UNDERSTANDING BINDING-INDUCED DISORDER-TO-ORDER AND CONFORMATIONAL TRANSITIONS IN PROTEINS THAT REGULATE AUTOPHAGY

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

By

Karen Marie Glover

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

Major Department:
Chemistry and Biochemistry

June 2018

Fargo, North Dakota

North Dakota State University
Graduate School

Title

UNDERSTANDING BINDING-INDUCED DISORDER-TO-ORDER AND CONFORMATIONAL TRANSITIONS IN PROTEINS THAT REGULATE AUTOPHAGY

By

Karen Marie Glover

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. Gregory Cook

Dr. Stuart Haring

Dr. Penelope Gibbs

Approved:

26 June 2018	Dr. Gregory Cook
Date	Department Chair

ABSTRACT

Autophagy, a cellular homeostasis process that degrades and recycles cytosolic contents, is regulated at various stages by protein interactions. BECN1 regulates autophagy through interactions with diverse proteins via binding-induced conformational changes in its intrinsically disordered region (IDR) and flexible domains. Understanding the structure-function relationship of conformational transitions provides a basis for understanding these interactions and may be useful for targeting therapeutics for these proteins. We devised a method to identify IDRs likely to become helical upon binding, even in the absence of known binding partners, and analyzed BECN1 interactions to further understand how binding-induced conformational changes affect BECN1-mediated autophagy regulation.

We showed that the BECN1 coiled-coil domain (CCD) and β-α repeated, autophagy-specific domain (BARAD) share an overlapping helix (OH) region that transiently adopts mutually exclusive packing states and transiently packs against the CCD and BARAD, but predominantly packs against the BARAD in the homodimer, and demonstrated that mutation of OH residues that pack against both domains abrogates starvation-induced up-regulation of autophagy. Our results have important implications for the relative stability of autophagy-inactive and autophagy-active BECN1 complexes.

We also attempted to establish the mechanism by which various virus-encoded proteins regulate autophagy. Nef, a critical protein for HIV pathogenesis, interacts with BECN1 BARAD regions that pack against the OH. However, we show that Nef does not appear to interact directly with the BECN1 homodimer, suggesting that packing of the OH against the BARAD may inhibit the Nef interaction. Several anti-apoptotic viral BCL2 (vBCL2) proteins down-regulate autophagy by binding the BECN1 BH3 homology domain (BH3D) and mediating binding-induced helical transitions in the BH3D. Although protein purification issues prevented analysis of Kaposi's sarcoma herpesvirus (KSHV) vBCL2 interactions, we determined that BHRF1, an Epstein-Barr virus (EBV)-encoded vBCL2, down-regulates autophagy by binding a flexible region of BECN1 that includes the BH3D. Further, we solved the 2.6 Å crystal structure of BHRF1 bound to the BID BH3D, a pro-apoptotic BCL2, and showed that the BID BH3D binds ~400-times tighter than the BECN1 BH3D to BHRF1 and inhibits BHRF1-mediated down-regulation of autophagy.

Together, our studies highlight the importance of binding-induced conformational changes in proteins that regulate autophagy.

ACKNOWLEDGEMENTS

To my advisor, Dr. Sangita Sinha, I express my gratitude for her generous guidance, support, and mentorship.

I also thank my committee members: Dr. Gregory Cook, Dr. Stuart Haring, and Dr. Penelope Gibbs for their advice and direction.

I also extend thanks to our collaborator Dr. Christopher Colbert at NDSU for generously sharing his expertise, advice, and support, and our other collaborators, especially: Dr. Srinivas Chakravarthy at Bio-CAT, APS, Argonne.

I am further grateful to current and former members of the Sinha group, including Yang Mei, Minfei Su, Yue Li, Shreya Mukhopadhyay, Srinivas Dasanna, Elizabeth Bueno, and Samuel Wyatt, and the Colbert group, including Jaime Jensen, Shane Wyborny, Benjamin LeVahn, and Beau Jernberg. Thank you for constant advice, assistance, and encouragement.

This research was supported an funded by NIH grants RO3 NS090939, R15 GM122035 (S.S.), P20 RR015566 (S.S), R21 AI078198 (S.S), and R15 GM113227 (C.C.); a NSF grant MCB-1413525 (S.S.); a ND Dept. of Commerce Award #14-11-J1-73 (S.S), a NDSU Graduate School doctoral dissertation award (K.G.), and a ND EPSCoR doctoral dissertation award (K.G., 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. Work performed at NE-CAT was supported by NIH NIGMS P41 GM103403. This research used resources of the APS, a U.S. DOE Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

This research used resources of the Advanced Photon Source, a U.S. DOE Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. I also acknowledge the NDSU Core Biology Facility (funded by NIH grant P30 GM103332-01) for access to tissue culture facilities and the NDSU Advanced Imaging and Microscopy Core Laboratory for access to microscopy and imaging equipment. I also thank Dr. Pawel Borowicz for assistance with imaging and puncta quantification method development.

DEDICATION

For my husband Josh, who is my favorite.

For my daughter Indiana, who inspires me to be the best version of myself.

For my parents, who encourage me in all things great and small.

And for my daughter, who wrote this with me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3-AT .....3-Amino-1,2,4-Trizole.

Å.....Angstrom.

AD .....Activation Domain.

AFM .....Aromatic Finger Mutation.

AIDS.....Acquired Immune Deficiency Syndrome.

ATG.....Autophagy Related Protein.

BARAD.....β-α Repeated, Autophagy-Specific Domain.

BCL2 .....B Cell Lymphoma 2.

BCL2BD .....BCL2 Binding Domain.

BH3D .....BH Homology 3 Domain.

BHRF1 .....BamHI H Fragment Rightward Reading Frame 1.

BME .....β-mercaptoethanol.

BSA.....Buried Surface Area.

CCD .....Coiled-Coil Domain.

CD.....Circular Dichroism Spectroscopy.

DBD .....DNA Binding Domain.

D_{max} .....Maximum Particle Size.

DMEM .....Dulbecco Modified Eagle Medium.

DTT .....Dithiothreitol.

EBSS .....Earle's Balanced Salt Solution.

EBV.....Epstein-Barr Virus.

E. coli .....*Escherichia coli*.

EDTA .....Ethylenediaminetetraacetic Acid.

EOM.....Ensemble Optimization Method.

FBS .....Fetal Bovine Serum.

FHD.....Flexible Helical Domain.

FL.....Full Length.

ΔG	Change in Gibbs Free Energy.
GFP	Green Fluorescent Protein.
GST	Glutathione S-Transferase.
ΔH	Change in Enthalpy.
HEAT	Huntington, Elongation Factor 3 Protein Phosphatase 2A, Yeast Kinase TOR1.
His_6	Hexa-histadine.
HIV	Human Immunodeficiency Virus.
HSV	Herpes Virus.
I(0)	Scattering Intensity at Zero Angle.
I(q)	Scattering Intensity.
IDR	Intrinsically Disordered Region.
IPTG	Isopropyl Thio-β-D-Galactoside.
ITC	Isothermal Titration Calorimetry.
K_d	Dissociation Constant.
KSHV	Kaposi's Sarcoma Herpes Virus.
LC3	Light Chain 3.
MBP	Maltose Binding Protein.
MoRF	Molecular Recognition Feature.
MW	Molecular Weight.
Nef	Negative Factor.
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.

ΔS ..... Change in Entropy.

SAXS ..... Small Angle X-ray Scattering.

SDS ..... Sodium Dodecyl Sulfate.

SEC ..... Size Exclusion Chromatography.

SL ..... Short, Triple Alanine, Linker.

TAT ..... Transcriptional Activator Protein Transduction Domain.

TEV ..... Tobacco Etch Virus.

TFE ..... 2,2,2-Trifluoroethanol.

TM ..... Transmembrane.

T_m ..... Melting Temperature.

vBCL2 ..... Viral BCL2.

VPS ..... Vacuolar Protein Sorting.

WT ..... Wild-type.

Θ ..... Molar ellipticity.

CHAPTER 1. INTRODUCTION

1.1. BECN1 in autophagy

Macroautophagy, hereafter called autophagy, is a conserved cellular homeostasis pathway facilitating nutrient recycling via engulfment of long-lived, misfolded, aggregated, damaged, or mutated proteins, organelles, and pathogens in multi-layered vesicles called autophagosomes that fuse with lysosomes to enable degradation of the sequestered contents[1-5] (Figure 1.1). Thus, autophagy is a survival mechanism that enables cells to withstand nutrient deprivation, environmental stressors, aging, and infection by recycling intracellular debris to generate metabolic precursors, such as amino acids and ATP. Proteins that execute autophagy are conserved in all eukaryotes and many of these conserved autophagy-related effectors are called ATG proteins[5-8]. Multi-protein complexes mediate each stage of autophagy, which include 1) initiation signaling, 2) autophagosome nucleation; which initiates synthesis of the autophagic isolation membrane that eventually encloses the cytoplasmic cargo destined for degradation, 3) autophagosome expansion, which involves elongation of the isolation membrane resulting in an autophagosome that contains multiple phospholipid bilayer membranes surrounding the cytoplasmic cargo destined for degradation, and 4) autophagosome maturation, which involves docking and fusion with the lysosome, whose degradative enzymes catalyze the breakdown of the autolysosomal contents[5-8].

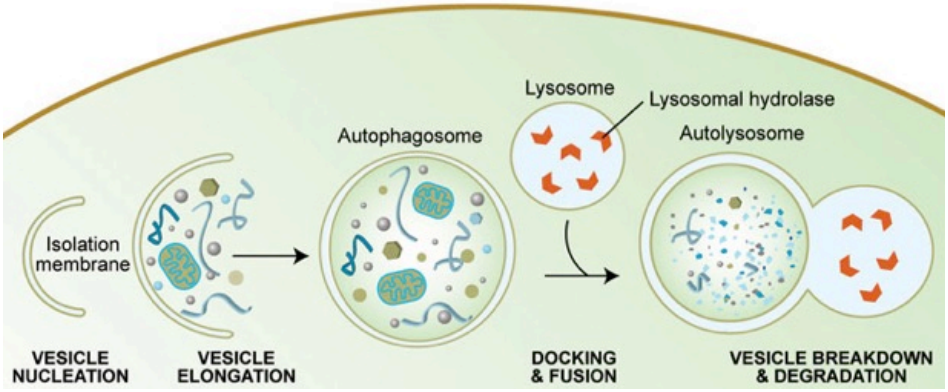

Figure 1.1. Steps of autophagy.
Upon autophagy initiation, the isolation membrane begins forming during the vesicle nucleation step. The autophagy machinery coordinates vesicle elongation and autophagosome formation to engulf cytoplasmic contents. The outer membranes of the autophagosome and lysosome fuse during the docking and fusion stage to generate the autolysosome. The lysosomal hydrolases then facilitate the breakdown and degradation of the sequestered material within the autolysosome. This figure is adapted from[9].

BECN1/Beclin 1/ATG6/VPS30, first identified as a B cell lymphoma 2 (BCL2)-interacting protein in humans[10] and amongst the first autophagy proteins to be identified in mammals[4], is highly conserved in

eukaryotes. As a core component of the class III phosphatidylinositol 3-kinase (PI3KC3) complexes essential for autophagosome nucleation and maturation[11-15], BECN1 forms a parallel heterodimer with either ATG14 or UVRAG/VPS38 via their coiled-coil domains (CCDs)[16-18]. Defective BECN1 or deficient expression levels are linked to various diseases, including neurodegenerative disorders[19-21], cancers[4, 22-25], and infectious diseases[10, 26-28].

BECN1 upregulates PI3KC3 activity and autophagy in response to diverse signals, although the exact mechanisms by which BECN1 performs this function remain unclear. BECN1 appears to be a major interaction hub for autophagy, interacting with at least twenty different cellular proteins as well as proteins encoded by at least five viruses (HIV, Influenza virus, α-, β-, and ɣ-herpesviruses) that inhibit autophagy to prevent virus degradation[28-30]. While the mechanisms by which BECN1 binds so many diverse partners are unknown, conformational flexibility often facilitates multivalent protein interactions, allowing one protein to interact with various binding partners to carry out diverse functions. Thus, understanding BECN1 conformational flexibility is key to understanding varied BECN1 interactions, which in turn, will improve our understanding of the role of BECN1 in autophagy. Ultimately this is essential to better understand how eukaryotic organisms respond to and survive cellular stressors.

BECN1 is a conformationally-flexible protein that has at least four domains[31]: an N-terminal intrinsically disordered region (IDR, residues 1-140)[30, 32], a flexible helical domain (FHD, residues 141-171)[33], a coiled-coil domain (CCD, residues 175-265,)[18, 34], and a β-α repeated, autophagy-specific domain (BARAD, residues 266-450)[35, 36]. The IDR and FHD both contain “Anchor regions” (sequences flanking or overlapping IDRs that are predicted by the program ANCHOR to nucleate binding-associated folding[37]). Four of the BECN1 Anchor regions have been identified as α-molecular recognition features (αMoRFs), i.e. regions that undergo disorder-to-helix transitions upon binding to partners[38, 39]. The BH3D, comprising residues 105-130 of the IDR and encompassing an Anchor region, was the first BECN1 αMoRF identified[30]. The BH3D is disordered, but folds into a helix upon binding to various BCL2 homologs[30, 40] or in the presence of 2,2,2-trifluoroethanol (TFE), a chemical that induces helicity in αMoRFs even in the absence of their binding partners[41]. Similar methodology shows that residues 76-105[41], which comprise an IDR Anchor region, and the FHD[42], which encompasses two Anchor regions, are likely also αMoRFs. Even within the well folded CCD[18, 34] and BARAD[35, 36], residues 248-265 adopt

mutually exclusive helical conformations in different crystal structures and undergo conformational changes in response to membrane binding[17, 31]. Thus, conformational flexibility is an over-riding feature of BECN1 and many BECN1 regions are now known to undergo binding-associated conformational changes.

1.1.1. BCL2 binding to BECN1

BECN1 was initially discovered as a BCL2-interacting protein[10] and binding of the anti-apoptotic BCL2 proteins to BECN1 are now known to down regulate autophagy[26, 43, 44]. BECN1 residues 88-150 were named the BCL2-binding domain (BCL2BD), as this region was sufficient for binding BCL2 and BCL2L1 (Bcl-X_L), and its deletion weakened BECN1 interaction with BCL2 in mammalian cells[10]. Sequence analysis and structural studies demonstrated that the BECN1 BCL2BD contains a BH3D, comprising residues 105-130[43-47]. A minimal BH3D consists of a four-turn, amphipathic α-helix bearing the sequence motif: ϕ-X-X-X-ϕ-K/R-X-X-Sm-D/E-X-ϕ (where ϕ: hydrophobic residues; Sm: small residues, typically glycine)[48]. In BECN1, the conserved hydrophobic residues of the BH3D are L112, L116, and F123. The BH3D is conserved amongst pro-apoptotic proteins and enables binding of anti-apoptotic BCL2 proteins[48].

Diverse BCL2 homologs bind the isolated BECN1 BH3D with moderate affinity[30, 43, 44, 46]. Mutagenesis studies in combination with diverse binding assays show that the BH3D is both necessary and sufficient for BECN1 interaction with several anti-apoptotic BCL2 proteins. Structures of the BECN1 BH3D bound to human Bcl-2[30], human Bcl-X_L[46, 47], or the ɣ-herpesvirus (HV) 68 BCL2 homolog M11[43, 44], show that, like BH3Ds from pro-apoptotic proteins, the BECN1 BH3D binds as a helix to a hydrophobic surface groove on the BCL2 proteins. In each structure, the interface buries conserved BECN1 BH3D hydrophobic residues: L112, L116, and F123. The BECN1 G120-D121 pair, which is highly conserved in all BH3Ds, interacts with a G-R pair (G86-R87 of M11 and G138-R139 of Bcl-X_L) that is highly conserved in anti-apoptotic BCL2 homologs. Thus, anti-apoptotic BCL2 proteins bind the BECN1 BH3D in a manner similar to the other BH3Ds.

Despite the overall similarity in the mode of binding of different BCL2 homologs to BECN1, there are significant differences. Variations in residues lining the hydrophobic groove of each BCL2 protein dictate a differential binding affinity for BECN1. Thus, the BECN1 BH3D binds to M11 with 1.1 μM affinity,

and to human Bcl-2 and Kaposi's Sarcoma HV (KSHV) Bcl-2 with weaker affinities of 8.0 μM and 13.3 μM, respectively[44]. Two mechanisms regulate BCL2-mediated inhibition of BECN1-dependent autophagy: competitive binding of other BH3D-containing proteins[45] and modulation of BECN1:BCL2 interaction by phosphorylation of either partner[49-51].

The mechanism by which BCL2 proteins bind the BECN1 BH3D is well established; however, the mechanism by which this interaction down regulates BECN1-mediated autophagy is not. BCL2 overexpression diminishes BECN1 co-immunoprecipitation with VPS34[26]. Further, in the presence of Bcl-X_L or KSHV Bcl-2, BECN1 binds UVRAG with 4-fold lower affinity[52]. Isothermal titration calorimetry (ITC) experiments indicate Bcl-X_L and KSHV Bcl-2 bind the BECN1 homodimer[52]. Therefore, BCL2 binding may stabilize the BECN1 homodimer and disrupt interactions between BECN1 and other components of the autophagy nucleation complex.

1.1.2. BECN1 in PI3KC3 complexes

BECN1 is a core component of PI3KC3/VPS34 complexes involved in autophagy and vesicle trafficking[11, 53]. BECN1 associates with either ATG14/BARKOR or UVRAG/VPS38 to form two mutually exclusive PI3KC3 complexes: i) complex I comprising BECN1:ATG14:PI3KC3:p150 or ii) complex II comprising BECN1:UVRAG:PI3KC3:p150[12, 14, 15]. While p150/PI3KR4/VPS15 is an obligate partner of PIK3C3, association of the other proteins up-regulates PI3KC3 activity to convert phosphatidylinositol (PI) to phosphatidylinositol 3-phosphate (PI3P), an essential signal for autophagosome formation[54-56]. Complex I nucleates autophagosomes while complex II mediates autophagosome maturation[11, 12, 14, 15, 57].

PI3KC3/VPS34 contains a N-terminal C2 domain, a central helical domain, and C-terminal lipid kinase domain[58, 59]. VPS34 forms a heterodimer with N-terminally myristoylated VPS15[11, 60] (p150 in humans), a protein Ser/Thr kinase, which consists of a N-terminal kinase domain, a central HEAT (Huntington, elongation factor 3 protein phosphatase 2A, yeast kinase TOR1) domain, and a C-terminal WD40 domain[58]. The VPS34 C-terminal lipid kinase domain interacts with the VPS15 N-terminal kinase domain[17]. Human ATG14 contains a N-terminal cysteine-rich repeat region required for targeting ATG14 to the endoplasmic reticulum[61], a central CCD, and a C-terminal BATS (Barkor/Atg14 autophagosome targeting sequence) functional domain responsible for sensing and preferentially binding highly curved,

PI3P-rich membranes[62]. Human UVRAG consists of a proline-rich sequence, a C2 domain, a CCD and a C-terminal domain[53].

The 28 Å CryoEM reconstructions of complex I and II reveal nearly identical overall V-shaped envelopes[16]. PI3KC3/VPS34 and p150/VPS15 comprise one arm of each V-envelope, while the other arm is comprised of BECN1 and either ATG14 or UVRAG. MBP-tag mapping of the CryoEM reconstructions reveals that the BECN1 BARAD is located at the tip of one arm of the V, with BECN1 in a parallel arrangement with either ATG14 or UVRAG, extending to the N-termini of each protein located at the base of the V. Further, the VPS15 HEAT domain is positioned to interact with N-terminal regions of BECN1 and ATG14 or UVRAG, at the base of the V-envelope, while the VPS15 WD40 domain is positioned to interact with C-terminal regions of BECN1 and ATG14 or UVRAG. A recent 4.4 Å crystal structure of PI3KC3 complex II[17] confirms this overall architecture (Figure 1.2).

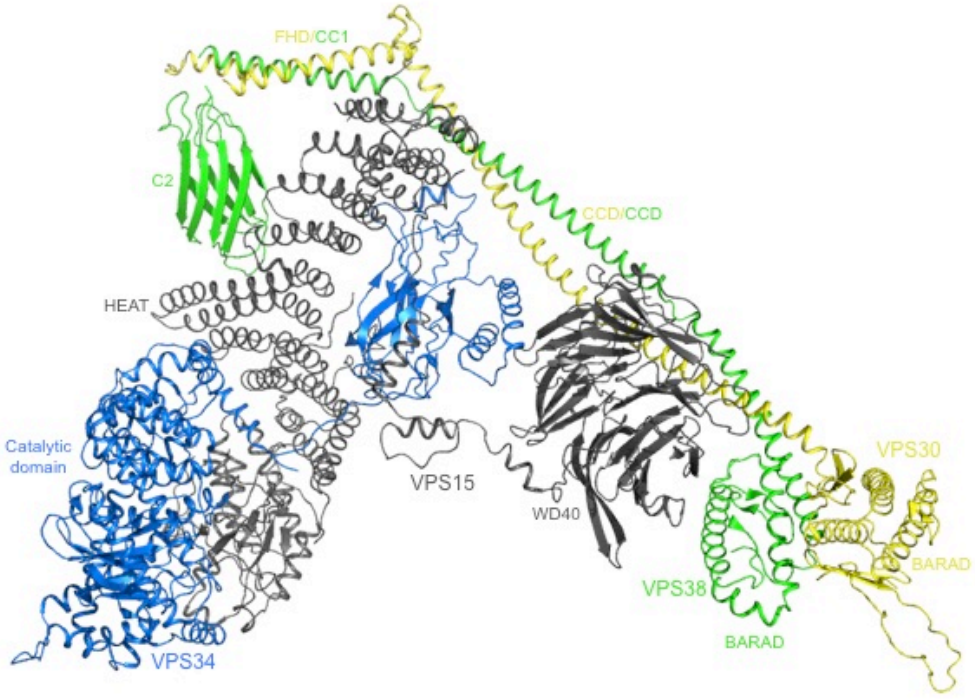

Figure 1.2. Domain architecture of the yeast PI3KC3 complex II.
All proteins are shown in ribbon, colored as VPS30, yellow; VPS38, green; VPS15, gray; and VPS34, blue. Protein domains are labeled.

1.2. Viral evasion and subversion of autophagy

Autophagy not only recycles defective or unnecessary macromolecules generated by the cell, but also contributes to innate and adaptive immunity by sequestering and degrading intracellular pathogens via a subtype of autophagy also referred to as xenophagy[63]. Enforced autophagy induction protects

against Sindbis virus, tobacco mosaic virus, and vesicular stomatitis virus infections[64, 65]. However, as obligate intracellular parasites, many viruses have evolved mechanisms to inhibit autophagic progression to evade degradation and even mechanisms to subvert of the autophagic machinery to facilitate viral replication and morphogenesis, thereby manipulating autophagy to serve as a pro-viral pathway[64]. While several viral proteins are implicated in interacting with the autophagy machinery[63, 64, 66-68], few interactions have been characterized.

Some viruses target BECN1 or BECN1-containing complexes to evade autophagic degradation, including the HIV[68, 69], influenza A virus[70], African swine fever virus[67], foot and mouth disease virus[71], αHV herpes simplex virus type 1 (HSV-1)[27], βHV human cytomegalovirus (HCMV)[72], and γHVs KSHV and γHV68[26, 43, 44], although the mechanisms of xenophagic evasion and subversion by these viruses are poorly understood. Elucidation of the structural features of these interactions will provide basic information required for the rational design of therapeutics to selectively inhibit interactions between viral proteins and the highly conserved autophagy machinery, allowing xenophagic degradation to proceed and clear virus from infected cells. Importantly, targeting these interactions could form the basis of therapeutic interventions that defy the development of viral resistance. The highly conserved nature of the autophagy machinery exerts evolutionary pressure on viruses to prevent mutations that abrogate autophagic down-regulation, since the ability to interact with and manipulate the autophagy machinery would be lost as well. However, it will be necessary to characterize the nature of these various interactions to design therapeutics that specifically inhibit viral evasion or subversion of xenophagy.

1.3. Overview of methods used in this research

1.3.1. Cellular autophagy assays

Cellular autophagy assays are used to investigate the impact of mutations or protein-protein interactions on autophagy induction. Since BECN1 is known to be required for autophagosome nucleation, cellular autophagy is evaluated by monitoring and comparing levels of puncta labeled with green fluorescent protein (GFP)-tagged light chain 3 (LC3), an autophagosome-specific marker, in cells grown in either nutrient-rich or starvation media. GFP-tagged LC3-II is commonly used as a marker of autophagosome formation in mammalian cells because the lipidated LC3-II localizes to autophagosome membranes, resulting in the formation of discreet green fluorescent puncta, while the non-lipidated

precursor (soluble LC3-I) generates a diffuse green fluorescence throughout the cytoplasm[73]. Therefore, GFP-fluorescent puncta can be observed under a fluorescent microscope and quantified. We use human breast adenocarcinoma MCF7 cells for all autophagy assays presented in this dissertation as they lack detectable endogenous expression of BECN1, resulting in very low basal levels of autophagy, even in starvation conditions unless BECN1 is ectopically expressed[4, 26, 44]. This allows the effect of BECN1 mutants and proteins that interact with BECN1 to be assayed in the absence of endogenous BECN1.

1.3.2. Isothermal titration calorimetry (ITC)

ITC, a useful technique for measuring the thermodynamics of binding reactions, directly monitors the heats of binding associated with chemical interactions. Protein-protein binding and dissociation events are associated with heat energy exchange with the environment, which makes ITC a convenient method for quantifying protein interaction thermodynamics since it does not require labeled protein or protein immobilization and requires small quantities of protein[74]. The ITC experiments described in this dissertation were performed on a Low Volume Nano ITC from TA instruments, which can detect heats from 0.05-5,000 μJ and determine dissociation constants from 10^{-9}-10^{-4} M[74]. The calorimeter contains two identical cells, a sample cell and reference cell, which are maintained at a constant temperature (Figure 1.3A). Titration of a molecule into the sample cell causes heat release or absorption upon binding and dissociation. In response, the calorimeter heats or cools the sample cell to maintain the sample and reference cells at the same temperature, resulting in a titration curve of the heater power as a function of time (Figure 1.3B)[75].

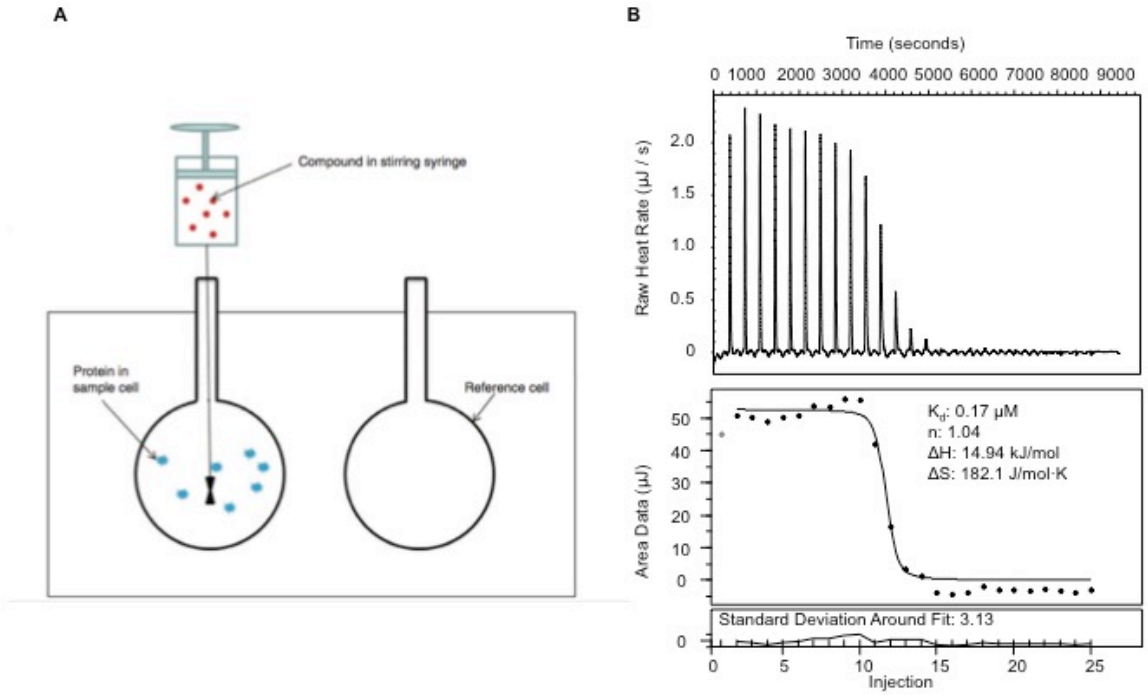

Figure 1.3. ITC experiments.
A) Experimental setup[75] and B) the raw titration data and data analysis for BID BH3D binding to BHRF1.

Proper sample preparation is imperative for reliable ITC results. Proteins should be highly pure to obtain good data since small impurities can affect the heats detected and protein concentration measurements should be highly reliable in order to calculate dependable thermodynamics parameters upon fitting the data to the appropriate binding model. All samples should be properly dialyzed against at least 2 L of the selected ITC buffer to ensure that they are fully buffer-matched in order to guarantee that all detected heats are derived from binding or dissociation events rather than from dilution of mismatched buffers. For protein binding interactions, the protein loaded into the syringe is titrated into the protein loaded in the cell until the protein in the cell is saturated. For protein dissociation experiments, the dimer is loaded into the syringe and titrated into dialysis buffer until the dimer no longer dissociates. In both cases, the area of each peak of the titration curve is integrated and plotted versus the molar ratio of the molecule in the syringe to the molecule in the cell. The resultant binding curve can then be fitted to a binding model through the use of programs such as the NanoAnalyze Software (TA Instruments, New Castle, US) to determine the enthalpy of interaction (ΔH), association constant (K_a), and stoichiometry (Figure 1.3B). The ΔH corresponds to the intercept of two asymptotic lines corresponding to the minimal and maximal heats on the binding curve, the slope at the inflection point indicates the K_a ($K_a = 1/K_d$), and

the molar ratio at the inflection point identifies the stoichiometry of the binding reaction. The change in entropy (ΔS) and change in Gibbs free energy (ΔG) can be derived from these data and the reaction temperature[76].

1.3.3. Circular dichroism (CD) spectroscopy

All amino acids, except glycine, are optically active due to their chiral nature and; therefore, rotate the polarized light. CD spectroscopy is commonly used to evaluate the secondary structure content, folding state, and binding properties of optically active molecules due to the unequal absorption of left-handed and right-handed circularly polarized light[77]. CD data for protein are commonly reported as molar ellipticity per residue in degrees·cm^2·$dmol^{-1}$ or degrees·M^{-1}·m^{-1}, where degrees ellipticity is defined as the angle whose tangent is the ratio of the minor axis to the major axis of the ellipse[77].

CD spectra may be used to estimate the secondary structure of proteins. In the far-UV spectra (wavelengths below 250nm) the optical transitions that arise from the amides of the protein backbone depending on their geometric arrangement in different secondary structures result in characteristic CD spectra for different types of secondary structures (Figure 1.4). Proteins with all α-helical content have negative bands at 222 nm and 208 nm and a positive band at 193 nm, while proteins with all β-strand content have a negative band at 218 nm and a positive band at 195 nm, and completely disordered proteins have minimal ellipticity above 210 nm and a negative band near 195 nm[77]. The CD spectra of proteins containing a mixture of secondary structures reflect a combination of the spectra that arise from individual secondary structure components. Therefore, CD spectra can be analyzed with programs such as the CDPro software package in order to estimate the secondary structure content of proteins[78].

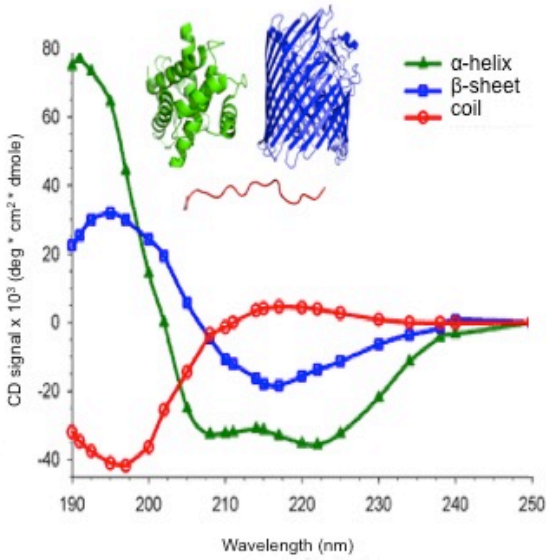

Figure 1.4. CD Spectra of polypeptides with representative secondary structures. Representative spectra for all α-helical (green), all β-strand (blue), and all coil (red) polypeptides are shown. This figure is adapted from the website: https://www.fbs.leeds.ac.uk/facilities/cd/.

CD spectroscopy is also a useful tool to assess the thermal stability of folded proteins. As folded proteins are heated, they begin to lose the characteristic CD spectra that arise from their α-helical or β-strand content, and begin to generate spectra that resemble disordered proteins[79]. Since disordered protein CD spectra lack ellipticity at 222 nm and 218 nm (Figure 1.4), the significant negative bands for α-helical and β-strand proteins respectively, thermal unfolding can be observed by monitoring the loss of molar ellipticity at these wavelengths as a function of temperature. If unfolding is reversible, the melting temperature (T_m) of a protein can be determined by fitting the melting curve with an appropriate Boltzmann algorithm to calculate the midpoint of the sigmoid. Even if unfolding is irreversible due to protein aggregation, information about relative stability can still be obtained[79].

1.3.4. Small angle X-ray scattering (SAXS)

SAXS permits low-resolution (50-10 Å) tertiary structure characterizations of macromolecules in solution, which are useful for understanding macromolecular folding, unfolding, extended conformations, flexibly linked domains, shape, conformation, and oligomerization[80]. Therefore, SAXS is an especially useful method for obtaining solution structure information about flexible proteins, which typically do not crystallize, and proteins or protein assemblies whose sizes are unsuitable for nuclear magnetic resonance (NMR) or electron microscopy studies.

In a SAXS experiment, purified protein in solution is exposed to the X-ray source and the scattering signal is derived from the difference in average electron density of protein molecules of interest and bulk solvent. Subtraction of the buffer scattering from sample scattering generates an isotropic scattering curve due to the random orientation of the sample tumbling in solution. The scattered intensity (I(q)) is a function of the magnitude of the scattering vector q. $q = (4\pi\sin\Theta)\lambda$, where 2Θ is the scattering angle and λ is the wavelength of the incident X-ray beam[80]. The scattering curves are typically plotted as log(I) versus q^2 (Figure 1.5). Protein homogeneity, radius of gyration (R_g), pairwise-distance distribution function (P(r)), maximum particle dimension (D_{max}), and folding state can be assessed by analysis of the scattering curve.

The Guinier plot (ln(I) versus q^2) is used to assess sample homogeneity and R_g. A linear Guinier plot in low q regions, where $q \times R_g \leq 1.3$ for globular proteins or $q \times R_g \leq 0.8$ for elongated proteins, is indicative of a homogenous, monodispersed sample while curvature of the Guinier plot in the low q regions is indicative of aggregation (Figure 1.5)[80]. The R_g is the average root-mean-square distance to the center of density from the surface of the molecule weighted by the scattering length and is calculated from the slope of the linear fit of the Guinier plot where slope = $-R_g^2/3$ [81]. Therefore, between two proteins of the same molecular weight, the more compact protein will have a smaller R_g than that of a more elongated protein.

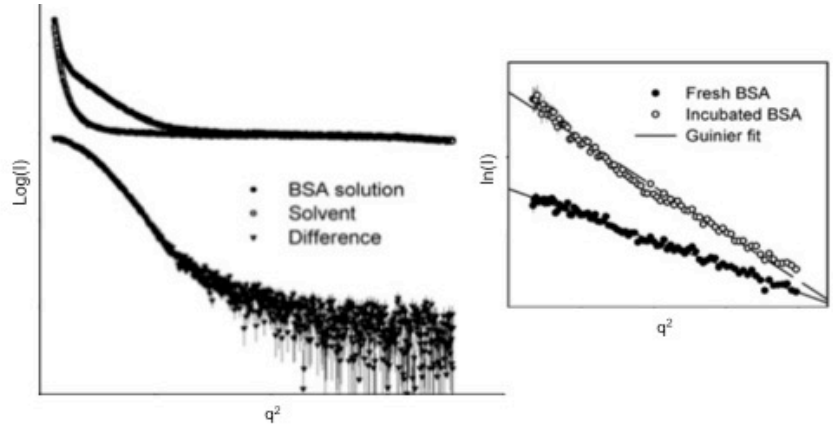

Figure 1.5. SAXS scattering data and Guinier plot.
Log(I) versus q^2 plot represents scattering from BSA in solvent, solvent, and the buffer subtracted scattering. The ln(I) versus q^2 plot represents fresh (homogenous) BSA and incubated (aggregated) BSA. This figure is adapted from[82].

The P(r), calculated through a Fourier transform of the scattering curve, provides the distribution of distances between electrons in the sample. The P(r) plot provides the D_{max} of the molecule, which is

the radius of the longest dimension of the protein and is defined by the point where the value of P(r) approaches zero. The P(r) plot provides the D_{max} of the molecule, which is the radius of the longest dimension of the protein and is defined by the point where the value of P(r) approaches zero. (Figure 1.6), with globular proteins displaying a bell-curve with a maximum at $D_{max}/2$, elongated proteins displaying an asymmetric curve with a maximum at small distances corresponding to the radii in the short dimensions, and dumbbell-shaped proteins consisting of two well-separated globular domains displaying two maxima[82].

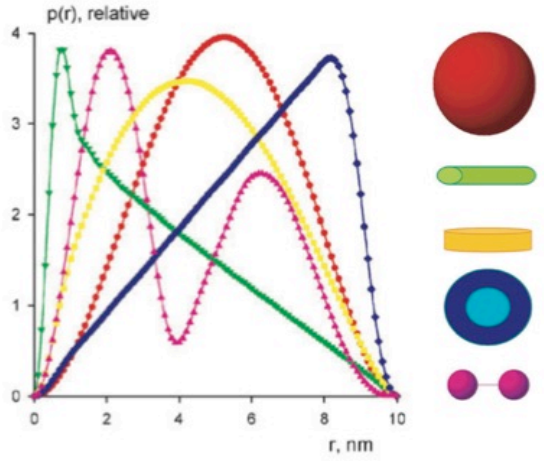

Figure 1.6. P(r) curves correlated with protein shape.
Representative curves for globular (red), elongated (green), disk shaped (yellow), hollow sphere (blue), and dumbbell-shaped (magenta) molecules. This figure is adapted from[82].

The Kratky plot ($I(q) \times q^2$ versus q) permits a qualitative assessment of protein flexibility (Figure 1.7). A bell-shaped curve with a maximum at low q and convergence at high q indicates a compact molecule, a bell-shaped curve with a clear maximum at low q but incomplete convergence at high q indicates a partially folded molecule, and a hyperbolic or plateau shaped curve indicates a disordered molecule. Although the Kratky is good at identifying unfolded proteins, it cannot clearly distinguish fully folded proteins and partially unfolded proteins with significant structured regions. The dimensionless Kratky plot, wherein the intensity I(q) is normalized to the forward scattering intensity I(0) and q is normalized to the R_g, plots $(q/R_g)^2 I(q)/I(0)$ versus q/R_g and permits comparison of globular and extended proteins regardless of size to permit semi-quantitative assessments of flexibility[83]. The normalized Kratky maxima for a globular protein is 1.1 at $q/R_g = \sqrt{3}$ while the maxima for increasingly unfolded proteins increases as disorder increases and also shifts towards higher q/R_g values.

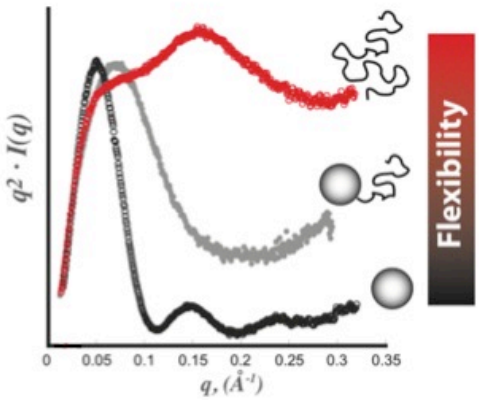

Figure 1.7. Sample Kratky plot corresponding to different folding states.
Representative curves for disordered (red), partially folded (gray), and globular (black) molecules. This figure is adapted from[84].

The scattering profiles can also be used to reconstruct *ab initio* molecular envelopes of proteins. These low resolution three-dimensional structures are generated via applications of programs such as DAMMIF[85] with subsequent analysis by DAMSEL, DAMSUP, DAMAVER and DAMFILT[86] wherein a bead model yielding a scattering pattern that fits the experimental data is generated from an random initial model followed by several rounds of simulated annealing. Proteins with known atomic structures can be fitted into these molecular envelopes either manually or automatically through programs such as SUPCOMB[87] to obtain information about the oligomerization state or domain arrangement of multi-domain proteins. For intrinsically disordered proteins or multi-domain proteins with flexible linker regions, the Ensemble Optimization Method (EOM) can be used to determine an ensemble of conformations from a random pool of conformations whose averaged theoretical scattering best describes the experimental SAXS data in order to generate a model that includes flexible or disordered regions.

1.3.5. Protein X-ray crystallography

Protein X-ray crystallography permits atomic resolution structural characterizations of macromolecules. The function of a protein and the mechanism of protein-protein or and protein-small molecule interactions can often be deduced from the atomic resolution three dimensional structures determined by X-ray crystallography. Therefore, this technique is a very useful tool for probing the structure-function relationship of proteins. However, X-ray crystallography only provides structural information about the low-energy conformation trapped during crystallization; therefore, this technique is

limited to well-ordered proteins as flexibility typically prevents acquisition of crystals that diffract to atomic resolution.

Protein crystallization requires a highly pure and homogenous protein sample. Crystallization is induced by bringing the protein solution to supersaturation, a thermodynamically unstable state that induces a portion of the protein to come out of solution either in the form of crystal nuclei that grow into larger crystals or precipitate. Protein concentration, precipitant type, precipitant concentration, buffer type, buffer pH, temperature, crystallization technique, ionic strength, and additives that promote crystallization can all affect the crystallization process[88]. Therefore, determining the appropriate conditions to promote growth of high quality protein crystals may require several rounds of optimization. Vapor diffusion is the most widely used crystallization technique, wherein a drop composed of a mixture of protein sample and reagent (mixture of precipitant, buffer, salt, crystallization additive, etc.) are placed in vapor equilibrium with a reservoir of reagent. Typically, the drop contains a lower reagent concentration than the reservoir so that, in order to attain equilibrium, water vapor evaporates from the drop, thereby causing the drop to become more concentrated in protein and reagent[88]. The increased concentrations in the drop may induce the state of supersaturation that promotes crystal nucleation and growth.

X-ray diffraction data is recorded by exposing the protein crystal to a monochromatic X-ray beam generated from accelerating electrons in a synchrotron storage ring or from electrons striking a copper anode. Synchrotron sources typically facilitate the attainment of a higher signal-to-noise ratio of the diffraction image and shorter exposure times due to their more intense and better focused X-ray beams and are therefore preferred for obtaining high-resolution diffraction data[89]. Diffraction data is usually collected from frozen crystals at 100 K to facilitate the collection of a complete data set from a single crystal by mitigating radiation damage[90]. X-ray waves are scattered by the electrons within the crystal resulting in a diffraction pattern that register as spots of different intensities, called reflections, on the detector[88].

Diffraction data processing permits determination of the space group, unit cell parameters, and crystal orientation in the beam[88]. Once the cell and orientation information is known, the reflections can be indexed and assigned their Miller indices, which consist of three integers: *h, k, and l*, to determine if a complete data set can be obtained[91]. The next stage of data processing involves measurement of the

intensity of the reflections. A full diffraction data set typically consists of many diffraction images, therefore equivalent reflections from each image must be merged and scaled so that the intensities of all of the images can be related, permitting all symmetry-related reflections to be merged into a unique set of reflections[88]. The Pearson correlation coefficient between random half-data sets ($CC_{1/2}$) is used to assess data quality and determine the high-resolution cutoff for data processing[92]. $CC_{1/2}$ is related to the signal-to-noise ratio of the data and values of 0.5 and above are considered acceptable.

The intensity of each reflection is a result of both the amplitude of the diffracted waves and the phase difference, expressed as an angle, between the diffracted waves. Both the amplitudes and phases of the diffracted X-rays are required for calculation of the electron density map. The electron density $\rho(x,y,z)$ is calculated using a Fourier transformation[93]: $\rho(x,y,z) = (1/Vz)\sum_h\sum_k\sum_l F_{(h,k,l)}\exp[-2\pi i(hx + ky + lz)]$, where V is the volume of the unit cell; $F_{(h,k,l)}$ is the structure factor for the Bragg reflection with Miller indices *h*, *k*, and *l*; x,y, and z are the fractional coordinates of the unit cell, and the triple summation is over all values of the Miller indices. $F_{(h,k,l),}$ the structure factor, is defined by the amplitude and phase of the reflection. While the amplitude can be calculated from the measured intensity, the phases cannot be determined directly.

Several methods exist for phase determination, including multiple isomorphous replacement, multi-wavelength anomalous dispersion, single-wavelength anomalous dispersion, and molecular replacement (MR)[94]. All structures presented in this dissertation were solved by MR. MR utilizes the phases of the structure factors from a known structure of a homologous protein to estimate the phases of the unknown structure[95].

Once the amplitudes and phases of the reflections are identified, new structure factors can be calculated and used to calculate the electron density map into which the new protein structure can be built. The newly built structure model must then be refined in order to improve its agreement with the electron density map. The *PHENIX*[96] software suite can be used to complete the MR, model building, and refinement steps of protein crystal structure solution. The R factor is a useful quality indicator to evaluate the agreement between an electron density map and a model[97]. R factor = $\sum_{hkl}||F_{obs}| - |F_{calc}||/\sum_{hkl}|F_{obs}|$, where $|F_{obs}(hkl)|$ are the measured amplitudes and $|F_{calc}(hkl)|$ are the amplitudes measured from the model. In order to permit validation of the final structure, a percentage of reflections are omitted during

the modeling and refinement process[98]. The R_{free} refers to how well the observed structure factors from the set of reflections omitted during the modeling and refinement process agree with the structure factors calculated from the model. The R_{work} (calculated from the reflections used during modeling and refinement) should be within about 5% of the R_{free}.

1.4. Specific aims of this research

Given that a growing number of autophagy proteins are recognized to bear IDRs and flexible ordered regions, it is important to understand how conformational flexibility affects protein function. Therefore, the goal of this dissertation is to identify and understand conformational and binding-induced disorder-to-order transitions in proteins that regulate autophagy. The specific aims of this research are:

1. To devise a diagnostic method to identify IDRs likely to become helical upon binding, even in the absence of known binding partners.
2. To investigate conformational flexibility within BECN1, which contains two structured domains sharing an overlapping region that has been crystallized in two mutually exclusive states.
3. To investigate the interaction between BECN1 and Nef, a protein that is critical for HIV pathogenesis and modulation of autophagy.
4. To investigate the binding of the anti-apoptotic KSHV vBcl-2 to the BH3Ds of the NOXA, a pro-apoptotic BCL2, and BECN1.
5. To determine the role in autophagy regulation by BHRF1, a vBCL2 encoded by EBV, and characterize its binding to BH3Ds.

CHAPTER 2. IDENTIFYING INTRINSICALLY DISORDERED PROTEIN REGIONS LIKELY TO UNDERGO BINDING-INDUCED HELICAL TRANSITIONS

2.1. Introduction

The canonical view that protein structure dictates function requires a more nuanced understanding in light of the growing number of proteins recognized to bear IDRs. IDRs lack the ability to form stable secondary structures such as α-helices and β-strands, and well-packed tertiary structures[99, 100]. IDRs tend to have sequences of lower complexity and an increased number of polar or charged residues rather than hydrophobic residues likely to pack in the interior of a protein[99, 101]. IDRs also tend to have low sequence conservation[30, 102]. They appear to be especially prevalent in eukaryotes, where at least 35-50% of proteins are predicted to bear IDRs[103, 104]. This prevalence implies that IDRs must confer some important evolutionary advantage in advanced organisms.

IDRs are increasingly recognized to mediate many functions essential for life. They appear to regulate protein lifetime by providing targets for proteolysis[105, 106]. Perhaps their most common function is to mediate macromolecular interactions, including protein-protein, protein-DNA, and protein-RNA interactions[100]. The low sequence conservation and conformational flexibility inherent in IDRs may provide IDRs with the evolutionary flexibility required to facilitate multiple protein-protein interactions[100, 104]. Consequently, IDRs are common among interaction hub proteins (regulatory and signaling proteins that bind a variety of diverse partners)[107]. IDRs often also serve as sites of phosphorylation, and in turn, phosphorylation appears to modulate interactions with other macromolecules[108-110]. Therefore, it is important to understand the structure-function relationship of IDRs and IDR-mediated binding.

Often, binding to appropriate partners appears to stabilize secondary structure elements within IDRs[99, 111], although it has recently been shown that up to one third of all residues are disordered in protein complexes[112]. Regions that have experimentally been shown to undergo disorder-to-order transitions have been called Molecular Recognition Features (MoRFs)[38, 39]. MoRFs have been further classified as α-MoRFs, β-MoRFs or σ-MoRFs, depending on whether they fold into α-helices, β-strands or stable coils, respectively. Understanding the structures of these regions in both unbound and bound states is important for understanding the mechanism of biological function of these regions, and ultimately for rational design of therapeutics targeting these interactions[111]. However, binding partners remain

unknown for most IDRs, complicating experimental elucidation of structural transitions within IDRs. We hypothesized that it would be possible to identify α-MoRFs by bioinformatics analyses using programs such as ANCHOR[113] and DisoRDPbind[114] combined with experimental evaluation of chemically-induced disorder-to-helix transitions.

Therefore, in this study our goal was to develop a method to identify α-MoRFs in the absence of known binding partners. We first performed a bioinformatics analysis using the programs IsUnstruct[115], PONDR-FIT[116] and IUPred[117] to predict IDRs, then used the programs ANCHOR[113] and DisoRDPbind[114] to identify IDRs that are likely to be stabilized in ordered structures upon binding. JPred4[118] was used to predict potential secondary structure. We used CD spectroscopy to monitor change in secondary structure content of diverse IDRs induced by 2,2,2-trifluoroethanol (TFE), which is known to stabilize α-helical structure in disordered amino acid sequences that have some native tendency to form helices[119-123]. Our results demonstrate that IDRs that are not predicted by ANCHOR to fold upon binding do not become helical in the presence of TFE. IDRs that are predicted to fold fall into two groups. One group becomes more helical upon addition of TFE and includes all the known α-MoRFs that were included in our study. The other group does not become helical and includes examples of β-MoRFs. Therefore, we propose that predictions using programs such as ANCHOR combined with experimental evaluation using TFE may provide a general method to identify IDRs that undergo a binding-induced disorder-to-helix transition.

2.2. Materials and methods

2.2.1. Production of IDRs

Peptides corresponding to the selected IDRs (Table 2.1) were chemically synthesized, HPLC purified to >90% purity, and purity confirmed by electrospray mass spectrometry (RS Synthesis or EZBioLabs). For each peptide, a 1 mM stock solution in 10 mM phosphate pH 7.4, 100 mM ammonium sulfate was prepared.

Table 2.1. IDR sequences used in TFE study.

Protein	Accession #	Peptide Sequence
TRAP220	Q925J9	^{637}GNTKNHPMLMNLLKDNPA^{654}
BECN1	Q14457	^{105}DGGTMENLSRRLKVTGDLFDIMSGQT^{130}
GRIM	Q24570	^{1}MAIAYFIPDQAQLLA^{15}
p53	P04637	^{367}SHLKSKKGQSTSRHKKLMFKTE^{388}
GOPC	Q9HD26	^{248}RHKTVIRACRGRNDLKRPMQAPPGHDQDSL^{277}
ATG3	Q9NT62	^{129}ENKDNIRLQDCSALCEEEEDEDEGEAADME^{158}
ATG9A	Q7Z3C6	^{810}AEDGQSASRHPEPVPEEGSEDELPPQVHKV^{839}
BECN1	Q14457	^{76}DGVSRRFIPPARMMSTESANSFTLIGEASD^{105}
AMBRA1	Q9C0C7	^{527}TQQAQEMLNNNIESE^{541}
GOPC	Q9HD26	^{255}ACRGRNDLKRPMQAPPGHDQDSLKK^{279}
BECN1	Q14457	^{50}AQAKPGETQEEETNSGEEPFIETPRQDGV^{78}
ATG16L1	Q676U5	^{59}EKHDVPNRHEISPGHDGTWNDNQLQE^{84}
ATG2A	Q7Z3C6	^{1244}HPPPRPPSPTEIAGQKLSESPASLPSCPPV^{1273}

Table note. Accession # corresponds to the full-length amino acid sequence downloaded from NCBI RefSeq protein database (http://www.ncbi.nlm.nih.gov/refseq).

2.2.2. Prediction of disorder and likelihood of folding upon binding

The amino acid sequences of the full-length proteins containing the thirteen selected IDRs were downloaded from the NCBI RefSeq protein database (http://www.ncbi.nlm.nih.gov/refseq) (Table 2.1). These sequences were analyzed with IsUnstruct (http://bioinfo.protres.ru/IsUnstruct/)[115], PONDR-FIT (http://www.disprot.org/metapredictor.php)[116], and IUPred (http://iupred.enzim.hu)[117] to predict whether the selected peptide sequences were intrinsically disordered (Table 2.2). The sequences were also analyzed with ANCHOR (http://anchor.enzim.hu)[113] and DisoRDPbind (http://biomine-ws.ece.ualberta.ca/DisoRDPbind/index.php)[114] to predict whether the selected IDRs were likely to undergo binding-associated structure stabilization (Table 2.2). Finally, these sequences were analyzed with Jpred4 (http://www.compbio.dundee.ac.uk/jpred/)[118] to predict secondary structure.

2.2.3. CD spectroscopy

Peptides were diluted to 25-100 μM in 10 mM phosphate pH 7.4, 100 mM ammonium sulfate and either 0%, 25%, or 40% v/v TFE. After a 1 hour incubation on ice, continuous scanning CD spectra were recorded from 190 to 240 nm at 4 °C in a 300 μL quartz cell (0.1 cm path length) on a Jasco J-815 spectropolarimeter equipped with a Peltier thermoelectric temperature control. Secondary structure content was estimated by analyzing the CD data using three analysis programs, SELCON3, CDSSTR,

and CONTIN, from the CDpro suite within the Jasco software[124-126]. For each single peptide set (i.e. 0%, 25% and 40% TFE), a consistent reference protein database was used for SELCON3, CDSSTR, and CONTIN estimations. For each spectrum, the average secondary structure was calculated by averaging estimated secondary structure content obtained from the three CD data analysis programs (Tables 2.3-2.5).

2.3. Results

2.3.1. Selection of IDRs for analysis

We selected thirteen diverse IDRs for examination (Table 2.1), focusing on IDRs or fragments of IDRs that were no longer than 30 residues, to ensure that we did not include more than one Anchor region within the IDR being examined. IDRs that have experimentally been shown to be either α-MoRFs or β-MoRFs with known binding partners were identified by a literature search and served as our control set (Table 2.2). The known α-MoRFs selected included TRAP220 residues 637-654 (PDB 1XDK) and BECN1 residues 105-130 (PDB 3DVU)[30, 44, 127]. A β-MoRF comprising GRIM residues 1-15 (PDB 1SE0, 1JD5) was used as a negative control for the analysis[128, 129]. We also included p53 residues 367-388, which can form either a α-helix (PDB 1DT7) or a β-strand (PDB 1MA3) depending on the binding partner[130, 131].

IDRs that have not been experimentally examined were selected from our prior bioinformatics analyses wherein we identified consensus IDRs, defined as regions of at least 25 consecutive residues predicted to be disordered by at least three of four disorder prediction programs[104]. The experimentally unverified IDRs selected included GOPC residues 248-277, ATG3 residues 129-158, ATG9A residues 810-839, BECN1 residues 76-105, AMBRA1 residues 527-541, GOPC residues 255-279, BECN1 residues 50-78, ATG16L1 residues 59-84, and ATG2A residues 1244-1273 (Table 2.1).

We assessed disorder in all selected sequences by analyzing the full-length protein sequences using the programs IsUnstruct[115], PONDR-FIT[116] and IUPred[117]. IsUnstruct estimates the energy of each residue in either the ordered or disordered state to predict whether each residue will be ordered or disordered based upon amino acid sequence[115]. Our IsUnstruct analyses (Table 2.2) indicate that all selected peptides are completely disordered except short stretches comprising GRIM residues 4-8, BECN1 residues 115-130, ATG3 residues 129-136, and GOPC residues 255-258.

Table 2.2. Predictions of disorder and binding regions.

Protein name (residues)	Residues predicted to be intrinsically disordered			Predicted binding region of IDR		Predicted secondary structure (Jpred4)
	IsUnstruct	PONDR-FIT	IUPRED	Anchor	DisoRDPbind (predicted binding partner)	
Known MoRFs; contain Anchor regions						
TRAP220 (637-654)	637-654	637-654	637-654	637-654	None	α: 643-648
BECN1 (105-130)	105-114	105-109	105, 108-112, 124, 128	117-127	None	α: 111-124
GRIM (1-15)	1-3, 9-15	1-15	None	1-15	1-15 (RNA) 4-15 (protein)	α: 10-15
p53 (367-388)	367-388	None	367-388	367-388	360-383 (DNA)	None
IDRs with unidentified binding partners; contain Anchor regions						
GOPC (248-277)	258-277	248-277	255-277	248-257	269, 272-277 (DNA)	β: 251-254
ATG3 (129-158)	137-158	129-158	139-158	136-143	139-158 (protein)	None
ATG9A (810-839)	810-839	810-839	810-839	831-839	819, 821, 833-839 (protein)	None
BECN1 (76-105)	76-105	76-105	76-105	79-103	60-84 (protein)	None
AMBRA1 (527-541)	527-541	527-541	527-541	530-538	527-541 (protein)	α: 527-535
IDRs with unidentified binding partners; lack Anchor regions						
GOPC (255-279)	258-279	255-279	255-279	None	269, 272-279 (DNA)	None
BECN1 (50-78)	50-78	50-78	50-78	None	59-78 (protein)	None
ATG16L1 (59-84)	59-84	None	59-84	None	None	α: 65-66, 80-84
ATG2A (1244-1273)	1244-1273	1244-1273	1244-1273	None	1257-1260 (protein)	None

PONDR-FIT predicts per-residue disorder based on a combination of multiple features such as the ratio of net charge and hydrophobicity, amino acid frequency in disordered regions, and sequence complexity; with accuracy of prediction increasing with the length of the predicted disordered region[116].

Our PONDR-FIT analyses (Table 2.2) indicate that all peptides we selected are completely disordered except for p53 residues 367-388, ATG16L1 residues 59-84, and a stretch comprising BECN1 residues 110-130.

IUPred, a program which is also incorporated into the PONDR-FIT suite, calculates a disorder probability for each residue by using the amino acid composition to estimate formation of energetically favorable interactions, thereby identifying residues that lack such interactions and do not contribute to a stable folded structure[132]. The long-disorder algorithm, a context-independent prediction that takes 100 flanking residues into consideration[117], predicts all selected peptides to be completely disordered, except for GRIM residues 1-15 and shorter stretches comprising BECN1 residues 113-123, GOPC residues 248-254, and ATG3 residues 129-138 (Table 2.2).

Thus, except for stretches of BECN1 residues 105-130, all of the protein regions included in this study are predicted to be disordered by either IsUnstruct, PONDR-FIT, or IUPred; and most of them are predicted to be disordered by at least two programs. However, since BECN1 residues 105-130 have previously been experimentally shown to be disordered even within larger protein fragments[30], we included it in our study.

2.3.2. Prediction of IDRs likely to fold upon binding

Next we used the programs ANCHOR[113], DisoRDPbind[114] and JPred4[118] to analyze full-length protein sequences to predict the ability of the selected IDRs to form folded structures (Table 2.2). The ANCHOR program analyzes amino acid sequence to predict short regions (that we term Anchor regions) within or flanking IDRs that likely stabilize the IDR in an ordered state upon binding. Of the selected IDRs, Anchor regions are predicted (Table 2.2) in all the known α-MoRFs and β-MoRFs: TRAP220 residues 637-654, BECN1 residues 105-130, GRIM residues 1-15 and p53 residues 367-388; as well as in five experimentally untested IDRs: GOPC residues 248-277, ATG3 residues 129-158, ATG9A residues 810-839, BECN1 residues 76-105 and AMBRA1 residues 527-541 (Table 2.2). Lastly, four experimentally untested IDRs were predicted to not contain Anchor regions: GOPC residues 255-279, BECN1 residues 50-78, ATG16L1 residues 59-84, and ATG2A residues 1244-1273. Although GOPC residues 255-279 contain three Anchor residues (255-257) (Table 2.2), this is insufficient to be considered an Anchor

region, which must contain at least five residues[113]. Therefore, the IDR comprising GOPC residues 255-279 was included in the experimental set lacking Anchor regions.

DisoRDPbind predicts RNA, DNA, and protein binding regions within IDRs based on amino acid content, sequence complexity, predicted structure and disorder, and sequence alignments[114]. Of the known MoRFs that we selected, all are predicted to contain Anchor regions; however, only GRIM is predicted by DisoRDPbind to contain a protein-binding site (Table 2.2). Among the five experimentally unverified IDRs bearing Anchor regions (GOPC residues 248-277, ATG3 residues 129-158, ATG9A residues 810-839, BECN1 residues 76-105, and AMBRA1 residues 527-541), only GOPC residues 248-277 are not predicted by DisoRDPbind to contain a protein-binding region. Conversely however, two of the IDRs lacking Anchor regions, BECN1 residues 50-78 and ATG2A residues 1244-1273, are predicted by DisoRDPbind to contain protein-binding sites. Thus, the binding regions predicted by ANCHOR and DisoRDPbind are not consistent.

Full-length protein sequences were also analyzed by JPred4[118] to predict secondary structure (Table 2.2). JPred4 analysis (Table 2.2) indicates some secondary structure in seven of the IDRs, however; only five of these predicted secondary structures were of sufficient length for the formation of stable α-helices or β-strands.

Three of these IDRs, BECN1 residues 105-130, TRAP220 residues 637-654 and GRIM residues 1-15 are known MoRFs, with known binding partners, and were predicted by JPred4 to contain at least six contiguous α-helical residues (Table 2.2). While TRAP220 residues 637-654 and BECN1 residues 105-130 are α-MoRFs, GRIM residues 1-15 form a β-MoRF, with residues 3-6 forming a β-strand that is hydrogen-bonded to one end of a β-sheet from its binding partner DIAP1 (Drosophila melanogaster inhibitor of apoptosis protein)[128, 129]. Only the prediction for TRAP220 (Table 2.2) is entirely consistent with available experimental data, as TRAP220 residues 643-648 have been shown to be become helical upon binding[127]. While Jpred4 only predicts that BECN1 residues 111-124 are helical (Table 2.2), BECN1 residues 108-126 have been shown to become helical upon binding[30, 44, 46], thus; JPred4 does not predict the complete extent of the helical structure adopted by this IDR upon binding. Finally, for p53, residues 383-384 have been shown to form a β-strand with two residues of the Sir2 enzyme[130], and residues 378-384 have been shown to become helical upon binding Ca^{2+}-bound S100B (ββ)[131], but neither region is

predicted by JPred4 to be structured (Table 2.2). Thus the JPred predictions do not appear to completely accurately predict the secondary structure induced upon binding to known partners.

2.3.3. All the selected IDRs lack α-helical structure in solution

CD spectra reflect the average signal from the entire molecular population over the time course of the measurement. Therefore, it is impossible to distinguish between the average secondary structures maintained in the entire population of the sample, versus more extensive secondary structure maintained in a smaller sub-population of the sample with a corresponding decrease of that secondary structure in the remainder of the population. Although the flexible structure of IDRs enables rapid (<millisecond) conformational exchange amongst multiple conformations including fully disordered and transient, partially-folded conformers in solution[133, 134], there is no reason to expect that a sub-population within a chemically identical set of molecules in a chemically homogeneous environment would be more ordered on average than the rest of the population. Therefore, since the molecules in solution are identical in terms of conformational behavior, we can apply the first interpretation, which requires a definition of minimal length to define a stable secondary structure, rather than the second interpretation, which requires a definition of minimal percentage of the population that should adopt a given secondary structure to be considered a physiologically relevant structural state. The first interpretation has been commonly used in classical literature, and definitions of minimal length to define secondary structure are well established[135]. A minimal helix is defined by the formation of two consecutive hydrogen bonds, therefore; a minimal α-helix requires at least six residues in a helical conformation[135, 136]. A β-strand cannot exist unless stabilized by hydrogen bonds to another strand. Stable β-structure is defined by four consecutive hydrogen bonds in either a parallel or anti-parallel bonding pattern, thereby involving at least six residues (three residues in each of two strands)[135].

