Regulatory domains of the Arf GEF Sec7 scaffold autoinhibited and active conformations

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Bryce Arthur Brownfield

August 2023

Regulatory domains of the Arf GEF Sec7 scaffold autoinhibited and active conformations

Bryce Brownfield, Ph.D.

Cornell University

The Golgi apparatus is a highly modular organelle that mediates the secretory pathway in eukaryotic cells. The final compartment, the trans-Golgi Network (TGN), connects the Golgi to the endomembrane system and balances the flow of membranes and cargo throughout the cell. Virtually all sorting and transport at the TGN is dependent on the small GTPase Arf1 which drives vesicle formation and membrane modification following its activation by the Arf GEF Sec7. A network of inter- and intramolecular interactions of Sec7 gives robust spatiotemporal specificity to the potent Arf GEF. The Sec7 homodimer is autoinhibited, and interactions with Rab1, Rab11, Arl1, Arf1 and the membrane recruit Sec7 and stimulate activity. These interactions have been mapped biochemically to the regulatory domains of Sec7, but the precise mechanism of autoregulation has not been established.

To identify the intramolecular interactions regulating Sec7, I determined the structure of full length Sec7 to 3.7 Å by cryo electron microscopy. This structure reveals an interaction between the GEF and HDS2 domains occludes the catalytic surface, and validate this interaction *in vitro* and *in vivo*. Serendipitously, the AlphaFold predicted structure placed the GEF domain adjacent to the DCB/HUS domain, and we demonstrate this is the active conformation. In this conformation the GEF domain displaces a loop in the DCB/HUS domain we refer to as the “D-loop”. Removal of the D-loop results in a modest increase in GEF activity *in vitro*, and when combined with HDS2 mutations synergistically increases activity and disrupts physiological function coincident with the severity of the HDS2 disruption. The structure also reveals a hydrophobic surface in the HDS4 domain is the dimerization interface. We show dimerization, an amphipathic helix in the HDS1-2 linker, and switching to the active conformation modulates membrane avidity. We therefore conclude the GEF domain is in equilibrium between active and autoinhibited conformations governed by the affinity of the GEF-HDS2 interaction, the D-loop, and the membrane.

Biographical Sketch

Bryce Brownfield was born in Pheonix Arizona on May 21 1993, and was raised in Larkspur Colorado from the age of three. The small town in Colorado fostered a love of nature, and his parents encouraged a passion for science and hard work. When not outside, Bryce spent his formative years at his work bench tinkering with anything he could get his hands on with delusions of building a submarine, a jet, or ending climate change, projects he maintains hope for. He graduated from Castle View High School in 2011, where he was active in theater (with roles as East in Almost Maine and Hortensio in the Taming of the Shrew), tennis, and working on Jeeps and enjoying nature with his Dad. Bryce was an average student, ironically struggling with chemistry. After a semester at Red Rocks Community College he went on to earn a B.S. in biochemistry from the University of Colorado at Colorado Springs, graduating summa cum laude in 2015.

While at UCCS, Bryce excelled in chemistry and was a student leader, tutoring chemistry and biology courses in the Excel Science Center and acting as vice president in the Student Members of the American Chemical Society for 3 years. His first research experience was working with Dr. Anatoliy Pinchuk in the BioFrontiers center developing microwave assisted synthesis of silver nanoparticles. In his 2nd year, Bryce began working in the lab of Dr. Wendy Haggren, where he designed a recombinant cellulase system to augment the masters work of Morgan Pinto, promising to give Saccharomyces cerevisiae the ability to digest both starch and cellulose from Buffalo Gourd to produce biofuel. Although never published, Bryce developed a passion for science and yeast. A major theme of this work was the increased activity of natural cellulase systems afforded by scaffolding enzymes into massive complexes called cellulosomes. For a recombinant cellulosome to function in yeast, these massive complexes needed to be secreted, so Bryce became fascinated with the complex machinery of the eukaryotic membrane transport systems and the secretory pathway.

After graduating from UCCS, Bryce spent a year working at a small lab in Colorado Springs establishing methods for clinical drug testing via homogeneous enzyme immunoassay and mass spectrometry while eagerly applying to graduate school. Bryce was accepted into the Cornell Biochemistry, Molecular and Cell Biology P.h.D program in 2017, and joined the lab of Dr. Chris Fromme the following year, where he delved into the field of membrane trafficking, and structural biology. Bryce was enthralled by the ease of imaging molecules with an electron microscope, and determined structures of several proteins by cryo em. These structures revealed the mechanism for autoregulation of the large Arf GEF Sec7, clarified substrate specificity for a bacterial NrnC-type RNases, and identified new mysterious oligomeric states of Ribonuclease PH.

Dedication

This work is dedicated to the many who inspire me, especially:

My dad, Jay Arthur Brownfield, my mom, Terry Brownfield, my amazing sister Laura Hansen, my fiance Maria, and Rubi.

Acknowledgements

To say I stand on the shoulders of giants is both a cliche and an understatement. I remain in disbelief that I was awarded the opportunity to study at Cornell under the tutelage of Dr. Chris Fromme. Chris's patience and stoic passion for science is truly remarkable, and continually raises the bar for the scientist I want to be higher than I thought achievable. I would not have overcome the many challenges faced during graduate school (especially through a pandemic) without his support, and I cannot express how thankful I am. My committee members Dr. Toshi Kawate and Dr. Gunther Hollopeter have also been instrumental, motivating me to ask the hard questions and keep an eye on the future. I left all of our meetings energized and full of new ideas to take to the lab. Dr. Holger Sondermann, who was on my committee until leaving Cornell for an appointment at Deutsches Elektronen-Synchrotron in Hamburg Germany, also deserves a special acknowledgement. I was able to collaborate with Dr. Sondermann on several projects, and he sponsored a number of trips to Germany to visit my fiance disguised as productive ventures to use the amazing facilities at CSSB. I'm incredibly grateful to have learned so much from all of the faculty at Cornell, but these giants in particular must have sore shoulders.

Cornell houses world class research facilities, which enhanced every aspect of my work. The Weill institute fosters the best labs on earth, and depends on so many people. The Director for most of my time Dr. Scott Emr is amazing in every sense, always excited about something and making science fun. I also must thank Jeff Ceria, Paige Reeves, Ellen Miller, and Bill Loftus, the beating heart of the institute. Lab support from Dave McDermitt, Revelin Bonomo, and for a time Nadia Nikulin (who is now a graduate student in the Dörr lab) simply cannot be over appreciated. The Cornell Center for Materials Research has been my second lab home, the cryo em masterminds Dr. Katherin Spoth and Mariena Silvestry Ramos taught me tremendously. Your patience with my late night/weekend texts and calls from the Arctica make you heroes. Your counterparts in Germany were equally heroic during my brief time there, Dr. Carolin Seuring and Cornelia Cazey, I can't thank you enough for welcoming me to the incredible Centre for Structural Systems Biology facility.

During my undergraduate career at the University of Colorado, Colorado Springs, Dr. Wendy Haggren set me on my current path. To this day her joyful "try it and see what happens" rings in my ears through every experiment. I originally chose to study biochemistry with the grand aspiration of brewing beer in my dorm room until she shared her work at S.M.A.C.S (Student Members of the American Chemical Society) and sparked my passion for research. Another mentor at UCCS, Dr. Janel Owens, actually did teach me how to brew (thanks Dr. Owens! Cheers!). I cannot list all of the amazing idols I had at UCCS, but Dr. Laura Zimmerman also holds a special place. Dr. Z was an amazing professor, and had the best life advice. I would not have pursued this path if not for these are the giants who first lifted me, and I am forever grateful to all of you.

I also have so many amazing friends who have made this possible. Current and former lab mates all helped so much both in and out of the lab. Thank you Carolyn Highland, Arnold Muccini, Saket Bagde, Ryan Feathers, Kaitlyn Manzer, Laura Thomas, Steve Halaby, Brian Richardson, Aaron Joiner, Shamar Wallace, Ari Broad, Anna Allen, Claudia Chambliss, Isaac Lamptey, Ryan Vignogna, and all the others who passed through the lab. I was excited to join a lab with peers who were more intelligent and talented than I, but you all made work a little less work too. I've made so many friends outside the lab who I want to thank, Sara Downs especially who helped me through the lowest points, and Erika Beyrent, Jacqueline Ehrlich, Alan Sulpizio, Chris Furman, Karim Omar, Ethan Sanford, and so many others. The BMCB community is amazing.

For any success I may have, I owe a huge debt of gratitude to my family. Growing up as the “mad scientist” set the stage well. My dad, who has always inspired me to work hard, do things the right way, enjoy life, and tinker made that possible. My mom, who has somehow managed to keep me well fed from 1,700 miles away, and always modeled the passion for taking care of the world I aspire to. My sister, who I can’t even write about without cracking a smile and laughing (“silver bells…”). My fiance Maria gave me the courage to come to Ithaca in the first place, and continues to show me how to be a better person. I truly could not have done this without Maria - I love you, and I don’t know how I swindled you into marrying me. To my dear Rubi, who is missed greatly, and my little dummy Palmer, my best friends. And Tanner Westlake, friends from Kindergarten through 23rd grade.

I am grateful for so many people. If you are reading this, you are one of them.

Table of Contents

List of Figures

List of Tables

CHAPTER 1 - Introduction

1.1 – The Foundations of Life

The essence of life is imbalance. Life is a constant climb away from thermodynamic equilibrium, maintaining some potential to do work. Ultimately, this is driven by natural reservoirs of low entropy, mainly sunlight, coupled to reactions that capture it. Shifting electrons from water onto less favorable carriers provides the potential energy to generate useful chemical energy. In all extant life known, this requires a delineation between a cell and its surroundings via a lipid bilayer (membrane) in water. A selectively permeable membrane allows small molecules and ions to be selectively concentrated in, or excluded from, the cell and hence capture entropy.

Many properties of membranes can be related to electrical components. As selectively permeable, they act as an insulator or resistor. When positive ions accumulate on one side of the membrane, negative ions cluster on the other side much like a capacitor. The utility of a membrane in controlling electrochemical gradients is essential to all the mechanisms life uses to generate energy. Known as the electron transport chain, electrons are transferred from high energy carriers back to the environment to move positively charged protons across a membrane to store energy. Even Alkaliphiles which live in pools of negative ions are enabled by the capacitive properties of a membrane to drive a proton motive force that generates ATP (J. W. Lee 2019). As Dr. Ian Malcolm explained in the 1993 film Jurassic Park - “life finds a way”. Membranes were fundamental to the evolution of life, and eukaryotes are the greatest innovators. Evolving a diverse set of membrane remodeling machines, eukaryotic cells developed the capacity to generate and maintain intracellular membrane limited compartments (organelles).

Organelles are a defining characteristic of Eukarya, allowing more complex and diverse systems to function (Gabaldón and Pittis 2015). They vary in composition, allowing for segregation of incompatible or toxic intermediates and tailoring conditions (such as pH and redox potential) to a suite of biochemical tasks. The lysosome (vacuole in yeast) maintains a low pH required for degradation of material via acid hydrolases, which would be incompatible with the alkaline lumen of the mitochondrial matrix. Similarly the ER maintains a slightly oxidized lumen to promote disulfide bond formation in folding proteins, incompatible with the slightly reducing cytosol. Some membrane limited compartments can be further subdivided by other functional characteristics. The rough and smooth ER (which are also continuous with the nuclear membrane) are defined by the functional components that cluster to them. The rough ER is defined by the presence of ribosomes giving a bumpy surface to protein manufacturing regions and distinct from morphologically smooth regions where lipid synthesis occurs. The inner and outer membranes of the mitochondria also form two distinct lumens that isolate the electron transport chain from the rest of the cell. Each compartment is highly specialized, and thus depends on rigorous homeostatic control over constant import and export of material (Tokarev, Alfonso, and Segev 2013).

Maintaining a Goldilocks balance between delivery of freshly synthesized cargos and recycling of damaged components is a complex task. Identifying, concentrating, and transporting cargos to specific destinations are all entropically unfavorable. Known mechanisms controlling these processes are diverse and multifaceted. As is the case with energy generation, virtually all of them depend directly on membranes, and can be elegantly simple.

Perhaps the simplest level of regulation is the physical properties of a membrane’s lipid constituents. The length and saturation of the hydrophobic tails defines both the rigidity and thickness of a membrane, and has been proposed as a model for selectively concentrating transmembrane proteins, based on the length of their transmembrane domains, and lipids by phase separation (Bretscher and Munro 1993; Welch and Munro 2019). The surface properties of a membrane depend primarily on the headgroups of the lipids, and can play a direct role in selectively recruiting proteins to a compartment. The most well-studied example of this is the phosphatidylinositol phosphates (PIPs), which can be selectively recognized by proteins based on the position of the phosphate group, and are generally recognized as barcodes to identify a membrane (Falkenburger et al. 2010). For example, PI(3)P is enriched in endosomes and autophagosomes, PI4P is enriched at the TGN, and PI(4,5)P_2 is a hallmark of the plasma membrane. A well established and tightly regulated group of enzymes modify PIPs, and many functional proteins bind them (De Matteis and Godi 2004). Headgroup chemistry also defines the overall surface characteristics of a bilayer. Primarily charge, but also surface packing defects along with rigidity (and curvature) can influence binding of peripheral proteins.

Membrane contact sites are increasingly recognized as important sites for regulating membranes and transferring material. First identified in the 1950s as “tubules with a precise orientation with respect to the mitochondria” (Copeland and Dalton 1959), membrane contact sites have now been identified between virtually all organelles, many with poorly defined functions. ER contact sites were initially described with roles in lipid and Ca^{2+} homeostasis, but it’s now clear they can transfer various small molecules, transmit signaling information or force and facilitate enzymatic membrane remodeling in *trans* (Feyder et al. 2015; Scorrano et al. 2019). However, intracellular vesicle carriers play a well established role in the transport of cargos for secretion, cell surface remodeling, targeted intracellular transport, and internalization of extracellular cargos. A hallmark of the eukaryotic cell, the Golgi, is the foundation for the secretory system and the majority of vesicular trafficking pathways. Remarkably, the few eukaryotes identified that lack cytologically discernible Golgi (i.e. M. *balamuthi*) or both Golgi and mitochondria (*Monocercomonoides* sp.) retain Golgi membrane trafficking machinery (Barlow et al. 2018; Karnkowska et al. 2016), highlighting the underlying importance of regulated membrane trafficking. In light of this, perhaps the Golgi should be considered the central dogma of eukaryotic biology.

1.2 – The Golgi apparatus

The Golgi was first described by Camillo Golgi in 1898, when he noticed fine intracellular structures of Purkinje cells stained by his “reazione nera” silver-osmium stain for which he was awarded the Nobel Prize in 1906. However, as many could not replicate his complicated staining technique, the structure was not generally accepted until the introduction of the electron microscope in the 1950s. The work of Dalton and Felix, among others, in 1954 solidified the Golgi as a *bona fide* organelle (Dröscher 1998; Dalton and Felix 1954; Farquhar and Palade 1981). Dalton and Felix described the “Golgi substance” of intact epididymis epithelial cells as stacked lamellae, and importantly observed association of small structures on their surface.

Already half a century old, the function of the Golgi was not entirely clear until the late 1960s, when a vectoral flow of material from the ER through the Golgi was demonstrated by George Emil Palade by EM autoradiography and cell fractionation of pancreatic cells. These foundational experiments demonstrated biochemically and morphologically the now archetypal stacked Golgi are polarized cisternae, carrying cargo from the ER to secretory granules (Figure 1.1 a) (Farquhar and Palade 1998; Jamieson and Palade 1967). Around this time several reports identified sugars, enzymes, and sulfates are enriched in specific Golgi compartments, providing evidence for the function of a multilamellar Golgi (Figure 1.1 b-d) (Morre, Merlin, and Keenan 1969; Neutra and Leblond 1966; Whur, Herscovics, and Leblond 1969; Godman and Lane 1964). Considered by many the dawn of modern molecular cell biology, this work solidified the Golgi as the primary sorting hub of the secretory pathway, connecting material synthesis at the ER to other compartments and performing functional modifications.

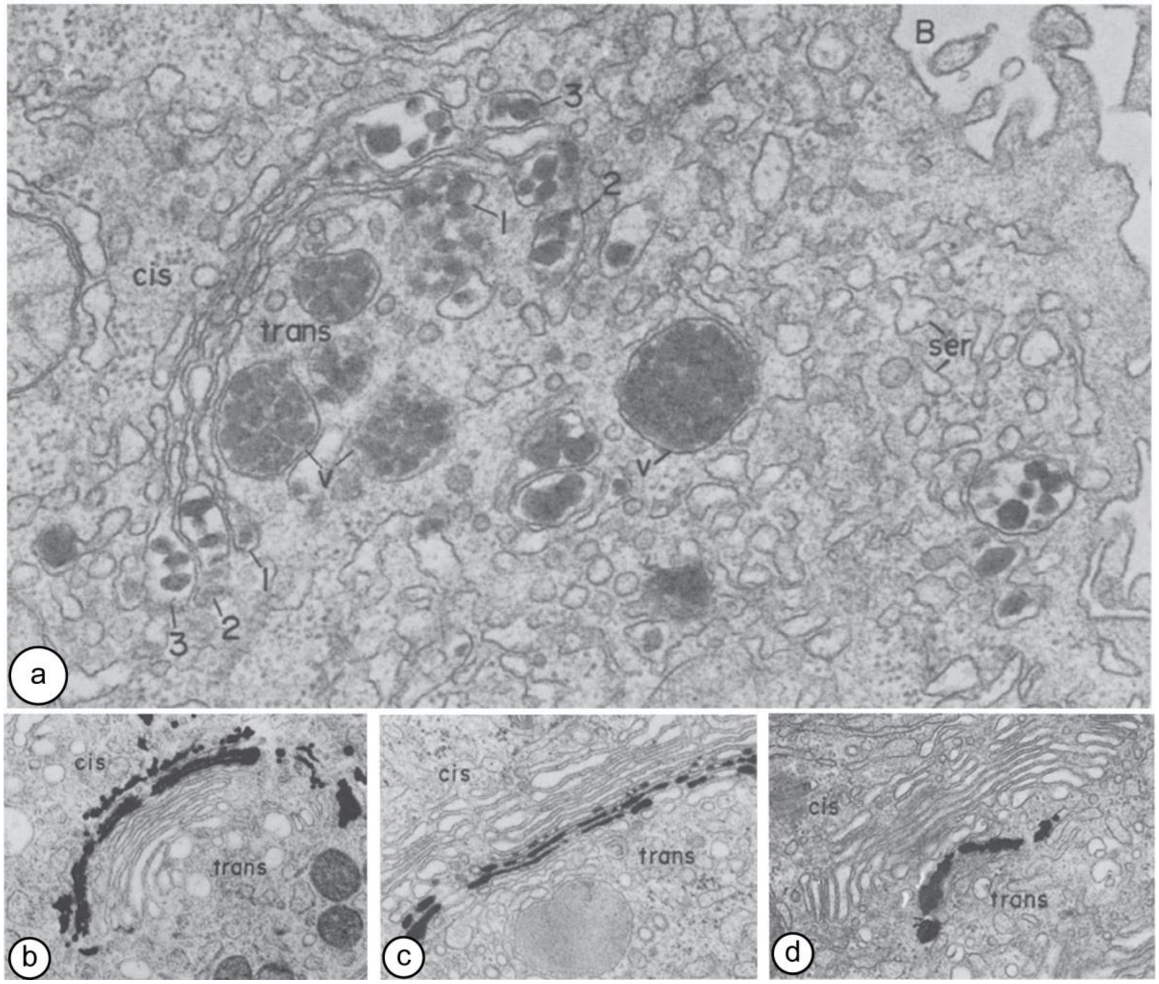


Figure 1.1 – Early thin section EM micrographs of mammalian Golgi. (a) A rat hepatocyte cell showing three distinguishable Golgi cisternae, and multiple electron dense secretory granules (labeled secretory vacuoles, 'v') between the TGN and the bile capillaries (labeled 'B'). (b-d) Epididymis rodent cells stained by osmium tetroxide (b) or cytochemically to identify thiamine pyrophosphatase (c) and acid phosphatase (d), highlighting the polarized distribution of osmiophilic lipids and enzymes. Taken directly with permission from (Farquhar and Palade 1981).

In the eukaryotic cell, approximately 20-30% of proteins and nearly all lipids are produced at the ER and transported through the Golgi (Gomez-Navarro and Miller 2016). Many of these proteins are hormones or enzymes that require modifications to become functional, and are considered pro-proteins or pre-pro-proteins. Two major modifications that occur along the secretory pathway are proteolysis and glycosylation, with major implications in disease pathology (Yarwood et al. 2020). Translocation into the ER occurs through the Sec61 complex, initiated by recognition of a signal sequence by the Signal Recognition Peptide (SRP) (Mandon, Trueman, and Gilmore 2013). Upon entering the ER, this signal peptide is removed by the Signal Peptidase (SP), and many are further cleaved in the Golgi. The causative agent in Alzheimer’s disease remains enigmatic, however disrupted processing of the Amyloid Precursor Protein (APP) by endosomal proteases is generally accepted to be involved (Jiang et al. 2014). Insulin, the first peptide hormone discovered, undergoes multiple proteolytic cleavage steps between the ER and secretory granules (M. Liu et al. 2014). Glycosylation is a modification where sugars and/or polysaccharides are added to a proteins or lipid. This alters their exposed surface, changing their solubility, stability, and interactome (Kattla et al. 2011). Recent advances in protein therapeutics utilize glycosylation to improve stability in pharmacological environments and develop new therapeutic modalities in the lab (M. Liu et al. 2014; Clausen et al., n.d.). Glycosylation can involve complex branched structures which would require many unique enzymes to build sequentially from sugar monomers if not for ‘cooperative sequential specificity’, successive additions by an assembly line of enzymes (Kleene and Berger 1993; Schoberer et al. 2010). The multi-lamellar morphology of the Golgi allows complex glycans to be constructed from a minimal set of enzymes. The successive compartments of the Golgi acts as a conductor for an orchestra of modifying enzymes.

Many models have been proposed to explain the dynamics of membrane trafficking through the Golgi, but none completely satisfy all observations. Conflicting reports of the contents of vesicle carriers, transport kinetics of various cargos, and differences between Golgi morphology in different species have all obfuscated a general model that explains all transport that occurs in every Golgi (Benjamin S. Glick and Luini 2011). It is well established however that:

(1) Cargos exit from ER exit sites (ERES) in COPII coated vesicles

(2) COPII vesicles are then transported to the cis face of the Golgi (or ER Golgi Intermediate Compartment (ERGIC) in metazoans), likely forming a new cisternae by homotypic fusion

(3) COPI coated vesicles bud from Golgi cisternae to recycle material back to the ER, or between Golgi cisternae, and cargos move anterograde from *cis* to *trans*/TGN (Figure 1.2) (Rothman 1994; Morin-Ganet et al. 2000; Papanikou and Glick 2009; Preuss et al. 1992; Bannykh and Balch 1997; Hammond and Glick 2000; Hughes et al. 2009; Paczkowski, Richardson, and Fromme 2015; Casler et al. 2019, 2022).

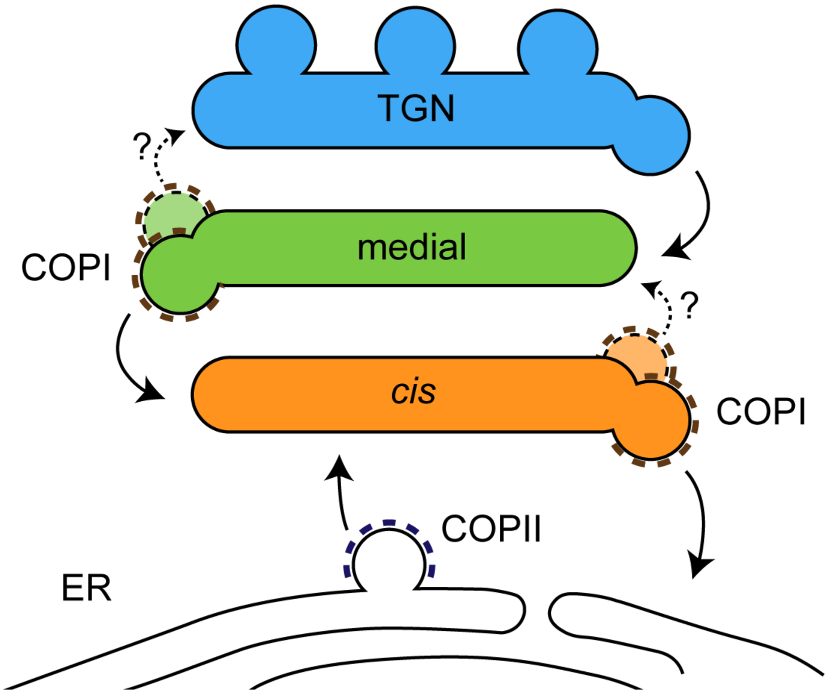


Figure 1.2 – Cartoon depiction of a Golgi. Cargos exit the ER in COPII vesicles, and fuse/form the *cis* Golgi. COPI vesicles recycle escaped ER proteins from the cis Golgi, and resident Golgi proteins from later to earlier Golgi cisternae. COPI vesicles may also facilitate anterograde transport of secretory cargo in some conditions. Cisternal maturation model posits no anterograde COPI transport, while the stable compartment model posits anterograde COPI transport carries cargo. Adapted with permission from Carolyn Highland's dissertation.

Whether COPI coated vesicles function in anterograde trafficking at the Golgi remains debated. The most widely accepted model, cisternal maturation, postulates COPI vesicles recycle Golgi resident proteins from later to earlier cisternae. A cisternae carries its secretory cargos through the Golgi while the resident enzymes "mature" from cis to trans. The main opposing model, stable compartments, postulates Golgi resident proteins are immobile while secretory cargos are carried from cis to trans in COPI vesicles (Benjamin S. Glick and Luini 2011). Cisternal maturation is regularly observed in yeast, both by the transition of markers at a compartment from cis to trans (Losev et al. 2006; Matsuura-Tokita et al. 2006; Gustafson and Fromme 2017; Highland and Fromme 2021), and by synthetic cargos trapped in the ER until released by a stimulus (Casler et al. 2020; Casler and Glick 2020). Cargos larger than the COPI coat (i.e. cell surface scales and algae and procollagen in mammals) are thought to require cisternal maturation (B. S. Glick and Malhotra 1998). Furthermore, retrieval of escaped ER proteins from the Golgi via COPI vesicles is well established (Gomez-Navarro and Miller 2016), and a similar process mechanistically extending through the Golgi seems most reasonable. However, small cargos appear to exit the Golgi much faster than would be expected by cisternal maturation, and experiments have shown cargos can transverse artificially immobilized Golgi compartments (Dancourt et al. 2016; Dunlop et al. 2017). Strangely, Golgi compartments containing an artificial large molecular aggregate that “staples” two sides of a cisternae together have been reported to immobilize the flat regions while cargos move through the rims, dubbed “rim progression” (Lavieu, Zheng, and Rothman 2013). If a single unified model can describe all Golgi trafficking, work remains to define it. It seems most likely from the available evidence that COPI vesicles can carry cargo in either direction, but generally act as retrograde transporters recycling material.

An important consideration is the process by which compartments are identified, and maturation is defined. The presence of fluorescently tagged known protein markers is a common and generally accepted way to identify the “age” of a cisternae. This is appropriate when maturation is delineated by the progression of trafficking pathways active at a compartment (i.e. COPI at early compartments and clathrin at later/TGN compartments). However, there are many functionally relevant markers with considerable overlap and variable residence times. The classically delineated “*cis*”, “medial”, and “*trans*” does not adequately represent the complexity of maturation (Pantazopoulou and Glick 2019), and can cause confusion. For example, the SNARE Tlg1 cycles between the PM and internal compartments assumed to be post Golgi endosomes in static images (Lewis et al. 2000). However, when tracked over time Tlg1 consistently arrives at compartments ~10-20 s before the TGN marker Sec7 (Day, Casler, and Glick 2018). Thus, restraint is warranted when interpreting these results, especially in non-physiological conditions.

1.3 – Machinery of the Golgi

Although the Golgi was first described in mammalian cells, much of the factors governing membrane trafficking were discovered and subsequently characterized in *Saccharomyces cerevisiae*. Foundational screens of *S. cerevisiae* mutants identified genes of the ***Sec***retory (SEC) and ***V***acuolar ***P***rotein ***S***orting (VPS) pathway (Novick, Field, and Schekman 1980; Robinson et al. 1988). Crucially, these factors are highly conserved, and there are numerous advantages to yeast as a model system for membrane trafficking. Their genetic tractability and rapid simple culturing enabled the isolation of SEC mutants by Peter Novick and Randy Schekman. Even with the rapid throughput of isolating secretion deficient mutant clones by measuring exposed invertase, they were only able to isolate two alleles initially (Novick and Schekman 1979). By ingeniously consolidating secretion deficient cells based on their density in a Ludox gradient, a technique that would have been impossible without the yeast cell wall, they accelerated this even further to identify 23 complementation groups. Furthermore, S. cerevisiae lost the tightly stacked ribbon structure of metazoan Golgi, making individual cisterna discernable *in vivo* without treatment with microtubule destabilizers (nocodazole) or superresolution microscopy (Losev et al. 2006). This does however introduce a caveat that yeast also lack the extensive tubules interconnecting cisternae in other species. Additionally, yeast appear to have a streamlined secretory pathway compared to metazoans, consolidating much of the endosomal system into the TGN (Figure 1.3) (Day, Casler, and Glick 2018). To highlight the utility of the yeast system, two forms of the adaptor complex AP-1 (AP-1A and AP-1B) differing only by the incorporation of the medium subunit (micro1A or micro1B) facilitate TGN intra Golgi recycling in yeast (Valdivia et al. 2002), and bidirectional TGN-RE transport in mammalian cells (Fölsch et al. 2003). Initially, the two forms in mammalian cells were shown to facilitate opposing transport pathways between the TGN and RE, recognizing specific PIPs. However, their expression is cell-type specific (only AP-1A is ubiquitously expressed) and they now appear to largely overlap (Guo et al. 2013; S. Y. Park and Guo 2014). Put another way, a decade after the role of AP-1 was demonstrated the function of the mammalian variant complexes were revised to probably fill redundant roles. Various combinations of the five mammalian AP subunits results in up to twelve AP-1 complexes, four AP-2, eight AP-3, one AP-4, and one AP-5, most thought to be redundant or fill cell type specific roles (Mattera et al. 2011; S. Y. Park and Guo 2014). While understanding the minutia of these subcomplexes is a fascinating prospect, yeast are the ideal model system to uncover the fundamental principles of membrane trafficking.

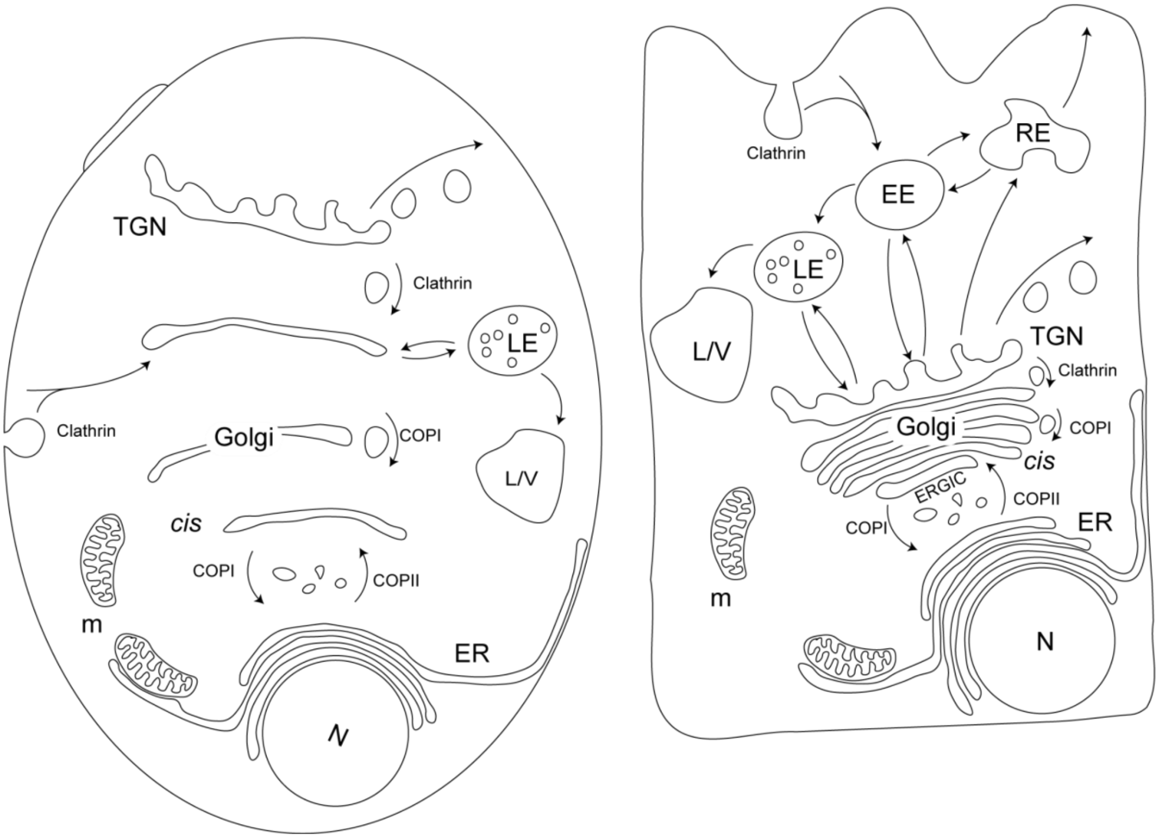


Figure 1.3 – Cartoon comparison of the yeast and mammalian endomembrane systems. L/V - lysosome/vacuole, LE - late endosome, EE - early endosome, RE - recycling endosome, ER - endoplasmic reticulum, m - mitochondria. Major differences include the dispersed Golgi, and the lack of the early endosome, recycling endosome, and ERGIC. In yeast (left), early endosome and recycling endosome markers overlap with the TGN marker Sec7, and endocytosed material marked by the lipophilic dye FM464 passes through the TGN first on route to the vacuole (Day, Casler, and Glick 2018). Mammalian cells often contain other cell-type specific post Golgi intermediates/compartments that are not depicted, such as melanosomes or insulin secretory granules.

Vesicular trafficking is a multistep process (Figure 1.4). There are a number of different factors that coordinate vesiculation, but the fundamental steps are remarkably similar. First, activation of a GTPases begins enriching cargos and deforming the membrane at the donor compartment. Coat proteins and cargos are continually enriched as this membrane ‘bud’ grows until the vesicle is fully formed and breaks free. The first GTPases is removed, causing the coat to disassemble, and a second GTPase is activated to mark the vesicle for transport along the cytoskeleton until it is captured by a tether coordinated by a third GTPase at the target compartment. A plethora of factors then coordinate the SNARE mediated membrane fusion, releasing the cargo. Vesicles are typically categorized by the coat protein, COPI, COPII, or clathrin, but there are important distinctions for some coats which can act in multiple pathways. The following section outlines the protein machinery underlying vesicle trafficking in the secretory pathway and endolysosomal system. Unless otherwise specified, the factors presented follow the yeast nomenclature.

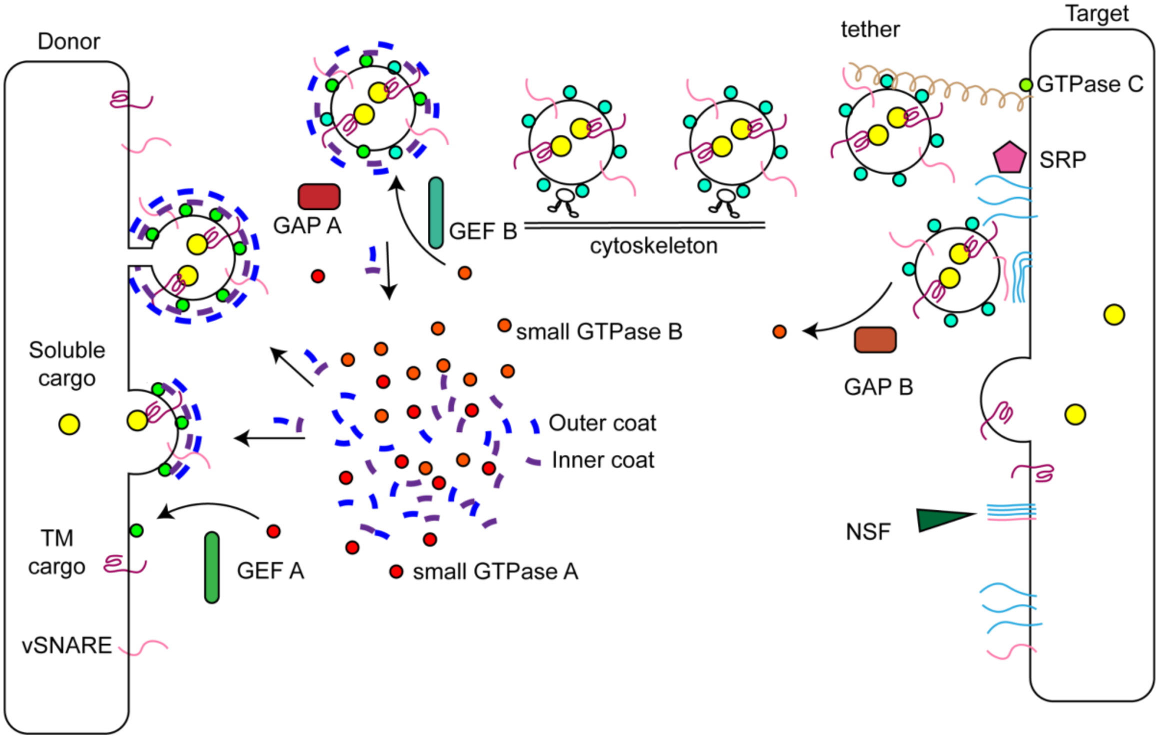


Figure 1.4 – General principles in vesicle trafficking. Vesicular transport starts at the donor compartment, where a GTP is activated by its GEF to initiate vesicle formation. Interactions between the GTPase, cargos, adaptor proteins and coats concentrates cargo and begins to deform the membrane to form a bud. Once the vesicle is fully formed, it breaks free from the donor compartment, and downstream GEFs and GAPs facilitate a transition between GTPase A to GTPase B. The vesicle is then transported along cytoskeletal networks, typically linked to a motor protein through GTPase B. Once near the target compartment, the vesicle is recognized by a tethering factor recruited to the target membrane by GTPase C. The tether brings the vesicle close to the membrane surface to be recognized by the cognate SNAREs to form the trans-SNARE bundle, facilitated by the SNARE regulatory Protein (SRP). The trans-SNARE bundle then “zips” to the cis-SNARE complex, providing the energy required to fuse the membranes. Cargos are released, and the SNARE bundle is reset through the action of NSF proteins, and recycled for another round of trafficking. Adapted from (Bonifacino 2014).

1.4 - GTPases

Regulation of every step in the secretory pathway depends on Ras superfamily GTPases, specifically of the Arf and Rab subfamilies. Each acts at specific subcellular locations within the cell to coordinate a diverse set of operations along the life cycle of a vesicle. In mammals, there are 6 Arf family GTPases (further subdivided into class I, II, and III), 20 Arl (Arf-like) proteins and over 60 members of the Rab family, many with overlapping, redundant, or cell-type specific roles (Donaldson and Jackson 2011; Hutagalung and Novick 2011). In the simplified yeast system, there are only 7 Arf/Arl and 11 Rab proteins. Members of the Arf family primarily coordinate vesicle biogenesis at the ER (Sar1) and the Golgi (Arf1/2). Notably, all vesiculation at the TGN is regulated by a pair of redundant paralogous which arose from the whole genome duplication in yeast (Arf1 and Arf2, 96% identical). For simplicity, I will refer to Arf1/2 as simply Arf1, as Arf1 contributes 90% of the total Arf1/2 expressed (Stearns, Kahn, et al. 1990). The role of Arl1 and Arl3 are less well understood, but they localize to medial/late compartments and seem to functionally overlap with Ypt6 (Panic, Whyte, and Munro 2003). Rabs play an important and well characterized role in distinguishing membrane identity, localizing to specific target compartments and vesicles to coordinate tethering, and interacting with motor proteins to facilitate active transport (Munro 2005; F. A. Barr 2009). The Rabs Ypt1, Ypt6, and Ypt31/Ypt32 (redundant homologs) in yeast coordinate trafficking at the early/medial Golgi, medial Golgi, and late/TGN respectively (Kim et al. 2016). Other important Rabs include Vps21 (Rab5), which coordinates late endosome maturation, Ypt7 (Rab7) which localizes primarily to the vacuole, and Sec4 (Rab8), which localizes to secretory vesicles and coordinates polarized secretion (Salminen and Novick 1987).

GTPases act as molecular switches, with a characteristic globular guanosine nucleotide domain adopting distinct conformations when bound to GDP or GTP. Regions known as the “switch 1”, “switch 2”, and “inter switch” regions adopt dramatically different conformations in the two states, altering the GTPase behavior significantly. Because they bind the nucleotide ligand extremely tightly and have very low intrinsic hydrolysis rates, an assortment of factors known as Guanine nucleotide Exchange Factors (GEFs) and GTPase Activating Proteins (GAPs) mediate the transition between the two states (Figure 1.5) (Roger S. Goody and Hofmann-Goody 2002). Control over this switch allows for spatiotemporal regulation of their effectors. In the inactive state, Arfs are are mostly soluble and are not known to bind any effectors, but upon activation they associate strongly with the membrane (Franco et al. 1995; B. Antonny et al. 1997), and recruit many different types of effectors, including coat proteins, adaptor complexes, tethering factors, and enzymes. The Rabs undergo a similar regulatory process, associating strongly with the membrane in the active state, but recruit different effectors including tethering factors and cytoskeletal motor proteins. Appendix 1 lists known Arf/Rab effectors in yeast. Crucially, the Arf1 effectors localize to only subregions of the Golgi, and thus must rely on some additional interactions (Munro 2005).

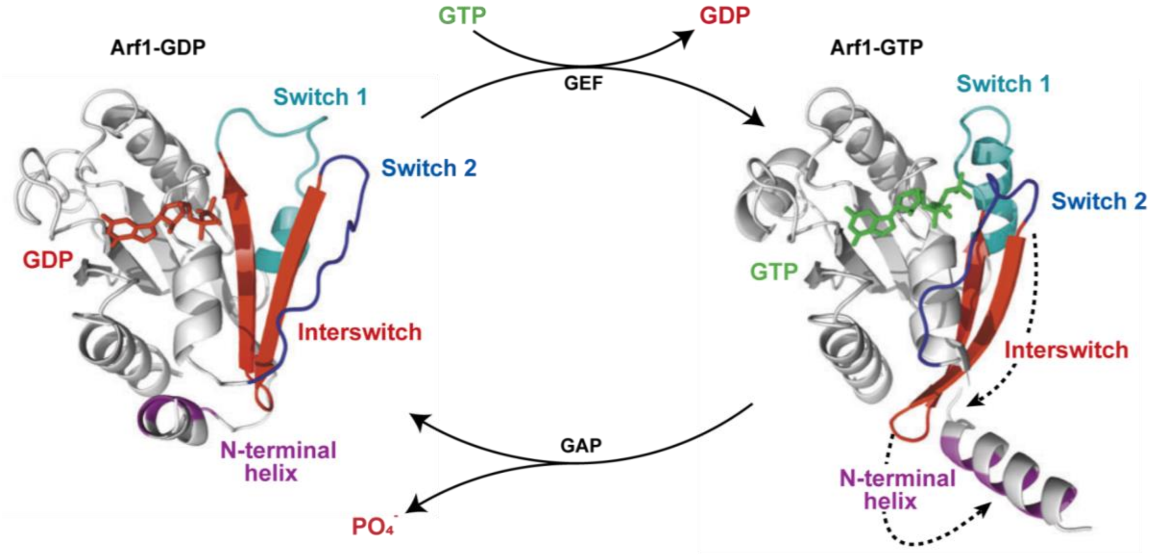


Figure 1.5 – The GTPase cycle. Arf1 transitions between inactive GDP bound (left) and active GTP bound states, through the action of GEFs and GAPs. Major conformational changes occur in the switch 1, switch 2, and interswitch regions, which for Arf1 displace an amphipathic helix. Adapted from (Gillingham and Munro 2007).

The GTP binding domains of Arfs and Rabs are remarkably similar, however their interaction with the membrane differs significantly. Rabs interact with the membrane through a long and variable C-terminal tail, while Arfs interact via an amphipathic N-terminal helix immediately preceding the globular domain. Both are modified by lipidation, either two C-terminal geranylgeranylations or a single N-terminal myristoylation for Rabs and Arfs respectively (B. Antonny et al. 1997; R. S. Goody, Rak, and Alexandrov 2005). An important exception is Arl3, which is instead modified by acetylation at the N-terminus and seems to require a small transmembrane protein Sys1 for its localization (Figure 1.6) (Setty et al. 2004). After modification, the hydrophobic geranylgeranyl groups of the Rabs are solubilized by a chaperone called GDI (GDP Dissociation Inhibitor), while the amphipathic helix and myristoyl group of Arf proteins are stabilized by a hydrophobic pocket within their GTP binding domain. This interaction in Arfs is disrupted by the interswitch region upon activation, directly coupling activation and membrane association. The bulky helix inserted into the outer leaflet contributes to membrane deformation in vesicle formation. This is in contrast with the Rabs, which are delivered to a membrane by the GDIs. Although early studies suggested another factor Yip3 (Pra1 in mammals) may act as a GDI displacement factor (GDF) (Sivars, Aivazian, and Pfeffer 2003; Dirac-Svejstrup, Sumizawa, and Pfeffer 1997), they are non-essential and no other GDFs have been identified. It is generally assumed the Rabs are somewhat non-specifically delivered to membranes in the inactive state, and subsequent activation prevents re-extraction (Figure 1.6). The lipid modification contributes to the Rab(GDP):GDI interaction suggesting extraction is favored when the Rab is inactive (Ignatev et al. 2008). Single molecule fluorescence has observed Rab5:GDP partitioning onto supported lipid bilayers with a short residence time (Bezeljak et al. 2020), but a GEF for Rab7 has also been reported to disrupt the GDI interaction, possibly serving to impart some specificity for the delivery (Y.-W. Wu et al. 2010). The GDI acts as a “proof-reader”, extracting Rabs from membranes where they are not active (Müller and Goody 2018). This difference in membrane delivery of Arfs and Rabs is relevant to their regulation and function.

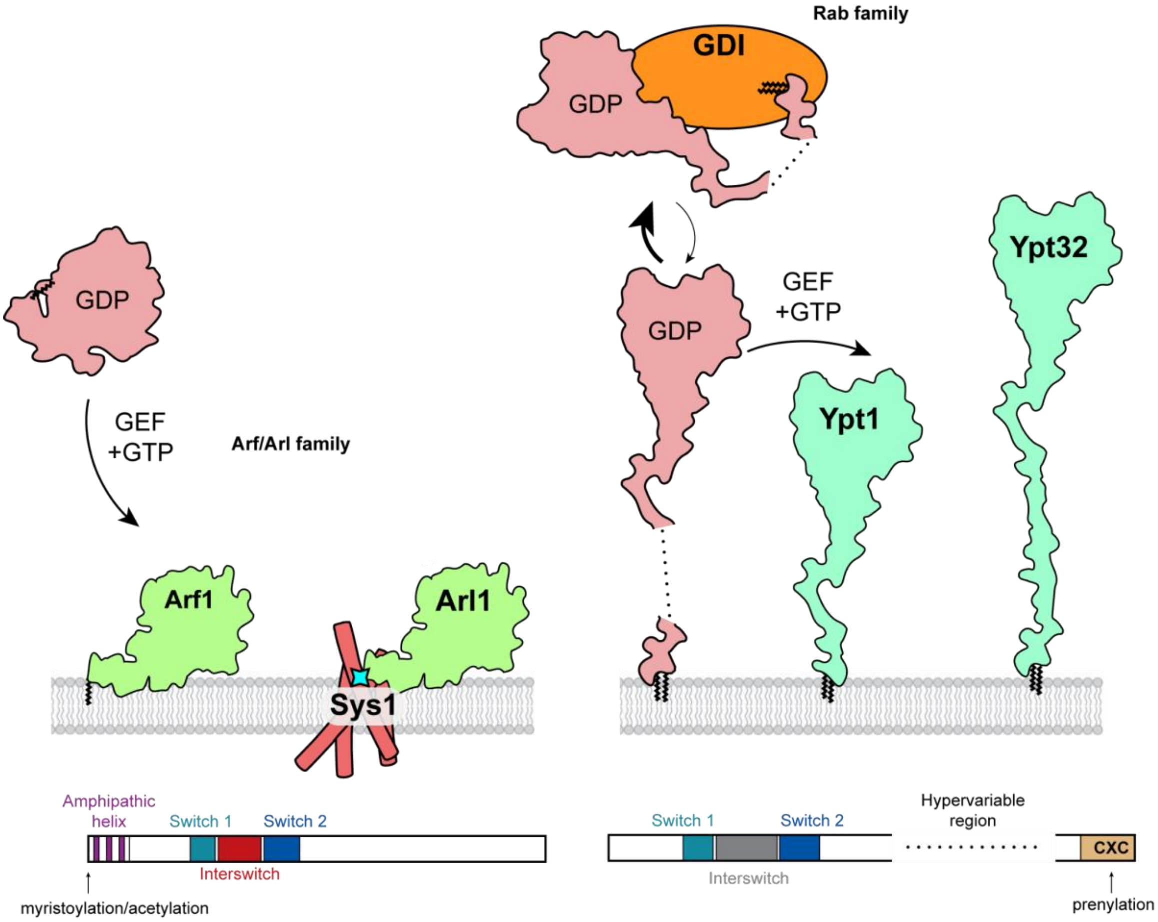


Figure 1.6 – Comparison of Arf family and Rab family GTPases on a membrane. Arf and Arl family GTPases (left) are myristoylated or acetylated respectively (black zig zag or turquoise star) at their N terminus. An intramolecular interaction stabilizing the membrane binding motif is disrupted by GEF activation, positioning the GTPase directly on the membrane surface. The Rabs (right) are prenylated at the C-terminus, at the end of a longer hypervariable linker. This is stabilized by an intermolecular interaction with the GDI, which transiently delivers the Rab to the membrane. GEF activation disrupts the Rab:GDI interaction, stabilizing them on the membrane.

1.5 – Regulation of Vesicle Trafficking at the Golgi

The Arf and Rab GEFs vary considerably in size, structure, and sequence, but mechanistically interact very similarly with their substrate (Müller and Goody 2018; Sztul et al. 2019). Nucleotide exchange is stimulated by transiently stabilizing a nucleotide free intermediate, which is replaced by GTP due to higher cytosolic concentrations of GTP than GDP. Although not rigorously tested empirically, this mechanism is not nucleotide specific and in general must not be for the exchange to occur favorably (Figure 1.7). A GEF directly competing with the nucleotide for binding would only do so appreciably if the GTPase had a higher affinity for the GEF, favoring a stable ternary complex in the nucleotide free state. Similarly, if an exchange factor were to maximally avoid displacing GTP (i.e., preferentially displace GDP), it would need equal affinity for both the GTP and nucleotide free GTPase, and therefore GTP binding would not displace the GEF. Higher cytosolic GTP concentrations allow moderate preference for displacing GDP, and this has been documented for the Ran GTPase GEF RCC1 (Roger S. Goody and Hofmann-Goody 2002), and Sec12 (A. M. N. Joiner and Fromme 2021).

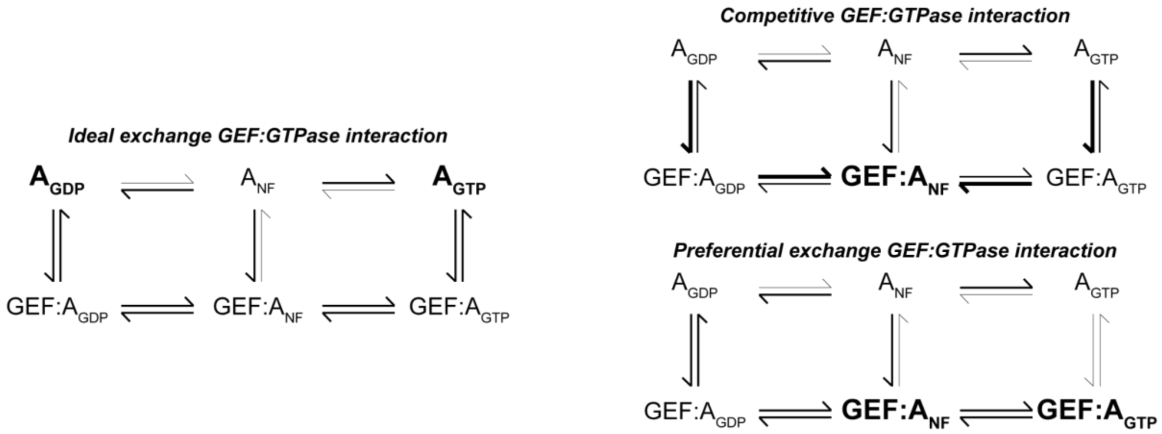


Figure 1.7 – GEF exchange reaction diagram. A GTPase “A” does not readily transition to a nucleotide free intermediate (A_{NF}) without the exchange factor. An ideal exchange factor displaces both GDP and GTP bound to “A” equally, allowing the reaction scheme to progress through A_{NF} (left). Competitive exchange factor stalls in the GEF:A_{NF} transition state (right top). To maximally prefer GDP displacement, the GEF must minimally displace GTP (GEF has equal affinity for A_{NF} and A_{GTP}), so GTP binding does not induce dissociation of the ternary complex.

Most Arf GEFs share a conserved catalytic domain, named the Sec7 domain after the founding member. Crystal structures of the domain bound to Arf1 reveal the mechanism of exchange is mediated by exposed hydrophobic grooves in the Sec7 domain interacting with the switch 1 and switch 2 regions, while a strictly conserved glutamate (E923 in S cerevisiae Sec7) displaces the coordinating Mg^{2+} and beta phosphate of the nucleotide diphosphate from the binding pocket of Arf1, stimulating the exchange (Beraud-Dufour et al. 1998; Goldberg 1998). Outside this domain, the Arf GEFs vary considerably, and can be classified as “large” (>100 kd) and “small” (<100 kd), and further divided into the GBF/BIG, cytohesin, EFA6, BRAG, and F-box families (Figure 1.4) (C. L. Jackson and Casanova 2000; Wright, Kahn, and Sztul 2014; Casanova 2007). Of these, the cytohesin, EFA6, BRAG, and FBX families are only found in metazoans, and all seem to play a role with metazoan specific functions at the PM.

The large Arf GEFs of the GBF/BIG family are the only conserved Arf GEFs that regulate trafficking at the Golgi. Yeast also express a fungal specific Sec7 containing GEF Syt1 that has been proposed to be the GEF for Arl1 (K.-Y. Chen et al. 2010), however this was not reproducible (See chapter 3), and no other GEF has been rigorously established for the Arls. Interestingly, plants contain only the GBF/BIG family GEFs (in arabidopsis there are three GBF- like proteins and five BIG-like proteins). Unlike the other metazoan subfamilies, the large Arf GEFs do not contain a recognizable membrane binding domain, however they do localize specifically to different compartments. In yeast, these are Gea1, Gea2, and Sec7 which localize to the early, medial, and late Golgi/TGN respectively (Gustafson and Fromme 2017; Bui, Golinelli-Cohen, and Jackson 2009; Franzusoff et al. 1991; Peyroche, Paris, and Jackson 1996; Spang et al. 2001). At the ER, Sar1 is activated by a unique GEF Sec12 that lacks the Sec7 domain and is anchored to the membrane by a C-terminal transmembrane domain. The exchange mechanism is similar however, Sec12 destabilizes the nucleotide by disrupting the switch regions and displacing the nucleotide and Mg^{2+} with a hydrophobic isoleucine (A. Joiner and Fromme 2020; A. M. N. Joiner and Fromme 2021).

The yeast Rabs are activated by multi-subunit complexes, TRAPPII and TRAPPIII, the Mon1-Ccz1 complex, the dimer Ric1-Rgp1, Vps9, and Sec2. The TRAPPII and TRAPPIII complexes share a catalytic core composed of 7 protomers, which induces nucleotide exchange by displacing the nucleotide with the C-terminus of Bet3 (Y. Cai et al. 2008). It was unclear how TRAPPII and TRAPPIII selectively activate Ypt31/32 and Ypt1 (Rab11 and Rab1 in humans) respectively until recently. The length of the hypervariable C-terminal tail is much shorter in Ypt1, which acts as a steric gate for access to the catalytic core (Thomas, van der Vegt, and Fromme 2019). Cryo em structures of TRAPPII and TRAPPIII revealed TRAPPII specific subunits lift the core off the membrane, and provide additional interactions with Ypt31/32 to overcome electrostatic clashes between the Rab and the catalytic subunit preventing TRAPIII activation (Bagde and Fromme 2022; A. M. Joiner et al. 2021). The Ric1/Rgp1 complex is the GEF for Ypt6 (Rab6), coordinating recycling at medial Golgi compartments. Post-Golgi Rabs, Vps21, Ypt7, and Sec4 are activated by Vps9, Mon1-Ccz1, and Sec2 respectively (Walch-Solimena, Collins, and Novick 1997; F. Barr and Lambright 2010; Nordmann et al. 2010).

Once a GTPase has fulfilled its function, GTP hydrolysis is induced to switch back to the GDP bound state. This is regulated by GAPs, which complete the catalytic hydrolysis pocket of the GTPase (generally with an Arginine) and convert bound GTP to GDP. The direction of this cycle is fundamental to the function of these GTPases. They are not simply nucleotide binding proteins, and would not be as effective regulators if hydrolysis activated them. The energy released by GTP hydrolysis allows for significantly less variance in the lifetime of the GTP bound state, improving the fidelity of a GTPase as a molecular timer (Li and Qian 2002). There are 4 Arf GAPs in yeast, Gcs1, Glo3, Age1, and Age2, which are conserved (ARFGAP1, ARFGAP2/3, ACAP1/3 and SMAP1/2 in humans), but not essential. They localize to distinct Golgi compartments, with decreasing abundance matching an increase of Arf1 from early to late compartments, however they may have partly overlapping/redundant roles (Yorimitsu, Sato, and Takeuchi 2014; Poon et al. 2001) (Kaitlyn Manzer, in progress). Importantly, coat complexes can stimulate GAP activity, as is the case for Sar1. Following assembly of the complete COPII coat, the outer coat (Sec13/31) complex stimulates the GAP activity of Sec23 (part of the inner coat Sec23/24 complex) to deactivate Sar1 (Bi, Mancias, and Goldberg 2007; Matsuoka et al. 1998). Deactivation of Arf GTPases is required for uncoating the vesicle, and hence progression through the vesicle trafficking pathway. Some evidence suggests that GTP hydrolysis is functionally relevant to cargo sorting. In the presence of non-hydrolyzable

GTP analogues or GTP locked Arf1 (Arf1[Q71L]), COPI vesicles form more readily but contain less cargo (Nickel et al. 1998; Lanoix et al. 1999). Further, at the TGN, Arf1 recruits the AP-1 adaptor and it remains bound after GTP hydrolysis. The Rab GAPs do not seem to be specific *in vitro*, however their localization *in vivo* suggests the relevant Rab GAPs for Ypt1, Ypt6, and Ypt31/32 are Gyp1, Gyp6, and Gyp2. The GAPs appear somewhat redundant, and coordinate one of the only processes in vesicle trafficking that is obviously favorable energetically.

Coat proteins play a direct role in reshaping the membrane surface and concentrating cargos. Coats are recruited by interactions between the coat, GTPases, and adaptor complexes (Bonifacino and Glick 2004). There are three well established protein coat complexes, clathrin, COP-I, and COP-II. Clathrin was originally described in EM micrographs of mosquito oocytes as a “20 mμ bristle coat” involved in endocytosis and formation of the yoke (T. F. Roth and Porter 1964). Later, clathrin isolated from pig brain isolates was characterized by SDS-PAGE and verified to be protein (Pearse 1975). Clathrin has since been attributed to vesicles budding from the PM, the TGN, and endosomes (Edeling, Smith, and Owen 2006). Clathrin is composed of a heavy and light chain, which oligomerize into triskelia that wrap around the membrane, forming a characteristic “soccer ball” pattern. The formation and composition of clathrin coated vesicles depends on the donor compartment on which they form. Some details of clathrin mediated endocytosis are still unresolved, however at the PM clathrin is recruited primarily by the adaptor complex AP-2, which is facilitated by mucins and requires the main phosphoinositide at the PM PI(4,5)P_2 (Partlow et al. 2022; Godlee and Kaksonen 2013). Importantly, its recruitment does not involve a GTPase. At the TGN, clathrin is recruited by interactions with Arf1 and the adaptor complex AP-1 or the epsin-related adaptor Ent5 to mediate TGN intra-Golgi recycling (Casler et al. 2019, 2022). The GGAs (Golgi-localized, γ-ear-containing, Arf (ADP-ribosylation factor)-binding proteins) GGA1 and GGA2 are clathrin adaptors that mediate TGN to late endosome transport . The next coat identified was COPI (COatamer Protein I) (Malhotra et al. 1989). Originally thought to carry all intra Golgi and Golgi to ER cargo, COPI vesicles are now recognized as the carrier for all early Golgi and Golgi to ER cargo recycling. Interestingly, mammalian and plant cells appear to have two classes of morphologically and compositionally distinct COPI vesicles, which have been proposed as anterograde and retrograde carriers (Orci et al. 1997; Malsam et al. 2005; Moelleken et al. 2007; Benjamin S. Glick and Luini 2011). COPI is a heptameric complex, which is recruited *en bloc* by Arf1 to membranes (Hara-Kuge et al. 1994). The fully assembled vesicle is composed of three COPI heptamers and six Arf1 proteins forming a triad structure (Dodonova et al. 2015; Faini et al. 2012). The COPII coat is composed of the inner Sec23/24 coat, which is recruited to the membrane by Sar1 and acts as a cargo adaptor, and the Sec13/31 outer coat, which is recruited as a heterotetramer to form a cage-like lattice (Stagg et al. 2006; Bi, Corpina, and Goldberg 2002). Importantly, the geometry of the inner and outer coats can vary to produce structures ranging from small vesicles to large tubules, allowing for transport of larger cargo (Stagg et al. 2008; Zanetti et al. 2013).

Some vesicular transport pathways do not seem to require a coat. The “ALP” pathway, named for the cargo alkaline phosphatase used in the genetic screen to identify it, concentrates cargos in AP-3 vesicles (Cowles et al. 1997; Stepp, Huang, and Lemmon 1997). The formation of secretory vesicles also remains somewhat of a mystery. Clathrin has been suggested (De Matteis and Luini 2008; Deborde et al. 2008), however clathrin is not essential for yeast, and does not disrupt secretion (Payne and Schekman 1985). In yeast, the exomer complex coordinates with Arf1 to sort secretory cargo and deform membranes (Paczkowski and Fromme 2014; C.-W. Wang et al. 2006), however it is only required for a fraction of secretory cargo, and is not conserved in other eukaryotes. Phase separation of the TGN membrane has also been suggested to concentrate secretory cargo (Parchure et al. 2022), and a kinetic analysis of a GFP tagged cargo (VSVG-GFP) suggests microtubule motors pulling on the membrane are sufficient to form secretory vesicles (Hirschberg et al. 1998)). A combination of phase separation and fragmentation by microtubule motors would be a simple model for “adaptor free” secretory vesicle formation, but this has not been directly investigated.

Fission of vesicles from the membrane is not completely understood. Clathrin vesicles require dynamin, a large GTPase that constricts the bud neck until fission occurs (Bruno Antonny et al. 2016). For COPI and COPII vesicles, a specific fission factor has not been identified. Reconstitution of COPI vesicle formation on Golgi membranes suggests the membrane composition may allow COPI vesicle fission, and binding of the Arf GAP may further stimulate fission (S.-Y. Park et al. 2019; J.-S. Yang et al. 2005). Importantly, the addition of “cone” shaped lipids, lipids with a smaller head group and a wider tail, seem to promote fission. Given the similarity between the cis Golgi and the ER, it is reasonable to assume COPII vesicles may similarly require only the formation of a complete coat to break free from the donor membrane.

After fission and before fusion, the vesicle must uncoat, be transported to its destination, and tethered. The precise timing of uncoating is not entirely clear. As coats are recruited by Arfs, uncoating immediately following GTP hydrolysis seems reasonable, but there is some evidence this is not the case. Fluorescence Recovery After Photobleaching (FRAP) of COPI and Arf1 components suggest COPI is stabilized on the membrane after GTP hydrolysis (Presley et al. 2002). Furthermore, GTP hydrolysis almost certainly never occurs simultaneously for every GTPases on a vesicle. Transport and tethering of a vesicle involves cytoskeletal motors (myosin, dynein, and kinesin) and tethers interacting with Rabs on the vesicle or target compartment (Hammer and Wu 2002).

Two classes of tethers, coiled coils (golgins) and multisubunit tethering complexes (MTC), capture vesicles at target compartments. For many tethers it is not clear exactly how they capture vesicles or bring them near the membrane for fusion. Ectopic localization of tethering factors at the mitochondria has been shown to divert traffic (Wong and Munro 2014) or induce Golgi-mitochondria stacking (Barlow et al. 2018), so some golgins are thought to play a role in maintaining the stacked Golgi in mammalian cells. The golgins are predicted to form long α-helical structures, and are classified by their known membrane binding domain typically at the C-terminus. Crucially, this domain can interact with either the vesicle or the target compartment, and presumably the other termini recognizes some feature on the other membrane. The GRAB and GRIP domains recognize GTPases, primarily Ypt1 and Arl1. Recognition of the cognate vesicle/target compartment has not been established for several golgins, although it is clear they recognize specific types of vesicles (Wong and Munro 2014). In yeast there are 5 golgins: Uso1, Rud3, Bug1, Coy1, and Imh1, functioning across the Golgi. Uso1 is recruited by Ypt1 to COPII vesicles (Sapperstein et al. 1996). The coiled coil tether Imh1 is recruited to medial Golgi membranes by Arl1 (Panic, Whyte, and Munro 2003). The MTCs in yeast are Dsl1 (ER), COG (intra-Golgi), GARP (intra-Golgi), Hops (late endosome/vacuole), CORVET (late endosome/vacuole), and exocyst (PM). There is evidence that both golgins and MTCs function in series to capture vesicles and prepare them for fusion (Anderson and Barlowe 2019).

Finally, fusion of vesicles requires the activity of SNARE proteins. Membrane fusion is a substantially energetically demanding process, requiring somewhere between 123.8 kJ/mol - 247.6 kJ/ml to break and reform the bilayers (Cohen and Melikyan 2004).

The SNARE proteins provide significant energy by “zipping” a four helix bundle spanning the two membranes. This zipping is mediated by Sec1/Munc18-like (SM) proteins that clamp the SNARE bundle, and are commonly associated with the tethering factors. Ultimately, this energy is derived from ATP hydrolysis coupled to unfolding the trans-SNARE bundle by N-ethylmaleimide-sensitive factor (NSF), resetting the SNAREs for another round of membrane fusion. The four helix bundle is composed of unique proteins, termed R-, Qa-, Qb-, and Qc- for the residue in the SNARE motif contributing to the zero ionic layer in the center of the bundle. There are 24 SNARE proteins in yeast, but out of hundreds of combinations only about 12 are fusogenic (Burri and Lithgow 2004; Südhof and Rothman 2009). These SNARE combinations are pathway specific, giving a final layer of regulation to the vesicle trafficking pathway, ensuring cargos are delivered to the correct target compartment

Specific localization of the regulatory factors has been of great interest, since it seems fundamental to Golgi maturation, and raises many intriguing questions. Why are GEFs with similar or identical catalytic domains active on a specific substrate and at a specific compartment? How do similar or identical GTPases recruit specific effectors, and why do they localize to specific compartments? How is directionality of membrane transport achieved? Overarching themes have emerged in answering these questions.

1.6 - Golgi membranes

As is the case with every intracellular compartment, membrane composition is critical, providing identity to a compartment and setting the foundation for complex functions to arise. Interactions with specific lipids are less prevalent at the Golgi, with the exception of PI4P at the TGN. The PI4 kinase Pik1 is recruited by Arf1 to generate PI4P and is essential to the final stages of the Golgi (Highland and Fromme 2021; De Matteis and Luini 2008). PI3P is generated by Vps34 primarily at the vacuole, and is required for efficient delivery of endocytic cargo (Wurmser and Emr 1998; Gillooly et al. 2000). Membrane charge, thickness, and rigidity do differ considerably across the Golgi, generally thought to be a gradient of ER-like at the cis Golgi to PM-like at the TGN.

Membranes are composed of many different lipids, however sphingomyelin (SM) and sterols are considered major contributors to membrane rigidity/order (along with acyl chain saturation), and phosphatidylserine (PS) is an anionic lipid and a major determinant of surface charge (along with PI4P, and others). A detailed description of Golgi cisterna lipidomics has been elusive, however as the Golgi acts as the primary intermediate between the ER and the PM, some general conclusions can be drawn from the available evidence (Bigay and Antonny 2012).Originally, Bretscher and Munro made the observation that the length of the transmembrane spanning helices is substantially longer in PM proteins compared to Golgi resident proteins (Bretscher and Munro 1993), and later analysis with more available sequences confirmed that this can be predictive of a proteins localization (Sharpe, Stevens, and Munro 2010). By EM, the ER membrane does appear thinner than the PM, and mass spec lipidomics finds PM is enriched in saturated lipids (Schneiter et al. 1999). The ER is also a relatively neutral membrane, while the PM contains more PS on the cytoplasmic leaflet (Holthuis and Levine 2005). Lipidomics of isolated COPI vesicles suggests SM and sterols are excluded by the highly curved membrane, and TGN/endosome membranes are enriched in anionic phospholipids, sphingolipids, saturated glycerolipids, and sterols (Brügger et al. 2000; Klemm et al. 2009). Somewhat paradoxically, secretory vesicles were also found to be enriched in sterols by Klemm et. al., an apparent contradiction with their depletion in COPI vesicles. Cholesterol has been shown to have different effects on membrane fluidity at the surface/center of a bilayer, decreasing fluidity at the surface and increasing fluidity near the center

(Subczynski et al. 2017). Thus it is perhaps not surprising that sterols behave differently in vesicles budding from membranes of different thickness. Importantly, leaflets of the membrane bilayer are not necessarily identical. Most lipid synthesis occurs on the cytosolic face of the ER membrane, and requires transmembrane proteins known as “flippases” or “scramblases” to evenly distribute lipids between leaflets (Buton et al. 1996; Coleman and Bell 1978). Surprisingly, more recent experiments suggest the anionic lipid PS is mainly in the lumenal leaflet of the ER and early Golgi, and the cytoplasmic leaflet of the TGN and PM (Fairn et al. 2011). Together, the available evidence suggests that the Golgi membrane transitions between ER-like (thin, neutral, more packing defects) to PM-like (thick, anionic, more ordered). The transition has been suggested to be even more abrupt than would seem by bulk lipid analysis, resulting in two “zones” (Bigay and Antonny 2012). This is primarily due to membrane modifying enzymes at the late Golgi. Two flippases, Neo1 and Drs2, transfer PS or PS and PE from the lumenal to cytoplasmic leaflets of the late Golgi and endosomes (C. Y. Chen et al. 1999; Wicky, Schwarz, and Singer-Krüger 2004). Lipid modifying enzymes such as phospholipase D are also activated by Arfs (Hong et al. 1998), further increasing the anionic charge of the TGN membrane by converting neutral phosphatidylcholine to anionic phosphatidic acid.

It’s worth noting that in the absence of direct empirical data for early Golgi membranes, the cisternal maturation model contributes significantly to the assumption that Golgi membranes transition somewhat smoothly. However, the available evidence is largely incompatible with the stable compartments model. If cisternae were static, it would be reasonable to expect the TGN membrane composition to be relatively comparable to the ER, notwithstanding leaflet asymmetry, and likely somewhat depleted of SM and sterols since they seem excluded from COPI vesicles. The observed composition of the TGN membrane is more consistent with a model where PM lipids are successively diluted in intracellular compartments, i.e. cisternal maturation, but further analysis could be informative.

The charge and packing defects of the Golgi membrane are surprisingly significant to the localization of some proteins. Two G proteins Rac1 and Rac2 interact with PM and endosome membranes respectively, differentiated by the number of basic residues in their membrane binding domain, and the anionic surface seems necessary for the localization of K-Ras, Rac1, and c-Src to phagosomes (Ueyama et al. 2005; Yeung et al. 2006). After myristoylation, Arf1 in the GDP bound state also weakly interacts with anionic membranes *in vitro*, which could contribute partly to its specificity (Franco et al. 1995; B. Antonny et al. 1997). However, Arf1 can be activated by a GEF on neutral membranes, so this is only a minor contribution. Two major membrane curvature sensing domains have also been identified, the BAR domain and Alps motif, which are commonly found in GEF and GAP proteins. Importantly, the BAR domain localizes to specific compartments, not by curvature recognition but by nonspecific electrostatic interactions (Peter et al. 2004). The specific interactions of PI4P have not been fully demonstrated, although it is clearly an important lipid for the formation of secretory vesicles and localization of the AP-1 cargo adapter (Highland and Fromme 2021; Y. J. Wang et al. 2003).

1.7 – Molecular logic of the Golgi

Membrane composition alone is not sufficient to account for the specificity of most regulatory factors. A network of interactions between GEFs, GTPases, GAPs, and effectors is increasingly apparent. Most prominent are the cascades of GEFs and GAPs; one GTPase recruits the GAP for an upstream GTPase and/or the GEF for a downstream GTPase (Mizuno-Yamasaki, Rivera-Molina, and Novick 2012; Pfeffer 2012; Stalder and Antonny 2013). This concept was coined by Ortiz et al. in 2002, when they identified the

Rab Ypt31/32 recruits the Sec4 GEF Sec2 to secretory vesicles (Ortiz et al. 2002). Since this conception many "crosstalk" interactions between pathways have been elucidated. The network of Ypt31/32 interactions has expanded to include the late Golgi Arf GEF Sec7, and GAPs for earlier Rabs, Gyp1 and Gyp6 (McDonold and Fromme 2014; Rivera-Molina and Novick 2009; Suda et al. 2013). However, more recently the recruitment of Gyp1 and Gyp6 by Arf1 and TRAPPII respectively has been demonstrated to be more relevant as they precede Ypt31/32 (Thomas and Fromme 2016; Thomas, Joiner, and Fromme 2018; Thomas, Highland, and Fromme 2021). Sec7 is recruited by other GTPases, including Arl1, Ypt1, and Arf1 creating positive feedback (McDonold and Fromme 2014; Richardson and Fromme 2012). The early/medial Arf GEFs Gea1 and Gea2 are also recruited by Ypt1, but not the others (Gustafson and Fromme 2017). Arl1 and Arl3 are also thought to act as a cascade, as Arl3 is required for the localization of Arl1, however the details of this have not been clearly established (Panic, Whyte, and Munro 2003; Setty et al. 2004). In humans, the endosomal Rab5 GEF Rabex-5 is recruited by both an upstream Rab, Rab22 and positive feedback through a Rab5 effector Rabaptin-5 (Zhu, Liang, and Li 2009; Horiuchi et al. 1997; Lippé et al. 2001).

Evidence suggests the network of GEF/GAP functional interactions extends beyond the GTPase cycle. Classically GTPases have been considered the primary regulators of trafficking machinery, and GEFs and GAPs function through their substrates. However, the early Golgi Arf GEF Gea1 (GBF1 in mammals) has also been suggested to recruit the COPI coat to membranes, increasing the local concentration before Arf1 activation (Deng et al. 2009; Catherine L. Jackson 2014). The TRAPP catalytic core has also been shown to interact directly with the COPII coat (H. Cai et al. 2007; Sacher et al. 2001; Tan et al. 2013; Zhao et al. 2017), most likely to facilitate Ypt1 activation by TRAPPIII. The significance of these interactions is less clear, but there is significant interplay between all regulatory elements at the Golgi. This coincidence detection forms a network of logic gates which together behave like a “Golgi algorithm”.

1.8 The Arf GEF Sec7

The TGN is the intersection between the endolysosomal trafficking pathways and the Golgi. Morphologically, it is a dynamic compartment that varies even between cell-types (Clermont, Rambourg, and Hermo 1995). In general, the TGN forms a highly tubular structure, often appearing to “peel off” of the Golgi stack in mammalian cells (Rambourg, Clermont, and Hermo 1979; J. Roth et al. 1985; Griffiths and Simons 1986). Originally the TGN was thought to only form from the last trans cisternae of a mammalian Golgi stack, however 3D tomography suggests the earlier two cisternae can tubulate as well. Notably, the last cisternae has distinct trafficking pathways, indicated by the presence of clathrin coated buds (De Matteis and Luini 2008). Tomography also revealed structures annotated vesicles by TEM can actually be continuous with the highly tubular structure (Ladinsky et al. 1994). The required trafficking at the TGN is distinct from the earlier Golgi, as secretory/endosomal cargos must be sorted into export carriers. Virtually all of these pathways require Arf1, activated by its GEF at the TGN Sec7.

The Sec7 name originated in the foundational sec screen, where two alleles were noted to cause enlarged double membrane structures termed “Berkeley bodies” to form (Novick, Field, and Schekman 1980). These structures were later found to be marked with components of the cytoplasm to vacuole targeting (Cvt) pathway, a pathway that mediates sorting of vacuolar hydrolases such as Ape1 that are synthesized in the cytosol, and shares most of the machinery in autophagy (Reggiori et al. 2004; Harding et al. 1995; Scott et al. 1996). Further characterization of Sec7 found it was involved in secretion of nearly all PM proteins (Novick and Schekman 1983; Tschopp, Esmon, and Schekman

1984). Soon after, Arf1 was identified as a small GTPase regulating the Golgi and interacted genetically with Sec7 (Stearns, Willingham, et al. 1990). A fungal toxin, brefeldin A (BFA) was soon found to disrupt the formation of COPI coats by preventing Arf1 activation, and used to isolate Sec7 as a BFA sensitive Arf GEF from bovine brain (Morinaga et al. 1996; Helms and Rothman 1992). Both Sec7 and the early Golgi GEFs Gea1 and Gea2 are BFA sensitive, and the mammalian Gea homologue was isolated from a BFA resistant cell line in 1999 (Claude et al. 1999). Interestingly, while yeast have two early/medial Arf GEFs (Gea1 and Gea2) and one late Golgi Arf GEF (Sec7), mammals have a single Gea homologue Gbf1, and two Sec7 homologues Big1 and Big2. A third Sec7 homologue in mammals, Big3, lacks the catalytic glutamate in the GEF domain and so is predicted to function in some other capacity. In yeast Gea1 and Gea2 localize to unique regions of the early Golgi, however neither are essential and expression of a single Gea from both promoters suppresses phenotypes of a single deletion (Gustafson and Fromme 2017). Neither can rescue a *sec7*Δ null, indicating they play distinct roles at the Golgi. Interestingly, the filamentous fungi *Aspergillus nidulans* contains only two large Arf GEFs, GeaA and HypB corresponding to Gea1/2 and Sec7 respectively, and a single residue substitution in a region identified as essential for Gea2 membrane localization allows GeaA to suppress *hypB*Δ (Arst et al. 2014; Muccini, Gustafson, and Fromme 2022).

Most of the recognized domains are shared between Sec7 and Gea1/2 (Sec7 shares 36.23% and 35.94% identity with Gea1 and Gea2 respectively), however biochemically they differ considerably. In total, there are 6 shared domains: the Dimerization and Cyclophilin Binding and Homology Upstream of Sec7 (DCB/HUS) domains, the catalytic Sec7 domain, and three Homology Downstream of Sec7 domains (HDS1-3). An additional C terminal domain is present in Sec7 (HDS4) (Figure 1.8 A). Domains outside the catalytic Sec7 domain are thought to play regulatory roles. The domain architecture originally described in 2005 by Mouratou et al. has been used extensively to characterize Sec7 (Mouratou et al. 2005). I will outline here previous work to characterize these domains, however as detailed in Chapter 2 several important regions lie in the unannotated regions of the original domain delineation, so I will also highlight these discrepancies. Additionally, previous analysis has used “functional Sec7” ($Sec7_f$), which lacks only the first 202 residues that are not conserved or essential *in vivo* (Richardson and Fromme 2012).

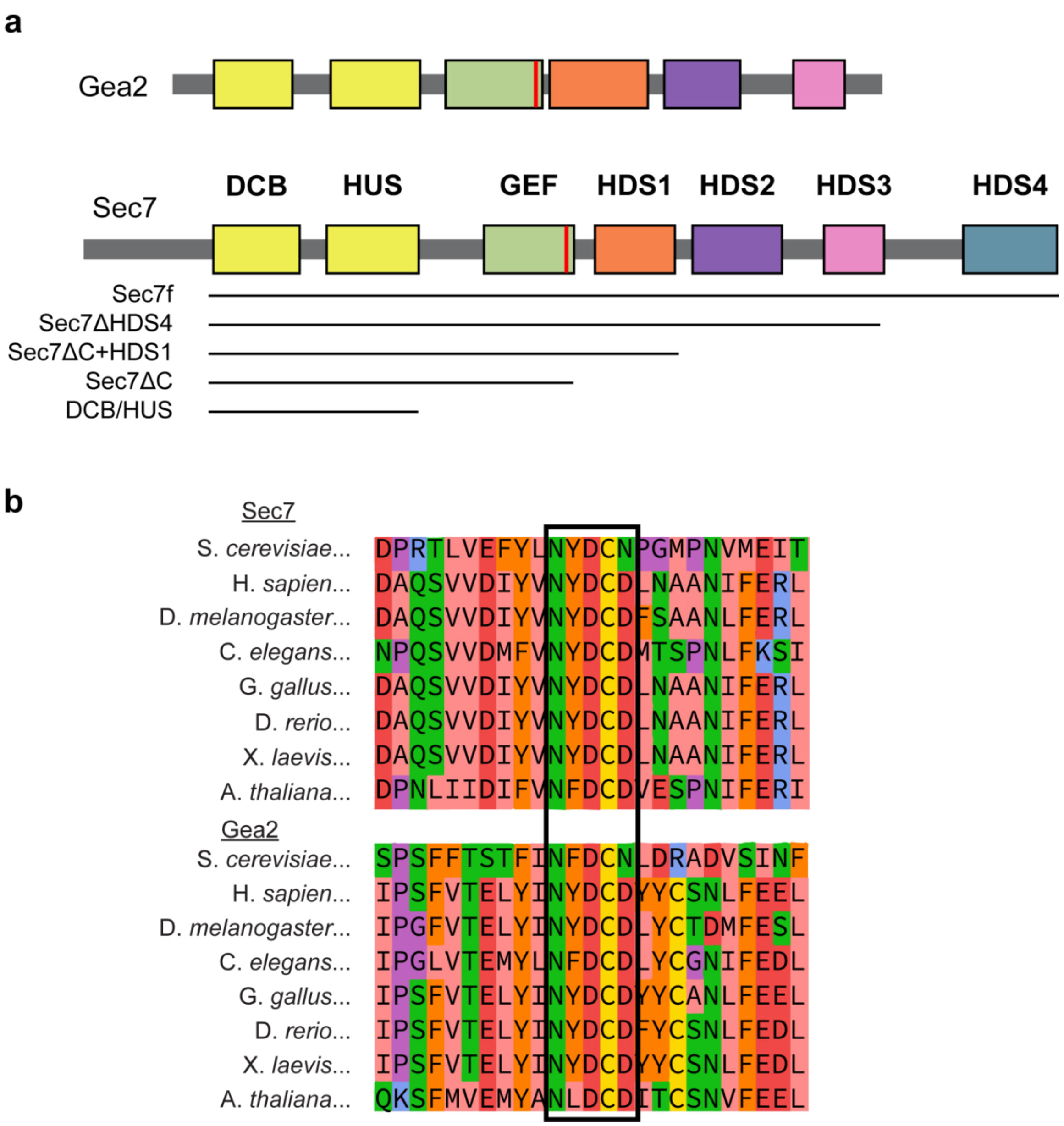


Figure 1.8 – domain architecture of Sec7 and the HUS box. Domain architecture of S *cerevisiae* Sec7 and Gea2 (a). Sequence alignment of the HUS box, showing the invariant N(Θ)DC(D/N), Θ denotes hydrophobic. Colored by physicochemical properties of the side chain (Zappo coloring scheme).

The DCB/HUS domains form a single structural unit implicated in homodimerization of Gea1/2 (Ramaen et al. 2007), and membrane recruitment for both Gea1/2 and Sec7 (Bhatt et al. 2016; McDonold and Fromme 2014; Richardson et al. 2016; R. Wang et al. 2016; Galindo et al. 2016). It is well established that Gea2 dimerizes via the DCB/HUS domain, however in Sec7 this interface interacts with Arl1, recruiting Sec7 to the membrane. The *sec7*-1 allele encodes a single mutation in the DCB domain (S402L) that disrupts membrane localization, although it is not clear why this substitution has this effect. As discussed in Chapter 2, this mutation does occur on the surface of the DCB domain predicted to interact with the membrane, however it is tightly packed in the DCB domain and there is no obvious reason a serine would interact considerably with an anionic membrane. It is likely this substitution results in some local folding defect.

In Sec7, the DCB/HUS domain confers a significant enhancement to GEF activity compared to the isolated GEF domain (Richardson et al. 2016). Importantly, addition of the isolated DCB/HUS domain to the DCB/HUS-GEF (Sec7ΔC) reaction inhibits GEF activity. This suggests that the GEF domain could interact with the DCB/HUS domain in a productive way on the membrane, and additional DCB/HUS competes for this interaction. This Sec7ΔC construct has lower activity than WT Sec7, however the addition of the HDS1 domain (Sec7ΔC+HDS1) results in a construct with higher activity than WT Sec7. Strangely, the Sec7ΔC+HDS1 construct was less active than Sec7ΔC in solution, and so the HDS1 domain was reported to be autoinhibitory in solution and activating on the membrane (Richardson and Fromme 2012). There are two important considerations in interpreting these results. First, the Sec7ΔC construct results in a dimeric construct, likely to be an artifact of truncation (Richardson et al. 2016). In the full Sec7 structure presented in chapter 2, a horseshoe shaped loop at the end of the DCB/HUS domain covers this dimerization interface. Second, the full Sec7 structure also shows the linker between the

GEF and HDS1 domains stabilizes the active conformation, similar to Gea2 (Muccini, Gustafson, and Fromme 2022). In light of these results, it seems probable that the Sec7ΔC is unable to effectively adopt the active conformation, and this might be required for the Arf1-Sec7 interaction or contribute some interaction with the membrane that increases avidity. The latter could also explain the apparent stimulation of the HDS1 domain on membranes but not in solution.

The HUS domain also contains a highly conserved region known as the HUS box, a characteristic N(Θ)DC(D/N) found in all BFA sensitive large Arf GEFs (Figure 1.8 B). The HUS box was originally described as an important element for folding/stability, but this supposition was based on a stabilizing effect in truncated Sec7 constructs (Mansour et al. 1999). It has been shown to be critical for both Gea2 and Sec7 function *in vivo*, single residue substitutions are not tolerated but do not affect protein expression (S.-K. Park, Hartnell, and Jackson 2005). The structure of full length Gea2 revealed that the HUS-box interacts with the linker regions between the GEF and HUS domains (Muccini, Gustafson, and Fromme 2022). This linker is not included in the HUS domain delineated by Mouratou, and therefore constructs lacking the HUS domain may also lack a crucial element. Mutations in the HUS Box also disrupt stimulation by Ypt31, and membrane binding (Halaby and Fromme 2018).

The HDS4 domain initially appeared to be the primary autoinhibitory domain. Removal of the HDS4 domain, or HDS2-4 domain results in constructs with similar activation rates (McDonold and Fromme 2014). This however also should be interpreted with caution, as the HDS1-HDS2 linker contains an amphipathic helix important for membrane binding of Gea2 (Muccini, Gustafson, and Fromme 2022), which was included in the truncation of HDS2-4. Four GTPases have been demonstrated to recruit and/or stimulate GEF activity of Sec7, Arl1, Arf1, Ypt1, and Ypt31/32. Removal of the HDS4 domain does not disrupt stimulation by Ypt1 or Ypt31/32, however the HDS2-4 domains are required, suggesting Ypt1 and Ypt31/32 bind to the HDS2/3 domains. Importantly, at the restrictive temperature, the *ypt1-3* allele disrupts Sec7 localization, while the GEF defective *sec7-4* is suppressed by overexpression of Ypt31/32, suggesting the Rabs are primary regulators of localization and activation respectively (McDonold and Fromme 2014). Based on the Sec7ΔC/Sec7ΔC+HDS1 results discussed earlier, Arf1 appears to require the HDS1 domain (Richardson and Fromme 2012). The mammalian homologs Big1/Big2 are known to bind Arl1 via the DCB domain (R. Wang et al. 2016; Galindo et al. 2016). Curiously, this interaction is essential for localization of Big1/Big2 in mammals, but does not seem essential in yeast (Christis and Munro 2012; McDonold and Fromme 2014).

The transition between soluble Arf1:GDP and membrane bound Arf1:GTP remains a standing question of the Arf GEFs. The Sec7 domain isolated from ARNO was found to require membranes even to perform a futile [^{3}H]GDP/GDP exchange, suggesting the membrane is involved even if Arf1 does not transition to a membrane bound state (Béraud-Dufour et al. 1999). Arf1 has also been observed binding to membranes via the myristoyl group before activation (B. Antonny et al. 1997), and Arf1 without the myristoylation is a less suitable GEF substrate (Franco et al. 1995). However, Gea2 readily associates with Arf1 in solution (Muccini, Gustafson, and Fromme 2022), and BFA stabilizes the Arf1:GEF complex independent of the N-terminal helix (Renault, Guibert, and Cherfils 2003; Mossessova, Corpina, and Goldberg 2003), suggesting the GEF-Arf1 interaction occurs absent the membrane. Further studies with Arf1 and its GEFs on membranes are required to delineate the path between soluble and membrane bound Arf1.

1.9 – The Sys1-Arl3-Arl1 cascade

The redundant nature of the secretory system presages some fundamental machinery may not be essential for viability. For example, cargos that erroneously escape the Golgi can be recycled via PM to endosome pathways (Valdivia et al. 2002; K. Liu et al. 2008; Casler et al. 2022). Of the GTPases in yeast, Arl1 and Arl3 are relatively poorly studied. While not essential in yeast, they are essential in metazoan development (Torres, Rosa-Ferreira, and Munro 2014; Mueller et al. 2002).

Arl1 has a clear role in recruiting Sec7 and the golgin Imh1 to the medial Golgi, in a pathway potentially overlapping with Ypt6. Two different tethering factors, Imh1 and GARP, have been clearly demonstrated for Arl1 and Ypt6 respectively. Neither Arl1 or Ypt6 are essential, however *arl1*Δ*ypt6*Δ is not viable. It is possible that Imh1 and GARP coordinate the same pathway, potentially forming a golgin-MTC pair. A GEF for Arl1 has not been clearly demonstrated. A Sec7 domain containing GEF Syt1 has been suggested (K.-Y. Chen et al. 2010), however Syt1 is not conserved in metazoans, and this was not reproducible (see Chapter 3). Additionally, Arl3 is required for Arl1 localization, but it is not clear why this is the case. As other GEF cascades have been demonstrated, it is assumed Arl3 recruits the GEF for Arl1, but absent a definitive Arl1 GEF this is unclear. No GEF for Arl3 has been demonstrated, however a small transmembrane protein Sys1 is required for Arl3 localization in vivo. No other functions for Arl3 have been described.

CHAPTER 2 - Structure of Sec7 reveals autoinhibitory GEF-HDS2 and activating GEF-DCB/HUS intramolecular interactions

---

2.1 - Abstract

The late stages of Golgi maturation involve a series of sequential trafficking events in which cargo-laden vesicles are produced and targeted to multiple distinct subcellular destinations. Each of these vesicle biogenesis events requires activation of an Arf GTPase by the Sec7/BIG guanine nucleotide exchange factor (GEF). Sec7 localization and activity is known to be regulated by autoinhibition, recruitment and stimulation by upstream GTPases, and positive feedback. Although these mechanisms have been characterized biochemically, we lack a clear picture of how GEF localization and activity is modulated by integrating these upstream signals. Here we report the cryoEM structure of full-length Sec7 in its autoinhibited form, revealing the architecture of its multiple regulatory domains. We use functional experiments to determine the basis for autoinhibition and use structural predictions to produce a model for an active conformation of the GEF that is also supported empirically. This study therefore elucidates the conformational changes that Sec7 undergoes to become active on the organelle membrane surface.

2.2 – Introduction

The Golgi apparatus is the central organelle of the secretory pathway in eukaryotic cells, connecting synthesis of proteins and lipids at the endoplasmic reticulum (ER) with the plasma membrane (PM) and endolysosomal system. Approximately 30% of the eukaryotic proteome requires the Golgi for post translational modifications, sorting, and transport to their functional destinations. Vectoral flow of cargo through the Golgi requires timely and accurate packaging into transport vesicles and modification of the membrane lipid constituents. Numerous pathway- specific factors are recruited to and activated on the Golgi membrane at a precise time and location. The final stage of the Golgi, the trans-Golgi network (TGN), connects the Golgi to the PM and endolysosomal system. Significant changes to the membrane itself occur at the TGN, and multiple trafficking pathways originate from this dynamic compartment.

Membrane trafficking is regulated by Ras-related small GTPases of the Arf and Rab families (Barr 2009; Donaldson and Jackson 2011). Transitioning between soluble GDP-bound and membrane-anchored GTP-bound conformations acts as a molecular switch to recruit and activate effectors on organelle and vesicle membranes. The intrinsic conversion between these two states is essentially negligible, occurring less frequently than once every 10^5 seconds (Goody and Hofmann-Goody 2002; Kahn and Gilman 1986). GEFs facilitate GDP to GTP exchange by displacing bound nucleotide, driven by a higher cytosolic GTP concentration, and ‘GTPase activating proteins’ (GAPs) inactivate GTPases by stimulating GTP hydrolysis to convert bound GTP to GDP. Coordination between GEFs, GAPs, GTPases, and effectors at the Golgi gives rise to an ordered progression through a series of trafficking pathways referred to as Golgi maturation (Pantazopoulou and Glick 2019).

At the Golgi, Arf1 and its close paralogs (Arf1-5 in mammals, Arf1-2 in yeast) are essential for vesicle biogenesis by recruiting cargo adaptors, membrane modifying enzymes, and coat proteins. Arf GTPases can also directly induce curvature by inserting a bulky helix into the outer leaflet of the membrane (Fig. 1a) (Donaldson and Jackson 2011; Gillingham and Munro 2007; Adarska, Wong-Dilworth, and Bottanelli 2021). In the budding yeast (Saccharomyces cerevisiae) model system, three “large” Arf-GEFs activate Arf1 at the early, medial, and late Golgi/TGN: Gea1, Gea2, and Sec7 respectively (Gustafson and Fromme 2017; Franzusoff et al. 1991; Spang et al. 2001). In mammals, GBF1 is homologous to Gea1/2 and the BIG/ARFGEF proteins are homologous to Sec7 (Claude et al. 1999; Togawa et al. 1999).

The Sec7 Arf-GEF is regulated by autoinhibition and a network of GTPase interactions including positive feedback (Richardson, McDonold, and Fromme 2012; McDonold and Fromme 2014). Sec7 possesses six identified regulatory domains (Mouratou et al. 2005) that mediate its localization and regulation. Although these domains are conserved across Sec7 homologs in all eukaryotes, they are not found in other proteins and therefore it has not been possible to infer their function from homology alone. Previous work from our laboratory reported the Sec7 HDS1-4 domains are autoinhibitory, the DCB-HUS domains are important for activating Arf1 on the membrane surface, the HUS domain is important for allosteric stimulation of Sec7, and that the HDS1 domain is important for positive feedback (Mouratou et al. 2005; Richardson, McDonold, and Fromme 2012; McDonold and Fromme 2014; Richardson et al. 2016; Halaby and Fromme 2018). However, how these different regulatory behaviors are integrated by Sec7 is not known.

To understand how the regulatory domains of Sec7 cooperate to regulate its localization and activity, we report the structure of full length Sec7 determined by cryogenic electron microscopy (cryoEM) to 3.7 Å average resolution. The structure reveals the overall architecture of the regulatory and GEF domains within the context of the Sec7 homodimer. The arrangement of the Sec7 dimer is strikingly different than that of the related Arf-GEF Gea2. Whereas Gea2 dimerizes via its N-terminal DCB-HUS domains, Sec7 dimerizes via its C-terminal HDS4 domain. The structure also indicates that the primary mechanism for autoinhibition is an interaction between the GEF and HDS2 domains that occludes the catalytic surface of the GEF domain. An alternative conformation of the GEF domain is predicted by AlphaFold (Jumper et al. 2021) and resembles the positioning of the GEF domain in Gea2 (Muccini, Gustafson, and Fromme 2022). As Gea2 is not autoinhibited, we infer that the predicted structural conformation represents an active state of Sec7. We also performed extensive *in vitro* and *in vivo* functional experiments to test and validate both the cryoEM and predicted structural models. Our findings enable us to provide a structural explanation for how the Sec7 regulatory domains function in both the autoinhibited and active states.

2.3 – Results: CryoEM structure of Sec7 from T. terrestris

To determine how Arf activation at the TGN is regulated by the Arf-GEF regulatory domains (Fig. 2.1a,b), we sought the structure of an intact Sec7/BIG homolog. We therefore purified full-length Sec7 from the thermophilic yeast Thielavia terrestris using overexpression in Pichia pastoris, verified that it possessed *in vitro* GEF activity towards Arf1, and determined its structure by single-particle cryoEM (Fig. 2.1c, Table 2.1, Figure 2.2).

Exploiting the C2 symmetry of the Sec7 dimer allowed us to generate a 5.3 Å resolution map (Fig. 2.2d), however significant flexibility across the thin ~40 nm molecule limited the resolution, especially at the distal ends of the dimer. Heterogeneous reconstruction analysis by cryoDRGN (Kinman et al. 2023) suggested each end of the molecule can flex as much as 28° relative to the central dimerization domain (Fig. 2.3). This flexibility is apparently continuous, as intermediate states were poorly reconstructed and standard real space 3D classification was unable to isolate discrete states.

Subsequent symmetry expansion, particle subtraction, and 3D classification of individual monomers enabled us to generate a significantly improved reconstruction of the monomer with an average 0.143 FSC of 3.7Å (Fig. 2.4). Focused refinements of finer sub-volumes did not further improve resolution, so the monomer map was used to build and refine an atomic model of a Sec7 monomer, which was then combined and refined into the dimer map to produce the full model of a Sec7 homodimer (Fig 2.1c, Fig. 2.3).

The architecture of the Sec7 dimer is a long continuous alpha-solenoid consisting of tandem pairs of α-helices forming an elongated ladder structure we refer to as the backbone (Fig 2.1c). The six previously described regulatory domains of Sec7 (DCB, HUS, HDS1-4) comprise the backbone. The catalytic GEF domain, which lies in between the HUS and HDS1 domains in the primary sequence, is extruded from the backbone such that the HUS and HDS1 domains directly interact with each other. The GEF domain is bound to the surface of the HDS2 domain. Homo-dimerization is mediated by the HDS4 domain, consistent with conclusions from previous studies (Mariño-Ramírez and Hu 2002; Richardson et al. 2016) but distinct from the dimerization mode of the related Arf-GEF Gea2 (Muccini, Gustafson, and Fromme 2022; Grebe et al. 2000). As a result of dimerization Sec7 adopts a flattened ‘W’ architecture when viewed from the side, and a sinuous partial corkscrew when viewed from above (Fig 2.1c,d).

Table 2.1 – CryoEM data collection, processing, and model statistics.

	Monomer	Dimer (composite map)
Nominal magnification	63,000	
Voltage (kV)	200	
Total dose ($e^-/Å^2$)	55.6	
Defocus range (μm)	-0.6 to – 2.2	
Pixel size (Å)	1.29	
Symmetry imposed	N/A	C2
Particle images	196,888	296,177
Map resolution, 0.143-FSC (Å)	3.7	5.3
Map sharpening B factor	-102.268	-93.48
No. of atoms	9,921	19,842
No. of protein residues	1237	2474
B factor, protein	63.08	78.98
RMSD, bond length (Å)	0.002	0.003
RMSD, bond angles (°)	0.541	0.610
MolProbity score	1.56	1.78
Clashscore	7.55	11.23
Poor rotamers (%)	0.9	0.32
Ramachandran plot		
Favored (%)	97.21	96.62
Allowed (%)	2.7	3.3
Disallowed (%)	0.08	0.08

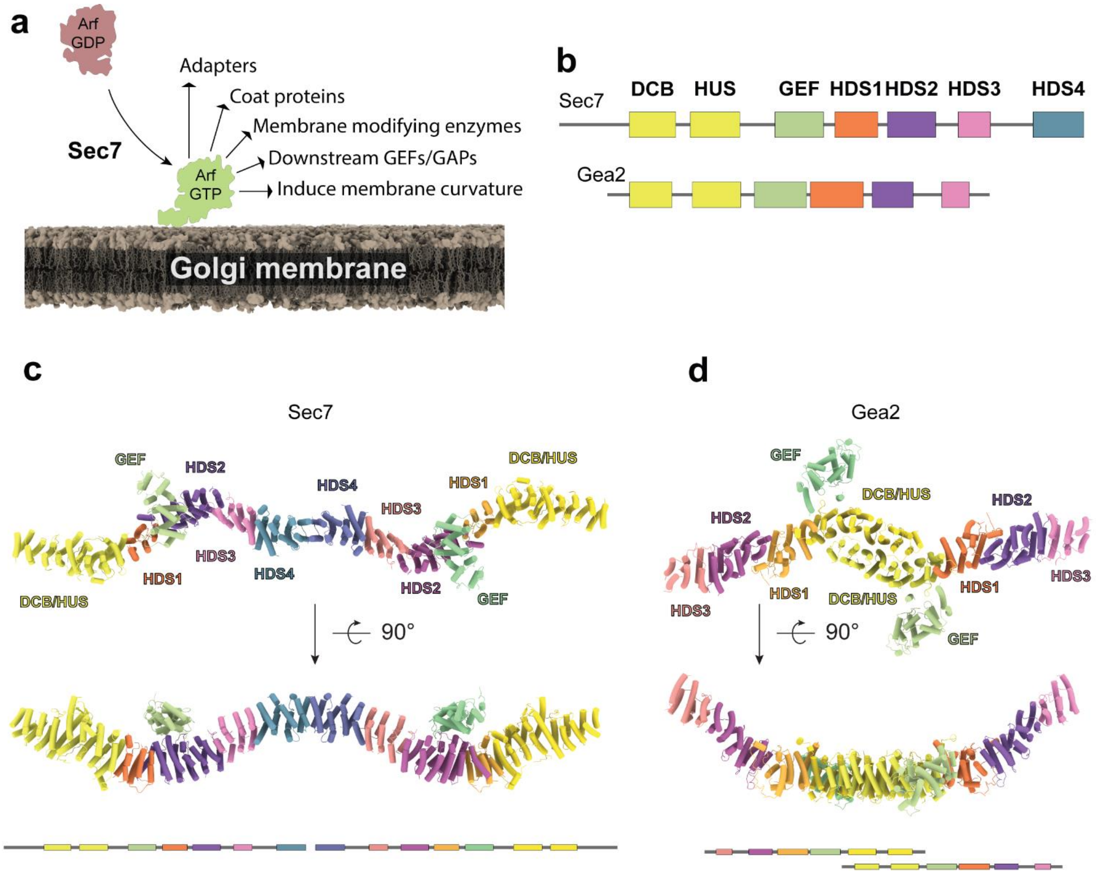


Figure 2.1 – CryoEM structure of Sec7. a, Active Arf1 on the TGN membrane facilitates many vesicle trafficking functions. **b**, Domain architecture of Sec7 and Gea2, colored as in c and d. **c**, Model for the full Sec7 dimer. **d**, Model of the Gea2 “open” dimer. Note the difference in dimerization via the HDS4 domain (Sec7, c) and DCB/HUS domains (Gea2, d).

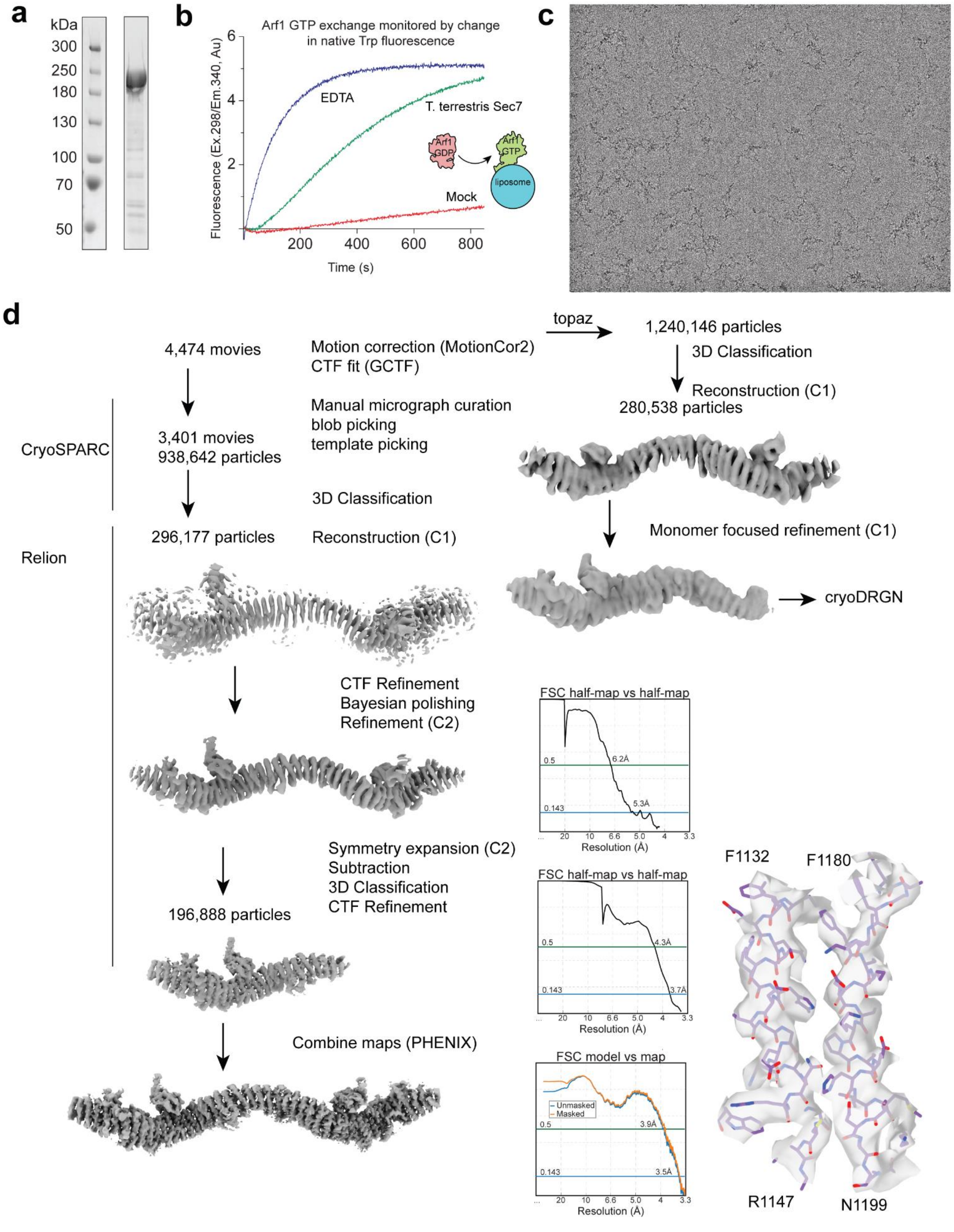


Figure 2.2 – *T. terrestris.* Sec7 purification and Cryo EM data processing workflow. Full figure legend on following page

Figure 2.2 – *T. terrestris.* Sec7 purification and Cryo EM data processing workflow. **a**, SDS-PAGE of purified *T. terrestris* Sec7 expressed in *P. pastoris*. **b**, The purified protein has GEF activity on *S. cerevisiae* Arf1. Exchange was performed on TGN liposomes with myristoylated Arf1 (see methods). Note that the observed rate and sigmoidal shape of the curve are consistent with autoinhibitory behavior previously determined for *S. cerevisiae* Sec7 (REF 2012). **c**, Representative cryoEM micrograph. **d**, CryoEM data-processing workflow. FSC curves are shown for the initial dimer, focused monomer, and combined half maps, and a region of representative density with model side chains fit shows agreement between map and model.

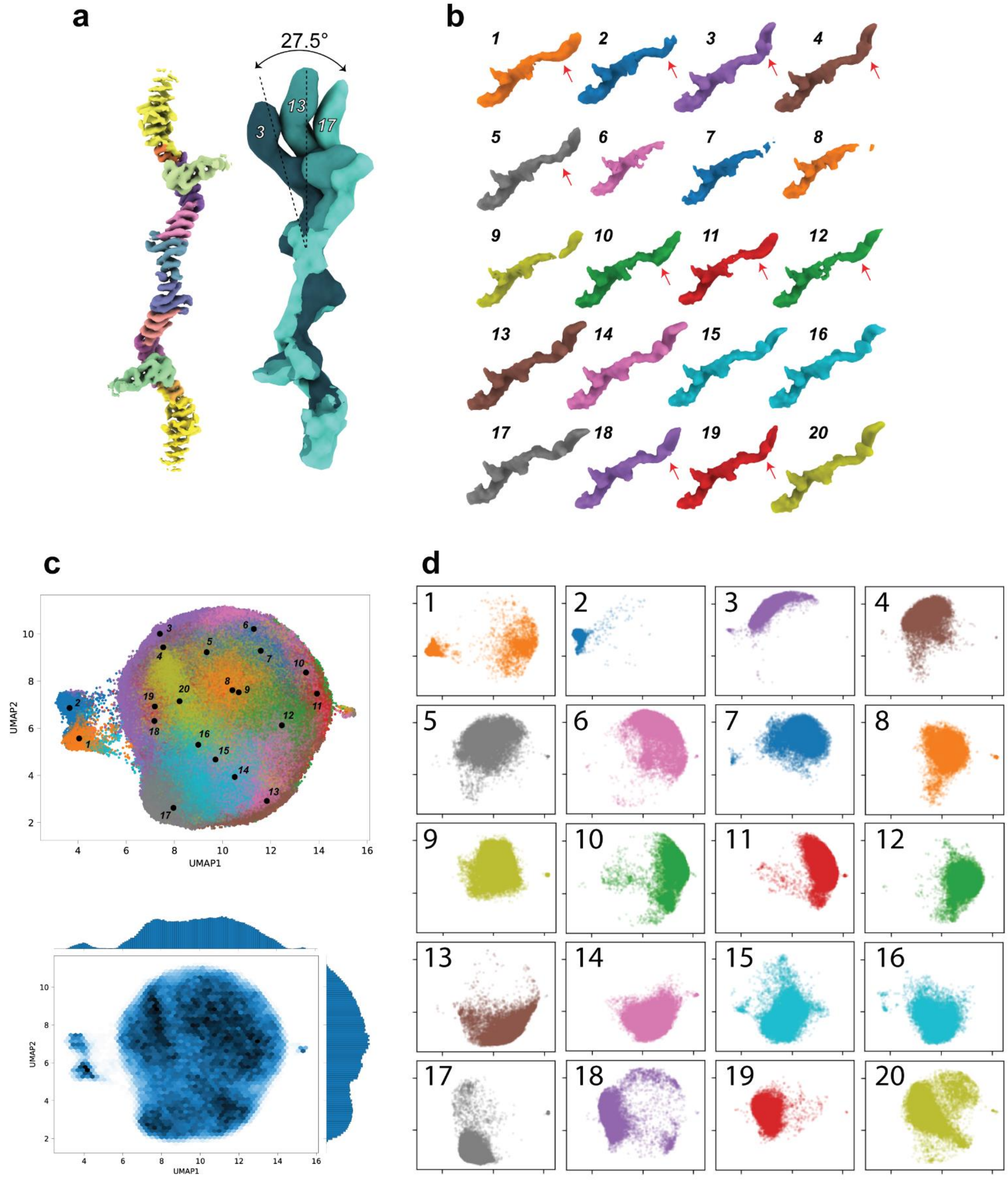


Figure 2.3 – Heterogeneous Reconstruction and UMAP analysis by cryoDRGN. Full figure legend on following page.

Figure 2.3 – Heterogeneous Reconstruction and UMAP analysis by cryoDRGN. Dimer particles were focus-refined on a monomer (without symmetry expansion or subtraction) and used to train a cryoDRGN variability model. **a**, Cryo EM map of the Sec7 dimer reconstruction at 5.3Å overall resolution (left), and the cryoDRGN generated maps with the greatest observable flexibility, showing at least 27.5° of flexibility. **b**, All maps generated by cryoDRGN. Maps 6-10 significantly lack density for a second monomer, likely due to flexibility. Red arrows indicate maps with little/no density for the GEF domain bound to HDS2. We note that this only occurs in the monomer that was not focused on during refinement, and therefore could be artifactual. **c**, Plots of UMAP analysis colored by reconstruction weight of the maps in b (top) or by density of particles (bottom). **d**, Isolated UMAP distribution plots of each map in b

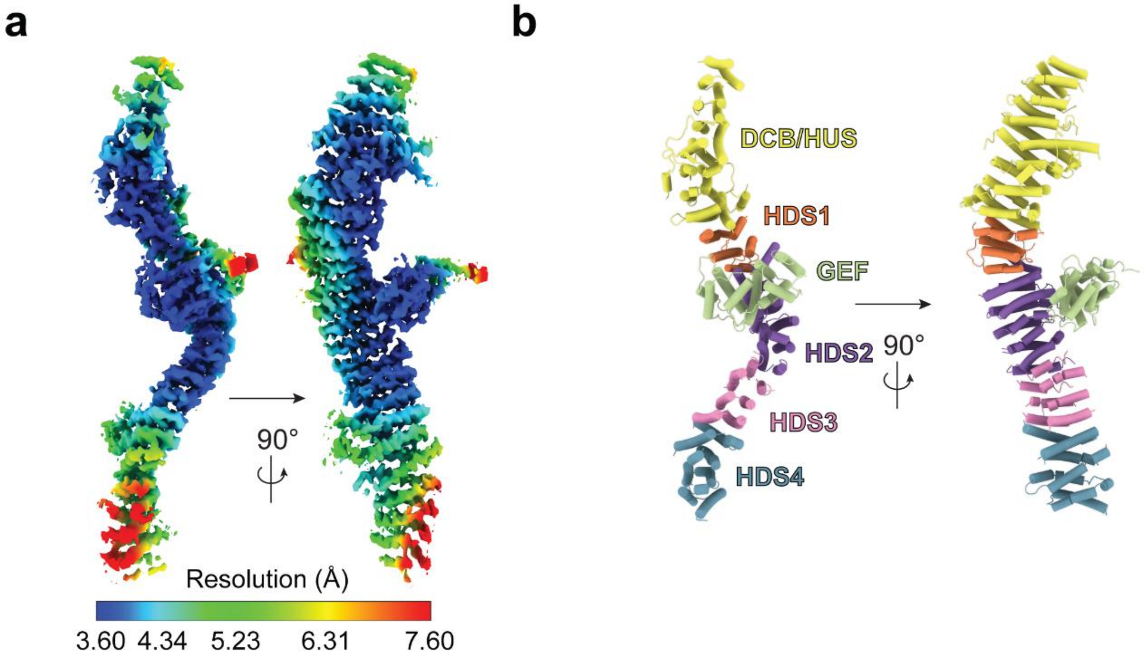


Figure 2.4 – Focused refinement of a Sec7 monomer. a, Map generated by focused refinement of an entire single monomer using symmetry expanded particles gave the highest resolution, 3.7 Å overall. Local resolution determined in Phenix was used to color the map as indicated. **b**, Model built and refined using the map in a, colored by domains as indicated.

2.4 – Results: Autoinhibition of Sec7 via the HDS2 domain

The cryoEM structure reveals the HDS2 domain interacts with the catalytic surface of the GEF domain in a manner that is mutually exclusive with Arf1 binding. When the crystal structure of nucleotide-free Arf1 bound to the Gea2 GEF domain (Goldberg 1998) is superimposed on the GEF domain in Sec7 cryoEM structure, a significant steric clash is observed between Arf1 and the HDS2 domain (Fig. 2.5a) This indicates that the observed position of the GEF domain, bound to the HDS2 domain, prevents Arf1 activation, and provides a clear mechanism for autoinhibition. We previously reported that truncations of the C terminal HDS2-HDS4 domains results in a construct with higher *in vitro* GEF activity (Gustafson and Fromme 2017; McDonold and Fromme 2014; Richardson, McDonold, and Fromme 2012), consistent with the observed HDS2-GEF domain interaction serving to mediate autoinhibition.

The arrangement of the GEF domain bound to the HDS2 domain is quite different than that observed in the cryoEM structure of Gea2 (Muccini, Gustafson, and Fromme 2022) (Fig 2.1c,d), in which the GEF domain is positioned adjacent to the DCB/HUS domain in a manner that does not interfere with Arf1 binding. Unlike Sec7, Gea1/2 are not autoinhibited (Gustafson and Fromme 2017) and are not regulated by positive feedback (Richardson, McDonold, and Fromme 2012). We compared the sequences of Sec7 and Gea2 in the region of HDS2 contacting the GEF domain the Sec7 cryoEM structure and found these sequences are well conserved across Sec7 homologs but are less well conserved in Gea homologs (Fig. 2.5b, Fig. 2.6a). On the GEF domain-binding surface of HDS2, two loops of predominantly hydrophobic conserved residues flank a weakly basic groove that are complementary to similarly well conserved regions in the GEF domain (Fig. 2.5b,c, Fig. 2.6b). We note this interaction is different from the autoinhibitory interaction of the cytohesin Grp1 in which a strongly basic patch in the PH domain C-terminus blocks the catalytic surface of the Grp1 GEF domain (DiNitto et al. 2007) (Fig 2.6c). Although there is less electrostatic potential at the Sec7 autoinhibitory interface, it appears more extensive, burying 623 Å2 compared to 555 Å2 in Grp1.

To test the importance of the HDS2-GEF domain interaction for the autoinhibitory behavior of Sec7, we designed a mutation in S. cerevisiae Sec7 to disrupt the exposed hydrophobic surface of the HDS2 domain and repel acidic residues of the GEF domain based on the equivalent T. terrestris residues. We purified this L1376D mutant S. cerevisiae Sec7 protein, along with the wild-type, and tested its ability to activate Arf1 on a membrane surface using an established *in vitro* GEF activity assay (Richardson and Fromme 2015; Richardson, McDonold, and Fromme 2012). A change in native tryptophan fluorescence of Arf1 can be used to monitor exchange kinetics of Arf family proteins (Higashijima et al. 1987; Beraud-Dufour et al. 1998; Futai et al. 2004; DiNitto et al. 2007). Using myristoylated-Arf1 and synthetic TGN liposomes (see Methods), we observed that the purified L1376D Sec7 mutant construct exhibited significantly higher GEF activity than the wild-type (Fig. 2.5d,e, Fig. 2.6d). This result supports our interpretations that the HDS2-GEF domain interface mediates autoinhibition of Sec7, and that the cryoEM structure represents the autoinhibited conformation of Sec7.

To determine the *in vivo* significance of this autoinhibitory interaction, we tested the ability of Sec7 constructs harboring mutations in L1376 to complement the loss of SEC7 gene function in S. cerevisiae (Fig. 2.5f). We determined that in a sensitized *arf1*Δ background, in which only 10% of the total Arf1/2 protein remains (Stearns et al. 1990), the L1376D mutant was inviable at 38 °C, and very sick at 30 °C, even though it was expressed comparably to wild-type (Fig 2.6d). In contrast, the L1376K and L1376S mutations did not have significant impacts on growth. Wild-type Sec7 localizes to Golgi and TGN compartments which appear as puncta distributed in the cytosol (Losev et al.

2006; Franzusoff et al. 1991; Richardson, McDonold, and Fromme 2012). We visualized the localization of the L1376D mutant Sec7 protein by fluorescence microscopy and verified the mutations did not disrupt its localization to punctate Golgi structures (Fig 2.6e). Interestingly, the morphology of the Sec7 labeled Golgi compartments in the mutant cells appeared significantly more variable in size and intensity, suggesting an impact on Golgi function. Taken together, although the L1376D mutation was tolerated in the wild-type ARF1 strain, these results indicate that disruption of the autoinhibitory HDS2-GEF domain interface is deleterious *in vivo*.

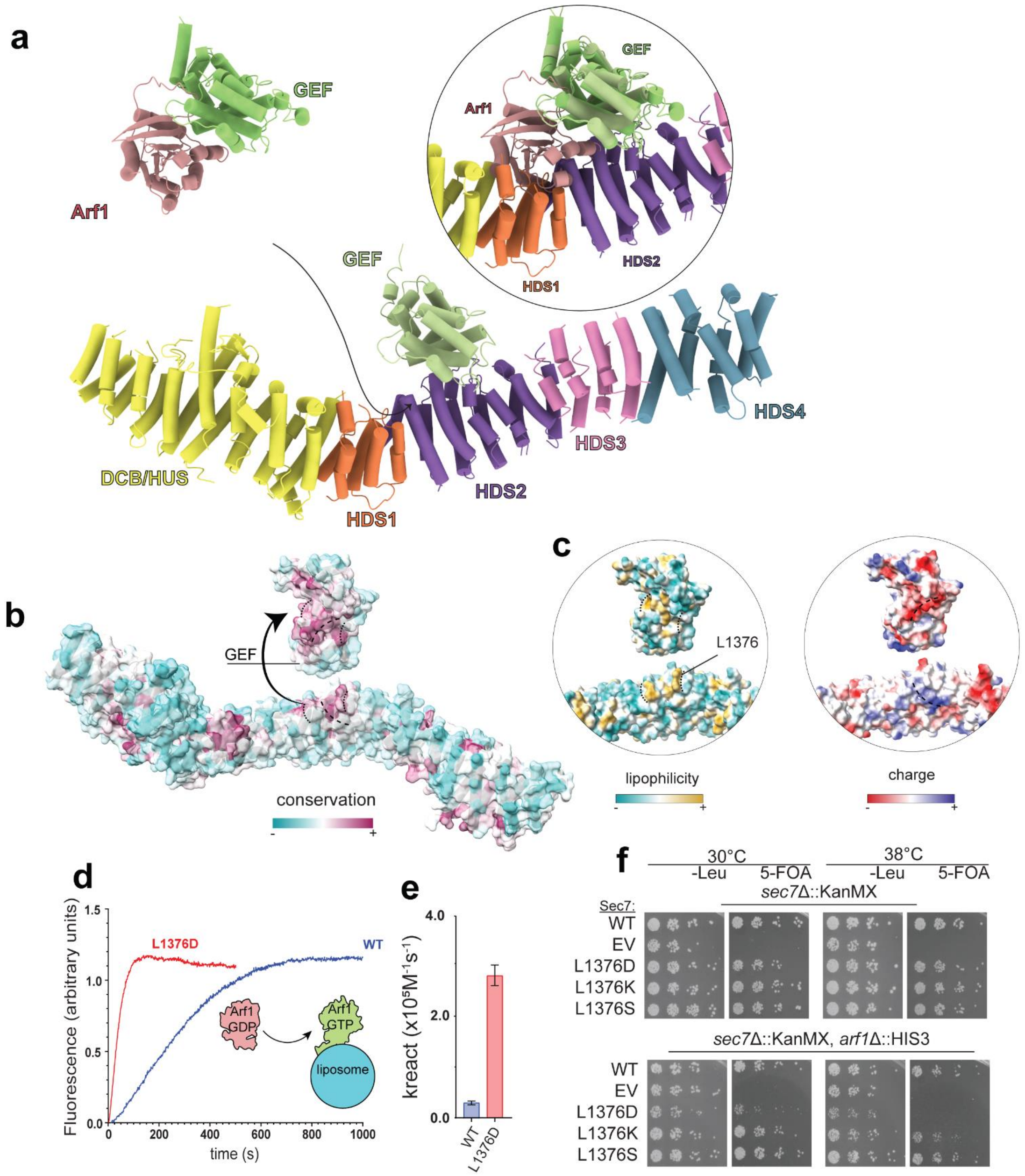


Figure 2.5 – The GEF:HDS2 interaction interface is autoinhibitory and physiologically important. a, Crystal structure of the Gea2 Sec7 domain bound to Arf1 (Goldberg 1998) docked into the full Sec7 structure (inset) shows the HDS2 domain prevents Arf1 from interacting with the GEF domain. **b**, Sec7 model surface colored by consurf analysis showing the HDS2:GEF interfaces are well conserved. **c**, Surface maps of the model as shown in b colored by lipophilicity (left) or electrostatic potential (right). The location of the mutation shown in d-f is marked according to *S. cerevisiae* numbering, *T. terrestris* equivalent residue is L1177 **d,** Representative trace of measured Trp fluorescence after addition of GTP (see methods) monitoring Arf activation. **e**, Average quantified kreact of three replicate assays as shown in d. **f**, 5-FOA plasmid shuffling assay shows hyperactive Sec7 is able to complement the *sec7*Δ null, but not the sensitized *sec7*Δ*arf1*Δ null. Substituting the hydrophobic residue with an acidic, but not basic or hydrophilic, residue causes the disruption.

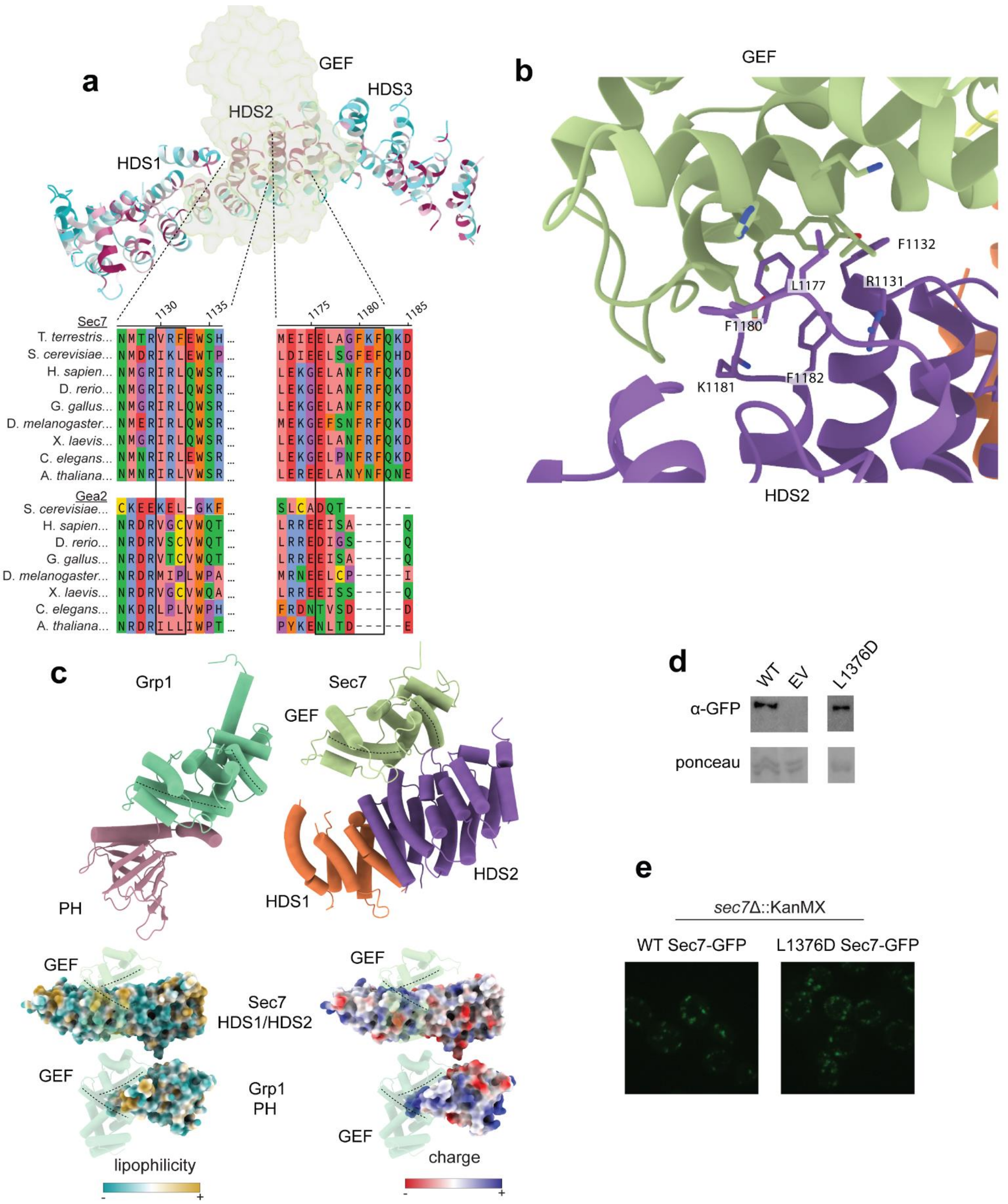


Figure 2.6 – Detailed view of HDS2 interface. a, Multiple sequence alignment of the HDS2 interface in Sec7 and Gea2 homologs. Boxed regions indicate the surface exposed loops that contact the GEF domain. **b**, Key residues of the HDS2 interface are shown. The basic patch largely arises from a single arginine (R1131) residue, nearby aromatic residues (F1180/F1182) could transmit the charge through a t-shaped charge-pi interaction. Residues F1132, L1177, and F1180 primarily form the hydrophobic surface. (**Continued…**)

Figure 2.6 – Detailed view of HDS2 interface. c, Comparison with the autoinhibited Grp1 crystal structure. Grp1 autoinhibition involves a strongly basic patch in the PH domain that interacts with the GEF domain in a similar orientation. **d**, Western blot for α-GFP-Sec7, showing the L1376D construct is expressed similarly to WT after shuffling. **b**, Microscopy of GFP-Sec7 constructs after shuffleing in *sec7*Δ::KanMX strain background. Hyperactivity does not appear to dramatically alter localization.

2.5 – Results: AlphaFold predicts an active conformation of Sec7

The ability to accurately predict protein structures with machine learning tools such as AlphaFold is revolutionizing structural biology (Jumper et al. 2021). We compared the AlphaFold prediction of S. cerevisiae Sec7 to our experimental cryoEM model, and found a striking difference in the position of the GEF domain. Rather than bound to the HDS2 domain as we observed in the experimental structure, AlphaFold predicted the GEF domain to occupy a position that closely resembles the “open” conformation observed in the cryoEM structure of Gea2 (Muccini, Gustafson, and Fromme 2022) (Fig. 2.7a). Importantly, in the predicted Sec7 conformation the GEF domain is available to bind Arf1 without steric hindrance (Fig. 2.8a). Furthermore, linker regions between the GEF and backbone (HUS and HDS1) domains make extensive contact with the “HUS box” (Fig. 2.8b), a conserved motif near the C-terminal end of the HUS domain implicated in Sec7 function (Mouratou et al. 2005; Halaby and Fromme 2018). We therefore considered the prediction to represent a plausible model for the active conformation of the Sec7 GEF domain.

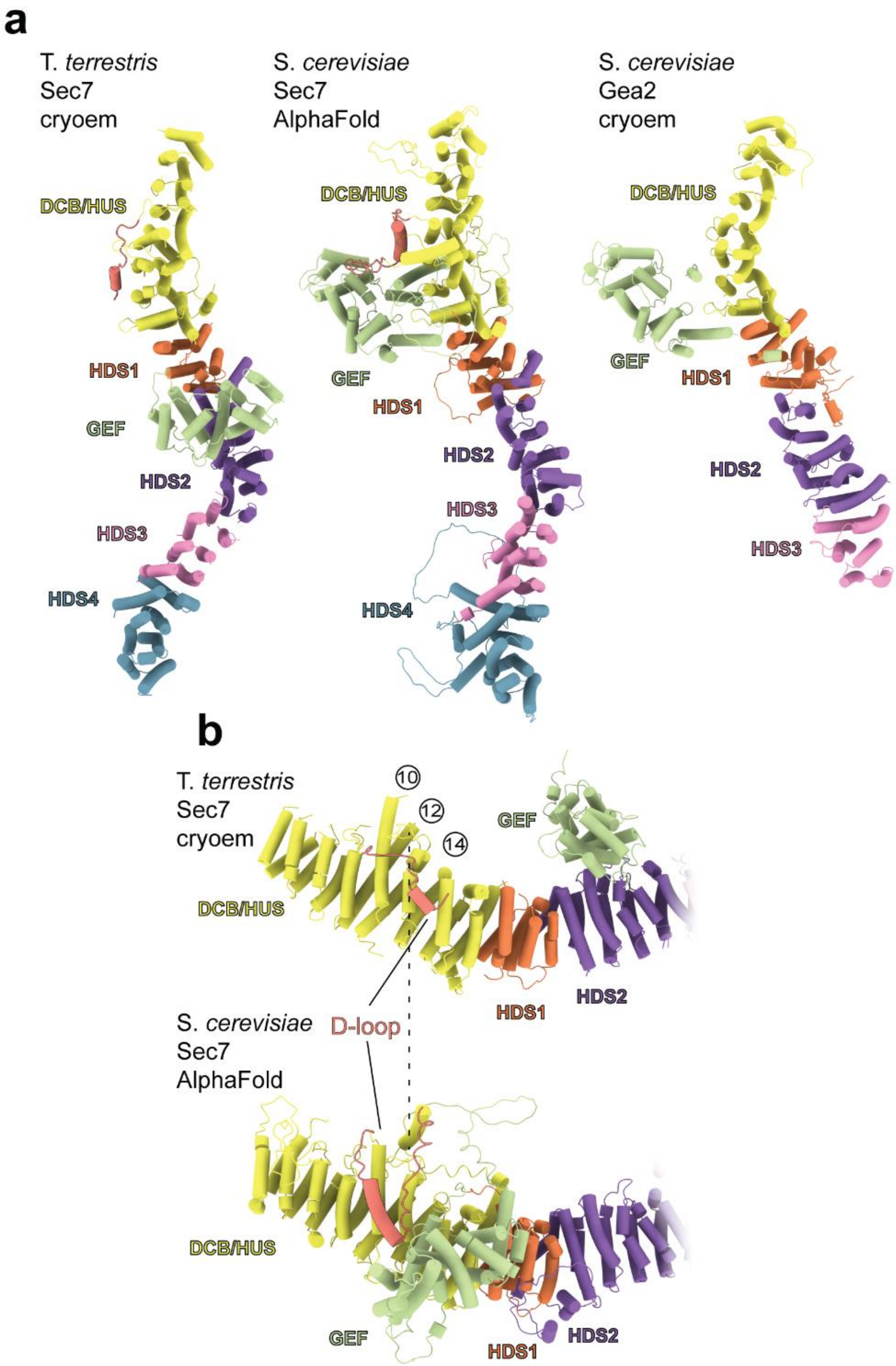


Figure 2.7 – AlphaFold predicts the active conformation of Sec7. a, Side by side comparison of cryoEM T. terrestris Sec7, AlphaFold predicted S. cerevisiae Sec7, and the cryoEM structure of S. cerevisiae Gea2 in the “open” conformation (PDB: 7UTH). **b**, The “D-loop” (colored salmon) binds to a surface between helices 12 and 14 in the cryoEM structure, but shifts to helices 10 and 12 in the AlphaFold structure, displaced by the GEF domain.

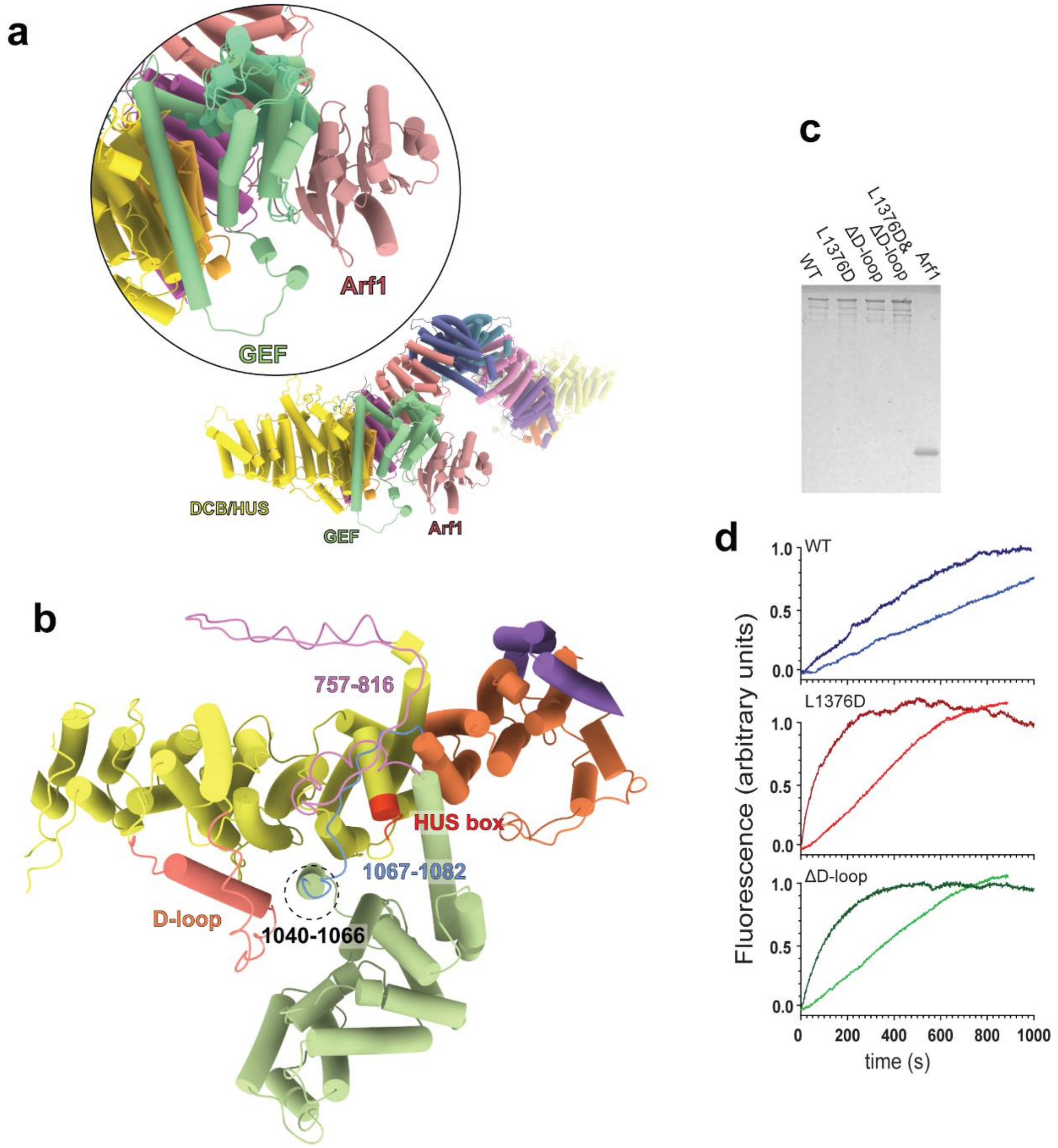


Figure 2.8 – Active Sec7, detailed view a, AlphaFold prediction with Arf1 positioned on the GEF domain as in the Gea2-Arf1 crystal structure (Goldberg 1998) shows the catalytic surface of the GEF domain is accessible in this conformation. Inset shows the crystal structure of the Gea2 GEF domain bound to Arf1 docked into the GEF domain of the AlphaFold structure. **b**, Positions of the GEF linkers (757-816 and 1067-1082 in S. cerevisiae) surround the HUS box in the AlphaFold prediction, and physically connect the GEF domain to the activating surface of the DCB/HUS domains. **c**, SDS-PAGE of purified components used in Fig. 2d,e and Fig. 4a-d. **f**, Representative trace of GEF assay shown in Fig 6c.

In this conformation the GEF domain has displaced a partially ordered loop within the DCB-HUS domains that we refer to as the ‘D-loop’ (Fig. 2.7b, Fig. 2.8b). In the autoinhibited cryoEM structure, the D-loop interacts with the surface of the DCB-HUS domains on helices 12 and 14, but appears to shift position in the predicted active state to instead interact with the surfaces of helices 10 and 12. The position of the D-loop on the surface of the HUS domain in the autoinhibited Sec7 structure appears mutually exclusive with the predicted location of the GEF domain in the predicted active structure. We therefore hypothesized that the D-loop may contribute to autoinhibition by competing with the GEF domain for binding to this surface on the DCB-HUS domains.

To test this hypothesis and to characterize the relative roles of the D-loop and the HDS2-GEF domain interface in Sec7 autoinhibition we investigated the impact of purified proteins harboring relevant mutations (Fig. 2.8c) on *in vitro* GEF activity. We observed that the truncation of the D-loop (‘ΔD-loop’, removal of residues 444-486 in S. cerevisiae Sec7) resulted in a ~8-fold increase in GEF activity towards myristoylated Arf1 on liposome membranes compared to the wild-type (Fig 2.9a,b). In comparison, the L1376D mutation described above caused a ~20-fold increase in GEF activity relative to the wild-type. A construct in which both of these mutations was combined exhibited a ~50-fold increase in GEF activity relative to the wild-type (Fig 2.9a,b). This set of observations indicates that combining disruption of the HDS2-GEF domain interface together with truncation of the D-loop has an additive effect, and therefore both the D-loop and the HDS2-GEF domain interface contribute to Sec7 autoinhibition.

To test whether the loss of autoinhibition caused by these mutations requires the involvement of the membrane surface, we also performed GEF assays in the absence of membranes, using the ΔN17-Arf1 substrate which does not require the presence of membranes for its activation (Kahn et al. 1992; Antonny et al. 1997). Under these conditions, the same three mutant Sec7 constructs displayed the same trend of hyperactivation, albeit with lower overall magnitude (Fig. 2.9c,d). This indicates that both on membranes and in solution, the D-loop and HDS2-GEF domain interactions are important for enforcing autoinhibition of Sec7. These results, which validate the importance of the D-loop in mediating autoinhibition, also indicate the predicted structure of the Sec7 active conformation can be used to make accurate functional predictions.

We next sought to determine the relative importance of the D-loop and HDS2-GEF domain autoinhibitory interactions *in vivo*. We tested mutants in cells that were otherwise wild- type as well as in a sensitized strain lacking Age2, a TGN-localized Arf-GAP (Poon et al. 2001). We reasoned that decreased inactivation of Arf1 due to loss of Age2 might exacerbate the effects of excessive Arf1 activation that may arise from loss of Sec7 autoinhibition. We observed a gradient of phenotypes when the D-loop truncation was combined with different HDS2-GEF interface mutations (Fig. 2.9e). We found that removal of the D-loop had no impact on growth, consistent with a previous study from our lab that first observed the D-loop bound to the surface of the isolated DCB-HUS crystal structure (Richardson et al. 2016). The L1331S/L1376S/F1379S mutant, which perturbs multiple hydrophobic residues at HDS2-GEF interface (but does not significantly alter the charge of the HDS2 surface), also displayed no growth phenotype. However, combining the L1331S/L1376S/F1379S and ΔD-loop mutations resulted in a mutant construct that exhibited a growth defect in the *age2*Δ background. The more severe L1376D mutant also exhibited a growth defect in the *age2*Δ background. Finally, combining this more severe L1376D mutation with the D-loop truncation resulted in a mutant construct that was inviable in both the *age2*Δ and the AGE2 backgrounds (Fig. 2.9e). Taken together, these results indicate that additive roles of both the D-loop and HDS2-GEF interfaces are important for Sec7 regulation *in vivo*.

Remarkably, the *in vitro* GEF activities exhibited by this series of mutants as described above (Fig. 2.9a-d) inversely correlated with the magnitude of their *in vivo* growth defects (Fig 2.9e). This indicates that increasing loss of Sec7 autoinhibition *in vitro* correlates with increasing deleterious effects on Sec7 function *in vivo*.

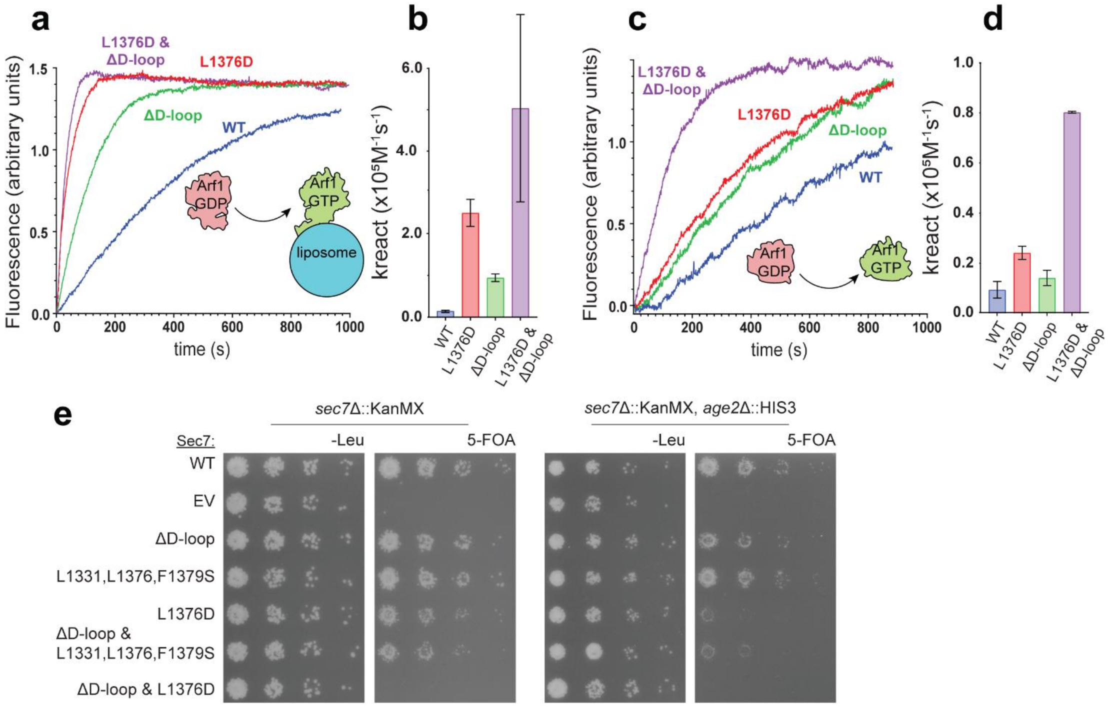


Figure 2.9 – The Sec7 D-loop and GEF-HDS2 domain interface exert additive effects on autoinhibition *in vitro* and are important *in vivo*. a-d, GEF exchange of myristoylated Arf1 on TGN like liposomes (a and b) or ΔN17Arf1 (c and d) monitored by native Trp fluorescence. 100 nM GEF was added for all reactions. The D-loop and L1376D are hyperactivating, and synergistically enhance GEF activity when combined in the presence and absence of membranes. **e**, 5-FOA plasmid shuffling assay in the indicated strains. Plates were grown at 30°C for 3 days.

2.6 – Results: Dimerization via a hydrophobic patch in the HDS4 domain

Sec7 was previously found to dimerize via the HDS4 domain (Richardson et al. 2016), and this domain was also found to be essential for Sec7 function *in vivo* and to have a potential role in autoinhibition (McDonold and Fromme 2014; Richardson, McDonold, and Fromme 2012; Richardson et al. 2016). The cryoEM structure reveals the Sec7 dimerization interface is composed of 6 hydrophobic residues on each monomer (Fig. 2.10a). With this new detailed structural information, we sought to directly characterize the importance of dimerization by disrupting the dimer interface while keeping the remainder of the HDS4 domain intact. We therefore generated a mutant in which the six hydrophobic residues at the dimer interface are mutated to serine residues (Y1975, I1979, L1982, V1986, L1998, V2001 in S. cerevisiae). The purified mutant protein was monomeric *in vitro* as determined by SEC-MALS analysis (Fig. 2.11a,b).

To test whether dimerization is important for Sec7 function *in vivo*, we tested whether the monomeric mutant could complement a *sec7*Δ mutant. We observed very slight temperature sensitivity in an otherwise wild-type background, but complete loss of viability in the sensitized *arf1*Δ background (Fig. 2.10b). In contrast, removal of the entire HDS4 domain results in inviability in a wild-type ARF1 background (Richardson et al. 2016). The differences in magnitude of these growth phenotypes suggest that there may be a role for the HDS4 domain in addition to dimerization. However we cannot rule out the possibility that loss of the HDS4 domain impacts protein folding or alternatively that the dimerization mutant does not completely disrupt dimerization *in vivo*. In either case, our results establish that dimerization is important for Sec7 function *in vivo*.

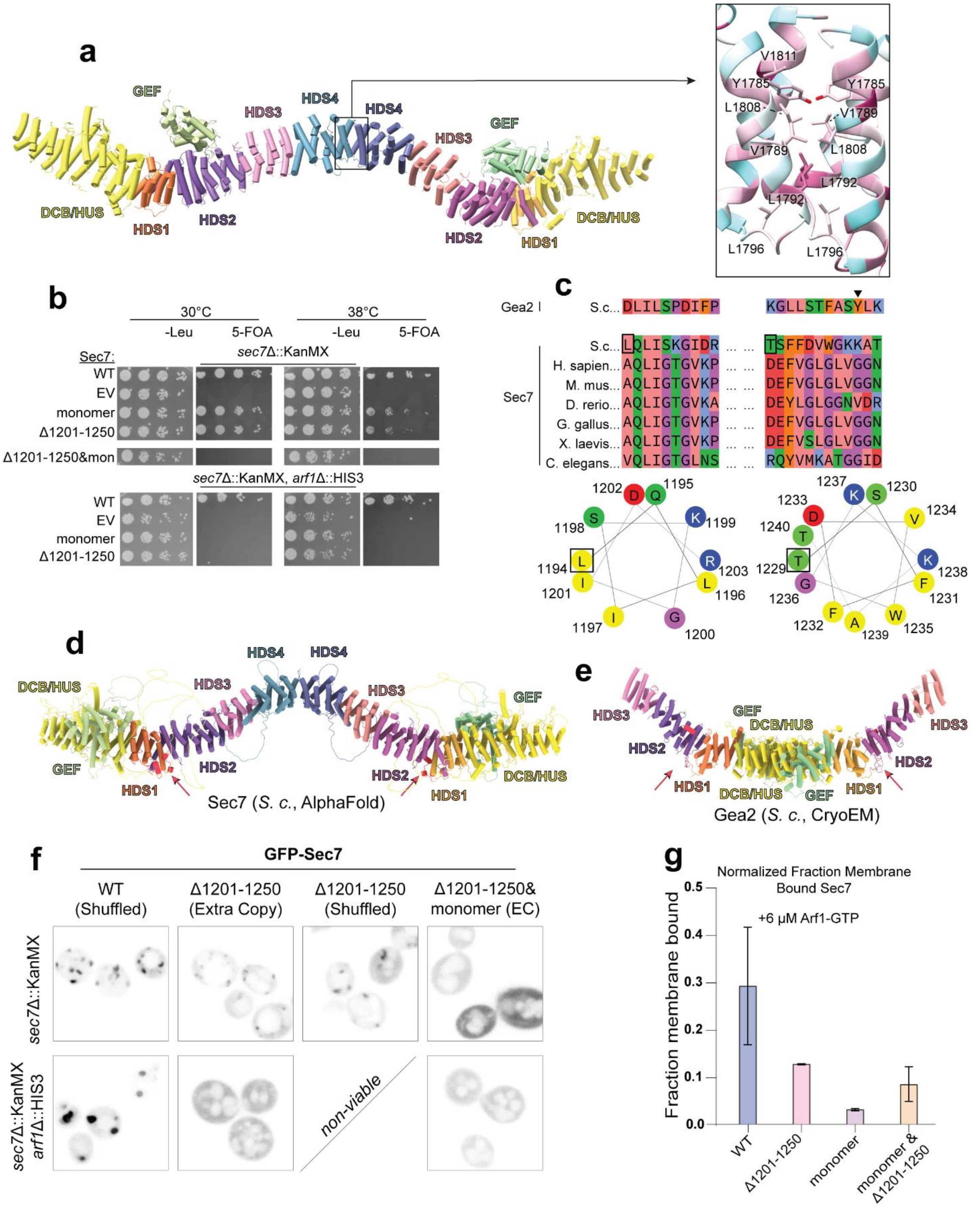
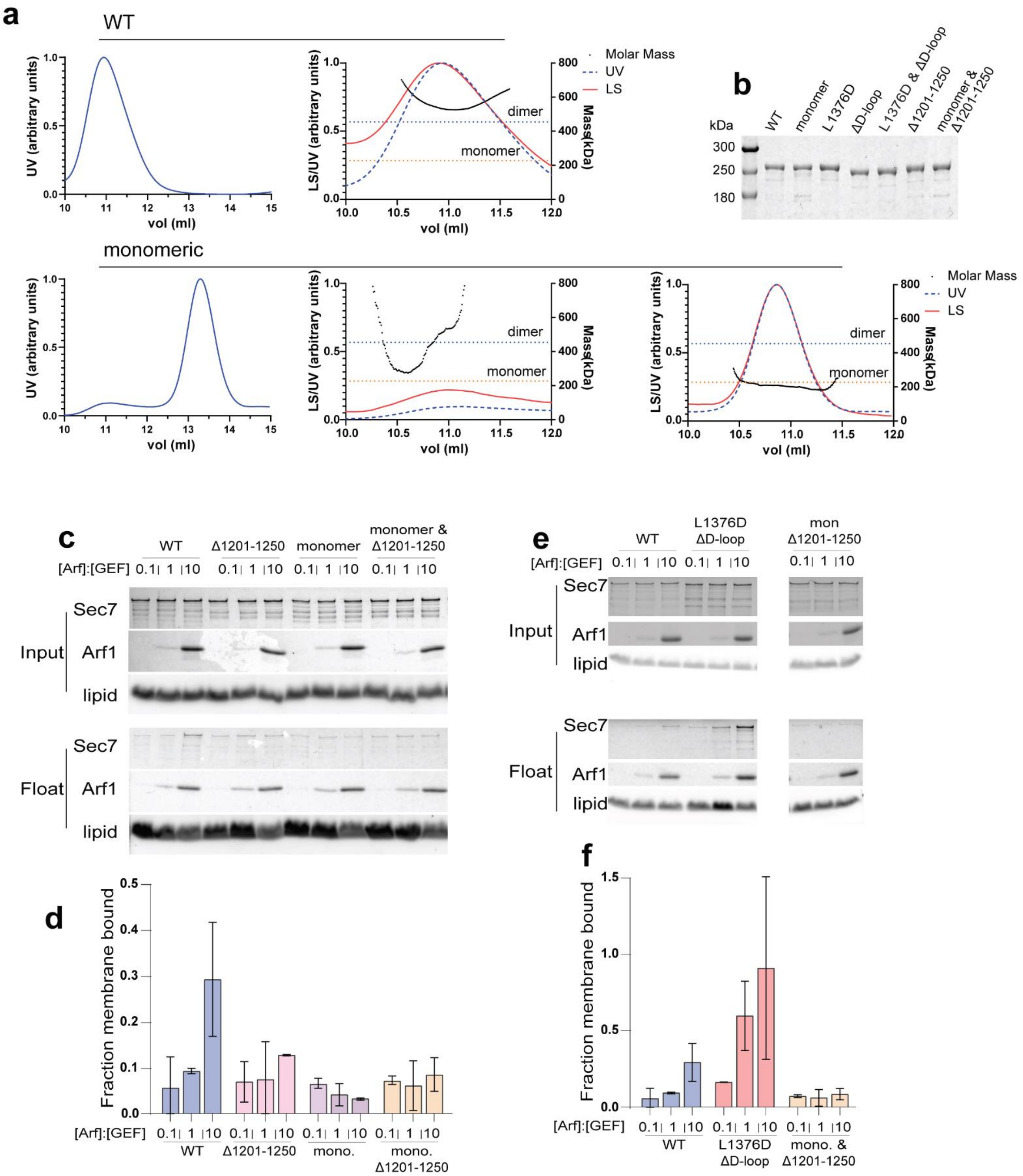


Figure 2.10 – Dimerization and an amphipathic sequence are important for Sec7 membrane binding and localization *in vivo.* Full figure legend on following page

Figure 2.10 – Dimerization and an amphipathic sequence are important for Sec7 membrane binding and localization *in vivo*. a, CryoEM Sec7 dimer model (left) with the HDS4 dimer interface showing hydrophobic residues contributing to the dimerization pocket (boxed inset). **b,** 5-FOA plasmid shuffling assay showing no growth defect in *sec7*Δ null with monomeric or Δ1201-1250 constructs but complete lack of viability in combination. Neither construct is viable in the *sec7*Δ*arf1*Δ. **c**, Multiple sequence alignment and helical wheel projections of the HDS1-2 linker regions that could form amphipathic helices. Sequences shown are for S. cerevisiae (S. c.) Gea2 or various Sec7 homologs (in species with multiple homologs Big1 is shown). We note that several homologues of Sec7 contain longer stretches of residues which could embed in the membrane in the second region, but a more conserved amphipathic helix in the first. **d**,**e**, Sec7 (left) and Gea2 (right) dimer models. Putative membrane binding region in Sec7 and corresponding residues (where present) are highlighted in red and indicated by arrows. **f**, Fluorescence microscopy of strains from c, with the indicated construct expressed as an extra copy (EC) when necessary for growth. **g**, 0.6 µM Sec7 construct was incubated with TGN liposomes pre-loaded with 6 µM Arf1 and isolated by ultracentrifugation. Shown is normalized recovery of Sec7 construct by lipid recovery and input intensity of triplicate experiments (see methods).



Extended Data Figure 2.11 – SEC/MALS of monomeric Sec7 construct and membrane binding. a, Purified monomeric Sec7 eluted ~2.5 ml later than WT Sec7 over a Superose 6 10/300 column (left), and was verified by MALS to be monomeric (right). Interestingly a small fraction of the monomeric construct was not shifted, and MALS suggests this might contain a mixture of monomer/dimer. **b**, Purified constructs used in GEF and membrane binding assays. **c**,**d** Representative SDS-PAGE gel of membrane float assay reported in Fig. 2.10 (c), and quantification of triplicate experiments (d). **e,f** Representative SDS-PAGE gel of membrane float assay similar to Fig 2.10 (e), and quantification of triplicate experiments (f).

2.7 – Results: The HDS1-HDS2 loop is important for Sec7 membrane-binding

Recruitment of Sec7 and its metazoan homologs to the TGN has been found to involve multiple protein-protein interactions, including Rab1, Rab11, Arl1, and a positive feedback interaction with Arf1 (McDonold and Fromme 2014; Richardson, McDonold, and Fromme 2012; Christis and Munro 2012). Several of the Sec7 regulatory domains have been implicated in TGN membrane binding (McDonold and Fromme 2014; Christis and Munro 2012; Galindo et al. 2016; Richardson et al. 2016)

The membrane binding mechanism for the related Arf-GEF Gea2 was recently identified as an amphipathic helix that is essential for membrane-binding and Golgi-localization of Gea2 (Fig. 2.10c-e) (Muccini, Gustafson, and Fromme 2022). This helix is located in a loop between the HDS1 and HDS2 domains, and is not ordered in the Gea2 cryoEM structure. We observed that Sec7 also possesses a disordered loop between the HDS1 and HDS2 domains, and this loop contains two stretches of residues which could each form amphipathic helices (Fig. 2.10c-e). The first lies at the C-terminal end of the HDS1 domain (residues 1194-1202 in S. cerevisiae Sec7), and the second is in the middle of the linker (residues 1229-1240 in S. cerevisiae Sec7) corresponding to the amphipathic helix in Gea2. Sec7 homologs in other species also possess one or two amphipathic sequence regions within this loop, however we note that both differ considerably from Gea2 (Fig. 2.10c). Importantly, while neither region is ordered in the Sec7 cryoEM structure, both lie near the putative membrane interacting surface in the AlphaFold prediction, while the first region is ordered in the Gea2 cryoEM structure bound to the HDS2 domain far from the membrane surface. Further, in many Sec7 homologues the second region is more hydrophobic and less amphipathic.

To test if the HDS1-HDS2 linker region is important for the function of Sec7 *in vivo*, we generated a Sec7 construct lacking residues 1201-1250 to disrupt both potential amphipathic helices. We found that this mutant was viable in an otherwise wild-type background but was inviable in the sensitized *arf1*Δ strain (Fig. 2.10b). Furthermore, combining the loss of the HDS1- HDS2 linker and loss of dimerization caused inviability in otherwise wild-type cells.

To determine whether the HDS1-HDS2 linker region is important for Golgi membrane association of Sec7 *in vivo*, we examined the localization of these mutants using live-cell imaging (Fig. 2.10f). Sec7 Δ1201-1250 localization to puncta appeared only slightly diminished, either when co-expressed with wild-type Sec7 or as the only copy of Sec7 in cells. In the absence of Arf1, TGN compartments become enlarged. Notably, although wild-type Sec7 is enriched on these swollen Golgi compartments, we found that Sec7 Δ1201-1250 becomes completely cytosolic in *arf1*Δ cells (Fig 2.10f). As we were unable to image this mutant as the sole copy of Sec7 in the absence of Arf1 due to the inviability of that genetic combination, we cannot distinguish whether the localization phenotype is due to the lower overall Arf1/2 levels or else to competition with the endogenous copy of Sec7.

As a dimer, Sec7 possesses two copies of the HDS1-HDS2 linker. We therefore tested the impact on localization of combining linker truncation with loss of dimerization. We found that the monomeric Sec7 construct lacking residues 1201-1250 is completely mislocalized to the cytoplasm in otherwise wild-type cells (Fig. 2.10f, construct visualized as an extra copy). This result suggests avidity due to dimerization also plays a role in Sec7 localization to the TGN.

Taken together, these results indicate that the HDS1-HDS2 linker contributes to, but is not essential for, Sec7 interaction with Golgi compartments *in vivo*. In the absence of this linker, it is likely that the known protein-protein interactions of Sec7 and other potential membrane- interacting regions of Sec7 are sufficient for Golgi localization.

To test whether the HDS1-HDS2 linker is directly involved in lipid-membrane binding we purified Δ1201-1250, monomeric, and Δ1201-1250/monomeric Sec7 constructs for *in vitro* biochemical characterization. To directly test the membrane-binding ability of these mutants, we performed liposome flotation experiments (Highland, Thomas, and Fromme 2023). We incubated these constructs with TGN liposomes pre-loaded with myristoylated Arf1-GTP, because Sec7 binds membranes poorly in the absence of stabilizing GTPases (Richardson, McDonold, and Fromme 2012; Christis and Munro 2012), and isolated membrane bound liposomes by sucrose gradient centrifugation (Fig. 2.10g, Fig. 2.11b-d). Remarkably, both the Δ1201-1250 and monomeric disruptions significantly disrupted membrane binding. We next measured the GEF activity of these purified proteins in the presence and absence of membranes as described above. Although the monomeric and Δ1201-1250 mutants behaved similarly to WT, the double mutant was dramatically less active on myristoylated-Arf1 in the presence of liposomes membranes (Fig. 2.12a,b). In contrast, all three constructs exhibited GEF activity similar to the wild-type when activating ΔN17-Arf1 in solution (Fig. 2.12c). Taken together these results indicate the Δ1201-1250/monomeric Sec7 double mutant is specifically impaired in its ability to activate Arf1 on the membrane surface. This result correlates well with the loss of Golgi localization of this construct.

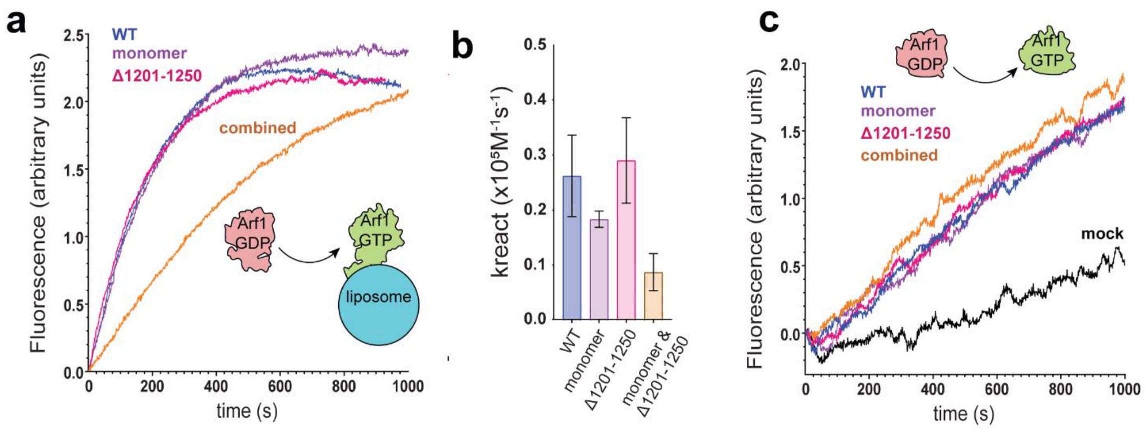


Figure 2.12 – GEF activity of dimerization and membrane binding mutants. a, Representative Trp fluorescence traces of myristoylated Arf1 exchange in the presence of synthetic TGN liposomes showing no significant difference for WT, monomeric, or Δ1201-1250 constructs, but significantly less activity when combined. **b**, Quantification of triplicate measurements in a. **c**, Representative Trp fluorescence traces of ΔN17-Arf1 exchange showing no difference in activity in solution.

2.8 – Results: The relationship between Sec7 autoinhibition and stimulation by GTPase regulators

Sec7 membrane binding and GEF activity is known to be stimulated by interaction with the active forms of multiple GTPases. A positive feedback interaction with its product Arf1-GTP stabilizes Sec7 on the membrane, interactions with Arl1 and Rab1 can also recruit Sec7 to the membrane surface, and an interaction with Rab11 appears to allosterically stimulate GEF activity (Richardson, McDonold, and Fromme 2012; McDonold and Fromme 2014; Halaby and Fromme 2018). To further characterize the relationship between Sec7 autoinhibition and stimulation by these regulators we performed additional *in vitro* GEF assays with hyperactive Sec7 constructs.

We first tested the impact of stimulation of GEF activity by Rab11 (yeast Ypt31 protein). In the presence of membrane bound Rab11-GTP, the L1376D mutant, the ΔD-loop mutant, and the combined mutant were all stimulated by Rab11-GTP to a greater extent than the wild-type (Fig. 2.13a-c). We then tested positive feedback stimulation of GEF activity by the presence of membrane-bound Arf1-GTP. We observed that not only are the basal GEF activities of the L1376D and ΔD-loop mutants significantly higher than that of the wild-type protein, their GEF activity was robustly stimulated by Arf1-GTP (Fig. 2.13d, Fig. 2.8d) Taken together, these results indicate that the effects of GEF activity stimulation by regulatory GTPases and loss of autoinhibition are additive.

Given the additive nature of autoinhibition and stimulation by membrane-bound GTPases, we wondered whether loss of autoinhibition would result in more stable membrane binding. We therefore performed a liposome flotation assay in which we titrated the amount of pre-loaded active Arf1-GTP (Fig. 2.13e,f, Fig. 2.11e,f). We found that the hyperactive L1376D/ΔD-loop double mutant was recruited to liposomes by Arf1-GTP much more robustly than was the wild-type. This result indicates the active conformation of Sec7 binds membranes more stably, suggesting that membrane-bound Sec7 is more likely to adopt the active conformation.

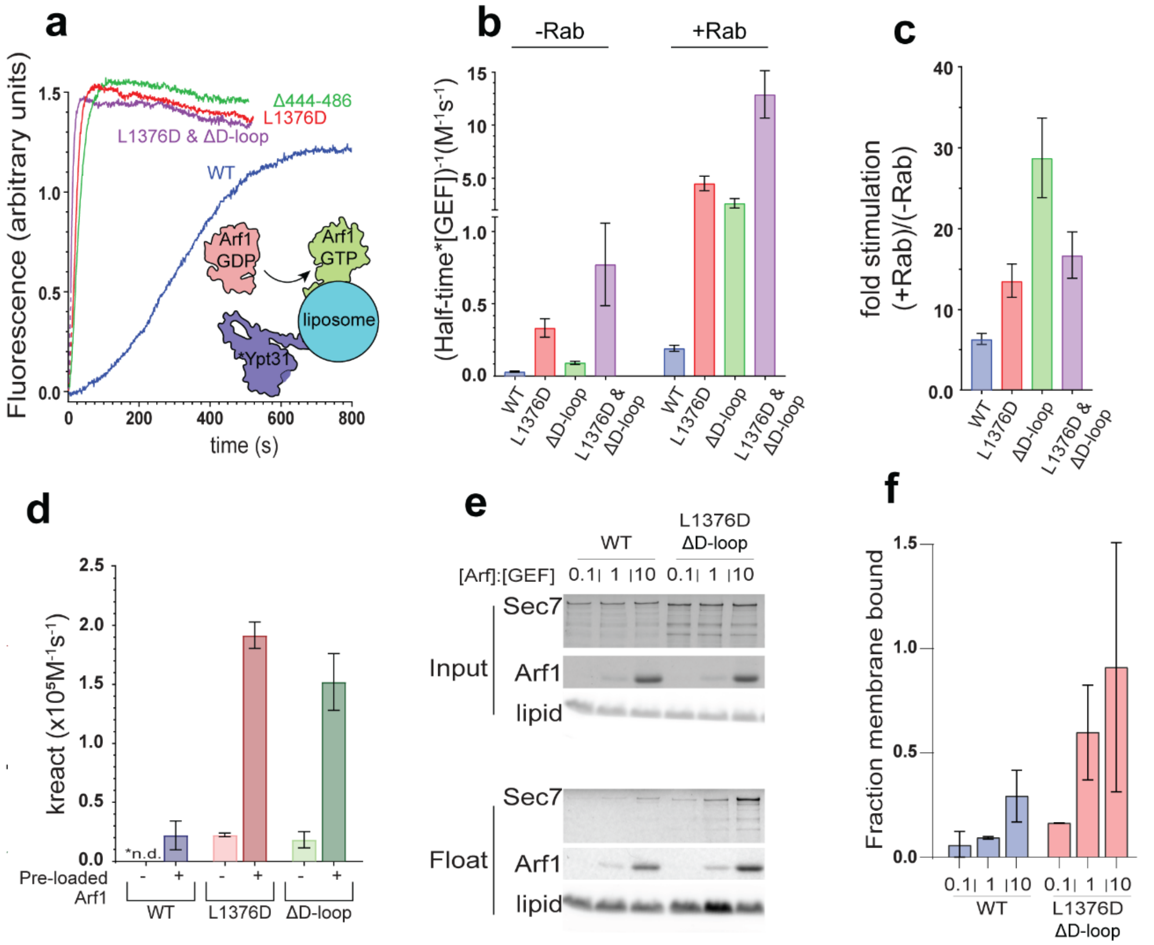


Figure 2.13 – Hyperactive Sec7 is still activated by GTPases. a, Representative traces of Arf1 exchange measured by Trp fluorescence on synthetic TGN liposomes pre-loaded with 500 nM Ypt31. **b**, Quantification of triplicate measurements shown in a, (right) shown side by side with quantification in Fig. 4a (left). **c,** Ratio of kreact constants shown in b. **d**, Quantification of exchange rates of Arf1 by the designated constructs (20 nM GEF) prior to and following exchange of 1 µM Arf1. After the first trace stabilized, a second addition of 1 µM Arf1 was added and the exchange monitored as before. **e**, SDS-PAGE gel of liposome floatation experiment with 600 nM GEF and the indicated molar ratio of pre-loaded Arf1 on 250 µM liposomes. **f**, Densitometry of band intensity in e normalized by lipid recovery and input GEF.

2.9 – Results: Model for the transition between autoinhibited and active states

To model how Sec7 would stably bind membranes, we began with the assumption that in the fully active, membrane-bound state, the DCB-HUS domains of each monomer should be close to the membrane. This assumption is based on several pieces of evidence. The crystal structure of the human BIG1 DCB domain bound to Arl1, which is anchored to the membrane by its N-terminal amphipathic helix, indicates that the “bottom” surface of the DCB-HUS domains lie proximal to the membrane (Galindo et al. 2016). Furthermore, a temperature-sensitive mutant allele (sec7-1), comprising a single amino acid substitution near this same surface of the DCB- HUS domain, is mislocalized to the cytoplasm at the restrictive temperature (McDonold and Fromme 2014). Finally, the similar positions of the GEF domain in both the AlphaFold Sec7 structural prediction and the cryoEM structures of Gea2 are such that membrane insertion of Arf1 during the nucleotide exchange reaction (Muccini, Gustafson, and Fromme 2022) would also require the same surface of the DCB-HUS domain to be located close to the membrane surface (Fig. 2.14a,b).

In our autoinhibited cryoEM structure of Sec7, the overall twisted-W shape of the dimer prevents both DCB-HUS domains from being close to the membrane (Fig. 2.14b,c). This suggests that in addition to the occlusion of the active site caused by the HDS2-GEF domain interaction, the autoinhibited conformation of Sec7 is also not compatible with stable membrane binding.

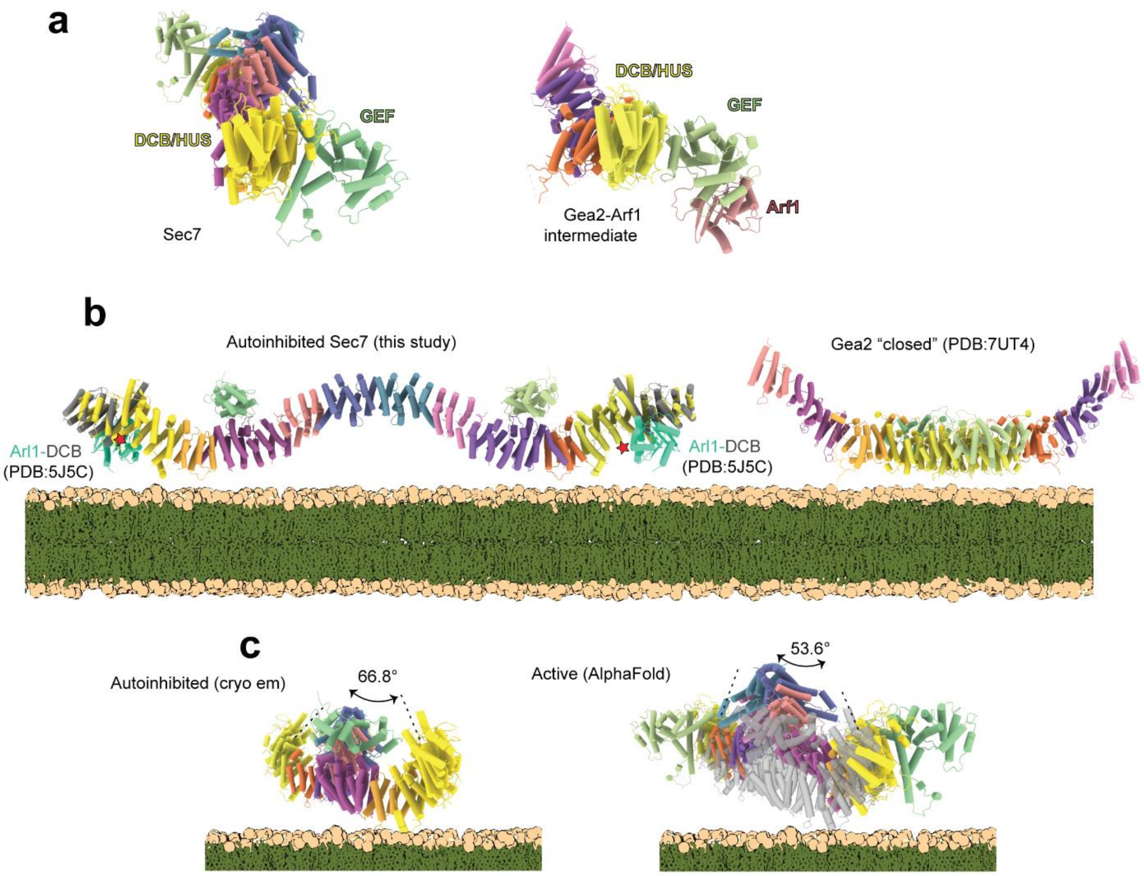


Figure 2.14 – Orientation of Sec7 on the membrane during Arf1 exchange. a, Comparison of Sec7 and Gea2-Arf1 structures, showing they likely share a similar membrane interacting surface on the 'bottom' of the DCB/HUS domains. **b**, Comparison of Sec7 and Gea2 structures on the membrane surface, with the crystal structure of BIG1 DCB:Arl1 docked. The Orientation is in agreement with the orientation in a. **c**, the "twist" imparted by dimerization induces 66.8° of rotation between the DCB domains in the autoinhibited structure (left), which is partially relieved in the AlphaFold prediction (right), with only 53.6°.

We noticed that AlphaFold predicted less bending and twisting of the Sec7 backbone than we observed in the cryoEM structure. To model a possible conformation of an activated dimer, we docked two copies of the AlphaFold prediction into the cryoEM density of the dimerization interface (Fig. 2.15, Fig. 2.14c). In this model, the overall twist and bend of the dimer have both decreased relative to the autoinhibited dimer. Importantly, the DCB-HUSdomains are closer to the membrane and both GEF domains are located such that both could Insert Arf1 into the membrane during the nucleotide exchange reaction. Although in this model the DCB-HUS domains are still not perfectly positioned to bind the membrane simultaneously, based on the observed flexibility of the Sec7 backbone we think it is possible that a small amount of additional bending along this trajectory is possible and would result in stable membrane association of Sec7.

We are therefore able to model the structural transitions that Sec7 needs to undergo to switch from the autoinhibited state to the active, membrane bound state (Fig. 2.15). In the active state, all of the known interactions would be satisfied. Because the HDS1-2 linker extends below the backbone it may contact the membrane first and loosely tether Sec7 to the surface. Near the membrane the conformational switch to the active state is likely promoted by interactions with the GTPase regulators.

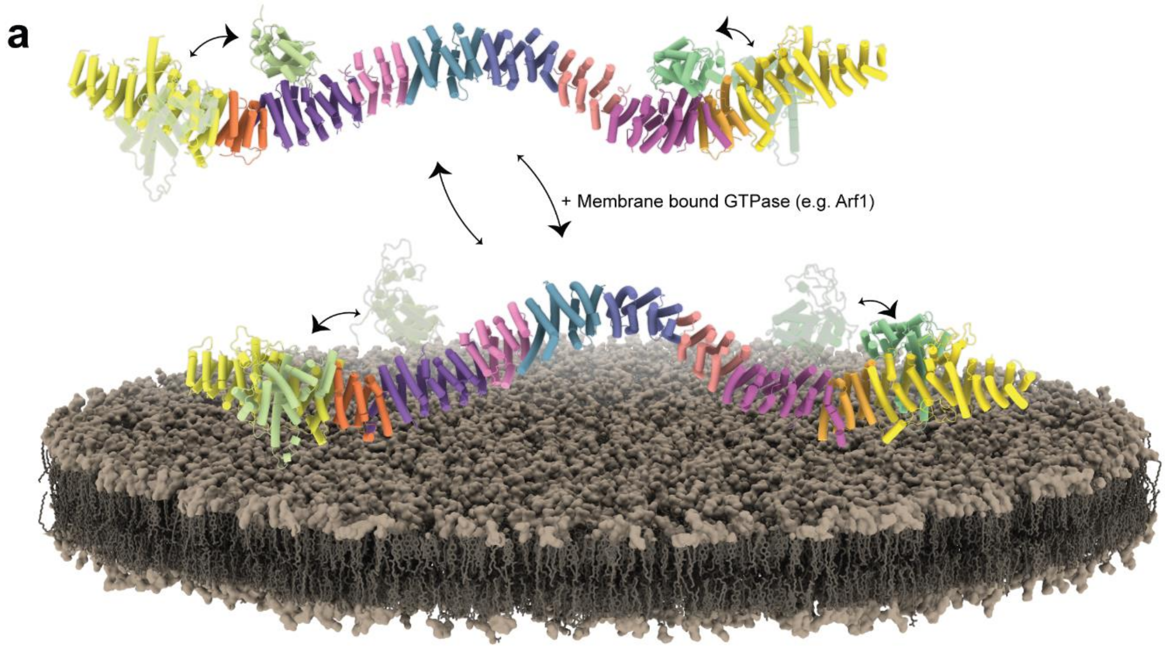


Figure 2.15 – Model for Sec7 activation on the membrane. The GEF domain is in equilibrium between the two conformations. In solution, affinity for the HDS2 domain and competition with the D loop favors the GEF domain in the autoinhibited conformation. The HDS1-2 loop weakly tethers Sec7 to the membrane, which subsequently stabilizes the GEF domain in the active conformation. Other factors on the membrane are then able to bind, further stabilizing the membrane bound state and/or preventing the GEF domain from switching back to the autoinhibited conformations.

2.10 – Discussion

The experiments reported here provide structural models for Sec7 in both its inhibited and active states. Our structural and functional data reveal how the arrangement and interaction of Sec7 regulatory and GEF domains enables Sec7 to switch from an autoinhibited conformation in solution to an active conformation on the membrane surface.

Regulatory mechanisms governing vesicle trafficking involve a complex network of interactions. Crosstalk between GTPases, GEFs and GAPs gives spatiotemporal specificity to trafficking pathways (Chen et al. 2011; Chesneau et al. 2012; Kobayashi and Fukuda 2012; D'Souza et al. 2014). When coupled with positive feedback and autoinhibition, this gives regulatory pathways unique and complex properties. For example, spontaneous nucleation of Rab5 activation by the Rab5 GEF Rabex5 propagates as a wave on a supported lipid bilayer in the presence of a GAP (Bezeljak et al. 2020). A similar dynamic behavior has also been described for the Ras GEF SOS (Iversen et al. 2014). Autoinhibition has been documented for GEFs and GAPs of GTPases from various families including Ras, Rho, Arf, and Rab. Allosteric occlusion of the catalytic site is the typical mechanism (Yadav and Bar-Sagi 2010; Moskwa et al. 2005; Bircher et al. 2022), however more complex mechanisms have been suggested (Lauer et al. 2019). The network of interactions governing Arf1 activation by its GEF at the TGN has been documented by our group and others (Galindo et al. 2016; McDonold and Fromme 2014; Richardson, McDonold, and Fromme 2012; Christis and Munro 2012), but the intramolecular interactions governing Sec7 have not been established.

Autoinhibition is a common regulatory mechanism among both GEFs (DiNitto et al. 2007; Richardson et al. 2016; Sondermann et al. 2004) and GAPs (Canagarajah et al. 2004), which is often coupled with positive feedback and coincidence detection to enhance specificity (Cherfils and Zeghouf 2013). A recent study identified an allele of the C. elegans Sec7 homologue AGEF-1 which encodes a single residue substitution in the HDS2 domain at the GEF interface we identified. This allele results in enlarged late endosomes/lysosomes, and suppresses the Vul phenotype of the *let-23(sy97)* allele. This provides evidence that the autoinhibitory HDS2-GEF interaction is broadly conserved and is broadly relevant to cell and developmental biological processes.

We found that the autoinhibition of Sec7 arises from an interaction between the catalytic GEF domain and the HDS2 regulatory domain. In the predicted active state, the GEF domain is instead poised on the membrane surface by interaction with the DCB-HUS domains. We found that a hyperactive construct interacted more robustly with membranes *in vitro*, suggesting switching to the active conformation is associated with stable membrane binding *in vitro* and *in vivo*. However it is also possible that Sec7 could be membrane bound yet remain autoinhibited, and this idea is supported by the fact that both Arf1-GTP and Rab11-GTP can recruit Sec7 to liposomes *in vitro* but Rab11-GTP has a significantly stronger stimulatory effect on Sec7 GEF activity.

Our group previously reported the HDS1 and HDS4 domains played autoinhibitory roles, and that the HDS1 domain switches to an activating conformation on membranes (Richardson, McDonold, and Fromme 2012; McDonold and Fromme 2014). In light of these results, several observations reconcile previous interpretations and provide additional support for our structural model of Sec7 autoregulation. First, constructs used previously to test the function of the HDS1 domain (Richardson, McDonold, and Fromme 2012; Nawrotek, Zeghouf, and Cherfils 2016; Walton, Leier, and Sztul 2020) considered the linker between the GEF and HDS1 domain as part of the HDS1 domain. These residues appear required to position the GEF relative to the DCB-HUS domains in the predicted active conformation of Sec7 and the Gea2 cryo-EM structure (Muccini, Gustafson, and Fromme 2022) and Fig 2.8b). The HUS-box element is also located adjacent to the GEF-HDS1 linker in these active states, explaining its important role in Arf1-activation on the membrane (Halaby and Fromme 2018). This GEF-HDS1 linker was absent in constructs lacking the HDS1 domain, perhaps explaining why the HDS1 domain was required for robust GEF activity on membranes (Richardson, McDonold, and Fromme 2012). Second, the HDS1 domain promotes membrane binding in constructs comprising the DCB-HUS, GEF, and HDS1 domains (Richardson, McDonold, and Fromme 2012; Halaby and Fromme 2018; Bouvet et al. 2013) These findings are consistent with a membrane-binding role for the HDS1-HDS2 linker region, because the HDS1-containing constructs used previously were truncated at residue 1220 and therefore include the conserved amphipathic helix present at the beginning of this linker. Taken together these findings also explain why the HDS1 domain does not stimulate GEF activity in solution (Richardson, McDonold, and Fromme 2012).

It is unclear why removal of the HDS4 domain resulted in loss of autoinhibition to a similar extent as loss of the HDS1-4 domains (McDonold and Fromme 2014), but we surmise that truncation of the HDS4 domain somehow perturbed the ability of the HDS2 domain to interact with the GEF domain. Our cryoEM structure of Sec7 enabled us to interfere with dimerization more surgically, resulting in our finding that the primary function of dimerization is to provide avidity for membrane-binding. Disruption of dimerization also enabled us to probe the role of the HDS1-HDS2 linker in membrane binding.

Although we have a predicted model for the active conformation of Sec7, the experimental structure of Sec7 in its active state on a membrane will require additional investigation. There may be more than one path between inactive/freely diffusing and active/membrane bound Sec7. For example, if the DCB-Arl1 interaction occurs on one side, membrane interactions might propagate down the dimer to untwist the dimer, but the specific order of binding events associated with full activation of Sec7 remains to be determined. Future studies are needed to determine precisely how Sec7 interacts with its regulatory GTPases on the membrane surface.

2.11 – Materials and methods

Cloning of expression and purification constructs

All constructs presented here were cloned with NEB HiFi gibson master mix (Cat. no. E2621) (see Table 2.2 and 2.3). All fragments were generated by PCR using a linearized template. Constructs used for plasmid shuffling were cloned into the pRS415 vector with the Sec7 promoter and terminator. All constructs are N-terminally GFP tagged, and otherwise only modified as indicated. Purification constructs were cloned into the pPICZB integration vector. The T. terrestris construct used for cryo em was cloned from a plasmid containing terrestris Sec7 with introns removed our group generated in a previous study (Richardson et al. 2016), subcloned into pPICZB with preScission cleavable 6xHis tags on the N and C termini. The S. cerevisiae constructs were cloned into the same vector, but with an N terminal 12xHis-GFP-Hrv3 tag. All constructs were verified by sequencing.

Strain modification

(See Table 2.3). To generate the *sec7*Δ*age2*Δ strain, BY4741 *sec7*Δ::KanMX (CFY409) was transformed with a longtine AGE2Δ::HIS3 cassette by standard LiOAc transformation. Pichia pastoris expression strains were generated by transforming the KM71H strain by electroporation as described previously after linearizing the pPICZB-Sec7 cassette with PmeI and ethanol precipitation.

Purification of Sec7 constructs

T. terrestris for cryo EM-

After transformation, 5 colonies were patched together on a fresh Zeocin plate, and cultured with autoinduction media as described (J. Y. Lee et al. 2017). Cells were collected by centrifugation after 48hrs, and yeast cell paste was flash frozen in liquid nitrogen in small aliquots. Cell paste was then lysed in a speX cryogenic mill (6875) for 15 cycles, 15 cps, 2 min rest between cycles, and powder was stored at -80 until further processing. Lysate powder was thawed rapidly by adding 18 ml room temperature thaw buffer (55 mM Hepes, pH 7.4, 495 mM KOAc, 11% glycerol, 44 mM Imidazole, 1.1 mM DTT, 1.1 mM AEBSF, and 1.1 x Roche cOmplete Protease Inhibitor Cocktail) to 20 g of powder/L, mixed to a slushy consistency, and sonicated with a macro tip sonicator, 100% power, 1s on 1s off, 20s cycles, until mostly thawed (~2 min). Then, 2 ml of 10% CHAPS was added and sonicated for an additional round to solubilize membranes. This lysate was clarified by centrifugation (20 min, 10,000 rpm in AV10 rotor, and 30 min 20,000 rpm in ss34 rotor), and loaded onto a pre-equilibrated HisTrap column (cytiva) with a syringe. Wash and elution was performed with an AKTA pure with two buffers: buffer A (50 mM Hepes, pH 7.4, 450 mM NaCl, 5% glycerol, 40 mM Imidazole, 1 mM DTT, and 0.1% CHAPS) buffer B (50 mM Hepes, pH 7.4, 300 mM NaCl, 10% glycerol, 500 mM Imidazole, 1 mM DTT, and 0.1% CHAPS). Column was washed with 25 CVs of 4% B, followed by a gradient up to 35% B over 10 CVs to elute. Elution fractions were checked by SDS-PAGE, pooled, concentrated, flash frozen in liquid nitrogen, and stored at -80 for later use.

S cerevisiae for biochemical analysis-

S. cerevisiae Sec7 constructs were grown/induced/collected as described above, but we found a batch purification protocol with NiNTA resin that had been fragmented by sonication (washed in miliQ H2O 3 times, resuspended in water to a 20% slurry, then sonicated at 80% power for 3 minutes with a macro tip, 20s on 10s off) improves yield (Fig. 2.16). This is likely due to the large hydrodynamic radius of Sec7 preventing access to the resin interior, and by fragmenting the resin we are sacrificing flow characteristics of the resin to increase the accessible area. Following sonication we equilibrated the “fined” resin 5x with buffer at a lower speed (1000 rpm) to ensure all resin that had fragmented too much to be efficiently collected by centrifugation was removed. Then the clarified lysate was incubated on the fined resin for 2 hours at 4 °C with rotation to bind. After binding the fined resin was washed 5x with 10 ml wash buffer, transferred to a fresh tube after the 3rd wash. Sec7 was eluted from the fined NiNTA resin by PreScission cleavage overnight at 4 °C. This elution was further purified by size exclusion chromatography using a Sepharose 6 increase 10/300 (SEC buffer: 25 mM Hepes, pH 7.4, 250 mM NaCl, 5% glycerol, 1 mM DTT). SEC fractions were analyzed by SDS-PAGE, pooled, concentrated, flash frozen in liquid nitrogen, and stored at -80 for later use.

ΔN17-Arf1, myristoylated Arf1, and Ypt31 were purified as previously described (Richardson and Fromme 2015; Richardson, McDonold, and Fromme 2012; Thomas and Fromme 2016)

Figure 2.16 – Fragmented NiNTA resin improves purification yield. a, DIC microscopy image (4x magnification) showing thermofisher HisPur resin (Cat. no. 88221) before (left) and after (right) 3 minutes of sonication. **b**, Parallel purifications using standard or fragmented resin.

CryoEM sample preparation and data collection

Grids were prepared using a Vitro-bot Mark IV. Quantifoil R1.2/1.3 grids were glow discharged in a Pelco easyGlow for 60s, 10 mA, under 9:1 argon:oxygen, and 3 ul of T. terrestris Sec7 in NiNTA elution buffer with 2 mM fluorinated fos-choline 8 (Anatrace, cat# F300F) was applied, blotted for 3 s, and immediately plunged into liquid ethane. Cryo EM data was collected at 63kx nominal magnification on a Talos Arctica operating at 200 kV equipped with a K3 detector and a BioQuantum energy filter. In total, 4,474 movies were collected, with 129 frames and a total dose of 55.6 e/A2. Automated data collection was done with SerialEM (Mastronarde 2005) using Aberration-free image shift to collect multi-shot movies.

CryoEM data processing

Movies were motion corrected and dose-weighted using MotionCor2 (Zheng et al. 2017), and CTF parameters were estimated using GCTF (Zhang 2016). Micrographs were manually inspected, and culled to 3,401 usable micrographs, which were imported into CryoSPARC for particle picking (Punjani et al. 2017). CTF parameters were also estimated with patch-CTF estimation in CryoSPARC. CryoSPARC blob picker was used to pick an initial set of particles, and generate templates for CryoSPARC template picker. An initial stack of 938,642 particles was iteratively refined by heterogeneous 3D refinement in CryoSPARC to a final stack of 296,177 particles. These particles were re-extracted in RELION 3.1 (Zivanov, Nakane, and Scheres 2020) using GCTF estimated CTF parameters. Iterative rounds of CTF refinement and Bayesian polishing improved the resolution from 7.14 A to 5.3 A. 3D classification was attempted, but no separation was attainable for the dimer particles. For focused refinement on a monomer (Nakane et al.

2018), a mask including a single full monomer and a small portion of the other (corresponding to approximately the HDS4 domain) was used for particle subtraction after symmetry expansion. A second monomer mask containing only the symmetrical regions of the map was used for refinement. After several iterations of CTF refinement, fixed angle 3D classification isolated 196,888 symmetry expanded particles which then improved further with iterative CTF refinement to the final map. For cryoDRGN analysis, we used TOPAZ to increase the likelihood of rare particles (Bepler et al. 2019) (this did not improve resolution of the monomer). Starting with 1,240,146 topaz picked particles heterogeneous 3D classification in CryoSPARC generated a final stack of 280,528 particles that were then used to generate a dimer map with C2 symmetry imposed (6.3 A), followed by a focused monomer refinement without subtraction before cryoDRGN training and analysis (Extended Data Fig 1) (Kinman et al. 2023).

Yeast complementation assay

Yeast expression plasmids with Sec7 constructs were transformed into the indicated strain (*sec7*Δ strain CFY409, *sec7*Δ*arf1*Δ strain CFY863 *sec7*Δ*age2*Δ strain). When colonies appeared (~2 days) single colonies were scraped into a clear bottom 96 well tray with selection media, and OD600 was measured with a plate reader to normalize the cell density, serial diluted and pinned on indicated media. Plates were incubated for three days at the indicated temperature (or 30 °C if not otherwise specified) before imaging.

Fluorescence microscopy

Colonies from the complementation assay were grown overnight at 30 °C in liquid selection media (-Leu) to an OD of 0.6. Cells were allowed to settle on a coverslip dish (MatTek) for 10 min, and washed with fresh media. All Imaging was done using a DeltaVision Elite system equipped with an Olympus IX-71 inverted microscope, a DV Elite complementary metal-oxide semiconductor camera, a ×100/1.4 NA oil objective, and a DV Light SSI 7 Color illumination system with Live Cell Speed Option with DV Elite filter sets. Exposure and laser power were adjusted according to intensity, but are the same for all mutant constructs, and the brightness/intensity was normalized across the constructs using ImageJ.

Liposome preparation

Synthetic TGN liposomes were prepared as described previously (Richardson and Fromme 2015). In brief - lipid stocks in chloroform were mixed in a pear shaped flask according to the mixture of TGN lipids identified by (Klemm et al. 2009) (Table 2.4). Chloroform was evaporated slowly in a rotary evaporator heated to ~37 °C, then rehydrated in HK buffer (20mM HEPES pH 7.5, 150 mM KOAc) at 37 °C overnight. After gentle resuspension, the mixture was extruded through 100 nm filters 21 times and stored at 4 °C for no more than 1 month.

Liposome flotation membrane binding assay

Liposome flotation was performed as described (Highland, Thomas, and Fromme 2023). Briefly - 100 nm liposomes were loaded with Arf1 by EDTA exchange of GMPPNP for a final concentration of 250 µM lipid and the indicated Arf1 concentration. Sec7 constructs were added to a final concentration of 550 nM, and incubated at room temperature for 1 hour. Then 2.5 M sucrose in HK (20mM HEPES pH 7.5, 150 mM KOAc) was added to 1 M final concentration, and 80 µl was transferred to a polycarbonate tube. Then a 100 ul layer of 0.75 M Sucrose was added, followed by 20 µl of HK. This was centrifuged at 20 oC in a TLA100 rotor for 30 minutes, and the top layer collected for SDS-PAGE analysis (12% acrylamide gel).

In vitro Arf activation assay.

GEF activity was determined by measuring the native Trp fluorescence of Arf1, as described previously (Richardson and Fromme 2015). Briefly, to a final volume of 150 µl, 200 µM 100 nm TGN liposomes, 20-100 nM GEF (as indicated), 1 µM myristoylated Arf1, and 200 µM GTP were added to a quartz cuvette, diluted to the final volume in HKM (20mM HEPES pH 7.5, 150 mM KOAc, 2 mM $MgCl_2$, 1 mM DTT). Tryptophan fluorescence (297.5 nm excitation, 340 nm emission) was measured using a fluorometer, and between additions the reaction was well mixed and the fluorescence was allowed to stabilize. The order listed is the order components were added, except for reactions with preloaded Ypt31. For these reactions, 500 nM prenylated Ypt31 was loaded onto liposomes by EDTA exchange of GMPPNP at 30 oC for 30 minutes. The presence of GMPPNP in the liposome mixture dictated that Arf1 was added last. Curves were fit in GraphPad PRISM.

Table 2.2 – Recombinant DNA, Chapter 2

Name	Description	Vector Backbone	Source
pPICZB	*P. pastoris* integration plasmid		Thermo Fischer Cat # V19020
	pPICZB with N terminal 12xHis-GFP-HrV 3C	pPICZB	This Study
pBB01271	pPICZB *T. terrestris* Sec7 (6x-His-Sec7-6xHis)	pPICZB	This Study
pBB01793	pPICZB *S. cerevisiae* Sec7 WT (12xHis-GFP-HR3 VC-Sec7)	pPICZB	This Study
pBB01921	pPICZB *S. cerevisiae* Sec7 monomer (12xHis-GFP-HR3 VC-Sec7 Y1975S, I1979S, L1982S, V1986S, L1998S, V2001S)	pPICZB	This Study
pBB01922	pPICZB *S. cerevisiae* Sec7 L1376D (12xHis-GFP-HR3 VC-Sec7 L1376D)	pPICZB	This Study
pBB01923	pPICZB *S. cerevisiae* Sec7 ΔD-loop (12xHis-GFP-HR3 VC-Sec7 Δ444-486)	pPICZB	This Study
pBB01924	pPICZB *S. cerevisiae* Sec7 L1376 & ΔD-loop (12xHis-GFP-HR3 VC-Sec7 L1376, Δ444-486)	pPICZB	This Study
pBB01925	pPICZB *S. cerevisiae* Sec7 Δ1201-1250 (12xHis-GFP-HR3 VC-Sec7 Δ1201-1250)	pPICZB	This Study
pBB01926	pPICZB *S. cerevisiae* Sec7 monomer & Δ1201-1250 (12xHis-GFP-HR3 VC-Sec7 Δ1201-1250, Y1975S, I1979S, L1982S, V1986S, L1998S, V2001S)	pPICZB	This Study
pRS415	pRS415 vector (centromeric LEU2 plasmid)		(Sikorski and Hieter 1989)
pCF1084	pRS415;GFP-Sec7		(Richardson, McDonold, and Fromme 2012)
pBB01602	pRS415;GFP-Sec7 L1376D		This Study
pBB01761	pRS415;GFP-Sec7 L1376K		This Study
pBB01762	pRS415;GFP-Sec7 L1376S		This Study

pBB01871	pRS415;GFP-Sec7 ΔD-loop (Δ444-486)		This Study
pBB01863	pRS415;GFP-Sec7 L1331S, L1376S, F1379S		This Study
pBB01874	pRS415;GFP-Sec7 ΔD-loop (Δ444-486) & L1376D		This Study
pBB01876	pRS415;GFP-Sec7 ΔD-loop (Δ444-486) & L1331S, L1376S, F1379S		This Study
pBB01691	pRS415;GFP-Sec7 monomer (Y1975S, I1979S, L1982S, V1986S, L1998S, V2001S)		This Study
pBB01753	pRS415;GFP-Sec7 Δ1201-1250		This Study
pBB01867	pRS415;GFP-Sec7 monomer & Δ1201-1250 (Δ1201-1250, Y1975S, I1979S, L1982S, V1986S, L1998S, V2001S)		This Study
pArf1	Bacterial Arf1 Expression plasmid	pET3C	(Randazzo, Weiss, and Kahn 1992)
pBB131	Bacterial Nmt1 expression plasmid	pCYC	(Duronio et al. 1990)
pCF1053	Arf1ΔN17	pET28	(Richardson, McDonold, and Fromme 2012)
pLT72	Full-length Ypt31 with cleavable N-terminal GST tag	pGEX-6P	(Thomas and Fromme 2016)
pLT40	Gdi1 with cleavable N-terminal GST tag	pGEX-6P	(Thomas and Fromme 2016)
pLT35	Mrs6 with cleavable N-terminal His 6 tag	pET28	(Thomas and Fromme 2016)
pLT41	Bet2 with cleavable N-terminal His 6 tag and Bet4	pCDF-Duet-1	(Thomas and Fromme 2016)

Table 2.3 – Bacterial and yeast strains, Chapter 2

Name	Genotype	Source
Rosetta2, *E. coli*		Novagen Cat # 71400
DH5α, *E. coli*		New England Biolabs Cat # C2987I
BY4742	*MATα his3-Δ1 leu2-Δ0 lys2-Δ0 ura3-Δ0*	(Brachmann et al. 1998)
CFY409	BY4742 *sec7Δ::KANMX* + pCF1043	(Richardson, McDonold, and Fromme 2012)
CFY863	CFY409 *arf1Δ::HIS3* + pCF1043	(Richardson, McDonold, and Fromme 2012)
CFY4969	CFY409 *age2Δ::HIS3* + pCF1043	This study
KM71H (*P. pastoris*)		ThermoFisher Scientific Cat # C18200
CFY4970	KM71H *pAOX1*::6x-His-Sec7(*T. terrestris*)-6xHis::*BleoR*	This study
CFY4971	KM71H *pAOX1*::12xHis-GFP-HR3 VC-Sec7 (*S. cerevisiae*)::*BleoR*	This study
CFY4972	KM71H *pAOX1*::12xHis-GFP-HR3 VC-Sec7 monomer (*S. cerevisiae*, Y1975S, I1979S, L1982S, V1986S, L1998S, V2001S)::*BleoR*	This study
CFY4973	KM71H *pAOX1*::12xHis-GFP-HR3 VC-Sec7 L1376D (*S. cerevisiae*)::*BleoR*	This study
CFY4974	KM71H *pAOX1*::12xHis-GFP-HR3 VC-Sec7 ΔD-loop (*S. cerevisiae*, Δ444-486)::*BleoR*	This study
CFY4975	KM71H *pAOX1*::12xHis-GFP-HR3 VC-Sec7 ΔD-loop & L1376D (*S. cerevisiae*, Δ444-486, L1376D)::*BleoR*	This study
CFY4976	KM71H *pAOX1*::12xHis-GFP-HR3 VC-Sec7 Δ1201-1250 (*S. cerevisiae*)::*BleoR*	This study
CFY4977	KM71H *pAOX1*::12xHis-GFP-HR3 VC-Sec7 monomer & Δ1201-1250 (*S. cerevisiae*, Δ1201-1250, Y1975S, I1979S, L1982S, V1986S, L1998S, V2001S)::*BleoR*	This study

Table 2.3 - TGN liposome composition

Lipid	mol%
DOPC	24
POPC	6
DOPE	7
POPE	3
DOPS	1
POPS	2
DOPA	1
POPA	2
Liver or Soy PI	30
PI(4)P	1
CDP-DAG	2
PO-DAG	4
DO-DAG	2
Ceramide (C18)	5
Cholesterol	10
DiR	1

CHAPTER 3 – The Sys1, Arl3, Arl1, cascade

3.1 - Introduction

The Arf-like GTPases are less well understood than the Arfs. This is partly because in humans there are over 20 Arl proteins, and they share only 40-60% sequence identity with each other or any Arf protein (Jochum et al. 2002). In yeast, there are two Arf-like proteins, Arl1 and Arl3 (Arl1 and Arfrp1 in humans respectively). They localize to the late Golgi, but there is little information about their function. Both are nonessential genes in yeast, but Arl1 is essential for Drosophila melanogaster development past late L3 larval stages or pupae (Torres, Rosa-Ferreira, and Munro 2014), ard *arfrp1* knockout causes embryonic lethality in mice (Mueller et al. 2002). In both cases, this is thought to be caused by disruption of Arl1 effectors.

In yeast, two clear effectors have been established for Arl1: the coiled coil tether Imh1 and Sec7. Crystal structures of both interactions have been solved (M. Wu et al. 2004; Galindo et al. 2016). Arl1 has also been reported to interact with Mon2 (Jochum et al. 2002), however the relevance of this interaction is unclear. Localization of Sec7 requires Arl1 in mammalian cells (Christis and Munro 2012), but not in yeast (McDonold and Fromme 2014), allowing for more direct interrogation of Arl1 function. In *arl1*Δ null yeast, vacuole sorting and bulk phase endocytosis is disrupted. No specific effectors have been established to date for Arl3, although it is thought to recruit the GEF for Arl1.

Two studies published simultaneously in 2003 first reported the Arl3-Arl1 cascade (Panic, Whyte, and Munro 2003; Setty et al. 2003). The primary evidence for this is the dependence of Arl1 on Arl3 for Golgi localization, suggesting Arl3 acts upstream.

However, to date a rigorous Arl3 - Arl1 GEF interaction has not been identified. In part, this is because a GEF has not been rigorously demonstrated for either Arl3 or Arl1.

A poorly conserved Sec7-domain-containing protein in yeast, Syt1, has been suggested to be the GEF for Arl1 (K.-Y. Chen et al. 2010). This was reported based on partial mislocalization of Arl1 in *syt1*Δ yeast, catalytic activity of its GEF domain, and interaction by yeast-2-hybrid. However, there are several reasons to be skeptical of this conclusion. First, Syt1 does not have a clear homologue in metazoans, and so another factor must be relevant. Second, *syt1*Δ does not phenocopy *arl1*Δ, and is not essential for Arl1 suppression of *ypt6*Δ (I.-H. Wang et al. 2017; Y.-T. Chen et al. 2019). Finally, Syt1 has been shown to catalyze exchange of other Arf family GTPases, suggesting it may have some promiscuous activity (Jones et al. 1999). Shown below is data further refuting this claim.

Localization of Arl3 to membranes differs from other Arf family GTPases. Rather than a lipid modification to the amphipathic helix, Arl3 is acetylated and requires a small transmembrane protein Sys1 for its localization (Setty et al. 2004; Behnia et al. 2004). This interaction has not been extensively characterized. Sys1 was originally identified in a screen for genetic interactions with Ypt6 (Tsukada and Gallwitz 1996). It is now clear this interaction is due to parallel, partially overlapping functions of Ypt6 and Arl3/Arl1, in recruiting the GARP or Golgin tethering factors respectively. Subsequent studies showed that Arl3 is acetylated by the NatC N^{α}-acetyltransferase complex and this modification is required for its association with Sys1, suggesting Sys1 acts as a receptor for Arl3. However, this interaction has not been characterized biochemically *ex vivo*. As nucleotide exchange of Arf family GTPases is coupled to the displacement of the amphipathic helix, it is possible this interaction partially facilitates nucleotide exchange.

Presented here is evidence that the Sys1-Arl3-Arl1 cascade is more complex than previously realized. Ectopically localized Sys1 at the ER causes relocalization of Arl3, but not Arl1, suggesting Sys1 is sufficient for Arl3 activation and the Sys1-Arl3 complex is necessary but not sufficient for the Arl1 GEF. Constructs of truncated Sys1 identify a portion of the C terminus that may act catalytically to activate Arl3. AlphaFold predictions of the Arl3-Sys1 complex validate this hypothesis, however future work is required to establish Sys1 as a bonafide Arl3 GEF.

3.2 – Results: The C-terminus of Sys1 is important for its function in the Arl3-Arl1 cascade

Sys1 is a short integral membrane protein with only four predicted transmembrane helices separated by short linkers (Fig. 3.1a). Both termini extend into the cytosol, and the final C-terminal transmembrane helix is predicted to extend slightly farther than the others, presumably above the membrane surface, followed by an unstructured loop, with a non-conserved C-terminal extension. To test if the C-terminus of Sys1 is essential for Arl3 localization, we generated truncated Sys1 expression constructs and tested their ability to restore localization of endogenously mNeonGreen tagged Arl1 and Arl3 in a *sys1*Δ null strain background (Fig. 3.1b). To test if the non-conserved extension is required, Sys1t (Sys1 1-180) removes the poorly conserved residues but retains some unstructured loop following the C-terminal helix. To probe if the C-terminal region exposed to the cytosol is required, Sys1ΔC (Sys1 1-155) removes all residues that extend beyond the predicted transmembrane domain. We found that Sys1t is sufficient for both Arl3 and Arl1 localization, while Sys1ΔC is sufficient for Arl3 to localize to punctate structures, while Arl1 remains cytosolic (Fig 3.1c). We surmised that this is due to inactive Arl3 localizing to Golgi puncta, which are not able to recruit the Arl1 GEF. This is supported by a 2020 report that the N-terminal helix of Arfrp1 is sufficient for Golgi localization in mammalian cells (F. Yang et al. 2020).

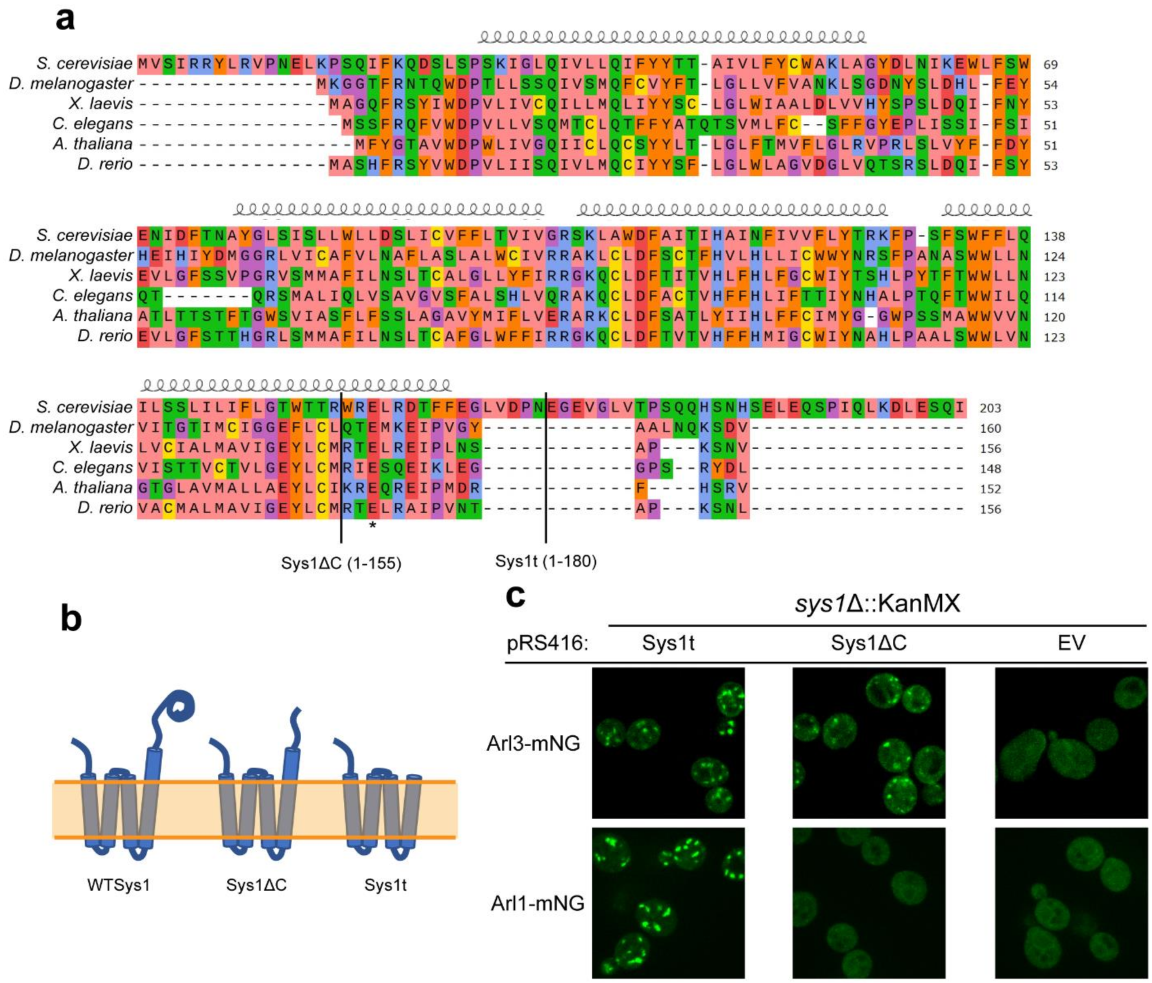


Figure 3.1 - The C-terminus of Sys1 is important for its function in the Arl3-Arl1 cascade. a, Multiple sequence alignment of Sys1 homologues, colored by physicochemical property of the side chain (zappo). Truncation constructs described in the main text are marked, and E157 is indicated with "*". **b**, Cartoon depiction of Sys1 and truncation constructs used in c. **c**, Microscopy of Arl3-mNeonGreen and Arl1-mNeonGreen endogenously tagged in a *sys1*Δ null background after transformation with the indicated Sys1 construct.

3.3 – Results: Evidence for Sys1 as an Arl3 GEF

We noted within the C-terminus of Sys1 is a strictly conserved glutamate, E157 in *S. cerevisiae* (Fig 3.1a). If Sys1 is an Arl3 GEF, it may share a similar mechanism to Sec7. Two interactions between Sec7 and its substrate mediate nucleotide exchange. Hydrophobic grooves in the GEF domain stabilize the switch regions in an intermediate state, while a glutamate displaces the bound nucleotide and Mg^{2+} ((Beraud-Dufour et al. 1998; Goldberg 1998)). We hypothesized that the established interaction between the N-terminus of Arl3 and Sys1 could be analogous mechanistically to the hydrophobic GEF-switch interactions, while this glutamate displaces the nucleotide. To test this hypothesis, we generated expression constructs with a 3xHA epitope tag and an E157A substitution and tested their ability to restore Arl3 and Arl1 localization as before (Fig 3.2a). Surprisingly, the addition of the 3xHA tag resulted in a significant vacuolar localization of Arl3, but not Arl1. This resulted in less punctate localization of Arl3 when Sys1ΔC is expressed, likely due to less available Sys1 at the Golgi. Remarkably, the single residue substitution (Sys1t E157A) resulted in a construct with nearly identical Arl3 and Arl1 localization as Sys1ΔC. Equal expression of all constructs was verified by western blot (Fig. 3.2b). These data suggest that this single residue is required for Arl3 activation *in vivo*, and supports the hypothesis that Sys1 is both a receptor and a GEF for Arl3.

We next wondered if Sys1 was sufficient for Arl3 and Arl1 localization. To test this, we generated expression constructs of Sys1 with a C-terminal KKSL sequence (Fig. 3c). This motif is recognized by the KDEL receptor and causes Sys1 to remain trapped in the ER, and has been used in mammalian cells (Pelham 1989; Behnia et al. 2004). Remarkably, this results in a redistribution of Arl3 to the ER, but not Arl1. This result indicates that Sys1 is sufficient for Arl3 localization however it is unclear if it is sufficient for activation. It is possible that other factors promote specificity for Arl1 to late Golgi compartments, or that Arl3 is not activated by Sys1-KKSL.

To test if the acetylated N-terminus is essential for Arl3 activation, we generated expression constructs of chimeric Arl1 and Arl3 (Fig. 3.2e,f). The acetylated N-terminus of Arl3 is required for the Arl3-Sys1 interaction (Setty et al. 2004; Behnia et al. 2004), so we reasoned swapping the acetylated N-terminus with the myristoylated Arl1 amphipathic helix would disrupt this interaction while preserving the membrane association. Remarkably, both Arl3 and Arl1 with the other helix localize to puncta in WT cells (Fig. 3.2e). However, when expressed in a *sys1*Δ null strain, the Arl3 GTPase with Arl1 N-terminus was cytosolic (Fig. 3.2f). This suggests the Arl3-Sys1 interaction does play a role in activation of Arl3, however further study is required to verify these *in vivo* results.

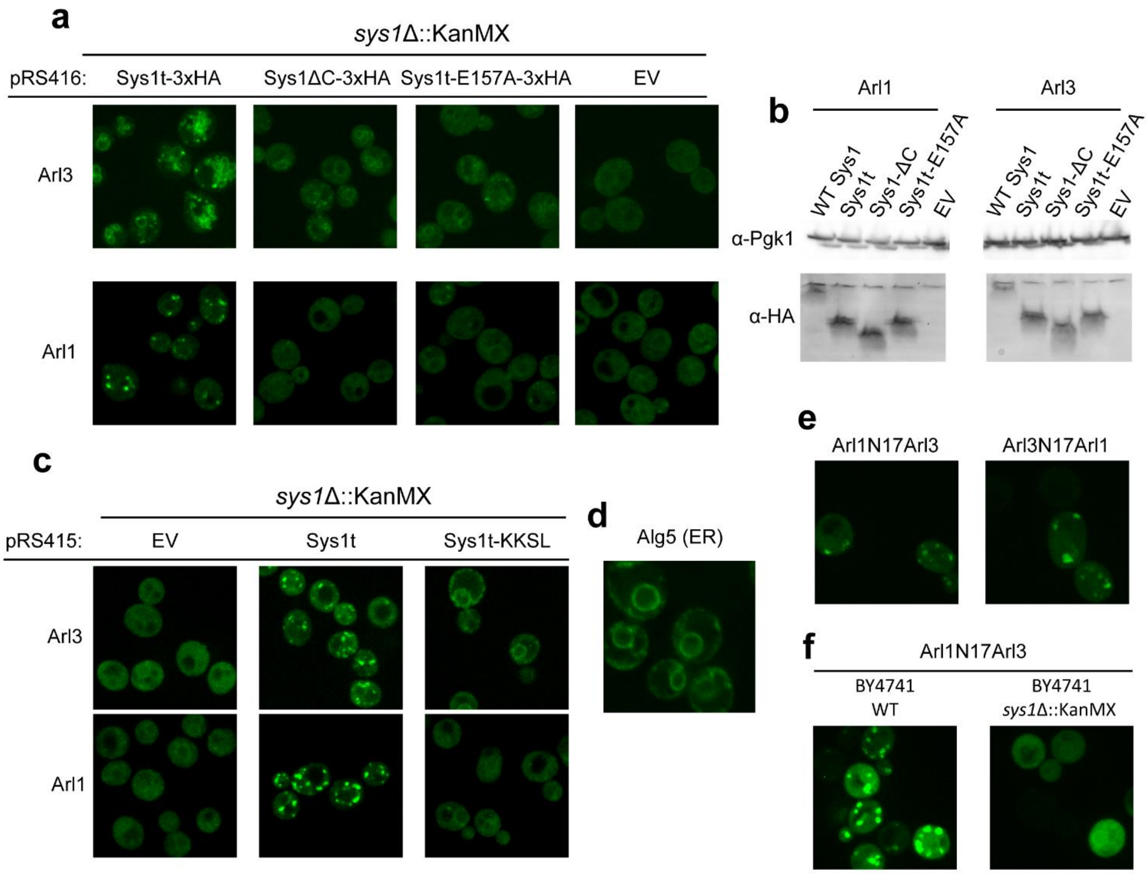


Figure 3.2 – Evidence for Sys1 acting as an Arl3 GEF. a, Fluorescence microscopy of endogenously tagged Arl3-mNeonGreen and Arl1-mNeonGreen when the indicated Sys1 construct is the sole copy of Sys1. The Sys1 E157A mutation is nearly identical to the Sys1ΔC. **b**, Western blot of TCA precipitation of cells shown in a. Pgk1 is a loading control, 3xHA has a nonspecific band in all lanes. All Sys1 constructs are comparably expressed. **c**, fluorescence microscopy as in 'a' but with a Sys1 construct containing a C-terminal KKSL motif to block its export from the ER. **d**, A representative image of the ER marked by Alg5-mNeonGreen. **e**, Chimeric constructs of Arl3 and Arl1 with swapped N-terminal helices (residues 1-17) and mNeonGreen tag. Arl1N17Arl3 indicates Arl3 with the Arl1 N-terminus. Arl3NArl1 indicates Arl1 with the Arl3 N-terminus. These were expressed in wild-type SEY6210. **f**, Microscopy of the Arl1N17Arl3 construct expressed in BY4741 wild-type (left) and sys*1*Δ null (right).

3.4 – Results: Syt1 is likely not the only GEF for Arl1

To further interrogate the Sys1-Arl3-Arl1 cascade, we first wanted to validate the reported Arl1 GEF, Syt1. We generated a strain of endogenously tagged Arl1-mNeonGreen with *syt1*Δ null, and checked the localization of Arl1 by fluorescence microscopy (Fig. 3.3). We found Arl1 was localized to puncta, in contradiction with the 2010 report (K.-Y. Chen et al. 2010), and verified this result in three independent clones. We can only speculate why we were unable to replicate the 2010 study, however one possibility is a difference in strain background. Our strain was generated in SEY6210, while Chen et al. utilized the BY4741 background. Regardless, our finding is more in line with other published data as discussed above: no clear Syt1 homologue, differences in genetic interactions of Syt1 and Arl1, and known promiscuous activity of Sys1.

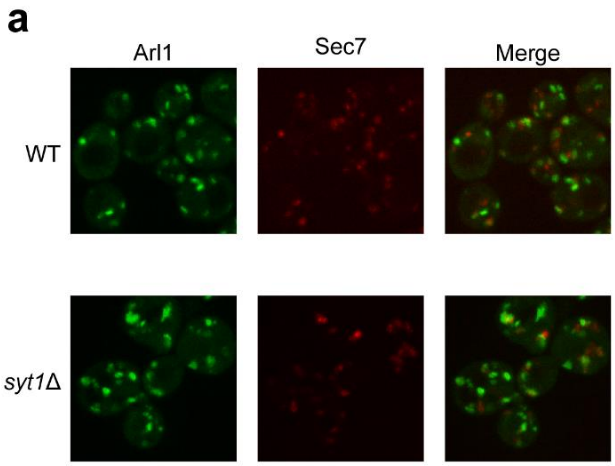


Figure 3.3 – Arl1 localizes to punctate structures in the absence of Syt1. a, Strains harboring endogenously tagged Arl1-mNeonGreen and Sec7-Mars were imaged as described in methods. In the absence of Syt1 (*syt1*Δ), Arl1 localizes to punctate structures that partially overlap with Sec7, as in an otherwise wild-type background. Representative image of 3 independent clones of *syt1*Δ.

3.5 – Future directions

Presented above is some evidence providing more detail about the Sys1-Arl3 interaction, and may indicate that Sys1 facilitates nucleotide exchange of Arl3. However, further study is required to directly test for GEF activity of Sys1. Reconstitution of the Arl3 activation including Sys1 *in vitro* would provide strong evidence to validate the *in vivo* data, however this has been challenging due to the recalcitrance of Sys1 for purification. During her rotation, Jacqueline Ehlrich performed a detergent screen in Toshi Kawate's laboratory by fluorescence-detection size-exclusion chromatography, and found the detergent C12E8 as a good candidate for purification, however all detergents tested solubilized Sys1 to some degree. Optimizing this purification to perform this assay is a top priority to complete this work.

The next steps to delineating the Sys1-Arl3-Arl1 cascade are to identify the GEF for Arl1. Presented here is evidence contradictory to the proposed Arl1 GEF, but more in line with other genetic and biochemical data. It is possible the Sys1-Arl3 complex itself induces Arl1 exchange, perhaps involving the glutamate residue discussed above (E157). Reconstitution of the Sys1-Arl3 complex *in vitro* would allow this hypothesis to be tested. Otherwise, a SILAC based approach may prove useful to identifying novel Arl1 interactors.

3.6 – Materials and Methods

Cloning

All constructs presented here were cloned with NEB HiFi gibson master mix (Cat. no. E2621) (see Table 2.2 and 2.3). All fragments were generated by PCR using NEB Phusion polymerase (Cat. No. M0530).

Strain modification

(See Table 3.2). To generate the fluorescent tagged strains with sys1Δ, SEY6210 was transformed with the *sys1Δ*::KANMX deletion cassette PCR amplified from BY4741 *sys1*Δ::KanMX (CFY954) by standard LiOAc transformation. Plasmids were transformed within 1 week of microscopy.

Fluorescence microscopy

Single colonies from freshly transformed Colonies from the complementation assay were grown overnight at 30 °C in liquid selection media (-Leu) to an OD of 0.6. Cells were allowed to settle on a coverslip dish (MatTek) for 10 min, and washed with fresh media. All Imaging was done using a DeltaVision Elite system equipped with an Olympus IX-71 inverted microscope, a DV Elite complementary metal-oxide semiconductor camera, a ×100/1.4 NA oil objective, and a DV Light SSI 7 Color illumination system with Live Cell Speed Option with DV Elite filter sets. Exposure and laser power were adjusted according to intensity, but are the same for all mutant constructs, and the brightness/intensity was normalized across the constructs using ImageJ.

Expression test

Cultures used for microscopy were diluted in 50 ml liquid culture and grown to mid-log phase. 5 OD600 mls of cells were collected by centrifugation, washed in cold water, resuspended in 1 ml cold 10% TCA and allowed to rest for 30 minutes. The precipitate was collected by centrifugation at 14 krpm, 4 C for 3 min, and washed twice with acetone. After drying in a speedvac, the pellet was resuspended in boiling buffer (50 mM Tris pH 7.5, 1 mM EDTA, 1%SDS) and ~100 ul of glass beads were added, and vortexed for 3 min. Beads were removed, samples were diluted in 2x urea sample buffer (150 mM Trip pH 6.8, 6M urea, 6% SDS, bromo blue added by eye for color) and samples were run on a 15 % SDS-PAGE gel. Bands were transferred to a PVDF membrane using a BioRad TurboBlot, and developed with the indicated antibodies.

Table 3.1 – Recombinant DNA, Chapter 3

Name	Description	Vector Backbone	Source
pRS416	pRS416 vector (centromeric URA plasmid)		(Sikorski and Hieter 1989)
pRS415	pRS415 vector (centromeric LEU2 plasmid)		(Sikorski and Hieter 1989)
pBB	pRS416;Sys1	pRS416	This study
	pRS416;Sys1t	pRS416	This study
	pRS416;Sys1ΔC	pRS416	This study
	pRS416;Sys1t-3xHA	pRS416	This study
	pRS416;Sys1ΔC-3xHA	pRS416	This study
	pRS416;Sys1E157A-3xHA	pRS416	This study
	pRS416;Sys1t-KKSL	pRS416	This study
	pRS451-Arl1N17Arl3	pRS415	This study
	pRS451-Arl3N17Arl1	pRS415	This study

Table 2.3 – Bacterial and yeast strains, Chapter 2

Name	Genotype	Source
SEY6210	*MATα Ura3-52, his3-Δ200 leu2-3, 112 lys2-801, trp1-Δ901*	(Robinson et al. 1988)
BY4742	*MATα his3-Δ1 leu2-Δ0 lys2-Δ0 ura3-Δ0*	(Brachmann et al. 1998)
CFY954	BY4741 *sys1Δ*::KANMX	Dharmacon deletion collection
CFY4977	SEY6210 Arl1-mNeonGreen::HIS3 *sys1Δ::KANMX*	This study
CFY4978	SEY6210 Arl3-mNeonGreen::HIS3 *sys1Δ::KANMX*	This study
CFY3580	SEY6210 Alg5-mNeonGreen::HIS3	Fromme Lab (unpublished)

CHAPTER 4 – Structural Characterization of NrnC Identifies Unifying Features of Dinucleases

4.1 – Summary of work presented

This chapter is the product of a multi-institution collaboration between the laboratories of Chris Fromme, Holger Sondermann, and Vincent Lee. The full authors list includes: Justin D. Lormand, Soo-Kyoung Kim, George A. Walters-Marrah, Bryce A. Brownfield, J. Christopher Fromme, Wade C. Winkler, Jonathan R. Goodson, Vincent T. Lee, Holger Sondermann. Bryce A. Brownfield's contribution includes all cryoEM experiments, including sample preparation, data collection, and analysis. Justin D. Lormand provided all purified protein and ligand samples, and was responsible for connecting the cryoEM work to the rest of the project. Figures 4.12 - 4.15 and the cryoEM methods section were the primary contributions of Bryce A. Brownfield.

4.2 – Abstract

RNA degradation is fundamental for cellular homeostasis. The process is carried out by various classes of endolytic and exolytic enzymes that together degrade an RNA polymer to monoribonucleotides. Within the exoribonucleases, nanoRNases play a unique role as they act on the smallest breakdown products and hence catalyze the final steps in the process. We recently showed that oligoribonuclease (Orn) acts as a dedicated diribonucleotidase, defining the ultimate step in RNA degradation that is crucial for cellular fitness (Kim et al., 2019). Whether such a specific activity exists in organisms that lack Orn-type exoribonucleases remained unclear. Through quantitative structure-function analyses we show here that NrnC-type RNases share this narrow substrate preference with Orn. Although NrnC employs similar structural features that distinguish these two classes as diribonucleotidases from other exoribonucleases, these key determinants for diribonucleotidase activity are realized through distinct structural scaffolds. The structures together with comparative genomic analyses of the phylogeny of DEDD-type exoribonucleases implicate convergent evolution as the mechanism of how diribonucleotidase activity emerged repeatedly in various organisms. The evolutionary pressure to maintain diribonucleotidase activity further underlines the important role these analogous proteins play for cell growth.

4.3 – Introduction

Traditionally, nanoRNases – enzymes that act on short, typically 2 to 7 residue-long RNA substrates – have occupied a distinct role in RNA metabolism as they catalyze the final steps in RNA degradation. To date, three main enzyme families have been assigned this function (Liao et al., 2018): Oligoribonucleases (Orn) (Cohen et al., 2015; Orr et al., 2015; Zhang et al., 1998), nanoRNase A and B (NrnA and NrnB) (Fang et al., 2009; Mechold et al., 2007), and nanoRNase C (NrnC) (Liu et al., 2012). NrnA is comprised of DHH-DHHA1 domains and is suggested to act as a bidirectional exonuclease that cleaves short RNA fragments from 3’ to 5’ and longer substrates from 5’ to 3’ (Mechold et al., 2007; Schmier et al., 2017). Orn and NrnC are 3’-5’ exonucleases with a DnaQ fold containing the catalytic DEDD motif (Chin et al., 2006; Yuan et al., 2018), a domain that is common in enzymes that act on nucleic acids. Notably, Orn (and its eukaryotic ortholog REXO2) or NrnC activity is critical for cellular growth, rendering them unique amongst the exoribonucleases known to date (Ghosh & Deutscher, 1999; Liu et al., 2012; Nicholls et al., 2019).

Despite its classification as a nanoRNase, we recently reported that Orn acts primarily as a diribonucleotidase, assigning it a highly specific and unique function in clearing the diribonucleotide pool as the terminal step in RNA degradation. Substrate-bound structures of bacterial and human orthologs, Orn and REXO2, respectively, revealed the scissile bond of the dinucleotide surrounded by the conserved, catalytic DEDDh motif that is involved in divalent cation coordination and catalysis. More importantly, key determinants for the RNA length preference of Orn and REXO2 include a leucine residue that wedges between the two bases of the diribonucleotide and a phosphate cap – invariable residues that coordinate the 5’-phosphate on the substrate and limit substrate length to two nucleotides (Kim et al., 2019). An independent study on

REXO2 confirmed our structural analysis and established the human enzyme as a diribonucleotidase in mitochondria, where its activity alters gene expression (Nicholls et al., 2019), a function that relates to the role of diribonucleotides in transcription initiation (Druzhinin et al., 2015; Goldman et al., 2011; Nicholls et al., 2019; Vvedenskaya et al., 2012).

Our previous work on Orn-type RNases was motivated by three main considerations: Orn’s essential role for growth in many bacteria, its role in cyclic dinucleotide signaling, and a lack of understanding how substrate specificity towards short RNA substrates, and in turn lack of activity towards longer RNAs, is achieved (Ghosh & Deutscher, 1999; Kamp et al., 2013; Kim et al., 2019; Orr et al., 2015; Palace et al., 2014). The latter is a basic question as the unique substrate profile is the defining characteristic of this class of enzymes. We demonstrated that Orn has a much higher preference for diribonucleotides compared to 3-7 residue-long RNAs than anticipated previously. This selectivity is due to an active site that is exquisitely suited for diribonucleotides with a 5’ phosphate. The enzyme’s diribonucleotidase activity is required for normal bacterial growth and clearance of c-di-GMP breakdown products. Notably, a knock-out of Pseudomonas aeruginosa Orn can not only be complemented by various Orn orthologs, but also by the other three nanoRNases: NrnA, NrnB, and NrnC (Orr et al., 2018). However, it is not clear whether this functional complementary correlates with a narrow substrate specificity for dinucleotides.

A recent structural study of NrnC from the soil bacterium Agrobacterium tumefaciens identified a homo-octameric assembly as the enzyme’s functional unit (Yuan et al., 2018). The unit can be divided into two stacked rings, each composed of four NrnC monomers, which together form a central channel lined with eight active sites. Mutagenesis of key active-site residues confirmed the requirement of the conserved

DEDDy motif for catalysis and identified positively charged residues lining the channel, which are also important for function. Activity assays indicated single-stranded RNA and double-stranded DNA as potential substrates of NrnC, however the structural basis for substrate specificity has not been established.

Here, we ask the fundamental question: What is the substrate specificity of NrnC-type enzymes, an RNase that is essential for the growth of gram-negative bacteria such as Caulobacter crescentus, Bartonella henselae and Brucella abortus (Christen et al., 2011; Liu et al., 2012; Sternon et al., 2018). We present crystal structures of NrnC from B. henselae and Brucella melitensis in their substrate-bound and apo states. The structures confirm an octameric assembly as a conserved feature of NrnC-type RNases. The substrate-bound states reveal, similar to Orn, a narrow active site that appears optimized for dinucleotides. This preference is reflected also in the enzyme’s activity profile. A comparative genomics analysis indicates that Orn and NrnC, despite using a common DnaQ fold, evolved separately as isofunctional enzymes (Galperin & Koonin, 2012; Omelchenko et al., 2010). Considering also the distribution of the structurally unrelated, yet functionally overlapping NrnA- and NrnB-type RNases predominantly in organisms that lack Orn and NrnC underlines the importance to maintain diribonucleotidase activity for cellular function in the bacterial and eukaryotic domains of life.

4.4 – Results and Discussion: Overall structure of diribonucleotide-bound NrnC

The earlier observation that NrnC expression is able to complement a P. aeruginosa orn deletion strain indicates that both enzymes function on diribonucleotides (Orr et al., 2018), consistent with their initial classification as nanoRNases (Datta & Niyogi, 1975; Liu et al., 2012; Niyogi & Datta, 1975; Yu & Deutscher, 1995). However, while we showed that Orn acts as a dedicated diribonucleotidase (Kim et al., 2019), the structural basis for NrnC's substrate preference remained not well defined. Specifically, it was not clear how NrnC distinguishes between short RNAs and longer polymers. To answer this question, we determined crystal structures of wildtype B. henselae and B. melitensis NrnC (NrnCBh and NrnCBm, respectively) bound to pGG and, in the case of NrnCBm, also in the substrate-free state. NrnCBh forms a homo-octameric assembly comprised of two C4-symmtric rings (Figure 4.1A) as observed with the previously determined substrate-free, orthologous protein from A. tumefaciens (69% sequence identity compared to NrnCBh, monomer/octamer all-atom rmsd 0.4/1.3 Å; PDB:5ZO3, (Yuan et al., 2018)). The two rings stack with the same face, tail-to-tail, forming a D4-symmetric octamer. The contacts between the subunits within each ring are dominated by a few polar interactions spanning between 441 to 512 Å2 (Figure 4.1 and 4.2 A-B) (determined by PISA (Krissinel & Henrick, 2007)). In contrast, pairwise, homotypic interactions between the two rings involve an extensive hydrophobic interface of 1204 Å2 via an antiparallel packing of the last helix of the NrnC fold. This mode of ring stacking positions the C-terminus of one monomer so that it reaches into the active site of the adjacent monomer in the other ring (Figure 4.1 and 4.2 A and C).

The diribonucleotide pGG is bound to all eight active sites of the NrnCBh octameric assembly (Figure 4.1B). The active sites face the center of the central cavity formed by the NrnC octamer, positioned mid-way of each ring. In this crystallographic state, the residues of the catalytic DEDDy signature motif (D25 , E27, D84, D155, Y151) are primed for accepting divalent metal ions for nucleotide hydrolysis, further supported by the observation that side chain of Y151 coordinates a water molecule, which likely serves as the attacking nucleophile in the reaction (Figure 4.1C).

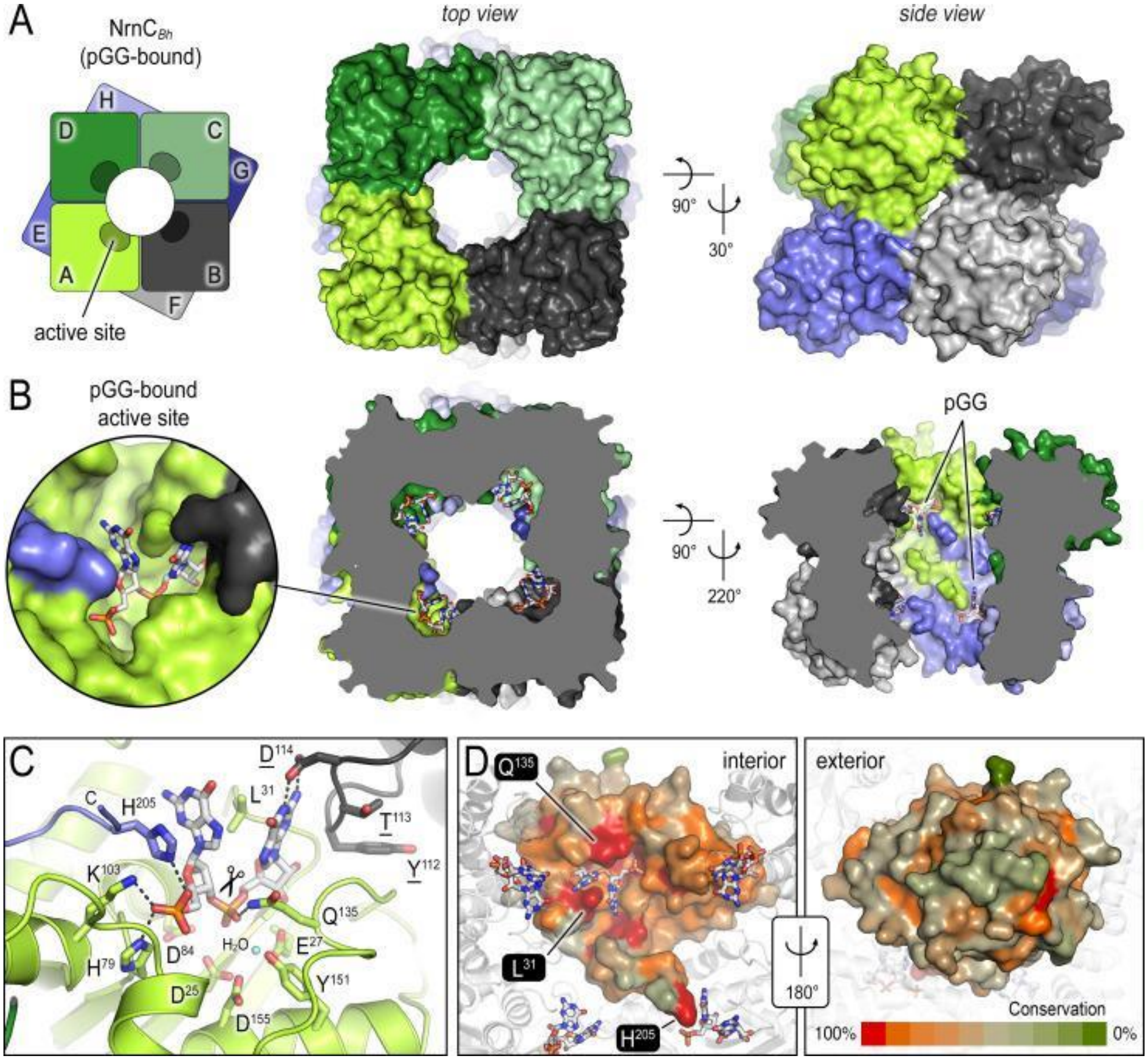


Figure 4.1 – The crystal structures of *B. henselae* NrnC bound to pGG reveals motifs defining substrate specificity. (**A**) The octameric assembly. NrnCBh is shown as surface representation in two views. Each monomer is shown in a distinct color. The cartoon illustrates the stacking of the two tetrameric NrnC rings that form the octamer with a central, round opening. (**B**) Active-site position. Each monomer contributes one active site, here bound to the substrate pGG, facing toward NrnC's central pore. Each active site includes a C-terminal tail of a subunit from an adjacent ring. (**C**) Substrate coordination. The catalytic DEDDy motif and residues coordinating each moiety of pGG contacts are shown as sticks, with carbon residues colored according to monomer identity. Residue Y151 coordinates water molecule near the scissile bond. (**D**) Conservation mapping on a surface representation of a NrnC monomer. Conservation scores were calculated based on a multisequence alignments (MUSCLE; (Edgar, 2004)) of NrnC homologs identified using a sequence search on the EggNOG resource, version 5.0.0 (Huerta-Cepas et al., 2019) and the sequence of NrnCBh as the input. Outliers were identified based on sequence length and non-consensus insertions, resulting in a final collection of 560 sequences of putative NrnC orthologs. The two views, separated by a 180º rotation, show the cavity-facing (interior, left) and outer-facing (exterior, right) surface regions.

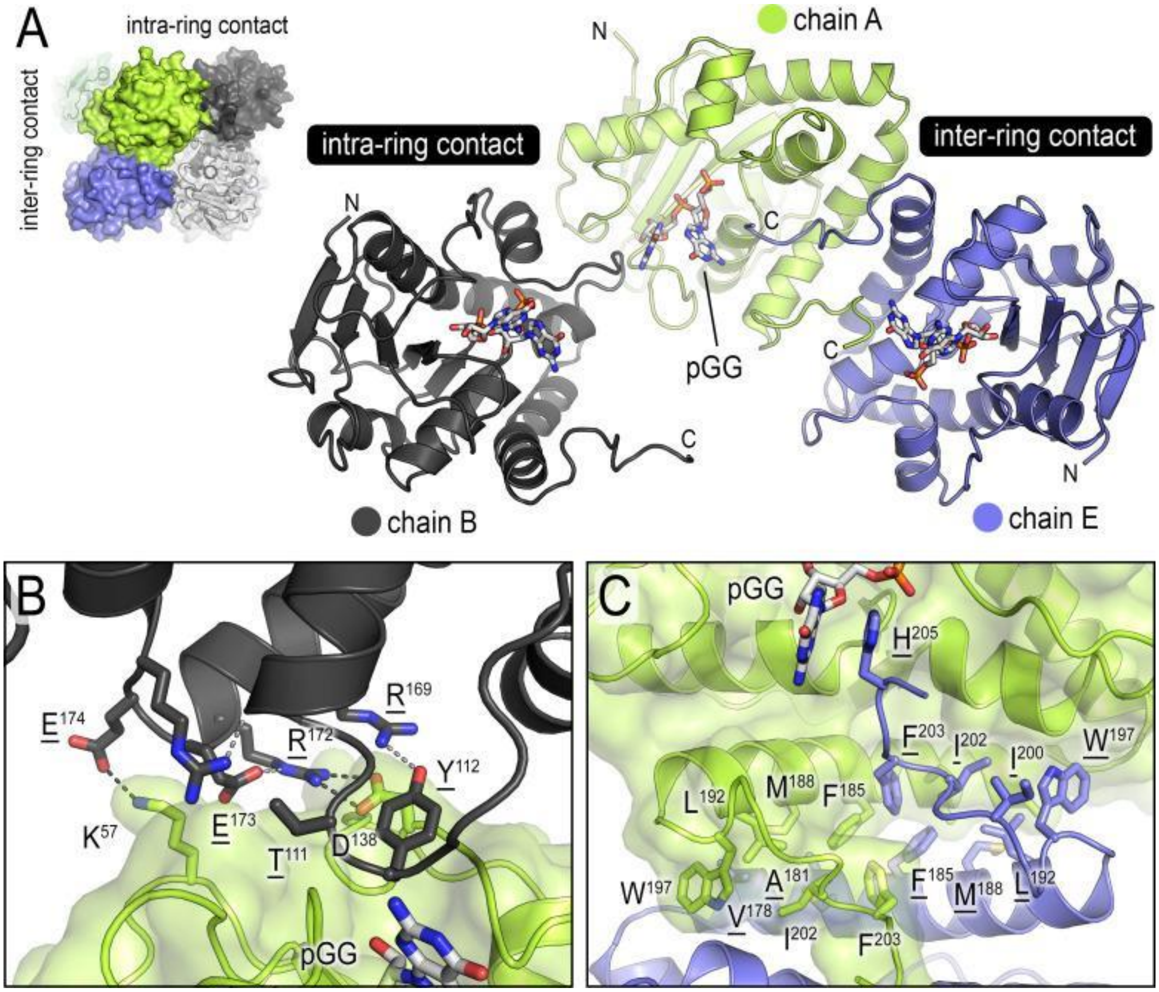


Figure 4.2 – Inter- and intra-ring contacts in the NrnCBh octamer. (**A**) Overview of three pGG-bound NrnCBh monomers. The monomers are shown in cartoon representation The chain colored in dark gray forms an intra-ring contact with the central, green-colored chain, whereas the chain colored in blue forms a representative inter-ring contact within the octameric nano-RNase C (NrnC). (**B**) Detailed intra-ring interface. (**C**) Detailed inter-ring interface. Residues contributing direct interactions between monomers are shown as sticks. Representative hydrogen bonds are shown as dashed lines.

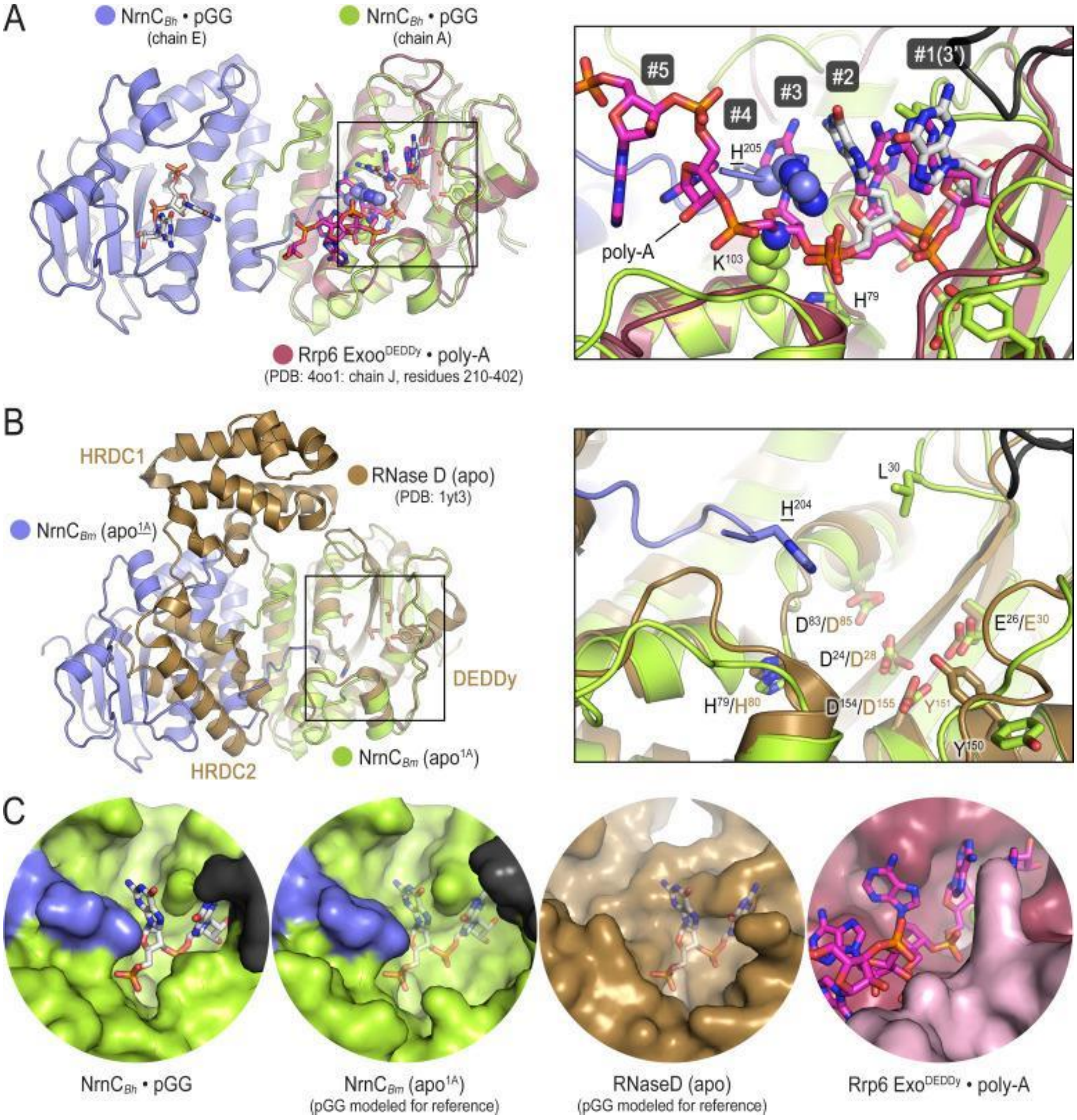


Figure 4.3 – Comparison of nano-RNase C (NrnC) to structurally related proteins reveals the constricted nature of NrnC's active site. (**A**) A pGG-bound NrnCBh inter-ring dimer (slate and green chains, with pGG carbon atoms shown in white) was superimposed on a structure of the exosome's Rrp6 exonuclease (purple chain) bound to poly-A (pink carbon atoms; PDB:4oo1; Wasmuth et al., 2014). The inset shows a detailed view of the superimposed active sites. Numbering refers to the residues in the substrate. Phosphate cap residues of NrnC that block the path of poly-A substrate are shown as spheres. The black protein chain in top-right corner stems from an adjacent intra-ring monomer. (**B**) A substrate-free NrnCBm inter-ring dimer (slate and green chains) was superimposed on a structure of apo-RNase D (Zuo et al., 2005), highlighting conservation of the catalytic DEDDy motif and differences in regions around NrnC's L-wedge and phosphate cap (inset). (**C**) Surface views of the active sites of NrnC, RNaseD, and Rrp6 accentuate the constraint of the NrnC active site. Translucent pGG represents modeled substrate as opposed to co-crystallized substrate.

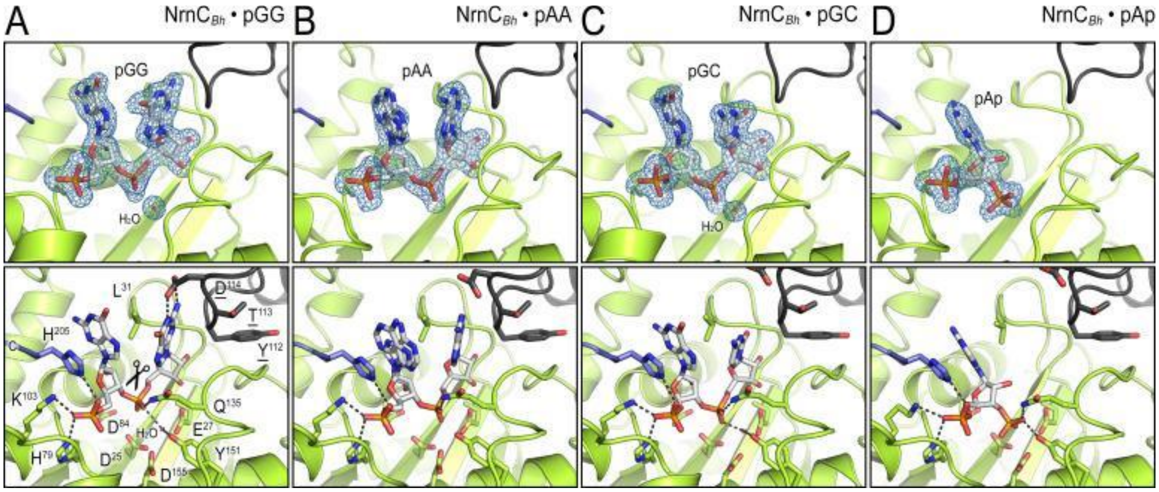


Figure 4.4 Structural comparison of NrnCBh bound to various ribonucleotides. (**A**) Diribonucleotide pGG-bound NrnCBh, identical to the structure shown in Figure 1. (**B**) Diribonucleotide pAA-bound NrnCBh. (**C**) Diribonucleotide pGC-bound NrnCBh. (**D**) Adenosine-3′,5′-bisphosphate (pAp)-bound NrnCBh. The top row of images shows nano-RNase C (NrnC) as a cartoon representation with the nucleotide substrate represented as sticks with carbon atoms colored white. Polder omit maps for each substrate are shown as blue mesh. The bottom row shows detailed views of the active site residues contacting each ribonucleotide.

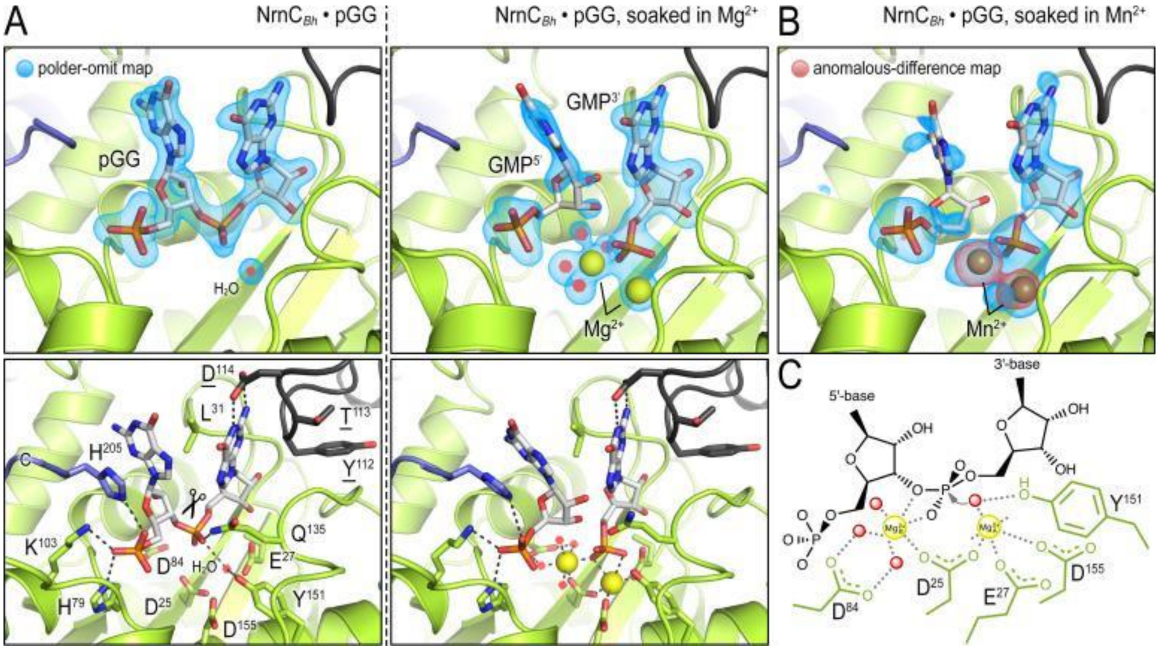


Figure 4.5 – In crystallo catalysis indicates a two-metal mechanism of NrnCBh activity. (**A**) Active sites of NrnCBh•pGG, before and after soaking crystals in a solution containing Mg2+ prior to data collection. (**B**) Active site of NrnCBh•pGG, after soaking crystals in a solution containing Mn2+ prior to data collection. Top panels show polder omit maps highlighting nucleotide and metal density. The red density in (B) represents an anomalous-difference map calculated from data collected at the Mn2+ absorption edge. The bottom panels show specific active site contacts between protein, nucleotide, ions, and water molecules. (**C**) Schematic overview of two-metal coordination at the active site of nano-RNase C (NrnC).

The structure also reveals the molecular basis for substrate coordination. Reminiscent of Orn's active site, the bases of the diribonucleotide are splayed apart by a leucine residue, or Lwedge (L31) (Figure 4.1C). Continuing with the parallels to Orn, the 5' phosphate of pGG is coordinated by several residues forming a 'phosphate cap', in this case basic residues H79, K103, and H205, the latter being the second to last residue of the protein, contributed from a subunit from the adjacent ring. The specific motifs coordinating the substrate are invariable in NrnC orthologs, in contrast to the exterior surface with overall lower conservation (Figure 4.1D), and culminate in a length-restricted active site that appears optimized for diribonucleotides. The comparison with the structurally related exosome subunit Rrp6 bound to a longer RNA substrate supports this notion as residues K103 and H205 of NrnC's phosphate cap directly block the path for longer substrates (Figure 4.1 and 4.3 A and C) (Wasmuth et al., 2014). Similarly, RNase D, a homologous exoribonuclease that processes longer and stable RNA molecules, presents a more expansive, open active site, although the lack of a substrate-bound structure prevents a more direct comparison of this state (Figure 4.1 and 4.3 B and C) (Zuo et al., 2005).

Similar binding poses to pGG at NrnC were observed with pAA and pGC (Figure 4.1 and 4.3 A-C), suggesting that most if not all diribonucleotides can be accommodated by NrnC. A co-crystal structure with the di-phosphorylated mononucleotide pAp, a metabolite described as an inhibitor of NrnC (Liu et al., 2012), shows the ligand predominantly occupying the 5' site of the active site with the 3' phosphate engaging the catalytic motif and the 5' phosphate being coordinated by the phosphate cap of NrnC (Figure 4.1 and Figure 4.4 D).

To establish the functional relevance of the crystallized state of NrnC, we initiated in crystallo catalysis by soaking NrnCBh•pGG co-crystals in solutions with divalent cations, magnesium (Mg2+) or manganese (Mn2+). The enzyme became catalytically active with the addition of either cation, with the resulting electron densities showing a broken phosphodiester bond (Figure 4.1 and 4.5). The experiments also confirmed a two-metal mechanism first by interpretation of electron density upon Mg2+ soaking (Figure 4.1–4.5A). The placement of active site metal atoms was subsequently confirmed using anomalous data collected on Mn2+-soaked crystals (see anomalous difference map, (Figure 4.1 and 4.5B). In these post-catalysis structures, the 5’ GMP appears to leave the active site first, as suggested by weaker electron density indicative of lower mononucleotide occupancy at the 5’ site compared to the 3’ site.

4.5 – Results and Discussion: The characteristic active-site motifs of NrnC$_{Bh}$ contribute to diribonucleotidase activity

The structural analysis revealed the molecular basis for substrate binding to NrnC, identifying molecular features that constrain the active site. To assess their relevance for NrnC’s catalytic activity, we tested tag-less, purified NrnCBh and structure-based point mutants thereof in an *in vitro* activity assay. All mutant proteins retained their quaternary structure and purified as octamers, indistinguishable from wild-type NrnC (Figure 4.6 and 4.7). 5’- 32Pradiolabeled pGG was incubated with wild-type or mutant enzymes in the presence of divalent cations at physiological ionic strength. Quenched reactions were resolved via urea-denaturing PAGE to observe nucleolytic cleavage over time (Figure 4.6A). With wild-type NrnC$_{Bh}$ the majority of pGG was processed already by the first timepoint at 30 seconds, and complete cleavage of pGG to GMP was achieved by 3 minutes. Nucleolytic activity on pGG was completely inhibited by alanine mutation of the catalytic DEDDy motif residues D25 and Y151. Intermediate cleavage kinetics were evident with proteins with disrupted leucine wedge (NrnCBh-L31A) as well as the phosphate cap (NrnCBh-H79A, -K103 A, or -H205A). NrnC binding to and activity on pGG were 82 slightly inhibited by pAp (Figures 4.6B and C). In contrast, the diguanylate compound GG, which lacks a 5' phosphate, did not inhibit NrnC's binding to or activity on pGG (Figures 4.6B and C), further elaborating on the importance of the interaction between phosphate cap residues and the 5' phosphate for NrnC function.

We previously proposed a model of cellular fitness in which the loss of orn leads to toxic diribonucleotide accumulation detrimental to the cell. The orn deletion in P. aeruginosa manifests as small colony growth, which is reversible by complementation with orn or other nanoRNases expressed from a plasmid (Figure 4.6D) (Kim et al., 2019; Orr et al., 2018). Here we used the rescue of the small colony phenotype as a readout of NrnC diribonucleotidase activity in cells, quantified as colony size. Complementation of the deletion strain with wild-type NrnCBh (with a C-terminal HA tag for detection) restored normal colony size, while NrnC alleles containing mutations within the DEDDy motif failed to complement (Figure 4.6D). Further, NrnC alleles containing mutations in the L-wedge or the phosphate cap showed reduced complementation effects (L31A, K103A, K132A, or H205A; Figure 4.6D). NrnCBh-H79A failed to complement the orn deletion. Western blot analysis established protein expression for all mutant and wild-type NrnC variants, with the exception of NrnCBh-H79A, which expressed poorly in P. aeruginosa, preventing a distinction between failure to rescue because of the mutation or protein levels, or both (Figure 4.6 and 4.8). Together, these data confirm the importance of the motifs identified in the substrate-bound NrnC structures for the enzyme's diribonucleotidase activity *in vitro* and in cells.

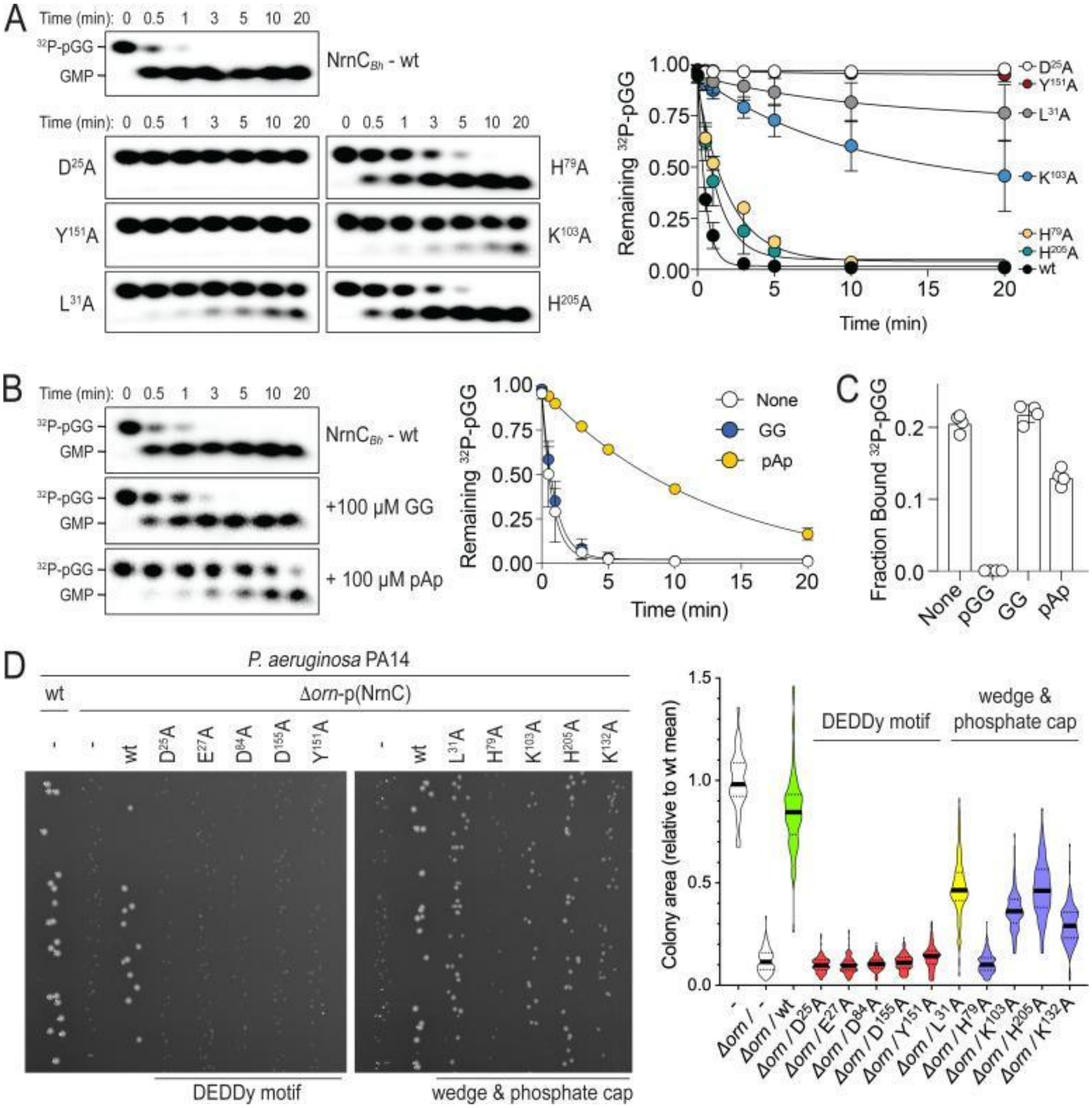


Figure 4.6 – Phosphate cap and L-wedge contribute to NrnC's diribonucleotidase activity. (**A**) Degradation of 32P-pGG (1 µM total) by purified wild-type NrnCBh or variants with alanine substitutions (5 nM). Samples were stopped at the indicated times (min) and analyzed by denaturing 20% PAGE. Representative gels are shown (left). The graph (right) shows the means and SD of three independent experiments. (**B**) Effect of a dinucleotide lacking the 5' phosphate (GG) and pAp on NrnC catalysis. pGG processing was assessed as in (A) but in the presence or absence of 100-fold excess (over 32P-pGG) GpG or pAp. Representative gels (left) and quantification from three independent experiments (right) are shown. Means and SD are plotted. (**C**) Competition binding studies - Fraction bound of 32P-pGG to 200 nM purified NrnCBh in presence of 100 µM competitor is plotted as individual data, means, and SD of four independent experiments. (**D**) Complementation of the small-colony phenotype of P. aeruginosa ∆orn by wild-type and mutant NrnCBh. Bacterial cultures were diluted and dripped on LB agar plates. After overnight incubation, representative images of the plates were taken (left). Experiments were performed in triplicate. (**Continued…**)

Figure 4.6. Phosphate cap and L-wedge contribute to NrnC's diribonucleotidase activity. (D, Continued) Quantification of respective colony sizes are shown as violin plots (right). Panels A,B,C: data and illustration provided by Soo-Kyoung Kim (Lee lab, University of Maryland).

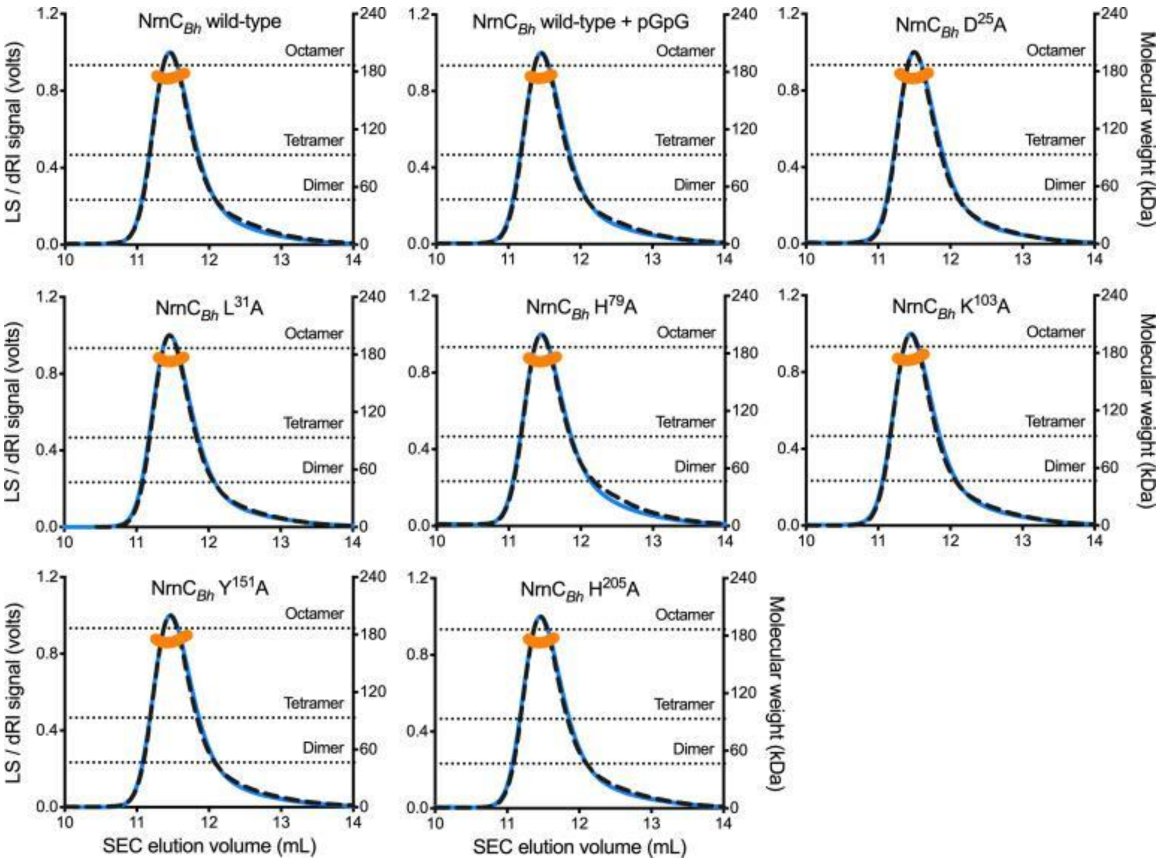


Figure 4.7 – SEC-MALS of NrnCBh wild-type and mutant variants. Molecular weight determination indicates that pGG binding does not impact oligomerization, and that purified NrnCBh point mutants remain octameric in solution. Absolute molecular weights of nano-RNase C (NrnC) are shown as orange data points across elution peaks plotted on the right axis. Theoretical oligomerization states are shown as dashed horizontal lines. 90°-light scattering: blue solid lines; refractive index signal: black dashed lines; plotted on left axis.

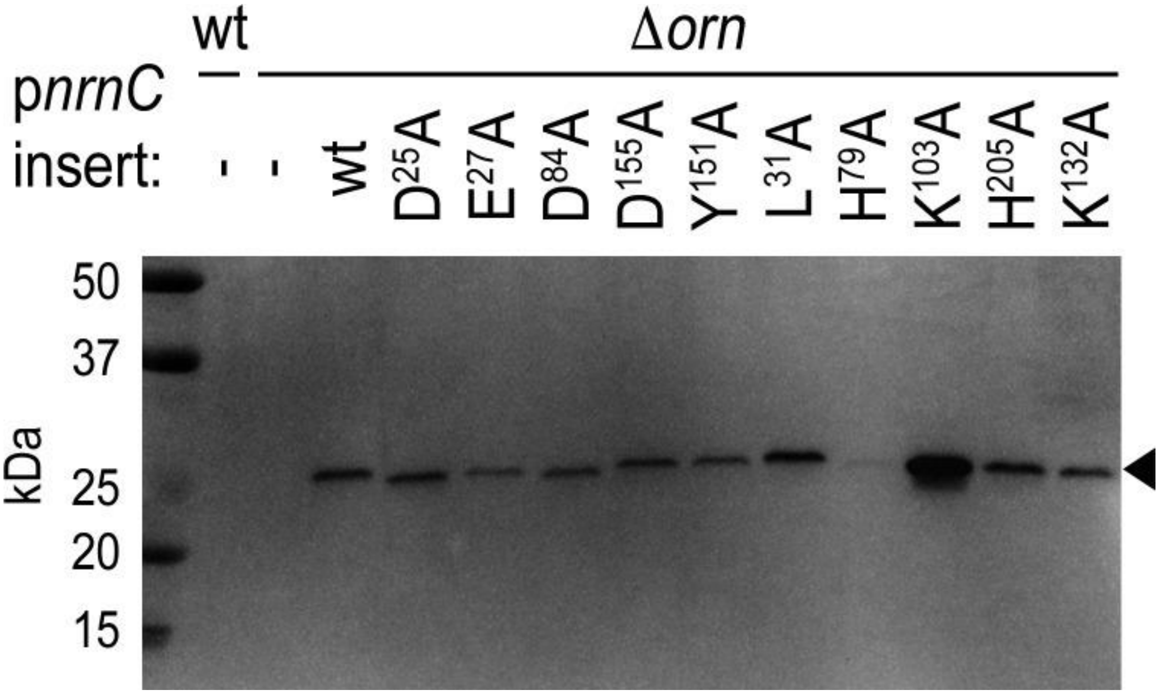


Figure 4.8 – Expression of NrnCBh wild-type and mutant variants in P. aeruginosa ∆orn. Cell lysates were analyzed by western blotting, detecting the C-terminal HA-tag in recombinantly expressed NrnCBh.

4.6 – Results and Discussion: Structural comparison of NrnC substrate-bound states reveals an active site optimized for diribonucleotides

To further understand the structural basis of NrnC's substrate preference, we determined structures of B. melitensis NrnC (NrnCBm) in several apo and partially-pGG-bound states. The overall structure of pGG-bound NrnCBm is virtually identical to NrnCBh, namely a homo-octameric assembly with eight active sites pointing toward the central channel (Figures 4.9 A and B). The acidic active site residues (D24, E26, D83, D154, and Y150) as well as the L-wedge (L30) and phosphate cap (H78, K102, and H204) are structurally and functionally conserved. NrnCBm incubated with pGG crystallized with two molecules per asymmetric unit, only one of which was bound to the diribonucleotide. The resulting octameric assembly contains alternating apo- and pGG-bound subunits per tetrameric ring (Figure 4.9B). This mixed-state structure allowed us to propose features modulating substrate binding. In substrate-bound monomers (including in the structures of NrnCBh), the DEDDy residue Y150 points inward toward the scissile phosphodiester bond, coordinating the attacking water. In contrast, the Y150 side chain points away from the active site in substrate-free NrnC (Figures 4.9A and B). A loop from an adjacent subunit that mediates inter-ring contacts between the monomers and that includes residue D113 buttresses the 3' base of the substrate. While the octameric assembly remains in the absence of substrate, this loop moves outward from the active site and D113 rotates away from the substrate (Figure 4.9D). The most drastic conformational change however is attributed to a flexible loop spanning residues 130SKQQQS135. This loop is ordered and positioned in contact with the diribonucleotide in both ortholog structures (Figures 4.1C, 4.9B, and 4.9C). In this state, residue Q134 of NrnCBm (or Q135 in NrnCBh) contacts the scissile phosphate via hydrogen bond; K131 (or K132 in NrnCBh) points towards the 5' phosphate, thus contributing to the phosphate cap. In 85 contrast, in the substrate-free state NrnCBm, captured in the mixed-state structure or a homogeneous apo-state structure, this loop swings away from the active site or is completely disordered, leading to an overall widening of the active site (Figure 4.9A; Figure 4.2C).

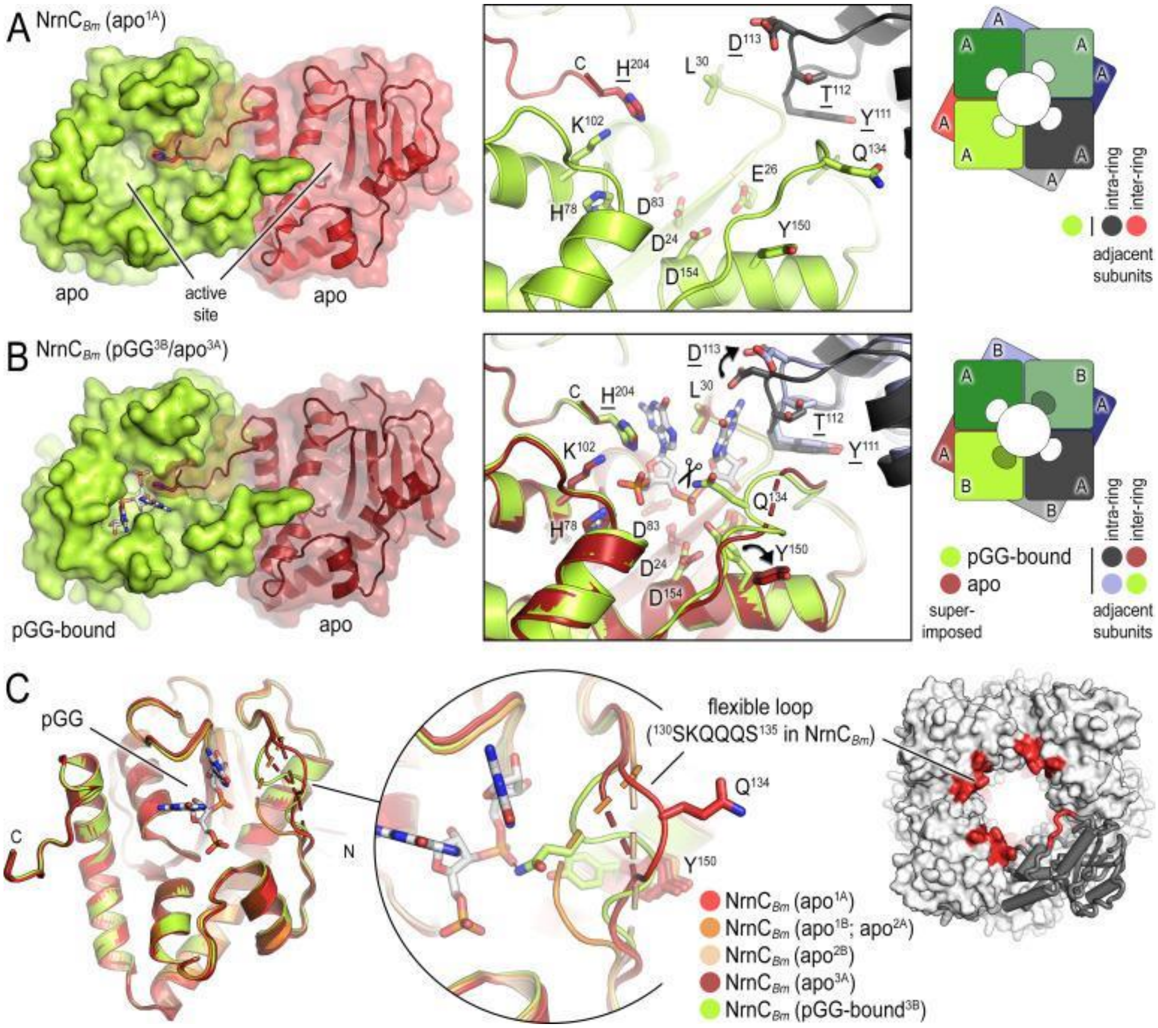


Figure 4.9 – Brucella melitensis NrnC crystal structures reveal a flexible loop that constraints the enzyme's active site. (**A**) Crystal structure of apo-NrnCBm. A crystallographic dimer as part of the octameric assembly is shown as surface presentation (left) and close-up of the active site (middle). The diagram (right) depicts the octamer and the spatial relationship of the monomers shown. (**B**) Crystal structure of NrnCBm with alternating substrate-bound and empty active sites. The close-up (middle) shows a superposition of the two monomers in the asymmetric unit, depicting their conformational difference and adjacent monomers, with intra- and inter-ring neighbors colored as shown in the diagram (right). (**C**) Superposition of four apo-NrnCBm conformations based on three independent crystal forms, compared to the pGG-bound conformation of the same protein shown in (B). The position of the flexible loop (red) in the NrnC octamer is shown (right panel).

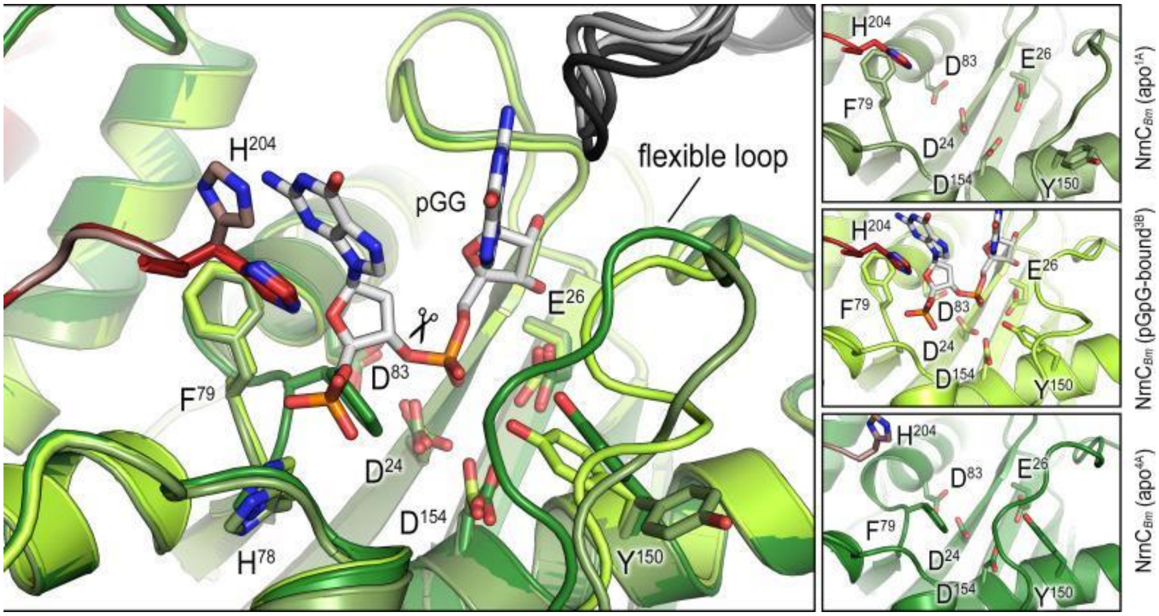


Figure 4.10 – Overlay of an alternative crystallographic apo-NrnCBm state with the apo- and pGG-bound states observed in the crystal structure shown in Figure 4.9. The large panel shows three structures superimposed. The smaller panels on the right show each active site isolated with residues of interest labeled and shown as sticks.

In a second apo-NrnCBm crystal form, we captured an alternate conformation in which the conserved hydrophobic residue F79 moves into the active site (Figure 4.10). The movement of this residue, which is adjacent to the phosphate cap residue H78, is realized through a flip of the peptide backbone and would introduce a clash with the nucleotide. In the same structure the flexible SKQQQS loop is collapsed into the active site, trapping the catalytic residue Y150 in an intermediate conformation, which would introduce further clashes with the nucleotide substrate. Simultaneously, a rotamer change of the phosphate cap residue H204 pivots its sidechain away from the active site, opening it for access to substrates (Figure 4.10). This conformation may depict a post-hydrolysis state, suggesting a mechanism for product release.

4.7 – Results and Discussion: Cryo-electron microscopic analysis supports a narrow substrate preference of NrnC.

Efforts to crystallize NrnC bound to longer substrates yielded structures with either empty active sites or with only pGG being resolved in the resulting electron densities. Residual RNase activity over the course of the crystallization or an impact of the longer substrates on crystal packing could contribute to the inability to resolve longer substrates, assuming they bind in the first place. As an alternative to crystallography, we determined NrnCBh structures bound to pGG, pAGG and pAAAGG by cryo-electron microscopy (cryoEM), a technique that can visualize complexes formed after short equilibration periods and does not rely on proteins packing in a lattice. Considering that NrnC forms an octamer as the biological unit, we processed each data set with C1 and D4 symmetry, with the resulting models consisting of 8 independent or an averaged chain, respectively (resolutions range from 2.72 to 3.27 Å) (Figures 4.11-4.19). Processing with C1 symmetry preserves the individuality of each monomer allowing the observation of differences e.g., between active sites within the octamer. Applying D4 87 symmetry during processing averages all 8 monomers, yielding a consensus model; however, regions with conformational differences between individual monomers may contribute to apparent disorder in the electron-density maps. In an additional experiment with the 5-mer pAAAGG as the substrate, CaCl2 was added to the buffer to probe whether the addition of Ca2+ ions, which prevent catalysis but can support substrate binding, has a qualitative effect on the binding of longer RNAs. The structure of pGG-bound NrnCBh confirmed all active site features described above based on the crystallographic data, namely a narrow active site, a phosphate cap coordinating the 5’ phosphate of the substrate, a L-wedge splaying apart the two bases, and an active-site facing, well-resolved SKQQQS loop (Figure 4.11A, Figure 4.20A). Refinement with lower and higher symmetry resulted in comparable density maps, indicating a consensus state with 8 nearly identical active sites. The density maps of NrnC bound to any of the longer substrates resolved invariably only a diribonucleotide at the active site (with poorer density indicating the position of the ribose of the third 5'-substrate residue) (Figure 4.11B). The remainder of the longer substrates appeared disordered. Using D4-symmetry-averaged data, we also noticed consistently disorder of the SKQQQS loop (Figure 4.11B). Inspection of symmetry-less (C1) density maps revealed different loop conformations in the individual monomers of the octamers that, when symmetry-averaged, result in the apparent disorder (Figure 4.11, 4.16-4.19). Notably, the majority of monomers contain a disengaged loop conformation, leaning away from the active site. Addition of Ca2+ with the pAAAGG substrate results in increased ordering of the loop in a disengaged state similar to that observed in the substrate-free crystal structures of NrnCBm (Figures 4.9E, Figure 4.20). Together, these results suggest that longer substrates may bind NrnC, but only the first two 3' residues are well coordinated at the active site.

Furthermore, RNAs with more than two residues in length increase conformational variability at the active site, likely impacting catalytic activity towards those substrates. In summary, the combined structural data indicate NrnC is optimized for diribonucleotide processing over longer substrates, mirroring our analysis of Orn and REXO2 (Kim et al., 2019; Nicholls et al., 2019).

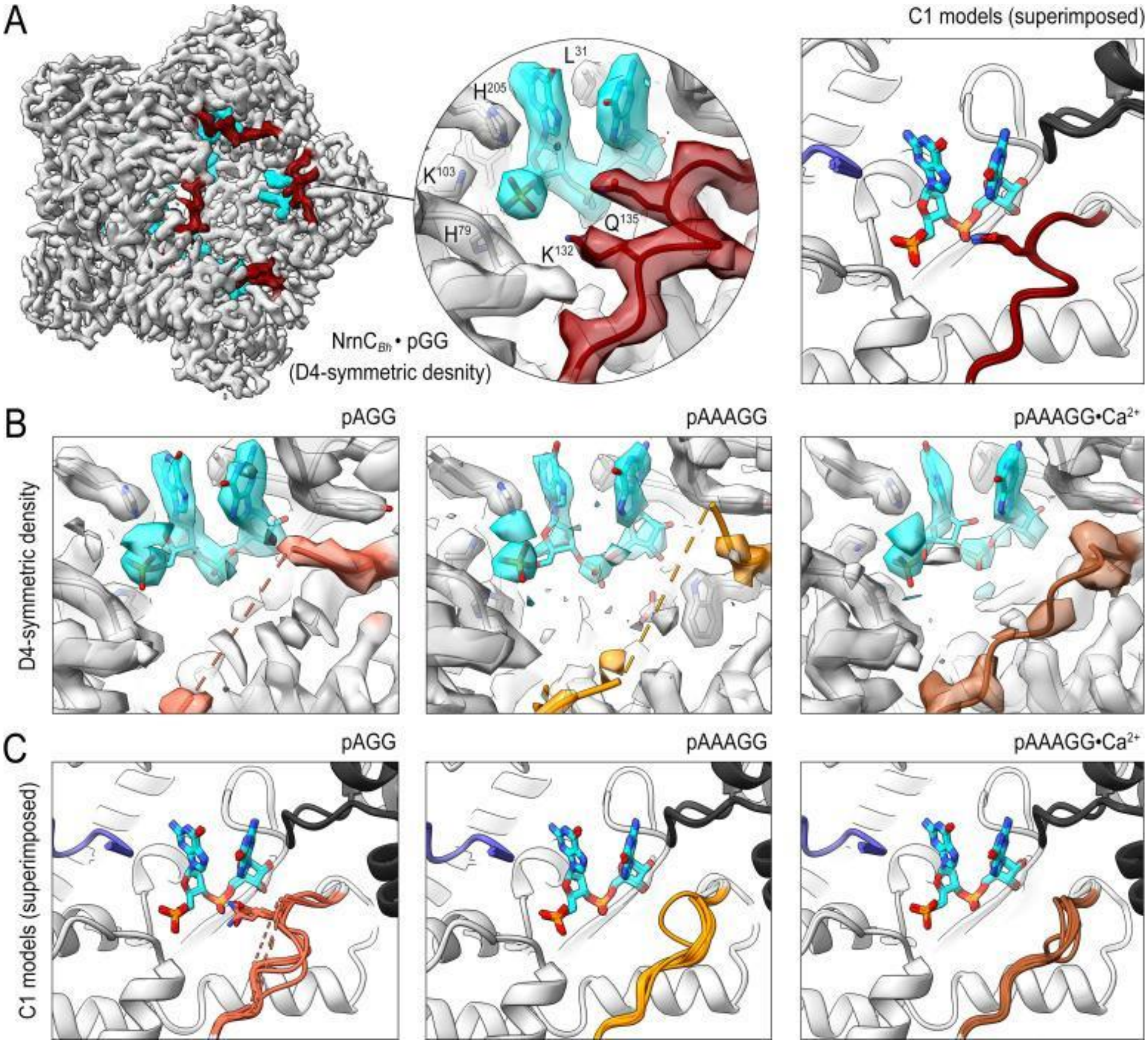


Figure 4.11 – Cryo-EM structures of NrnCBh with 2-, 3-, 5-mer RNA substrates show substrate lengthdependent active site conformations. (**A**) Electron density map of a NrnCBh octamer in complex with pGG. D4 symmetry was applied during final map refinement. pGG molecule and density are colored cyan. The SKQQS-containing loops (residues 130-137) are colored maroon. Superposition of all eight active sites from a reconstruction with C1 symmetry (right panel) shows consensus order in the loop when bound to pGG. (**B**) Active site images shown for NrnCBh incubated with 3-mer and 5-mer (with or without Ca2+) RNA substrates. Regions corresponding to those shown in (A) are shown in color. D4-symmetric maps are shown. (**C**) Superposition of all eight active sites from octamer reconstructions based on respective C1- symmetric maps for each RNA substrate.

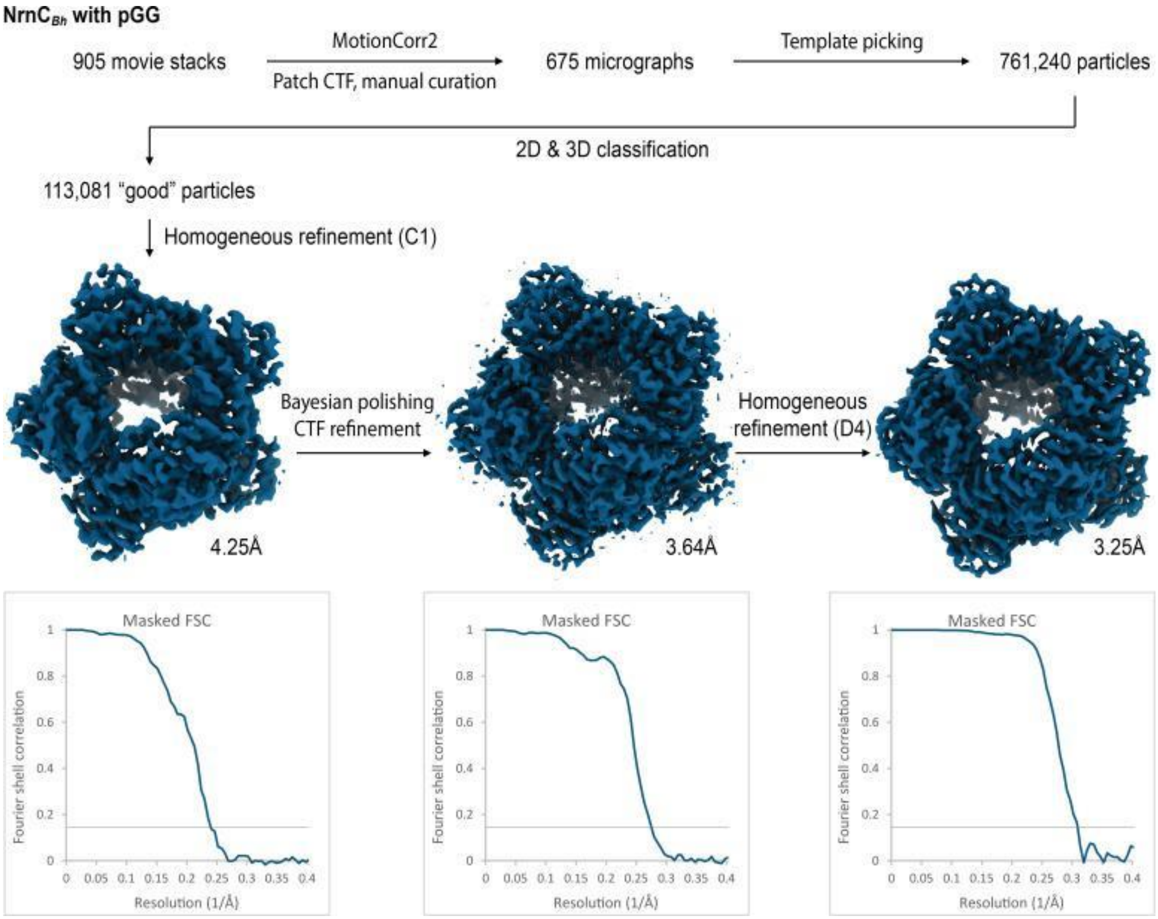


Figure 4.12 – Cryo-electron microscopy (cryo-EM) workflow and resolution for NrnCBh•pGG.

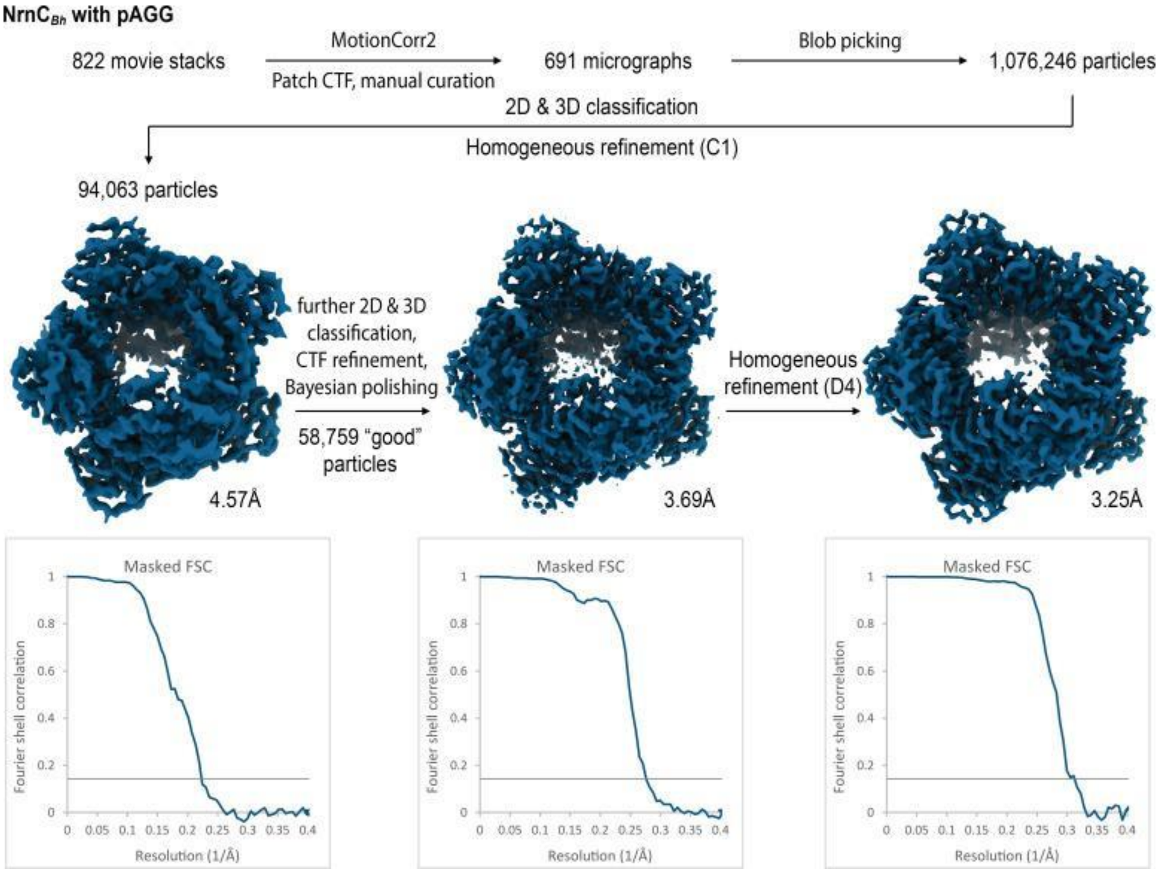


Figure 4.13 – Cryo-electron microscopy (cryo-EM) workflow and resolution for NrnCBh•pAGG.

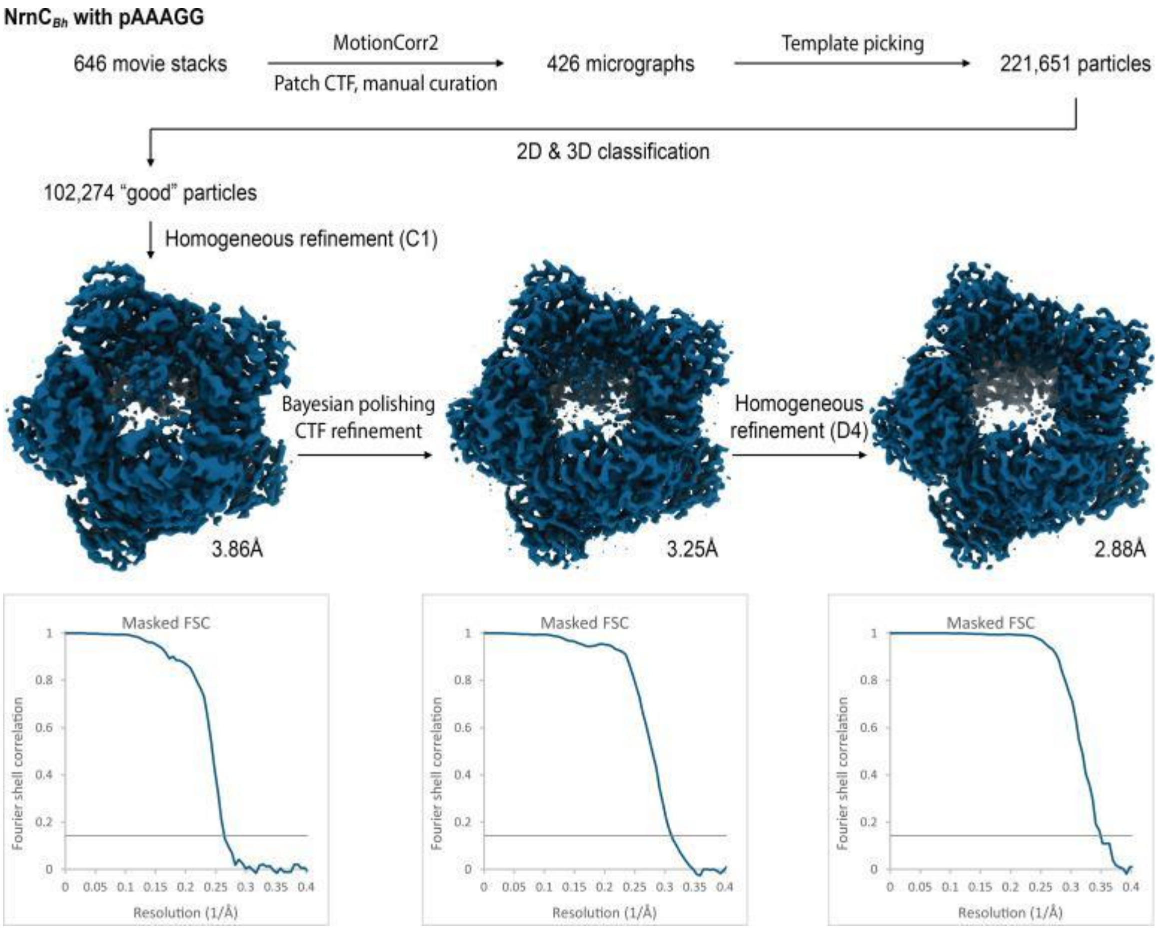


Figure 4.14 – Cryo-electron microscopy (cryo-EM) workflow and resolution for NrnCBh•pAAAGG.

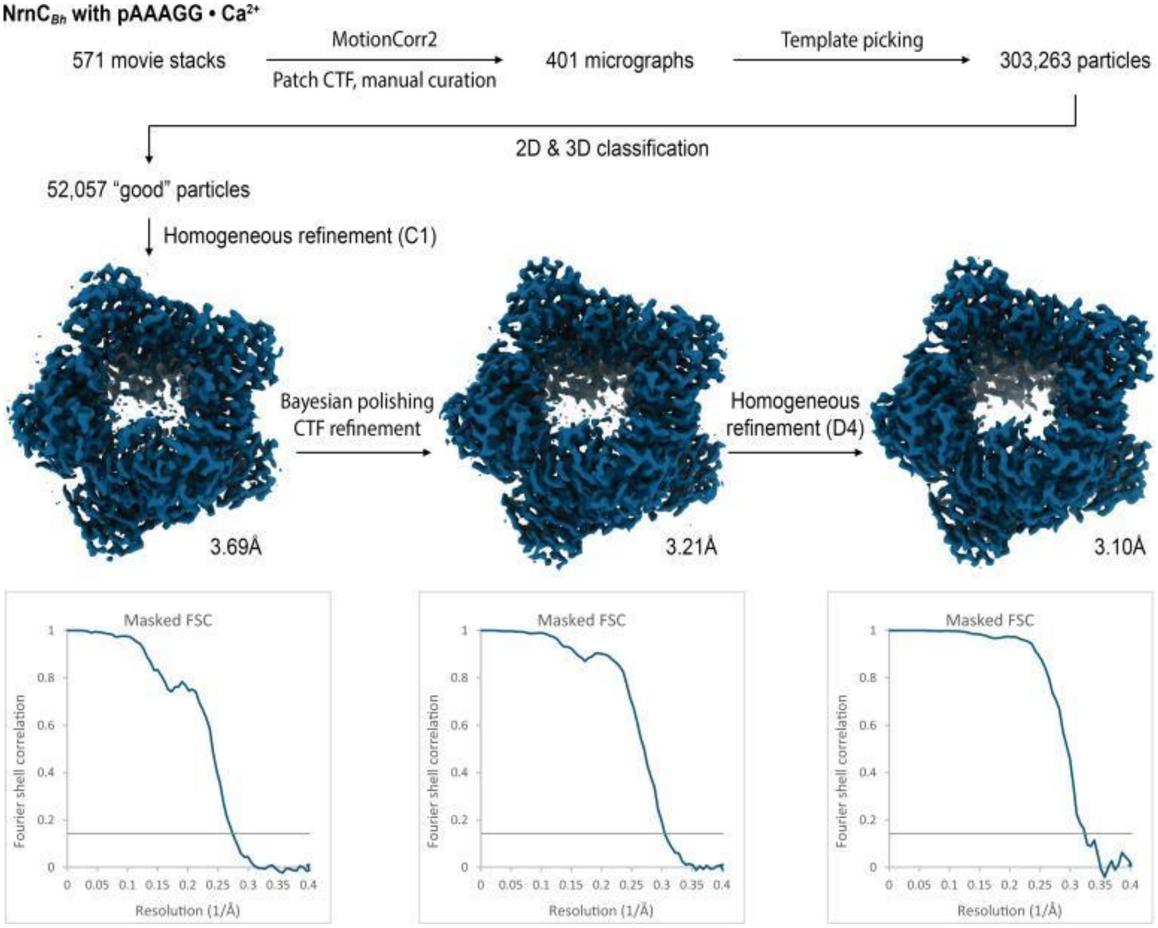


Figure 4.15 – Cryo-electron microscopy (cryo-EM) workflow and resolution for NrnCBh•pAAAGG in the presence of Ca2+ ions.

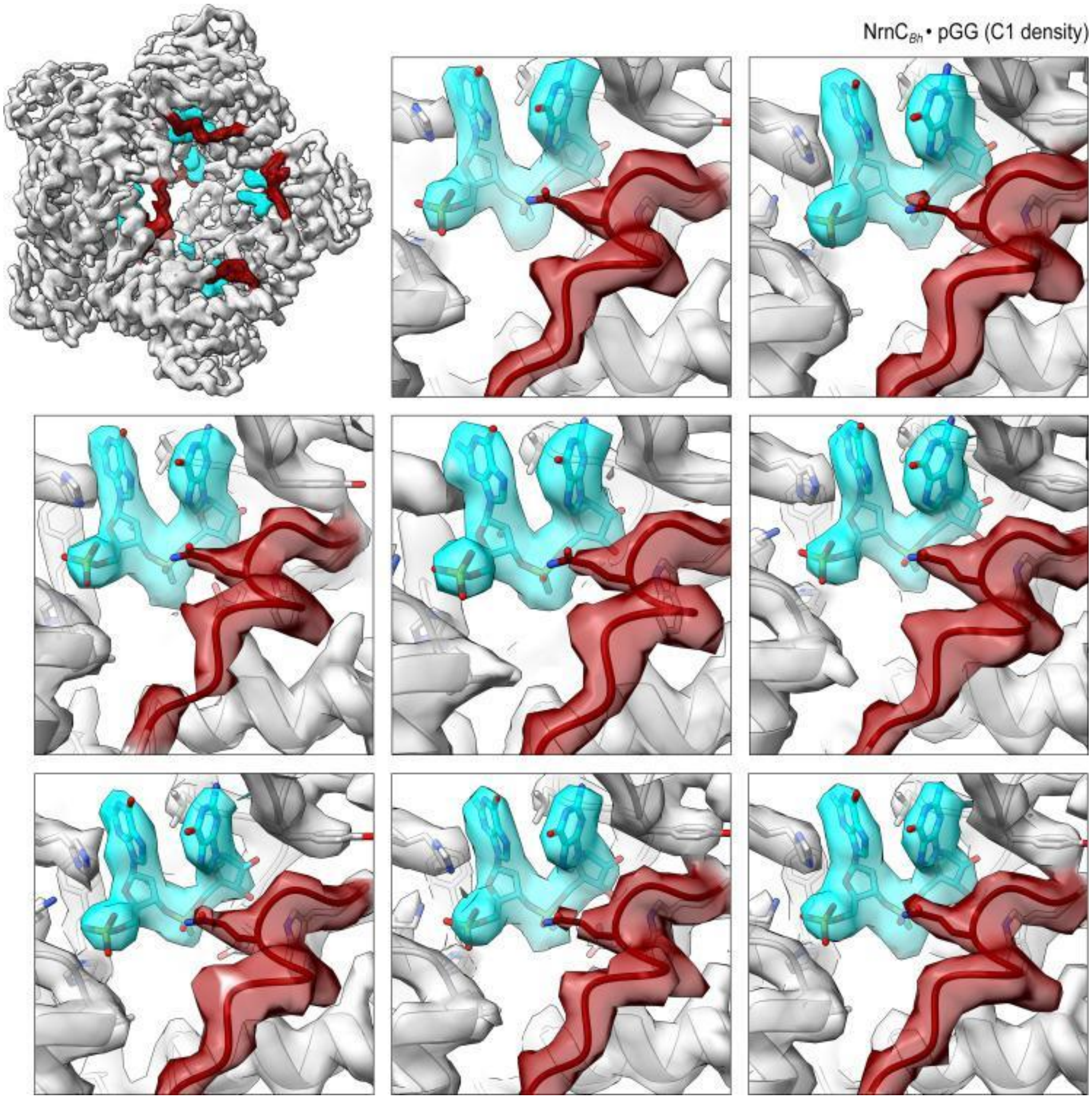


Figure 4.16 – Overall and individual active site electron density of a NrnCBh•pGG octamer after refinement with C1 symmetry.

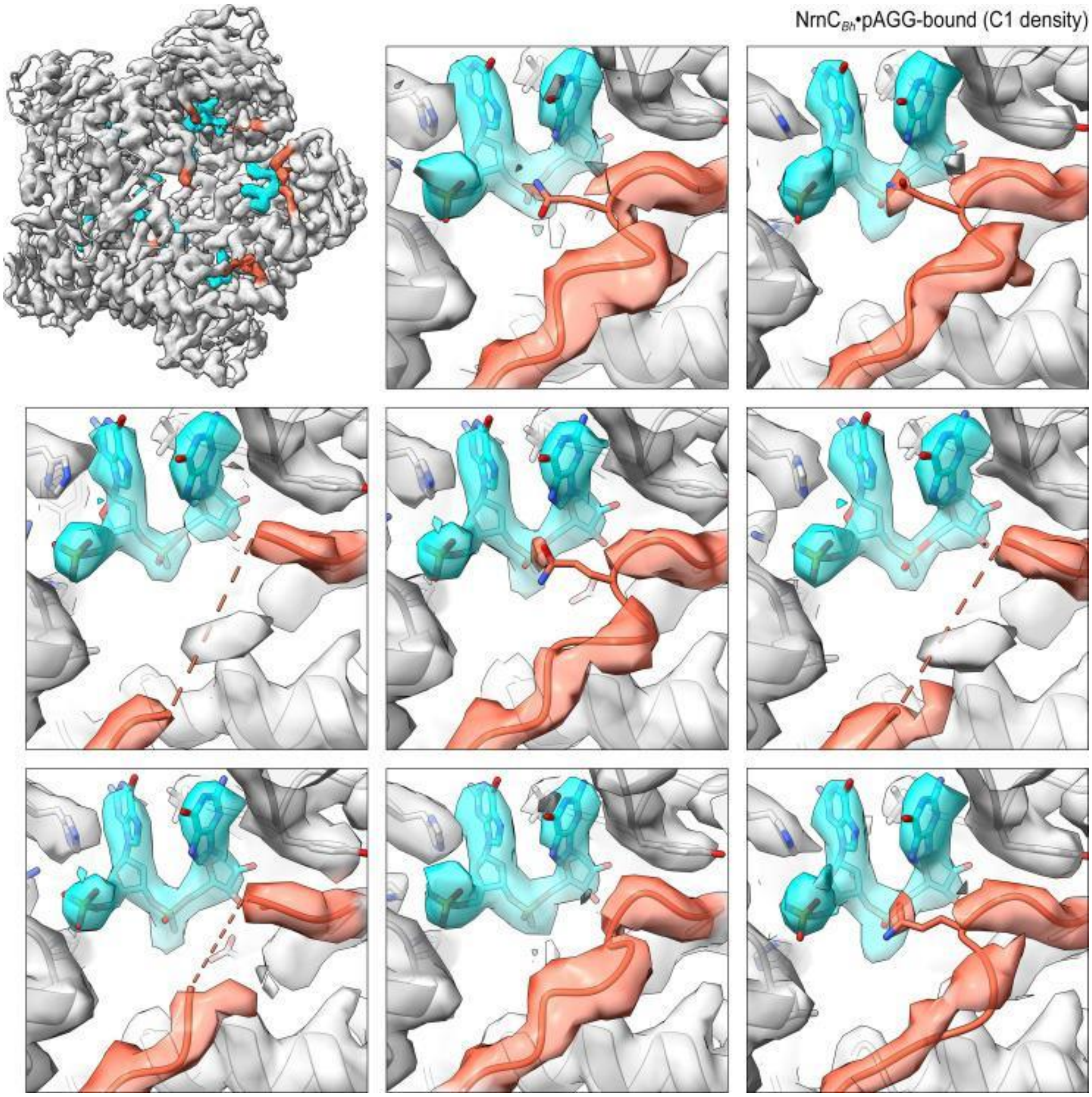


Figure 4.17 – Overall and individual active site electron density of a NrnCBh•pAGG octamer after refinement with C1 symmetry.

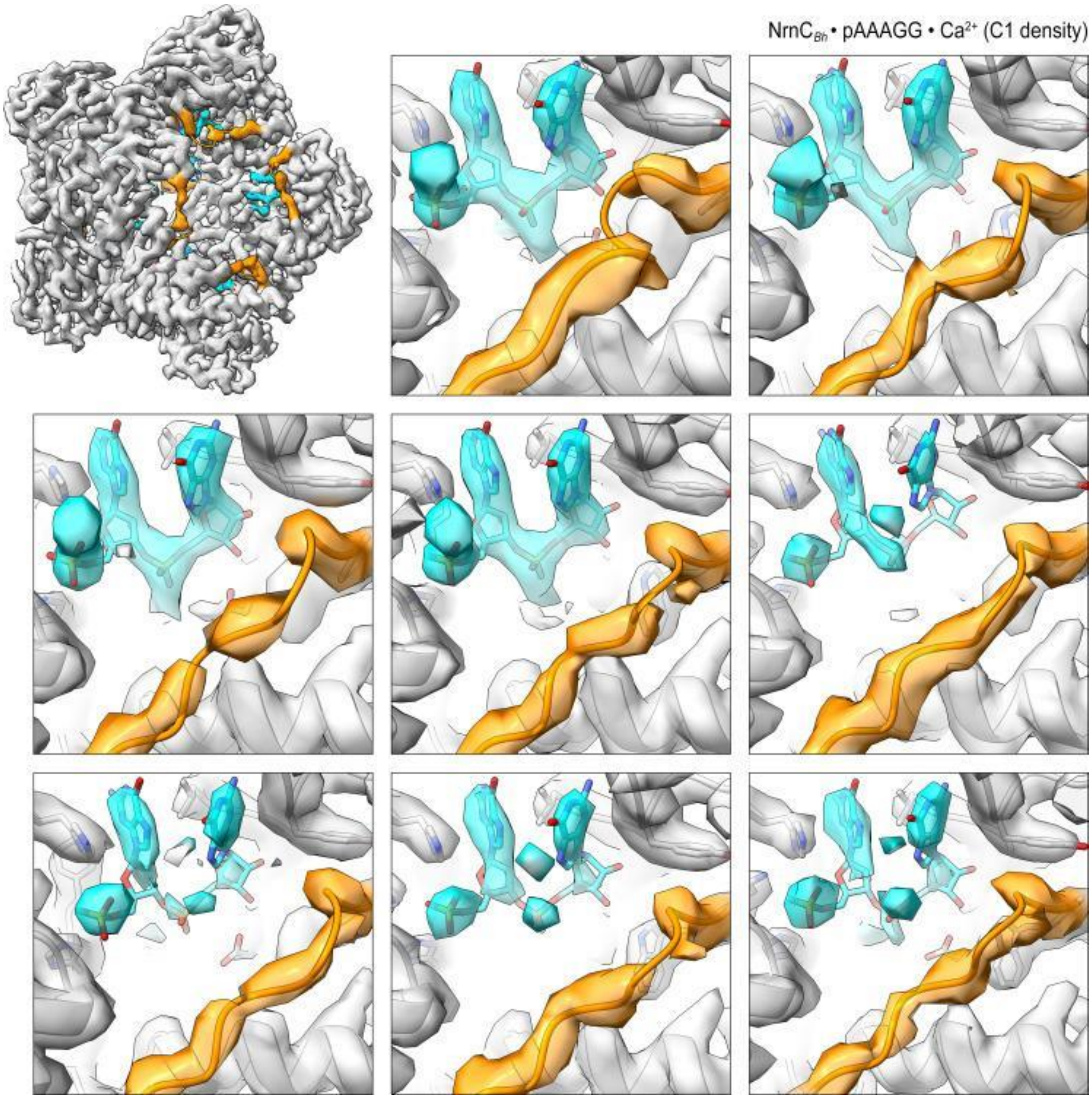


Figure 4.18 – Overall and individual active site electron density of a NrnCBh•pAAAGG octamer after refinement with C1 symmetry.

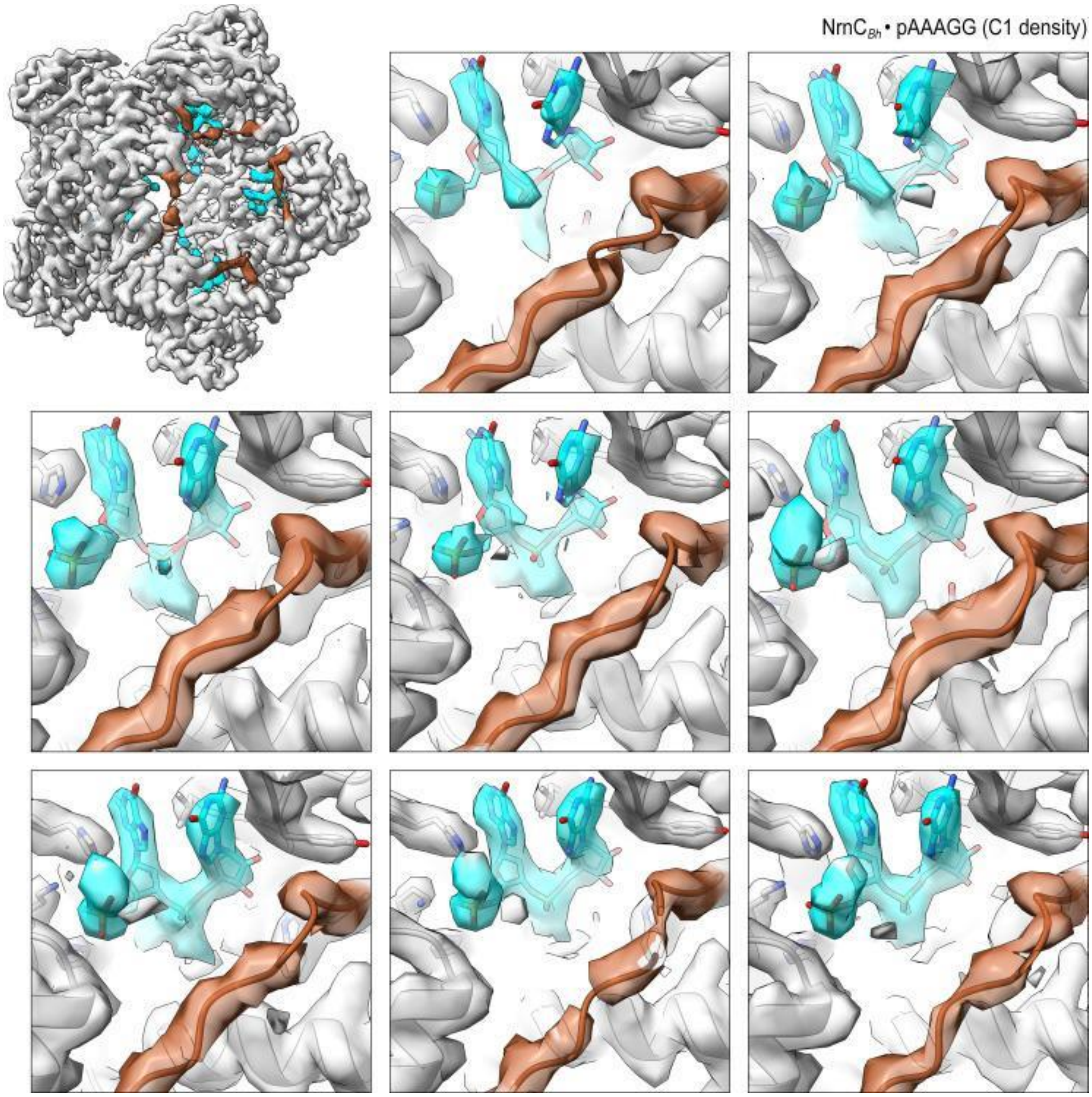
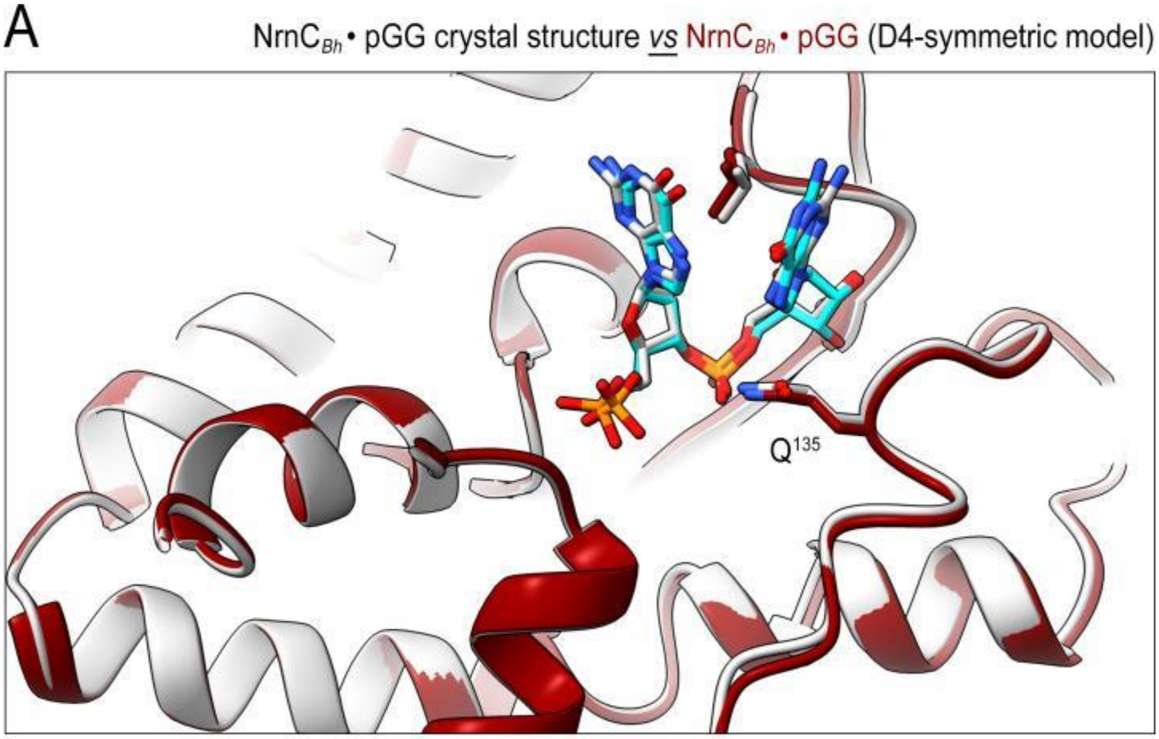
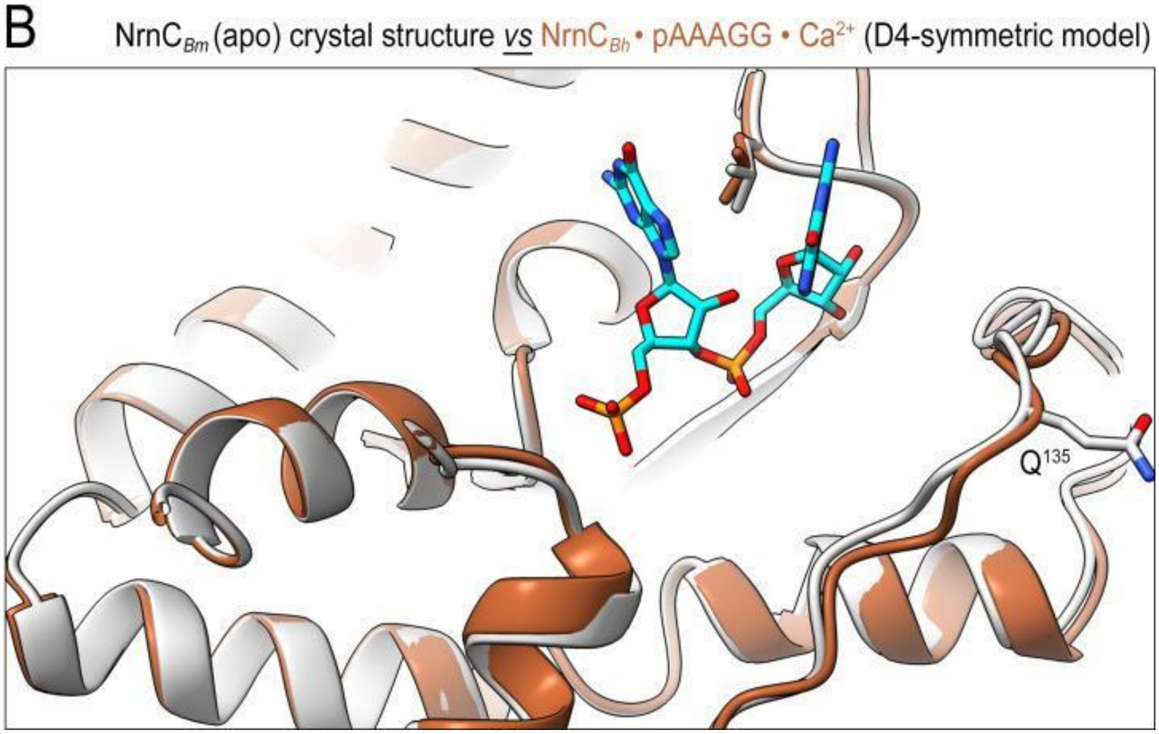


Figure 4.19 – Overall and individual active site electron density of a NrnCBh•pAAAGG octamer in the presence of Ca2+ ions after refinement with C1 symmetry.



NrnC$_{Bh}$ • pGG crystal structure *vs* NrnC$_{Bh}$ • pGG (D4-symmetric model)



NrnC$_{Bm}$ (apo) crystal structure *vs* NrnC$_{Bh}$ • pAAAGG • Ca^{2+} (D4-symmetric model)

Figure 4.20 The conformation of nano-RNase C (NrnC) bound to substrates with more than two bases resembles the crystallographic apo-state. (**A**) Comparison of the crystal structure of NrnCBh-pGG with the corresponding cryo-electron microscopy (cryo-EM) structure shows agreement between the solution and crystalline state of the protein with a well-ordered conformation of the loop residues 130–137 engaging the substrate. (**Continued…**)

Figure 4.20 The conformation of nano-RNase C (NrnC) bound to substrates with more than two bases resembles the crystallographic apo-state. (**B**) Comparison of the crystal structure of apo-NrnCBm with the cryo-EM structure of NrnCBm-pAAAGG. The superposition indicates that longer substrates may bind the active site but only the first full residues appear ordered, resulting in a conformation of nano-RNase C (NrnC) similar to the inactive state observed in the apo-state crystal structures.

4.8 – Results and Discussion: NrnC$_{Bh}$ acts preferentially on diribonucleotides.

Although the structural analysis revealed a narrow active site akin to that of Orn, NrnC's substrate-length preference has not been formally proven in this new context. To address this, we conducted kinetic experiments with RNAs of increasing length as substrates, following protocols established for Orn (Kim et al., 2019). In the initial assay, RNAs with two or more residues were in 200-fold abundance over NrnC, a condition where NrnC turned over the entire pGG pool within 1 min (Figure 4.21A). In comparison to NrnC's expedient activity on pGG, increasing the substrate length by only a single residue (pAGG) resulted in a striking decrease of nucleolytic cleavage under otherwise identical conditions. For RNAs with four and more residues, the substrate was incompletely processed and a band indicative of a cleavage product that was one base shorter at the 3' end slowly increased over the course of the experiment (Figures 3.5A and B). Increasing the concentration of NrnC increased activity on the longest substrate tested, an RNA with 7 bases (5'- 32P-labeled AAAAAGG, pAAAAAGG), but full conversion to mononucleotides required a ratio of 1:1 NrnC:RNA, indicating a comparatively inefficient, and likely less physiological mechanism (Figure 4.21B). Furthermore, and similar to the kinetics observed with Orn (Kim et al., 2019), a diribonucleotide intermediate was undetectable with enzyme concentrations that were required to observe the successive breakdown of the longer RNA, revealing the rapid turnover of 90 dinucleotides, proceeding at much faster timescales than with any other intermediate that could be readily observed (Figure 4.21B).

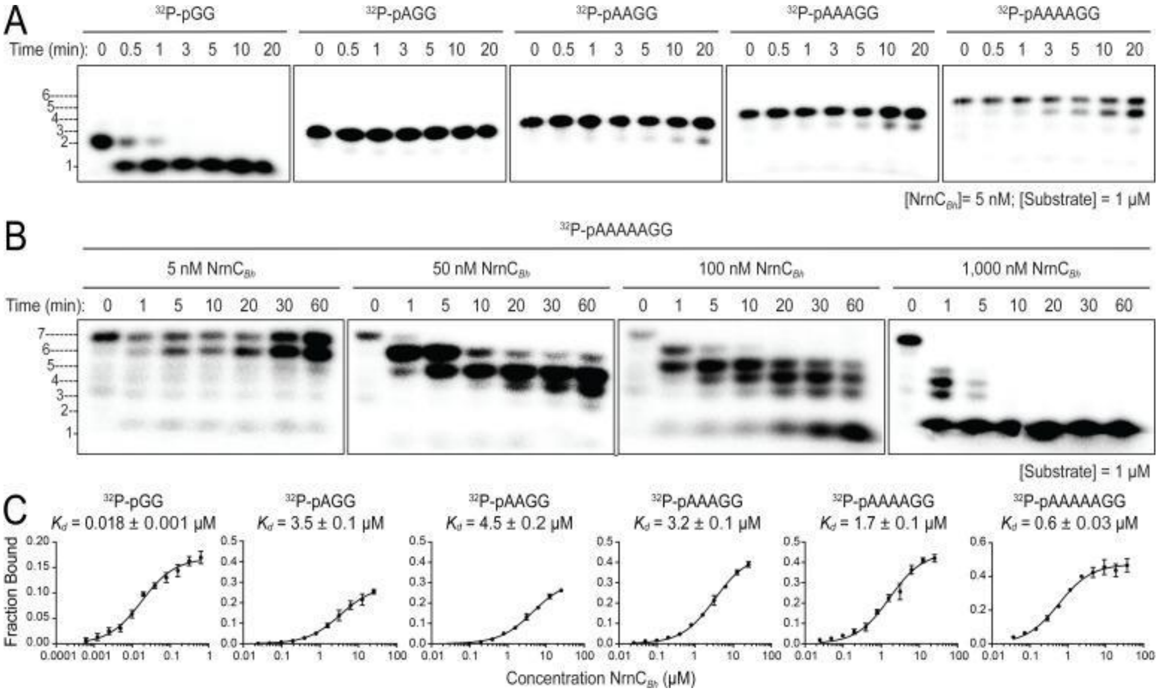


Figure 4.21 – NrnC shows a strong preference for substrates with two residues in length. (**A**) and (**B**). RNase assays. Experiments are similar to those in Figure 3.2 but were performed with radiolabeled substrates from 2 to 7 residues in length. Representative gels of at least two independent experiments are shown. In (B) enzyme concentration was varied from 5-1,000 nM (1:200 to 1:1 enzyme:substrate ratio). Substrate length-dependent binding studies. (**C**) Affinity of NrnC for RNA with different lengths. Fraction bound of radiolabeled substrates of increasing length was assessed at different NrnC concentrations and is plotted as means and SD from three independent experiments. Figure and data provided by Soo-Kyoung Kim (Lee laboratory, University of Maryland).

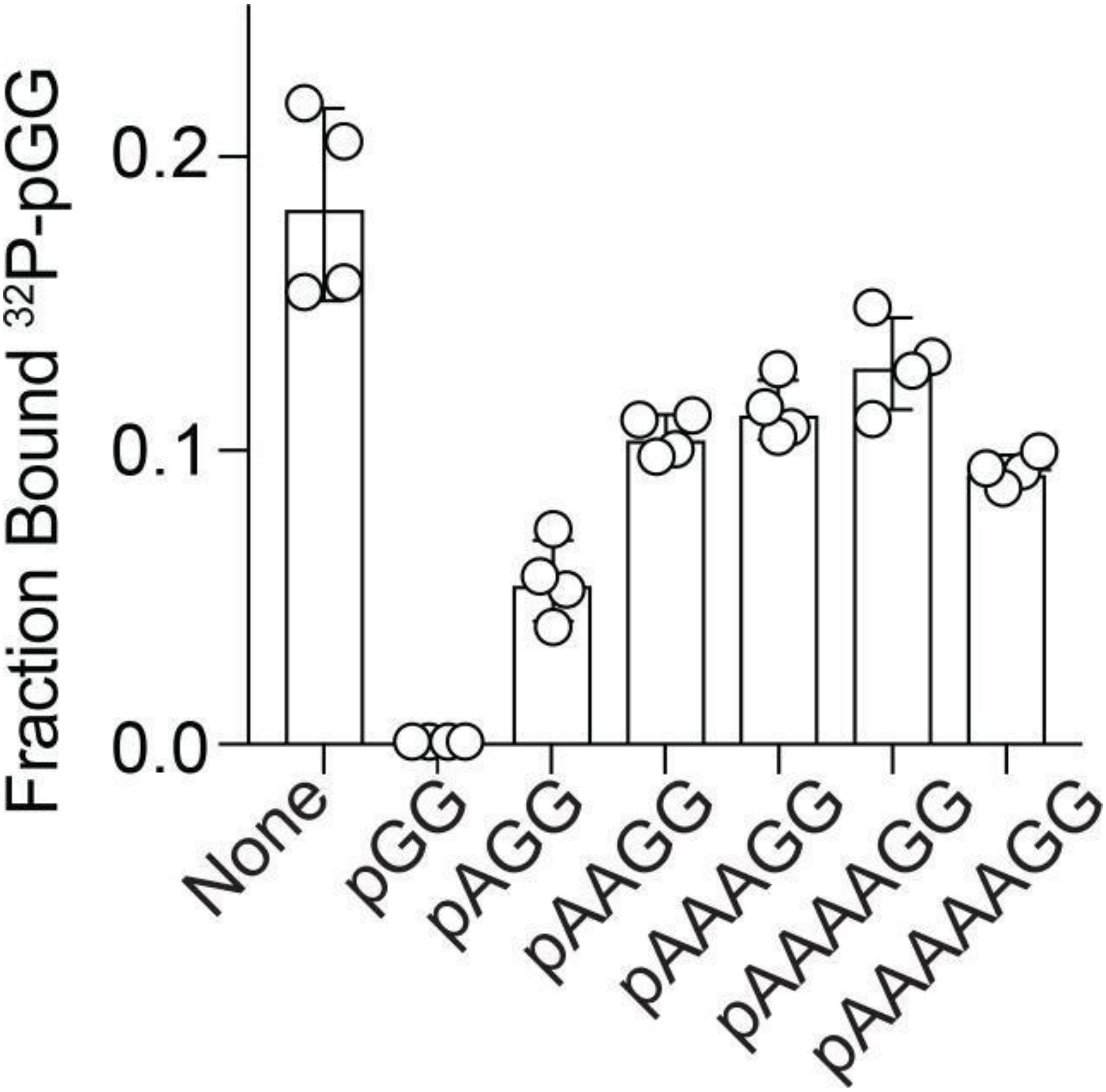


Figure 4.22 – Competition binding studies. Fraction bound of 32P-pGG to 200 nM purified NrnCBh in the presence of no competitor or 100 µM unlabeled RNA as indicated. Individual data, means, and SD of four independent experiments are plotted.

Quantification of NrnC’s binding to radiolabeled RNA substrates of different length agreed with the preference of NrnC to cleave diribonucleotides. Affinities of NrnC for radiolabeled substrates (2-7-mer RNA) were determined at physiological ionic strength and in the presence of Ca2+ to inhibit any potential residual catalysis (Rosta et al., 2014). NrnCBh bound to pGG with a 91 Kd = 17.7 nM. Increasing substrate length by just one additional residue resulted in a nearly 200- fold decrease in affinity (Kd pAGG = 3.49 µM) (Figure 4.21C). RNA substrates of four, five, or six residues showed similar decreases in affinity, while a 7-mer RNA substrate showed intermediate binding strength with a 32-fold decrease from pGG (Kd pAAAAAGG = 576 nM). As another method to assess substrate preference, competition experiments were performed by incubating NrnCBh with 32P-pGG with or without unlabeled RNAs of different length. While unlabeled pGG was able to displace radiolabeled pGG quantitatively on NrnC, longer substrates were less potent competitors under otherwise identical conditions (Figure 4.21 and 4.22). Together, the combined structural and biochemical results argue for a strong preference of NrnC towards the shortest species of RNAs, diribonucleotides with a 5’ phosphate.

4.9 – Results and Discussion: NrnC$_{Bh}$ processes DNA under specific experimental conditions.

The Agrobacterium tumefaciens NrnC octamer was previously interpreted as a conduit for long, polymeric substrates, in particular single-stranded RNA and single-stranded or doublestranded DNA (dsDNA), based on the octamer’s channel dimensions and positioning of the active sites (Yuan et al., 2018). DNase activity was proposed to allow NrnC octamers to act on opposite ends of dsDNA to completely unwind and degrade it by passing the strand through the central channel. Here, we asked whether NrnCBh could act on DNA substrates (Figure 4.23). Under near-physiological ionic strength and in the presence of either Mg2+ or Mg2+ ions, NrnCBh failed to degrade a 1.5 kb-long dsDNA fragment (Figure 4.23 B and C). Degradation of dsDNA was only observed at low ionic strength (~0-60 mM NaCl) and only in the presence of Mn2+ ions, conditions that match those used for Agrobacterium tumefaciens NrnC. The requirement for Mn2+ for activity on DNA substrates mirrors the previously reported 92 phenomenon observed with the 3’-5’ exonuclease EXD2 that acts on both ribo- and deoxyribonucleotides (Park et al., 2019). To examine whether potential NrnC activity on dsDNA was dependent on the presence of a 5’ phosphate or a 5’ overhang, dsDNA fragments were treated either with restriction enzymes, T4 polynucleotide kinase (PNK), or calf intestinal alkaline phosphatase (CIP) as indicated (Figure 4.23 A and D). With any of these modifications and similar to the parent, blunt dsDNA, degradation was only observed in combination of the absence of NaCl and presence of Mn2+ ions. These experiments call into question a general DNase activity of NrnC, although such an activity under specific cellular conditions remains plausible.

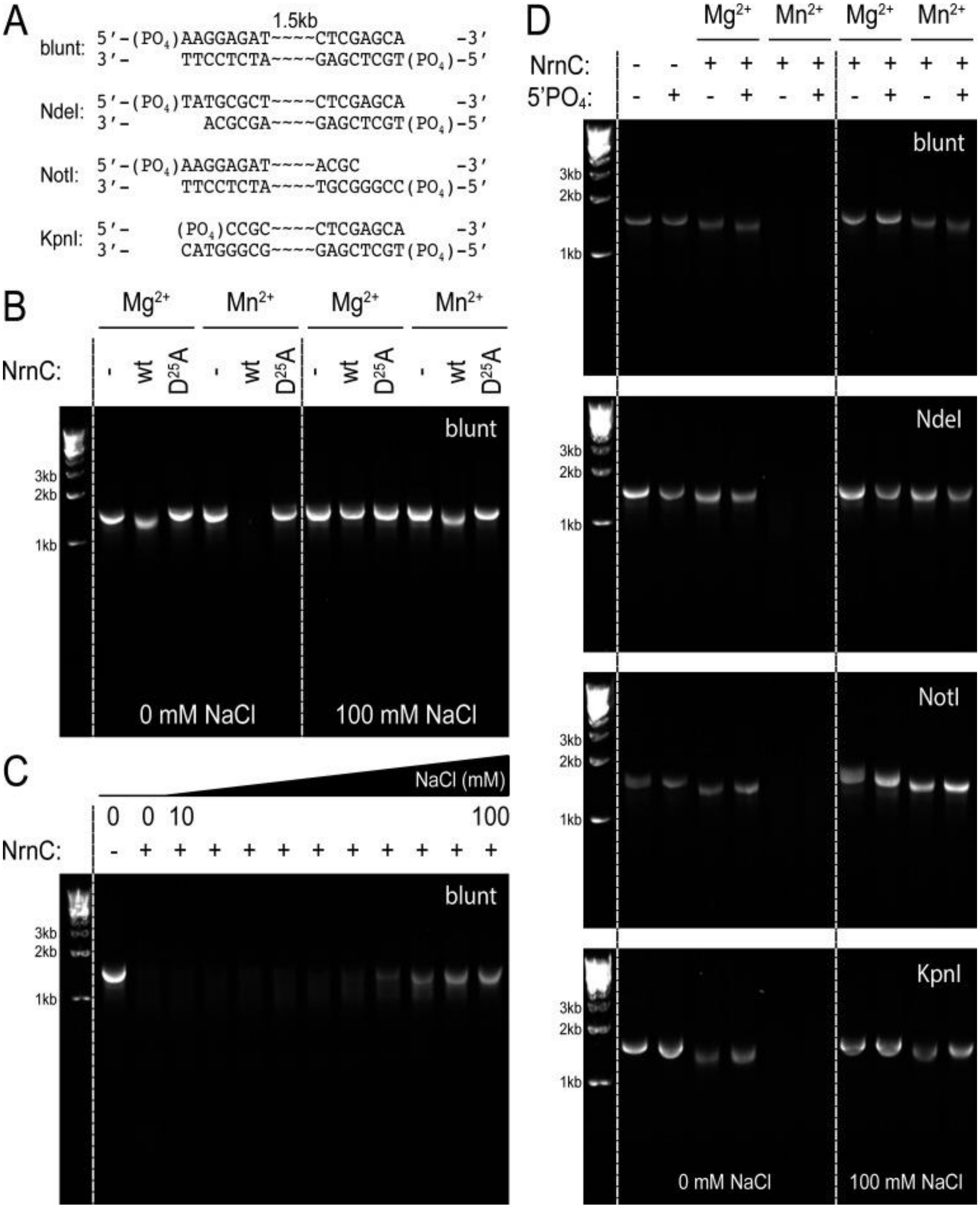


Figure 4.23 – NrnCBh degrades long DNA fragments under distinct conditions. (**A**) DNA fragments tested in this assay. (**B**) DNase activity of wild-type NrnCBh and a catalytically inactive mutant variant on blunt dsDNA in the presence of either Mg2+ or Mn2+ and in the absence or presence of NaCl. (**C**) NaCl titration on blunt dsDNA using wild-type NrnCBh. (**Continued…**)

Figure 4.23 NrnCBh degrades long DNA fragments under distinct conditions. (**D**) Nano-RNase C (NrnC) activity on various dsDNA substrates with or without a 5′-PO4 and in the presence of Mg2+ of Mn2+. Representative agarose gels are shown from at least two independent experiments.

4.10 – Results and Discussion: Phylogenetic analysis indicates repeated evolution of critical diribonucleotidase activity.

After having elucidated the mechanisms that signify NrnC-type and Orn-type RNases as dedicated diribonucleotidases, we reinvestigated the evolutionary distribution of these enzymes, relative to their structural homologs RNase D and RNase T, respectively. We included NrnA and NrnB in this analysis as the only other general RNases unrelated to NrnC and Orn, which also could stand in for Orn in a P. aeruginosa deletion strain, as shown previously (Orr et al., 2018). We identified homologs of the aforementioned proteins in the full UniprotKB database (version 2020_03), correlated the appearance of homologs in each species, and used the results to identify the spread of each type of nuclease at the order level (Figure 4.24). Orn was widespread in all Eukaryotic groups, although in bacteria it was infrequently found outside of the Proteobacteria and Actinobacteria phyla. The structurally related RNase T is largely limited to Proteobacteria, with a few exceptions. Of the other major nucleases included in this analysis, NrnA was most frequently found in many bacteria. In contrast, NrnB occurs more sparsely distributed, potentially suggesting a more specialized function in individual organisms. NrnC is primarily found in most Alphaproteobacteria and Cyanobacteria, often overlapping with the occurrence of the homologous RNase D (with respect to the catalytic-domain sequence), although RNase D is present in many more bacterial genomes than NrnC. The three proteins that act as effective diribonucleotidases (Orn, NrnC, NrnA/NrnB) were largely – although not always – mutually exclusive; most bacterial taxa in this analysis had only one of the three, with the notable exception of Actinobacteria that frequently contained both Orn and NrnC (Figure 4.24). We used the identified sequences of DnaQ family ribonucleases, each family curated individually, to create multiple sequence alignments and ultimately a combined phylogenetic tree of representative sequences of Orn, RNase T, NrnC, and RNase D family proteins (Figure 4.25). Similar to a previous analysis (Zuo & Deutscher, 2001), the DEDDh sequences (Orn and RNase T) segregated from the DEDDy sequences (RNase D and NrnC), highlighting the distinct evolutionary background of NrnC and Orn. Emanating from the ancestorial sequence, the first branch separated RNase T and Orn from RNase D and NrnC. This ancient split was soon followed by a split between RNase T and Orn. In the other lineage, NrnC and RNase D diverged from each other after a longer effective evolutionary time, presumably following a duplication from the more closely related RNase D (Figure 4.25A). Due to this relatively recent split, and as it frequently cooccurs with RNase D, NrnC appears to have arisen from a more recent specialization event.

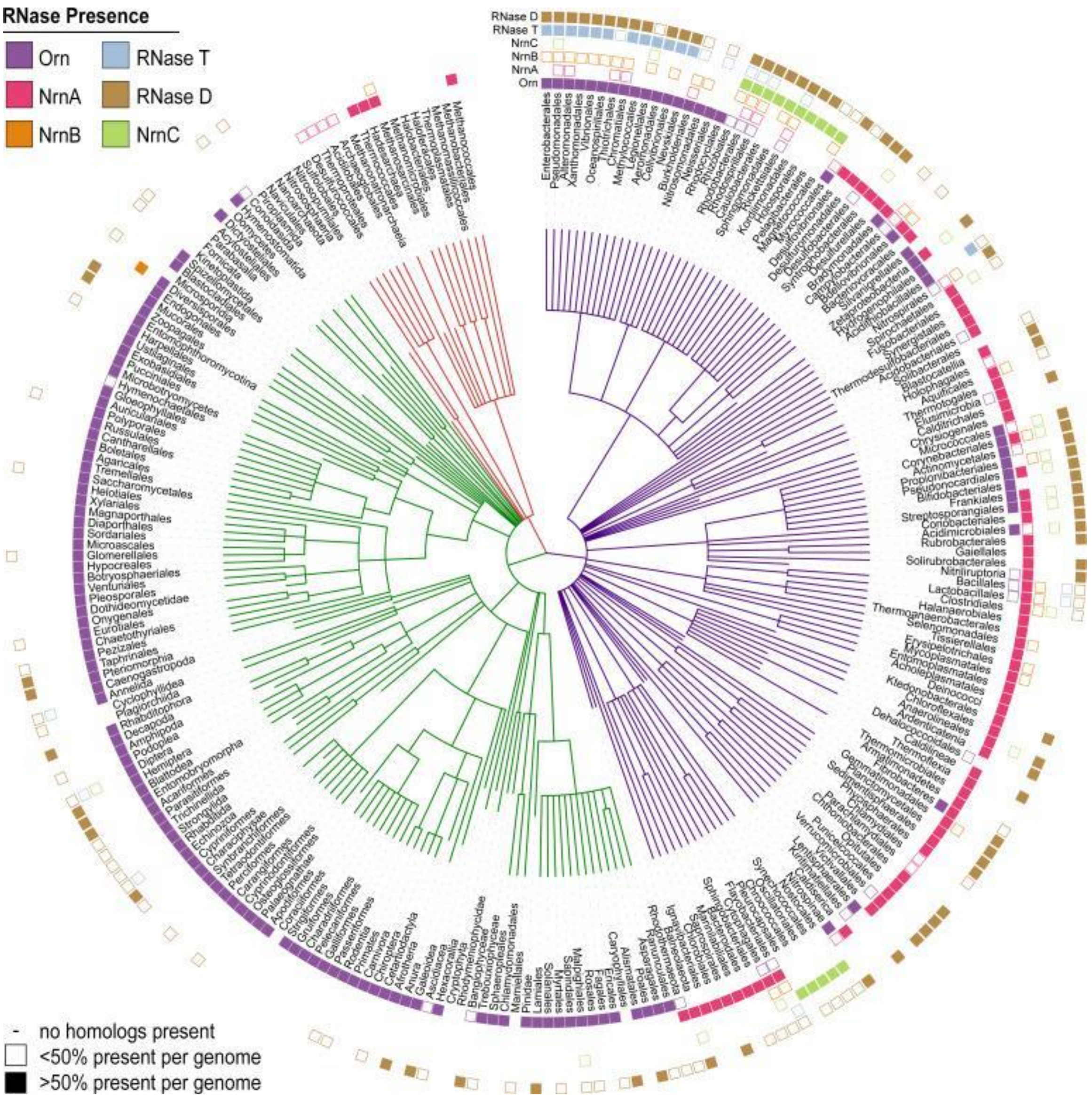


Figure 4.24 – Presence of RNase homologs across sequenced organism classes. Shown is a "Tree of Life" with all taxonomic groups at the class level with at least one substantially-complete proteome available in the dataset. The tree is based on the structure of the NCBI Taxonomy database, with bacterial taxa shown with purple lines, eukaryotic taxa shown with green lines, and archaeal taxa shown with red lines. The presence of each RNase homolog as a proportion of the total proteins in that taxonomic group is shown as either a filled square (>50% presence of a homolog per genome) or an empty square (<50% presence of a homolog per genome). Lack of a square indicates no homologs for that family were present in genomes of that class. Figure and data provided by Jonathan Goodson (University of Maryland).

Catalytic, L-wedge, and phosphate cap residues are strictly conserved in many NrnC orthologs (Figures 4.1D and 3.25B). RNase D, the closest relative to NrnC, shares the catalytic DEDDy motif with NrnC, but lacks the L-wedge and phosphate cap (Figure 3.25B), suggesting that these features distinguish NrnC (and Orn) from other RNases and DnaQ-fold enzymes. The phylogenetic analysis also identified a unique group of NrnC-like enzymes in Actinobacteria, which shares most of the characteristic NrnC features, including conservation of the active site, most of the phosphate cap and wedge residues. However, subtle but specific changes in important residues (i.e., a Q-to-R change in the SKQQQS loop and the phosphate caps H205, which is replaced by a tyrosine residue) hint at the possibility of distinct function of the NrnC orthologs in this subgroup (Figure 4.25). Some Actinobacteria also encode an additional Orn and/or NrnA in the same genome, which could suggest either redundant functions or further specialization of these enzymes in these organisms (Figure 4.24).

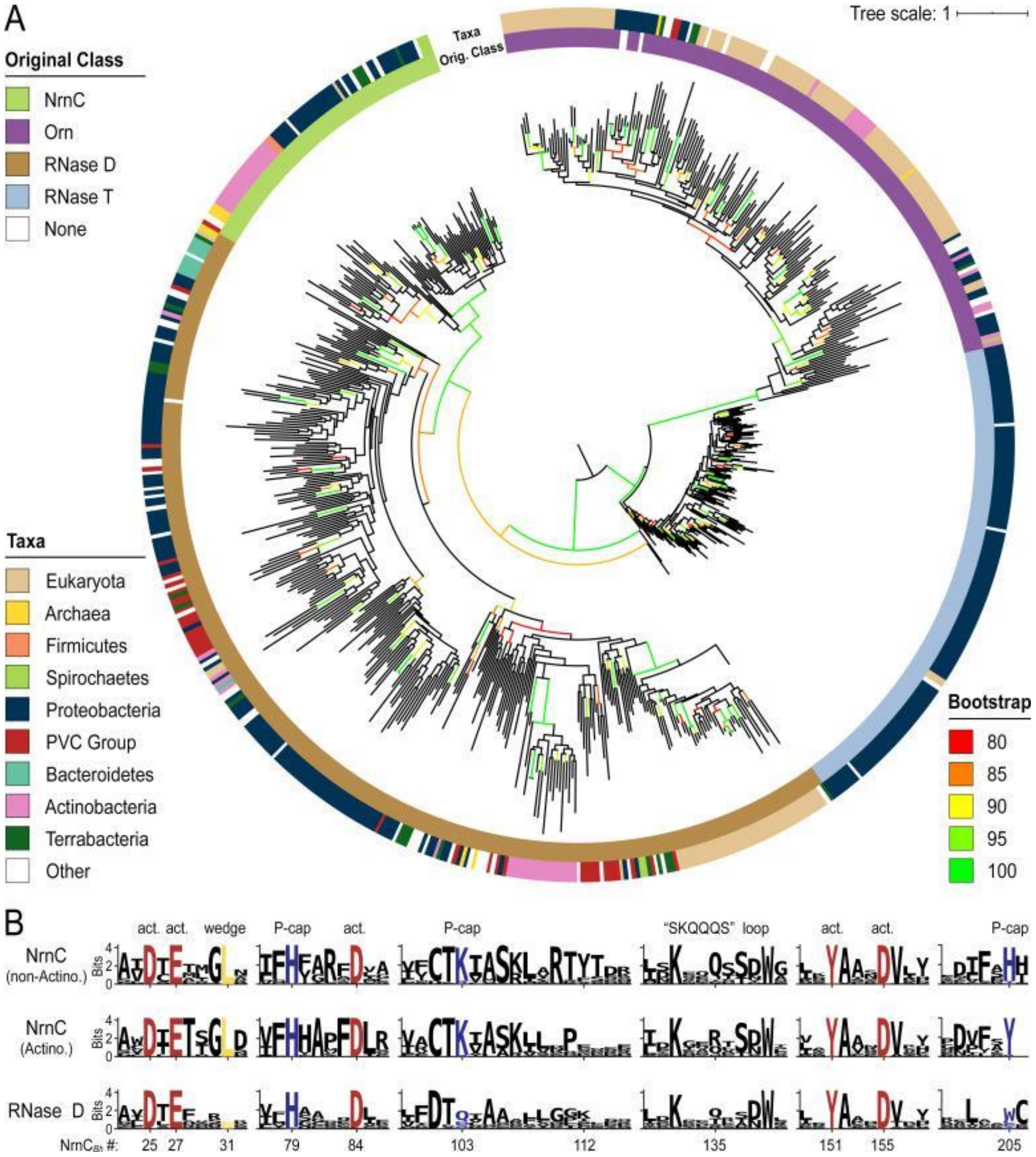


Figure 4.25 – Phylogenetic tree of four DnaQ-fold RNase families. (**A**) Phylogenetic tree of 669 representatives of the RNase T, RNase D, Orn, and NrnC families of RNase proteins. The inner ring represents the original classification of each sequence by HMM analysis. The outer ring represents the highlevel taxonomic classification of the organism the protein is found in. The color of the branch represents the UFBoot bootstrap value, where black branches are <80%, red is 80%, orange is 85%, yellow is 90%, light green is 95%, and bright green is 100%. Bootstrap values >90% indicate high-confidence splits. (**B**) Sequence logos of RNase D and NrnC subgroups. Sequence logos showing the relative entropy (information content) at selected positions in RNase D as well as the Actinobacterial and nonActinobacterial subsets of NrnC. Sequence numbering is relative to Bartonella birtlessi NrnC (G4VUY7). Active site residues are shown in red, phosphate cap residues in dark blue, and the L-wedge in yellow. Figure and data provided by Jonathan Goodson (University of Maryland).

4.11 – Conclusions

The establishment of NrnC as a dedicated diribonucleotidase led us to compare and contrast the structural features of NrnC and Orn, the other enzyme with such a specific activity (Kim et al., 2019). Because NrnC and Orn evolved independently from two different families of RNases, RNase D and RNase T, respectively, their shared substrate-length specificity is particularly noteworthy. On the level of a single subunit, NrnC and Orn contain a conserved catalytic core comprised of a β-sheet and adjacent α-helices, which harbors the residues for divalent cation coordination, the functionally important tyrosine or histidine residue (‘DEDDy’ or ‘DEDDh’), and a wedge residue (L-wedge) that separates the substrate’s bases (Figure 4.26A). In addition, the position of two residues contributing to the respective phosphate caps is conserved, but the identity of the residues varies between the two enzyme groups. For both enzyme families, the active site involves residues from two monomers requiring minimally a dimeric protein (Figure 4.26B). Thus, the two enzyme families are characterized by a functionally conserved active site optimized to accommodate the shortest RNA substrates, which distinguishes the diribonucleotidases characterized to date from other RNases that act on longer substrates.

Despite these remarkable, function-defining commonalities between NrnC and Orn/Rexo2, their different evolutionary histories have led to distinct implementation of some of these important features. Particularly, NrnC and Orn differ in secondary structure motifs at the periphery of their conserved catalytic core. The enzyme family-specific parts include, for NrnC, additional phosphate cap residues, an N-terminal β-strand and an additional C-terminal α-helix; and for Orn, short helices La, Lb, and Lc forming a lobe that coordinates the 3’ base of the dinucleotide substrate (Figure 4.26A). On the quaternary-structure level, both NrnC and Orn form dimers (Figure 4.26B). However, the specific dimer arrangements vary between NrnC and Orn. The C-terminal α- 99 helix of NrnC, which is absent in Orn, serves as the major dimerization interface in this family of enzymes, yielding a 2-fold-symmetric dimer. Orn and Orn-related enzymes (such as Rexo2) also form a 2-fold symmetric unit, but via a distinct rotation axis and the interface involves elements of the central core fold (Figure 4.26B). In contrast to Orn/Rexo2, whose biological unit is the dimeric form, four dimers of NrnC assemble further to the final octameric assembly. Lateral interfaces between dimers of NrnC within the octamer replace the 3' lobe of Orn in coordinating the 3' base of the substrate. Despite these major differences in the architectural composition and unique features of the two enzyme families, the elements that we identified as characteristic for diribonucleotidases align nearly perfectly in space (Figures 4.26B and C).

Of the NrnC-specific motifs, the N-terminal β-strand and near-C-terminal α-helix are conserved in RNase D (Figure 4.3B). NrnC's C-terminal extension including H205, placed in the active site of an adjacent subunit within a NrnC dimer, is involved in 5' base stacking and completes the phosphate cap. Mutations in NrnC's unique phosphate cap retained residual activity, suggesting a role mainly in restricting the length of substrates accommodated at the active site (Figure 4.6). In contrast, corresponding mutations at the phosphate cap of Orn abolished catalytic activity, suggesting these residues contribute directly to the catalytic mechanism in addition to imposing a substrate-length restriction at the active site (Kim et al., 2019). One possibility is that the more ancient Orn evolved more stringent diribonucleotide preference compared to the more recently occurring NrnC-type enzymes. Taken together, the phosphate cap is a unifying feature of Orn and NrnC, though arisen independently in the two enzyme classes. In general, evolution of such specialized active sites that only degrade diribonucleotides would also allow for rapid turnover of this specific RNA pool since competition 100 from longer RNAs would be suppressed. Whether similar motifs have evolved in other enzymes to restrict substrate length remains to be established.

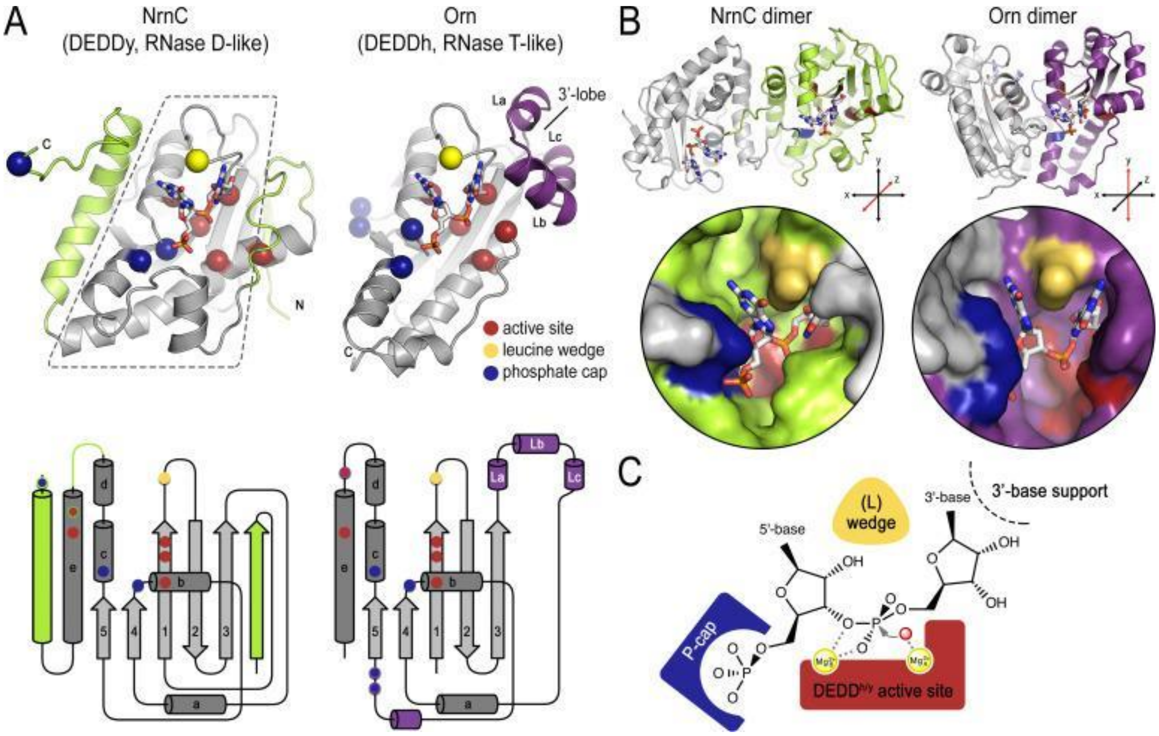


Figure 4.26 – Structural comparison NrnC and Orn. (**A**) Fold topology. pGG-bound NrnC and Orn monomers are shown in a similar orientation as cartoons (top) or schematic topology diagrams (bottom). Conserved catalytic core elements are colored in grey. NrnC and Orn-specific features are colored in green and purple, respectively. Other color codes mark the positions of the DEDDy/h motif (red spheres), Lwedge (yellow sphere), and phosphate cap residues (dark blue spheres). (**B**) Comparison of dimer units of NrnC and Orn (top) with close-ups of the composite active sites of the enzymes (bottom). A NrnC monomer is colored green and an Orn monomer is colored purple, with adjacent monomers in the biological assemblies colored in light grey. Specific residues are colored as in (A). Coordinate systems indicate the 2- fold symmetry axis of the enzyme dimers, with the colored monomers shown in a similar orientation. (**C**) Structurally and functionally conserved features common among NrnC- and Orn-type diribonucleotidases.

Together, our bioinformatics analysis revealed the unique histories of NrnC and Orn, two nucleases that arose independently to fulfill the crucial role of diribonucleotide degradation. Their activities are at the confluence of RNA metabolism and bacterial second messenger signaling, signifying their importance for cellular homeostasis and regulation. Indeed, functional NrnC or Orn, and consequently the clearance of the cellular diribonucleotide pool, are necessary for proper growth, being essential in many organisms (Ghosh & Deutscher, 1999; Kim et al., 2019; Liu et al., 2012; Sternon et al., 2018), which could provide an avenue for targeted antimicrobial intervention. Specifically, as Bartonella and Brucella species are important pathogens (Dehio, 2005; Greenfield et al., 2002; Pappas et al., 2006), understanding the implications of NrnC function and failure could offer insight and effective strategies to battle the pernicious impacts of these organisms. Our structure-function studies present blueprints for such endeavors by revealing the specific activesite architectures and activity profiles of NrnC and Orn/Rexo2, their similarities and differences, as well as the general features that distinguish diribonucleotidases from other 3'-5' exoribonucleases

4.12 – Materials and methods

Expression constructs and mutagenesis. For protein expression in E. coli, codon optimized NrnC genes from Bartonella henselae (BH02530) and Brucella melitensis (BMEI1828) were synthesized by Geneart (Life Technologies). Genes were cloned into a modified pET28a vector (Novagen) between BamHI and NotI sites using InFusion cloning (Takara Bio). The resulting fusion proteins expressed from these plasmids contained an N-terminal His6-tagged small ubiquitin-like modifier (SUMO) cleavable by recombinant Ulp-1 protease.

For the arabinose-inducible expression and detection of NrnC in P. aeruginosa, we used a modified pJN105 vector (Newman & Fuqua, 1999). The vector pJGHA was constructed by inserting a coding sequence for monomeric superfolder GFP (msfGFP)-

HA epitope fusion between the NheI and XbaI sites in pJN105. The plasmid allows insertion of genes of interest 102 between novel NdeI and EcoRI sites, and their expression results in proteins with C-terminal msfGFP-HA. The vector pJHA was constructed by digesting pJGHA with HindIII (New England Biolabs) to remove the msfGFP coding sequence. Following re-ligation of the gel-purified restriction digest, the coding sequence for the HA epitope remained, allowing proteins of interest to be expressed with a C-terminal HA epitope. For expression in P. aeruginosa, the codonoptimized NrnCBh sequences were amplified by PCR and inserted between NdeI and EcoRI sites of the pJHA vector using InFusion cloning. A QuikChange II site-directed mutagenesis kit (Agilent) was used for the introduction of point mutations in nrncBh following the manufacturer's instructions. All mutations were verified by DNA sequencing.

Protein expression and purification.

E. coli BL21 T7 Express cells (New England Biolabs) were transformed with pET28a plasmids encoding His6-SUMO-NrnCBh or -NrnCBm and grown in Terrific Broth (TB) supplemented with 50 µg/ml kanamycin at 37°C to an OD600 of ~1.0. Induction was carried out at 18°C with 0.5 mM IPTG for 16 hours. Cells were harvested by centrifugation, resuspended in minimal volume of Ni-NTA binding buffer (25 mM Tris-Cl pH 8.5, 500 mM NaCl, 20 mM imidazole), frozen in liquid nitrogen, and stored at -80 °C.

Cell pellets were thawed followed by lysis through sonication and centrifugation. Supernatants were incubated on ice with Ni-NTA resin (Qiagen) equilibrated with Ni-NTA binding buffer for 1 hour with gentle agitation. The NrnC-bound resin was washed three times with 10 column volumes of Ni-NTA binding buffer by gravity flow. Bound NrnC was eluted in 6 column volumes of Ni-NTA elution buffer (25 mM Tris-Cl pH 8.5, 500 mM NaCl, 400 mM 103 imidazole). Eluates were buffer exchanged into gel filtration buffer (25 mM

Tris-Cl, 150 mM NaCl, pH 7.5) via a HiPrep 26/10 desalting column (GE Healthcare) and incubated overnight with Ulp-1-His6 to cleave the His6-tagged SUMO moiety from NrnC. Following Ulp-1 cleavage, untagged NrnC protein was recovered in the flow through of a HisTrap Ni-NTA column (GE Healthcare), separated from His6-SUMO, and ULP1-His6. EDTA at a final concentration of 10 mM was added to NrnC before concentration via an Amicon Ultra 10K concentrator (Merck Millipore). Concentrated NrnC was injected onto a HiLoad 16/60 Superdex 200 gel filtration column (GE Healthcare) equilibrated in gel filtration buffer. Fractions containing NrnC were concentrated, frozen in liquid nitrogen, and stored at -80°C.

Crystallization, data collection and structure refinement.

NrnC-RNA complexes (pGG and pAA from Biolog Life Science Institute, other nucleotides from Dharmacon) were formed prior to crystallization by mixing 1:2 molar ratio of protein:nucleotide in gel filtration buffer, followed by 30 min incubation at the crystallization temperature. Protein concentrations used in crystallization ranged from 2.5 – 10 mg/ml (NrnCBh) and 1.0 – 8.0 mg/ml (NrnCBm). Crystals were grown via hanging-drop vapor diffusion by mixing equal volumes (0.8 µl) of sample with reservoir solution. NrnCBh crystals grew at 20°C over a reservoir solution containing 0.1 M succinic acid (pH 6.5), 15 – 20% PEG 3350, and 20% xylitol. NrnCBm crystals grew at 4°C and 20°C, in reservoir solutions containing 0.1 M Tris-Cl (pH 7), 2.0 - 2.4 M ammonium sulfate or 1.4 M sodium-potassium phosphate, and 20% xylitol. Prior to freezing crystals in liquid nitrogen, they were soaked in reservoir solution with up to 25% xylitol. Data were collected by synchrotron radiation on frozen crystals at Cornell High Energy Synchrotron Source (CHESS) and NE-CAT 24ID-C and 24ID-E beamlines at the Advanced 104 Photon Source (APS) at Argonne National Laboratory. Diffraction data sets were processed using XDS, Pointless and Scala, (Evans, 2006; Kabsch, 2010). The initial structures were solved by molecular replacement using the software package Phenix (Adams et al., 2010) and MrBUMP in the ccp4 suite (Keegan et al., 2018; Winn et al., 2011) with the coordinates of E. coli RNase D (PDB:1yt3, (Zuo et al., 2005)) as the search model. Manual model building and refinement were carried out with Coot (Emsley et al., 2010) and Phenix. Illustrations were prepared in Pymol (Version 2.0 Schrödinger, LLC). All software packages were accessed through SBGrid (Morin et al., 2013). All data collection and refinement statistics are summarized in Table 4.1.

Table 4.1 – Data collection and refinement statistics.[a]

Name	NrnC•pGpG	NrnC•pApA	NrnC•pGpC	NrnC•pAp	NrnC•pGpG/Mg^{2+}	NrnC•pGpG/Mn^{2+}
Organism	*B. henselae*	*B. henselae*	*B. henselae*	*B. henselae*	*B. henselae*	*B. henselae*
Ligand	pGpG	pApA	pGpC	pAp	2GMP/Mg2+	2GMP/Mn2+
Data collection						
X-ray source	APS 24-ID-E	CHESS 7B2	CHESS 7B2	CHESS 7B2	APS 24-ID-C	APS 24-ID-C
Wavelength	0.9792	1.000	1.000	1.000	0.9792	1.8917
Resolution range (Å)	64.4-1.80	42.7-1.95	47.2-1.94	47.2-1.95	127-2.00	128-2.35
Space group	P21	P21	P212121	P212121	P21	P21
Unit cell a/b/c (Å)	71.2 128.3 129.3	70.8 127.5 128.6	100.1 142.4 148.8	99.7 141.7 147.9	70.9 126.7 127.1	71.7 127.8 128.8
Unit cell a/b/g (°)	90 94.7 90	90 95.3 90	90 90 90	90 90 90	90 94.3 90	90 94.9 90
Total reflections	889162 (80298)	615957 (61235)	1031986 (104247)	1121813 (108055)	616220 (52575)	589748 (79845)
Unique reflections	212143 (20933)	164518 (16396)	155943 (15333)	152553 (15073)	148330 (14541)	174471 (12956)
Multiplicity	4.2 (3.8)	3.7 (3.7)	6.6 (6.8)	7.4 (7.2)	4.2 (3.6)	6.6 (6.2)
Completeness (%)	99.21 (98.12)	99.68 (99.53)	99.71 (99.13)	99.89 (99.90)	98.14 (95.95)	98.8 (98.5)
I/sigma(I)	17.50 (1.43)	11.98 (2.36)	9.08 (1.72)	14.04 (2.46)	18.57 (1.94)	11.0 (1.9)
Wilson B-factor	30.72	26.37	24.26	25.11	33.48	46.67
CC1/2	0.999 (0.602)	0.997 (0.775)	0.998 (0.778)	0.999 (0.877)	0.992 (0.444)	0.994 (0.709)
Refinement						
Reflections (R-free)	10608 (1073)	1990 (198)	1999 (195)	1990 (198)	2004 (197)	4387 (449)
R-work	15.09	14.51	19.41	16.43	15.80	16.19
R-free	17.11	17.96	22.50	18.81	18.72	20.63
Non-hydrogen atoms	15456	15812	15868	15307	14792	14235
macromolecules	13590	13766	13556	13516	13224	13261
ligands	0	0	0	0	424	352
solvent	1866	2046	2312	1791	1144	622
Protein residues	1639	1638	1641	1644	1641	1639
Rms (bonds)	0.006	0.006	0.006	0.012	0.006	0.007
Rms (angles)	0.78	0.77	0.78	0.97	0.84	0.87
Ramachandran						
favored (%)	99.51	99.51	99.08	99.26	99.45	99.45
allowed (%)	0.49	0.49	0.92	0.74	0.55	0.55
outliers (%)	0	0	0	0	0	0
Rotamer outliers (%)	0.90	0.69	1.66	0.95	1.17	2.00
Clashscore	1.48	1.98	2.08	2.04	1.44	2.81
Average B-factor	37.17	30.72	27.52	29.31	41.77	48.70
macromolecules	36.20	29.26	25.72	27.90	40.90	47.68
ligands					59.22	81.66
solvent	45.18	40.16	38.14	39.25	45.46	48.76
PDB ID	7MPL	7MPM	7MPN	7MPO	7MPP	7MPQ

[a] Statistics for the highest-resolution shell are shown in parentheses.

Table 4.1 – Data collection and refinement statistics (Continued).[a]

Name	NrnC	NrnC	NrnC	NrnC•pGpG
Organism	*B. melitensis*	*B. melitensis*	*B. melitensis*	*B. melitensis*
Ligand	-	-	-	pGpG (partial)
Data collection				
X-ray source	APS 24-ID-C	APS 24-ID-C	APS 24-ID-C	APS 24-ID-E
Wavelength	0.9795	0.9795	0.9795	0.9792
Resolution range (Å)	74.56-1.42	111-1.45	93-1.75	54-1.72
Space group	P422	P4	P422	I41 2 2
Unit cell a/b/c (Å)	93.3 93.3 124.0	110.7 110.7 110.2	93.3 93.3 124.0	170.9 170.9 85.3
Unit cell a/b/g (°)	90 90 90	90 90 90	90 90 90	90 90 90
Total reflections	1580788 (109208)	2355312 (194722)	888616 (81508)	1095551 (107578)
Unique reflections	103308 (10058)	233668 (23187)	55784 (5443)	66925 (6617)
Multiplicity	15.3 (10.9)	10.1 (8.4)	15.9 (15.0)	16.4 (16.3)
Completeness (%)	99.84 (98.78)	99.91 (99.40)	99.75 (99.07)	99.92 (99.43)
I/sigma(I)	30.58 (1.91)	21.06 (1.86)	26.42 (2.02)	22.49 (3.04)
Wilson B-factor	18.16	17.29	24.89	21.79
CC1/2	1 (0.782)	1 (0.702)	0.999 (0.758)	0.999 (0.878)
Refinement				
Reflections (R-free)	2000 (195)	2016 (194)	1998 (195)	2000 (198)
R-work	16.53	15.27	16.59	15.51
R-free	18.10	16.64	18.00	18.38
Non-hydrogen atoms	4019	8041	3823	3991
macromolecules	3303	6553	3253	3352
ligands	15	125	9	30
solvent	701	1363	561	609
Protein residues	410	815	407	410
Rms (bonds)	0.004	0.006	0.006	0.006
Rms (angles)	0.72	0.86	0.77	0.81
Ramachandran				
favored (%)	99.26	98.88	99.25	99.00
allowed (%)	0.74	1.12	0.75	1.00
outliers (%)	0	0	0	0
Rotamer outliers (%)	0.55	0.42	0.28	0
Clashscore	1.21	1.72	0.77	1.34
Average B-factor	25.26	23.93	31.19	26.69
macromolecules	22.57	20.96	29.22	24.39
ligands	69.90	39.38	47.55	51.50
solvent	36.95	36.79	42.37	37.88
PDB ID	7MPR	7MPS	7MPT	7MPU

[a]Statistics for the highest-resolution shell are shown in parentheses.

Structure determination by cryo-EM.

Purified NrnCBh was diluted in buffer (25 mM Tris-Cl pH 7.5, 150 mM NaCl) to 4, 5, or 6 mg/ml (for incubation with pGG, pAGG, and pAAAGG, respectively). RNA substrate was added at 3-fold molar excess. After 15 min at room temperature, NP40 was added at 0.01% v/v and samples were placed on ice for an additional 15 min. Alternatively, NrnC was diluted to 7.5 mg/ml and incubated with 3-fold excess pAAAGG; after 15 min, 1/5 volume of buffer with CaCl2 and NP40 (25 mM Tris-Cl pH 7.5, 150 mM NaCl, 5 mM CaCl2, 0.05% NP40) was added and incubations proceeded for another 15 min on ice. All cryo-EM samples were prepared with 107 Quantifoil R1.2/1.3 300-mesh grids after glow discharging in a PELCO easiGlow (60 sec glow, 10 mA current, 80% Ar/20% O2) using a FEI Mark IV Vitrobot (4 °C, 100% humidity, 2.5 sec blot) to plunge grids into liquid nitrogen-cooled ethane immediately after blotting. Data were collected using the Cornell CCMR facility Thermo Fisher Scientific Talos Arctica with a Gatan K3 detector and BioQuantum energy filter operated at 200 kV at a nominal magnification of 63kX (1.24 Å/pixel), 20 eV slit width, and 0.5X binning (super-resolution). Movies were collected with a total dose of 50 e/Å2 , fractionated into either 50 or 75 frames.

Data processing was performed using RELION 3.1 (Zivanov et al., 2018) and cryoSPARC (Punjani et al., 2017). Super-resolution movie exposures were aligned, dose-weighted, and Fourier-cropped to the physical pixel size in RELION using MotionCor2 (Zheng et al., 2017), and defocus values were estimated using GCTF (Zhang, 2016). Micrographs were then imported into cryoSPARC for manual curation, particle picking and classification. Particles were picked using the cryoSPARC “blob” and template picking and initially extracted with Fourier-cropping to a nominal pixel size of 2.89 Å. This particle stack was cleaned with 2D and 3D classifications in cryoSPARC, then re-extracted in RELION (1.24 Å/pixel) for 3D refinement, CTF refinement, and Bayesian polishing. Further 2D and 3D classification of CTF-refined particles in RELION was performed for the pAGG dataset. For each dataset, no symmetry was imposed during processing until a final refinement was performed imposing D4 symmetry. The crystal structure of NrnCBh was docked into the reconstructed cryo-EM density maps using the program package Phenix (Liebschner et al., 2019) and the models were refined in Coot (Emsley et al., 2010), ChimeraX (Pettersen et al., 2021), and the real-space refinement module for cryo-EM in Phenix (Afonine et al., 2018). Illustrations were prepared with ChimeraX and show the density of the sharpened maps. All data collection and refinement statistics are summarized in Table 4.2.

Table 4.2 – Cryo-EM model validation statistics

Model	**NrnC_{Bh}**	**NrnC_{Bh}**	**NrnC_{Bh}**	**NrnC_{Bh}**
Ligand	**pGG**	**pAGG**	**pAAAGG**	**pAAAGG/Ca^{2+}**
Symmetry	D4	D4	D4	D4
Composition (#)				
Protein Chains	8	8	8	8
Nucleotide Chains	8	8	8	8
Atoms	12760	11968	12296	12744
Protein Residues	1640	1592	1608	1640
Nucleotide	8 pGG	8 pGG	8 pGG	8 pGG
Bonds (RMSD)				
Length (Å) (# > 4σ)	0.007 (0)	0.003 (0)	0.006 (0)	0.007 (0)
Angles (°) (# > 4σ)	0.721 (0)	0.619 (0)	0.717 (0)	0.810 (0)
MolProbity score	1.03	0.97	0.8	1.01
Clash score	2.48	1.98	1.01	2.29
Ramachandran plot (%)				
Outliers	0.00	0.00	0.00	0.00
Allowed	0.49	1.03	0.51	0.99
Favored	99.51	98.97	99.49	99.01
Rama-Z: Ramachandran plot Z-score (RMSD)				
whole (N = 1754)	1.28 (0.20)	1.04 (0.21)	1.24 (0.21)	0.95 (0.20)
helix (N = 842)	0.73 (0.19)	0.79 (0.19)	0.74 (0.19)	0.78 (0.19)
sheet (N = 231)	0.72 (0.23)	0.50 (0.22)	0.56 (0.23)	0.48 (0.21)
loop (N = 681)	1.33 (0.26)	0.94 (0.28)	1.35 (0.28)	0.78 (0.26)
Rotamer outliers (%)	0.66	0	0	0
Cβ outliers (%)	0	0	0	0
CaBLAM outliers (%)	0.5	0.52	0.52	0.5
ADP (B-factors)				
Protein (min/max/mean)	31.65/64.33/43.25	23.08/54.02/35.37	22.76/59.16/38.31	21.29/58.88/31.57
Nucleotide (min/max/mean)	35.21/37.37/36.55	33.19/35.42/34.40	40.11/41.46/40.95	34.75/36.69/35.59
Resolution Estimates (Å)				
d FSC (half maps; 0.143)	3.2	3.2	2.9	3.0
d 99 (full/half1/half2)	3.3/4.4/4.4	3.2/4.0/4.0	2.7/2.4/2.5	3.0/3.9/3.9
d model	3.3	3.3	2.8	3.2
d FSC model (0/0.143/0.5)	3.0/3.1/3.2	2.8/3.0/3.2	2.5/2.6/2.9	2.8/2.8/3.1
Model vs. Data				
CC mask	0.86	0.83	0.87	0.86
CC box	0.81	0.75	0.72	0.79
CC peaks	0.77	0.72	0.69	0.75
CC volume	0.84	0.8	0.83	0.83
PDB ID	7MQB, EMD-23941	7MQD, EMD-23943	7MQF, EMD-23945	7MQH, EMD-23947

Table 4.2 – Cryo-EM model validation statistics (Continued).

Model	**NrnC$_{Bh}$**	**NrnC$_{Bh}$**	**NrnC$_{Bh}$**	**NrnC$_{Bh}$**
Ligand	**pGG**	**pAGG**	**pAAAGG**	**pAAAGG/Ca^{2+}**
Symmetry	C1	C1	C1	C1
Composition (#)				
Chains	8	8	8	8
Nucleotide Chains	8	8	8	8
Atoms	12556	12171	12678	12704
Protein Residues	1640	1629	1640	1640
Nucleotide Ligands	8 pGG	8 pGG	8 pGG	8 pGG
Bonds (RMSD)				
Length (Å) (# > 4σ)	0.007 (0)	0.008 (0)	0.006 (0)	0.006 (0)
Angles (°) (# > 4σ)	0.806 (0)	0.842 (0)	0.781 (0)	0.793 (0)
MolProbity score	0.86	0.9	1.08	1.06
Clash score	1.31	1.58	2.87	2.71
Ramachandran plot (%)				
Outliers	0.00	0.00	0.00	0.00
Allowed	0.49	1.74	1.17	1.35
Favored	99.51	98.26	98.83	98.65
Rama-Z: Ramachandran plot Z-score (RMSD)				
whole (N = 1754)	1.03 (0.20)	0.67 (0.20)	0.73 (0.20)	0.80 (0.20)
helix (N = 842)	0.57 (0.18)	0.34 (0.18)	0.27 (0.18)	0.50 (0.19)
sheet (N = 231)	0.44 (0.24)	0.22 (0.24)	0.55 (0.22)	0.59 (0.22)
loop (N = 681)	1.22 (0.26)	0.98 (0.27)	1.07 (0.28)	0.81 (0.26)
Rotamer outliers (%)	0.61	0.77	0	0.59
Cβ outliers (%)	0	0	0	0
CaBLAM outliers (%)	0.5	0.76	0.75	0.81
ADP (B-factors)				
Protein (min/max/mean)	30.00/71.38/49.55	30.00/80.81/50.16	19.92/70.75/36.57	31.09/95.82/52.94
Nucleotide (min/max/mean)	40.96/46.54/43.68	42.36/57.99/49.65	31.36/44.78/38.54	56.80/69.16/62.23
Resolution Estimates (Å)				
d FSC (half maps; 0.143)	3.6	3.7	3.2	3.2
d 99 (full/half1/half2)	3.1/3.0/3.0	3.2/2.5/2.5	3.0/2.6/2.6	3.0/2.5/2.5
d model	3.4	3.5	3.1	3.2
d FSC model (0/0.143/0.5)	2.9/3.1/3.6	2.9/3.1/3.6	2.8/2.9/3.2	2.7/2.8/3.2
Model vs. Data				
CC mask	0.86	0.81	0.82	0.81
CC box	0.81	0.75	0.72	0.67
CC peaks	0.77	0.69	0.67	0.62
CC volume	0.84	0.79	0.79	0.77
PDB ID	7MQC, EMD-23942	7MQE, EMD-23944	7MQG, EMD-23946	7MQI, EMD-23948

Complementation analysis in P. aeruginosa.

Deletion of orn in P. aeruginosa UCBPP-PA14 was performed using two-step allelic exchange as described by Hmelo and colleagues (Hmelo et al., 2015). Briefly, deletion alleles were created by overlap extension PCR, and delivered on a pEX18 suicide plasmid to the P. aeruginosa host strain by conjugation with E. coli donor strain S17.1 leading to the removal of the gene from the genome. To test for complementation, genes were introduced into the P. aeruginosa orn deletion strains by using electroporation (Choi et al., 2006). Briefly, P. aeruginosa cells were grown overnight, centrifuged, then washed with and resuspended in 300 mM sucrose. Expression 110 plasmids based on the pJHA vector were mixed with 100 µl of resuspended cells and electroporated using a Micropulser (Biorad) followed by recovery in 1 ml of lysogeny broth (LB), shaking at 250 rpm for 1 hour at 37 °C. Recovered cells were plated on LB plates containing 60 µg/ml gentamicin. Individual colonies were used for subsequent experiments.

Drip Assay.

The indicated P. aeruginosa strains harboring expression plasmids were grown overnight with shaking at 37 °C in LB supplemented with 60 µg/ml gentamicin. The cells were adjusted to CFU = 10,000 in fresh LB medium and applied to LB plates supplemented with 60 µg/ml gentamicin and 0.2% arabinose in 20 µl drops. The plates were inverted allowing the culture to drip down the length of the plate, followed by incubation overnight at 37 °C. The plates were imaged using a Chemidoc MP imager (Biorad) with a 0.2 s exposure time. The colony-measurer Python program (https://github.com/gwmarrah/colony-measurer) was employed to quantify the size of bacterial colonies by pixel measurement. Images were prepared for size quantification by cropping each lane of a drip plate as an individual 8-bit image. ImageJ (Schneider et al.,

2012) and colSizeMeasurer.py were used to determine the background pixel intensity and minimum/maximum colony sizes to be measured, respectively. These values were used to refine the parameters in colSizeAnalyzer.py for accurate measurement.

Immunoprecipitation and Western blot.

P. aeruginosa strains containing plasmid-borne NrnCBh-HA were grown overnight, followed by dilution to an OD600 = 0.1 in fresh LB supplemented with 60 µg/ml gentamicin. Cultures were allowed to grow to an OD600 = 0.8. Arabinose was added to a final concentration of 111 0.2% to induce protein expression for 2 hours at 37 °C. Following induction, cultures were normalized by OD, pelleted, and flash frozen in liquid nitrogen. Cells were resuspended in lysis buffer (150 mM NaCl, 25 mM Tris-Cl pH 7.5) followed by sonication. Anti-HA resin (Sigma) was prewashed with lysis buffer. Resin was added to the cleared lysate and incubated with rotation for 1 hour at 4 °C. Following binding, the HA resin was washed with lysis buffer, boiled in SDS loading buffer, and resolved by SDS-PAGE. Western blot transfer to a PVDF membrane proceeded for 90 minutes at constant 0.25 A, followed by overnight blocking with superblock (ThermoFisher) at 4 °C. Rabbit anti-HA primary antibody (Takara Bio) was diluted to 1:100 in TBS-T and incubated with the membrane for 1 hour at 20°C. Following washes with TBS-T, an HRP-conjugated, goat anti-rabbit antibody (GE Life Sciences) was diluted to 1:5000 in TBS-T and incubated with the membrane for 30 minutes at 20 °C. The membrane was washed with TBST before treating with SuperSignal West Femto (ThermoFisher) ECL reagent, followed by imaging with a BioRad Chemidoc system.

Size-exclusion chromatography-coupled multiangle light scattering (SEC-MALS).

Purified, wild-type or mutant variant NrnCBh at 2 mg/ml (85 µM) was injected onto a Superdex 200 Increase 10/300 gel filtration column (GE Healthcare) equilibrated with gel filtration buffer (25 mM Tris-Cl, pH 7.5, 150 mM NaCl). Size-exclusion chromatography was coupled to an in-line, static 18-angle light scattering detector (DAWN HELEOS-II, Wyatt Technology) and a refractive index detector (Optilab T-rEX, Wyatt Technology). Data were collected every second. Data analysis was performed with Astra 6.1 (Wyatt Technology) yielding the molar mass and mass distribution (polydispersity) of the sample. Monomeric BSA (Sigma; 5 mg/ml) was used as a control sample and to normalize the light scattering detectors.

Measuring dissociation constant (Kd) and binding specificity by DRaCALA.

To measure Kd, the purified protein was serially diluted in binding buffer (10 mM TrisHCl, pH 8, 100 mM NaCl, and 5 mM CaCl2). Each dilution was mixed with 32P-lableled substrate and spotted onto nitrocellulose. The dried nitrocellulose was exposed to a phosphorimager screens, scanned and analyzed as previously described (Roelofs et al., 2011). The fraction bound was plotted against protein concentration using the program Prism. For competition experiments to determine binding specificity, 100 µM of unlabeled nucleotides were mixed with 32P-labeled pGG; the mixtures were added 200 nM of purified NrnC.

Biochemical assay of RNase activity.

The reactions were performed by adding the indicated concentration of enzyme to the indicated concentration of substrate containing a tracer, consisting of 5’-end 32P-labeled substrate with the same length and sequence, in reaction buffer (10 mM Tris-HCl, pH 8, 100 mM NaCl, and 5 mM MgCl2). Reactions are stopped at indicated time points by the addition of 0.2 M EDTA. The samples were mixed with loading buffer (4 M Urea, 20% sucrose, 0.1% SDS, 0.05% bromophenol blue, 0.05% xylene cyanole FF, and 1x TBE), and separated by electrophoresis on 20% polyacrylamide gels.

DNase activity measurement.

DNase activity was assessed using an unspecific 1.5 kb PCR fragment, either with blunt ends or after restriction digestion with either KpnI, NdeI, or NotI (New England Biolabs). CIP (New England Biolabs) was used to dephosphorylate overhangs, and PNK (New England Biolabs) was used to phosphorylate blunt PCR products. NaCl, MgCl2, and MnCl2 were added to the 113 concentration indicated in the figures, with concentrations of NrnC and DNA at 1 µM and 20 nM, respectively, except where indicated otherwise in the figure. Reactions were incubated for 30 min or the indicated time at 37 °C. Reactions were stopped by the addition of a stop buffer containing proteinase K (Qiagen) and EDTA (JT Baker) to a final concentration of 0.1 mAU proteinase K and 10 mM EDTA, followed by incubation at 50 °C for 30 minutes. Samples were resolved by electrophoresis in a 1% agarose gels containing GelRed stain (Biotium) and imaged by UV visualization in a GelDocXR system (BioRad).

Identification of RNase homologs.

For each group of RNases, a list of seed protein sequences was manually prepared with Uniprot entry names: Orn (ORN_PSEAE, ORN_ECOLI, ORN_HUMAN, ORN_STRGR, ORN_CORDI, ORN_BURMA, ORN_YEAST, ORN_VIBCH), NrnA (NRNA_BACSU, NRNA_MYCPN, NRNA_THET8, NRNA_MYCTU, A0A3R9HUU0_STRSA), NrnB (NRNB_BACSU, A0A5C5X7X8_9BACT), NrnC (A9CG28_AGRFC, G4VUY7_9RHIZ, A1UU18_BARBK), RNase T (RNT_ECOLI, RNT_VIBCH, RNT_BUCAP, RNT_HAEIN, RNT_XYLFA), and RNase D (RND_ECOLI, Q9ZD81_RICPR, RND_HAEIN, I6XF17_MYCTU).

Initially a seed alignment was prepared for each RNase family using T-COFFEE (Notredame et al., 2000) in the Expresso structural alignment mode (Armougom et al., 2006). A search on the 2020_03 release of UniprotKB was performed with an HMM prepared from these alignments with hmmsearch from HMMER 3.3 (Eddy, 2011). Results were filtered by two criteria, a hit score above 125 and a template/query length ratio between 0.8 and 1.2. Search hits were used to construct a new multiple sequence alignment with MAFFT in E-INS-i mode (Katoh et al., 2005). 114 The hmmsearch was repeated with an HMM constructed from this expanded alignment and results were filtered with the same hit score cutoff but a more generous template/query length ratio between 0.6 and 1.5. The resulting hits were considered to be the final list for each RNase family, with any sequences found in multiple categories assigned to the category for which it had a higher hit score.

To determine whether a particular taxa contained homologs for each RNase group, the total number of proteins for each taxon in the NCBI taxon database present in our dataset were counted (Federhen, 2012). The average number of proteins per genome for each taxon was determined for all genomes annotated as a descendent taxon available in the NCBI genome database. Finally, the number of homologs found in each taxonomic category was calculated as a fraction of the total proteins in that taxon, multiplied by the average genome size to get an average presence-pergenome. For visualization purposes, values above 0.5 are considered present and non-zero values below 0.5 are potentially or partially present. The minimal species tree was extracted from the NCBI taxonomy database using ETE3 (Huerta-Cepas et al., 2016), followed by visualization of the resulting tree and annotation with iTOL (Letunic & Bork, 2019). All commands and code from this analysis were constructed as a reproducible SnakeMake pipeline (Köster & Rahmann,

2018) available at https://github.com/jgoodson/rnases (commit 66b2664 used in current versions of the figures).

Phylogenetics of DnaQ-family RNases.

To construct a phylogenetic tree of the DnaQ-fold RNases, the final sequences identified from the previous analysis for each family were clustered by sequence identity with MMSeqs2 (Steinegger & Söding, 2017). Targeting a final sequence count of 600, the sequence identity 115 threshold was determined for each family necessary to approximately maintain the original proportion of each family in the final representatives (30% for Orn, 45% for NrnC, 50% for RNase T, and 30% for RNase D). From these cluster representatives, a multiple sequence alignment was constructed using MAFFT in E-INS-i mode using DASH to obtain additional structural homologs (Rozewicki et al., 2019). The alignment was trimmed by removing the additional DASH sequences and columns with more than 90% gaps. The most appropriate evolutionary model was determined with IQ-TREE ModelFinder (LG+R8) and a phylogenetic tree was constructed using IQ-TREE 2.1.1 with 10,000 UFBoot replicates, and some modified parameters for expanded tree search (additional UFBoot NNI search, initial SPR search radius 20, 500 initial trees, initial search on 100 best trees, maintenance of the 50 best trees, and 500 iterations without improvement as stopping criteria) (Hoang et al., 2018; Kalyaanamoorthy et al., 2017; Minh et al., 2020). The tree was rooted with midpoint rooting on the long internal branch between RNase T/Orn and RNase D/NrnC. Sequence logos were created from monophyletic subgroup alignments using Logomaker (Tareen & Kinney, 2020).

Data deposition.

The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID codes 7MPL, 7MPM, 7MPN, 7MPO, 7MPP, 7MPQ,

7MPR, 7MPS, 7MPT, 7MPU, 7MQB/EMD-23941, 7MQD/EMD-23943, 7MQF/EMD-23945, 7MQH/EMD23947, 7MQC/EMD-23942, 7MQE/EMD-23944, 7MQG/EMD-23946, 7MQI/EMD-23948).

4.13 – Acknowledgements

This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). The Eiger 16M detector on 24-ID-E is funded by a NIH-ORIP HEI grant (S10OD021527). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Additional crystallographic research was conducted at the Center for High Energy X-ray Sciences (CHEXS), Cornell High Energy Synchrotron Source (CHESS), which is supported by the NSF under award DMR-1829070. The MacCHESS resource is supported by NIGMS award 1-P30-GM124166-01A1 and NYSTAR. This work made use of the Cornell Center for Materials Research Shared Facilities, which are supported through the NSF MRSEC program (DMR-1719875). The work was supported by NIH awards R01AI142400 (to. V.T.L.), R01GM123609 (to H.S.), and R35GM136258 (to J.C.F. and B.A.B.)

References (Chapters 1 - 3)

Anderson, Nadine S., and Charles Barlowe. 2019. “Conserved Juxtamembrane Domains in the Yeast Golgin Coy1 Drive Assembly of a Megadalton-Sized Complex and Mediate Binding to Tethering and SNARE Proteins.” *The Journal of Biological Chemistry* 294 (25): 9690–9705.

Antonny, B., S. Beraud-Dufour, P. Chardin, and M. Chabre. 1997. “N-Terminal Hydrophobic Residues of the G-Protein ADP-Ribosylation Factor-1 Insert into Membrane Phospholipids upon GDP to GTP Exchange.” *Biochemistry* 36 (15): 4675–84.

Antonny, Bruno, Christopher Burd, Pietro De Camilli, Elizabeth Chen, Oliver Daumke, Katja Faelber, Marijn Ford, et al. 2016. “Membrane Fission by Dynamin: What We Know and What We Need to Know.” *The EMBO Journal* 35 (21): 2270–84.

Arst, Herbert N., Jr, Miguel Hernandez-Gonzalez, Miguel A. Peñalva, and Areti Pantazopoulou. 2014. “GBF/Gea Mutant with a Single Substitution Sustains Fungal Growth in the Absence of BIG/Sec7.” *FEBS Letters* 588 (24): 4799–4806.

Bagde, Saket R., and J. Christopher Fromme. 2022. “Structure of a TRAPPII-Rab11 Activation Intermediate Reveals GTPase Substrate Selection Mechanisms.” *Science Advances* 8 (19): eabn7446.

Bannykh, Sergei I., and William E. Balch. 1997. “Membrane Dynamics at the Endoplasmic Reticulum–Golgi Interface.” *The Journal of Cell Biology* 138 (1): 1–4.

Barlow, Lael D., Eva Nývltová, Maria Aguilar, Jan Tachezy, and Joel B. Dacks. 2018. “A Sophisticated, Differentiated Golgi in the Ancestor of Eukaryotes.” *BMC Biology* 16 (1): 27.

Barr, Francis A. 2009. “Rab GTPase Function in Golgi Trafficking.” *Seminars in Cell & Developmental Biology* 20 (7): 780–83.

Barr, Francis, and David G. Lambright. 2010. “Rab GEFs and GAPs.” *Current Opinion in Cell Biology* 22 (4): 461–70.

Behnia, Rudy, Bojana Panic, James R. C. Whyte, and Sean Munro. 2004. “Targeting of the Arf-like GTPase Arl3p to the Golgi Requires N-Terminal Acetylation and the Membrane Protein Sys1p.” *Nature Cell Biology* 6 (5): 405–13.

Bepler, Tristan, Andrew Morin, Micah Rapp, Julia Brasch, Lawrence Shapiro, Alex J. Noble, and Bonnie Berger. 2019. “Positive-Unlabeled Convolutional Neural Networks for Particle Picking in Cryo-Electron Micrographs.” *Nature Methods* 16 (11): 1153–60.

Beraud-Dufour, Sophie, Sylviane Robineau, Pierre Chardin, Sonia Paris, Marc Chabre, Jacqueline Cherfils, and Bruno Antonny. 1998. “A Glutamic Finger in the Guanine Nucleotide Exchange Factor ARNO Displaces Mg2 and the B-Phosphate to Destabilize GDP on ARF1.” *Planning Perspectives: PP* 3651: 3659.

Béraud-Dufour, S., S. Paris, M. Chabre, and B. Antonny. 1999. “Dual Interaction of ADP Ribosylation Factor 1 with Sec7 Domain and with Lipid Membranes during Catalysis of Guanine Nucleotide Exchange.” *The Journal of Biological Chemistry* 274 (53): 37629–36.

Bezeljak, Urban, Hrushikesh Loya, Beata Kaczmarek, Timothy E. Saunders, and Martin Loose. 2020. “Stochastic Activation and Bistability in a Rab GTPase Regulatory Network.” *Proceedings of the National Academy of Sciences of the United States of America* 117 (12): 6540–49.

Bhatt, Jay M., Ekaterina G. Viktorova, Theodore Busby, Paulina Wyrozumska, Laura E. Newman, Helen Lin, Eunjoo Lee, et al. 2016. “Oligomerization of the Sec7 Domain Arf Guanine Nucleotide Exchange Factor GBF1 Is Dispensable for Golgi Localization and Function but Regulates Degradation.” *American Journal of Physiology. Cell Physiology* 310 (6): C456–69.

Bigay, Joëlle, and Bruno Antonny. 2012. “Curvature, Lipid Packing, and Electrostatics of Membrane Organelles: Defining Cellular Territories in Determining Specificity.” *Developmental Cell* 23 (5): 886–95.

Bi, Xiping, Richard A. Corpina, and Jonathan Goldberg. 2002. “Structure of the Sec23/24-Sar1 Pre-Budding Complex of the COPII Vesicle Coat.” *Nature* 419 (6904): 271–77.

Bi, Xiping, Joseph D. Mancias, and Jonathan Goldberg. 2007. “Insights into COPII Coat Nucleation from the Structure of Sec23.Sar1 Complexed with the Active Fragment of Sec31.” *Developmental Cell* 13 (5): 635–45.

Bonifacino, Juan S. 2014. “Vesicular Transport Earns a Nobel.” *Trends in Cell Biology* 24 (1): 3–5.

Bonifacino, Juan S., and Benjamin S. Glick. 2004. “The Mechanisms of Vesicle Budding and Fusion.” *Cell* 116 (2): 153–66.

Brachmann, C. B., A. Davies, G. J. Cost, E. Caputo, J. Li, P. Hieter, and J. D. Boeke. 1998. “Designer Deletion Strains Derived from Saccharomyces Cerevisiae S288C: A Useful Set of Strains and Plasmids for PCR-Mediated Gene Disruption and Other Applications.” *Yeast* 14 (2): 115–32.

Bretscher, M. S., and S. Munro. 1993. “Cholesterol and the Golgi Apparatus.” *Science* 261 (5126): 1280–81.

Brügger, B., R. Sandhoff, S. Wegehingel, K. Gorgas, J. Malsam, J. B. Helms, W. D. Lehmann, W. Nickel, and F. T. Wieland. 2000. “Evidence for Segregation of Sphingomyelin and Cholesterol during Formation of COPI-Coated Vesicles.” *The Journal of Cell Biology* 151 (3): 507–18.

Bui, Quynh Trang, Marie-Pierre Golinelli-Cohen, and Catherine L. Jackson. 2009. “Large Arf1 Guanine Nucleotide Exchange Factors: Evolution, Domain Structure, and Roles in Membrane Trafficking and Human Disease.” *Molecular Genetics and Genomics: MGG* 282 (4): 329–50.

Burri, Lena, and Trevor Lithgow. 2004. “A Complete Set of SNAREs in Yeast.” *Traffic* 5 (1): 45–52.

Buton, X., G. Morrot, P. Fellmann, and M. Seigneuret. 1996. “Ultrafast Glycerophospholipid-Selective Transbilayer Motion Mediated by a Protein in the Endoplasmic Reticulum Membrane.” *The Journal of Biological Chemistry* 271 (12): 6651–57.

Cai, Huaqing, Sidney Yu, Shekar Menon, Yiying Cai, Darina Lazarova, Chunmei Fu, Karin Reinisch, Jesse C. Hay, and Susan Ferro-Novick. 2007. “TRAPPI Tethers COPII Vesicles by Binding the Coat Subunit Sec23.” *Nature* 445 (7130): 941–44.

Cai, Yiying, Harvey F. Chin, Darina Lazarova, Shekar Menon, Chunmei Fu, Huaqing Cai, Anthony Sclafani, et al. 2008. “The Structural Basis for Activation of the Rab Ypt1p by the TRAPP Membrane-Tethering Complexes.” *Cell* 133 (7): 1202–13.

Casanova, James E. 2007. “Regulation of Arf Activation: The Sec7 Family of Guanine Nucleotide Exchange Factors.” *Traffic* 8 (11): 1476–85.

Casler, Jason C., and Benjamin S. Glick. 2020. “A Microscopy-Based Kinetic Analysis of Yeast Vacuolar Protein Sorting.” *eLife* 9 (June). https://doi.org/10.7554/eLife.56844.

Casler, Jason C., Natalie Johnson, Adam H. Krahn, Areti Pantazopoulou, Kasey J. Day, and Benjamin S. Glick. 2022. “Clathrin Adaptors Mediate Two Sequential Pathways of Intra-Golgi Recycling.” *The Journal of Cell Biology* 221 (1). https://doi.org/10.1083/jcb.202103199.

Casler, Jason C., Effrosyni Papanikou, Juan J. Barrero, and Benjamin S. Glick. 2019. “Maturation-Driven Transport and AP-1-Dependent Recycling of a Secretory Cargo in the Golgi.” *The Journal of Cell Biology* 218 (5): 1582–1601.

Casler, Jason C., Allison L. Zajac, Fernando M. Valbuena, Daniela Sparvoli, Okunola Jeyifous, Aaron P. Turkewitz, Sally Horne-Badovinac, William N. Green, and Benjamin S. Glick. 2020. “ESCargo: A Regulatable Fluorescent Secretory Cargo for Diverse Model Organisms.” *Molecular Biology of the Cell* 31 (26): 2892–2903.

Chen, C. Y., M. F. Ingram, P. H. Rosal, and T. R. Graham. 1999. “Role for Drs2p, a P-Type ATPase and Potential Aminophospholipid Translocase, in Yeast Late Golgi Function.” *The Journal of Cell Biology* 147 (6): 1223–36.

Chen, Kuan-Yu, Pei-Chin Tsai, Jia-Wei Hsu, Hsin-Chia Hsu, Chiung-Ying Fang, Lin-Chun Chang, Yueh-Tso Tsai, Chia-Jung Yu, and Fang-Jen S. Lee. 2010. “Syt1p Promotes Activation of Arl1p at the Late Golgi to Recruit Imh1p.” *Journal of Cell Science* 123 (Pt 20): 3478–89.

Chen, Yan-Ting, I-Hao Wang, Yi-Hsun Wang, Wan-Yun Chiu, Jen-Hao Hu, Wen-Hui

Chen, and Fang-Jen S. Lee. 2019. “Action of Arl1 GTPase and Golgin Imh1 in Ypt6-Independent Retrograde Transport from Endosomes to the Trans-Golgi Network.” *Molecular Biology of the Cell* 30 (8): 1008–19.

Christis, Chantal, and Sean Munro. 2012. “The Small G Protein Arl1 Directs the Trans-Golgi-Specific Targeting of the Arf1 Exchange Factors BIG1 and BIG2.” *The Journal of Cell Biology* 196 (3): 327–35.

Claude, A., B. P. Zhao, C. E. Kuziemsky, S. Dahan, S. J. Berger, J. P. Yan, A. D. Armold, E. M. Sullivan, and P. Melançon. 1999. “GBF1: A Novel Golgi-Associated BFA-Resistant Guanine Nucleotide Exchange Factor That Displays Specificity for ADP-Ribosylation Factor 5.” *The Journal of Cell Biology* 146 (1): 71–84.

Clausen, Henrik, Hans H. Wandall, Matthew P. DeLisa, Pamela Stanley, and Ronald L. Schnaar. n.d. “Glycosylation Engineering.” In *Essentials of Glycobiology*, edited by Ajit Varki, Richard D. Cummings, Jeffrey D. Esko, Pamela Stanley, Gerald W. Hart, Markus Aebi, Debra Mohnen, et al. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

Clermont, Y., A. Rambourg, and L. Hermo. 1995. “Trans-Golgi Network (TGN) of Different Cell Types: Three-Dimensional Structural Characteristics and Variability.” *The Anatomical Record* 242 (3): 289–301.

Cohen, F. S., and G. B. Melikyan. 2004. “The Energetics of Membrane Fusion from Binding, through Hemifusion, Pore Formation, and Pore Enlargement.” *The Journal of Membrane Biology* 199 (1): 1–14.

Coleman, R., and R. M. Bell. 1978. “Evidence That Biosynthesis of Phosphatidylethanolamine, Phosphatidylcholine, and Triacylglycerol Occurs on the Cytoplasmic Side of Microsomal Vesicles.” *The Journal of Cell Biology* 76 (1): 245–53.

Copeland, D. E., and A. J. Dalton. 1959. “An Association between Mitochondria and the Endoplasmic Reticulum in Cells of the Pseudobranch Gland of a Teleost.” *The Journal of Biophysical and Biochemical Cytology* 5 (3): 393–96.

Cowles, C. R., G. Odorizzi, G. S. Payne, and S. D. Emr. 1997. “The AP-3 Adaptor Complex Is Essential for Cargo-Selective Transport to the Yeast Vacuole.” *Cell* 91 (1): 109–18.

Dalton, A. J., and M. D. Felix. 1954. “Cytologic and Cytochemical Characteristics of the Golgi Substance of Epithelial Cells of the Epididymis in Situ, in Homogenates and after Isolation.” *The American Journal of Anatomy* 94 (2): 171–207.

Dancourt, Julia, Hong Zheng, Francesca Bottanelli, Edward S. Allgeyer, Joerg Bewersdorf, Morven Graham, Xinran Liu, James E. Rothman, and Grégory Lavieu. 2016. “Small Cargoes Pass through Synthetically Glued Golgi Stacks.” *FEBS Letters* 590 (12): 1675–86.

Day, Kasey J., Jason C. Casler, and Benjamin S. Glick. 2018. “Budding Yeast Has a Minimal Endomembrane System.” *Developmental Cell* 44 (1): 56–72.e4.

Deborde, Sylvie, Emilie Perret, Diego Gravotta, Ami Deora, Susana Salvarezza, Ryan Schreiner, and Enrique Rodriguez-Boulan. 2008. "Clathrin Is a Key Regulator of Basolateral Polarity." *Nature* 452 (7188): 719–23.

De Matteis, Maria Antonietta, and Anna Godi. 2004. "PI-Loting Membrane Traffic." *Nature Cell Biology* 6 (6): 487–92.

De Matteis, Maria Antonietta, and Alberto Luini. 2008. "Exiting the Golgi Complex." *Nature Reviews. Molecular Cell Biology* 9 (4): 273–84.

Deng, Yi, Marie-Pierre Golinelli-Cohen, Elena Smirnova, and Catherine L. Jackson. 2009. "A COPI Coat Subunit Interacts Directly with an Early-Golgi Localized Arf Exchange Factor." *EMBO Reports* 10 (1): 58–64.

Dirac-Svejstrup, A. Barbara, Tomoyuki Sumizawa, and Suzanne R. Pfeffer. 1997. "Identification of a GDI Displacement Factor That Releases Endosomal Rab GTPases from Rab–GDI." *The EMBO Journal* 16 (3): 465–72.

Dodonova, S. O., P. Diestelkoetter-Bachert, A. von Appen, W. J. H. Hagen, R. Beck, M. Beck, F. Wieland, and J. A. G. Briggs. 2015. "VESICULAR TRANSPORT. A Structure of the COPI Coat and the Role of Coat Proteins in Membrane Vesicle Assembly." *Science* 349 (6244): 195–98.

Donaldson, Julie G., and Catherine L. Jackson. 2011. "ARF Family G Proteins and Their Regulators: Roles in Membrane Transport, Development and Disease." *Nature Reviews. Molecular Cell Biology* 12 (6): 362–75.

Dröscher, A. 1998. "Camillo Golgi and the Discovery of the Golgi Apparatus." *Histochemistry and Cell Biology* 109 (5-6): 425–30.

Dunlop, Myun Hwa, Andreas M. Ernst, Lena K. Schroeder, Derek K. Toomre, Grégory Lavieu, and James E. Rothman. 2017. "Land-Locked Mammalian Golgi Reveals Cargo Transport between Stable Cisternae." *Nature Communications* 8 (1): 432.

Duronio, R. J., E. Jackson-Machelski, R. O. Heuckeroth, P. O. Olins, C. S. Devine, W. Yonemoto, L. W. Slice, S. S. Taylor, and J. I. Gordon. 1990. "Protein N-Myristoylation in Escherichia Coli: Reconstitution of a Eukaryotic Protein Modification in Bacteria." *Proceedings of the National Academy of Sciences of the United States of America* 87 (4): 1506–10.

Edeling, Melissa A., Corinne Smith, and David Owen. 2006. "Life of a Clathrin Coat: Insights from Clathrin and AP Structures." *Nature Reviews. Molecular Cell Biology* 7 (1): 32–44.

Faini, Marco, Simone Prinz, Rainer Beck, Martin Schorb, James D. Riches, Kirsten Bacia, Britta Brügger, Felix T. Wieland, and John A. G. Briggs. 2012. "The Structures of COPI-Coated Vesicles Reveal Alternate Coatomer Conformations and Interactions." *Science* 336 (6087): 1451–54.

Fairn, Gregory D., Nicole L. Schieber, Nicholas Ariotti, Samantha Murphy, Lars Kuerschner, Richard I. Webb, Sergio Grinstein, and Robert G. Parton. 2011. "High-

Resolution Mapping Reveals Topologically Distinct Cellular Pools of Phosphatidylserine." *The Journal of Cell Biology* 194 (2): 257–75.

Falkenburger, Björn H., Jill B. Jensen, Eamonn J. Dickson, Byung-Chang Suh, and Bertil Hille. 2010. "Phosphoinositides: Lipid Regulators of Membrane Proteins." *The Journal of Physiology* 588 (Pt 17): 3179–85.

Farquhar, M. G., and G. E. Palade. 1981. "The Golgi Apparatus (complex)-(1954-1981)-from Artifact to Center Stage." *The Journal of Cell Biology* 91 (3 Pt 2): 77s – 103s.

———. 1998. "The Golgi Apparatus: 100 Years of Progress and Controversy." *Trends in Cell Biology* 8 (1): 2–10.

Feyder, Serge, Johan-Owen De Craene, Séverine Bär, Dimitri L. Bertazzi, and Sylvie Friant. 2015. "Membrane Trafficking in the Yeast Saccharomyces Cerevisiae Model." *International Journal of Molecular Sciences* 16 (1): 1509–25.

Fölsch, Heike, Marc Pypaert, Sandra Maday, Laurence Pelletier, and Ira Mellman. 2003. "The AP-1A and AP-1B Clathrin Adaptor Complexes Define Biochemically and Functionally Distinct Membrane Domains." *The Journal of Cell Biology* 163 (2): 351–62.

Franco, M., P. Chardin, M. Chabre, and S. Paris. 1995. "Myristoylation of ADP-Ribosylation Factor 1 Facilitates Nucleotide Exchange at Physiological Mg2+ Levels." *The Journal of Biological Chemistry* 270 (3): 1337–41.

Franzusoff, A., K. Redding, J. Crosby, R. S. Fuller, and R. Schekman. 1991. "Localization of Components Involved in Protein Transport and Processing through the Yeast Golgi Apparatus." *The Journal of Cell Biology* 112 (1): 27–37.

Gabaldón, Toni, and Alexandros A. Pittis. 2015. "Origin and Evolution of Metabolic Sub-Cellular Compartmentalization in Eukaryotes." *Biochimie* 119 (December): 262–68.

Galindo, Antonio, Nicolas Soler, Stephen H. McLaughlin, Minmin Yu, Roger L. Williams, and Sean Munro. 2016. "Structural Insights into Arl1-Mediated Targeting of the Arf-GEF BIG1 to the Trans-Golgi." *Cell Reports* 16 (3): 839–50.

Gillingham, Alison K., and Sean Munro. 2007. "The Small G Proteins of the Arf Family and Their Regulators." *Annual Review of Cell and Developmental Biology* 23: 579–611.

Gillooly, D. J., I. C. Morrow, M. Lindsay, R. Gould, N. J. Bryant, J. M. Gaullier, R. G. Parton, and H. Stenmark. 2000. "Localization of Phosphatidylinositol 3-Phosphate in Yeast and Mammalian Cells." *The EMBO Journal* 19 (17): 4577–88.

Glick, Benjamin S., and Alberto Luini. 2011. "Models for Golgi Traffic: A Critical Assessment." *Cold Spring Harbor Perspectives in Biology* 3 (11): a005215.

Glick, B. S., and V. Malhotra. 1998. "The Curious Status of the Golgi Apparatus." *Cell*.

Godlee, Camilla, and Marko Kaksonen. 2013. "Review Series: From Uncertain

Beginnings: Initiation Mechanisms of Clathrin-Mediated Endocytosis." *The Journal of Cell Biology* 203 (5): 717–25.

Godman, G. C., and N. Lane. 1964. "ON THE SITE OF SULFATION IN THE CHONDROCYTE." *The Journal of Cell Biology* 21 (3): 353–66.

Goldberg, J. 1998. "Structural Basis for Activation of ARF GTPase: Mechanisms of Guanine Nucleotide Exchange and GTP-Myristoyl Switching." *Cell* 95 (2): 237–48.

Gomez-Navarro, Natalia, and Elizabeth Miller. 2016. "Protein Sorting at the ER-Golgi Interface." *The Journal of Cell Biology* 215 (6): 769–78.

Goody, Roger S., and Waltraud Hofmann-Goody. 2002. "Exchange Factors, Effectors, GAPs and Motor Proteins: Common Thermodynamic and Kinetic Principles for Different Functions." *European Biophysics Journal: EBJ* 31 (4): 268–74.

Goody, R. S., A. Rak, and K. Alexandrov. 2005. "The Structural and Mechanistic Basis for Recycling of Rab Proteins between Membrane Compartments." *Cellular and Molecular Life Sciences: CMLS* 62 (15): 1657–70.

Griffiths, G., and K. Simons. 1986. "The Trans Golgi Network: Sorting at the Exit Site of the Golgi Complex." *Science* 234 (4775): 438–43.

Guo, Xiaoli, Rafael Mattera, Xuefeng Ren, Yu Chen, Claudio Retamal, Alfonso González, and Juan S. Bonifacino. 2013. "The Adaptor Protein-1 μ1B Subunit Expands the Repertoire of Basolateral Sorting Signal Recognition in Epithelial Cells." *Developmental Cell* 27 (3): 353–66.

Gustafson, Margaret A., and J. Christopher Fromme. 2017. "Regulation of Arf Activation Occurs via Distinct Mechanisms at Early and Late Golgi Compartments." *Molecular Biology of the Cell* 28 (25): 3660–71.

Halaby, Steve L., and J. Christopher Fromme. 2018. "The HUS Box Is Required for Allosteric Regulation of the Sec7 Arf-GEF." *The Journal of Biological Chemistry* 293 (18): 6682–91.

Hammer, John A., 3rd, and Xufeng S. Wu. 2002. "Rabs Grab Motors: Defining the Connections between Rab GTPases and Motor Proteins." *Current Opinion in Cell Biology* 14 (1): 69–75.

Hammond, A. T., and B. S. Glick. 2000. "Dynamics of Transitional Endoplasmic Reticulum Sites in Vertebrate Cells." *Molecular Biology of the Cell* 11 (9): 3013–30.

Hara-Kuge, S., O. Kuge, L. Orci, M. Amherdt, M. Ravazzola, F. T. Wieland, and J. E. Rothman. 1994. "En Bloc Incorporation of Coatomer Subunits during the Assembly of COP-Coated Vesicles." *The Journal of Cell Biology* 124 (6): 883–92.

Harding, T. M., K. A. Morano, S. V. Scott, and D. J. Klionsky. 1995. "Isolation and Characterization of Yeast Mutants in the Cytoplasm to Vacuole Protein Targeting Pathway." *The Journal of Cell Biology* 131 (3): 591–602.

Helms, J. B., and J. E. Rothman. 1992. “Inhibition by Brefeldin A of a Golgi Membrane Enzyme That Catalyses Exchange of Guanine Nucleotide Bound to ARF.” *Nature* 360 (6402): 352–54.

Highland, Carolyn M., and J. Christopher Fromme. 2021. “Arf1 Directly Recruits the Pik1-Frq1 PI4K Complex to Regulate the Final Stages of Golgi Maturation.” *Molecular Biology of the Cell* 32 (10): 1064–80.

Highland, Carolyn M., Laura L. Thomas, and J. Christopher Fromme. 2023. “Methods for Studying Membrane-Proximal GAP Activity on Prenylated Rab GTPase Substrates.” *Methods in Molecular Biology* 2557: 507–18.

Hirschberg, K., C. M. Miller, J. Ellenberg, J. F. Presley, E. D. Siggia, R. D. Phair, and J. Lippincott-Schwartz. 1998. “Kinetic Analysis of Secretory Protein Traffic and Characterization of Golgi to Plasma Membrane Transport Intermediates in Living Cells.” *The Journal of Cell Biology* 143 (6): 1485–1503.

Holthuis, Joost C. M., and Tim P. Levine. 2005. “Lipid Traffic: Floppy Drives and a Superhighway.” *Nature Reviews. Molecular Cell Biology* 6 (3): 209–20.

Hong, J. X., F. J. Lee, W. A. Patton, C. Y. Lin, J. Moss, and M. Vaughan. 1998. “Phospholipid- and GTP-Dependent Activation of Cholera Toxin and Phospholipase D by Human ADP-Ribosylation Factor-like Protein 1 (HARL1).” *The Journal of Biological Chemistry* 273 (25): 15872–76.

Horiuchi, H., R. Lippé, H. M. McBride, M. Rubino, P. Woodman, H. Stenmark, V. Rybin, et al. 1997. “A Novel Rab5 GDP/GTP Exchange Factor Complexed to Rabaptin-5 Links Nucleotide Exchange to Effector Recruitment and Function.” *Cell* 90 (6): 1149–59.

Hughes, Helen, Annika Budnik, Katy Schmidt, Krysten J. Palmer, Judith Mantell, Chris Noakes, Andrew Johnson, et al. 2009. “Organisation of Human ER-Exit Sites: Requirements for the Localisation of Sec16 to Transitional ER.” *Journal of Cell Science* 122 (Pt 16): 2924–34.

Hutagalung, Alex H., and Peter J. Novick. 2011. “Role of Rab GTPases in Membrane Traffic and Cell Physiology.” *Physiological Reviews* 91 (1): 119–49.

Ignatev, Alexander, Sergey Kravchenko, Alexey Rak, Roger S. Goody, and Olena Pylypenko. 2008. “A Structural Model of the GDP Dissociation Inhibitor Rab Membrane Extraction Mechanism.” *The Journal of Biological Chemistry* 283 (26): 18377–84.

Jackson, Catherine L. 2014. “GEF-Effector Interactions.” *Cellular Logistics* 4 (2): e943616.

Jackson, C. L., and J. E. Casanova. 2000. “Turning on ARF: The Sec7 Family of Guanine-Nucleotide-Exchange Factors.” *Trends in Cell Biology* 10 (2): 60–67.

Jamieson, J. D., and G. E. Palade. 1967. “Intracellular Transport of Secretory Proteins in the Pancreatic Exocrine Cell. II. Transport to Condensing Vacuoles and Zymogen

Granules." *The Journal of Cell Biology* 34 (2): 597–615.

Jiang, Shangtong, Yanfang Li, Xian Zhang, Guojun Bu, Huaxi Xu, and Yun-Wu Zhang. 2014. "Trafficking Regulation of Proteins in Alzheimer's Disease." *Molecular Neurodegeneration* 9 (January): 6.

Jochum, Alexandra, David Jackson, Heinz Schwarz, Rüdiger Pipkorn, and Birgit Singer-Krüger. 2002. "Yeast Ysl2p, Homologous to Sec7 Domain Guanine Nucleotide Exchange Factors, Functions in Endocytosis and Maintenance of Vacuole Integrity and Interacts with the Arf-Like Small GTPase Arl1p." *Molecular and Cellular Biology* 22 (13): 4914–28.

Joiner, Aaron, and Chris Fromme. 2020. "Structural Details of GEF-mediated Sar1 Activation." *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 34 (S1): 1–1.

Joiner, Aaron M. N., and J. Christopher Fromme. 2021. "Structural Basis for the Initiation of COPII Vesicle Biogenesis." *Structure* 29 (8): 859–72.e6.

Joiner, Aaron Mn, Ben P. Phillips, Kumar Yugandhar, Ethan J. Sanford, Marcus B. Smolka, Haiyuan Yu, Elizabeth A. Miller, and J. Christopher Fromme. 2021. "Structural Basis of TRAPPIII-Mediated Rab1 Activation." *The EMBO Journal* 40 (12): e107607.

Jones, S., G. Jedd, R. A. Kahn, A. Franzusoff, F. Bartolini, and N. Segev. 1999. "Genetic Interactions in Yeast between Ypt GTPases and Arf Guanine Nucleotide Exchangers." *Genetics* 152 (4): 1543–56.

Karnkowska, Anna, Vojtěch Vacek, Zuzana Zubáčová, Sebastian C. Treitli, Romana Petrželková, Laura Eme, Lukáš Novák, et al. 2016. "A Eukaryote without a Mitochondrial Organelle." *Current Biology: CB* 26 (10): 1274–84.

Kattla, J. J., W. B. Struwe, M. Doherty, B. Adamczyk, R. Saldova, P. M. Rudd, and M. P. Campbell. 2011. "3.41 - Protein Glycosylation." In *Comprehensive Biotechnology (Third Edition)*, edited by Murray Moo-Young, 501–20. Oxford: Pergamon.

Kim, Jane J., Zhanna Lipatova, Uddalak Majumdar, and Nava Segev. 2016. "Regulation of Golgi Cisternal Progression by Ypt/Rab GTPases." *Developmental Cell* 36 (4): 440–52.

Kinman, Laurel F., Barrett M. Powell, Ellen D. Zhong, Bonnie Berger, and Joseph H. Davis. 2023. "Uncovering Structural Ensembles from Single-Particle Cryo-EM Data Using cryoDRGN." *Nature Protocols* 18 (2): 319–39.

Kleene, R., and E. G. Berger. 1993. "The Molecular and Cell Biology of Glycosyltransferases." *Biochimica et Biophysica Acta* 1154 (3-4): 283–325.

Klemm, Robin W., Christer S. Ejsing, Michal A. Surma, Hermann-Josef Kaiser, Mathias J. Gerl, Julio L. Sampaio, Quentin de Robillard, et al. 2009. "Segregation of Sphingolipids and Sterols during Formation of Secretory Vesicles at the Trans-Golgi Network." *The Journal of Cell Biology* 185 (4): 601–12.

Ladinsky, M. S., J. R. Kremer, P. S. Furcinitti, J. R. McIntosh, and K. E. Howell. 1994. "HVEM Tomography of the Trans-Golgi Network: Structural Insights and Identification of a Lace-like Vesicle Coat." *The Journal of Cell Biology* 127 (1): 29–38.

Lanoix, J., J. Ouwendijk, C. C. Lin, A. Stark, H. D. Love, J. Ostermann, and T. Nilsson. 1999. "GTP Hydrolysis by Arf-1 Mediates Sorting and Concentration of Golgi Resident Enzymes into Functional COP I Vesicles." *The EMBO Journal* 18 (18): 4935–48.

Lavieu, Gregory, Hong Zheng, and James E. Rothman. 2013. "Stapled Golgi Cisternae Remain in Place as Cargo Passes through the Stack." *eLife* 2 (June): e00558.

Lee, James Weifu. 2019. "Electrostatically Localized Proton Bioenergetics: Better Understanding Membrane Potential." *Heliyon* 5 (7): e01961.

Lee, Jonas Y., Hao Chen, Alan Liu, Benjamin M. Alba, and Ai Ching Lim. 2017. "Auto-Induction of Pichia Pastoris AOX1 Promoter for Membrane Protein Expression." *Protein Expression and Purification* 137 (September): 7–12.

Lewis, M. J., B. J. Nichols, C. Prescianotto-Baschong, H. Riezman, and H. R. Pelham. 2000. "Specific Retrieval of the Exocytic SNARE Snc1p from Early Yeast Endosomes." *Molecular Biology of the Cell* 11 (1): 23–38.

Li, Guangpu, and Hong Qian. 2002. "Kinetic Timing: A Novel Mechanism That Improves the Accuracy of GTPase Timers in Endosome Fusion and Other Biological Processes." *Traffic* 3 (4): 249–55.

Lippé, R., M. Miaczynska, V. Rybin, A. Runge, and M. Zerial. 2001. "Functional Synergy between Rab5 Effector Rabaptin-5 and Exchange Factor Rabex-5 When Physically Associated in a Complex." *Molecular Biology of the Cell* 12 (7): 2219–28.

Liu, Ke, Kavitha Surendhran, Steven F. Nothwehr, and Todd R. Graham. 2008. "P4-ATPase Requirement for AP-1/clathrin Function in Protein Transport from the Trans-Golgi Network and Early Endosomes." *Molecular Biology of the Cell* 19 (8): 3526–35.

Liu, Ming, Jordan Wright, Huan Guo, Yi Xiong, and Peter Arvan. 2014. "Proinsulin Entry and Transit through the Endoplasmic Reticulum in Pancreatic Beta Cells." *Vitamins and Hormones* 95: 35–62.

Losev, Eugene, Catherine A. Reinke, Jennifer Jellen, Daniel E. Strongin, Brooke J. Bevis, and Benjamin S. Glick. 2006. "Golgi Maturation Visualized in Living Yeast." *Nature* 441 (7096): 1002–6.

Malhotra, Vivek, Tito Serafini, Lelio Orci, James C. Shepherd, and James E. Rothman. 1989. "Purification of a Novel Class of Coated Vesicles Mediating Biosynthetic Protein Transport through the Golgi Stack." *Cell* 58 (2): 329–36.

Malsam, Jörg, Ayano Satoh, Laurence Pelletier, and Graham Warren. 2005. "Golgin Tethers Define Subpopulations of COPI Vesicles." *Science* 307 (5712): 1095–98.

Mandon, Elisabet C., Steven F. Trueman, and Reid Gilmore. 2013. “Protein Translocation across the Rough Endoplasmic Reticulum.” *Cold Spring Harbor Perspectives in Biology* 5 (2). https://doi.org/10.1101/cshperspect.a013342.

Mansour, S. J., J. Skaug, X. H. Zhao, J. Giordano, S. W. Scherer, and P. Melançon. 1999. “p200 ARF-GEP1: A Golgi-Localized Guanine Nucleotide Exchange Protein Whose Sec7 Domain Is Targeted by the Drug Brefeldin A.” *Proceedings of the National Academy of Sciences of the United States of America* 96 (14): 7968–73.

Mastronarde, David N. 2005. “Automated Electron Microscope Tomography Using Robust Prediction of Specimen Movements.” *Journal of Structural Biology* 152 (1): 36–51.

Matsuoka, K., L. Orci, M. Amherdt, S. Y. Bednarek, S. Hamamoto, R. Schekman, and T. Yeung. 1998. “COPII-Coated Vesicle Formation Reconstituted with Purified Coat Proteins and Chemically Defined Liposomes.” *Cell* 93 (2): 263–75.

Matsuura-Tokita, Kumi, Masaki Takeuchi, Akira Ichihara, Kenta Mikuriya, and Akihiko Nakano. 2006. “Live Imaging of Yeast Golgi Cisternal Maturation.” *Nature* 441 (7096): 1007–10.

Mattera, Rafael, Markus Boehm, Rittik Chaudhuri, Yogikala Prabhu, and Juan S. Bonifacino. 2011. “Conservation and Diversification of Dileucine Signal Recognition by Adaptor Protein (AP) Complex Variants.” *The Journal of Biological Chemistry* 286 (3): 2022–30.

McDonold, Caitlin M., and J. Christopher Fromme. 2014. “Four GTPases Differentially Regulate the Sec7 Arf-GEF to Direct Traffic at the Trans-Golgi Network.” *Developmental Cell* 30 (6): 759–67.

Mizuno-Yamasaki, Emi, Felix Rivera-Molina, and Peter Novick. 2012. “GTPase Networks in Membrane Traffic.” *Annual Review of Biochemistry* 81 (March): 637–59.

Moelleken, Jörg, Jörg Malsam, Matthew J. Betts, Ali Movafeghi, Ingeborg Reckmann, Ingrid Meissner, Andrea Hellwig, et al. 2007. “Differential Localization of Coatomer Complex Isoforms within the Golgi Apparatus.” *Proceedings of the National Academy of Sciences of the United States of America* 104 (11): 4425–30.

Morinaga, N., S. C. Tsai, J. Moss, and M. Vaughan. 1996. “Isolation of a Brefeldin A-Inhibited Guanine Nucleotide-Exchange Protein for ADP Ribosylation Factor (ARF) 1 and ARF3 That Contains a Sec7-like Domain.” *Proceedings of the National Academy of Sciences of the United States of America* 93 (23): 12856–60.

Morin-Ganet, M. N., A. Rambourg, S. B. Deitz, A. Franzusoff, and F. Képès. 2000. “Morphogenesis and Dynamics of the Yeast Golgi Apparatus.” *Traffic* 1 (1): 56–68.

Morre, J., L. M. Merlin, and T. W. Keenan. 1969. “Localization of Glycosyl Transferase Activities in a Golgi Apparatus-Rich Fraction Isolated from Rat Liver.” *Biochemical and Biophysical Research Communications* 37 (5): 813–19.

Mossessova, Elena, Richard A. Corpina, and Jonathan Goldberg. 2003. “Crystal Structure of ARF1*Sec7 Complexed with Brefeldin A and Its Implications for the Guanine Nucleotide Exchange Mechanism.” *Molecular Cell* 12 (6): 1403–11.

Mouratou, Barbara, Valerie Biou, Alexandra Joubert, Jean Cohen, David J. Shields, Niko Geldner, Gerd Jürgens, Paul Melançon, and Jacqueline Cherfils. 2005. “The Domain Architecture of Large Guanine Nucleotide Exchange Factors for the Small GTP-Binding Protein Arf.” *BMC Genomics* 6 (February): 20.

Muccini, Arnold J., Margaret A. Gustafson, and J. Christopher Fromme. 2022. “Structural Basis for Activation of Arf1 at the Golgi Complex.” *Cell Reports* 40 (9): 111282.

Mueller, A. G., M. Moser, R. Kluge, S. Leder, M. Blum, R. Büttner, H-G Joost, and A. Schürmann. 2002. “Embryonic Lethality Caused by Apoptosis during Gastrulation in Mice Lacking the Gene of the ADP-Ribosylation Factor-Related Protein 1.” *Molecular and Cellular Biology* 22 (5): 1488–94.

Müller, Matthias P., and Roger S. Goody. 2018. “Molecular Control of Rab Activity by GEFs, GAPs and GDI.” *Small GTPases* 9 (1-2): 5–21.

Munro, Sean. 2005. “The Golgi Apparatus: Defining the Identity of Golgi Membranes.” *Current Opinion in Cell Biology* 17 (4): 395–401.

Nakane, Takanori, Dari Kimanius, Erik Lindahl, and Sjors H. W. Scheres. 2018. “Characterisation of Molecular Motions in Cryo-EM Single-Particle Data by Multi-Body Refinement in RELION.” *eLife* 7 (June): e36861.

Neutra, M., and C. P. Leblond. 1966. “Synthesis of the Carbohydrate of Mucus in the Golgi Complex as Shown by Electron Microscope Radioautography of Goblet Cells from Rats Injected with Glucose-H3.” *The Journal of Cell Biology* 30 (1): 119–36.

Nickel, W., J. Malsam, K. Gorgas, M. Ravazzola, N. Jenne, J. B. Helms, and F. T. Wieland. 1998. “Uptake by COPI-Coated Vesicles of Both Anterograde and Retrograde Cargo Is Inhibited by GTPgammaS in Vitro.” *Journal of Cell Science* 111 (Pt 20) (October): 3081–90.

Nordmann, Mirjana, Margarita Cabrera, Angela Perz, Cornelia Bröcker, Clemens Ostrowicz, Siegfried Engelbrecht-Vandré, and Christian Ungermann. 2010. “The Mon1-Ccz1 Complex Is the GEF of the Late Endosomal Rab7 Homolog Ypt7.” *Current Biology: CB* 20 (18): 1654–59.

Novick, P., C. Field, and R. Schekman. 1980. “Identification of 23 Complementation Groups Required for Post-Translational Events in the Yeast Secretory Pathway.” *Cell* 21 (1): 205–15.

Novick, P., and R. Schekman. 1979. “Secretion and Cell-Surface Growth Are Blocked in a Temperature-Sensitive Mutant of Saccharomyces Cerevisiae.” *Proceedings of the National Academy of Sciences of the United States of America* 76 (4): 1858–62.

———. 1983. “Export of Major Cell Surface Proteins Is Blocked in Yeast Secretory Mutants.” *The Journal of Cell Biology* 96 (2): 541–47.

Orci, L., M. Stamnes, M. Ravazzola, M. Amherdt, A. Perrelet, T. H. Söllner, and J. E. Rothman. 1997. “Bidirectional Transport by Distinct Populations of COPI-Coated Vesicles.” *Cell* 90 (2): 335–49.

Ortiz, Darinel, Martina Medkova, Christiane Walch-Solimena, and Peter Novick. 2002. “Ypt32 Recruits the Sec4p Guanine Nucleotide Exchange Factor, Sec2p, to Secretory Vesicles; Evidence for a Rab Cascade in Yeast.” *The Journal of Cell Biology* 157 (6): 1005–15.

Paczkowski, Jon E., and J. Christopher Fromme. 2014. “Structural Basis for Membrane Binding and Remodeling by the Exomer Secretory Vesicle Cargo Adaptor.” *Developmental Cell* 30 (5): 610–24.

Paczkowski, Jon E., Brian C. Richardson, and J. Christopher Fromme. 2015. “Cargo Adaptors: Structures Illuminate Mechanisms Regulating Vesicle Biogenesis.” *Trends in Cell Biology* 25 (7): 408–16.

Panic, Bojana, James R. C. Whyte, and Sean Munro. 2003. “The ARF-like GTPases Arl1p and Arl3p Act in a Pathway That Interacts with Vesicle-Tethering Factors at the Golgi Apparatus.” *Current Biology: CB* 13 (5): 405–10.

Pantazopoulou, Areti, and Benjamin S. Glick. 2019. “A Kinetic View of Membrane Traffic Pathways Can Transcend the Classical View of Golgi Compartments.” *Frontiers in Cell and Developmental Biology* 7 (August): 153.

Papanikou, Effrosyni, and Benjamin S. Glick. 2009. “The Yeast Golgi Apparatus: Insights and Mysteries.” *FEBS Letters* 583 (23): 3746–51.

Parchure, Anup, Meng Tian, Danièle Stalder, Cierra K. Boyer, Shelby C. Bearrows, Kristen E. Rohli, Jianchao Zhang, et al. 2022. “Liquid-Liquid Phase Separation Facilitates the Biogenesis of Secretory Storage Granules.” *The Journal of Cell Biology* 221 (12). https://doi.org/10.1083/jcb.202206132.

Park, Sang Yoon, and Xiaoli Guo. 2014. “Adaptor Protein Complexes and Intracellular Transport.” *Bioscience Reports* 34 (4). https://doi.org/10.1042/BSR20140069.

Park, Sei-Kyoung, Lisa M. Hartnell, and Catherine L. Jackson. 2005. “Mutations in a Highly Conserved Region of the Arf1p Activator GEA2 Block Anterograde Golgi Transport but Not COPI Recruitment to Membranes.” *Molecular Biology of the Cell* 16 (8): 3786–99.

Park, Seung-Yeol, Jia-Shu Yang, Zhen Li, Pan Deng, Xiaohong Zhu, David Young, Maria Ericsson, et al. 2019. “The Late Stage of COPI Vesicle Fission Requires Shorter Forms of Phosphatidic Acid and Diacylglycerol.” *Nature Communications* 10 (1): 3409.

Partlow, Edward A., Kevin S. Cannon, Gunther Hollopeter, and Richard W. Baker. 2022. “Structural Basis of an Endocytic Checkpoint That Primes the AP2 Clathrin Adaptor for Cargo Internalization.” *Nature Structural & Molecular Biology* 29 (4): 339–47.

Payne, G. S., and R. Schekman. 1985. “A Test of Clathrin Function in Protein Secretion and Cell Growth.” *Science* 230 (4729): 1009–14.

Pearse, B. M. 1975. “Coated Vesicles from Pig Brain: Purification and Biochemical Characterization.” *Journal of Molecular Biology* 97 (1): 93–98.

Pelham, H. R. 1989. “Control of Protein Exit from the Endoplasmic Reticulum.” *Annual Review of Cell Biology* 5: 1–23.

Peter, Brian J., Helen M. Kent, Ian G. Mills, Yvonne Vallis, P. Jonathan G. Butler, Philip R. Evans, and Harvey T. McMahon. 2004. “BAR Domains as Sensors of Membrane Curvature: The Amphiphysin BAR Structure.” *Science* 303 (5657): 495–99.

Peyroche, A., S. Paris, and C. L. Jackson. 1996. “Nucleotide Exchange on ARF Mediated by Yeast Gea1 Protein.” *Nature* 384 (6608): 479–81.

Pfeffer, Suzanne R. 2012. “Rab GTPase Localization and Rab Cascades in Golgi Transport.” *Biochemical Society Transactions* 40 (6): 1373–77.

Poon, P. P., S. F. Nothwehr, R. A. Singer, and G. C. Johnston. 2001. “The Gcs1 and Age2 ArfGAP Proteins Provide Overlapping Essential Function for Transport from the Yeast Trans-Golgi Network.” *The Journal of Cell Biology* 155 (7): 1239–50.

Presley, John F., Theresa H. Ward, Andrea C. Pfeifer, Eric D. Siggia, Robert D. Phair, and Jennifer Lippincott-Schwartz. 2002. “Dissection of COPI and Arf1 Dynamics in Vivo and Role in Golgi Membrane Transport.” *Nature* 417 (6885): 187–93.

Preuss, D., J. Mulholland, A. Franzusoff, N. Segev, and D. Botstein. 1992. “Characterization of the Saccharomyces Golgi Complex through the Cell Cycle by Immunoelectron Microscopy.” *Molecular Biology of the Cell* 3 (7): 789–803.

Punjani, Ali, John L. Rubinstein, David J. Fleet, and Marcus A. Brubaker. 2017. “cryoSPARC: Algorithms for Rapid Unsupervised Cryo-EM Structure Determination.” *Nature Methods* 14 (3): 290–96.

Ramaen, Odile, Alexandra Joubert, Philip Simister, Naïma Belgareh-Touzé, Maria Conception Olivares-Sanchez, Jean-Christophe Zeeh, Sophie Chantalat, et al. 2007. “Interactions between Conserved Domains within Homodimers in the BIG1, BIG2, and GBF1 Arf Guanine Nucleotide Exchange Factors.” *The Journal of Biological Chemistry* 282 (39): 28834–42.

Rambourg, A., Y. Clermont, and L. Hermo. 1979. “Three-Dimensional Architecture of the Golgi Apparatus in Sertoli Cells of the Rat.” *The American Journal of Anatomy* 154 (4): 455–76.

Randazzo, P. A., O. Weiss, and R. A. Kahn. 1992. “Preparation of Recombinant ADP-Ribosylation Factor.” *Methods in Enzymology* 219: 362–69.

Reggiori, Fulvio, Chao-Wen Wang, Usha Nair, Takahiro Shintani, Hagai Abeliovich, and Daniel J. Klionsky. 2004. “Early Stages of the Secretory Pathway, but Not Endosomes, Are Required for Cvt Vesicle and Autophagosome Assembly in Saccharomyces Cerevisiae.” *Molecular Biology of the Cell* 15 (5): 2189–2204.

Renault, Louis, Bernard Guibert, and Jacqueline Cherfils. 2003. “Structural Snapshots of the Mechanism and Inhibition of a Guanine Nucleotide Exchange Factor.” *Nature* 426 (6966): 525–30.

Richardson, Brian C., and J. Christopher Fromme. 2012. “Autoregulation of Sec7 Arf-GEF Activity and Localization by Positive Feedback.” *Small GTPases* 3 (4): 240–43.

———. 2015. “Biochemical Methods for Studying Kinetic Regulation of Arf1 Activation by Sec7.” *Methods in Cell Biology* 130 (June): 101–26.

Richardson, Brian C., Steve L. Halaby, Margaret A. Gustafson, and J. Christopher Fromme. 2016. “The Sec7 N-Terminal Regulatory Domains Facilitate Membrane-Proximal Activation of the Arf1 GTPase.” *eLife* 5 (January). https://doi.org/10.7554/eLife.12411.

Richardson, Brian C., Caitlin M. McDonold, and J. Christopher Fromme. 2012. “The Sec7 Arf-GEF Is Recruited to the Trans-Golgi Network by Positive Feedback.” *Developmental Cell* 22 (4): 799–810.

Rivera-Molina, Félix E., and Peter J. Novick. 2009. “A Rab GAP Cascade Defines the Boundary between Two Rab GTPases on the Secretory Pathway.” *Proceedings of the National Academy of Sciences of the United States of America* 106 (34): 14408–13.

Robinson, J. S., D. J. Klionsky, L. M. Banta, and S. D. Emr. 1988. “Protein Sorting in Saccharomyces Cerevisiae: Isolation of Mutants Defective in the Delivery and Processing of Multiple Vacuolar Hydrolases.” *Molecular and Cellular Biology* 8 (11): 4936–48.

Roth, J., D. J. Taatjes, J. M. Lucocq, J. Weinstein, and J. C. Paulson. 1985. “Demonstration of an Extensive Trans-Tubular Network Continuous with the Golgi Apparatus Stack That May Function in Glycosylation.” *Cell* 43 (1): 287–95.

Rothman, J. E. 1994. “Mechanisms of Intracellular Protein Transport.” *Nature* 372 (6501): 55–63.

Roth, T. F., and K. R. Porter. 1964. “YOLK PROTEIN UPTAKE IN THE OOCYTE OF THE MOSQUITO AEDES AEGYPTI. L.” *The Journal of Cell Biology* 20 (2): 313–32.

Sacher, M., J. Barrowman, W. Wang, J. Horecka, Y. Zhang, M. Pypaert, and S. Ferro-Novick. 2001. “TRAPP I Implicated in the Specificity of Tethering in ER-to-Golgi Transport.” *Molecular Cell* 7 (2): 433–42.

Salminen, A., and P. J. Novick. 1987. “A Ras-like Protein Is Required for a Post-Golgi Event in Yeast Secretion.” *Cell* 49 (4): 527–38.

Sapperstein, S. K., V. V. Lupashin, H. D. Schmitt, and M. G. Waters. 1996. “Assembly of the ER to Golgi SNARE Complex Requires Uso1p.” *The Journal of Cell Biology* 132 (5): 755–67.

Schneiter, R., B. Brügger, R. Sandhoff, G. Zellnig, A. Leber, M. Lampl, K. Athenstaedt, et al. 1999. “Electrospray Ionization Tandem Mass Spectrometry (ESI-MS/MS) Analysis of the Lipid Molecular Species Composition of Yeast Subcellular Membranes Reveals Acyl Chain-Based Sorting/remodeling of Distinct Molecular Species En Route to the Plasma Membrane.” *The Journal of Cell Biology* 146 (4): 741–54.

Schoberer, Jennifer, John Runions, Herta Steinkellner, Richard Strasser, Chris Hawes, and Anne Osterrieder. 2010. “Sequential Depletion and Acquisition of Proteins during Golgi Stack Disassembly and Reformation.” *Traffic* 11 (11): 1429–44.

Scorrano, Luca, Maria Antonietta De Matteis, Scott Emr, Francesca Giordano, György Hajnóczky, Benoît Kornmann, Laura L. Lackner, et al. 2019. “Coming Together to Define Membrane Contact Sites.” *Nature Communications* 10 (1): 1287.

Scott, S. V., A. Hefner-Gravink, K. A. Morano, T. Noda, Y. Ohsumi, and D. J. Klionsky. 1996. “Cytoplasm-to-Vacuole Targeting and Autophagy Employ the Same Machinery to Deliver Proteins to the Yeast Vacuole.” *Proceedings of the National Academy of Sciences of the United States of America* 93 (22): 12304–8.

Setty, Subba Rao Gangi, Marcus E. Shin, Atsuko Yoshino, Michael S. Marks, and Christopher G. Burd. 2003. “Golgi Recruitment of GRIP Domain Proteins by Arf-like GTPase 1 Is Regulated by Arf-like GTPase 3.” *Current Biology: CB* 13 (5): 401–4.

Setty, Subba Rao Gangi, Todd I. Strochlic, Amy Hin Yan Tong, Charles Boone, and Christopher G. Burd. 2004. “Golgi Targeting of ARF-like GTPase Arl3p Requires Its Nalpha-Acetylation and the Integral Membrane Protein Sys1p.” *Nature Cell Biology* 6 (5): 414–19.

Sharpe, Hayley J., Tim J. Stevens, and Sean Munro. 2010. “A Comprehensive Comparison of Transmembrane Domains Reveals Organelle-Specific Properties.” *Cell* 142 (1): 158–69.

Sikorski, R. S., and P. Hieter. 1989. “A System of Shuttle Vectors and Yeast Host Strains Designed for Efficient Manipulation of DNA in Saccharomyces Cerevisiae.” *Genetics* 122 (1): 19–27.

Sivars, Ulf, Dikran Aivazian, and Suzanne R. Pfeffer. 2003. “Yip3 Catalyses the Dissociation of Endosomal Rab–GDI Complexes.” *Nature* 425 (6960): 856–59.

Spang, A., J. M. Herrmann, S. Hamamoto, and R. Schekman. 2001. “The ADP Ribosylation Factor-Nucleotide Exchange Factors Gea1p and Gea2p Have Overlapping, but Not Redundant Functions in Retrograde Transport from the Golgi to the Endoplasmic Reticulum.” *Molecular Biology of the Cell* 12 (4): 1035–45.

Stagg, Scott M., Cemal Gürkan, Douglas M. Fowler, Paul LaPointe, Ted R. Foss, Clinton S. Potter, Bridget Carragher, and William E. Balch. 2006. “Structure of the Sec13/31 COPII Coat Cage.” *Nature* 439 (7073): 234–38.

Stagg, Scott M., Paul LaPointe, Abbas Razvi, Cemal Gürkan, Clinton S. Potter, Bridget Carragher, and William E. Balch. 2008. “Structural Basis for Cargo Regulation of COPII Coat Assembly.” *Cell* 134 (3): 474–84.

Stalder, Danièle, and Bruno Antonny. 2013. “Arf GTPase Regulation through Cascade Mechanisms and Positive Feedback Loops.” *FEBS Letters* 587 (13): 2028–35.

Stearns, T., R. A. Kahn, D. Botstein, and M. A. Hoyt. 1990. “ADP Ribosylation Factor Is an Essential Protein in Saccharomyces Cerevisiae and Is Encoded by Two Genes.” *Molecular and Cellular Biology* 10 (12): 6690–99.

Stearns, T., M. C. Willingham, D. Botstein, and R. A. Kahn. 1990. “ADP-Ribosylation Factor Is Functionally and Physically Associated with the Golgi Complex.” *Proceedings of the National Academy of Sciences of the United States of America* 87 (3): 1238–42.

Stepp, J. D., K. Huang, and S. K. Lemmon. 1997. “The Yeast Adaptor Protein Complex, AP-3, Is Essential for the Efficient Delivery of Alkaline Phosphatase by the Alternate Pathway to the Vacuole.” *The Journal of Cell Biology* 139 (7): 1761–74.

Subczynski, Witold K., Marta Pasenkiewicz-Gierula, Justyna Widomska, Laxman Mainali, and Marija Raguz. 2017. “High Cholesterol/Low Cholesterol: Effects in Biological Membranes: A Review.” *Cell Biochemistry and Biophysics* 75 (3-4): 369–85.

Suda, Yasuyuki, Kazuo Kurokawa, Ryogo Hirata, and Akihiko Nakano. 2013. “Rab GAP Cascade Regulates Dynamics of Ypt6 in the Golgi Traffic.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (47): 18976–81.

Südhof, Thomas C., and James E. Rothman. 2009. “Membrane Fusion: Grappling with SNARE and SM Proteins.” *Science* 323 (5913): 474–77.

Sztul, Elizabeth, Pei-Wen Chen, James E. Casanova, Jacqueline Cherfils, Joel B. Dacks, David G. Lambright, Fang-Jen S. Lee, et al. 2019. “ARF GTPases and Their GEFs and GAPs: Concepts and Challenges.” *Molecular Biology of the Cell* 30 (11): 1249–71.

Tan, Dongyan, Yiying Cai, Juan Wang, Jinzhong Zhang, Shekar Menon, Hui-Ting Chou, Susan Ferro-Novick, Karin M. Reinisch, and Thomas Walz. 2013. “The EM Structure of the TRAPPIII Complex Leads to the Identification of a Requirement for COPII Vesicles on the Macroautophagy Pathway.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (48): 19432–37.

Thomas, Laura L., and J. Christopher Fromme. 2016. “GTPase Cross Talk Regulates TRAPPII Activation of Rab11 Homologues during Vesicle Biogenesis.” *The Journal of Cell Biology* 215 (4): 499–513.

Thomas, Laura L., Carolyn M. Highland, and J. Christopher Fromme. 2021. “Arf1 Orchestrates Rab GTPase Conversion at the Trans-Golgi Network.” *Molecular Biology of the Cell* 32 (11): 1104–20.

Thomas, Laura L., Aaron M. N. Joiner, and J. Christopher Fromme. 2018. “The TRAPPIII Complex Activates the GTPase Ypt1 (Rab1) in the Secretory Pathway.” *The Journal of Cell Biology* 217 (1): 283–98.

Thomas, Laura L., Solveig A. van der Vegt, and J. Christopher Fromme. 2019. “A Steric Gating Mechanism Dictates the Substrate Specificity of a Rab-GEF.” *Developmental Cell* 48 (1): 100–114.e9.

Tokarev, Andrei A., Aixa Alfonso, and Nava Segev. 2013. *Overview of Intracellular Compartments and Trafficking Pathways*. Landes Bioscience.

Torres, Isabel L., Cláudia Rosa-Ferreira, and Sean Munro. 2014. “The Arf Family G Protein Arl1 Is Required for Secretory Granule Biogenesis in Drosophila.” *Journal of Cell Science* 127 (Pt 10): 2151–60.

Tschopp, J., P. C. Esmon, and R. Schekman. 1984. “Defective Plasma Membrane Assembly in Yeast Secretory Mutants.” *Journal of Bacteriology* 160 (3): 966–70.

Tsukada, M., and D. Gallwitz. 1996. “Isolation and Characterization of SYS Genes from Yeast, Multicopy Suppressors of the Functional Loss of the Transport GTPase Ypt6p.” *Journal of Cell Science* 109 (Pt 10) (October): 2471–81.

Ueyama, Takehiko, Mika Eto, Keiichiro Kami, Toshihiko Tatsuno, Toshihiro Kobayashi, Yasuhito Shirai, Michelle R. Lennartz, Ryu Takeya, Hideki Sumimoto, and Naoaki Saito. 2005. “Isoform-Specific Membrane Targeting Mechanism of Rac during Fc Gamma R-Mediated Phagocytosis: Positive Charge-Dependent and Independent Targeting Mechanism of Rac to the Phagosome.” *Journal of Immunology* 175 (4): 2381–90.

Valdivia, Raphael H., Daniel Baggott, John S. Chuang, and Randy W. Schekman. 2002. “The Yeast Clathrin Adaptor Protein Complex 1 Is Required for the Efficient Retention of a Subset of Late Golgi Membrane Proteins.” *Developmental Cell* 2 (3): 283–94.

Walch-Solimena, C., R. N. Collins, and P. J. Novick. 1997. “Sec2p Mediates Nucleotide Exchange on Sec4p and Is Involved in Polarized Delivery of Post-Golgi Vesicles.” *The Journal of Cell Biology* 137 (7): 1495–1509.

Wang, Chao-Wen, Susan Hamamoto, Lelio Orci, and Randy Schekman. 2006. “Exomer: A Coat Complex for Transport of Select Membrane Proteins from the Trans-Golgi Network to the Plasma Membrane in Yeast.” *The Journal of Cell Biology* 174 (7): 973–83.

Wang, I-Hao, Yi-Jie Chen, Jia-Wei Hsu, and Fang Jen Lee. 2017. “The Arl3 and Arl1 GTPases Co-Operate with Cog8 to Regulate Selective Autophagy via Atg9 Trafficking.” *Traffic* 18 (9): 580–89.

Wang, Rong, Zhijing Wang, Kaikai Wang, Tianlong Zhang, and Jianping Ding. 2016. “Structural Basis for Targeting BIG1 to Golgi Apparatus through Interaction of Its DCB Domain with Arl1.” *Journal of Molecular Cell Biology* 8 (5): 459–61.

Wang, Ying Jie, Jing Wang, Hui Qiao Sun, Manuel Martinez, Yu Xiao Sun, Eric Macia, Tomas Kirchhausen, Joseph P. Albanesi, Michael G. Roth, and Helen L. Yin. 2003. “Phosphatidylinositol 4 Phosphate Regulates Targeting of Clathrin Adaptor AP-1 Complexes to the Golgi.” *Cell* 114 (3): 299–310.

Welch, Lawrence G., and Sean Munro. 2019. “A Tale of Short Tails, through Thick and Thin: Investigating the Sorting Mechanisms of Golgi Enzymes.” *FEBS Letters* 593 (17): 2452–65.

Whur, P., A. Herscovics, and C. P. Leblond. 1969. “Radioautographic Visualization of the Incorporation of Galactose-3H and Mannose-3H by Rat Thyroids in Vitro in Relation to the Stages of Thyroglobulin Synthesis.” *The Journal of Cell Biology* 43 (2): 289–311.

Wicky, Sidonie, Heinz Schwarz, and Birgit Singer-Krüger. 2004. “Molecular Interactions of Yeast Neo1p, an Essential Member of the Drs2 Family of Aminophospholipid Translocases, and Its Role in Membrane Trafficking within the Endomembrane System.” *Molecular and Cellular Biology* 24 (17): 7402–18.

Wong, Mie, and Sean Munro. 2014. “Membrane Trafficking. The Specificity of Vesicle Traffic to the Golgi Is Encoded in the Golgin Coiled-Coil Proteins.” *Science* 346 (6209): 1256898.

Wright, John, Richard A. Kahn, and Elizabeth Sztul. 2014. “Regulating the Large Sec7 ARF Guanine Nucleotide Exchange Factors: The When, Where and How of Activation.” *Cellular and Molecular Life Sciences: CMLS* 71 (18): 3419–38.

Wu, Mousheng, Lei Lu, Wanjin Hong, and Haiwei Song. 2004. “Structural Basis for Recruitment of GRIP Domain Golgin-245 by Small GTPase Arl1.” *Nature Structural & Molecular Biology* 11 (1): 86–94.

Wurmser, A. E., and S. D. Emr. 1998. “Phosphoinositide Signaling and Turnover: PtdIns(3)P, a Regulator of Membrane Traffic, Is Transported to the Vacuole and Degraded by a Process That Requires Lumenal Vacuolar Hydrolase Activities.” *The EMBO Journal* 17 (17): 4930–42.

Wu, Yao-Wen, Lena K. Oesterlin, Kui-Thong Tan, Herbert Waldmann, Kirill Alexandrov, and Roger S. Goody. 2010. “Membrane Targeting Mechanism of Rab GTPases Elucidated by Semisynthetic Protein Probes.” *Nature Chemical Biology* 6 (7): 534–40.

Yang, Feng, Tiantian Li, Ziqing Peng, Yang Liu, and Yusong Guo. 2020. “The Amphipathic Helices of Arfrp1 and Arl14 Are Sufficient to Determine Subcellular Localizations.” *The Journal of Biological Chemistry* 295 (49): 16643–54.

Yang, Jia-Shu, Stella Y. Lee, Stefania Spanò, Helge Gad, Leiliang Zhang, Zhongzhen Nie, Matteo Bonazzi, Daniela Corda, Alberto Luini, and Victor W. Hsu. 2005. “A Role for BARS at the Fission Step of COPI Vesicle Formation from Golgi Membrane.” *The EMBO Journal* 24 (23): 4133–43.

Yarwood, Rebecca, John Hellicar, Philip G. Woodman, and Martin Lowe. 2020. “Membrane Trafficking in Health and Disease.” *Disease Models & Mechanisms* 13 (4). https://doi.org/10.1242/dmm.043448.

Yeung, Tony, Mauricio Terebiznik, Liming Yu, John Silvius, Wasif M. Abidi, Mark Philips, Tim Levine, Andras Kapus, and Sergio Grinstein. 2006. “Receptor Activation Alters

Inner Surface Potential during Phagocytosis.” *Science* 313 (5785): 347–51.

Yorimitsu, Tomohiro, Ken Sato, and Masaki Takeuchi. 2014. “Molecular Mechanisms of Sar/Arf GTPases in Vesicular Trafficking in Yeast and Plants.” *Frontiers in Plant Science* 5 (August): 411.

Zanetti, Giulia, Simone Prinz, Sebastian Daum, Annette Meister, Randy Schekman, Kirsten Bacia, and John A. G. Briggs. 2013. “The Structure of the COPII Transport-Vesicle Coat Assembled on Membranes.” *eLife* 2 (September): e00951.

Zhang, Kai. 2016. “Gctf: Real-Time CTF Determination and Correction.” *Journal of Structural Biology* 193 (1): 1–12.

Zhao, Shan, Chun Man Li, Xiao Min Luo, Gavin Ka Yu Siu, Wen Jia Gan, Lin Zhang, William K. K. Wu, Hsiao Chang Chan, and Sidney Yu. 2017. “Mammalian TRAPPIII Complex Positively Modulates the Recruitment of Sec13/31 onto COPII Vesicles.” *Scientific Reports* 7 (February): 43207.

Zheng, Shawn Q., Eugene Palovcak, Jean-Paul Armache, Kliment A. Verba, Yifan Cheng, and David A. Agard. 2017. “MotionCor2: Anisotropic Correction of Beam-Induced Motion for Improved Cryo-Electron Microscopy.” *Nature Methods* 14 (4): 331–32.

Zhu, Huaiping, Zhimin Liang, and Guangpu Li. 2009. “Rabex-5 Is a Rab22 Effector and Mediates a Rab22-Rab5 Signaling Cascade in Endocytosis.” *Molecular Biology of the Cell* 20 (22): 4720–29.

Zivanov, Jasenko, Takanori Nakane, and Sjors H. W. Scheres. 2020. “Estimation of High-Order Aberrations and Anisotropic Magnification from Cryo-EM Data Sets in RELION-3.1.” *IUCrJ* 7 (Pt 2): 253–67.

References (Chapter 4)

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr, 66(Pt 2), 213-221. https://doi.org/10.1107/s0907444909052925

Afonine, P. V., Poon, B. K., Read, R. J., Sobolev, O. V., Terwilliger, T. C., Urzhumtsev, A., & Adams, P. D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr D Struct Biol, 74(Pt 6), 531-544. https://doi.org/10.1107/s2059798318006551

Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., & Notredame, C. (2006). Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. Nucleic Acids Res, 34(Web Server issue), W604-608. https://doi.org/10.1093/nar/gkl092

Chin, K. H., Yang, C. Y., Chou, C. C., Wang, A. H., & Chou, S. H. (2006). The crystal structure of XC847 from Xanthomonas campestris: a 3'-5' oligoribonuclease of DnaQ fold family with a novel opposingly shifted helix. Proteins, 65(4), 1036-1040. https://doi.org/10.1002/prot.21148

Choi, K. H., Kumar, A., & Schweizer, H. P. (2006). A 10-min method for preparation of highly electrocompetent Pseudomonas aeruginosa cells: application for DNA fragment transfer between chromosomes and plasmid transformation. J Microbiol Methods, 64(3), 391-397. https://doi.org/10.1016/j.mimet.2005.06.001

Christen, B., Abeliuk, E., Collier, J. M., Kalogeraki, V. S., Passarelli, B., Coller, J. A., Fero, M. J., McAdams, H. H., & Shapiro, L. (2011). The essential genome of a bacterium. Mol Syst Biol, 7, 528. https://doi.org/10.1038/msb.2011.58

Cohen, D., Mechold, U., Nevenzal, H., Yarmiyhu, Y., Randall, T. E., Bay, D. C., Rich, J. D., Parsek, M. R., Kaever, V., Harrison, J. J., & Banin, E. (2015). Oligoribonuclease is a central feature of cyclic diguanylate signaling in Pseudomonas aeruginosa. Proc Natl Acad Sci U S A, 112(36), 11359-11364. https://doi.org/10.1073/pnas.1421450112

Datta, A. K., & Niyogi, K. (1975). A novel oligoribonuclease of Escherichia coli. II. Mechanism of action. J Biol Chem, 250(18), 7313-7319. Dehio, C. (2005). Bartonella-host-cell interactions and vascular tumour formation. Nat Rev Microbiol, 3(8), 621-631. https://doi.org/10.1038/nrmicro1209

Druzhinin, S. Y., Tran, N. T., Skalenko, K. S., Goldman, S. R., Knoblauch, J. G., Dove, S. L., & Nickels, B. E. (2015). A Conserved Pattern of Primer-Dependent Transcription

Initiation in Escherichia coli and Vibrio cholerae Revealed by 5' RNA-seq. PLoS Genet, 11(7), e1005348. https://doi.org/10.1371/journal.pgen.1005348

Eddy, S. R. (2011). Accelerated Profile HMM Searches. PLoS Comput Biol, 7(10), e1002195. https://doi.org/10.1371/journal.pcbi.1002195

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res, 32(5), 1792-1797. https://doi.org/10.1093/nar/gkh340

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr, 66(Pt 4), 486-501. https://doi.org/10.1107/s0907444910007493

Evans, P. (2006). Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr, 62(Pt 1), 72-82. https://doi.org/10.1107/s0907444905036693

Fang, M., Zeisberg, W. M., Condon, C., Ogryzko, V., Danchin, A., & Mechold, U. (2009). Degradation of nanoRNA is performed by multiple redundant RNases in Bacillus subtilis. Nucleic Acids Res, 37(15), 5114-5125. https://doi.org/10.1093/nar/gkp527

Federhen, S. (2012). The NCBI Taxonomy database. Nucleic Acids Res, 40(Database issue), D136-143. https://doi.org/10.1093/nar/gkr1178

Galperin, M. Y., & Koonin, E. V. (2012). Divergence and convergence in enzyme evolution. J Biol Chem, 287(1), 21-28. https://doi.org/10.1074/jbc.R111.241976

Ghosh, S., & Deutscher, M. P. (1999). Oligoribonuclease is an essential component of the mRNA decay pathway. Proc Natl Acad Sci U S A, 96(8), 4372-4377. https://doi.org/10.1073/pnas.96.8.4372

Goldman, S. R., Sharp, J. S., Vvedenskaya, I. O., Livny, J., Dove, S. L., & Nickels, B. E. (2011). NanoRNAs prime transcription initiation in vivo. Mol Cell, 42(6), 817-825. https://doi.org/10.1016/j.molcel.2011.06.005

Greenfield, R. A., Drevets, D. A., Machado, L. J., Voskuhl, G. W., Cornea, P., & Bronze, M. S. (2002). Bacterial pathogens as biological weapons and agents of bioterrorism. Am J Med Sci, 323(6), 299-315. https://doi.org/10.1097/00000441-200206000-00003

Hmelo, L. R., Borlee, B. R., Almblad, H., Love, M. E., Randall, T. E., Tseng, B. S., Lin, C., Irie, Y., Storek, K. M., Yang, J. J., Siehnel, R. J., Howell, P. L., Singh, P. K., Tolker-Nielsen, T., Parsek, M. R., Schweizer, H. P., & Harrison, J. J. (2015). Precision-engineering the Pseudomonas aeruginosa genome with two-step allelic exchange. Nat Protoc, 10(11), 820-1841. https://doi.org/10.1038/nprot.2015.115

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol Biol Evol, 35(2), 518-522. https://doi.org/10.1093/molbev/msx281

Huerta-Cepas, J., Serra, F., & Bork, P. (2016). ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. Mol Biol Evol, 33(6), 1635-1638. https://doi.org/10.1093/molbev/msw046

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., Mende, D. R., Letunic, I., Rattei, T., Jensen, L. J., von Mering, C., & Bork, P. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res, 47(D1), D309- d314. https://doi.org/10.1093/nar/gky1085

Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr D Biol Crystallogr, 66(Pt 2), 133-144. https://doi.org/10.1107/s0907444909047374

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods, 14(6), 587-589. https://doi.org/10.1038/nmeth.4285

Kamp, H. D., Patimalla-Dipali, B., Lazinski, D. W., Wallace-Gadsden, F., & Camilli, A. (2013). Gene fitness landscapes of Vibrio cholerae at important stages of its life cycle. PLoS Pathog, 9(12), e1003800. https://doi.org/10.1371/journal.ppat.1003800

Katoh, K., Kuma, K., Toh, H., & Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res, 33(2), 511-518. https://doi.org/10.1093/nar/gki198

Keegan, R. M., McNicholas, S. J., Thomas, J. M. H., Simpkin, A. J., Simkovic, F., Uski, V., Ballard, C. C., Winn, M. D., Wilson, K. S., & Rigden, D. J. (2018). Recent developments in MrBUMP: better search-model preparation, graphical interaction with search models, and solution improvement and assessment. Acta Crystallogr D Struct Biol, 74(Pt 3), 167-182. https://doi.org/10.1107/s2059798318003455

Kim, S. K., Lormand, J. D., Weiss, C. A., Eger, K. A., Turdiev, H., Turdiev, A., Winkler, W. C., Sondermann, H., & Lee, V. T. (2019). A dedicated diribonucleotidase resolves a key bottleneck for the terminal step of RNA degradation. Elife, 8. https://doi.org/10.7554/eLife.46313

Köster, J., & Rahmann, S. (2018). Snakemake-a scalable bioinformatics workflow engine. Bioinformatics, 34(20), 3600. https://doi.org/10.1093/bioinformatics/bty350

Krissinel, E., & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J Mol Biol, 372(3), 774-797. https://doi.org/10.1016/j.jmb.2007.05.022

Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res, 47(W1), W256-w259. https://doi.org/10.1093/nar/gkz239

Liao, H., Liu, M., & Guo, X. (2018). The special existences: nanoRNA and nanoRNase. Microbiol Res, 207, 134-139. https://doi.org/10.1016/j.micres.2017.11.014

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J., & Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol, 75(Pt 10), 861- 877. https://doi.org/10.1107/s2059798319011471

Liu, M. F., Cescau, S., Mechold, U., Wang, J., Cohen, D., Danchin, A., Boulouis, H. J., & Biville, F. (2012). Identification of a novel nanoRNase in Bartonella. Microbiology (Reading), 158(Pt 4), 886-895. https://doi.org/10.1099/mic.0.054619-0

Mechold, U., Fang, G., Ngo, S., Ogryzko, V., & Danchin, A. (2007). YtqI from Bacillus subtilis has both oligoribonuclease and pAp-phosphatase activity. Nucleic Acids Res, 35(13), 4552- 4561. https://doi.org/10.1093/nar/gkm462

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol, 37(5), 1530-1534. https://doi.org/10.1093/molbev/msaa015

Morin, A., