification. There was considerable protein degradation starting at the affinity stage, and the protein showed aggregation as well as degradation later during SEC, therefore this protein was not used further.

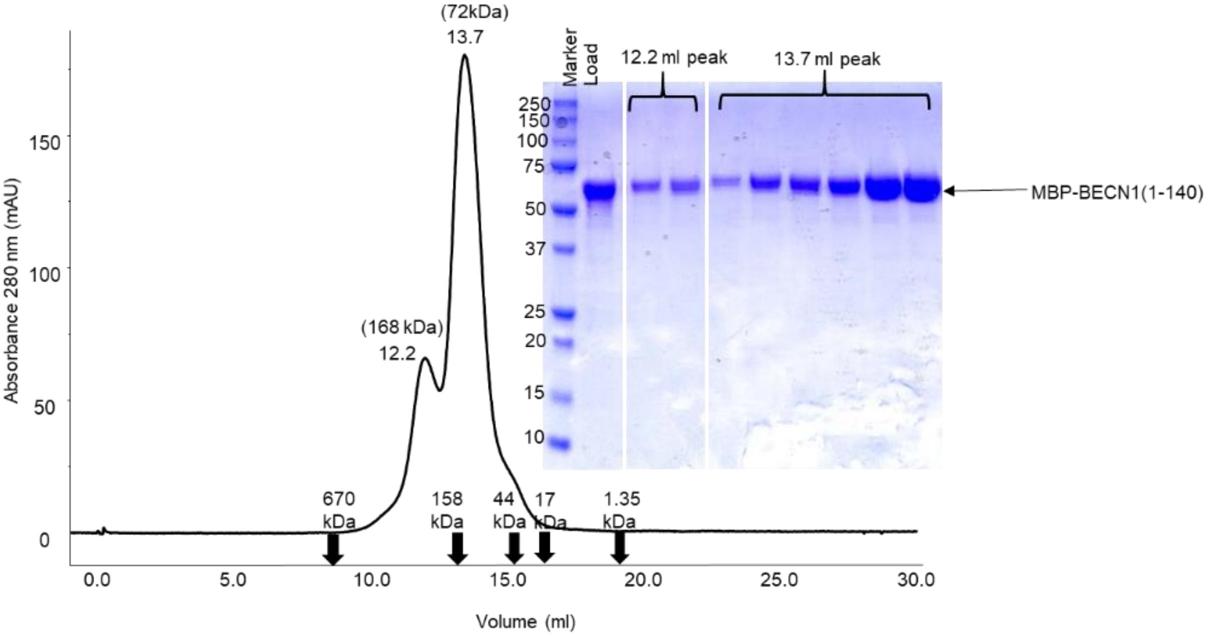


Figure 4.13: Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-BECN1(1-140). The elution positions of SEC standards are indicated by arrows**.**

4.3.1.6. Purification of WT and mutant Strep-SUMO-BECN1(1-143)

The SEC chromatogram from a Superdex 200 10/300 column of Strep-SUMO-BECN1(1-143) (Figure 4.14) has two peaks: a void-volume peak at 8.3 ml, and a very broad, major peak at 12.7 ml, with an overlapping 15.5 ml shoulder on the trailing edge. The void-volume peak likely represents aggregated protein. SDS PAGE indicates that both the peaks predominantly contain

Strep-SUMO-BECN1(1-143), with the major peak, including the 15.5 ml shoulder also containing a 20kDa contaminant, presumably degraded Strep-SUMO-BECN1(1-143). The net yield of the purified Strep-SUMO-BECN1(1-143) protein obtained for the major peak is 2 mg/L of bacterial culture. The molecular mass estimated from the peak elution volume is 109 kDa; however, the theoretical molecular mass of Strep-SUMO-BECN1(1-143) is 28.8 kDa. This significantly larger than estimated molecular mass is likely because Strep-SUMO-BECN1(1-143) is a partially disordered protein, allowing the protein to have an extended conformation.

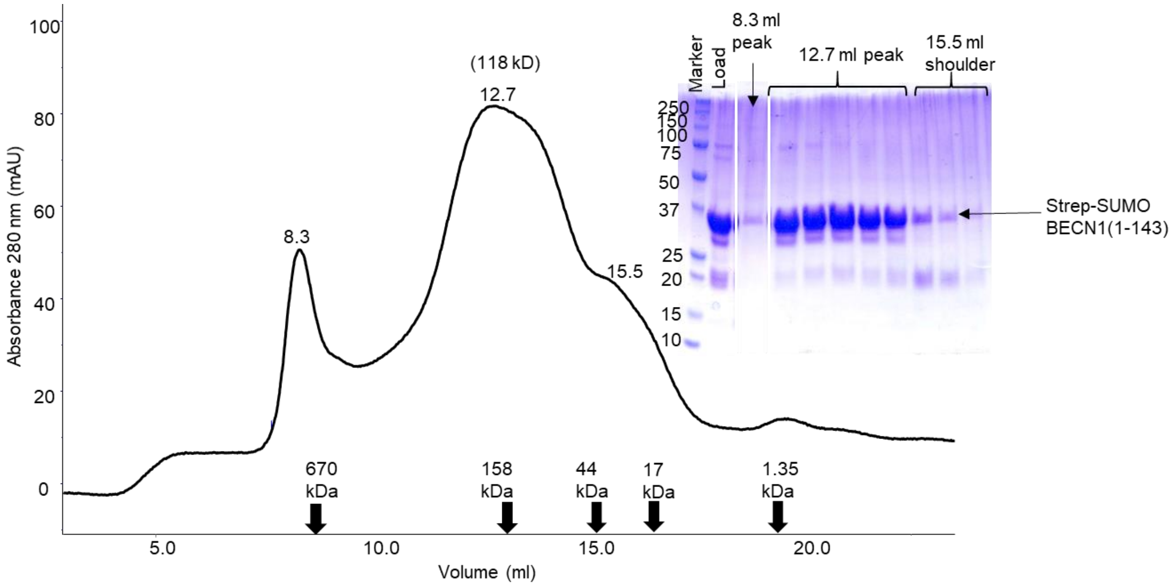


Figure 4.14: Size exclusion chromatogram and the corresponding SDS-PAGE of Strep-SUMO BECN1(1-143). The elution positions of SEC standards are indicated by arrows**.**

The SEC chromatogram from Superdex 200 10/300 column for Strep-SUMO-BECN1(1-143)^{CysTM} (Figure 4.15) has 4 peaks: a well-separated void-volume peak at 8.1 ml, a broad major peak at 13.7 ml, with an overlapping peak on the trailing edge at 15.3 ml, and a small low molecular mass peak at 19.9 ml. SDS PAGE does not show any protein in the void-volume peak, indicating that this peak corresponds to a small amount of aggregated protein. The major peak contains pure Strep-SUMO BECN1(1-143)^{CysTM} protein, while the overlapping trailing peak primarily contains degraded protein. The last peak appears to correspond to a contaminant protein. The net yield of the purified Strep-SUMO-BECN1(1-143)$^{\mathrm{CysTM}}$ protein obtained for the major peak is 0.25 mg/L of bacterial culture. The molecular mass estimated from the elution volume of the major peak is 67.4 kDa; however the theoretical molecular mass of Strep-SUMO-BECN1(1-143)$^{\mathrm{CysTM}}$, is 28.8 kDa. This significantly larger estimated molecular mass is likely because Strep-SUMO-BECN1(1-143)$^{\mathrm{CysTM}}$ is a partially disordered protein, allowing the protein to have an extended conformation.

This protein purification of the WT and CysTM Strep-SUMO BECN1(1-143) constructs had multiple problems. There was considerable protein degradation starting at the affinity stage, and the protein also did not elute cleanly from the Streptactin resin, but rather leached off gradually. Both these phenomena contributed to substantial loss of protein during purification, therefore, this construct was not used further.

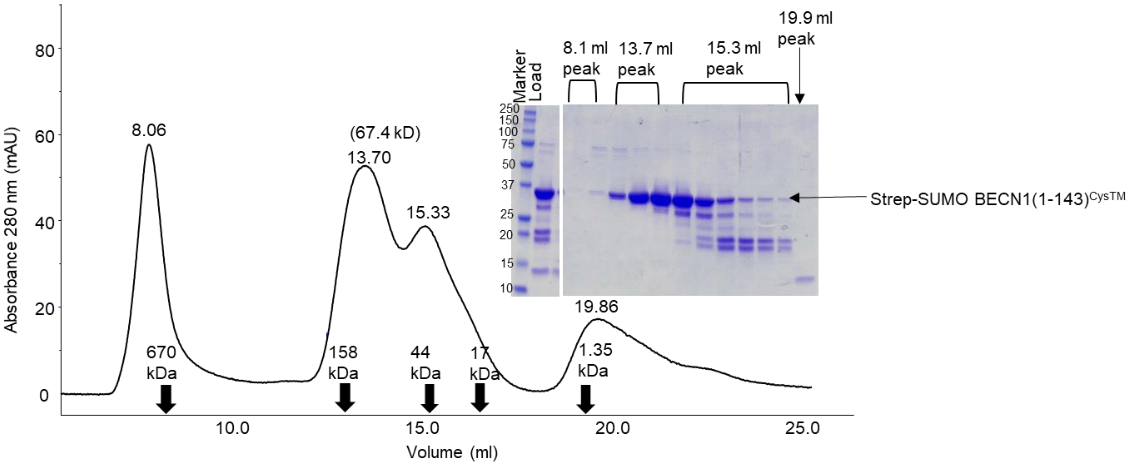


Figure 4.15: Size exclusion chromatogram and the corresponding SDS-PAGE of Strep-SUMO BECN1(1-143)$^{\mathrm{CysTM}}$. The elution positions of SEC standards are indicated by arrows.

4.3.2. BECN1 bears two invariant CxxC motifs responsible for binding Zn^{2+}:

Although the BECN1 IDR is poorly conserved, it has two invariant CxxC motifs (Figure 4.1) whose role is unknown. Since CxxC motifs are often responsible for binding metal, especially Zn^{2+}, we used quantitative ICP-MS to ident ify and quantify any element bound. All protein used for these experiments were purified to 90-95% homogeneity as verified by SDS-PAGE. The significance of differences in levels of Zn^{2+} bound by proteins was evaluated using a two-tailed heteroscedastic student's *t* test, wherein $p \leq 0.05$ is considered significant.

4.3.2.1. Element Analyses for MBP-BECN1 and MBP-BECN1(Δ31-123) constructs

ICP-MS data for various MBP-BECN1 constructs and the MBP-BECN1(Δ31-123) construct are summarized in Table 4.6.

Table 4.6: Elemental content as determined by ICP-MS for various MBP-BECN1 and MBP-BECN1(Δ31-123) constructs

Elements screened		**MBP-BECN1**		**MBP BECN1^{CysTM}**		**MBP-BECN1 $_{CysDM1}$**		**MBP-BECN1 $_{CysDM2}$**		**MBP-BECN1(Δ31-123)**	
Element / Isotope	Atomic wt (g/mole)	Avg	SD	Avg	SD	Avg	SD	Avg	SD	Avg	SD
		(μmole/μmole)									
Li/ 7	6.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
B/ 11	10.81	3.17	0.02	1.11	0.06	3.63	0.07	2.59	0.16	0.77	0.01
Na/ 23	22.99	0.00	0.05	0.00	0.00	42.77	108.0	0.04	0.11	1677.96	9.13
Mg/ 24	24.31	0.00	0.00	0.00	0.00	0.03	0.01	0.00	0.01	0.01	0.00
P/ 31	30.97	0.00	0.00	0.18	0.31	0.00	0.00	0.00	0.00	0.28	0.00
S/ 34	32.07	78.7	2.11	41.9	1.26	62.5	3.05	0.06	0.00	23.08	20.22
K/ 39	39.10	0.15	0.53	0.00	0.00	0.28	0.74	0.00	0.00	0.37	0.02
Ca/ 40	40.08	0.00	0.00	0.00	0.00	0.15	0.03	0.00	0.00	0.14	0.01
V/ 51	50.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	not tested	not tested
Cr/ 52	52.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	not tested	not tested
Mn/55	54.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.04
Fe/ 56	55.85	0.00	0.00	0.00	0.00	0.05	0.01	0.00	0.01	0.06	0.00
Co/ 59	58.93	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ni/ 60	58.69	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
Cu/ 63	63.55	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.17	0.00
Zn/ 66	65.39	1.03	0.08	0.01	0.00	0.55	0.09	0.52	0.05	1.08	0.00
As/ 75	74.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Se/ 78	78.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mo/ 95	95.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cd/ 111	112.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

ICP-MS analyses indicated that WT, DM1, DM2 MBP-BECN1 proteins and MBP-BECN1(Δ31-123) that contain CxxC motifs bind Zn. These constructs along with TM of MBP-BECN1 also showed to contain other elements in either considerable or trace quantities. Na, B, S, K, and Ca were found in considerable quantities and Mg, Fe, Ni and Cu as trace contaminants, even though elements in the corresponding buffer were subtracted, perhaps because the buffer used for subtraction was the SEC buffer rather than the SEC column flow-through. Of the elements found Na and B are common contaminants from glassware and S may be released from the cysteines and methionine residues during the initial HNO_3 treatment or due to the use of 2 mM DTT in the elution buffer for amylose affinity chromatography. MBP-BECN1^{CysDM2} showed only negligible amount of sulfur compared to the other MBP-BECN1 proteins, perhaps because the sample may have precipitated due to HNO_3 treatment before completely releasing the covalently bonded sulfur atoms. Mg, K, Cu, Ca, and Ni are likely contaminants in the water, or the salt used during buffer preparation (Figure 4.16, Table 4.6).

EDTA was used during the initial purification steps of MBP-BECN1 proteins to prevent protein degradation. EDTA treatment may result in the loss of Zn^{2+} bound by the protein. Therefore, after the affinity column was washed to remove contaminating proteins, the protein bound to the affinity column was washed with a Zn^{2+}-containing buffer to replace any Zn^{2+} that may have been stripped out. Subsequent washes using a Zn^{2+}-free buffer removed any unbound Zn^{2+}. Zn^{2+} was also not included in buffers used for all subsequent purification steps, ensuring that any Zn^{2+} detected by ICP-MS was clearly associated with the protein.

Protein purified using this protocol bound Zn in a 1:1 molar ratio (Figure 4.16, Table 4.6) in MBP-BECN1. Strikingly, mutation of either of the CxxC motifs reduces Zn binding by ~50% in FL BECN1 ($p = 0.002$ for MBP-BECN1^{CysDM1} versus MBP-BECN1; $p = 0.006$ for MBP-BECN1^{CysDM2} versus MBP-BECN1), while Zn binding is completely abrogated when both CxxC motifs are mutated ($p = 0.002$ for MBP-BECN1^{CysTM} versus MBP-BECN1). We also determined that deletion of BECN1 IDR residues 31-123 did not impact Zn binding ($p = 0.69$ for MBP-BECN1(Δ31-123) versus MBP-BECN1) (Table 4.6, Figure 4.16). Thus, of the 20 elements tested by quantitative ICP-MS, Zn was the only ion found to be dependent on the presence of the BECN1 CxxC motifs, presumably because of binding.

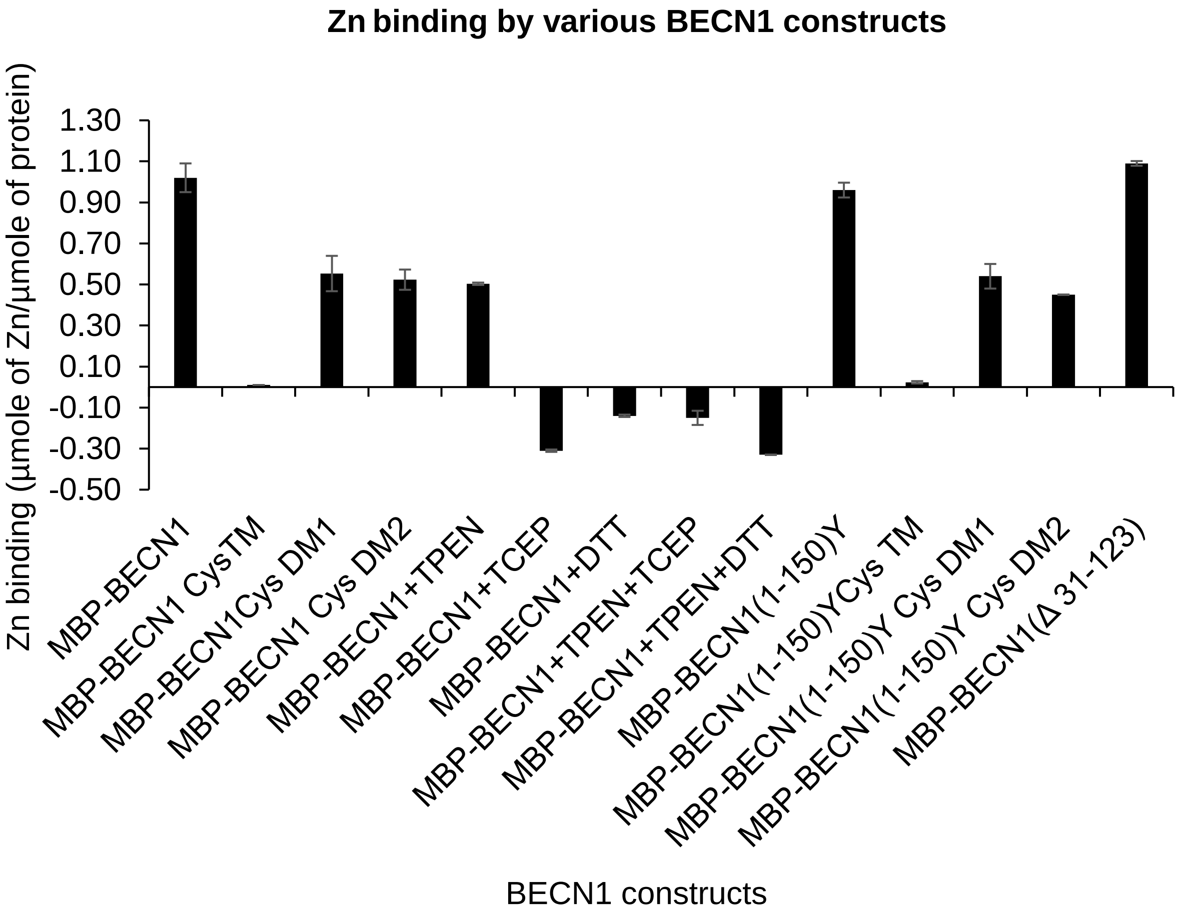


Figure 4.16: ICP-MS quantification of Zn-bound by WT or mutant BECN1. The abscissa indicates the various BECN1 mutants and the treatments tested, while the ordinate quantifies the µmoles of Zn bound / µmole of protein. All proteins used in these experiments were MBP-tagged.

4.3.2.2. Element Analyses for His_6-BECN1(1-150)Y constructs

The ICP-MS analyses indicated that WT, DM1 and DM2 His_6-BECN1(1-150)Y proteins that contain CxxC motifs bind Zn. These constructs along with TM of His_6-BECN1(1-150)Y (Table 4.7) also showed to contain other elements in either considerable or trace quantities. B and S were found in significant quantities and other elements such as Ca, Fe and Ni (Table 4.7) were found in trace amounts. In addition to the element sources listed in section 4.3.2.1, P was found in trace amounts, which may have been introduced into the sample from the water or salt used during buffer preparation. Ni may have also leached from the resin used for the affinity purification.

Table 4.7: ICP-MS results for His$_6$-BECN1(1-150)Y constructs

Elements screened		**His$_6$-BECN1(1-150)Y**		**His$_6$-BECN1(1-150)Y^{CysTM}**		**His$_6$-BECN1(1-150)Y^{CysDM1}**		**His$_6$-BECN1(1-150)Y^{CysDM2}**	
Element / Isotope	Atomic wt (g/mole)	Avg	SD	Avg	SD	Avg	SD	Avg	SD
		(µmole/µmole)							
Li / 7	6.94	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
B / 11	10.81	0.53	0.01	0.48	0.01	1.19	0.07	0.49	0.02
Na / 23	22.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mg / 24	24.31	0.01	0.00	0.00	0.00	0.02	0.00	0.01	0.00
P / 31	30.97	0.04	0.01	0.03	0.00	0.10	0.01	0.03	0.00
S / 34	32.07	6.63	0.44	3.40	0.06	4.86	0.50	5.46	0.38
K / 39	39.10	0.00	0.00	0.00	0.00	0.21	0.01	0.00	0.00
Ca / 40	40.08	0.02	0.00	0.01	0.00	0.01	0.00	0.02	0.00
V / 51	50.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cr / 52	52.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mn/55	54.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Fe / 56	55.85	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Co / 59	58.93	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ni / 60	58.69	0.00	0.00	0.01	0.00	0.02	0.00	0.01	0.00
Cu / 63	63.55	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Zn / 66	65.39	0.70	0.01	0.06	0.00	0.58	0.06	0.69	0.04
As / 75	74.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Se / 78	78.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mo / 95	95.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cd / 111	112.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

ICP analyses of His$_6$-BECN1(1-150)Y indicates that the sample contains Zn in significant quantities (Figure 4.16, Table 4.7). CxxC motifs are primarily implicated in binding Zn (Castagnetto et al., 2002). His$_6$-BECN1(1-150)Y binds 0.7 µmole of Zn/ µmole of protein (Figure 4.17, Table 4.7). The lack of 1:1 Zn-binding as seen for the MBP-BECN1 protein may be due to the use of β-mercaptoethanol in the buffer during purification by Ni-affinity and anion exchange chromatography, which may reduce the cysteine residues resulting in the loss of Zn binding. Mutation of the first CxxC motif in the His$_6$-BECN1(1-150)Y construct decreases Zn binding by 12% ($p = 0.06$ for His$_6$-BECN1(1-150)Y^{CysDM1} versus His$_6$-BECN1(1-150)Y), while there was no significant difference in Zn binding by His$_6$-BECN1(1-150)Y^{CysDM2} (p = 0.78 for His$_6$-BECN1(1-150)Y^{CysDM2} versus His$_6$-BECN1(1-150)Y). However, as expected, Zn binding is completely abrogated when both CxxC motifs are mutated (p = 2.453 x 10-6 for His$_6$-BECN1(1-150)Y^{CysTM} versus His$_6$-BECN1(1-150)Y). Thus, of the 20 elements tested by quantitative ICP-MS, Zn was the only element found to be dependent on the presence of the BECN1 CxxC motifs, presumably because of binding (Figure 4.17, Table 4.7). Since, unlike the MBP-tagged FL proteins, His$_6$-BECN1(1-150)Y did not bind Zn in equimolar ratios, we purified MBP-BECN1(1-150)Y constructs to test for Zn binding.

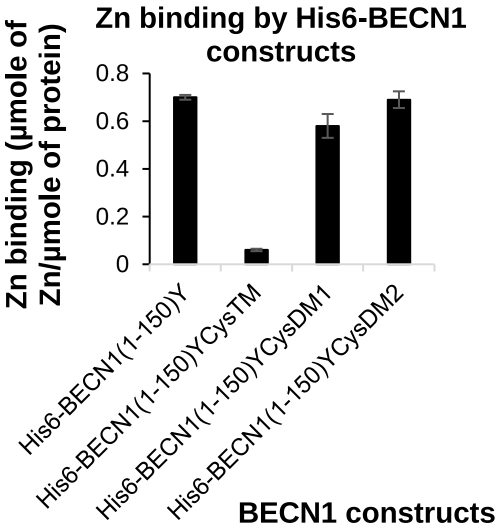


Figure 4.17: ICP-MS quantification of Zn-bound by His$_6$-tagged WT or mutant BECN1. The abscissa indicates the various BECN1 mutants and the treatments tested, while the ordinate quantifies the µmoles of Zn bound / µmole of protein.

4.3.2.3. Element analyses for MBP-BECN1(1-150)Y constructs

ICP-MS analyses indicated that WT, DM1 and DM2 MBP-BECN1(1-150)Y proteins that contain CxxC motifs bind Zn. These constructs along with TM of MBP-BECN1(1-150)Y also showed presence of other elements such as Na, B, S, K in considerable quantities and P, Ca, Mg, Se, Ni, Cu, Fe in trace quantities (Figure 4.16, Table 4.8), likely for the same reasons outlined in sections 4.3.2.1 and 4.3.2.2. In addition to the elements listed before, Li found in these proteins was likely from glassware, and trace amounts of Cd may be from the water and salt used for buffer preparation.

Table 4.8: ICP-MS results for MBP-BECN1(1-150)Y constructs

Elements screened		**MBP BECN1(1-150)Y**		**MBP BECN1(1-150)Y^{CysTM}**		**MBP BECN1(1-150)Y^{CysDM1}**		**MBP BECN1(1-150)Y^{CysDM2}**	
Isotope of elements	Atomic wt (g/mole)	Avg	SD	Avg	SD	Avg	SD	Avg	SD
		µmole/µmole							
Li/ 7	6.94	0.11	0.23	0.00	0.00	0.01	0.02	0.00	0.01
B/ 11	10.81	0.16	0.01	0.42	0.01	0.04	0.02	0.01	0.01
Na/ 23	22.99	0.00	0.00	0.00	0.00	0.35	0.34	0.00	0.00
Mg/ 24	24.31	0.01	0.00	0.02	0.00	0.00	0.00	0.03	0.00
P/ 31	30.97	0.04	0.11	0.00	0.00	0.00	0.00	0.00	0.00
S/ 34	32.07	4.90	5.30	1.55	0.28	1.20	0.21	0.00	0.00
K/ 39	39.10	0.38	0.34	0.38	0.05	0.00	0.00	0.12	0.02
Ca / 40	40.08	0.04	0.01	0.10	0.00	0.00	0.00	0.11	0.01
V / 51	50.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cr / 52	52.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mn/55	54.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Fe / 56	55.85	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.00
Co / 59	58.93	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ni / 60	58.69	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00
Cu / 63	63.55	0.01	0.00	0.01	0.00	0.00	0.00	0.01	0.00
Zn / 66	65.39	0.96	0.04	0.02	0.00	0.55	0.07	0.45	0.00
As / 75	74.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Se / 78	78.96	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00
Mo / 95	95.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cd / 111	112.40	0.00	0.00	0.00	0.00	0.04	0.02	0.00	0.00

Unlike the His-BECN1(1-150)Y constructs, where to avoid chelation of Ni^{2+} during Ni-NTA affinity purification, EDTA was not used to prevent degradation by metalloproteases; and similar to the FL MBP-BECN1 constructs (section 4.3.2.1), EDTA was used in buffers for the initial purification stages to prevent degradation of the MBP-BECN1(1-150)Y proteins. Subsequently, Zn^{2+} that was chelated by EDTA was restored by washing the affinity column with Zn^{2+} containing buffer. Excess Zn^{2+} was removed by washing with Zn^{2+} free buffer prior to elution, and exclusion of Zn^{2+} from buffers used for subsequent purification steps, ensuring that Zn^{2+} detected by ICP-MS was actually protein-bound.

Protein purified using this protocol bound Zn in a 0.96:1 molar ratio (Figure 4.16) of MBP-BECN1(1-150)Y, which is close to 1:1 molar ratio of Zn to protein binding obtained for MBP-BECN1 ($p = 0.29$ for MBP-BECN1(1-150)Y versus MBP-BECN1). This indicates that that BECN1 domains beyond the IDR do not impact Zn-binding. Similar to MBP-BECN1 CxxC mutation constructs, mutation of either of the CxxC motifs reduces Zn binding by ~50% in BECN1(1-150)Y ($p = 0.002$ for MBP-BECN1(1-150)Y$^{\text{CysDM1}}$ versus MBP-BECN1(1-150)Y; $p = 0.002$ for MBP-BECN1(1-150)Y$^{\text{CysDM2}}$ versus MBP-BECN1(1-150)Y), while Zn binding is completely abrogated when both CxxC motifs are mutated ($p = 0.0005$ for MBP-BECN1(1-150)Y$^{\text{CysTM}}$ versus MBP-BECN1(1-150)Y). Thus, of the 20 elements tested by quantitative ICP-MS, Zn was the only ion found to be dependent on the presence of the BECN1 CxxC motifs.

4.3.2.4. Element analyses for MBP-BECN1 constructs treated with TPEN, TCEP or DTT

Next, we assessed the impact of chelating agents and reducing agents on Zn-binding (Table 4.9, Figure 4.16).

Table 4.9: ICP-MS results for MBP-BECN1 constructs treated with TPEN, TCEP or DTT

Elements screened		**MBP-BECN1 +TPEN**		**MBP-BECN1 +TCEP**		**MBP-BECN1 +DTT**		**MBP-BECN1 +TPEN+ TCEP**		**MBP-BECN1 +TPEN+DTT**	
Element / Isotope	Atomic wt (g/mole)	Avg	SD	Avg	SD	Avg	SD	Avg	SD	Avg	SD
		(µmole/µmole)									
Li/7	6.94	0.06	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
B/11	10.81	0.01	0.01	0.00	0.01	0.00	0.01	0.01	0.00	0.01	0.01
Na/23	22.99	712.99	52.34	0.11	0.10	820.65	97.26	2100.78	76.33	2810.34	255.61
Mg/24	24.31	0.00	0.00	0.00	0.00	0.00	0.00	0.39	0.08	0.00	0.00
P/31	30.97	0.11	0.00	0.16	0.01	0.00	0.00	9.06	0.17	0.00	0.00
S/34	32.07	0.00	0.00	0.00	0.00	625.27	41.45	0.00	0.00	82.64	42.90
K/39	39.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ca/40	40.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
V/51	50.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cr/52	52.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mn/55	54.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Fe/56	55.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Co/59	58.93	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ni/60	58.69	56.50	47.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cu/63	63.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zn/ 66	65.39	0.50	0.00	-0.32	0.00	-0.14	0.01	-0.15	0.03	-0.33	0.00
As/ 75	74.92	0.00	0.00	1.66	1.62	0.00	0.00	0.00	0.00	0.00	0.00
Se / 78	78.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mo/ 95	95.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cd / 111	112.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

The samples were found to contain considerable quantities of elements such as Na, Mg, S, P, Ni and As, whereas Li and B were found in trace quantities (Table 4.9). In addition to the reasons outlined in sections 4.3.2.1 and 4.3.2.3, S was found in really high amount for samples treated with 50mM DTT, Ni and Mg was found to be high in MBP-BECN1 sample treated with either TPEN or TPEN with TCEP likely because of chelation of ions by TPEN. P was mainly found in considerable concentrations in samples treated with TCEP and in trace quantities in samples treated with only TPEN. TCEP treatment may also have introduced As (Figure 4.16, Table 4.9).

Treatment of the WT protein with 20-fold molar excess of the Zn-chelating agent TPEN, reduces Zn-binding by 50% ($p = 0.099 \times 10^{-2}$ for MBP-BECN1+TPEN versus MBP-BECN1). Treatment with reducing agents such as 208-fold molar excess of TCEP or 4166-fold molar excess DTT removes all Zn ($p = 0.032 \times 10^{-2}$ and $p = 0.039 \times 10^{-2}$ in MBP-BECN1+ TCEP or DTT respectively versus MBP-BECN1). Simultaneous treatment with TPEN and either reducing agent produced results similar to treatment by reducing agent alone ($p = 0.002 \times 10^{-2}$ and $p = 0.045 \times 10^{-2}$ for MBP-BECN1+TPEN, in presence of TCEP or DTT respectively versus MBP-BECN1). As the element concentrations in samples are determined after buffer subtraction, the negative Zn concentrations for samples treated with TCEP, DTT, TPEN+TCEP and TPEN+DTT indicates that the dialysis buffer used for buffer subtraction contained ~ 3-fold more Zn than the samples, even after three buffer exchanges.

4.3.3. Investigating secondary structure content of the BECN1 IDR using CD

We used CD to determine if Zn^{2+}-binding impacts the secondary structure content of the BECN1 IDR. As IDRs often fold upon binding to binding partners/ligands(Habchi *et al.*, 2014), we also used CD to investigate changes in the secondary structure content of the BECN1 IDR in the presence of increasing concentrations of 2,2,2-trifluroethanol (TFE), which is often used to mimic hydrophobic interaction environments and to probe the helical propensity of IDRs.

4.3.3.1. WT and mutant His_6-BECN1(1-150)Y proteins

We evaluated the impact of Zn^{2+}-binding on the secondary structure content of the BECN1 IDR by comparing CD spectra of Zn^{2+}-bound His_6-BECN1(1-150)Y and Zn^{2+}-depleted His_6-BECN1(1-150)Y^{CysTM} at 4°C and at 20°C (Figure 4.18 and Table 4.10 and Table 4.11) .

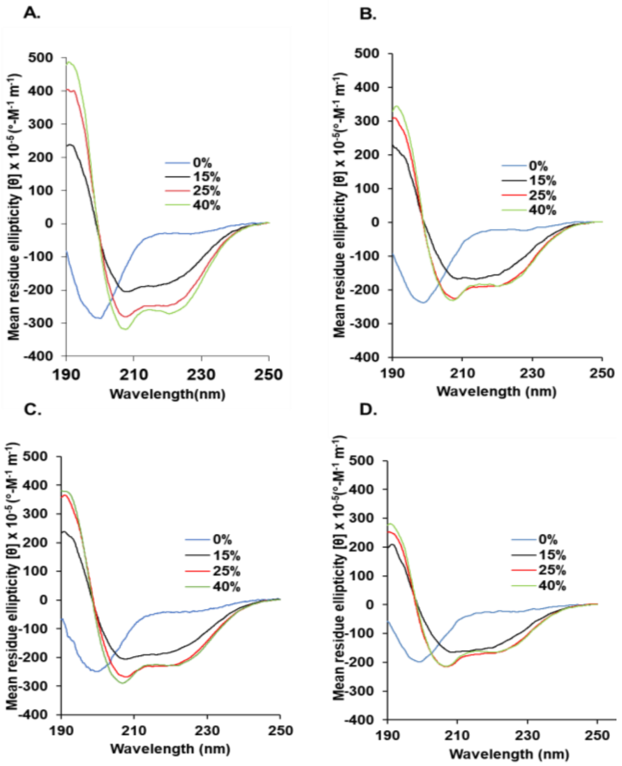


Figure 4.18: Secondary structure content of the BECN 1 IDR and its modulation upon addition of TFE. CD spectra were recorded at 4°C (top panels) and 20 °C (bottom panels) in the presence of different TFE concentrations as indicated, for: A,C) His$_6$-BECN1(1-150)Y and B,D) His$_6$-BECN1(1-150)Y^{CysTM}

Table 4.10: Secondary structural estimation of His$_6$-BECN1(1-150)Y constructs at 4°C

Protein (165 residues)	**TFE treatment**	**Secondary structure estimation (%)**		
		Helix	**Strand**	**Coil**
His$_6$-BECN1(1-150)Y	0%	5±1	4±6	91±7
	15%	46±5	11±4	43±5
	25%	78±0	6±0	17±0
	40%	78±10	2±1	20±12
His$_6$-BECN1(1-150)Y^{CysTM}	0%	3±0	6±9	91±3
	15%	38±7	16±0	46±7
	25%	53±6	6±2	42±8
	40%	57±6	5±2	37±8

Table 4.11: Secondary structural estimation of His$_6$-BECN1(1-150)Y constructs at 20°C

Protein (165 residues)	**TFE treatment**	**Secondary structure estimation (%)**		
		Helix	**Strand**	**Coil**
His$_6$-BECN1(1-150)Y	0%	7±0	4±6	88±6
	15%	47±6	10±6	43±11
	25%	62±7	5±7	34±9
	40%	71±14	3±1	25±15
His$_6$-BECN1(1-150)Y^{CysTM}	0%	4±1	12±19	83±3
	15%	37±6	17±2	46±7
	25%	50±6	6±2	44±8
	40%	52±5	7±4	41±9

A quantitative estimation of the secondary structure elements from the CD spectra at 4°C (Figure. 4.18 A) for His$_6$-BECN1(1-150)Y fragment indicates it has 8 residues in a helical conformation which can form two stable helical turns, 7 in extended or β-conformation, and 150 in coil conformation (Table 4.10). His$_6$-BECN1(1-150)Y^{CysTM} (Figure. 4.18 B), which lacks bound Zn^{2+}, has 5 residues in helical conformation which can form only one stable helical turn, while there are 10 residues in the extended β-conformation and 149 residues in the coil conformation respectively. Thus, helicity of the BECN1 IDR is marginally increased when Zn^{2+} is bound by the CxxC motifs.

Addition of as little as 15% TFE to His$_6$-BECN1(1-150)Y fragment dramatically increases the number of residues in helical conformation from 8 to 76, at the expense of residues in coil conformation which is reduced to 71. The number of residues in the extended, β-conformation also increases to 18. Further addition of TFE to 25% increases helicity of the His$_6$-BECN1(1-150)Y fragment to 128, but there is no further increase when the concentration of TFE was increased to 40%; and in both cases there is a concomitant decrease in number of residues in the β-conformation to 10 and 4 respectively, as well as a decrease in the number of residues in coil conformation to 28 and 32, respectively. Similarly, addition of 15% TFE to the His$_6$-BECN1(1-150)Y^{CysTM} fragment also increases the number of residues in helical conformation from 5 to 62, also at the expense of residues in the coil conformation, which are reduced in number to 76. The number of residues in the extended, β-conformation also increase to 27. Further addition of TFE to 25% and 40% increases helicity of the His$_6$-BECN1(1-150)Y^{CysTM} fragment further to 87 and 95 residues respectively, with a concomitant decrease in the number of residues in β-conformation to 10 and 9 respectively, and a decrease in the number of residues in the coil conformation is decreased to 69 and 61 respectively.

Similar quantitative estimation of the secondary structure elements from the CD spectra at 20°C (Figure. 4.18 C) for His$_6$-BECN1(1-150)Y fragment indicates it has 11 residues in a helical conformation which can form also form three stable helical turns and 7 residues in extended or β-conformation, and 146 in coil conformation (Table 4.11). His$_6$-BECN1(1-150)Y^{CysTM} (Figure. 4.18 D), which lacks bound Zn^{2+}, has 7 residues in helical conformation which can form two stable helical turns; while there are 21 residues in the extended β-conformation and 137 residues in the coil conformation respectively. Thus, helicity of the BECN1 IDR is marginally increased when Zn^{2+} is bound by the CxxC motifs.

Addition of as little as 15% TFE to His$_6$-BECN1(1-150)Y (fragment dramatically increases the number of residues in helical conformation, from 11 to 77, at the expense of residues in coil conformation which is reduced to 72. The number of residues in the extended, β-conformation also increases to 18. Further addition of TFE to 25% and 40% increases helicity of the His$_6$-BECN1(1-150)Y fragment to 102 and 117 residues respectively, with a concomitant decrease in number of residues in the β-conformation to 8 and 5 respectively, and a decrease in the number of residues in coil conformation to 55 and 42, respectively. Addition of 15% TFE to the His$_6$-BECN1(1-150)Y^{CysTM} fragment also increases the number of residues in helical conformation, from 7 to 61, also at the expense of residues in the coil conformation which are reduced in number to 75. The number of residues in the extended, β-conformation also increases to 28. Further addition of TFE to 25% and 40% increases helicity of the His$_6$-BECN1(1-150)Y^{CysTM} fragment further to 61 and 82 residues respectively, with a concomitant decrease in the number of residues in β-conformation to 10 and 11 respectively, and a decrease in the number of residues in the coil conformation to 73 and 68, respectively.

Thus, we find that while TFE increases helicity of both, the His$_6$-BECN1(1-150)Y and the His$_6$-BECN1(1-150)Y^{CysTM} fragments, elimination of Zn^{2+}-binding appears to reduce binding-induced disorder-to-helix transitions.

4.3.3.2. WT and mutant BECN1(1-150)Y proteins

His$_6$-BECN1(1-150)Y does not bind Zn^{2+} in equimolar ratios, unlike MBP-BECN1(1-150)Y. Therefore, we repeated the CD experiments using untagged proteins purified using the MBP tag (Figure 4.19 and Table 4.12 and Table 4.13).

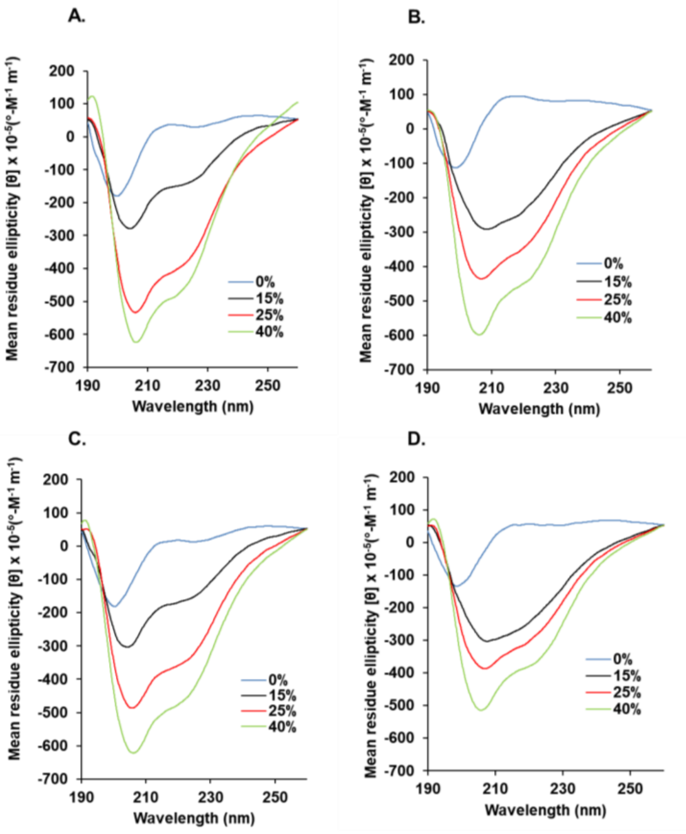


Figure 4.19: Secondary structure content of the BECN 1 IDR and its modulation upon addition of TFE. CD spectra were recorded at 4°C (top panels) and 20 °C (bottom panels) in the presence of different TFE concentrations as indicated, for: A,C) BECN1(1-150)Y and B,D) BECN1(1-150)Y^{CysTM}.

