Structural Investigation of mTORC1 Regulation at the Lysosomal Membrane

By

Rachel Marie Jansen

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular & Cellular Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor James H Hurley, Chair
Professor Andreas Martin
Professor Roberto Zoncu
Assistant Professor Eunyong Park

Spring 2024

Abstract

Structural Investigation of mTORC1 Regulation at the Lysosomal Membrane

by

Rachel Marie Jansen

Doctor of Philosophy in Molecular & Cellular Biology

University of California, Berkeley

Professor James H Hurley, Chair

The mechanistic target of rapamycin complex 1 (mTORC1) regulates cell growth and catabolism by phosphorylating key substrates in response to nutrient fluctuations. Recruitment of mTORC1 to the lysosomal membrane by the Rag-Ragulator complex is important for activation of its kinase activity and interaction with select substrates. The Rag-Ragulator complex is comprised of two small GTPases, Rag A/B and Rag C/D, and the pentameric scaffold protein complex Ragulator. The Rags GTPases exist predominately in an active ($RagA^{GTP}$:$RagC^{GDP}$) or inactive ($RagA^{GDP}$:$RagC^{GTP}$) state, with mTORC1 binding to the active state. The Rag GTPase's nucleotide state and activity are modulated by two GAPs, FLCN and GATOR1, transitioning the complex between active and inactive forms. Additional protein complexes (KICSTOR, GATOR2) and nutrient sensors (SLC38A9, Castor1, Sestrin2 and Samtor) regulate the activity of these GAPs in response to nutrient levels ultimately regulating the activity of mTORC1. In this work, structural analysis and in vitro reconstitution were used to elucidate the mechanisms of FLCN and GATOR1 regulation. The network of GATOR1 regulation was dissected to understand how the KICSTOR complex regulates GATOR1 and how nutrient sensors Castor1 and Sestrin2 relay information to the GATOR2 complex.

Table of Contents

Chapter 1: Introduction *1*

Mechanistic Target of Rapamycin Complex 1 and Lysosomal Regulation **2**

The Rag-Ragulator Complex **2**

Regulation of the GTPase Activating Protein FLCN **3**

Regulation of the GTPase Activating Protein GATOR1 ..... **4**

Nutrient Sensors and the GATOR complexes **4**

Focus of this Work **5**

Figures: **7**

Chapter 2: Structural Basis of FLCN RagC GAP Activation *15*

Introduction **16**

Results **17**

Discussion **19**

Materials and Methods **20**

Figures **25**

Chapter 3: Structural Basis for GATOR2 inhibition by Sestrin2 and Castor1 ..... *45*

Introduction **45**

Small Scale Expression Tests for Gator2 ..... **46**

Large Scale Purification of Gator2 ..... **48**

Cryo-EM of GATOR2 ..... **48**

Low Resolution Initial Cryo-EM for Castor1 + GATOR2 Complex **49**

Arctica Data Set 49

Experimentally Guided Alphafold Prediction for Castor1 + GATOR2 Complex **51**

Interface #1: Castor1-MIOS 51

Interface #2: Castor1-WDR59 52

Sample Optimization for GATOR2-Castor1 Complex **53**

High Resolution Data Collection for GATOR2-Castor1 Complex **54**

Reconstituting the GATOR1 and GATOR2 System In Vitro **54**

Design of Promega GTPase Glo Assay 55

Design of HPLC Based Assay 55

Chapter 4: Regulation of GATOR1 and GATOR2 by Membrane Interacting Protein Complex KICSTOR *82*

Introduction **82**

Designing DNA Constructs for KICSTOR **83**

Small Scale Expression Tests for KICSTOR **83**

Large Scale Purification of KICSTOR **84**

Method #1: Purification of Full-Length KICSTOR using GST-SZT2 84

Method #2: Purification of Full-Length KICSTOR using STEP-SZT2 84

Method #3: Purification of GST-KICSTOR using STREP-KPTN, STREP-ITFG2 and STREP-KICS2 85

Purification of ITFG2-KPTN Dimer 85

Single Particle Cryo-EM for Full-Length KICSTOR Complex **86**

Phase Plate Dataset for Full-Length STREP-KICSTOR 87

Structural Understanding of KICSTOR with AlphaFold2 ..... **87**

Membrane Binding Ability of KICSTOR **88**

Interaction of KICSTOR with GATOR Components **89**

Membrane Recruitment Assay for KICSTOR and GATOR Components **90**

Cryo-EM for KICSTOR-GATOR1-GATOR2 Complex **91**

KICSTOR-GATOR1-GATOR2 Complex Formation 91

Data processing for KICSTOR-GATOR1-GATOR2 Complex 91

List of Figures:

Figure 1.1 Activity of the Rag-Ragulator complex determined by nutrient availability. 7

Figure 1.2. Schematic representation of the regulation of Rag-Ragulator in different nutrient conditions. 8

Figure 1.3 Dual binding modes of FLCN and GATOR1. 9

Figure 1.4 GATOR1 interaction with Rag-Ragulator in active mode and GAP mechanism. 10

Figure 1.5 Regulation of GATOR1 and the corresponding structures. 11

Figure 2.1 Active FLCN complex formation. 25

Figure 2.2 Purification of individual active FLCN components. 26

Figure 2.3 Cryo-EM data processing pipeline. 27

Figure 2.4 AFC local resolution estimation. 29

Figure 2.5 AFC map-model fit. 30

Figure 2.6 Cryo-EM structure of active FLCN complex. 31

Figure 2.8 Structural comparison between RagC:FLCN and other small GTPases:GAP. 33

Figure 2.9 Structural investigation of RagC:SMCR8 and Arf1:FLCN interface. 34

Figure 2.10 Hydrophobic residues in the FLCN:RagC interface necessary for GAP activity..... 36

Figure 2.11 GAP activity of FLCN:SLC-FNIP2. 37

Figure 2.12 Immunofluorescence for FLCNR164. 38

Figure 2.13 Immunoblot analysis for mTORC1 targets. 39

Figure 3.1 Schematic of Nutrient Dependent GATOR2 regulation of GATOR1. 57

Figure 3.2 Constructs designed for GATOR2 subunits. 58

Figure 3.3 Initial small scale expression tests for GATOR2 with STREP and MIOS tags. 59

Figure 3.4 Initial small scale expression tests for GATOR2 with FLAG-MIOS..... 60

Figure 3.5 Initial small scale expression tests for GATOR2 using FLAG-tagged subunits..... 61

Figure 3.6 Large scale purification of STREP-GATOR2. 62

Figure 3.7 Cryo-EM analysis of GATOR2 sample. 63

Figure 3.8 Purification of FLAG-Sestrin2 and FLAG-Castor1 from HEK 293 GnTI. (..... 64

Figure 3.9 Castor1 Purification from E.Coli. 65

Figure 3.10 Sestrin2 Purification from E. coli. 66

Figure 3.11 GATOR2-Castor1 complex formation. 67

Figure 3.12 Data Processing for GATOR2-Castor1 initial map. 68

Figure 3.13 Improved resolution of GATOR2-CASTOR1 complex using local refinement. 69
Figure 3.14 Castor1-Mios Interface Analysis. 70
Figure 3.15 GATOR2-Castor1 Release Mechanism for Castor1-MIOS interface. 71
Figure 3.16 Castor1-WDR59 Interface Analysis. 72
Figure 3.17 GATOR2- Castor1 Release Mechanism at Castor1-WDR59 interface. 73
Figure 3.18. Data processing for GATOR2-Castor1 high resolution map. 74
Figure 3.19 Overview of Promega GTPase Glo Assay. 75
Figure 3.20 GTPase Glo GATOR1 GAP activity towards Rag GTPases. 76
Figure 3.21 GTPase Glo GATOR1 GAP activity towards Rag GTPases is altered by GATOR2.... 77
Figure 3.22 Castor1 restores GATOR1 GAP activity in presence of GATOR2 in GTPase Glo Assay. 78
Figure 3.23 HPLC Based GAP Activity Assay for GATOR1. 79
Figure 4.1 Schematic of GATOR1 regulation. 93
Figure 4.2 Constructs designed for KICSTOR subunits 94
Figure 4.3 Small scale expression tests for KICSTOR complex. 95
Figure 4.4 Mass spectrometry analysis of SDS-PAGE bands correlating with KICSTOR components. 96
Figure 4.5 Representative large-scale purification of KICSTOR complex using GST-SZT2 (method #1 – no SEC). 97
Figure 4.6 Representative large-scale purification of KICSTOR complex using STREP-SZT2. 98
Figure 4.7 Representative large-scale purification of GST-KICSTOR complex using STREP tagged smaller subunits. 99
Figure 4.9. Comparison of KICSTOR cryo-EM datasets with Alphafold. 101
Figure 4.10 Data collection and processing workflow for STREP-KICSTOR. 102
Figure 4.11 Phase plate data collection for KICSTOR. 103
Figure 4.12 Alphafold2 generated model of full-length KICSTOR complex. 104
Figure 4.13 KICSTOR complex flexibility seen in Alphfold2 model. 105
Figure 4.14 Proposed membrane binding site in KICS2 subunit of KICSTOR complex. 106
Figure 4.15 GUV assay for WT KICSTOR-GFP membrane binding. 107
Figure 4.16 GUV assay for KICSTOR-GFP mutant 4D and mutant 10D membrane binding. ... 108
Figure 4.17 Pulldown assay with GST-KICSTOR and GATOR complexes. 109
Figure 4.18 Purification of GATOR1-mCherry..... 110

Figure 4.19 GUV assay for GATOR1-mCherry membrane binding. 111
Figure 4.20 Purification of GATOR2-GFP. 112
Figure 4.21 GUV assay for GATOR2-GFP membrane binding. 113
Figure 4.22 Attempt #1 - large scale complex formation for KICSTOR, GATOR1 and GATOR2. 114
Figure 4.23 Attempt #2 - large scale complex formation for KICSTOR, GATOR1 and GATOR2. 115
Figure 4.24 Cryo-EM for KICSTOR and GATOR components. 116

Acknowledgements

The completion of my graduate studies and thesis stands as a testament to the unwavering support and guidance I've received from my family, friends, lab mates, and mentors. Their collective support has been instrumental in shaping my academic journey, and I am deeply grateful to each and every one of them.

First, I would like to thank my family. To my parents, your support throughout my graduate studies was the foundation for my sucess. From nurturing my scientific curiosity with frequent trips to science museums during my childhood to introducing me to research in high school, your guidance has been invaluable. Despite the distance, you've always been there for me with a kind message or FaceTime call for encouragement. To my younger sisters, Laura and Sarah, thank you for keeping my spirits high. Your FaceTime calls and visits to Berkeley were cherished moments of joy when research was tough. And to my grandma, I've loved every one of our phone calls. You've taught me to work hard and push myself to be the best I can be. I'm thankful for the time we had during your visit and all our fun adventures in Berkeley.

To my incredible partner, Dillon McCutcheon, thank you for your encouragement and support over the years. Your consistent calls after my long lab days, your willingness to adjust schedules for my lab work, and your thoughtful gestures like bringing me dinner during late nights in the lab mean more to me than words can express. You always know how to lift my spirits, often with food, and your presence on the phone during weekend lab sessions has been a huge sense of comfort. Your ability to make me smile after a challenging day is truly a gift, and I'm grateful to have you by my side.

To my wonderful college friends, Lilly Voke, Becca Leduc, and Sashrika Saini, thank you for being my source of love, support, and encouragement. Lilly, your belief in me during both my highs and lows means the world. Your intuitive understanding and hilarious personality make every moment with you special. Becca, your visits to Berkeley brought a piece of home to me, and I cherished every adventure we had together. Your constant check-ins, enthusiasm, and morning FaceTime workouts kept me motivated and inspired. And to Sashrika, thank you for showing me research is an enjoyable journey. Your support during my move from MA to CA was a lifeline, and our adventures in Napa will always be treasured memories.

To my incredible roommate and quarantine pal, Alena Bishop, thank you for your friendship since day one. I'll always remember how you came to the airport to pick me up with my two pieces of luggage, then took me straight to IKEA without even having met me yet. You stayed with me until my whole bed was put together, and that's the moment I knew we would be the best of friends. I feel incredibly fortunate that we decided to room together because I couldn't have imagined quarantining with anyone else. From being study buddies during our first year to teaching me how to cook tofu deliciously, from rushing me to the ER when I sliced open my finger to being a listening ear after a tough day, you've been there for it all. Our adventures to Yosemite and Lake Tahoe, your encouragement to go out when I wanted to stay in, and maybe not so accidentally wearing the same outfits—every moment has been memorable because of you.

To my first friend at UC Berkeley, Adam Begeman, thank you for being there from the very beginning. Our connection during the interview day housing tour was a stroke of luck, and I'm grateful for the bond we formed before even starting the program. Thank you for taking me out for my first legal drink when we arrived in Berkeley. Our pre-290 lunches, pickleball matches, and long chats over cocktails have been highlights of my time here. But beyond that, I am so thankful for introducing me to your incredible fiancé, Annie Maloney. Annie, your friendship has been a wonderful gift. From our walks and wine nights to gardening days and shopping sprees, every moment with you has been cherished. Thank you for being my sounding board, guiding me through both personal and professional challenges, and helping me become the most confident version of myself.

To my structural biology buddy, Trinity Cookis, thank you for always being eager to geek out about structures with me. Our hangouts, girls' nights, hikes, and boba outings have been highlights of my graduate experience. Your friendship has added so much joy and laughter to my journey. To all my amazing friends from the Best Berkeley Crew and MCB Softball Team, I want to give a huge thanks for your friendship and community. All of you have made my Berkeley experience truly unforgettable. From our nights out to the championships we've won and the summer lake trips we've shared, every moment with you made all the tough challenges of graduate school worthwhile. To all my other friends in my cohort, thank you. Despite our non-traditional start to the Berkeley experience, we made the most of it.

Thank you to the former Hurley Lab postdocs, Simon Fromm and Adam Yokom, for their pivotal roles in my cryo-EM training. Simon, your guidance during my rotation and your introduction to cryo-EM were essential in shaping my skills in this area. Adam, thank you for your support and assistance with cryo-EM, despite us working on different projects. Your willingness to help me collect, process, and analyze data was immensely appreciated.

Thank you to my other former lab mates, King Cada, Lisa Strong, Liv Jensen, and Emma Garst, who made the Hurley Lab such a positive and enjoyable working environment. King, your sense of humor never failed to brighten up the break room, and your jokes always brought a smile to my face. Lisa, your infectious positive energy constantly lifted my spirits, and I'm grateful for your support and guidance in lab. Liv, thank you for your career guidance and for showing me the ropes in lab. And Emma, your introduction to crosswords added an extra layer of enjoyment to our lab days, and I'm thankful for the fun times we shared. Also, a special thank you for letting me spend some extra time with Michelle when I needed a doggy break.

To the current postdocs in the Hurley Lab, Rick Hooy, Chen Cui, Aaron Joiner, Minghao Chen, Sankalp Shukla, and Kevin Larsen, I want to thank you for your contributions and support. Rick, sharing a bay with you has been a blast, and I'm thankful for all your IT support whenever I've struggled with the workstations. Chen, your willingness to drop everything to help me with lysosome-related questions and your encouragement throughout my projects have been incredibly valuable. I've learned so much from you about sample optimization and data processing. Aaron, thank you for your support with my in vitro assays and for taking the time to show me how to use new equipment and pass along protocols. Minghao, your assistance with cryo-EM microscope support and guidance with processing difficult EM datasets have been indispensable. Sankalp, collaborating with you on the KICSTOR project has been a pleasure, and

I appreciate your help with performing the GUV assays for KICSTOR. And Kevin, thank you for your support with my various in vitro assays and always turning on the dehumidifiers. Most importantly, thank you to you all for being people I can chat with to bounce ideas off of. I feel very fortunate to have worked with such supportive and collaborative people.

Thank you to my fellow graduate students who joined the Hurley Lab at the same time as me, Kevin Rose, Zeke Cook, and Dan Tudorica. Kevin, thank you for the countless break room chats and the memorable nights out. Zeke, it's been great learning more about autophagy from you, and your expertise has enriched our discussions. Dan, I appreciate your willingness to collaborate and share ideas, and I look forward to our continued collaboration as we progress in our research.

Thank you to my undergraduates Albert Tran, Karla Tapia and Selina Wu. Working with each of you has been an absolute pleasure. Thank you for your support on my projects and assistance with generating mutants, protein purifications and cell culture.

Most importantly, thank you to our lab manager Snow Ren. Snow, learning from your vast expertise in molecular biology and biochemistry has been nothing short of amazing. Thank you for your support in various aspects of the KICSTOR project, from generating constructs for SZT2 to helping me refine the purification protocols for KICSTOR and GATOR2. Your assistance in making mutants for the GATOR2 subunits and your generosity in sharing lab resources, including buffers, have been invaluable.

A special thank you to Catherine Dea for her assistance with all my scheduling and administrative needs. Catherine, your dedication and organization help keep our lab running smoothly, and I'm grateful for your support. I also want to express my thanks to the members of Cal Cryo, Jon Remis, Dan Toso, Ravi Thakkar, and Paul Tobias, for their support with my data collections and for ensuring the smooth operation of the microscope.

A huge thank you to my collaborators in the Zoncu Lab for their support with the cell biology experiments that have complemented my structural biology work. Working with you has been an enriching experience, and I am grateful for your guidance and collaboration. Roberta Peruzzo, it has been a pleasure working alongside you. Thank you for patiently teaching me the cell biology techniques and for your support as I learned them for the first time.

To my thesis committee members, Roberto Zoncu, Andreas Martin, and Eunyoung Park, I am so grateful for your scientific support and guidance throughout my thesis journey. Roberto, your expertise in mTORC1 and cell biology has been incredibly helpful in shaping my structural studies. Andy, I appreciate your challenging questions and encouragement to think critically about my in vitro assays, which have greatly enriched my research. Eunyoung, your insights and expertise in cryo-EM have been instrumental in advancing my work in this area. Thank you all for your contributions to my academic and scientific growth.

To my mentor, Jim, your guidance, and mentorship throughout my graduate studies has been invaluable. Your constant push for growth and challenge has shaped me into the scientist I am today. Thank you for supporting my scientific and professional aspirations and for placing

me on projects that align with my interests while pushing me to excel. Your ongoing support during our frequent 1x1 meetings and the opportunities you've provided for me to develop my professional skills have been instrumental. I'm truly grateful for your mentorship and I'm excited for our continued collaboration.

Chapter 1: Introduction

Portions of this chapter were published as:

Longin domain GAP complexes in nutrient signaling, membrane traffic, and neurodegeneration

Rachel M. Jansen[1,2] and James H. Hurley[1,2,3*]

[1]Department of Molecular and Cell Biology, University of California Berkeley; Berkeley CA 94720, USA.
[2]California Institute for Quantitative Biosciences, University of California, Berkeley, CA, 94720, USA
[3]Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA 94720, USA

*Corresponding author

Citation:

Jansen RM and Hurley JH (2022) Longin domain GAP complexes in nutrient signaling, membrane traffic, and neurodegeneration. FEBS Letters

Mechanistic Target of Rapamycin Complex 1 and Lysosomal Regulation

Lysosomes are highly dynamic organelles that play a crucial role in various cellular functions, including gene regulation, metabolic signaling, cell motility, and immunity. These organelles use a diverse array of integral membrane proteins and lysosome-associated proteins to integrate environmental signals and adjust cellular processes accordingly. Appropriately, the lysosome has been termed a regulatory hub for cellular homeostasis and is implicated in various diseases associated with cellular imbalance, such as neurogenerative disorders and cancer [1].

The mechanistic target of rapamycin complex 1 (mTORC1), a notable lysosome-associated protein, plays a central role in the response to fluctuations in nutrients, growth factors, and energy in cells by altering cell growth. Upregulation and overactivation of the mTORC1 pathway have been implicated in numerous types of cancer. Similarly, inhibition of mTORC1 has been shown to increase autophagic flux leading to the clearance of aggregate prone proteins seen in many neurodegenerative diseases [2]. The criticality of mTORC1 regulation in human disease has made mTORC1 a highly sought-after drug target. Ever since the discovery of rapamycin, the first mTOR inhibitor, it has been a long-standing goal to develop equally efficacious inhibitors with less toxic off-target effects. As a result, mTORC1, its regulation and its involvement in various signaling pathways has been widely studied.

As a member of the phosphoinositide 3-kinase (PI3K)-related kinase family, mTORC1 integrates and propagates signals through phosphorylation of multiple substrates. Many of these substrates share a mTORC1 recognition sequence such as 4E-BP1 and S6K1 that function to regulate cell growth, transcription and translation[3]. However, some transcription factors, such as transcription factor EB and E3 (TFEB/TFE3), that function to regulate lysosome biogenesis and autophagy do not share the same binding motif. Instead, the interaction between TFEB/TEF3 and mTORC1 requires co-localization to the nutrients-sensing, lysosome-associated Rag-Ragulator complex.

The Rag-Ragulator Complex

The Rag GTPases, a unique subgroup of the Ras superfamily, localize to the lysosomal membrane and function as indirect nutrient sensors by adapting their nucleotide state in response to nutrient levels. The Rag GTPases function as an obligate heterodimer containing Rag A or B in complex with Rag C or D. The inactive dimer (RagA or B^{GDP}:RagC or D^{GTP}) is present under low nutrient conditions, and the active dimer under high nutrient conditions (RagA or B^{GTP}:RagC or D^{GDP}) [4]. Two longin domain-containing GTPase activating proteins (GAPs) complexes, folliculin: folliculin interacting protein 2 (FLCN:FNIP2) and GATOR1, convert the Rag GTPases between two stable states [5,6] (Figure 1.1). Unlike most small GTPases, the Rags lack the lipid modifications required to directly interact with the membrane. Instead, they are localized to the membrane through the scaffold protein complex, Ragulator. Ragulator is comprised of five subunits, LAMTOR 1-5, and interacts with the Rags through LAMTOR1 and LAMTOR2 to form the Rag-Ragulator complex [7-9]. Through the Rag-Ragulator complex, mTORC1 is localized to the lysosomal membrane under high nutrient conditions [9-11]. Tight regulation of the Rag GTPase nutrient state through FLCN:FNIP2 and GATOR1 is critical for proper localization and activity of mTORC1.

Regulation of the GTPase Activating Protein FLCN

The FLCN:FNIP complex functions as a GAP for RagC/D, transitioning the Rag-Ragulator complex from an inactive to active state under high nutrient conditions [5]. Two forms of FNIP are expressed in most cell types, FNIP1 and FNIP2, and both complexes are active against RagC/D [5]. FLCN and RagC/D^{GDP} are uniquely important for a subset of mTORC1 substrates, the MiT-TFE transcription factors [12,13]. The transcription factors EB and E3 (TFEB/TFE3) are the best known members of this family. The phosphorylation of other well-known mTORC1 substrates, as well as mTORC1 localization to lysosomes, is essentially independent of RagC/D nucleotide state [12,13] (Figure 1.2B). Phosphorylation of TFEB/TFE3 promotes its sequestration to the cytosol, preventing transcription of genes important for autophagy and lysosome biogenesis (Figure 1.2B). Thus, FLCN and RagC/D^{GDP} selectively regulate a "non-canonical" subset of mTORC1 substrates.

Cryo-EM structures have provided snapshots of FLCN:FNIP2 interacting with Rag-Ragulator to elucidate the GAP mechanism and nuances of FLCN:FNIP2 regulation (Figure 1.3A). Although FLCN:FNIP2 functions as a RagC/D GAP, presumably at lysosomes, under high nutrient conditions, paradoxically, it localizes most strongly to the lysosome under conditions of amino acid starvation [14] (Figure 1.2A). The cryo-EM structure of FLCN in complex with Rag-Ragulator, termed the lysosomal folliculin complex (LFC), demonstrated that this binding mode was a GAP inhibitory binding mode where FLCN rests in the cleft between RagA and RagC making substantial contacts with both Rag GTPases [12]. In this binding mode, the catalytic Arg finger residue, Arg-164, is positioned more than 20 Å away from the RagC nucleotide binding pocket, confirming FLCN is inactive. To accommodate FLCN in this position, when inactive, the Rags undergo a conformational change, rotating to increase the size of the cleft by nearly double that of the active state [12]. The lack of FLCN lysosomal localization under high nutrients suggested that the GAP active binding mode for FLCN was transient. This led to challenges in determining a cryo-EM structure of the active FLCN complex and left a gap in knowledge about the mechanisms of FLCN GAP activity.

It is known that the transition between the inhibitory binding state and active binding mode is triggered by the presence of cellular nutrients. A nutrient sensor in the mTORC1 pathway, the amino acid transporter solute carrier family 38 member 9 (SLC38A9), was shown to stimulate this conversion upon increase in cellular nutrients [15,16]. When lysosomal arginine levels are low, the N-terminal tail of SLC38A9 is nestled in the transmembrane portion of the transporter. As arginine levels increase, binding of arginine to SLC38A9 outcompetes binding of the N-terminal tail (SLC38A9^{NT}) and liberates it on the cytosolic face of the lysosomal membrane. The liberated SLC38A9^{NT} then binds to the Rag dimer inter-domain cleft [17] (Figure 1.2B). SLC38A9^{NT} binds in such a way that it competes with FLCN as it binds to the inactive LFC, so breaking up the LFC [17]. It remains elusive if the binding of SLC38A9^{NT} is compatible with the active GAP interaction and if SLC38A9 passively or actively engages with FLCN to promote the active binding mode.

Regulation of the GTPase Activating Protein GATOR1

RagA/B^{GTP} makes the major direct contact with mTORC1 involved in its lysosomal localization, and as such, it is critical for activation. The GAP for RagA/B is the GATOR1 complex. GATOR1 consists of three subunits NPRL2, NPRL3 and DEPDC5. GATOR1 inactivates the Rag dimer by converting RagA/B^{GTP} to the RagA/B^{GDP} under low nutrient conditions. Under these conditions, mTORC1 is inactive with respect to all of its substrates, canonical and non-canonical (Figure 1.2A).

In a striking parallel to the FLCN:FNIP complex, cryo-EM structures of GATOR1 bound to Rag dimers have revealed both a GAP inhibitory and an active GAP binding mode [18,19] (Figure 1.3B). In the inhibitory binding mode, GATOR1 interacts with RagA predominately through the non-catalytic DEPDC5 domain [19]. In this conformation, there is no interaction between the catalytic finger Arg-78 of NPRL2 and RagA [20]. The Rag GTPases remain in their active form with the cleft in its open conformation. The GAP active binding mode occurs when GATOR1 is reoriented such that the catalytic domains NPRL2-NPRL3 interact with RagA and the catalytic Arg-78 is positioned in the nucleotide binding pocket [18] (Figure 1.4A and B). In this binding mode, GATOR1 DEPDC5 domain also makes substantial contacts with RagC. This contrast FLCN, where in its GAP active binding mode, it only interacts with the GTPase it is stimulating hydrolysis of, RagC. The stimuli that trigger the conversion between the inhibitory and active binding modes have yet to be determined.

The answer may lie within the identified regulators that interact directly with GATOR1. A handful of GATOR1 regulators have been uncovered such as GATOR2, KICSTOR and the nutrient sensor SAMTOR [6,21,22]. GATOR2 is a negative regulator of GATOR1. The GATOR2 complex is massive, consisting of five subunits (WDR24, MIOS, WDR59, SEH1L and SEC13) that adopt an open cage-like structure with an octagonal scaffold [23]. Extending from the scaffold are eight pairs of WD40 β-propellers that could mediate interactions between GATOR1, KICSTOR and additional nutrients sensors in the pathway [23] (Figure 1.5).

The membrane-associated protein complex KICSTOR is comprised of a large scaffold-like protein SZT2 and three additional subunits KPTN, ITFG2, and KICS2 [22] (Figure 1.5). The large size and limited structural homology of SZT2 with other proteins has posed challenges for gaining structural insight into the nature of this complex. GATOR1 is thought to interact with the complex through SZT2, but all four KICSTOR subunits are necessary for GAP activity [22,24]. It remains unclear if GATOR2 is required to mediate the interaction between GATOR1 and KICSTOR. It is possible that KICSTOR functions as a nutrient sensor or interacts with GATOR2 and/or additional nutrient sensors to trigger the conversion between the GATOR1 binding modes. Additional studies into the interactions between GATOR1, GATOR2 and KICSTOR are necessary for determining the mechanism of each within the pathway in response to fluctuations in nutrients.

Nutrient Sensors and the GATOR complexes

The activity of the GATOR complexes depends on cellular nutrient levels, specifically amino acids, although direct amino acid binding to these complexes has not been demonstrated.

Instead, smaller proteins in the pathway serve as nutrient sensors, interacting with GATOR complexes to convey signals about nutrient availability. Three known nutrient sensors, SAMTOR, Sestrin2, and Castor1, each recognize specific amino acids and integrate those signals into the pathway (Figure 1.5). SAMTOR binds to SAM, derived from methionine and ATP, while Sestrin2 binds leucine, and Castor1 binds arginine.

The nutrient sensor SAMTOR, is the only identified sensor that directly interacts with and alters the activity of GATOR1 and KICSTOR. Under low-nutrient conditions it binds to the interface between KICSTOR and GATOR1 facilitating GATOR1 GAP activity. Upon binding to SAM, SAMTOR releases from the complexes transitioning GATOR1 to an inhibitory mode. Structural studies of *Drosophila* SAMTOR MTase domain suggest SAM binding induces a conformational change blocking SAMTOR's interaction with GATOR1 and KICSTOR[25]. On the other hand, Castor1 and Sestrin2 relay information about nutrient levels through interaction with GATOR2. In low-nutrient conditions, both nutrient sensors interact with GATOR2 subunits, preventing GATOR2 from inhibiting GATOR1 GAP function. It's been proposed that Castor1 and Sestrin2 interacts with the GATOR2 subunits MIOS and WDR24 respectively[26,27]. Upon increases in nutrient levels, these sensors dissociate, allowing GATOR2 to interact with GATOR1 in a manner that inhibits GAP activity.

Structural information about Castor1 and Sestrin2 have provided insights into their nutrient dependent regulation of GATOR2 activity. Castor1 forms a functional homodimer with each monomer binding to one arginine[27-29]. The arginine molecule is buried inside the Castor1 structure and located at the interface of the N-terminal domain (NTD) and C-terminal domain (CTD). Visualization of the location of the arginine molecule in initial crystal structures suggested Castor1 potentially underwent an open-close mechanism upon arginine binding[27]. However, a later structure of apo Castor1 revealed that the global architecture of the structure remained the same as the arginine bound state[30]. This left questions about how Castor1 interacts with GATOR2 in a nutrient dependent manner given there was no obvious conformation change in the structure. The crystal structure of Sestrin2, revealed it was a monomeric globular protein that binds a single leucine molecule in the CTD. The leucine molecule is position in the interior of Sestrin2 cover by a "lid" formed by two alpha helices[31]. The determination of this structure suggested Sestrin2 may transmit signals to GATOR2 through a "lid-latch" mechanism [31]. However, attempts to determine a structure of apo-Sestrin2 to verify this mechanism have been unsuccessful. As a result, there are many questions still persist including how Sestrin2 interact with GATOR2 to relay information about nutrient availability.

Focus of this Work

This work aims to fill gaps in knowledge in the regulation of FLCN and GATOR1. The first part focuses on the FLCN, elucidating its structure when bound to Rag-Ragulator in its active mode and determining the mechanism of GAP activity. This allows for a more complete picture of how FLCN interacts with and regulates the Rag-Ragulator complex in a nutrient dependent manner. The second part delves into the GATOR1 axis, examining the role of the KICSTOR and GATOR2 in the regulation of GATOR1. Specifically focusing on structural and functional insights into the KICSTOR as well as the interaction of the nutrient sensors Castor1

and Sestrin2 with GATOR2. This works has been a collaborative effort with members of the Zoncu Lab and integrates structural insights with cellular biology understanding.

Figures:

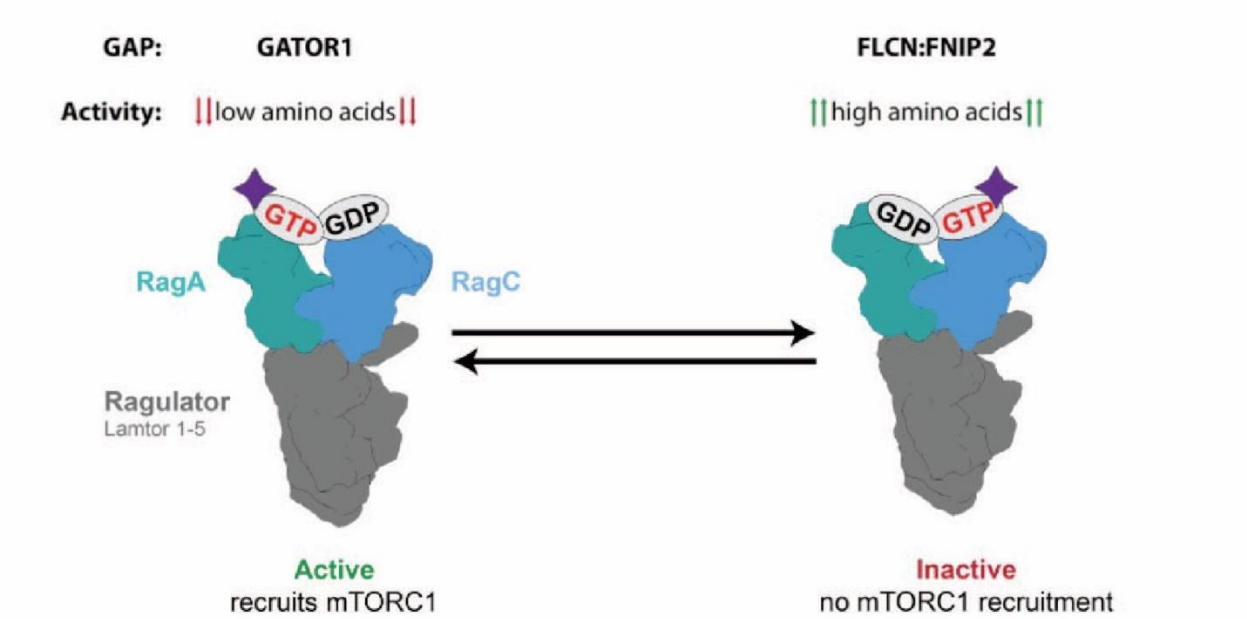


Figure 1.1 Activity of the Rag-Ragulator complex determined by nutrient availability. The Rag-Ragulator complex exists primarily in two stable states and active and inactive complex. The two states are cycled between with the help of the GAPs GATOR1 and FLCN:FNIP2 that function under low and high nutrient conditions respectively. This figure was adapted from Simon Fromm.

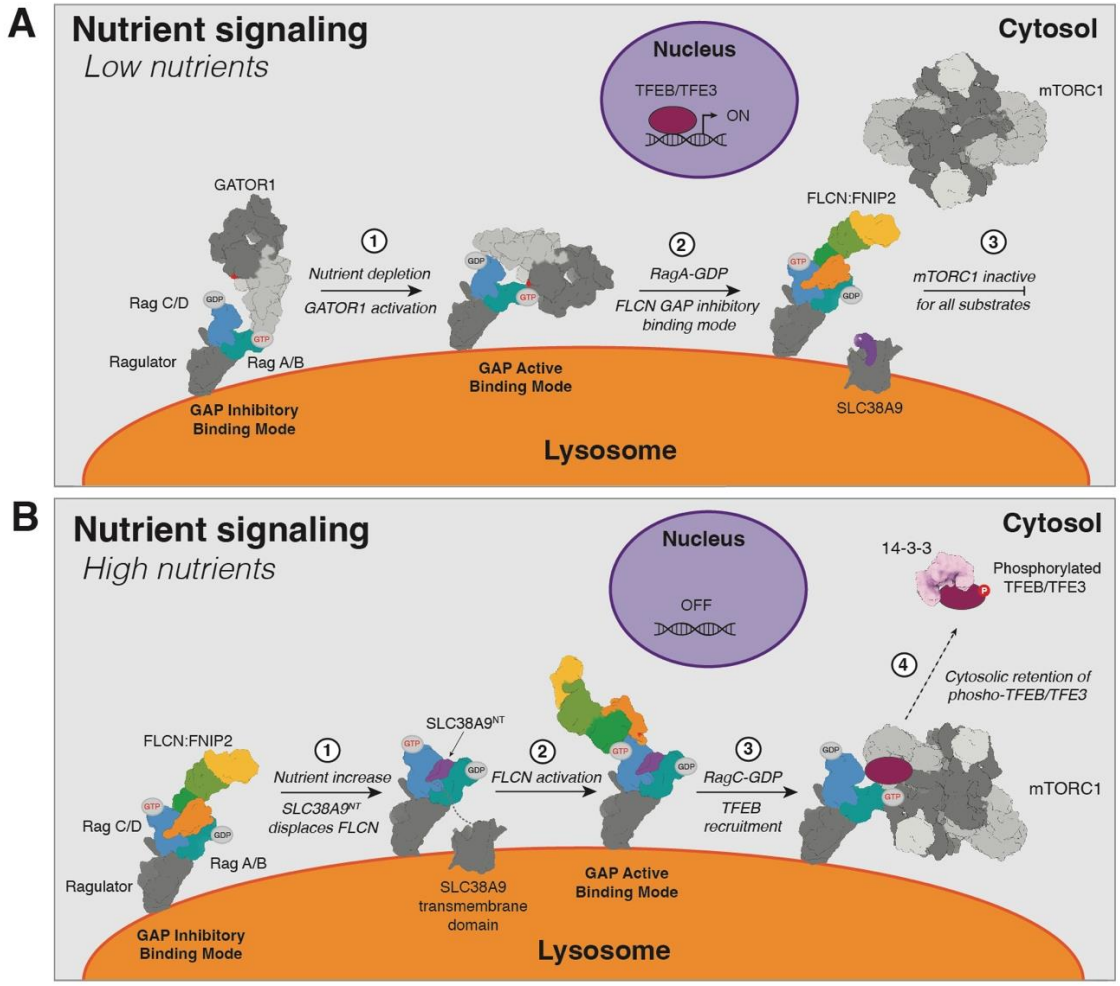


Figure 1.2. Schematic representation of the regulation of Rag-Ragulator in different nutrient conditions. (A) *(1)* Upon depletion of nutrients, GATOR1 transitions from its GAP inhibitory binding to its GAP active binding mode. *(2)* GATOR1 stimulates hydrolysis of RagA-GTP to form RagA-GDP. RagC spontaneously exchanges GDP for GTP generating the inactive Rags ($RagA^{GDP}$:$RagC^{GTP}$). FLCN:FNIP2 binds to inactive Rags in its GAP inhibitory binding mode. *(3)* Inactivation of Rags prevents mTORC1 recruitment and activation leading to no phosphorylation of all mTORC1 substrates. (B) *(1)* As cellular nutrients increase, the N-terminal of SLC38A9 ($SLC38A9^{NT}$) (purple) is freed from the transmembrane domain (dark grey) and displaces FLCN (orange):FNIP2 (green) in its GAP inhibitory binding mode. *(2)* FLCN:FNIP2 reorients to interact with the Rag-Ragulator complex (blue and grey) in its GAP active binding mode. *(3)* FLCN:FNIP2 stimulates the hydrolysis of RagC-GTP forming RagC-GDP. RagA-GDP undergoes spontaneous exchange to RagA-GTP generating active Rags ($RagA^{GTP}$:$RagC^{GDP}$). Active Rags recruit mTORC1 to the lysosomal membrane. *(4)* TFEB/TFE3 co-localizes with mTROC1 at the lysosomal membrane allowing mTORC1 to phosphorylate TFEB/TFE3. Phosphorylated TFEB/TFE3 is retained in the cytosol through interaction with 14-3-3 proteins. Not shown, following GATOR1 GAP inactivation, GDP to GTP exchange on RagA activates mTORC1 phosphorylation of canonical substrates, which include 4E-BP1, S6K, and ULK1.

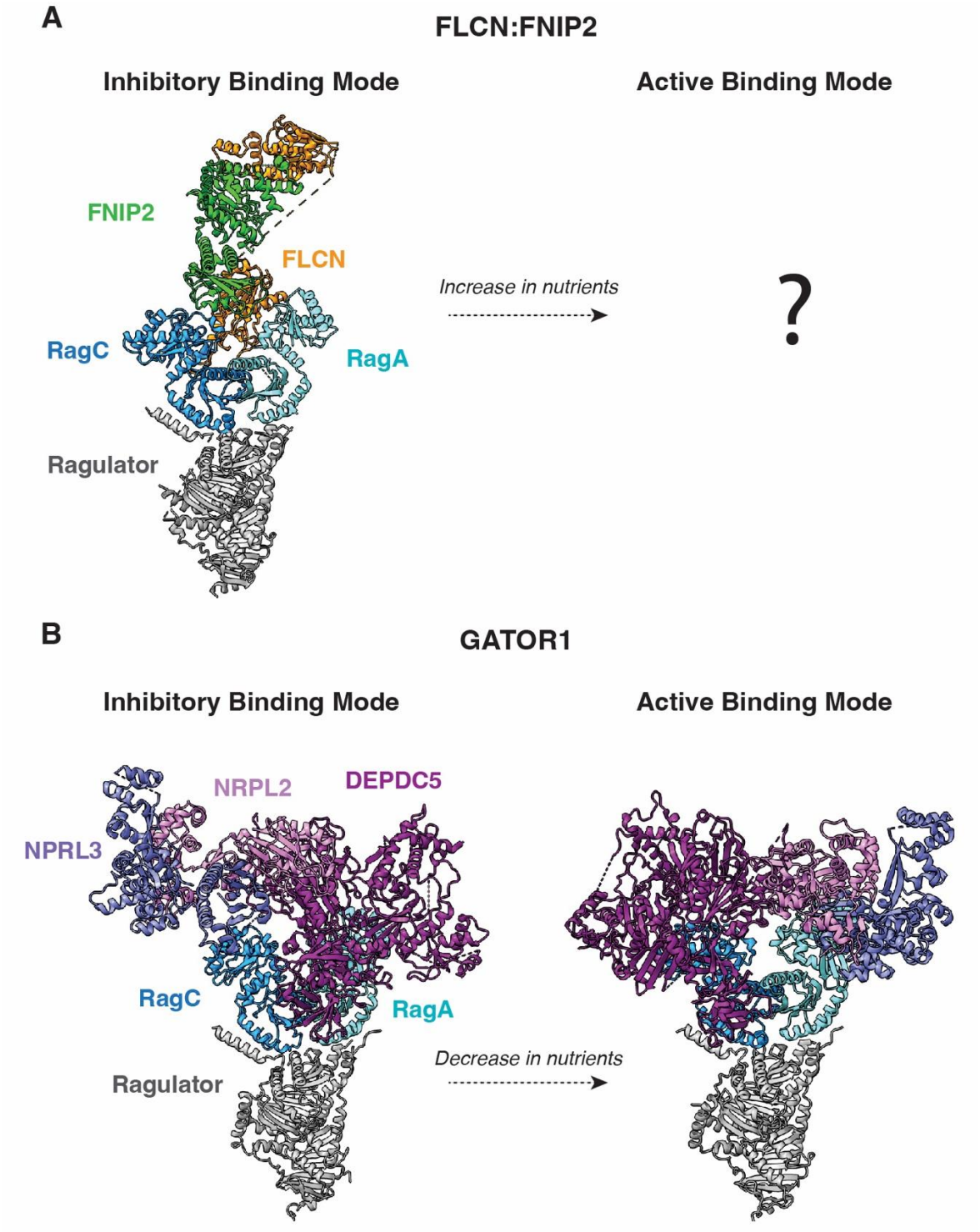


Figure 1.3 Dual binding modes of FLCN and GATOR1. (A) Cryo-EM structure of FLCN:FNIP2 (orange and green) bound to the Rag-Ragulator complex (blue and grey) reveals inhibitory binding mode (PDB: 6NZD). (B) Cryo-EM structures of GATOR1 complex containing NRPL2, NRPL3 and DEPDC5 (purple) bound to the Rag-Ragulator complex in two alternative binding modes (PDB: 7T3A and 7T3B).

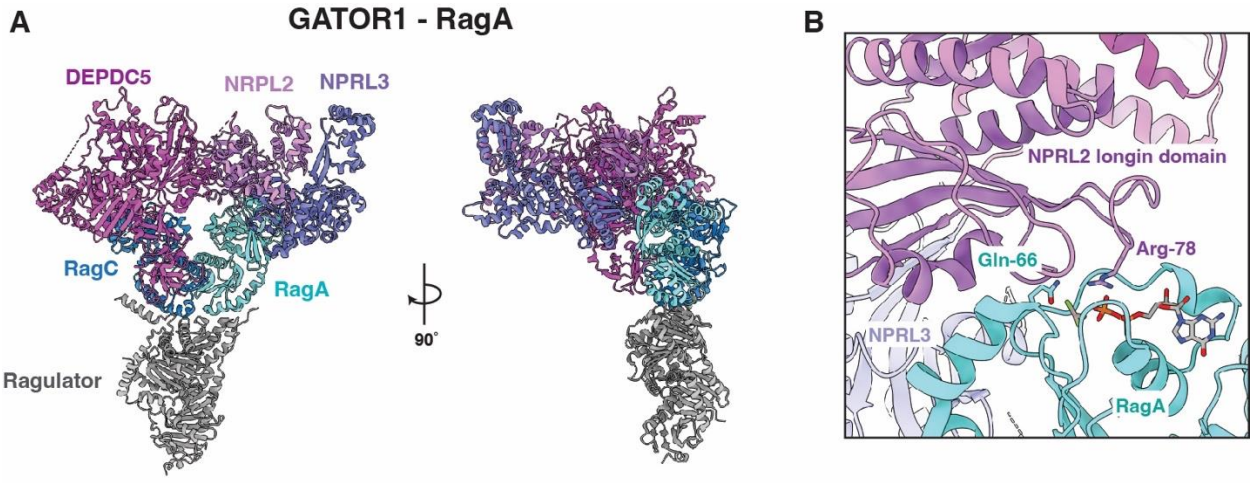


Figure 1.4 GATOR1 interaction with Rag-Ragulator in active mode and GAP mechanism. (A) Overall cryo-EM structure of GATOR1 complex containing NPRL2 (light purple), NRPL3 (blue-purple) and DEPDC5 (magenta) in complex with Rag-Ragulator (blue and grey) in active GAP binding mode (PDB: 7T3B). (B) Close-up view of GATOR1 interaction with RagA nucleotide binding site.