Secondary structure content estimated by analysis of CD spectra for each of the thirteen IDRs showed that none contained more than two residues in a helical conformation (Figures 2.1-2.3, Table 2.3-2.5), which is insufficient for the formation of a stable α-helix requiring a minimum of six residues. Formation of a minimal beta-structure requires six residues[135], and seven of the thirteen IDRs (Table 2.3-2.5) appear to have between six to eight residues in extended or β-conformation which may be sufficient for forming short stable β-hairpins (Table 2.3-2.5). Thus, while none of the selected IDRs had stable

helical structure, several may have short β-strands. However, the marginal secondary structure content suggests that all the peptides tested are IDRs.

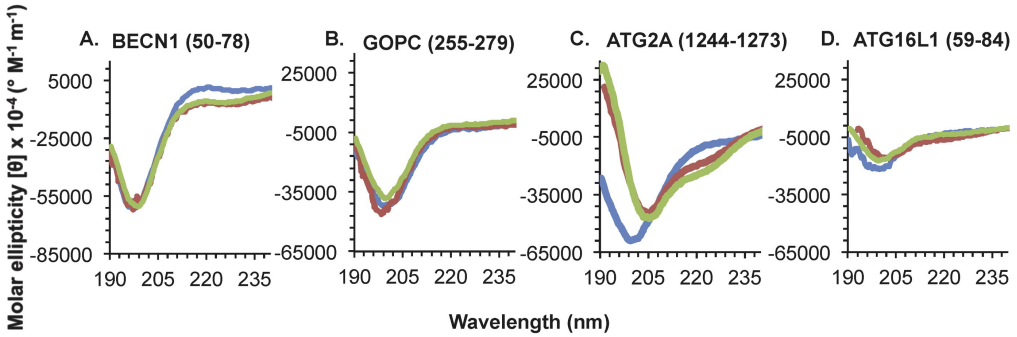

Figure 2.1. CD spectra of IDRs lacking Anchor regions.
Different spectra correspond to 0% TFE (blue), 25% TFE (red), or 40% TFE (green).

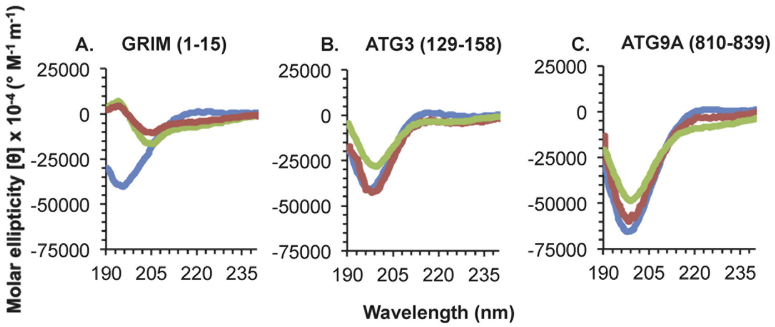

Figure 2.2. CD spectra of IDRs with Anchor regions that do not undergo a TFE-induced helical transition. Different spectra correspond to 0% TFE (blue), 25% TFE (red), or 40% TFE (green). (A) A known β-MoRF. (B-C) IDRs with experimentally unverified structural transitions.

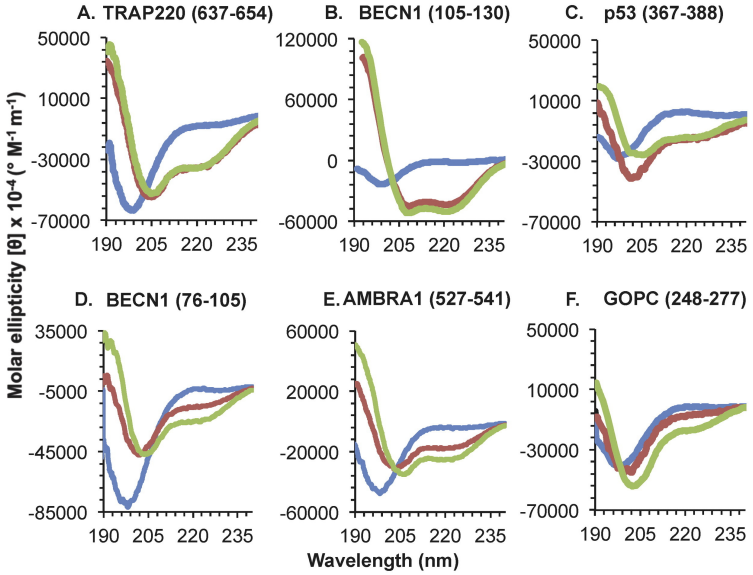

Figure 2.3. CD spectra of IDRs with Anchor regions that undergo a TFE-induced helical transition. Different spectra correspond to 0% TFE (blue), 25% TFE (red), or 40% TFE (green). (A-C) Known α-MoRFs. (D-F) IDRs with experimentally unverified structural transitions.

Table 2.3. Secondary structure content in consensus IDRs lacking Anchor regions.

Protein (residues in peptide)		BECN1 (50-78)	GOPC (255-279)	ATG2A (1244-1273)	ATG16L1 (59-84)
		Average estimated secondary structure content (%)			
Helix (%)	No TFE	1 ± 2	6 ± 2	5 ± 2	4 ± 2
	25% TFE	7 ± 6	12 ± 8	18 ± 1	17 ± 1
	40% TFE	1 ± 1	6 ± 3	16 ± 1	3 ± 2
Strand (%)	No TFE	16 ± 10	32 ± 1	19 ± 5	16 ± 6
	25% TFE	14 ± 11	28 ± 4	24 ± 2	17 ± 2
	40% TFE	8 ± 3	35 ± 4	25 ± 1	35 ± 2
Coil (%)	No TFE	81 ± 8	61 ± 1	76 ± 2	79 ± 8
	25% TFE	80 ± 17	59 ± 4	58 ± 2	65 ± 3
	40% TFE	89 ± 3	58 ± 1	58 ± 2	60 ± 6

Table 2.4. Secondary structure content in known MoRFs.

Protein (residues in peptide)		GRIM (1-15)	TRAP220 (637-654)	BECN1 (105-130)	p53 (367-388)
Type of MoRF		β-MoRF	α-MoRF	α-MoRF	α- & β- MoRF
		Average estimated secondary structure content (%)			
Helix (%)	No TFE	0 ± 1	7 ± 3	2 ± 1	6 ± 4
	25% TFE	0 ± 4	62 ± 10	48 ± 2	34 ± 10
	40% TFE	7 ± 7	50 ± 3	55 ± 2	31 ± 1
Strand (%)	No TFE	9 ± 8	29 ± 5	28 ± 7	31 ± 2
	25% TFE	19 ± 14	6 ± 3	12 ± 1	12 ± 4
	40% TFE	22 ± 16	7 ± 6	8 ± 1	18 ± 2
Coil (%)	No TFE	89 ± 7	62 ± 6	65 ± 6	63 ± 2
	25% TFE	72 ± 7	32 ± 11	40 ± 2	54 ± 7
	40% TFE	70 ± 20	43 ± 10	36 ± 2	50 ± 4

Table 2.5. Secondary structure content in IDRs with Anchor regions with unidentified binding partners.

Protein (residues in peptide)		ATG3 (129-158)	ATG9A (810-839)	BECN1 (76-105)	AMBRA1 (527-541)	GOPC (248-277)
		Average estimated secondary structure content (%)				
Helix (%)	No TFE	3 ± 5	3 ± 2	1 ± 1	2 ± 1	3 ± 1
	25% TFE	4 ± 1	5 ± 4	20 ± 3	37 ± 1	4 ± 1
	40% TFE	-3 ± 6	4 ± 0	33 ± 3	50 ± 4	20 ± 1
Strand (%)	No TFE	27 ± 8	20 ± 4	4 ± 2	4 ± 4	19 ± 4
	25% TFE	26 ± 3	28 ± 2	8 ± 6	10 ± 5	17 ± 6
	40% TFE	22 ± 9	18 ± 5	7 ± 4	8 ± 5	20 ± 4
Coil (%)	No TFE	63 ± 15	76 ± 2	95 ± 4	95 ± 5	79 ± 4
	25% TFE	67 ± 5	59 ± 15	53 ± 5	53 ± 5	78 ± 4
	40% TFE	76 ± 10	78 ± 5	43 ± 8	43 ± 9	60 ± 4

2.3.4. IDRs lacking Anchor regions fail to undergo disorder-to-helix transitions

The secondary structure estimated from the CD spectra indicates that addition of TFE causes varying fluctuations in β-content of the different IDRs lacking Anchor regions (Table 2.3). In BECN1 residues 50-78, addition of 40% TFE, results in a decrease in the number of residues in β-conformation from five to two, clearly insufficient for the formation of stable β-strands. However, GOPC residues 255-279 and ATG2A residues 1244-1273 show marginal increases of residues in β-content at 25% and 40% TFE, that would enable them to form short stable β-sheets. Finally, β-content in ATG16L (residues 59-84) remains constant at four residues upon addition of 25% TFE, but in 40% TFE, the number of residues in β -conformation increases to nine residues, which is sufficient for formation of stable β-strands.

Strikingly, the secondary structure content estimated from the CD spectra indicates that addition of TFE does not induce stable helical structure in the four IDRs lacking Anchor regions (Figure 2.1, Table 2.3). Helical content peaks in the presence of 25% TFE and is maintained or slightly decreased in 40% TFE; however, neither condition results in the six-residue minimum required for a stable helix. Thus, the changes in coil content estimated from the CD spectra of the four IDRs lacking Anchor regions appear to simply reflect fluctuations in helix and strand content (Table 2.3).

2.3.5. IDRs that bear Anchor regions, yet do not undergo TFE-induced helical transitions, may be β- or σ- MoRFs

GRIM residues 1-15, which include a known β-MoRF and bear an Anchor region, do not become helical in the presence of TFE (Figure 2.2A, Table 2.4). The helical content, none at 0% and 25% TFE, and one residue at 40% TFE, is insufficient to stabilize a minimal helix, which requires at least six residues. However, while the number of residues in β-conformations increases marginally in 25% and 40% TFE, the three residues estimated to be in extended conformation in the presence of TFE, are also insufficient to form a stable β-strand, which requires six residues. These results demonstrate that TFE does not induce helicity in β-MoRFs, nor does it appear to support formation of stable β-structure.

Two of the Anchor region-containing IDRs for which binding partners have not yet been identified, ATG3 residues 129-158 and ATG9A residues 810-839, also do not undergo a disorder-to-helix transition in either 25% or 40 % TFE (Figure 2.2). These IDRs contain two or fewer helical residues (Table 2.5) even in 25% and 40% TFE, which is insufficient for the formation of a single stable helical turn. This

failure to induce helicity indicates that ATG3 residues 129-158 and ATG9A residues 810-839 are not α-MoRFs. In contrast, in both these IDRs, there is sufficient β-content in TFE concentrations ranging from 0%-40% to permit the formation of β-strand structure; with the β-content increasing marginally at 25% TFE, then decreasing marginally at 40% TFE. Thus, TFE causes a marginal increase in the β-content, similar to that observed for the known β-MoRF, GRIM residues 1-15. However, as this increase is marginal, and given the limited number of known β-MoRFs available for testing, we cannot suggest that this method is a reliable way of identifying β-MoRFs.

2.3.6. IDRs that bear Anchor regions and undergo TFE-induced helical transitions are α-MoRFs

TRAP220 residues 637-654, BECN1 residues 105-130, and p53 residues 367-388 are all known α-MoRFs with known binding partners. Consistent with this, addition of TFE induces helicity in each IDR (Figure 2.3 A-C). Although none of these IDRs have a helical content sufficient for formation of a stable helix in the absence of TFE, addition of TFE induces a significant disorder-to-helix transition in all three IDRs (Table 2.4). TFE-induced helical content peaks upon addition of 25% TFE, with 40% TFE inducing slight fluctuations. In TRAP220 residues 637-653, a maximum of eleven helical residues are stabilized at 25% and nine are maintained at 40% TFE. In BECN1 residues 105-130, 25% and 40% TFE induces the stabilization of thirteen and fourteen helical residues respectively. And finally, in p53 residues 367-388, 25% and 40% TFE stabilize eight and seven helical residues respectively. In each of these IDRs, the number of residues in β- and coil conformations decreases upon addition of TFE, suggesting that TFE destabilizes β-structure in these α-MoRFs.

All known α-MoRFs become helical in the presence of 25% and 40% TFE; however, the maximum number of helical residues induced is not always comparable to the number of helical residues actually observed in the corresponding crystal structures. Of the three known α-MoRFs, only in p53 does TFE induce a number of helical residues (Table 2.4) comparable to that observed in the crystal structure. Residues 378-384 (seven residues) of p53 become helical upon binding Ca^{2+}-bound S100B(ββ) [131], which is equal to those stabilized by 40% TFE and just one less than the maximum observed (eight helical residues) in 25% TFE. However, in BECN1 residues 105-130, a maximum of fourteen residues are stabilized in helical conformations by 40% TFE, while upon binding to BCL2 homologs, BECN1 residues 108-126 (nineteen residues) form a helix[43, 44, 46], indicating that for this IDR the number of residues that

undergo helical transformation may be underestimated by this method. Conversely, this method appears to slightly overestimate the number of TRAP220 residues that become helical upon binding. Only TRAP220 residues 643-648 (six residues) form a helix upon binding the retinoid receptor (RAR/RXR) ligand binding domain heterodimer[127]; however, eleven and nine helical residues are stabilized by 25% and 40% TFE respectively. Therefore, although this method stabilizes helical structure in α-MoRFs, the number of helical residues induced upon binding cannot be directly inferred from the secondary structure content estimations.

Amongst the Anchor-containing IDRs with unknown binding partners (Table 2.2), three IDRs: BECN1 residues 76-105, AMBRA1 residues 527-541, and GOPC residues 248-277, appear to undergo a TFE-induced disorder-to-helix transition. Helical content in BECN1 residues 76-105 increases from none in the absence of TFE, to six and ten, in 25% and 40% TFE respectively (Table 2.5, Figure 2.3D-F). In AMBRA1 residues 527-541, helical content increases from nothing in the absence of TFE to five and eight, in 25% and 40% TFE respectively. This increase occurs at the expense of residues in coil conformation, as both contain just one residue in β-conformation in the absence of TFE. Like the known α-MoRFs, the increase in helical content of BECN1 residues 76-105 and AMBRA1 residues 527-541 is observed in both 25% and 40% TFE. However, unlike the known α-MoRFs and BECN1 residues 76-105, for AMBRA1 residues 527-541, a stable helix comprising at least six residues is induced only in 40% TFE (Table 2.5); while in GOPC residues 248-277, 25% TFE is insufficient for inducing the disorder-to-helix transition, which occurs only upon addition of 40% TFE (Figure 2.3F, Table 2.5). Further, unlike the known α-MoRFs, the increase in helicity occurs only at the expense of residues in coil conformation, and residues in β-conformation do not appear to be destabilized by the addition of TFE (Table 2.4-2.5).

2.3.7. Effect of amino acid content on response to TFE

Amino acid composition of the various IDRs used in this study was analyzed (Figure 2.4). Relative to the average frequency of 37% hydrophobic residues and 51% polar or charged amino acids in proteins[137], all but one of the selected IDRs have a lower content of hydrophobic residues and a greater content of polar or charged residues, as is expected for typical intrinsically disordered regions (Figure 2.4). The known β-MoRF, GRIM residues 1-15, is an exception as 67% of this IDR is comprised of hydrophobic residues, which is higher than average; while only 30% of the residues are polar or charged,

which is lower than average. All but two of the selected IDRs also have a lower proline content (Figure 2.4), compared to the average frequency of 5% found in proteins[137]. While the proline content of GRIM residues 1-15 is similar to the average; ATG2A residues 1244-1273 contain 33% prolines, which is substantially higher than the average proline content. Neither of these IDRs undergoes a transition to helix. Lastly, the glycine content of all the selected IDRs is approximately 7% (Figure 2.4), similar to that of average proteins[137]. However, while the amino acid composition of these peptides is consistent with these regions being IDRs, there does not appear to be a correlation between the amino acid composition of the IDRs and the TFE-induced helical transitions, beyond the absence of prolines in IDRs that do undergo a helical transition. This is consistent with findings that indicate that the accuracy of ANCHOR in predicting disordered binding regions is independent of amino acid content[113], and may also explain the ambiguity in the JPred predictions for these regions. Thus, beyond the lack of proline residues, the amino acid composition does not appear to be a direct predictor of potential transitions.

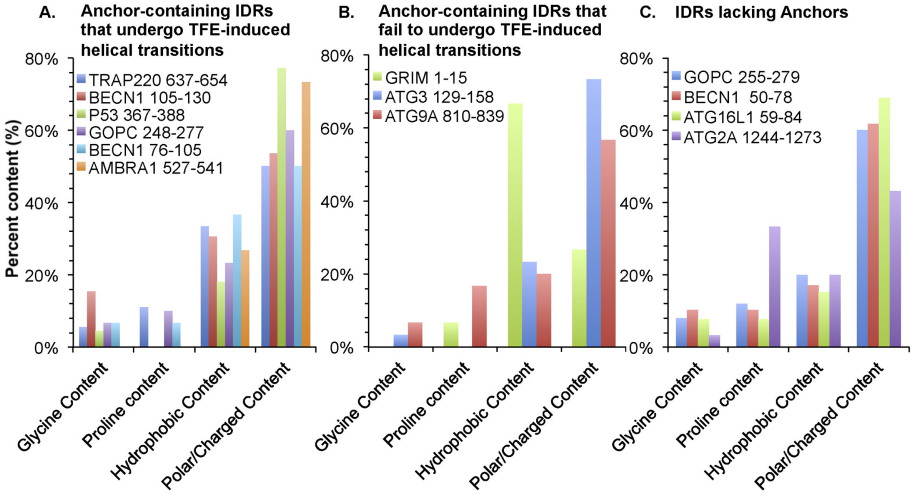

Figure 2.4. Amino acid content of the selected IDRs.
Amino acid content of: (A) Anchor-containing IDRs that undergo TFE-induced helical transitions, (B) Anchor-containing IDRs that fail to undergo TFE-induced helical transitions, and (C) IDRs lacking Anchors.

2.4. Discussion and conclusions

All but one of the sequences selected for this study were predicted to be entirely disordered by at least one, and often all three of the disorder predictors, IUPred, IsUnstruct, or PONDR-FIT, as none of these sequences contained contiguous non-disordered stretches that were long enough to form stable secondary structure. The only exception was BECN1 105-130, which has short regions that are not predicted to be disordered, but this sequence was included in our analysis as BECN1 residues 105-130

have previously been experimentally shown to be disordered even in the context of adjacent BECN1 domains[30]. Additionally, many of the IDRs selected for this study are preceded by long regions, which are also predicted to be disordered[30, 127-131]. Lastly, our CD analysis verifies that, in solution, each of these IDRs is largely disordered.

We used ANCHOR and DisoRDPbind to predict regions that may fold upon binding but found little agreement between predictions by these different programs. DisoRDPbind also did not accurately predict binding regions within known MoRFs. Interestingly, we find that JPred4 often predicts secondary structure of sufficient length to permit formation of stable secondary structure that overlap or lie within predicted Anchor regions, rather than predicting these regions as disordered. However, the actual predictions of the type of secondary structure were not always accurate, nor did the lengths of the predicted folded structures always agree with experimental observations, perhaps due to the conformational flexibility of these regions.

Our results indicate that the combined approach described in this report, of performing sequence-based predictions using programs like ANCHOR followed by experimental assessment of structural transitions in the presence of TFE, is suitable for identifying IDRs that are α-MoRFs. Strikingly, we show that while TFE induces helicity in known α-MoRFs, it does not induce disorder-to-helix transitions in IDRs lacking Anchor regions, or in known β-MoRFs. This combined predictive and experimental approach did not produce any false positives or false negatives, i.e. we were able to use this approach to successfully identify all the α-MoRFs we tested, and no known β-MoRFs were misidentified as α-MoRFs. Thus, this approach appears to reliably identify α-MoRFs, and distinguish between these and β- or σ-MoRFs.

Although p53 residues 367-388 can be either a α- or β-MoRF depending on its binding partner [130, 131], it behaves as an α-MoRF in the presence of TFE where a helical transition is induced at the expense of β- and coil content. This is distinct from GRIM, the β-MoRF, in which β-content slightly increases while helical content is not increased by TFE. Therefore, this method can identify IDRs that are α-MoRFs even if they can adopt different conformations upon binding to different partners. Conversely, identification of an IDR as an α-MoRF does not preclude it from undergoing other disorder-to-order transitions upon binding to a different partner.

We have used the method proposed in this study to identify three new potential α-MoRFs: BECN1 residues 76-105, AMBRA1 residues 527-541, and GOPC residues 248-277; from amongst the IDRs for which binding partners have not yet been identified. Like the known α-MoRFs, the helical content of each of these IDRs increases in the presence of TFE. However, unlike the known α-MoRFs, in GOPC residues 248-277 an increase in helicity is seen only in 40% TFE and in AMBRA1 residues 527-541, although an increase in helical content is observed in 25% TFE, stable helical content is achieved only in 40% TFE. BECN1 residues 76-105 and AMBRA1 residues 527-541 lack significant β-content, hence their TFE-induced helical transition occurs at the expense of coil conformation. GOPC residues 248-277 contain β-structure; however, unlike the known α-MoRFs with β-content, the TFE-induced increase in helicity in GOPC residues 248-277 occurs at the expense of the coil-content, and the β-content is not affected by TFE. Therefore, although GOPC residues 248-277 may become helical upon binding to one partner, perhaps the lack of destabilization of β-structure, which is similar to that seen for β-MoRFs, suggests that it may adopt β-conformations upon binding to other partners.

We also identify two potential Anchor-containing IDRs that do not undergo helical transition in the presence of TFE. It is possible that these might be β- or σ-MoRFs. Based on the similarity in TFE-induced behavior to the known β-MoRFs, it is tempting to speculate that ATG3 residues 129-158 and ATG9A residues 810-839 are also β-MoRFs.

In this study we included two overlapping IDRs from GOPC: the first comprised of residues 248-277 and the second comprising residues 255-279. GOPC residues 248-257 comprise a ten-residue Anchor region, thus the first GOPC IDR includes the Anchor while the second contains only the last three residues of the Anchor. Strikingly, loss of the first seven Anchor residues abrogates the ability of the latter GOPC IDR to undergo the TFE-induced helical transition observed in GOPC residues 248-277. Identification of the Anchor region-containing GOPC residues 248-277 as a α-MoRF, demonstrates that the TFE-induced helical transition requires the presence of an Anchor region. This is reminiscent of an earlier study wherein residues within the Anchor region was demonstrated to be essential for nucleation of a binding-induced disorder-to-helix transition in the case of BECN1 residues 105-130[104].

Thus, we propose that ANCHOR predictions combined with assessment of TFE-induced helicity serves to identify α-MoRFs. However, we cannot use this method to generally exclude the possibility that

these regions may also form either σ- or β-MoRFs in the context of other binding partners. It would be interesting to eventually compare the predictions in this report to experimental structural transitions upon binding to different partners.

It is useful to have a convenient experimental diagnostic to identify IDRs that are α-MoRFs as that would help better identify disordered regions with a potential biological function in mediating macromolecular interactions. Further, a tool to identify α-MoRFs may help identify a potential binding partner for the IDR from a wider list of potential interaction partners for the full-length protein, by providing information about the IDR conformation and residues most likely to be involved in the interaction. This, in turn, would lead to a better mechanistic understanding of the function of proteins involved in IDR-mediated interactions. For example, within the BECN1 residues 76-105 shown by this study to be a α-MoRF, residue S90 has previously been shown to be phosphorylated by MK2 (members of the p38 mitogen-activated protein kinase signaling pathway PAPKAPK2), which binds helical substrates and is essential for the tumor suppressor function of BECN1[138]. Identification of BECN1 residues 76-105 as a α-MoRF supports the hypothesis that MK2 binds directly to this region to phosphorylate and activate BECN1. This information may permit the analysis of MK2 structure to determine the binding determinants for interaction with, and phosphorylation of, BECN1.

Information regarding binding-associated structural transitions in IDRs would not only provide a basis for understanding these interactions but may also be useful for developing therapeutics targeting these proteins. For instance, a potential therapeutic was recently discovered that specifically binds p27, an IDR-containing protein that regulates the cell cycle via interaction with Cdk2/cyclinA. This molecule traps the disordered kinase binding domain of p27 in a conformation that prevents folding and binding to Cdk2/cyclin A[139].

An example of an IDR that was shown to be an α-MoRF and subsequently targeted for therapeutic development is the BECN1 BH3 domain (BECN1 105-130). An Anchor region within the BECN1 BH3 domain nucleates a helical transition upon binding BCL2 homologs, resulting in down-regulation of autophagy, a cellular homeostasis pathway important for innate immune defenses[30, 140]. An understanding of these structural transitions and the role of the Anchor region, enabled development of a peptide that selectively inhibits the interaction of BECN1 with the murine γ-herpesvirus68 BCL2 M11, to

prevent viral down-regulation of autophagy, while leaving cellular regulation unaffected[140]. However, unlike p27 and the BECN1 BH3 domain, binding partners for many Anchor regions remain unknown.

In a separate, parallel study, we used X-ray crystallography, CD spectroscopy, small angle X-ray scattering (SAXS), and electron paramagnetic resonance (EPR) to demonstrate that the BECN1 flexible helical domain (FHD) is partially disordered and partially helical (forming a 2.5 turn helix), when not in complex with other proteins, even in larger BECN1 fragments containing adjacent BECN1 domains[33]. Further, we found that although a binding partner for this region had not been identified, the FHD contains an Anchor region and undergoes TFE-induced helical transitions suggesting that it is a α-MoRF. Subsequently, in a 4.4 Å crystal structure (PDB 5DFZ) of the full-length yeast VPS34:VPS15:VPS30:VPS38 complex (VPS30 is the yeast homolog of BECN1), it was shown that VPS38, which binds primarily via the VPS30/BECN1 coiled-coil domain, also interacts in a coiled-coil manner with the BECN1 FHD (labeled CC1 in this structure)[17]. Therefore, in this complex the FHD is completely helical, forming an eight-turn helix, verifying our prediction that the FHD is a α-MoRF. Thus, our method has successfully identified an IDR to be a α-MoRF; which was then confirmed by studies wherein a binding partner for this region was established, and the co-complex structure determined.

Identification of IDRs that are α-MoRFs by our method may enable development of therapeutics targeting interactions mediated by these IDRs by similar methods. Many of the IDR-containing proteins included in this study are implicated in diverse diseases. For instance, GOPC, which binds and modulates membrane proteins, including several oncogenes, is associated with glioblastoma and other cancers, as well as cystic fibrosis[141, 142]. Although a binding partner for the selected GOPC IDR remains to be identified, the GOPC α-MoRF identified here could potentially be targeted for therapeutic treatment by designing molecules that prevent helix formation or bind very tightly to the helical conformation of the GOPC α-MoRF, preventing interactions that facilitate oncogenesis. Thus, this study provides a useful tool toward better understanding IDRs and their conformational transitions.

CHAPTER 3. STRUCTURAL TRANSITIONS IN CONSERVED, ORDERED BECN1 DOMAINS ESSENTIAL TO REGULATING AUTOPHAGY

3.1. Introduction

Co-immunoprecipitation assays show that the BECN1 homodimer exists in cells[143]. BECN1 self-associates via its CCD to form a homodimer in the absence of other binding partners. However, the last 18 residues of the CCD (residues 248-265) form a helical region that we have named the "overlap helix" (OH), because it has been crystallized in two differently-packed states: (i) as part of the anti-parallel CCD homodimer[33, 34], constituting the C-terminal four turns of each CCD; or (ii) packed against the BARAD[36]. Notably, in the structure of the quaternary complex of yeast VPS34/PI3KC3, VPS15, VPS30/BECN1 and VPS38/UVRAG, VPS30 residues 304-327 (equivalent to residues 248-271 of the human BECN1 OH and BARAD) are part of the VPS30/BECN1 CCD and pack against the CCD of VPS38/UVRAG, stabilizing the parallel CCD heterodimer[17]. This led to the assumption that the OH is part of the CCD in the context of full-length BECN1.

The OH and BARAD appear to undergo conformational changes upon binding membranes. The BARAD bears an aromatic finger comprised of residues 359-361 (residues 430-432 of yeast VPS30) that is required for BARAD binding to membranes[36]. Hydrogen-deuterium exchange mass spectrometry (HDX-MS) experiments show that this aromatic finger likely directly interacts with membranes when BECN1 is part of the PI3KC3 complex[17]. The HDX-MS experiments also indicate that VPS30 residues 312-327 (equivalent to BECN1 residues 256-271), which comprise the latter half of the OH, undergo membrane-binding induced conformational changes but do not directly bind membranes[17]. Inclusion of the OH as part of the anti-parallel CCD homodimer would greatly increase stability of the autophagy-inactive homodimer. However, it has not been shown whether the BECN1 OH is part of the CCD or the BARAD when BECN1 is not part of PI3KC3 complexes. Therefore, different conformational states of the BECN1 OH likely play important roles in regulating protein:membrane or protein:protein interactions. Thus, understanding BECN1 conformational flexibility is key to understanding varied BECN1 interactions, which in turn, will improve our understanding of the role of BECN1 in autophagy. Ultimately this is essential to better understand how eukaryotic organisms respond to and survive cellular stressors.

Therefore, in order to determine the packing state of the OH in the absence of heterologous binding partners, we investigated the solution structure, stability, and self-dissociation of purified CCD, OH+BARAD, and a two-domain CCD-BARAD BECN1 fragment, as well as the impact the impact of mutating OH residues that pack against either the partner helix of the CCD or against the BARAD. The solution structure was established using CD spectroscopy and size-exclusion chromatography (SEC) in tandem with SAXS, self-dissociation using ITC, and stability of the CCD, OH+BARAD, and CCD-BARAD by monitoring thermal denaturation by CD. Lastly, we assessed the impact of the OH mutations on cellular autophagy by evaluating the change in the number of autophagosomes per cell.

3.2. Materials and methods

3.2.1. Structural superimposition and calculation of buried surface area

The crystal structures of the human BECN1 CCD homodimer (PDB ID: 5HHE) and the human BECN1 BARAD containing the OH (OH+BARAD) (PDB ID: 4DDP) were superimposed using the align function of the PyMOL Molecular Graphics System (Schrodinger, LLC)[144]. The PDBePISA[145] server (http://www.ebi.ac.uk/pdbe/pisa/) was used to calculate interface buried surface areas.

3.2.2. Plasmid preparation

Human BECN1 residues 175-265 (CCD), residues 175-450 (CCD-BARAD), or residues 248-450 (OH+BARAD) were cloned between the *NcoI* and *NotI* restriction enzyme sites of the pMBP-Parallel-1 bacterial expression vector[146]. The BECN1 OH+BARAD was also cloned between the *XhoI* and *BamHI* restriction enzyme sites of the pET-15b bacterial expression vector encoding an N-terminal His_6 tag (Table 3.1). The OH Tetrad mutation (V250A, M254A, L261A, and L264A) and AFM (F359D, F360D, and W361D) were generated via site-directed mutagenesis. A pcDNA3.1 mammalian expression vector encoding full-length, human wild-type (WT) BECN1 was subjected to site-directed mutagenesis to generate the OH Tetrad mutation (V250A, M254A, L261A, and L264A).

Table 3.1. Protein constructs used in this study.

	Label	Description	BECN1 Residues
1	MBP-BECN1 CCD-BARAD	MBP-BECN1 CCD-BARAD fusion protein	175-450
2	MBP-BECN1 CCD-BARAD^{TETRAD}	MBP-BECN1 CCD-BARAD fusion protein OH Tetrad Mutant	175-450 (V250A,M254A,L261A,L264A)
3	BECN1 OH+BARAD	His_6-BECN1 OH+BARAD	248-450
4	MBP-BECN1 OH+BARAD	MBP-BECN1 OH+BARAD fusion protein	248-450
5	MBP-BECN1 OH+BARAD^{TETRAD}	MBP-BECN1 OH+BARAD fusion protein OH Tetrad Mutant	248-450, (V250A,M254A,L261A,L264A)
6	MBP-BECN1 CCD	MBP-BECN1 CCD fusion protein	175-265
7 *	BECN1 CCD	BECN1 CCD	175-265
8	MBP-BECN1 CCD^{TETRAD}	MBP-BECN1 CCD fusion protein OH Tetrad Mutant	175-265, (V250A,M254A,L261A,L264A)
9 *	BECN1 CCD^{TETRAD}	BECN1 CCD OH Tetrad Mutant	175-265, (V250A,M254A,L261A,L264A)
10	MBP-BECN1 CCD-BARAD^{AFM}	MBP-BECN1 CCD-BARAD fusion protein Aromatic Finger Mutant	175-450 (F359D,F360D,W361D)
11 *	BECN1 CCD-BARAD^{AFM}	BECN1 CCD-BARAD Aromatic Finger Mutant	175-450, (F359D,F360D,W361D)
12	MBP-BECN1 CCD-BARAD^{AFM,TETRAD}	MBP-BECN1 CCD-BARAD fusion protein Aromatic Finger Mutant +OH Tetrad Mutant	175-450 (V250A,M254A,L261A,L264A)
13 *	BECN1 CCD-BARAD^{AFM,TETRAD}	BECN1 CCD-BARAD Aromatic Finger Mutant + OH Tetrad Mutant	175-450, (F359D,F360D,W361D), (V250A,M254A,L261A,L264A)
14	MBP-BECN1 OH+BARAD^{AFM}	MBP-BECN1 OH+BARAD fusion protein Aromatic Finger Mutant	248-450, (F359D,F360D,W361D)
15 *	BECN1 OH+BARAD^{AFM}	His_6-BECN1 OH+BARAD Aromatic Finger Mutant	248-450, (F359D,F360D,W361D)
16	MBP-BECN1 OH+BARAD^{AFM,TETRAD}	MBP-BECN1 OH+BARAD fusion protein Aromatic Finger Mutant + OH Tetrad Mutant	248-450, (F359D,F360D,W361D), (V250A,M254A,L261A,L264A)
17 *	BECN1 OH+BARAD$^{AFM, TETRAD}$	His_6-BECN1 OH+BARAD Aromatic Finger Mutant + OH Tetrad Mutant	241-450, (F359D,F360D,W361D), (V250A,M254A,L261A,L264A)

* Denotes constructs used for our final published analyses since the AFM dramatically reduces aggregation and permits large-scale purification of high-quality, homogenous, BARAD-containing constructs, and also does not appear to impact secondary or tertiary structure of these fragments as described in subsequent sections.

For preliminary analyses with maltose binding protein (MBP)-tagged fusion protein, the flexible long linker (LL) between the MBP-tag and N-terminus of the BECN1 domains was replaced with a short, triple alanine linker (SL) via site-directed mutagenesis (primer: 5' GCC CTG AAA GAC GCG CAG ACT AAT GCA GCA GCA CTG TAT TTT CAG GGC GCC 3') in order to reduce flexibility of the linker region to facilitate analyses of the effect of the Tetrad mutant on protein structure and flexibility.

3.2.3. Protein expression and purification

For the BECN1 CCD and CCD^{TETRAD}, *Escherichia coli* (*E. coli*) BL21(DE3)·pLysS cells were transformed and grown in LB media with 100 µg/mL ampicillin at 37 °C to an OD_{600} of ~0.8 prior to equilibrating the temperature to 20 °C. For WT and mutant BECN1 CCD-BARAD and His_6-BECN1 OH+BARAD, *E. coli* Arctic Express cells were transformed and grown in LB media with 100 µg/mL ampicillin at 30 °C to an OD_{600} of ~0.6 prior to equilibrating the temperature to 10 °C. In each case, cells were grown to an OD_{600} of 0.8-1.2. Expression of the CCD and CCD-BARAD was induced by addition of 0.5 mM isopropyl thio-β-D-galactoside (IPTG). His_6-OH+BARAD constructs were induced with 0.3 mM IPTG. All constructs were expressed overnight. One protease inhibitor tablet (Bimake, Houston, TX, US) per liter of cell culture was added during lysis for all protein constructs. At each stage of purification, protein purity 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.

For all cleaved BECN1 CCD and BECN1 CCD-BARAD constructs, soluble MBP-tagged fusion protein was purified from clarified crude cell lysate by amylose affinity chromatography on Amylose Resin (NEB BioLabs, Ipswich, US). The column was washed with 10 column volumes of Wash Buffer 1 and then 10 column volumes of Wash Buffer 2 (Wash Buffer 1: 25 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM EDTA, 2 mM Dithiothreitol (DTT), 10% glycerol; Wash Buffer 2: 25 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM EDTA, 2 mM DTT) and subjected to overnight, on-column cleavage by glutathione S-transferase – tobacco etch virus (GST-TEV) protease at 4 °C. Cleaved protein was washed from the column with Wash Buffer 2 and further purified by anion exchange chromatography on a 5/50 GL MonoQ (Pharmacia Biotech – GE Life Sciences, Pittsburgh, US) (MonoQ Buffer A: 50mM Tris-HCl pH 7.5, 2mM DTT; MonoQ Buffer B: 50mM Tris-HCl pH 7.5, 2mM DTT, 1M NaCl). BECN1 protein fragments were purified to homogeneity by size exclusion chromatography (SEC), using a 10/30 Superdex 200 (GE Life Sciences, Pittsburgh, US) column (SEC buffer: 25 mM Tris-HCl pH 7.5, 300 mM NaCl, 2 mM DTT). The cleaved proteins were concentrated in a 10 kD Amicon Ultra-0.5 centrifugal concentrator (EMD Millipore, Billerica, MA, US).

For WT His_6-BECN1 OH+BARAD, fusion protein was purified from clarified crude cell lysate by nickel immobilized affinity chromatography (IMAC) using a HisTrap column (GE Life Sciences, Pittsburgh,

US). The column was washed with 10 column volumes of Wash Buffer (Wash Buffer: 25 mM Tris HCl pH 8.0, 150 mM NaCl, 10% v/v glycerol, 25 mM imidazole) and protein was eluted with Wash Buffer containing 350 mM imidazole. WT BECN1 OH+BARAD was subsequently purified by SEC, using a 10/30 Superdex 200 in tandem with a 10/30 Superdex 75 column (GE Life Sciences, Pittsburgh, US) (SEC buffer: 25 mM Tris-HCl pH 7.5, 300 mM NaCl, 2 mM DTT) and concentrated to a maximum of 2 mg/mL in a 10 kD Amicon Ultra-0.5 centrifugal concentrator (EMD Millipore, Billerica, MA, US). 0.4 % (w/v) β-octyl glucoside was added during lysis and concentration stages (except final protein concentration) to prevent aggregation.

For His_6-BECN1 OH+BARAD^{AFM} constructs, fusion protein was purified from clarified crude cell lysate by nickel immobilized affinity chromatography (IMAC) using Ni Sepharose High Performance nickel-charged IMAC resin (GE Life Sciences, Pittsburgh, US). The column was washed with 10 column volumes of Wash Buffer (Wash Buffer: 50 mM HEPES acid pH 7.5, 300 mM NaCl, 10% v/v glycerol, 25 mM imidazole) and protein was eluted with Wash Buffer containing 350 mM imidazole. BECN1 OH+BARAD constructs were subsequently purified to homogeneity by SEC, using a 10/30 HR Superdex 200 column (GE Life Sciences, Pittsburgh, US) (SEC buffer: 25 mM Tris-HCl pH 7.5, 300 mM NaCl, 2 mM DTT) and concentrated to a maximum of 6 mg/mL in a 10 kD Amicon Ultra-0.5 centrifugal concentrator (EMD Millipore, Billerica, MA, US).

For preliminary analyses with MBP-tagged fusion proteins, *E. coli* Arctic Express(DE3) cells were transformed by the WT MBP-SL-BECN1 CCD-BARAD, MBP-SL-BECN1 CCD-BARAD^{AFM}, MBP-SL-BECN1 CCD-BARAD^{TETRAD}, MBP-SL-BECN1 OH+BARAD^{AFM}, MBP-SL-BECN1 OH+BARAD^{TETRAD}, or MBP-SL-BECN1 OH+BARAD^{AFM,TETRAD} expression vectors and grown in LB media with 100 µg/mL ampicillin at 30 °C to an OD600 of ~0.6 prior to allowing the temperature to equilibrate to 10 °C. Recombinant protein expression was then induced by addition of 0.3 mM IPTG upon an OD_{600} of 1.0-1.2 and expressed overnight. *E. coli* BL21(DE3)pLysS cells were transformed by the MBP-SL-BECN1 CCD and MBP-SL-BECN1 CCD^{TETRAD} expression vectors and grown in LB media with 100 µg/mL ampicillin at 37 °C to an OD600 of ~0.8 when recombinant protein expression was induced by addition of 0.5 mM IPTG and expressed for two hours.

For WT and mutant MBP-BECN1 CCD-BARAD constructs, fusion protein was purified from clarified crude cell lysate by amylose affinity chromatography on Amylose Resin (NEB BioLabs, Ipswich, US). The column was washed with 10 column volumes of Wash Buffer (25 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% v/v glycerol, 1 mM EDTA, 2 mM DTT) and eluted with Wash Buffer plus 20 mM maltose. Subsequently, the WT protein was loaded onto a 5/50 GL MonoQ (Pharmacia Biotech - GE Life Sciences, Pittsburgh, US) at 50 mM NaCl and eluted at 300 mM NaCl, which concentrated the protein to 7mg/mL. Lastly, the protein was purified to homogeneity by SEC, using a 16/60 Superdex 200 or 10/30 Superdex 200 (GE Life Sciences, Pittsburgh, PA, US) column (SEC buffer: 25 mM Tris-HCl pH 8.0, 300 mM NaCl, 2 mM DTT) and concentrated to 5 mg/mL in a 50 kD Amicon Ultra-0.5 centrifugal concentrator (EMD Millipore, Billerica, MA, US). The amylose affinity elutions of the remaining MBP-SL fusion constructs were directly purified to homogeneity by SEC, using a 16/60 Superdex 200 or 10/30 Superdex 200 (GE Life Sciences, Pittsburgh, PA, US) column (SEC buffer: 25 mM Tris-HCl pH 8.0, 300 mM NaCl, 2 mM DTT).

I would like to thank Yue Li for her significant contribution to the expression and purification of the CCD, CCD-BARAD^{AFM}, CCD-BARAD^{AFM,TETRAD}, MBP-CCD-BARAD^{AFM,TETRAD}, MBP-OH+BARAD^{AFM}, MBP-OH+BARAD^{TETRAD}, MBP-OH+BARAD^{AFM,TETRAD}.

3.2.4. CD spectroscopy

Continuous scanning CD spectra were recorded from 190 to 240 nm at 4 °C in a 300 µL quartz cell (0.1 cm path length) on a Jasco J-815 spectropolarimeter equipped with a Peltier thermoelectric temperature control for each protein sample at concentrations of 3-20 µM in 5 mM Na phosphate dibasic, 5 mM Na phosphate monobasic pH 7.4, 100 mM $(NH_4)_2SO_4$. Secondary structure content was estimated by analyzing CD data using the SDP48 reference protein database and three programs, SELCON3, CDSSTR, and CONTIN, from the CDpro suite within the Jasco software[78, 125, 126]. For each spectrum, the average secondary structure was calculated by averaging estimated secondary structure content obtained from the three CD data analysis programs.

Variable temperature CD spectra were recorded at 222 nm at 1 °C intervals, with a ramp rate of 1 °C/min after dilution of protein samples to 2-20 µM in 5 mM Na phosphate dibasic, 5 mM Na phosphate monobasic pH 7.4, 100 mM $(NH_4)_2SO_4$. Spectra were recorded for BECN1 OH+BARAD and BECN1

CCD-BARAD proteins from 4-75 °C, and for BECN1 CCD proteins from 4-55 °C or 55-4 °C for reverse measurements. Data were analyzed using OriginPro 8 (OriginLab), where mean residue molar ellipticity was plotted against temperature, and the T_m was calculated by fitting the data to the Boltzmann or Double Boltzmann algorithm included in OriginPro 8.

3.2.5. ITC

BECN1 CCD, CCD^{TETRAD}, $CCD\text{-}BARAD^{AFM}$, and $CCD\text{-}BARAD^{AFM,TETRAD}$ protein samples were dialyzed against 50 mM HEPES base pH 8.0, 150 mM NaCl, and 2 mM β-mercaptoethanol (BME). ITC experiments were performed in triplicate at 20 °C using a Low Volume Gold Nano isothermal titration calorimeter (TA Instruments, New Castle, US). 65 μM $CCD\text{-}BARAD^{AFM}$, 100 μM $CCD\text{-}BARAD^{AFM,TETRAD}$, 350 μM CCD, and 100 μM CCD^{TETRAD} were separately titrated into dialysis buffer using 20 injections of 2.5 μL each. Data were analyzed with the NanoAnalyze Software (TA Instruments, New Castle, US), using a dimer dissociation model to calculate the dimer dissociation constants.

3.2.6. SAXS data collection and analysis

SAXS data were recorded at the BioCAT beamline (ID18) (Advanced Photon Source, Argonne, IL, US) on a Pilatus 3 1M detector at a sample to detector distance of 3.5 m at a wavelength of 1.03 Å. For all constructs, SEC was performed in tandem with SAXS to ensure that all SAXS data were collected from a homogenous sample eluting from the SEC column (10/300 GL Superdex 200 Increase). SEC-SAXS data were recorded by exposing the column eluate to the X-ray beam for 1 second with a periodicity of 3 seconds. Scattering data were normalized to the incident X-ray beam intensity and scattering from buffer was subtracted using the BioCAT beamline pipeline based on the ATSAS suite of programs, prior to further analysis. Data were analyzed with the ATSAS program suite[147], including PRIMUS[148] to calculate Guinier extrapolations to estimate the R_g and Kratky plots to evaluate disorder within the samples. The P(r) function plotted from the Fourier inversion of the scattering intensity, I(q) using GNOM[149], was used to calculate the R_g and D_{max}, and to reconstruct *ab initio* envelopes by applying ten cycles in DAMMIF[85] with subsequent analysis by DAMSEL, DAMSUP, DAMAVER and DAMFILT[86] to compare and identify the most probable model, align all models to the most probable model, average aligned models and compute a probability map, and filter the averaged model.

The EOM, as encoded within EOM 2.1[150], was used to generate an ensemble of models that included the flexible and disordered elements for each protein fragment. For the CCD^{TETRAD}, we modeled the OH and its partner helix (BECN1 residues 175-193) as flexible using PDB ID: 5HHE, after removing OH and partner helix residues (ShortCCD), as a rigid body. For the $BARAD^{AFM}$ and $BARAD^{AFM,TETRAD}$, we modeled the flexible His_6 tag using PDB ID: 4DDP (OH+BARAD) as a rigid body. For the CCD-$BARAD^{AFM}$, we generated two-domain models using i) the ShortCCD and the OH+BARAD as rigid bodies, or ii) PDB ID: 5HHE and 4DDP, after removing OH residues (BARAD), as rigid bodies. For the CCD-$BARAD^{AFM,TETRAD}$, we generated two-domain models using i) the ShortCCD and the OH+BARAD as rigid bodies, or ii) the ShortCCD and BARAD as rigid bodies. Models were obtained using native-like chain options and P2 symmetry was imposed on CCD-containing fragments by fixing the original coordinates of the PDB ID: 5HHE or ShortCCD. The theoretical scattering data calculated from each model in the ensemble was fitted to the averaged scattering data using CRYSOL [151]. The EOM model that best agreed with the experimentally derived R_g and D_{max} and best fit the averaged scattering curves was then fitted into the molecular envelopes using SUPCOMB[87].

3.2.7. Autophagy assays

Cellular autophagy levels were quantified by monitoring cellular localization of GFP-tagged LC3 protein[152]. Each chamber of an 8-well culture slide (Nunc™ Lab-Tek™ Chambered Coverglass) was seeded with 3.0 x 105 MCF7 cells and cultured overnight in Dulbecco modified eagle medium (DMEM) with 10% fetal bovine serum (FBS) until 80-90% confluent. The cells were co-transfected with 500 ng total plasmid per well, comprising of 200 ng GFP-LC3 and 300 ng FLAG-BECN1 WT or mutant expression plasmids, using Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. After transfection and incubation at 37 °C for 24 hours, the cells were cultured in either rich (DMEM, 10% FBS, 2x essential amino acids and 2x nonessential amino acids) or starvation (Earle's balanced salt solution) media for 4 hours. 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 fluorescent microscope (Zeiss AxioObserver Z1) and 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. I would like to thank Shreya Mukhopadhyay for performing these autophagy assays.

Expression levels of WT FLAG-BECN1 and mutants were verified by Western blot using commercial mouse monoclonal anti-FLAG M2-peroxidase antibody (MilliporeSigma). As a loading control, Actin levels in MCF7 cell lysates were detected with mouse anti-Actin (MilliporeSigma).

3.3. Results

3.3.1. The BECN1 OH cannot simultaneously pack against the CCD homodimer partner helix and the BARAD

Superimposition of the human BECN1 OH of the OH+BARAD (PDB ID: 4DDP) and one chain of the CCD homodimer (PDB ID: 5HHE) reveals the OH cannot simultaneously be within the CCD and packed against the BARAD, as there are extensive steric conflicts between the BARAD and the partner helix of the CCD (Figure 3.1). Further, the same five OH residues; V250, M254, A257, L261, and L264, involved in the CCD homodimer interface also pack against the BARAD (Figure 3.1). Thus, the OH in the CCD and OH+BARAD crystal structures represent mutually exclusive conformations of BECN1.

Approximately 1359 $Å^2$ of surface area are buried upon OH packing against the BARAD, significantly more than the 937 $Å^2$ buried when the OH packs within the CCD. Within the CCD, OH residues are only involved in paired interactions. However, OH residues pack against the first β-sheet and third α-helix of the BARAD (Figure 3.1) resulting in more extensive packing.

Each of the OH interface residues is largely conserved from yeast to human (Figure 3.2A). The position equivalent to human Met^{254} is least well conserved, although all examined vertebrates have an M or V at equivalent positions. In comparison to the OH, residues of the partner BECN1 CCD that pack against the OH are not as well conserved (Figure 3.2B). Based on this preliminary analysis, we hypothesized that simultaneous alanine mutagenesis of four of the common OH interface residues; V250A, M254A, L261A, and L264A (Tetrad mutant) would differently impact structure and stability of the CCD and BARAD. Structural and thermodynamic analyses of these differences would allow us to

determine whether the OH is part of the CCD, packs against the BARAD, or transitions between the two states.

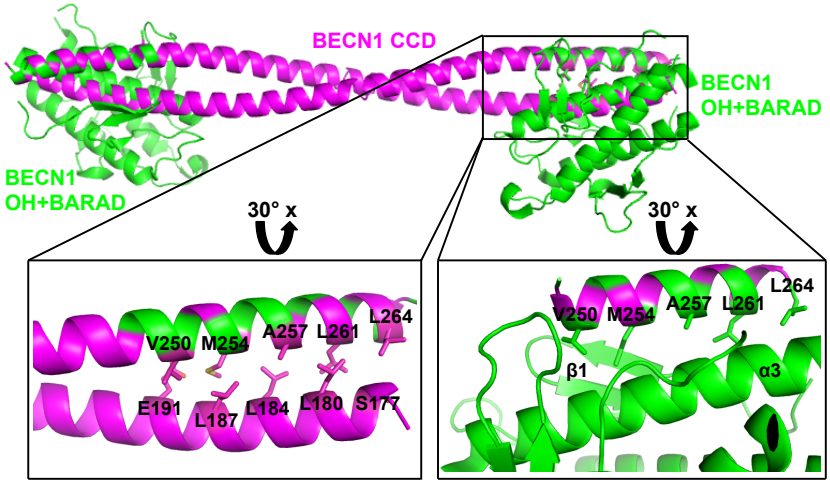

Figure 3.1. The OH has two different packing states.
The BECN1 CCD (magenta) and BARAD (green) are shown in ribbon. The OHs in the CCD and BARAD structures are superimposed. Boxes indicate regions that are rotated 30° about the x-axis and enlarged to show the mutually exclusive packing arrangement of 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. This and other molecular figures were made using PyMOL (v.1.5.0.2; Schrödinger)[144].

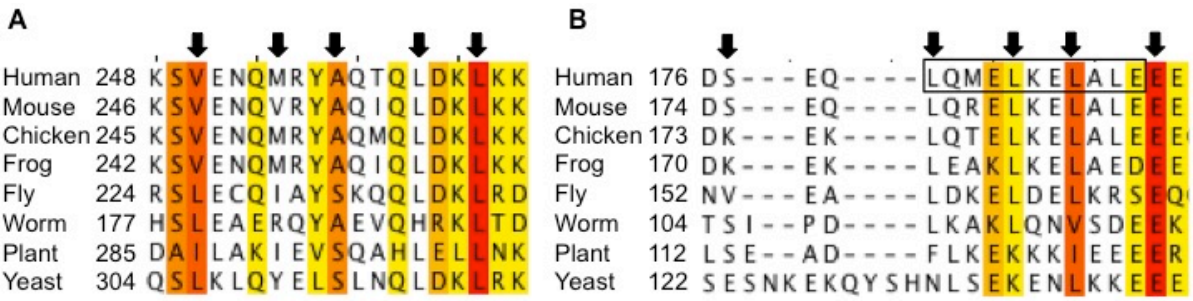

Figure 3.2. Sequence alignment of the OH and partner helix in BECN1 orthologs.
Increasing background color intensity corresponds to increasing residue conservation: with red corresponding to invariant residues of the A) Overlap helix and B) Partner helix with NES boxed. Interface residues are indicated with black arrows.

3.3.2. Expression and purification results of BECN1 protein fragments

In order to obtain enough pure and homogenous WT and OH mutant BECN1 protein fragments for usage in the various biophysical and structural analyses in this study, we attempted to express untagged and MBP fusion protein constructs of the BECN1 CCD, CCD^{TETRAD}, OH+BARAD, OH+BARAD^{TETRAD}, CCD-BARAD, and CCD-BARAD^{TETRAD} (Table 3.1). Representative purification results are reported below, which highlight the difficulty in obtaining high quality, homogenous proteins containing the WT BARAD.

For MBP-BECN1 CCD-BARAD, MBP-SL-BECN1 CCD-BARAD was expressed in *E. coli* Arctic Express cells, and soluble protein in the cell lysate from a 6 L expression culture was purified to homogeneity by affinity chromatography using a 3 mL amylose resin gravity column, followed by anion exchange chromatography using a 1mL 5/50 GL MonoQ column (load and wash at 50 mM NaCl, step elution at 300 mM NaCl), and finally by SEC using a 10/30 HR Superdex 200 column. The single peak on SEC and single band on SDS-PAGE indicate the MBP-SL-BECN1 CCD-BARAD was pure and homogenous (Figure 3.3). The apparent molecular weight of MBP-SL-BECN1 CCD-BARAD calculated based on the elution volume from SEC is 245.4 kD, which is 1.7-times greater than the theoretical molecular weight of 147.2 kD for the MBP-SL-BECN1 CCD-BARAD dimer calculated from the amino acid sequence. This is consistent with the MBP-SL-BECN1 CCD-BARAD forming rod-shaped dimer. The final yield of the purified MBP-SL-BECN1 CCD-BARAD obtained from one liter of bacterial culture was 0.3 mg. The MBP tag; however, could not be cleaved from the fusion protein without causing severe precipitation.

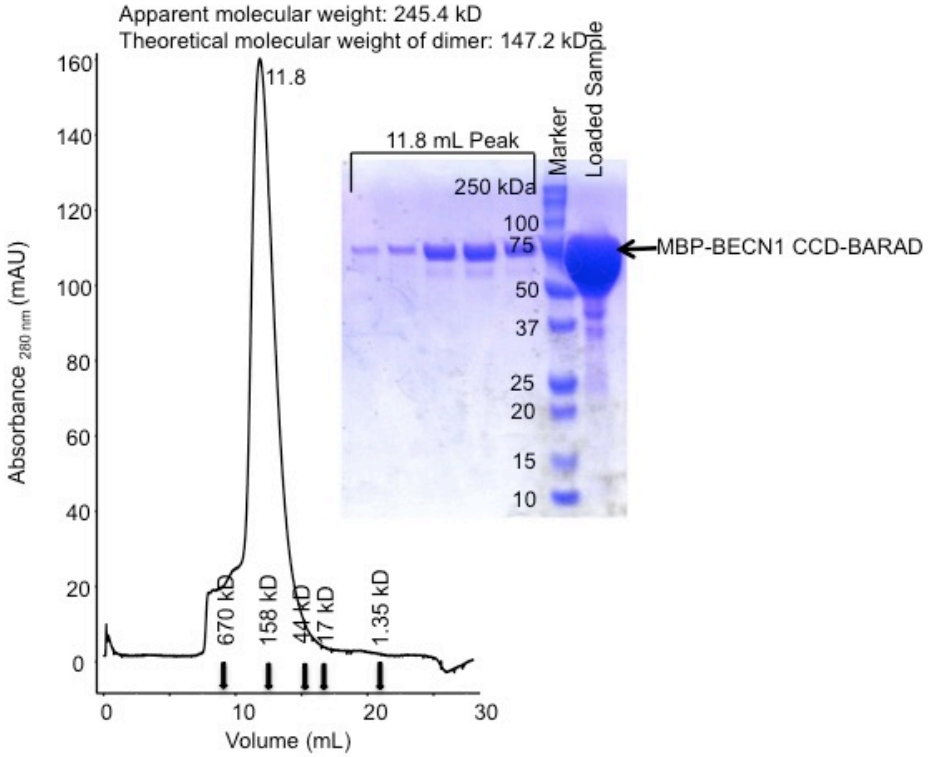

Figure 3.3. Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-SL-BECN1 CCD-BARAD.
The elution positions of SEC standards are indicated by arrows.

For MBP-BECN1 CCD-BARAD^{TETRAD}, MBP-SL-BECN1 CCD-BARAD^{TETRAD} was expressed in *E. coli* Arctic Express cells, and soluble protein in the cell lysate from a 6 L expression culture was purified to homogeneity by affinity chromatography using a 4 mL amylose resin gravity column followed by SEC using a 10/30 HR Superdex 200 column. The single peak on SEC and single band on SDS-PAGE indicate the MBP-SL-BECN1 CCD-BARAD^{AFM} was pure and homogenous (Figure 3.4). The apparent molecular weight of MBP-SL-BECN1 CCD-BARAD^{AFM} calculated based on the elution volume from SEC is 114.0 kD, which is similar to the theoretical molecular weight of 146.8 kD for the MBP-SL-BECN1 CCD-BARAD^{TETRAD} dimer calculated from the amino acid sequence. This is consistent with the MBP-SL-BECN1 CCD-BARAD^{TETRAD} forming dimer. The final yield of the purified MBP-SL-BECN1 CCD-BARAD^{TETRAD} obtained from one liter of bacterial culture was 4 mg. The MBP tag; however, could not be cleaved from the fusion protein without causing severe precipitation.

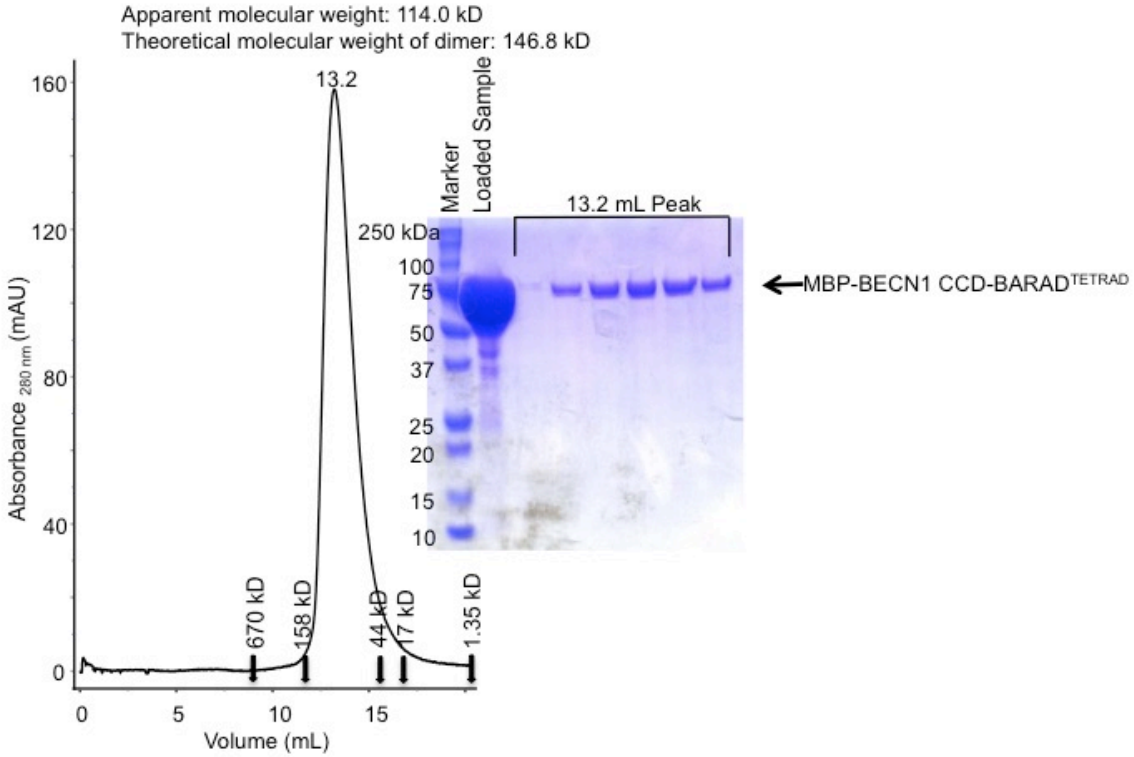

Figure 3.4. Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-SL-BECN1 CCD-BARAD^{TETRAD}.
The elution positions of SEC standards are indicated by arrows.

For BECN1 OH+BARAD, His_6-BECN1 OH+BARAD was expressed in *E. coli* Arctic Express cells. Soluble protein in the cell lysate from a 2 L expression culture was purified by affinity chromatography using a 5 mL HisTrap column, followed by anion exchange chromatography using a 1mL 5/50 GL MonoQ column, and finally SEC using a 10/30 Superdex 200 in tandem with a 10/30 Superdex 75 column. Most of the BECN1 OH+BARAD and lingering contaminants eluted from SEC in the void volume (16.4 and 18.3 mL peaks) (Figure 3.5), which is indicative of protein aggregation. A small peak with an elution volume of 26.1 mL containing BECN1 OH+BARAD with an apparent molecular weight of 44.7 kD, which is 1.7-times larger than the theoretical molecular weight of 26.1 kD for the BECN1 OH+BARAD monomer calculated from the amino acid sequence, was obtained. However, the 44.7 kD apparent molecular weight BECN1 OH+BARAD could not be concentrated beyond 1 mg/mL without precipitating. The final yield of the purified OH+BARAD obtained from one liter of bacterial culture was 0.5 mg. Therefore, pure and homogenous protein of sufficient quantity and concentrations could not be obtained to permit the various biophysical and structural analyses in this study.

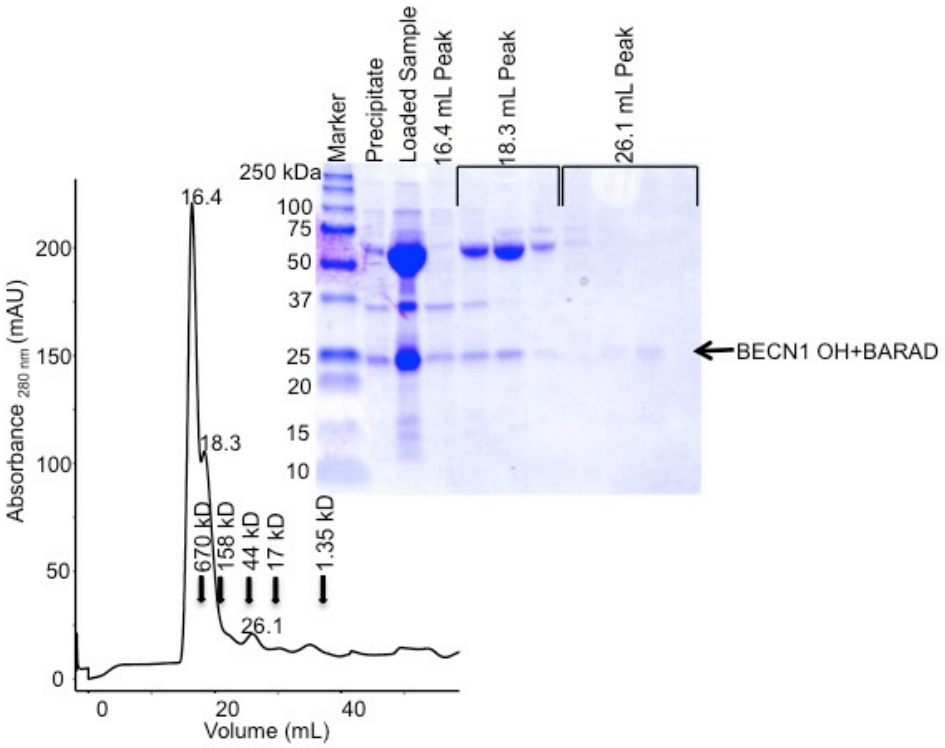

Figure 3.5. Size exclusion chromatogram and the corresponding SDS-PAGE of BECN1 OH+BARAD. The elution positions of SEC standards are indicated by arrows.