Table 4.12: Secondary structural estimation of BECN1(1-150)Y constructs at 4°C

Protein (156 residues)	**TFE treatment**	**Secondary structure estimation (%)**		
		Helix	**Strand**	**Coil**
BECN1(1-150)Y	0%	12±7	10±14	77±22
	15%	40±15	14±7	44±5
	25%	64±0	7±8	29±7
	40%	74±12	7±12	19±2
BECN1(1-150)Y^{CysTM}	0%	0	65	34
	15%	40±2	15±17	46±17
	25%	60±12	9±12	32±17
	40%	70±6	7±9	23±3

Table 4.13: Secondary structural estimation of BECN1(1-150)Y constructs at 20°C

Protein (156 residues)	**TFE treatment**	**Secondary structure estimation (%)**		
		Helix	**Strand**	**Coil**
BECN1(1-150)Y	0%	9±12	37±6	54±18
	15%	47±10	6±7	46±17
	25%	66±2	8±11	26±13
	40%	73±8	7±10	20±2
BECN1(1-150)Y^{CysTM}	0%	0±0	43±6	58±8
	15%	50±15	8±6	42±20
	25%	57±14	9±9	34±22
	40%	66±1	8±4	26±5

A quantitative estimation of the secondary structure elements from the CD spectra at 4°C (Figure. 4.19 A) for BECN1(1-150)Y fragment indicates it has 19 residues in a helical conformation, which can form 5 stable helical turns, 16 in extended or β-conformation, and 121 in coil conformation (Table 4.12). BECN1(1-150)Y^{CysTM} (Figure. 4.19 B), which lacks bound Zn^{2+}, has no residues in helical conformation, while there are 102 residues in the extended β-conformation and 54 residues in the coil conformation, respectively. Thus, abrogation of Zn^{2+}-binding decreases helicity of the BECN1 IDR.

Addition of as little as 15% TFE to the BECN1(1-150)Y fragment dramatically increases the number of residues in helical conformation from 19 to 62, at the expense of residues in coil conformation, which are reduced to 68. The number of residues in the extended, β-conformation also increases to 22. Further addition of TFE to 25% and 40% increases helicity of the BECN1(1-150)Y fragment further to 100 and 115 residues respectively, with a concomitant decrease in number of residues in the β-conformation to 11 in both cases respectively, and a decrease in the number of residues in coil conformation to 46 and 30 respectively. Similarly, addition of 15% TFE to the BECN1(1-150)Y^{CysTM} fragment also increases the number of residues in helical conformation from 0 to 62, also at the expense of residues in extended, β-conformation which are decreased to 24 and also residues in coil conformation in which are reduced in number to 72. Further addition of TFE to 25% and 40% increases helicity of the

BECN1(1-150)Y fragment further to 93 and 109 residues respectively, with a concomitant decrease in the number of residues in β-conformation to 13 and 10 respectively, and a decrease in the number of residues in coil conformation to 50 and 36 respectively. Thus, we find that while TFE increases helicity of both, the BECN1(1-150)Y and the BECN1(1-150)Y^{CysTM} fragments, elimination of Zn^{2+}-binding appears to reduce binding-induced disorder-to-helix transitions.

A similar quantitative estimation of the secondary structure elements from the CD spectra at 20°C (Figure 4.19 C) for BECN1(1-150)Y fragment indicates it has 14 residues in a helical conformation, which can form 4 stable helical turns, 58 in extended or β-conformation, and 84 in coil conformation (Table 4.13). BECN1(1-150)Y^{CysTM} (Figure. 4.19 D) which lacks bound Zn^{2+}, has no residues in helical conformation, while there are 67 residues in the extended β-conformation and 91 residues in the coil conformation. Thus, abrogation of Zn^{2+}-binding decreases helicity of the BECN1 IDR.

Addition of as little as 15% TFE to the BECN1(1-150)Y fragment dramatically increases the number of residues in helical conformation from 14 to 74, at the expense of residues in the extended, β-conformation and coil conformation which are reduced to 10 and 72 respectively. Further addition of TFE to 25% and 40% increases helicity of the BECN1(1-150)Y fragment further to 103 and 114 residues respectively, with a concomitant decrease in number of residues in the β-conformation to 12 and 11 respectively, and decrease in the number of residues in coil conformation to 41 and 31 respectively. Addition of 15% TFE to the BECN1(1-150)Y^{CysTM} fragment also increases the number of residues in helical conformation from 0 to 78, also at the expense of residues in extended, β-conformation and coil conformation, which are reduced to 13 and 66 respectively. Further addition of TFE to 25% and 40% increases helicity of the

BECN1(1-150)Y fragment further to 90 and 103 residues respectively, with a concomitant decrease in the number of residues in β-conformation to 14 and 12 respectively, and a decrease in the number of residues in the coil conformation to 53 and 40 respectively. Thus, we find that while TFE increases helicity of both, the BECN1(1-150)Y and the BECN1(1-150)Y^{CysTM} fragments, elimination of Zn^{2+}-binding appears to reduce binding-induced disorder-to-helix transitions. This conclusion is similar to that obtained for the His_6-tagged proteins.

A comparison of the secondary structure elements of BECN1(1-150)Y with the His_6-BECN1(1-150)Y reveals that at 0% TFE both constructs have similar helicity, but compared to His_6-BECN1(1-150)Y, BECN1(1-150)Y has a higher number of residues in the β-extended conformation and fewer residues in the coil conformation. Similarly, BECN1(1-150)Y^{CysTM} has helical content comparable to His_6-BECN1(1-150)Y^{CysTM}, but has more residues in the extended β- conformation and fewer residues in the coil conformation than His_6-BECN1(1-150)Y^{CysTM}. Notably, both BECN1(1-150)Y and His_6-BECN1(1-150)Y constructs show comparable binding-induced helicity in the presence and absence of Zn^{2+}-binding by CxxC motifs.

4.3.4. SAXS analysis of the BECN1 IDR constructs

4.3.4.1. His_6-BECN1(1-150)Y

The integrated scattering intensity for His_6-BECN1(1-150)Y plotted across scattering frames (Figure 4.20 A) shows a profile similar to the SEC elution profile (Figure 4.2) including the presence of a shoulder preceding the peak, where the peak and shoulder correspond to 82% and 18% of the total protein. Notably however, the R_g distribution decreases across the scattering frames (Figure 4.20 A), suggesting that the shoulder contains a species with higher R_g. We hypothesized that the leading edge may correspond to either an oligomer or an extended conformation, followed by either a monomeric or more compact species within the peak. Due to the possibility of having different species in the SEC elution, the SAXS scattering profile of His_6-BECN1(1-150)Y sample was deconvoluted using EFA.

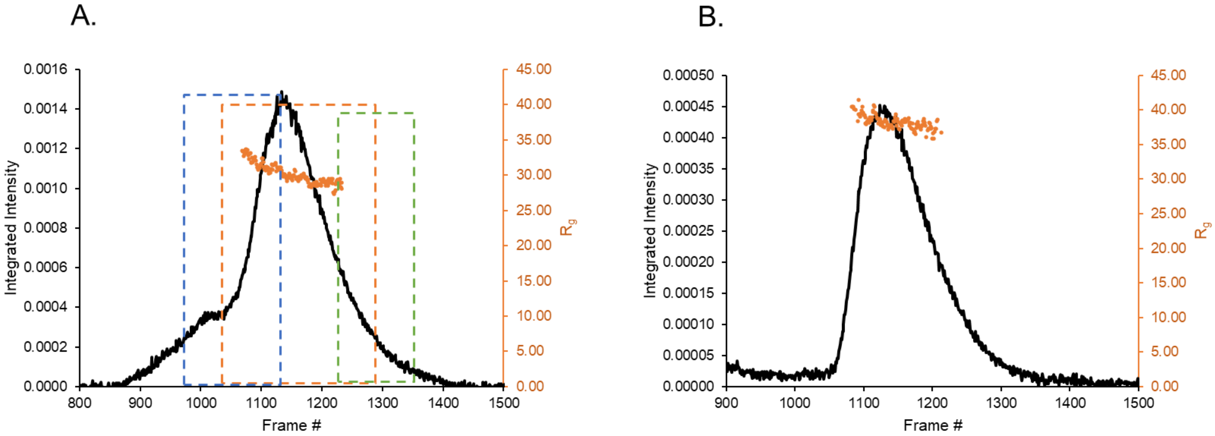


Figure 4.20: EFA Analysis of His_6-BECN1(1-150)Y and total SAXS scattering intensity of His_6-BECN1(1-150)Y^{CysTM}. A, B) Plots showing total scattered intensity (left y-axis) and Guinier R_g (right y-axis) across subsequently measured scattering profiles (frames) of His_6-ECN1(1-150)Y and His_6-BECN1(1-150)Y^{CysTM} respectively. BioXTAS RAW (version 1.6.4) was used to select EFA data ranges for His_6-BECN1(1-150)Y Components 0, 1, and 2 (blue, orange and green respectively) which are indicated by dashed boxes.

Three potential component species were identified in the scattering profile of His_6-BECN1(1-150)Y namely Component 0, 1 and 2. Component 0 was identified from the leading edge shoulder of the His_6-BECN1(1-150)Y peak comprising scattering frames 965 to 1129 (Figure 4.20 A), Component 1 was identified from the peak comprising of scattering frames 1035 to 1298 (Figure 4.20 A), Component 2 was identified from the trailing edge of the peak comprising of scattering frames 1233 to 1334 (Figure 4.20 A). Component 0 and Component 1 were analyzed further, however the third EFA component, Component 2 displayed a low signal to noise as shown by negative intensities at low q (Figure 4.21 C), indicative of a very low concentration, or small molecular mass or both, which prevented further analysis.

The P(r) distribution curve for Component 0 has an extended tail characteristic of disordered proteins (Kikhney & Svergun, 2015), and provides R_g and D_{max} estimates of ~39.7 ±

0.7 and 180 Å respectively (Figure 4.21 E, H and Table 4.14). Similarly, the P(r) distribution curve of Component 1 is more symmetrical, but also shows an extended tail characteristic of disorder, with R_g and D_{max} estimates of ~30.1 Å ± 0.1 and 160 Å respectively (Figure 4.21 F, H and Table 4.14). The higher R_g and D_{max} for Component 0 compared to Component 1 indicates that Component 0 is larger in size than Component 1, and a comparison of the shape of P(r) distribution curves indicates a more extended structure for Component 0. Dimensionless Kratky plots (Fig 4.21 I) shows that both Component 0 and Component 1 show a shallower trend than expected for completely disordered proteins, which is indicative of partial disorder (Figure 4.21 I).

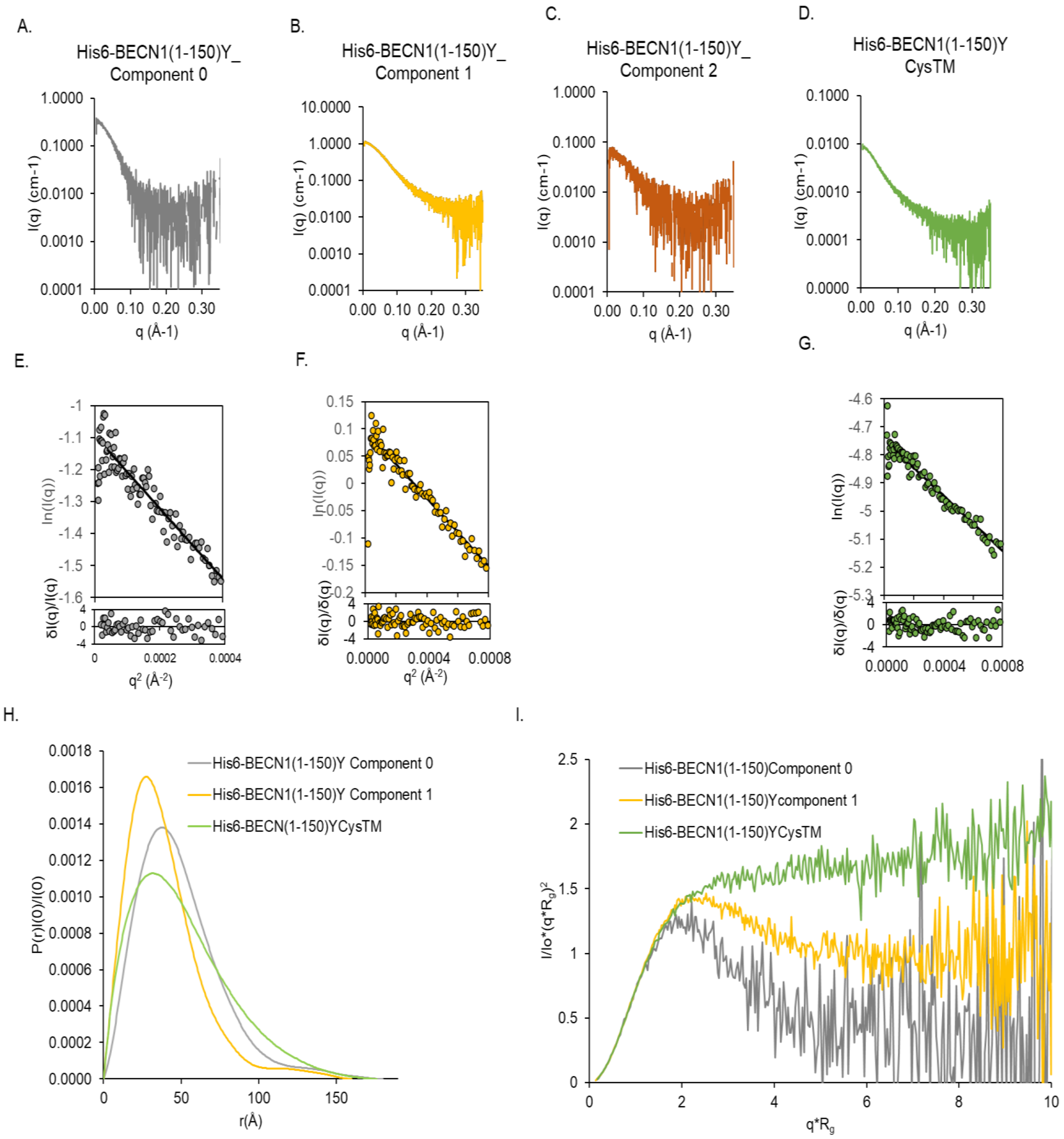


Figure 4.21: SAXS Analyses for His$_6$-BECN1(1-150)Y Component 0 (gray data and curves), His$_6$-BECN1(1-150)Y Component 1 (yellow data and curves), Component 2 (brown curve) and His$_6$-BECN1(1-150)Y$^{\text{CysTM}}$ (green data and curves). A-D) Scattering profiles, E-G) Guinier Plots. Top panel: the black lines represent fits to the data represented by gray, yellow and green points respectively; Bottom panel: corresponding fit residuals. H) P(r) function. I) Kratky Plot. Dashed lines indicate the peak expected for a compact globular protein

Table 4.14: SEC-SAXS Data Analysis for His$_6$-BECN1(1-150)Y constructs

	His$_6$-BECN1 (1-150)Y Component 0	**His$_6$-BECN1 (1-150)Y Component 1**	**His$_6$-BECN1(1-150)Y^{CysTM}**
Guinier Analysis			
I(0) (cm^{-1})	0.335 ± 0.002	1.102 ± 0.002	0.00874 ± 0.004 x 10^{-2}
R_g (Å)	36.6 ± 0.4078	30.6 ± 0.1731	39.3 ± 0.336
q-range (Å^{-1})	0.01019-0.03284	0.00651-0.03114	0.00821-0.03058
qR_g max	1.20	0.952	1.20
r^2 fit	0.934	0.977	0.975
M from V_c (ratio to theoretical) (kDa)	81.5 (4.45)	24.1 (1.32)	23.1 (1.26)
P(r) analysis			
I(0) (cm^{-1})	0.3411 ± 0.0025	1.100 ± 0.002	0.0071 ± 0.3526
R_g (Å)	39.7 ± 0.7548	30.1 ± 0.125	41.9 ± 0.4558
dmax (Å)	180	160	176
q-range (Å^{-1})	0.0042-0.3526	0.0042-0.3526	0.0042-0.3526
Total Estimate	0.7154	0.7184	0.7029
χ^2	2.63	2.51	1.16

Molecular mass estimation from V_c (Rambo & Tainer, 2013) indicates that His$_6$-BECN1(1-150)Y Component 0 has a molecular mass of 81.5 kDa which would correspond to an oligomer and Component 1 has a molecular mass of 24.1 kDa which corresponds to a monomer. However, molecular mass estimations from V_c are notoriously inaccurate for disordered proteins (Hajizadeh *et al.*, 2018). Further, V_c estimation is dependent on having an overall high signal to noise ratio in the scattering intensity profile. Unfortunately, Component 0 has a low signal to noise ratio in the scattering intensity profile, likely because of a low concentration at the leading edge shoulder as it constitutes only 18% of the total protein eluted from SEC, and noisy due to the influence of small amount of oligomeric species at the void volume, that may have impacted the scattering. Therefore, the V_c calculation, and any inference of the oligomeric state of Component 0 is not reliable. This suggests that Component 0 either corresponds to an oligomer, or more likely exists in a more extended state relative to Component 1.

EOM was used to generate a random pool of conformers representing Component 0 of His$_6$-BECN1(1-150)Y, and select an ensemble of conformers that best fit the scattering data. The selected ensemble of His$_6$-BECN1(1-150)Y Component 0 conformers has an average R_g of 37.6

Å and D_{max} of 115 Å, while the randomly-generated EOM pool has an average R_g of 36.6 Å and D_{max} of 113 Å, indicating that the selected conformers are as large and extended as all possible conformations in the pool. The selected conformer ensemble fits the experimental ensemble with a χ^2 of 2.56, and the normalized residual is flat but less randomly distributed at the high q region, indicative of a poor fit. Indeed, the conformational flexibility of Component 0 cannot be reliably modeled in the absence of information about the oligomeric state of Component 0, and therefore this data is not useful for any further analysis.

EOM was also used to generate a random pool of conformers representing Component 1 of His_6-BECN1(1-150)Y, and select an ensemble of conformers that best fit the scattering data. The selected ensemble of His_6-BECN1(1-150)Y Component 1 conformers has a narrower average R_g and D_{max} distribution compared to that of the randomly-generated EOM pool, indicating that Component 1 conformations (Figure 4.22A, C and Table 4.15) are more compact relative to all possible conformations. R_{flex} for the Component 1 His_6-BECN1(1-150)Y is ~ 70.1% for the selected ensemble, significantly smaller than the 89.2 % for the pool, which suggests that the conformer ensemble is significantly less flexible than would be expected for a completely disordered protein. Additionally, the fit of the EOM selected conformer ensemble to the experimental data for His_6-BECN1(1-150)Y Component 1 has a χ^2 of 2.49 (Figure 4.22 E, Table 4.15). While the χ^2 is somewhat large, the normalized residual is flat and randomly distributed, indicating a good fit. Perhaps the uncertainty from the EFA deconvolution is underestimated, resulting in a larger than expected χ^2.

Thus, together the SEC and the SAXS EFA and EOM analyses show that the His_6-BECN1(1-150)Y protein is a partially-disordered monomer, with an extended conformation, and appears to transition to a minor state where it occupies a larger volume, which likely represents a more extended molecule, although an oligomer cannot be completely ruled out. For reasons outlined in sections 4.3.1 and 4.3.2, the His$_6$-BECN1(1-150)Y SAXS data and analyses were not reported in the manuscript submitted for publication.

4.3.4.2. His$_6$-BECN1(1-150)Y^{CysTM}

We also used SEC-SAXS to investigate the impact of mutating the CxxC motifs and abrogating Zn^{2+}-binding on the size and shape of the BECN1 IDR (Figure 4.20 B). The integrated scattering intensity across different scattering frames for His$_6$-BECN1(1-150)Y^{CysTM} (Figure 4.20 B) trends similar to the SEC profile (Figure 4.3), with no distinct shoulders on the main peak. Strikingly unlike His$_6$-BECN1(1-150)Y (Figure 4.20 A), His$_6$-BECN1(1-150)Y^{CysTM} has a uniform R$_g$ distribution across scattering frames (Figure 4.20 B) suggesting it consists of a single homogeneous species. Therefore, EFA deconvolution was not required for this sample, and an averaged SAXS scattering profile (Figure 4.21 D) was used for SAXS analyses. The His$_6$-BECN1(1-150)Y^{CysTM} P(r) distribution shows an extended tail which is a characteristic of disordered proteins(Kikhney & Svergun, 2015), and this distribution is more asymmetric than either Component 0 or 1 of His$_6$-BECN1(1-150)Y. The R$_g$ and D$_{max}$ estimated from the P(r) distribution of His$_6$-BECN1(1-150)Y^{CysTM} are ~ 41.9 Å ± 0.4 and 176 Å respectively (Figure 4.21 G, H and Table 4.14), which is 39% larger than R$_g$ and only 10% larger than D$_{max}$ estimates of Component 1 of His$_6$-BECN1(1-150)Y (Table 4.14), and 5.5% larger than R$_g$ and only 2.2% larger than D$_{max}$ estimates of Component 0 (Table 4.14). Thus, both Component 0 and 1 of His$_6$-BECN1(1-150)Y are less extended and smaller than His$_6$-BECN1(1-150)Y^{CysTM}. Further, unlike the shallow trend in the Kratky plots for His$_6$-BECN1(1-150)Y (section 4.3.4.1), indicative of partially folded proteins, His$_6$-BECN1(1-150)Y^{CysTM} has a dimensionless Kratky plot that clearly plateaus near 1.5, indicative of increased flexibility in the absence of Zn^{2+}-binding

(Figure 4.21 I). The theoretical molecular mass of His$_6$-BECN1(1-150)Y^{CysTM} is 18.3 kDa, while that estimated from V$_c$ calculations is 23.1 kDa, consistent with the fact that V$_c$ is known to somewhat over-estimate the mass of disordered proteins (Hajizadeh *et al.*, 2018).

EOM was also used to attempt to model conformational flexibility of His$_6$-BECN1(1-150)Y^{CysTM}. The selected ensemble of conformers for His$_6$-BECN1(1-150)Y^{CysTM} that best fit the scattering data has average R$_g$ and D$_{max}$ distributions similar to the randomly generated EOM pool, suggesting a completely flexible set of conformations (Figure 4.22 B, D and Table 4.15). The best fit ensemble of His$_6$-BECN1(1-150)Y^{CysTM} has an R_{flex} of ~88.1% compared to 87.1% for the pool. This R_{flex} value is very similar to that of the pool, indicating that the selected ensemble flexibility is not significantly different from pool flexibility. Additionally, the fit of the selected conformer ensemble to experimental ensemble has a χ^2 of 1.103 (Figure 4.22 F, Table 4.15). The χ^2 is close to 1 indicating a good fit between the selected ensemble and data. The normalized residual is also flat and randomly distributed indicating a good fit.

Thus, together the SEC, SAXS EFA and EOM analyses indicate that the presence of CxxC motifs Zn^{2+}-binding makes the BECN1 IDR more conformationally compact. However, for reasons outlined in sections 4.3.1 and 4.3.2, the His$_6$-BECN1(1-150)Y^{CysTM} SAXS data and analyses were not reported in the manuscript submitted for publication.

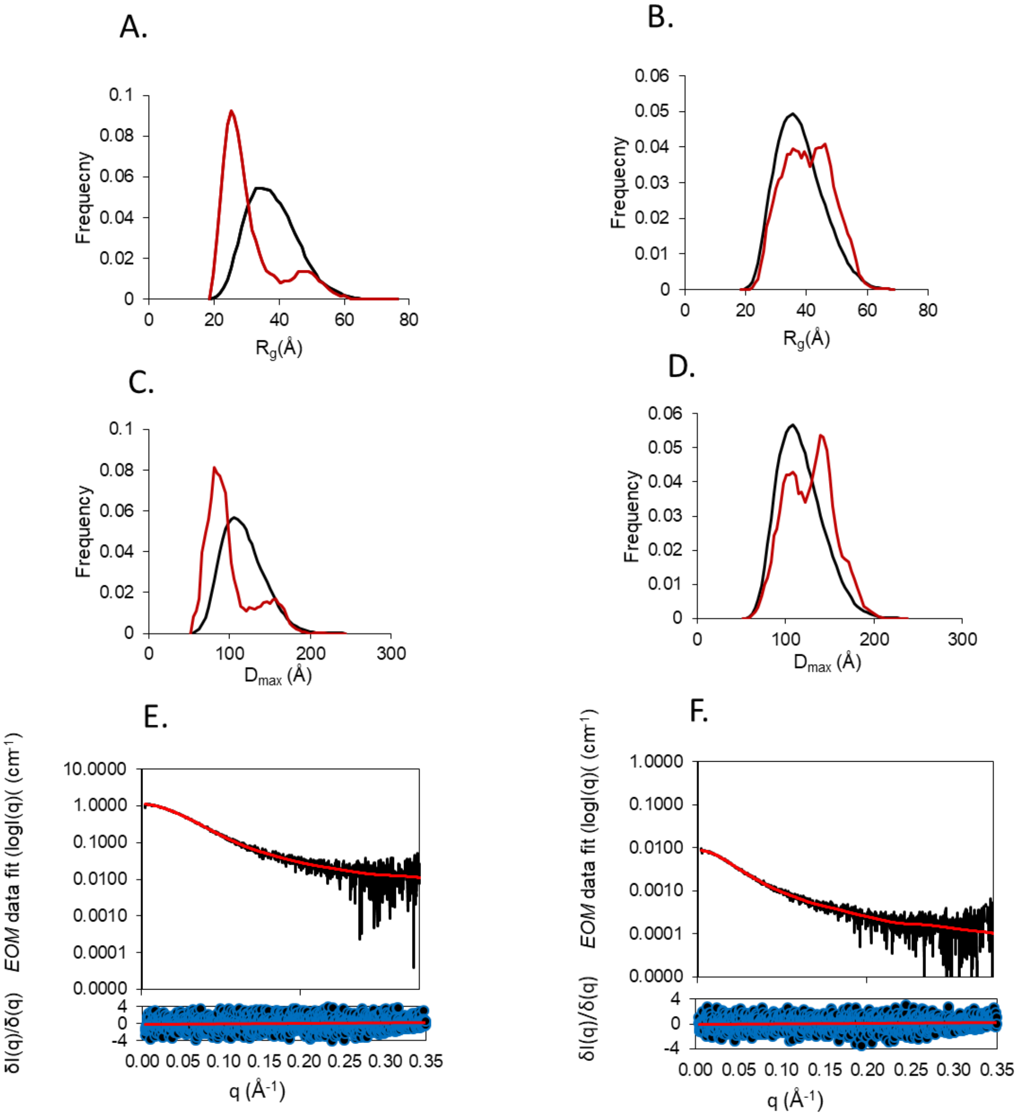


Figure 4.22: EOM Analysis for His$_6$-BECN1(1-150)Y (left panels) and His$_6$-BECN1(1-150)Y^{CysTM} (right panels). A, B) R$_g$ distribution for the EOM-generated conformer pool (black) and the selected conformer ensemble (red). C,D) D$_{max}$ distribution for the pool (black) and selected ensemble (red). E,F) Top plot: Experimental scattering profile (black) with the fit of the selected conformer ensemble (red); Bottom plot: the normalized fit residual (blue).

Table 4.15: Atomistic modelling by EOM for His$_6$-BECN1(1-150)Y constructs

	His$_6$-BECN1(1-150)Y Component 1	**His$_6$-BECN1(1-150)Y ^{CysTM}**
q-range for fitting (Å^{-1})	0.0042-0.3526	0.0042-0.3526
Symmetry assumptions	None	None
χ^2	2.486	1.104
Constant subtraction	- 0.007	-0.000
No. of representative structures	5	11
Ensemble (pool) average R_g (Å)	31.3 (37.8)	40.9 (37.9)
Ensemble (pool) average D_{max} (Å)	101.1 (116.5)	128.8 (116.6)
Ensemble (pool) average volume (Å^3)	24903.9 (32071.6)	33108.2 (32034.4)
Ensemble (pool) average Ca(N)-Ca(C) distance (Å)	70.1 (89.2)	104.89 (89.1)
R_{flex} ensemble (pool)	~ 79.94% (~84.14%)	~ 88.12% (~87.13%)
R_σ	0.82	1.07
Domain/subunit coordinates and contacts, regions of presumed flexibility as appropriate	Only the sequence was provided, so the entire protein was assumed to be flexible.	

4.3.4.3. BECN1(1-150)Y

The integrated scattering intensity for BECN1(1-150)Y plotted across scattering frames (Figure. 4.23A) shows a profile similar to the SEC elution profile (Figure. 4.10), including the presence of a leading-edge shoulder preceding the peak, where the peak and shoulder correspond to 83% and 17% of the total protein. Similar to the His$_6$-BECN1(1-150)Y construct, the R_g distribution decreases across the scattering frames (Figure.4.23A), suggesting that the shoulder contains a species with higher R_g. As such the shoulder likely corresponds to either an oligomer or an extended conformation, followed by either a monomeric or more compact species within the peak. Due to the possibility of having different species in the SEC elution, the SAXS scattering profile of BECN1(1-150)Y sample was deconvoluted using EFA.

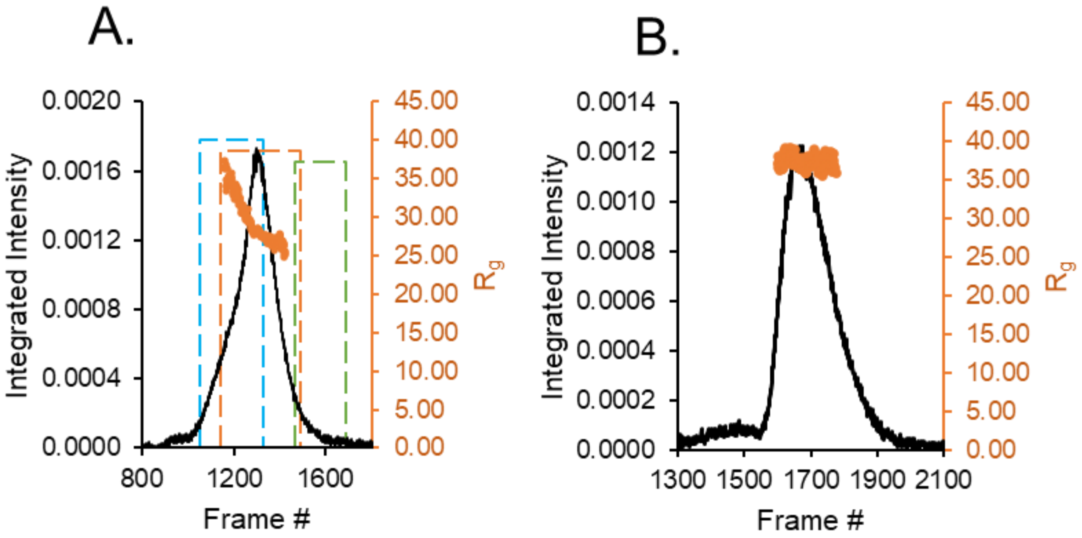


Figure 4.23: EFA Analysis of BECN1(1-150)Y and total SAXS scattering intensity of BECN1(1-150)Y^{CysTM}. A,B) Plots showing total scattered intensity (left y-axis) and Guinier R_g (right y-axis) across subsequently measured scattering profiles (frames) of BECN1(1-150)Y and BECN1(1-150)Y^{CysTM} respectively. BioXTAS RAW (version 1.6.4) was used to select EFA data ranges for BECN1(1-150)Y Components 0,1,and 2 (blue, orange and green respectively) which are indicated by dashed boxes.

Three potential component species were identified in the scattering profile of BECN1(1-150)Y, namely Components 0, 1 and 2. Component 0 was identified from the leading edge of the BECN1(1-150)Y peak, comprising scattering frames 1052 to 1304 (Figure 4.23 A), Component 1 was identified from the peak comprising of scattering frames 1125 to 1519 (Figure 4.23 A), Component 2 was identified from the trailing edge of the peak comprising of scattering frames 1481 to 1686 (Figure 4.23 A). Component 0 and Component 1 were analyzed further, however the third EFA component, Component 2 displayed a low signal to noise as shown by negative intensities at low q (Figure 4.24 C), indicative of a very low concentration, or small molecular mass or both, which prevented further analysis.

The P(r) distribution curve for Component 0 has an extended tail characteristic of disordered proteins (Kikhney & Svergun, 2015), and provides R_g and D_{max} estimates of ~46.3 Å

± 0.7 and 200 Å respectively (Figure. 4.24 E, H and Table 4.16). Similarly, the P(r) distribution curve of Component 1 is more symmetrical, but also shows an extended tail characteristic of disorder, with R_g and D_{max} estimates of ~28.7 Å ± 0.1 and 120 Å respectively (Figure. 4.24 F, H and Table 4.16). The higher R_g and D_{max} for Component 0 compared to Component 1 indicates that Component 0 is larger in size compared to Component 1, and a comparison of the shape of P(r) distribution curves indicates a more extended structure for Component 0. The dimensionless Kratky plot for Component 0 (Figure. 4.24 I) plateaus near 1.5 with an increase in qR_g, that indicates that Component 0 of BECN1(1-150) is completely unfolded, whereas the Kratky plot for Component 1 trends somewhat shallower relative to that of Component 0 (Figure. 4.24 I), indicating that these molecules are only partially disordered.

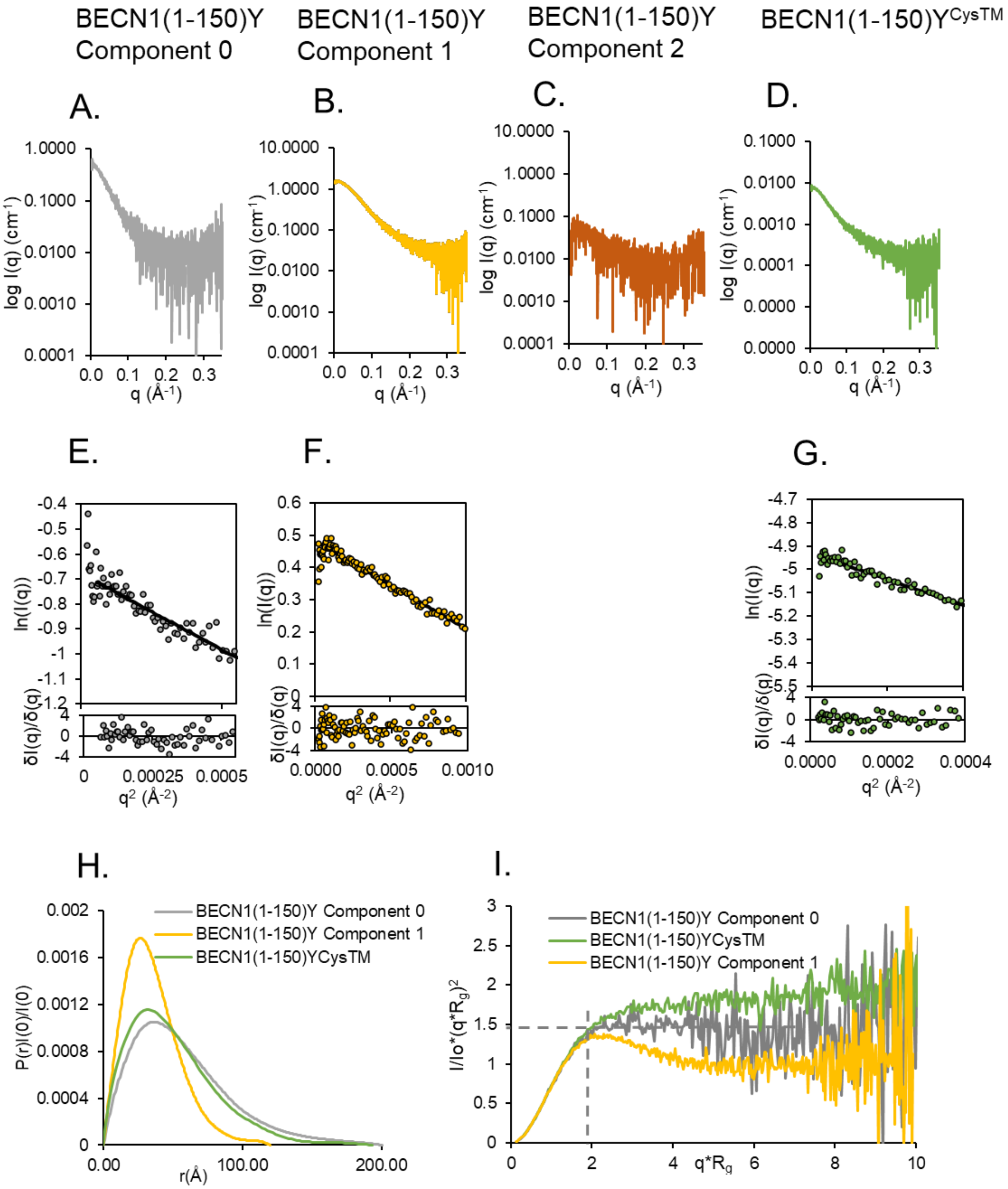


Figure 4.24: SAXS Analyses for BECN1(1-150)Y Component 0 (gray data and curves), BECN1(1-150)Y Component 1 (yellow data and curves), Component 2 (brown curve) and BECN1(1-150)Y^{CysTM} (green data and curves). A-D) Scattering profiles, E-G) Guinier Plots. Top panel: the black lines represent fits to the data represented by gray, yellow and green points respectively; Bottom panel: corresponding fit residuals. H) P(r) function. I) Kratky Plot. Dashed lines indicate the peak expected for a compact globular protein

Table 4.16: SEC-SAXS Data Analysis for BECN1(1-150)Y constructs

	BECN1(1-150)Y Component 0	**BECN1(1-150)Y Component 1**	**BECN1(1-150)Y$^{\mathbf{CysTM}}$**
Guinier Analysis			
I(0) (cm^{-1})	0.522 ± 0.003	1.618 ± 0.002	0.00720 ± 0.003 x10^{-2}
R_g (Å)	42.5 ± 0.4	28.3 ± 0.1	40.5 ± 0.6
q-range (Å^{-1})	0.00425-0.02916	0.00651-0.03341	0.00708-0.02237
qR_g max	1.24	0.945	0.907
r^2 fit	0.903	0.982	0.945
M from *Vc (ratio to theoretical)* (kDa)	34.3 (2.0)	20.5 (1.19)	22.7 (1.3)
P(*r*) analysis			
I(0) (cm^{-1})	0.522 ± 0.003	1.618 ± 0.002	0.00723 ± 0.028 x10^{-3}
R_g (Å)	46.3 ± 0.7	28.7 ± 0.1	41.8 ± 0.5
d_{max} (Å)	200	120	193
q-range (Å^{-1})	0.0042-0.3526	0.0042-0.3526	0.0071-0.3526
Total Estimate	0.71	0.73	0.69
χ^2	2.43	2.64	1.04

Molecular mass estimations from V_c (Rambo & Tainer, 2013) indicate that BECN1(1-150)Y Component 0 has a molecular mass of 34.3 kDa which would correspond to an oligomer, and Component 1 has a molecular mass of 20.5 kDa which corresponds to a monomer. As mentioned before (section 4.3.5.1), molecular mass estimations from V_c are notoriously inaccurate for disordered proteins and dependent on having an overall high signal to noise ratio in the scattering intensity profile (Hajizadeh *et al.*, 2018). Unfortunately, Component 0 has a low signal to noise ratio in the scattering intensity profile. The low signal is likely because of a low concentration at the leading-edge shoulder as it constitutes only 17% of the total eluted protein from SEC, and noisy due to the influence of small amount of oligomeric species at the void volume, that may have impacted the scattering. Therefore, the V_c calculation, and any inference of the oligomeric state of Component 0 is not reliable. This suggests that Component 0 either corresponds to an oligomer, or more likely, exists in a more extended state relative to Component 1.

EOM was used to generate a random pool of conformers for the Component 0 of BECN1(1-150)Y, and select an ensemble of conformers that best fit the scattering data. The selected ensemble of BECN1(1-150)Y Component 0 conformers has an average R_g of 44.9Å and D_{max} of 140 Å compared to the EOM generated random pool of conformers with R_g of 36.6Å and D_{max} of 112 Å. This indicates that the selected conformer ensembles are much larger and extended compared to the average of the random pool of conformers. The selected conformer ensemble fits the experimental data has a χ^2 of 2.92. The normalized residual is flat but less randomly distributed at the high q region, indicative of a poor fit. Indeed, the conformational flexibility of Component 0 cannot be reliably modeled in the absence of information about the oligomeric state of Component 0, and therefore this data is not useful for any further analysis.

EOM was also used to generate a random pool of conformers representing Component 1 of BECN1(1-150)Y, and select an ensemble of conformers that best fit the scattering data. The selected ensemble of BECN1(1-150)Y Component 1 conformers has a narrower R_g and D_{max} distribution compared to that of the randomly-generated EOM pool, indicating that Component 1 conformations (Figure 4.25 A, C and Table 4.17) are more compact relative to all possible conformations. R_{flex} for the Component 1 BECN1(1-150)Y is ~ 76.8% for the selected ensemble, significantly smaller than the 84.9 % for the pool, which suggests that the conformer ensemble is significantly less flexible than would be expected for a completely disordered protein. When the R_{flex} of the ensemble is less than R_{flex} of the pool, R_σ is expected to be less than 1. The R_σ for Component 1 is 0.94, as expected. Additionally, the fit of the EOM selected conformer ensemble to the experimental ensemble for BECN1(1-150)Y Component 1 has a χ^2 of 3.36 (Figure 4.25 E, Table 4.17). While the χ^2 is somewhat large, the normalized residual is flat and randomly distributed, indicating a good fit. It may be that the uncertainty from the EFA deconvolution is underestimated, resulting in a larger than expected χ^2.