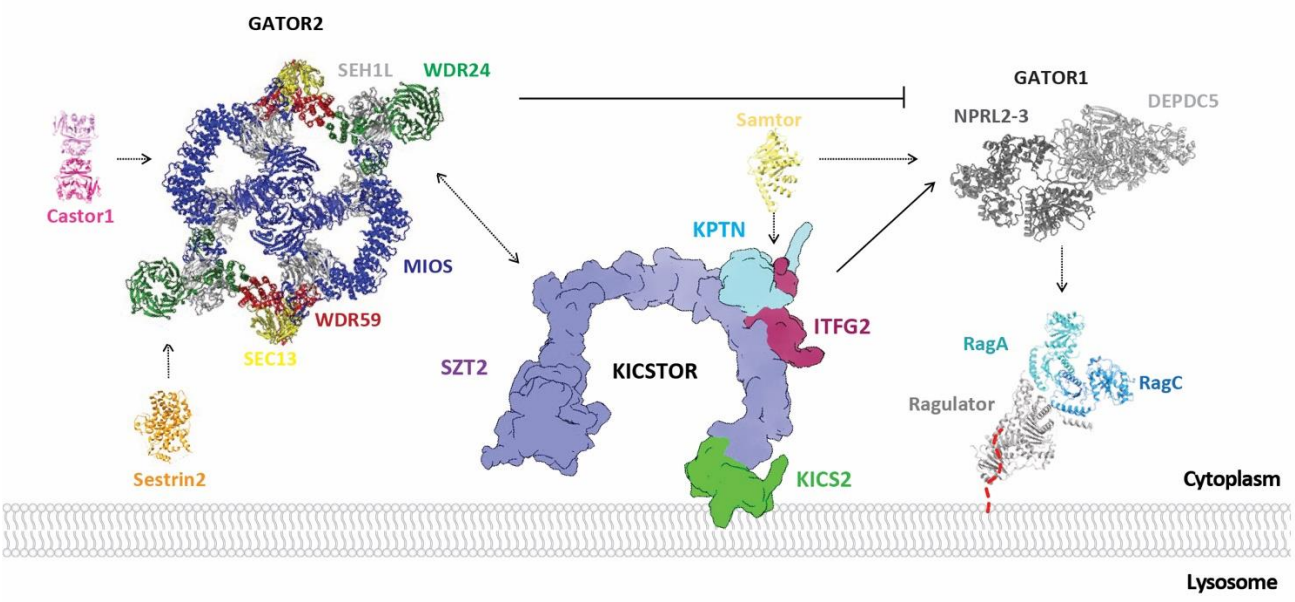


Figure 1.5 Regulation of GATOR1 and the corresponding structures. RagA GTPase activating protein (GAP), GATOR1 is regulated by a complex network of protein complexes and nutrient sensors. GATOR2 is proposed to be a negative regulator of GATOR1 while KICSTOR is hypothesized to be an activator. The core large complexes, GATOR1, GATOR2 and KICSTOR have been shown to interact although the extent and conditions surrounding these interactions remains elusive. Nutrient signals are integrated into the pathway through nutrient sensors Castor1 and Sestrin2 which interact with GATOR2 and SAMTOR which interacts with GATOR1 and KICSTOR.

References

1. Ballabio, A., and Bonifacino, J.S. (2020). Lysosomes as dynamic regulators of cell and organismal homeostasis. Nat Rev Mol Cell Biol *21*, 101-118. 10.1038/s41580-019-0185-4.
2. Saxton, R.A., and Sabatini, D.M. (2017). mTOR Signaling in Growth, Metabolism, and Disease. Cell *168*, 960-976. 10.1016/j.cell.2017.02.004.
3. Battaglioni, S., Benjamin, D., Walchli, M., Maier, T., and Hall, M.N. (2022). mTOR substrate phosphorylation in growth control. Cell *185*, 1814-1836. 10.1016/j.cell.2022.04.013.
4. Shen, K., Choe, A., and Sabatini, D.M. (2017). Intersubunit Crosstalk in the Rag GTPase Heterodimer Enables mTORC1 to Respond Rapidly to Amino Acid Availability. Mol Cell *68*, 552-565 e558. 10.1016/j.molcel.2017.09.026.
5. Tsun, Z.Y., Bar-Peled, L., Chantranupong, L., Zoncu, R., Wang, T., Kim, C., Spooner, E., and Sabatini, D.M. (2013). The folliculin tumor suppressor is a GAP for the RagC/D GTPases that signal amino acid levels to mTORC1. Mol Cell *52*, 495-505. 10.1016/j.molcel.2013.09.016.
6. Bar-Peled, L., Chantranupong, L., Cherniack, A.D., Chen, W.W., Ottina, K.A., Grabiner, B.C., Spear, E.D., Carter, S.L., Meyerson, M., and Sabatini, D.M. (2013). A Tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. Science *340*, 1100-1106. 10.1126/science.1232044.
7. Yonehara, R., Nada, S., Nakai, T., Nakai, M., Kitamura, A., Ogawa, A., Nakatsumi, H., Nakayama, K.I., Li, S., Standley, D.M., et al. (2017). Structural basis for the assembly of the Ragulator-Rag GTPase complex. Nat Commun *8*, 1625. 10.1038/s41467-017-01762-3.
8. Sekiguchi, T., Hirose, E., Nakashima, N., Ii, M., and Nishimoto, T. (2001). Novel G proteins, Rag C and Rag D, interact with GTP-binding proteins, Rag A and Rag B. J Biol Chem *276*, 7246-7257. 10.1074/jbc.M004389200.
9. Sancak, Y., Bar-Peled, L., Zoncu, R., Markhard, A.L., Nada, S., and Sabatini, D.M. (2010). Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. Cell *141*, 290-303. 10.1016/j.cell.2010.02.024.
10. Kim, E., Goraksha-Hicks, P., Li, L., Neufeld, T.P., and Guan, K.L. (2008). Regulation of TORC1 by Rag GTPases in nutrient response. Nat Cell Biol *10*, 935-945. 10.1038/ncb1753.
11. Sancak, Y., Peterson, T.R., Shaul, Y.D., Lindquist, R.A., Thoreen, C.C., Bar-Peled, L., and Sabatini, D.M. (2008). The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. Science *320*, 1496-1501. 10.1126/science.1157535.
12. Lawrence, R.E., Fromm, S.A., Fu, Y., Yokom, A.L., Kim, D.J., Thelen, A.M., Young, L.N., Lim, C.Y., Samelson, A.J., Hurley, J.H., and Zoncu, R. (2019). Structural mechanism of a Rag GTPase activation checkpoint by the lysosomal folliculin complex. Science *366*, 971-977. 10.1126/science.aax0364.
13. Napolitano, G., Di Malta, C., Esposito, A., de Araujo, M.E.G., Pece, S., Bertalot, G., Matarese, M., Benedetti, V., Zampelli, A., Stasyk, T., et al. (2020). A substrate-specific mTORC1 pathway underlies Birt-Hogg-Dubé syndrome. Nature *585*, 597-602. 10.1038/s41586-020-2444-0.

14. Petit, C.S., Roczniak-Ferguson, A., and Ferguson, S.M. (2013). Recruitment of folliculin to lysosomes supports the amino acid-dependent activation of Rag GTPases. J Cell Biol *202*, 1107-1122. 10.1083/jcb.201307084.
15. Rebsamen, M., Pochini, L., Stasyk, T., de Araujo, M.E., Galluccio, M., Kandasamy, R.K., Snijder, B., Fauster, A., Rudashevskaya, E.L., Bruckner, M., et al. (2015). SLC38A9 is a component of the lysosomal amino acid sensing machinery that controls mTORC1. Nature *519*, 477-481. 10.1038/nature14107.
16. Wang, S., Tsun, Z.Y., Wolfson, R.L., Shen, K., Wyant, G.A., Plovanich, M.E., Yuan, E.D., Jones, T.D., Chantranupong, L., Comb, W., et al. (2015). Metabolism. Lysosomal amino acid transporter SLC38A9 signals arginine sufficiency to mTORC1. Science *347*, 188-194. 10.1126/science.1257132.
17. Fromm, S.A., Lawrence, R.E., and Hurley, J.H. (2020). Structural mechanism for amino acid-dependent Rag GTPase nucleotide state switching by SLC38A9. Nat Struct Mol Biol *27*, 1017-1023. 10.1038/s41594-020-0490-9.
18. Egri, S.B., Ouch, C., Chou, H.T., Yu, Z., Song, K., Xu, C., and Shen, K. (2022). Cryo-EM structures of the human GATOR1-Rag-Ragulator complex reveal a spatial-constraint regulated GAP mechanism. Mol Cell. 10.1016/j.molcel.2022.03.002.
19. Shen, K., Huang, R.K., Brignole, E.J., Condon, K.J., Valenstein, M.L., Chantranupong, L., Bomaliyamu, A., Choe, A., Hong, C., Yu, Z., and Sabatini, D.M. (2018). Architecture of the human GATOR1 and GATOR1-Rag GTPases complexes. Nature *556*, 64-69. 10.1038/nature26158.
20. Shen, K., Valenstein, M.L., Gu, X., and Sabatini, D.M. (2019). Arg-78 of Nprl2 catalyzes GATOR1-stimulated GTP hydrolysis by the Rag GTPases. J Biol Chem *294*, 2970-2975. 10.1074/jbc.AC119.007382.
21. Gu, X., Orozco, J.M., Saxton, R.A., Condon, K.J., Liu, G.Y., Krawczyk, P.A., Scaria, S.M., Harper, J.W., Gygi, S.P., and Sabatini, D.M. (2017). SAMTOR is an S-adenosylmethionine sensor for the mTORC1 pathway. Science *358*, 813-818. 10.1126/science.aao3265.
22. Wolfson, R.L., Chantranupong, L., Wyant, G.A., Gu, X., Orozco, J.M., Shen, K., Condon, K.J., Petri, S., Kedir, J., Scaria, S.M., et al. (2017). KICSTOR recruits GATOR1 to the lysosome and is necessary for nutrients to regulate mTORC1. Nature *543*, 438-442. 10.1038/nature21423.
23. Valenstein, M.L., Rogala, K.B., Lalgudi, P.V., Brignole, E.J., Gu, X., Saxton, R.A., Chantranupong, L., Kolibius, J., Quast, J.P., and Sabatini, D.M. (2022). Structure of the nutrient-sensing hub GATOR2. Nature *607*, 610-616. 10.1038/s41586-022-04939-z.
24. Peng, M., Yin, N., and Li, M.O. (2017). SZT2 dictates GATOR control of mTORC1 signalling. Nature *543*, 433-437. 10.1038/nature21378.
25. Tang, X., Zhang, Y., Wang, G., Zhang, C., Wang, F., Shi, J., Zhang, T., and Ding, J. (2022). Molecular mechanism of S-adenosylmethionine sensing by SAMTOR in mTORC1 signaling. Sci Adv *8*, eabn3868. 10.1126/sciadv.abn3868.
26. Wolfson, R.L., Chantranupong, L., Saxton, R.A., Shen, K., Scaria, S.M., Cantor, J.R., and Sabatini, D.M. (2016). Sestrin2 is a leucine sensor for the mTORC1 pathway. Science *351*, 43-48. 10.1126/science.aab2674.
27. Saxton, R.A., Chantranupong, L., Knockenhauer, K.E., Schwartz, T.U., and Sabatini, D.M. (2016). Mechanism of arginine sensing by CASTOR1 upstream of mTORC1. Nature *536*, 229-233. 10.1038/nature19079.

28. Gai, Z., Wang, Q., Yang, C., Wang, L., Deng, W., and Wu, G. (2016). Structural mechanism for the arginine sensing and regulation of CASTOR1 in the mTORC1 signaling pathway. Cell Discov *2*, 16051. 10.1038/celldisc.2016.51.
29. Xia, J., Wang, R., Zhang, T., and Ding, J. (2016). Structural insight into the arginine-binding specificity of CASTOR1 in amino acid-dependent mTORC1 signaling. Cell Discov *2*, 16035. 10.1038/celldisc.2016.35.
30. Zhou, Y., Wang, C., Xiao, Q., and Guo, L. (2019). Crystal structures of arginine sensor CASTOR1 in arginine-bound and ligand free states. Biochem Biophys Res Commun *508*, 387-391. 10.1016/j.bbrc.2018.11.147.
31. Saxton, R.A., Knockenhauer, K.E., Wolfson, R.L., Chantranupong, L., Pacold, M.E., Wang, T., Schwartz, T.U., and Sabatini, D.M. (2016). Structural basis for leucine sensing by the Sestrin2-mTORC1 pathway. Science *351*, 53-58. 10.1126/science.aad2087.

Chapter 2: Structural Basis of FLCN RagC GAP Activation

Portions of this chapter were published as:

Structural basis for FLCN RagC GAP activation in MiT-TFE substrate-selective mTORC1 regulation

Rachel M. Jansen[1,2], Roberta Peruzzo[1], Simon A. Fromm[1,2,†], Adam L. Yokom[1,2] Roberto Zoncu[1,2] and James H. Hurley[1,2,3*]

[1]Department of Molecular and Cell Biology, University of California Berkeley; Berkeley CA 94720, USA.
[2]California Institute for Quantitative Biosciences, University of California, Berkeley, CA, 94720, USA
[3]Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA 94720, USA
†Present address: Imaging Centre, European Molecular Biology Laboratory, Heidelberg, Germany

*Corresponding author: jimhurley@berkeley.edu

Citation:

Jansen RM, Peruzzo R, Fromm SA, Yokom AL, Zoncu R, and Hurley JH (2022) Structural basis for FLCN RagC GAP activation in MiT-TFE substrate-selective mTORC1 regulation. Science Advances

Introduction

The mechanistic target of rapamycin complex 1 (mTORC1) plays a central role in the response to fluctuations in nutrients, growth factors, and energy in cells by altering the balance between cell growth and catabolism. In nutrient-rich environments, mTORC1 activates anabolic processes including protein and lipid synthesis while inhibiting catabolic ones such as autophagy [1,2]. mTORC1 regulates its pro-anabolic programs through phosphorylation of key targets, some of which contain a conserved TOS motif [3,4]. The microphthalmia (MiT-TFE) family of transcription factors, including similarly regulated transcription factors EB and E3 (TFEB and TFE3), that activate transcription of genes involved in autophagy [5] and lysosome biogenesis [6], and are regulated by mTORC1 [7-9]. Among well-known mTORC1 substrates, MiT-TFE family members lack the TOS motif responsible for docking some other substrates onto mTORC1 via its Raptor subunit. Also, in contrast to other mTORC1 substrates, MiT-TFE phosphorylation and negative regulation are completely dependent on the activity of the RagC GAP, FLCN [10-12].

The Rag-Ragulator complex consists of two Rag guanosine triphosphatases (commonly referred to as Rags) that function as obligate heterodimers and are anchored to the lysosomal membrane by the pentameric Ragulator [13,14]. The Rags respond to lysosomal and cytoplasmic nutrient levels, which regulate the proportion of inactive (RagA or B^{GDP}:RagC or D^{GTP}) and active (RagA or B^{GTP}:RagC or D^{GDP}) Rag dimers [15]. Conversion between active and inactive states is regulated by the RagA/B-specific GAP, GATOR1, and by the RagC/D-specific GAP, folliculin: folliculin interacting protein 2 (FLCN:FNIP2) [16,17]. In response to amino acid levels, as sensed by the transceptor SLC38A9 among other mechanisms [18-20], FLCN stimulates the conversion of Rags to their active form enabling mTORC1 recruitment and interaction with MiT-TFE members. Upon phosphorylation, MiT-TFE members bind to 14-3-3 proteins and is sequestered in the cytosol where they remain inactive. The regulation of TOS motif-containing mTORC1 substrates [3], including regulators of mRNA translation and cell proliferation, such as 4E-BP1 and S6K1 is unaffected [10,11,21]. Taken together, FLCN directly and selectively controls mTORC1 regulation of MiT-TFE members but not other substrates [10,11]. These findings support the existence of separate mechanisms for control of mTORC1-dependent anabolic versus catabolic programs [22], and highlight a therapeutic potential for FLCN to be targeted as a means of mTOR inhibition for treating LSDs and neurodegenerative diseases, with fewer toxic effects on cell proliferation and translation.

The transient nature of the GAP interaction with Rag-Ragulator has proven challenging for visualizing details necessary for deducing the mechanism of GAP activity. Previous structural investigation into the basis for FLCN localization to lysosomes during amino acid starvation [23] revealed a stable inhibitory FLCN-Rag-Ragulator complex, the lysosomal folliculin complex (LFC) [11,24]. This begged the question as to how FLCN binds to RagC to promote nucleotide hydrolysis, and what are the interfaces in FLCN essential for stimulating catalysis? We sought to understand how these residues interact with the RagC nucleotide binding pocket, and what makes FLCN specific for RagC. We previously found that the amino acid transporter, SLC38A9, promotes FLCN GAP activity by inserting its cytosolic domain in the cleft between the RagA and C G-domains and thereby disrupting the LFC [19]. Here, we trapped FLCN in its GAP-competent binding mode using a fusion construct with the N-terminal of SLC38A9

(SLC38A9^{NT}), determined the cryo-EM structure, and deduced the mechanism of FLCN GAP activity essential for interrogating this interaction during drug development.

Results

To provide a structural framework for understanding the mechanism of FLCN GAP activity, we stabilized the transient complex and determined the cryo-EM structure, hereafter referred to as the Active Folliculin Complex (AFC). We began by generating a stable complex of FLCN:FNIP2, Rags, Ragulator and SLC38A9^{NT}, utilizing a mutant FLCN^{F118D} known to disrupt the LFC without affecting GAP activity [11]. To tether FLCN to the Rag-Ragulator complex, we designed a fusion construct containing a 10-residue glycine-serine linker between the C-terminus of SLC38A9^{NT} and the N-terminus of FNIP2 (Figure 2.1). The C-terminal tail of SLC38A9^{NT} contains ~20 disordered residues, effectively increasing the linker length to ~30 residues allowing for up to 90 Å in length. We used a phosphate analog (BeF_3) during complex formation to capture FLCN^{F118D} bound to Rag-Ragulator [25]. Rags were loaded into a state mimicking starvation conditions (RagA^{GDP}:RagC^{GTP}) and incubated with Ragulator to form a stable complex. Next, FLCN^{F118D} and the SLC38A9^{NT}-FNIP2 (SLC-FNIP2) fusion protein were added to RagA^{GDP}:RagC^{GTP} – Ragulator immediately followed by addition of BeF_3 (Figure 2.2). Incubating the complex overnight, we generated a nucleotide ground-state analog that trapped FLCN^{F118D} in an active conformation and allowed for isolation of a stable active FLCN complex via size exclusion chromatography (Figure 2.1).

The cryo-EM structure of the AFC was determined to an overall resolution of 3.5 Å (Table 2.1, Figure 2.3). The cryo-EM density was of sufficient quality to model the ordered mass of the complex constituting the AFC (Figure 2.4 and 2.5). In the AFC, the longin domains of FLCN^{F118D} and FNIP2 interact with the RagC G-domain such that the FLCN^{F118D}:FNIP2 approaches Rag-Ragulator from the side (Figure 2.6). The SLC38A9^{NT} sits in the cleft between the Rags and does not interact directly with FLCN^{F118D}:FNIP2, suggesting it does not play a role in positioning the GAP. The overall dimensions of the complex are ~120 Å by ~180 Å (Figure 2.6). To interact with RagC in a GAP-competent conformation, FLCN^{F118D}:FNIP2 reorients 90 degrees and transitions from its elongated position in the LFC to bind the side of RagC in the AFC. In the LFC, FLCN:FNIP2 interacts with both RagC and RagA extensively, with 870 Å^2 and 645 Å^2 interfaces, respectively. In contrast, in the AFC conformation, FLCN^{F118D}:FNIP2 interacts solely with RagC, burying 1,141 Å^2 of surface area with RagC, but none with RagA.

To visualize the AFC catalytic complex in as much detail as possible, we used local refinement at the FLCN-RagC interface to increase the resolution to 3.16 Å and allow for clear placement of the nucleotide analog and previously identified arginine finger (Arg-164) in RagC nucleotide-binding domain (NBD) [11] (Figure 2.5 and 2.7). As bound to GDP-BeF^-_3, RagC adopts an architecture analogous to GTP-bound RagC and Gtr2 [26,27] (Figure 2.6). In this state, the switch I region is ordered and forms a lid at the top of the binding pocket. Additional density in the RagC NBD, adjacent to the BeF^-_3, aligned with the placement of the highly conserved Arg-164. The position of FLCN^{F118D} Arg-164 resembled the geometry of previously identified arginine fingers in Rab1 GAP structures trapped with BeF^-_3 nucleotide analogs [28,29] (Figure 2.8). The nucleotide binding pocket is composed of primarily charged residues with a 534 Å^2 surface area, while the overall volume of the larger surrounding pocket is 865 Å^3 [30] (Figure 2.7).

To determine the basis for FLCN specificity for RagC over RagA, we modeled the GAP interaction with RagA-GTP. The $FLCN^{F118D}$ loop containing residues 10-16 was accommodated when interacting with RagC (Figure 2.7). In contrast, in the RagA model, $FLCN^{F118D}$ (Phe^{16}, His^{14}, Pro^{16}) clashed with RagA at the interface (Figure 2.7). We sought to further explore FLCN specificity for RagC by understanding why GATOR1 does not function as a GAP for RagC. We modelled RagC with GATOR1 and observed residues in GATOR1 (Thr^{18} and Pro^{19}) that clashed with RagC at the interface (Figure 2.9). We also compared the RagC structure to that of Arf1, which is a substrate of another longin domain GAP, SMCR8 [31,32]. In the Arf1 model, $FLCN^{F118D}$ (Leu^{13} and His^{14}) clash with Arf1 residues (Figure 2.9), rationalizing specificity for RagC. To determine why other longin domain GAPs do not function to stimulate nucleotide hydrolysis of RagC, we modelled RagC with SMCR8, a longin GAP that acts on Arf GTPases. The binding between SMCR8 and RagC is incompatible due to clashes between SMCR8 (Glu^{62} and Phe^{63}) and RagC (Figure 2.9).

We analyzed the interface whereby $FLCN^{F118D}$:FNIP2 interacts with RagC. Two hydrophobic residues FLCN Phe10 and FNIP2 Val146 form prominent contacts in the RagC binding interface interacting with Thr96, Lys78, Met121 and Phe128 (Figure 2.10). We hypothesized that these residues were critical for GAP activity. To validate the importance of these residues, we created two mutants $FLCN^{F118D/F10D}$ (F10D:Phe^{10}→Asp) and $FNIP2^{V146E}$ (V146E:Val^{146}→Glu). We used a tryptophan-based GAP activity assay to monitor the activity of $FLCN^{F118D}$:FNIP2 and mutants $FLCN^{F118D/F10D}$:FNIP2 and $FLCN^{F118D}$:$FNIP2^{V146E}$ toward inactive Rags ($RagA^{GDP}$:$RagC^{GTP}$) [11,19]. The GAP activity of the $FLCN^{F118D}$:SLC-FNIP2 fusion protein was identical to the $FLCN^{F118D}$:FNIP2 without the SLC38A9 N-terminus fusion (Figure 2.11). Both $FLCN^{F118D/F10D}$:FNIP2 and $FLCN^{F118D}$:$FNIP2^{V146E}$ abolished GAP activity as compared to $FLCN^{F118D}$:FNIP2 confirming their importance in facilitating GAP binding (Figure 2.10).

We sought to investigate whether these GAP-inactive mutants, $FLCN^{F10D}$ and $FNIP2^{V146E}$, affected the ability of Rag-Ragulator to activate mTORC1 with respect to its various substrates. We used cell-based assays to image the localization of MiT-TFE family transcription factor TFE3. In amino-acid deprived cells reconstituted with either WT FLCN ($FLCN^{WT}$) or WT FNIP2 ($FNIP2^{WT}$), both $FLCN^{WT}$ and $FNIP2^{WT}$ were localized to the lysosome and TFE3 was present in the nucleus (Figure 2.10). Cells expressing $FLCN^{F10D}$ and $FNIP2^{V146E}$ during starvation resulted in lysosomal localization and nuclear localization of the MiT-TFE family member TFE3 (Figure 2.10). When cells were restimulated with amino acids, those reconstituted with $FLCN^{WT}$ and $FNIP2^{WT}$ resulted in translocation of TFE3 to the cytoplasm. In contrast, in cells expressing $FLCN^{F10D}$ and $FNIP2^{V146E}$ TFE3 remained nuclear, phenocopying the catalytic loss-of-function mutant $FLCN^{R164}$ [11] (Figure 2.12) and was not sequestered to the cytoplasm despite high cellular nutrients (Figure 2.10). These results further supported that the mutants $FLCN^{F10D}$ and $FNIP2^{V146E}$ eliminated GAP activity and confirmed that loss of GAP activity resulted in mis-regulation of mTORC1 and TFE3. Immunoblot analysis confirmed that the mutants did not have any effect on the phosphorylation of the TOS-motif containing substrates S6K and 4E-BP1 (Figure 2.13).

Discussion

The cryo-EM structure presented in this work reveals that FLCN:FNIP2 undergoes a dramatic re-orientation to convert from the inactive LFC to the active AFC. The AFC structure confirmed the prediction that FLCN activates nucleotide hydrolysis of RagC through the catalytic arginine finger Arg-164 [11,33]. We identified features on the FLCN surface that are complementary to RagC and confer the high degree of specificity that is seen biologically. Our structure revealed that the loop containing residues 10-16 in FLCN make it incompatible as a GAP for other small GTPases. In the AFC structure, $SLC38A9^{NT}$ sits in the cleft between the Rags, while FLCN:FNIP2 binds only RagC. This confirms the previously proposed model that $SLC38A9^{NT}$ displaces inactive FLCN:FNIP2 in the LFC to allow for binding in the GAP competent mode [19]. $SLC38A9^{NT}$ does not, however, play a role in GAP activity through interaction with FLCN or either Rag nucleotide binding pocket in the AFC, as further supported by the near-identical RagC-GAP activity of $FLCN^{F118D}$:FNIP2 with or without the $SLC38A9^{NT}$ fusion.

RagA is regulated by a different GAP complex, GATOR1. A recent cryo-EM structure of the GAP-competent interaction between Rag-Ragulator and GATOR1 suggested GATOR1 is stabilized by both Rag GTPases in its GAP mode [34]. The overall geometry and the role of the Arg finger in the active GATOR1 complex resembles that of $FLCN^{F118D}$:FNIP2 seen here, as well as that of another longin-domain based GAP, the C9orf72-SMCR8 complex with ARF1 [31]. In comparison to GATOR1, the structure of the AFC revealed that FLCN solely interacts with RagC to stimulate GTP hydrolysis. This finding is consistent with observations that while GATOR1 RagA GAP activity depends on the RagC nucleotide state, the RagC GAP activity of the AFC is independent of the RagA nucleotide state [19].

The past few years have seen a paradigm shift in understanding how the MiTF family of transcription factors is regulated downstream of mTORC1 in a uniquely FLCN and $RagC^{GDP}$-dependent manner [10,11]. Numerous studies corroborate the discovery of a "non-canonical" mTORC1 pathway that specifically regulates MiT-TFE transcription factors [12,21,35-38]. The new knowledge of FLCN structural mechanism provided by the AFC structure can potentially be exploited for therapeutic benefit. The AFC presents itself as an attractive avenue for the long-standing goal of developing substrate selective mTORC1 inhibitors which can upregulate autophagy and lysosomal capacity without toxic consequences for immune function, epithelial renewal and other physiological processes that require mTORC1-dependent anabolic programs. Such inhibitors could be highly specific for MiT-TFE upregulation, in turn promoting autophagy and increasing lysosomal capacity. The latter could be of potential benefit for treating LSDs, while both could potentially benefit neurodegenerative diseases involving intracellular protein aggregates. On the other hand, excessive upregulation of MiT-TFE transcription factors in FLCN loss of function has been shown to cause kidney cystogenesis, and tumorigenesis and induce melanoma [10,35]. Therapeutics that promote FLCN activity, for instance through inhibition of the LFC, might have potential for treating these conditions. An important consideration in therapeutic use of FLCN GAP antagonists would be to avoid the complete shutdown of FLCN activity and so avoid promoting tumorigenesis. The structural insights into the AFC presented here fulfill a key step necessary to leverage FLCN as a drug target and develop selective mTORC1 inhibitors or activators for MiT-TFE-dependent transcriptional upregulation.

Materials and Methods

Cloning and Protein Purification:

The pCAG-GST-SLC38A9^{NT}-FNIP2 fusion construct was designed to contain a ten-residue glycine-serine linker. Codon-optimized DNA coding for SLC38A9^{NT} and the C-terminal linker was amplified by polymerase chain reaction (PCR). DNA containing SLC38A9^{NT} and the C-terminal linker was subcloned into the pCAG-GST vector containing codon-optimized DNA coding for FNIP2. The pCAG-GST-FNIP2 vector was linearized using KpnI and XhoI restriction sites. DNA containing SLC38A9^{NT} and C-terminal linker was inserted N-terminal to FNIP2 by Gibson assembly. HEK293-GNTI cells were transfected with 1mg DNA at 2:1 ratio (FLCN: SLC38A9^{NT}-FNIP2) and 3mg P.E.I per 1L of cells at 2E6 cells/ml. Cells were pelleted at 2000 xg for 20 min at 4 °C and purified as described previously. [11].

Mutants FLCN F10D and FNIP2 V146E were generated by site-directed mutagenesis using KAPA HiFi HotStart ReadyMix (Roche) from pCAG-FLCN and pCAG-GST-FNIP2 respectively. HEK293-GNTI cells were transfected with 1mg DNA at 2:1 ratio (FLCN:FNIP2) 3mg P.E.I per 1L of cells at 2E6 cells/ml. Cells were pelleted at 2000 xg for 20 min at 4 °C and purified as described previously [11].

RagA:RagC^{D181N} and Ragulator were expressed and purified from Sf9 cells via baculovirus infection as described [11]. GATOR1 was expressed and purified from HEK293-GNTI cells as described [11].

Nucleotide Loading:

Purified RagA:RagC^{D181N} heterodimers were diluted 1:10 (v/v) in 1X phosphate-buffered saline (PBS) containing 5mM EDTA and 0.5 mM TCEP for 10 minutes at 25°C. Desired tri-phosphate nucleotides were added at a 10-fold molar excess over Rags and incubated for 30 minutes at 25°C. Next, $MgCl_2$ was added to a final concentration of 20 mM and incubated for 10 minutes at 25°C. Excess nucleotides were removed using a PD-10 desalting column (Cytiva) and wash buffer containing 25 mM HEPES, 130 mM NaCl, 2.5 mM $MgCl_2$, 2 mM EGTA, pH 7.4 and 0.5mM EDTA. To obtain the inactive state for Rag heterodimers, RagA^{GTP}:RagC^{XTP} was incubated with GATOR1 at a 1:100 GAP:Rag molar ratio at 37 °C for 30 minutes.

Active FLCN Complex Assembly:

To assemble the active FLCN complex, RagA:RagC^{D181N} heterodimers in desired nucleotide state were incubated for 1 hour at 25°C with a 1.2x molar excess of Ragulator. FLCN: SLC38A9^{NT}-FNIP2 was added to Rag-Ragulator complex at a 2-fold molar excess followed immediately by the addition of 0.5 M BeF_3. BeF_3 was prepared by mixing $BeSO_4$ and NaF to final concentration of 0.5mM and 5mM respectively and incubating for 20 minutes at 25°C prior to addition to complex. The complex containing Rags, Ragulator, FLCN: SLC38A9^{NT}-FNIP2 and BeF_3 was incubated at 25°C for 30 minutes, then at 4°C for 16 hours. Following incubation, the complex was centrifuged at 17,000 xg for 10 minutes and then loaded onto a Superose 6 (GE Healthcare) column equilibrated in wash buffer containing 25 mM HEPES, 130 mM NaCl, 2.5 mM $MgCl_2$, 2 mM EGTA, pH 7.4 and 0.5mM EDTA. Fractions were analyzed via SDS PAGE and those containing assembled active FLCN complex were collected, concentrated to 1 mg/ml and immediately frozen onto grids for imaging via cryo-EM.

Cryo-EM Grid Preparation and Imaging:

For data collection of active FLCN complex, 3 µl sample (1 mg/ml protein in 25 mM HEPES, 130 mM NaCl, 2.5 mM $MgCl_2$, 2 mM EGTA, 0.5 mM TCEP, pH 7.4) was deposited onto freshly glow-discharged (PELCO easiGlow, 45 s in air at 20 mA and 0.4 mbar) holey carbon grids (C-flat: 2/1-3Cu-T). FEI Virobot Mark IV was used to blot grids for 2 seconds with a blot force of 20 (Whatman 597 filter paper) at 4℃ and 90-100 % humidity and subsequentially plunged into liquid ethane. The Titan Krios G3i microscope equipped with a Gatan Quantum energy filter (slit width 20 eV) and a K3 summit camera at a defocus of -1.0 to -2.0 µm was used to record 4968 movies. Automated image acquisition was performed using SerialEM [39] recording four movies per 2 µm hole with image shift. Image parameters are summarized in Table 1.1

Cryo-EM Data Processing:

The data processing workflow is summarized in Figure S1. In short, Relion-3.1.1 wrapped MotionCor2 program was used to motion-correct gain-corrected movies [40,41]. Motion-corrected micrographs were imported into cryosparc2 v3.3.1 [42]. Patch CTF estimated (multi) was used for CTF determination and two rounds of cryosparc2 blob picker with a diameter range of 150Å-220Å was used to generate 1,165,732 and 1,025,735 particles. Particles were extracted with a box size of 440x440 pixels in cryosparc2. A series of 2D classifications followed by an ab-initio-reconstruction was used to generate three reference maps. The resulting 3D maps were used for consecutive rounds of heterogenous refinement from the initial particle sets following a round of 2D classification to remove obvious 'junk'. The resulting particle sets from round 1 and round 2 of picks were merged and duplicates were removed. The combined particle set contained 177,018 particles. A final round of homogenous refinement resulted in a 3.47Å map at 0.143 FSC. A mask was generated surrounding the interface between RagC and FLCN:FNIP2 using UCSF Chimera and imported into cryosparc2 v3.3.1 where it was lowpass filtered and dilated [43]. The mask was used for subsequent local refinement and resulted in a focused map reaching a resolution of 3.16Å.

Atomic Model Building and Refinement:

The coordinates from RagC were based on GTP-bound crystal structure (3LLU), the coordinates for FLCN and FNIP2 were based off the LFC (6NZD) and all five Ragulator, RagA and $SLC38A9^{NT}$ were based on Rag-Ragulator-$SLC38A9^{NT}$ structure (6WJ2). All coordinates were rigid body fitted separately into the density map using UCSF Chimera. The interface was refined using a map generated from local refinement separately by iterative rounds of Phenix real-space refinement and manual correction in coot [44]. Secondary structure restrains were enabled during real-space refinement. The interface was combined with the remaining components of the complex in coot and refined using iterative round of Phenix real-space refinement and ISOLDE [45]. The final model was validated using MolProbity [46]. Coordinate models can be found in the Protein Data Bank (PDB) with accession code 8DHB. Summarized in Table 2.2.

Tryptophan Fluorescence RagC XTPase Assay:

Tryptophan fluorescence experiments were performed in triplicate according to *Fromm et al.* (*16*).

Cell Culturing:
HEK293T and HEK293A cells were grown in Dulbecco's modified Essential Media (DMEM, Thermo Fisher Scientific) supplemented with 10% (v/v) fetal bovine serum (FBS, VWR), 100 U/ml penicillin and 100 U/ml streptomycin (Life Technologies) and were maintained at 37 °C and 5% CO_2.

Reagents and Antibodies:
The following antibodies were employed in this study: phospho-T389 S6K1 (9234S), S6K1 (2708S), phospho-S65 4EBP (9451S), 4EBP (9644S and 9452S), FLCN (3697), TFE3 (14779), TFEB (37785), V5 (13202), FLAG (14793) from Cell Signaling Technology; LAMP2 (sc-18822) from Santa Cruz Biotechnology.

RNAi
The control pLKO.1-LUC shRNA vector and the pLKO.1-FLCN lentiviral shRNA vector were obtained respectively from Addgene (Plasmid #30324) and the RNAi Consortium (TRCN0000237886). Target sequences are:

shLUC: TCCTAAGGTTAAGTCGCCCTCG.
shFLCN: GATGGAGAAGCTCGCTGATTT

RNAi experiments were prepared as described below.

Lentiviral Infection and Stable Cell Line Generation:
Lentivirus was generated using the PEI method. 5 µg of a lentiviral vector (desired construct in pLJM1-PURO) was combined with 3.75 µg of psPAX2 and 1.25 µg pCMG.2 viral packaging plasmids in 500 µl optiMEM (Thermofisher), mixed, and then combined with 60 µl 1 mg/mL P.E.I. This solution was incubated at room temperature for 30 minutes, then added to a 10 cm dish containing 2 million recently plated 293T cells.

24-48 hours after that, media containing virus was collected from cells, spun down at 1,300 x rpm for 5 minutes to remove cell debris, and sterile filtered. Cells were plated at a density of 200,000-300,000 cells per well in a 6-well plate in the presence of Polybrene transfection reagent (Sigma-Aldrich) and infected with 10% of the total virus from one 10 cm plate of the indicated constructs. 24 hours later, media was supplemented with 200 nM puromycin.

Live Cells Treatments:
HEK293T cells were seeded and allowed to attach overnight in complete DMEM medium. 24h after seeding, cells were rinsed with PBS and incubated in amino acid-free RPMI for 2 hours (-AA) and restimulated with amino acids for 15 minutes or overnight with complete media (+AA). Amino acid solutions were prepared from powders and the final concentration of amino acids in the media was the same as in commercial RPMI. Torin1 treatment was performed for 4 hours in complete media.

Western Blotting:
HEK-293T cells were plated in a 6-well plate at 1,000,000 cells/well in complete media. After the treatments (see above), the medium was removed and 150 µl of Lysis Buffer (40 mM HEPES [pH 7.4], 4 mM EDTA, 1% Triton X-100, 10 mM sodium β-glycerol phosphate, 10 mM sodium pyrophosphate, and one tablet of EDTA-free protease inhibitors per 50 ml) were added to each

well. Total extracts were collected from each well after scraping, rotated over-end for 10 minutes, and then centrifuged at 13,000 × g for 10 min at 4 °C. Samples were normalized to a total concentration of 1 mg/mL protein and combined with protein sample buffer, then boiled for 5 minutes at 95 °C.
Ten micrograms of protein were loaded into each lane of Tris-Glycine precast gels (Thermo Fisher Scientific). PVDF membranes were blocked in 5% skim milk dissolved in Tris-buffered saline with Tween 20 for 1 hour at RT. Membranes were incubated overnight at 4 °C with the indicated primary antibodies, all diluted in 5% skim milk prepared in TBST. Secondary antibodies conjugated to horseradish peroxidase were used for protein detection. β-Actin was used as loading controls.

Immunofluorescence:
HEK-293T cells were plated on fibronectin-coated glass coverslips in 12-well plates, at 300,000 cells/well. The following day, cells were subjected to amino acid depletion/or complete media restimulation (see above) and then fixed in 4% paraformaldehyde (in PBS) for 15 min at RT. The coverslips were rinsed twice with PBS and cells were permeabilized with 0.1% (w/v) saponin in PBS for 10 min. After rinsing twice with PBS, the slides were incubated overnight at 4°C with the indicated primary antibody in 5% normal donkey serum, then rinsed with PBS, and incubated with fluorophore-conjugated secondary antibodies produced in goat or donkey (Life Technologies, diluted 1:500 in 5% normal donkey serum) for 45 min at room temperature in the dark. Coverslips were then washed three times in PBS and mounted on glass slides using Vectashield Antifade Mounting Medium (Vector Laboratories) containing DAPI stain. All images were collected on a Nikon Ti-E inverted microscope (Nikon Instruments, Melville, NY) equipped with a Plan Apo 60x oil objective. Images were acquired using a Zyla 5.5 sCMOS camera (Andor Technology), using iQ3 acquisition software (Andor Technology).

Live Imaging:
HEK293A cells were plated onto glass-bottom 35mm culture dishes and transfected with the Indicated FLCN-FLAG-eGFP, HA-FNIP2, and LAMP2-RFP-FLAG constructs. The following day, cells were subjected to amino acid depletion/or amino acids restimulation (see above). Starved cells were transferred to imaging buffer (10 mM HEPES, 136 mM NaCl, 2.5 mM KCl, 2 mM CaCl2, 1.2 mM MgCl2, pH 7.4), while restimulated cells were transferred to imaging buffer supplemented with amino acids, 5mM glucose, and 1% dialyzed FBS (+AA) and imaged using the Nikon Ti-E inverted microscope (Nikon Instruments, Melville, NY) as explained above.

Quantitation of Lysosomal Enrichment Score:
For immunofluorescence and live imaging datasets, a home-built MATLAB (Mathworks) script was used to determine the lysosomal enrichment of both V5-FNIP2 and FLAG-FLCN stains as previously described (*9*). Briefly, a single cell was manually selected in the lamp channel and its nucleus was excluded from further analysis. Then, a mask was created in the LAMP2 channel to segment cellular pixels into LAMP2 (lysosomal) or non-LAMP2 (cytosolic) pixels. This mask was then applied to the non-LAMP2 channel. The Lysosomal Enrichment Score was determined by dividing the average intensity of pixels in the lysosomal region by the average intensity of the pixels in the cytosolic region. For each condition, at least twenty cells were analyzed from different multi-channel images.

Quantitation of Cytoplasmic:Nuclear TFE3 Ratio:
To measure TFE3 nuclear localization, the Pearson's coefficient of at least 10 different fields was measured using the ImageJ plugin JACoP [47].

Acknowledgements: We thank Zhicheng Cui and Aaron Joiner for discussions, Andrea Ballabio for comments on the manuscript, Rick Hooy for workstation support and Dan Toso and Jonathan Remis for electron microscope support.

Funding: This work was supported by the NIH grants R01 GM111730 (J.H.H.) and R01 GM130995 (R.Z.), a National Science Foundation Graduate Research Fellowship (R.M.J.), and an Italian Cancer Research Association (AIRC) postdoctoral fellowship (R.P.).

Author contributions: R.M.J. and J.H.H. conceived and designed research, R.M.J. and R.P. carried out research, A.L.Y. and S.F. trained R.M.J. in cryo-EM, R.Z. and J.H.H. supervised research, R.M.J. and J.H.H. wrote the first draft, and all authors edited the manuscript.

Competing interests: J.H.H. is a co-founder and shareholder of Casma Therapeutics and receives research funding from Casma Therapeutics, Genentech, and Hoffmann-La Roche. R.Z. is a cofounder and shareholder of Frontier Medicines and receives research funding from Genentech. The authors declare no other competing interests.

Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Coordinates and density are deposited in the RCSB (PDB ID 8DHB) and EMDB (EMD-27435), respectively. The fusion construct used to generate the purified AFC is deposited in Addgene.org (189869). The DNA constructs used for FNIP2 localization will be provided on request pending a completed material transfer agreement. Requests for the FNIP2 DNA constructs should be submitted to Roberto Zoncu at rzoncu@berkeley.edu.

Figures

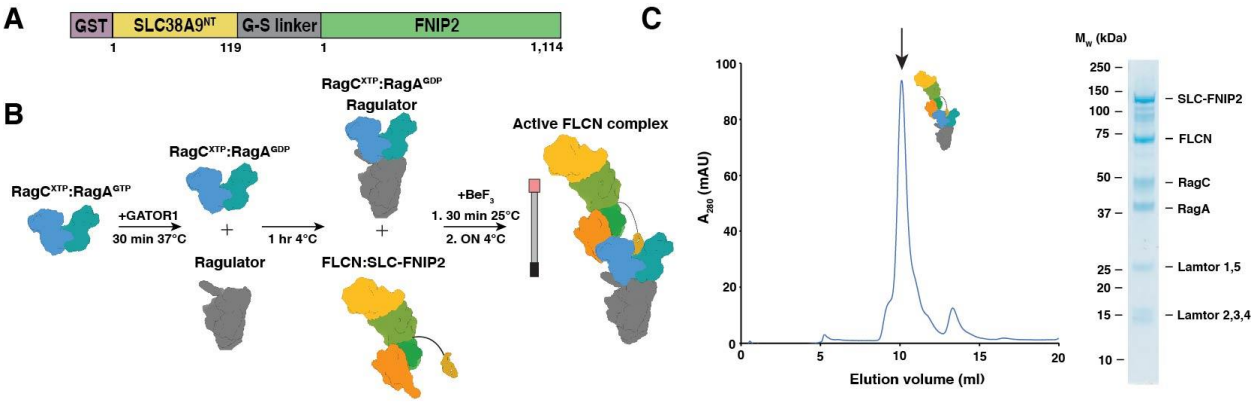


Figure 2.1 Active FLCN complex formation. (A) Construct design for SLC38A9^{NT}-FNIP2 fusion construct. (B) Complex formation workflow. (C) Size exclusion profile for active FLCN complex. Fraction containing all components of complex is labelled with a black arrow and visualized on adjacent gel.