For MBP-BECN1 OH+BARAD, MBP-SL-BECN1 OH+BARAD was expressed in *E. coli* Arctic Express cells, and soluble protein in the cell lysate from a 1 L expression culture was purified by affinity chromatography using a 2 mL amylose resin gravity column, followed by anion exchange chromatography using a 1mL 5/50 GL MonoQ column, and finally by SEC using a 10/30 HR Superdex 200 column. Most of the MBP-BECN1 OH+BARAD eluted from SEC in the void volume (8.2 mL peak) (Figure 3.6), which is indicative of protein aggregation. The fusion protein was intended for comparison with the MBP-tagged WT and OH mutant BECN1 protein fragments. However, MBP-SL-BECN1 OH+BARAD pure and homogenous protein of sufficient quantity and concentrations could not be obtained to permit the various biophysical and structural analyses in this study.

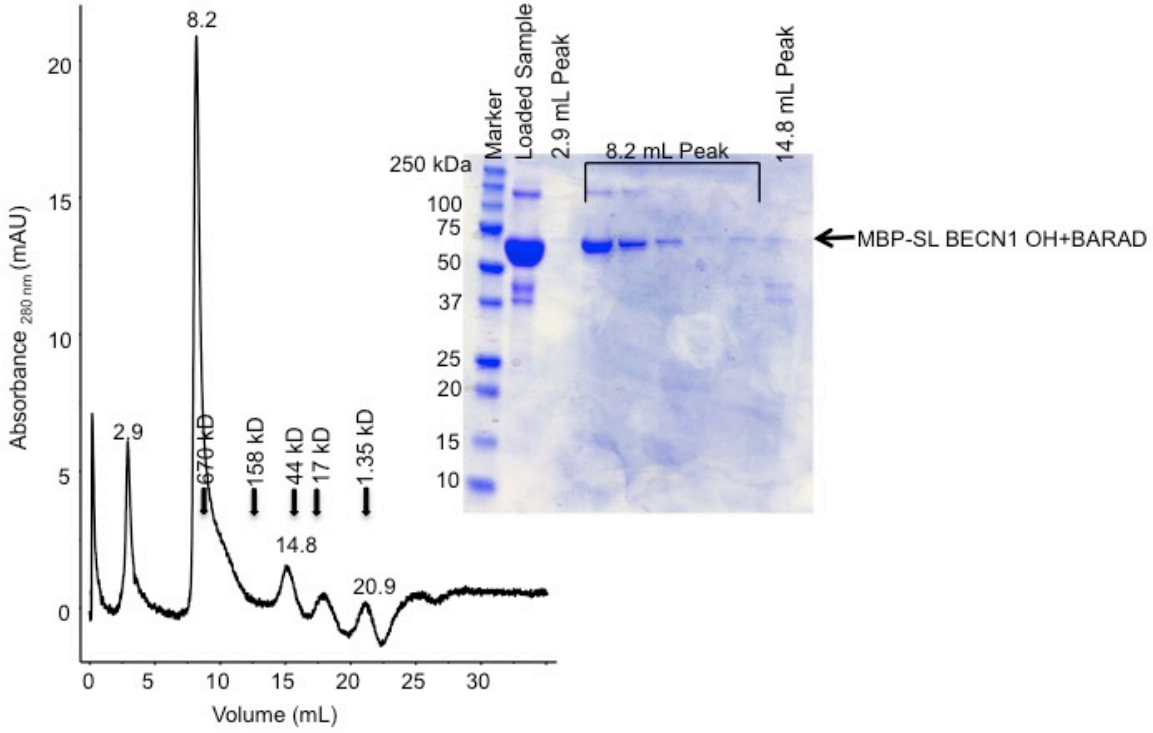

Figure 3.6. Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-SL-BECN1 OH+BARAD.
The elution positions of SEC standards are indicated by arrows.

For MBP-BECN1 OH+BARAD^{TETRAD}, MBP-SL-BECN1 OH+BARAD^{TETRAD} was expressed in *E. coli* Arctic Express cells, and soluble protein in the cell lysate from a 2.5 L expression culture was purified to homogeneity by affinity chromatography using a 10 mL amylose resin gravity column followed by SEC

using a 16/60 HiLoad Superdex 200 column. The single peak on SEC and single band on SDS-PAGE indicates the MBP-SL-BECN1 OH+BARAD^{TETRAD} was pure and homogenous (Figure 3.7). The apparent molecular weight of MBP-SL-BECN1 OH+BARAD^{TETRAD} calculated based on the elution volume from SEC is 37.2 kD, which is 1.7-times smaller than the theoretical molecular weight of 62.8 kD for the MBP-SL-BECN1 OH+BARAD^{TETRAD} monomer calculated from the amino acid sequence. This is consistent with the MBP-SL-BECN1 OH+BARAD^{TETRAD} forming a monomer. The final yield of the purified MBP-SL-BECN1 OH+BARAD^{TETRAD} obtained from one liter of bacterial culture was 4.5 mg. The fusion protein was included for comparison with the MBP-tagged WT and OH mutant BECN1 protein fragments.

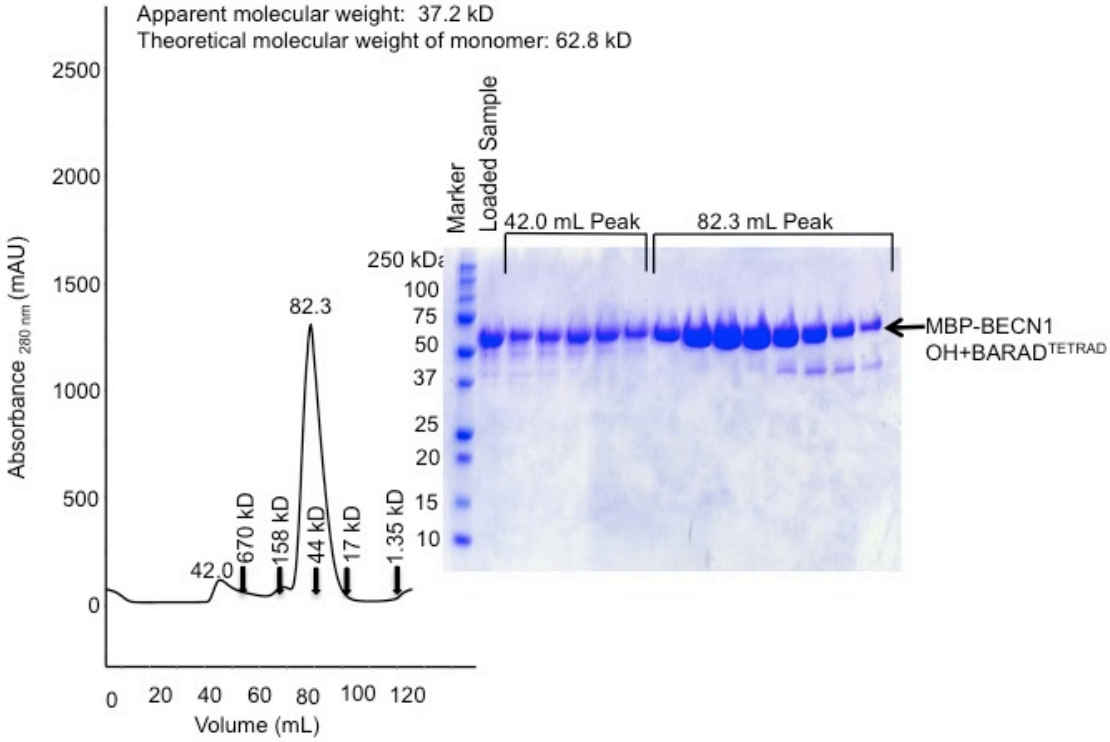

Figure 3.7. Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-SL-BECN1 OH+BARAD^{TETRAD}.
The elution positions of SEC standards are indicated by arrows.

For MBP-BECN1 CCD, MBP-SL-BECN1 CCD was expressed in *E. coli* BL21(DE3)pLysS cells, and soluble protein in the cell lysate from a 1 L expression culture was purified to homogeneity by affinity chromatography using a 5 mL amylose resin gravity column followed by SEC using a 16/60 HiLoad Superdex 200 column. The single peak on SEC and single band on SDS-PAGE indicates the MBP-SL-BECN1 CCD was pure and homogenous (Figure 3.8). The apparent molecular weight of the MBP-SL-

BECN1 CCD calculated based on the elution volume from SEC is 110.6 kD, which is similar to the theoretical molecular weight of 104.3 kD for the MBP-SL-BECN1 CCD dimer calculated from the amino acid sequence. This is consistent with the MBP-SL-BECN1 CCD forming a dimer. The final yield of the purified MBP-SL-BECN1 CCD obtained from one liter of bacterial culture was 8 mg. The fusion protein was intended for comparison with the MBP-tagged WT and OH mutant BECN1 protein fragments.

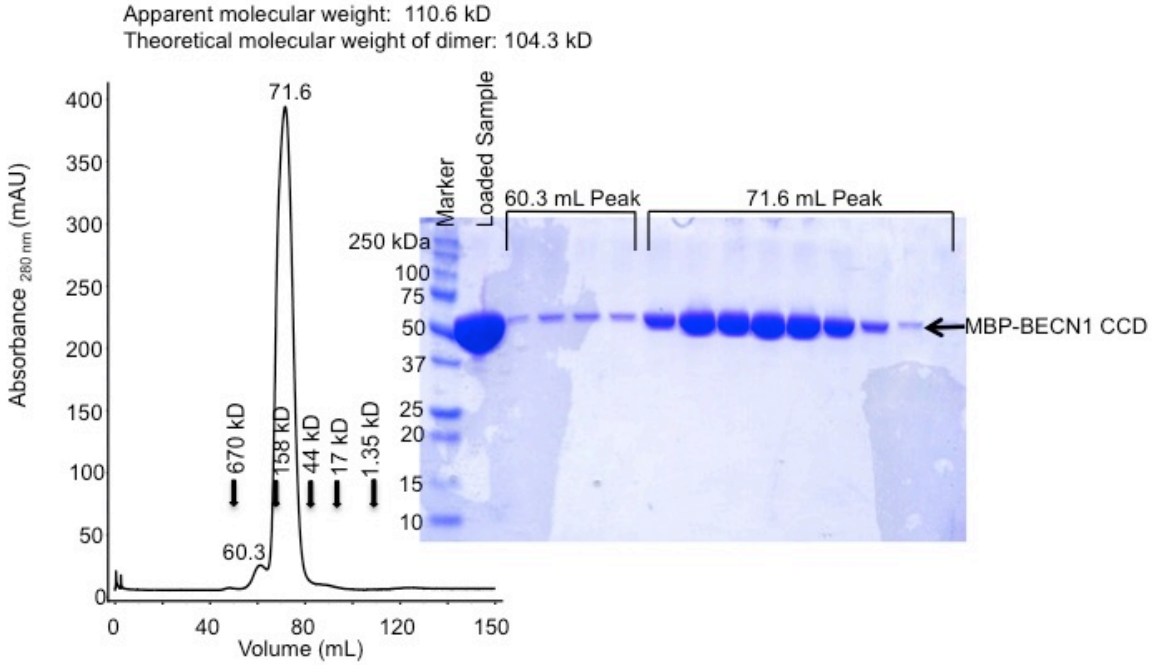

Figure 3.8. Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-SL-BECN1 CCD. The elution positions of SEC standards are indicated by arrows.

For BECN1 CCD, MBP-BECN1 CCD was expressed and purified in a manner similar to that previously reported[33], except that an anion exchange chromatography step was added to remove contaminating MBP after off-column cleavage, which afforded improved cleavage and yield compared to on-column cleavage. The MBP-BECN1 CCD was expressed in *E. coli* BL21(DE3)·pLysS cells, and soluble fusion protein in the cell lysate from a 6 L expression culture was purified to homogeneity by affinity chromatography using a 10 mL amylose resin gravity column, followed by off-column cleavage using 16 mg GST-TEV protease, to remove the MPB tag. The cleaved protein was further purified to homogeneity by anion exchange chromatography using a 1mL 5/50 GL MonoQ column, and finally by SEC using a 10/30 HR Superdex 200 column. The single peak on SEC and single band on SDS-PAGE

indicate the CCD was pure and homogenous (Figure 3.9). The apparent molecular weight of CCD calculated based on the elution volume from SEC is 39.8 kD, which is 1.8-times greater than the theoretical molecular weight of 22.3 kD for the CCD dimer calculated from the amino acid sequence. This is consistent with the CCD forming rod-shaped dimer. The final yield of the purified CCD obtained from one liter of bacterial culture was 0.83 mg.

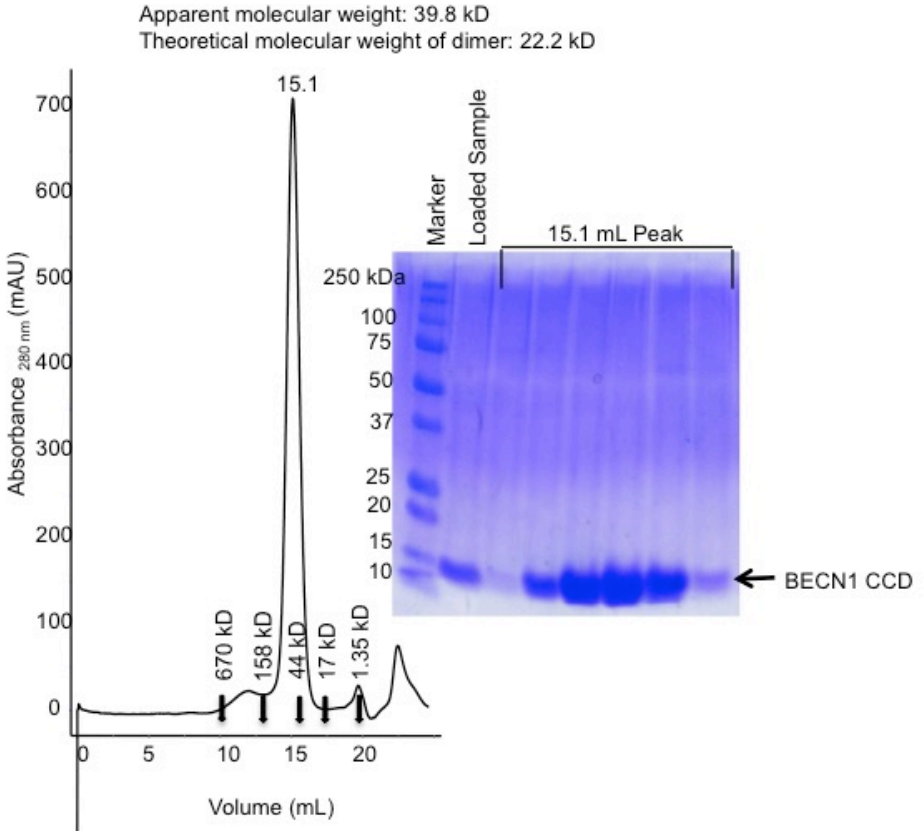

Figure 3.9. Size exclusion chromatogram and the corresponding SDS-PAGE of BECN1 CCD. The elution positions of SEC standards are indicated by arrows.

For MBP- BECN1 CCD^{TETRAD}, MBP-SL-BECN1 CCD^{TETRAD} was expressed in *E. coli* BL21(DE3)pLysS cells, and soluble protein in the cell lysate from a 2 L expression culture was purified to homogeneity by affinity chromatography using a 2 mL amylose resin gravity column followed by anion exchange chromatography using a 5/50 GL MonoQ column (gradient elution from 50-500 mM NaCl over 100 mL) where MBP-SL-BECN1 CCD^{TETRAD} eluted from MonoQ in two peaks, one at 150mM NaCl and another at 200mM NaCl. Each MonoQ elution peak was loaded separately onto SEC using a 10/30 HE Superdex 200 column. Both MonoQ elution peaks eluted from SEC as a single peak and yielded a single

band on SDS-PAGE, indicating that the MBP-SL-BECN1 CCD^{TETRAD} was pure and homogenous (Figure 3.10). The apparent molecular weight of MBP-SL-BECN1 CCD^{TETRAD} calculated based on the elution volume from SEC is 96.1 kD, which is similar to the theoretical molecular weight of 104.0 kD for the MBP-SL-BECN1 CCD^{TETRAD} dimer calculated from the amino acid sequence. This is consistent with the MBP-SL-BECN1 CCD^{TETRAD} forming a dimer. The final yield of the purified MBP-SL-BECN1 CCD obtained from one liter of bacterial culture was 0.2 mg. The fusion protein was intended for comparison with the MBP-tagged WT and OH mutant BECN1 protein fragments.

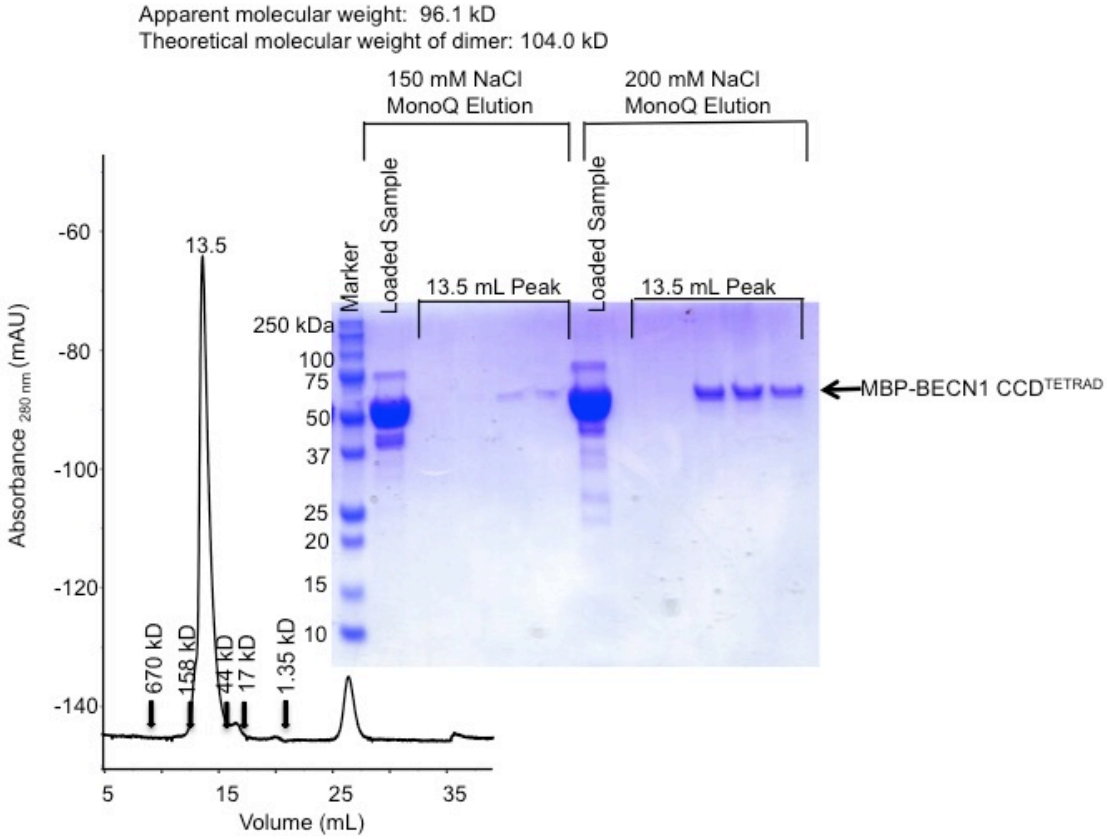

Figure 3.10. Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-SL-BECN1 CCD^{TETRAD}.
The elution positions of SEC standards are indicated by arrows.

For BECN1 CCD^{TETRAD}, MBP-BECN1 CCD^{TETRAD} was expressed in *E. coli* BL21(DE3)·pLysS cells, and soluble fusion protein in the cell lysate from a 2 L expression culture was purified to homogeneity by affinity chromatography using a 5 mL amylose resin gravity column with on-column cleavage using 6.6 mg GST-TEV protease, to remove the MBP tag. The cleaved protein was further

purified to homogeneity by anion exchange chromatography using a 1mL 5/50 GL MonoQ column and finally by SEC using a 10/30 HR Superdex 200 column. The single peak on SEC and single band on SDS-PAGE indicate the CCD^{TETRAD} was pure and homogenous (Figure 3.11). The apparent molecular weight of CCD^{TETRAD} calculated based on the elution volume from SEC is 55.4 kD, which is 2.5-times greater than the theoretical molecular weight of 21.8 kD for the CCD^{TETRAD} dimer calculated from the amino acid sequence. This is consistent with the CCD^{TETRAD} forming a rod-shaped dimer. The final yield of the purified CCD^{TETRAD} obtained from one liter of bacterial culture was 1 mg.

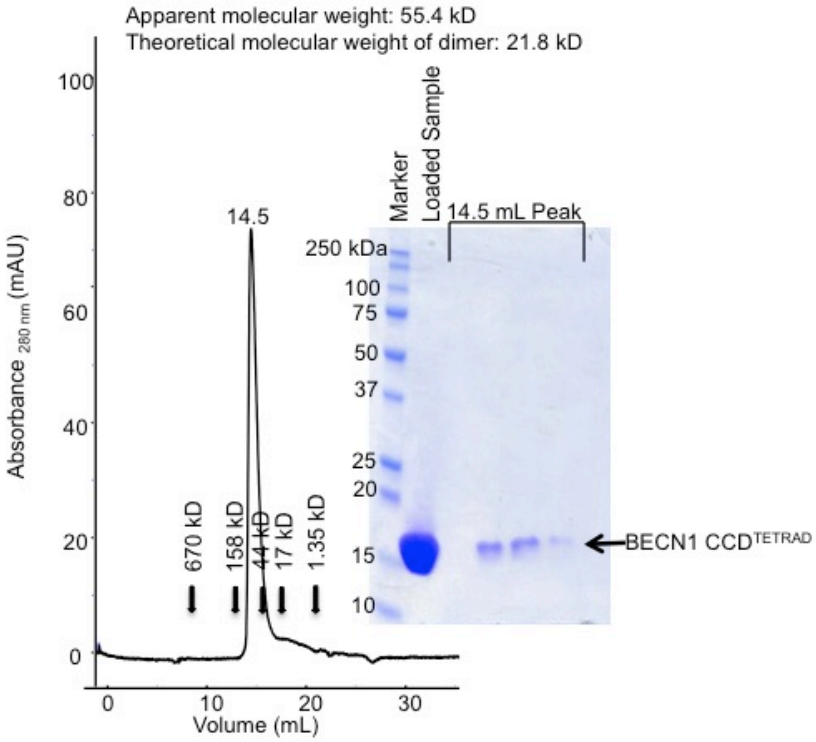

Figure 3.11. Size exclusion chromatogram and the corresponding SDS-PAGE of BECN1 CCD^{TETRAD}. The elution positions of SEC standards are indicated by arrows.

3.3.3. MBP affects homodimerization of the BECN1 CCD

Since the MBP tag could not be removed from the MBP-BECN1 CCD-BARAD, we sought to characterize homodimerization of the WT and OH mutant CCD-containing constructs as MBP-fusion proteins. The presence of MBP appears to affect homodimerization of BECN1 via the CCD. Dissociation was not detected at both 10 and 20 °C for the MBP-BECN1 CCD (Table 3.2), unlike the untagged CCD, which has a K_d of ~48 μM[18] at 20 °C. This could indicate either that the MBP-BECN1 CCD fails to form a homodimer, or that the K_d of the MBP-BECN1 CCD homodimer is tighter than the ~10 nM detection limit.

The presence of MBP also affects the homodimerization of the CCD^{TETRAD} wherein the K_d of the MBP-BECN1 CCD^{TETRAD} is ~28 μM at 10 °C and ~113 μM at 20 °C, while that of the untagged CCD^{TETRAD} is ~27 μM at both 10 and 20 °C (Table 3.2).

Table 3.2. Thermodynamics of dissociation of MBP-BECN1 fusion proteins and BECN1 proteins at variable temperatures.

BECN1 Fragment	Temperature (°C)	K_d (μM)	ΔH (kJ/mol)	ΔS (J/K·mol)
MBP-CCD	10 & 20	No dissociation detected		
MBP-CCD^{TETRAD}	10	28	46	248
MBP-CCD^{TETRAD}	20	113	137	541
CCD^{TETRAD}	10	26	35	210
CCD^{TETRAD}	20	29	129	526

Performed in 25 mM Tris HCl pH 8.0, 300mM NaCl, 2mM BME. Replicates were not performed.

3.3.4. Secondary structure and melting temperature analyses of the fusion proteins are complicated by the presence of MBP

Since the MBP tag could not be removed from the MBP-BECN1 CCD-BARAD, we also sought to characterize the structure and thermal stability of the WT and OH mutant BECN1 fragments as MBP-fusion proteins. The secondary structure content and melting temperatures of MBP-BECN1 fusion proteins were estimated from CD spectra analyses of cleaved MBP-LL, MBP-LL-CCD, MBP-SL-CCD, and MBP-SL-BECN1 OH+$BARAD^{AFM}$. Since maltose has been reported to stabilize MBP[153], 50 mM maltose was used to increase the T_m of MBP and permit separation of the MBP melting curve from that of the BECN1 OH+$BARAD^{AFM}$.

CD spectra analysis for the MBP-LL after cleavage indicates that the MBP-LL contains ~129 α-helical residues, ~94 β-strand residues, and ~169 coil residues in the absence of maltose and ~277 α-helical residues, ~26 β-strand residues, and ~90 coil residues (Figure 3.12A, Table 3.3). Therefore, maltose increases helical content by ~98 residues, decreases strand content by 68, and decreases coil content by 79 residues. The cooperative, although irreversible, melting curves indicate that the T_m of ~57 °C of MBP increases to ~71 °C in the presence of 50 mM maltose (Figure 3.12B, Table 3.3), which is indicative of maltose stabilizing MBP.

A comparison of the secondary structure content estimated from CD spectra analysis for the MBP-SL-BECN1 CCD and MBP-LL-BECN1 CCD indicates that the MBP-SL-BECN1 CCD contains ~249

α-helical residues, ~47 β-strand residues, and ~168 coil residues while the MBP-LL-BECN1 CCD contains ~202 α-helical residues, ~68 β-strand residues, and ~220 coil residues (Figure 3.12C, Table 3.3). Therefore, despite containing 18 fewer total residues, the MBP short linker increases α-helical content by ~52 residues while reducing β-strand content by ~23 residues and coil content by a concomitant ~47 residues as compared to the MBP long linker. The MBP-SL-BECN1 CCD in 50 mM maltose contains ~241 α-helical residues, ~43 β-strand residues, and ~182 coil residues (Figure 3.12C, Table 3.3); therefore, maltose does not significantly alter the secondary structure of MBP-SL-BECN1 CCD.

Although the MBP-LL-BECN1 CCD contains ~202 α-helical residues, ~68 β-strand residues, and ~220 coil residues (Table 3.3), CD spectra analysis for the MBP-LL and BECN1 after cleavage indicates that the MBP-LL contains ~129 α-helical residues, ~94 β-strand residues, and ~169 coil residues (Figure 3.12A, Table 3.3) while the BECN1 CCD contains ~83 α-helical residues, ~2 β-strand residues, and ~8 coil residues (Figure 3.31B, Table 3.3). This indicates that the MBP-LL-BECN1 CCD fusion protein contains ~10 fewer α-helical, ~28 fewer β-strand, and ~43 more coil residues than the sum of each type of secondary structure content in the cleaved MBP-LL and BECN1 CCD proteins.

The MBP-SL-BECN1 CCD displays a cooperative melting curve from 5-55 °C that indicates a T_m of ~31 °C in the presence of 50 mM maltose (Figure 3.12D, Table 3.3), which corresponds to that of the cleaved CCD (T_m = 33 °C) (Table 3.5 and 3.8). CD data were only collected up to 55 °C; therefore, the MBP transition that begins at 65 °C in the presence of 50 mM maltose (Figure 3.12B) is not observed. The MBP is therefore likely still folded at 55 °C based on the lingering ellipticity observed at 222 nm (Figure 3.12D).

A comparison of the secondary structure content estimated from CD spectra analysis for the MBP-SL-BECN1 OH+BARAD^{AFM} indicates that the MBP-SL-BECN1 OH+BARAD^{AFM} contains ~215 α-helical residues, ~79 β-strand residues, and ~287 coil residues in the absence of 50 mM maltose but contains ~ 300 α-helical residues, ~ 39 β-strand residues, and ~ 248 coil residues in the presence of 50 mM maltose (Figure 3.12E, Table 3.3).

As expected for a two-domain protein, the cooperative melting curve for the MBP-SL-BECN1 OH+BARAD^{AFM} contains two well separated transitions (Figure 3.12F), indicating the presence of two

independently folded domains. The higher T_m of ~71°C corresponds well with that of MBP in the presence of maltose (T_m = 71 °C); while the lower T_m of ~46°C corresponds to that of the OH+BARAD^{AFM} (T_m = 48 °C) (Table 3.3 and 3.8).

Table 3.3. CD analysis: Estimated secondary structure content and melting temperatures of MBP-BECN1 fusion proteins.

Protein	# residues	Average estimated secondary structure content (%)			Melting temperature (°C)
		Helix	Strand	Coil	Forward
MBP-LL	393	32.7 ± 5.4	23.9 ± 3.1	42.9 ± 2.7	57
MBP-LL + Maltose	393	70.6 ± 6.8	6.7 ± 2.9	23.0 ± 9.0	71
MBP-LL-BECN1 CCD	484	41.7 ± 7.4	14.0 ± 6.6	45.5 ± 8.9	Not collected
MBP-SL-BECN1 CCD	466	53.4 ± 6.9	10.0 ± 5.2	36.1 ±6.8	Not collected
MBP-SL-BECN1 CCD + Maltose	466	51.7 ± 1.7	9.3 ± 7.4	39.1 ± 9.1	31
MBP-SL-BECN1 OH+BARAD^{AFM}	578	37.2 ± 1.9	13.6 ± 1.4	49.6 ± 1.4	Not collected
MBP-SL-BECN1 OH+BARAD^{AFM} + Maltose	578	51.9 ± 3.3	6.8 ± 1.5	42.9 ± 4.9	46 & 71

Proteins precipitated; therefore, reverse T_m could not be determined.

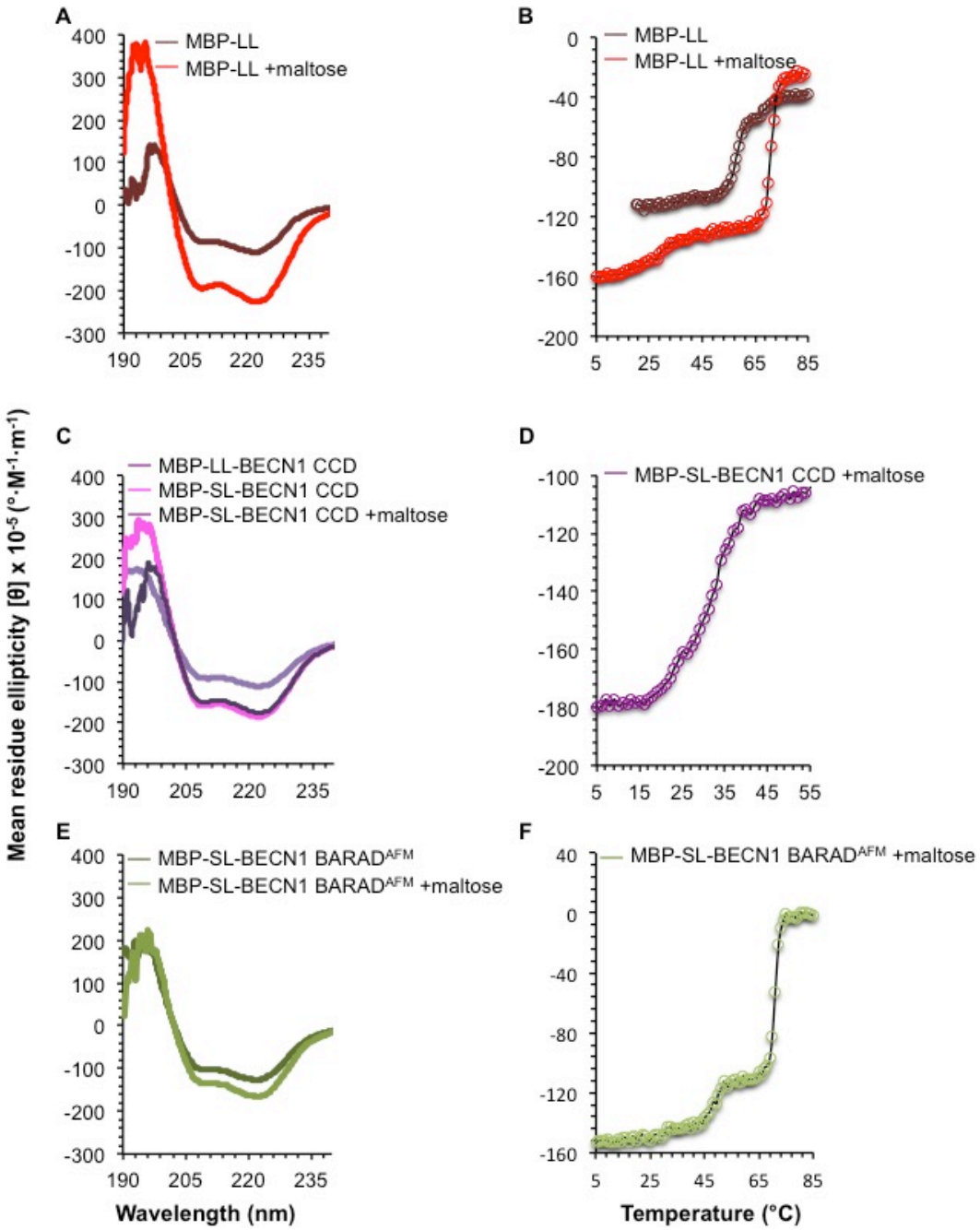

Figure 3.12. MBP-BECN1 fusion protein CD spectra and melting curves.
A) MBP (dark red) and MBP plus maltose (red). B) MBP (dark red) and MBP plus maltose (red) melting curve. C) MBP-LL-BECN1 CCD (light purple), MBP-SL-BECN1 CCD (magenta), MBP-SL-BECN1 CCD plus maltose (dark purple). D) MBP-SL-BECN1 CCD plus maltose (dark purple) melting curve. E) MBP-SL-BECN1 BARAD^{AFM} (dark green) and MBP-SL-BECN1 BARAD^{AFM} plus maltose (light green). F) MBP-SL-BECN1 BARAD^{AFM} plus maltose (light green) melting curve.

3.3.5. Tertiary structure analyses of the fusion proteins are complicated by the presence of MBP

We used SEC-SAXS to investigate the tertiary and quaternary structure of MBP-LL, MBP-SL-BECN1 CCD, MBP-SL-BECN1 CCD^{TETRAD}, and MBP-SL-BECN1 OH+BARAD^{TETRAD} (Figures 3.13-3.16, Table 3.4).

Table 3.4. Summary of SEC-SAXS data for MBP, MBP-BECN1 CCD, and MBP-BECN1 BARAD proteins.

Protein	MW_T (kD)	Guinier	P(r)	
		R_g (Å)	R_g (Å)	D_{max} (Å)
MBP-LL	43.1 (monomer)	24	25	86
MBP-SL-BECN1 CCD	104.4 (dimer)	62	69	245
MBP-SL-BECN1 CCD^{TETRAD}	104.0 (dimer)	40	43	165
MBP-SL-BECN1 OH+BARAD	Insufficient protein available for data collection			
MBP-SL-BECN1 OH+BARAD^{TETRAD}	64.6 (monomer)	39	42	200

For MBP-LL, the R_g estimated from the Guinier plot and P(r) distribution is ~25 Å, (Figure 3.13A and B, Table 3.4) and the shape of the P(r) distribution curves and the Kratky plots indicates that it is a well-folded, globular protein (Figure 3.13B and C).

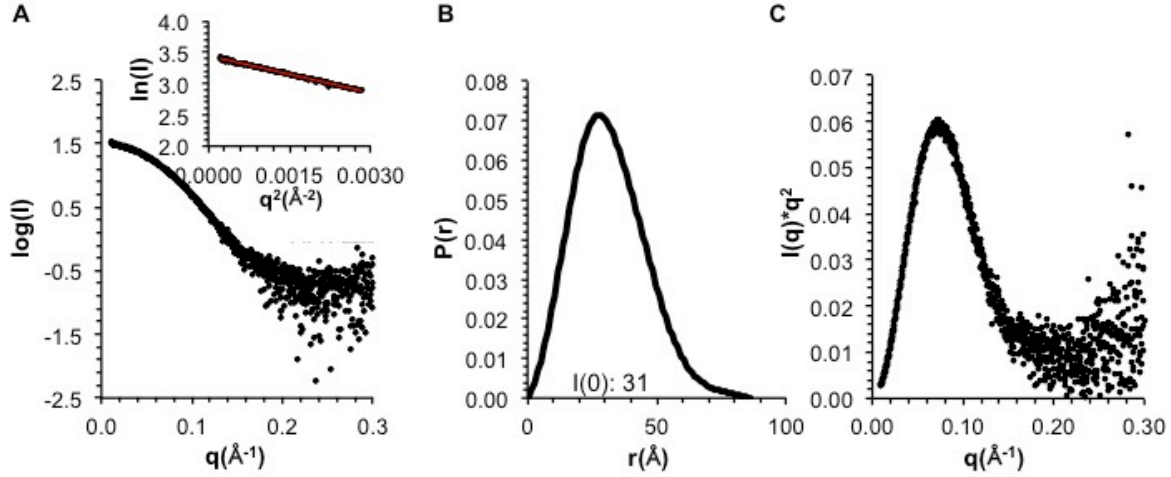

Figure 3.13. SAXS analysis of MBP-LL.
A) Guinier plot. B) P(r) pairwise distribution. C) Kratky plot.

For MBP-SL-BECN1 CCD, the R_g estimated from the Guinier plot and P(r) distribution is ~66 Å, (Figure 3.14A and B, Table 3.4) and the shape of the P(r) distribution curves and the Kratky plots indicates that it is a well-folded, dumbbell-shaped protein (Figure 3.14B and C).

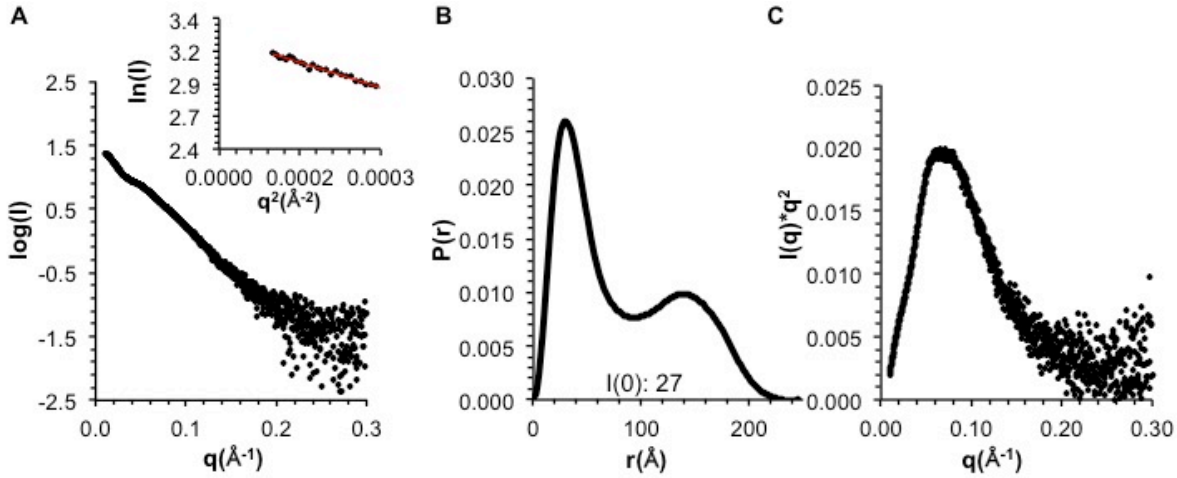

Figure 3.14. SAXS analysis of MBP-SL-BECN1 CCD.
A) Guinier plot. B) P(r) pairwise distribution. C) Kratky plot.

For the MBP-SL-BECN1 CCD^{TETRAD}, the R_g estimated from the Guinier plot and P(r) distribution is ~42 Å, (Figure 3.15A and B, Table 3.4) and the shape of the P(r) distribution curves and the Kratky plots indicates that it is a well-folded, rod-shaped protein (Figure 3.15B and C). Comparison of the R_g and D_{max} of the MBP-SL-BECN1 CCD and MBP-SL-BECN1 CCD^{TETRAD} indicates that the MBP-SL-BECN1 CCD^{TETRAD} is more compact than the WT. This is unexpected since the ITC data that indicates that the MBP-SL-BECN1 CCD^{TETRAD} homodimerizes more weakly than the MBP-SL-BECN1 CCD at 20 °C (Table 3.2), which suggests that the MBP-SL-BECN1 CCD^{TETRAD} would be more flexible than the WT. As a more flexible protein, the MBP-SL-BECN1 CCD^{TETRAD} would be expected to have a larger R_g and D_{max} than the WT. However, one possible explanation is that the MBP is positioned around the flexible regions of the MBP-SL-BECN1 CCD^{TETRAD} in such a manner that its overall tertiary structure is more compact.

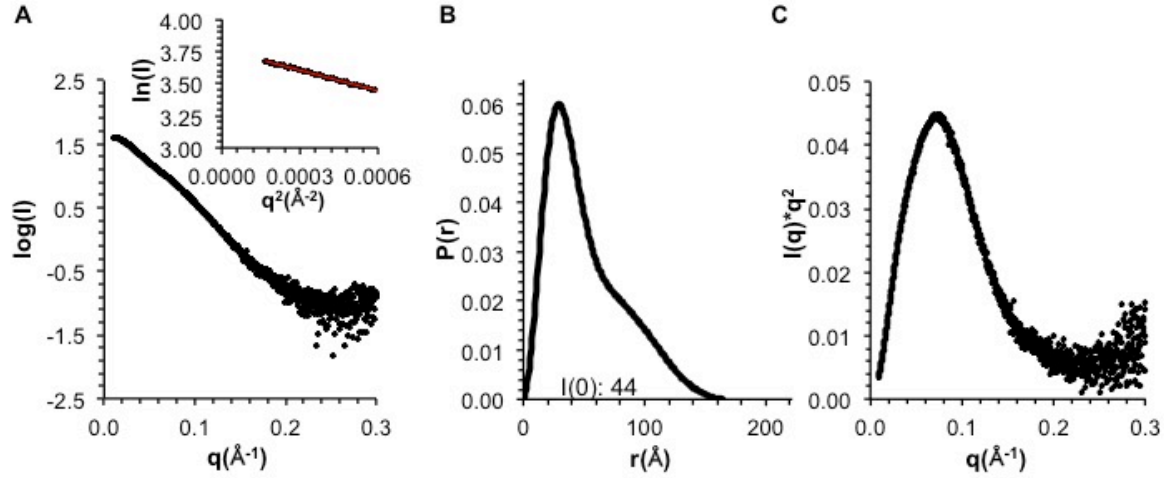

Figure 3.15. SAXS analysis of MBP-SL-BECN1 CCD^{TETRAD}.
A) Guinier plot. B) P(r) pairwise distribution. C) Kratky plot.

For the MBP-SL-OH+BARAD^{TETRAD}, the R_g estimated from the Guinier plot and P(r) distribution is ~41 Å, (Figure 3.16A and B, Table 3.4) and the shape of the P(r) distribution curves and the Kratky plots indicates that it is a well-folded, globular protein (Figure 3.17B and C). This fusion protein fragment was intended for comparison to the MBP-SL-OH+BARAD in order to assess the effect of the OH mutant on tertiary structure of the BARAD; however, attempts to purify adequate WT MBP-SL-OH+BARAD for SEC SAXS analyses were unsuccessful.

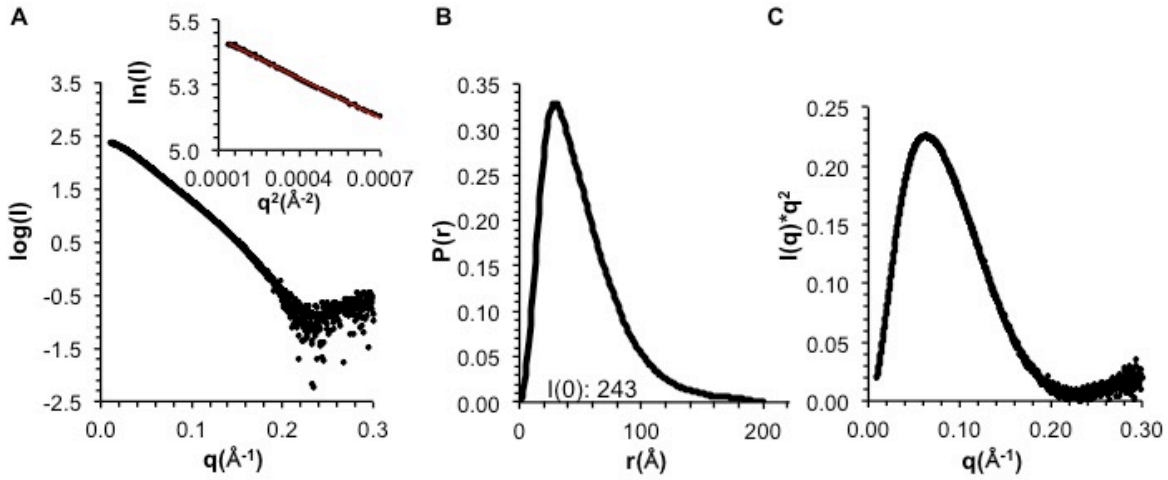

Figure 3.16. SAXS analysis of MBP-SL-BECN1 OH+BARAD^{TETRAD}.
A) Guinier plot. B) P(r) pairwise distribution. C) Kratky plot.

In summary, the presence of the MBP tag on the BECN1 fusion protein fragments renders the analysis of the effect of the OH on protein structure and stability ambiguous. Furthermore, the presence of the MBP tag is insufficient to obtain sufficient quantities of high quality protein for each fragment (*i.e.* the WT MBP-SL-OH+BARAD) required for biophysical analysis.

3.3.6. An Aromatic Finger in the BARAD may decrease solubility of BARAD-containing BECN1 fragments

Based on the above described purification results, we find that the WT BECN1 OH+BARAD and CCD-BARAD proteins aggregate during purification. However, the structural and stability analyses described above indicate that the presence of the MBP tag, which aids in purification of soluble protein, affects the structure and stability of the multi-domain protein constructs and; therefore, hinders comparison of the WT and OH mutant BECN1 protein fragments. Therefore, it was imperative to identify a different method to permit expression and purification of pure, homogenous protein without affecting protein structure and stability. Analysis of the OH+BARAD crystal structure (PDB ID: 4DDP) indicates that

the OH+BARAD molecules are arranged in a head-to-tail manner in the crystal lattice, stabilized by the aromatic finger, consisting of Phe^{359}, Phe^{360}, and Trp^{361}, of one OH+BARAD molecule being buried within a hydrophobic pocket formed partly by the OH in the next OH+BARAD molecule [36] (Figure 3.17). We hypothesized that this interaction may be the cause of protein aggregation during purification of WT OH+BARAD-containing constructs. Therefore, we decided to create an aromatic finger mutant (AFM), wherein the aromatic finger residues (Phe^{359}, Phe^{360}, and Trp^{361}) are mutated to Asp.

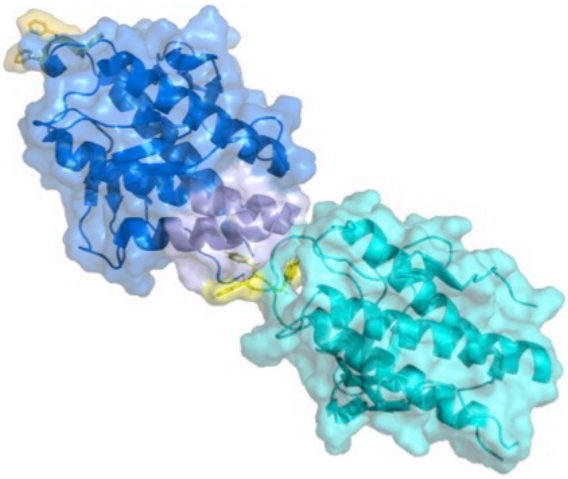

Figure 3.17. The aromatic finger of the OH+BARAD buried in the hydrophobic pocket of another molecule. The aromatic finger is shown in yellow and hydrophobic pocket is shown in grey.

3.3.7. Expression and purification results of BECN1 protein fragments with the AFM mutation

In order to obtain enough pure and homogenous WT OH and OH mutant BECN1 protein fragments for usage in the various biophysical and structural analyses in this study, we expressed AFM-containing, untagged and MBP fusion protein constructs of the BECN1 OH+BARAD, $OH+BARAD^{TETRAD}$, CCD-BARAD, and $CCD\text{-}BARAD^{TETRAD}$ (Table 3.1). Representative purification results are reported below, which demonstrate that the AFM permits acquisition of high quality, homogenous proteins containing the BARAD.

For MBP-BECN1 $CCD\text{-}BARAD^{AFM}$, MBP-SL-BECN1 $CCD\text{-}BARAD^{AFM}$ was expressed in *E. coli* Arctic Express cells, and soluble protein in the cell lysate from a 4 L expression culture was purified to homogeneity by affinity chromatography using a 3 mL amylose gravity column followed by SEC using a 10/30 HR Superdex 200 column. The single band on SDS-PAGE indicates the MBP-SL-BECN1 $CCD\text{-}BARAD^{AFM}$ was pure (Figure 3.18). The apparent molecular weight of MBP-SL-BECN1 $CCD\text{-}BARAD^{AFM}$

calculated based on the elution volume from SEC is 249.5 kD, which is 1.7-times greater than the theoretical molecular weight of 146.8 kD for the MBP-SL-BECN1 CCD-BARAD dimer calculated from the amino acid sequence. This is consistent with the MBP-SL-BECN1 CCD-BARAD^{AFM} forming rod-shaped dimer. The final yield of the purified MBP-SL-BECN1 CCD-BARAD obtained from one liter of bacterial culture was 2 mg. This fusion protein was used for structural comparisons with the WT MBP-BECN1 CCD-BARAD to verify that the AFM does not affect the secondary and tertiary structure of the CCD-BARAD.

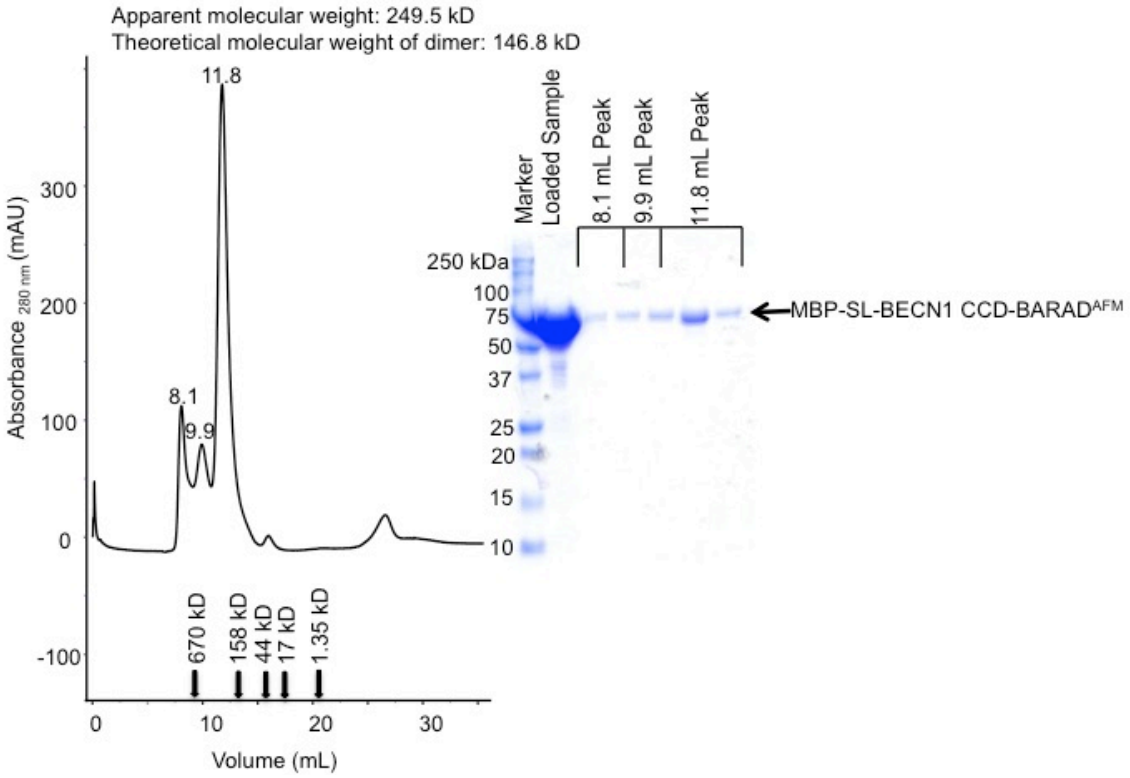

Figure 3.18. Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-SL-BECN1 CCD-BARAD^{AFM}.
The elution positions of SEC standards are indicated by arrows.

For BECN1 CCD-BARAD^{AFM}, MBP-BECN1 CCD-BARAD^{AFM} was expressed in *E. coli* Arctic Express cells, and soluble fusion protein in the cell lysate from a 6 L expression culture was purified to homogeneity by affinity chromatography using a 10 mL amylose resin gravity column, followed by off-column cleavage using 30 mg GST-TEV protease, to remove the MBP tag. The cleaved protein was

further purified to homogeneity by anion exchange chromatography using a 1mL 5/50 GL MonoQ column, and finally by SEC using a 10/30 HR Superdex 200 column. The single peak on SEC and single band on SDS-PAGE indicate the CCD-BARAD^{AFM} was pure and homogenous (Figure 3.19). The apparent molecular weight of CCD-BARAD^{AFM} calculated based on the elution volume from SEC is 120.0 kD, which is 1.9-times greater than the theoretical molecular weight of 64.6 kD for the CCD-BARAD^{AFM} dimer calculated from the amino acid sequence. This is consistent with the CCD-BARAD^{AFM} forming a long rod-shaped dimer. The final yield of the purified CCD-BARAD^{AFM} obtained from one liter of bacterial culture was 2.2 mg.

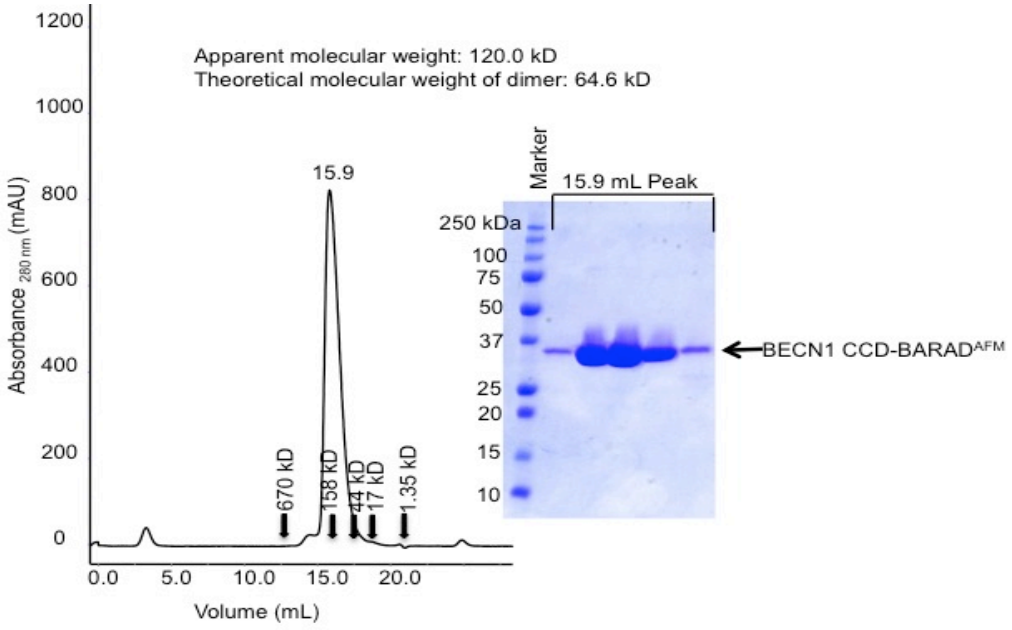

Figure 3.19. Size exclusion chromatogram and the corresponding SDS-PAGE of BECN1 CCD-BARAD^{AFM}.
The elution positions of SEC standards are indicated by arrows.

For MBP-BECN1 CCD-BARAD^{AFM,TETRAD}, MBP-SL-BECN1 CCD-BARAD^{AFM,TETRAD} was expressed in *E. coli* Arctic Express cells, and soluble protein in the cell lysate from a 4 L expression culture was purified to homogeneity by affinity chromatography using a 3 mL amylose gravity column followed by SEC using a 10/30 HR Superdex 200 column. The single band on SDS-PAGE indicates the MBP-SL-BECN1 CCD-BARAD^{AFM,TETRAD} was pure (Figure 3.20). The apparent molecular weight of MBP-SL-BECN1 CCD-BARAD^{AFM} calculated based on the elution volume from SEC is 208.0 kD, which is 1.4-times greater than

the theoretical molecular weight of 146.5 kD for the MBP-SL-BECN1 CCD-BARAD dimer calculated from the amino acid sequence. This is consistent with the MBP-SL-BECN1 CCD-BARAD^{AFM,TETRAD} forming rod-shaped dimer. The final yield of the purified MBP-SL-BECN1 CCD-BARAD obtained from one liter of bacterial culture was 4 mg. This fusion protein was used for structural comparisons with the MBP-BECN1 CCD-BARAD^{TETRAD} to verify that the AFM does not affect the secondary and tertiary structure of the CCD-BARAD^{TETRAD}.

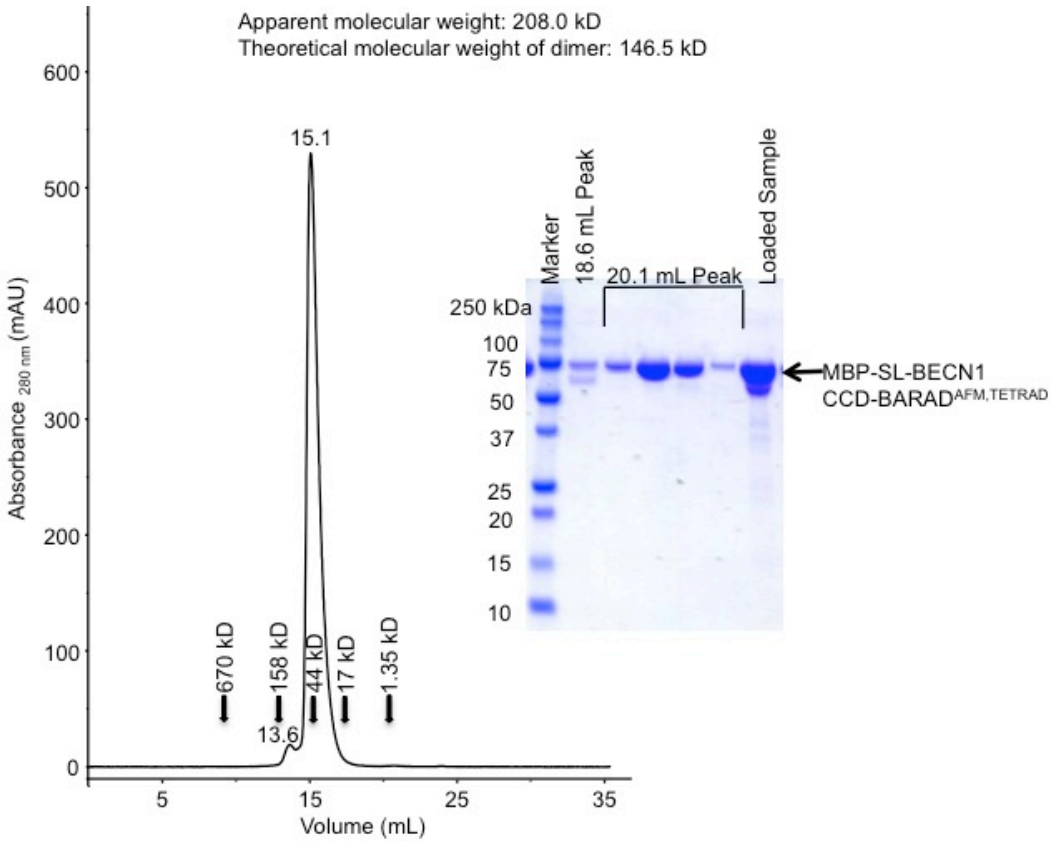

Figure 3.20. Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-SL-BECN1 CCD-BARAD^{AFM,TETRAD}.
The elution positions of SEC standards are indicated by arrows.

For BECN1 CCD-BARAD^{AFM,TETRAD}, MBP-BECN1 CCD-BARAD^{AFM,TETRAD} was expressed in *E. coli* Arctic Express cells, and soluble fusion protein in the cell lysate from a 6 L expression culture was purified to homogeneity by affinity chromatography using a 10 mL amylose resin gravity column, followed by off-column cleavage using 30 mg GST-TEV protease, to remove the MBP tag. The cleaved protein was further purified to homogeneity by anion exchange chromatography using a 1mL 5/50 GL MonoQ

column, and finally by SEC using a 10/30 HR Superdex 200 column. The single peak on SEC and single band on SDS-PAGE indicate the CCD-BARAD^{AFM,TETRAD} was pure and homogenous (Figure 3.21). The apparent molecular weight of CCD-BARAD^{AFM,TETRAD} calculated based on the elution volume from SEC is 69.0 kD, which is similar to the theoretical molecular weight of 64.8 kD for the CCD-BARAD^{AFM,TETRAD} dimer calculated from the amino acid sequence. This is consistent with the CCD-BARAD^{AFM,TETRAD} forming a dimer. The final yield of the purified CCD-BARAD^{AFM,TETRAD} obtained from one liter of bacterial culture was 1.8 mg.

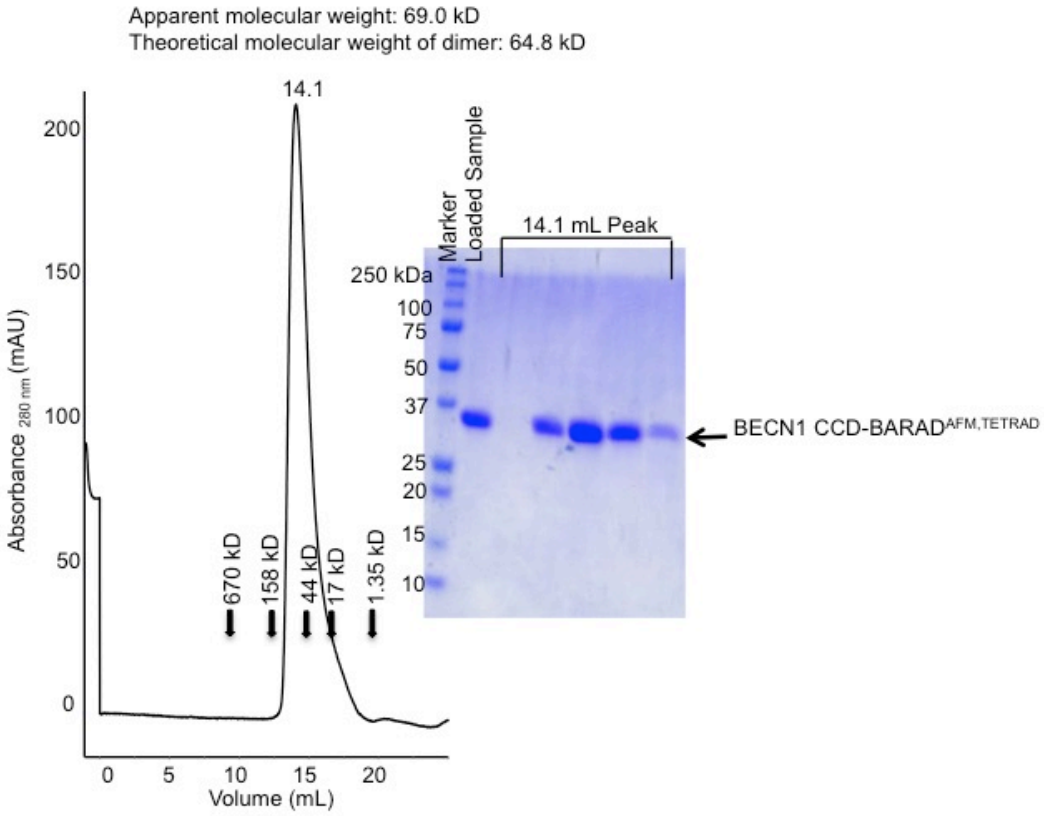

Figure 3.21. Size exclusion chromatogram and the corresponding SDS-PAGE of BECN1 CCD-BARAD^{AFM,TETRAD}.
The elution positions of SEC standards are indicated by arrows.

For MBP-BECN1 OH+BARAD^{AFM}, MBP-SL-BECN1 OH+BARAD^{AFM} was expressed in *E. coli* Arctic Express cells, and soluble protein in the cell lysate from a 3 L expression culture was purified to homogeneity by affinity chromatography using a 10 mL amylose resin gravity column followed by SEC using a 10/30 HR Superdex 200 column. The single peak on SEC and single band on SDS-PAGE

indicates the MBP-SL-BECN1 OH+BARAD^{AFM} was pure and homogenous (Figure 3.22). The apparent molecular weight of MBP-SL-BECN1 OH+BARAD^{AFM} calculated based on the elution volume from SEC is 61.8 kD, which is similar to the theoretical molecular weight of 62.9 kD for the MBP-SL-BECN1 OH+BARAD^{AFM} monomer calculated from the amino acid sequence. This is consistent with the MBP-SL-BECN1 OH+BARAD^{AFM} forming a monomer. The final yield of the purified MBP-SL-BECN1 OH+BARAD^{AFM} obtained from one liter of bacterial culture was 9.2 mg. This fusion protein was used for structural comparisons to verify that the AFM does not affect the secondary and tertiary structure of the OH+BARAD.

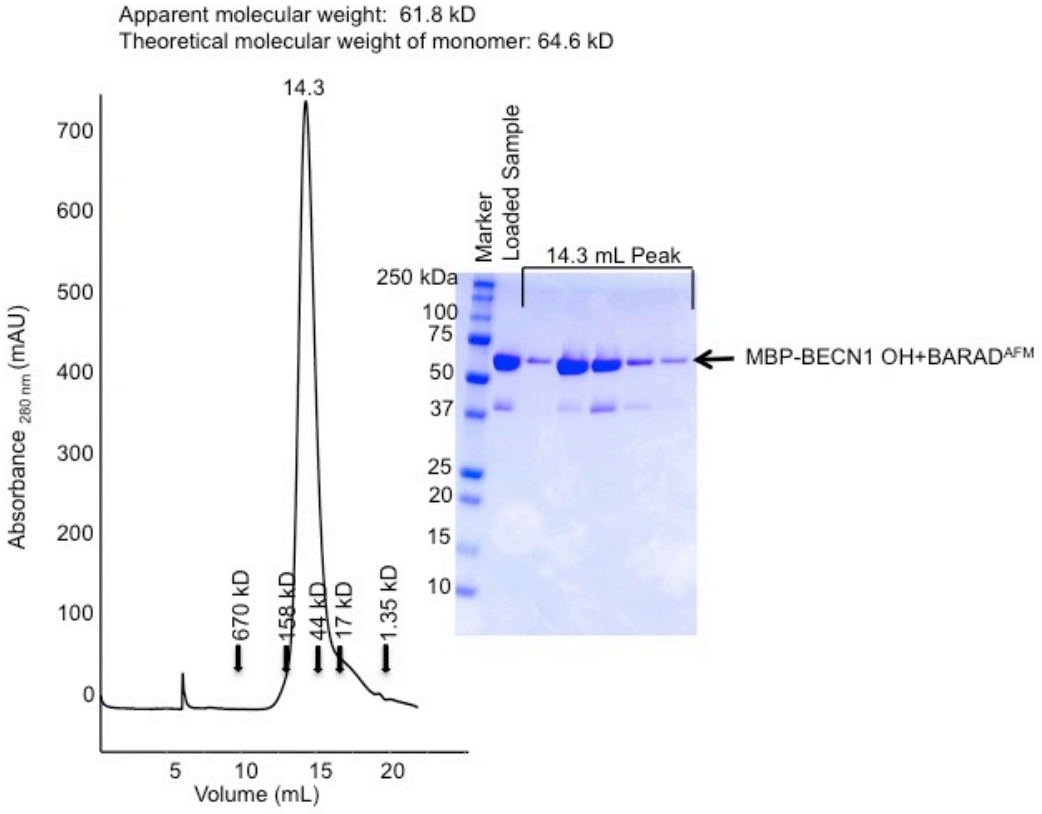

Figure 3.22. Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-SL-BECN1 OH+BARAD^{AFM}.
The elution positions of SEC standards are indicated by arrows.

For BECN1 OH+BARAD^{AFM}, His_6-BECN1 OH+BARAD^{AFM} was expressed in *E. coli* Arctic Express cells, and soluble protein in the cell lysate from a 3 L expression culture was purified to homogeneity by affinity chromatography using a 3 mL nickel affinity resin gravity column followed by SEC using a 10/30

HR Superdex 200 column. The single peak on SEC and single band on SDS-PAGE indicate the OH+BARAD^{AFM} was pure and homogenous (Figure 3.23). The apparent molecular weight of OH+BARAD^{AFM} calculated based on the elution volume from SEC is 20.6 kD, which is similar to the theoretical molecular weight of 26.0 kD for the OH+BARAD^{AFM} monomer calculated from the amino acid sequence. This is consistent with the OH-BARAD^{AFM} forming a globular monomer. The final yield of the purified OH+BARAD^{AFM} obtained from one liter of bacterial culture was 2 mg.

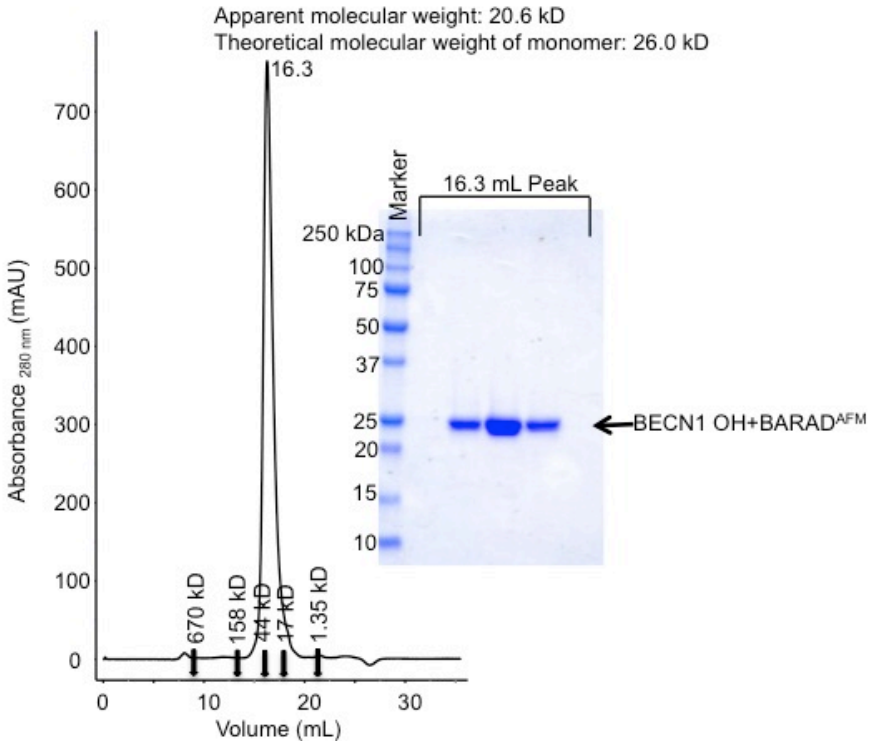

Figure 3.23. Size exclusion chromatogram and the corresponding SDS-PAGE of BECN1 OH+BARAD^{AFM}. The elution positions of SEC standards are indicated by arrows.

For MBP-BECN1 OH+BARAD^{AFM,TETRAD}, MBP-SL-BECN1 OH+BARAD^{AFM,TETRAD} was expressed in *E. coli* Arctic Express cells, and soluble protein in the cell lysate from a 3 L expression culture was purified to homogeneity by affinity chromatography using a 10 mL amylose resin gravity followed by SEC using a 10/30 HR Superdex 200 column. The single peak on SEC and single band on SDS-PAGE indicates the MBP-SL-BECN1 OH+BARAD^{AFM,TETRAD} was pure and homogenous (Figure 3.24). The apparent molecular weight of MBP-SL-BECN1 OH+BARAD^{AFM,TETRAD} calculated based on the elution volume from SEC is 35.3 kD, which is 1.8-times smaller than the theoretical molecular weight of 62.7 kD

for the MBP-SL-BECN1 OH+BARAD^{AFM,TETRAD} monomer calculated from the amino acid sequence. This is consistent with the MBP-SL-BECN1 OH+BARAD^{AFM,TETRAD} forming a monomer. The final yield of the purified MBP-SL-BECN1 OH+BARAD^{AFM,TETRAD} obtained from one liter of bacterial culture was 6.5 mg. This fusion protein was used for structural comparisons to verify that the AFM does not affect the secondary and tertiary structure of the OH+BARAD.

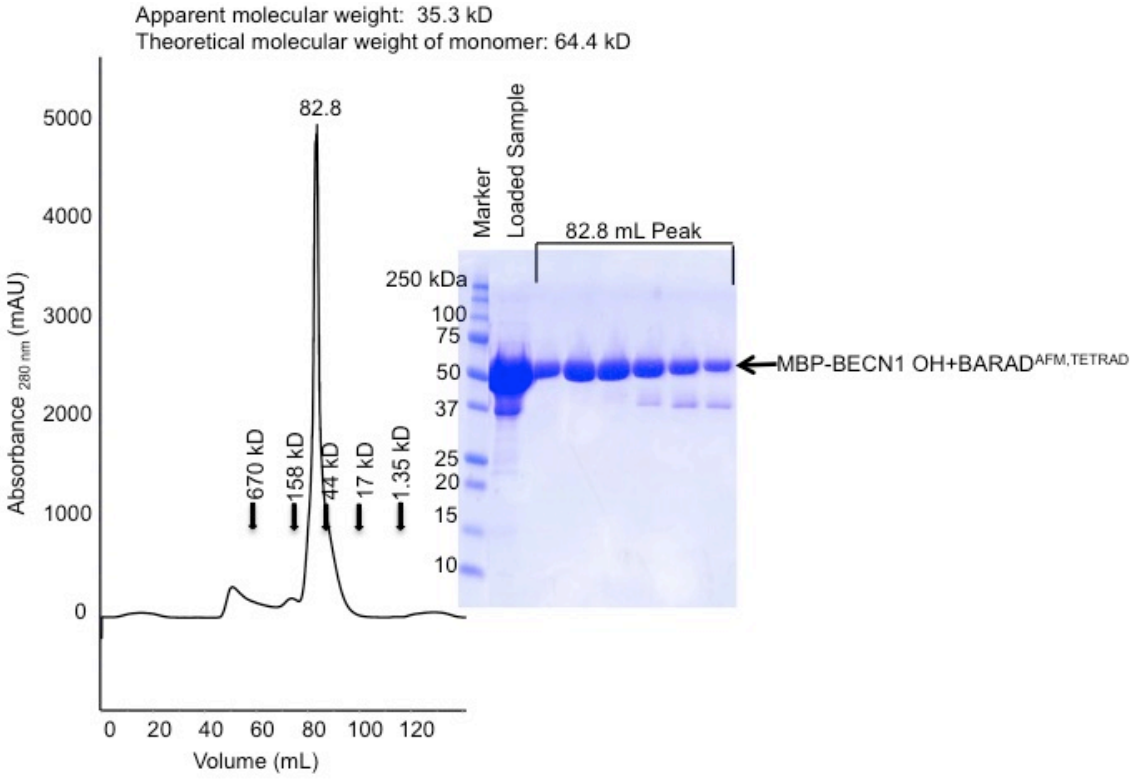

Figure 3.24. Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-SL-BECN1 OH+BARAD^{AFM,TETRAD}.
The elution positions of SEC standards are indicated by arrows.

For BECN1 OH+BARAD^{AFM,TETRAD}, His_6-BECN1 OH+BARAD^{AFM,TETRAD} was expressed in *E. coli* Arctic Express cells, and soluble protein in the cell lysate from a 4 L expression culture was purified to homogeneity by affinity chromatography using a 3 mL nickel affinity resin gravity column followed by SEC using a 10/30 HR Superdex 200 column. The single peak on SEC and single band on SDS-PAGE indicate the OH+BARAD^{AFM,TETRAD} was pure and homogenous (Figure 3.25). The apparent molecular weight of OH+BARAD$^{AFM.TETRAD}$ calculated based on the elution volume from SEC is 19.6 kD, which is 1.4

times smaller than the theoretical molecular weight of 26.7 kD for the OH+BARAD^{AFM,TETRAD} monomer calculated from the amino acid sequence. This is consistent with the OH+BARAD^{AFM,TETRAD} forming a globular monomer. The final yield of the purified OH+BARAD^{AFM,TETRAD} obtained from one liter of bacterial culture was 2 mg.

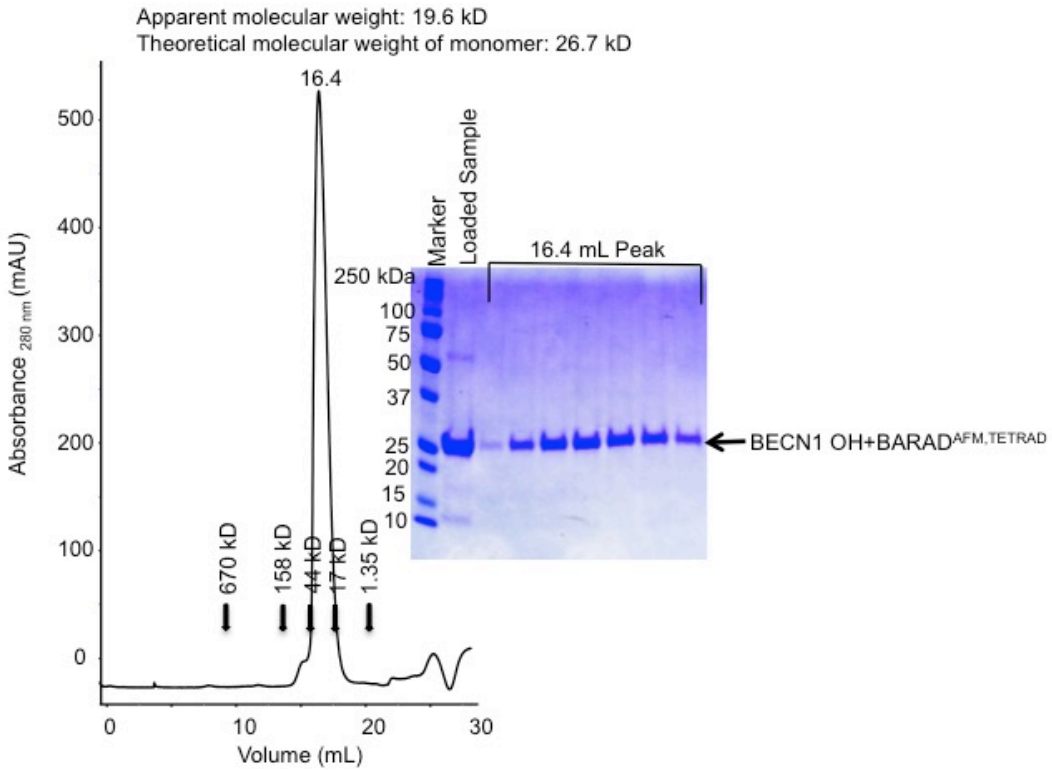

Figure 3.25. Size exclusion chromatogram and the corresponding SDS-PAGE of BECN1 OH+BARAD^{AFM,TETRAD}.
The elution positions of SEC standards are indicated by arrows.

3.3.8. The AFM does not impact secondary and tertiary structure of the CCD+BARAD constructs

The AFM, wherein the aromatic finger residues (F359, F360, and W361) are mutated to Asp, dramatically reduces aggregation and permits large-scale purification of homogenous, high-quality protein. Therefore, we next sought to demonstrate that the AFM has no effect on the secondary structure content estimated from CD of the CCD+BARAD constructs (Figure 3.26, Table 3.5). Secondary structure content estimated by analysis of CD spectra for the MBP-BECN1 CCD-BARAD and the MBP-BECN1 CCD-BARAD^{AFM} indicates that the AFM does not change the secondary structure content of the BECN1 CCD-BARAD, regardless of the presence of the Tetrad mutation (Figure 3.26, Table 3.5). The MBP-

BECN1 CCD-BARAD contains 269 helical, 82 strand, and 296 coil residues while the MBP-BECN1 CCD-BARAD^{AFM} contains 263 helical, 89 strand, and 299 coil residues and the MBP-BECN1 CCD-BARAD^{TETRAD} contains 300 helical, 74 strand, and 282 coil residues while the MBP-BECN1 CCD-BARAD^{AFM,TETRAD} contains 293 helical, 73 strand, and 288 coil residues (Table 3.5). Since for each class of secondary structure, secondary structure content does not vary beyond 10% between MBP-BECN1 CCD-BARAD and MBP-BECN1 CCD-BARAD^{AFM}, we conclude the AFM does not alter the secondary structure of these proteins.

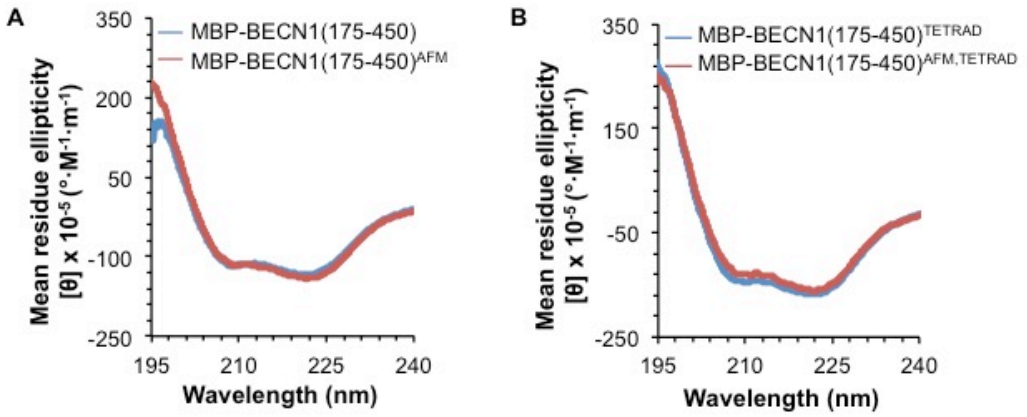

Figure 3.26. The AFM does not alter CD spectra of MBP-BECN1 CCD-BARAD.
A) MBP-BECN1(175-450) and B) MBP-BECN1(175-450)^{TETRAD}.

Table 3.5. Average Estimated Secondary Structure Content in MBP-BECN1 CCD-BARAD Proteins.

Protein	# Residues	Helix (%)	Beta (%)	Coil %
MBP-BECN1(175-450)	651	41.3 ± 2.9	12.6 ± 1.0	45.5 ± 1.6
MBP-BECN1(175-450)^{AFM}	651	40.5 ± 1.3	13.7 ± 0.9	46.0 ± 0.9
MBP-BECN1(175-450)^{TETRAD}	651	46.0 ± 2.9	11.4 ± 0.8	43.3 ± 2.1
MBP-BECN1(175-450)^{AFM,TETRAD}	651	45.1 ± 2.2	11.2 ± 1.6	44.3 ± 0.6

The AFM also has no effect on the SEC-SAXS profiles of the CCD-BARAD constructs (Figure 3.27, Table 3.6). SEC-SAXS data were recorded for the MBP-BECN1 CCD-BARAD (Figure 3.27A) and MBP-BECN1 CCD-BARAD^{AFM} proteins (Figure 3.27E). An R_g of 68 ± 2 Å (Table 3.6) is estimated from the Guinier plots and P(r) plots of MBP-BECN1 CCD-BARAD (Figure 3.27B and D), as well as MBP-BECN1 CCD-BARAD^{AFM} (Figure 3.27F and H). Similarly, a D_{max} of 250 Å is estimated from the P(r) distribution for both samples (Figure 3.27D and H; Table 3.6). Further, the P(r) plots (Figure 3.27D and H) and Kratky plots (Figure 3.27C and G) indicate that both the MBP-BECN1 CCD-BARAD and MBP-

BECN1 CCD-BARAD^{AFM} are well folded. The P(r) plots (Figure 3.27D and H) also suggest a dumbbell-shaped envelope for both samples. This similarity in size and shape of the WT and AFM mutant MBP-BECN1 CCD-BARAD indicates that the tertiary structure of the BARAD is unaffected by the AFM.

Table 3.6. Summary of SEC-SAXS data for MBP-BECN1 CCD-BARAD proteins.

Protein	**MW_T**	**Guinier**	**P(r)**	
	(KD)	**R_g (Å)**	**R_g (Å)**	**D_{max} (Å)**
MBP-BECN1(175-450)	147.2 (dimer)	66	71	250
MBP-BECN1(175-450)^{AFM}	146.9 (dimer)	63	71	250

MW_T, theoretical molecular weight.

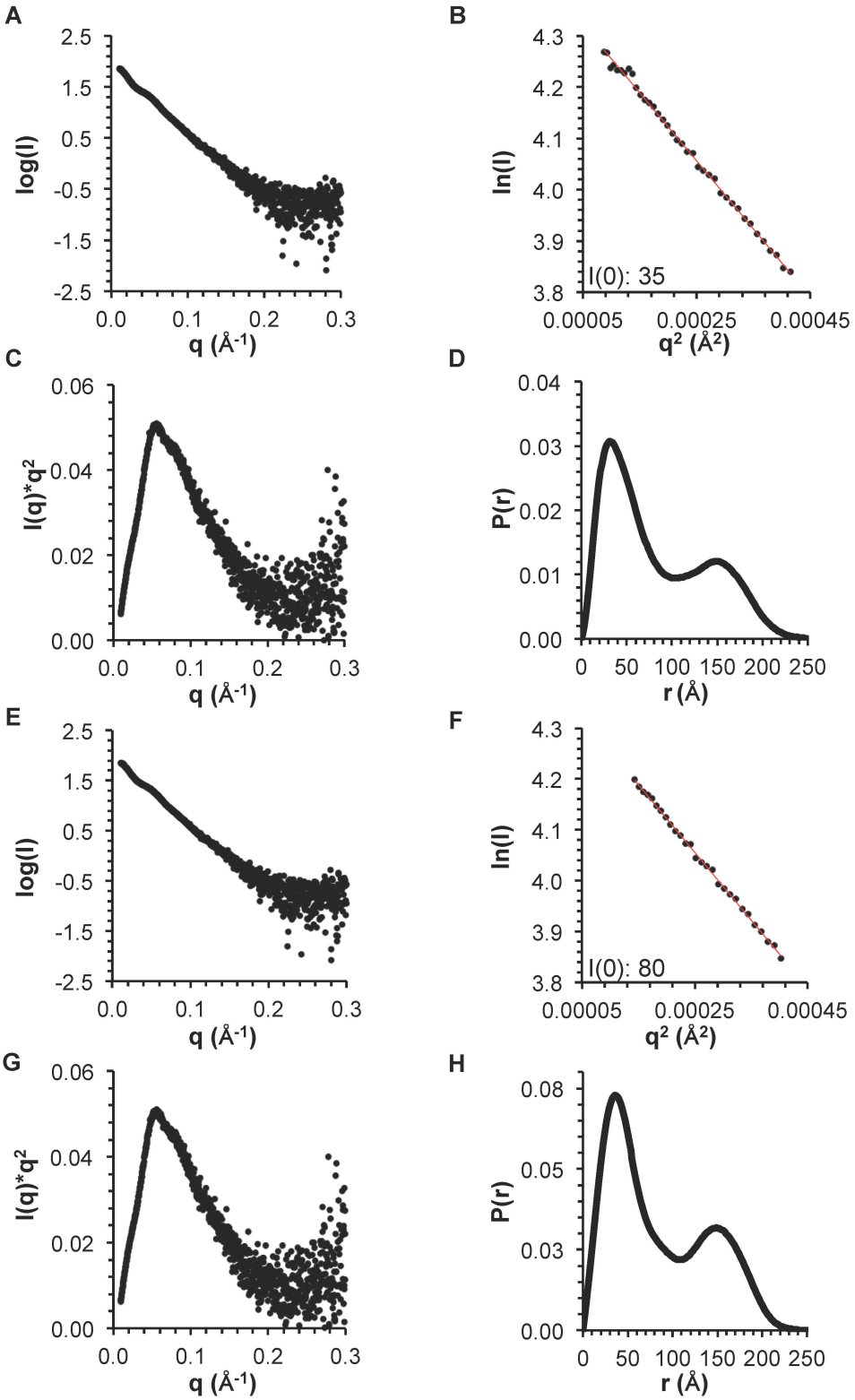

Figure 3.27. SEC-SAXS analysis of MBP-BECN1 CCD-BARAD proteins.
For the MBP-BECN1 CCD-BARAD, the A) scattering data, B) Guinier, C) Kratky, and D) P(r) pairwise distribution. For the MBP-BECN1 CCD-BARAD^{AFM}, the E) scattering data, F) Guinier, G) Kratky, and H) P(r) pairwise distribution.

For the MBP-SL-BECN1 OH+BARAD^{AFM}, the R_g estimated from the Guinier plot and P(r) distribution is ~30 Å, (Figure 3.28A and B, Table 3.7) and the shape of the P(r) distribution curves and the Kratky plots indicates that it is a well-folded, globular protein (Figure 3.28B and C). This fusion protein was intended for comparison to the MBP-SL-OH+BARAD in order to assess the effect of the AFM mutant on tertiary structure of the BARAD. However, as explained in section 3.3.2, attempts to purify adequate WT MBP-SL-OH+BARAD for SEC SAXS analyses were unsuccessful.

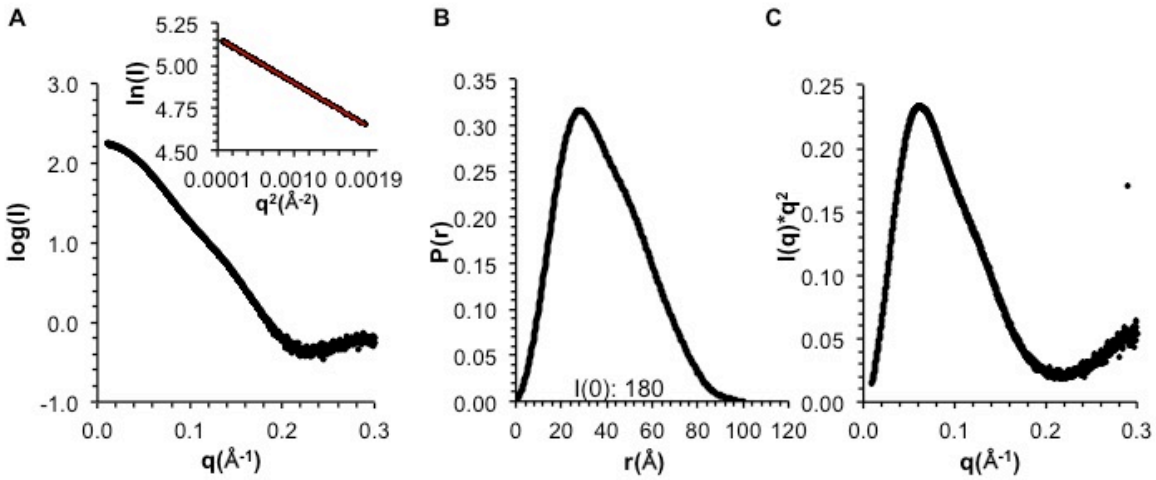

Figure 3.28. SAXS analysis of MBP-SL-BECN1 OH+BARAD^{AFM}.
A) Guinier plot. B) P(r) pairwise distribution. C) Kratky plot.

The AFM also does not appear to significantly affect the tertiary structure of the MBP-SL-BECN1 OH+BARAD^{TETRAD}. SEC-SAXS data were recorded for the MBP-BECN1 OH+BARAD^{TETRAD} (Figure 3.16, Table 3.4) and the MBP-BECN1 OH+BARAD^{AFM,TETRAD} (Figure 3.29, Table 3.7). An R_g of 37 ± 2 Å (Table 3.4) is estimated from the Guinier plots and P(r) plots of MBP-BECN1 OH+BARAD^{TETRAD} (Figure 3.16A and B), as well as MBP-BECN1 OH+BARAD^{AFM,TETRAD} (Figure 3.28A and B). Similarly, a D_{max} of 163 ± 37 Å is estimated from the P(r) distribution for both samples (Figures 3.16B and 3.29B; Table 3.4 and 3.7). Further, the P(r) plots (Figures 3.16B and 2.28B) and Kratky plots (Figure 3.16C and 3.28C) indicate that both the MBP-BECN1 OH+BARAD^{TETRAD} and MBP-BECN1 OH+BARAD^{AFM,TETRAD} are well folded. The asymmetric curve P(r) plots with a maximum at distances less than $D_{max}/2$ (Figures 3.16B and 3.28B) also both suggest a slightly elongated protein, which is consistent with two globular domains linked by a short linker. This similarity in size and shape of the WT and AFM mutant MBP-BECN1 OH+BARAD^{TETRAD} indicates that the tertiary structure of the BARAD^{TETRAD} is unaffected by the AFM.

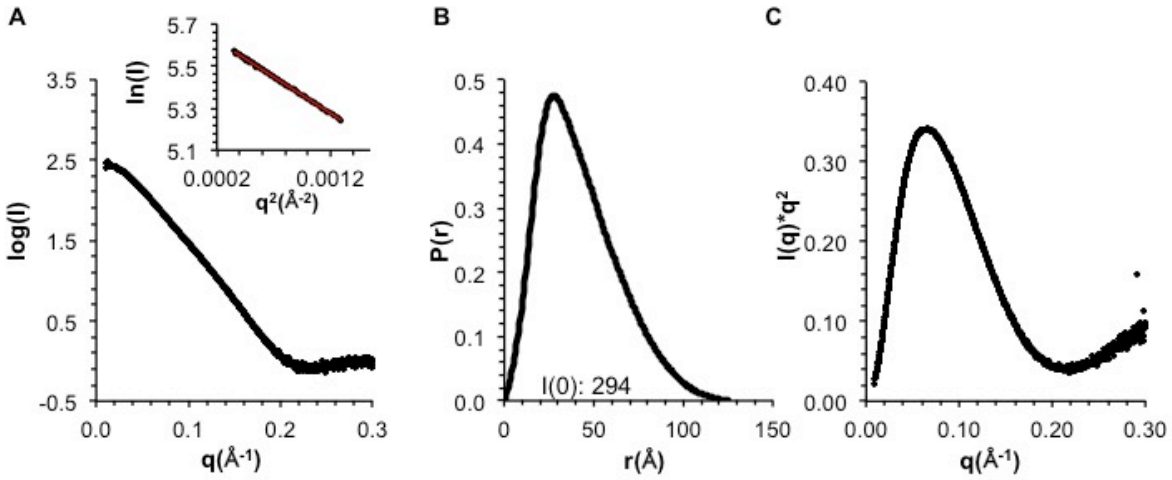

Figure 3.29. SAXS analysis of MBP-SL-BECN1 OH+BARAD^{AFM,TETRAD}.
A) Guinier plot. B) P(r) pairwise distribution. C) Kratky plot.

Table 3.7. Summary of SEC-SAXS data for MBP-BECN1 OH+BARAD proteins.

Protein	MW_T	Guinier	P(r)	
	(KD)	R_g (Å)	R_g (Å)	D_{max} (Å)
MBP-SL-BECN1 OH+BARAD^{AFM}	64.6 (monomer)	30	30	100
MBP-SL-BECN1 OH+BARAD^{AFM,TETRAD}	64.4 (monomer)	32	33	125

MW_T, theoretical molecular weight.

These results indicating that the AFM does not affect the secondary structure content or SEC-SAXS profiles of the CCD-BARAD constructs are consistent with previously published analyses that showed that the AFM does not change the overall secondary structure or thermal stability of the BECN1 OH+BARAD[36]. Thus, since the AFM dramatically improves the yield of soluble homogeneous protein for both the OH+BARAD and the CCD+BARAD fragments and also does not appear to impact secondary or tertiary structure of these fragments; all BARAD-containing proteins used for subsequent analysis bear the AFM.

3.3.9. BECN1 CCD structure is destabilized by the OH Tetrad mutation

The BECN1 CCD interface is composed of 26 interacting pairs comprised of 13 unique pairs related by two-fold homodimer symmetry. Only six of these thirteen unique pairs comprise residue pairs that form good hydrophobic interactions; while for each of the remaining seven pairs, one residue has a polar and/or bulky side-chain resulting in non-ideal packing [33, 34]. At the termini of each set of 13 interacting pairs within the CCD homodimer, each OH contributes to five pairs of interactions; V250:E191,

M254:L187, A257:L184, L261:L180, and L264:S177 (Figure 3.1). This includes three of the six good hydrophobic interacting pairs. Thus, the OH contributes half the hydrophobic interactions that stabilize the BECN1 CCD homodimer.

In order to assess the effect of the Tetrad mutation on CCD homodimerization, we compared dimer self-dissociation K_d of the CCD and CCD^{TETRAD}, as quantified by ITC. At 20 °C, self-association of the CCD^{TETRAD} is ~1.5 fold weaker than that of the WT CCD (Figure 3.30, Table 3.8). A comparison of the secondary structure content estimated by analysis of CD spectra for the CCD and CCD^{TETRAD} indicates that the Tetrad mutation reduces helicity by ~9.5 residues with a concomitant increase of ~9 residues in coil content of the CCD but does not impact β-strand content (Figure 3.31A, Table 3.9). Therefore, the Tetrad mutation destabilizes CCD secondary structure. The cooperative, reversible melting curves indicate that the melting temperature (T_m) of ~33.5 °C of the WT CCD (Figure 3.31B, Table 3.9) decreases substantially to 22.5 °C for the CCD^{TETRAD} (Figure 3.31C, Table 3.9), which is also indicative of the Tetrad mutation destabilizing the CCD structure.

Table 3.8. Thermodynamics of self-dissociation of CCD-containing proteins.

BECN1 Fragment	**K_d (μM)**	**ΔH (kJ/mol)**	**ΔS (J/K mol)**
CCD	44.8 ± 2.0	45.0 ± 10.5	236.6 ± 36.2
CCD^{TETRAD}	70.7 ± 3.1	157.0 ± 8.8	615.1 ± 30.0
CCD-BARAD^{AHM}	61.0 ± 8.1	207.1 ± 21.5	787.1 ± 73.0
CCD-BARAD^{AHM,TETRAD}	134.2 ± 4.9	143.0 ± 10.3	561.9 ± 34.9

Performed at 20 °C in 25 mM HEPES salt pH 8.0, 150mM NaCl, 2mM BME.

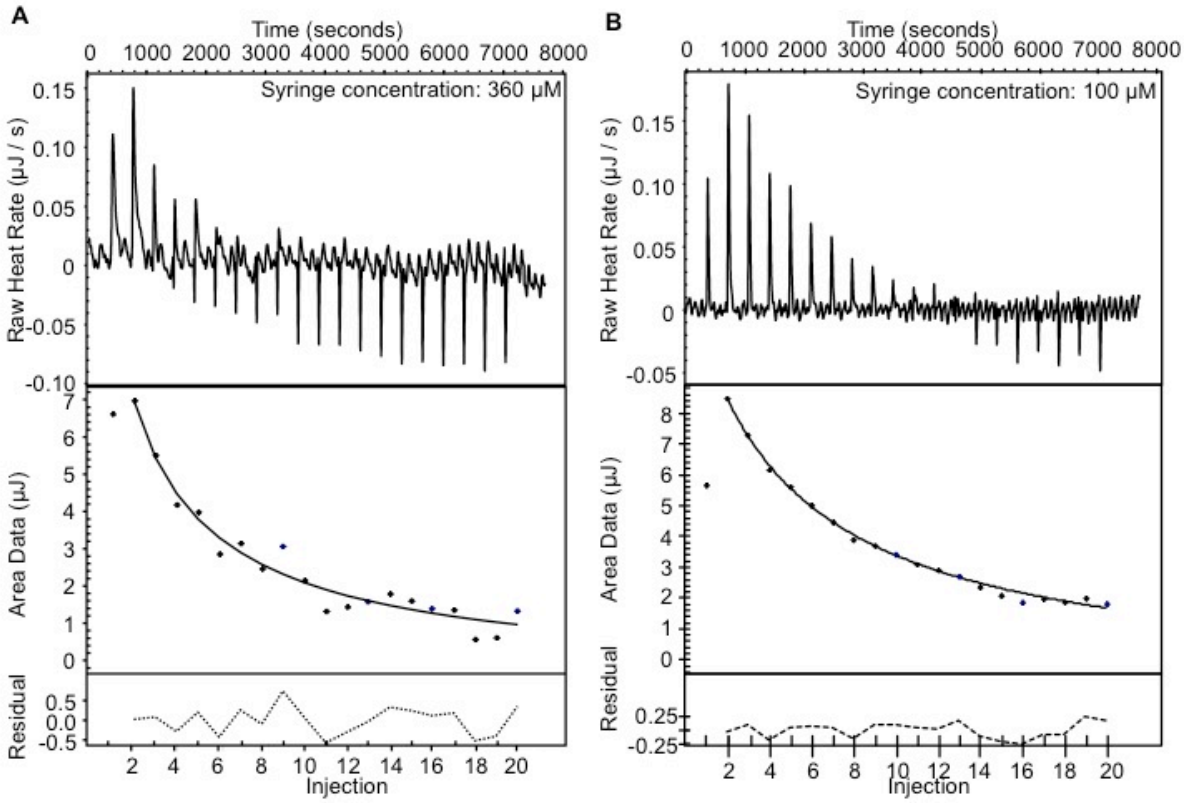

Figure 3.30. Homodimer dissociation ITC profiles of the BECN1 CCD and CCD^{TETRAD}.
A) BECN1 CCD. B) BECN1 CCD^{TETRAD}.

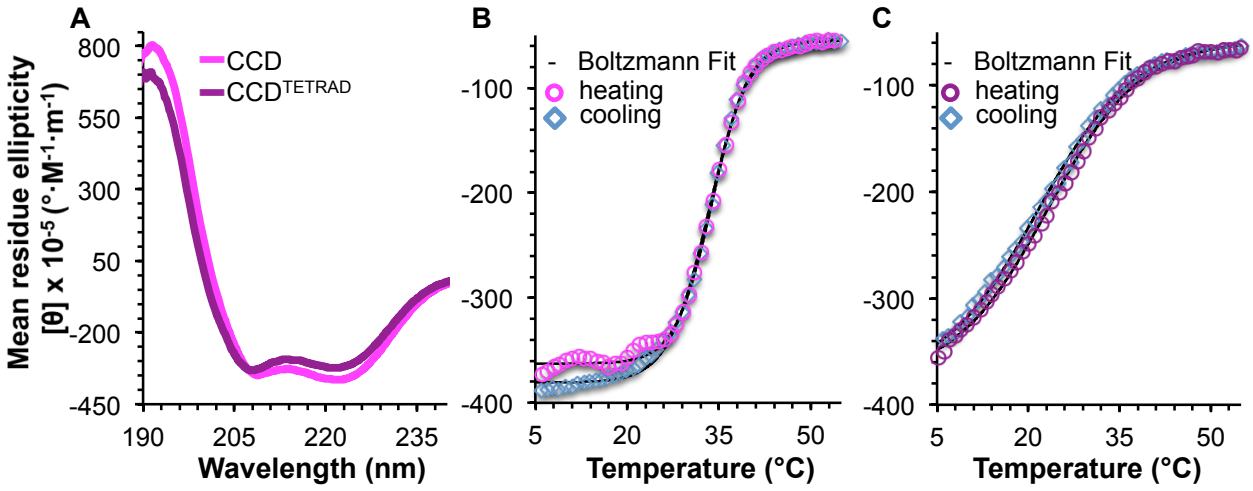

Figure 3.31. BECN1 CCD CD spectra and melting curves.
A) BECN1 CCD (magenta) and BECN1 CCD^{TETRAD} (purple). B) BECN1 CCD melting curve. C) BECN1 CCD^{TETRAD} melting curve. The Boltzmann fits of the melting curves are shown in black.

Table 3.9. CD analysis: Estimated secondary structure content and melting temperatures.

Protein	Average estimated secondary structure content (%)			Melting temperature (°C)	
	Helix	Strand	Coil	Forward	Reverse
BECN1 CCD	89.8 ± 6.1	1.9 ± 1.1	8.6 ± 6.4	33.9 ± 0.1	33.3 ± 0.1
BECN1 CCD^{TETRAD}	80.4 ± 4.8	1.9 ± 1.6	17.6 ± 6.8	23.2 ± 0.2	21.6 ± 0.2
BECN1 OH+BARAD^{AFM}	46.1 ± 7.9	20.2 ± 3.0	34.2 ± 9.8	47.7 ± 0.1	NA[a]
BECN1 OH+BARAD^{AFM,TETRAD}	46.3 ± 3.8	10.6 ± 2.6	44.0 ± 3.2	41.4 ± 0.1	NA
BECN1 CCD-BARAD^{AFM}	65.2 ± 6.7	5.3 ± 1.6	29.9 ± 7.4	23.9 ± 0.6; 49.0 ± 0.3	NA
BECN1 CCD-BARAD^{AFM,TETRAD}	56.2 ± 4.1	7.5 ± 0.2	36.7 ± 3.7	22.1 ± 0.6; 45.1 ± 0.2	NA

[a] NA, not available; Proteins containing the BARAD precipitated, even upon heating just to the melting temperature; therefore, reverse T_m could not be determined.

We used SEC-SAXS to investigate the impact of the Tetrad mutation on size, shape, and structure of the CCD. The R_g of the CCD^{TETRAD}, estimated from the Guinier plot and P(r) distribution, is ~39 Å (Figure 3.32A and B; Table 3.10), which is similar to the R_g of ~36 Å reported for the WT CCD[33]. However, the D_{max} of the CCD^{TETRAD} obtained from the P(r) distribution is 155 Å, which is about 30% longer than the D_{max} reported for the CCD[33] (Figure 3.32B, Table 3.10). Further, the Kratky plot indicates that the CCD^{TETRAD} is largely unfolded (Figure 3.32C), in contrast to the CCD which is well-folded[33]. The increase in the D_{max} observed in the CCD^{TETRAD} is consistent with the destabilization of the CCD by the Tetrad mutation, observed in secondary structure estimations and melting temperature calculations by CD, since increased disorder would result in polypeptides transiently sampling a larger volume in solution.

Table 3.10. SAXS data analysis of OH-containing BECN1 protein fragments.

		BECN1 Domain					
		CCD [33]	CCD TETRAD	OH+BARAD AFM	OH+BARAD AFM,TETRAD	CCD-BARAD AFM	CCD-BARAD AFM,TETRAD
From sequence	MW_T (KD)	24.4 (dimer)	21.8 (dimer)	26.0 (monomer)	26.7 (monomer)	64.6 (dimer)	64.8 (dimer)
SAXS parameters	Rg (Å)	34	37	20	22	45	47
	P(r) R_g (Å)	37	40	20	23	53	53
	Dmax (Å)	121	155	77	92	195	229

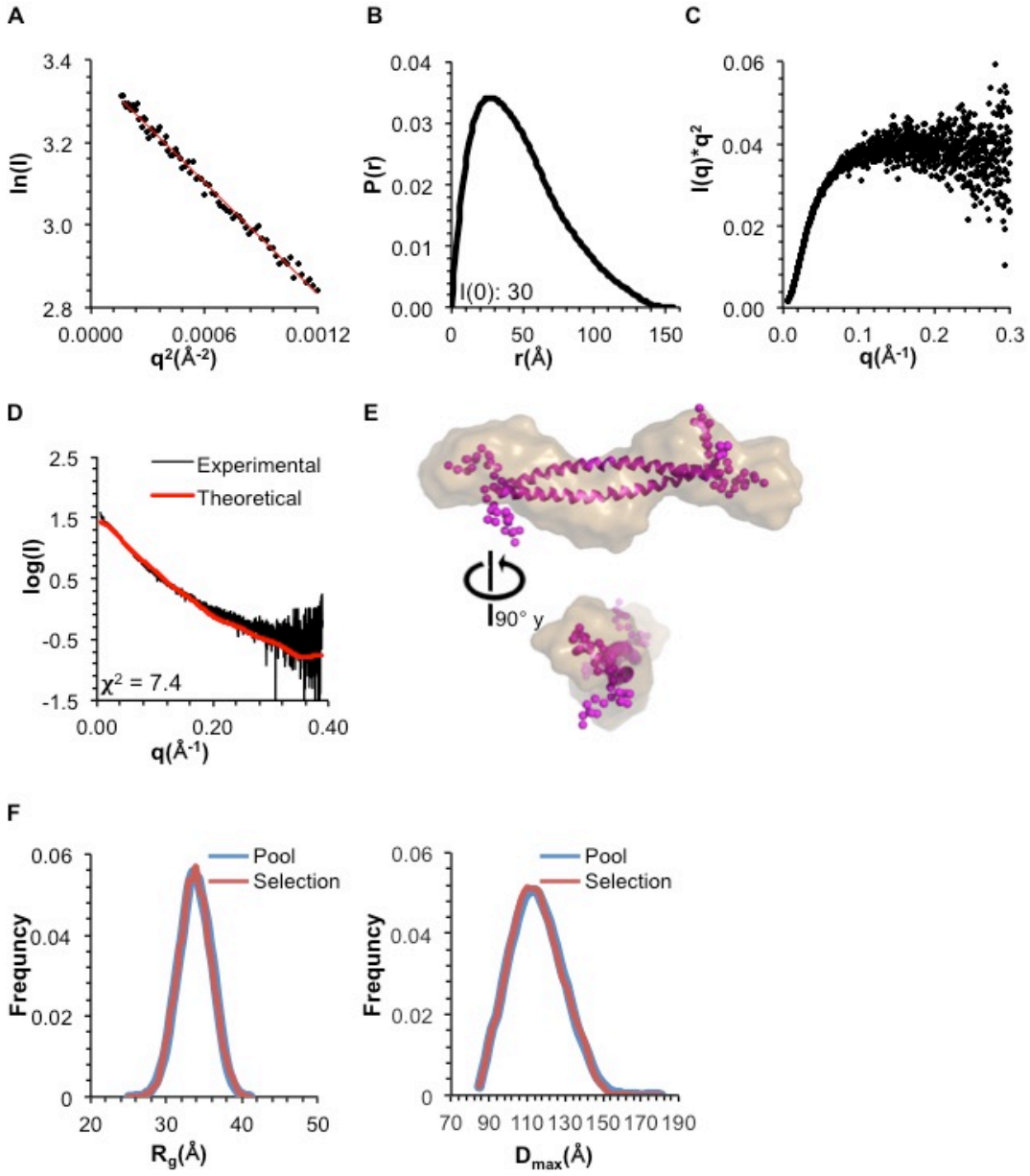

Figure 3.32. SAXS analysis of BECN1 CCD^{TETRAD}.
A) Guinier plot. B) P(r) pairwise distribution. C) Kratky plot. D) Theoretical scattering curve calculated from the selected EOM model of the BECN1 CCD with flexible OH fitted to the experimental data. E) Selected EOM model (magenta cartoon, flexible residues as spheres) superimposed into the filtered SAXS envelope (sand). F) R_g (left panel) and D_{max} (right panel) distribution of initial pool and selected ensemble of models.

The theoretical scattering data calculated from the selected EOM model of the CCD^{TETRAD} fits the experimental SAXS data poorly with a χ^2 of 7.4 (Figure 3.32D and E); which supports the conclusion that the CCD^{TETRAD} is highly flexible and may further indicate that the Tetrad mutation destabilizes more of the CCD than just the OH since the Kratky plot suggests the CCD^{TETRAD} is largely unfolded (Figure 3.32C). In

contrast, the theoretical scattering calculated from the WT CCD crystal structure fits the WT SAXS data very well[33]. Analysis using EOM[150] further confirms that the CCD^{TETRAD} is highly flexible. EOM generates a random pool of models and selects an ensemble of models from the pool that best fit the scattering data, providing an R_g and D_{max} distribution for both the pool and selected ensemble of models[150]. A narrow R_g and D_{max} distribution for the ensemble compared to the pool suggests that the molecule is rigid; whereas an R_g and D_{max} distribution for the ensemble that is as broad as the pool suggests that the molecule is flexible[154]. The R_g and D_{max} distribution for the CCD^{TETRAD} pool and ensemble are equally broad (Figure 3.32F), suggesting that the CCD^{TETRAD} is highly flexible.

Taken together; the secondary structure and melting temperature estimations by CD, homodimer association affinity quantified by ITC, solution size and shape parameter determinations by SEC-SAXS, and assessment of flexibility by the Kratky plot and EOM, all indicate that CCD secondary, tertiary and quaternary structure stability is significantly disrupted by the OH Tetrad mutation.

3.3.10. The OH Tetrad mutation impacts the BECN1 OH+BARAD structure less than that of the CCD

The crystal structure of the human BARAD[36] includes the OH packed against the BARAD, but the OH was not included in the yeast VPS30 BARAD crystal structure[35]. In each case, the crystal structure as well as solution data indicate that the BARAD is a monomer. Therefore, the OH does not mediate homodimerization of the BARAD, nor does it appear to be required for stability of the BARAD structure in all VPS30/BECN1 homologs.

A comparison of the secondary structure content estimated from CD spectra analysis for the $OH+BARAD^{AFM}$ and $OH+BARAD^{AFM,TETRAD}$ indicates that the Tetrad mutation reduces β-strand content by ~21 residues with a concomitant increase of ~25 residues in coil content of the OH+BARAD, but does not impact α-helical content (Figure 3.33A, Table 3.9). This suggests that although the Tetrad mutation does not disrupt the OH itself, it does distort BARAD secondary structure. Reversible melting curves were not obtained for any construct containing the BARAD, as the proteins precipitated even when heated only to the Tm. However, melting is cooperative for the $OH+BARAD^{AFM}$ and $OH+BARAD^{AFM,TETRAD}$ (Figure 3.33B and C), enabling us to assess T_m. The Tetrad mutation decreases the T_m of the BARAD by only 7 °C (Table 3.9). Thus, the BARAD is less affected by the Tetrad mutation than the CCD.

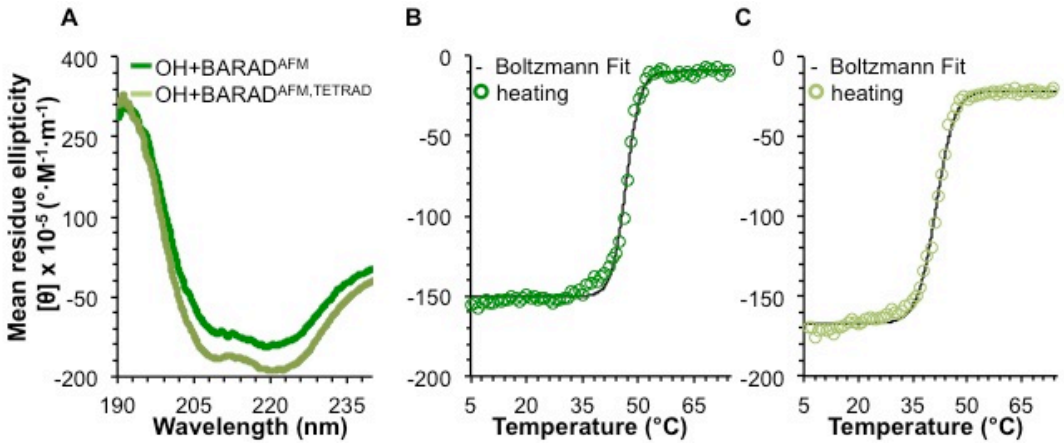

Figure 3.33. BECN1 OH+BARAD CD spectra and melting curves.
A) BECN1 OH+BARAD^{AFM} (green) and BECN1 OH+BARAD^{AFM,TETRAD} (light green). B) BECN1 OH+BARAD^{AFM} melting curve. C) OH+BARAD^{AFM,TETRAD} melting curve. The Boltzmann fits of the melting curves are shown in black.

Next, we used SEC-SAXS to investigate the impact of the Tetrad mutation on size, shape, and structure of the OH+BARAD. The R_g of the OH+BARAD^{AFM} estimated from the Guinier plot and P(r) distribution is ~20 Å (Figure 3.34A and B, Table 3.10). Further, the shape of the P(r) distribution curve and the Kratky plots indicates that the OH+BARAD^{AFM} is a well-folded, globular domain (Figure 3.34B and C). EOM was used to generate an ensemble of OH+BARAD models to account for the flexible, N-terminal His_6 tag. The best fit of the experimental SAXS data to the theoretical scattering data calculated from an EOM model had a χ^2 of 1.3 (Figure 3.34D and E), indicative of a very good fit.

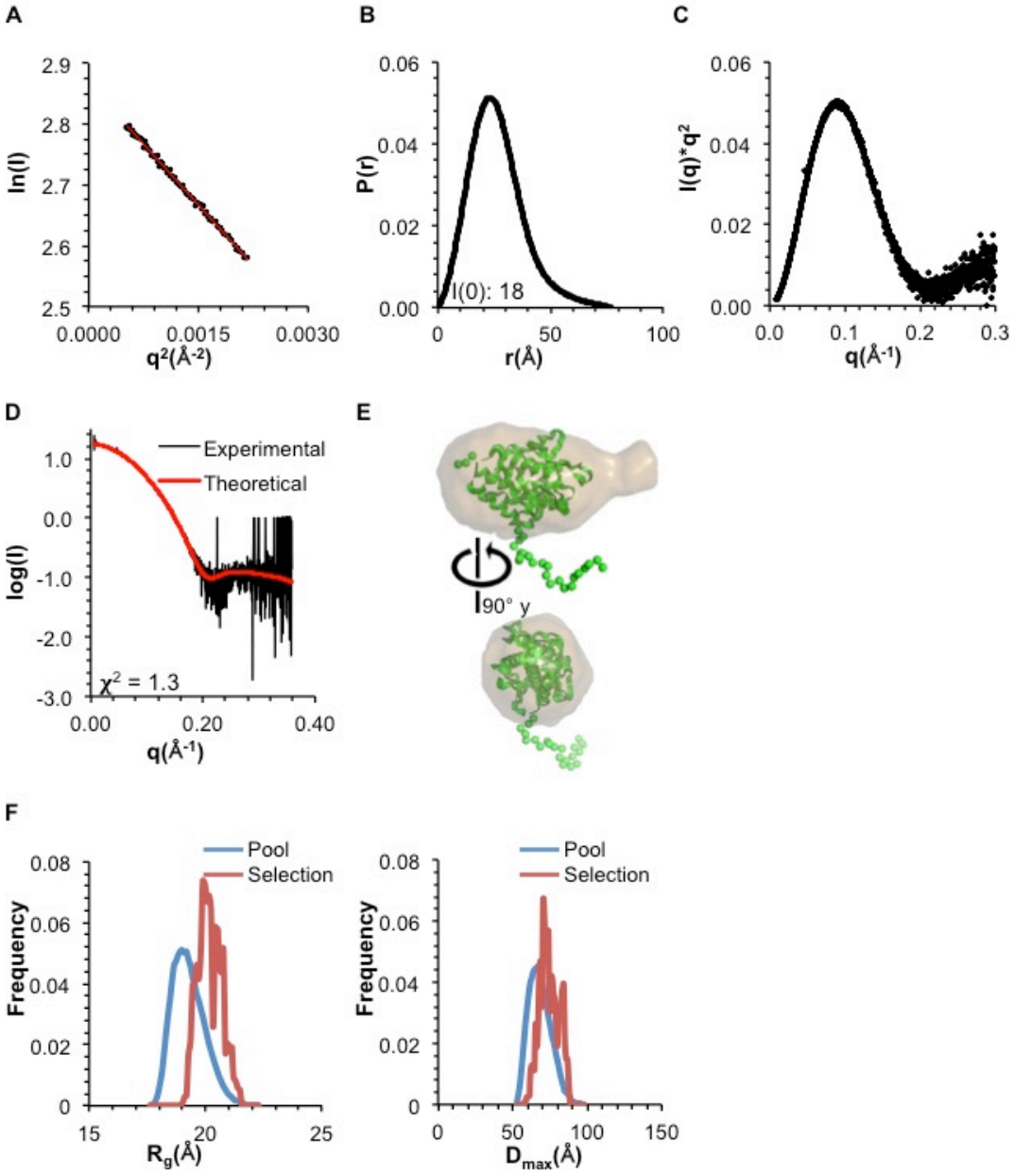

Figure 3.34. SAXS analysis of BECN1 OH+BARAD^{AFM}.
A) Guinier plot. B) P(r) pairwise distribution. C) Kratky plot. D) Theoretical scattering curve calculated from the selected EOM model of BECN1 OH+BARAD^{AFM}, with OH modeled against the BARAD and a flexible His_6 tag, fitted to the experimental data. E) Selected EOM model (green cartoon, flexible residues as spheres) superimposed into the filtered SAXS envelope (sand). F) R_g (left panel) and D_{max} (right panel) distribution of initial pool and selected ensemble of models.

The R_g of the OH+BARAD^{AFM,TETRAD} estimated from the Guinier plot and P(r) distribution is ~23 Å (Figure 3.35A and B, Table 3.10), which is similar to that of the OH+BARAD^{AFM}. Similarly, the shape of the P(r) distribution curve and the Kratky plot reveals that the OH+BARAD^{AFM,TETRAD} is also well-folded and globular (Figure 3.35B and C). However, the 15 Å increase in D_{max} and the extended tail on the P(r)

distribution of the OH+BARAD^{AFM,TETRAD}, (Figure 3.35B) relative to the OH+BARAD^{AFM} (Figure 3.34B) suggest the presence of a more flexible region in the OH+BARAD^{AFM,TETRAD}. Indeed, the normalized Kratky [83] shows that although the OH+BARAD^{AFM,TETRAD} is well folded, it is slightly more flexible than the OH+BARAD^{AFM} (Figure 3.36). This is consistent with the increased coil content estimated by CD.

EOM was used to generate an ensemble of OH+BARAD^{AFM,TETRAD} models to account for the flexible, N-terminal His_6 tag of the OH+BARAD fragments. The best fit of the experimental SAXS data to the theoretical scattering data calculated from an EOM model had a χ^2 of 3.2 (Figure 3.35D and E). Lastly, for both the OH+BARAD^{AFM} and OH+BARAD^{AFM,TETRAD}, EOM analysis shows that the R_g and D_{max} distribution for the selected ensemble is narrower than that of the initial pool of ensembles generated (Figure 3.34F and 3.35F), indicating that both constructs are relatively conformationally homogeneous in solution.

Taken together, the secondary structure and T_m estimations by CD, solution size and shape parameter determinations by SEC-SAXS analyses, and assessment of flexibility by the Kratky plot and EOM, all demonstrate that the OH Tetrad mutation impacts OH+BARAD structure and stability significantly less than it does the CCD.

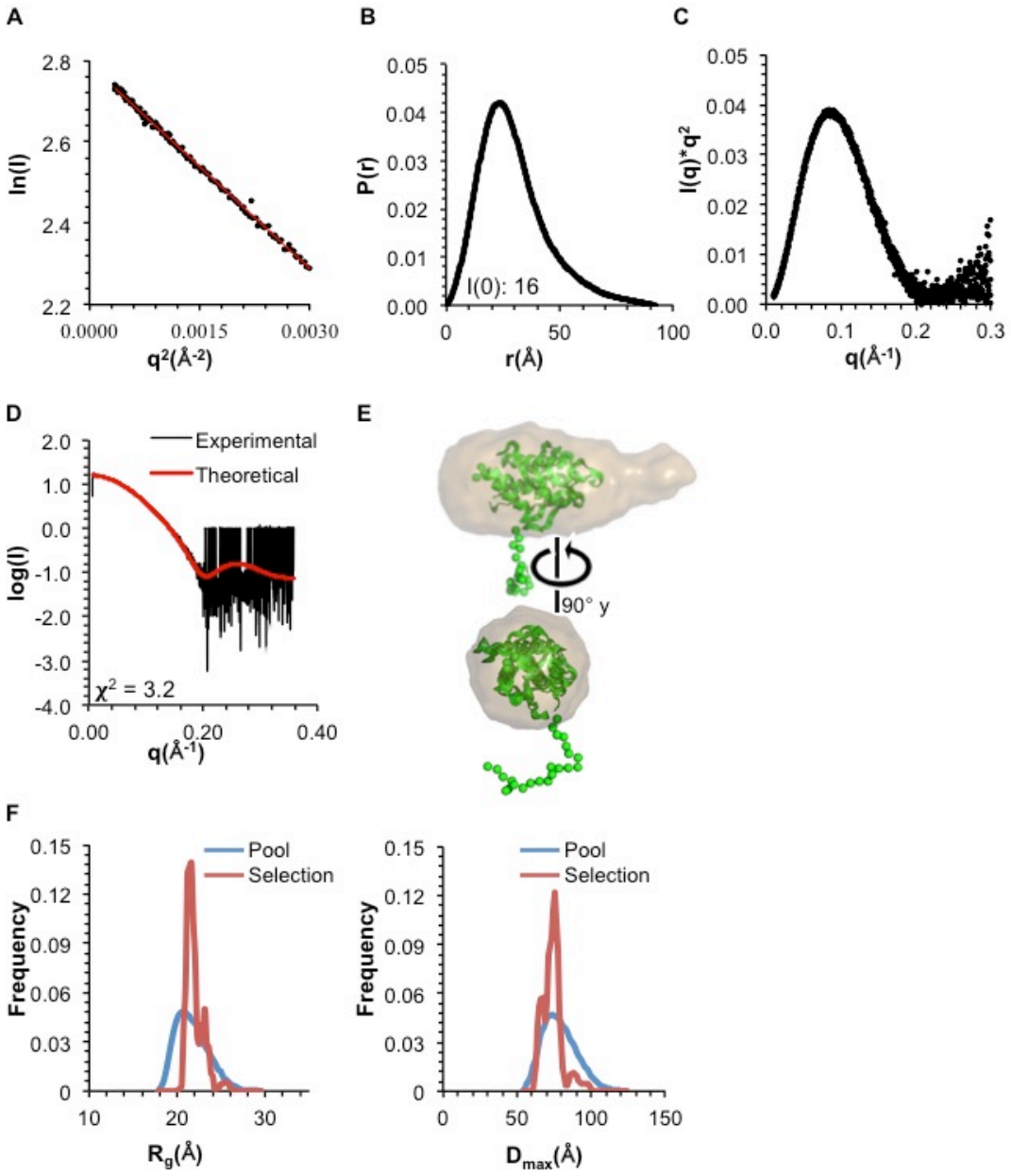

Figure 3.35. SAXS analysis of BECN1 OH+BARAD^{AFM,TETRAD}.
A) Guinier plot. B) P(r) pairwise distribution. C) Kratky plot. D) Theoretical scattering curve calculated from the selected EOM model of BECN1 OH+BARAD^{AFM,TETRAD}, with OH modeled against the BARAD and a flexible His_6 tag, fitted to the experimental data. E) Selected EOM model (green cartoon, flexible residues as spheres) superimposed into the filtered SAXS envelope (sand). F) R_g (left panel) and D_{max} (right panel) distribution of initial pool and selected ensemble of models.

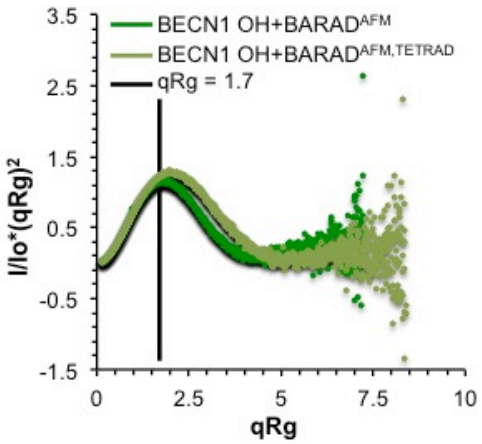

Figure 3.36. The BECN1 OH+BARAD^{AFM,TETRAD} is more flexible than the BECN1 OH+BARAD^{AFM}. Normalized Kratky plot of the OH+BARAD^{AFM} (green) and BECN1 OH+BARAD^{AFM,TETRAD} (light green).

3.3.11. The OH preferentially packs against the BARAD rather than with the CCD homodimer, but transiently samples both conformations

A comparison of the secondary structure content estimated from the CD spectra for the BECN1 CCD-BARAD^{AFM} and CCD-BARAD^{AFM,TETRAD} indicates that the Tetrad mutation decreases helical content by 25 residues and increases strand and coil content by ~4 and ~9 residues respectively (Figure 3.37A, Table 3.9). Thus, secondary structure of the two-domain construct is impacted by the Tetrad mutation.

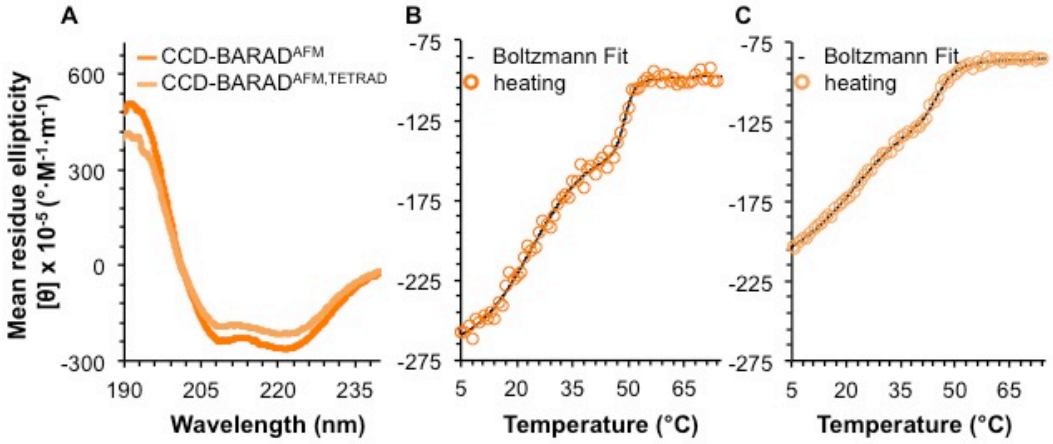

Figure 3.37. BECN1 CCD-BARAD CD spectra and melting curves.
A) CCD-BARAD^{AFM} (orange) and CCD-BARAD^{AFM,TETRAD} (light orange). B) CCD-BARAD^{AFM} melting curve. C) CCD-BARAD^{AFM,TETRAD} melting curve. The Boltzmann fits of the melting curves are shown in black.

In order to determine whether the CCD-BARAD also forms a homodimer like the CCD, we used ITC to quantify the dimer K_d. At 20 °C, self-association of the CCD-BARAD^{AFM} is 1.4-fold weaker than that of the WT CCD, but slightly tighter than that of the CCD^{TETRAD} (Figure 3.30, Figure 3.38A, Table 3.8).

Notably, self-association of the CCD-BARAD^{AFM,TETRAD} is approximately 2-fold weaker than the CCD-BARAD^{AFM} (Figure 3.38, Table 3.8). Taken together, these ITC data indicate that the Tetrad mutation destabilizes both the CCD and CCD-BARAD homodimer. Equally importantly, compared to the WT CCD homodimer, the CCD-BARAD homodimer is substantially dissociated, with a K_d comparable to the CCD^{TETRAD} rather than the CCD, suggesting that the OH helix does not stabilize the homodimer as much in the two-domain fragment.

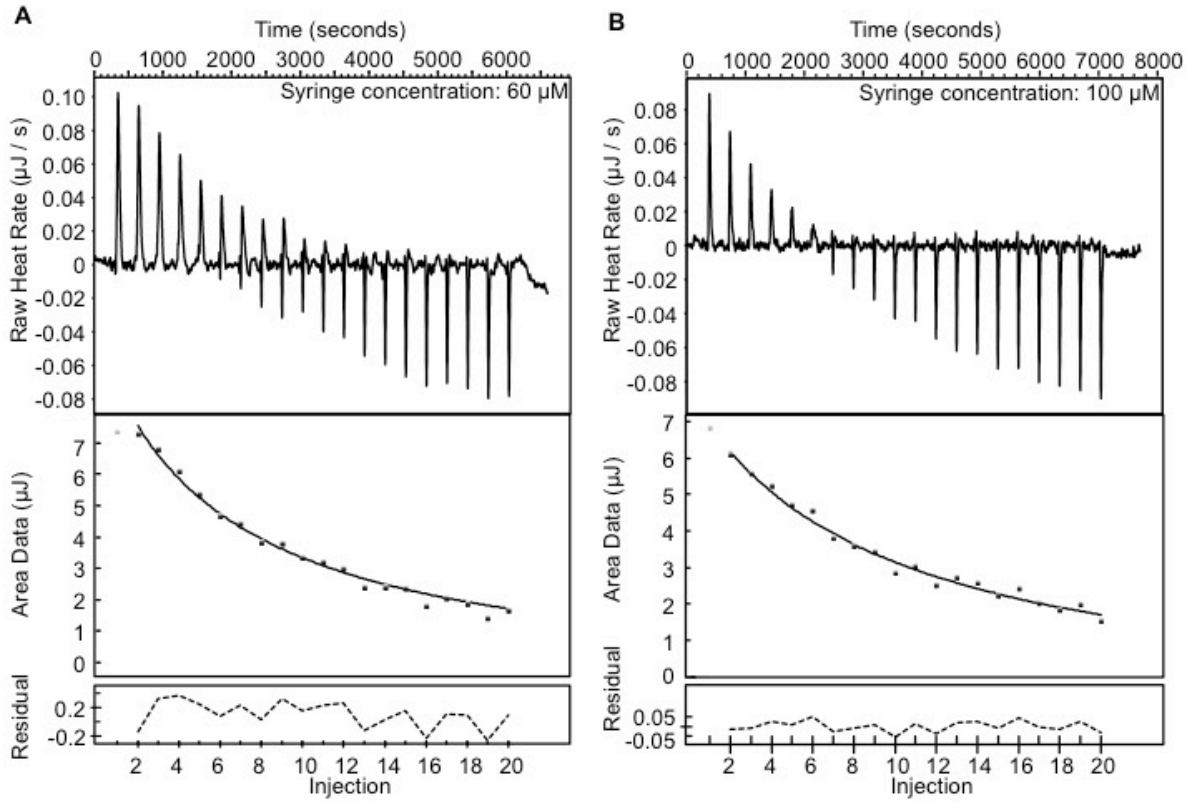

Figure 3.38. Homodimer dissociation ITC profiles of the BECN1 CCD-BARAD and CCD-BARAD^{TETRAD}. A) CCD-BARAD^{AFM}. B) CCD-BARAD^{AFM,TETRAD}.

In order to assess the thermostability of the CCD-BARAD, we used CD to measure the melting temperature of the two-domain fragment. Although protein precipitation prevented acquisition of reversible curves, melting curves for both the CCD-BARAD^{AFM} and CCD-BARAD^{AFM,TETRAD} contain two well-separated transitions (Figure 3.37B and C), indicating that each construct contains two independently folded domains. The higher T_m of 49 °C for the CCD-BARAD^{AFM} corresponds well with that of the OH+BARAD^{AFM} (T_m = 48 °C); but notably, the lower T_m of 24 °C corresponds to that of the CCD^{TETRAD} (T_m = 23 °C) rather than the WT CCD (T_m = 33 °C) (Table 3.9). This suggests that in the CCD-BARAD construct, the OH is not part of the CCD homodimer, but rather, packs against the BARAD.

This is further corroborated by the melting temperatures of 22 °C and 45 °C observed for the CCD-BARAD^{AFM,TETRAD} (Table 3.9), as the 22 °C transition temperature is comparable to that of the lower Tm of the CCD-BARAD^{AFM} as well as that of the CCD^{TETRAD}, whereas the 45 °C transition is a little less than the second T_m of 49 °C for the CCD-BARAD^{AFM} or the T_m of 48 °C for the OH+BARAD^{AFM}, albeit slightly higher than the T_m of 41 °C for the OH+BARAD^{AFM,TETRAD} (Table 3.9). Therefore, the Tetrad mutation does not impact stability of the CCD in the CCD-BARAD construct, but rather appears to marginally destabilize the OH+BARAD, consistent with the disruption of packing of the OH against the BARAD in the two-domain construct.

Thus, the melting temperature analysis indicates that the OH packs against the BARAD in the BECN1 homodimer. However, self-dissociation measurements by ITC indicate that the CCD-BARAD homodimerizes slightly better than the CCD^{TETRAD} and that disruption of the OH packing interface weakens CCD-mediated homodimerization of the CCD-BARAD, although not as much as in the CCD alone. Taken together, these data suggest that in solution, the OH transitions between packing as part of the CCD or against the BARAD, with the predominant state comprising the OH packed against the BARAD.

3.3.12. SAXS indicates the OH packs against the BARAD in the CCD-BARAD protein fragment

We used SEC-SAXS to further investigate the conformation of the OH in the CCD-BARAD^{AFM} and CCD-BARAD^{AFM,TETRAD}. In each case, the R_g estimated from the Guinier plot and P(r) distribution is ~50 Å, (Figure 3.39A and B, Figure 3.40A and B; Table 3.10) and the shape of the P(r) distribution curves and the Kratky plots indicates that both are well-folded, dumbbell-shaped proteins (Figure 3.39B and C, Figure 3.40B and C).

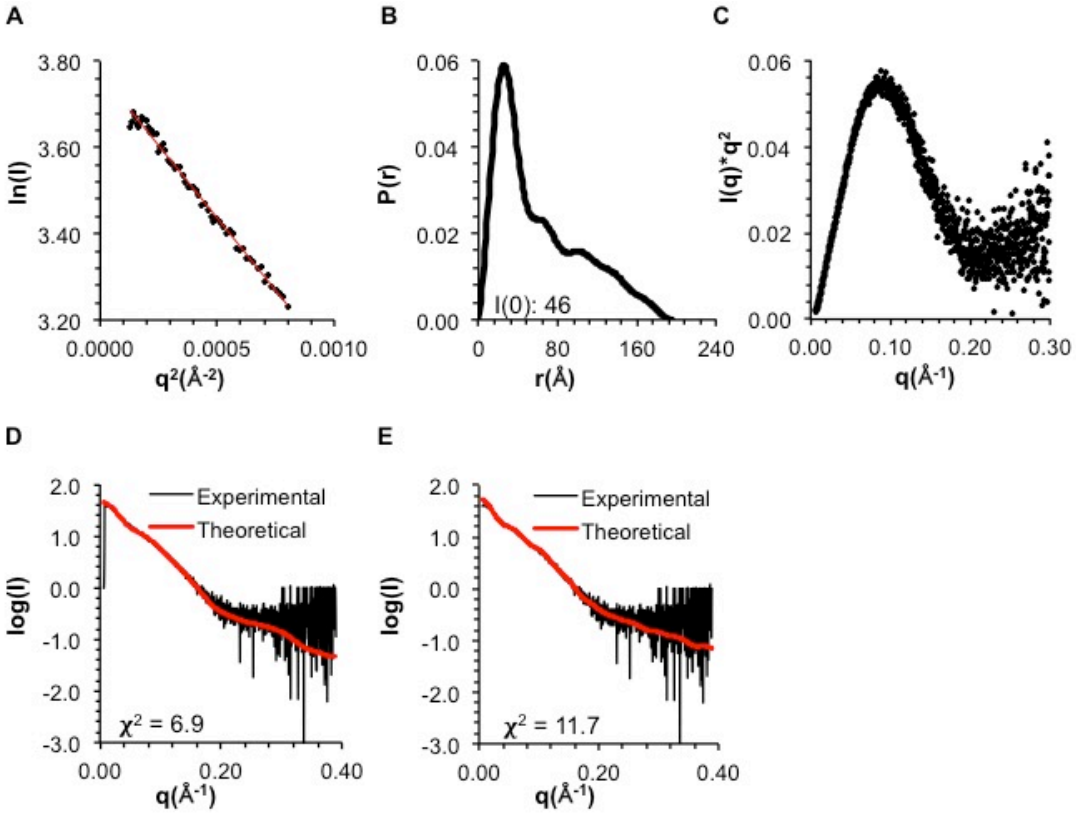

Figure 3.39. SAXS analysis of BECN1 CCD-BARAD^{AFM}.
A) Guinier plot. B) P(r) pairwise distribution. C) Kratky plot. D) Theoretical scattering data calculated from the selected EOM model of the BECN1 CCD-BARAD^{AFM}, with the OH modeled as packed against the BARAD, fitted into the experimental data. E) Theoretical scattering data calculated from the selected EOM model of the BECN1 CCD-BARAD^{AFM}, with the OH modeled as part of the CCD, fitted into the experimental data.

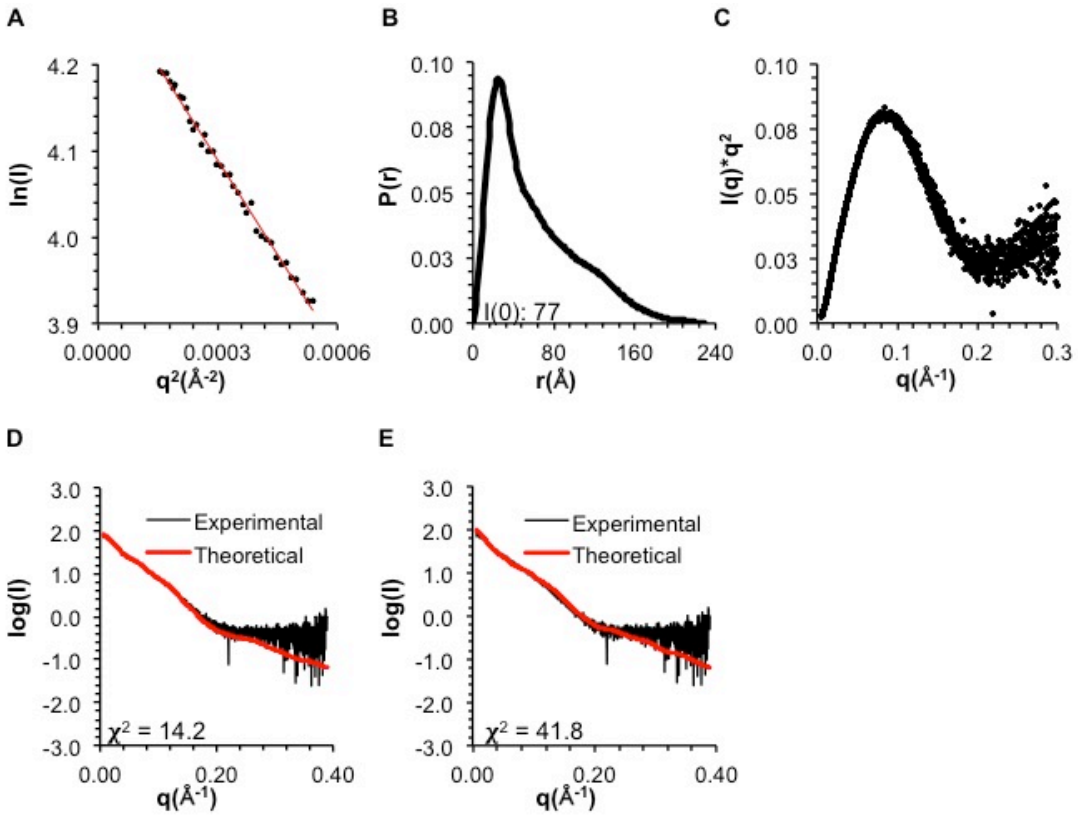

Figure 3.40. SAXS analysis of BECN1 CCD-BARAD^{AFM,TETRAD}.
A) Guinier plot. B) P(r) pairwise distribution. C) Kratky plot. D) Theoretical scattering data calculated from the selected EOM model of the BECN1 CCD-BARAD^{AFM,TETRAD}, with the OH modeled as packed against the BARAD, fitted into the experimental data. E) Theoretical scattering data calculated from the selected EOM model of the BECN1 CCD-BARAD^{AFM,TETRAD}, with the overlap region as flexible loop, fitted into the experimental data.

EOM was used to generate the best ensemble of CCD-BARAD^{AFM} homodimer models. Two separate sets of ensembles were generated: one with the OH modeled within the CCD and the other with the OH modeled as packed against the BARAD. The average R_g of the ensemble wherein the OH is packed against the BARAD (53 Å) is more similar to the experimental data (~50 Å) (Table 3.10) than the ensemble wherein the OH is part of the CCD (65 Å). The theoretical scattering curves calculated from each selected EOM model were also fit to the experimental SAXS data. The theoretical scattering curves calculated from the selected EOM model wherein the OH is packed against the BARAD agrees better with the experimental SAXS data (χ^2 = 6.9) (Figure 3.39D), compared to the theoretical scattering curves calculated from the selected EOM model wherein the OH is part of the CCD (χ^2 = 11.7) (Figure 3.39E), although neither ensemble fits particularly well. Thus, while the EOM analysis also suggests that the OH

likely packs against the BARAD rather than within the CCD in the BECN1 CCD-BARAD homodimer, the poor fits of both modeled states suggest that different conformations are transiently occupied in solution.

EOM was used to generate the best ensemble of CCD-BARAD^{AFM,TETRAD} homodimer models. Two separate sets of ensembles were generated: one with the OH modeled packed against the BARAD and the other with the OH modeled as completely flexible. The theoretical scattering curves calculated from each selected EOM model were also fit to the experimental SAXS data. The theoretical scattering curve calculated from the selected EOM model wherein the OH is packed against the BARAD agrees better with the CCD-BARAD^{AFM,TETRAD} experimental SAXS data (χ^2 = 14.2) (Figure 3.40D), compared to the theoretical scattering curves calculated from the selected EOM model wherein the OH is completely flexible (χ^2 = 41.8) (Figure 3.40E), but worse compared to the CCD-BARAD^{AFM} experimental SAXS data (χ^2 = 6.9) (Figure 3.40D). These results suggest that the Tetrad mutation partially destabilizes the OH in the CCD-BARAD homodimer.

3.3.13. The Tetrad mutant decreases starvation-induced autophagy

Lastly, 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 the Tetrad mutant (V250A, M254A, L261A, and L264A) 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 mutant BECN1 was comparable in starvation and nutrient-rich conditions (Figure 3.41A). We used human breast adenocarcinoma MCF7 cells as they lack detectable endogenous expression of BECN1, resulting in very low basal levels of autophagy [4, 26, 44] even in starvation conditions unless BECN1 is ectopically expressed (Figure 3.41B and C). This allows the effect of BECN1 mutants to be assayed in the absence of endogenous BECN1.

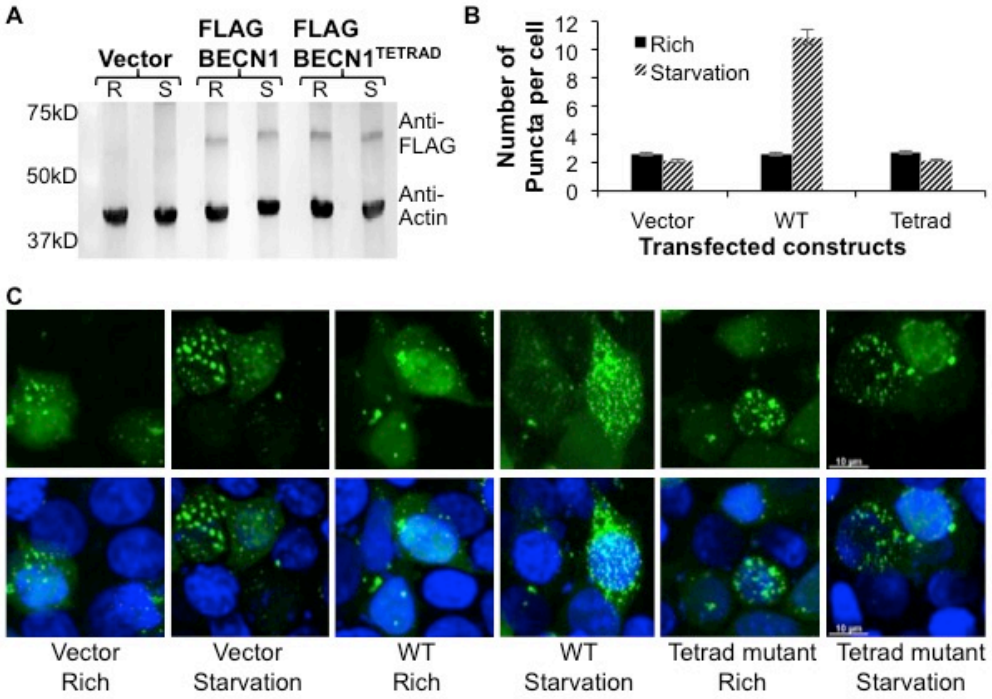

Figure 3.41. Effect of the OH Tetrad mutation on autophagy.
A) Western blot of MCF7 cell extracts showing comparable expression levels of WT and mutant FLAG-BECN1 in nutrient rich (R) and starvation (S) conditions, with Actin as a loading control. B) Light microscopy quantification of discrete GFP-LC3 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) and DAPI (blue) staining in cells grown in starvation or nutrient rich media and transfected with mutant FLAG-BECN1 as indicated.

Transient expression of BECN1 in MCF7 cells did not increase autophagy levels in nutrient-rich conditions ($p = 0.951$ for BECN1 expression versus no expression; Figure 3.41B); but led to a marked increase in autophagy upon starvation ($p = 0.005$ for starved versus nutrient-rich cells; Figure 3.41B). We find that the BECN1 Tetrad mutation does not impact autophagy levels in nutrient rich media ($p = 0.836$ for mutant versus WT; Figure 3.41B); consistent with the lack of impact of BECN1 expression on basal autophagy levels in nutrient-rich conditions. Strikingly however, the BECN1 Tetrad mutation decreases autophagy levels in starvation media ($p = 0.004$ versus WT; Figure 3.41B), 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[18, 33].

3.4. Discussion and conclusions

Our observation that the short linker consisting of three alanines promotes helical content as compared to the long linker in the MBP-BECN1 CCD fusion proteins is consistent with previous observations that the short linker can adopt a helical structure in MBP-fusion proteins[155, 156]. Since this short linker directly precedes the N-terminus of the CCD, which can pack against the conformationally flexible OH, it is possible that the presence of a helical short linker stabilizes helical content of the CCD within the MBP-SL-BECN1 CCD fusion protein. Thus, since the AFM dramatically reduces aggregation and permits large-scale purification of high-quality, homogenous, BARAD-containing constructs that do not require the MBP tag, and also does not appear to impact secondary or tertiary structure of these fragments; all experiments for our final published analyses used AFM-containing BARAD fragments. Utilization of the AFM-containing BARAD constructs eliminated the ambiguities in structural analyses caused by the presence of MBP.

Regarding the main objectives of our study, based on the biophysical and structural analyses reported here, we conclude that the BECN1 C-terminal domains, the CCD and BARAD, are linked by a conformationally-labile helix. Structural superimposition unambiguously demonstrates that the OH residues present in the crystal structure of both the CCD[18, 34] and OH+BARAD[36], cannot simultaneously pack as part of the CCD and against the BARAD. Melting temperature measurements by CD spectroscopy and analyses by SEC-SAXS indicate that mutation of OH residues that are key components of the interface with the CCD partner helix and the BARAD, disrupts CCD structure more than the BARAD. Indeed, these mutations substantially diminish CCD homodimer association. Furthermore, the comparative melting temperature analyses, together with the weakened self-association of the CCD-BARAD relative to the CCD, indicate that the OH preferentially packs against the BARAD in the two-domain CCD-BARAD protein fragment. This preference is supported by our buried surface area calculations, which identify a larger interface between the OH and BARAD than between the OH and the partner helix within the CCD. However, self-dissociation measurements by ITC indicate that the OH also contributes to CCD-BARAD homodimerization. Therefore, contrary to currently established ideas, the OH preferentially packs against the BARAD in the absence of heterologous binding partners but appears to also transiently sample conformations where it forms part of the CCD and contributes to

homodimerization. This is consistent with the recent discovery in full-length BECN1 that mutations analogous to the Tetrad mutant weaken the homodimer, indicating that the OH also contributes to homodimerization of full-length BECN1[157].

We further show that the BECN1 homodimer is unstable at physiological temperatures. Our melting temperature analyses show that nearly half the CCD is unfolded in the two-domain CCD-BARAD fragment at 37 °C. Together, the weak self-association and instability of the CCD-BARAD homodimer suggest that BECN1 likely does not exist in a stable homodimeric state at physiological temperatures, unless stabilized in that state by other protein partners such as the BCL2 proteins. BECN1 homodimerization likely serves as an important means of regulating autophagy in cells. Notably, this may be a regulatory mechanism that evolved in higher eukaryotes such as vertebrates, as yeast do not encode BCL2 proteins, and it is unknown whether yeast VPS30 homodimerizes.

This is significant because upregulation of autophagy by BECN1 requires either ATG14 or UVRAG to displace one molecule of the BECN1 homodimer, which would be facilitated by a weak BECN1 homodimer. The 4.4 Å crystal structure (PDB ID: 5DFZ) of full-length yeast VPS30/BECN1 in complex with VPS34/PI3KC3, VPS15/p150, and VPS38/UVRAG[17], shows that the OH is part of a parallel VPS30:VPS38 CCD heterodimer essential for stabilizing this complex. The 28 Å reconstruction of the quaternary complex of human PI3KC3, p150, BECN1 and ATG14[16], indicates that BECN1 and ATG14 interact in an analogous manner. A SAXS data-constrained model of the BECN1:ATG14 heterodimer indicates that four of the OH interface residues identified in the homodimer also contribute to the heterodimer coiled-coil interface (V250:I165, M254:N169, A257:L172, L261:V176)[18]. Notably, these OH residues are largely conserved, especially in vertebrates (Figure 3.2A); and are predicted to pack against ATG14 residues that are also conserved. Indeed amongst these interacting pairs, the least conserved BECN1 residue, M254, is paired with N169, the least conserved paired residue in ATG14[18]. Thus, the OH contributes critical binding determinants to the parallel CCD heterodimers required for formation of VPS34/PI3KC3 complexes essential for autophagy. Consistent with this, we show that the Tetrad mutation of the OH interface residues abrogates the starvation-induced up-regulation of autophagy but has no effect on basal levels of autophagy.

Our experimental evidence showing that the OH preferentially packs against the BARAD in the BECN1 homodimer suggests a mechanism wherein homodimerization via the CCD results in BECN1 conformations that prevents the BARAD aromatic finger from associating with membranes, while BECN1 heterodimerization with ATG14 or UVRAG disrupts this inhibitory conformation. The crystal structure and HDX-MS experiments of the yeast PI3KC3:p150:BECN1:UVRAG ternary complex show that the aromatic finger interacts with membranes when the OH participates in the CCD interface of the heterodimer [17]. However, formation of a BECN1 homodimer, wherein the OH preferentially packs against the BARAD, would shorten the CCD and involve reorientation of the BARAD relative to that observed in the yeast PI3KC3:p150:BECN1:UVRAG ternary complex. This would reposition the BARAD such that the aromatic finger is no longer able to interact with membranes, thereby inhibiting BECN1-mediated autophagy.

Another important implication of the OH being preferentially packed against the BARAD in the homodimer is that this would release the N-terminal region of the partner helix within the anti-parallel CCD homodimer. The N-terminal region of the human BECN1 CCD (Figure 3.2B) contains a leucine-rich nuclear export signal (NES) that binds the chromosomal protein 1 (CRM1) to enable nuclear export of BECN1[158, 159]. A functional NES is comprised of the sequence motif: $Lx_{(2-3)}Lx_{(2-3)}LxL$. NES residues L184 and L187 are essential for BECN1 nuclear export, BECN1-mediated autophagy, and tumor suppression[158, 159]. In the CCD homodimer, NES residues L184 and L187 pack against OH interface residues, A257 and M254 respectively[18]. Therefore, in order to bind CRM1, the BECN1 NES cannot be packed against the BECN1 OH. Thus, our results showing that the OH preferentially packs against the BARAD in the homodimeric state provide a mechanism of how the NES may be exposed for CRM1-binding to enable nuclear export. Indeed, interaction with CRM1 may stabilize BECN1 conformations wherein the OH is packed against the BARAD. Interestingly, unlike vertebrates, lower eukaryotes do not have a NES located at the N-terminus of the CCD (Figure 3.2B). Therefore, this conformational switch is likely not needed for nuclear export in yeast, as the NES is located between yeast VPS30 residues 12-21[158], which is part of the IDR, rather than the CCD.

In summary, previous studies have established the BECN1 CCD and BARAD as structurally well-defined domains, but here we show that these domains include a region, now termed the overlap helix (OH), that can adopt two distinct, mutually-exclusive packing states, emphasizing overall BECN1

conformational flexibility. Further, in the absence of heterologous partners, the OH transitions between these states with a preference to packing against the BARAD, rather than the CCD as has been commonly assumed. This region likely serves to regulate BECN1 function in autophagy and nuclear transport, and interactions with different partners likely stabilize one packing state or the other to promote these different functions. Understanding how different interactions stabilize different BECN1 conformational states is the next critical step toward understanding the regulation of BECN1-mediated autophagy.

CHAPTER 4. INVESTIGATING THE HIV NEF INTERACTION WITH BECN1

4.1. Introduction

The human immunodeficiency virus (HIV), the virus responsible for the global acquired immune deficiency syndrome (AIDS) epidemic, is highly resilient to therapeutic intervention due to its remarkable ability to mutate and adapt to evolutionary pressure[160, 161]. Patients infected with Negative Factor (Nef)-defective strains of HIV-1 tend to either not progress, or progress very slowly, towards AIDS[66]. This critical role in HIV-1 pathogenesis makes Nef a particularly interesting therapeutic target.

Nef is a small, polymorphic in length (200-215 residues), myristoylated protein primarily localized in the paranuclear region that has been demonstrated to subvert a variety of host cell proteins to promote viral replication[162, 163]. Nef molecules incorporated into virion progeny are commonly cleaved by the viral protease, freeing the C-terminal Core domain from the myristoylated N-terminus[164, 165].

Notably, cellular assays in human cells indicate that Nef is essential for enhancing HIV virion yields in response to autophagy induction, causes accumulation of early autophagic markers, inhibits autophagosome maturation, blocks autophagic degradation of HIV, co-localizes with autophagosome maturation factors including BECN1, and immunoprecipitates from cell lysates with BECN1-containing complexes[68, 69]. This suggests that Nef subverts the autophagy machinery to promote HIV replication.

Immunoprecipitation assays from cell lysates with Nef mutants revealed that Nef residues D174 and D175 in the C-terminal Core domain are required for interaction with BECN1 and autophagosome maturation inhibition[68]. Separate immunoprecipitation assays from cell lysates indicate that BECN1 residues 267-284 of the BARAD are crucial for the BECN1:Nef interaction, but are not required for the BECN1:PI3KC3 interaction[166]. However, the mechanism of the Nef interaction with BECN1-containing complexes is unknown. Therefore, we utilized affinity pull-down assays and ITC with purified protein to determine if BECN1 is sufficient for a direct interaction with Nef.

4.2. Materials and methods

4.2.1. Plasmid preparation

Human BECN1 residues 175-450 (CCD-BARAD) were cloned between the *NcoI* and *NotI* restriction enzyme sites of the pMBP-Parallel-1 bacterial expression vector[146]. The full length Nef SF2 allele residues 1-210 (FL Nef) and Nef SF2 allele residues 62-210 Core domain (Nef Core) prepared as

published previously[167], were kindly provided by Dr. John Jeff Alvarado (University of Pittsburgh School of Medicine, Pittsburgh, USA).

4.2.2. Protein expression and purification

FL Nef and Nef Core were expressed and purified as previously described[167]. This involved expression of FL Nef in *Rosetta(DE3)pLysS E. coli* cells. Soluble protein in the cell lysate (Lysis Buffer: 25 mM Tris pH 8.0, 1 mM EDTA, 8% (v/v) glycerol, 2 mM DTT, 50 mM NaCl, 0.4 % (w/v) β-octyl glucoside) from a 1 L expression culture was purified to homogeneity using a 10/100 MonoQ anion exchange column chromatography (MonoQ Buffer A: 25 mM Tris pH 8.0, 1 mM EDTA, 8% (v/v) glycerol, 2 mM DTT, 50 mM NaCl; MonoQ Buffer B: MonoQ Buffer A plus 1 mM NaCl). 0.4 % (w/v) β-octyl glucoside was subsequently added to minimize aggregation; protein was then concentrated to 5 mg/mL, dialyzed against MonoS Buffer A (MonoS Buffer A: 25 mM HEPES sodium salt pH 7.5, 1 mM EDTA, 1 mM DTT, 8% (v/v) glycerol; MonoS Buffer B: MonoS Buffer A plus 1 mM NaCl), and further purified by cation exchange chromatography on a 10/100 GL MonoS column. 0.4 % (w/v) β-octyl glucoside was subsequently added; then protein was concentrated to 5 mg/mL and purified by SEC using a 16/60 HiLoad Superdex 200 column in 20 mM Tris HCl pH 8.0, 150 mM NaCl, 8% (v/v) glycerol, 2 mM DTT.

Nef Core was expressed in *Rosetta(DE3)pLysS E. coli* cells. Soluble protein in the cell lysate (Lysis Buffer: 25 mM Tris pH 8.0, 1 mM EDTA, 8% (v/v) glycerol, 2 mM DTT, 50 mM NaCl, 0.2 % (w/v) β-octyl glucoside) from a 1 L expression culture was purified using a 10/100 MonoQ anion exchange column (MonoQ Buffer A: 25 mM Tris pH 8.0, 1 mM EDTA, 8% (v/v) glycerol, 2 mM DTT, 50 mM NaCl; MonoQ Buffer B: MonoQ Buffer A plus 1mM NaCl). 0.2 % (w/v) β-octyl glucoside was subsequently added to minimize aggregation; then protein was concentrated to 5 mg/mL and purified to homogeneity by SEC using a 16/60 HiLoad Superdex 200 column in 20 mM Tris HCl pH 8.0, 150 mM NaCl, 8% (v/v) glycerol, 2 mM DTT.

For the MBP-LL-BECN1 CCD-BARAD, *E. coli* Arctic Express DE3 cells were transformed and grown in LB medium with 100 µg/ml ampicillin at 37 °C to an A_{600} of ~ 0.7 before equilibrating the temperature to 13 °C. Once the cells achieved an A_{600} of ~ 1.2, they were induced by the addition of 0.5 mM IPTG and expressed overnight. Fusion protein in the cell lysate (Lysis buffer: 25 mM Tris HCl pH 8, 150 mM NaCl, 10% (v/v) glycerol, 1mM EDTA, 2 mM DTT, and 1 cOmplete, EDTA-free protease inhibitor

cocktail tablet (MilliporeSigma)) from a 1 L expression culture was purified by amylose affinity chromatography using a 10 mL amylose resin gravity column, washed with 10 column volumes of wash buffer (wash buffer: 25 mM Tris HCl pH 8, 150 mM NaCl, 10% (v/v) glycerol, 1mM EDTA, 2 mM DTT), and eluted with wash buffer plus 20 mM maltose. The protein was then purified by anion exchange chromatography on a 5/50 GL MonoQ column (GE Life Sciences, Pittsburgh, US) (MonoQ Buffer A: 25 mM Tris HCl pH 8, 10% (v/v) glycerol, 1 mM EDTA, 1mM DTT; MonoQ Buffer B: MonoQ Buffer A plus 1M NaCl). Protein was eluted from the MonoQ column via step elutions at 190 mM and 220 mM NaCl. The protein was then purified to homogeneity by SEC using a 16/60 Superdex 200 column (GE Life Sciences, Pittsburgh, US) in 20 mM Tris HCl pH 8, 150 mM NaCl, 10% glycerol, 2mM DTT.

4.2.3. Pull-down

For a 1:1 molar ratio, 0.9 mg of FL Nef and 0.3 mg MBP-BECN1 CCD-BARAD were added to 250 µL amylose resin, washed with Wash Buffer (25 mM Tris pH 8.0, 50 mM NaCl, 5% (v/v) glycerol, 1 mM EDTA, 2 mM DTT) until the A_{280} = 0, and finally eluted with wash buffer plus 20 mM maltose. Fractions were evaluated by SDS-PAGE stained with Coomassie Blue. As a control, 0.9 mg of FL Nef was added to 250 µL amylose resin and treated in the same manner.

4.2.4. ITC

Protein samples were co-dialyzed against 25 mM HEPES sodium salt pH 7.5, 150 mM NaCl, 10 % (v/v) glycerol, and 2 mM TCEP. ITC experiments were performed at 20 °C using a Low Volume Gold Nano isothermal titration calorimeter (TA Instruments, New Castle, US). 100 µM Nef Core was titrated into 40 µM MBP-BECN1 CCD-BARAD. Data were analyzed with the NanoAnalyze Software (TA Instruments, New Castle, US).

4.3. Results

4.3.1. Protein expression and purification

FL Nef was purified by sequential anion exchange chromatography, cation exchange chromatography, and SEC. The single peak on SEC and single band on SDS-PAGE indicate the FL Nef was pure and homogenous (Figure 4.1). The apparent molecular weight of FL Nef calculated based on the elution volume from SEC is 20 kD, which is 1.2-times smaller than the theoretical molecular weight of 24 kD for the FL Nef monomer calculated from the amino acid sequence. This is consistent with the FL

Nef forming a monomer. The final yield of the purified FL Nef obtained from one liter of bacterial culture was 6 mg.

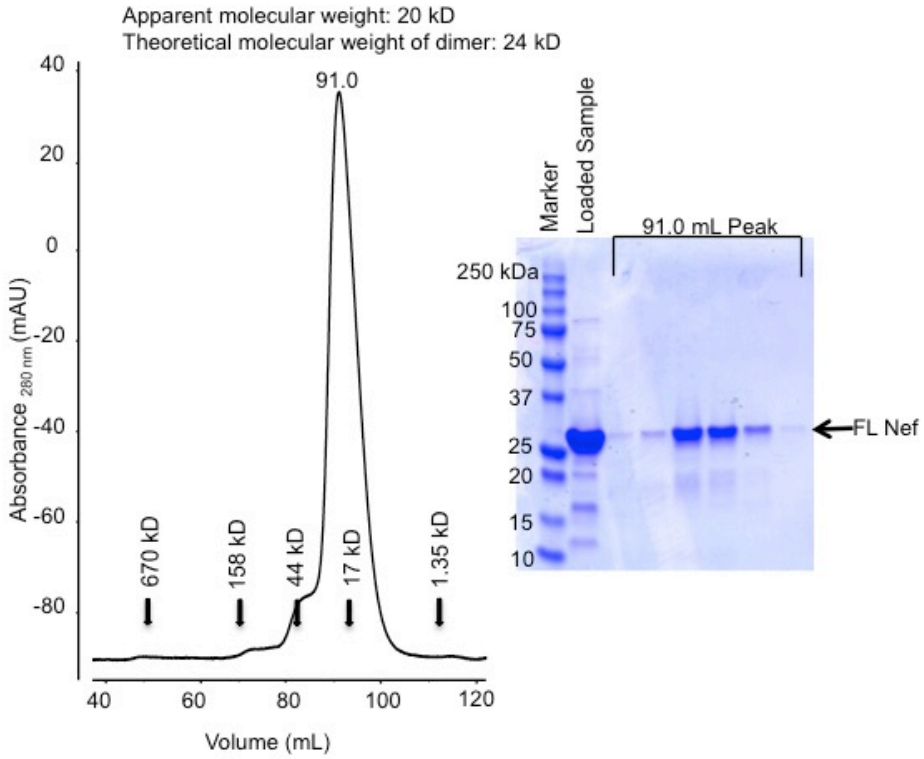

Figure 4.1. Size exclusion chromatogram and the corresponding SDS-PAGE of FL Nef.
The elution positions of SEC standards are indicated by arrows.

Nef Core was purified by sequential anion exchange chromatography and SEC. The apparent molecular weight of Nef Core calculated based on the elution volume from SEC of the 71.9 mL peak is 106 kD, which is ~6 times greater than the theoretical molecular weight of 18 kD for the Nef Core monomer calculated from the amino acid sequence (Figure 4.2), indicating oligomerization. The final yield of the purified Nef Core obtained from one liter of bacterial culture was 5 mg.

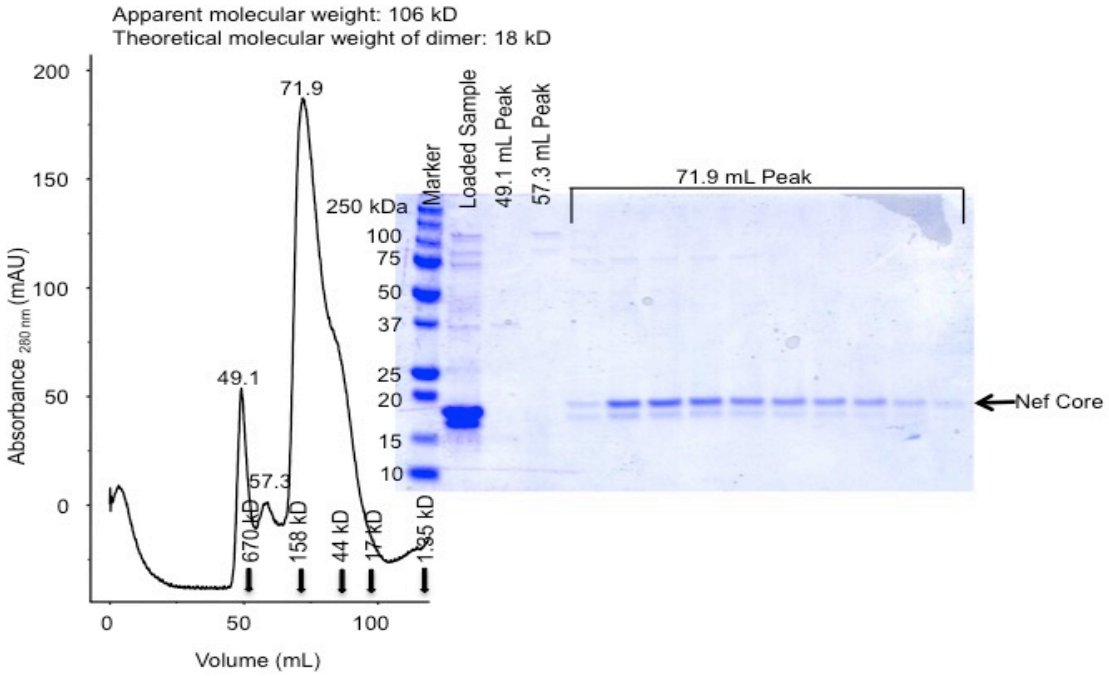

Figure 4.2. Size exclusion chromatogram and the corresponding SDS-PAGE of Nef Core. The elution positions of SEC standards are indicated by arrows.

MBP-LL-BECN1 CCD-BARAD was purified by sequential amylose chromatography, anion exchange chromatography, and SEC. The single peak on SEC and single band on SDS-PAGE indicate the MBP-LL-BECN1 CCD-BARAD was pure and homogenous (Figure 4.3). The apparent molecular weight of MBP-LL-BECN1 CCD-BARAD calculated based on the elution volume from SEC is 178 kD, which is 1.2-times greater than the theoretical molecular weight of 151 kD for the MBP-LL-BECN1 CCD-BARAD dimer calculated from the amino acid sequence. This is consistent with the MBP-LL-BECN1 CCD-BARAD forming rod-shaped dimer. The final yield of the purified MBP-LL-BECN1 CCD-BARAD obtained from one liter of bacterial culture was 4 mg.

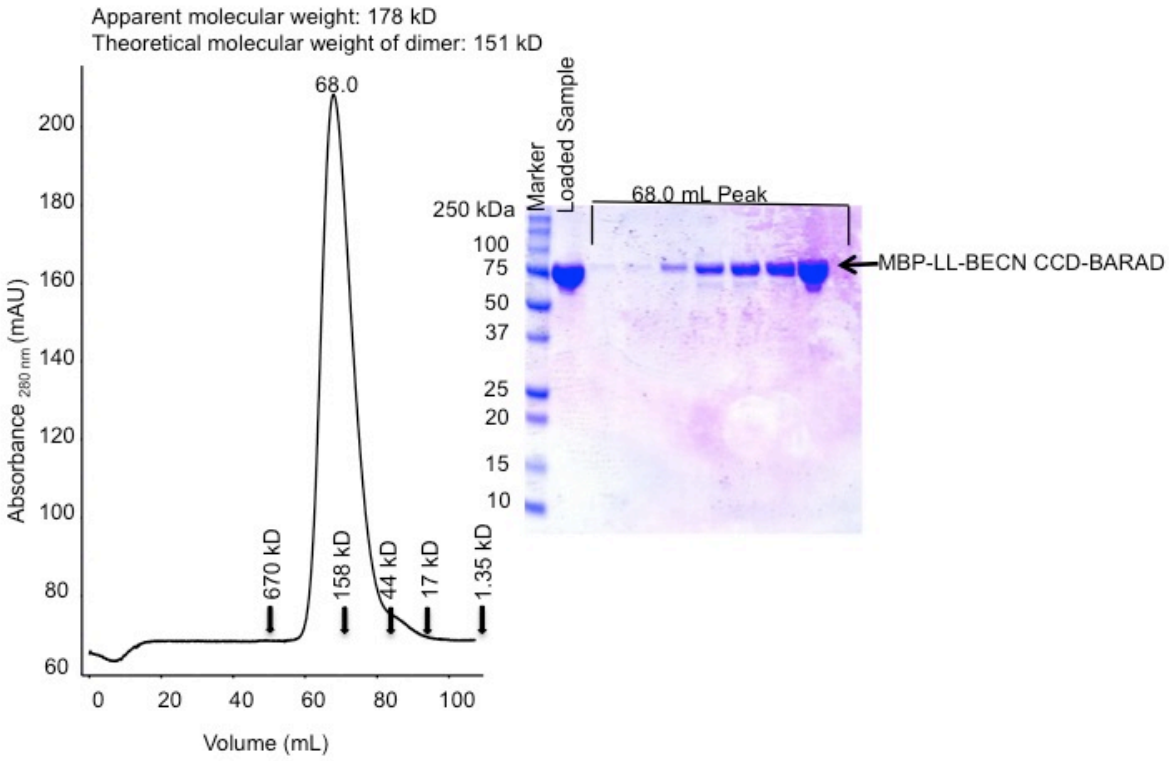

Figure 4.3. Size exclusion chromatogram and the corresponding SDS-PAGE of MBP-LL-BECN1 CCD-BARAD.
The elution positions of SEC standards are indicated by arrows.

4.3.2. Nef and BECN1 do not directly interact via affinity pull-down assays

In order to determine if Nef interacts directly with BECN1, we performed amylose affinity pull-downs of purified FL Nef and MBP-BECN1 CCD-BARAD. We first confirmed that FL Nef does not bind amylose resin in the absence of an MBP-tagged binding partner. FL Nef from two separate protein preps (FL Nef #1 and FL Nef #2) did not bind the amylose resin in the absence of MBP-tagged BECN1 (Figure 4.4A).

MBP-BECN1 CCD-BARAD does not pull down FL Nef (Figure 4.4B). MBP-BECN1 CCD-BARAD binds well to the amylose resin, with only a small fraction coming off in the flow through, whereas most of the FL Nef comes off in the flow through. The remainder of the FL Nef washes off of the amylose resin and does not co-elute with MBP-BECN1 CCD-BARAD. This suggests either that Nef and BECN1 do not interact directly or that BECN1 is insufficient for a direct interaction with Nef.

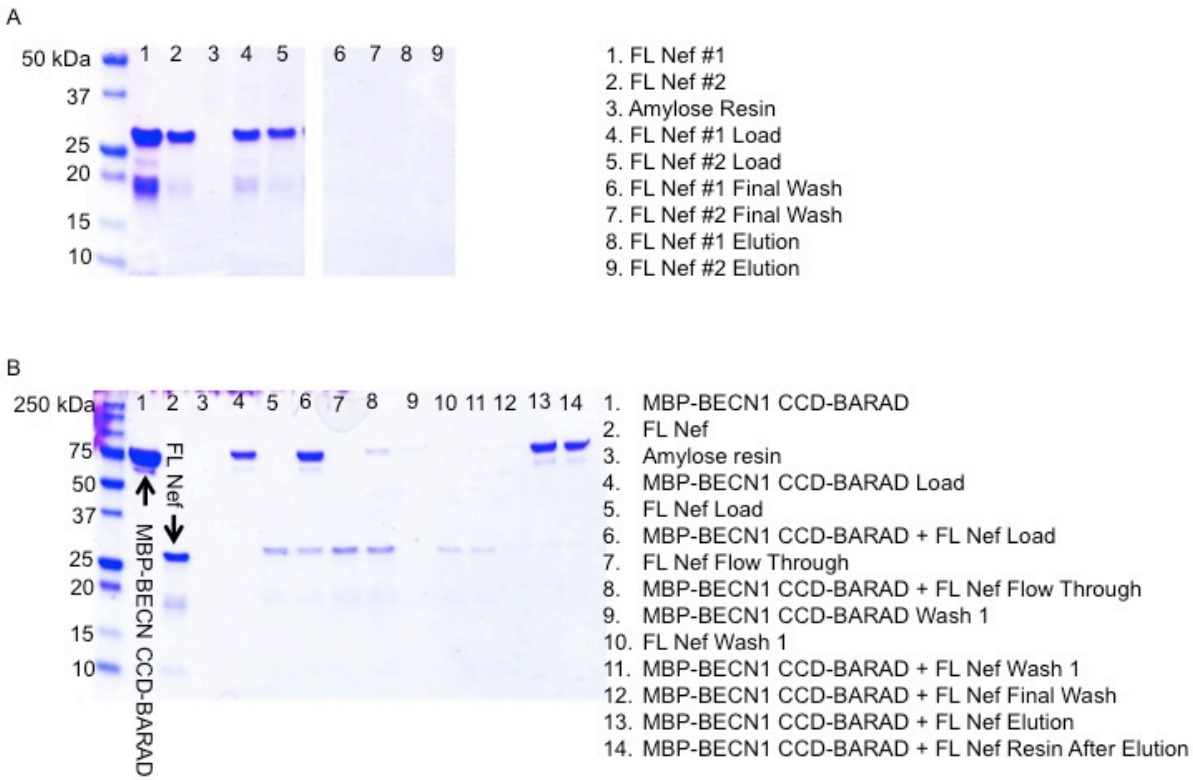

Figure 4.4. Pull-down assay to assess binding between FL Nef and BECN1 CCD-BARAD.
A) SDS-PAGE of FL Nef pull-down on amylose resin. FL Nef from two separate protein preps was tested (labeled FL Nef #1 and FL Nef #2). B) SDS-PAGE of FL Nef and MBP-BECN1 CCD-BARAD pull-down on amylose resin.

4.3.3. ITC of MBP-BECN1 CCD-BARAD and Nef Core

Since Nef residues Asp^{174} and Asp^{175} in the C-terminal Core domain were found to be required for the BECN1 interaction and autophagosome maturation inhibition[68], we utilized ITC to determine if a direct interaction could be detected between Nef Core and BECN1 by titrating Nef Core into MBP-LL-BECN1 CCD-BARAD; however, no binding was detected (Figure 4.5). Therefore, consistent with affinity pull-down assays performed with purified protein, ITC confirms that Nef and BECN1 either do not interact directly or that BECN1 is insufficient for a direct interaction with Nef

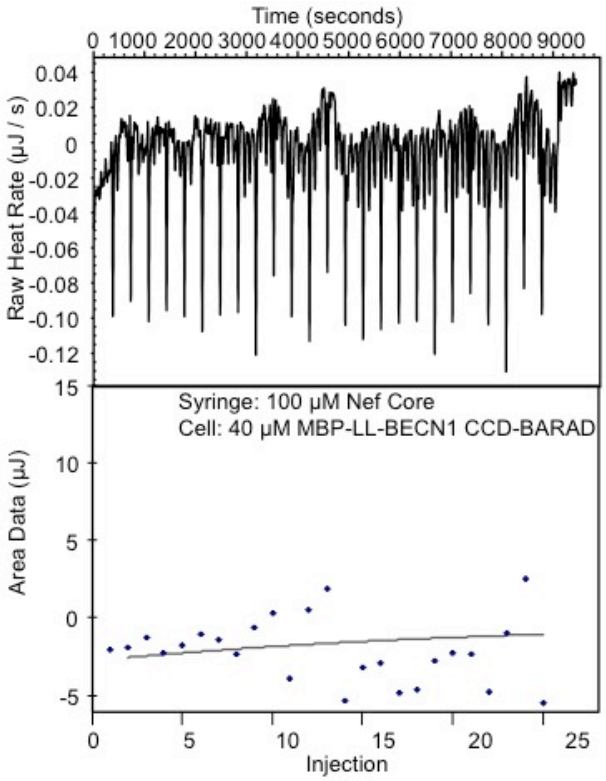

Figure 4.5. ITC profiles of Nef Core titrated into MBP-LL-BECN1 CCD-BARAD.

4.4. Discussion and conclusions

Based on the *in vitro* analyses with purified protein reported here, we conclude that the BECN1 BARAD in the context of the BECN1 homodimer is insufficient to mediate a direct interaction with Nef. Previously published cellular co-immunoprecipitation assays indicated that C-terminal residues of the Nef Core mediate the interaction with BECN1-containing complexes[68] and that BECN1 residues 267-284 of the BARAD are crucial for the BECN1:Nef interaction[166]. However, our affinity pull-down assays and ITC indicate that neither FL Nef nor the Nef Core interacts with the MBP-BECN1 CCD-BARAD homodimer.

Previous studies indicate that Nef is essential for HIV progression towards HIV[66] and the observed augmentation of HIV virion yields in response to autophagy induction[68]. Notably, Nef is required for accumulation of the early autophagic marker LC3-II, which is typically degraded in autolysosomes, and prevents acidification of the autolysomes[68]. This role in arresting maturation suggests that Nef interacts with BECN1 in the context of PI3KC3 Complex II, which consists of BECN1 (VPS30 in yeast),

PI3KC3 (VPS34 in yeast), p150 (VPS15 in yeast), and UVRAG (VPS38 in yeast) since Complex II mediates autophagosome maturation[11, 12, 14, 15, 57].

BECN1 residues 267-284 (residues 323-340 in yeast VPS30) crucial for the BECN1:Nef interaction[166], are located at the N-terminus of the BARAD immediately following the OH. The 4.4 Å crystal structure (PDB ID: 5DFZ) of full-length yeast VPS30 in Complex II[17] shows that the OH is part of a parallel VPS30:VPS38 CCD heterodimer such that the critical Nef-binding BECN1 (VPS30) residues are largely solvent exposed as part of a coil immediately following the OH and the first β sheet of the BARAD. However, a 1.6 Å crystal structure (PDB ID: 4DDP) of the OH+BARAD[36] shows that the OH packs against one face of the first β sheet of the BARAD. We have shown that the BECN1 homodimer transitions between these two mutually exclusive states with a preference to packing against the BARAD, rather than the CCD[168]. It is possible that the BECN1 homodimer preference for packing the OH against the BARAD prevents Nef from interacting with the BECN1 homodimer. Therefore, we cannot rule out the possibility that Nef interacts directly with the BECN1 BARAD in the context of the BECN1:UVRAG heterodimer wherein the Nef-binding BECN1 residues are less occluded.

Alternatively, perhaps the Nef:BECN1 interaction requires stabilization by interactions with other PI3KC3 Complex II proteins in addition to BECN1. The UVRAG (VPS38) BARA2 and p150 (VPS15) WD40 domains in Complex II[17] create a pocket with a diameter of about 25 Å on the side of the first β sheet of the BARAD opposite to where the OH packs in the homodimer. The 3.0 Å crystal structures of unbound Nef Core (PDB ID: 1AVV) and the Nef Core bound to the SH3 domain of FYN (PDB ID: 1AVZ)[163], as well as the 1.9 Å crystal structure (PDB ID: 4U5W) of the Nef Core bound to the SH3-SH2 domains of the kinase Hck[167] show that Nef Core is a small globular protein with a diameter of about 30 Å comprising an N-terminal ordered coil, followed by 2 α-helices, an ordered coil, a β sheet consisting of two strands, and two short C-terminal α-helices. The N-terminal coil and the first and second α-helices of the Nef Core interact with the SH3 domains in Nef:FYN and Nef:Hck complexes. The Nef residues Asp^{174} and Asp^{175} in the C-terminal Core domain determined to be required for interaction with BECN1 and for autophagosome maturation inhibition[68] are located within a region of 31 disordered residues between the strands of the β sheet. Therefore, it is also possible that a direct Nef:BECN1 interaction requires stabilization by residues the p150 WD40 domain within the PI3KC3 Complex II.

In summary, previous studies have established that residues from the BECN1 BARAD and Nef Core are essential for the BECN1:Nef interaction, but here we show that these domains are insufficient to mediate a direct interaction. Therefore, Nef likely interacts with BECN1 in the context of the PI3CK3 complex II; although, we cannot rule out a direct interaction between Nef and the BECN1 heterodimer observed in this complex. Understanding the mechanism by which Nef interacts with BECN1 in the PI3CK3 complex II is the next critical step toward understanding how Nef subverts the autophagy machinery to enhance HIV pathogenesis.

CHAPTER 5. KSHV V-BCL2 INTERACTION WITH BECN1 AND NOXA BH3DS

5.1. Introduction

KSHV is a human γHV associated with Kaposi's sarcoma (KS), primary effusion lymphomas, and Castleman's disease[169]. KSHV encodes a vBcl-2 mimic that is expressed in KSHV lesions[170], which promotes survival of KSHV-infected cells and may contribute to tumor development[169]. Like host cell anti-apoptotic Bcl-2s, γHV Bcl-2s inhibit apoptosis[171]. Further like host cell and other γHV anti-apoptotic Bcl-2s[43-47, 172], KSHV Bcl-2 also appears to down-regulate autophagy by inhibiting autophagosome formation by binding the BH3D of BECN1[26]. The solution structure of the KSHV vBcl-2 indicates that the overall fold is similar to cellular Bcl-2 proteins[173]. However, binding affinities for various BH3Ds by the KSHV vBcl-2 vary widely from the affinities for the same BH3Ds by host cell Bcl-2 proteins[44, 174]. This suggests that the mechanism of binding by the KSHV vBcl-2 to BH3Ds differs from that of host cell proteins.

A BECN1 BH3D G120E/D121A double mutant has been shown to selectively bind the mouse γHV Bcl-2 homolog M11, but not cellular Bcl-2 proteins, to down-regulate M11-mediated inhibition of autophagy[140]. Although the KSHV vBcl-2 has previously been demonstrated to bind the wild-type BECN1 BH3D[44], we show via ITC that the BECN1 BH3D G120E/D121A double mutant does not bind the KSHV vBcl-2. This suggests that the determinants of binding specificity for BH3Ds differ for the KSHV vBcl-2 and M11. However, to date, there is no crystal structure deposited in the PDB of KSHV vBcl-2 bound to any BH3D. Therefore, solving the KSHV vBcl-2 structure in complex with a pro-apoptotic BH3D may permit identification of the molecular determinants that define binding specificity for the KSHV vBcl-2.

Since mouse NOXA (mNOXA) has been shown to tightly bind KSHV vBcl-2, Mcl-1, and A1, but binds weakly or not at all to the other human anti-apoptotic Bcl-2 paralogs[174, 175], we quantified the thermodynamics of binding between the KSHV vBcl-2 and mNOXA via ITC. However, protein precipitation issues hindered acquisition of sufficient quantities of high quality, homogenous protein to perform ITC replicates required to obtain reliable thermodynamics parameters and to perform other biophysical and structural analyses of the interaction.

5.2. Materials and methods

5.2.1. Plasmid preparation and peptide synthesis

KSHV residues 1-146 were cloned between the *NcoI* and *NotI* restriction enzyme sites of the pET21d(+) bacterial expression vector such that the construct included a C-terminal His_6 tag. Site-directed mutagenesis was used to introduce N67D and V117A point mutations to improve protein solubility[173]. Peptides corresponding to the mouse NOXA (mNOXA) BH3D and the BECN1 BH3D G120E/D121A mutant were chemically synthesized, HPLC purified to >90% purity, and purity confirmed by electrospray mass spectrometry (RS Synthesis). For each peptide, a 3 mM stock solution in 25 mM Tris HCl pH 7.8, 2mM BME was prepared. The mNOXA sequence was RAELPPEFAAQLRKIGDKVYCTWSAP (2.9 kD, ε_{280} = 6,990 M^{-1} cm^{-1}). The BECN1 BH3D G120E/D121A sequence was DGGTMENLSRRLKVTEALFDIMSGQT (2.9 kD, ε_{280} = 0 M^{-1} cm^{-1}).

5.2.2. Protein expression and purification

E. coli BL21-Gold(DE3) cells were transformed and grown in LB medium with 100 µg/mL ampicillin at 37 °C to an A_{600} of 0.350 before equilibrating the temperature to 30 °C. Once the cells achieved an A_{600} of ~1.2, they were induced by the addition of 0.5 mM IPTG and expressed overnight at 20 °C. Soluble protein in the cell lysate (Lysis Buffer: 20 mM Tris HCl pH 7.8, 150 mM NaCl, 2mM BME, 40 mM Imidazole, and 1 cOmplete, EDTA-free protease inhibitor cocktail tablet (MilliporeSigma)) from a 1 L expression culture was by nickel IMAC using a 5 mL HisTrap column (GE Life Sciences, Pittsburgh, US), washed with 10 column volumes of nickel wash buffer consisting of 20 mM Tris HCl pH 7.8, 150 mM NaCl, 2mM BME, and 40 mM Imidazole, and then eluted with wash buffer containing 200 mM imidazole. The protein was then desalted on a HiPrep 16/10 desalt column (GE Life Sciences, Pittsburgh, US) in 50 mM Tris HCl pH 7.8, 2mM BME and then purified to homogeneity by SEC on a 10/30 Superdex 200 linked in tandem to a 10/30 Superdex 75 column (GE Life Sciences, Pittsburgh, US) in the same buffer.

5.2.3. Complex SEC and crystal screening

300 µL of 170 µM KSHV vBcl-2 was injected onto a 10/30 Superdex 200 linked in tandem to a 10/30 Superdex 75 SEC column, which was run with SEC buffer consisting of 50 mM Tris HCl pH 7.8, 2mM BME. Purified KSHV vBcl-2 and mNOXA were mixed with a 1:2 molar ratio and incubated at 4 °C for one hour. 300 µL of the mixture (170 µM complex) was injected onto the same tandem SEC columns using

the same SEC buffer. The KSHV vBcl-2:mNOXA complex peak from SEC was concentrated and screened for crystallization at 20 °C against Microlytic screens MCSG 1-4 (Anatrace, Maumee, OH, USA) using a 1:1, 2:1, or 3:1 mixture of protein stock (1 mg/mL complex in 50 mM Tris HCl pH 7.8 and 2mM BME) and well solution using a Gryphon LCP crystal screen dispenser (Art Robbins Instruments).

5.2.4. ITC

Protein and peptide samples were co-dialyzed against 50 mM Tris HCl pH 7.8 and 2 mM BME. In order to assess the thermodynamics of binding of the KSHV vBcl-2 to the BH3D peptides, ITC experiments were performed in duplicate at 14 °C using a Low Volume Gold Nano isothermal titration calorimeter (TA Instruments, New Castle, US). 2 mM BECN1 BH3D G120E/D121A or 2 mM mNOXA BH3D were separately titrated into 200-250 μM KSHV vBcl-2. Data were blanked against a peptide-into-buffer titration curve and analyzed with the NanoAnalyze Software (TA Instruments, New Castle, US) using an independent binding model.

5.3. Results

5.3.1. KSHV vBcl-2 purification

KSHV vBcl-2 was purified to homogeneity using a 5 mL HisTrap column; however, the protein precipitated severely if left overnight in nickel elution buffer. Therefore, elution fractions were immediately desalted on a desalt column to minimize precipitation and subsequently purified by SEC. KSHV vBcl-2 itself elutes as two peaks, one at 16.0 mL likely corresponding to aggregated protein, and another peak at 27.2 mL (Figure 5.1). The apparent molecular weight of KSHV vBcl-2 calculated based on the elution volume of the second SEC peak is 22.2 kD (Figure 5.1), which is 1.3-times greater than the theoretical molecular weight of 17.7 kD for the KSHV vBcl-2 monomer calculated from the amino acid sequence. This is inconsistent with any oligomerization, indicating the KSHV vBcl-2 is a monomer. The 27.2 mL SEC peak could be concentrated to a maximum of 2.5 mg/mL before precipitation was observed. The final yield obtained from one liter of bacterial culture was 9.5 mg.

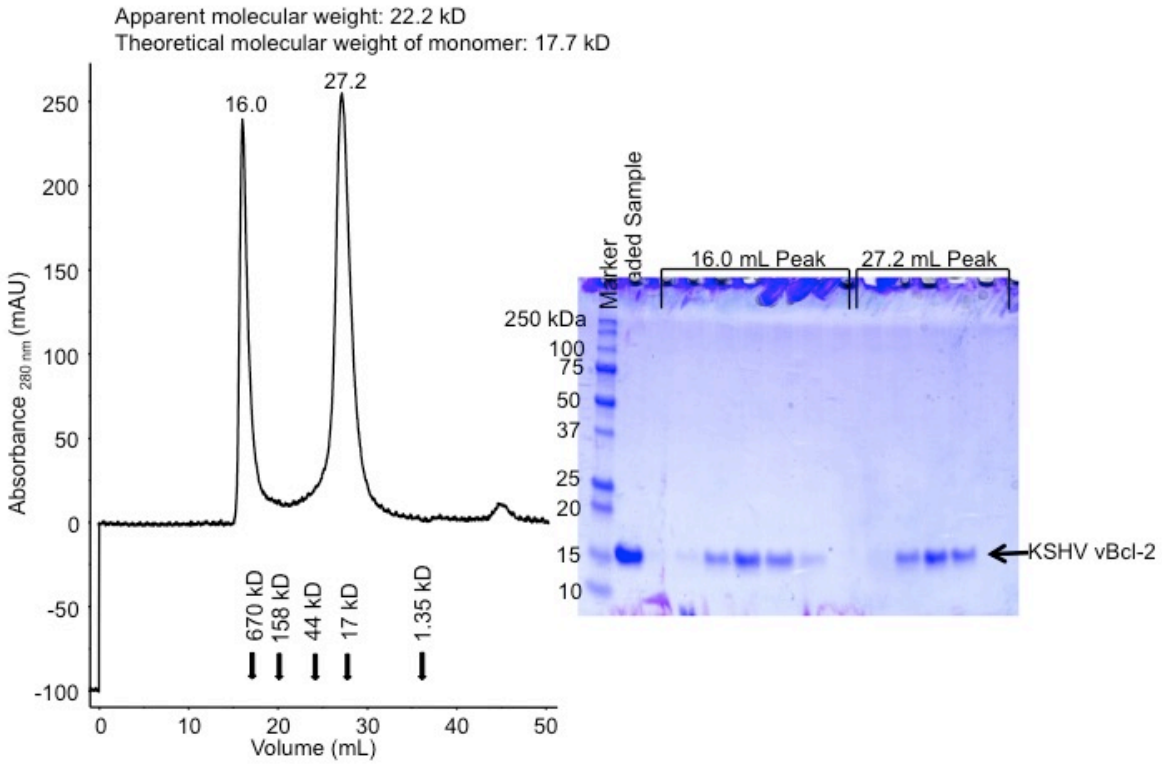

Figure 5.1. Size exclusion chromatogram and the corresponding SDS-PAGE of KSHV vBcl-2. The elution positions of SEC standards are indicated by arrows.