Thus, together the SEC- SAXS EFA and EOM analyses indicates that the BECN1(1-150)Y protein is a partially-disordered monomer with an extended conformation, and appears to transition to a minor state where it occupies a larger volume, which likely represents a more extended molecule, although an oligomer cannot be completely ruled out.

4.3.4.4. BECN1(1-150)Y^{CysTM}

Similar to His_6-BECN1(1-150)Y proteins, we also investigated the impact of mutating the CxxC motifs and abrogating Zn^{2+} binding, on the size and shape of the BECN1 IDR (Figure 4.23). The integrated scattering intensity across different scattering frames for BECN1(1-150)Y^{CysTM} (Figure. 4.23 B) trends similar to the SEC profile (Figure 4.11), with no distinct shoulders on the main peak. Strikingly unlike BECN1(1-150)Y (Figure 4.23A), BECN1(1-150)Y^{CysTM} has a uniform R_g distribution across scattering frames (Figure. 4.23 B), indicating it consists of a single homogeneous species. Therefore, EFA deconvolution was not required for this sample, and an averaged SAXS scattering profile (Figure 4.24 C) was used for the SAXS analysis. The BECN1(1-150)Y^{CysTM} P(r) distribution shows an extended tail which is a characteristic of disordered proteins(Kikhney & Svergun, 2015), and this distribution is more asymmetric than either Component 0 or 1 of BECN1(1-150)Y. The R_g and D_{max} estimated from the P(r) distribution of BECN1(1-150)Y^{CysTM} are ~ 41.8 Å ± 0.5 and 193 Å respectively (Figure. 4.24 G, H and Table 4.16), which is 46% and 61% larger than R_g and D_{max} estimates of Component 1 of BECN1(1-150)Y (Table 4.16) and 9.7% smaller and only 3.5% smaller in R_g and D_{max} estimates of Component 0 (Table 4.16). Thus, BECN1(1-150)Y^{CysTM} is slightly smaller than Component 0 and significantly larger than Component 1 of BECN1(1-150)Y. BECN1(1-150)Y^{CysTM} has a dimensionless Kratky plot that clearly plateaus near 1.5 (Figure. 4.24 I), indicative of complete disorder. Notably, BECN1(1-150)Y^{CysTM} and Component 0 of BECN1(1-

150)Y have very similar Kratky plots, indicating that both are similarly unfolded. In contrast, the Kratky plot for Component 1 of BECN1(1-150)Y (Figure 4.24 I), indicates it is partially folded. Thus, the BECN1 IDR is more disordered and extended in the absence of Zn^{2+}-binding. The theoretical molecular mass of BECN1(1-150)Y^{CysTM} is 17.1 kDa, while that estimated from V_c calculations is 22.7 kDa which is consistent with a monomer, given that V_c is known to somewhat over-estimate the mass of disordered proteins (Hajizadeh *et al.*, 2018).

Next, EOM was used to model conformational flexibility of BECN1(1-150)Y^{CysTM} by generating a random pool of conformers and selecting an ensemble of conformers that best fit the scattering data. The selected ensemble has R_g and D_{max} distributions similar to the randomly generated EOM pool, which suggests that flexibility of the conformers in the selected ensemble (Figure. 4.25 B, D and Table 4.17) is similar to that of the randomly generated pool. This R_{flex} value is very similar to that of the pool, indicating that the selected ensemble flexibility is not significantly different from pool flexibility. When R_{flex} of the selected conformer ensemble is similar to the pool, we expect the R_σ to be close to 1. Consistent with this, the R_σ is 1.22. Additionally, the fit of the selected conformer ensemble to experimental data has a χ^2 of 0.993 (Figure. 4.25 F, Table 4.17), indicative of a good fit to experimental scattering. The normalized residual is also flat and randomly distributed indicating a good fit.

Therefore, together the SEC, SAXS EFA and EOM results suggest that the mutation of CxxC motifs and abrogation of Zn^{2+} binding makes the BECN1 IDR more conformationally flexible and extended.

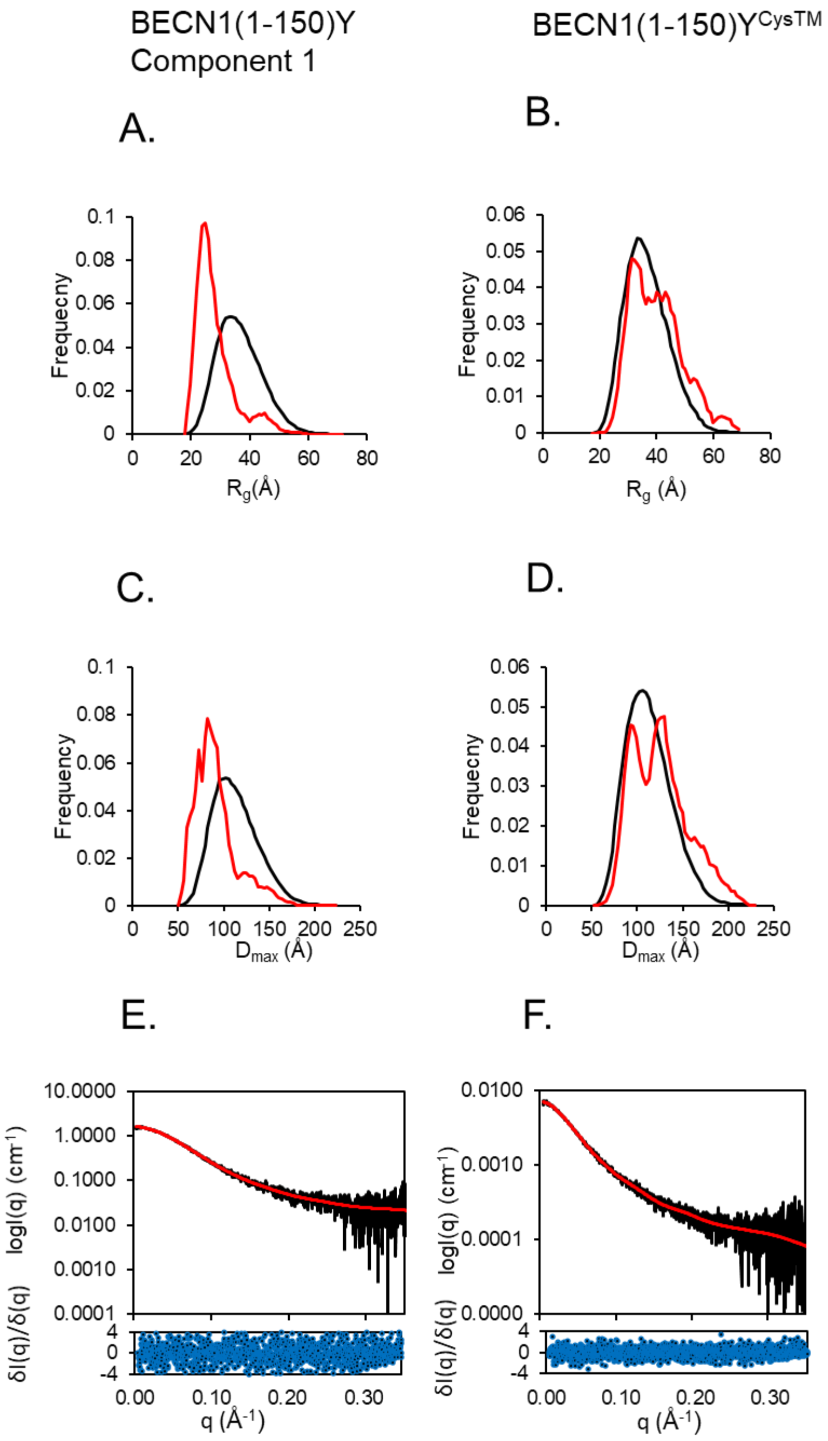


Figure 4.25: EOM Analysis for BECN1(1-150)Y (left panels) and BECN1(1-150)Y^{CysTM} (right panels). A, B) R_g distribution for the EOM-generated conformer pool (black) and the selected conformer ensemble (red). C,D) D_{max} distribution for the pool (black) and selected ensemble (red). E,F) Top plot: Experimental scattering profile (black) with the fit of the selected conformer ensemble (red); Bottom plot: the normalized fit residual (blue).

Table 4.17: Atomistic modelling by EOM for BECN1(1-150)Y constructs

	BECN1(1-150)Y Component 1	**BECN1(1-150)Y$^{\text{CysTM}}$**
q-range for fitting (Å^{-1})	0.0042-0.3526	0.0042-0.3526
Symmetry assumptions	None	None
χ^2	3.36	0.993
Constant subtraction	- 0.008	-0.000
No. of representative structures	6	11
Ensemble (pool) average R_g (Å)	28.6 (36.6)	40.7 (37.1)
Ensemble (pool) average D_{max} (Å)	92.6 (112.8)	126.9 (114.5)
Ensemble (pool) average volume (Å^3)	22020.9 (29434.7)	31941.6 (31542.2)
Ensemble (pool) average Ca(N)-Ca(C) distance (Å)	66.3 (86.5)	104.6 (88.8)
R_{flex} ensemble (pool)	~ 76.76%(~84.88%)	~ 89.10% (~85.61%)
R_σ	0.94	1.22
Domain/subunit coordinates and contacts, regions of presumed flexibility as appropriate	Only the sequence was provided, so the entire protein was assumed to be flexible.	

4.3.5. MALS analysis of the BECN1 IDR constructs

4.3.5.1. His$_6$-BECN1(1-150)Y

The EFA analysis of the His$_6$-BECN1(1-150)Y SAXS data indicated this sample has two significant species that correspond to His$_6$-BECN1(1-150)Y monomer and either oligomers or molecules with very extended conformations. Therefore, we attempted to estimate the molecular mass of the proteins using MALS. We expected that the high sensitivity of MALS would enable more accurate determination of molecular mass than SAXS analysis. Since the R_g distribution decreases across the SAXS scattering profile (Figure 4.20A), we decided to analyze the His$_6$-BECN1(1-150)Y MALS scattering profile in sections and calculate the approximate average molecular mass for each section (Figure 4.26 A). The light scattering data within the major peak has a high signal to noise ratio, indicative of a high concentration of homogenous protein. The estimated average molecular mass is 22 kDa (Figure 4.26A), which corresponds well to that from SAXS analyses and the theoretical monomeric molecular mass estimated from sequence analysis (Table 4.14) and is consistent with His$_6$-BECN1(1-150)Y being a monomer. The trailing edge of the peak shows presence of relatively small molecular mass sections that may represent low concentration of degraded proteins coeluting with the monomeric His$_6$-BECN1(1-150)Y.

Unfortunately, however, the MALS data corresponding to the leading-edge shoulder of the major peak has a poor signal to noise ratio preventing robust data deconvolution. The MALS analysis suggests that the leading edge of the peak has different oligomeric species in small concentrations that are not well-resolved due to a low signal to noise ratio, which makes molecular mass estimation and corresponding inferences of oligomerization rather unreliable. Thus, MALS analysis of His_6-BECN1(1-150) indicates that the BECN1 IDR is predominantly a monomer.

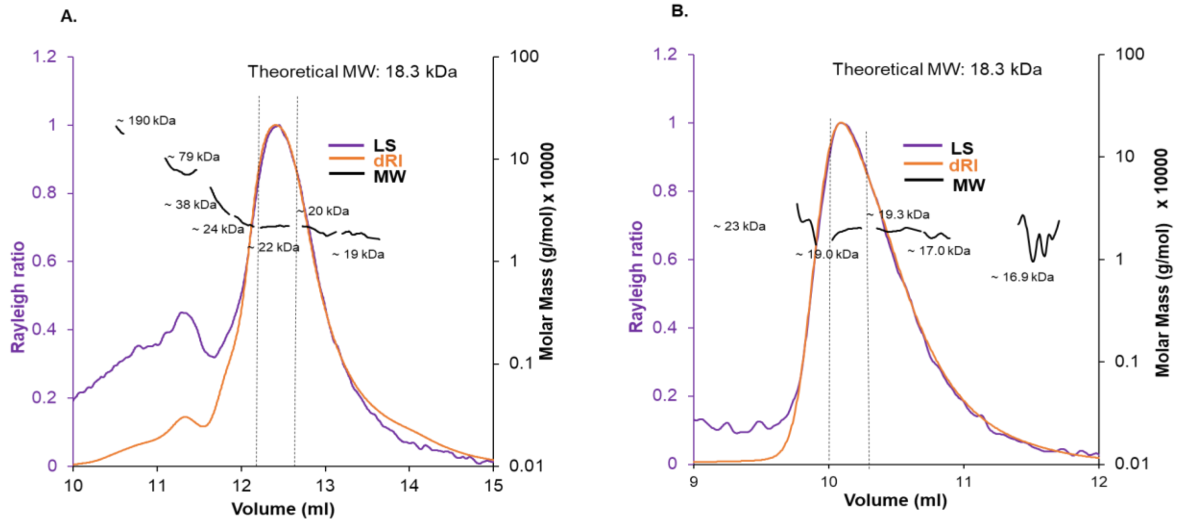


Figure 4.26: MALS estimation of molecular mass of His_6-BECN1(1-150)Y constructs. MALS data showing light scattering (LS) in purple (left y-axis), differential refractive index (dRI) in orange and molar mass in black (right y-axis) calculated for the SEC elution volumes (x-axis) for A) His_6-BECN1(1-150)Y and B) His_6-BECN1(1-150)Y^{CysTM}. The dashed lines indicate the peak that correlate with the SAXS data.

4.3.5.2. His_6-BECN1(1-150)Y^{CysTM}

MALS analysis of His_6-BECN1(1-150)Y^{CysTM} (Figure 4.26 B) indicates the sample is largely homogenous and the corresponding molecular mass estimation of this construct indicates the presence of a single species with a molecular mass of 18.4 kDa, which corresponds well to the molecular mass estimation corresponds from the SAXS analyses and the theoretical molecular mass estimated from sequence analysis (Table 4.14), and is consistent with this protein being a monomer. Similar to His$_6$-BECN1(1-150)Y, at the training edge there are relatively small molecular mass sections that may represent low concentration of degraded proteins coeluting with the monomeric His$_6$-BECN1(1-150)Y^{CysTM}. For reasons outlined in sections 4.3.1 and 4.3.2, the His$_6$-BECN1(1-150)Y and His$_6$-BECN1(1-150)Y ^{CysTM} MALS data and analyses are not reported in the manuscript.

4.3.5.3. BECN1(1-150)Y

Similar to the His$_6$-BECN1(1-150)Y construct, we also performed MALS analysis for BECN1(1-150)Y to further verify the oligomeric state and molecular mass of BECN1(1-150)Y. The EFA analysis of the BECN1(1-150)Y SAXS data indicated this sample has at least two significant species. The leading edge of the major peak may correspond to BECN1(1-150)Y molecules with extended conformations or potential oligomers. Since the R_g distribution decreases across the SAXS scattering profile (Figure 4.23 A), we decided to analyze the BECN1(1-150)Y MALS scattering profile in sections and calculate the approximate average molecular mass for each section (Figure 4.27 A). The major peak in the MALS data has a high signal to noise ratio, indicative of a high concentration of homogenous protein. The average molecular mass estimated is 19.9 kDa (Figure 4.27 A), which corresponds well with SAXS analyses and the theoretical monomeric molecular mass estimated from sequence analysis (Table 4.16) and is consistent with BECN1(1-150)Y being a monomer. The trailing edge of the peak shows presence of relatively small molecular mass sections that may represent low concentration of degraded proteins coeluting with the monomeric BECN1(1-150)Y. Unfortunately, however, the leading edge of the major MALS peak has a poor signal to noise ratio preventing robust data deconvolution. The MALS analysis suggests that the leading edge of the peak has different oligomeric species in small concentrations that are not well resolved due to the low signal to noise ratio. This makes molecular mass estimation and corresponding inferences of oligomerization rather unreliable. Thus, MALS analysis of BECN1(1-150) indicates that the BECN1 IDR is predominantly a monomer.

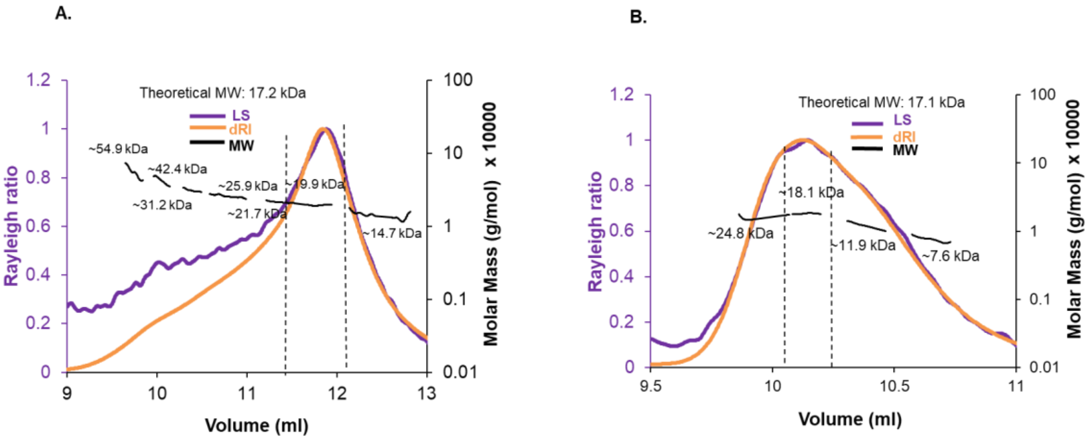


Figure 4.27: MALS estimation of molecular mass of BECN1(1-150)Y constructs. MALS data showing light scattering (LS) in purple (left y-axis), differential refractive index (dRI) in orange and molar mass in black (right y-axis) calculated for the SEC elution volumes (x-axis) for A) BECN1(1-150)Y and B) BECN1(1-150)Y^{CysTM}. The dashed lines indicate the peak that correlate with the SAXS data.

4.3.5.4. BECN1(1-150)Y^{CysTM}

MALS analysis of BECN1(1-150)Y^{CysTM} (Figure 4.27 B) indicates the sample is largely homogenous and the corresponding molecular mass estimation indicates the presence of a single species with a monomeric molecular mass (18.1 kDa). The trailing edge of the peak shows presence of relatively small molecular mass sections that may represent low concentration of degraded proteins coeluting with the monomeric BECN1(1-150)Y^{CysTM}. This molecular mass estimation corresponds well with the SAXS analyses and the theoretical molecular mass estimated from sequence analysis (Table 4.16), and is consistent with BECN1(1-150)Y^{CysTM} being a monomer.

4.3.6. ITC experiments suggest that the BECN1 IDR does not homodimerize

The SAXS EFA analyses and analyses of the MALS data corresponding to Component 0 of BECN1(1-150)Y suggest that the BECN1 IDR may self-associate. Therefore, we used ITC to attempt to quantify the thermodynamics of BECN1(1-150)Y self-dissociation. The protein concentrations used for ITC were similar to that loaded on the SEC column for SEC-SAXS-MALS (approx. 294 µM or 5 mg/ml) analyses. As homodimerization appeared to occur in the SEC eluate, which has protein concentrations relatively lower than the load, we expected that the equilibrium would be shifted to a oligomeric state in the ITC experiments, when performed using 116 µM (2 mg/ml) and 355 µM (6 mg/ml) BECN1(1-150)Y (Figure. 4.28 B, C). This would facilitate quantification of thermodynamic parameters of IDR self-dissociation. However, we did not observe any heat of self-dissociation when the protein was titrated into buffer in the cell (Figure. 4.28 B, C), suggesting that the IDR does not self-associate. Prior to these experiments, we had performed these experiments using MBP-BECN1(1-150)Y at 258 µM (15.5 mg/ml) concentration, as the protein was better behaved and it was easier to obtain sufficient protein for ITC experiments, but, as for the untagged protein, no heat of self-dissociation was observed when the protein was titrated into buffer in the cell (Figure. 4.28 A).

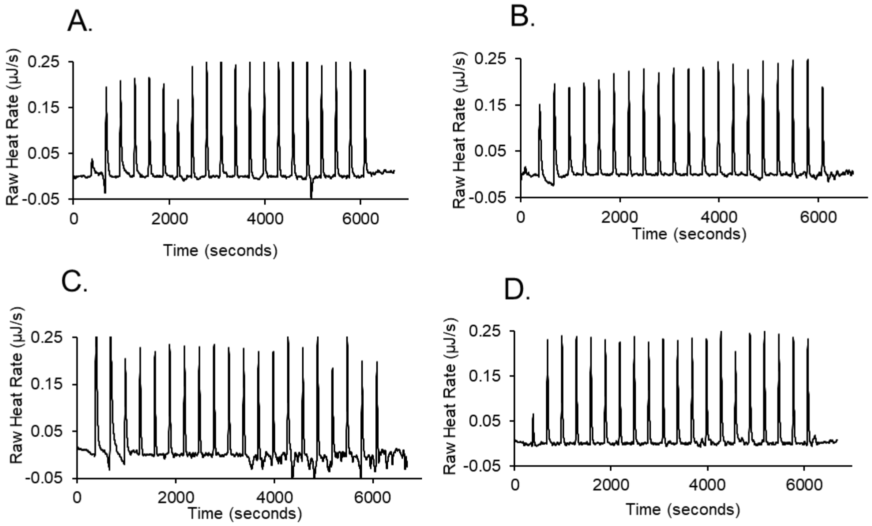


Figure 4.28: Raw ITC self-titration data for BECN1 IDR constructs. A) 258 µM MBP-BECN1(1-150)Y, B) 116 µM BECN1(1-150)Y, C) 355 µM BECN1(1-150)Y and D) 243 µM MBP-BECN1(1-150)Y^{CysTM}.

Thus, while a potential oligomeric species was suggested by the SEC-SAXS and MALS analyses of BECN1(1-150)Y; the ambiguity of that analysis due to the low quality of the data in those regions of curve together with the lack of detectable BECN1(1-150)Y self-dissociation heats indicates that the protein is more likely an extended monomeric species; although we cannot rule out BECN1(1-150)Y oligomerization being mediated by covalent disulfide bonds that prevent oligomer dissociation in the ITC experiments, or perhaps that self-dissociation involves very small heats that are not detectable above noise. Similar self-dissociation experiment at concentration of MBP-BECN1(1-150)Y^{CysTM}, similar to that used during SEC-SAXS and MALS analyses also showed no heat of dissociation (Figure. 4.28 D) and is consistent with the SEC-SAXS and MALS analyses, which indicate the mutant protein exists only as a monomer.

4.3.7. BECN1 CxxC motifs have no impact on M11 interaction

As M11 binds to BECN1 via its BH3D (Sinha, Colbert, Becker, Wei & Levin, 2008), we investigated if Zn^{2+}-binding by the BECN1 CxxC motifs impacts M11 binding. As SDS-PAGE indicates that some MBP-BECN1(1-150)Y degrades during the course of the ITC experiment, a local incompetent fraction parameter was included in our binding model for isotherm analysis. The local incompetent fraction estimate ranged between 42% - 50% among the triplicate measurements. These analyses indicate that MBP-BECN1(1-150)Y interacts with M11 in a 1:1 stoichiometry with binding affinity (K_d) of 0.33 µM, with a 68.3% confidence interval of [0.22, 0.48 µM] (values in square brackets indicate a 68.3% confidence interval ±1 standard deviation for the mean value presented) (Figure 4.29A, Table 4.18). Similarly, MBP-BECN1(1-150)Y^{CysTM} was also tested for interaction with M11. Consistent with reduced degradation of MBP-BECN1(1-150)Y^{CysTM}, the local incompetent fraction was estimated between 13% - 14% among the triplicate measurements. These analyses indicate that similar to MBP-BECN1(1-150)Y, MBP-BECN1(1-150)Y^{CysTM} also interacts with M11 in a 1:1 stoichiometry, with K_d of 0.40 µM, with a 68.3% confidence interval of [0.23, 0.68 µM] (Figure. 4.29B, Table 4.18). These results indicate that Zn^{2+} binding by BECN1 CxxC motifs does not significantly impact the M11 interaction (Table 4.18).

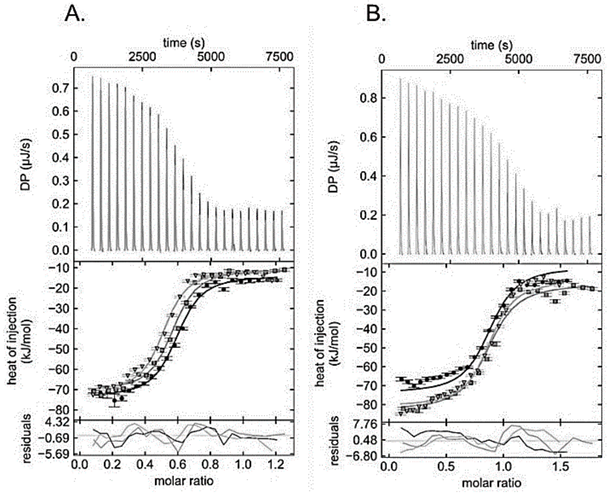


Figure 4.29: Global analyses of ITC isotherms for M11 titrated into BECN1 IDR constructs. A) MBP-BECN1(1-150)Y , B) MBP-BECN1(1-150)Y^{CysTM}. The heats of binding (top panel), and the isotherms with the curve fits using a global model (middle panel) and residuals of the global model fit (bottom panel) for triplicate experiments are shown in black, gray, and dark gray.

Table 4.18: Thermodynamic parameters of the BECN1 IDR: M11 interaction as determined from ITC data using a global analysis in SEDPHAT

Titration	**[BECN1] (μM)**	**[M11] (μM)**	**LIF^a**	**K_d (μM)**	**ΔH^b (kJ/mol)**	**ΔS^b (J/mol.K)**	**ΔG^b (kJ/mol)**
MBP-BECN1 (1-150)Y to M11	34.6	124	0.42	0.33 [0.22, 0.48]	-59.74 [-63.22, -56.67]	-79.65 [-88.21, -72.33]	-36.39 [-37.36, -35.47]
	34.9	122	0.50				
	35.2	131	0.44				
MBP-BECN1 (1-150)Y^{CysTM} to M11	26.6	123	0.13	0.40 [0.23, 0.68]	-66.10 [-72.8, -60.91]	-103.1 [-121.4, -89.70]	-35.88 [-37.21, -34.61]
	23.3	123	0.14				
	30.2	134	0.14				

[a] Local incompetent fraction (LIF)

[b] Values in square brackets indicate a 68.3% confidence interval (1 S.D.) for the mean value presented.

Mean values were determined from a global fit to a set of three ITC experiments.

4.3.8. Using Hydrogen/Deuterium Exchange Mass Spectrometry to assess solvent accessibility of BECN1 IDR residues

Both, His$_6$-BECN1(1-150)Y and His$_6$-BECN1(1-150)Y^{CysTM} showed 99.4% peptide coverage with a total of 70 unique peptides, largely with high confidence match (in green) and one with medium confidence match (in yellow) (Figure 4.30). The average size of each peptide was 19 residues, with a redundancy of 8.2. Redundancy refers to the average number of times each position of the protein sequence is covered by the unique peptides that enable deuteration analysis at that position. A common set of peptides were analyzed for both, His$_6$-BECN1(1-150)Y and His$_6$-BECN1(1-150)Y^{CysTM}, to eliminate inconsistencies due to a different choice of peptides for the two constructs. This allowed direct comparison of deuteration levels as an indicator of conformational flexibility. Deuteration uptake was consistently slightly increased for His$_6$-BECN1(1-150)Y^{CysTM} relative to His$_6$-BECN1(1-150)Y. The heat maps represent the computed deuteration level at each residue, which is the best least-squares match to the individual peptide measurements. Both His$_6$-BECN1(1-150)Y and His$_6$-BECN1(1-150)Y^{CysTM} constructs have very similar deuteration patterns with overall deuteration trends ranging from less than 50% deuteration to greater than 90% deuteration (Figure 4.30). Consistent with the deuteration uptake plots, the heat maps also show that the His$_6$-BECN1(1-150)Y^{CysTM} is slightly more deuterated, and this is sustained across the 4 different time points. A comparison of the representative deuteration uptake plots of the peptides around each of the two CxxC motifs (residues 30 to 44 and residues 139 to 159) shows that mutation of the CxxC motifs results in more deuteration compared to wild type (Figure 4.31 A,B).

Thus, the deuterium uptake plots and heat maps show that mutation of the BECN1 CxxC motifs correlates with slight increases in deuterium uptake by the IDR, indicating increased flexibility. This may indicate that Zn^{2+} binding by the CxxC motifs makes the BECN1 IDR less conformationally flexible and more compact. This is consistent with the CD and SAXS results described in previous sections.

A major problem for this experiment was the consistent decrease in the deuterium uptake between the 120 secs and the 300 secs time points. This decrease in deuterium uptake could be because of increased aggregation of proteins during later time points, or deuterium uptake reached its highest capacity by the first time point itself. In this case, shorter time points such as 0.1 s, 1 s, 10 s and 100 s or millisecond timepoints may have shown a much clearer trend in the deuterium uptake (Masson *et al.*, 2019). This shorter time scale may have also prevented possible aggregation during higher time points. Thus, more experimental optimization may improve the results, but this experiment is extremely cost intensive, which limited more possible experimental optimization and also estimation of conformational flexibility in BECN1(1-150)Y constructs.

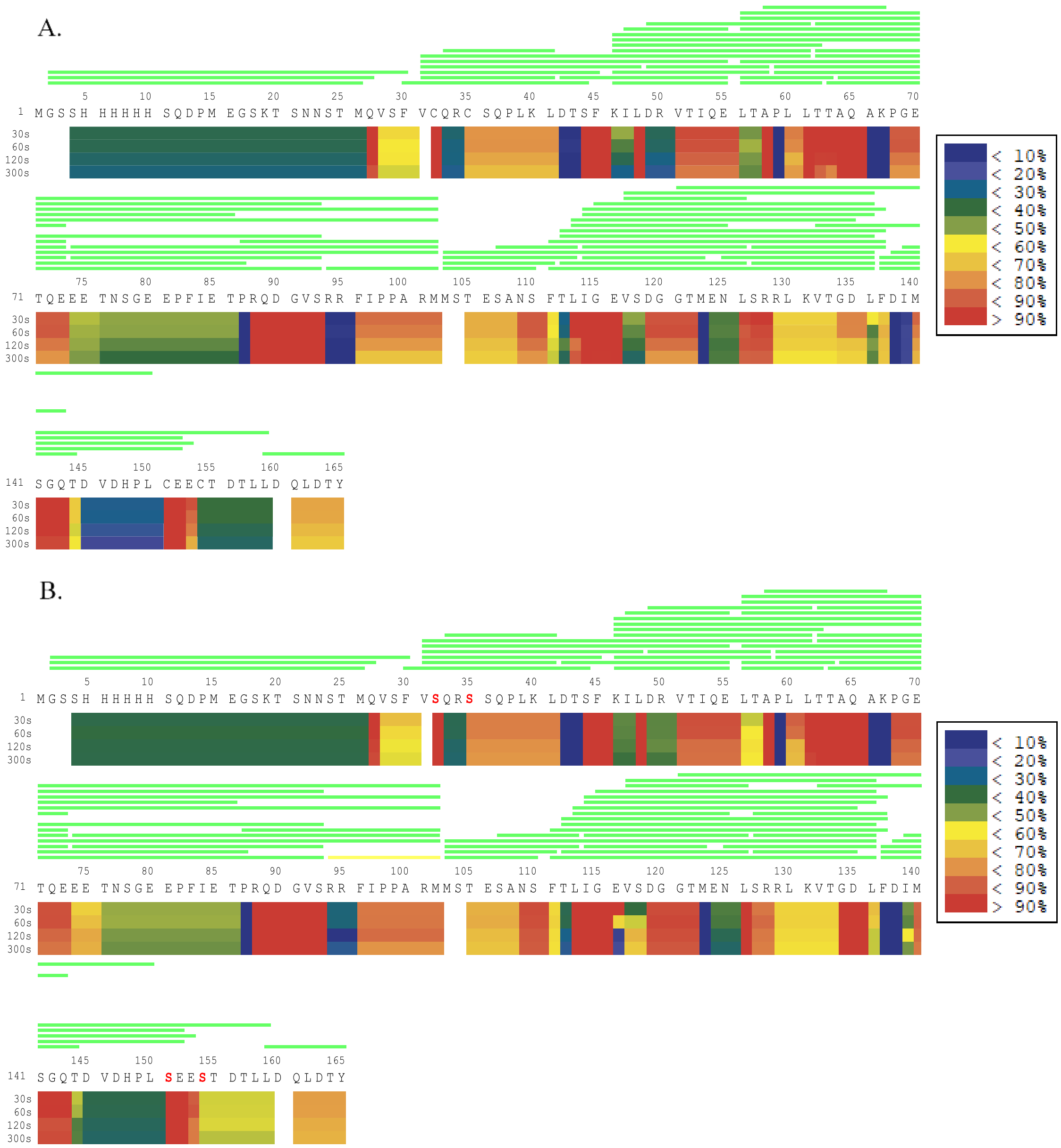


Figure 4.30: Heat Maps for: A) His$_6$-BECN1(1-150)Y, B) His$_6$-BECN1(1-150)Y$^{\mathrm{CysTM}}$. The unique high confidence peptides are shown in green above the sequences. The relative deuterium uptakes at different time points (30s, 60s, 120s, 300s) are indicated by rainbow colors, with purple to red representing the least (<10%) to most (>90%) deuteration.

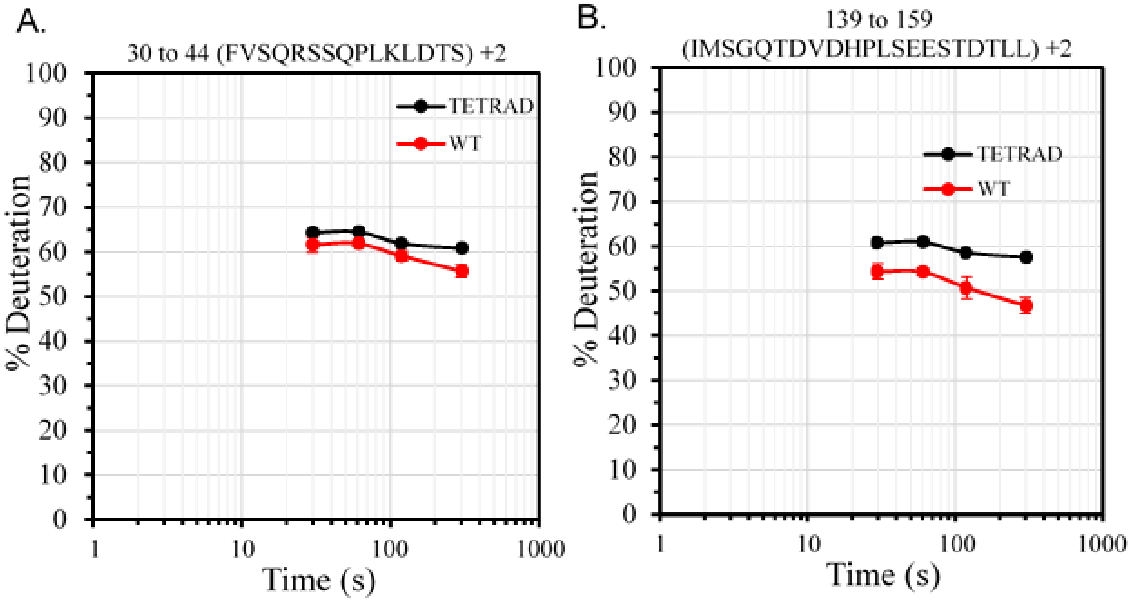


Figure 4.31: % Deuterium Uptake Plots showing for peptides including the A) first CxxC motif and B) second CxxC motif. % Deuteration of the amino acids in a peptide (on y-axis) is plotted against time in secs, shown on a logarithmic scale ($\log_{10}$) (on x-axis). The 4 data points for each curve represent different time points (30s, 60s, 120s, 300s), for either His$_6$-BECN1(1-150)Y$^{\mathrm{CysTM}}$ in black or His$_6$-BECN1(1-150)Y in red.

4.3.9. The BECN1 IDR and invariant CxxC motifs are critical for starvation-induced autophagy

Lastly, we investigated the impact of the BECN1 IDR and Zn^{2+}-binding by the invariant CxxC motifs on cellular autophagy levels. Since, BECN1 is known to be required for induction of autophagosome nucleation, we evaluated starvation induced cellular autophagy by monitoring and comparing the levels of puncta labeled with GFP-tagged LC3, an autophagosome specific marker for both. Cells were grown on nutrient rich medium and starvation medium, upon exogenous expression of either WT FL or a BECN1 construct lacking the IDR, (FL$^{\Delta(31\text{-}123)}$) or cysteine double mutants (FL$^{\mathrm{DM1}}$ and FL$^{\mathrm{DM2}}$) or cysteine tetrad mutant (FL$^{\mathrm{TM}}$) of BECN1. To determine the effect of BECN1 in mammalian cells, we used human breast adenocarcinoma MCF7 cells for these experiments because they have undetectable levels of endogenous BECN1 expression, resulting in very low levels of basal autophagy under both nutrient rich and starvation conditions, unless BECN1 is ectopically expressed (Figure 4.32 A). This allows us to assess the impact of WT or mutated BECN1 on cellular autophagy levels.

Transient expression of WT FL BECN1 in MCF7 cells did not increase autophagy levels in nutrient-rich conditions, p = 0.77 for WT expression as compared to no expression (Figure. 4.32 B) but facilitated a significant increase in the autophagy in starvation conditions, p = 0.004 x 10^{-3} for FL expression in starvation condition to nutrient-rich condition. Remarkably, autophagy levels facilitated by a BECN1 (Δ31-123) ($FL^{\Delta(31\text{-}123)}$) construct were similar to WT (p = 0.79 for $FL^{\Delta(31\text{-}123)}$ versus WT FL), and showed only a slight decrease in starvation induced autophagy (p = 0.77 x10^{-6} for $FL^{\Delta(31\text{-}123)}$ versus WT FL) (Figure. 4.32 B). Similar to WT FL, FL^{DM1}, FL^{DM2} and FL^{TM} do not impact autophagy levels in the nutrient-rich conditions (p = 0.52 for FL^{DM1} mutant versus WT FL; p = 0.88 for FL^{DM2} mutant versus WT FL; and p = 0.75 for FL^{TM} versus WT FL). Additionally, FL^{DM1} and FL^{DM2} also shows no significant difference in autophagy when compared to no expression conditions under both nutrient rich (p = 0.52 for FL^{DM1} mutant versus no expression and p = 0.93 for FL^{DM2} mutant versus and no expression) and starvation conditions (p = 0.92 for FL^{DM1} mutant versus no expression and p = 0.85 for FL^{DM2} mutant versus and no expression). Strikingly however, autophagy levels are significantly decreased upon starvation when either of the Cys double mutants or the Cys tetrad mutant is expressed (p = 0.71 x 10^{-2} for FL^{DM1} versus FL; p = 0.10 x 10^{-3} for FL^{DM2} versus FL; and p = 0.25 x 10^{-3} for FL^{TM} versus FL).

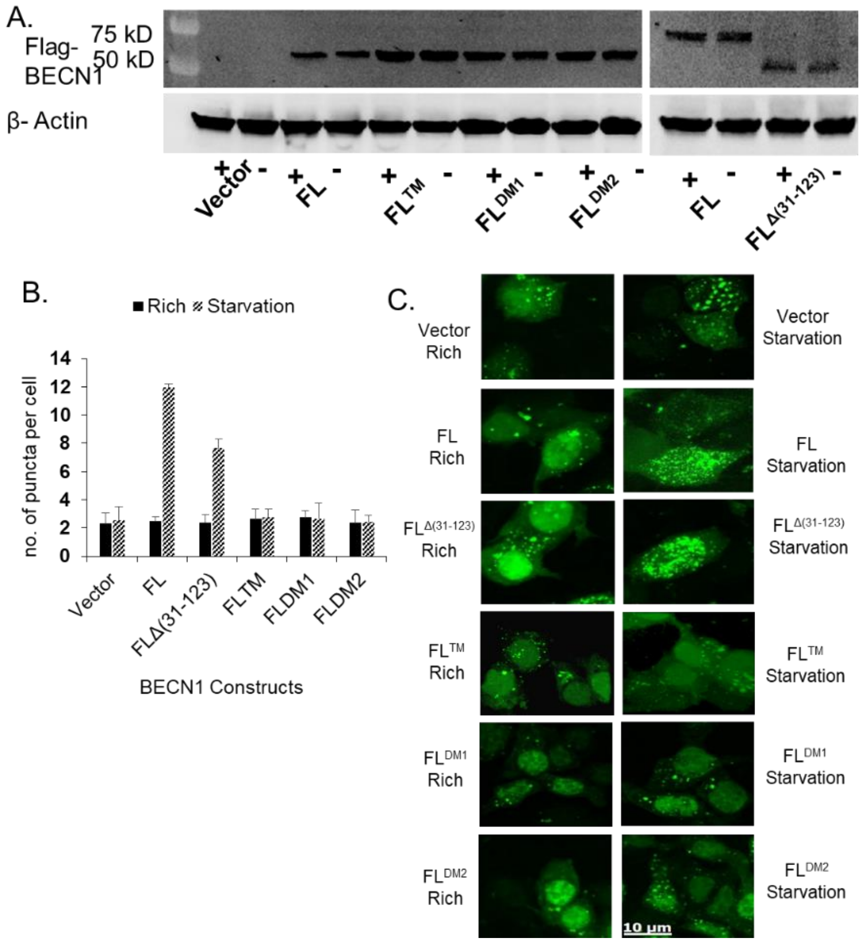


Figure 4.32: Effect of CxxC mutation on autophagy. A) Western blot of MCF7 cell lysates showing comparable expression of WT and mutant FLAG-BECN1 in nutrient-rich (+) and starvation (-) conditions, with actin as loading control. B) Light microscopy quantification of discrete GFP-LC3 puncta per cell in GFP-positive MCF7 cells co-transfected with GFP-LC3 and WT or mutant BECN1 as indicated below the x-axis. Bars represent the number of puncta per cell. Error bars represent standard deviation. C) Representative images of GFP-LC3 (green) fluorescence in cells grown in starvation or nutrient rich medium and transfected with mutant Flag-BECN1.

4.4. Discussions and conclusions

BECN1 is shown to bind Zn^{2+} in a 1:1 to molar ratio and the two invariant CxxC motifs are now shown to mediate this binding as tetrad mutation of these residues completely knocks out Zn2+ binding. Additionally, mutant proteins with either of the two CxxC motifs retain ~50% Zn^{2+}, indicating that each of the two CxxC motifs are equally responsible for this binding. When CxxC motifs were treated in molar excess of a Zn^{2+} chelating agent TPEN, it still retains significant amount of Zn^{2+}, whereas these CxxC motifs when treated with excess of reducing agent either DTT or TCEP, abrogates Zn^{2+} binding.

We also studied the binding induced helicity in BECN1 IDR by treating the BECN1 IDR using an increasing concentration of 2,2,2-trifluroethanol (TFE), that facilitates increased coil to helix transition upon increase in its concentration. We observed increased helical content in the BECN1 presence of both CxxC motifs compared to the cysteine tetrad mutants, which may imply that Zn^{2+} binding by BECN1 IDR makes BECN1 more helical. Additionally, when TFE concentration was increased to 15%, 25% and 40%, the helical content increased at the expense of the residues in the extended β and coil conformation. Additionally, in presence of the CxxC motifs an increase in TFE concentration increased the helical content in BECN1 IDR marginally, but showed a quicker transition compared to the cysteine tetrad mutant. BECN1 IDR contains three complete Anchor regions, sequences that are predicted to become helical upon binding to a binding partner. These Anchor regions comprises of residues 13-49, 79-103, 116-127 in the BECN1 IDR (Mei, Glover*, et al.*, 2016b) and a fourth Anchor region comprising of residues 137-145 extends into the FHD. Apart from these, we already know of two α-MORFs, which are disordered regions that become helical upon binding to its binding partner. One of those two α-MORFs is BECN1 BH3D (residues 105-130) which becomes helical upon binding to antiapoptotic BCL2 protein M11 (Sinha, Colbert, Becker, Wei & Levin, 2008) and residues 76-105, which includes the second Anchor region, is likely also an α-MORF as it becomes helical upon treatment with TFE (Glover et al., 2016).

We performed conformational studies using Small Angle X-ray Scattering (SAXS). The SAXS data shows that BECN1 IDR exists in a more compact conformation which is likely due to Zn^{2+} binding by BECN1 IDR. In the absence of Zn^{2+} binding by BECN1 IDR, in the CxxC tetrad mutant a clear trend of a disordered conformation is observed with increased size, shape and folded nature, compared to smaller size, compact shape and partially folded nature in the BECN1 IDR construct that binds Zn^{2+} via CxxC motifs. Additionally, from hydrogen- deuterium exchange mass spectrometry we further observed that CxxC motifs keep the BECN1 IDR marginally more conformationally compact, thus allowing increased deuterium uptake in the absence of the CxxC motifs.

From the SAXS result there was an indication that BECN1 IDR in presence of the CxxC motifs may have a oligomeric species at the leading edge of the SEC. The molecular mass estimated from the Volume of Correlation (Vc) from SEC-SAXS, relies upon having a high signal to noise ratio at q of scattering plot. As this molecular mass estimation was calculated from a scattering data with low signal to noise in the scattering plot, this oligomeric molecular mass estimation was not reliable.

Furthermore, in order to test any possibility of having an oligomeric molecular species of BECN1 IDR we performed a self-dissociation of MBP- tagged BECN1 IDR in addition to an untagged BECN1 IDR, which was used for SEC-SAXS experiment. Both the two constructs showed no heat of self-dissociation, which may indicate that BECN1 IDR is monomeric or the heat of dissociation was too low to be detected or the BECN1 IDR was too tight to be dissociated by ITC. Additionally, we currently have no evidence of believe that BECN1 IDR is an oligomer, as the previously discovered crystal structures (Li *et al.*, 2012a, Mei, Su, *et al.*, 2016b) of BECN1 shows that BECN1 CCD forms an anti-parallel homodimer and so if BECN1 IDR takes part in dimerization then one chain of BECN1 will have to extend all the way across the structured FHD and CCD to pair with the IDR of the opposite molecule. Therefore, since this dimerization event is physiologically irrelevant, we can only conclude that BECN1 IDR is a monomer.

Further, as we observed binding induced helicity in BECN1 IDR, we tested the importance of BECN1 CxxC motifs in interaction with such γ-HV68 anti-apoptotic BCL2 protein M11, as we know from previous studies that BECN1 BH3D becomes helical upon binding to the hydrophobic groove of M11(Sinha, Colbert, Becker, Wei & Levin, 2008). So, we hypothesized that BECN1 CxxC motifs may have a role in this disorder to helix transition. However, this study indicates that in presence of the CxxC motifs in BECN1 IDR, there was no improved interaction with M11, compared to a BECN1 IDR that do not have CxxC motifs. So, we concluded that CxxC motifs are not critical for M11 interaction and may need other residues in BECN1 IDR that may be essential for facilitating binding induced helicity.

Lastly, we performed autophagy assay to test the importance of BECN1 CxxC motifs and intervening IDR region (residues 31-123) in starvation induced autophagy. We found out that the BECN1 CxxC motifs are critical, but the intervening IDR is not essential for the function of BECN1 IDR in starvation induced autophagy. Further it can be implied that presence of Zn^{2+} binding may have some role in the regulation of the autophagy. Previous studies have shown that when the MCF-7 cells were treated with a cell permeable Zn^{2+} chelator for 72 hours, this process led to suppression of both basal and tamoxifen (a chemotherapeutic used to selectively bind to estrogen receptor, which helps in decreasing the growth of breast cancer cells) induced autophagy(Hwang et al., 2010). But when the cells were further replenished with excess Zn^{2+}, the cells showed enhanced basal and induced autophagy.

Additionally, starvation may induce generation of Reactive Oxygen Species (ROS) due to malfunctioning of mitochondrial electron transport chain, which may lead to leakage of electrons from NADH-Coenzyme Q oxidoreductase and Coenzyme Q-Cytochrome c oxidoreductase complex, resulting in partial oxygen reduction to superoxide radicals ($O2•^{-}$), the precursor of most other reactive oxygen species (ROS). This oxidative stress can then oxidize the cysteine thiols and thus promote Zn^{2+} coordination by the CxxC motifs and thereby lead to starvation induced autophagy. It has been shown that a nuclear protein High Mobility Group Box 1(HMGB1) translocate from nucleus to cytoplasm in response to ROS. HMGB1 may be involved in regulation of Bcl-2 phosphorylation by ERK/MAPK pathway(Tang et al., 2010), which then downregulates Bcl-2-BECN1 interaction and thereby upregulate starvation induced autophagy. Since our ICP-MS result shows that BECN1 IDR is reduction sensitive, it could be implied that BECN1 may only inefficiently bind to Zn^{2+} during basal cellular conditions, but upon starvation induced autophagy ROS can be generated, which may facilitate BECN IDR to bind to Zn^{2+} efficiently and show starvation induced autophagy, which may be upregulated due to a potential BECN1-HMGB1 interaction. In future it will be invaluable to delineate the importance of Zn^{2+} binding by BECN1 IDR in mediating potential protein-protein interaction that may show binding induced disorder to order transition. Additionally, it will be interesting to understand the mechanism of Zn^{2+} binding by BECN1 IDR under starvation induced upregulation of autophagy.

CHAPTER 5: CONCLUSIONS AND FUTURE DIRECTIONS

The work presented in this dissertation aims to understand the importance of the conformational flexibility of BECN1 and its impact on its function in starvation-induced autophagy. The focus of Chapter 2 is to investigate the role of BECN1 BH3 domain and the overlap helix (OH) in starvation induced autophagy. Chapter 3 is focused on delineating the proteins and protein regions required for interaction of CRM1 on BECN1. Chapter 4 focuses on investigating the importance of invariant CxxC motifs on the structure and conformational transitions of the BECN1 IDR, and also its function in autophagy.

In Chapter 2 we use cellular autophagy assays to show that deletion of the BECN1 BH3D does not impact autophagy levels in nutrient-rich or starvation media. This indicates that even though anti-apoptotic BCL2 proteins bind to BECN1 BH3D to downregulate autophagy, other domains in BECN1 may also be important for BCL2-mediated downregulation of autophagy. Other biophysical experiments conducted by Dr. Yue Li indicate that BECN1 regions outside BH3D, especially the IDR and FHD improve binding to M11, and in addition to the BH3D, the BH3D-FHD linker, the FHD, and the center of BECN1 CCD becomes more dynamic upon binding to M11. BECN1 homodimers represent the autophagy-inactive state of BECN1 as in this state BECN1 cannot interact with other autophagy proteins that bind via the CCD such as ATG14 or UVRAG. Therefore, it appears that BCL2 binding to the BECN1 BH3D stabilizes autophagy-inactive oligomeric state and prevents interactions of other BECN1 domains that are required for BECN1 function in autophagy. Other BECN1 binding partners such as VMP1 that bind via the BH3D, AMBRA1 that bind via the FHD, ATG14 and UVRAG that bind to the CCD, are known to upregulate autophagy. Additionally, there are other CCD containing proteins that down-regulate autophagy; as well as the nuclear exporter, CRM1. Thus, while the BH3D is not directly required for autophagy, it appears to enable binding of autophagy regulators that impact structure, oligomerization, and function of other BECN1 domains. Future studies may establish whether M11-induced conformational changes in the BH3D modulate conformations of the FHD and NES, and the mechanism behind these conformational changes.

In Chapter 2 we also use cellular autophagy assays to investigate the importance of OH interface residues responsible for stabilizing both of two mutually-exclusive packing states: against the partner helix in a CCD dimer or the BARAD. We show that alanine mutagenesis of four OH interface residues, V^{250}, M^{254}, L^{261} and L^{264}, causes complete abrogation of starvation induced autophagy, while basal autophagy in nutrient-rich conditions remains unaffected, indicating that these residues mediate interactions important for autophagy. This is likely because these OH interface residues help stabilize the parallel CCD heterodimer complexes, BECN1:ATG14 or BECN1:UVRAG, required for formation of the ternary PI3K/VPS34 complexes required for autophagy. A SAXS data-constrained model of the BECN1-ATG14 homodimer, indicates that four OH residues contribute to heterodimer coiled coil interface (V250:I165, M254:N169, A257:L172, L261:V176)(Mei, Su, *et al.*, 2016b). Notably, these OH residues are largely conserved, especially in vertebrates; and are predicted to pack against ATG14 residues that are also conserved. Amongst these interacting pairs, the least conserved BECN1 residue is M254, which is paired with N169, the least conserved paired residue in ATG14. Similarly, these OH residues likely also contribute to the parallel CCD heterodimer with UVRAG (Rostislavleva *et al.*, 2015a). Other biophysical experiments by Dr. Karen Glover (Glover *et al.*, 2017), showed that, in the absence of these other binding partners, the BECN1 OH packs preferentially with BECN1 BARAD in the autophagy-inactive BECN1 homodimer, weakening the homodimer, while simultaneously likely also preventing the BARAD from productive membrane interactions. Thus, the BECN1 OH probably serves as an important conformational switch-point regulating autophagy in cells.

In Chapter 3, we showed that purified CRM1 binds to full-length BECN1 and the BECN1 CCD-BARAD, but only in the presence of RanGTP. This indicates that a cooperative binding of RanGTP to CRM1 facilitated binding of BECN1 NES containing constructs. This is consistent with previous studies (Monecke *et al.*, 2013) which show that, in its inactive state, CRM1 exists in an extended toroid shape, with the C-terminal helix blocking the volume that is occupied by RanGTP in the active state of CRM1, which is required for the successful transport of cargo proteins from nucleus to cytoplasm. Strikingly, the BECN1 CCD does not bind to CRM1 even in the presence of RanGTP, suggesting that the BECN1 BARAD may play a role in this interaction. However, it is unclear whether the BARAD interacts directly with CRM1. Indeed, it is possible that as the BECN1 CCD forms an antiparallel homodimer (Li *et al.*, 2012a, Mei, Su, *et al.*, 2016b), within which the OH of one BECN1 molecule packs against the NES sequence of the other, the NES is sterically prevented from interacting with CRM1. As reported in our previous research presented in Chapter 2, in BECN1 fragments containing the BARAD, the OH packs against the BARAD (Glover *et al.*, 2017), releasing the NES sequence for other interactions such as with the CRM1. These results lay the groundwork for future studies investigating the interaction of BECN1 and CRM1. Thermodynamic analyses of the interaction will elucidate how different interactions compete. Structural studies will elucidate the BECN1 conformation that permits interactions with CRM1, the role of RanGTP in this interaction, whether BECn1 domains beyond the NES interact with CRM1 and the CRM1 and BECN1 residues at the interaction interface. Ultimately, this may add insights to the transport of BECN1 to the cytoplasm, and consequently to a unique mechanism of regulating autophagy. Lastly, it may even provide clues to the role of BECN1 in the nucleus.

In Chapter 4 we investigate the role of invariant CxxC motifs that flank the BECN1 IDR. We show that BECN1 binds Zn^{2+} in a 1:1 to molar ratio. Mutation of either one of the two invariant CxxC motifs, or treatment with molar excess of chelating agents reduces Zn^{2+}-binding. Mutation of both these invariant CxxC motifs or treatment with molar excess reducing agents such as DTT or TCEP abrogate Zn^{2+}-binding.

On the other hand, we also study the binding induced helicity in BECN1 IDR by treating the BECN1 IDR using an increasing concentration of 2,2,2-trifluroethanol (TFE), that allows increased coil to helix transition upon increase in its concentration. The hydrophobic nature of the compound mimics the hydrophobic packing in the proteins. We observed increased helical content in the BECN1 presence of both CxxC motifs compared to the cysteine tetrad mutants, which may imply that Zn^{2+} binding by BECN1 IDR makes BECN1 more helical. Additionally, when TFE concentration was increased to 15%, 25% and 40%, the residues in the helical content increased at the expense of the residues in the extended β and coil conformation. Additionally, in presence of the CxxC motifs an increase in TFE concentration increased the helical content in BECN1 IDR marginally, but showed quicker transition compared to the cysteine tetrad mutant. BECN1 IDR contain three complete Anchor regions, which are sequences that are predicted to become helical upon binding to a binding partner.

These Anchor regions comprises of residues 13-49, 79-103, 116-127 in the BECN1 IDR (Mei, Glover*, et al.*, 2016b) and a fourth Anchor region comprising of residues 137-145 extends into the FHD. Apart from these, we already know of two α-MORFs, which are disordered regions that become helical upon binding to its binding partner. One of those two α-MORFs is

BECN1 BH3D (residues 105-130) which becomes helical upon binding to antiapoptotic BCL2 protein M11 (Sinha, Colbert, Becker, Wei & Levin, 2008) and residues 76-105, which includes the second Anchor region, is likely also an α-MORF as it becomes helical upon treatment with TFE (Glover *et al.*, 2016).

We performed conformational dynamics studies using Small Angle X-ray Scattering (SAXS). The SAXS data shows that BECN1 IDR exist in a compact conformation which is likely due to Zn^{2+} binding by BECN1 IDR. In the absence of Zn^{2+} binding by BECN1 IDR, in the CxxC tetrad mutant a clear trend of a disordered conformation is observed with increased size, shape and folded nature, compared to smaller size, compact shape and partially folded nature in the BECN1 IDR construct that binds Zn^{2+} via CxxC motifs. Additionally, from hydrogen- deuterium exchange mass spectrometry we further observed that CxxC motifs in keep BECN1 IDR marginally more conformationally compact, thus allowing increased deuterium uptake in the absence of the CxxC motifs.

From the SAXS result there was an indication that BECN1 IDR in presence of the CxxC motifs may have a dimeric species at the leading edge of the SEC. The molecular mass estimated from the Volume of Correlation (V_c) from SEC-SAXS, relies upon having a high signal to noise ratio at q of scattering plot. As this molecular mass estimation was calculated from a scattering data with low signal to noise at the high q in the scattering plot, this dimeric molecular mass estimation was not reliable.

Furthermore, in order to test any possibility of having a dimeric molecular species of BECN1 IDR we performed a self-dissociation of MBP- tagged BECN1 IDR in addition to an untagged BECN1 IDR, which was used for SEC-SAXS experiment. We hypothesized that if there was a dimeric species of BECN1 IDR present at a relatively low concentration during

SAXS experiment, an increase in the concentration may push the equilibrium from a monomeric to a dimeric state, which can then be tested for self-dissociation using ITC. In addition to the MBP-fusion protein, we had also used untagged BECN1 IDR in order to eliminate any possibility of interference from MBP in BECN1 IDR self-dissociation. Both the two constructs showed no heat of self-dissociation, which may indicate that BECN1 IDR is monomeric or the heat of dissociation was too low to be detected or the BECN1 IDR was too tight to be dissociated by ITC. Additionally, we currently have no evidence of believe that BECN1 IDR is a dimer, as the previously discovered crystal structures (Li *et al.*, 2012a, Mei, Su, *et al.*, 2016b) of BECN1 shows that BECN1 CCD forms an anti-parallel homodimer and so if BECN1 IDR takes part in dimerization then one chain of BECN1 will have to extend all the way across the structured FHD and CCD to pair with the IDR of the opposite molecule. Therefore, since this dimerization event is physiologically irrelevant, we can only conclude that BECN1 IDR is a monomer.

Further, as we observed binding induced helicity in BECN1 IDR, we tested the importance of BECN1 CxxC motifs in interaction with such γHV68 anti-apoptotic BCL2 protein M11, as we know from previous studies that BECN1 BH3D becomes helical upon binding to the hydrophobic groove of M11(Sinha, Colbert, Becker, Wei & Levin, 2008). So, we hypothesized that BECN1 CxxC motifs may have a role in this disorder to helix transition. However, this study indicates that in presence of the CxxC motifs in BECN1 IDR, there was no improved interaction with M11, compared to a BECN1 IDR that do not have CxxC motifs. So, we concluded that CxxC motifs are not critical for M11 interaction and may need other residues in BECN1 IDR that may be essential for facilitating binding induced helicity.

Lastly, we performed autophagy assay to test the importance of BECN1 CxxC motifs and intervening IDR region (residues 31-123) in starvation induced autophagy. We found out that the

BECN1 CxxC motifs are critical, but the intervening IDR is not essential for the function of BECN1 IDR in starvation induced autophagy. Further it can be implied that presence of Zn^{2+} binding may have some role in the regulation of the autophagy. Previous studies have shown that when the MCF-7 cells were treated with a cell permeable Zn^{2+} chelator for 72 hours, this process led to suppression of both basal and tamoxifen (a chemotherapeutic used to selectively bind to estrogen receptor, which helps in decreasing the growth of breast cancer cells) induced autophagy(Hwang *et al.*, 2010). But when the cells were further replenished with excess Zn^{2+}, the cells showed enhanced basal and induced autophagy.

Additionally, starvation may induce generation of Reactive Oxygen Species (ROS) due to malfunctioning of mitochondrial electron transport chain, which may lead to leakage of electrons from NADH-Coenzyme Q oxidoreductase and Coenzyme Q-Cytochrome c oxidoreductase complex, resulting in partial oxygen reduction to superoxide radicals ($O_2^{\bullet -}$), the precursor of most other reactive oxygen species (ROS). This oxidative stress can then oxidize the cysteine thiols and thus promote Zn^{2+} coordination by the CxxC motifs and thereby lead to starvation induced autophagy. It has been shown that a nuclear protein High Mobility Group Box 1(HMGB1) translocate from nucleus to cytoplasm in response to ROS. HMGB1 may be involved in regulation of Bcl-2 phosphorylation by ERK/MAPK pathway(Tang *et al.*, 2010), which then downregulates Bcl-2-BECN1 interaction and thereby upregulate starvation induced autophagy. Since our ICP-MS result shows that BECN1 IDR is reduction sensitive, it could be implied that BECN1 may only inefficiently bind to Zn^{2+} during basal cellular conditions, but upon starvation induced autophagy ROS can be generated, which may facilitate BECN IDR to bind to Zn^{2+} efficiently and show starvation induced autophagy, which may be upregulated due to a potential BECN1-HMGB1 interaction. In future it will be invaluable to delineate the importance of Zn^{2+} binding by BECN1 IDR in mediating potential protein-protein interaction that may show binding induced disorder to order transition. Additionally, it will be interesting to understand the mechanism of Zn^{2+} binding by BECN1 IDR under starvation induced upregulation of autophagy.

REFERENCES

Akey, C. W. & Radermacher, M. (1993). *Journal of Cell Biology* **122**, 1-19.

Andreini, C., Banci, L., Bertini, I. & Rosato, A. (2006). *Journal of Proteome Research* **5**, 196-201.

Baskaran, S., Carlson, L., Stjepanovic, G., Young, L., Kim, D., Grob, P., Stanley, R., Nogales, E. & Hurley, J. (2014). *Elife* **3**.

Beauchemin, D. (2017). *Encyclopedia of Spectroscopy and Spectrometry (Third Edition)*, edited by J. C. Lindon, G. E. Tranter & D. W. Koppenaal, pp. 236-245. Oxford: Academic Press.

Beck, M., Förster, F., Ecke, M., Plitzko, J. M., Melchior, F., Gerisch, G., Baumeister, W. & Medalia, O. (2004). *Science* **306**, 1387.

Beck, M. & Hurt, E. (2017). *Nature Reviews Molecular Cell Biology* **18**, 73-89.

Bernado, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. (2007). *J Am Chem Soc* **129**, 5656-5664.

Bernado, P. & Svergun, D. I. (2012). *Methods in molecular biology (Clifton, N.J.)* **896**, 107-122.

Bohgaki, T., Bohgaki, M. & Hakem, R. (2010). *Genome Integr*, p. 15.

Botti, J., Djavaheri-Mergny, M., Pilatte, Y. & Codogno, P. (2006). *Autophagy* **2**, 67-73.

Brautigam, C. A., Zhao, H., Vargas, C., Keller, S. & Schuck, P. (2016). *Nature Protocols* **11**, 882-894.

Burri, L. & Lithgow, T. (2004). *Traffic (Copenhagen, Denmark)* **5**, 45-52.

Callan, H. G. & Tomlin, S. G. (1950). *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)* **137**, 367-378.

Campobasso, N. & Huddler, D. (2015). *Bioorganic & Medicinal Chemistry Letters* **25**, 3771-3776.

Castagnetto, J. M., Hennessy, S. W., Roberts, V. A., Getzoff, E. D., Tainer, J. A. & Pique, M. E. (2002). *Nucleic Acids Res* **30**, 379-382.

Culik, R. M., Abaskharon, R. M., Pazos, I. M. & Gai, F. (2014). *J Phys Chem B* **118**, 11455-11461.

Darsow, T., Rieder, S. E. & Emr, S. D. (1997). *The Journal of cell biology* **138**, 517-529.

Diella, F., Haslam, N., Chica, C., Budd, A., Michael, S., Brown, N. P., Trave, G. & Gibson, T. J. (2008). *Front Biosci* **13**, 6580-6603.

Dong, X., Biswas, A. & Chook, Y. M. (2009). *Nature structural & molecular biology* **16**, 558-560.

Dong, X., Biswas, A., Süel, K. E., Jackson, L. K., Martinez, R., Gu, H. & Chook, Y. M. (2009). *Nature* **458**, 1136-1141.

Dunker, A. K. & Obradovic, Z. (2001). The protein trinity—linking function and disorder, Vol. 19, Nature Biotechnology, pp. 805-806. Nature Publishing Group.

Durand, D., Vivès, C., Cannella, D., Pérez, J., Pebay-Peyroula, E., Vachette, P. & Fieschi, F. (2010). *Journal of Structural Biology* **169**, 45-53.

Dyson, H. J. & Wright, P. E. (2005). *Nat Rev Mol Cell Biol* **6**, 197-208.

Fischer, U., Schäuble, N., Schütz, S., Altvater, M., Chang, Y., Faza, M. B. & Panse, V. G. (2015). *Elife* **4**.

Franke, D., Jeffries, C. M. & Svergun, D. I. (2015). *Nature methods* **12**, 419-422.

Franke, D., Petoukhov, M. V., Konarev, P. V., Panjkovich, A., Tuukkanen, A., Mertens, H. D. T., Kikhney, A. G., Hajizadeh, N. R., Franklin, J. M., Jeffries, C. M. & Svergun, D. I. (2017). *Journal of Applied Crystallography* **50**, 1212-1225.

Freyer, M. W. & Lewis, E. A. (2008). *Methods in cell biology* **84**, 79-113.

Fung, H. Y. J., Fu, S.-C., Brautigam, C. A. & Chook, Y. M. (2015). *eLife* **4**, e10034.

Gall, J. G. (1967). *The Journal of cell biology* **32**, 391-399.

Glatter, O. (1977). *Journal of Applied Crystallography* **10**, 415-421.

Glick, D., Barth, S. & Macleod, K. F. (2010). *The Journal of pathology* **221**, 3-12.

Glover, K., Li, Y., Mukhopadhyay, S., Leuthner, Z., Chakravarthy, S., Colbert, C. L. & Sinha, S. C. (2017). *Journal of Biological Chemistry*.

Glover, K., Mei, Y. & Sinha, S. (2016). *Biochimica Et Biophysica Acta-Proteins and Proteomics* **1864**, 1455-1463.

Güttler, T., Madl, T., Neumann, P., Deichsel, D., Corsini, L., Monecke, T., Ficner, R., Sattler, M. & Görlich, D. (2010). *Nature Structural & Molecular Biology* **17**, 1367-1376.

Habchi, J., Tompa, P., Longhi, S. & Uversky, V. N. (2014).

Hajizadeh, N. R., Franke, D., Jeffries, C. M. & Svergun, D. I. (2018). *Scientific Reports* **8**, 7204.

Hopkins, J. B., Gillilan, R. E. & Skou, S. T. (2017). *Journal of applied crystallography*.

Houtman, J. C., Brown, P. H., Bowden, B., Yamaguchi, H., Appella, E., Samelson, L. E. & Schuck, P. (2007). *Protein Sci* **16**, 30-42.

Huang, W., Choi, W., Hu, W., Mi, N., Guo, Q., Ma, M., Liu, M., Tian, Y., Lu, P., Wang, F.-L., Deng, H., Liu, L., Gao, N., Yu, L. & Shi, Y. (2012). *Cell Res* **22**, 473-489.

Hung, H.-H., Huang, W.-P. & Pan, C.-Y. (2013). *Cell Biology and Toxicology* **29**, 415-429.

Hwang, J. J., Kim, H. N., Kim, J., Cho, D. H., Kim, M. J., Kim, Y. S., Kim, Y., Park, S. J. & Koh, J. Y. (2010). *Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine* **23**, 997-1013.

Itakura, E. & Mizushima, N. (2009). *Autophagy* **5**, 534-536.

Jones, S. & Thornton, J. M. (1996). *Proc Natl Acad Sci U S A* **93**, 13-20.

Kabeya, Y., Mizushima, N., Ueno, T., Yamamoto, A., Kirisako, T., Noda, T., Kominami, E., Ohsumi, Y. & Yoshimori, T. (2000). *Embo J* **19**, 5720-5728.

Kawamata, T., Horie, T., Matsunami, M., Sasaki, M. & Ohsumi, Y. (2017). *The Journal of biological chemistry* **292**, 8520-8530.

Kikhney, A. G. & Svergun, D. I. (2013). *FEBS Lett* **589**, 2570-2577.

Kikhney, A. G. & Svergun, D. I. (2015). *FEBS Letters* **589**, 2570-2577.

Kirby, N., Cowieson, N., Hawley, A. M., Mudie, S. T., McGillivray, D. J., Kusel, M., Samardzic-Boban, V. & Ryan, T. M. (2016). *Acta Crystallogr D Struct Biol* **72**, 1254-1266.

Kırlı, K., Karaca, S., Dehne, H. J., Samwer, M., Pan, K. T., Lenz, C., Urlaub, H. & Görlich, D. (2015). *eLife* **4**, e11466.

Klionsky, D. J., Baehrecke, E. H., Brumell, J. H., Chu, C. T., Codogno, P., Cuervo, A. M., Debnath, J., Deretic, V., Elazar, Z., Eskelinen, E.-L., Finkbeiner, S., Fueyo-Margareto, J., Gewirtz, D. A., Jäättelä, M., Kroemer, G., Levine, B., Melia, T. J., Mizushima, N., Rubinsztein, D. C., Simonsen, A., Thorburn, A., Thumm, M. & Tooze, S. A. (2011). *http://dx.doi.org/10.4161/auto.7.11.17661*.

Kochanczyk, T., Drozd, A. & Krezel, A. (2015). *Metallomics : integrated biometal science* **7**, 244-257.

Konermann, L., Pan, J. & Liu, Y. (2011). *Chemical Society Reviews* **40**, 1224-1234.

Kroemer, G., Marino, G. & Levine, B. (2010). *Molecular cell* **40**, 280-293.

Ku, B., Woo, J.-S., Liang, C., Lee, K.-H., Hong, H.-S., E, X., Kim, K.-S., Jung, J. U. & Oh, B.-H. (2008). *PLoS Pathog* **4**, e25-e25.

Kumar, M., Gouw, M., Michael, S., Sámano-Sánchez, H., Pancsa, R., Glavina, J., Diakogianni, A., Valverde, J. A., Bukirova, D., Čalyševa, J., Palopoli, N., Davey, N. E., Chemes, L. B. & Gibson, T. J. (2020). *Nucleic Acids Research* **48**, D296-D306.

Lee, E., Perugini, M., Pettikiriarachchi, A., Evangelista, M., Keizer, D., Yao, S. & Fairlie, W. (2016). *Autophagy* **12**, 460-471.

Lee, S.-J. & Koh, J.-Y. (2010). *Molecular Brain* **3**, 30.

Levine, B. & Klionsky, D. J. (2004). *Dev Cell* **6**, 463-477.

Levine, B. & Kroemer, G. (2008). *Cell* **132**, 27-42.

Li, X., He, L., Che, K. H., Funderburk, S. F., Pan, L., Pan, N., Zhang, M., Yue, Z. & Zhao, Y. (2012a). *Nature communications* **3**, 662.

Li, X., He, L., Che, K. H., Funderburk, S. F., Pan, L., Pan, N., Zhang, M., Yue, Z. & Zhao, Y. (2012b). *Nature communications* **3**, 662-662.

Liang, C., Feng, P., Ku, B., Dotan, I., Canaani, D., Oh, B.-H. & Jung, J. U. (2006). *Nature Cell Biology* **8**, 688-698.

Liang, X. H., Jackson, S., Seaman, M., Brown, K., Kempkes, B., Hibshoosh, H. & Levine, B. (1999). *Nature* **402**, 672-676.

Liang, X. H., Yu, J., Brown, K. & Levine, B. (2001). *Cancer Research* **61**, 3443.

Liuzzi, J. P., Guo, L., Yoo, C. & Stewart, T. S. (2014). *Biometals* **27**, 1087-1096.

Liuzzi, J. P. & Yoo, C. (2013). *Biological trace element research* **156**, 350-356.

Lobley, A., Whitmore, L. & Wallace, B. A. (2002). *Bioinformatics* **18**, 211-212.

Luis, G. a.-F., Ramiro, Téllez, S., Indalecio, Q.-S. & Carmen, B. n. (2011).

Luo, P. & Baldwin, R. L. (1997). *Biochemistry* **36**, 8413-8421.

Masson, G. R., Burke, J. E., Ahn, N. G., Anand, G. S., Borchers, C., Brier, S., Bou-Assaf, G. M., Engen, J. R., Englander, S. W., Faber, J., Garlish, R., Griffin, P. R., Gross, M. L., Guttman, M., Hamuro, Y., Heck, A. J. R., Houde, D., Iacob, R. E., Jørgensen, T. J. D., Kaltashov, I. A., Klinman, J. P., Konermann, L., Man, P., Mayne, L., Pascal, B. D., Reichmann, D., Skehel, M., Snijder, J., Strutzenberg, T. S., Underbakke, E. S., Wagner, C., Wales, T. E., Walters, B. T., Weis, D. D., Wilson, D. J., Wintrode, P. L., Zhang, Z., Zheng, J., Schriemer, D. C. & Rand, K. D. (2019). *Nature methods*, pp. 595-602.

Matsunaga, K., Saitoh, T., Tabata, K., Omori, H., Satoh, T., Kurotori, N., Maejima, I., Shirahama-Noda, K., Ichimura, T., Isobe, T., Akira, S., Noda, T. & Yoshimori, T. (2009). *Nature cell biology* **11**, 385-396.

Maul, G. G. (1971). *Journal of Cell Biology* **51**, 558-563.

Mayer, A. & Wickner, W. (1997). *The Journal of cell biology* **136**, 307-317.

Mei, Y., Glover, K., Su, M. & Sinha, S. C. (2016a). *Protein Sci* **25**, 1767-1785.

Mei, Y., Glover, K., Su, M. F. & Sinha, S. C. (2016b). *Protein Sci* **25**, 1767-1785.

Mei, Y., Ramanathan, A., Glover, K., Stanley, C., Sanishvili, R., Chakravarthy, S., Yang, Z., Colbert, C. L. & Sinha, S. (2016). *Biochemistry* **55**, 1945-1958.

Mei, Y., Su, M., Sanishvili, R., Chakravarthy, S., Colbert, C. & Sinha, S. (2016a). *Biochemistry* **55**, 4239-4253.

Mei, Y., Su, M., Sanishvili, R., Chakravarthy, S., Colbert, C. L. & Sinha, S. C. (2016b). *Biochemistry* **55**, 4239-4253.

Mei, Y., Su, M., Soni, G., Salem, S., Colbert, C. & Sinha, S. (2013). *Proteins-Structure Function and Bioinformatics* **82**, 565-578.

Meisburger, S. P., Taylor, A. B., Khan, C. A., Zhang, S., Fitzpatrick, P. F. & Ando, N. (2016). *Journal of the American Chemical Society* **138**, 6506-6516.

Mészáros, B., Simon, I. & Dosztányi, Z. (2009). *PLOS Computational Biology* **5**, e1000376.

Miyamoto, Y., Yoneda, Y. & Oka, M. (2018). *Nuclear Architecture and Dynamics*, edited by C. Lavelle & J.-M. Victor, pp. 387-403. Boston: Academic Press.

Mizushima, N. (2010). *Curr Opin Cell Biol* **22**, 132-139.

Mizushima, N., Yoshimori, T. & Levine, B. (2010). *Cell* **140**, 313-326.

Monecke, T., Haselbach, D., Voß, B., Russek, A., Neumann, P., Thomson, E., Hurt, E., Zachariae, U., Stark, H., Grubmüller, H., Dickmanns, A. & Ficner, R. (2013). *Proceedings of the National Academy of Sciences* **110**, 960.

Nemergut, M. E. & Macara, I. G. (2000). *Journal of Cell Biology* **149**, 835-850.

Núñez, S., Venhorst, J. & Kruse, C. G. (2012). *Drug discovery today* **17**, 10-22.

Oberstein, A., Jeffrey, P. & Shi, Y. (2007). Crystal Structure of the Bcl-XL-Beclin 1 Peptide Complex

Pace, N. J. & Weerapana, E. (2014). *Biomolecules*, pp. 419-434.

Pattingre, S., Tassa, A., Qu, X. P., Garuti, R., Liang, X. H., Mizushima, N., Packer, M., Schneider, M. D. & Levine, B. (2005). *Cell* **122**, 927-939.

Popelka, H., Uversky, V. N. & Klionsky, D. J. (2014). *Autophagy* **10**, 1093-1104.

Porcheron, G., Garénaux, A., Proulx, J., Sabri, M. & Dozois, C. M. (2013). *Frontiers in Cellular and Infection Microbiology* **3**, 90.

Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. (2007). *Q Rev Biophys* **40**, 191-285.

Rambo, R. P. & Tainer, J. A. (2013). *Nature* **496**, 477-481.

Rostislavleva, K., Soler, N., Ohashi, Y., Zhang, L., Pardon, E., Burke, J., Masson, G., Johnson, C., Steyaert, J., Ktistakis, N. & Williams, R. (2015a). *Science* **350**.

Sato, T. K., Darsow, T. & Emr, S. D. (1998). *Mol Cell Biol* **18**, 5308-5319.

Scheuermann, T. H. & Brautigam, C. A. (2015). *Methods (San Diego, Calif.)* **76**, 87-98.

Shao, D., Oka, S., Liu, T., Zhai, P., Ago, T., Sciarretta, S., Li, H. & Sadoshima, J. (2014). *Cell metabolism* **19**, 232-245.

Sheffield, P., Garrard, S. & Derewenda, Z. (1999). *Protein Expression and Purification* **15**, 34-39.

Singh, A. K. (2016). *Engineered Nanoparticles*, edited by A. K. Singh, pp. 77-123. Boston: Academic Press.

Sinha, S., Colbert, C. L., Becker, N., Wei, Y. & Levin, B. (2008). *Autophagy* **4**, 989-997.

Sinha, S. & Levine, B. (2008). *Oncogene* **27**, S137-S148.

Sinha, S. C., Li, Y., Mukhopadhyay, S., Wyatt, S. & Dasanna, S. (2019). *Handbook of Famine, Starvation, and Nutrient Deprivation: From Biology to Policy*, edited by V. R. Preedy & V. B. Patel, pp. 2045-2065. Cham: Springer International Publishing.

Sreerama, N., Venyaminov, S. & Woody, R. (2000). *Analytical Biochemistry* **287**, 243-251.

Sreerama, N., Venyaminov, S. & Woody, R. (2001). *Analytical Biochemistry* **299**, 271-274.

Sreerama, N. & Woody, R. (2000). *Analytical Biochemistry* **287**, 252-260.

Stoffler, D., Fahrenkrog, B. & Aebi, U. (1999). *Current Opinion in Cell Biology* **11**, 391-401.

Su, M., Mei, Y., Sanishvili, R., Levine, B., Colbert, C. L. & Sinha, S. (2014). *J Biol Chem*, pp. 8029-8040.

Svergun, D. I. (1992). *Journal of Applied Crystallography* **25**, 495-503.

Tang, D., Kang, R., Livesey, K. M., Cheh, C. W., Farkas, A., Loughran, P., Hoppe, G., Bianchi, M. E., Tracey, K. J., Zeh, H. J. & Lotze, M. T. (2010). *J Cell Biol*, pp. 881-892.

Tria, G., Mertens, H. D., Kachala, M. & Svergun, D. I. (2015). *IUCrJ* **2**, 207-217.

van der Lee, R., Buljan, M., Lang, B., Weatheritt, R. J., Daughdrill, G. W., Dunker, A. K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D., Kim, P. M., Kriwacki, R., Oldfield, C. J., Pappu, R. V., Tompa, P., Uversky, V. N., Wright, P. & Babu, M. M. (2014). *Chem Rev* **114**, 6589-6631.

Wei, Y., Thyparambil, A. A. & Latour, R. A. (2014). *Biochimica et biophysica acta* **1844**, 2331-2337.

Wen, W., Harootunian, A. T., Adams, S. R., Feramisco, J., Tsien, R. Y., Meinkoth, J. L. & Taylor, S. S. (1994). *Journal of Biological Chemistry* **269**, 32214-32220.

Whitmore, L. & Wallace, B. A. (2004). *Nucleic Acids Res* **32**, W668-673.

Whitmore, L. & Wallace, B. A. (2008). *Biopolymers* **89**, 392-400.

Whitten, A. E., Cai, S. & Trewhella, J. (2008). *Journal of Applied Crystallography* **41**, 222-226.

Wiseman, T., Williston, S., Brandts, J. F. & Lin, L.-N. (1989). *Analytical Biochemistry* **179**, 131-137.

Wright, P. E. & Dyson, H. J. (1999). *J Mol Biol* **293**, 321-331.

Wyatt, P. (1993). *Analytica Chimica Acta* **272**, 1-40.

Xu, F., Fang, Y., Yan, L., Xu, L., Zhang, S., Cao, Y., Xu, L., Zhang, X., Xie, J., Jiang, G., Ge, C., An, N., Zhou, D., Yuan, N. & Wang, J. (2017). *Scientific Reports* **7**, 45385.

Yang, Q., Rout, M. P. & Akey, C. W. (1998). *Molecular Cell* **1**, 223-234.

Young, L. N., Cho, K., Lawrence, R., Zoncu, R. & Hurley, J. H. (2016). *Proceedings of the National Academy of Sciences* **113**, 8224.

Young, L. N., Goerdeler, F. & Hurley, J. H. (2019). *Proceedings of the National Academy of Sciences* **116**, 21508.

Zaffagnini, G. & Martens, S. (2016). *Journal of Molecular Biology* **428**, 1714-1724.

Zhao, H., Piszczek, G. & Schuck, P. (2015). *Methods (San Diego, Calif.)* **76**, 137-148.

Zhao, H. & Schuck, P. (2015). *Acta crystallographica. Section D, Biological crystallography* **71**, 3-14.

Zheng, A., Kallio, A. & Härkönen, P. (2007). *Endocrinology* **148**, 2764-2777.