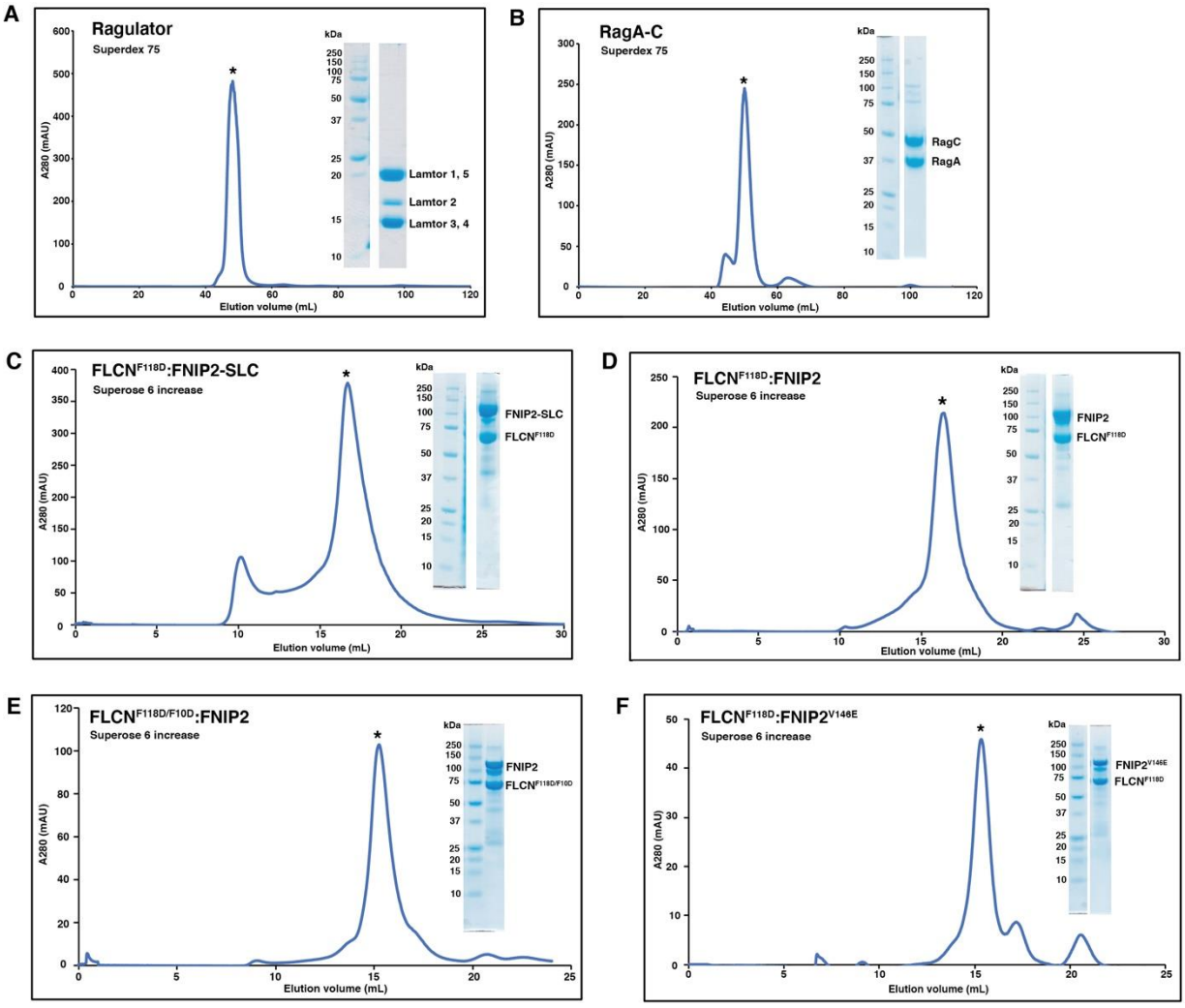


Figure 2.2 Purification of individual active FLCN components. (A) Superdex 75 SEC elution profile for Ragulator and Coomassie blue stained SDS-PAGE analysis of protein sample collected from peak indicated with asterisk. (B) Superdex 75 SEC elution profile for Rag GTPases and Coomassie blue stained SDS-PAGE analysis of protein sample collected from peak indicated with asterisk. (C) Superose 6 increase SEC elution profile for FLCN^{F118D}:FNIP2-SLC fusion and Coomassie blue stained SDS-PAGE analysis of protein sample collected from peak indicated with asterisk. (D) Superose 6 increase SEC elution profile for FLCN^{F118D}:FNIP2 and Coomassie blue stained SDS-PAGE analysis of protein sample collected from peak indicated with asterisk. (E) Superose 6 increase SEC elution profile for FLCN$^{F118D/F10D}$:FNIP2 and Coomassie blue stained SDS-PAGE analysis of protein sample collected from peak indicated with asterisk. (F) Superose 6 increase SEC elution profile for FLCN^{F118D}:FNIP2^{V146E} and Coomassie blue stained SDS-PAGE analysis of protein sample collected from peak indicated with asterisk.

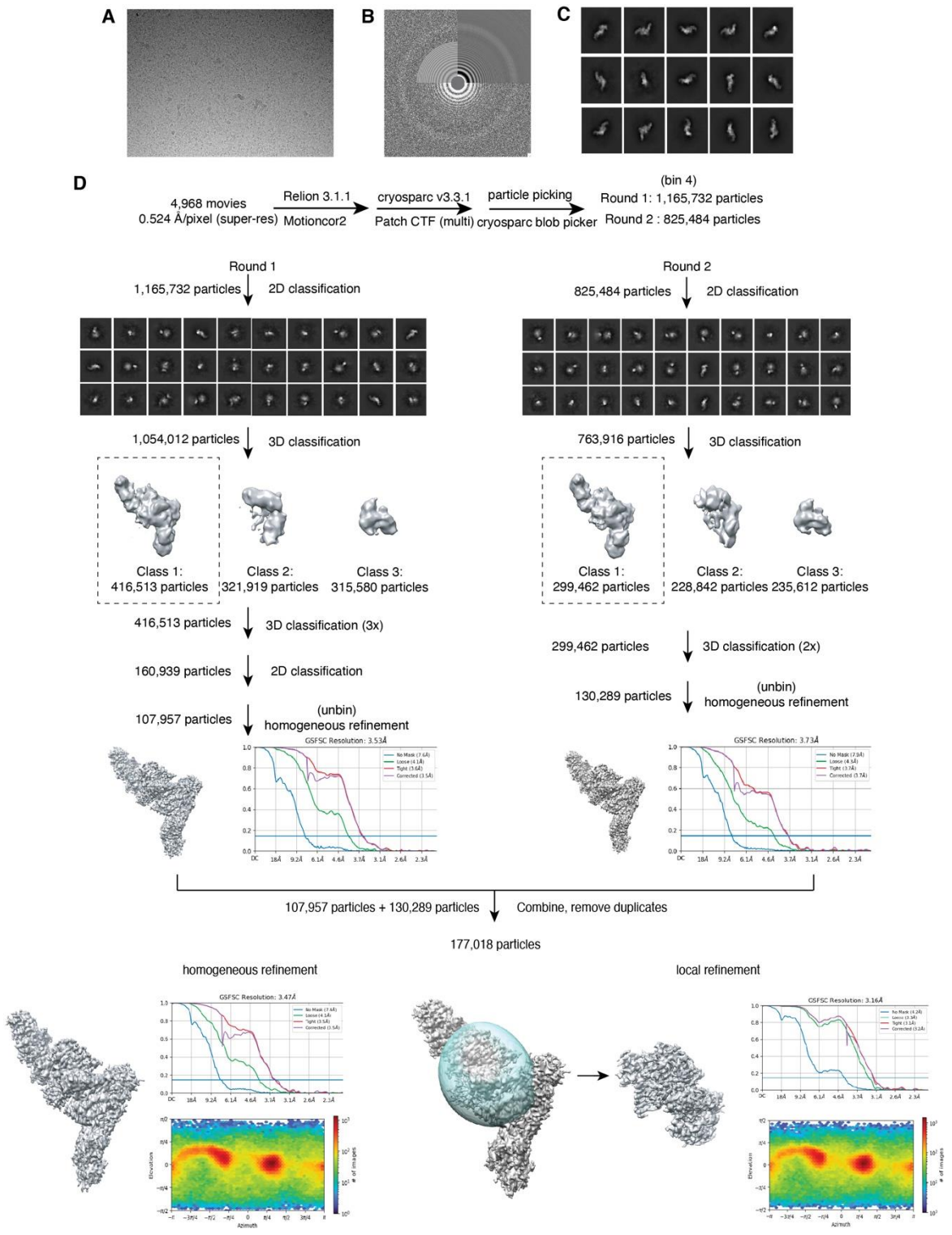


Figure 2.3 Cryo-EM data processing pipeline. (A) Representative cryo-EM micrograph (B) Power spectrum and CTF estimation of micrograph shown in (A). Exemplary 2D class averages for the active FLCN complex. (D) Data processing pipeline for final map determination.

Table 2.1 Cryo-EM data acquisition and image processing.

	Active FLCN complex
Data acquisition	
Microscope	Titan Krios
Voltage (kV)	300
Camera	Quantum-K3 Summit
Magnification	165,000
Pixel size (Å)	0.524 (super-resolution)
Cumulative exposure ($e^-/Å^2$)	50
Energy filter slit width (eV)	20 eV
Defocus range (μm)	-1.0 to -2.0
Automation software	SerialEM
Exposure navigation	Image Shift
Number of movies	4968
Image processing	
Initial picked particles (no.)	Round 1: 1,165,732 Round 2: 825,484
Final refined particles (no.)	177,018
Map resolution (Å)	Overall: 3.53 Interface: 3.16
FSC threshold	0.143

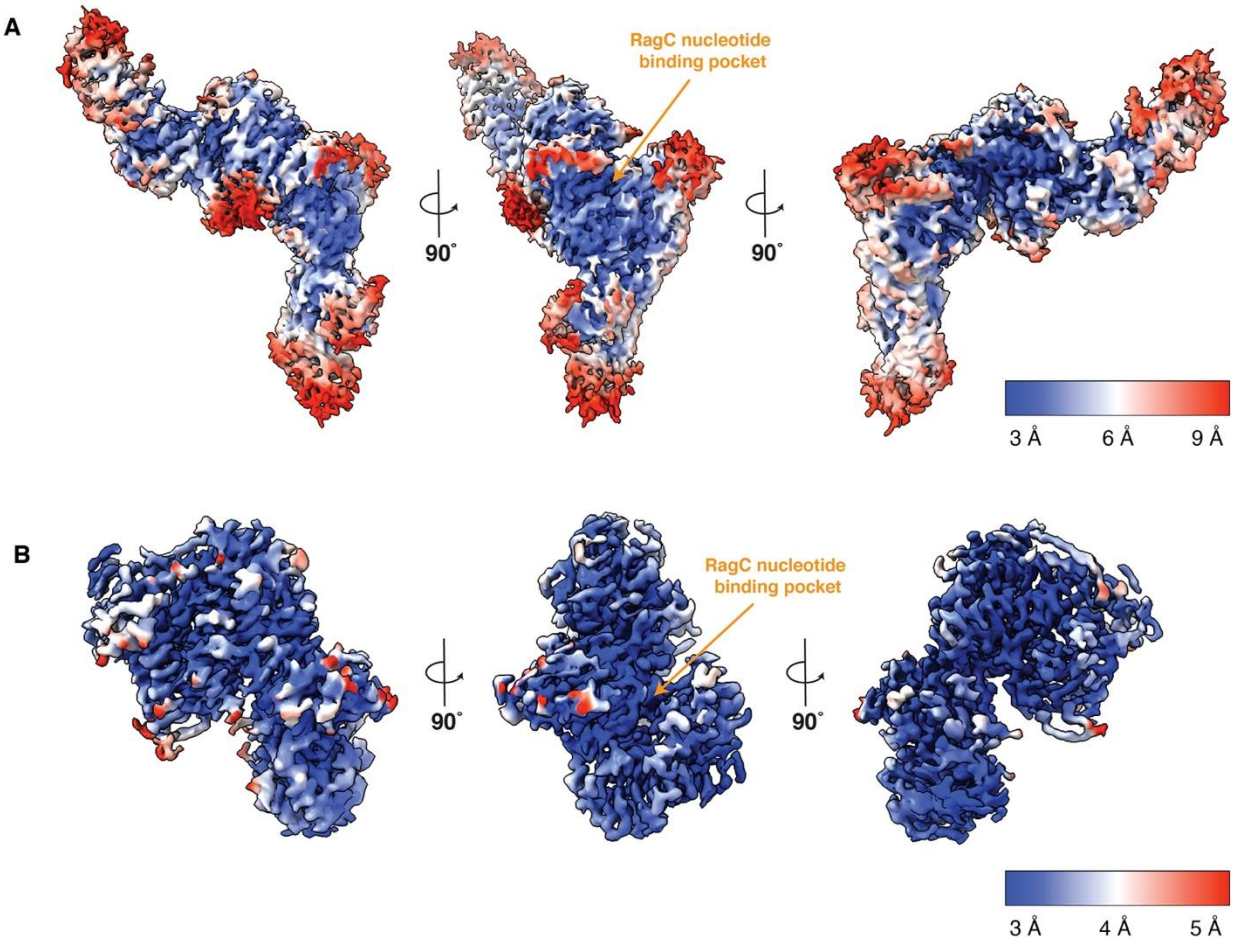


Figure 2.4 AFC local resolution estimation. (A) Overall cryo-EM map for AFC colored according to local resolution. Resolution ranges from 3 Å to 9 Å. (B) Cryo-EM map of RagC – $FLCN^{login}$:$FNIP2^{longin}$ interface generated from focused refinement colored according to local resolution. Resolution ranges from 3 Å to 6 Å.

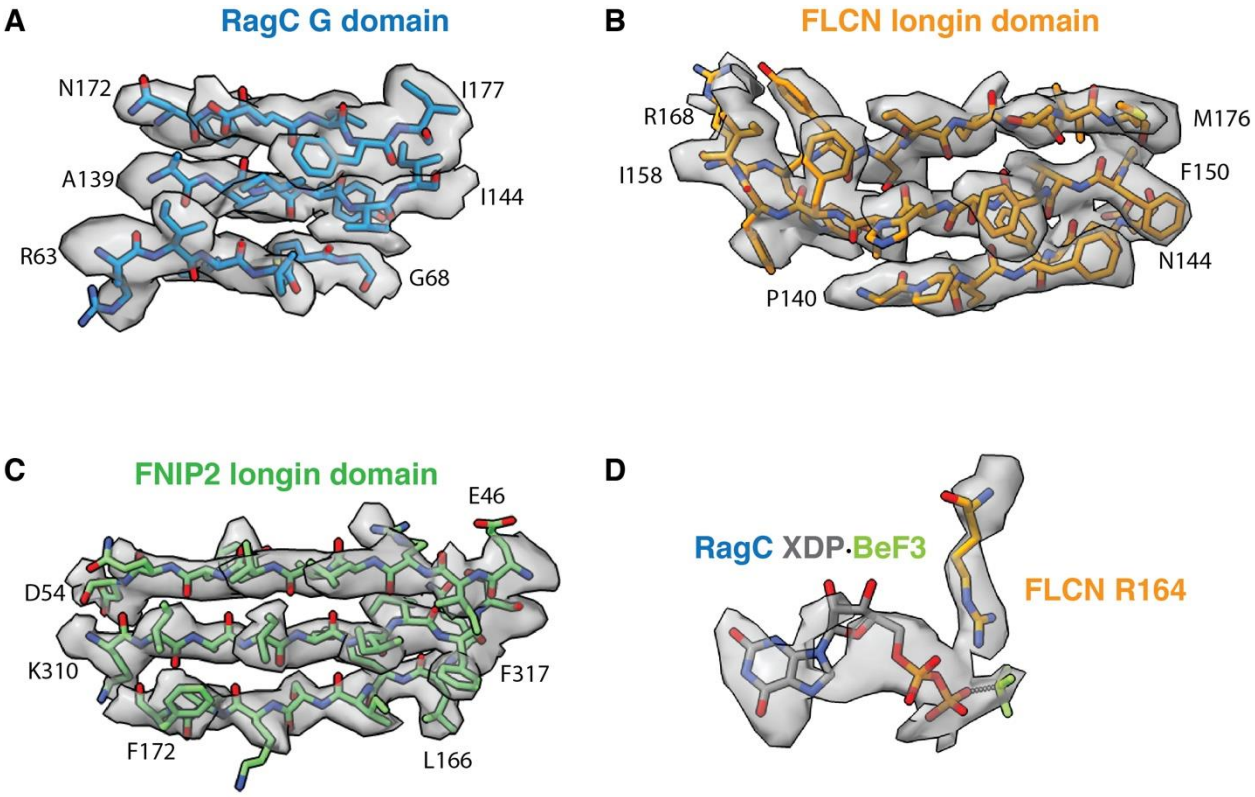


Figure 2.5 AFC map-model fit. Representative refined coordinate model fit in cryo-EM density for (A) RagC (B) FLCN (C) FNIP2 (D) RagC nucleotide and FLCN arginine finger.

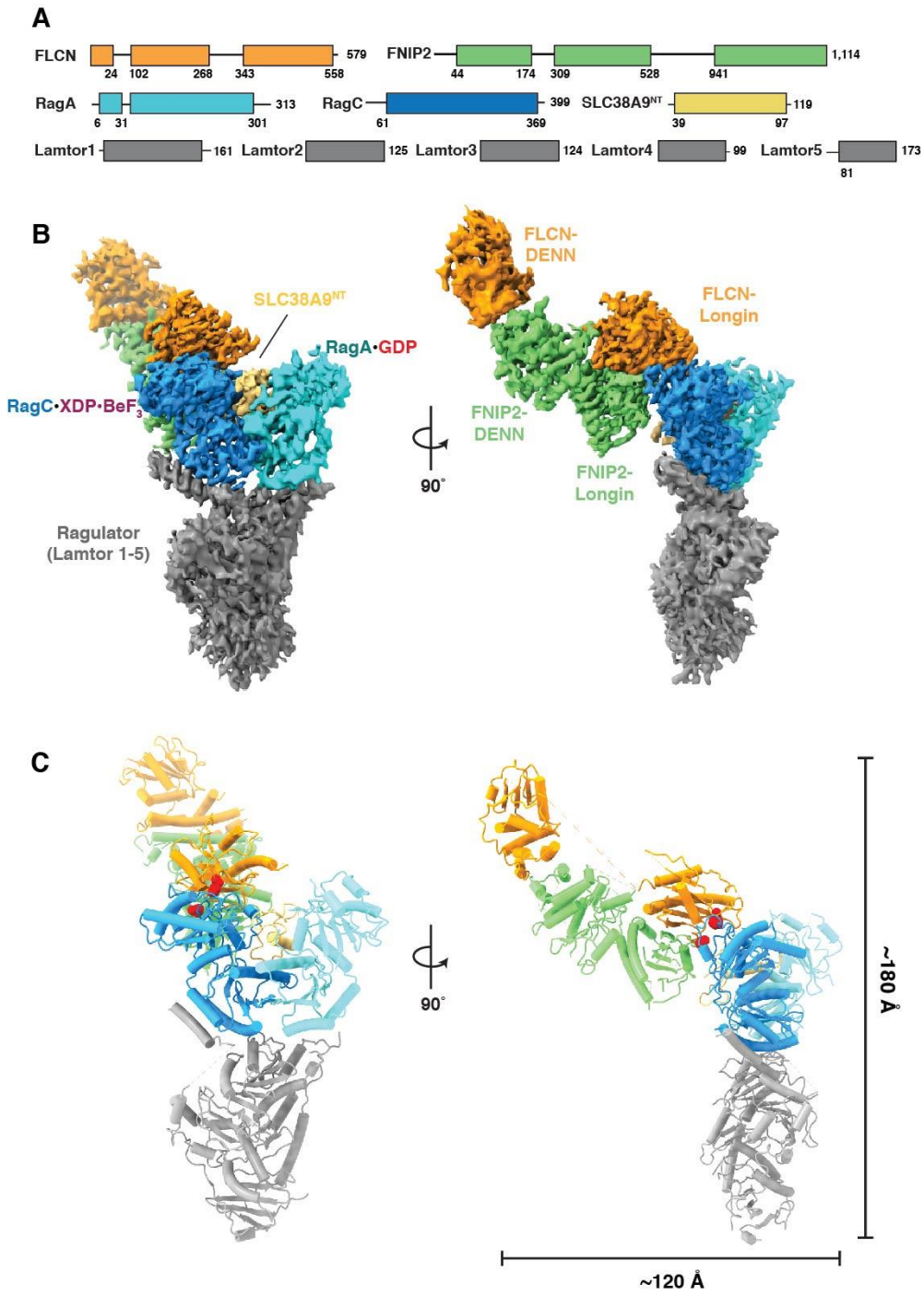


Figure 2.6 Cryo-EM structure of active FLCN complex. (A) Domain organization for components of active FLCN complex. (B) Composite map for complex. Focused map of interface was combined with overall map to produce composite map containing highest resolution information for each component. (C) Reconstructed model for AFC with overall dimensions of the side view of the complex. Dotted regions in model represent unstructured regions in FLCN connecting the two domains. Red spheres indicated locations of mutations in FLCN F10D and FNIP2 V146E.

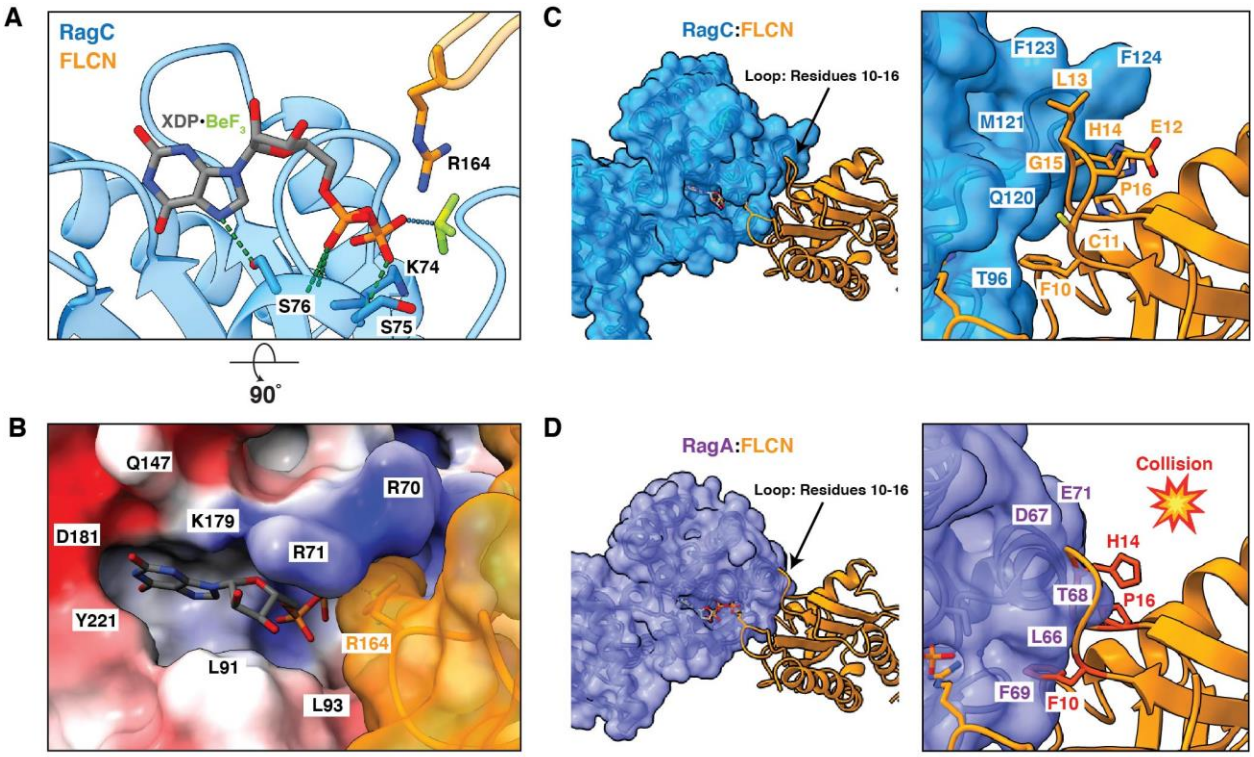


Figure 2.7 Specificity of FLCN-RagC GAP interaction in AFC. (A) Nucleotide binding pocket of RagC (blue) in AFC containing XDP-BeF_3 nucleotide analog. Location of FLCN (orange) arg-finger indicated. Hydrogen bonds between nucleotide and RagC residues (blue sticks) in surrounding pocket represented by dashed lines (green). (B) Surface representation of RagC binding site with FLCN arg-finger (orange). RagC colored based on electrostatic potential. (C) RagC (blue) FLCN (orange) interface in AFC. Loop residues 10-16 in FLCN indicated as sticks (black labels). (D) RagA (purple) FLCN interface modelled with GTP-bound RagA (PDB:6S6D). Residues in FLCN loop that clash with RagA are shown in red.

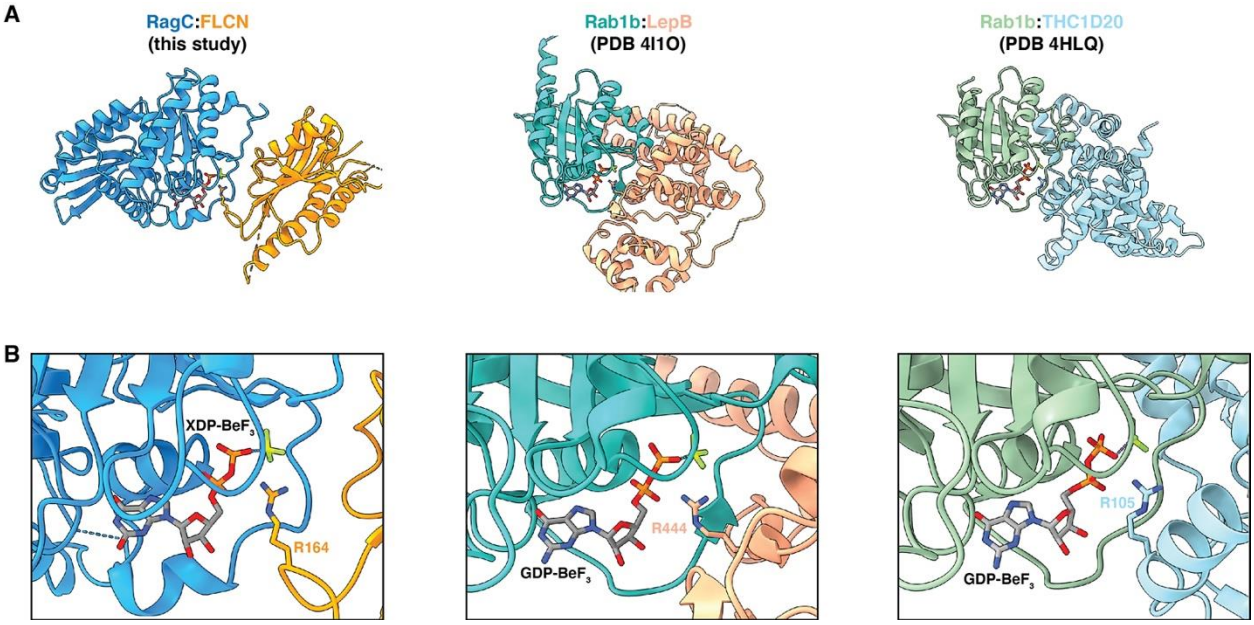


Figure 2.8 Structural comparison between RagC:FLCN and other small GTPases:GAP. (A) Rab1b:LepB and Rab1b:THC1D20. (A) Structural comparison of RagC:FLCN (AFC structure), Rab1b:LepB (PDB 4I1O) and Rab1b:THC1D20 (PDB 4HLQ). Structures are shown in the same orientation. (B) Close-up of the nucleotide binding pocket with catalytic arginine-finger residue positioned in active site.

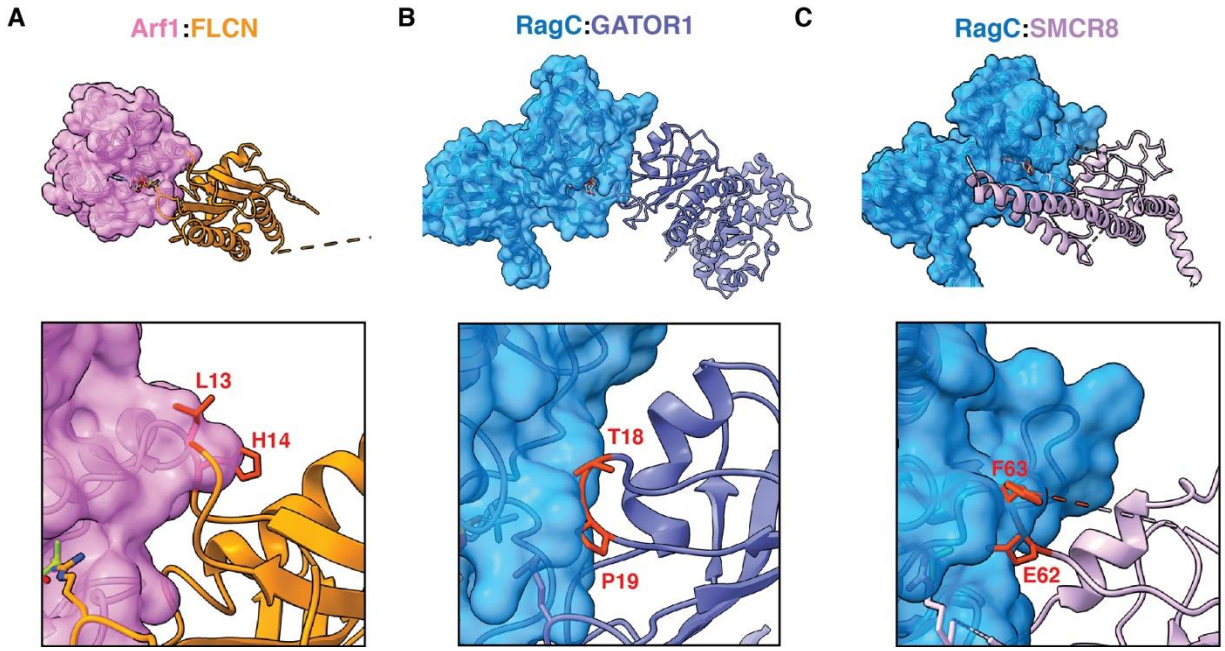


Figure 2.9 Structural investigation of RagC:SMCR8 and Arf1:FLCN interface. (A) Model of Arf1 (pink) and FLCN (orange) interface generated by aligning Arf1 from Arf1:SMCR8 structure (PDB: 7MGE) with RagC in AFC structure (*27*). Residues in FLCN at interface that clash are indicated and labelled in red. (B) Model of RagC (pink) and SMCR8 (orange) interface generated by aligning SMCR8 from Arf1:SMCR8 structure (PDB: 7MGE) with FLCN in AFC structure (*27*). Residues in SMCR8 at interface that clash are indicated and labelled in red.

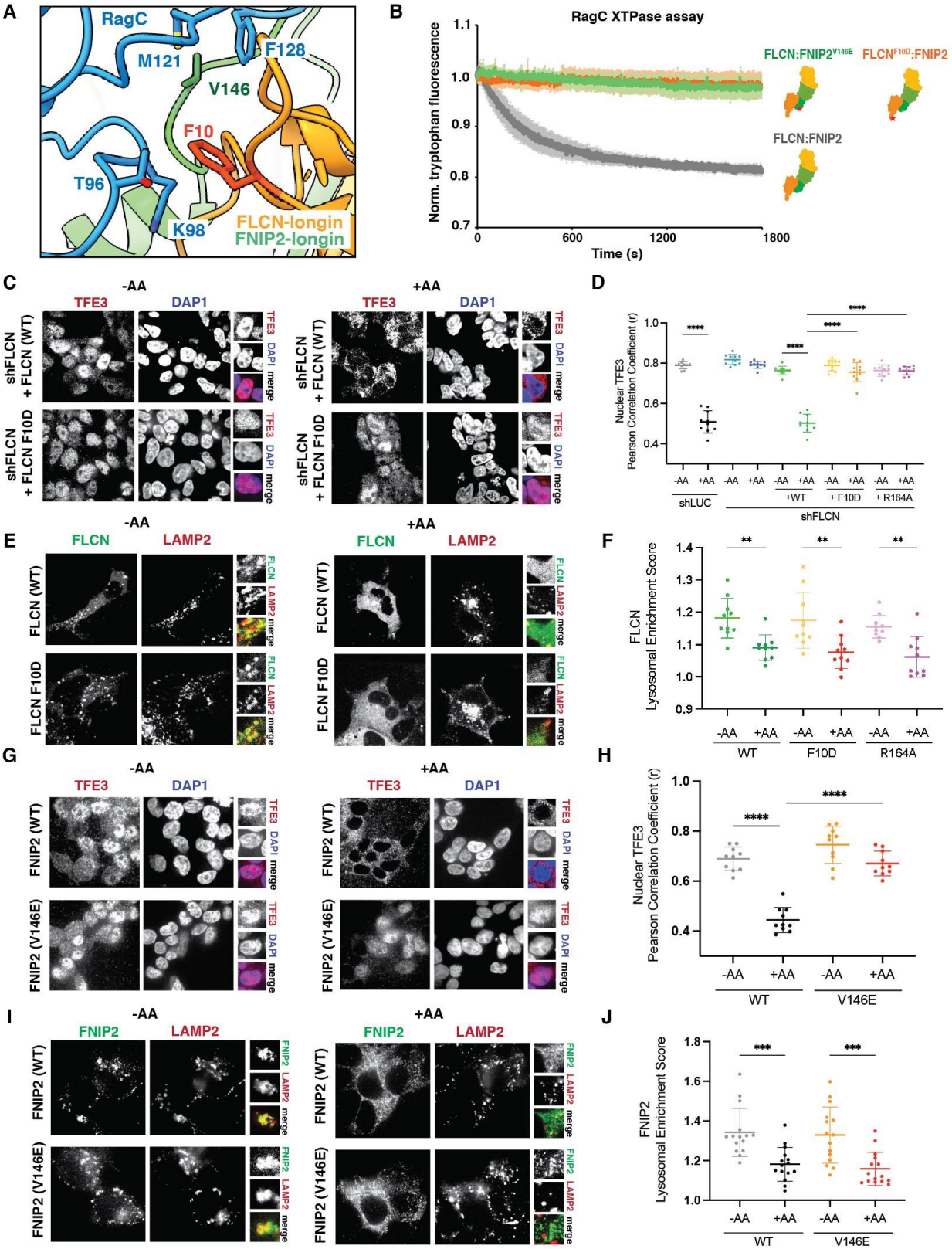

A
RagC
M121
F128
V146
F10
T96
K98
FLCN-longin
FNIP2-longin
B
RagC XTPase assay
Norm. tryptophan fluorescence
Time (s)
FLCN:FNIP2V146E
FLCNF10D:FNIP2
FLCN:FNIP2
C
-AA
+AA
TFE3
DAP1
shFLCN + FLCN (WT)
shFLCN + FLCN F10D
D
Nuclear TFE3 Pearson Correlation Coefficient (r)
shLUC
shFLCN
+WT
+ F10D
+ R164A
E
FLCN
LAMP2
FLCN (WT)
FLCN F10D
F
FLCN Lysosomal Enrichment Score
WT
F10D
R164A
G
FNIP2
FNIP2 (WT)
FNIP2 (V146E)
H
V146E
I
J
FNIP2 Lysosomal Enrichment Score

Figure 2.10 Hydrophobic residues in the FLCN:RagC interface necessary for GAP activity. (A) RagC (blue) and FLCN (orange):FNIP2 (green) interface in AFC. FLCN:FNIP2 residues mutated at interface indicated (red). (B) Tryptophan fluorescence XTPase assay with FLCN:FNIP2 mutants. $FLCN^{WT}$:FNIP2 (gray), $FLCN^{F10D}$:FNIP2 (orange), or FLCN:$FNIP2^{V146E}$ (green) was incubated with $RagA^{GDP}$:$RagC^{XTP}$. Plotted are means ± SEM. n=3 replicates. (C) Immunofluorescence images of HEK293T cells stably expressing shRNAs targeting the indicated genes and FLCN rescue constructs. Cells were starved for amino acids (-AA) or restimulated with complete DMEM (+AA). (D) Quantification of TFE3 nuclear localization for immunofluorescence images and the positive control $FLCN^{164}$ identified previously [11] shown in Figure 2.12 (means ± SD, n=10 fields). ****p<0.0001. (E) Fluorescence images of HEK293A cells expressing the indicated FLCN:eGFP construct along with HA-FNIP2 and the lysosomal marker LAMP2:RFP. Cells were starved for amino acids (-AA) or starved and then restimulated with amino acids for 15 min (+AA). (F) Quantification of the lysosomal localization of over-expressed FLCN for immunofluorescence images (means ± SD, n=15 cells) **p<0.01. (G) Immunofluorescence images of HEK293T cells stably over-expressing the $FNIP2^{WT}$-eGFP-V5 or $FNIP2^{V146E}$-eGFP-V5 constructs. Cells were starved for amino acids (-AA) or maintained in complete media (+AA). (H) Quantification of TFE3 nuclear localization for immunofluorescence images (means ± SD, n=10 fields) ****p<0.0001. (I) Immunofluorescence images of HEK293T cells stably over-expressing the indicated FNIP2-eGFP-V5 constructs. LAMP2 was used as lysosomal marker. Cells were starved for amino acids (-AA) or maintained in complete media (+AA). (J) Quantification of the lysosomal localization of over-expressed FNIP2 for immunofluorescence images (means ± SD, n=15 cells) ***p<0.001.

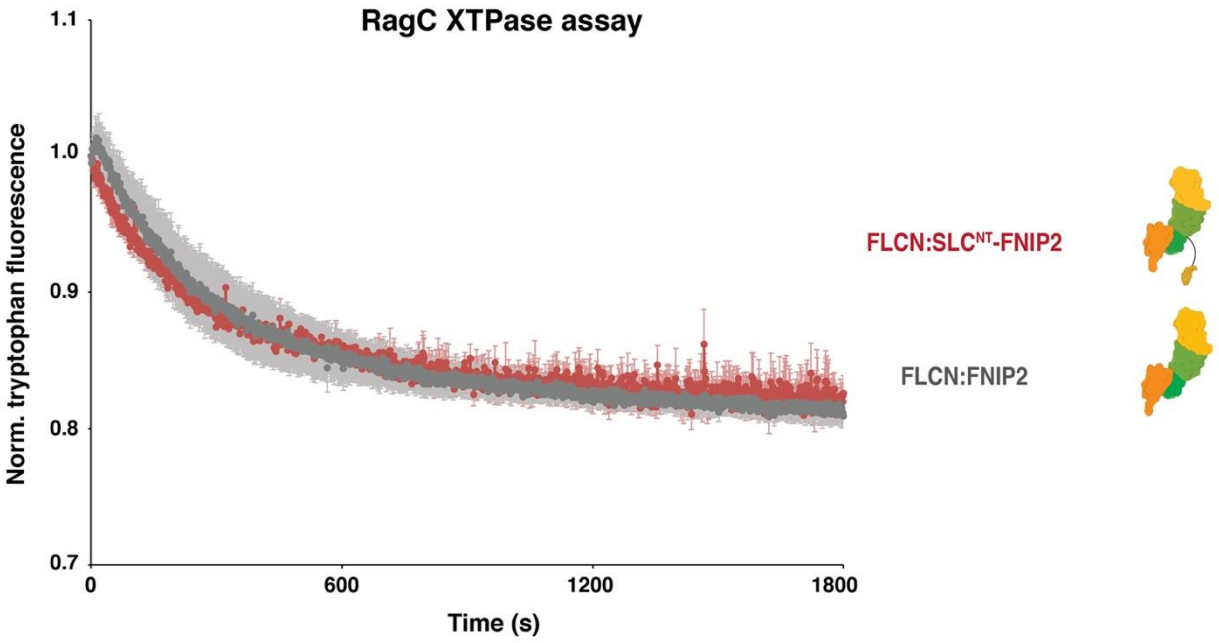


Figure 2.11 GAP activity of FLCN:SLC-FNIP2. Tryptophan fluorescence XTPase assay with FLCN:SLC-FNIP2 fusion construct. $FLCN^{F118D}$: FNIP2 (gray), $FLCN^{F118D}$: SLC-FNIP2 (red), was incubated with $RagA^{GDP}$:$RagC^{XTP}$. Plotted are means ± SEM. n=3 replicates.

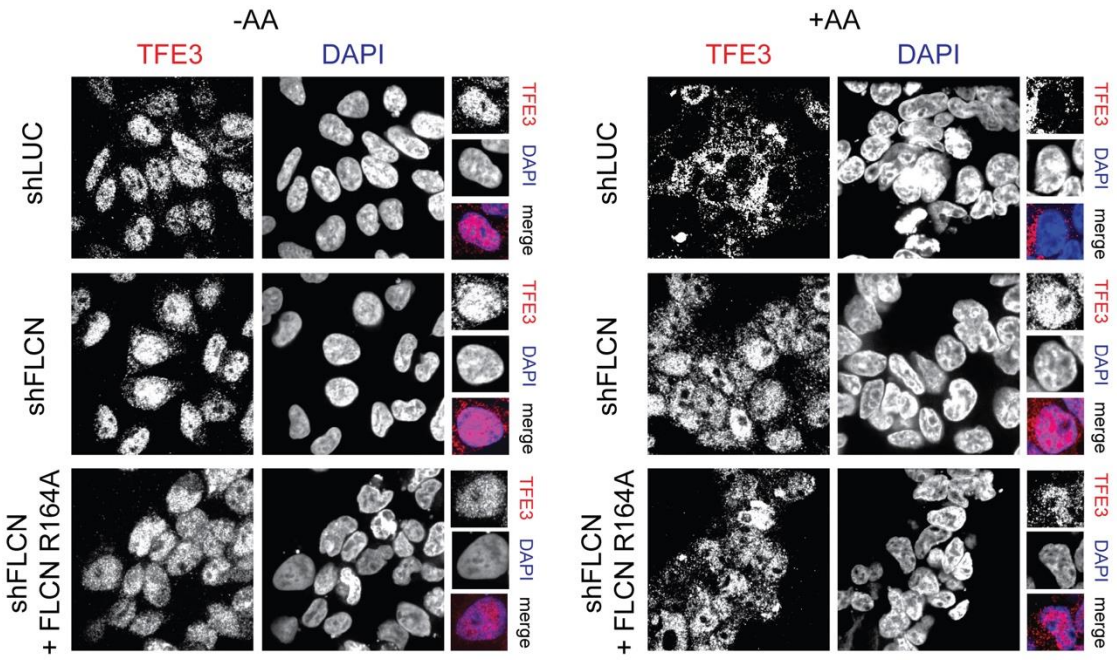


Figure 2.12 Immunofluorescence for FLCNR164. Immunofluorescence images of human embryonic kidney 293T (HEK293T) cells stably expressing the indicated genes shRNA and FLCN rescue constructs. Cells were starved for 2 hours for amino acids (-AA) or restimulated with complete DMEM for 2 hours (+AA).

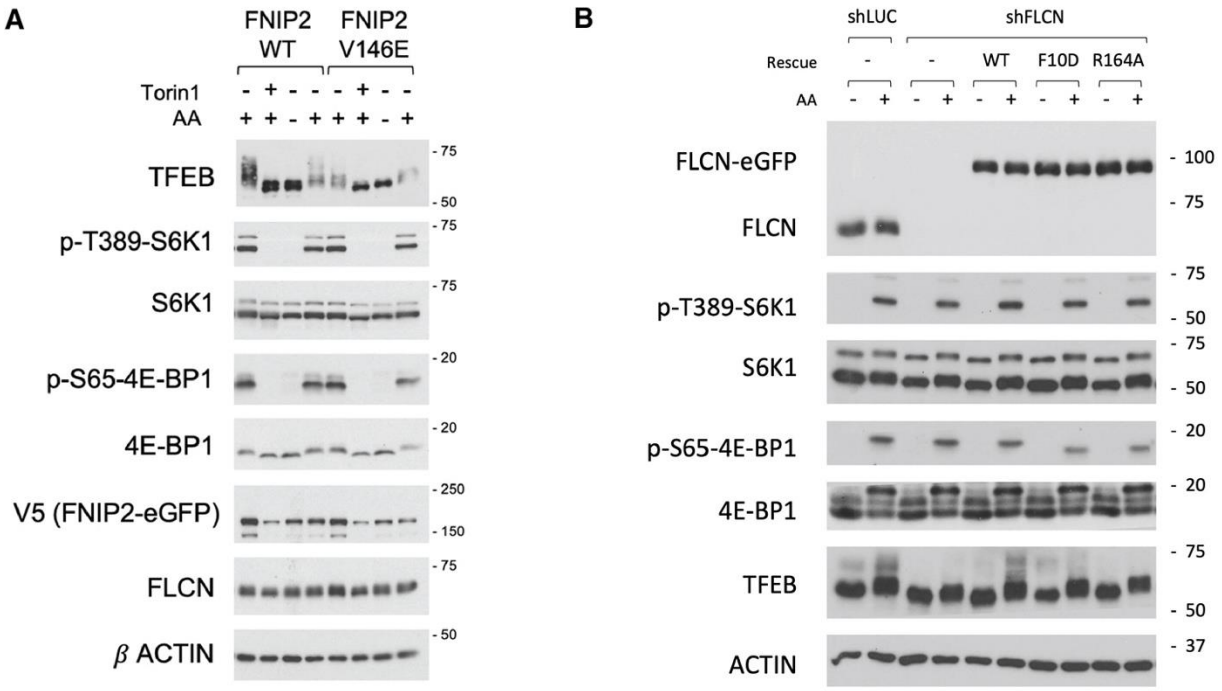


Figure 2.13 Immunoblot analysis for mTORC1 targets. (A) HEK-293T cells stably expressing $FNIP2^{WT}$ or $FNIP2^{V146E}$ were either untreated or treated with Torin1, or starved of amino acids for 2 hours, or starved and then restimulated with amino acids for 15 minutes. Cells were lysed, followed by immunoblotting for the indicated proteins and phospho-proteins. (B) HEK-293T cells stably expressing shRNAs targeting the indicated genes were starved of amino acids for 2 hours, or starved and then restimulated with amino acids for 15 minutes. Cells were lysed, followed by immunoblotting for the indicated proteins and phospho-proteins.

Table 2.2 AFC coordinate model refinement and assembly.