5.3.2. The KSHV vBcl-2 binds mNOXA but not the BECN1 BH3D G120E/D121A double mutant

It has been demonstrated that the KSHV vBcl-2 binds the BECN1 BH3D[44] and that the BECN1 BH3D G120E/D121A double mutant selectively binds the mouse γHV Bcl-2 homolog M11, but not cellular Bcl-2 proteins, to down-regulate M11-mediated inhibition of autophagy[140]. We therefore, used ITC to determine if the BH3D G120E/D121A double mutant likewise binds the KSHV vBcl-2. However, unlike M11, ITC reveals that the BECN1 BH3D G120E/D121A double mutant does not bind the KSHV vBcl-2 (Table 5.1, Figure 5.2A).

Since mNOXA has been shown to tightly bind KSHV vBcl-2, Mcl-1, and A1 but weakly or not at all to the other anti-apoptotic Bcl-2 paralogs[174, 175], we used ITC to quantify the thermodynamics of binding between the KSHV vBcl-2 and mNOXA. The KSHV vBcl-2 binds mNOXA with a K_d of ~64.5 μM (Table 5.1, Figure 5.2B). Although this confirms a direct interaction of moderate affinity, the affinity we determined by ITC is significantly less than the reported 0.163 μM affinity determined by fluorescence polarization assays[174].

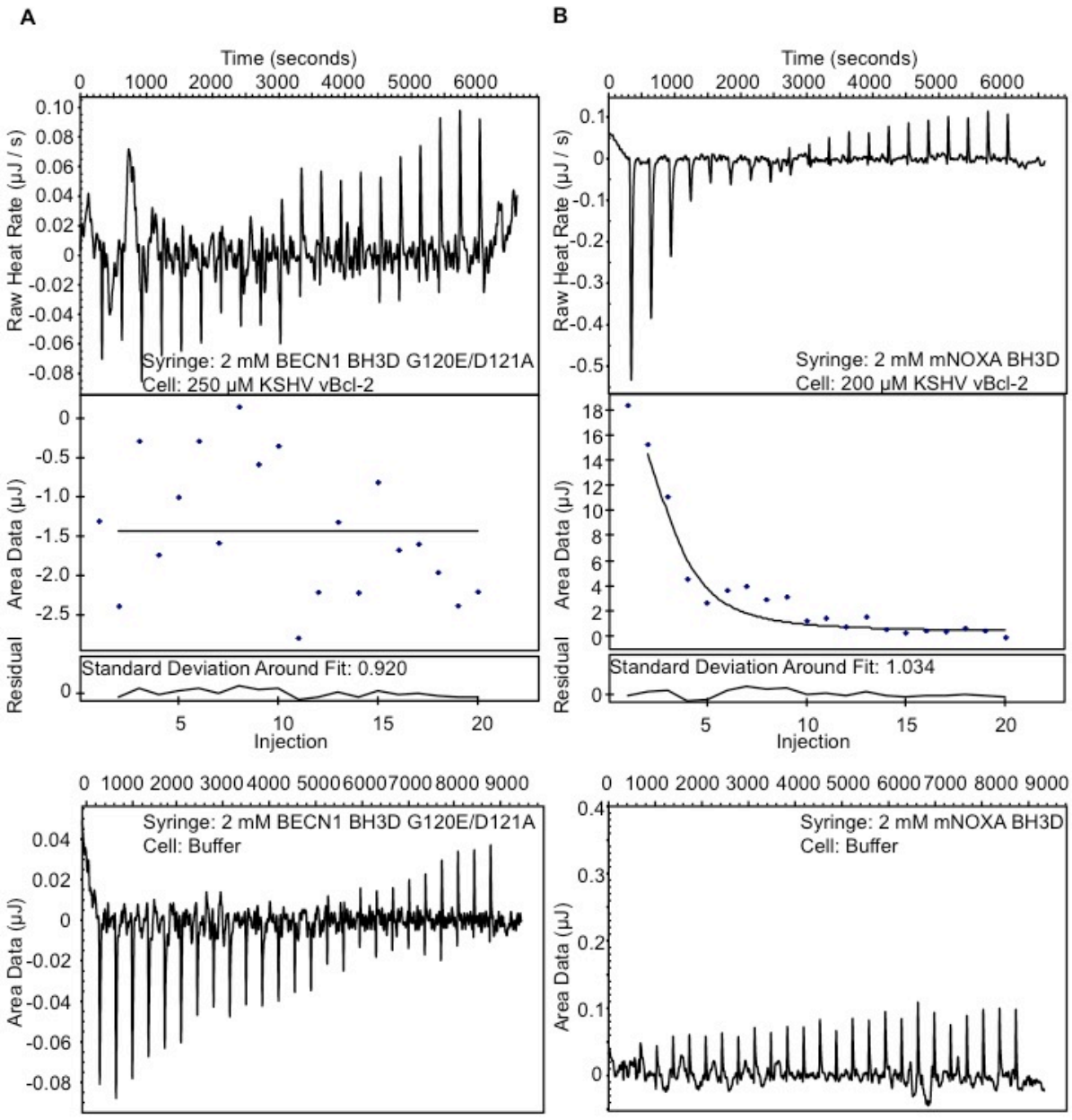

Figure 5.2. ITC profiles of KSHV vBcl-2 binding to BH3Ds.
KSHV vBcl-2 binding to the A) BECN1 BH3D G120E/D121A and B) mNOXA BH3D.

Table 5.1. Thermodynamics of binding of KSHV vBcl-2 to BH3D peptides.

BH3D	**K_d (µM)**	**ΔH (kJ/mol)**	**ΔS (J/K mol)**
BECN1 BH3D G120E/D121A		No binding detected	
mNOXA BH3D	64.5 ± 27.9	2.4 ± 0.5	89.0 ± 1.2

5.3.3. SEC and crystal screening of the KSHV vBcl-2:mNOXA complex

We used SEC to characterize the oligomeric state of the KSHV vBcl-2:mNOXA complex in solution, using a 10/30 Superdex 200 column linked in tandem with a 10/30 Superdex 75 SEC column to determine the elution volume of the complex. KSHV vBcl-2 itself elutes as two peaks, one at 16.0 mL corresponding to aggregated protein and another at 27.2 mL corresponding to the monomer (Figure 5.3). Four peaks were observed in the SEC chromatogram after KSHV vBcl-2 was incubated with mNOXA, two of which correspond to the peaks of the SEC chromatogram of the isolated KSHV vBcl-2 (Figure 5.3). The aggregation peak (16mL) is diminished, but still present; while the 27.1 mL peak represents the KSHV vBcl-2 monomer with an apparent molecular weight of 22.9 kD (theoretical MW = 17.7 kD). The two new peaks include a 33.2 mL peak with an apparent molecular weight of 3.0 kD that likely represents the mNOXA BH3D monomer (theoretical MW = 2.9 kD) and a 24.2 mL peak with an apparent molecular weight of 60.4 kD. The theoretical molecular weight of the KSHV vBcl-2:mNOXA heterodimer complex is 20.6 kD, so it is unclear what this peak corresponds to, but it may represent a homotrimer of KSHV vBcl-2:mNOXA BH3D heterodimers. Thus, the KSHV vBcl:mNOXA BH3D mixture appears to elute as a combination of complex and monomer species (Figure 5.3).

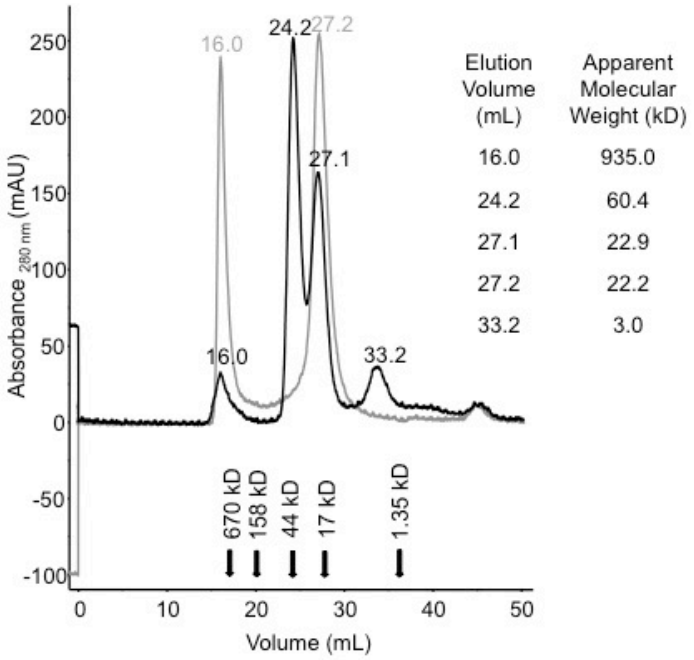

Figure 5.3. Size exclusion chromatogram of KSHV vBcl-2 (gray) and the KSHV vBcl-2:mNOXA mixture (black).
The elution positions of SEC standards are indicated by arrows.

We screened the 24.2 mL SEC peak for crystallization in an attempt to determine conditions suitable for crystallization of the KSHV vBcl-2:mNOXA complex. However, the 24.2 mL SEC peak could only be concentrated to a maximum of 1 mg/mL before precipitation was observed. Robotic screening of the 1 mg/mL KSHV vBcl-2:mNOXA did not yield an crystal hits after observation over the course of two months.

5.4. Discussion and conclusions

Protein precipitation during nickel IMAC and concentration stages of purification resulted in the acquisition of low concentrations of KSHV vBcl-2, wherein a maximum concentration of 2.5 mg/mL, was obtained, and KSHV vBcl-2:mNOXA complex, wherein a maximum concentration of 1 mg/mL, was obtained. We used these low concentration samples to perform a single ITC experiment to quantify the binding affinity for a BECN1 BH3D G120E/D121A double mutant and mNOXA BH3D peptide, as well as assess the oligomeric state of the KSHV vBcl-2:mNOXA complex by SEC and screen the complex for crystallization. Our ITC results indicate that, unlike M11[140], the KSHV vBcl-2 does not bind the BECN1 BH3D G120E/D121A double mutant. However, consistent with previous reports[174], our binding affinity measurements by ITC and elution volume determinations by SEC indicate that the KSHV vBcl-2 forms a complex with a mNOXA BH3D peptide. Our crystallization screening experiments did not yield any crystal hits; however, this may be due to the low concentration of complex (1 mg/mL).

Our experimental evidence showing that the BECN1 BH3D G120E/D121A double mutant does not bind the KSHV vBcl-2 indicates that this peptide is not a suitable candidate for selective inhibition of KSHV vBcl-2-mediated autophagy inhibition. The absence of binding suggests that, as expected, there are differences in the residues contributing to KSHV vBcl-2 binding to BH3Ds compared to other host cell and ɣHV Bcl-2 proteins. This is consistent with the observation that the binding affinity of KSHV vBcl-2 for other BH3Ds varies greatly compared to those of other anti-apoptotic Bcl-2 proteins[44, 174]. However, due to the precipitation issues during protein purification, we were unable to obtain sufficient quantities of KSHV vBcl-2 at concentrations high enough to permit biophysical and structural analyses to determine which residues contribute to binding specificity for the KSHV vBcl-2 by BH3Ds.

CHAPTER 6. BHRF1, AN EPSTEIN-BARR VIRUS BCL2 MIMIC, DOWN-REGULATES AUTOPHAGY BY NON-CANONICAL BINDING OF BECN1

6.1. Introduction

Epstein-Barr virus (EBV) is a widespread γHV, infecting more than 95% of the adult population worldwide[176]. EBV, first isolated over 50 years ago in Burkitt's lymphoma patients[177] and the cause of infectious mononucleosis, is associated with a variety of other malignancies including post-transplant lymphoproliferative disease, AIDS-related central nervous system lymphoma, AIDS-related leiomyosarcoma, AIDS-associated peripheral lymphomas, AIDS-associated primary effusion lymphomas, nasopharyngeal carcinoma, Hodgkin's lymphoma, and gastric carcinoma[178]. Like all γHVs, EBV is an obligate intracellular pathogen that establishes latent and lytic infections[178]. During latent infection, only a limited number of viral proteins are expressed, allowing the EBV genome to replicate without killing the host cell. Upon reactivation, EBV enters the lytic cycle wherein many genes are expressed to facilitate viral replication and release of infectious virus progeny, typically resulting in death of the host cell. Latent and lytic EBV genes have both been linked to tumorigenesis[179]. All γHVs encode mimics of the anti-apoptotic cellular BCL2 proteins[180, 181], suggesting that vBCL2s play an important role in the pathogenesis of these viruses.

The first BCL2 protein discovered, Bcl-2, was first identified as an oncogene that blocks apoptotic cell death[182, 183], but has since been established as a member of a family of proteins that also regulates autophagy, unfolded protein response, cell cycle progression, calcium signaling, glucose homeostasis, mitochondrial dynamics, and transcription[184, 185]. BCL2 family members are categorized based on their Bcl-2 homology (BH) domains and include anti-apoptotic and pro-apoptotic proteins[184]. Mammalian anti-apoptotic BCL2 proteins (Bcl-2, Bcl-X_L, Bcl-w, Mcl-1, and A1) contain four BH domains (BH1-4) and down-regulate apoptosis by binding the BH3D of pro-apoptotic BCL2 proteins, inhibiting the ability of pro-apoptotic Bcl-2s to induce apoptosis. All pro-apoptotic proteins (BAD, BAK, BAX, BID, BIM, PUMA) contain, at minimum, a single BH3 domain; although, BAX and BAK contain BH1-3[184, 186]. EBV encodes a vBCL2 called BamH1 H fragment rightward reading frame 1 (BHRF1)[171]. BHRF1 has been implicated in the development of various EBV-associated cancers. BHRF1 expression has been detected in clinical isolates of nasopharyngeal carcinomas[187, 188] and oral hairy leukoplakia[188], and promotes survival and

proliferation of lytically infected cells by inhibiting apoptotic cell death[189]. Although BHRF1 is controlled by the early lytic cycle BamHI W promoter, it is also constitutively expressed as a latent protein in some EBV-positive Burkitt's lymphoma cell lines and may contribute to Burkitt lymphomagenesis *in vivo*[190-192]. Therefore, BHRF1 is likely important for both the latent and lytic cycles of EBV. BHRF1 contains three BH domains[193, 194], which have together also been termed the early antigen protein R domain[195], comprising residues 1-160 and a C-terminal transmembrane (TM) helix comprising residues 161-191 that, like other BCL2 homologs, localizes BHRF1 to mitochondrial, endoplasmic reticulum, and nuclear membranes[193, 194, 196].

Human cellular anti-apoptotic BCL2 paralogs share about 50% sequence identity and have similar three-dimensional structures, consisting of a central hydrophobic α-helix surrounded by six or seven amphipathic helices[197]. The amphipathic, α-helical BH3Ds of pro-apoptotic proteins bind to a hydrophobic surface groove on the anti-apoptotic BCL2 paralogs such that the hydrophobic face of the BH3D helix is buried in the hydrophobic groove on the BCL2 surface[198-204]. BHRF1 shares about 24% sequence identity with human anti-apoptotic BCL2 proteins. Although the solution structure of the ligand-free BHRF1 lacks a hydrophobic surface groove[205], significant conformational changes occur upon binding BAK and BIM[206] to create a surface hydrophobic groove that binds the amphipathic α-helical BH3D of BAK or BIM. Different viral and human anti-apoptotic BCL2 homologs have a wide range of affinities for the BH3Ds of pro-apoptotic proteins[44, 174, 175], although the molecular binding determinants that dictate these variable specificities are poorly understood.

Several anti-apoptotic BCL2 homologs also inhibit autophagy by binding the BH3D of BECN1 to inhibit autophagosome formation[26, 43, 44, 46, 47, 172]. Since autophagy contributes to innate immune defenses by sequestering and degrading intracellular pathogens, several cellular and γHV BCL2 homologs, including the KSHV vBCL2 and the γHV 68 M11, inhibit autophagosome formation by binding the BH3D of BECN1 in order to evade autophagic degradation[26, 43, 44]. However, the role of BHRF1 in autophagy regulation is relatively poorly understood. Cellular autophagy assays have demonstrated that EBV initially inhibits autophagy after induction of lytic cycle, then later up-regulates autophagy and incorporates autophagic membranes into the final envelope of EBV[207, 208] but the mechanism by which EBV modulates

autophagy is unknown. The initial inhibition of autophagy could be mediated by an interaction between BHRF1 and BECN1.

We hypothesized that, like other BCL2 homologs, BHRF1 would bind BECN1 and down-regulate starvation-induced autophagy. Therefore, we used cellular assays to demonstrate that BHRF1 down-regulates BECN1-dependent, starvation-induced autophagy and that a cell-permeable version of the BID BH3D inhibits BHRF1-mediated down-regulation of autophagy in mammalian cells. We also used yeast two-hybrid assays, affinity pull-down assays, and ITC to identify the minimal region of BECN1 required for interaction with BHRF1 and quantify the thermodynamics of binding. Further, based on the differential affinities between BHRF1 and cellular BCL2s for BID, BIM, PUMA[174, 175], we expected that determining and comparing the thermodynamic contributions and atomic binding determinants to the BHRF1:BID BH3D interaction would illuminate differences in how BHRF1 versus cellular BCL2s bind BID. Therefore, we determined the 2.6 Å X-ray crystal structure of the BID BH3D peptide bound to BHRF1, as well as the thermodynamics of binding by ITC, to elucidate the mechanism by which the BID BH3D peptide binds to BHRF1. Together, these results show for the first time that BHRF1 down-regulates autophagy, that a flexible region of BECN1 comprising 15 residues of the IDR preceding the BH3D, the BH3D, and FHD is necessary and sufficient for binding to BHRF1, and that a BID BH3D can inhibit BHRF1-mediated down-regulation of autophagy. This information might facilitate the design of a peptide capable of displacing BECN1 in order to bind BHRF1 and selectively inhibit down-regulation of autophagy.

6.2. Materials and methods

6.2.1. Expression plasmid preparation

EBV BHRF1 residues 1-160, which lack the C-terminal transmembrane helix (BHRF1ΔTM) (Figure 6.1A), were cloned between the *NcoI* and *NotI* restriction enzyme sites of the pET-21d(+) bacterial expression vector such that the expression construct included a C-terminal His_6 tag. Human BECN1 residues 1-450 (FL), residues 1-171, or residues 90-171 were cloned between the *NcoI* and *NotI* restriction enzyme sites of the pMBP-Parallel-1 bacterial expression vector such that the expression construct encoded an N-terminal MBP tag. Site directed mutagenesis was used to generate the FL $BECN1^{AFM}$, which contained a triple mutation (F359D, F360D, and W361D) that improves protein solubility[168]. The bacterial expression construct for Bcl-X_L residues 1-208 N52D/N66D, which lacks the C-

terminal transmembrane helix and contains a double mutant that improves proteins solubility (Bcl-$X_L\Delta$TM), was generated as previously described[140].

For the yeast two-hybrid assays, human ɣHV BHRF1 residues 1-191 (FL) or residues 1-160 (BHRF1ΔTM) were cloned between the *SalI* and *XhoI* restriction enzyme sites of the pGAD-GH yeast two-hybrid vector in frame with the Gal4 activation domain (AD). UVRAG, another key autophagy protein that is an established binding partner of BECN1, was used as a positive control for the yeast two-hybrid assays. FL human UVRAG was cloned between the *BamHI* and *XhoI* restriction enzyme sites of the pGAD-GH yeast two-hybrid vector in frame with the Gal4 AD. Different human BECN1 fragments were cloned between the *BamHI* and *PstI* restriction enzyme sites of the pGBT8 yeast two-hybrid vector in frame with the Gal4 DNA binding domain (DBD). These BECN1 fragments were: BECN1 residues 1-450 (FL), residues 1-171 (IDR-FHD), residues 90-171, residues 105-265 (BH3D-FHD-CCD), residues 1-140 (IDR), or residues 1-265 (IDR-FHD-CCD) (Table 6.1, Figure 6.1B).

Table 6.1. BECN1 constructs used for yeast two-hybrid assays with BHRF1.

Label	BECN1 Residues
FL BECN1	1-450
BECN1 IDR-FHD	1-171
BECN1 residues 90-171	90-171
BECN1 BH3D-FHD-CCD	105-265
BECN1 IDR	1-140
BECN1 IDR-FHD-CCD	1-265

Figure 6.1. The domain architecture of BHRF1 and BECN1.
A) BHRF1 includes an early antigen protein R domain and transmembrane (TM) helix, and B) BECN1 includes an N-terminal intrinsically disordered region (IDR), Bcl-2 homology 3 domain (BH3D), flexible helical domain (FHD), coiled-coil domain (CCD), and β-α repeated, autophagy-specific domain (BARAD). The BARAD contains an aromatic finger comprised of three aromatic residues (FFW).

Lastly, for the cellular autophagy assays, FL human BECN1 or FL human γHV BHRF1 were cloned between the *NcoI* and *XhoI* restriction enzyme sites of the pcDNA3.1 mammalian expression vector, such that the expressed protein would include an N-terminal FLAG-tag.

6.2.2. Autophagy assays

Cellular autophagy levels were quantified by monitoring cellular localization of GFP-tagged LC3 protein[209]. Each chamber of an 8-well culture slide (Nunc™ Lab-Tek™ Chambered Coverglass) was seeded with 3.0 x 10^5 MCF7 cells and cultured overnight in Dulbecco modified eagle medium (DMEM) with 10% fetal bovine serum (FBS) until 80-90% confluent. The cells were co-transfected with 500 ng total plasmid per well. For transfection with BECN1, cells were transfected with 200 ng GFP-LC3 and 300 ng FLAG-BECN1, using Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. For co-transfection with BHRF1, cells were co-transfected with 200 ng GFP-LC3, 100 ng FLAG-BECN1, and 200 ng FLAG-BHRF1. After transfection and incubation at 37 °C for 24 hours, the cells were cultured in either rich (DMEM, 10% FBS, 2x essential amino acids and 2x nonessential amino acids) or starvation (Earle's balanced salt solution) medium for 4 hours. Cells were counterstained with DAPI to visualize nuclei, 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 fluorescent microscope (Zeiss AxioObserver Z1) and 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. The effect of a potential inhibitory TAT-BID BH3D peptide was investigated by comparing autophagy levels in the absence or presence of 30 μM peptide. A 2 mM peptide stock was prepared in water and stored at -80 °C until use. The peptide was diluted to 30 μM directly into the rich or starvation medium.

Expression levels of FLAG-BECN1 and FLAG-BHRF1 in MCF7 cells were verified by Western blotting using commercial mouse monoclonal anti-FLAG M2-peroxidase antibody (Sigma). As a loading control, actin levels in MCF7 cell lysates were detected using commercial mouse monoclonal anti-Actin-peroxidase antibody (Sigma).

6.2.3. Yeast two-hybrid assays

Several plasmid combinations were co-transformed into the *Saccharomyces cerevisiae* HF7c using a yeast transformation kit (Sigma). A single yeast colony obtained in the co-transformation procedure was resuspended in 1 mL of sterile deionized H_2O and the optical density of each cell suspension was determined at a wavelength A_{600}. A cell suspension of each co-transformant was prepared at an A_{600} value of 0.01 such that the final volume, diluted in sterile deionized H_2O, was 500 µL. This initial cell suspension (OD_{600} = 0.01) was serially diluted to 1:5 and 1:25. Five µL of the 1:1, 1:5, and 1:25 dilutions were spotted on SD/-Leu/-Trp dropout medium to assess transformation efficiency and SD/-Leu-Trp-His selection medium to test for potential interactions. 1-9 mM 3-amino-1,2,4-trizole (3-AT) was used in the SD/-Leu-Trp-His plates to reduce false positives. The culture plates were incubated at 30 °C for 2–3 days. Co-transformation of BECN1 and UVRAG served as a positive interaction control [53]. Each experiment was performed in triplicate. I would like to thank Srinivas Dasanna for performing these yeast two-hybrid assays.

6.2.4. Protein expression and purification

All proteins were expressed in *E. coli* cells and purified by chromatography methods. At each state of purification, protein purity 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.

For BHRF1ΔTM, BL21(DE3)pLysS cells were transformed and grown in LB medium with 100 µg/ml ampicillin at 37 °C to an A_{600} of ~0.8 before equilibrating the temperature to 20 °C. Once the cells achieved an A_{600} of ~1.4, they were induced by the addition of 0.5 mM IPTG and expressed overnight. Soluble protein in the cell lysate (Lysis buffer: 25 mM Tris HCl pH 7.5, 100 mM NaCl, 40 mM imidazole, 2 mM BME, and 1 cOmplete, EDTA-free protease inhibitor cocktail tablet (MilliporeSigma) per liter of cell culture) from a 3 L expression culture was purified by nickel IMAC using 10 mL of Ni Sepharose High Performance nickel-charged IMAC resin (GE Life Sciences, Pittsburgh, US), washed with 10 column volumes of Wash Buffer (25 mM Tris HCl pH 7.5, 100 mM NaCl, 40 mM imidazole, and 2 mM BME), and eluted with Wash Buffer containing 250 mM imidazole. The protein was subsequently purified by SEC using a 16/60 Superdex 200 column (GE Life Sciences, Pittsburgh, US) in SEC buffer consisting of 25

mM HEPES sodium salt pH 7.5, 100 mM NaCl, and 2 mM BME. The BHRF1ΔTM-His_6 was quantified via UV/Vis using an ε_{280} of 33,460 M^{-1} cm^{-1} and a molecular mass of 19,119 Da.

For $BECN1^{AFM}$, Arctic Express(DE3) cells were transformed and grown in LB medium with 100 µg/ml ampicillin at 37 °C to an A_{600} of ~0.8 before equilibrating the temperature to 13 °C. Once the cells achieved an A_{600} of ~1.4, they were induced by the addition of 0.5 mM IPTG and expressed overnight. Soluble protein in the cell lysate (Lysis buffer: 50 mM Tris HCl pH 8.0, 300 mM NaCl, 2 mM DTT, and 1 cOmplete, EDTA-free protease inhibitor cocktail tablet (MilliporeSigma) per liter of cell culture) from a 3 L expression culture was purified by amylose affinity chromatography on a 10 mL Amylose Resin (NEB BioLabs, Ipswich, US) gravity column and washed with 10 column volumes of amylose Wash Buffer (50 mM Tris HCl pH 8.0, 300 mM NaCl, 2 mM DTT, and 1mM EDTA). Fusion protein was subjected to overnight, on-column cleavage using 8 mg GST-TEV at 4 °C to remove the MBP tag. Cleaved protein was washed from the column with Wash Buffer and further purified to homogeneity by SEC using a 10/30 Superose 6 Increase column (GE Life Sciences, Pittsburgh, US) in SEC buffer consisting of 50 mM HEPES sodium salt pH 8.0, 150 mM NaCl, and 2 mM BME. The $BECN1^{AFM}$ was quantified via UV/Vis using an ε_{280} of 49,390 M^{-1} cm^{-1} and a molecular mass of 52,176 Da.

For BECN1(90-171), BL21(DE3)pLysS cells were transformed and grown in LB medium with 100 µg/ml ampicillin at 37 °C to an A_{600} of ~0.8 before equilibrating the temperature to 15 °C. Once the cells achieved an A_{600} of ~1.4, they were induced by the addition of 0.5 mM IPTG and expressed overnight. Soluble protein in the cell lysate (Lysis buffer: 50 mM Tris HCl pH 7.5, 150 mM NaCl, 2 mM BME, and 1 cOmplete, EDTA-free protease inhibitor cocktail tablet (MilliporeSigma) per liter of cell culture) from a 6 L expression culture was purified by amylose affinity chromatography on a 10 mL Amylose Resin (NEB BioLabs, Ipswich, US) gravity column and washed with 10 column volumes of amylose Wash Buffer (50 mM Tris HCl pH 7.5, 150 mM NaCl, 2 mM BME, and 1mM EDTA). Fusion protein was subjected to overnight, on-column cleavage using 8 mg of GST-TEV protease at 4 °C to remove the MBP tag. Cleaved protein was washed from the column with wash buffer and further purified by ion exchange chromatography on a 5/50 GL MonoQ column (GE Life Sciences, Pittsburgh, US) (MonoQ Buffer A: 50 mM Tris HCl pH 7.5, 2 mM BME; MonoQ Buffer B: 50 mM Tris HCl pH 7.5, 2 mM BME, 1 M NaCl). The protein was loaded and washed at 50 mM NaCl, and then eluted with a gradient elution where it eluted at

approximately 100 mM NaCl. The protein was then purified by SEC using a 10/30 Superdex 200 column (GE Life Sciences, Pittsburgh, US) in SEC buffer consisting of 25 mM HEPES sodium salt pH 7.5, 150 mM NaCl, 2 mM BME. The BECN1(90-171) was quantified via UV/Vis using an ε_{280} of 1,490 M^{-1} cm^{-1}, as well as using Bradford reagent (BioRad, Hercules, CA) according to the manufacturer's instructions, and a molecular mass of 9,432 Da.

Bcl-X_LΔTM-His_6 was expressed and purified as described previously[140]. This involved transformation of BL21(DE3)pLysS and growth in LB medium with 35 µg/ml kanamycin at 37 °C to an A_{600} of ~0.6 before equilibrating the temperature to 20 °C. Once the cells achieved an A_{600} of ~0.8, they were induced by the addition of 0.5 mM IPTG and expressed overnight. Soluble protein in the cell lysate (Lysis buffer: 25 mM Tris HCl pH 7.5, 100 mM NaCl, 40 mM imidazole, 2 mM BME, and 1 cOmplete, EDTA-free protease inhibitor cocktail tablet (MilliporeSigma) per liter of cell culture) from a 3 L expression culture was purified by nickel IMAC on a 10 mL HisTrap nickel IMAC column (GE Life Sciences, Pittsburgh, US), washed with 10 column volumes of Wash Buffer (25 mM Tris HCl pH 7.5, 100 mM NaCl, 40 mM imidazole, and 2 mM BME), and eluted with Wash Buffer containing 250 mM imidazole. The protein was then purified by anion exchange chromatography using an 8 mL 10/10 HR MonoQ column (GE Life Sciences, Pittsburgh, US) (MonoQ Buffer A: 25 mM Tris HCl pH 7.5, 2 mM BME; MonoQ Buffer B: 50 mM Tris HCl pH 7.5, 2 mM BME, 1 M NaCl) using a step elution with 300 mM NaCl. Finally, the protein was purified by SEC using a 16/60 Superdex 200 column (GE Life Sciences, Pittsburgh, US) in SEC buffer consisting of 25 mM HEPES sodium salt pH 7.5, 100 mM NaCl, and 2 mM BME. Bcl-X_LΔTM-His_6 was quantified via UV/Vis using an ε_{280} of 41,940 M^{-1} cm^{-1} and molecular mass of 24,035 Da.

6.2.5. Pull-down assays

The clarified lysate from a 1 L expression culture of MBP-BECN1(90-171) and MBP-BECN1(1-171) was separately bound to 0.5 mL of Amylose Resin (NEB BioLabs, Ipswich, US) and washed with 10 column volumes of wash buffer consisting of 50 mM Tris HCl pH 7.5, 150 mM NaCl, 2 mM BME, and 1mM EDTA. 6 mg of purified BHRF1ΔTM-His_6 or Bcl-X_LΔTM-His_6 was separately incubated with the MBP-BECN1(90-171)-bound or MBP-BECN1(1-171)-bound amylose resin for 1 hour at 4 °C and washed with wash buffer. Bound protein was eluted with wash buffer plus 20 mM maltose. The amylose elution fractions were then loaded onto 100 µL of nickel immobilized metal ion chromatography resin and washed

with 10 column volumes of wash buffer. Bound protein was eluted with wash buffer plus 250 mM imidazole. Fractions were evaluated by SDS-PAGE stained with Coomassie Blue.

6.2.6. Peptide synthesis

Peptides corresponding to BECN1 BH3D residues 105-130, BID BH3D residues 79-104, and BID BH3D-derived mutant peptides (Table 6.2) were chemically synthesized, HPLC-purified to >90% purity, and purity confirmed by electrospray mass spectrometry (EZBiolab, Carmel, IN). The BID BH3D mutants were designed by comparing the sequences of the BID, BAK, and BIM BH3D, which bind BHRF1, and the BECN1 BH3D, which does not bind BHRF1, in an effort to discover a peptide that would selectively bind BHRF1 but not Bcl-X_L. To make the BID BH3D peptide cell-permeable, the transactivating HIV-1 transcriptional activator protein transduction domain (TAT), which is a cell-penetrating peptide previously shown to facilitate entry of extended poly-peptides into mammalian cells under the conditions used here[166], was attached via a diglycine linker to the N terminus of the BID BH3D peptide (TAT-BID BH3D).

Table 6.2. Peptide sequences synthesized for BHRF1 interaction analyses.

Peptide	Amino Acid Sequence
BECN1 BH3D	AGGTMENLSRRLKVTGDLFDIMSGQT
BID BH3D	QEDIIRNIARHLAQVGDSMDRSIPPG
BID BH3D G94A	QEDIIRNIARHLAQVADSMDRSIPPG
BID BH3D Mutant 1	DTMRNIARQLARIGDEFDRSYGRT
BID BH3D Mutant 2	DTMRNIARHLARIGDEFDRSIGRT
BID BH3D Mutant 3	DTMRNIARHLAQVGDSMDRSIGRT
TAT-BID BH3D	YGRKKRRQRRR-GG-QEDIIRNIARHLAQVGDSMDRSIPPG

6.2.7. ITC

Protein and peptide samples were dialyzed against 50 mM HEPES sodium salt pH 8.0, 150 mM NaCl, and 2 mM BME. In order to assess the thermodynamics of binding of BHRF1 to BECN1 fragments, ITC experiments were performed in triplicate at 20 °C. 1mM BHRF1ΔTM-His_6 was separately titrated into 25 μM FL BECN1 and 50 μM BECN1 (90-171). 1 mM BECN1(105-130) was titrated into 100 mM BHRF1ΔTM. In order to assess the thermodynamics of binding of the BID BH3D and BID BH3D mutants to BHRF1 and Bcl-X_L, ITC experiments were performed in duplicate at 10 °C. 2 mM peptides were separately titrated into 200 mM BHRF1ΔTM-His_6 and Bcl-X_L(1-208, N52D/N66D)-His_6. All ITC

experiments were performed using a Low Volume Gold Nano isothermal titration calorimeter (TA Instruments, New Castle, US) and data were analyzed with the NanoAnalyze Software (TA Instruments, New Castle, US) using an independent binding model.

6.2.8. Crystallization

The purified BHRF1ΔTM-His_6 and BID BH3D peptide were mixed with a 1:1 molar ratio and incubated at 4 °C for two hours. The BHRF1:BID BH3D complex was crystalized by hanging drop vapor diffusion from a 3:1 mixture of protein stock (5 mg/mL complex in 25 mM HEPES sodium salt pH 7.5, 100 mM NaCl, 2 mM BME) and well solution (0.1 mM $NaC_2H_3O_2$ pH 5.0, 1.6 mM HCOONa). Rod-shaped crystals were harvested and cryoprotected in a cryosolution consisting of 0.1 M $NaC_2H_3O_2$ pH 5.0 and 25% glycerol and then immediately flash-frozen in liquid N_2.

6.2.9. X-ray Data collection, structure solution, and refinement

Diffraction data were collected at 100 K at a wavelength of 0.97943 Å in one 360° sweep from a single crystal on NE-CAT beamline 24-ID-C at the Advanced Photon Source, Argonne National Laboratory, Argonne, IL, USA. Diffraction intensities were recorded on a 4 x 4 tiled MAR Mosaic CCD detector (Rayonix) at a crystal-to-detector distance of 250 mm over a 0.5° crystal rotation with 1 s exposure per image. Diffraction data were processed using the RAPD automated processing web server (https://rapd.nec.aps.anl.gov/rapd/). The structure was solved by molecular replacement using *Phaser*[210] in the *PHENIX* suite[96], using as a model the BHRF1:BIM BH3D complex structure (PDB entry 2WH6) modified to remove coordinates pertaining to the BIM BH3D. The structure was refined to 2.61 Å resolution using the *PHENIX* refinement pipeline[96],combined with iterations of manual model building using *Coot*[211].

6.2.10. Protein structure analysis

Since the asymmetric unit contained 5 BHRF1:BID complexes, simultaneous superimposition of multiple structures was performed with THESEUS[212], which applies a maximum-likelihood method, via the PSSweb server (http://pssweb.org)[213]. The quality of the geometry and fit to the electron density of each chain was analyzed via the wwPDB Validation Service (https://validate-rcsb-1.wwpdb.org). The complex composed of chains A and B was identified via this validation as having the highest quality. Therefore, pairwise structural superimposition of each of the other four complexes was performed against

BHRF1:BID chains A and B using *PyMOL* (v.1.3; Schrödinger). All subsequent analyses were performed using BHRF1:BID chains A and B. The BHRF1:BID interface was analyzed in detail using *PISA*[145]. Pairwise structural superimpositions of the BHRF1:BID complex with other complexes were performed using *PyMOL* (v.1.3; Schrödinger). Molecular figures were created using *PyMOL* (v.1.3; Schrödinger). Multiple sequence alignments of the BH3Ds were performed using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/).

6.3. Results

6.3.1. BHRF1 down-regulates starvation-induced autophagy

We investigated the impact of exogenously expressed BHRF1 on cellular autophagy levels in human breast adenocarcinoma MCF7 cells (Figure 6.2). MCF7 cells lack detectable endogenous expression of BECN1, resulting in very low basal levels of autophagy[26, 44, 214] even in starvation conditions unless BECN1 is ectopically expressed (Figure 6.2B,C). Expression of BECN1 and BHRF1 was comparable in starvation and nutrient-rich conditions (Figure 6.2A). 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 medium (Figure 6.2B,C).

Transient expression of BECN1 in MCF7 cells did not increase autophagy levels in nutrient-rich conditions ($p = 0.7800$ for BECN1 expression versus no expression; Figure 6.2B); but led to a marked increase in autophagy upon starvation ($p = 0.0002$ for starved versus nutrient-rich cells; Figure 6.2B). This is consistent with the established importance of BECN1 in starvation-triggered autophagy, rather than basal autophagy[215, 216]. Strikingly, starvation-induced, BECN1-dependent autophagy is significantly down-regulated by co-expression of BHRF1 (Figure 6.2B,C). Furthermore, co-expression of BHRF1 returns autophagy levels to those consistent with basal autophagy levels ($p = 0.2500$ for starved BHRF1 co-expressed with BECN1 versus nutrient-rich BECN1). Therefore, we find that BHRF1 abrogates BECN1-mediated, starvation-induced autophagy.

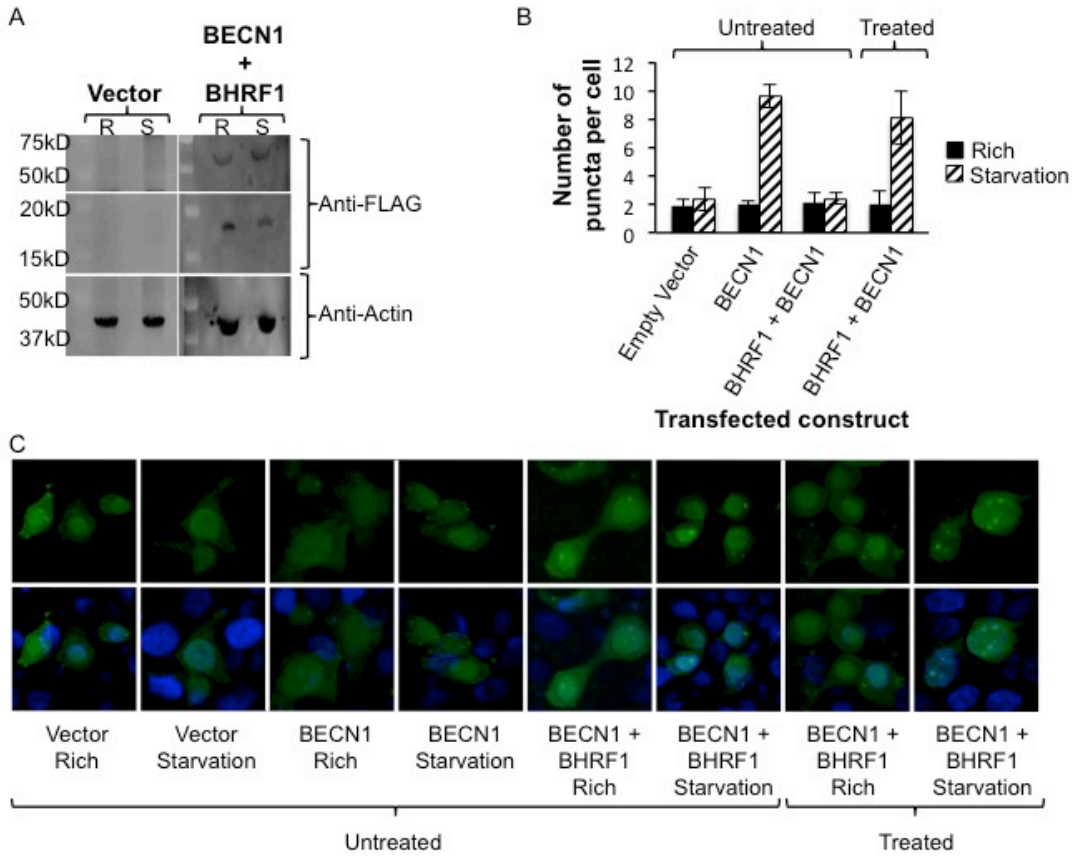

Figure 6.2. The effect of BHRF1 and TAT-BID peptide treatment on autophagy.
A) Western blot of MCF7 cell extracts showing comparable expression levels of FLAG-BECN1 and FLAG-BHRF1 in nutrient rich (R) and starvation (S) conditions, with Actin as a loading control. B) Light microscopy quantification of discrete GFP-LC3 puncta per cell in GFP-positive MCF7 cells co-transfected with GFP-LC3 and BECN1 or BHRF1, as indicated below the x-axis, and then treated with either no peptide or TAT-BID peptide, as indicated above the graph. Bars represent the number of puncta per cell. Error bars represent standard deviation. C) Representative images of GFP-LC3 (green) and DAPI (blue) staining in cells grown in starvation or nutrient rich media and transfected with mutant FLAG-BECN1 or FLAG-BHRF and treated with TAT-BID peptide, as indicated below the images.

6.3.2. BECN1 residues 90-171 are required for binding BHRF1

Since several cellular and γHV BCL2 homologs, including the KSHV vBCL2 and the γHV 68 M11, inhibit autophagosome formation by binding the BH3D of BECN1[26, 43, 44], we used yeast two-hybrid assays to determine whether BECN1 and BHRF1 interact and to identify the minimal BECN1 region required for binding to BHRF1. *Saccharomyces cerevisiae* HF7c cells were co-transformed with BHRF1 fused to the AD and various combinations of BECN1 domains fused to the DBD.

Growth is not observed in the 1:5 and 1:25 dilutions under 1 mM 3-AT with empty GAD-GH or GBT8 but is observed with FL BHRF1 and FL BECN1 (Figure 6.3), indicating that the reporter vectors do not self-associate but that FL BHRF1 and FL BECN1 interact. Failure to grow in the presence of higher

concentrations of 3-AT is likely due to association of the expressed proteins with membranes due to the presence of the BHRF1 TM helix and the BECN1 aromatic finger, since these assays are not optimized for use with membrane-associated proteins. Therefore, subsequent yeast two-hybrid assays tested binding of the BHRF1ΔTM fused to the AD, to various combinations of soluble BECN1 domains fused to the DBD.

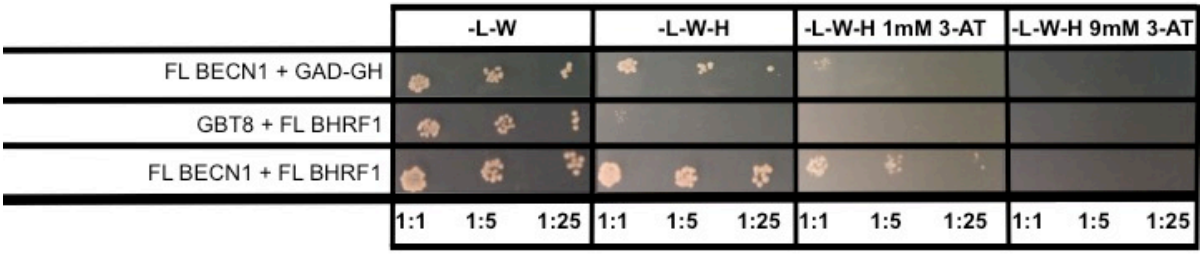

Figure 6.3. BHRF1 interacts with BECN1.
Yeast two-hybrid analysis in *Saccharomyces cerevisiae* HF7c. Labels on the left indicate BECN1 fused to the DBD (GBT8) that was co-transformed with BHRF1 fused to the AD (GAD-GH), and spotted onto SD/-Leu/-Trp dropout medium and SD/-Leu/-Trp/-His selection medium, as indicated above the image. Each co-transformation was serially diluted as indicated below the image.

Growth is not observed under 3 or 9 mM 3-AT with empty GAD-GH or GBT8 but is observed with BECN1 and UVRAG (Figure 6.4). This indicates that reporter vectors do not self-associate, and that interaction by known binding partners, BECN1 and UVRAG[53], is detected under the conditions used here. Growth under 3 or 9 mM 3-AT of co-transformants of BHRF1ΔTM and the BECN1 IDR-FHD-CCD fragment, which contains all non-membrane interacting BECN1 domains, as well as with the BECN1 IDR-FHD fragment, indicates that the BECN1 BARAD and CCD are not required for interaction with BHRF1. However, co-transformation with BHRF1ΔTM and the BECN1 IDR, which includes the BH3D, did not enable growth indicating that unlike other interactions between BECN1[44, 140] and pro-apoptotic BCL2 proteins[199-201, 206], wherein the BH3D is sufficient for interaction, the BECN1 FHD is also required for interaction with BHRF1. However, co-transformants of the BHRF1ΔTM and the BECN1 BH3D-FHD-CCD also did not grow, indicating that the BH3D and FHD together are also insufficient for interaction with BHRF1 and additional regions of the IDR may be required. Indeed, growth is enabled by co-transformation of BHRF1ΔTM and a BECN1 fragment that includes 15 residues from the BECN1 IDR immediately preceding the BH3D, i.e. BECN1 residues 90-171 (Figure 6.4). Thus, we find that BECN1 residues 90-171 are required and sufficient for interaction with BHRF1.

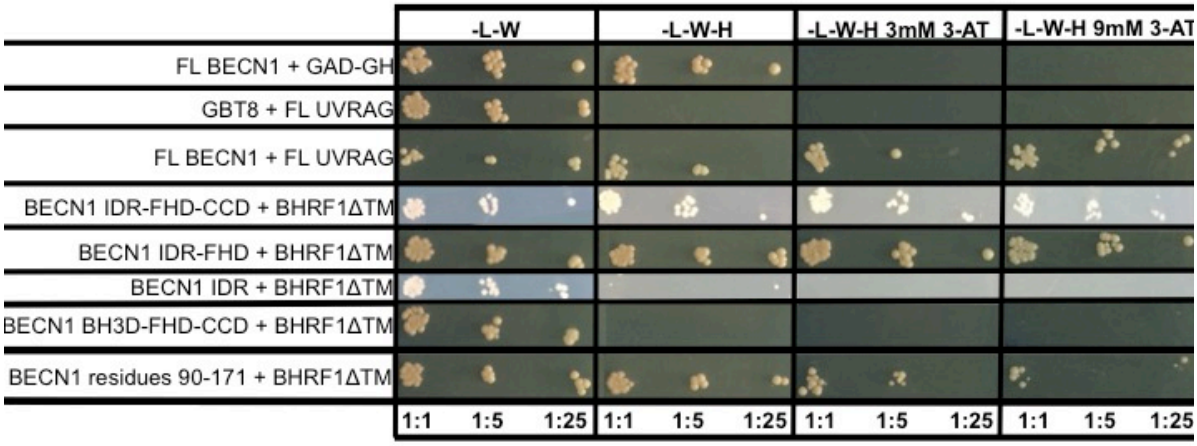

Figure 6.4. Determination of the minimal BECN1 region required for BHRF1 binding.
Yeast two-hybrid analysis in *Saccharomyces cerevisiae* HF7c. Labels on the left indicate the various BECN1 domains fused to the DBD (GBT8) that were co-transformed with BHRF1 fused to the AD (GAD-GH), and spotted onto SD/-Leu/-Trp dropout medium and SD/-Leu/-Trp/-His selection medium, as indicated above the image. Each co-transformation was serially diluted as indicated below the image.

6.3.3. Purification of BHRF1ΔTM

BHRF1ΔTM-His_6 was purified by sequential nickel IMAC and SEC. The single peak on SEC and single band on SDS-PAGE indicate the BHRF1ΔTM-His_6 was pure and homogenous (Figure 6.5). The apparent molecular weight of the BHRF1ΔTM-His_6 calculated based on the elution volume from SEC is 12.9 kD, which is 1.4-fold smaller than the theoretical molecular weight of 19.1 kD for the BHRF1ΔTM-His_6 monomer calculated from the amino acid sequence. This is inconsistent with any oligomerization, indicating the BHRF1ΔTM-His_6 forms a monomer. The final yield of the purified BHRF1ΔTM-His_6 obtained from one liter of bacterial culture was 20 mg.

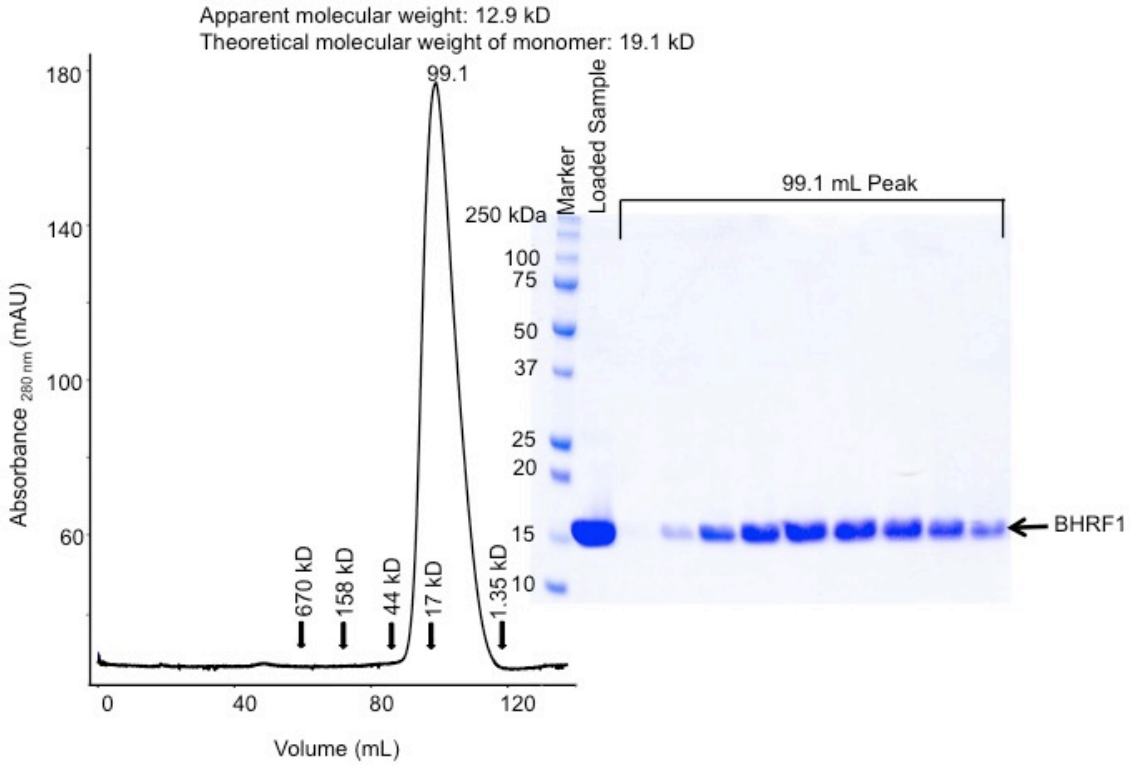

Figure 6.5. Size exclusion chromatogram and the corresponding SDS-PAGE of BHRF1ΔTM. The elution positions of SEC standards are indicated by arrows.

6.3.4. Purification of BECN1^{AFM}

MBP-BECN1^{AFM} fusion protein was purified by amylose affinity chromatography followed by on-column cleavage using GST-TEV protease to remove the MBP tag. BECN1^{AFM} then purified by sequential anion exchange chromatography and SEC. The SEC elution profile contains three peaks: an 8.6 mL peak in the void volume that likely contains aggregated protein, a 14.3 mL peak that contains BECN1^{AFM}, and a 17.8 mL peak that contains lower molecular weight contaminants (Figure 6.6). The apparent molecular weight of BECN1^{AFM} calculated based on the 14.3 mL elution volume from SEC is 535.4 kD, which is 10-fold larger than the theoretical molecular weight of 52.4 kD for the BECN1^{AFM} monomer calculated from the amino acid sequence. This is consistent with the BECN1^{AFM} forming a rod-shaped dimer via the CCD as well as containing a large IDR. The final yield of the purified BECN1^{AFM} obtained from one liter of bacterial culture was 1.8 mg.

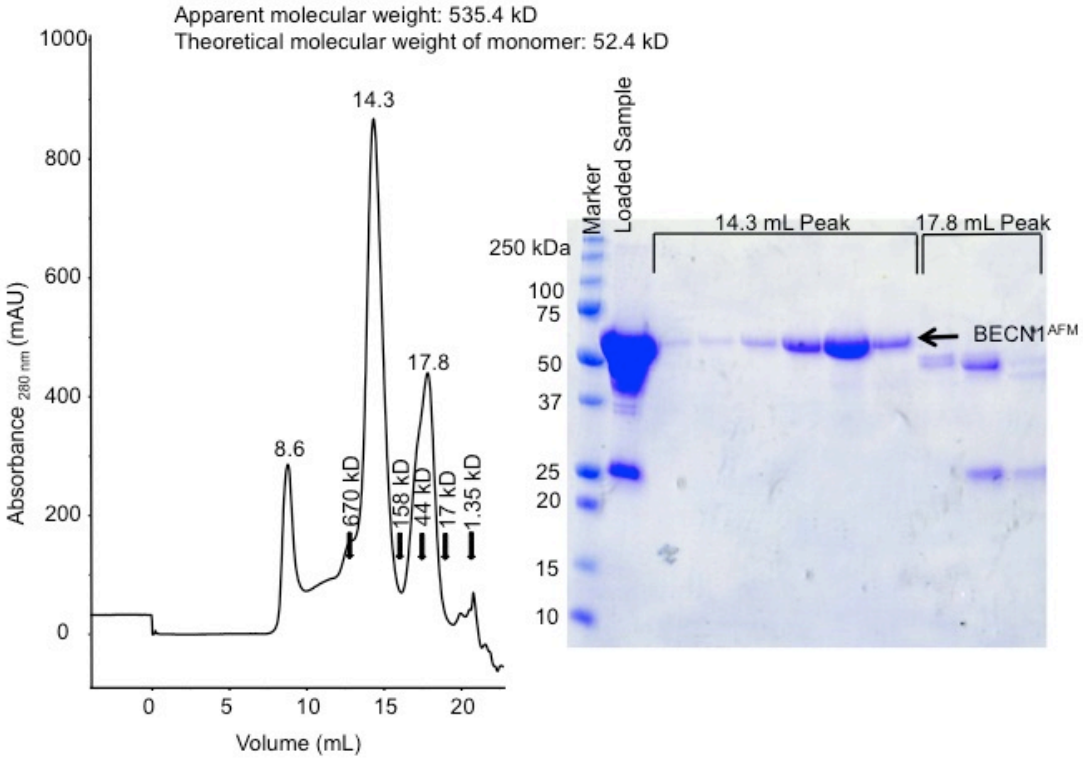

Figure 6.6. Size exclusion chromatogram and the corresponding SDS-PAGE of BECN1^{AFM}. The elution positions of SEC standards are indicated by arrows.

6.3.5. Purification of BECN1 residues 90-171

MBP-BECN1(90-171) fusion protein was purified by affinity chromatography followed by on-column cleavage using GST-TEV protease to remove the MBP tag. BECN1(90-171) was purified by sequential anion exchange chromatography and SEC. The single peak on SEC and single band on SDS-PAGE indicate that BECN1(90-171) was pure and homogenous (Figure 6.7). The apparent molecular weight of BECN1(90-171) calculated based on the elution volume from SEC is 24.2 kD, which is 2.6-fold larger the theoretical molecular weight of 9.4 kD for the BECN1 residues 90-171 monomer calculated from the amino acid sequence. This is inconsistent with the formation of a dimer or trimer and could indicate that the protein is disordered or partially disordered. The final yield of the purified BECN1 residues 90-171 obtained from one liter of bacterial culture was 0.3 mg.

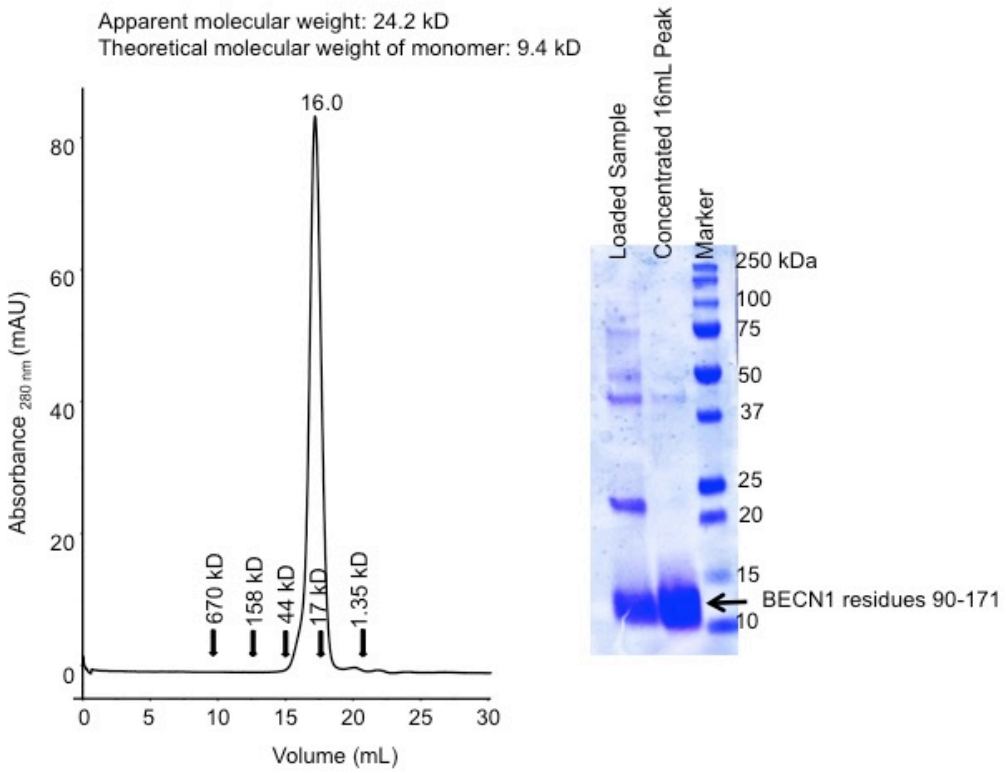

Figure 6.7. Size exclusion chromatogram and the corresponding SDS-PAGE of BECN1 residues 90-171. The elution positions of SEC standards are indicated by arrows.

6.3.6. Purification of Bcl-X_LΔTM

In order to compare binding of BHRF1 and a mammalian BCL2 protein by BH3Ds, Bcl-X_L was purified. Bcl-X_LΔTM-His_6 was purified by sequential nickel IMAC, anion exchange chromatography, and SEC. The SEC elution profile contains two peaks: a 46.6 mL peak in the void volume that contains high molecular weight contaminants and a 90.1 mL peak that contains Bcl-X_LΔTM-His_6 (Figure 6.8). The apparent molecular weight of the Bcl-X_LΔTM-His_6 calculated based on the 90.1 mL SEC elution volume is 16.9 kD, which is 1.4-fold smaller than the theoretical molecular weight of 24.4 kD for the Bcl-X_LΔTM-His_6 monomer calculated from the amino acid sequence. This is inconsistent with any oligomerization, indicating the Bcl-X_LΔTM-His_6 forms a monomer. The final yield of the purified Bcl-X_LΔTM-His_6 obtained from one liter of bacterial culture was 6 mg.

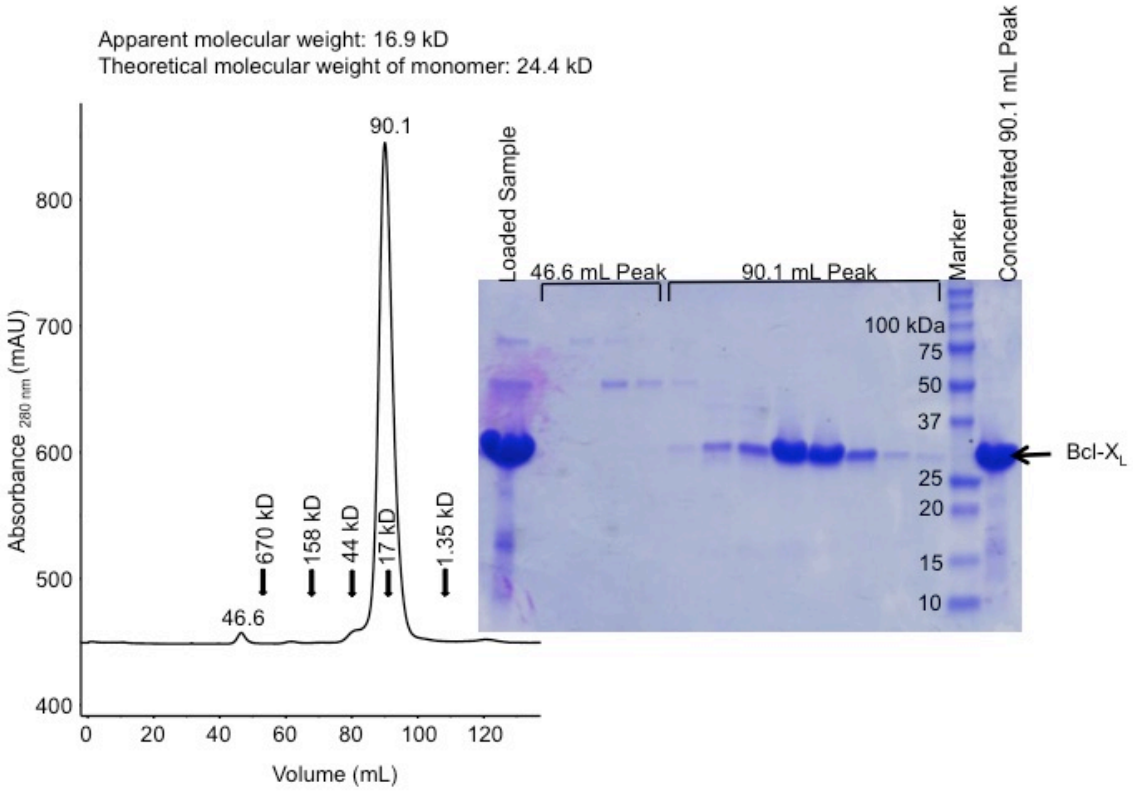

Figure 6.8. Size exclusion chromatogram and the corresponding SDS-PAGE of Bcl-X_L.
The elution positions of SEC standards are indicated by arrows.

6.3.7. Purified BHRF1 and BECN1 fragments interact weakly

We performed affinity pull-downs using His_6-tagged BHRF1ΔTM and MBP-tagged BECN1 fragments to confirm a direct interaction. We used Bcl-X_LΔTM which is known to interact with the BECN1 BH3D [46] as a positive control. Amylose affinity resin-bound MBP-BECN1(1-171) or MBP-BECN1(90-171) pull-down of Bcl-X_LΔTM-His_6 indicates that Bcl-X_L interacts directly with both of these BECN1 fragments (Figure 6.9A). This direct interaction was further confirmed by nickel affinity pull-down of the amylose elution fractions (Figure 6.9B). Similarly, amylose affinity resin-bound MBP-BECN1(1-171) or MBP-BECN1(90-171) pull-down of BHRF1ΔTM-His_6 (Figure 6.9A) indicates that BHRF1 interacts directly with both BECN1 fragments. This direct interaction was further confirmed reverse pull-down experiments wherein MBP-BECN1(1-171) or MBP-BECN1(90-171) were pulled down by BHRF1ΔTM-His_6 bound to nickel affinity resin (Figure 6.9B).

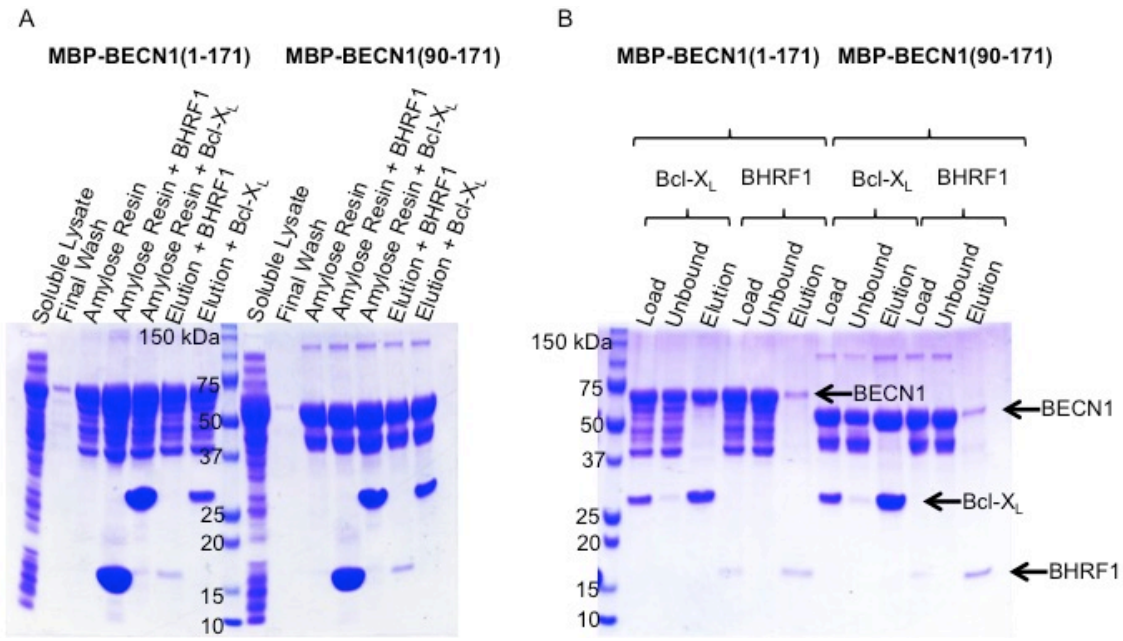

Figure 6.9. BECN1 interacts directly with BHRF1.
A) SDS-PAGE of an amylose affinity pull-down of purified BHRF1ΔTM-His_6 or Bcl-X_L(1-208, N52D/N66D)-His_6 by MBP-BECN1(1-171) or MBP-BECN1(90-171), as indicated above the gel. Amylose resin indicates the MBP-BECN1 bound resin after the final wash step. B) SDS-PAGE of nickel affinity pull-down of amylose affinity elution fractions from Figure 6.9A.

We used ITC to quantify and compare binding of BECN1 protein fragments to BHRF1ΔTM. FL $BECN1^{AFM}$ binds weakly to BHRF1ΔTM with a K_d of ~118 μM (Table 6.3, Figure 6.10). BECN1(90-171), which is the smallest BECN1 fragment shown by yeast two-hybrid and affinity pull-down assays to bind BHRF1, binds with a similar K_d of ~127 μM. For both interactions, the favorable ΔG is due to ΔS rather than due to ΔH, which are positive each case (Table 6.3). Finally, consistent with yeast two-hybrid assays, no binding was detected by ITC between the BECN1 BH3D and BHRF1ΔTM. Therefore, qualitative assessment of binding by yeast two-hybrid and affinity pull-downs, as well as quantification of binding affinity by ITC, all indicate that BECN1 residues 90-171 are required and sufficient for a direct, albeit weak, interaction with BHRF1.

Table 6.3. Thermodynamics of binding of BHRF1 to BECN1 fragments.

	BHRF1ΔTM			
BECN1 Fragment	K_d (μM)	ΔH (kJ/mol)	ΔS (J/K·mol)	ΔG (kJ/mol)
$BECN1^{AFM}$	117.7 ± 15.3	3.6 ± 0.5	87.5 ± 1.4	-22.1 ± 0.2
BECN1 residues 90-171	126.5 ± 4.6	7.2 ± 2.0	99.3 ± 7.2	-21.9 ± 0.1
BECN1 BH3D	No binding detected			

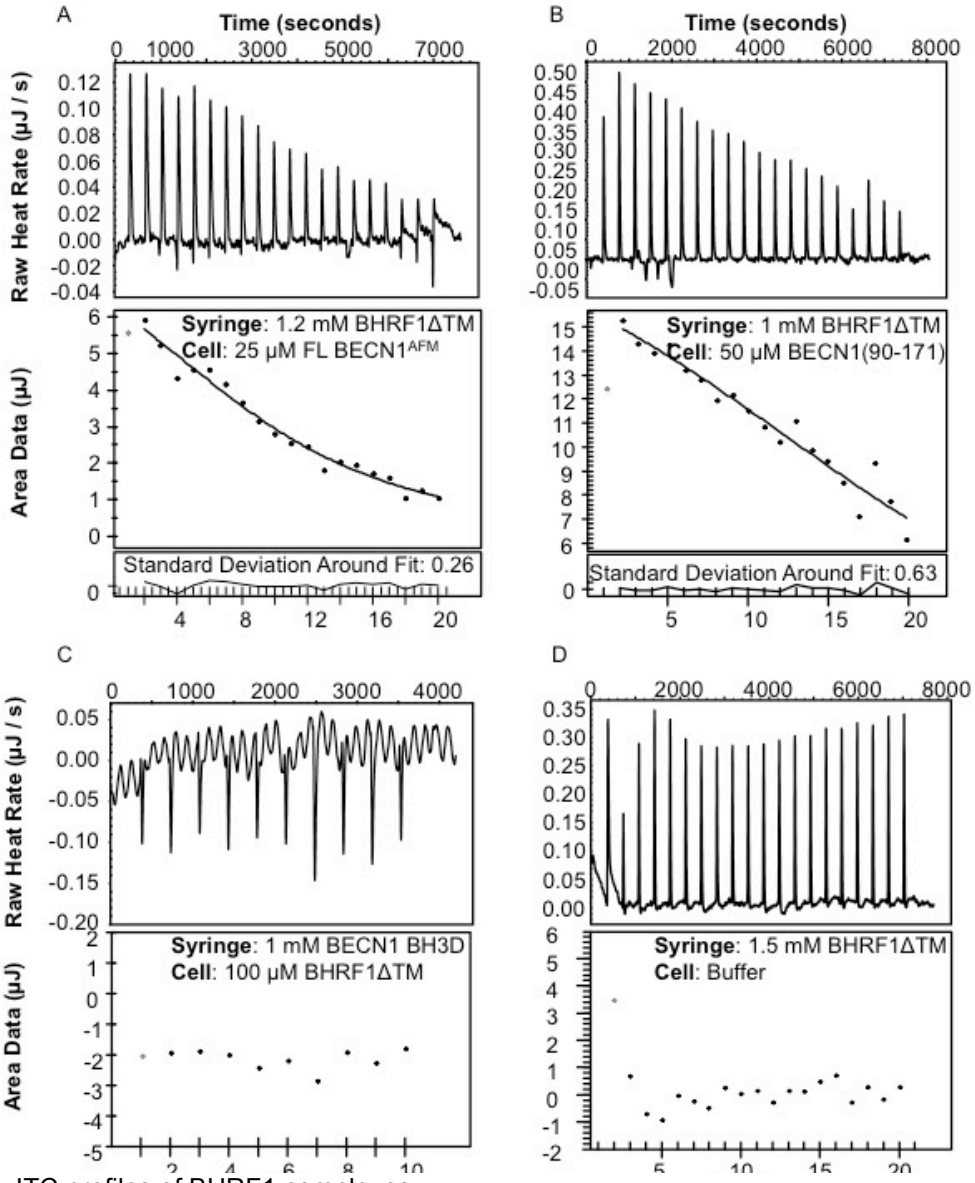

Figure 6.10. ITC profiles of BHRF1 complexes.
A) BHRF1ΔTM binding to BECN1^{AFM}, B) BHRF1ΔTM binding to BECN1 residues 90-171, C) BHRF1ΔTM binding to BECN1 BH3D, D) BHRF1ΔTM titrated into buffer.

6.3.8. BID binds tightly to BHRF1 and Bcl-X_L

ITC and fluorescence polarization binding assays have previously demonstrated that BHRF1 binds the pro-apoptotic BH3Ds of BID, BIM, and PUMA[174, 206] and competitive binding assays have demonstrated that the BID BH3D binds cellular BCL2 proteins (Bcl-2, Bcl-X_L, Bcl-w, Mcl-1, and A1) with weaker affinity compared to BIM and PUMA[175]. Therefore, we selected the BID BH3D to serve as a template for designing a selective inhibitor to prevent BHRF1-mediated autophagy down-regulation

without inhibiting cellular BCL2-mediated autophagy down-regulation since the WT BID BH3D tightly binds BHRF1 but binds the cellular BCL2 proteins more weakly relative to the other pro-apoptotic BH3Ds.

We quantified the thermodynamics of binding of the BID BH3D to BHRF1ΔTM and Bcl-X_LΔTM via ITC. We find that the BID BH3D binds BHRF1 tightly with a K_d of ~0.3 μM, which is ~3.6-fold tighter than it binds Bcl-X_L (K_d ~ 1.1 μM) (Table 6.4), ~400-fold tighter than BHRF1 binding to BECN1 (K_d ~ 118 μM) (Table 6.3), and ~6.7-fold tighter than Bcl-X_L binding to BECN1 (K_d ~2.0 μM)[30]. For both the BHRF1:BECN1 interaction and the BHRF1:BID BH3D interaction, the favorable ΔG is due to ΔS rather than due to ΔH, both of which are positive (Table 6.3, Figure 6.11, and Table 6.4). For the Bcl-X_L:BID BH3D interaction, contributions by both ΔS, which is positive, and ΔH, which is negative, contribute to the favorable ΔG (Table 6.4). For the Bcl-X_L:BECN1 BH3D interaction, the favorable ΔG is due to ΔH rather than due to ΔS, both of which are negative[30]. The differing entropic and enthalpic contributions between BECN1 or BID binding to BHRF1 or Bcl-X_L is consistent with differential modes of binding. Furthermore, the tighter binding of BHRF1 to the BID BH3D compared to BECN1 suggests that the BID BH3D can displace BECN1 to bind BHRF1. Our measured BHRF1 affinity (Table 6.4) for the BID BH3D is consistent with the previously published affinities (K_d = 0.10 μM) determined by ITC[206] and fluorescence polarization binding assays[174], and is similar to the BHRF1 affinity for the BAK BH3D (K_d = 0.15 μM)[206] determined by ITC. The BHRF1 binding affinity for BID and BAK is ~10-fold weaker than the BHRF1 affinity for the BIM BH3D (K_d = 0.020 μM)[206] determined by ITC.

Table 6.4. Thermodynamics of binding of the BID BH3D to BHRF1 and Bcl-X_L.

	BHRF1ΔTM				**Bcl-X_LΔTM**			
BID Peptide	**K_d (μM)**	**ΔH (kJ/mol)**	**ΔS (J/K·mol)**	**ΔG (kJ/mol)**	**K_d (μM)**	**ΔH (kJ/mol)**	**ΔS (J/K·mol)**	**ΔG (kJ/mol)**
BID BH3D	0.3 ± 0.1	18 ± 3.1	189.3 ± 7.2	-35.6 ± 1.0	1.1 ± 0.0	-20.4 ± 0.7	42.5 ± 2.3	-32.4 ± 0.0

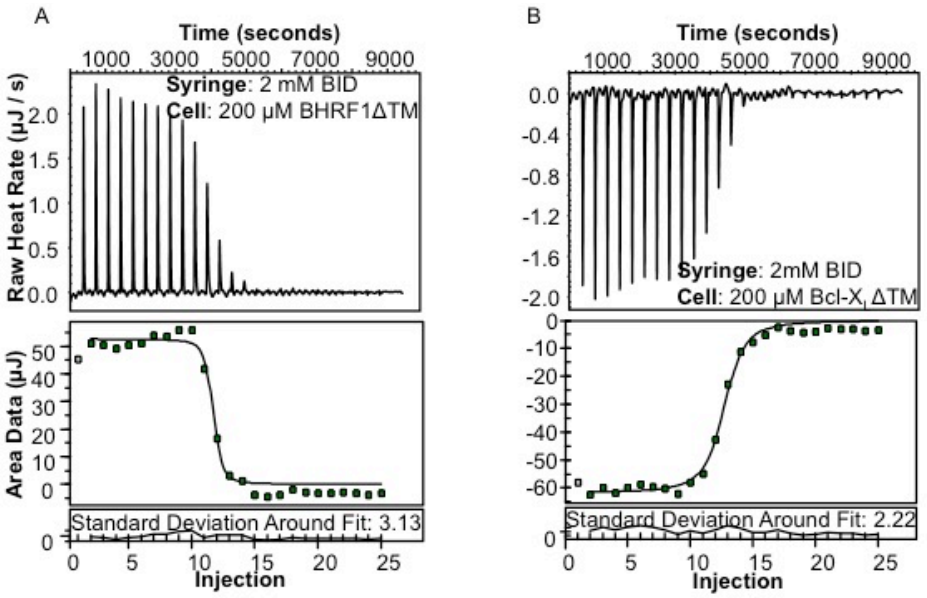

Figure 6.11. ITC profiles of BID BH3D binding to BHRF1 or Bcl-X_L.
A) BID BH3D binding to BHRF1ΔTM, B) BID BH3D binding to Bcl-X_LΔTM.

6.3.9. The hydrophobic surface groove of BHRF1 binds the BID BH3D

In order to improve our understanding of binding determinants for the BHRF1 interaction with the BID BH3D, we determined a 2.6 Å resolution X-ray crystal structure of the BHRF1ΔTM:BID BH3D complex (Figure 6.12A,B). The BHRF1:BID BH3D co-crystals belonged to the space group $P6_5$, with 5 complexes per asymmetric unit. Data collection and refinement statistics indicate a good structure solution was obtained (Table 6.5). Simultaneous superimposition of multiple structures by THESEUS[212] via the PSSweb server (http://pssweb.org)[213] indicates that all five complexes within the asymmetric unit superimpose well with an average RMSD of 0.30 ± 0.24 Å for the atomic Cartesian coordinates of the backbone atoms over 174 residues. Likewise, pairwise structural superimposition of each complex performed against BHRF1:BID BH3D chains A and B using *PyMOL* (v.1.3; Schrödinger) also indicates that each complex superimposes well against BHRF1:BID BH3D chains A and B (Table 6.6). Since all five complexes superimpose well, the complex composed of BHRF1:BID BH3D chains A and B was used for all structural analyses. The final crystal structure represents residues 2-156 of BHRF1 and residues 81-98 of BID. The total surface area buried upon formation of the BHRF1:BID BH3D complex is 1552 $Å^2$, wherein 857 $Å^2$ is contributed by the BID BH3D and 695 $Å^2$ by BHRF1, as calculated using PISA[145]. Helices α2-5 of BHRF1 form the ligand-binding groove (Figure 6.12A).

Table 6.5. X-ray data collection and structure refinement statistics.

Data collection statistics	
Space group	$P6_5$
Unit-cell parameters (Å)	a=b=186.8, c=70.5
Resolution (Å)	161.79-2.61
Wilson B factor (Å^2)	74.21
Unique reflections	42919 (4267)
Completeness (%)	99.9 (100.0)
Multiplicity	11.3 (11.5)
$\langle I/\sigma(I) \rangle$	19.6 (1.7)
$CC_{1/2}$	0.998 (0.611)
Refinement statistics	
Resolution range	80.89-2.61 (2.70-2.61)
Completeness (%)	99.85 (99.95)
R_{work}† (%)	0.1976 (0.3375)
R_{free}†‡ (%)	0.2597 (0.3959)
No. of protein residues	848
No. of waters	37
RMSD, bond lengths (Å)	0.009
RMSD, bond angles (°)	1.19
Ramachandran outliers (%)	0.0
Average B factor (Å^2)	
Overall	92.4
Macromolecules	92.5
Solvent	69.0

Values in parentheses pertain to the outermost shell of data.

† R factor = $\sum_{hkl}||F_{obs}| - |F_{calc}||/\sum_{hkl}|F_{obs}|$.

‡ The test set for R_{free} consisted of 5% of the data.

Table 6.6. Structural superimpositions of complexes in the ASU performed against BHRF1:BID BH3D chains A and B.

BHRF1:BID BH3D Complex	**RMSD**
BHRF1:BID BH3D chains C and D	0.213 Å over 152 C^{α} atoms
BHRF1:BID BH3D chains E and F	0.259 Å over 160 C^{α} atoms
BHRF1:BID BH3D chains G and H	0.336 Å over 162 C^{α} atoms
BHRF1:BID BH3D chains I and J	0.413 Å over 128 C^{α} atoms

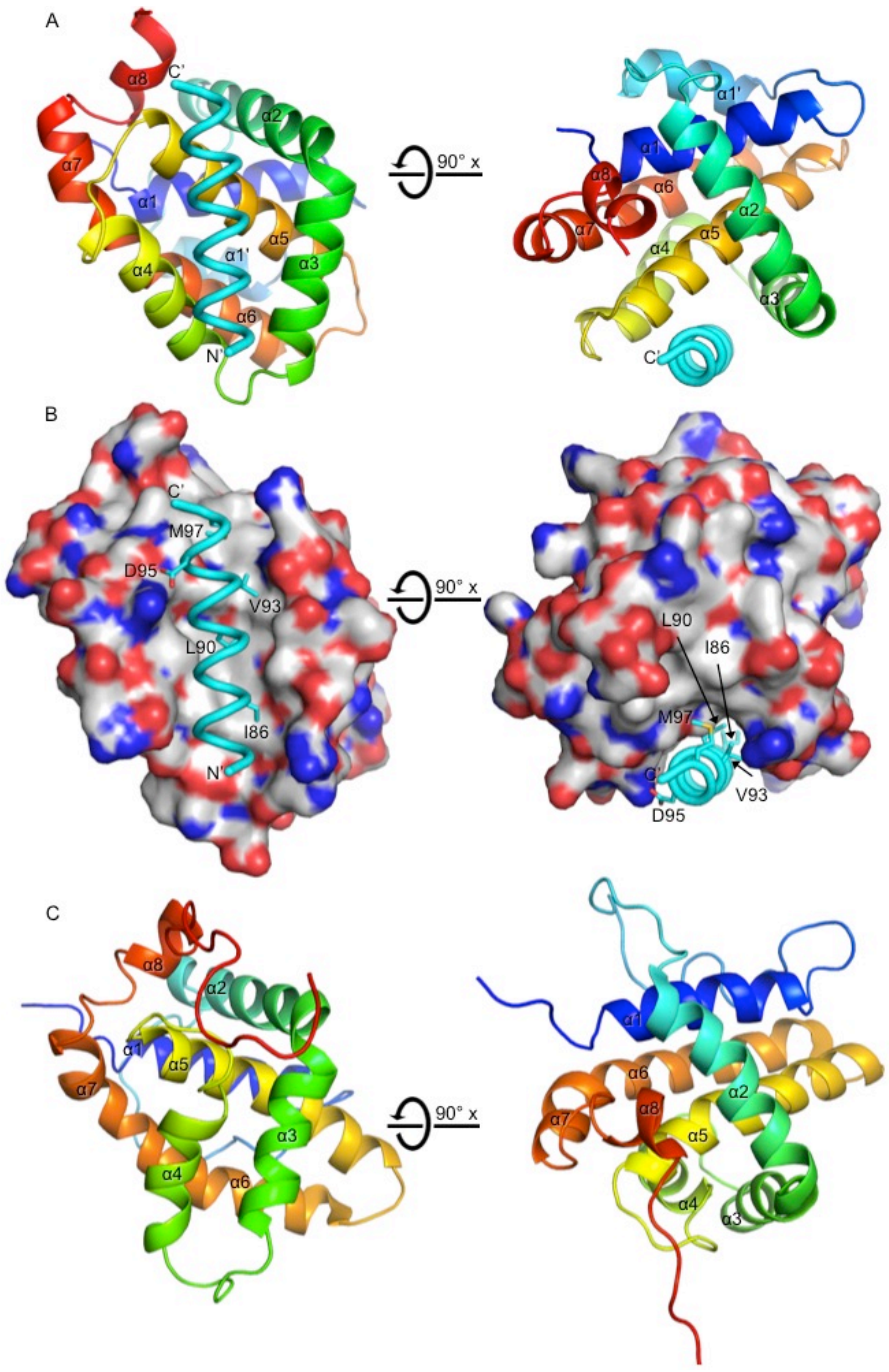

Figure 6.12. Crystal structure of the BHRF1:BID BH3D complex.
A) The BID BH3D backbone is traced in cyan ribbon while the BHRF1 backbone is traced in a ribbon colored in rainbow shades with blue at the N-terminus and red at the C-terminus. BHRF1 helices are labeled α1, α1', and α2-8. B) The BID BH3D backbone is as in (A), while the BHRF1 surface is colored by atom type: oxygen, red; nitrogen, blue, carbon, grey; sulfur, yellow. Conserved hydrophobic and salt bridge-forming BID BH3D residues that interact with BHRF1 are shown in stick and colored by atom type: oxygen, red; nitrogen, blue; carbon, cyan; sulfur, yellow. C) Ligand-free BHRF1[205] backbone is traced in a ribbon colored in rainbow shades with blue at the N-terminus and red at the C-terminus. Helices are labeled as α1-8. All molecular figures are displayed in superimposable views and were made using *PyMOL* (v.1.3; Schrödinger).

Analysis of the buried surface area (BSA) of interface residues and electrostatic interactions that stabilize the complex indicate that BID interacts with BHRF1 in a manner similar to other interactions between pro-apoptotic BH3 only proteins and anti-apoptotic BCL2s. The BSA calculated using PISA[145] indicates that the following BHRF1 residues contribute to the BID BH3D interface: I57, R60, N61, T64, F65, E67, T68, R71, F72, H75, D82, S85, V86, E89, I90, R93, S97, L98, G99, R100, L102, A104, W104, and W107. BID BH3D residues I82, I83, R84, I86, A87, H89, L90, A91, V93, G94, D95, M97, and D98 all contribute to the interface, as calculated using PISA[145]. Among the BID interface residues, the four hydrophobic residues that define the BH3D consensus sequence[217] (I86, L90, V93, and M97) are significantly buried by the interaction with BHRF1 (Figure 6.12B). Equivalent residues in the BH3D of various pro-apoptotic BCL2 proteins are also significantly buried by Bcl-X_L[201], A1[200], Mcl-1[199], and M11[44]. Three electrostatic interactions also stabilize the BHRF1:BID BH3D complex. These include a salt bridge between BHRF1 E89 and BID BH3D R84, a salt bridge BHRF1 R100 and BID BH3D D95, and a hydrogen bond between BHRF1 S97 and BID BH3D D98. Residues equivalent to those forming the salt bridge between BHRF1 R100 –and BID BH3D D95 (Figure 6.12B) form a salt bridge in Bcl-X_L complexes with the BH3D of BID[206], BIM[202], and BECN1[46]; Mcl-1 complexes with the BH3D of BID[199] and BIM[218]; A1 complexes with the BH3D of BID[200]; and M11 complexes with the BH3D of BECN1[44].

6.3.10. The BHRF1:BID BH3D complex is structurally similar to complexes of BHRF1 and either the BAK or BIM BH3Ds

The solution structure of ligand-free BHRF1ΔTM has been previously determined by NMR (PDB entry 1Q59)[205] and superimposes with the BHRF1:BID BH3D complex with a pairwise RMSD of 3.678 Å over 148 C^{α} atoms, indicating the structures are similar but have significant differences. BHRF1 has also been previously crystallized in complex with either the BAK or BIM BH3D[206]. Like the BHRF1:BAK (PDB entry 2XPX; 2.05 Å resolution) and BHRF1:BIM (PDB entry 2V6Q; 2.7 Å resolution) complexes[206], the BHRF1:BID complex has an all α-fold comprising 9 helices in BHRF1 with the BID BH3D forming a single helix within a hydrophobic surface groove of BHRF1 (Figure 6.12A,B). We labeled the helices α1, α1', and α2-8, consistent with the BHRF1:BAK and BHRF1:BIM complexes[206], because the BHRF1 complexes contain an extra helix between the α1 and α2 relative to the ligand free BHRF1[205] (Figure 6.12C) and Bcl-X_L:BID complex (PDB entry 4QVE)[201]. The BHRF1:BID BH3D complex superimposes with a pairwise

RMSD of 0.401 Å over 147 C^{α} atoms of the BHRF1:BAK complex, and with a pairwise RMSD of 0.462 Å over 155 C^{α} atoms of the BHRF1:BIM complex, indicating that all three of these complexes are very similar. The formation of these BHRF1:BH3D complexes requires significant rearrangement of α3 and α4 (Figure 6.12A), which obscure the binding groove in the ligand-free state (Figure 6.12C).

6.3.11. Interface residues contribute differentially to BHRF1 interactions with the BID, BAK, and BIM BH3Ds

Despite the overall structural similarity between our BHRF1:BID BH3D and the previously published BHRF1:BH3D complexes, BSA analysis by PISA[145] interface residues differentially contribute to the complexes. In our crystal structure, the BHRF1:BID BH3D interaction buries 857 $Å^2$ of the BID BH3D and 695 $Å^2$ of BHRF1, for a total of 1552 $Å^2$; with 13 BID residues and 24 BHRF1 residues contributing to the interface. The BHRF1:BAK BH3D interaction buries 956 $Å^2$ of the BAK BH3D and 783 $Å^2$ of BHRF1, for a total of 1739 $Å^2$; with 15 BAK residues and 25 BHRF1 residues contributing to the interface. The BHRF1:BIM BH3D interaction buries 1038 $Å^2$ of the BIM BH3D and 882 $Å^2$ of BHRF1, for a total of 1920 $Å^2$; with 17 BIM residues and 27 BHRF1 residues contributing to the interface. Of these three complexes, the interaction with the BIM BH3D buries the largest surface area of each partner, which is consistent with the observation reported above that BHRF1 binds to the BIM BH3D with the tightest affinity compared to the BID and BAK BH3Ds.

In order to improve our understanding of binding determinants for the BID, BAK, and BIM BH3Ds for BHRF1, we used PISA[145] to analyze the BSA of conserved interface residues. Among the residues conserved between BID, BAK, and BIM; nine residues (corresponding to BID I86, A87, H89, L90, V93, G94, D95, M97, D98) are involved in the interface with BHRF1 (Figure 6.13A). These residues are all similarly oriented in each complex except for the BIM F69 aromatic ring, which is rotated relative to the BID M97 and BAK I85 side chains (Figure 6.13B). BID M97 and BAK I85 interact with BHRF1 I57, L102, and A103. The BIM F69 maintains contact with BHRF1 I57 and L102 but also contacts with BHRF1 L53, R60, and L98. The BHRF1:BIM BH3D complex buries 111 $Å^2$ of BIM F69 and 181 $Å^2$ of BHRF1 L53, I57, R60, L98, L102, and A103. The BHRF1:BID BH3D complex buries 98 $Å^2$ of BID M97 and 96 $Å^2$ of BHRF1 L53, I57, R60, L98, L102, and A103. The BHRF1:BAK BH3D complex buries 55 $Å^2$ of BAK I85 and 129 $Å^2$ of BHRF1 L53, I57, R60, L98, L102, and A103. In total, these BH3D and BHRF1 residues contribute

292 Å^2 of the BHRF1:BIM complex BSA, compared to 194 Å^2 of the BHRF1:BID complex BSA and 184Å^2 of the BHRF1:BAK complex BSA. Therefore, F69 may be important for the improved affinity for BHRF1 of BIM relative to BID and BAK.

Additionally, among interface residues (Figure 6.13A), the BIM BH3D has an R in the position corresponding to BIM R63, which is moderately conserved among various pro-apoptotic BH3-only BCL2s[175] but is an A in the BH3D of BID and BAK. BSA and interface analysis by PISA[145] indicates that the BIM R63 contributes more to the BHRF1 interface than do the corresponding BID A91 and BAK A79. The BID A91 and BAK A79 contribute 37 Å^2 and 47 Å^2 BSA, respectively. The BIM R63 contributes 111 Å^2 BSA and forms two hydrogen bonds with the main chain of BHRF1 E89 via the BIM R63 side chain NΣ and NH2. The extra hydrogen bonds afforded by BIM R63 and greater buried surface area afforded by BIM F69 likely also contribute to the improved affinity of BIM for BHRF1 relative to BID and BAK.

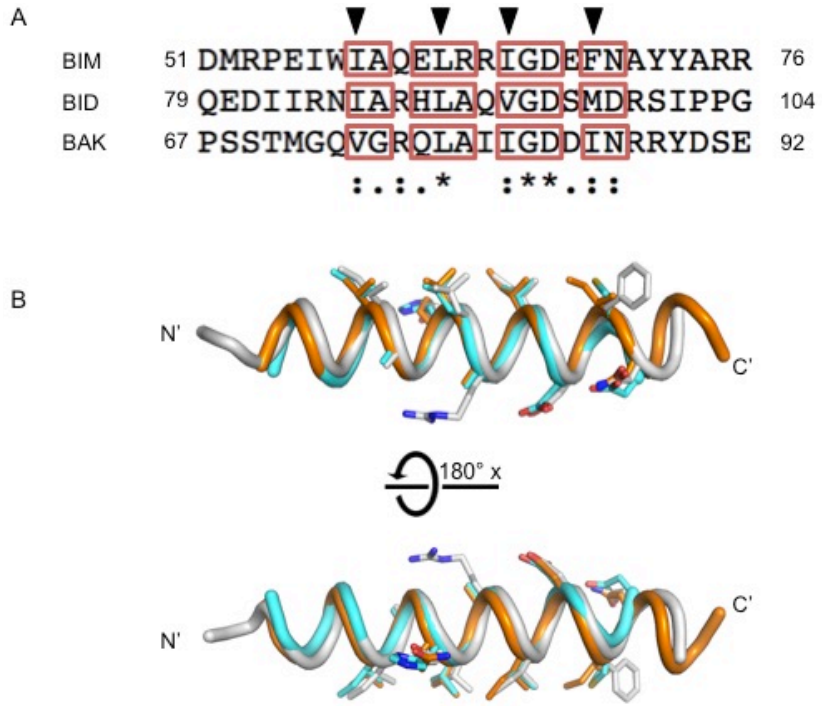

Figure 6.13. Comparison of pro-apoptotic BH3Ds that interact with BHRF1.
A) Sequence alignment of the BH3D of BID, BIM, and BAK. Symbols beneath the alignment indicate residue conservation: with an asterisk (*) corresponding to invariant, colon (:) corresponding to highly conserved, and period (.) corresponding to moderately conserved residues. The arrows above the sequence denote the four hydrophobic residues of the BH3 consensus sequence. Red boxes indicate residues within the BH3 consensus sequence that form the interface with BHRF1. B) Structural superimposition of the BID (cyan), BIM (gray) (PDB entry 2V6Q), and BAK (orange) (PDB entry 2XPX) BH3Ds from the BHRF1 complex. Interface residues boxed in red in Figure 6.13A are shown in stick colored by atom type: oxygen, red; nitrogen, blue; carbon, cyan/gray/orange; sulfur, yellow).

6.3.12. Interface residues contribute differentially to BID BH3D interactions with BHRF1 and cellular anti-apoptotic BCL2s

The BID BH3D has also been previously crystalized in the presence of human Bcl-X_L (PDB entry 4QVE; 2.05 Å resolution)[201], human Mcl-1 (PDB entry 5C3F; 1.43 Å resolution)[199], and mouse A1 (PDB entry 2VOI; 2.1 Å resolution)[200]. Our BHRF1:BID complex superimposes with the Bcl-X_L:BID BH3D complex with a pairwise RMSD of 2.352 Å over 124 C^{α} atoms, the Mcl-1:BID BH3D complex with a pairwise RMSD of 3.555 Å over 136 C^{α} atoms, and the AI:BID BH3D complex with a pairwise RMSD of 3.676 Å over 126 C^{α} atoms, indicating that the complexes are similar but have significant differences. However, the BID from our BHRF1:BID complex superimposes with the BID from the Bcl-X_L:BID structure with a pairwise RMSD of 0.500 Å over 17 C^{α} atoms, with the BID from the Mcl-1:BID structure with a pairwise RMSD of 0.404 Å over 16 C^{α} atoms, and with the BID from the A1:BID structure with a pairwise RMSD of 0.434 Å over 16 C^{α} atoms, indicating that the BID BH3D structures are similar in all four complexes.

In order to improve our understanding of binding determinants for the BID BH3D interactions with BCL2 proteins, we used PISA[145] to analyze the BSA of interface residues within the BH3D. The BH3 consensus sequence is $\phi_1\Sigma XX\phi_2XX\phi_3\Sigma DZ\phi_4\Gamma$, where ϕ is a hydrophobic residue, Σ is a small residue, Z is usually an acidic residue, and Γ is usually a hydrophilic residue[217]. In BID, the BH3D corresponds to residues 86-98. BID BH3D residues involved in all interfaces include I86, A87, H89, L90, A91, V93, G94, D95, M97, and D98. BID S96 is involved in the interface in the Bcl-X_L complex but is solvent exposed in the other three complexes. The Bcl-X_L interface buries 26 $Å^2$ of BID S96 and provides a hydrogen bond between the backbone carbonyl of BID S96 and the Bcl-X_L R100 NH1. Since BID S96 contributes to the Bcl-X_L interface, but not to the other interfaces, it is likely more important for binding specificity of the BID BH3D for Bcl-X_L than for BHRF1, Mcl-1, or A1.

Among the interface residues common to each complex, BSA analysis and structural superimposition of the BID BH3D from the four complexes demonstrate that the BID H89 and M97 side chains differ and have variable contributions to the BSA. The position of the Cγ of BID H89 differs by 2.3 Å among the four complexes, causing the imidazole to occupy different positions when bound to BHRF1, Bcl-X_L, Mcl-1, or A1 (Figure 6.14). BID H89 contributes 81 $Å^2$ BSA to the Bcl-X_L interaction, 74 $Å^2$ BSA to

mouse A1 interaction, 49 Å^2 BSA to BHRF1 interaction, and 41 Å^2 BSA to Mcl-1 interaction. BID H89 provides an additional salt bridge with A1 E47 but does not mediate a salt bridge with BHRF1, Bcl-X_L, or Mcl-1. BID H89 also provides a hydrogen bond between the imidazole of BID H89 and the hydroxyl of Bcl-X_L Y101 but does not mediate any hydrogen bonds with BHRF1 or Mcl-1. The increased BSA as well as the extra salt bridge and hydrogen bond mediated by BID H89 in the Bcl-X_L and A1 indicate that BID H89 is likely more important for binding specificity of the BID BH3D for Bcl-X_L and A1 than for BHRF1 and Mcl-1. The position of the C$^\gamma$ of BID M97 differs by 2.7 Å among the four complexes, causing the side chain to occupy different positions when bound to BHRF1, Bcl-X_L, Mcl-1, or A1 (Figure 6.14). BID M97 contributes 128 Å^2 BSA to the Bcl-X_L interaction, 120 Å^2 BSA to mouse A1 interaction, 98 Å^2 BSA to BHRF1 interaction, and 92 Å^2 BSA to Mcl-1 interaction. BID M97 provides a hydrogen bond between the backbone carbonyl of BID M97 and the hydroxyl of Bcl-X_L Y195 bud does not mediate a hydrogen bond with BHRF1, Mcl-1, or A1. The increased BSA of BID M97 in the Bcl-X_L and A1 complexes, as well as the extra hydrogen bond it provides in the Bcl-X_L complex, indicate that BID M97 is likely more important for binding specificity of the BID BH3D for Bcl-X_L and A1 than for BHRF1 and Mcl-1.

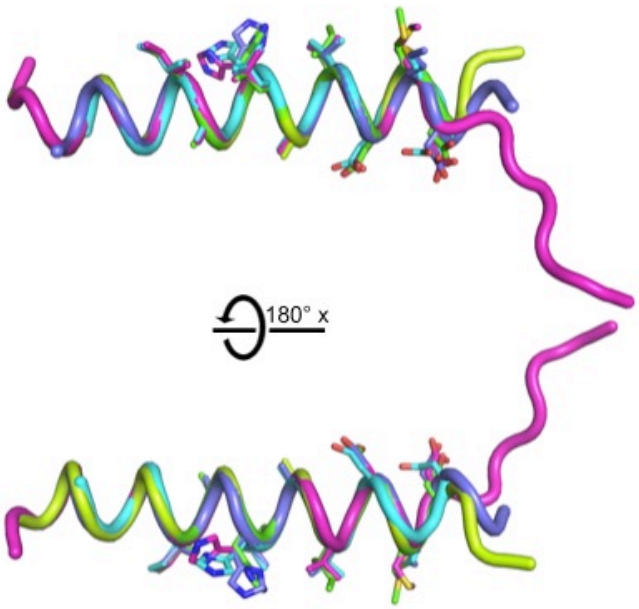

Figure 6.14. Structural superimposition of the BID BH3D in complex with anti-apoptotic BCL2 homologs. The BH3Ds of BID from the BHRF1 complex (cyan), Bcl-X_L complex (magenta) (PDB entry 4QVE), Mcl-1 complex (lavender) (PDB entry 5C3F), and A1 complex (green) (PDB entry 2VOI). Interface residues from the BH3 consensus motif are shown in stick colored by atom type: oxygen, red; nitrogen, blue; carbon, cyan/magenta/lavender/green; sulfur, yellow).

BID D95 corresponds to the conserved D in the BH3 consensus sequence. In each complex, BID D95 forms a salt bridge with the anti-apoptotic BCL2 (via BHRF1 R100, Bcl-X_L R139, Mcl-1 R263, and A1 R88). The carboxyl of BID D95 is also involved in a hydrogen bond to the amide of Asns; Bcl-X_L N136, Mcl-1 N260, or A1 N85. BID D95 contributes 46 $Å^2$ BSA to the Bcl-X_L, Mcl-1, and A1 interfaces and 30 $Å^2$ BSA to the BHRF1 interface. The increased BSA and extra hydrogen bonds of BID D95 in the Bcl-X_L, Mcl-1, and A1 complexes suggest that BID D95 may be more important for binding specificity of the BID BH3D for the cellular BCL2s than for BHRF1.

BID D98 forms a salt bridge with A1 K147, a hydrogen bond between the carboxyl of BID 98 and the backbone amine of A1 W86, a hydrogen bond between the carboxyl of BID D98 and the backbone amine of A1 G87, a hydrogen bond between the carboxyl of BID D98 and the backbone amine of Mcl-1 G262, and a hydrogen bond between the carboxyl of BID D98 and the hydroxyl of BHRF1 S97. BID D98 contributes 91 $Å^2$ BSA to the A1 interface, 74 $Å^2$ BSA to the Mcl-1 interface, 47 $Å^2$ BSA to the Bcl-X_L interface, and 38 $Å^2$ BSA to the BHRF1 interface. Since BID D98 mediates a salt bridge and three hydrogen bonds with A1, mediates a hydrogen bond with Mcl-1, and contributes more BSA to A1 and Mcl-1 compared to BHRF1 and Bcl-X_L, D98 is likely more important for binding specificity of the BID BH3D for A1 and Mcl-1 than for BHRF1 and Bcl-X_L.

Although BID E80 and R84 are located outside of the BH3 consensus sequence, these residues may also contribute differentially to binding specificity for cellular BCL2s and BHRF1. BID E80 forms two salt bridges with Mcl-1, one with Mcl-1 R248 and another with Mcl-1 H252, but does not mediate any salt bridges within the other complexes. Therefore, E80 is likely more important for binding specificity of the BID BH3D to Mcl-1 than to BHRF1, Bcl-X_L, or A1. BID R84 forms a salt bridge with BHRF1 E89 but does not mediate any salt bridges or hydrogen bonds with Bcl-X_L, Mcl-1, and A1. Therefore, BID R84 is likely more important for binding specificity of the BID BH3D to BHRF1 than to Bcl-X_L, Mcl-1 and A1.

6.3.13. Differences in BECN1 BH3D interface residues may destabilize the BHRF1 interaction

In order to better understand why the isolated BECN1 BH3D does not bind BHRF1, we assessed sequence conservation of interface residues between the BH3Ds of BECN1, which binds M11[44] and Bcl-X_L[46], and BIM, BAK, and BID (Figure 6.15). Among the residues conserved between BECN1, BID, BAK, and BIM; eight residues (corresponding to BECN1 L112, S113, R115, L116, G120, D121, F123, D124)

are involved in the interface with BHRF1, M11, or Bcl-X_L (Figure 6.15). BECN1 K117, which is equivalent to the BIM R63 identified above as contributing to the BHRF1 specificity and is moderately conserved among various pro-apoptotic BH3-only BCL2s[175], also participates in the interface with M11 and Bcl-X_L.

The position equivalent to BID V93, which was a conserved hydrophobic residue in BID, BIM, and BAK is a polar residue in BECN1 (T119). BECN1 F123, which is equivalent to the BIM F69 identified above as contributing to the BHRF1 specificity, participates in the interface with M11 and Bcl-X_L. Therefore, we performed structural superimposition of the M11:BECN1 BH3D[44], Bcl-X_L:BECN1 BH3D[46], BHRF1:BAK and BHRF1:BIM complexes[206] against our BHRF1:BID BH3D to analyze the different contributions of these residues to their respective interactions.

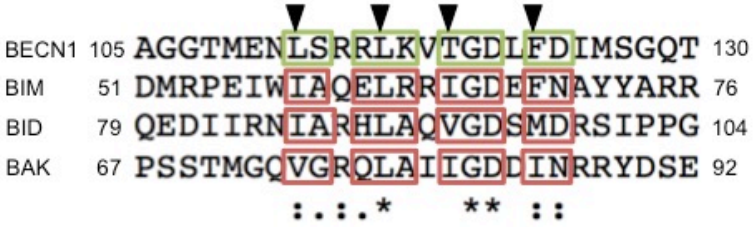

Figure 6.15. Sequence alignment of the BH3D of BECN1, BIM, BID, and BAK.
Symbols beneath the alignment indicate residue conservation: with an asterisk (*) corresponding to invariant, colon (:) corresponding to highly conserved, and period (.) corresponding to moderately conserved residues. The arrows above the sequence denote the four hydrophobic residues of the BH3 consensus sequence. Red boxes indicate residues within the BH3 consensus sequence that form the interface with BHRF1. Green boxes indicate residues within the BH3 consensus sequence that form the interface with M11[44] and Bcl-X_L[46].

BECN1 contains K117 at a position equivalent to the BIM R63 (Figure 6.15). BECN1 K117 forms two hydrogen bonds via its NZ with the backbone carbonyl and side chain hydroxyl of S77 on M11[44]. BHRF1 contains a glutamate (E89) at the position equivalent to M11 S77. The side chain of BHRF1 E89 is rotated away such that the BHRF1 E89 side chain carboxyls are positioned greater than 4 Å away from the NZ of BECN1 K117. Therefore, BECN1 K117 is unlikely able to mediate equivalents of either of the hydrogen bonds observed in the M11:BECN1 BH3D complex (with M11 S77) and the BHRF1: BIM BH3D complex (with BHRF1 E89 and BIM R63). BECN1 K117 forms salt bridges with Bcl-X_L E129 and D133[46]. Bcl-X_L E129 is positioned equivalent to BHRF1 E89, which is located at the C-terminus of α4. However, structural superimposition of the Bcl-X_L:BECN1 BH3D, BHRF1:BIM, and BHRF1:BAK complexes onto our BHRF1:BID BH3D complex demonstrates that the BHRF1 E89 main chain is rotated relative to Bcl-X_L E129 such that the Cα positions differ by 1 Å, pointing the BHRF1 E89 side chain away from the bound

BH3Ds such that the BHRF1 E89 side chain carboxyls are positioned greater than 5 Å away from the NZ of BECN1 K117. Therefore, BECN1 K117 is unlikely able to mediate a salt bridge with BHRF1. Bcl-X_L D133 is located in a loop equivalent to the loop linking α4 and α5 of BHRF1. The only negatively charged residue in this loop of BHRF1 is D95, which is disordered in the BHRF1:BIM BH3D complex, and rather than interact with BHRF1 D95, the BIM R63 (equivalent to BECN1 K117) forms two hydrogen bonds with the main chain carbonyl of BHRF1 E89. Therefore, BECN1 K117 is unlikely able to form equivalent salt bridges with BHRF1 due to the suboptimal position of the BHRF1 E89 side chain and the flexibility of the loop containing BHRF1 D95. However, since BID and BAK have an alanine at the position equivalent to BECN1 K117 / BIM R63 yet still interact with BHRF1, the fact that the BECN1 BH3D is insufficient for an interaction with BHRF1 cannot be solely attributed to the inability of BECN1 K117 to mediate a double hydrogen bond or salt bridges with BHRF1.

BECN1 T119, which was a conserved hydrophobic residue in BID, BIM, and BAK (Figure 6.15), occupies a position typically filled by a hydrophobic residue in BH3Ds. BH3D residues from pro-apoptotic BCL2 homologs corresponding to $\phi_{1\text{-}4}$ of the BH3 consensus sequence are significantly buried by Bcl-X_L[46, 201], A1[200], Mcl-1[199], and M11[44]. Although BECN1 contains a polar residue (T119) at the ϕ_3 position, equivalent residues of BID, BIM, and BAK bury similar amounts of surface area. Binding to BHRF1 buries 78 $Å^2$ of BID V93, 82 $Å^2$ of BIM I65, and 71 $Å^2$ of BAK I81, while binding to M11 buries 79 $Å^2$ of BECN1 T119, and binding of Bcl-X_L buries 80 $Å^2$ of BECN1 T119. A hydrophobic pocket, primarily formed by the aromatic rings of M11 F48 and Y52 or Bcl-X_L F97 and Y101, accommodates the aliphatic part of the BECN1 T119 side chain in the M11 and Bcl-X_L complexes. This pocket on BHRF1 is primarily formed by I57, N61, T64 Cγ, and F65. Superimposition of the BECN1 from the M11 and Bcl-X_L complexes onto the BHRF1:BID, BHRF1:BIM, and BHRF1:BAK complexes indicates that the BECN1 T119 Cγ would be positioned within ~2.8 Å of the BHRF1 N61 amide (BHRF1:BID complex), ~3.1 Å of the BHRF1 N61 amide (BHRF1:BIM complex), or ~2.9 Å of the N61 amide (BHRF1:BAK complex). Therefore, this pocket of the BHRF1 hydrophobic surface groove should be able to accommodate BECN1 T119. Further, since BH3D residues equivalent to BECN1 T119 bury equivalent amounts of surface area in the BHRF1 complexes, the presence of a polar residue in the ϕ_3 position of the BECN1 BH3D is inadequate to explain why the BECN1 BH3D is insufficient to mediate the interaction with BHRF1.

The position equivalent to BECN1 F123 is conserved between all four BH3Ds analyzed (Figure 6.15). However, since the equivalent BIM residue (F69) was identified above as a potential determinant for BIM specificity for BHRF1, we analyzed the structural superimposition of the BECN1 BH3D from the complex with M11[44] and Bcl-X_L[46] against BIM from the BHRF1 complex[206]. The BIM F69 Cα is 0.6 Å from BECN1 F123 in the Bcl-X_L complex and 1.1 Å from BECN1 F123 in the M11 complex, which alters the position of the aromatic ring, placing it within ~2 Å of BHRF1 I57 C**γ** on the surface groove of BHRF1. Therefore, the pocket on the hydrophobic groove of BHRF1 is too shallow to accommodate these positions of the BECN1 F123 aromatic ring (Figure 6.16). Since BECN1 F123 has previously been identified as important for the BECN1 interaction with M11 and Bcl-X_L[140], this difference in the BHRF1 hydrophobic groove likely destabilizes the BHRF1:BECN1 BH3D interaction. The destabilization caused by the orientation of BECN1 F123 combined with the absence of double hydrogen bond or salt bridges mediated by BECN1 K117 may be factors that contribute to the inability of the BECN1 BH3D to form a stable interaction with BHRF1.

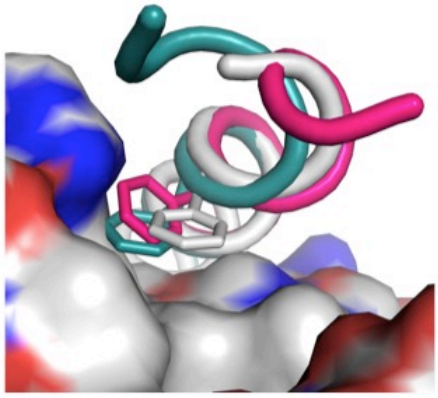

Figure 6.16. Comparison of BECN1 and BIM residues equivalent to BECN1 F123.
Structural superimposition of the BECN1 from the Bcl-X_L:BECN1 complex (magenta) (PDB entry 2P1L) and BECN1 from the M11:BECN1 complex (teal) (PDB entry 3DVU) against BIM (gray) from the BHRF1:BIM complex (PDB entry 2V6Q) with the BHRF1 surface colored by atom type: oxygen, red; nitrogen, blue, carbon, grey; sulfur, yellow. BH3D residues equivalent to BECN1 F123 are shown in stick.

6.3.14. Designing a peptide to inhibit down-regulation of autophagy by BHRF1

In an effort to discover a BHRF1-selective peptide, we also quantified the thermodynamics of binding to BHRF1ΔTM and Bcl-X_LΔTM by the BID BH3D mutant peptides via ITC (Table 6.7, Figure 6.17, 6.18). However, in all cases, the mutant BID BH3D peptides used here weakened binding 4 to 40-fold to BHRF1 and maintained or improved binding to Bcl-X_L (Table 6.7) relative to the WT BID BH3D (0.3 μM

for BHRF1 and 1.1 μM for Bcl-X_L) (Table 6.4). Unfortunately, therefore, the mutant BID BH3D peptides used here were not selective for BHRF1 but instead improved selectivity for Bcl-X_L.

Therefore, we decided to investigate whether a cell-permeable peptide derived from the WT BID BH3D would prevent BHRF1-mediated down-regulation of BECN1-dependent autophagy. Compared with untreated cells, TAT-BID peptide treatment markedly increases autophagy levels in starved cells that are transiently co-transfected with BHRF1 ($p = 0.0130$ for treated versus untreated cells co-transfected with BHRF1 and BECN1) (Figure 6.2B). This suggests that the TAT-BID peptide binds BHRF1 in mammalian cells and prevents BHRF1 from down-regulating BECN1-mediated autophagy, despite the low affinity of BHRF1 for the isolated BECN1 BH3D.

Table 6.7. Thermodynamics of binding of BID BH3D mutants to BHRF1 and Bcl-X_L.

	BHRF1ΔTM				**Bcl-X_LΔTM**			
BID Peptide	**K_d (μM)**	**ΔH (kJ/mol)**	**ΔS (J/K·mol)**	**ΔG (kJ/mol)**	**K_d (μM)**	**ΔH (kJ/mol)**	**ΔS (J/K·mol)**	**ΔG (kJ/mol)**
BID G94A	1.2 ± 0.0	43.9 ± 10.5	268.5 ± 36.8	-32.1 ± 0.0	0.1 ± 0.0	-17.1 ± 5.7	73.0 ± 19.0	-62.2 ± 13.3
BID Mutant 1	1.2 ± 0.2	-7.4 ± 0.9	87.8 ± 1.6	-32.3 ± 1.4	0.1 ± 0.0	-13.2 ± 1.1	86.4 ± 4.4	-37.7 ± 2.4
BID Mutant 2	12.2 ± 2.0	-1.8 ± 0.2	88.0 ± 0.6	-26.7 ± 0.0	0.5 ± 0.1	-7.0 ± 0.4	96.6 ± 0.1	-34.4 ± 0.4
BID Mutant 3	1.6 ± 0.6	8.9 ± 0.6	143.1 ± 5.4	-31.6 ± 0.9	0.9 ± 0.4	-31.4 ± 0.1	5.7 ± 4.3	-33.0 ± 1.3

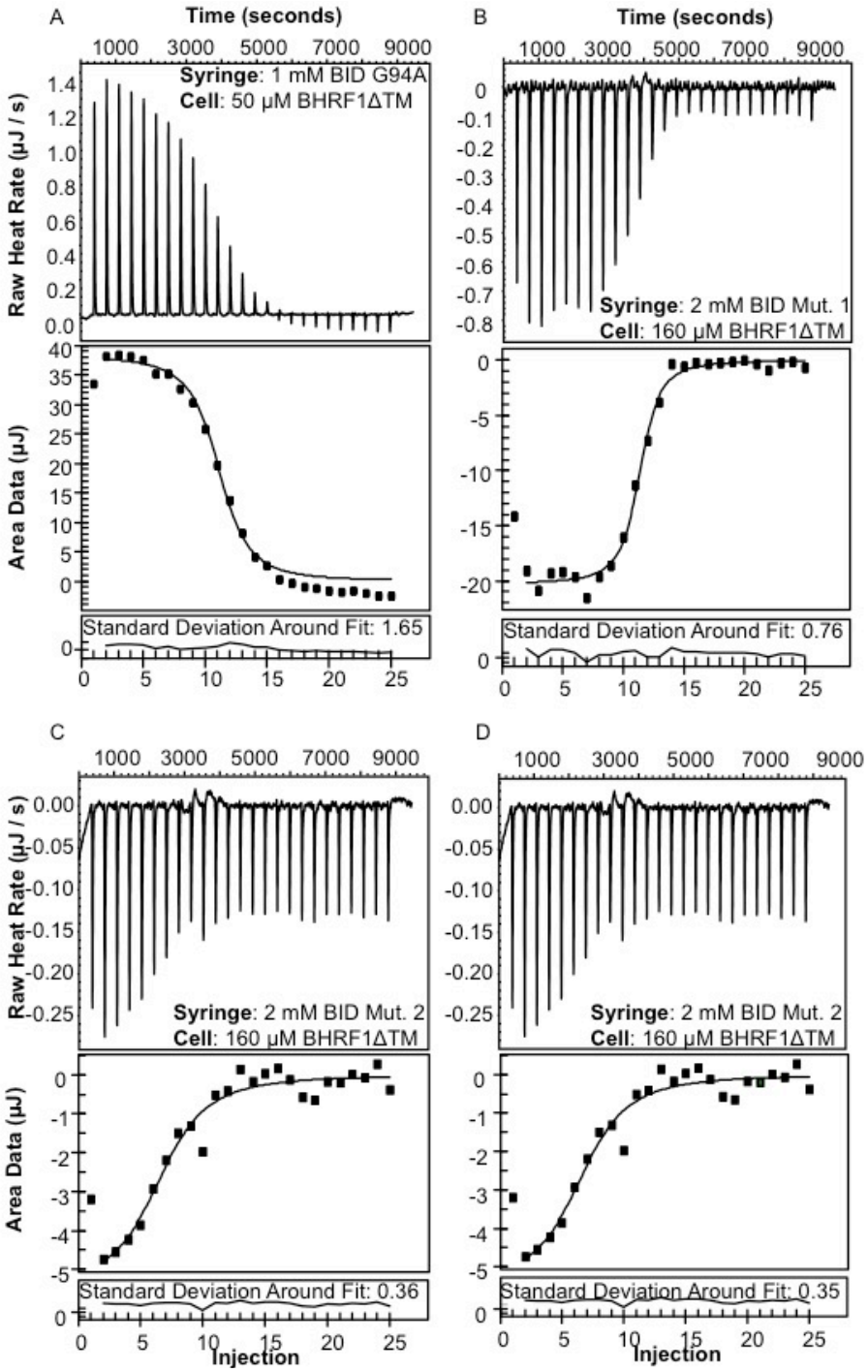

Figure 6.17. ITC profiles of BID BH3D Mutants binding to BHRF1.
A) BID G94A binding to BHRF1ΔTM, B) BID Mutant 1 binding to BHRF1ΔTM, C) BID Mutant 2 binding to BHRF1ΔTM, D) BID Mutant 3 binding to BHRF1ΔTM.

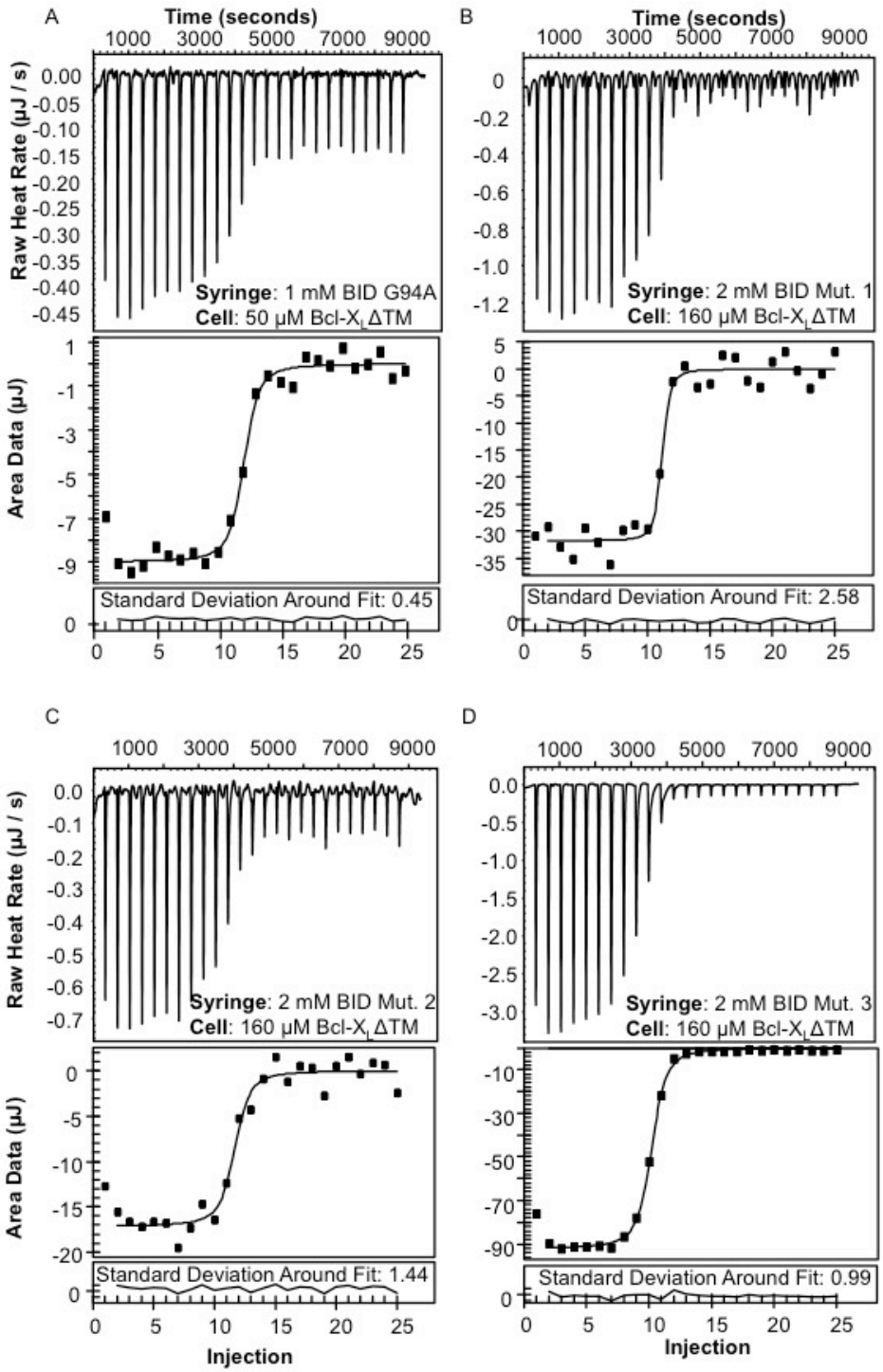

Figure 6.18. ITC profiles of BID BH3D Mutants binding to Bcl-X_L.
A) BID G94A binding to Bcl-X_LΔTM, B) BID Mutant 1 binding to Bcl-X_LΔTM, C) BID Mutant 2 binding to Bcl-X_LΔTM, D) BID Mutant 3 binding to Bcl-X_LΔTM.

6.4. Discussion and conclusions

Based on the cellular assays and biophysical analyses reported here, we conclude that the EBV vBCL2, BHRF1, down-regulates autophagy by non-canonical binding to BECN1. Autophagy assays in MCF7 cells demonstrate that BHRF1 abrogates BECN1-mediated, starvation-induced autophagy, but has no effect on basal levels of autophagy. Unlike BECN1 BH3D interactions with Bcl-2[30], Bcl-X_L[46], and M11[44]; as well as BID[174], BIM[174, 206], BAK[206], and PUMA[174] BH3D interactions with BHRF1; wherein the BH3Ds are sufficient for interaction, the BECN1 BH3D is insufficient to mediate the BHRF1:BECN1 interaction. Yeast two-hybrid assays, affinity pull-downs with purified protein, and binding affinity measurements by ITC all indicate that BECN1 residues 90-171 are required and sufficient for interaction with BHRF1. The region of BECN1 required for interaction with BHRF1 consists of 15 residues of the IDR preceding the BH3D, the BH3D, and the FHD. This suggests that BECN1 binds to BHRF1 via a non-canonical mode of binding.

Analysis of our 2.6 Å crystal structure of the BID BH3D bound to BHRF1 compared the BAK and BIM BH3Ds bound to BHRF1[206] suggests that BIM F69 contributes to the improved binding affinity of BIM for BHRF1 compared to BID, which has M97 at the equivalent BH3D position, and BAK, which has I85 at the equivalent BH3D position. Although BECN1 has F123 at the BH3D position equivalent to BIM F69, comparison of the M11:BECN1 BH3D[44], Bcl-X_L:BECN1 BH3D[46], and BHRF1:BIM complexes[206] indicates that the orientation of BECN1 F123 combined with the inability of BECN1 K117 to mediate a double hydrogen bond or salt bridges with BHRF1 likely destabilize the interaction between the BECN1 BH3D and BHRF1. BECN1 K117 and F123 have previously been shown[140] to contribute to the stability of BECN1 complexes with M11 and Bcl-X_L, wherein a BECN1 K117A mutant weakens binding 10-fold to Bcl-X_L, although it maintains similar binding M11, relative to WT BECN1. The BECN1 F123A mutant has a larger effect, as it weakens binding by 200-fold to Bcl-X_L and by 5-fold to M11 relative to WT BECN1. The weaker interaction between BHRF1 and the BECN1 BH3D relative to other BH3Ds is consistent with our yeast two-hybrid assays, affinity pull-downs with purified protein, and binding affinity measurements by ITC indicating a non-canonical mode of binding between BHRF1 and BECN1 residues 90-171.

BECN1 contains five Anchor regions likely to be energetically stabilized in ordered conformations upon binding[30, 31]. The minimal BHRF1-binding region (residues 90-171) contains one partial Anchor

region (residues 79-103) and three complete Anchor regions (residues 117-127, residues 137-145, and residues 162-169). The BECN1 BH3D Anchor region (residues 117-127) has previously been shown to nucleate concomitant disorder-to-helix transition and binding to Bcl-2[30]. The FHD Anchor regions (residues 137-145 and residues 162-169)[216] and the BECN1 IDR Anchor region (residues 79-103)[41] have been previously shown, in the absence of binding partners through the use of TFE, to be likely to undergo binding-associated disorder-to-helix transitions. In a 4.4 Å crystal structure (PDB 5DFZ) of the full-length yeast VPS34:VPS15:VPS30:VPS38 complex (VPS30 is the yeast homolog of BECN1) it was shown that the VPS30/BECN1 FHD (labeled CC1 in this structure) folds into a helix which is stabilized by binding VPS38 (the yeast homolog of UVRAG) in a coiled-coil manner[17]. Therefore, in addition to the BECN1 BH3D, the BHRF1:BECN1 interaction may involve additional binding-induced helical transitions within the IDR and FHD necessary to stabilize the non-canonical interaction.

Comparison of our 2.6 Å crystal structure of the BID BH3D bound to BHRF1 to the BID BH3D bound to human Bcl-X_L[201], human Mcl-1[199], and mouse A1[200] indicates that each BID BH3D interface residue contributes differentially to the interaction with BHRF1 and each cellular anti-apoptotic BCL2. Based on analyses of the BSA of interface residues and electrostatic interactions that stabilize the complexes, BID R84 is likely more important for binding specificity of the BID BH3D for BHRF1 than for the cellular BCL2 proteins. We also conclude that BID E80, H89, D95, S96, M97, and D98 are likely more important for binding specificity of the BID BH3D for cellular BCL2 proteins than for BHRF1. We have previously shown[140] that BECN1 D121, which is equivalent to BID D95, and BECN1 F123, which is equivalent to BID M97, are more important for the stability of BECN1 BH3D complexes with the cellular BCL2 Bcl-X_L than with the vBCL2 M11. In that study, a BECN1 D121A mutant abrogated binding to Bcl-X_L, but maintained a similar binding affinity to M11, relative to WT BECN1. Likewise, a BECN1 F123A mutant weakened binding relative to WT BECN1 for Bcl-X_L 40-fold more than for M11.

Lastly, we show that a cell-permeable BID BH3D peptide inhibits down-regulation of autophagy by BHRF1. This is consistent with our binding association measurements demonstrating that BHRF1 binds to the BID BH3D (~0.3 μM) approximately 400 times tighter than to BECN1 (~127 μM). Although the binding affinity for BECN1 appears to be sufficient for down-regulation of autophagy, it is substantially weaker than that for the BID BH3D. This suggests that the BID BH3D can displace BECN1 to inhibit

BHRF1-mediated down-regulation of autophagy. These results are significant because they suggest that if peptides that selectively bind to the BH3D-binding groove of BHRF1, but not to cellular BCL2s could be identified, as has been done by us for inhibition of M11 down-regulation of autophagy[140], these BHRF1-selective peptides could inform the rational design of therapeutics to prevent BHRF1 from down-regulating autophagy. Indeed, a selective, BH3D-derived 117-residue protein that binds tightly to BHRF1 has been designed and shown to induce apoptosis in EBV-infected cancer cells, slow tumor progression, and prolong survival in mice with human EBV-positive B cell lymphoma xenograft tumors[219]. However, a role for this inhibitor in autophagy regulation was not assessed.

In summary, previous studies have established that BHRF1 down-regulates apoptosis, but here we show that, like other γHV BCL2 proteins, it also down-regulates autophagy. However, unlike other BCL2s, the BECN1 BH3D is insufficient for binding BHRF1. Instead, a longer flexible BECN1 region that includes, but is not limited to, the BECN1 BH3D is required and sufficient for the BHRF1 interaction. This study also reports the atomic details of the BHRF1:BID BH3D interaction, which adds to the information available for elucidating binding determinants that enforce BHRF1 specificity. Lastly, we demonstrate that despite this non-canonical binding, BHRF1 down-regulation of BECN1-mediated autophagy can be inhibited through the use of a cell-permeable peptide derived from the BID BH3D that tightly binds to BHRF1 in a canonical manner via its BH3D. Understanding the factors that enforce binding specificity may be used for the structure-based design of small molecules that selectively inhibit the BHRF1:BECN1 interaction to prevent BHRF1-mediated down-regulation of autophagy. Such inhibitors would provide a useful tool to examine the role of EBV in regulating autophagy and may ultimately form the basis of novel therapeutics to treat EBV-associated cancers and infections.

CHAPTER 7. CONCLUSION AND FUTURE DIRECTIONS

The work presented in this dissertation aims to identify and understand binding-induced disorder-to-order and conformational transitions in proteins that regulate autophagy. First, we devised a method to identify IDRs likely to undergo helical transitions upon binding, even in the absence of known binding partners. We also investigated conformational flexibility within BECN1, which contains two structured domains sharing an overlapping region that appears to undergo conformational transitions. Lastly, we investigated the interaction between BECN1 and various viral proteins; including the HIV-encoded Nef and γHV vBCL2 homologs encoded by KSHV and EBV. These studies improve our structure-based understanding of the mechanism and biological function of these proteins in regulating autophagy.

In this study, we developed a method to identify IDRs that are likely to undergo helical transitions upon binding by combining bioinformatics analyses followed by CD spectroscopy to monitor TFE-induced changes in the secondary structure content of IDRs. Our results indicate that this method may provide a novel, convenient experimental diagnostic to identify IDRs that become helical upon binding. This tool may help identify potential binding partners for the IDR from a wider list of potential interaction partners for the full-length protein by providing information about the IDR conformation and amino acids most likely to be involved in the interaction. This, in turn, would lead to a better mechanistic understanding of the function of proteins involved in IDR-mediated interactions, which may also be useful for developing therapeutics targeting these proteins.

We also demonstrated that the overlapping helical region of BECN1, which we termed OH, shared by the CCD and BARAD appears to undergo conformational transitions. Without other binding partners, BECN1 self-associates via its CCD to form a homodimer. However, the last 18 CCD residues, composing the OH, have crystallized in two mutually exclusive states: a helix either as part of the CCD[34] or packed against the C-terminal BARAD[36]. We used CD spectroscopy, ITC, and SAXS to show that, in the homodimer, the OH transitions transiently between these two different packing states, with the predominant state comprising the OH packed against the BARAD. We confirmed this observation by comparing the impact of mutating four residues that mediate packing of the OH against both the CCD and BARAD on structure and stability of these domains. We also demonstrated that mutation of these OH-interface residues abrogates starvation-induced up-regulation of autophagy through use of cellular

assays. Previous studies had established the BECN1 CCD and BARAD as structurally well-defined domains, but we showed that these domains include an OH that can transiently adopt two distinct, mutually exclusive packing states, emphasizing overall BECN1 conformational flexibility. Previous studies had also established that the OH forms part of the CCD in BECN1 interactions with heterologous binding partners, but we demonstrated that the OH preferentially packs against the BARAD in the BECN1 homodimer. Our findings have important implications for the relative stability of autophagy-inactive and autophagy-active BECN1 complexes.

As discussed in Chapter 3, packing of the OH against the BARAD in the BECN1 homodimer would likely require reorientation of the BARAD relative to that observed in the yeast PI3KC3:p150:BECN1:UVRAG structure of Complex II. Additional structural studies, through the use of X-ray crystallization, CryoEM, or NMR, of the BECN1 homodimer are required to determine how this rearrangement of the BARAD would reposition the aromatic finger. Furthermore, the BECN1 homodimer and the BECN1-containing Complex II could be used to test binding to liposomes of different sizes, which may help determine if the repositioning of the aromatic finger in the BECN1 homodimer interferes with the ability of BECN1 to interact with membranes.

The conformational flexibility of the OH may also be partially responsible for the difficulty in obtaining an atomic resolution crystal structure of the BECN1 CCD-BARAD since the transitions between packing in the CCD and against the BARAD likely hinders the attainment of high-resolution crystal structures. Since we have identified a preference by the OH for packing against the BARAD, it may be useful to remove the N-terminal CCD residues against which the OH packs in the anti-parallel CCD homodimer[34] for further structural studies of the CCD-BARAD. This BECN1 fragment, consisting of residues 193-450, might minimize the flexibility of the CCD and BARAD and; therefore, improve crystallization and X-ray structure solution outcomes. A crystal structure of the homodimer state of BECN1 would provide the atomic details of the interaction of shortened CCD as well as the location of the re-oriented aromatic finger, which would improve our understanding of function of the likely autophagy-inactive homodimer state of BECN1.

Also discussed in Chapter 3, another important implication of the OH being preferentially packed against the BARAD in the homodimer is that this would release the N-terminal region of the partner helix

within the anti-parallel CCD homodimer, which contains an NES that binds CRM1 to enable nuclear export of BECN1[158, 159]. NES residues Leu[184] and Leu[187] are essential for BECN1 nuclear export, BECN1-mediated autophagy, and tumor suppression[158, 159]. Therefore, these NES residues are likely critical for binding CRM1. However, in the CCD homodimer, NES residues Leu[184] and Leu[187] pack against OH interface residues[18]. Therefore, in order to bind CRM1, the BECN1 NES cannot be packed against the BECN1 OH. Our results showing that the OH preferentially packs against the BARAD in the homodimeric state suggest a mechanism of how the NES may be exposed for CRM1-binding to enable nuclear export. Further studies to delineate the structural basis for the CRM1:BECN1 NES interaction are required to determine if and how CRM1 stabilizes BECN1 conformations wherein the OH is packed against the BARAD.

In the absence of other partners, the OH transitions between these states with a preference to packing against the BARAD, rather than the CCD as has been commonly assumed. The OH likely regulates BECN1 function in autophagy as interactions with different partners likely stabilize one packing state or the other to promote these different functions. It would be interesting to utilize HDX-MS experiments to generate a protein heat map showing the rate of deuterium exchange across the full-length human BECN1 homodimer to provide structural dynamics information regarding the conformational flexibility in the OH and the regions against which the OH appears to transiently pack (i.e. the NES on the CCD N-terminus and the first β-sheet and third α-helix of the BARAD). Although this experiment has been done in yeast VPS30[17], unlike vertebrates, lower eukaryotes lack a NES located at the N-terminus of the CCD. The yeast VPS30 NES comprises residues 12-21[158], which is part of the IDR, rather than the CCD so the conformational switch of the OH between packing against the CCD versus the BARAD is likely not needed for nuclear export in yeast. Therefore, obtaining the structural dynamics information in the full-length human BECN1 would improve our understanding of the importance of the conformational flexibility of the BECN1 OH.

Additionally, we investigated the interaction between BECN1 and the HIV protein Nef via *in vitro* analyses with purified protein. We concluded that, despite previous cellular co-immunoprecipitation assays demonstrating that C-terminal residues of the Nef Core mediate an interaction with BECN1-containing complexes[68] and that BECN1 residues 267-284 of the BARAD are crucial for this interaction[166],

the BECN1 BARAD in the context of the BECN1 homodimer is insufficient to mediate a direct interaction with Nef. As discussed in Chapter 4, it is possible that the preferential packing of the OH against the BARAD in the BECN1 homodimer occludes the Nef binding site. Furthermore, the role of Nef in autophagosome maturation inhibition suggests that Nef interacts with BECN1-containing heterodimers. Therefore, further binding assays and structural studies of the Nef interaction with the BECN1:UVRAG heterodimer or the full PI3KC3:p150:BECN1:UVRAG Complex II are needed to fully understand the mechanism of interaction. This would add to our current understanding of the mechanism by which Nef modulates autophagy in order to enhance HIV replication and pathogenesis. This information may be critical for designing inhibitors of Nef. The crucial role of Nef in disease progression to AIDS means that inhibition of Nef functions could represent a major breakthrough in HIV/AIDS treatment regimens.

In this study, we also attempted to obtain structural information regarding the KSHV vBcl-2 interaction with BH3Ds; however, we faced issues with the purification of the KSHV vBcl-2 due to protein precipitation. Although precipitation issues during protein purification prevented us from obtaining KSHV vBcl-2 at concentrations sufficient to permit biophysical and structural analyses to determine which residues contribute to binding specificity for the KSHV vBcl-2 by BH3Ds, we were able to test the binding affinity for a BECN1 BH3D G120E/D121A double mutant and mNOXA BH3D peptide, as well as assess the oligomeric state of the KSHV vBcl-2:mNOXA complex by SEC. Our experimental evidence showing that the BECN1 BH3D G120E/D121A double mutant does not bind the KSHV vBcl-2 binding suggests that KSHV vBcl-2 employs a distinct mechanism of binding to BH3Ds compared the γHV 68 vBCL2, M11. A unique binding mechanism is consistent with the observation that the binding affinity of KSHV vBcl-2 for other BH3Ds varies greatly compared to those of other anti-apoptotic BCL2 proteins[44, 174].

If a method for expression and purification of sufficient quantities of stable KSHV vBcl-2 could be developed, further structural studies of KSHV vBcl-2 interactions with the BECN1 and mNOXA BH3D would be required to elucidate the determinants of binding specificity for the KSHV vBcl-2 compared to other γHV BCL2 proteins. Although the moderate binding affinity (~13 μM) of the KSHV vBcl-2 for the BECN1 BH3D[44] may be too weak to obtain diffraction quality crystals the KSHV vBcl-2:BECN1 BH3D complex, an NMR solution structure and chemical shifts analysis might enable identification of binding

determinants this interaction. This information will be critical for improving our understanding of how KSHV modulates autophagy.

Our ITC and SEC co-elution results confirm the KSHV vBcl-2:mNOXA BH3D interaction. To date, there is no crystal structure deposited in the PDB of KSHV vBcl-2 bound to any BH3D. Determination of a high-resolution structure via X-ray crystallography should allow elucidation the atomic determinants of the KSHV vBcl-2:mNOXA BH3D interaction. Such information, especially if coupled with co-crystallization and structure determination of the Mcl-1:mNOXA and A1:mNOXA, would improve our understanding of the molecular basis for binding BH3Ds by the KSHV vBcl-2 and Bcl-2 homologs. Understanding these binding determinants may facilitate the rational design of a peptide that selectively inhibits KSHV vBcl-2, but not cellular Bcl-2, mediated inhibition of apoptosis or autophagy. However, development of a method for expression and purification of sufficient quantities of stable KSHV vBcl-2 would be required for these protein X-ray crystallography experiments.

Lastly, we identified a role in autophagy regulation for BHRF1, a vBCL2 mimic encoded by EBV. BHRF1 binds the BH3D of pro-apoptotic Bcl-2 proteins, down-regulating apoptosis and promoting survival of infected cells[174]. We showed that, unlike the KSHV vBcl-2[26] and M11[44], BHRF1 does not bind the isolated BECN1 BH3D. We used yeast two-hybrid assays to identify a minimal region of BECN1 (consisting of 15 IDR residues preceding the BH3D, the BH3D, and the FHD) required to bind BHRF1. We confirmed this to be a direct, albeit weak, interaction via affinity pull-down assays and ITC, and determined that BHRF1 down-regulates autophagy in mammalian cells. We also show that the BID BH3D, a pro-apoptotic Bcl-2 protein, binds ~400-times tighter than the BECN1 BH3D to BHRF1. To further understand the structural basis of BHRF1 specificity, we solved the 2.6 Å crystal structure of the BID BH3D bound to BHRF1. Finally, we show that a cell-permeable peptide derived from the BID BH3D also inhibits BHRF1-mediated down-regulation of autophagy.

Understanding the factors that enforce binding specificity may be used for the structure-based design of small molecules that selectively inhibit the BHRF1:BECN1 interaction to prevent BHRF1-mediated down-regulation of autophagy. The information obtained from our structure of the BHRF1:BID BH3D complex, in combination with further future binding assays and cellular autophagy assays, may be utilized to design peptides that selectively inhibit BHRF1 from down-regulating autophagy. It would also

be interesting to test the previously described 117 residue protein that selectively inhibits BHRF1 from down-regulating apoptosis[219] for an effect on autophagy regulation. Inhibitors of BHRF1-mediated down-regulation of autophagy would provide a useful tool to examine the role of EBV in autophagy regulation and may even ultimately form the basis of novel therapeutics to treat EBV-associated cancers and infections.

REFERENCES

[1] De Duve, C., and Wattiaux, R. (1966) Functions of lysosomes, *Annu Rev Physiol 28*, 435-492.

[2] Lawrence, B. P., and Brown, W. J. (1992) Autophagic vacuoles rapidly fuse with pre-existing lysosomes in cultured hepatocytes, *J Cell Sci 102 (Pt 3)*, 515-526.

[3] Baba, M., Takeshige, K., Baba, N., and Ohsumi, Y. (1994) Ultrastructural analysis of the autophagic process in yeast: detection of autophagosomes and their characterization, *The Journal of Cell Biology 124*, 903-913.

[4] Liang, X. H., Jackson, S., Seaman, M., Brown, K., Kempkes, B., Hibshoosh, H., and Levine, B. (1999) Induction of autophagy and inhibition of tumorigenesis by beclin 1, *Nature 402*, 672-676.

[5] Klionsky, D. J. (2010) The autophagy connection, *Dev Cell 19*, 11-12.

[6] Levine, B., and Klionsky, D. J. (2004) Development by self-digestion: molecular mechanisms and biological functions of autophagy, *Dev Cell 6*, 463-477.

[7] Xie, Z., and Klionsky, D. J. (2007) Autophagosome formation: core machinery and adaptations, *Nat Cell Biol 9*, 1102-1109.

[8] Mizushima, N., Yoshimori, T., and Ohsumi, Y. (2011) The role of Atg proteins in autophagosome formation, *Annu Rev Cell Dev Biol 27*, 107-132.

[9] Meléndez, A., and Levine, B. (2009) Autophagy in C. elegans, *WormBook*, 1-26.

[10] Liang, X. H., Kleeman, L. K., Jiang, H. H., Gordon, G., Goldman, J. E., Berry, G., Herman, B., and Levine, B. (1998) Protection against fatal Sindbis virus encephalitis by Beclin 1, a novel Bcl-2-interacting protein., *Journal of Virology 72*, 8586-8596.

[11] Kihara, A., Noda, T., Ishihara, N., and Ohsumi, Y. (2001) Two distinct Vps34 phosphatidylinositol 3-kinase complexes function in autophagy and carboxypeptidase Y sorting in Saccharomyces cerevisiae, *J Cell Biol 152*, 519-530.

[12] Itakura, E., Kishi, C., Inoue, K., and Mizushima, N. (2008) Beclin 1 Forms Two Distinct Phosphatidylinositol 3-Kinase Complexes with Mammalian Atg14 and UVRAG, *Molecular Biology of the Cell 19*, 5360-5372.

[13] Sun, Q., Fan, W., Chen, K., Ding, X., Chen, S., and Zhong, Q. (2008) Identification of Barkor as a mammalian autophagy-specific factor for Beclin 1 and class III phosphatidylinositol 3-kinase, *Proc Natl Acad Sci U S A 105*, 19211-19216.

[14] Itakura, E., and Mizushima, N. (2009) Atg14 and UVRAG: mutually exclusive subunits of mammalian Beclin 1-PI3K complexes, *Autophagy 5*, 534-536.

[15] 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., and Yoshimori, T. (2009) Two Beclin 1-binding proteins, Atg14L and Rubicon, reciprocally regulate autophagy at different stages, *Nature Cell Biology 11*, 385-U369.

[16] Baskaran, S., Carlson, L., Stjepanovic, G., Young, L., Kim, D., Grob, P., Stanley, R., Nogales, E., and Hurley, J. (2014) Architecture and Dynamics of the Autophagic Phosphatidylinositol 3-Kinase Complex, *Elife 3*.

[17] Rostislavleva, K., Soler, N., Ohashi, Y., Zhang, L., Pardon, E., Burke, J., Masson, G., Johnson, C., Steyaert, J., Ktistakis, N., and Williams, R. (2015) Structure and flexibility of the endosomal Vps34 complex reveals the basis of its function on membranes, *Science 350*, 178-U181.

[18] Mei, Y., Su, M., Sanishvili, R., Chakravarthy, S., Colbert, C. L., and Sinha, S. C. (2016) Identification of BECN1 and ATG14 coiled-coil interface residues important for starvation-induced autophagy, *Biochemistry 55*, 4239−4253.

[19] Shibata, M., Lu, T., Furuya, T., Degterev, A., Mizushima, N., Yoshimori, T., MacDonald, M., Yankner, B., and Yuan, J. (2006) Regulation of intracellular accumulation of mutant Huntingtin by Beclin 1, *J Biol Chem 281*, 14474-14485.

[20] Pickford, F., Masliah, E., Britschgi, M., Lucin, K., Narasimhan, R., Jaeger, P. A., Small, S., Spencer, B., Rockenstein, E., Levine, B., and Wyss-Coray, T. (2008) The autophagy-related protein beclin 1 shows reduced expression in early Alzheimer disease and regulates amyloid beta accumulation in mice, *J Clin Invest 118*, 2190-2199.

[21] Spencer, B., Potkar, R., Trejo, M., Rockenstein, E., Patrick, C., Gindi, R., Adame, A., Wyss-Coray, T., and Masliah, E. (2009) Beclin 1 gene transfer activates autophagy and ameliorates the neurodegenerative pathology in alpha-synuclein models of Parkinson's and Lewy body diseases, *J Neurosci 29*, 13578-13588.

[22] Qu, X., Yu, J., Bhagat, G., Furuya, N., Hibshoosh, H., Troxel, A., Rosen, J., Eskelinen, E. L., Mizushima, N., Ohsumi, Y., Cattoretti, G., and Levine, B. (2003) Promotion of tumorigenesis by heterozygous disruption of the beclin 1 autophagy gene, *J Clin Invest 112*, 1809-1820.
[23] Koneri, K., Goi, T., Hirono, Y., Katayama, K., and Yamaguchi, A. (2007) Beclin 1 gene inhibits tumor growth in colon cancer cell lines, *Anticancer Res 27*, 1453-1457.
[24] Miracco, C., Cosci, E., Oliveri, G., Luzi, P., Pacenti, L., Monciatti, I., Mannucci, S., De Nisi, M. C., Toscano, M., Malagnino, V., Falzarano, S. M., Pirtoli, L., and Tosi, P. (2007) Protein and mRNA expression of autophagy gene Beclin 1 in human brain tumours, *Int J Oncol 30*, 429-436.
[25] Li, Z., Chen, B., Wu, Y., Jin, F., Xia, Y., and Liu, X. (2010) Genetic and epigenetic silencing of the beclin 1 gene in sporadic breast tumors, *BMC Cancer 10*, 98.
[26] Pattingre, S., Tassa, A., Qu, X., Garuti, R., Liang, X., Mizushima, N., Packer, M., Schneider, M., and Levine, B. (2005) Bcl-2 antiapoptotic proteins inhibit Beclin 1-dependent autophagy, *Cell 122*, 927-939.
[27] Orvedahl, A., Alexander, D., Talloczy, Z., Sun, Q. H., Wei, Y. J., Zhang, W., Burns, D., Leib, D. A., and Levine, B. (2007) HSV-1 ICP34.5 confers neurovirulence by targeting the Beclin 1 autophagy protein, *Cell Host & Microbe 1*, 23-35.
[28] Levine, B., and Kroemer, G. (2008) Autophagy in the pathogenesis of disease, *Cell 132*, 27-42.
[29] He, C., and Levine, B. (2010) The Beclin 1 interactome, *Current Opinion in Cell Biology 22*, 140-149.
[30] Mei, Y., Su, M. F., Soni, G., Salem, S., Colbert, C. L., and Sinha, S. C. (2014) Intrinsically disordered regions in autophagy proteins, *Proteins-Structure Function and Bioinformatics 82*, 565-578.
[31] Mei, Y., Glover, K., Su, M., and Sinha, S. C. (2016) Conformational flexibility of BECN1: Essential to its key role in autophagy and beyond, *Protein Sci 25*, 1767-1785.
[32] Lee, E. F., Perugini, M. A., Pettikiriarachchi, A., Evangelista, M., Keizer, D. W., Yao, S., and Fairlie, W. D. (2016) The BECN1 N-terminal domain is intrinsically disordered, *Autophagy 12*, 460-471.
[33] Mei, Y., Ramanathan, A., Glover, K., Christopher , Stanley, C., Sanishvili, R., Chakravarthy, S., Yang, Z., Colbert, C. L., and Sinha, S. C. (2016) Conformational Flexibility Enables Function of a BECN1 Region Essential for Starvation-Mediated Autophagy, *Biochemistry 10.1021/acs.biochem.5b01264*.
[34] Li, X., He, L., Che, K., Funderburk, S., Pan, L., Pan, N., Zhang, M., Yue, Z., and Zhao, Y. (2012) Imperfect interface of Beclin1 coiled-coil domain regulates homodimer and heterodimer formation with Atg14L and UVRAG, *Nature Communications 3*.
[35] Noda, N. N., Kobayashi, T., Adachi, W., Fujioka, Y., Ohsumi, Y., and Inagaki, F. (2012) Structure of the novel C-terminal domain of vacuolar protein sorting 30/autophagy-related protein 6 and its specific role in autophagy, *J Biol Chem 287*, 16256-16266.
[36] Huang, W., Choi, W., Hu, W., Mi, N., Guo, Q., Ma, M., Liu, M., Tian, Y., Lu, P., Wang, F., Deng, H., Liu, L., Gao, N., Yu, L., and Shi, Y. (2012) Crystal structure and biochemical analyses reveal Beclin 1 as a novel membrane binding protein, *Cell Research 22*, 473-489.
[37] Dosztanyi, Z., Meszaros, B., and Simon, I. (2009) ANCHOR: web server for predicting protein binding regions in disordered proteins, *Bioinformatics 25*, 2745-2746.
[38] Mohan, A., Oldfield, C., Radivojac, P., Vacic, V., Cortese, M., Dunker, A., and Uversky, V. (2006) Analysis of molecular recognition features (MoRFs), *Journal Molecular Biology 362*, 1043-1059.
[39] Vacic, V., Oldfield, C., Mohan, A., Radivojac, P., Cortese, M., Uversky, V., and Dunker, A. (2007) Characterization of molecular recognition features, MoRFs, and their binding partners, *Journal of Proteome Research 6*, 2351-2366.
[40] Sinha, S., Colbert, C. L., Becker, N., Wei, Y., and Levine, B. (2008) Molecular basis of the regulation of Beclin 1-dependent autophagy by the γ-herpesvirus 68 Bcl-2 homolog M11, *Autophagy 4*, 989-997.
[41] Glover, K., Mei, Y., and Sinha, S. C. (2016) Identifying intrinsically disordered protein regions likely to undergo binding-induced helical transitions, *Biochim Biophys Acta*.
[42] Mei, Y., Ramanathan, A., Glover, K., Stanley, C., Sanishvili, R., Chakravarthy, S., Yang, Z., Colbert, C. L., and Sinha, S. C. (2016) Conformational flexibility enables function of a BECN1 region essential for starvation-mediated autophagy, *Biochemistry 55*, 1945–1958.
[43] Ku, B., Woo, J., Liang, C., Lee, K., Hong, H., Xiaofei, E., Kim, K., Jung, J., and Oh, B. (2008) Structural and biochemical bases for the inhibition of autophagy and apoptosis by viral BCL-2 of murine gamma-herpesvirus 68, *Plos Pathogens 4*.

[44] Sinha, S., Colbert, C., Becker, N., Wei, Y., and Levine, B. (2008) Molecular basis of the regulation of Beclin 1-dependent autophagy by the gamma-herpesvirus 68 Bcl-2 homolog M11, *Autophagy 4*, 989-997.
[45] Maiuri, M., Le Toumelin, G., Criollo, A., Rain, J., Gautier, F., Juin, P., Tasdemir, E., Pierron, G., Troulinaki, K., Tavernarakis, N., Hickman, J., Geneste, O., and Kroemer, G. (2007) Functional and physical interaction between Bcl-X-L and a BH3-like domain in Beclin-1, *Embo Journal 26*, 2527-2539.
[46] Oberstein, A., Jeffrey, P., and Shi, Y. (2007) Crystal structure of the Bcl-X-L-beclin 1 peptide complex - Beclin 1 is a novel BH3-only protein, *Journal of Biological Chemistry 282*, 13123-13132.
[47] Feng, W., Huang, S., Wu, H., and Zhang, M. (2007) Molecular basis of Bcl-xL's target recognition versatility revealed by the structure of Bcl-xL in complex with the BH3 domain of beclin-1, *Journal of Molecular Biology 372*, 223-235.
[48] Sinha, S., and Levine, B. (2008) The autophagy effector Beclin 1: a novel BH3-only protein, *Oncogene 27 Suppl 1*, S137-148.
[49] Wei, Y., Sinha, S., and Levine, B. (2008) Dual role of JNK1-mediated phosphorylation of Bcl-2 in autophagy and apoptosis regulation, *Autophagy 4*, 949-951.
[50] Maejima, Y., Kyoi, S., Zhai, P., Liu, T., Li, H., Ivessa, A., Sciarretta, S., Del Re, D. P., Zablocki, D. K., Hsu, C. P., Lim, D. S., Isobe, M., and Sadoshima, J. (2013) Mst1 inhibits autophagy by promoting the interaction between Beclin1 and Bcl-2, *Nat Med 19*, 1478-1488.
[51] Gurkar, A., Chu, K., Raj, L., Bouley, R., Lee, S., Kim, Y., Dunn, S., Mandinova, A., and Lee, S. (2013) Identification of ROCK1 kinase as a critical regulator of Beclin1-mediated autophagy during metabolic stress, *Nature Communications 4*.
[52] Noble, C., Dong, J., Manser, E., and Song, H. (2008) Bcl-x(L) and UVRAG cause a monomer-dimer switch in Beclin1, *Journal of Biological Chemistry 283*, 26274-26282.
[53] Liang, C., Feng, P., Ku, B., Dotan, I., Canaani, D., Oh, B., and Jung, J. (2006) Autophagic and tumour suppressor activity of a novel Beclin1-binding protein UVRAG, *Nature Cell Biology 8*, 688-U694.
[54] Axe, E. L., Walker, S. A., Manifava, M., Chandra, P., Roderick, H. L., Habermann, A., Griffiths, G., and Ktistakis, N. T. (2008) Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum, *J Cell Biol 182*, 685-701.
[55] Obara, K., and Ohsumi, Y. (2008) Dynamics and function of PtdIns(3)P in autophagy, *Autophagy 4*, 952-954.
[56] Obara, K., Noda, T., Niimi, K., and Ohsumi, Y. (2008) Transport of phosphatidylinositol 3-phosphate into the vacuole via autophagic membranes in Saccharomyces cerevisiae, *Genes Cells 13*, 537-547.
[57] Liang, C., Lee, J. S., Inn, K. S., Gack, M. U., Li, Q., Roberts, E. A., Vergne, I., Deretic, V., Feng, P., Akazawa, C., and Jung, J. U. (2008) Beclin1-binding UVRAG targets the class C Vps complex to coordinate autophagosome maturation and endocytic trafficking, *Nat Cell Biol 10*, 776-787.
[58] Backer, J. M. (2008) The regulation and function of Class III PI3Ks: novel roles for Vps34, *Biochem J 410*, 1-17.
[59] Miller, S., Tavshanjian, B., Oleksy, A., Perisic, O., Houseman, B. T., Shokat, K. M., and Williams, R. L. (2010) Shaping development of autophagy inhibitors with the structure of the lipid kinase Vps34, *Science 327*, 1638-1642.
[60] Stack, J. H., Herman, P. K., Schu, P. V., and Emr, S. D. (1993) A membrane-associated complex containing the Vps15 protein kinase and the Vps34 PI 3-kinase is essential for protein sorting to the yeast lysosome-like vacuole, *EMBO J 12*, 2195-2204.
[61] Matsunaga, K., Morita, E., Saitoh, T., Akira, S., Ktistakis, N., Izumi, T., Noda, T., and Yoshimori, T. (2010) Autophagy requires endoplasmic reticulum targeting of the PI3-kinase complex via Atg14L, *Journal of Cell Biology 190*, 511-521.
[62] Fan, W., Nassiri, A., and Zhong, Q. (2011) Autophagosome targeting and membrane curvature sensing by Barkor/Atg14(L), *Proceedings of the National Academy of Sciences of the United States of America 108*, 7769-7774.
[63] Cavignac, Y., and Esclatine, A. (2010) Herpesviruses and Autophagy: Catch Me If You Can!, *Viruses-Basel 2*, 314-333.

[64] Kudchodkar, S., and Levine, B. (2009) Viruses and autophagy, *Reviews in Medical Virology 19*, 359-378.
[65] Sumpter, R., and Levine, B. (2011) Selective autophagy and viruses, *Autophagy 7*, 260-265.
[66] Foster, J., and Garcia, J. (2008) HIV-1 Nef: at the crossroads, *Retrovirology 5*.
[67] Hernaez, B., Cabezas, M., Munoz-Moreno, R., Galindo, I., Cuesta-Geijo, M., and Alonso, C. (2013) A179L, a New Viral Bcl2 Homolog Targeting Beclin 1 Autophagy Related Protein, *Current Molecular Medicine 13*, 305-316.
[68] Kyei, G., Dinkins, C., Davis, A., Roberts, E., Singh, S., Dong, C., Wu, L., Kominami, E., Ueno, T., Yamamoto, A., Federico, M., Panganiban, A., Vergne, I., and Deretic, V. (2009) Autophagy pathway intersects with HIV-1 biosynthesis and regulates viral yields in macrophages, *Journal of Cell Biology 186*, 255-268.
[69] Gupta, M. K., Kaminski, R., Mullen, B., Gordon, J., Burdo, T. H., Cheung, J. Y., Feldman, A. M., Madesh, M., and Khalili, K. (2017) HIV-1 Nef-induced cardiotoxicity through dysregulation of autophagy, *Sci Rep 7*, 8572.
[70] Gannage, M., Dormann, D., Albrecht, R., Dengjel, J., Torossi, T., Ramer, P., Lee, M., Strowig, T., Arrey, F., Conenello, G., Pypaert, M., Andersen, J., Garcia-Sastre, A., and Munz, C. (2009) Matrix Protein 2 of Influenza A Virus Blocks Autophagosome Fusion with Lysosomes, *Cell Host & Microbe 6*, 367-380.
[71] Gladue, D. P., O'Donnell, V., Baker-Branstetter, R., Holinka, L. G., Pacheco, J. M., Fernandez-Sainz, I., Lu, Z., Brocchi, E., Baxt, B., Piccone, M. E., Rodriguez, L., and Borca, M. V. (2012) Foot-and-mouth disease virus nonstructural protein 2C interacts with Beclin1, modulating virus replication, *J Virol 86*, 12080-12090.
[72] Chaumorcel, M., Lussignol, M., Mouna, L., Cavignac, Y., Fahie, K., Cotte-Laffitte, J., Geballe, A., Brune, W., Beau, I., Codogno, P., and Esclatine, A. (2012) The Human Cytomegalovirus Protein TRS1 Inhibits Autophagy via Its Interaction with Beclin 1, *Journal of Virology 86*, 2571-2584.
[73] Klionsky, D. J., *et al*. (2016) Guidelines for the use and interpretation of assays for monitoring autophagy (3rd edition), *Autophagy 12*, 1-222.
[74] García-Fuentes, L., Téllez-Sanz, R., Quesada-Soriano, I., and Barón, C. (2011) Thermodynamics of Molecular Recognition by Calorimetry, *Thermodynamics - Physical Chemistry of Aqueous Systems*.
[75] Zhou, X., Kini, R. M., and Sivaraman, J. (2011) Application of isothermal titration calorimetry and column chromatography for identification of biomolecular targets, *Nat Protoc 6*, 158-165.
[76] Núñez, S., Venhorst, J., and Kruse, C. G. (2012) Target-drug interactions: first principles and their application to drug discovery, *Drug Discov Today 17*, 10-22.
[77] Greenfield, N. J. (2006) Using circular dichroism spectra to estimate protein secondary structure, *Nat Protoc 1*, 2876-2890.
[78] Sreerama, N., and Woody, R. (2000) Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set, *Analytical Biochemistry 287*, 252-260.
[79] Greenfield, N. J. (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions, *Nat Protoc 1*, 2527-2535.
[80] Putnam, C., Hammel, M., Hura, G., and Tainer, J. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution, *Quarterly Reviews of Biophysics 40*, 191-285.
[81] Kikhney, A. G., and Svergun, D. I. (2015) A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins, *FEBS Lett 589*, 2570-2577.
[82] Svergun, D., and Koch, M. (2003) Small-angle scattering studies of biological macromolecules in solution, *Reports on Progress in Physics 66*, 1735-1782.
[83] Veronique, R.-B., and Dominique, D. (2012) How Random are Intrinsically Disordered Proteins? A Small Angle Scattering Perspective, *Current Protein & Peptide Science 13*, 55-75.
[84] Rambo, R., and Tainer, J. (2011) Characterizing Flexible and Intrinsically Unstructured Biological Macromolecules by SAS Using the Porod-Debye Law, *Biopolymers 95*, 559-571.
[85] Franke, D., and Svergun, D. I. (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering, *Journal of Applied Crystallography 42*, 342-346.
[86] Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of ab initio shape determination in small-angle scattering, *Journal of Applied Crystallography 36*, 860-864.

[87] Kozin, M. B., and Svergun, D. I. (2001) Automated matching of high- and low-resolution structural models, *Journal of Applied Crystallography 34*, 33-41.
[88] Smyth, M. S., and Martin, J. H. (2000) x ray crystallography, *Mol Pathol 53*, 8-14.
[89] Helliwell, J. R., and Mitchell, E. P. (2015) Synchrotron radiation macromolecular crystallography: science and spin-offs, *IUCrJ 2*, 283-291.
[90] Pflugrath, J. W. (2015) Practical macromolecular cryocrystallography, *Acta Crystallogr F Struct Biol Commun 71*, 622-642.
[91] Sauter, N. K., Grosse-Kunstleve, R. W., and Adams, P. D. (2004) Robust indexing for automatic data collection, *J Appl Crystallogr 37*, 399-409.
[92] Karplus, P. A., and Diederichs, K. (2015) Assessing and maximizing data quality in macromolecular crystallography, *Curr Opin Struct Biol 34*, 60-68.
[93] Glusker, J. P., and Trueblood, K. (2010) *Crystal Structure Analysis: A Primer*, 3 ed., Oxford University Press, New York, NY.
[94] Taylor, G. L. (2010) Introduction to phasing, *Acta Crystallogr D Biol Crystallogr 66*, 325-338.
[95] Evans, P., and McCoy, A. (2008) An introduction to molecular replacement, *Acta Crystallogr D Biol Crystallogr 64*, 1-10.
[96] Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D Biol Crystallogr 66*, 213-221.
[97] Wlodawer, A., Minor, W., Dauter, Z., and Jaskolski, M. (2013) Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination, *FEBS J 280*, 5705-5736.
[98] Brünger, A. T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures, *Nature 355*, 472-475.
[99] Wright, P., and Dyson, H. (1999) Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm, *Journal of Molecular Biology 293*, 321-331.
[100] Oldfield, C., Dunker, A., and Kornberg, R. (2014) Intrinsically Disordered Proteins and Intrinsically Disordered Protein Regions, *Annual Review of Biochemistry, Vol 83 83*, 553-584.
[101] Ishida, T., and Kinoshita, K. (2007) PrDOS: prediction of disordered protein regions from amino acid sequence., *Nucleic Acids Research 35*, W460-W464.
[102] Brown, C. J., Johnson, A. K., Dunker, A. K., and Daughdrill, G. W. (2011) Evolution and disorder, *Curr Opin Struct Biol 21*, 441-446.
[103] Orosz, F., and Ovadi, J. (2011) Proteins without 3D structure: definition, detection and beyond, *Bioinformatics 27*, 1449-1454.
[104] Mei, Y., Su, M., Soni, G., Salem, S., Colbert, C., and Sinha, S. (2014) Intrinsically disordered regions in autophagy proteins., *PROTEINS: Structure, Function and Bioinformatics 82*, 565–578.
[105] Babu, M., van der Lee, R., de Groot, N., and Gsponer, J. (2011) Intrinsically disordered proteins: regulation and disease, *Current Opinion in Structural Biology 21*, 432-440.
[106] van der Lee, R., Lang, B., Kruse, K., Gsponer, J., de Groot, N., Huynen, M., Matouschek, A., Fuxreiter, M., and Babu, M. (2014) Intrinsically Disordered Segments Affect Protein Half-Life in the Cell and during Evolution, *Cell Reports 8*, 1832-1844.
[107] Dunker, A., Cortese, M., Romero, P., Iakoucheva, L., and Uversky, V. (2005) Flexible nets - The roles of intrinsic disorder in protein interaction networks, *Febs Journal 272*, 5129-5148.
[108] Wright, P., and Dyson, H. (2015) Intrinsically disordered proteins in cellular signalling and regulation, *Nat Rev Mol Cell Bio 16*, 18-29.
[109] Vuzman, D., and Levy, Y. (2012) Intrinsically disordered regions as affinity tuners in protein-DNA interactions, *Molecular Biosystems 8*, 47-57.
[110] Huang, K., Chadee, A., Chen, C., Zhang, Y., and Shyu, A. (2013) Phosphorylation at intrinsically disordered regions of PAM2 motif-containing proteins modulates their interactions with PABPC1 and influences mRNA fate, *Rna-a Publication of the Rna Society 19*, 295-305.
[111] Tompa, P. (2011) Unstructural biology coming of age, *Curr Opin Struct Biol 21*, 419-425.
[112] Fong, J. H., Shoemaker, B. A., Garbuzynskiy, S. O., Lobanov, M. Y., Galzitskaya, O. V., and Panchenko, A. R. (2009) Intrinsic disorder in protein interactions: insights from a comprehensive structural analysis, *PLoS Comput Biol 5*, e1000316.

[113] Meszaros, B., Simon, I., and Dosztanyi, Z. (2009) Prediction of protein binding regions in disordered proteins, *PLoS Computational Biology 5*, e1000376.
[114] Peng, Z., and Kurgan, L. (2015) High-throughput prediction of RNA, DNA and protein binding regions mediated by intrinsic disorder, *Nucleic Acids Research*.
[115] Lobanov, M. Y., and Galzitskaya, O. V. (2011) The Ising model for prediction of disordered residues from protein sequence alone, *Phys Biol 8*, 035004.
[116] Xue, B., Dunbrack, R., Williams, R., Dunker, A., and Uversky, V. (2010) PONDR-FIT: A meta-predictor of intrinsically disordered amino acids, *Biochimica Et Biophysica Acta-Proteins and Proteomics 1804*, 996-1010.
[117] Dosztányi, Z., Csizmok, V., Tompa, P., and Simon, I. n. (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content, *Bioinformatics 21*, 3433-3434.
[118] Drozdetskiy, A., Cole, C., Procter, J., and Barton, G. (2015) JPred4: a protein secondary structure prediction server, *Nucleic Acids Research 43*, W389-W394.
[119] Dyson, H., and Wright, P. (1993) Peptide Conformation and Protein-Folding, *Current Opinion in Structural Biology 3*, 60-65.
[120] Shiraki, K., Nishkawa, K., and Goto, Y. (1995) Trifluoroethanol-Induced Stabilization of the Alpha-Helical Structure of Beta-Lactoglobulin - Implication for Non-Hierarchical Protein-Folding, *Journal of Molecular Biology 245*, 180-194.
[121] Buck, M. (1998) Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins, *Quarterly Reviews of Biophysics 31*, 297-355.
[122] Roccatano, D., Colombo, G., Fioroni, M., and Mark, A. (2002) Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: A molecular dynamics study, *Proceedings of the National Academy of Sciences of the United States of America 99*, 12179-12184.
[123] Dunlap, T., Kirk, J., Pena, E., Yoder, M., and Creamer, T. (2013) Thermodynamics of binding by calmodulin correlates with target peptide alpha-helical propensity, *Proteins-Structure Function and Bioinformatics 81*, 607-612.
[124] Sreerama, N., and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set., *Analytical Biochemistry 287*, 252–260.
[125] Sreerama, N., Venyaminov, S., and Woody, R. (2000) Estimation of protein secondary structure from circular dichroism spectra: Inclusion of denatured proteins with native proteins in the analysis, *Analytical Biochemistry 287*, 243-251.
[126] Sreerama, N., Venyaminov, S. Y., and Woody, R. W. (2001) Analysis of protein circular dichroism spectra based on the tertiary structure classification, *Analytical biochemistry 299*, 271-274.
[127] Pogenberg, V., Guichou, J., Vivat-Hannah, V., Kammerer, S., Perez, E., Germain, P., de Lera, A., Gronemeyer, H., Royer, C., and Bourguet, W. (2005) Characterization of the interaction between retinoic acid receptor/retinoid X receptor (RAR/RXR) heterodimers and transcriptional coactivators through structural and fluorescence anisotropy studies, *Journal of Biological Chemistry 280*, 1625-1633.
[128] Yan, N., Wu, J., Chai, J., Li, W., and Shi, Y. (2004) Molecular mechanisms of DrICE inhibition by DIAP1 and removal of inhibition by Reaper, Hid and Grim, *Nature Structural & Molecular Biology 11*, 420-428.
[129] Wu, J., Cocina, A., Chai, J., Hay, B., and Shi, Y. (2001) Structural analysis of a functional DIAP1 fragment bound to grim and hid peptides, *Molecular Cell 8*, 95-104.
[130] Avalos, J., Celic, I., Muhammad, S., Cosgrove, M., Boeke, J., and Wolberger, C. (2002) Structure of a Sir2 enzyme bound to an acetylated p53 peptide, *Molecular Cell 10*, 523-535.
[131] Rustandi, R., Baldisseri, D., and Weber, D. (2000) Structure of the negative regulatory domain of p53 bound to S100B(beta beta), *Nature Structural Biology 7*, 570-574.
[132] Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005) The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins, *Journal of Molecular Biology 347*, 827-839.
[133] Schulenburg, C., and Hilvert, D. (2013) Protein conformational disorder and enzyme catalysis, *Top Curr Chem 337*, 41-67.

[134] Zhang, W., Ganguly, D., and Chen, J. (2012) Residual structures, conformational fluctuations, and electrostatic interactions in the synergistic folding of two intrinsically disordered proteins, *PLoS Comput Biol 8*, e1002353.
[135] Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers 22*, 2577-2637.
[136] Frishman, D., and Argos, P. (1995) Knowledge-based protein secondary structure assignment, *Proteins-Structure Function and Genetics 23*, 566-579.
[137] McCaldon, P., and Argos, P. (1988) Oligopeptide biases in protein sequence and their use in predicting protein coding regions in nucleotide sequences, *Proteins Strcture,Function and Genetics 4*, 99-122.
[138] Wei, Y., An, Z., Zou, Z., Sumpter, R., Su, M., Zang, X., Sinha, S., Gaestel, M., and Levine, B. (2015) The stress-responsive kinases MAPKAPK2/MAPKAPK3 activate starvation-induced autophagy through Beclin 1 phosphorylation, *Elife 4*.
[139] Iconaru, L., Ban, D., Bharatham, K., Ramanathan, A., Zhang, W., Shelat, A., Zuo, J., and Kriwacki, R. (2015) Discovery of Small Molecules that Inhibit the Disordered Protein, p27(Kip1), *Scientific Reports 5*.
[140] Su, M., Mei, Y., Sanishvili, R., Levine, B., Colbert, C. L., and Sinha, S. (2014) Targeting gamma-herpesvirus 68 Bcl-2-mediated down-regulation of autophagy, *J Biol Chem 289*, 8029-8040.
[141] Herrmann, S., Ninkovic, M., Kohl, T., and Pardo, L. (2013) PIST (GOPC) modulates the oncogenic voltage-gated potassium channel K(v)10.1, *Frontiers in Physiology 4*.
[142] Cheng, J., Moyer, B., Milewski, M., Loffing, J., Ikeda, M., Mickle, J., Cutting, G., Li, M., Stanton, B., and Guggino, W. (2002) A golgi-associated PDZ domain protein modulates cystic fibrosis transmembrane regulator plasma membrane expression, *Journal of Biological Chemistry 277*, 3520-3529.
[143] Adi-Harel, S., Erlich, S., Schmukler, E., Cohen-Kedar, S., Segev, O., Mizrachy, L., Hirsch, J., and Pinkas-Kramarski, R. (2010) Beclin 1 Self-Association Is Independent of Autophagy Induction by Amino Acid Deprivation and Rapamycin Treatment, *Journal of Cellular Biochemistry 110*, 1262-1271.
[144] Schrodinger, LLC. (2015) The PyMOL Molecular Graphics System, Version 1.8.
[145] Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state, *J Mol Biol 372*, 774-797.
[146] Sheffield, P., Garrard, S., and Derewenda, Z. (1999) Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors, *Protein Expr Purif 15*, 34-39.
[147] Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D., Konarev, P. V., and Svergun, D. I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis, *J Appl Crystallogr 45*, 342-350.
[148] Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis, *Journal of Applied Crystallography 36*, 1277-1282.
[149] Svergun, D. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria, *Journal of Applied Crystallography 25*, 495-503.
[150] Tria, G., Mertens, H. D., Kachala, M., and Svergun, D. I. (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering, *IUCrJ 2*, 207-217.
[151] Svergun, D., Barberato, C., and Koch, M. H. J. (1995) CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates, *Journal of Applied Crystallography 28*, 768-773.
[152] Kabeya, Y., Mizushima, N., Ueno, T., Yamamoto, A., Kirisako, T., Noda, T., Kominami, E., Ohsumi, Y., and Yoshimori, T. (2000) LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing, *EMBO J 19*, 5720-5728.
[153] Novokhatny, V., and Ingham, K. (1997) Thermodynamics of maltose binding protein unfolding, *Protein Sci 6*, 141-146.
[154] Mertens, H. D., and Svergun, D. I. (2010) Structural characterization of proteins and complexes using small-angle X-ray solution scattering, *J Struct Biol 172*, 128-141.
[155] Jin, T., Chuenchor, W., Jiang, J., Cheng, J., Li, Y., Fang, K., Huang, M., Smith, P., and Xiao, T. S. (2017) Design of an expression system to enhance MBP-mediated crystallization, *Sci Rep 7*, 40991.

[156] Liu, Y., Manna, A., Li, R., Martin, W. E., Murphy, R. C., Cheung, A. L., and Zhang, G. (2001) Crystal structure of the SarR protein from Staphylococcus aureus, *Proc Natl Acad Sci U S A 98*, 6877-6882.
[157] Ranaghan, M. J., Durney, M. A., Mesleh, M. F., McCarren, P. R., Garvie, C. W., Daniels, D. S., Carey, K. L., Skepner, A. P., Levine, B., and Perez, J. R. (2017) The Autophagy-Related Beclin-1 Protein Requires the Coiled-Coil and BARA Domains To Form a Homodimer with Submicromolar Affinity, *Biochemistry 56*, 6639-6651.
[158] Liang, X., Yu, J., Brown, K., and Levine, B. (2001) Beclin 1 contains a leucine-rich nuclear export signal that is required for its autophagy and tumor suppressor function, *Cancer Research 61*, 3443-3449.
[159] Zhang, K., Chen, Y., Zhang, Y., Yao, Q., and Zhuo, H. (2014) Novel CRM1 Inhibitor Prevent the Nuclear Export of Beclin-1 Blocks Autophagy in Mantle Cell Lymphoma, *Blood 124*.
[160] Korber, B., Gaschen, B., Yusim, K., Thakallapally, R., Kesmir, C., and Detours, V. (2001) Evolutionary and immunological implications of contemporary HIV-1 variation, *Br Med Bull 58*, 19-42.
[161] Wain-Hobson, S. (1995) AIDS. Virological mayhem, *Nature 373*, 102.
[162] Geyer, M., Fackler, O. T., and Peterlin, B. M. (2001) Structure--function relationships in HIV-1 Nef, *EMBO Rep 2*, 580-585.
[163] Arold, S. T., and Baur, A. S. (2001) Dynamic Nef and Nef dynamics: how structure could explain the complex activities of this small HIV protein, *Trends Biochem Sci 26*, 356-363.
[164] Pandori, M. W., Fitch, N. J., Craig, H. M., Richman, D. D., Spina, C. A., and Guatelli, J. C. (1996) Producer-cell modification of human immunodeficiency virus type 1: Nef is a virion protein, *J Virol 70*, 4283-4290.
[165] Welker, R., Kottler, H., Kalbitzer, H. R., and Kräusslich, H. G. (1996) Human immunodeficiency virus type 1 Nef protein is incorporated into virus particles and specifically cleaved by the viral proteinase, *Virology 219*, 228-236.
[166] Shoji-Kawata, S., Sumpter, R., Leveno, M., Campbell, G., Zou, Z., Kinch, L., Wilkins, A., Sun, Q., Pallauf, K., MacDuff, D., Huerta, C., Virgin, H., Helms, J., Eerland, R., Tooze, S., Xavier, R., Lenschow, D., Yamamoto, A., King, D., Lichtarge, O., Grishin, N., Spector, S., Kaloyanova, D., and Levine, B. (2013) Identification of a candidate therapeutic autophagy-inducing peptide, *Nature 494*, 201-206.
[167] Alvarado, J. J., Tarafdar, S., Yeh, J. I., and Smithgall, T. E. (2014) Interaction with the Src homology (SH3-SH2) region of the Src-family kinase Hck structures the HIV-1 Nef dimer for kinase activation and effector recruitment, *J Biol Chem 289*, 28539-28553.
[168] Glover, K., Li, Y., Mukhopadhyay, S., Leuthner, Z., Chakravarthy, S., Colbert, C. L., and Sinha, S. C. (2017) Structural transitions in conserved, ordered Beclin 1 domains essential to regulating autophagy, *J Biol Chem 292*, 16235-16248.
[169] Boshoff, C., and Chang, Y. (2001) Kaposi's sarcoma-associated herpesvirus: a new DNA tumor virus, *Annu Rev Med 52*, 453-470.
[170] Sarid, R., Sato, T., Bohenzky, R. A., Russo, J. J., and Chang, Y. (1997) Kaposi's sarcoma-associated herpesvirus encodes a functional bcl-2 homologue, *Nat Med 3*, 293-298.
[171] Altmann, M., and Hammerschmidt, W. (2005) Epstein-Barr virus provides a new paradigm: a requirement for the immediate inhibition of apoptosis, *PLoS Biol 3*, e404.
[172] Orvedahl, A., Alexander, D., Talloczy, Z., Sun, Q., Wei, Y., Zhang, W., Burns, D., Leib, D., and Levine, B. (2007) HSV-1ICP34.5 confers neurovirulence by targeting the Beclin 1 autophagy protein, *Cell Host & Microbe 1*, 23-35.
[173] Huang, Q., Petros, A. M., Virgin, H. W., Fesik, S. W., and Olejniczak, E. T. (2002) Solution structure of a Bcl-2 homolog from Kaposi sarcoma virus, *Proc Natl Acad Sci U S A 99*, 3428-3433.
[174] Flanagan, A. M., and Letai, A. (2008) BH3 domains define selective inhibitory interactions with BHRF-1 and KSHV BCL-2, *Cell Death Differ 15*, 580-588.
[175] Chen, L., Willis, S. N., Wei, A., Smith, B. J., Fletcher, J. I., Hinds, M. G., Colman, P. M., Day, C. L., Adams, J. M., and Huang, D. C. (2005) Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function, *Mol Cell 17*, 393-403.
[176] Rickinson, A. B., Long, H. M., Palendira, U., Münz, C., and Hislop, A. D. (2014) Cellular immune controls over Epstein-Barr virus infection: new lessons from the clinic and the laboratory, *Trends Immunol 35*, 159-169.

[177] Epstein, M. A., Achong, B. G., and Barr, Y. M. (1964) Virus particles in cultured lymphoblasts from Burkitt's lymphoma, *Lancet 1*, 702-703.
[178] Kenney, S. (2006) Theodore E. Woodward Award: development of novel, EBV-targeted therapies for EBV-positive tumors, *Trans Am Clin Climatol Assoc 117*, 55-73; discussion 73-54.
[179] Tsao, S. W., Tsang, C. M., To, K. F., and Lo, K. W. (2015) The role of Epstein-Barr virus in epithelial malignancies, *J Pathol 235*, 323-333.
[180] Gangappa, S., van Dyk, L. F., Jewett, T. J., Speck, S. H., and Virgin, H. W. (2002) Identification of the in vivo role of a viral bcl-2, *J Exp Med 195*, 931-940.
[181] Cuconati, A., and White, E. (2002) Viral homologs of BCL-2: role of apoptosis in the regulation of virus infection, *Genes Dev 16*, 2465-2478.
[182] Vaux, D. L., Cory, S., and Adams, J. M. (1988) Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells, *Nature 335*, 440-442.
[183] McDonnell, T. J., Deane, N., Platt, F. M., Nunez, G., Jaeger, U., McKearn, J. P., and Korsmeyer, S. J. (1989) Bcl-2-immunoglobulin transgenic mice demonstrate extended B cell survival and follicular lymphoproliferation, *Cell 57*, 79-88.
[184] Reed, J. C. (1998) Bcl-2 family proteins, *Oncogene 17*, 3225-3236.
[185] Danial, N. N., Gimenez-Cassina, A., and Tondera, D. (2010) Homeostatic functions of BCL-2 proteins beyond apoptosis, *Adv Exp Med Biol 687*, 1-32.
[186] Han, J., Flemington, C., Houghton, A. B., Gu, Z., Zambetti, G. P., Lutz, R. J., Zhu, L., and Chittenden, T. (2001) Expression of bbc3, a pro-apoptotic BH3-only gene, is regulated by diverse cell death and survival signals, *Proc Natl Acad Sci U S A 98*, 11318-11323.
[187] Liu, M. Y., Shih, Y. Y., Li, L. Y., Chou, S. P., Sheen, T. S., Chen, C. L., Yang, C. S., and Chen, J. Y. (2000) Expression of the Epstein-Barr virus BHRF1 gene, a homologue of Bcl-2, in nasopharyngeal carcinoma tissue, *J Med Virol 61*, 241-250.
[188] Nicholls, J., Kremmer, E., Meseda, C. A., Mackett, M., Hahn, P., Gulley, M. L., Brink, A., Swinnen, L. J., Greenspan, J., De Souza, Y., Grässer, F., Sham, J., Ng, M. H., and Arrand, J. R. (2001) Comparative analysis of the expression of the Epstein-Barr virus (EBV) anti-apoptotic gene BHRF1 in nasopharyngeal carcinoma and EBV-related lymphoid diseases, *J Med Virol 65*, 105-113.
[189] Dawson, C. W., Dawson, J., Jones, R., Ward, K., and Young, L. S. (1998) Functional differences between BHRF1, the Epstein-Barr virus-encoded Bcl-2 homologue, and Bcl-2 in human epithelial cells, *J Virol 72*, 9016-9024.
[190] Kelly, G. L., Long, H. M., Stylianou, J., Thomas, W. A., Leese, A., Bell, A. I., Bornkamm, G. W., Mautner, J., Rickinson, A. B., and Rowe, M. (2009) An Epstein-Barr virus anti-apoptotic protein constitutively expressed in transformed cells and implicated in burkitt lymphomagenesis: the Wp/BHRF1 link, *PLoS Pathog 5*, e1000341.
[191] Watanabe, A., Maruo, S., Ito, T., Ito, M., Katsumura, K. R., and Takada, K. (2010) Epstein-Barr virus-encoded Bcl-2 homologue functions as a survival factor in Wp-restricted Burkitt lymphoma cell line P3HR-1, *J Virol 84*, 2893-2901.
[192] Leao, M., Anderton, E., Wade, M., Meekings, K., and Allday, M. J. (2007) Epstein-barr virus-induced resistance to drugs that activate the mitotic spindle assembly checkpoint in Burkitt's lymphoma cells, *J Virol 81*, 248-260.
[193] Hickish, T., Robertson, D., Clarke, P., Hill, M., di Stefano, F., Clarke, C., and Cunningham, D. (1994) Ultrastructural localization of BHRF1: an Epstein-Barr virus gene product which has homology with bcl-2, *Cancer Res 54*, 2808-2811.
[194] Fu, Q., He, C., and Mao, Z. R. (2013) Epstein-Barr virus interactions with the Bcl-2 protein family and apoptosis in human tumor cells, *J Zhejiang Univ Sci B 14*, 8-24.
[195] Pearson, G. R., Luka, J., Petti, L., Sample, J., Birkenbach, M., Braun, D., and Kieff, E. (1987) Identification of an Epstein-Barr virus early gene encoding a second component of the restricted early antigen complex, *Virology 160*, 151-161.
[196] Henderson, S., Huen, D., Rowe, M., Dawson, C., Johnson, G., and Rickinson, A. (1993) Epstein-Barr virus-coded BHRF1 protein, a viral homologue of Bcl-2, protects human B cells from programmed cell death, *Proc Natl Acad Sci U S A 90*, 8479-8483.
[197] Muchmore, S. W., Sattler, M., Liang, H., Meadows, R. P., Harlan, J. E., Yoon, H. S., Nettesheim, D., Chang, B. S., Thompson, C. B., Wong, S. L., Ng, S. L., and Fesik, S. W. (1996) X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death, *Nature 381*, 335-341.

[198] Loh, J., Huang, Q., Petros, A. M., Nettesheim, D., van Dyk, L. F., Labrada, L., Speck, S. H., Levine, B., Olejniczak, E. T., and Virgin, H. W. (2005) A surface groove essential for viral Bcl-2 function during chronic infection in vivo, *PLoS Pathog 1*, e10.
[199] Miles, J. A., Yeo, D. J., Rowell, P., Rodriguez-Marin, S., Pask, C. M., Warriner, S. L., Edwards, T. A., and Wilson, A. J. (2016) Hydrocarbon constrained peptides - understanding preorganisation and binding affinity, *Chem Sci 7*, 3694-3702.
[200] Smits, C., Czabotar, P., Hinds, M., and Day, C. (2008) Structural plasticity underpins promiscuous binding of the prosurvival protein a1, *Structure 16*, 818-829.
[201] Rajan, S., Choi, M., Baek, K., and Yoon, H. S. (2015) Bh3 induced conformational changes in Bcl-Xl revealed by crystal structure and comparative analysis, *Proteins 83*, 1262-1272.
[202] Liu, X., Dai, S., Zhu, Y., Marrack, P., and Kappler, J. (2003) The structure of a Bcl-x(L)/Bim fragment complex: implications for bim function, *Immunity 19*, 341-352.
[203] Petros, A. M., Nettesheim, D. G., Wang, Y., Olejniczak, E. T., Meadows, R. P., Mack, J., Swift, K., Matayoshi, E. D., Zhang, H., Thompson, C. B., and Fesik, S. W. (2000) Rationale for Bcl-xL/Bad peptide complex formation from structure, mutagenesis, and biophysical studies, *Protein Sci 9*, 2528-2534.
[204] Sattler, M., Liang, H., Nettesheim, D., Meadows, R. P., Harlan, J. E., Eberstadt, M., Yoon, H. S., Shuker, S. B., Chang, B. S., Minn, A. J., Thompson, C. B., and Fesik, S. W. (1997) Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis, *Science 275*, 983-986.
[205] Huang, Q., Petros, A. M., Virgin, H. W., Fesik, S. W., and Olejniczak, E. T. (2003) Solution structure of the BHRF1 protein from Epstein-Barr virus, a homolog of human Bcl-2, *J Mol Biol 332*, 1123-1130.
[206] Kvansakul, M., Wei, A. H., Fletcher, J. I., Willis, S. N., Chen, L., Roberts, A. W., Huang, D. C., and Colman, P. M. (2010) Structural basis for apoptosis inhibition by Epstein-Barr virus BHRF1, *PLoS Pathog 6*, e1001236.
[207] Nowag, H., Guhl, B., Thriene, K., Romao, S., Ziegler, U., Dengjel, J., and Münz, C. (2014) Macroautophagy Proteins Assist Epstein Barr Virus Production and Get Incorporated Into the Virus Particles, *EBioMedicine 1*, 116-125.
[208] Granato, M., Santarelli, R., Farina, A., Gonnella, R., Lotti, L. V., Faggioni, A., and Cirone, M. (2014) Epstein-barr virus blocks the autophagic flux and appropriates the autophagic machinery to enhance viral replication, *J Virol 88*, 12715-12726.
[209] Kabeya, Y., Mizushima, N., Ueno, T., Yamamoto, A., Kirisako, T., Noda, T., Kominami, E., Ohsumi, Y., and Yoshimori, T. (2000) LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing., *The EMBO Journal 19*, 5720-5728.
[210] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J Appl Crystallogr 40*, 658-674.
[211] Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallogr D Biol Crystallogr 66*, 486-501.
[212] Theobald, D. L., and Wuttke, D. S. (2006) THESEUS: maximum likelihood superpositioning and analysis of macromolecular structures, *Bioinformatics 22*, 2171-2172.
[213] Gaillard, T., Stote, R. H., and Dejaegere, A. (2016) PSSweb: protein structural statistics web server, *Nucleic Acids Res 44*, W401-405.
[214] Liang, X. H., Jackson, S., Seaman, M., Brown, K., Kempkes, B., Hibshoosh, H., and Levine, B. (1999) Induction of autophagy and inhibition of tumorigenesis by Beclin 1, *Nature 402*, 672-676.
[215] Mei, Y., Su, M., Sanishvili, R., Chakravarthy, S., Colbert, C. L., and Sinha, S. C. (2016) Identification of BECN1 and ATG14 Coiled-Coil Interface Residues That Are Important for Starvation-Induced Autophagy, *Biochemistry 55*, 4239-4253.
[216] Mei, Y., Ramanathan, A., Glover, K., Stanley, C., Sanishvili, R., Chakravarthy, S., Yang, Z., Colbert, C. L., and Sinha, S. C. (2016) Conformational Flexibility Enables the Function of a BECN1 Region Essential for Starvation-Mediated Autophagy, *Biochemistry*.
[217] Lomonosova, E., and Chinnadurai, G. (2008) BH3-only proteins in apoptosis and beyond: an overview, *Oncogene 27 Suppl 1*, S2-19.
[218] Czabotar, P. E., Lee, E. F., van Delft, M. F., Day, C. L., Smith, B. J., Huang, D. C., Fairlie, W. D., Hinds, M. G., and Colman, P. M. (2007) Structural insights into the degradation of Mcl-1 induced by BH3 domains, *Proc Natl Acad Sci U S A 104*, 6217-6222.

[219] Procko, E., Berguig, G., Shen, B., Song, Y., Frayo, S., Convertine, A., Margineantu, D., Booth, G., Correia, B., Cheng, Y., Schief, W., Hockenbery, D., Press, O., Stoddard, B., Stayton, P., and Baker, D. (2014) A Computationally Designed Inhibitor of an Epstein-Barr Viral Bcl-2 Protein Induces Apoptosis in Infected Cells, *Cell 157*, 1644-1656.