PDB access code	8DHB
EMDB	EMD-27435
Refinement	
Software	Phenix 1.19, ISOLDE
Refinement target (Å)	3.2 (interface) 3.5 (overall)
Non-hydrogen atoms	15171
Residues	1898
Ligands	GDP (1) XDP-BeF^-_3 (1)
RagA, Ragulator, $SLC38A9^{NT}$ reference PDB	6WJ2
FLCN:FNIP2 reference PDB	6NZD
RagC reference PDB	3LLU
Map-model statistics	
R.M.S deviations	
Bond lengths (Å)	0.006
Bonds angles (Å)	1.037
Validation	
Molprobitity	1.93
Clash score	8.98
Rotamer ouliers (%)	0
Cβ outliers (%)	0
CaBLAM outliers (%)	3.95
Ramachandran	
Favored (%)	92.97
Allowed (%)	6.7
Outlier (%)	0.33
Final model composition	
Number of chains	10
Number of Residues	1898
Ligands	2
B-factors	
Protein (min/max/average)	0.27/55.15/27.92
Ligands	10.59

References

1. Saxton, R.A., and Sabatini, D.M. (2017). mTOR Signaling in Growth, Metabolism, and Disease. Cell *168*, 960-976. 10.1016/j.cell.2017.02.004.
2. Lawrence, R.E., and Zoncu, R. (2019). The lysosome as a cellular centre for signalling, metabolism and quality control. Nat Cell Biol *21*, 133-142. 10.1038/s41556-018-0244-7.
3. Schalm, S.S., Fingar, D.C., Sabatini, D.M., and Blenis, J. (2003). TOS motif-mediated raptor binding regulates 4E-BP1 multisite phosphorylation and function. Curr Biol *13*, 797-806. 10.1016/s0960-9822(03)00329-4.
4. Nojima, H., Tokunaga, C., Eguchi, S., Oshiro, N., Hidayat, S., Yoshino, K., Hara, K., Tanaka, N., Avruch, J., and Yonezawa, K. (2003). The mammalian target of rapamycin (mTOR) partner, raptor, binds the mTOR substrates p70 S6 kinase and 4E-BP1 through their TOR signaling (TOS) motif. J Biol Chem *278*, 15461-15464. 10.1074/jbc.C200665200.
5. Settembre, C., Di Malta, C., Polito, V.A., Garcia Arencibia, M., Vetrini, F., Erdin, S., Erdin, S.U., Huynh, T., Medina, D., Colella, P., et al. (2011). TFEB links autophagy to lysosomal biogenesis. Science *332*, 1429-1433. 10.1126/science.1204592.
6. Sardiello, M., Palmieri, M., di Ronza, A., Medina, D.L., Valenza, M., Gennarino, V.A., Di Malta, C., Donaudy, F., Embrione, V., Polishchuk, R.S., et al. (2009). A gene network regulating lysosomal biogenesis and function. Science *325*, 473-477. 10.1126/science.1174447.
7. Martina, J.A., Chen, Y., Gucek, M., and Puertollano, R. (2012). MTORC1 functions as a transcriptional regulator of autophagy by preventing nuclear transport of TFEB. Autophagy *8*, 903-914. 10.4161/auto.19653.
8. Settembre, C., Zoncu, R., Medina, D.L., Vetrini, F., Erdin, S., Erdin, S., Huynh, T., Ferron, M., Karsenty, G., Vellard, M.C., et al. (2012). A lysosome-to-nucleus signalling mechanism senses and regulates the lysosome via mTOR and TFEB. EMBO J *31*, 1095-1108. 10.1038/emboj.2012.32.
9. Roczniak-Ferguson, A., Petit, C.S., Froehlich, F., Qian, S., Ky, J., Angarola, B., Walther, T.C., and Ferguson, S.M. (2012). The transcription factor TFEB links mTORC1 signaling to transcriptional control of lysosome homeostasis. Sci Signal *5*, ra42. 10.1126/scisignal.2002790.
10. Napolitano, G., Di Malta, C., Esposito, A., de Araujo, M.E.G., Pece, S., Bertalot, G., Matarese, M., Benedetti, V., Zampelli, A., Stasyk, T., et al. (2020). A substrate-specific mTORC1 pathway underlies Birt-Hogg-Dubé syndrome. Nature *585*, 597-602. 10.1038/s41586-020-2444-0.
11. Lawrence, R.E., Fromm, S.A., Fu, Y., Yokom, A.L., Kim, D.J., Thelen, A.M., Young, L.N., Lim, C.Y., Samelson, A.J., Hurley, J.H., and Zoncu, R. (2019). Structural mechanism of a Rag GTPase activation checkpoint by the lysosomal folliculin complex. Science *366*, 971-977. 10.1126/science.aax0364.
12. Wada, S., Neinast, M., Jang, C., Ibrahim, Y.H., Lee, G., Babu, A., Li, J., Hoshino, A., Rowe, G.C., Rhee, J., et al. (2016). The tumor suppressor FLCN mediates an alternate mTOR pathway to regulate browning of adipose tissue. Genes Dev *30*, 2551-2564. 10.1101/gad.287953.116.
13. Sekiguchi, T., Hirose, E., Nakashima, N., Ii, M., and Nishimoto, T. (2001). Novel G proteins, Rag C and Rag D, interact with GTP-binding proteins, Rag A and Rag B. J Biol Chem *276*, 7246-7257. 10.1074/jbc.M004389200.

14. Sancak, Y., Bar-Peled, L., Zoncu, R., Markhard, A.L., Nada, S., and Sabatini, D.M. (2010). Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. Cell *141*, 290-303. 10.1016/j.cell.2010.02.024.
15. Shen, K., Choe, A., and Sabatini, D.M. (2017). Intersubunit Crosstalk in the Rag GTPase Heterodimer Enables mTORC1 to Respond Rapidly to Amino Acid Availability. Mol Cell *68*, 552-565 e558. 10.1016/j.molcel.2017.09.026.
16. Tsun, Z.Y., Bar-Peled, L., Chantranupong, L., Zoncu, R., Wang, T., Kim, C., Spooner, E., and Sabatini, D.M. (2013). The folliculin tumor suppressor is a GAP for the RagC/D GTPases that signal amino acid levels to mTORC1. Mol Cell *52*, 495-505. 10.1016/j.molcel.2013.09.016.
17. Bar-Peled, L., Chantranupong, L., Cherniack, A.D., Chen, W.W., Ottina, K.A., Grabiner, B.C., Spear, E.D., Carter, S.L., Meyerson, M., and Sabatini, D.M. (2013). A Tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. Science *340*, 1100-1106. 10.1126/science.1232044.
18. Wang, S., Tsun, Z.Y., Wolfson, R.L., Shen, K., Wyant, G.A., Plovanich, M.E., Yuan, E.D., Jones, T.D., Chantranupong, L., Comb, W., et al. (2015). Metabolism. Lysosomal amino acid transporter SLC38A9 signals arginine sufficiency to mTORC1. Science *347*, 188-194. 10.1126/science.1257132.
19. Fromm, S.A., Lawrence, R.E., and Hurley, J.H. (2020). Structural mechanism for amino acid-dependent Rag GTPase nucleotide state switching by SLC38A9. Nat Struct Mol Biol *27*, 1017-1023. 10.1038/s41594-020-0490-9.
20. Rebsamen, M., Pochini, L., Stasyk, T., de Araujo, M.E., Galluccio, M., Kandasamy, R.K., Snijder, B., Fauster, A., Rudashevskaya, E.L., Bruckner, M., et al. (2015). SLC38A9 is a component of the lysosomal amino acid sensing machinery that controls mTORC1. Nature *519*, 477-481. 10.1038/nature14107.
21. Li, K., Wada, S., Gosis, B.S., Thorsheim, C., Loose, P., and Arany, Z. (2022). Folliculin promotes substrate-selective mTORC1 activity by activating RagC to recruit TFE3. PLoS Biol *20*, e3001594. 10.1371/journal.pbio.3001594.
22. Napolitano, G., Di Malta, C., and Ballabio, A. (2022). Non-canonical mTORC1 signaling at the lysosome. Trends Cell Biol. 10.1016/j.tcb.2022.04.012.
23. Petit, C.S., Roczniak-Ferguson, A., and Ferguson, S.M. (2013). Recruitment of folliculin to lysosomes supports the amino acid-dependent activation of Rag GTPases. J Cell Biol *202*, 1107-1122. 10.1083/jcb.201307084.
24. Shen, K., Rogala, K.B., Chou, H.T., Huang, R.K., Yu, Z., and Sabatini, D.M. (2019). Cryo-EM Structure of the Human FLCN-FNIP2-Rag-Ragulator Complex. Cell *179*, 1319-1329 e1318. 10.1016/j.cell.2019.10.036.
25. Fisher, A.J., Smith, C.A., Thoden, J.B., Smith, R., Sutoh, K., Holden, H.M., and Rayment, I. (1995). X-ray structures of the myosin motor domain of Dictyostelium discoideum complexed with MgADP.BeFx and MgADP.AlF4. Biochemistry *34*, 8960-8972. 10.1021/bi00028a004.
26. Jeong, J.H., Lee, K.H., Kim, Y.M., Kim, D.H., Oh, B.H., and Kim, Y.G. (2012). Crystal structure of the Gtr1p(GTP)-Gtr2p(GDP) protein complex reveals large structural rearrangements triggered by GTP-to-GDP conversion. J Biol Chem *287*, 29648-29653. 10.1074/jbc.C112.384420.

27. Anandapadamanaban, M., Masson, G.R., Perisic, O., Berndt, A., Kaufman, J., Johnson, C.M., Santhanam, B., Rogala, K.B., Sabatini, D.M., and Williams, R.L. (2019). Architecture of human Rag GTPase heterodimers and their complex with mTORC1. Science *366*, 203-210. 10.1126/science.aax3939.
28. Mihai Gazdag, E., Streller, A., Haneburger, I., Hilbi, H., Vetter, I.R., Goody, R.S., and Itzen, A. (2013). Mechanism of Rab1b deactivation by the Legionella pneumophila GAP LepB. EMBO Rep *14*, 199-205. 10.1038/embor.2012.211.
29. Gavriljuk, K., Gazdag, E.M., Itzen, A., Kotting, C., Goody, R.S., and Gerwert, K. (2012). Catalytic mechanism of a mammalian Rab.RabGAP complex in atomic detail. Proc Natl Acad Sci U S A *109*, 21348-21353. 10.1073/pnas.1214431110.
30. Tian, W., Chen, C., Lei, X., Zhao, J., and Liang, J. (2018). CASTp 3.0: computed atlas of surface topography of proteins. Nucleic Acids Res *46*, W363-W367. 10.1093/nar/gky473.
31. Su, M.Y., Fromm, S.A., Remis, J., Toso, D.B., and Hurley, J.H. (2021). Structural basis for the ARF GAP activity and specificity of the C9orf72 complex. Nat Commun *12*, 3786. 10.1038/s41467-021-24081-0.
32. Su, M.Y., Fromm, S.A., Zoncu, R., and Hurley, J.H. (2020). Structure of the C9orf72 ARF GAP complex that is haploinsufficient in ALS and FTD. Nature *585*, 251-255. 10.1038/s41586-020-2633-x.
33. Shen, K., Valenstein, M.L., Gu, X., and Sabatini, D.M. (2019). Arg-78 of Nprl2 catalyzes GATOR1-stimulated GTP hydrolysis by the Rag GTPases. J Biol Chem *294*, 2970-2975. 10.1074/jbc.AC119.007382.
34. Egri, S.B., Ouch, C., Chou, H.T., Yu, Z., Song, K., Xu, C., and Shen, K. (2022). Cryo-EM structures of the human GATOR1-Rag-Ragulator complex reveal a spatial-constraint regulated GAP mechanism. Mol Cell. 10.1016/j.molcel.2022.03.002.
35. Alesi, N., Akl, E.W., Khabibullin, D., Liu, H.J., Nidhiry, A.S., Garner, E.R., Filippakis, H., Lam, H.C., Shi, W., Viswanathan, S.R., et al. (2021). TSC2 regulates lysosome biogenesis via a non-canonical RAGC and TFEB-dependent mechanism. Nat Commun *12*, 4245. 10.1038/s41467-021-24499-6.
36. Goodwin, J.M., Walkup, W.G.t., Hooper, K., Li, T., Kishi-Itakura, C., Ng, A., Lehmberg, T., Jha, A., Kommineni, S., Fletcher, K., et al. (2021). GABARAP sequesters the FLCN-FNIP tumor suppressor complex to couple autophagy with lysosomal biogenesis. Sci Adv *7*, eabj2485. 10.1126/sciadv.abj2485.
37. Gosis, B.S., Wada, S., Thorsheim, C., Li, K., Jung, S., Rhoades, J.H., Yang, Y., Brandimarto, J., Li, L., Uehara, K., et al. (2022). Inhibition of nonalcoholic fatty liver disease in mice by selective inhibition of mTORC1. Science *376*, eabf8271. 10.1126/science.abf8271.
38. Paquette, M., Yan, M., Ramirez-Reyes, J.M.J., El-Houjeiri, L., Biondini, M., Dufour, C.R., Jeong, H., Pacis, A., Giguere, V., Estall, J.L., et al. (2021). Loss of hepatic Flcn protects against fibrosis and inflammation by activating autophagy pathways. Sci Rep *11*, 21268. 10.1038/s41598-021-99958-7.
39. Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J Struct Biol *152*, 36-51. 10.1016/j.jsb.2005.07.007.
40. Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol *180*, 519-530. 10.1016/j.jsb.2012.09.006.

41. Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods *14*, 331-332. 10.1038/nmeth.4193.
42. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods *14*, 290-296. 10.1038/nmeth.4169.
43. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605-1612. 10.1002/jcc.20084.
44. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501. 10.1107/S0907444910007493.
45. Croll, T.I. (2018). ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. Acta Crystallogr D Struct Biol *74*, 519-530. 10.1107/S2059798318002425.
46. Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2018). MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci *27*, 293-315. 10.1002/pro.3330.
47. Bolte, S., and Cordelieres, F.P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. J Microsc *224*, 213-232. 10.1111/j.1365-2818.2006.01706.x.

Chapter 3: Structural Basis for GATOR2 inhibition by Sestrin2 and Castor1

Introduction

The GATOR2 complex functions as a negative regulator of GATOR1, inhibiting GATOR1 GAP activity in high nutrient cellular conditions[1]. The activity of GATOR2 is modulated by interactions with nutrient sensors, Sestrin2 and Castor1[2,3]. In a low nutrient state, Sestrin2 and Castor1 are not bound to their respective amino acids, enabling them to interact with GATOR2. This interaction between GATOR2 and the nutrients sensors impedes the ability of GATOR2 to inhibit GATOR1, allowing GATOR1 to function as an active GAP. When there is an increase in nutrient levels, Sestrin2 and Castor1 bind to leucine and arginine, respectively, triggering their release from the GATOR2 complex and facilitating the inhibitory interaction between GATOR1 and GATOR2[2,3]. Under these conditions, GATOR1 is inactive and does not perform its GAP function (Figure 3.1). Previous structural studies have been helpful to visualize the structures of key proteins in this pathway and better understand their individual function[4-10]. However, there are gaps in the understanding of how these proteins shift their interactions, in a nutrient-dependent manner, to regulate mTORC1 activity.

Castor1 and Sestrin2 are critical for GATOR2 nutrient-dependent regulation of GATOR1, yet their dynamic interactions with GATOR2 and regulatory mechanisms remain unclear. In 2016, the crystal structure of Castor1 in an arginine-bound state was solved, although the apo structure could not be determined at that time[5,6,9]. Castor1 was shown to form a homodimer with a single arginine ligand bound in the core of each monomer. Critical residues for the interaction with GATOR2 were identified through sequence conservation and mutagenesis studies[6]. One of the key residues (D121), identified as mediating the interaction with GATOR2, was in the interior of Castor1. This prompted the hypothesis that Castor1 uses "open- closed clamp" mechanism for binding and release from GATOR2. In this model, apo Castor1 would exists in an "open" state exposing residues necessary for interaction with GATOR2. Upon arginine binding, Castor1 would undergo a global conformational change clamping down on the arginine and transitioning to a "closed" state. In this conformation, residues critical for GATOR2 interaction would be blocked, facilitating Castor1 dissociation from GATOR2[6]. In 2019, a pseudo apo Castor1 structure was obtained by mutating key arginine binding residues (D304A and S111A)[10]. Overlay of this structure with the arginine-bound Castor1 structure revealed no significant structural rearrangement upon arginine binding, disproving the open-closed clamp hypothesis[10]. This left a critical gap in the understanding of the Castor1 regulatory mechanism as there was no clear explanation for how arginine binding to Castor1 triggers its release from GATOR2, a step that has been shown to be essential for the mTORC1 pathway to respond to arginine[3].

The crystal structure of Sestrin2 in a leucine bound state revealed Sestrin2 adopts a globular α-helical structure with a single leucine bound. The leucine molecule is buried in the core of the C-terminal domain (CTD) covered by a "lid" formed by a loop connecting two helices in the CTD[11]. The CTD lid makes key hydrogen bonds with residues in the N-terminal domain (NTD) that form a "latch" to lock the lid in place over leucine[11]. Structure guided mutagenesis studies and immunoprecipitation experiments identified residues mediating Sestrin2

interaction with GATOR2 and confirmed both the CTD and NTD of Sestrin2 were essential for this interaction[11]. These findings led to the hypothesis that Sestrin2 uses a lid-latch mechanism to regulate GATOR2 activity. In this model, apo Sestrin2 would be in a confirmation where the CTD "lid" is open and not secured by the NTD "latch". Upon leucine binding, Sestrin2 would undergo a confirmational change closing the "lid" to secure leucine in the CTD with the NTD "latch" and releasing Sestrin2 from GATOR2. Attempts were made to determine an apo structure of Sestrin2 to validate this regulatory mechanism. However, mutating residues critical for leucine binding was not sufficient to determine a true apo structure, as the solved structure still contained density for a bound leucine[12]. This left an ambiguity in the understanding of the Sestrin2 regulatory mechanism as there was no concrete explanation for how leucine binding to Sestrin2 triggers its release from GATOR2, a step that has been shown to be essential for the mTORC1 pathway to respond to leucine[2].

Biochemical experiments revealed Castor1 interacts with the MIOS β-propeller in GATOR2 and Sestrin2 with the WDR24 subunit[13], but failed to identify the exact GATOR2 residues mediating the binding of the nutrient sensors and how that translates to GATOR2 inhibition of GATOR1. A cryo-EM structure of human GATOR2 alone revealed GATOR2 adopts a large cage-like architecture composed of two WDR24-SEH1L, four MIOS-SEH1L and two WDR59-Sec13 heterodimers[8]. It is possible that these nutrient sensors allosterically alter the GATOR2-GATOR1 interaction by breaking or distorting the intricate cage. It is also possible that the nutrient sensors directly disrupt the interaction between GATOR1 and GATOR2 with competing binding sites. Additional structural insights into Castor1 and Sestrin2 in complex with GATOR2 are necessary to provide a more complete understanding of how these nutrient sensors regulate the activity of GATOR2 and GATOR1.

To understand if the nutrient sensors disrupt the GATOR1-GATOR2 interaction through direct binding competition, additional structural insights into the human GATOR1-GATOR2 complex are critical. A major limitation of the human GATOR2 structure was the portion of GATOR2, WDR59 β-propeller, shown to interact with GATOR1 was not resolved restricting the information that could be gathered about the GATOR1- GATOR2 interaction[8]. A cryo-EM structure of the yeast GATOR1-GATOR2 complex (SEA complex), which adopted a similar cage-like structure as the human GATOR2, provided some additional insights [13]. In this structure, the corresponding subunits for both WDR59 and GATOR1 were clearly resolved suggesting a potential interaction mode between human GATOR1 and GATOR2[13]. However, it is unclear if this interaction mode is conserved between yeast and humans because the yeast system lacks KICSTOR, a complex necessary for regulating GATOR1 and GATOR2 in humans[14]. Furthermore, given the ability of GATOR2 to enhance mTORC1 activity through GATOR1 inhibition, a detailed regulatory mechanism and understanding of the dynamic interactions between Castor1, Sestrin2, GATOR1 and GATOR2 would be necessary if this system were to be targeted therapeutically to tune mTORC1 activity in disease states.

Small Scale Expression Tests for Gator2

The GATOR complexes were purified to understand how GATOR2 is regulated by Sestrin2 and Castor1. While the GATOR1 complex had been routinely purified in the Hurley lab, GATOR2 purification had not been previously performed in the lab. The structure of GATOR2,

published in Nature in 2022, and studies on the yeast GATOR1 and GATOR2 complex, known as the SEA complex provided insights into possible purification methods[8]. Codon-optimized genes for the GATOR2 subunits were synthesized by Twist Biosciences and cloned into pCAG expression vectors with various tags, including GST, 2X STREP, and FLAG (Figure 3.2). Small-scale expression tests were conducted in HEK 293 GnTI cells to determine the optimal combination of plasmids and tags for expression and purification of the complex. Since purification of the yeast complex used a tagged MIOS subunit to pull down the rest of the complex endogenously, attempts were made to transfect MIOS alone, as well as co-transfecting MIOS with additional subunits.

Small scale expression tests were performed using 50 mL of cells transfected with 0.05 mg of DNA and polyethyleneimine (PEI) at a 1:4 ratio (DNA:PEI). The cells were harvested after 48 hours and resuspended in lysis buffer containing 1% TritonX-100 for rocked for 1 hour at 4°C. The cells were then centrifuged for 30 minutes at 15,000xg. The clarified lysate was incubated with 30 µL of equilibrated resin for 1 hour rocking at 4°C. The resin was washed 3X with 500 µL wash buffer, and the proteins were eluted for sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) visualization.

The first attempt to express and purify the full GATOR2 complex involved co-expression of all subunits and purification using a GST or STREP tags on the MIOS subunit. The results from these small-scale expression tests showed that not all the GATOR2 subunits were visualized using SDS-PAGE (Figure 3.3). Next, a FLAG-tagged MIOS construct was used in the expression to see if the full complex could purify when co-expressed with other subunits (Figure 3.4). It was possible to co-express FLAG-MIOS, WDR24, and WDR59, as well as FLAG-MIOS, WDR24, and SEH1L. However, the presence of the remaining endogenous subunits could not be confirmed using SDS-PAGE because of the similar molecular weights of the subunits. Then, tests were done with a FLAG tag on different GATOR2 subunits. First to see if the subunit could be purified independently of the other subunits and then to see if any subunit could pull down the full complex when co-expressed with all other subunits (Figure 3.5). The results from these tests indicated that a FLAG-tag on other subunits were not sufficient to pull down the full complex (Figure 3.5A). However, it appeared that the MIOS subunit and SEC13 could potentially be purified separately (Figure 3.5B).

To increase the amounts of protein obtained from the cells, harsher lysis conditions and additional lysis buffer additives were used to improve extraction and the stability of the protein. The cells were first dounce homogenized and then incubated for 1 hour incubation with 1% n-dodecyl-β-D-maltoside (DDM): cholesteryl hemisuccinate (CHS) (10:1). Additionally, an increased amount of NaCl (final concentration 500 mM), 10% glycerol, and 1 mM PMSF was added to the lysis buffer to stabilize the protein. These adjusted lysis conditions led to successful small-scale purification of GATOR2 after co-transfection of all five codon-optimized GATOR2 subunits and a STREP-tag on the MIOS subunit. The yields obtained were greater than those observed using non-codon optimized constructs from the GATOR2 structure paper[8]. This updated protocol was used for large scale purification of the complex.

Large Scale Purification of Gator2

The GATOR2 purification was scaled up to gather enough protein to use for structural studies. The large-scale method for purifying GATOR2 involved utilizing a STREP tag on the MIOS subunit and leaving the other subunits (WDR24, WDR59, SEH1L, Sec13) untagged. The purification protocol for STREP-GATOR2 began with co-transfection of plasmids (1 mg total) containing all five GATOR2 subunits into 1 L HEK 293 GnTI cells. It is important to note that for many activity assays and structural studies, 2 L of cells were commonly used per purification. This is because the sample frequently aggregated with concentration using Millipore Sigma Amicon centrifugal filters, so protein was used directly from the gel filtration.

The cells were harvested after 48 hours and resuspended in lysis buffer containing 25 mM 4-(2-hydryoxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5, 500 mM NaCl, 2 mM $MgCl_2$, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM tris(2-carboxyethyl)phosphine (TCEP) and protease inhibitor pellet (1 pellet per 50 mL buffer). The cells were homogenized via dounce homogenizing using the B pistol 20X. Next, 10% n-dodecyl-β-D-maltoside (DDM): cholesteryl hemisuccinate (CHS) (10:1) was added to the dounce homogenized cells for a final concentration of 1% DDM: CHS (10:1). The homogenized cells were incubated while rocking for 1 hour at 4 °C prior to centrifugation at 17,000 rpm for 35 minutes.

The clarified lysate was incubated with 5 mL pre-equilibrated STREP resin for a minimum of 3 hours (maximum overnight incubation). The resin was washed with 3-5 column volumes (CV) of high salt/adenosine triphosphate (ATP) wash buffer (25 mM HEPES pH 7.5, 500 mM NaCl, 2 mM $MgCl_2$, 1 mM TCEP, 1 mM ATP, 50 mM arginine, 50 mM glutamic acid) containing 0.03% DDM:CHS (10:1), 3-5 CV low salt/ATP wash buffer (25 mM HEPES pH 7.5, 200 mM NaCl, 2 mM $MgCl_2$, 1 mM TCEP, 1 mM ATP, 50 mM arginine, 50 mM glutamic acid) containing 0.03% DDM:CHS (10:1), 3-5 CV of low salt wash buffer containing 0.03% DDM:CHS (10:1) without ATP and 3-5 CV of low salt wash buffer without detergent.

The protein complex was eluted in ~20 mL of low salt wash buffer containing 4 mM desthiobiotin. Before elution, the resin was incubated for 30 minutes on ice with elution buffer. The eluted sample was concentrated to 1 mL using 100 kDa Millipore Sigma Amicon centrifugation filter for gel filtration using a Superose 6 increase column. Gel filtration buffer contained 25 mM HEPES pH 7.5, 200 mM NaCl, 2 mM $MgCl_2$, 1 mM TCEP. The protein complex eluted ~11.9 mL and the fractions were visualized on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 3.6). Another important note is that while the sample is best in the presence of detergent, detergent was not optimal for structural characterization. Therefore, it was removed in later wash steps and during gel filtration. If experiments allow, keeping low amounts of detergent present in sample could help with stability and preventing aggregation.

Cryo-EM of GATOR2

To validate the structural integrity of the purified GATOR2 complex using the new purification method, samples were prepared for cryo-EM analysis. The sample was diluted to

ensure individual particles were visible on the micrographs and frozen on CF-2/1-3Au-T grids using a blot time of 3 seconds, blot force of +20, temperature at 4C and 100% humidity. Five screening images were taken on the Arctica. Each image showed similar particle dispersion and structure. While a complete dataset was not collected, the particles observed in the screening images closely resembled the previously published 2D classes from the human GATOR2 structure (Figure 3.7). This suggested that the GATOR2 complex purified in our lab adopts the same conformation to the previously characterized structure.

Purification of Castor1 and Sestrin2

Small scale expression tests were conducted with Sestrin2 and Castor1 to obtain purified sample to use to dissect interactions between GATOR2 and the nutrient sensors (Sestrin2 and Castor1). Initially, wild-type (WT) Sestrin2 and Castor1 were expressed in HEK 293 GnTI cells using the pRK5 expression vector and FLAG tags (Figure 3.8). Previous literature has reported these proteins could be purified from both mammalian cells and *E. coli*. Therefore, to enhance yields, simplicity, and cost-effectiveness of the purification, subsequent purification of the protein used *E. coli.* To improve the yields of Castor1 and Sestrin2 in *E. coli*, codon-optimized genes were synthesized by Twist Bioscience and cloned into the E. coli expression vector pET28a(+), containing a N-terminal 6X-His tag. The purification protocol for Castor1 followed a previously published method[6]. Castor1 mutants were purified using the same protocol as WT. Example SDS-PAGE and chromatogram from Castor1 purification can be seen in Figure 3.9.

The purification protocol for Sestrin2 was adapted from a previously published protocol[11]. The modifications to the purification protocol for Sestrin2 included running the clarified lysate through Ni-NTA resin 3X, followed by ~150 mL wash with of lysis buffer and ~70 mL elution using the described buffers. The eluted sample was concentrated to 2 mL for gel filtration using a HiLoad Superdex 200 column. The gel filtration buffer contained 10 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.1 mM ethylenediamine tetraacetic acid (EDTA), and 1 mM dithiothreitol (DTT). To remove a prominent contaminating species, fractions containing Sestrin2 were passed through a cation-exchange chromatography column, and the flow-through (FT) was collected. The contaminant remained attached to the column resin. The FT was concentrated, and the sample was utilized for subsequent structural and functional studies. Example SDS-PAGE and chromatogram from the Sestrin2 purification can be seen in Figure 3.10.

Low Resolution Initial Cryo-EM for Castor1 + GATOR2 Complex

Arctica Data Set

Cryo-EM was used to visualize the interaction between Castor1 and GATOR2. To prepare the sample for cryo-EM, GATOR2 was purified using STREP tag on MIOS. After elution from the STREP resin, GATOR2 was concentrated to ~440 nM and incubated overnight at 4°C with 20-fold molar excess of "apo" Castor1, a mutant (D304A, S111A) lacking critical arginine binding residues. The complex was then run on gel filtration using a Superose 6 increase column, resulting in a chromatogram showing a heterogenous peak. The left peak, which eluted at ~12 mL, corresponded to GATOR2 alone, as confirmed by SDS-PAGE analysis of the

fractions. In contrast, the right shoulder, eluting at ~13 mL, corresponded to GATOR2 bound to Castor1 (Figure 3.11).

Fractions from the peak corresponding to the GATOR2-Castor1 complex were concentrated, and another 10-fold molar excess of Castor1 was added for complete saturation of binding sites. After overnight incubation at 4°C, the complex was frozen on CF-2/1-3Au-T grids and ~9,000 micrographs were collected over a two-day Arctica session. The grid freezing conditions included a blot time of 3 seconds, blot force of +20, temperature at 4°C and 100% humidity. The particles on the collected movies appeared evenly dispersed across the images (Figure 3.12A and B).

The data processing workflow began with motion correction in Relion using MotionCor2 before being imported for further processing in cryoSPARC[15-17]. The cryoSPARC blob picker was used to create an initial particle set, and further rounds of 2D classification produced a smaller subset of particles exhibiting a distinct cage-like shape (Figure 3.12C and D). This shape differed from that observed with GATOR2 alone, indicating a potential conformational change upon Castor1 binding. Particles from the best 2D classes were used to construct an ab initio model. Multiple additional rounds of particle picking using blob picker, template picker, and Topaz were conducted until all the best particles from the micrographs were exhausted [18]. The resulting ~60,000 best particles were used to generate a low-resolution map with a resolution of ~9 Å (Figure 3.12E and F).

The low resolution map was of sufficient quality to rigid body fit the known cryo-EM and crystal structure of GATOR2 and Castor1 (Figure 3.12E). In the map, there was a clear density for a single Castor1 dimer that revealed two interfaces with GATOR2. The first interface observed touched one of the MIOS β-propellers, while the second interface touched an unassigned density adjacent to the inner core of the complex. Additional local refinement was used to try to improve the resolution of the map. The map was segmented into small sections around regions with likely rigidity. Masks were created around these regions for use in local refinement. The resolution in these regions was improved to ~5-7 Å (Figure 3.13) although the maps were still noisy and only some portions appeared to accurately show secondary structure. The most important region at the GATOR2-Castor1 interfaces had considerable noise and it was hard to accurately fit secondary structure. Further improvement of this map will be needed to accurately reconstruct these interfaces. It is possible that the low resolution at the Castor1 interface is a result of weak binding of Castor1 to the complex. Based on the heterogeneity from the gel filtration of the complex during sample preparation it would suggest that Castor1 is only stably binding a portion of the GATOR2 sample.

Another ambiguity from the map is that given that there are four copies of the MIOS subunit, it is unclear why only a single monomer of Castor1 interact with GATOR2 in this map. One possibility is that there are multiple binding sites, but they are not occupied simultaneously, and the dataset was too small to capture all the different binding modes. Another hypothesis is that binding of a single Castor1 molecule constrains the cage in a conformation that makes the other subunits incompatible for binding to other Castor1 molecules. This could be why Sestrin2 is needed to also regulate the cage. It is possible that Castor1 interacts with half of the cage, while Sestrin2 regulates the other half. A higher resolution structure and investigation into the

interplay with Sestrin2 is needed to elucidate the exact mechanism underlying the interaction between Castor1 and GATOR2.

Experimentally Guided Alphafold Prediction for Castor1 + GATOR2 Complex

Alphafold2 was used to generate predictions for the structure of the GATOR2-Castor1 interface[19,20]. The low resolution cryo-EM revealed Castor1, a functional homodimer, interacts with the GATOR2 complex through two distinct interfaces, each involving one of the Castor1 monomers. Generation of the alphafold prediction models of GATOR2-Castor1 interfaces were guided by the experimental low-resolution map. This approach proved helpful in predicting interface mutations, prior to completing experimental structural work, allowing for a jump start on complimentary biochemical and cell-based experiments.

Interface #1: Castor1-MIOS

After performing rigid body fitting of the cryo-EM and crystal structures of GATOR2 and Castor1 into the low-resolution map density, it became evident that while Castor1 fit well into the density, the MIOS beta propeller interacting with Castor1 did not (Figure 3.14). This suggested the interface in the rigid body fit structure was inaccurate, hindering the dissection of critical residues. To address this, residues corresponding to the MIOS β-propeller (residue 1-355) were modeled using Alphafold2 with residues corresponding to a single full length Castor1 monomer[19,20]. A prediction was generated for the interface between the two protein chains with high confidence (ipTM 0.87) (Figure 3.14B). Upon fitting the Alphafold predicted structure into the cryo-EM map, using the well-fitting Castor1 monomer as a reference, the MIOS β-propeller snugly fit into the EM density (Figure 3.14A). The alignment between the Alphafold model and the EM map suggested improved accuracy at this interface.

Examination of the residues at the interface revealed electrostatic interactions facilitated the binding between Castor1 and MIOS. The MIOS subunit has multiple positively charged arginine and lysine residues (K64, K111, K134, R137, R206, R225 and R288) on its surface, while Castor1 displayed a highly negative surface composed of aspartic acid, glutamic acid, and tyrosine residues (E2, D121, E146, E150, E192, E234, Y236, D252, E261, D292, Y296, and E310) (Figure 3.14C). This charge complementarity was likely driving the binding between MIOS and Castor1 at the interface. Mutants targeting the MIOS and Castor1 interfaces were generated for follow biochemical validation experiments. Three residues on the Castor1 interface (D121, D292, and E261) had been previously identified as critical for Castor1 interaction with GATOR2[6,9], further supporting the predictions made by Alphafold2[19].

A fundamental open question regarding the GATOR2-Castor1 interaction is how Castor1 dissociates from the complex upon arginine binding. To address this, the Castor1 monomer from the Castor1-MIOS interface Alphafold model (resembling the "apo" bound state) was overlaid with a single monomer in the Castor1 arginine-bound crystal structure to identify changes between the two structures (Figure 3.15A). This overlay revealed a difference in the location of a critical loop in Castor1 formed by residues 83-93. In the Alphafold structure, the loop was predicted in an open conformation with low confidence suggesting it may be disordered when Castor1 is in an apo state and bound to MIOS (Figure 3.15A). Comparison with the apo crystal

structure of Castor1 revealed that this critical loop is in fact disordered in the apo crystal structure (Figure 3.15B). The key residue in Castor1 mediating the interaction between Castor1 and GATOR2, D121, is located behind the loop formed by residues 83-93 only in the arginine-bound state. This suggests a new regulatory mechanism where arginine binding induces the ordering of a critical loop to block the interaction between Castor1 and GATOR2. In this model, the highly negative pocket on Castor1 containing residue D121 is exposed in the apo state when the loop is disordered, allowing MIOS to interact with Castor1 (Figure 3.15C). In the arginine-bound state, the loop becomes ordered and forms a "lid" covering the negative pocket (Figure 3.15C). This turns the Castor1 surface more neutral charge and disrupts the electrostatic interactions with GATOR2 facilitating dissociation.

Interface #2: Castor1-WDR59

Analyzing the interface with the second Castor1 monomer posed more challenges. The density interacting with the second Castor1 monomer was unassigned in the cryo-EM map, and no portion of the original GATOR2 structure clearly fit (Figure 3.16A). From the rigid body fitting of GATOR2, it was evident that Castor1 binding causes the inner cage to become constrained since the diameter on the cage is narrower than the original structure. This suggests a potential reorientation of a portion of the inner cage upon Castor1 binding, which could account for the unassigned density. Initially, residues of the WDR59 and WDR24 subunits from the inner cage were ran simultaneously with a single full length Castor1 monomer in Alphafold2[19,20]. However, none of these Alphafold predictions, incorporating portions of WDR59 and WDR24 visible in the original GATOR2 structure, yielded predictions with high confidence that aligned with the experimental map.

Upon closer examination of the GATOR2 cryo-EM structure, it was observed that a significant portion of the WDR59 subunit was disordered. However, in the yeast structure, more residues of this subunit were ordered. The yeast SEA complex was rigid body fitted into the map, and the entire yeast WDR59 subunit was examined. A region of the yeast WDR59 that was disordered in the human GATOR2 structure seemed to adopt a shape like the unassigned density. Sequence alignment between the yeast WDR59 and human WDR59 subunit identified the residues corresponding to this region in the human WDR59 subunit. These WDR59 residues (residues 392-502) were modeled using Alphafold2 with a single full-length Castor1 monomer to generate a prediction. While the prediction was lower confidence (ipTM 0.22), the structure fit well in the EM density (Figure 3.16A and B). Iterative Alphafold runs helped assign the initially unassigned density to different regions of WDR59.

The Alphafold model revealed that the interface between the second Castor1 monomer and WDR59 was also mediated by electrostatic interactions (Figure 3.16C). Castor1 exhibited a highly negative surface (E2, Y118, Q119, E150, D190, E192, E234, Y236, D252), while WDR59 contained key arginine and lysine residues (R426 and K428) that formed a positive surface. Upon examination of all residues on the Castor1 surface, six residues (E2, E150, E234, Y236, E192, and D252) were found to interact with MIOS on the first monomer. In addition, two residues (Q119 and Y118) that had been previously identified as critical for binding but were not observed in the first interface were seen on this interface, and one residue (D190) that was not seen in either the first interface or in previous studies was identified.

To elucidate why Castor1 dissociates from the GATOR2 complex upon arginine binding, the surface charges of the Alphafold monomer (pseudo-"apo") and the arginine-bound crystal structure were compared (Figure 3.17). This comparison revealed that in the apo structure, there are two negatively charged pockets that interact with the positively charged residues K428 and R426 in the WDR59 subunit. However, in the arginine-bound crystal structure, small rearrangements in the structure, likely from side chains, changed the charge of pocket where WDR59 R426 binds neutral, potentially disrupting the interaction at this interface (Figure 3.17). While this interface appears to play a lesser role in Castor1 dissociation from GATOR2 compared to the Castor1-MIOS monomer, both sides may contribute to the process. Follow-up studies will be necessary to determine the extent to which each interface contributes to the dissociation of Castor1 from GATOR2.

Sample Optimization for GATOR2-Castor1 Complex

Four different approaches were used to optimize the GATOR2-Castor1 complex sample prep. Additional datasets were collected on the Kios G3 microscope in attempts to improve the resolution of the initial data collection. The approach that yielded a high-resolution map involved purifying GATOR2 and concentrating the gel filtration peak to ~360 nM. Castor1 was added to the GATOR2 sample at a 3-fold molar excess and incubated on ice for 45 minutes – 1 hour to facilitate complex formation. The sample was not further concentrated after incubation and was not subjected to gel filtration to remove excess Castor1. The sample was frozen directly on CF-2/1-3Cu-T grids using the following grid freezing conditions: blot time of 3 seconds, blot force of +20, temperature at 4°C and 100% humidity.

A brief overview of the additional optimization attempts that were tried and problems with the approaches are below.

1. One of the main problems with the original data set was low concentration of particles on the micrograph. In attempts to increase concentration, a GATOR2 sample was prepared and mixed with excess Castor1. After incubation, the sample was concentrated to obtain a higher sample concentration. Grids were frozen and a dataset was collected. After processing the data, it was clear that all the protein precipitated and the only particles left on the micrographs were contaminants (chaperone, ribosome, etc).

2. To improve the stability of the sample during concentration, detergent was kept present throughout the entirety of the GATOR2 prep and complex formation with Castor1. Detergent typically improves the stability and yield for the GATOR2 sample alone. However, the when the GATOR2-Castor1 complex was concentrated, soluble aggregates were observed on the grids when imaging on the microscope. In the sample with no concentration and detergent, the percentage of detergent was too high, and the protein was clustered at the rim of the holes on the grids.

3. To avoid concentrating the sample following GATOR2-Castor1 complex formation, 2 nm carbon coated grids were used. These allow for ~10X less protein in the final

concentration. After screening the grids, particles looked evenly dispersed. However, during data processing the 2 nm carbon background made it hard to align the particles limiting the data processing.

High Resolution Data Collection for GATOR2-Castor1 Complex

The Krios G3i microscope was used to collected ~12,000 micrographs for the GATOR2-Castor1 complex. The micrographs exhibited uniform particle dispersion with particles whose shape and size were indicative of GATOR2 (Figure 3.18A and B). The raw movies were imported into cryoSPARC for data processing[16]. The movies were motion corrected using Patch Motion Correction and the CTF was estimated using Patch CTF. To generate initial particle sets, the cryoSPARC blob picker was used with a particle diameter range of 200 – 280 Å. The particle picks were inspected, and ~1.7 million particles were extracted using a box size of 560. The particles were sorted initially in 2D to select the best classes for an ab initio reconstruction. A three-class ab initio revealed a class representing an intact GATOR2 cage and one Castor1 bound. The additional two classes represented a broken half cage and noise. Using the 3D classes generated from ab initio as well as 3D classes generated from the initial Arctica data collection, heterogenous refinement was used to re-sort all the particle picks (Figure 3.18C and D). A prominent class emerged using the volume generated by the ab initio reconstruction. After filtering the particles extensively, ~140,000 remained and these particles were used in a homogenous refinement to generate a map with an overall resolution of 3.89 Å (Figure 3.18E). Additional local refinement was used to increase the resolution in regions of the map to ~3.0 – 3.7 Å. The interface between MIOS and Castor1 was refined to a resolution of 3.24 Å allowing for clear visualization and placement of amino acids at the interface (Figure 3.18F).

The high-resolution map revealed a different binding mode from the low-resolution map of GATOR2-Castor1 obtained from the Arctica. In the high-resolution structure, Castor1 is bound symmetrically across the cage with each Castor1 monomer in the dimer engaging with a MIOS subunit. This is likely a result of differences in sample preparation. For the Arctica data collection, the sample was subjected to gel filtration prior to grid preparation. Since the gel filtration chromatogram for the GATOR2-Castor1 complex was heterogenous, it is likely there are two different binding states. By selecting only one of the peaks after gel filtration, the second higher resolution binding mode was filtered out. However, in the data set for the high-resolution structure no gel filtration as used to separate the two states. Therefore, all binding modes could be observed, revealing the symmetrical high resolution binding mode. The particles representing the low-resolution binding mode could be selected for but in low quantities. The binding mode in the low-resolution structure may function as a transition state and could be biologically relevant. Additional follow up studies will be needed to explore this additional binding mode.

Reconstituting the GATOR1 and GATOR2 System In Vitro

The regulation of GATOR1 by GATOR2 has been shown in cell-based experiments but has never been reconstituted with purified protein in vitro. To better understand the molecular mechanisms of GATOR2 inhibition of GATOR1 and the role Castor1 and Sestrin2 play in this regulation, efforts were made to reconstitute the system in vitro. Having a reconstituted system

provides an isolated system to dissect the effects of direct interactions between proteins without the complications of the full cellular environment.

Design of Promega GTPase Glo Assay

The Promega GTPase Glo is a commercially available assay for measuring the activity of GTPases and their corresponding GAPs. This was used to monitor GATOR1 GAP activity in vitro. For this assay, the GTPase/GAP reaction is set up in the presence of excess GTP. The reaction is allowed to proceed for a given time until a GTPase Glo Reagent is added to convert any remaining GTP in the sample to ATP. A detection reagent is then added, and the luminescence of the sample is read. The intensity of the luminescence corresponds with the amount of ATP (a proxy for remaining GTP) in the sample. Therefore, high luminescence is indicative of low GAP activity while low luminescence is high GAP activity (Figure 3.19).

The assay was first used to monitor GATOR1 activity towards the Rag GTPases. Rags (1μM) were incubated with GATOR1 (100 nM) for 1 hour in the presence of 0.5 mM GTP. Under these conditions, the luminescence for Rags + GATOR1 reaction was significantly lower compared to Rags alone reaction after 1 hour (Figure 3.20). GATOR2 was then added to the system to see if it would alter the activity of GATOR1. When using GATOR2 in these assays, GATOR2 was always prepared fresh and never freeze/ thawed to ensure the integrity of the protein complex. Initial tests using Rags (2 μM) + GATOR1 (100 nM) + GATOR2 (100 nM) , there was a noticeable difference in the resulting luminescence from the Rags (2 μM) + GATOR1 (100 nM) baseline (Figure 3.21). This difference was not altered by the presence of KICSTOR (100 nM). However, the signal for Rags (2 μM) + GTP (0.5 mM) alone was much higher than when GATOR2 was added suggesting GATOR2 was not a complete inhibitor. This could be a result of the concentrations or conditions in the reaction not being suitable for this system. Another possibility is contaminants in the sample prep altered the activity or other components in the reagents. For example, the sample of purified GATOR2 is often contaminated with chaperone proteins. This is unavoidable and they are hard to remove during sample prep. These chaperones could have reduced the ATP in solution leading to an inaccurate GTP readout for this assay.

While the assay was not fully sufficient to examine the system in vitro, initial tests for adding Castor1 were conducted. In the reaction with Rags (1 μM) + GATOR1 (6.25 nM) + GATOR2 (100 nM) + Castor1 (1 μM), the luminescence was restored to that of Rags (1 μM) + GATOR1 (6.25 nM), suggesting that Castor1 could reverse the inhibitory potential of GATOR2. However, the signal for Rags (1 μM) + GATOR1 (6.25 nM) + GATOR2 (100 nM) was still significantly less than that of Rags (1 μM) + GTP (0.5 mM) alone (Figure 3.22) making the result less conclusive due to the challenges described in the above paragraph. The problems with the GTPase Glo kit led to the decision to switch to an HPLC-based assay that produces a direct readout of nucleotides.

Design of HPLC Based Assay

A second attempt to reconstitute the system in vitro was made using an HPLC-based assay. The HPLC assay allows for a direct readout of nucleotides in solution. The reaction was

set up similarly to the GTPase Glo. Rags (2 μM) were incubated with an excess of GTP (0.5 mM) in the presence or absence of GATOR1 (20 nM). Initial tests indicated that GATOR1 was able to stimulate the hydrolysis of GTP in the Rags and there was an increase in the amount of GDP present in the sample (Figure 3.23A and B). When GATOR2 was added to the sample, the resulting amount of GDP decreased suggesting GATOR2 inhibited the GAP activity of GATOR1 (Figure 3.23C). Controls testing the activity of GATOR1 and GATOR2 alone showed no increase in GDP compared to the baseline of Rags alone (Figure 3.23D and E). These results are preliminary, and replicates of this experiment will need to be conducted to confirm the effect of GATOR2 on GATOR1 GAP activity in vitro. In addition, once this system is validated in vitro, Castor1 will be introduced into the system to observe the ability of the protein to inhibit GATOR2 and restore GATOR1 GAP activity.

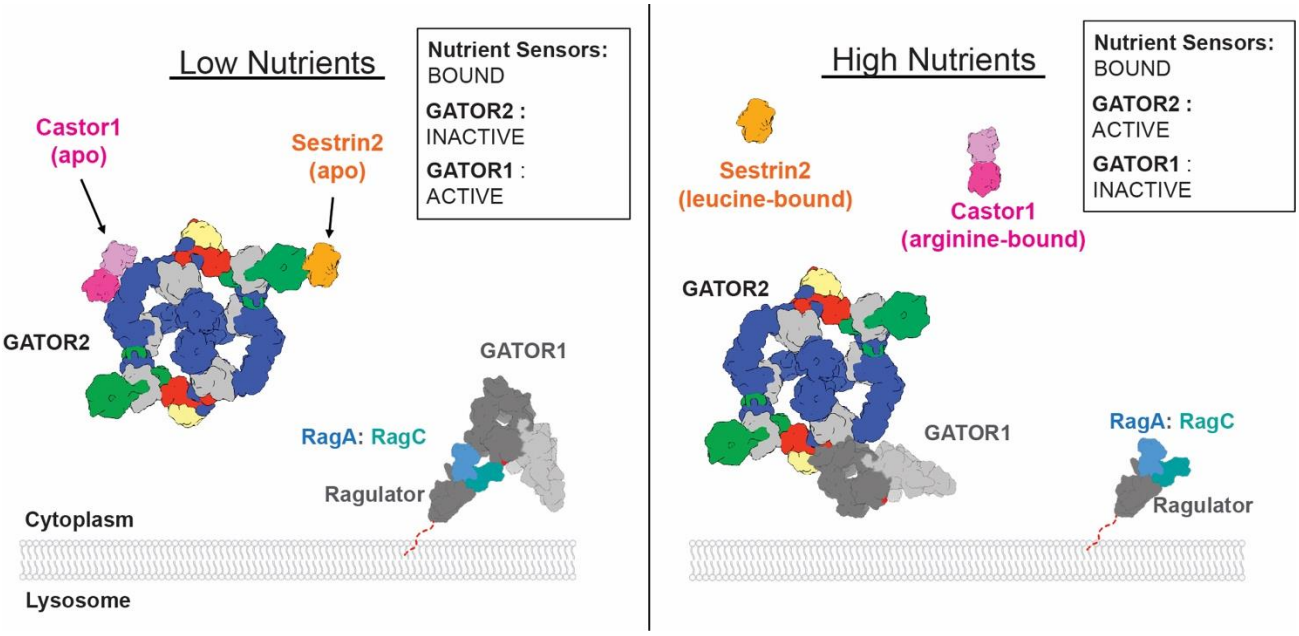


Figure 3.1 Schematic of Nutrient Dependent GATOR2 regulation of GATOR1. In low nutrients, nutrient sensors (Castor1 and Sestrin2) are bound to the GATOR2 complex. In this state, GATOR2 is "inactive" and unable to inhibit the GAP activity of GATOR1, allowing GATOR1 to function as an "active" GAP. In high nutrients, nutrient sensors (Castor1 and Sestrin2) are not bound to the GATOR2 complex. In this state, GATOR2 is "active" and able to inhibit the GAP activity of GATOR1.



Figure 3.2 Constructs designed for GATOR2 subunits. Different tags used for protein purification were cloned on to the different subunits of GATOR2.

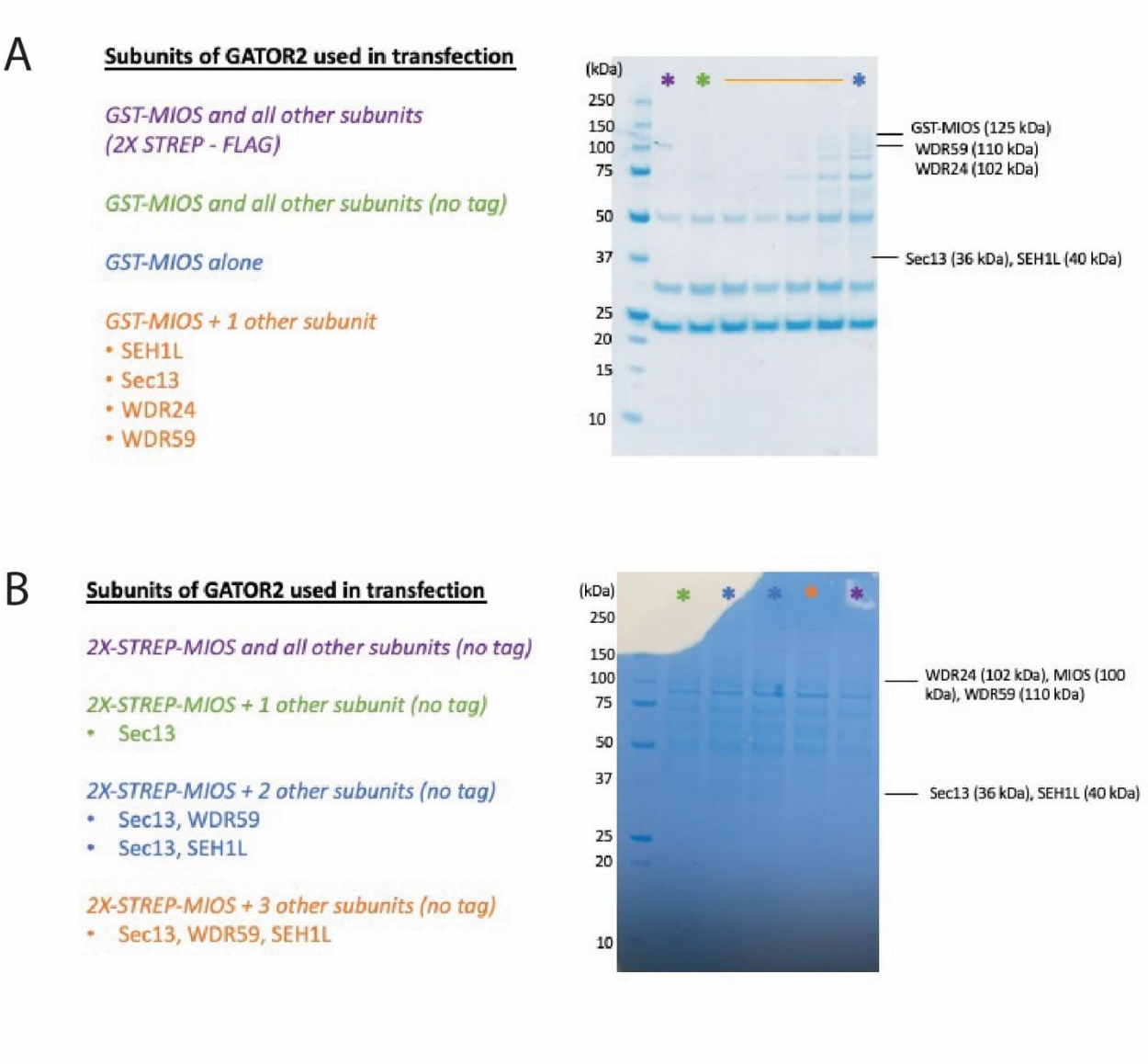


Figure 3.3 Initial small scale expression tests for GATOR2 with STREP and MIOS tags. (A) SDS-PAGE for small scale expression tests using GST tag. GST-MIOS was co-transfected with all other GATOR2 subunits containing 2X-STREP-FLAG tags (purple), all other GATOR2 subunits untagged (green), alone (blue) and with one other GATOR2 subunit (orange). (B) SDS-PAGE for small scale expression tests using 2X-STREP tag. 2X-STREP-MIOS was co-transfected with all other GATOR2 subunits untagged (purple), one other subunit (green), two other subunits (blue) and three other subunits (orange).

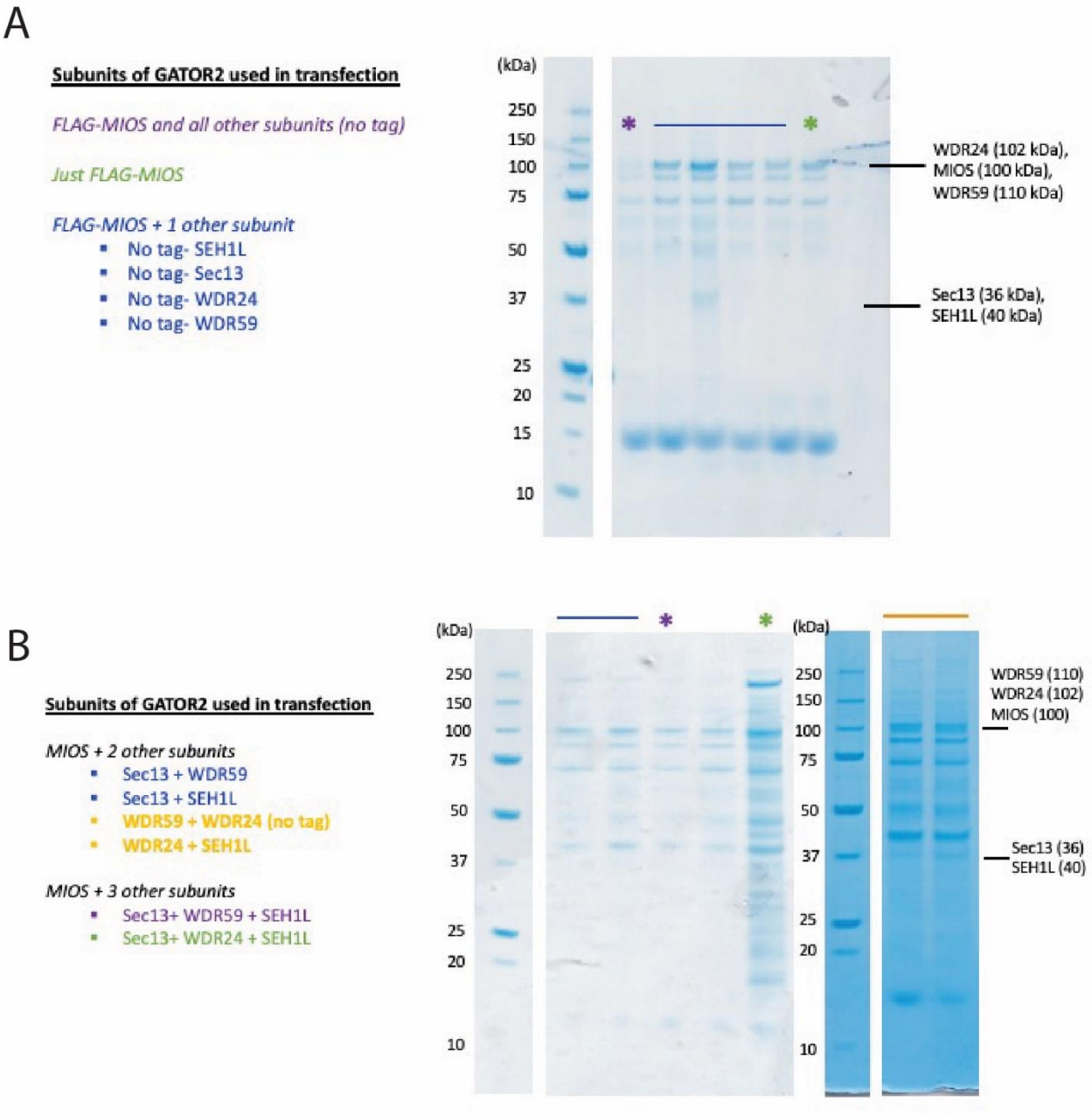


Figure 3.4 Initial small scale expression tests for GATOR2 with FLAG-MIOS. (A) SDS-PAGE for small scale expression tests using FLAG tag on MIOS. FLAG-MIOS was co-transfected with all other GATOR2 subunits untagged (purple), alone (green), with one other GATOR2 subunit (blue). (B) SDS-PAGE for FLAG-MIOS was co-transfected with two other GATOR2 subunit (blue and yellow) and three other subunits (purple and green).

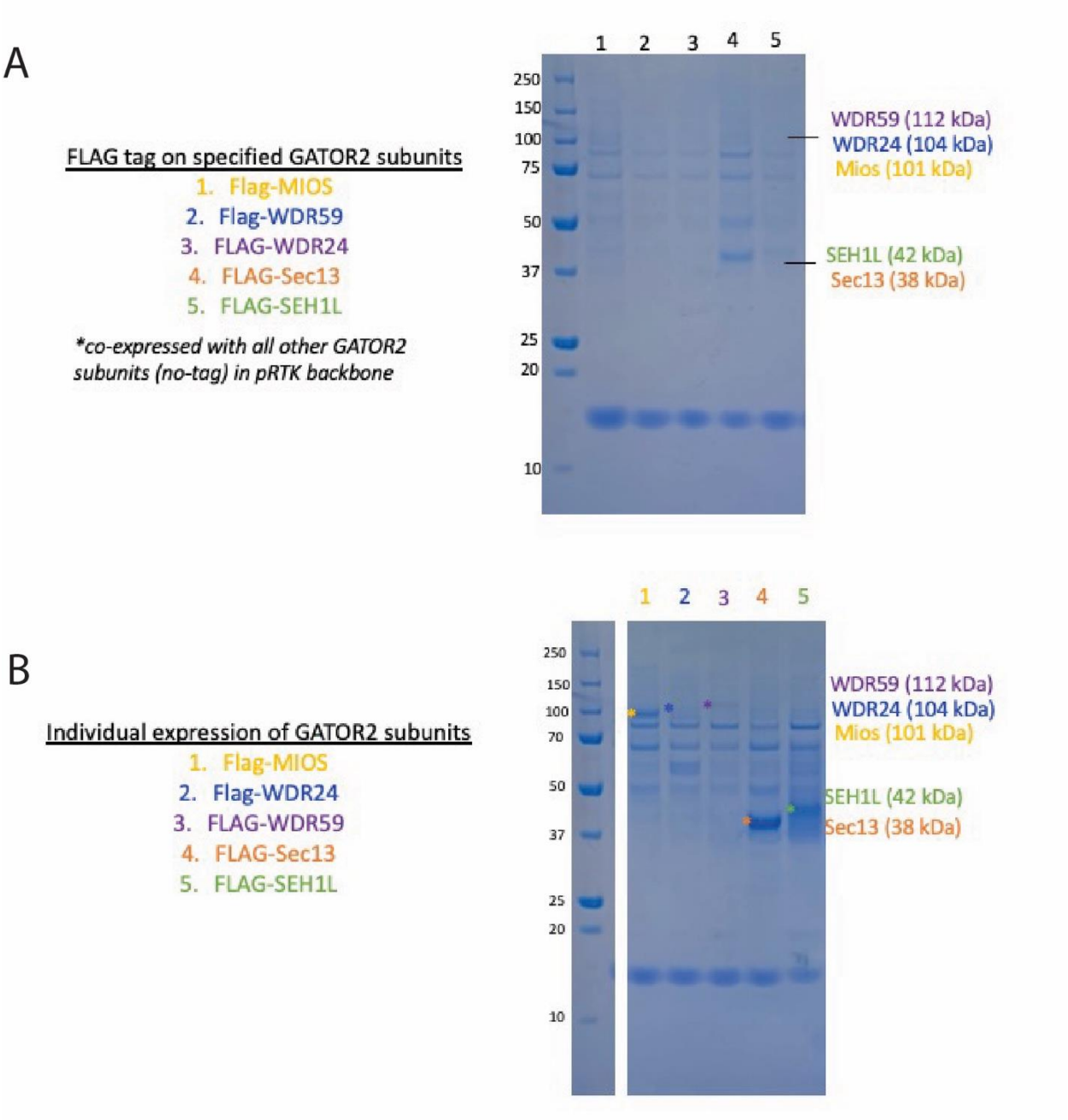


Figure 3.5 Initial small scale expression tests for GATOR2 using FLAG-tagged subunits. (A) SDS-PAGE for FLAG tagged GATOR2 subunits were co-expressed with all other GATOR2 subunits (untagged). (B) SDS-PAGE for FLAG tagged GATOR2 subunits were expressed alone.

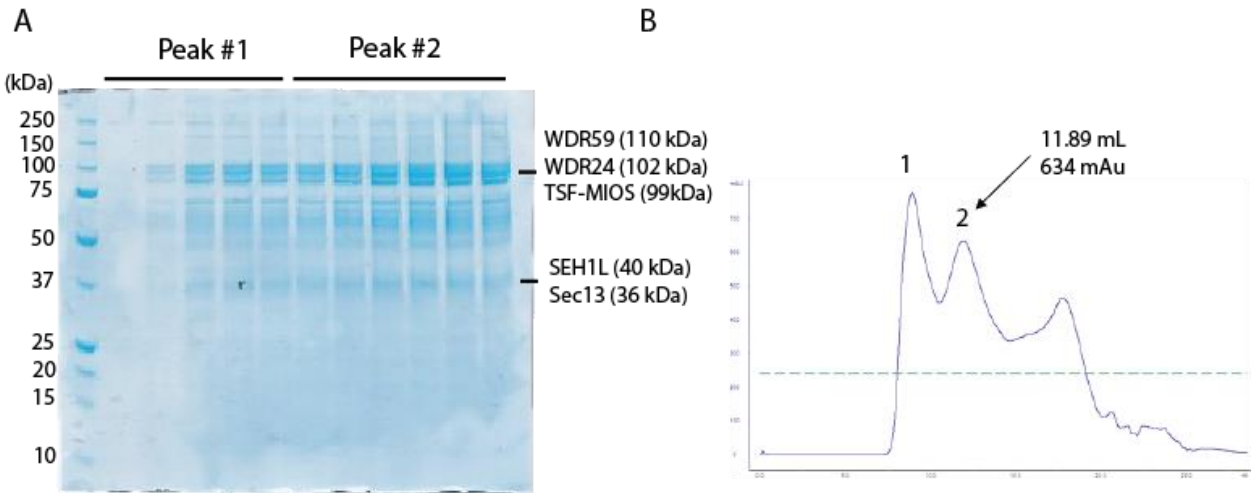


Figure 3.6 Large scale purification of STREP-GATOR2. (A) Representative SDS-PAGE and (B) Chromatogram from purification.

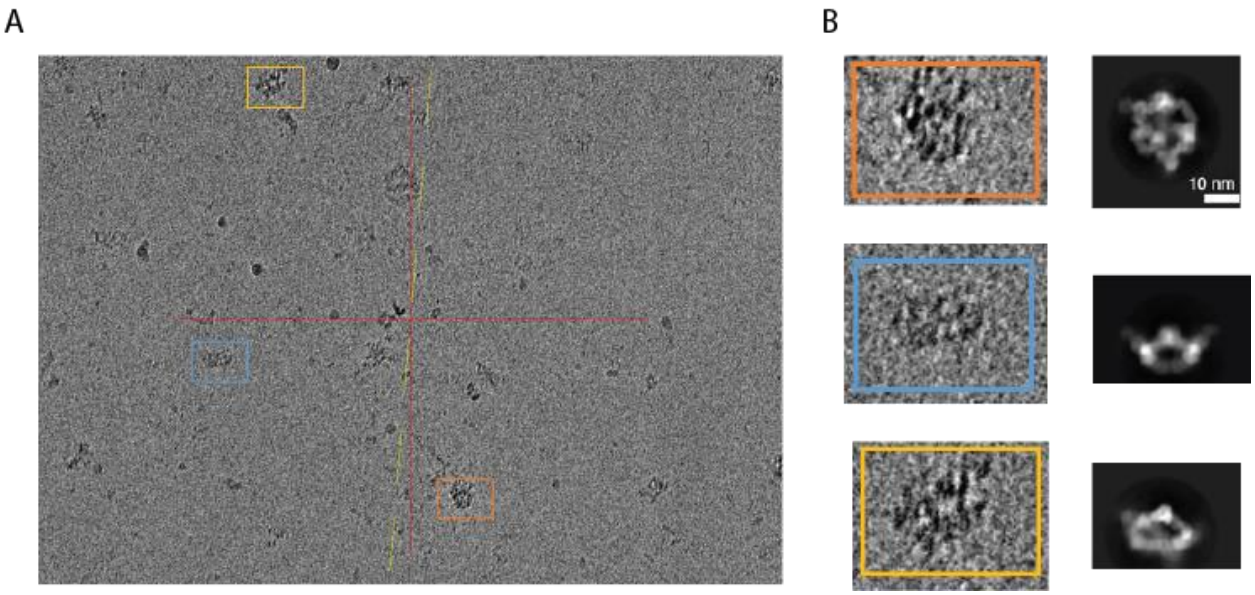


Figure 3.7 Cryo-EM analysis of GATOR2 sample. (A) Representative micrograph. (B) Individually picked particles compared to 2D class averages from the original Gator2 cryo-EM structure[8].

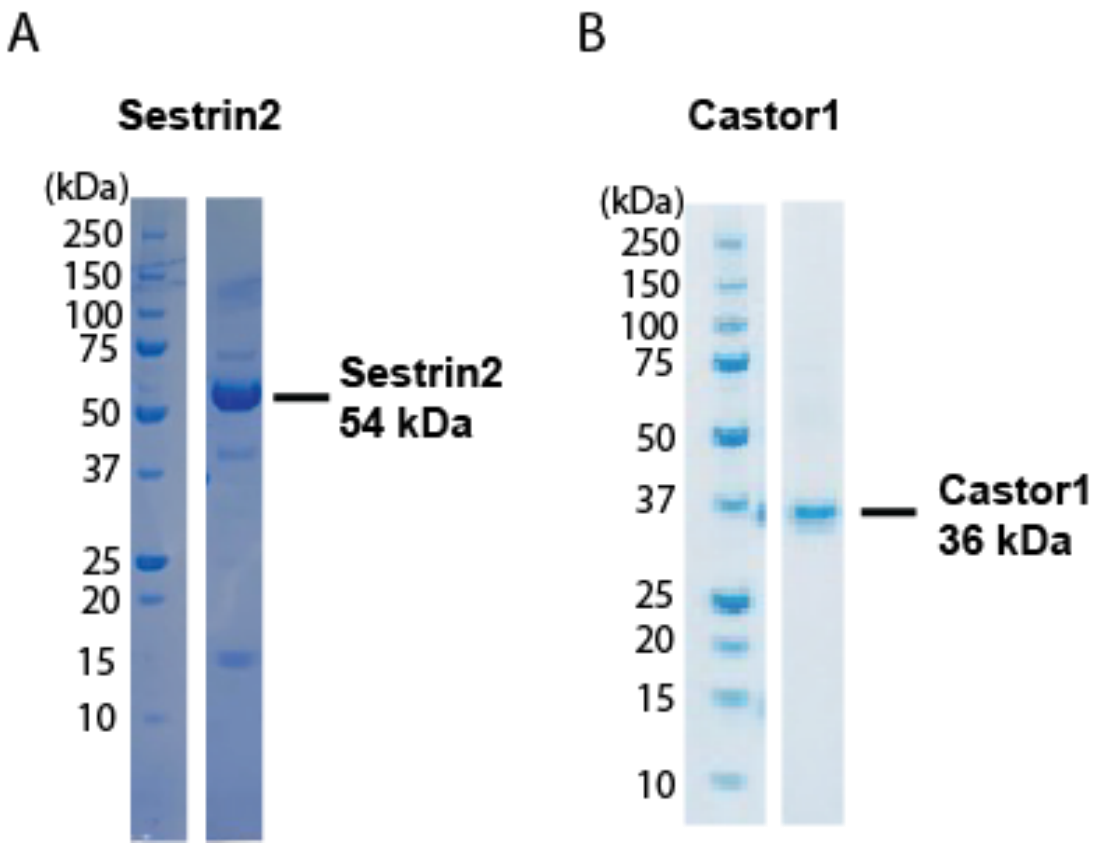


Figure 3.8 Purification of FLAG-Sestrin2 and FLAG-Castor1 from HEK 293 GnTI. (A) SDS-PAGE gels of Sestrin2 (A) and Castor1 (B) following elution from FLAG resin.

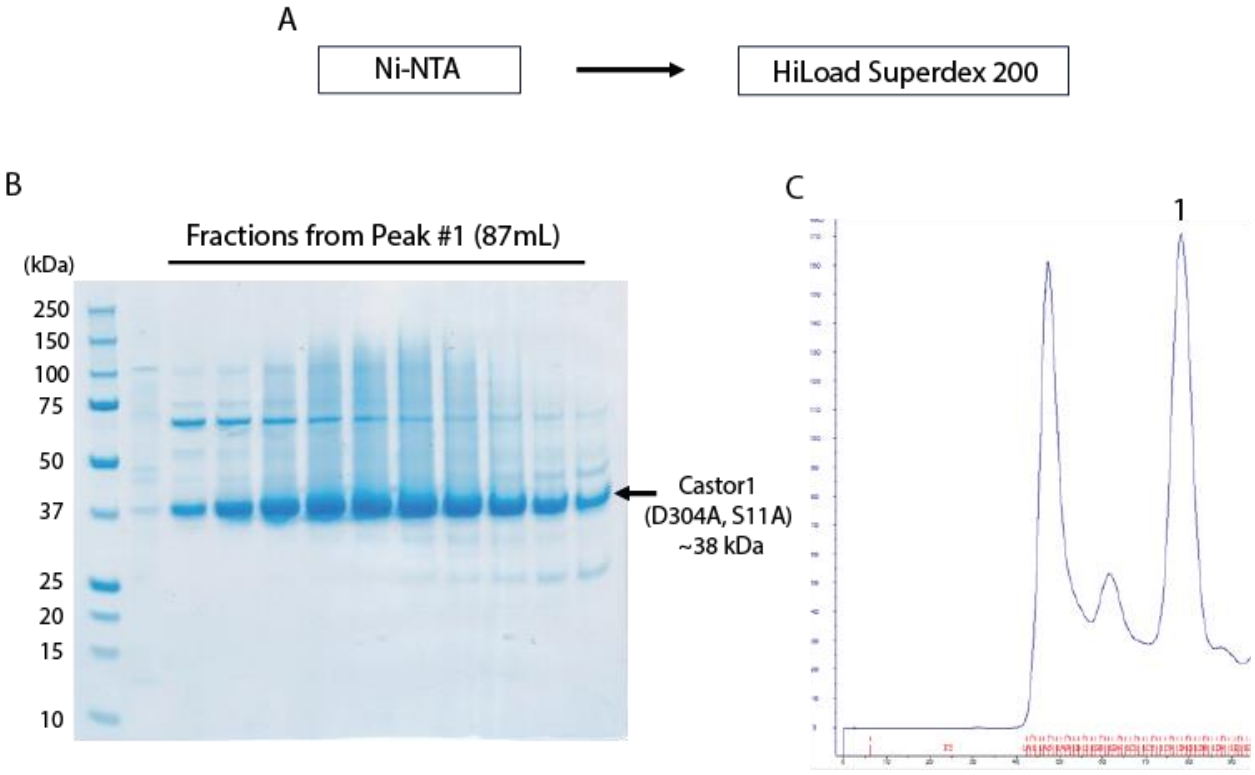


Figure 3.9 Castor1 Purification from E.Coli. (A) Representative SDS-PAGE gel and (B) Chromatogram from purification.

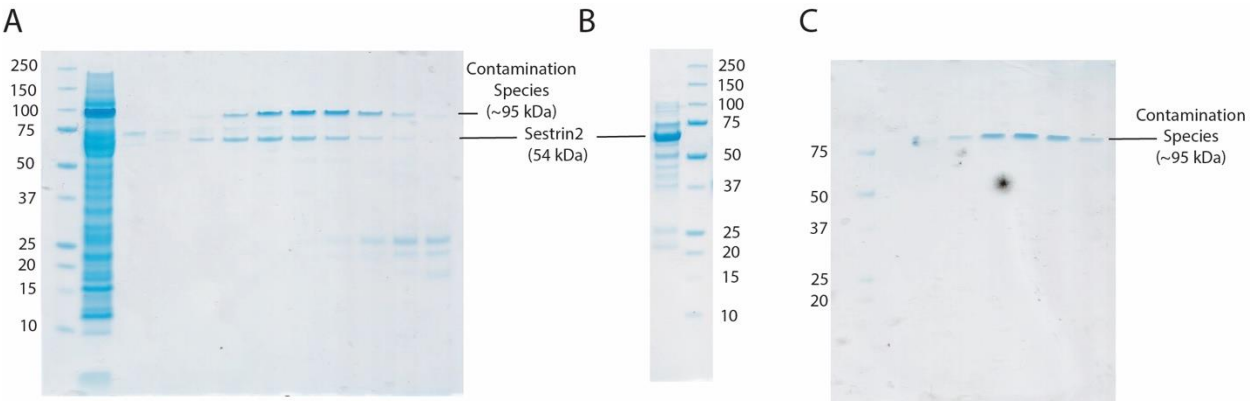


Figure 3.10 Sestrin2 Purification from *E. coli.* (A) Representative SDS-PAGE gel of fractions from HiLoad Superdex 200 and (B) Concentrated flow through (FT) from cation exchange column containing Sestrin2. (C) SDS-PAGE gel of fractions from cation exchange column step containing the contaminating species.

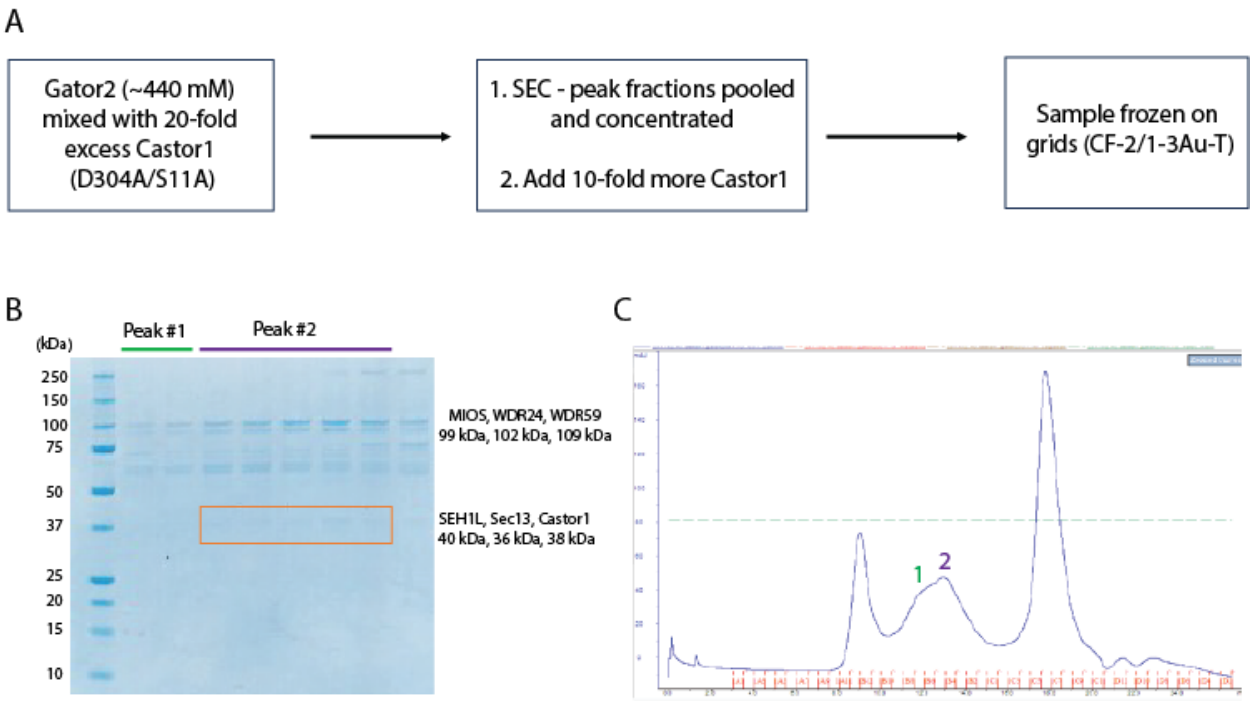


Figure 3.11 GATOR2-Castor1 complex formation. (A) Process of GATOR2-Castor1 complex formation. (B) SDS-PAGE gel from gel filtration using Superose6 increase. (C) Chromatogram from gel filtration.

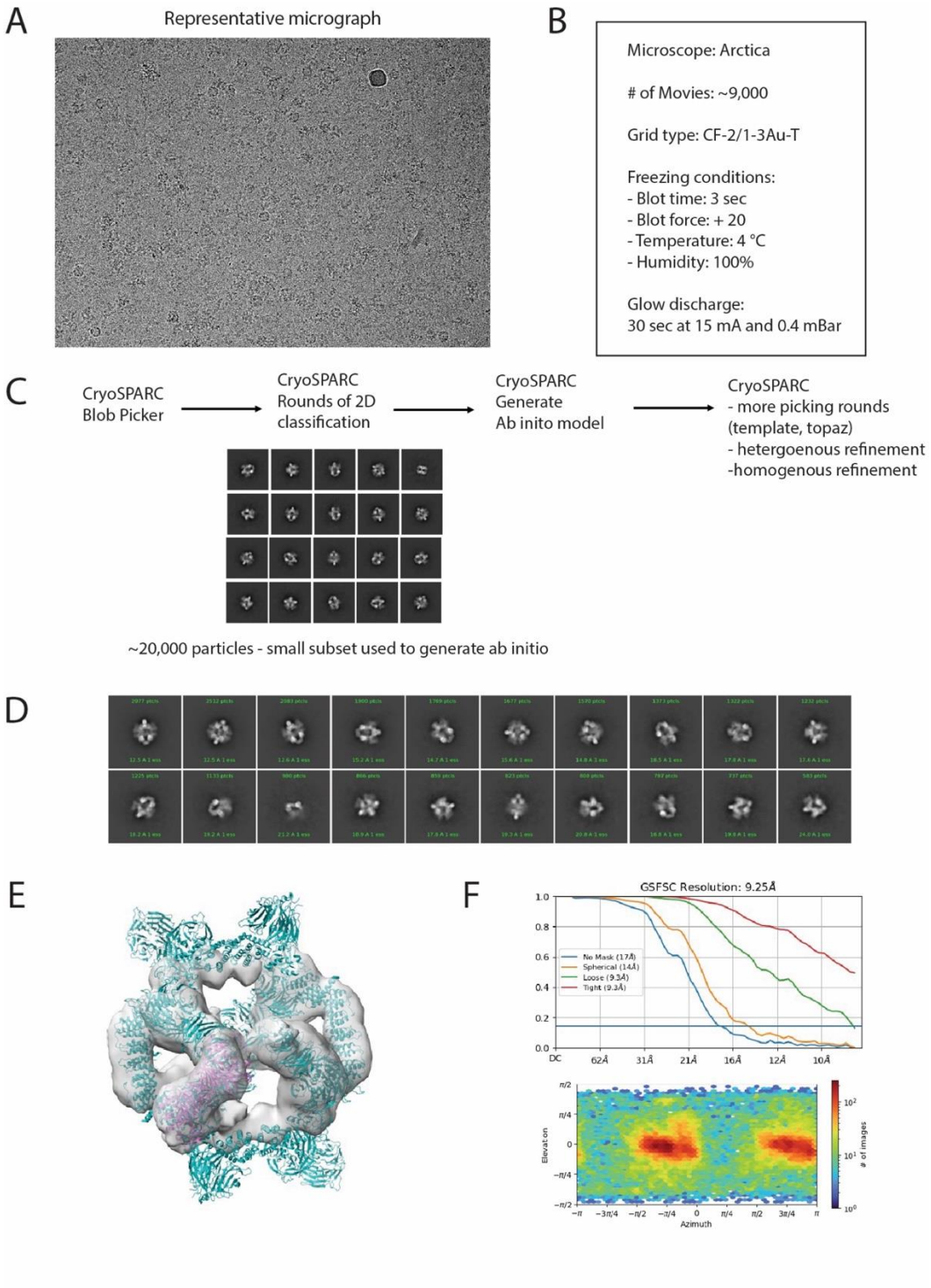


Figure 3.12 Data Processing for GATOR2-Castor1 initial map. (A) Representative micrograph from data collection. (B) Overview of data collection and grid preparation conditions. (C) Overview of data processing work from to generate particle picks. (D) Best 2D classes (E) Low resolution map of GATOR2-Castor1. Known cryo-EM and crystal structure of GATOR2 (turquoise) and Castor1 (pink) are rigid body fitted into map (grey) (F) FSC and orientation plots corresponding to map in E.

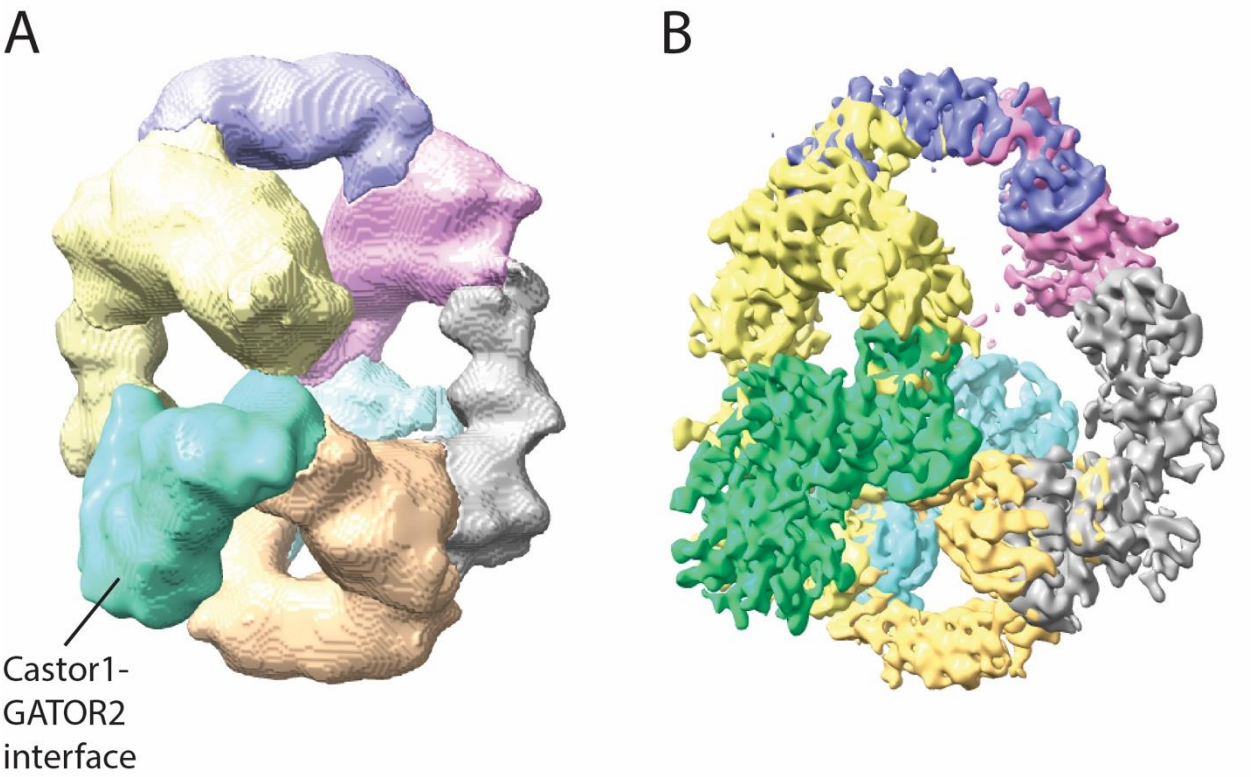


Figure 3.13 Improved resolution of GATOR2-CASTOR1 complex using local refinement. (A) Masks made from segments of the GATOR2 map. (B) Map after local refinement using mask in A. Colors correspond to the map generated using respective mask.

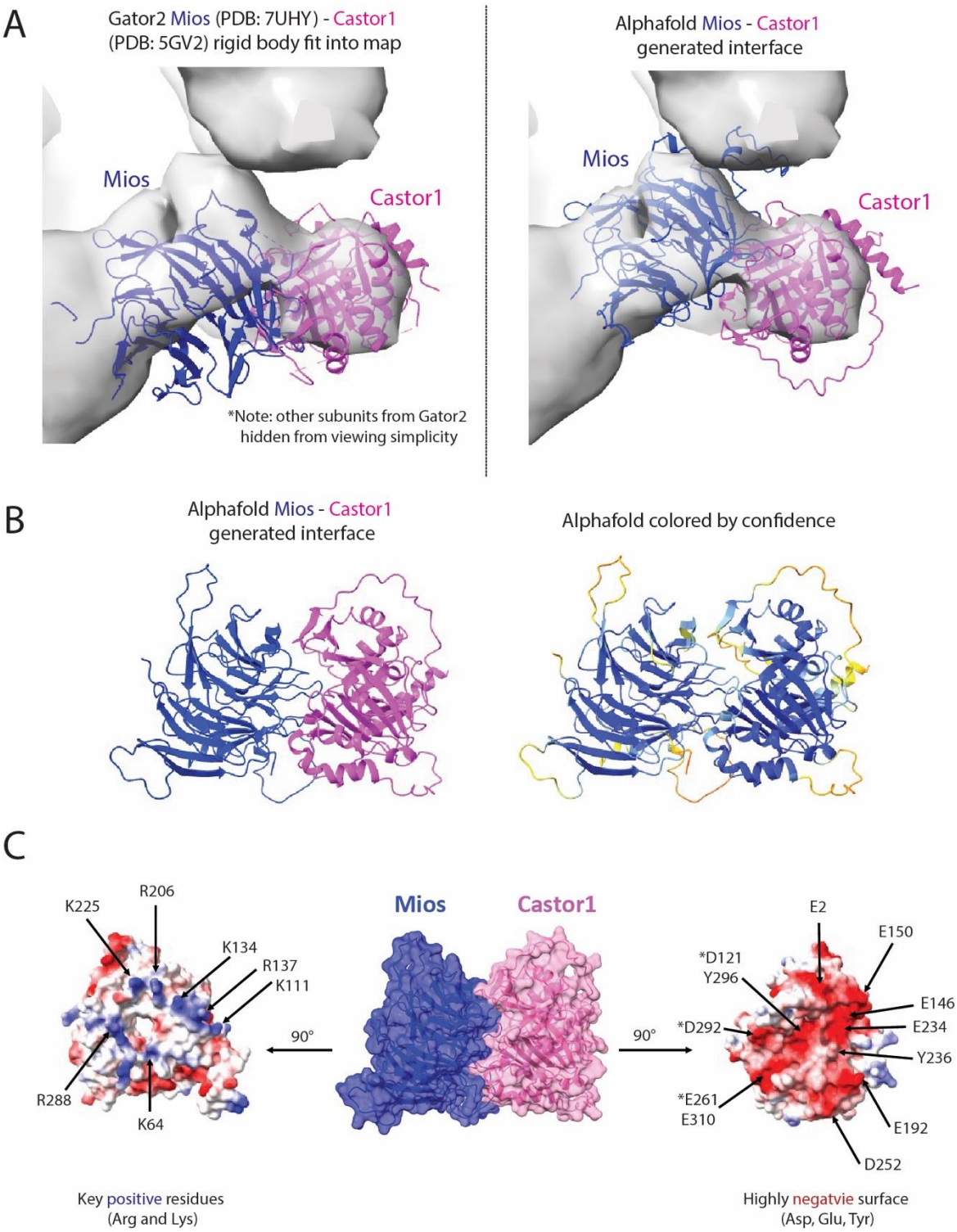


Figure 3.14 Castor1-Mios Interface Analysis. (A) Comparison of Castor1-Mios fit in cryo-EM map using rigid body fitting and Alphafold model. (B) Alphafold model generated between Castor1 monomer and MIOS beta propeller and confidence of the prediction. (C) Analysis of the electrostatics of the interface between MIOS and Castor1.

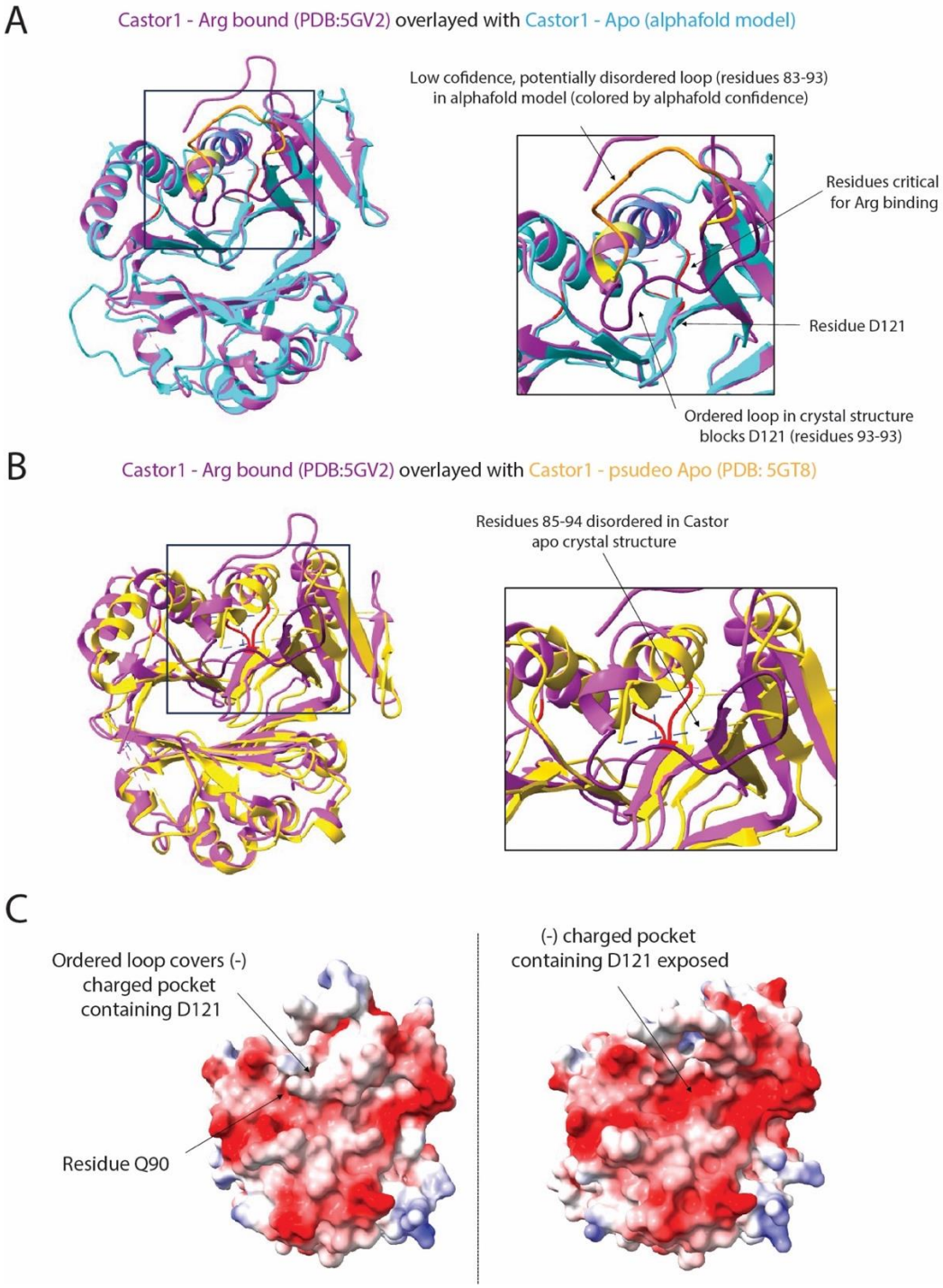


Figure 3.15 GATOR2-Castor1 Release Mechanism for Castor1-MIOS interface. (A) Overlay of a single monomer of Castor1 arginine-bound crystal structure with Castor1 monomer from alphafold model. The box depicts a difference in the maps between the two overlays. (A) Overlay of a single monomer of Castor1 arginine-bound crystal structure with monomer of Castor1 apo crystal structure. The box depicts a difference in the maps between the two overlays. (C) Comparison of the electrostatics of the Castor1 surface the interact with GATOR2 in the arg-bound (left) and apo (right) Castor1 states.

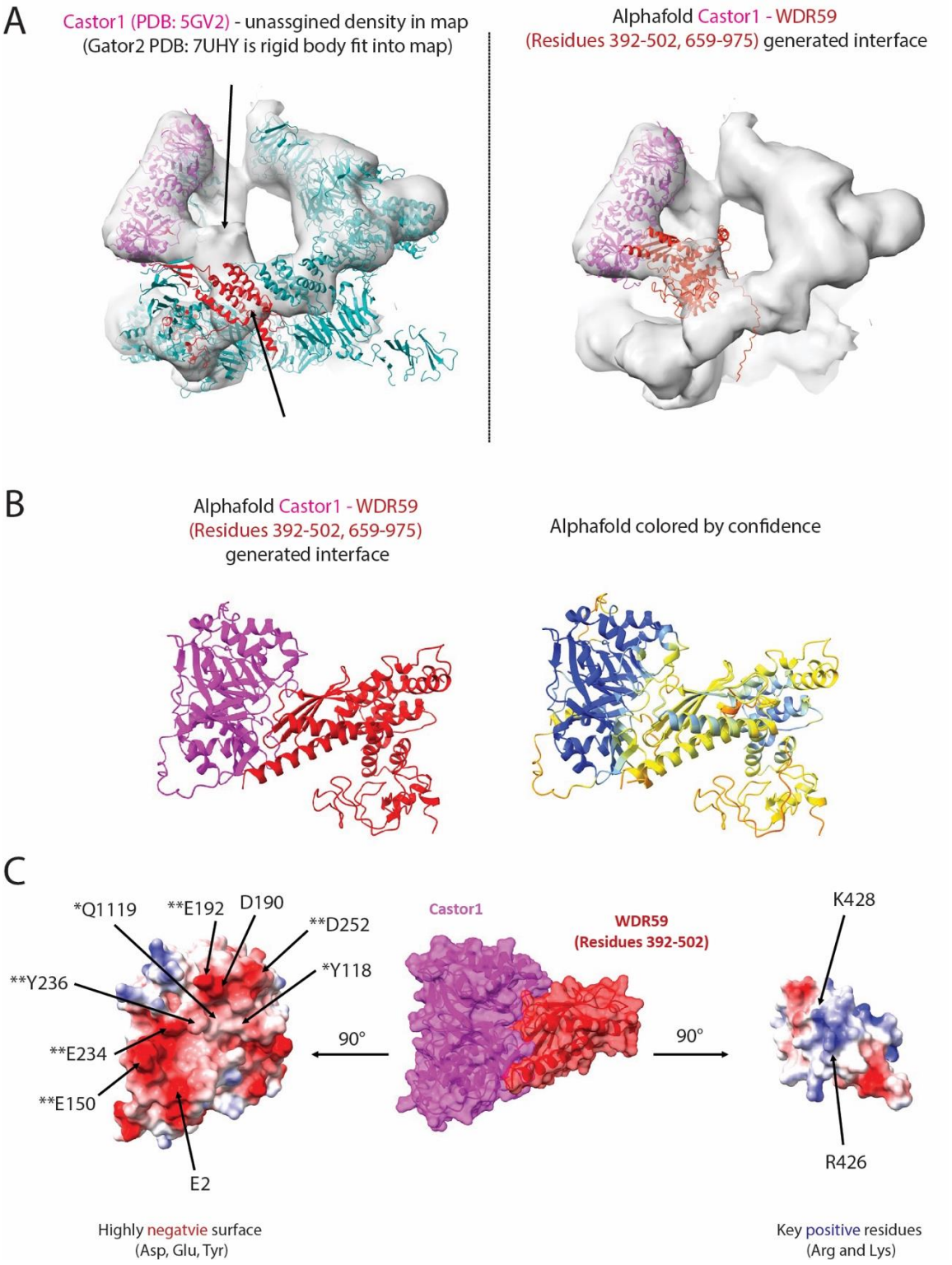


Figure 3.16 Castor1-WDR59 Interface Analysis. (A) Comparison of Castor1-WDR59 fit in cryo-EM map using rigid body fitting and Alphafold model. (B) Alphafold model generated between Castor1 monomer and portions of WDR59 and confidence of the prediction. (C) Analysis of the electrostatics of the interface between WDR59 and Castor1.

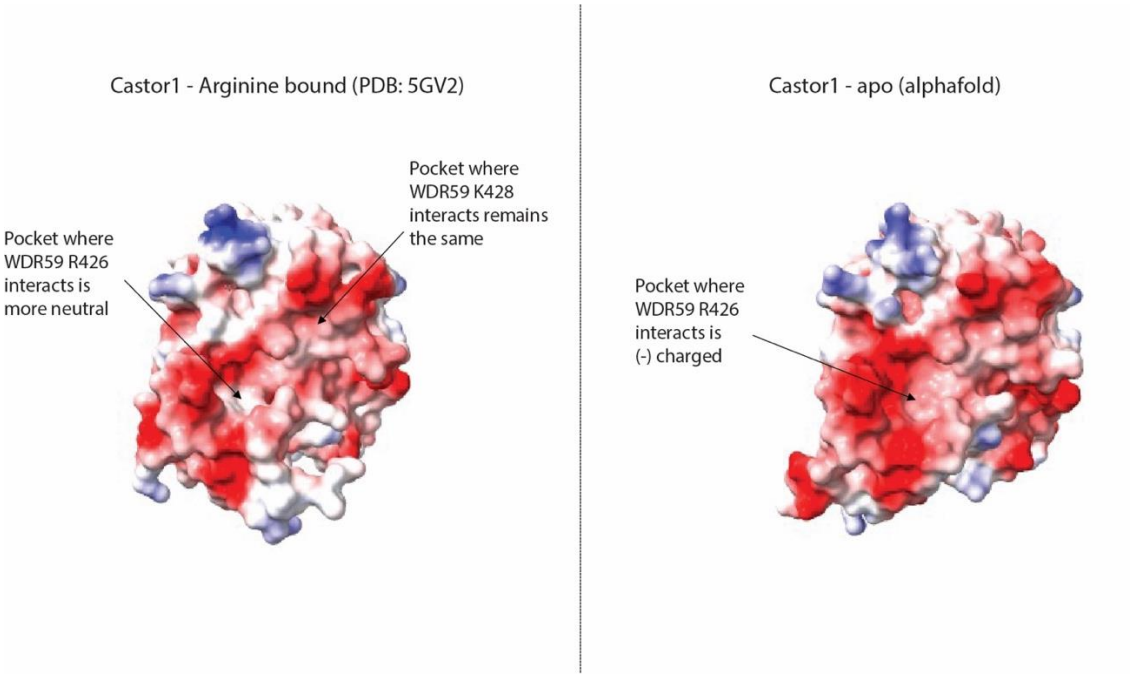


Figure 3.17 GATOR2- Castor1 Release Mechanism at Castor1-WDR59 interface.
Comparison of the electrostatics of the Castor1 surface the interact with GATOR2 in the Arg-bound (left) and apo (right) Castor1 states.

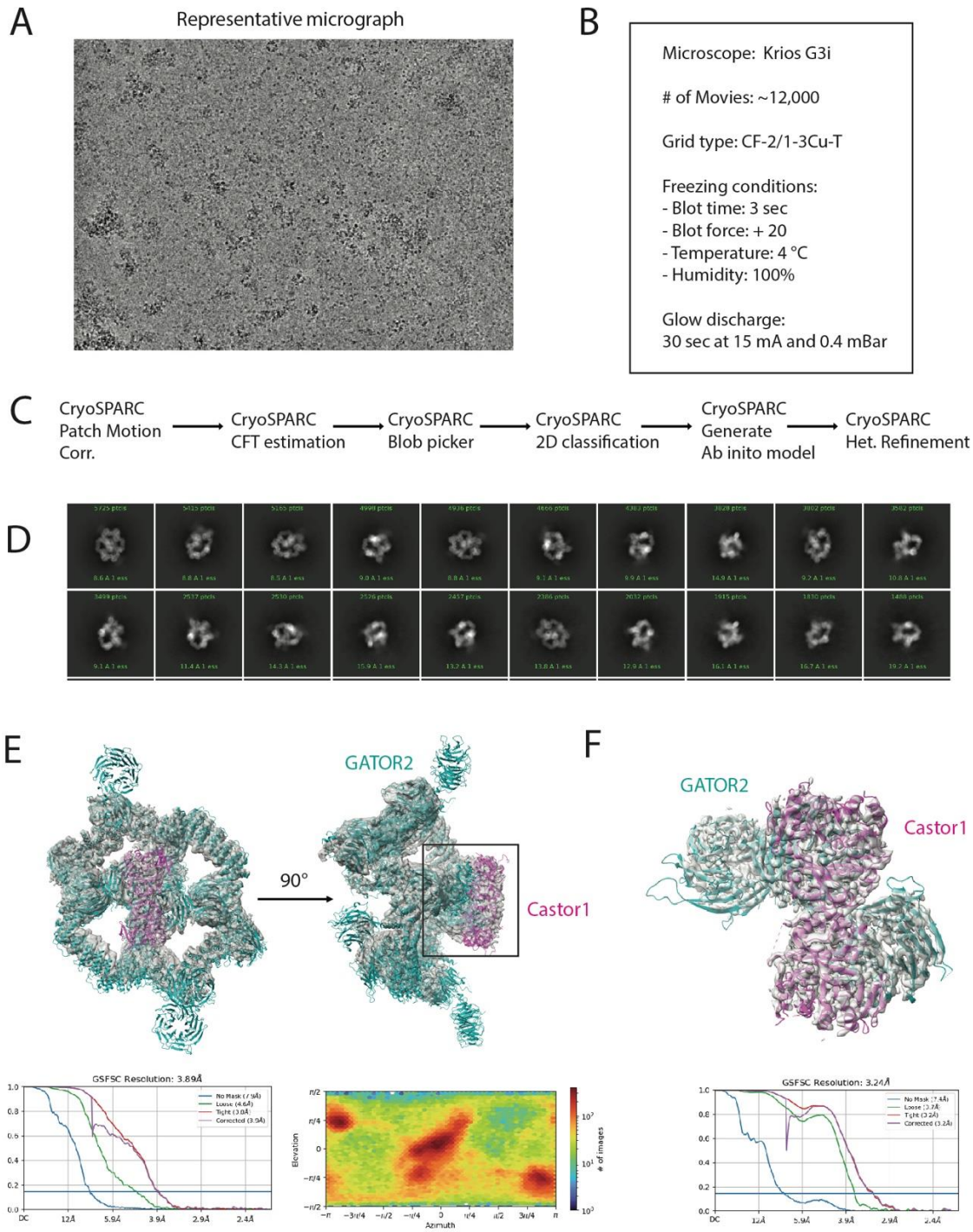


Figure 3.18. Data processing for GATOR2-Castor1 high resolution map. (A) Representative micrograph from data collection. (B) Overview of data collection and grid preparation conditions. (C) Overview of data processing work to generate best 3D map. (D) 2D classes used to generate best 3D map (E) Overall map of GATOR2-Castor1. Known cryo-EM and crystal structure of GATOR2 (turquoise) and Castor1 (pink) are rigid body fitted into map (grey). FSC and orientation plots corresponding to overall map. (F) Local refinement of GATOR2 -Castor1 interface.

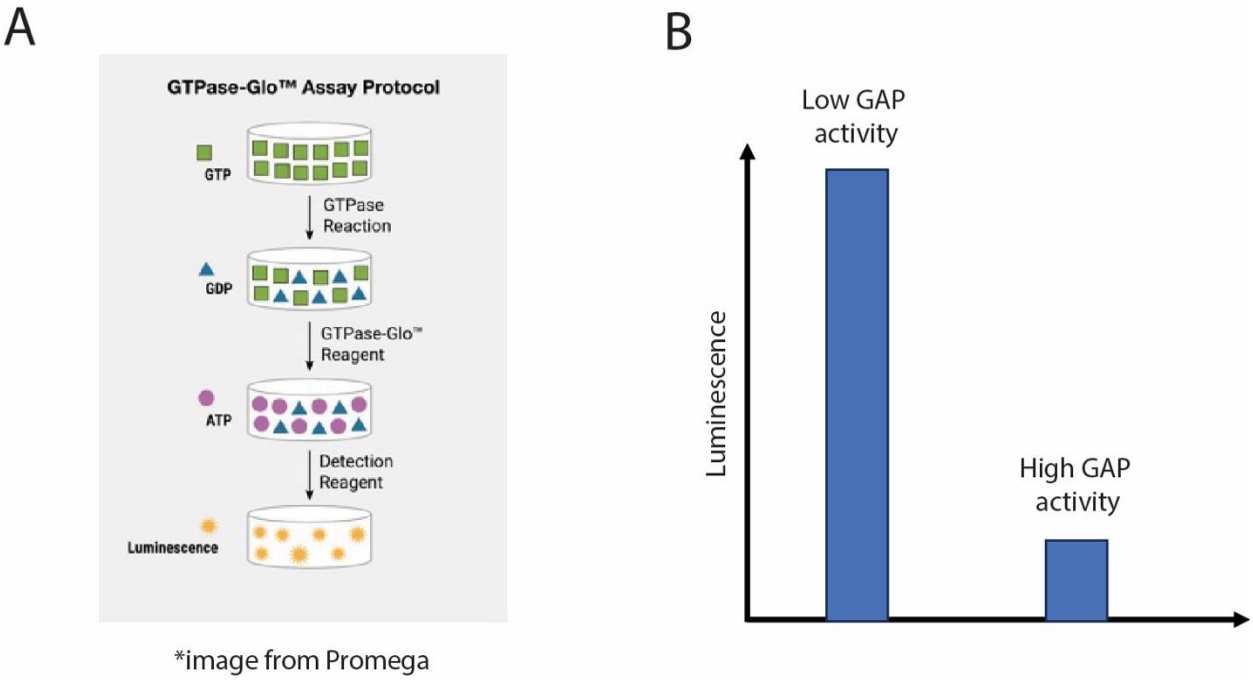


Figure 3.19 Overview of Promega GTPase Glo Assay. (A) Overview of the GTPase-Glo Assay protocol provide from Promega. (B) Example of the results from the GTPase Glo assay. High luminescence values correspond with low GAP activity while low luminescence correspond with high GAP activity.

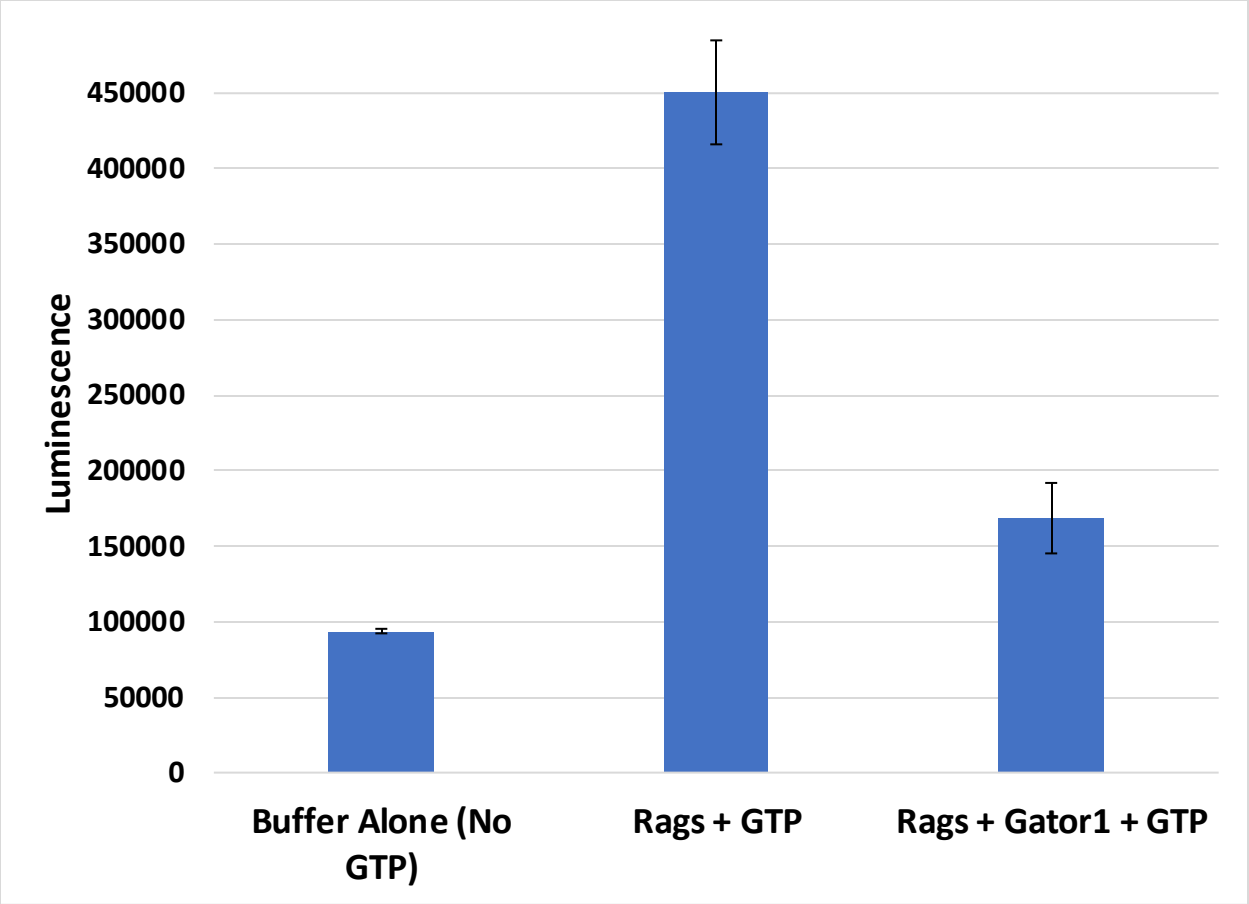


Figure 3.20 GTPase Glo GATOR1 GAP activity towards Rag GTPases. Reactions set up for GTPase Glo using 96 well plate. N=3 for each reaction. The error bars represent the SEM.

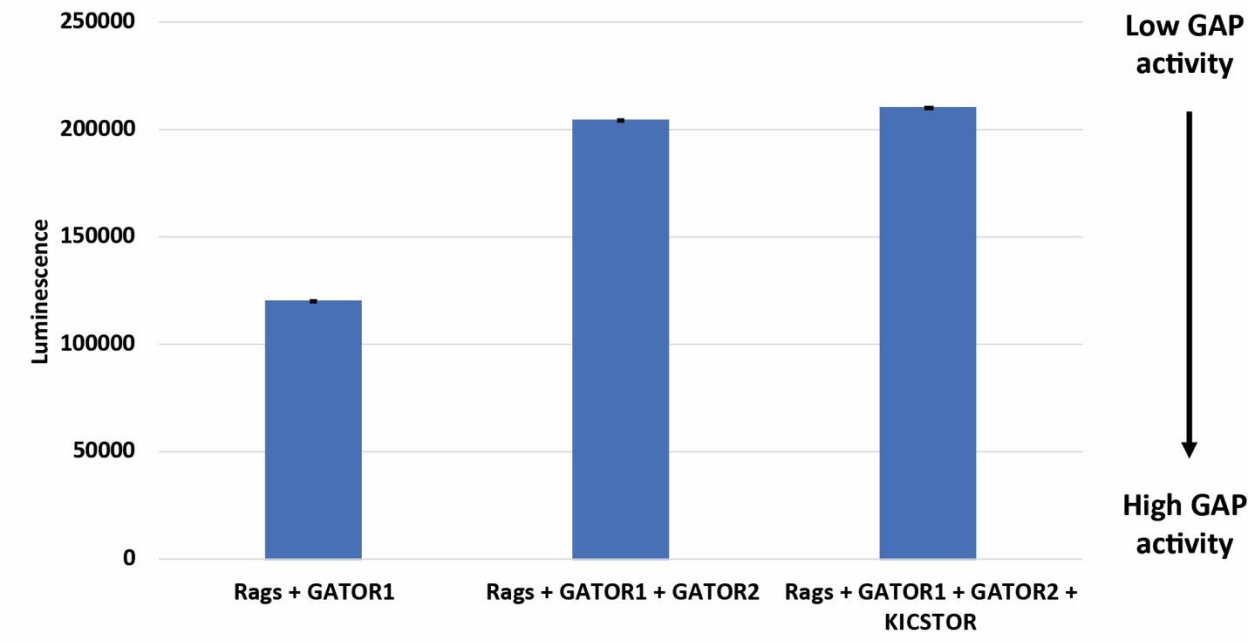


Figure 3.21 GTPase Glo GATOR1 GAP activity towards Rag GTPases is altered by GATOR2. Reactions set up for GTPase Glo using 96 well plate. N=3 for each reaction. The error bars represent the SEM.

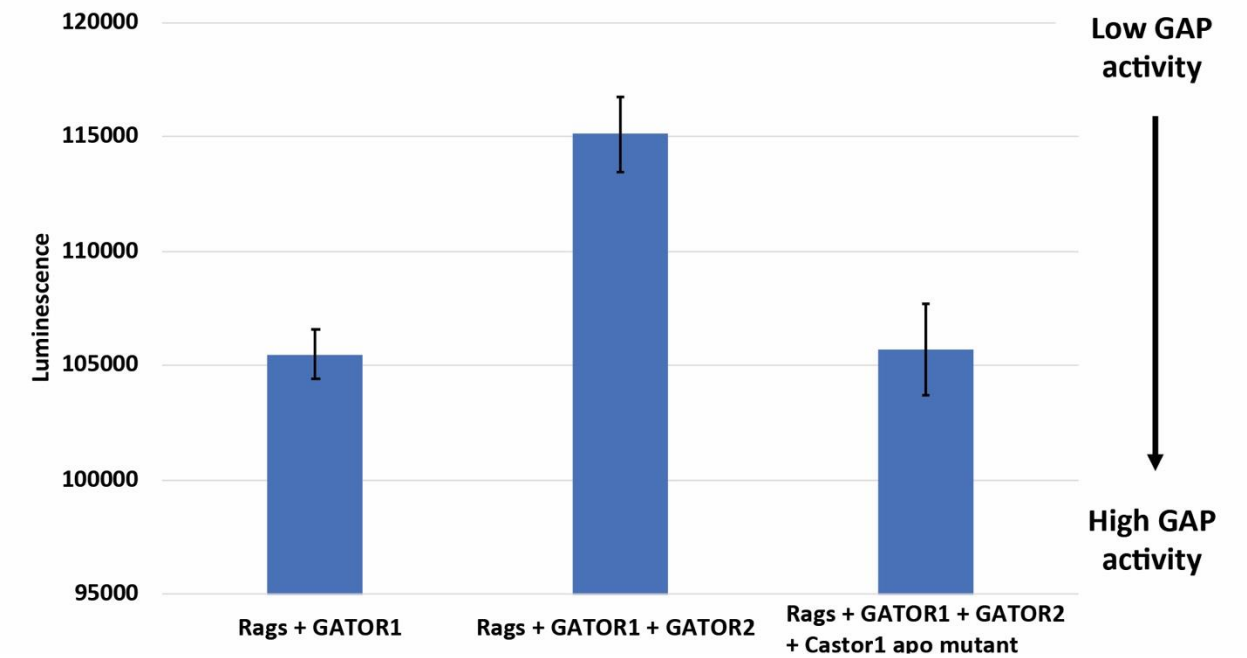


Figure 3.22 Castor1 restores GATOR1 GAP activity in presence of GATOR2 in GTPase Glo Assay. Reactions set up for GTPase Glo using 384 well plate. N=3 for each reaction. The error bars represent the SEM.

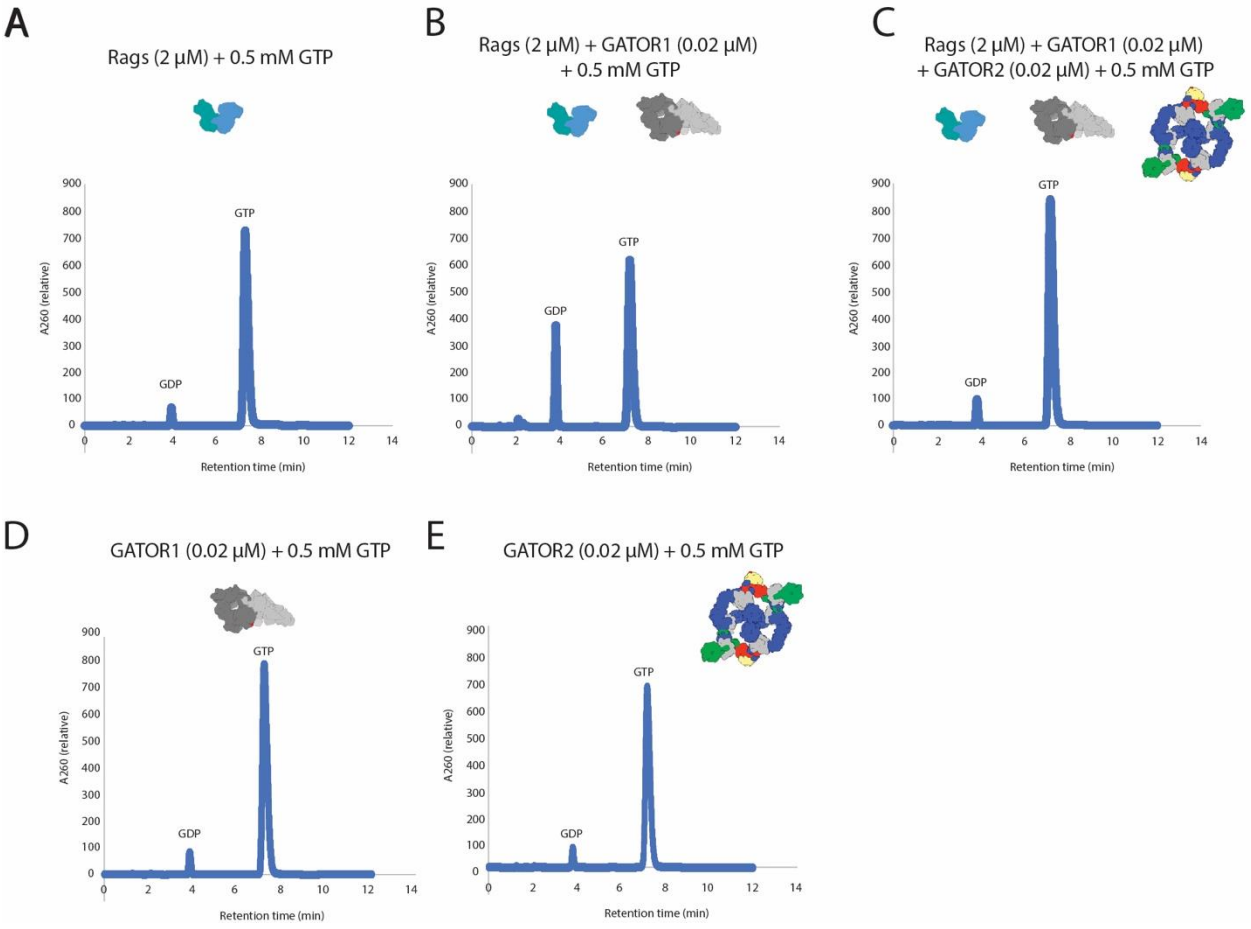


Figure 3.23 HPLC Based GAP Activity Assay for GATOR1. Resulting chromatograms from the HPLC for reactions containing (A) Rags alone (B) Rags + GATOR1 (C) Rags + GATOR1 + GATOR2 (D) GATOR1 alone (E) GATOR2 alone.

References

1. Bar-Peled, L., Chantranupong, L., Cherniack, A.D., Chen, W.W., Ottina, K.A., Grabiner, B.C., Spear, E.D., Carter, S.L., Meyerson, M., and Sabatini, D.M. (2013). A Tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. Science *340*, 1100-1106. 10.1126/science.1232044.
2. Chantranupong, L., Wolfson, R.L., Orozco, J.M., Saxton, R.A., Scaria, S.M., Bar-Peled, L., Spooner, E., Isasa, M., Gygi, S.P., and Sabatini, D.M. (2014). The Sestrins interact with GATOR2 to negatively regulate the amino-acid-sensing pathway upstream of mTORC1. Cell Rep *9*, 1-8. 10.1016/j.celrep.2014.09.014.
3. Chantranupong, L., Scaria, S.M., Saxton, R.A., Gygi, M.P., Shen, K., Wyant, G.A., Wang, T., Harper, J.W., Gygi, S.P., and Sabatini, D.M. (2016). The CASTOR Proteins Are Arginine Sensors for the mTORC1 Pathway. Cell *165*, 153-164. 10.1016/j.cell.2016.02.035.
4. Egri, S.B., Ouch, C., Chou, H.T., Yu, Z., Song, K., Xu, C., and Shen, K. (2022). Cryo-EM structures of the human GATOR1-Rag-Ragulator complex reveal a spatial-constraint regulated GAP mechanism. Mol Cell. 10.1016/j.molcel.2022.03.002.
5. Gai, Z., Wang, Q., Yang, C., Wang, L., Deng, W., and Wu, G. (2016). Structural mechanism for the arginine sensing and regulation of CASTOR1 in the mTORC1 signaling pathway. Cell Discov *2*, 16051. 10.1038/celldisc.2016.51.
6. Saxton, R.A., Chantranupong, L., Knockenhauer, K.E., Schwartz, T.U., and Sabatini, D.M. (2016). Mechanism of arginine sensing by CASTOR1 upstream of mTORC1. Nature *536*, 229-233. 10.1038/nature19079.
7. Shen, K., Huang, R.K., Brignole, E.J., Condon, K.J., Valenstein, M.L., Chantranupong, L., Bomaliyamu, A., Choe, A., Hong, C., Yu, Z., and Sabatini, D.M. (2018). Architecture of the human GATOR1 and GATOR1-Rag GTPases complexes. Nature *556*, 64-69. 10.1038/nature26158.
8. Valenstein, M.L., Rogala, K.B., Lalgudi, P.V., Brignole, E.J., Gu, X., Saxton, R.A., Chantranupong, L., Kolibius, J., Quast, J.P., and Sabatini, D.M. (2022). Structure of the nutrient-sensing hub GATOR2. Nature *607*, 610-616. 10.1038/s41586-022-04939-z.
9. Xia, J., Wang, R., Zhang, T., and Ding, J. (2016). Structural insight into the arginine-binding specificity of CASTOR1 in amino acid-dependent mTORC1 signaling. Cell Discov *2*, 16035. 10.1038/celldisc.2016.35.
10. Zhou, Y., Wang, C., Xiao, Q., and Guo, L. (2019). Crystal structures of arginine sensor CASTOR1 in arginine-bound and ligand free states. Biochem Biophys Res Commun *508*, 387-391. 10.1016/j.bbrc.2018.11.147.
11. Saxton, R.A., Knockenhauer, K.E., Wolfson, R.L., Chantranupong, L., Pacold, M.E., Wang, T., Schwartz, T.U., and Sabatini, D.M. (2016). Structural basis for leucine sensing by the Sestrin2-mTORC1 pathway. Science *351*, 53-58. 10.1126/science.aad2087.
12. Saxton, R.A., Knockenhauer, K.E., Schwartz, T.U., and Sabatini, D.M. (2016). The apo-structure of the leucine sensor Sestrin2 is still elusive. Sci Signal *9*, ra92. 10.1126/scisignal.aah4497.
13. Tafur, L., Hinterndorfer, K., Gabus, C., Lamanna, C., Bergmann, A., Sadian, Y., Hamdi, F., Kyrilis, F.L., Kastritis, P.L., and Loewith, R. (2022). Cryo-EM structure of the SEA complex. Nature *611*, 399-404. 10.1038/s41586-022-05370-0.
14. Wolfson, R.L., Chantranupong, L., Wyant, G.A., Gu, X., Orozco, J.M., Shen, K., Condon, K.J., Petri, S., Kedir, J., Scaria, S.M., et al. (2017). KICSTOR recruits GATOR1

to the lysosome and is necessary for nutrients to regulate mTORC1. Nature *543*, 438-442. 10.1038/nature21423.
15. Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol *180*, 519-530. 10.1016/j.jsb.2012.09.006.
16. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods *14*, 290-296. 10.1038/nmeth.4169.
17. Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods *14*, 331-332. 10.1038/nmeth.4193.
18. Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A.J., and Berger, B. (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. Nat Methods *16*, 1153-1160. 10.1038/s41592-019-0575-8.
19. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583-589. 10.1038/s41586-021-03819-2.
20. Mirdita, M., Schutze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. Nat Methods *19*, 679-682. 10.1038/s41592-022-01488-1.

Chapter 4: Regulation of GATOR1 and GATOR2 by Membrane Interacting Protein Complex KICSTOR

Introduction

Nutrient-dependent recruitment of mTORC1 to the lysosomal membrane is a critical step in the activation of mTORC1 kinase activity. This recruitment is facilitated by the Rag-Ragulator complex, comprised of two Rag GTPases and a membrane scaffold protein complex Ragulator. The Rag GTPases convert between two stable nucleotide states, inactive (RagA or B^{GDP}:RagC or D^{GTP}) in low nutrients and active (RagA or B^{GTP}:RagC or D^{GDP}) in high nutrients, dictating mTORC1 recruitment. The mTORC1 complex is recruited to the membrane by the active Rag-Ragulator complex. Two GTPase activating proteins (GAPs), GATOR1 and FLCN, facilitate the conversion between the two different states, stimulating the hydrolysis of GTP in RagA and RagC respectively. While significant process has been made to understand the activation of the complex by the RagC GAP FLCN, including that described in Chapter 2, inactivation of the complex by GATOR1 presents a more intricate and less understood scenario. Large protein complexes, KICSTOR and GATOR2, have been shown to interact with and alter the activity of GATOR1. In addition, signals about nutrient availability are integrated into this axis by smaller nutrient sensors, Sestrin2, Castor1 and SAMTOR that dynamically interact with the complexes (Figure 4.1)[1-6].

The KICSTOR complex, a core regulator of GATOR1, is composed of four subunits SZT2 (~350 kDa), KPTN (~50 kDa), KICS2 (or c12orf66) (~50 kDa), and ITFG2 (~50 kDa)[4]. The complex was shown to interact with GATOR1 and stimulate its GAP activity through facilitating GATOR1 recruitment to the lysosomal membrane[4,7]. However, the two papers that identified these finding presented conflicting evidence on the role of GATOR2 in the KICSTOR-GATOR1 membrane interaction. Wolfson et al. proposed that KICSTOR was sufficient to recruit GATOR1 to the membrane and the interaction with GATOR2 was indirect[4]. On the other hand, Peng et al. suggested the three complexes formed a synergistic higher order complex where all three complexes were necessary to mediate membrane recruitment[7]. While both papers put forth that KICSTOR localizes to lysosomes, it is unclear if KICSTOR localization is a result of direct or indirect interactions with the membrane, and if direct which subunit is mediating this interaction.

Despite these initial discoveries, no follow up in vitro investigations have been conducted to clarify the structure and function of KICSTOR. This is likely due to challenges with purifying and structurally characterizing the KICSTOR complex, which totals over 500 kDa. Structural studies will be helpful to reveal the potential membrane interaction site and to clarify which subunit is necessary for membrane recruitment. In addition, they may provide insights into the role of KICSTOR as a regulator of GATOR1 and GATOR2. Complimenting structural analysis with in vitro assays will clarify if KICSTOR directly or indirectly interacts with the GATOR complexes. In yeast, the GATOR1-GATOR2 interaction occurs independent of KICSTOR, which is not present[8]. It remains unclear if in humans KICSTOR is simply a scaffold protein, holding together GATOR1 and GATOR2, or if it plays a more active role in the pathway.

Designing DNA Constructs for KICSTOR

To address these unresolved questions, a viable, purified sample of the protein complex was obtained. This involved several steps, including determining DNA constructs for the complex, selecting an appropriate expression system, identifying purification tags, and determining a purification workflow. Other large protein complexes in this pathway, including FLCN and GATOR1, were purified using codon-optimized genes, the pCAG expression vector, and the HEK 293 GnTI expression system. This gave a good starting point for determining a KICSTOR purification scheme. Genes encoding KICSTOR subunits were synthesized by Twist Biosciences and codon-optimized for mammalian cell expression. These codon-optimized genes were cloned into the pCAG expression vector, incorporating various purification tags such as glutathione S-transferase (GST), STREP, and FLAG (Figure 4.2).

Small Scale Expression Tests for KICSTOR

Small-scale expression tests were conducted with the new constructs to determine optimal transfection conditions, including which subunits to co-transfect, the ideal tag and tag location for purification (Figure 4.3). To perform small scale expression tests, 50 mL of human embryonic kidney (HEK) 293 GnTI cells were transfected with 0.05 mg of DNA and polyethyleneimine (PEI) at a 1:4 ratio (DNA:PEI). The cells were harvested after 48 hours and resuspended in lysis buffer (25 mM HEPES, 130 mM NaCl, 2 mM EGTA, 2.5 mM $MgCl_2$, 0.5 mM TCEP and 1 protease inhibitor tablet per 50 mL) containing 1% TritonX-100 and rocked for 1 hour at 4°C. The cells were then centrifuged for 30 minutes at 15,000xg. The clarified lysate was incubated with 30 µL of equilibrated resin for 1 hour rocking at 4°C. The resin was washed 3X with 500 µL wash buffer (25 mM HEPES, 130 mM NaCl, 2 mM EGTA, 2.5 mM $MgCl_2$, 0.5 mM TCEP). Initially, the resin was heated to denature bound proteins before visualization using sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE). However, for later small-scales expression tests, the proteins were eluted from the resin prior to SDS-PAGE visualization to ensure the protein was not aggregated on the resin.

In the initial small-scale tests, individual expression of GST-KPTN, GST-ITFG2, and GST-KICS2 was performed to assess whether endogenous amounts of other subunits could be pulled down or if individual subunits could be purified independently. If the complex components were able to be purified independently it could be possible to form the full KICSTOR complex afterwards. The results from these tests showed that GST-KPTN could pull down endogenous ITFG2, and GST-ITFG2 could pull down endogenous KPTN (Figure 4.3A). It was unclear from these tests if endogenous SZT2 and KICS2 were pulled down as well. In addition, GST-KICS2 demonstrated good independent expression.

Follow up small scale tests involved co-transfecting KPTN and ITFG2 to assess if the proteins could be purified independently of the full KICSTOR complex. It was observed that GST-KPTN/ITFG2 and GST-ITFG2/KPTN co-expression yielded well-expressed dimers in stoichiometric amounts independent of the SZT2 and KICS2 subunits (Figure 4.3B). Later expression tests were conducted to include expression with the larger SZT2 subunit. For these tests the protein was eluted from the resin prior to visualization using SDS-PAGE. GST-SZT2 was co-transfected with KPTN, ITFG2, and KICS2 at a ratio of 1:1:1:1 and 2:1:1:1. The most

complex was expressed from cells that were co-transfected at a ratio of 2:1:1:1, along with the GST-tag on SZT2 (Figure 4.3C). To verify that all the subunits were present in the sample, the bands from the gel were analyzed via mass spectrometry confirming the largest band contained the SZT2 protein, and the thicker, smaller band contained the remaining three subunits (Figure 4.4).

Using a GST tag presented limitations in later experiments due to the additional cleavage and tag removal steps which prompted small scale expression tests using the SZT2 construct tagged with STREP. The 2X-STREP-FLAG SZT2 construct was co-transfected with untagged KPTN, ITFG2, and KICS2 at a ratio of 1:1:1:1 (Figure 4.3D). These transfection conditions led to full complex formation and was best for subsequent structural studies and activity assays, particularly for cryo-EM studies where leaving the GST tag on the protein was not ideal. The results from all small-scale expression tests can be found in Figure 4.3.

Large Scale Purification of KICSTOR

Method #1: Purification of Full-Length KICSTOR using GST-SZT2

Following initial small scale expression tests, full-length KICSTOR was purified using the GST tag on the SZT2 subunit. The additional three subunits (KPTN, ITFG2 and KICS2) contained a 2X STREP-FLAG tag but these tags were not utilized for this purification method. In latter purification attempts, described below, they were used. The purification workflow for GST-KICSTOR began with co-transfection of plasmids (1 mg total) containing the KICSTOR subunits (ratio of 2:1:1:1 → GST-SZT2:KPTN:ITFG2:KICS2) into 1 L HEK 293 GnTI cells. The cells were harvested after 72 hours and resuspended in lysis buffer containing 25 mM HEPES pH 7.4, 130 mM NaCl, 2.5 mM $MgCl_2$, 2 mM EGTA, 1% Triton-X, 0.5 mM TCEP and protease inhibitor pellet (1 pellet per 50 mL buffer). The cells were lysed via rocking for 1 hour at 4°C prior to centrifugation at 37,000xg for 30 minutes.

The clarified lysate was incubated with ~3 mL pre-equilibrated GST resin for a minimum of 3 hours (maximum overnight incubation). The resin was washed with 40 mL of wash buffer A (25 mM HEPES pH 7.4, 130 mM NaCl, 2.5 mM $MgCl_2$, 2 mM EGTA, 1% Triton-X, 0.5 mM TCEP), 40 mL of buffer B (25 mM HEPES pH 7.4, 500 mM NaCl, 2.5 mM $MgCl_2$, 2 mM EGTA, 1% Triton-X, 0.5 mM TCEP), 40 mL of buffer C (25 mM HEPES pH 7.4, 130 mM NaCl, 2.5 mM $MgCl_2$, 2 mM EGTA, 0.1% CHAPS, 0.5 mM TCEP) The protein complex was eluted in ~20 mL of elution buffer (25 mM HEPES pH 7.4, 130 mM NaCl, 2.5 mM $MgCl_2$, 2 mM EGTA, 0.1% CHAPS, 0.5 mM TCEP, 50 mM GSH). Before elution, the resin was incubated for 30 minutes on ice with elution buffer. The sample was incubated overnight with tobacco etch virus (TEV) protease to cleave the GST tag. The eluted sample was concentrated to 1 mL using 100 kDa Millipore Sigma Amicon centrifugal filter for grid preparation. (Figure 4.5). This purification method was used for cryo-EM analysis on the Arctica and Krios G3 microscopes. Improvements to this purification method were made to increase yield and purity of the sample. These updated methods are described below.

Method #2: Purification of Full-Length KICSTOR using STEP-SZT2

A second method for purifying full-length KICSTOR involved utilizing a STREP tag on the SZT2 subunit and leaving the smaller subunits (KPTN, ITFG2, and KICS2) untagged. The purification workflow for STREP-KICSTOR began with co-transfection of plasmids (1 mg total) containing the KICSTOR subunits into 1 L HEK 293 GnTI cells. The cells were harvested after 48 hours and resuspended in lysis buffer containing 25 mM HEPES pH 7.5, 500 mM NaCl, 2 mM $MgCl_2$, 10% glycerol, 1 mM TCEP and protease inhibitor pellet (1 pellet per 50 mL buffer). The cells were homogenized via douncing using the B pistol 20X. Next, 10% DDM:CHS (10:1) was added to the homogenized cells for a final concentration of 1% DDM:CHS (10:1). The homogenized cells were incubated while rocking for 1 hour at 4°C prior to centrifugation at 17,000 rpm for 35 minutes.

The clarified lysate was incubated with 3-5 mL pre-equilibrated STREP resin for a minimum of 3 hours (maximum overnight incubation). The resin was washed with 3-5 column volumes (CV) of high salt wash buffer (25 mM HEPES pH 7.5, 500 mM NaCl, 2 mM $MgCl_2$, 1 mM TCEP) containing 0.03% DDM:CHS (10:1), 3-5 CV low salt wash buffer (25 mM HEPES pH 7.5, 200 mM NaCl, 2 mM $MgCl_2$, 1 mM TCEP) containing 0.03% DDM:CHS (10:1), and 3-5 CV of low salt wash buffer without detergent. The protein complex was eluted in ~20-30 mL of low salt wash buffer containing 4 mM desthiobiotin. Before elution, the resin was incubated for 30 minutes on ice with elution buffer. The eluted sample was concentrated to 1 mL using 100 kDa Millipore Sigma Amicon centrifugal filter for gel filtration with a Superose 6 increase column. Gel filtration buffer contained 25 mM HEPES pH 7.5, 200 mM NaCl, 2 mM $MgCl_2$, and 1 mM TCEP. The protein complex eluted ~11.8 mL and the fractions were visualized with SDS-PAGE. Representative SDS-PAGE of fractions from Superose 6 and chromatogram are shown in Figure 4.6.

Method #3: Purification of GST-KICSTOR using STREP-KPTN, STREP-ITFG2 and STREP-KICS2

A second approach for purifying a full-length GST tagged KICSTOR complex was established to get higher yields and cleaner protein preps. Similar to the first method, SZT2 was tagged with a GST, while the remaining three smaller subunits (KPTN, ITFG2 and KICS2) contained 2X-STREP tag. This configuration enabled purification of the complex using STREP resin. The purification protocol for GST-KICSTOR followed the same steps as described above for the purification with STREP-SZT2. Representative SDS-PAGE of fractions from Superose 6 and chromatogram as shown in Figure 4.7. The GST-KICSTOR complex obtained through this method was used in pull-down assays to investigate the interaction between KICSTOR and GATOR complexes.

Purification of ITFG2-KPTN Dimer

Prior to the purification of the full KICSTOR complex, an attempt was made to purify the GST-ITFG2-KPTN dimer independently. Small scale expression tests indicated that these proteins could be co-expressed and purified together in high yields independent of the other two complex proteins. These proteins were purified as described in purification method #1 for KICSTOR. The GST-tag on the ITFG2 subunit was used to isolate the complex prior to TEV

cleavage and gel filtration. Representative SDS-PAGE of fractions from Superose 6 and chromatogram as shown in Figure 4.8.

Single Particle Cryo-EM for Full-Length KICSTOR Complex

Initial attempts to structurally characterize the KICSTOR complex used purification method #1. The yields from the initial purification method were low, posing challenges for gel filtration due to additional sample loss. To circumvent this issue, the sample was frozen on grids directly after elution from the GST resin and TEV cleavage. However, this approach resulted in a heterogenous sample comprising full-length KICSTOR, GST, TEV and a background of glutathione (GSH) in the buffer. To increase the number of particles on the grid, this sample was frozen on graphene grids prepared by Zhicheng Cui using a blot time of 6 seconds, wait time 30 seconds, blotch for +20, temperature 4°C and 100% humidity. The glow discharge conditions for the graphene grids were 20 seconds in air at 5 mA and 0.4 mbar. Two data sets were collected resulting in ~2,400 movies from the Arctica microscope and ~6,400 movies from the Krios G3 microscope. CryoSPARC was used for motion correction and CTF estimation[9]. Particles from the data sets were picked using cryoSPARC blob and template picker. The resulting 2D class averages were smaller than the predicted ordered mass of the complex and did not have clear secondary structure visible at this stage resulting in no clear 3D map.

Comparing the experimental 2Ds to those generated as a back projection from the Alphafold model, it was clear that they didn't represent the full complex. While the back projected 2D class averages resembled the experimental particles on the micrograph, adopting a thin U shape, this feature was not observed in the experimental 2D class averages (Figure 4.9). This observation suggested that the sample may possess intrinsic flexibility, making alignment challenging in 2D. In efforts to improve sample quality and reduce heterogeneity, an additional dataset was collected using freshly purified STREP-KICSTOR, eliminating the need for a TEV cleavage step. Purified STREP-KICSTOR complex was frozen on CF-2/1-3Cu-T grids using a blot time of 3 seconds, blot force of +20, temperature 4°C and 100% humidity. A dataset containing ~3,600 movies was collected using the Arctica microscope. Following motion correction in Relion with MotionCor2, further data processing was conducted in cryoSPARC[9-11]. Initial particle picking was performed using blob and manual picker methods, with templates selected from resulting 2D classes. These templates were then utilized to generate picks with cryoSPARC template picker. Subsequently, a Topaz picking model was trained, resulting in the extraction of approximately 450,000 particles for 2D classification[12].

The resulting 2D classes exhibited discernible secondary structure, notably featuring two adjacent β-propeller domains. These particles were used to generate an ab initio model, further refined through heterogeneous and homogeneous refinement processes. The refined map achieved a resolution of approximately 7.5 Å and demonstrated good alignment with one arm of the KICSTOR complex, specifically where the smaller subunits KPTN and ITFG2 bind to the larger SZT2 complex. The alignment to this region of the complex suggested that this portion of KICSTOR may be more rigid, while the remainder of the complex remains highly flexible. The processing workflow and resulting low resolution map are shown in Figure 4.10.

Phase Plate Dataset for Full-Length STREP-KICSTOR

A small dataset was collected using the phase plate in collaboration with the Holger Lab. STREP-KICSTOR was purified as described previously and frozen on R1.2/1.3 grids using a blot time of 3 seconds, blot force of +20, temperature 4°C and 100% humidity. ~240 movies were collected with the phase plate. The resulting data was processed in cryoSPARC using both blob picker and manual picker to pick particles. However, the resulting 2D images exhibited poor alignment and appeared blurry (Figure 4.11). This issue is likely attributed to either a lack of particles due to the low throughput of data collection with the phase plate or poor data/sample quality. Further optimization of data collection methods or sample preparation may be necessary to improve the quality of the dataset.

Structural Understanding of KICSTOR with AlphaFold2

Amid experimental challenges in characterizing the KICSTOR structure, the release of Alphafold2, an open-source structural prediction algorithm, offered an alternative avenue for investigation[13,14]. Alphafold2 operates by users inputting protein sequences, and the algorithm generates structural predictions as output along with an overall confidence score for each prediction. Given the complexity and size of the KICSTOR complex, computational limitations required breaking down the complex into smaller sections to predict its structure effectively.

Initially, full-length sequences of individual subunits were modeled using Alphafold2 to generate structural predictions[13,14]. The results indicated KPTN and ITFG2 may adopted a β-propeller structure, with KPTN featuring an additional alpha helix extending from the propeller and ITFG2 containing three smaller α-helices nestled together and extending from the propeller. The predicted structure for KICS2 was primarily composed of a large α-helical bundle, featuring a longin domain-like fold at the end, potentially serving as an interaction point for a small GTPase. In contrast, SZT2, when analyzed individually, exhibited a donut-shaped structure with a larger ordered base and two arms wrapping up to form a large pore. The collapsed nature of the SZT2 structure prediction indicated that this prediction alone may not be accurate for the full-length SZT2 when in complex with the smaller subunits.

To generate a more accurate model for the full-length SZT2 subunit, the protein sequence was divided into two segments with ~300 amino acids overlapping. The first segment encompassed residues 1-1750, while the second segment contained residues 1414- 3432. Residue 1414 was chosen as the starting point for the second segment due to its location at the beginning of a clearly ordered structural region, providing a suitable alignment point for overlapping the two prediction models. The predicted structures generated from this approach exhibited moderate to high confidence (pTM 0.62 and pTM 0.5 respectively). To construct a model for the full-length SZT2, the two models were overlaid at the overlap region in ChimeraX[15]. All five prediction models generated by Alphafold2 were overlapped and visualized. While the general structure of the models remained consistent, variations were observed in the exact angles of the arms of the structure, depending on the specific prediction model used. This variability suggested that KICSTOR may possess flexibility, with the ability to open and close in response to different

interactions or biological contexts. This confirmed why the cryo-EM data collection had been challenging and there was difficulty in particle alignment.

To pinpoint the regions where KPTN, ITFG2, and KICS2 interacts with SZT2, the SZT2 protein was dissected into smaller segments of ~1300 amino acid each. These segments were individually analyzed with each of the subunits. This iterative process allowed for identification of binding regions on SZT2 for the smaller subunits. The predictions revealed that KICS2 interacts with the C-terminus of SZT2, through an α-helical platform. KPTN and ITFG2 were shown to interact with the region of SZT2 consisting of residues 2040- 3314. Since KPTN and ITFG2 were shown to form an independent dimer and their binding site were adjacent to each other on the SZT2 protein, the SZT2 segment encompassing residues 2040 – 3314 was tested with both KPTN and ITFG2 simultaneously. This prediction corroborated the binding sites of these subunits with SZT2 as seen in previous predictions with SZT2 and the individual subunits. Additionally, it highlighted potential interactions between KPTN and ITFG2, suggesting that their interaction may occur through both the β-propellers and the additional α-helices extending from them.

The accuracy of the Alphafold model was further supported by findings from the original KICSTOR paper, which conducted immunoprecipitation experiments using three segments of SZT2 (1-1021, 1021-2241, and 2241-end) and KPTN, ITFG2, and KICS2. These experiments indicated that all three subunits interacted specifically with the C-terminal segment comprising residues 2241- 3432, consistent with the predictions made by Alphafold. However, there was a slight inconsistency with the immunoprecipitation results for KPTN, which showed its interaction with all portions of SZT2. This observation suggests there could be additional binding sites for KPTN on SZT2 that were not seen in the Alphafold predictions.

The predicted portions of SZT2 in the models containing KICS2, KPTN, and ITFG2 remained structurally consistent with the predictions of the full-length SZT2 generated by the two segments. This allowed for the creation of a model incorporating all four subunits (Figure 4.12). Regions of SZT2 predicted to interact with KPTN, ITFG2, and KICS2 were overlaid with the corresponding SZT2 regions in the full-length model, utilizing ChimeraX. Upon overlaying these models, it became evident that the binding of KPTN and ITFG2 to SZT2 conferred rigidity to that arm of the complex, while the portion of SZT2 not bound to additional subunits retained its flexibility (Figure 4.13). The region of SZT2 not interacting with KPTN, ITFG2, and KICS2 was previously identified to interact with GATOR1[4]. This observation raises questions regarding the dynamics of KICSTOR and its potential significance for interaction with GATOR1 in a biological context.

Membrane Binding Ability of KICSTOR

Since KICSTOR was identified as a lysosomal membrane localized protein, the full-length Alphafold predicted structure was carefully analyzed to identify potential membrane interaction sites. Particularly of interest were the α-helices extending from KPTN and ITFG2 as well as the charged interface of KICS2, which exhibited two hydrophobic loops extending from the plane. Upon closer examination, it was determined that neither the α helix of KPTN nor ITFG2 exhibited clear amphipathic properties with membrane-binding potential. However, the

analysis indicated that KICS2 had the most promising membrane-binding site (Figure 4.14). The charged interface and hydrophobic loops observed in KICS2 suggested its potential affinity for membrane interaction.

To assess the membrane binding potential of the KICSTOR complex, Guvacine Unilamellar Vesicles (GUV) recruitment assays were conducted. In these assays, a GFP tag was introduced to the KICS2 subunit of KICSTOR, while SZT2 was tagged with STREP and KPTN and ITFG2 were left untagged. The complex was then purified using the same protocol as the non-GFP-tagged KICSTOR complex and concentrated to a final concentration of 200 nM. Highly charged GUVs were utilized for initial membrane recruitment experiments, consisting of a lipid composition of DOPC (49.5%), DOPS (50%), and DOPE-Atto 488 (0.5%). The rehydration buffer for the GUVs contained 230 mM sucrose, and the final imaging buffer for all experiments consisted of 25 mM HEPES (pH 7.4), 137.5 mM NaCl, 0.1 mM $MgCl_2$, 46 mM sucrose, and 0.45 mM TCEP, with an osmolarity of 312 mOsm. During the assays, recruitment of KICSTOR to the membrane was monitored over a period of 24 minutes at 30-second intervals. A gradual increase in GFP signal around the GUVs was observed over the course of 24 minutes, with a strong signal detected at the end of the experiment, indicating robust binding of KICSTOR to the membrane (Figure 4.15).

To confirm the hypothesized membrane binding site on KICS2, mutants targeting the proposed site (Figure 4.14B) were generated. Two different mutants were created. The first mutant (mutant 4D) targeted the critical residues in loop #1 (K77D, F83D, F84D, and R86D) and the second mutatnt (mutant 10D) targeted the critical residues in both loop #1 and loop #2 (K77D, F83D, F84D, R86D, R234D, K237D, K238D, H239D, F241D, and K247D). These mutants were purified using the same protocol as the GFP-tagged KICSTOR complex and concentrated to 200 nM for utilization in GUV assays. The conditions used to perform the GUV assays for the mutants were consistent with those of WT KICSTOR. At the completion of the 30-minute monitoring period, robust membrane binding comparable to WT was observed for the 4D mutant, indicating that removal of loop #1 alone was insufficient to disrupt membrane binding. However, the 10D mutant showed a significant reduction in membrane binding, suggesting that both loops in KICS2 played a role in facilitating membrane binding to the GUV (Figure 4.16). There was still a minimal amount of membrane binding observed after 30 minutes for the 10D mutant. Given the large size of the KICSTOR complex, it's possible that an additional, weaker membrane-interacting site exists. Further experiments will be necessary with a third mutant, specifically targeting loop #2, to determine whether loop #2 is primarily responsible for membrane binding, or if loop #1 and loop #2 are capable of compensating for the loss of the other.

Interaction of KICSTOR with GATOR Components

The proposed biological role of KICSTOR has been to regulate the activity of GATOR1 through interaction with GATOR1 and GATOR2. However, besides the initial papers, no follow-up studies focused on these interactions[4,6]. To validate the interactions between KICSTOR, GATOR1 and GATOR2 in vitro, a pulldown assay was used. In these experiments, 100 µL of purified GST-tagged KICSTOR (1 µM) complex was incubated with purified 100 µL of

GATOR1 (1 µM) and/or 100 µL of purified GATOR2 (1 µM), along with 30 µL of GST resin. The resin was washed three times with 500 µL of wash buffer (25 mM HEPES pH 7.5, 200 mM NaCl, 2 mM $MgCl_2$, 1 mM TCEP) and eluted with 30 µL of wash buffer 50mM GSH. The eluted proteins were then visualized using SDS-PAGE. The results showed that GST-KICSTOR successfully pulled down both GATOR1 and GATOR2 individually. Moreover, GST-KICSTOR was capable of simultaneously pulling down GATOR1 and GATOR2, indicating that the binding sites for these complexes on KICSTOR are not mutually exclusive (Figure 4.17).

Membrane Recruitment Assay for KICSTOR and GATOR Components

Preliminary GUV assays suggested that the KICSTOR complex interacted with negatively charged membranes via the KICS2 subunit (as described above). Given this observation and the known interaction of KICSTOR with the GATOR complexes, GUV assays were employed to investigate whether KICSTOR played a role in recruiting the GATOR complexes to the membrane. First, to determine if GATOR1 and GATOR2 acted as membrane binding proteins independently of KICSTOR, these complexes were tagged with fluorescent tags and their membrane recruitment was analyzed using GUV assays under the same conditions as GFP-KICSTOR. While GATOR1 was not anticipated to interact with the membrane, the GATOR2 structure revealed a charged patch on the WDR24 β-propeller, suggesting a potential membrane interaction site[16].

To monitor GATOR1 membrane recruitment, an mCherry tag was engineered on the NPRL3 subunit of the GATOR1 complex. The purification process for GATOR1-mCherry followed the same protocol as for WT GATOR1. However, gel filtration yielded a more heterogeneous peak, with the smallest peak (#1) corresponding to the GATOR1-mCherry sample and an adjacent peak (#2) representing incomplete formation of the GATOR1 complex and degraded NPRL3-mCherry (Figure 4.18). This outcome may be attributed to too short linker between NPRL3 and mCherry that potentially disrupted proper GATOR1 complex formation. Future experiments should explore extending the linker between mCherry and NPRL3. Despite the presence of some broken complex, there was sufficient GATOR1-mCherry to conduct initial control GUV assays to investigate if GATOR1 interacts with charged membranes. After a 1-hour time course, no GATOR1 was observed to interact with the negatively charged membrane (Figure 4.19).

To monitor GATOR2 membrane recruitment, a GFP tag was added to the Sec13 subunit of the GATOR2 complex. Purification of the GFP-GATOR2 followed the same purification protocol as WT GATOR2. The chromatogram indicated the sample eluted ~12 mL and SDS-PAGE confirmed the presence of the GFP-GATOR2 complex (Figure 4.20). This sample was used to perform initial control GUV assays to assess the potential interaction of GATOR2 with a charged membrane. Following a 1-hour time course, no interaction between GATOR2 and negatively charged membranes was observed. There were some large clumps that interacted transiently with the membrane that can likely be attributed to partial aggregation of the complex throughout the duration of the assay (Figure 4.21).

To perform recruitment assays with all three protein complexes simultaneously, the tag for GATOR2 was switched from a GFP to a halo tag. GATOR2-halo was purified using the

same protocol as WT GATOR2 except prior to gel filtration the protein was incubated with halo ligand. However, through gel filtration, it was observed that the GATOR2 protein aggregated and was no longer present in useable concentrations for GUV assays. This purification will need to be optimized, or an alternative fluorescent tag may need to be added to the GATOR2.

Cryo-EM for KICSTOR-GATOR1-GATOR2 Complex

In attempts to 1) physiologically stabilize the flexible KICSTOR complex and 2) visualize the interactions between KICSTOR and GATOR complexes, the three protein complexes were purified to form a higher order complex.

KICSTOR-GATOR1-GATOR2 Complex Formation

In the first approach to form a higher-order complex, a pull-down method was utilized. GATOR1 and GST-KICSTOR were individually purified following established protocols, while GATOR2 was purified similarly to the method described in Chapter 3, with gel filtration omitted to preserve higher yields. The three protein complexes were incubated overnight at 4°C to promote complex formation. The following day, the GST tag on KICSTOR was utilized to separate the formed complex from excess GATOR1 and GATOR2 (Figure 4.22A). The complex was eluted from the GST resin and incubated with TEV protease for 10 hours at 4°C. The sample was concentrated to 1 mL for gel filtration. SDS-PAGE analysis of this sample indicated that all three complexes were present (Figure 4.22B). Gel filtration using buffer containing 25 mM HEPES pH 7.5, 200 mM NaCl, 2 mM $MgCl_2$, 1 mM TCEP showed a single peak at ~10.5 mL. This appeared to be a shift from the typical elution of ~12 mL for KICSTOR alone (Figure 4.22D). Fractions corresponding to the peak and the concentrated peak sample were visualized using SDS-PAGE revealing only KICSTOR were present in the sample (Figure 4.22C). It is possible that during gel filtration the complex broke apart and the peak shift was attributed to GST-KICSTOR that did not have the GST-tag properly cleaved.

In the next attempt, the pull-down step was eliminated since previous experiments indicated that the complex readily interacted. GATOR1, GATOR2 and STREP-KICSTOR were purified using established protocols and incubated overnight at 4°C to enable complex formation. Gel filtration of half of the sample showed a heterogenous profile with peaks corresponding to KICSTOR, GATOR2 and GATOR1 separately indicating that the complex likely fell apart during gel filtration (Figure 4.23). The second half of the sample that didn't undergo gel filtration was concentrated to freeze grids. The final target concentrations for the sample were 50 nM GATOR2, 200 nM GATOR1, and 200 nM KICSTOR. The grids were frozen using a blot time of 3 seconds, blot force of +20, temperature 4°C and 100% humidity on CF-2/1-3Au-T grids. The grids were glow discharged for 30 seconds at 15 mA and 0.5 mbar.

Data processing for KICSTOR-GATOR1-GATOR2 Complex

A test dataset was collected on the Arctica. The sample on the micrographs was very dilute with each micrograph only containing a few particles. It is possible that the sample aggregated during concentration prior to freezing or the actual concentrations were too low for open hole. This may be fixed using a carbon support layer. Nonetheless, movies from the data

collection were motion corrected in Relion[10] using MotionCor2[11] before being imported into cryoSPARC[9] for further data processing (Figure 4.24A). Particle picking was performed using the cryoSPARC blob picker with various particle diameters, including 130 Å – 150 Å, 180 Å – 250 Å, and 250 Å – 300 Å (Figure 4.24B). Particles corresponding to GATOR1 and KICSTOR were observed at smaller picking diameters, while larger diameters revealed new 2D classes. Upon further analysis, it was identified that these particles corresponded to the 60S and 40S ribosome subunits, a contaminate from GATOR2 sample preparation (Figure 4.24C). Additional datasets at a higher sample concentration would be necessary to determine if the higher order complex is stable enough to be visualized with cryo-EM.

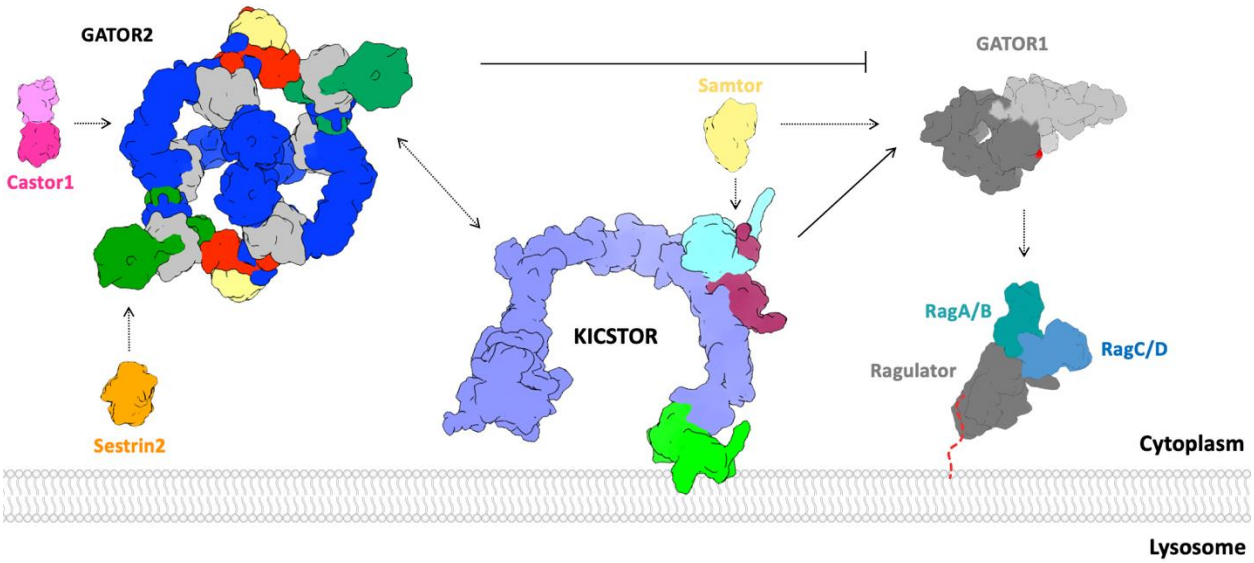
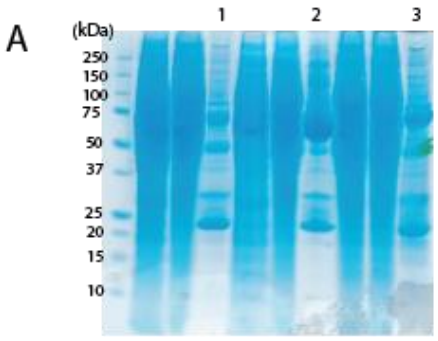


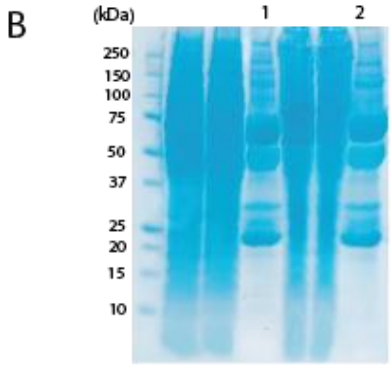
Figure 4.1 Schematic of GATOR1 regulation. RagA GTPase activating protein (GAP), GATOR1 is regulated by a complex network of protein complexes and nutrient sensors. GATOR2 is proposed to be a negative regulator of GATOR1 while KICSTOR is hypothesized to be an activator. The large complexes, GATOR1, GATOR2 and KICSTOR have been shown to interact although the extent and conditions surrounding these interactions remains elusive. Nutrient signals are integrated into the pathway through nutrient sensors Castor1 and Sestrin2 which interact with GATOR2 and SAMTOR which interacts with GATOR1 and KICSTOR.



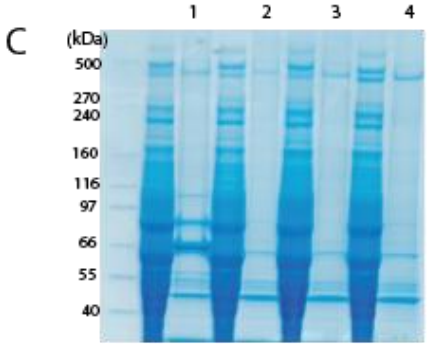
Figure 4.2 Constructs designed for KICSTOR subunits. All genes were ordered codon-optimized from Twist Bioscience and cloned into pCAG expression vectors with different tags including GST, 2X-STREP and FLAG. Subunits containing no tag were created for KPTN, ITFG2 and KICS2



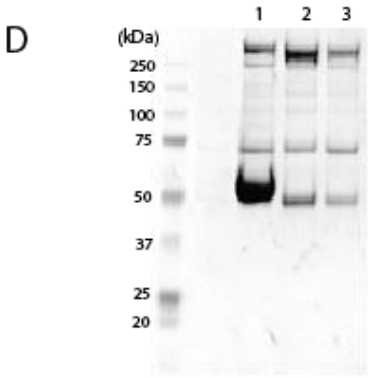
Lane	Subunits over-expressed
1	GST-KPTN
2	GST-ITFG2
3	GST-KICS2



Lane	Subunits over-expressed
1	GST-KPTN/ ITFG2
2	GST-ITFG2 / KPTN



Lane	Subunits over-expressed
1	GST-SZT2/ GST-KICS2
2	GST-SZT2/ KICS2
3	GST-SZT2/ KPTN/ ITFG2/ KICS2 (1:1:1:1)
4	GST-SZT2/ KPTN/ ITFG2/ KICS2 (2:1:1:1)



Lane	Subunits over-expressed
1	GST-SZT2/pCAG TSF-KPTN (48kDa) TSF-KICS2 (50kDa) TSF-ITFG2 (49kDa)
2	TSF-SZT2/pCAG (378kDa) KPTN (48kDa) KICS2 (50kDa) ITFG2 (49kDa)
3	TSF-SZT2/pCAG (378kDa) KPTN (48kDa) KICS2 (50kDa) ITFG2 (49kDa)

Figure 4.3 Small scale expression tests for KICSTOR complex. SDS-PAGE visualization of small-scale tests. (A) Lanes 1, 2, and 3 are heated resin samples from the small-scale purification. (B) Lanes 1, 2, and 3 are heated resin samples from the small-scale purification. (C) Lanes 1, 2, 3 and 4 are proteins eluted from GST resin in the small-scale purification. (D) Lanes 1, 2, and 3 are proteins eluted from STREP resin in the small-scale purification. The subunits of KICSTOR expressed in 50 mL HEK 293 GnTI cells is shown in the table to the right of all gels.

Locus	Sequence Count	Spectrum Count	Sequence Coverage	Length	MolWt	pI	Descriptive Name
sp|Q5T011|SZT2_HUMAN	63	103	22.50%	3432	378030	6.3	Protein SZT2 OS=Homo sapiens GN=SZT2 PE=1 SV=3

Locus	Sequence Count	Spectrum Count	Sequence Coverage	Length	MolWt	pI	Descriptive Name
sp|Q96MD2|CL066_HUMAN	48	110	65.60%	445	50415	8.9	UPF0536 protein C12orf66 OS=Homo sapiens GN=C12orf66 PE=2 SV=4
sp|Q969R8|ITFG2_HUMAN	35	94	65.60%	447	49313	5.2	Integrin-alpha FG-GAP repeat-containing protein 2 OS=Homo sapiens GN=ITFG2 PE=1 SV=1
sp|Q9Y664|KPTN_HUMAN	24	84	50.70%	436	48080	5.3	Kaptin OS=Homo sapiens GN=KPTN PE=1 SV=2

Figure 4.4 Mass spectrometry analysis of SDS-PAGE bands correlating with KICSTOR components.

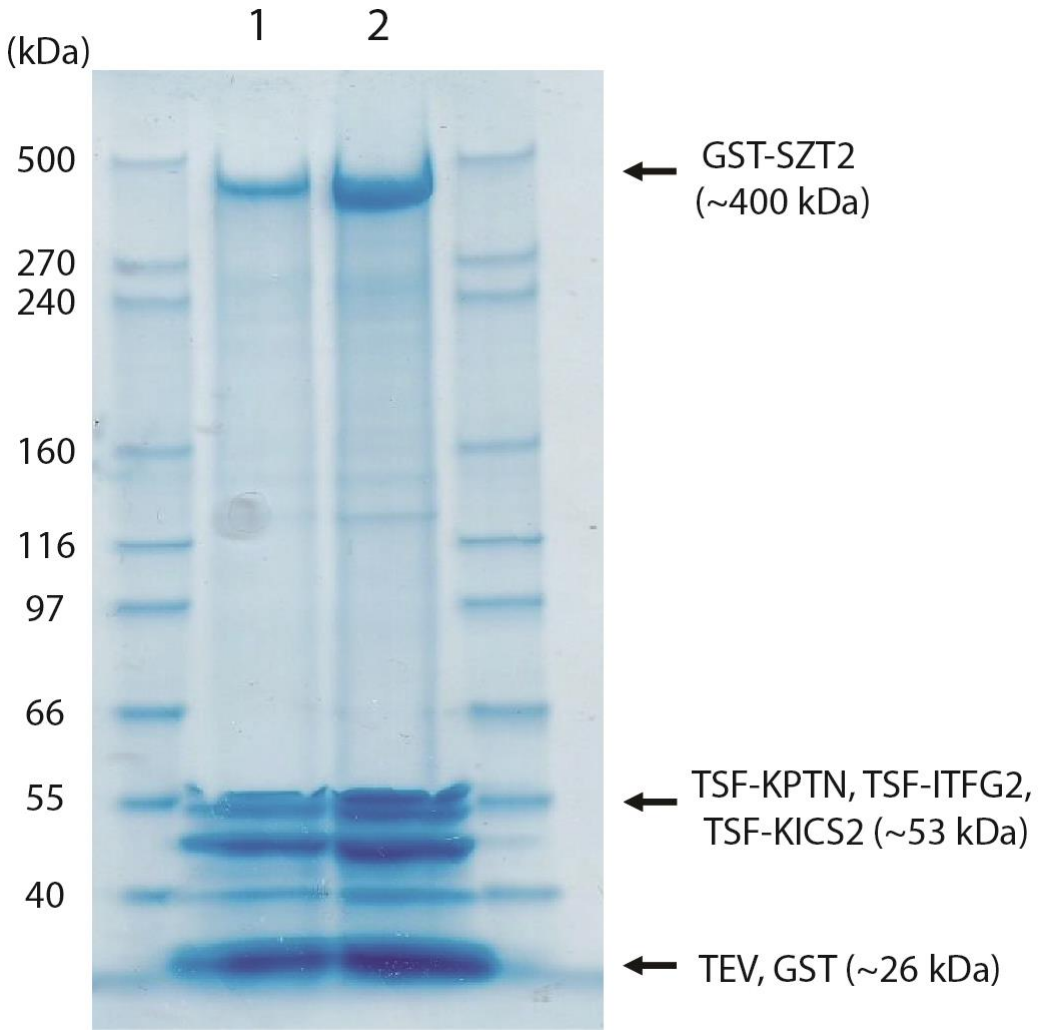


Figure 4.5 Representative large-scale purification of KICSTOR complex using GST-SZT2 (method #1 – no SEC). SDS-PAGE of elution from GST resin in lane 1. Lane 2 is the GST resin elution concentrated.

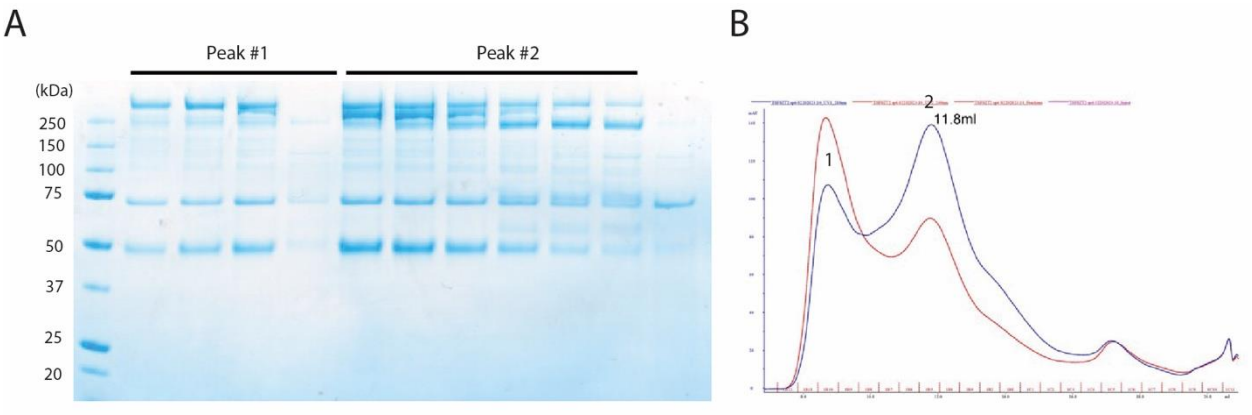


Figure 4.6 Representative large-scale purification of KICSTOR complex using STREP-SZT2. (A) SDS-PAGE of fractions from Superose 6 increase gel filtration. (B) Chromatogram from Superose 6 increase gel filtration.

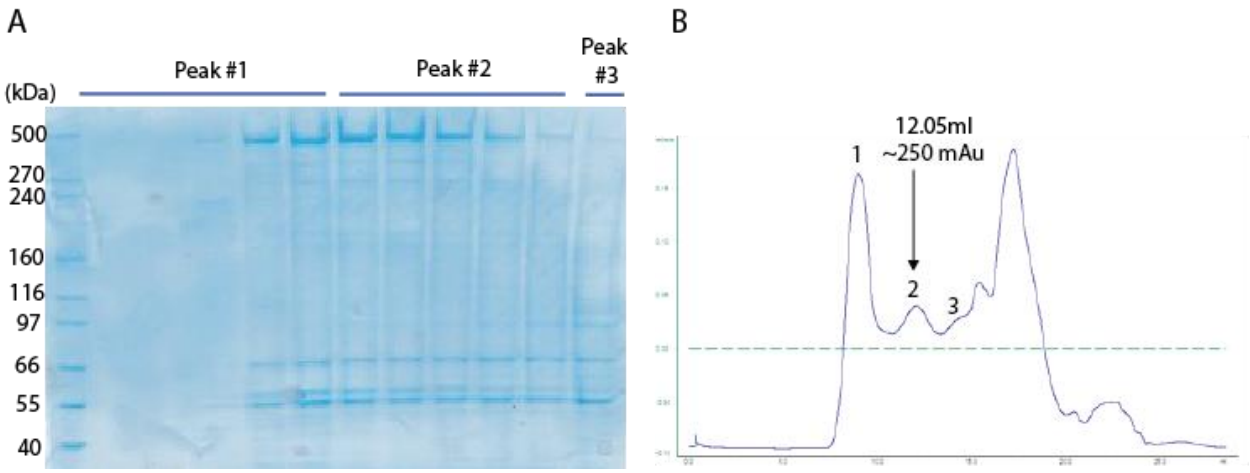


Figure 4.7 Representative large-scale purification of GST-KICSTOR complex using STREP tagged smaller subunits. (A) SDS-PAGE of fractions from Superose 6 increase gel filtration. (B) Chromatogram from Superose 6 increase gel filtration.

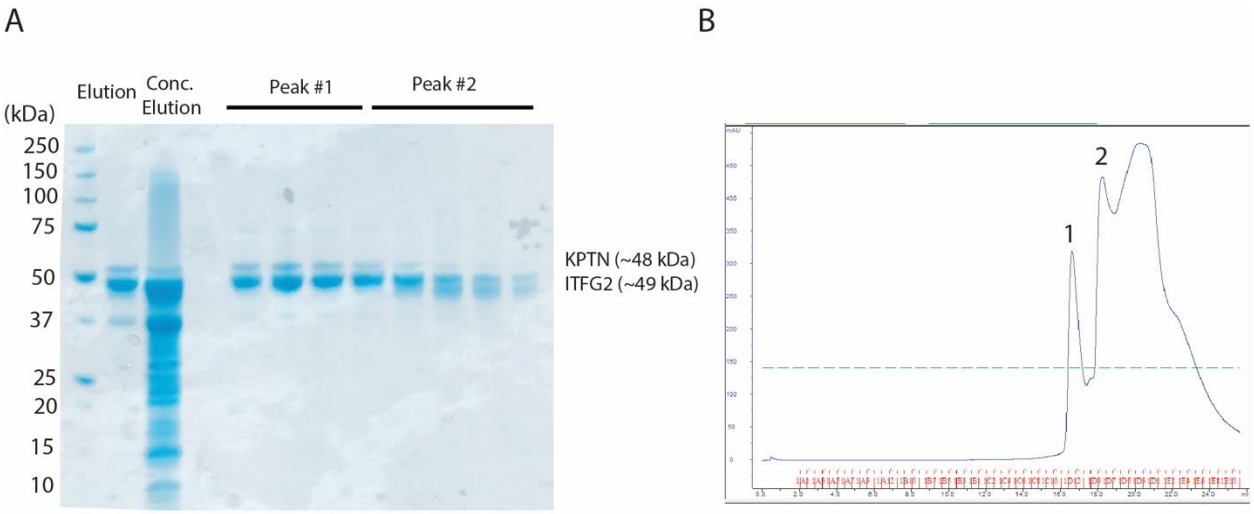


Figure 4.8 Large scale purification of ITFG2/KPTN dimer. (A) SDS-PAGE of fractions from Superose 6 increase gel filtration. (B) Chromatogram from Superose 6 increase gel filtration.

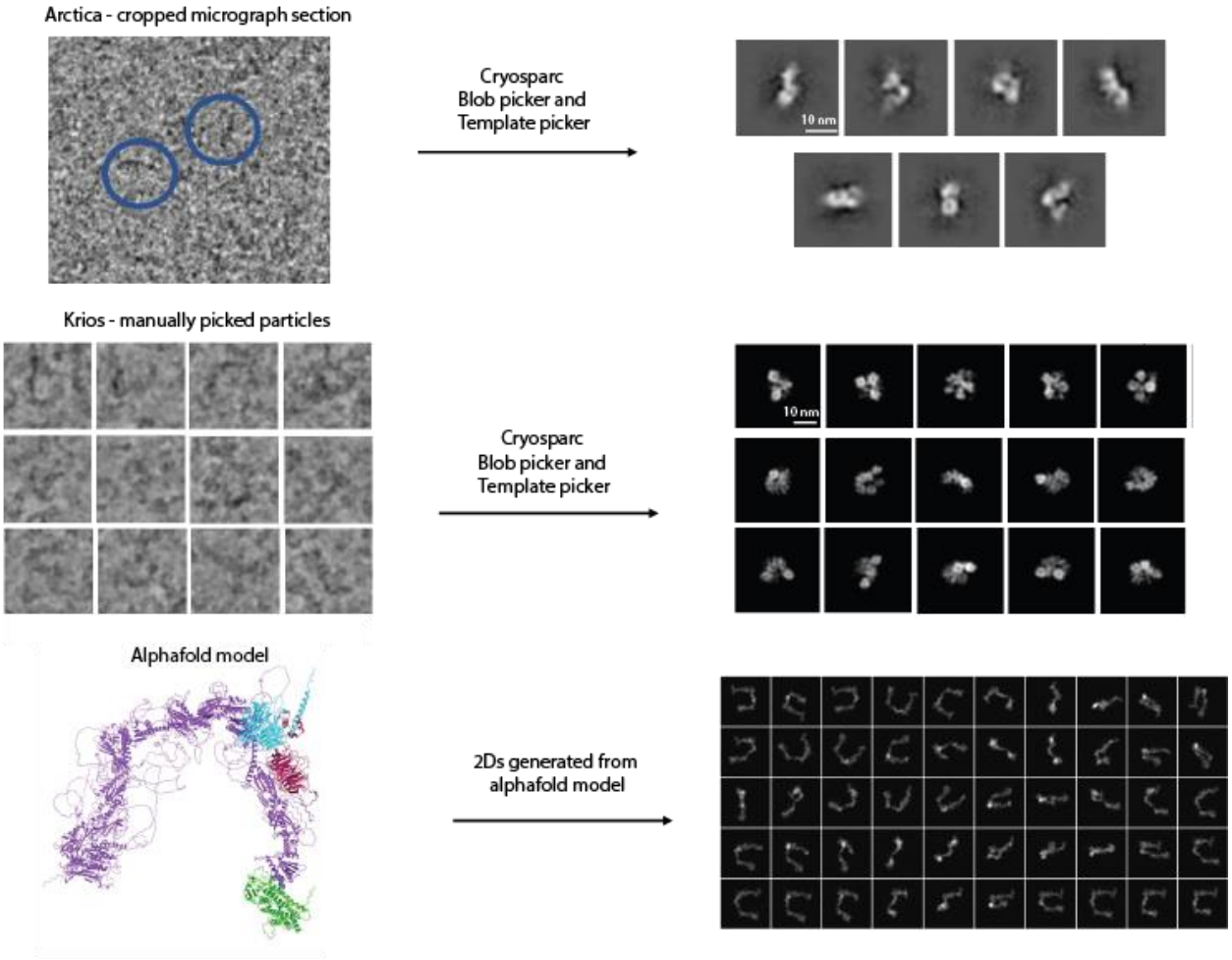


Figure 4.9. Comparison of KICSTOR cryo-EM datasets with Alphafold. Best particles and 2D class averages from each data set compared to Alphafold model generated 2D class averages.

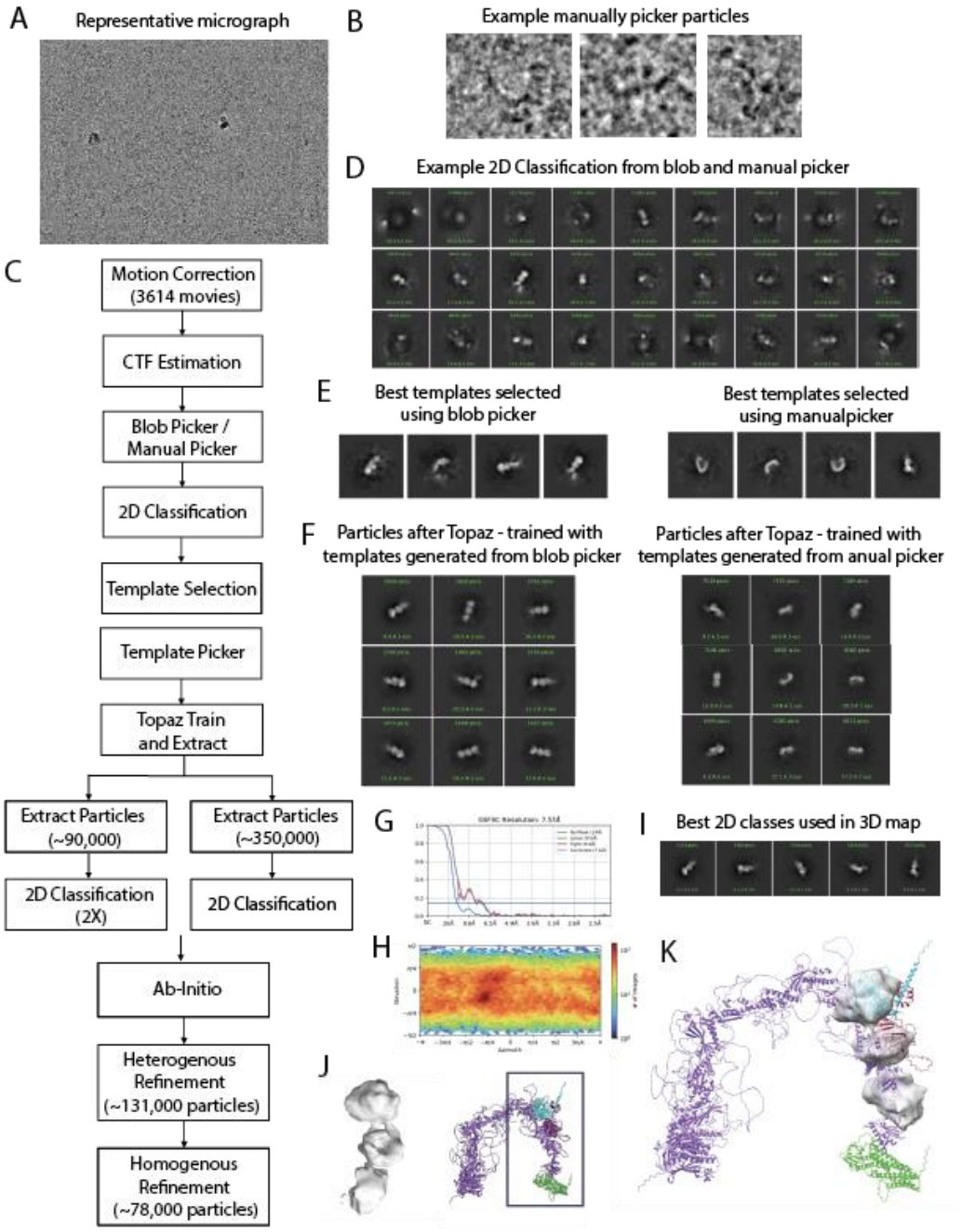


Figure 4.10 Data collection and processing workflow for STREP-KICSTOR. (A) Representative micrograph from the data collection. (B) Example particles generated from manual picking. (C) Data processing workflow performed in cryoSPARC. (D) Example of initial 2D class averages generated from blob and manual picker. (E) Best templates selected from block picker and manual picker. (F) Particles after Topaz was trained with particles picked using templates from blob picker and template generated from manual picking. (G) FSC graph from resulting map. (H) Orientation map for final 3D map. (J) Overview of 3D map generated after homogenous refinement and the arm of KICSTOR it corresponds to. (I) 2D classes used to generate 3D map (K) 3D map overlayed on the Alphafold generated KICSTOR structure.

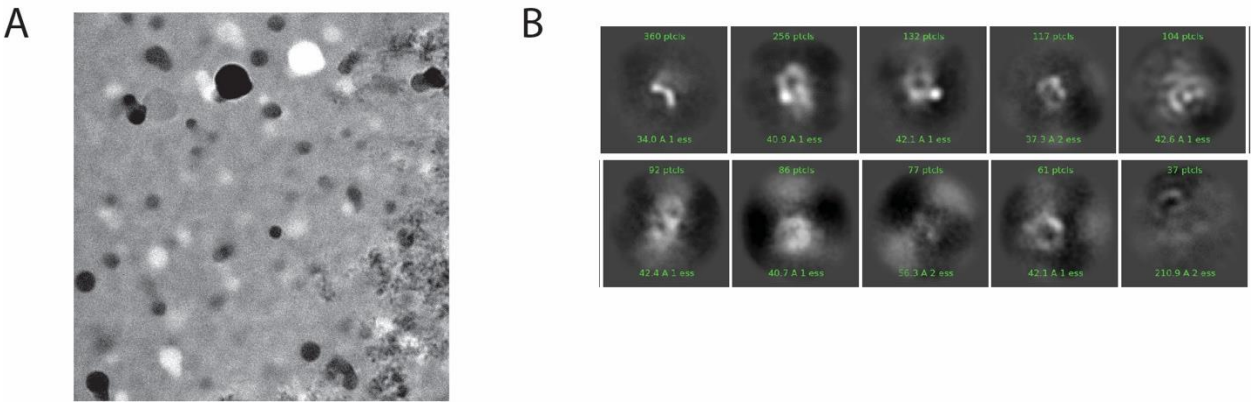


Figure 4.11 Phase plate data collection for KICSTOR. (A) Representative micrograph from the data collection. (B) Resulting 2D classes generated from blob and manual picking in cryoSPARC.

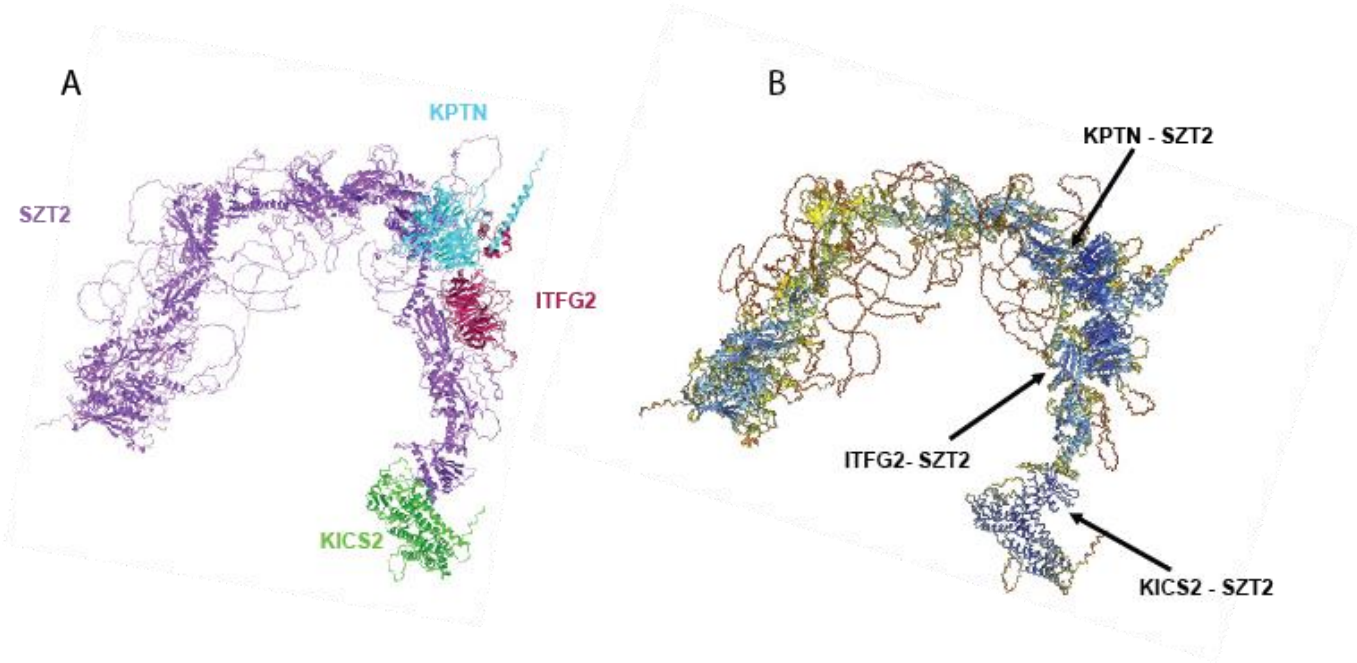


Figure 4.12 Alphafold2 generated model of full-length KICSTOR complex. (A) Overview of the whole complex containing SZT2 (purple), KPTN (blue), ITFG2 (pink) and KICS2 (green). (B) The Alphafold model colored by confidence.

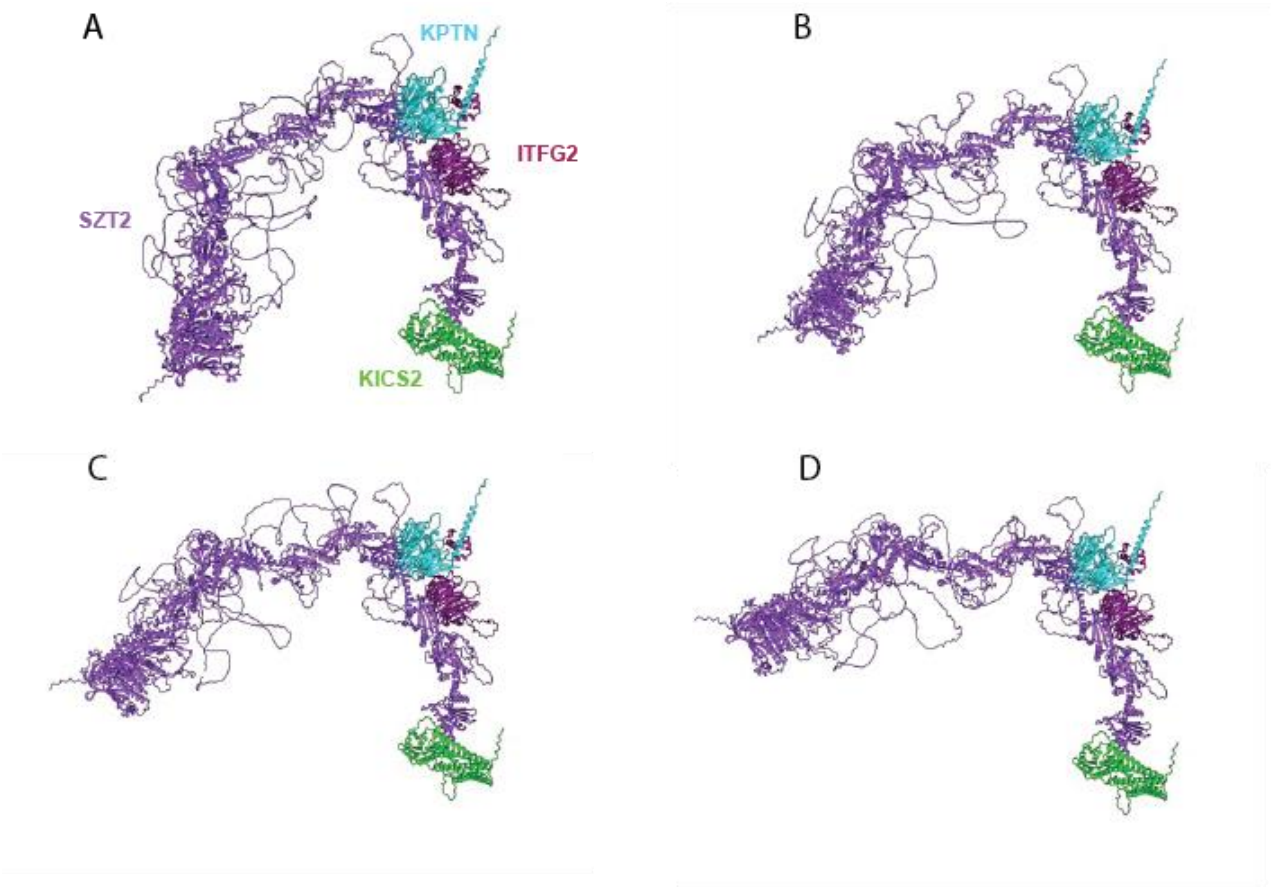


Figure 4.13 KICSTOR complex flexibility seen in Alphfold2 model. (A-D) Different confirmations of the KICSTOR complex based on the alignments and modeling of Alphafold predictions.

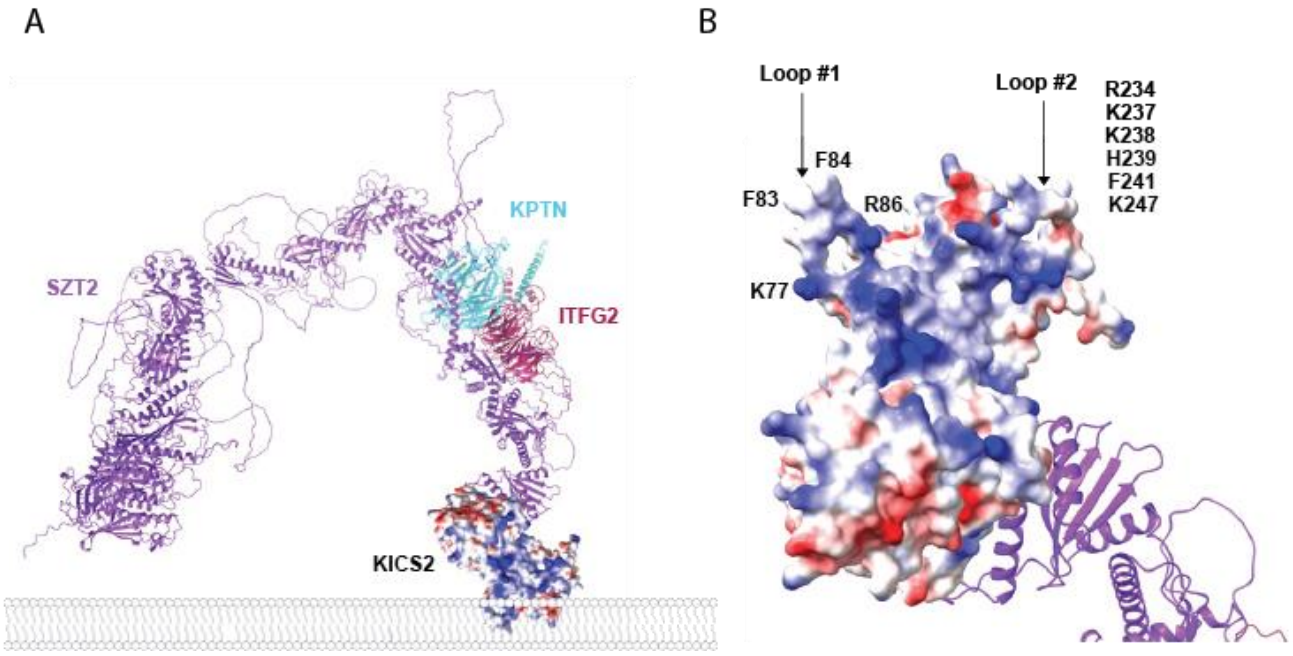


Figure 4.14 Proposed membrane binding site in KICS2 subunit of KICSTOR complex. (A) Visual representation of potential KICSTOR interaction with the lysosomal membrane. (B) Membrane binding site on KICS2 subunit. The interface is colored by electrostatics and the key membrane interaction residues are in loops #1 and #2 are annotated.

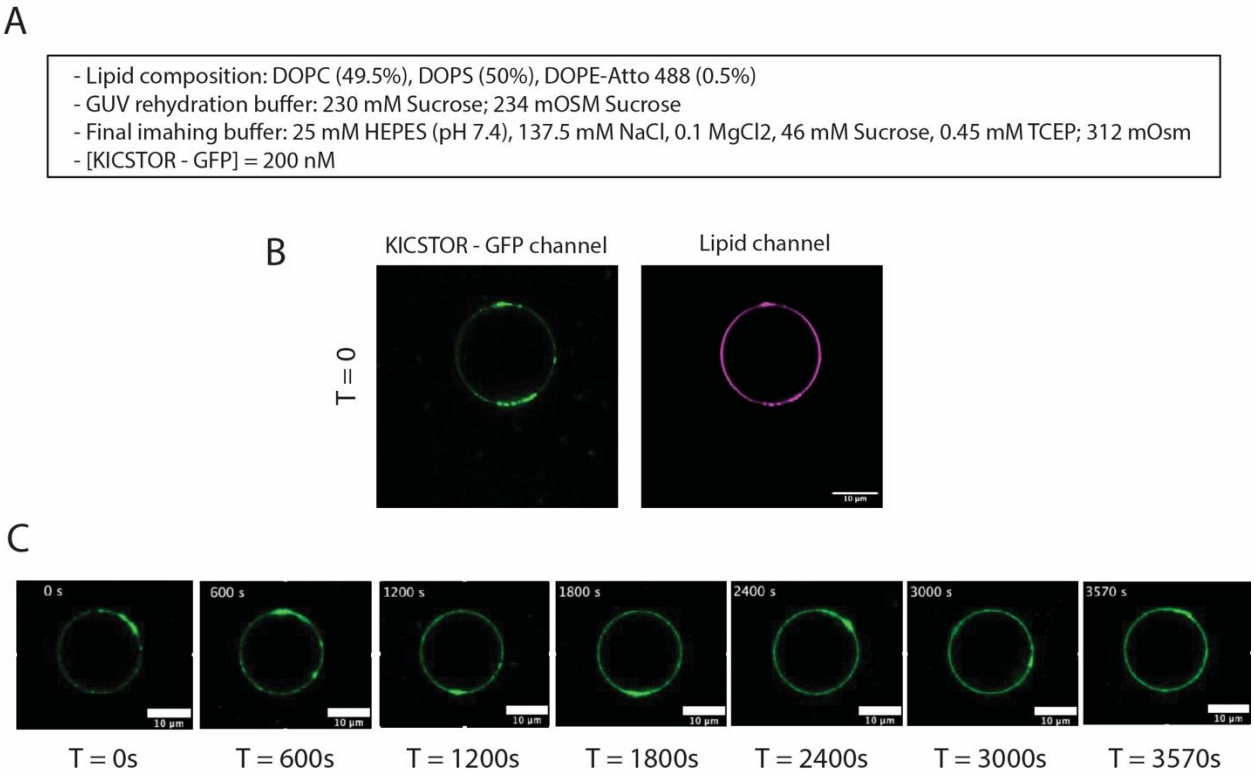


Figure 4.15 GUV assay for WT KICSTOR-GFP membrane binding. (A) Overview of conditions used for GUV assay. (B) KICSTOR – GFP and lipid channels (C) Snapshots from time course of KICSTOR-GFP interaction with GUV.

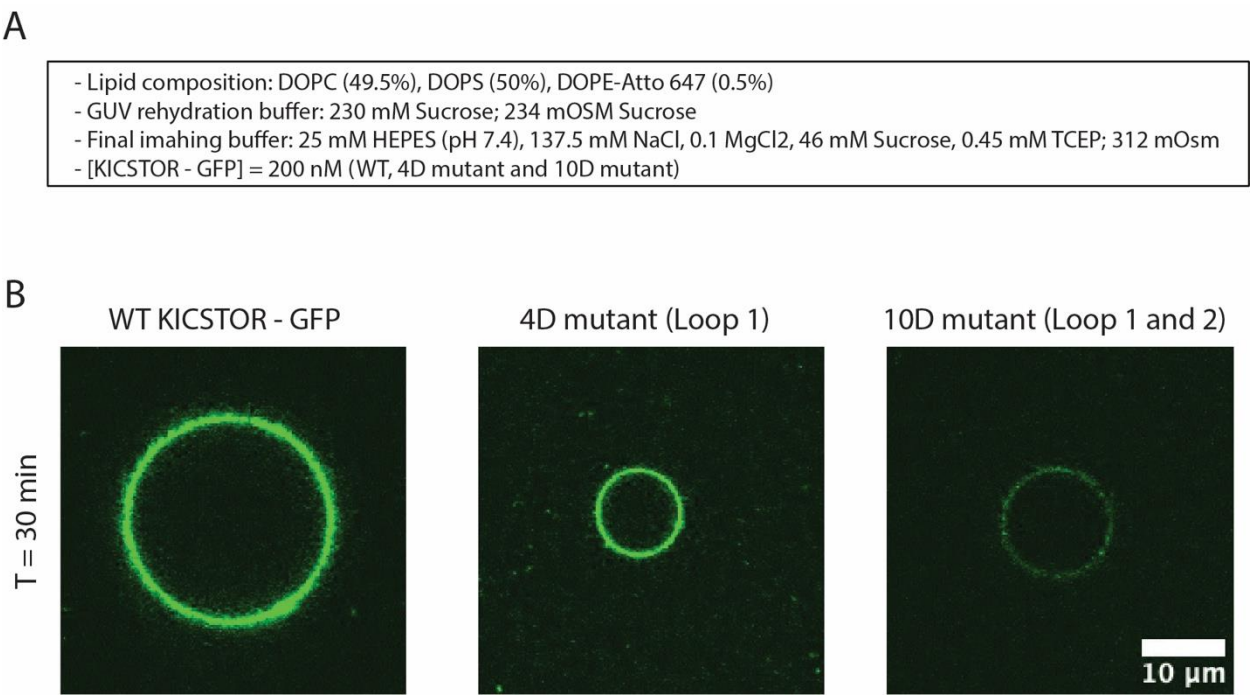
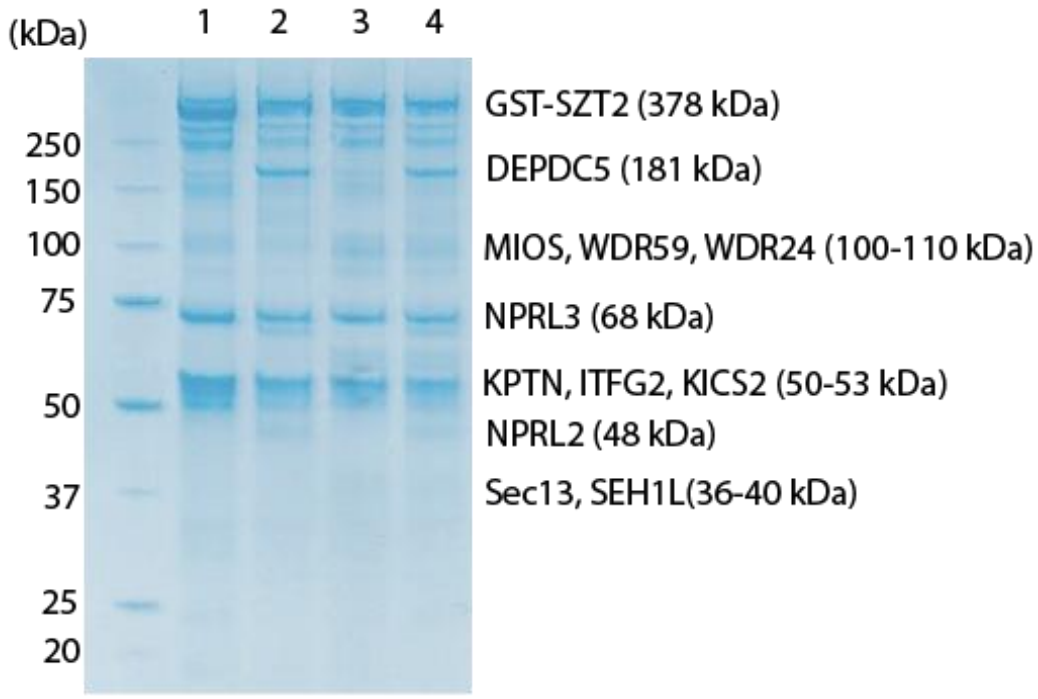


Figure 4.16 GUV assay for KICSTOR-GFP mutant 4D and mutant 10D membrane binding. (A) Overview of conditions used for GUV assay. (B) KICSTOR – GFP membrane binding comparison for WT, 4D mutant, and 10D mutant after 30-minute time course.



Lane	Proteins added
1	**GST-KICSTOR**
2	**GST-KICSTOR** + Gator1
3	**GST-KICSTOR** + Gator2
4	**GST-KICSTOR** + Gator1 + Gator2

Figure 4.17 Pulldown assay with GST-KICSTOR and GATOR complexes. GST resin was used to pull down GST-KICSTOR and interacting complexes. Table describes proteins added to each pull down reaction and corresponding lane on SDS-PAGE.

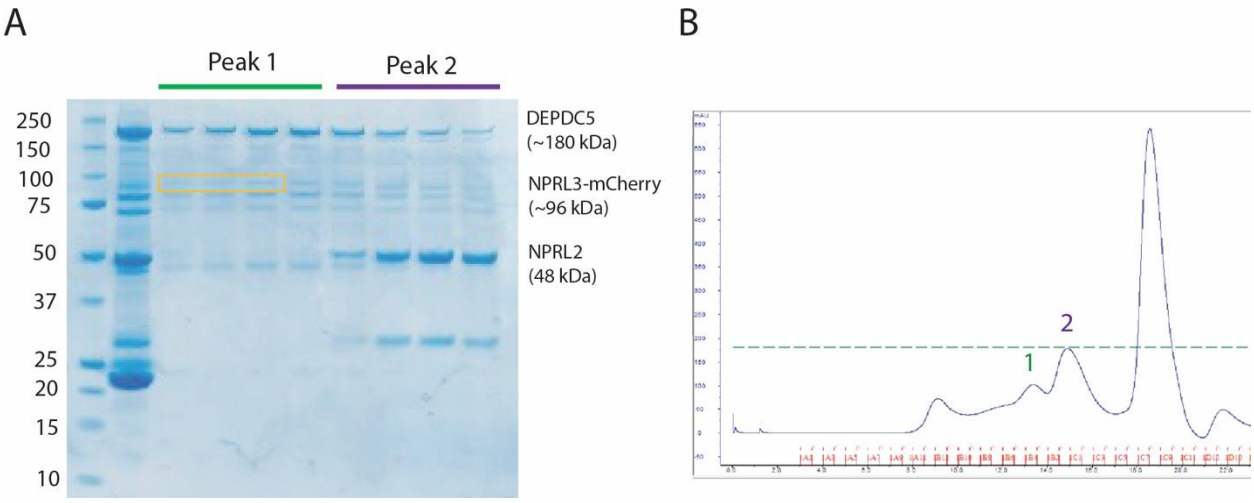


Figure 4.18 Purification of GATOR1-mCherry. (A) SDS-PAGE visualization of fractions from Superose 6 increase gel filtration. (B) Chromatogram from Superose 6 increase gel filtration.

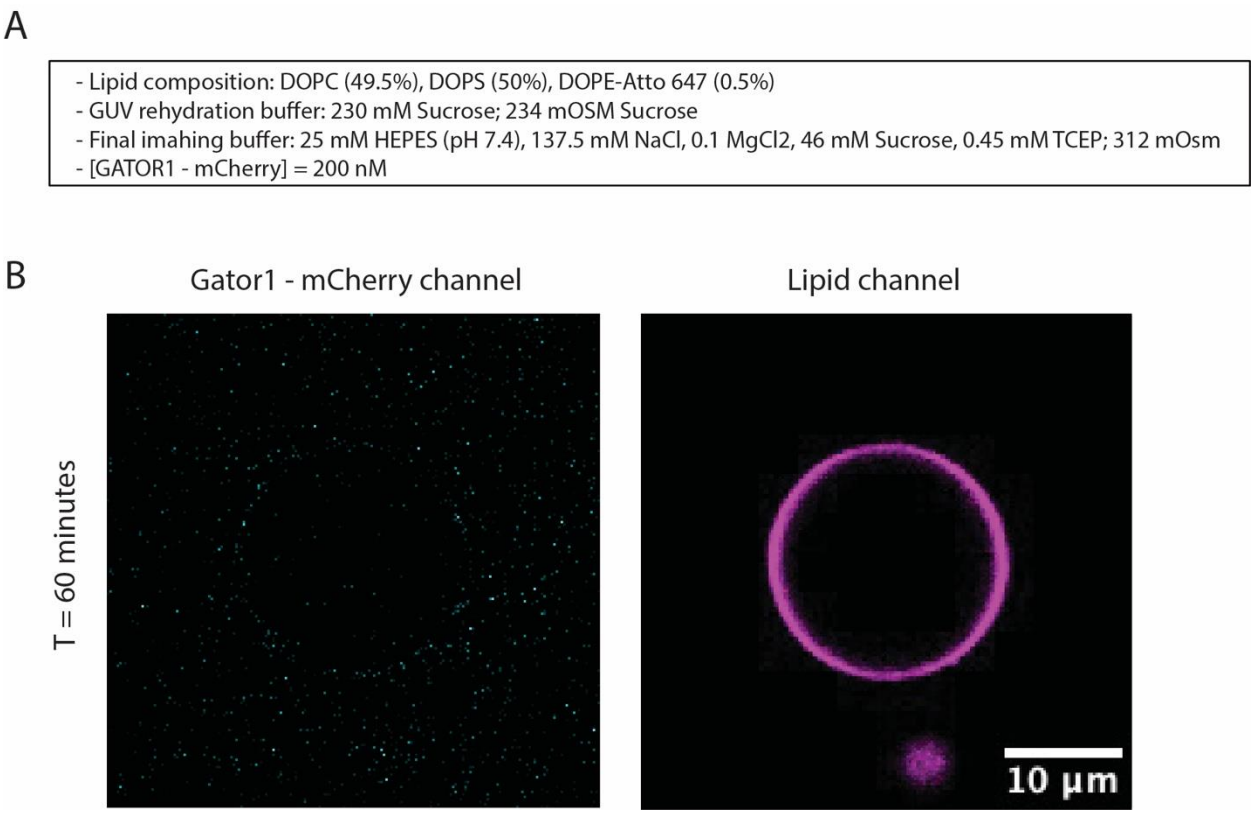


Figure 4.19 GUV assay for GATOR1-mCherry membrane binding. (A) Overview of conditions used for GUV assay. (B) GATOR1-mCherry and lipid channels after completion of the 1-hour time course.

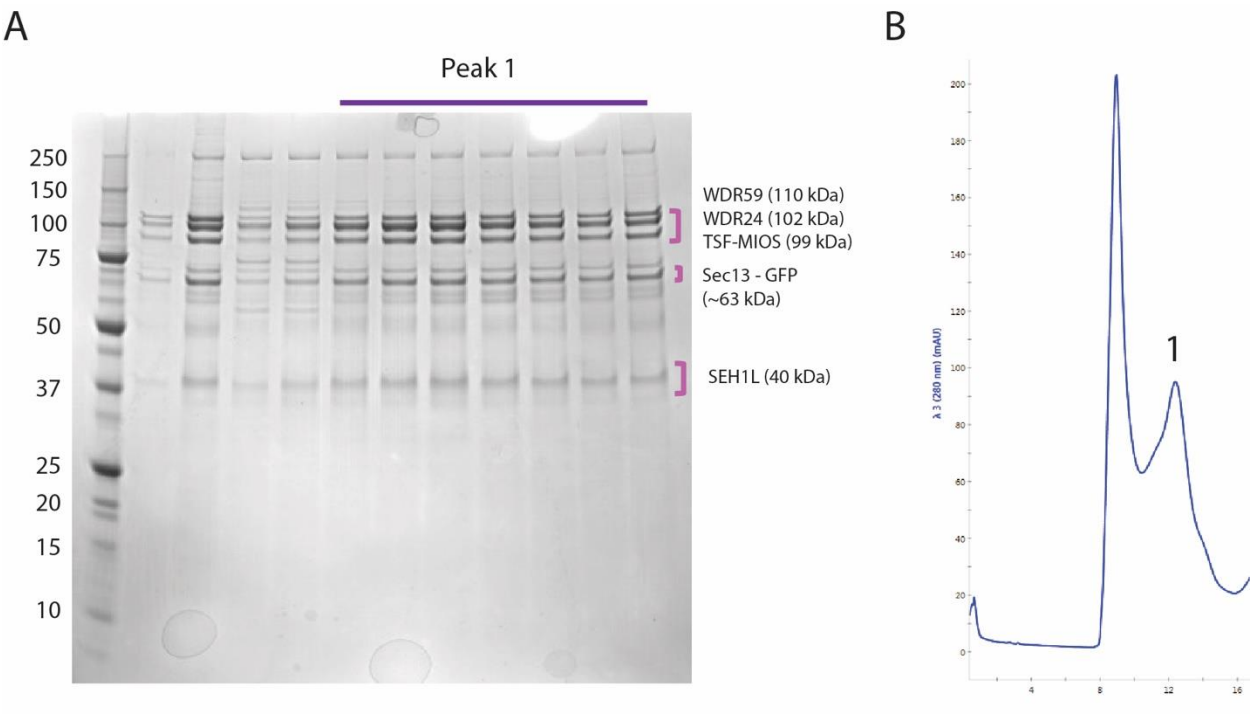


Figure 4.20 Purification of GATOR2-GFP. (A) SDS-PAGE gel of fractions from Superose 6 increase gel filtration. (B) Chromatogram from Superose 6 increase gel filtration.

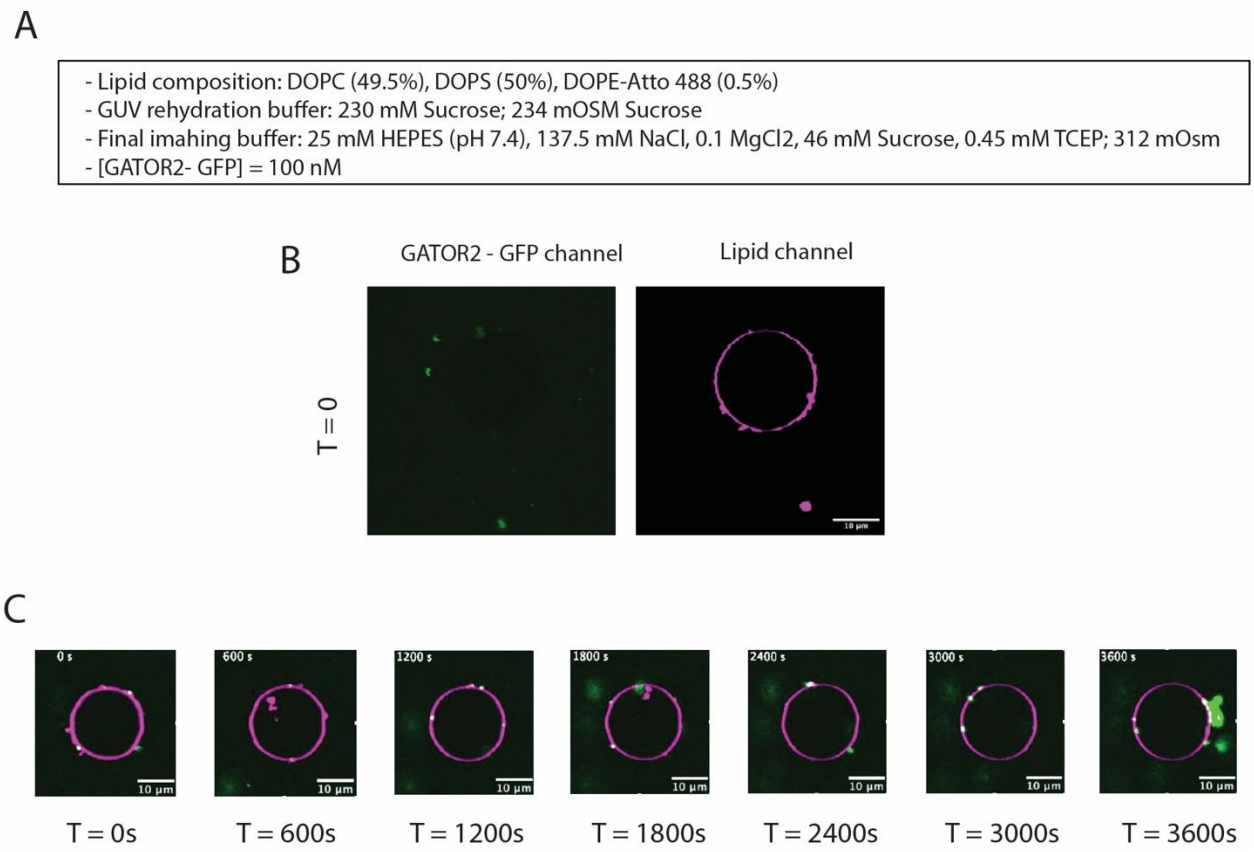


Figure 4.21 GUV assay for GATOR2-GFP membrane binding. (A) Overview of conditions used for GUV assay. (B) GATOR2 – GFP and lipid channels (C) Snapshots from time course of GATOR2-GFP interaction with GUV.

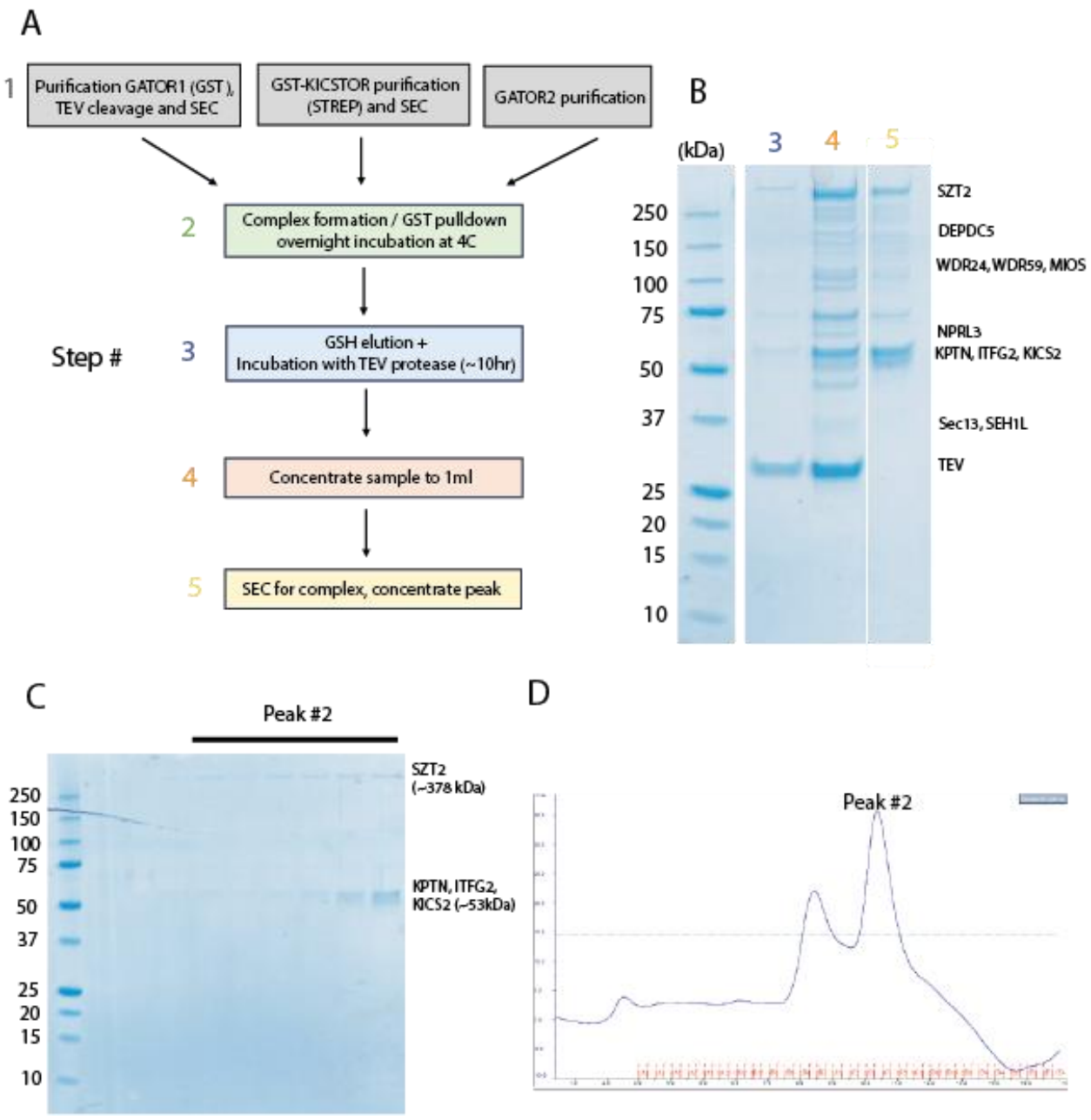


Figure 4.22 Attempt #1 - large scale complex formation for KICSTOR, GATOR1 and GATOR2. (A) Complex formation scheme for higher order complex. Each step in the complex formation is numbered and colored to match the SDS-PAGE in (B). (C) SDS-PAGE visualization of fractions from Superose 6 increase and (D) corresponding chromatogram.

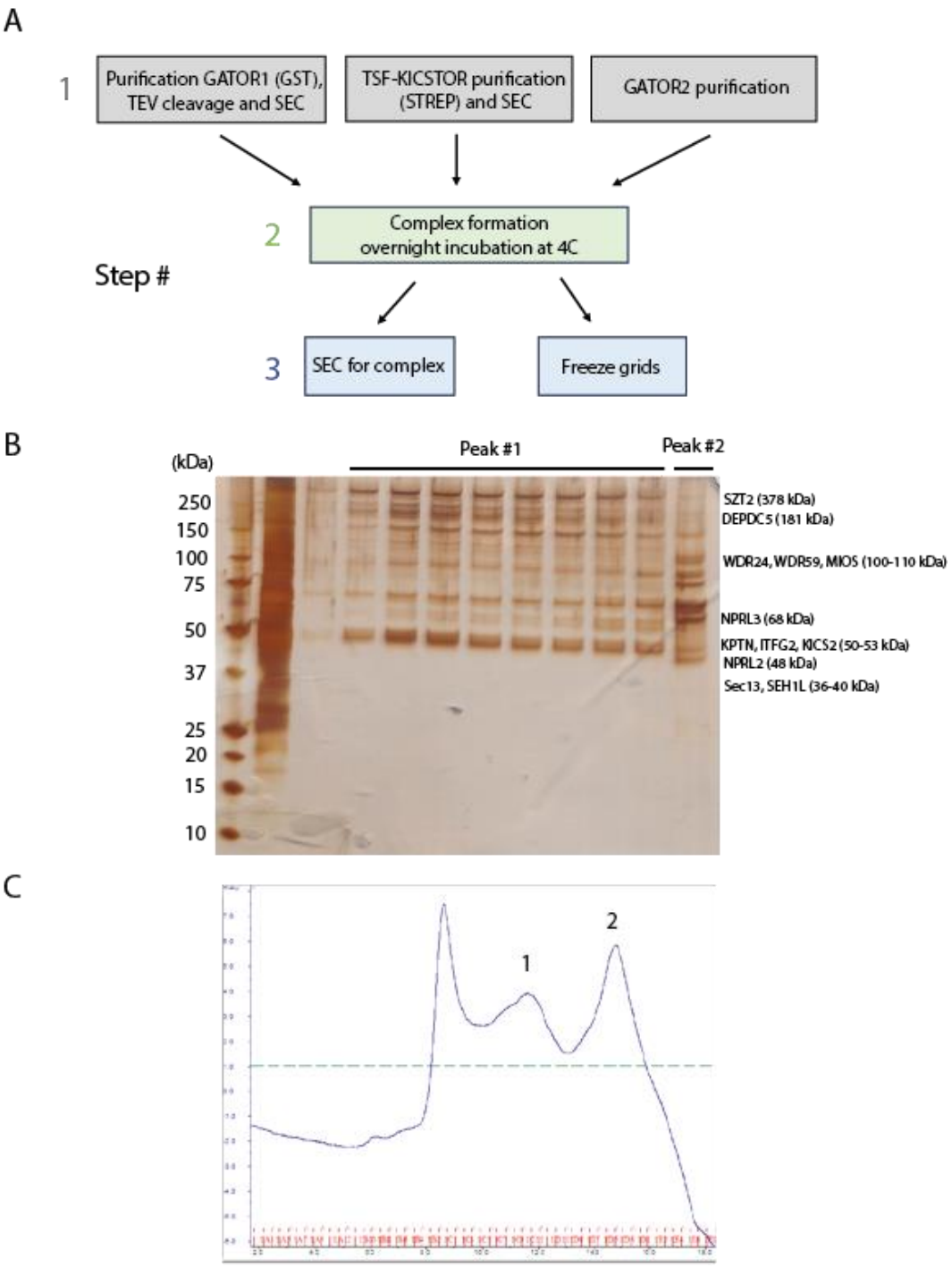


Figure 4.23 Attempt #2 - large scale complex formation for KICSTOR, GATOR1 and GATOR2. (A) Complex formation scheme for higher order complex. (B) SDS-PAGE visualization of fractions from Superose 6 increase and (C) corresponding chromatogram.

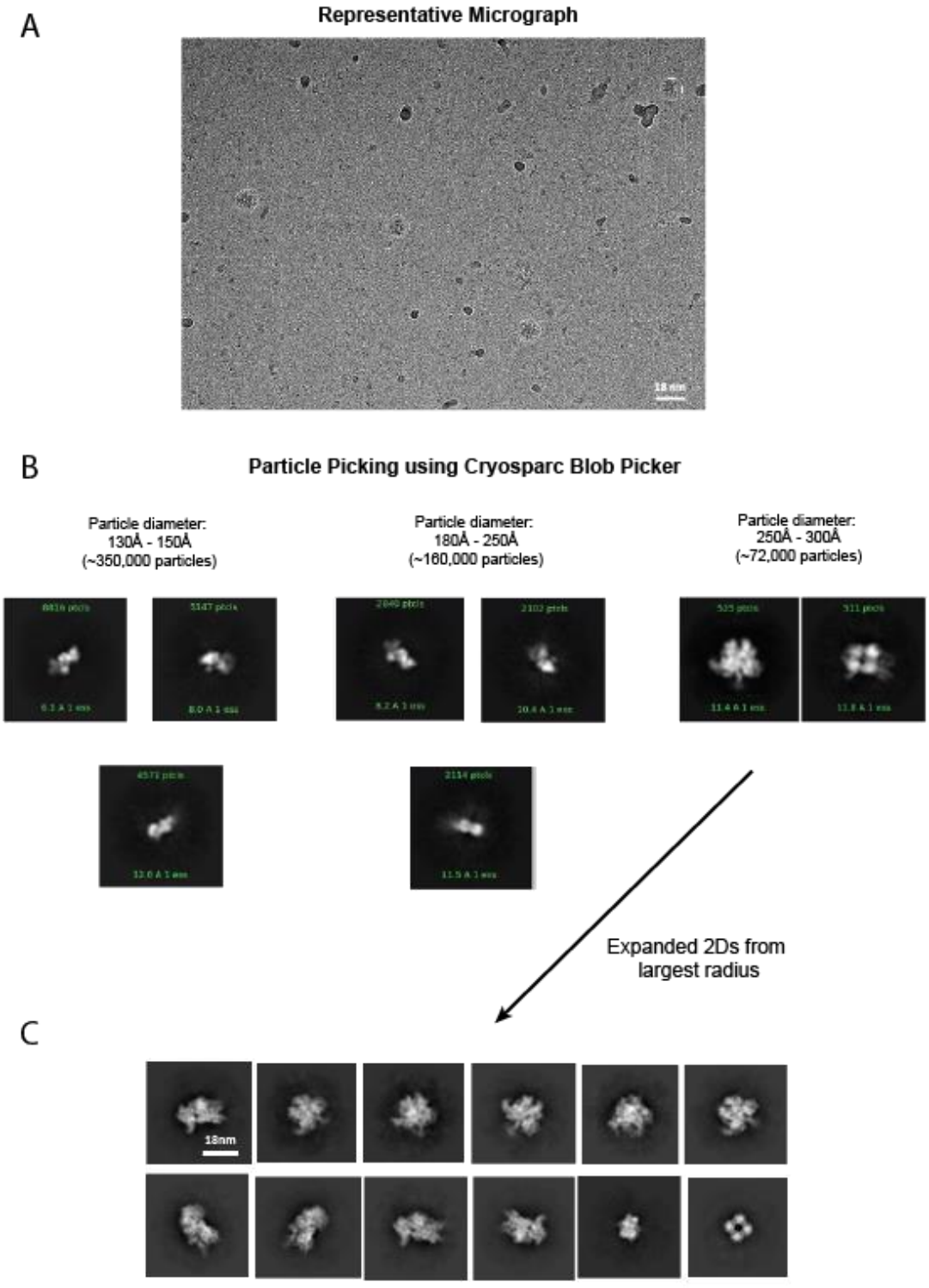


Figure 4.24 Cryo-EM for KICSTOR and GATOR components. (A) Representative micrograph from complex formation. (B) 2Ds generated from different picking diameters. (C) Particles from the larger diameter were expanded to generate more 2Ds.

References

1. Bar-Peled, L., Chantranupong, L., Cherniack, A.D., Chen, W.W., Ottina, K.A., Grabiner, B.C., Spear, E.D., Carter, S.L., Meyerson, M., and Sabatini, D.M. (2013). A Tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. Science *340*, 1100-1106. 10.1126/science.1232044.
2. Chantranupong, L., Wolfson, R.L., Orozco, J.M., Saxton, R.A., Scaria, S.M., Bar-Peled, L., Spooner, E., Isasa, M., Gygi, S.P., and Sabatini, D.M. (2014). The Sestrins interact with GATOR2 to negatively regulate the amino-acid-sensing pathway upstream of mTORC1. Cell Rep *9*, 1-8. 10.1016/j.celrep.2014.09.014.
3. Gu, X., Orozco, J.M., Saxton, R.A., Condon, K.J., Liu, G.Y., Krawczyk, P.A., Scaria, S.M., Harper, J.W., Gygi, S.P., and Sabatini, D.M. (2017). SAMTOR is an S-adenosylmethionine sensor for the mTORC1 pathway. Science *358*, 813-818. 10.1126/science.aao3265.
4. Wolfson, R.L., Chantranupong, L., Wyant, G.A., Gu, X., Orozco, J.M., Shen, K., Condon, K.J., Petri, S., Kedir, J., Scaria, S.M., et al. (2017). KICSTOR recruits GATOR1 to the lysosome and is necessary for nutrients to regulate mTORC1. Nature *543*, 438-442. 10.1038/nature21423.
5. Chantranupong, L., Scaria, S.M., Saxton, R.A., Gygi, M.P., Shen, K., Wyant, G.A., Wang, T., Harper, J.W., Gygi, S.P., and Sabatini, D.M. (2016). The CASTOR Proteins Are Arginine Sensors for the mTORC1 Pathway. Cell *165*, 153-164. 10.1016/j.cell.2016.02.035.
6. Tang, X., Zhang, Y., Wang, G., Zhang, C., Wang, F., Shi, J., Zhang, T., and Ding, J. (2022). Molecular mechanism of S-adenosylmethionine sensing by SAMTOR in mTORC1 signaling. Sci Adv *8*, eabn3868. 10.1126/sciadv.abn3868.
7. Peng, M., Yin, N., and Li, M.O. (2017). SZT2 dictates GATOR control of mTORC1 signalling. Nature *543*, 433-437. 10.1038/nature21378.
8. Algret, R., Fernandez-Martinez, J., Shi, Y., Kim, S.J., Pellarin, R., Cimermancic, P., Cochet, E., Sali, A., Chait, B.T., Rout, M.P., and Dokudovskaya, S. (2014). Molecular architecture and function of the SEA complex, a modulator of the TORC1 pathway. Mol Cell Proteomics *13*, 2855-2870. 10.1074/mcp.M114.039388.
9. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods *14*, 290-296. 10.1038/nmeth.4169.
10. Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol *180*, 519-530. 10.1016/j.jsb.2012.09.006.
11. Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods *14*, 331-332. 10.1038/nmeth.4193.
12. Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A.J., and Berger, B. (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. Nat Methods *16*, 1153-1160. 10.1038/s41592-019-0575-8.
13. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583-589. 10.1038/s41586-021-03819-2.

14. Mirdita, M., Schutze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. Nat Methods *19*, 679-682. 10.1038/s41592-022-01488-1.
15. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci *27*, 14-25. 10.1002/pro.3235.
16. Valenstein, M.L., Rogala, K.B., Lalgudi, P.V., Brignole, E.J., Gu, X., Saxton, R.A., Chantranupong, L., Kolibius, J., Quast, J.P., and Sabatini, D.M. (2022). Structure of the nutrient-sensing hub GATOR2. Nature *607*, 610-616. 10.1038/s41586-022-04939-z.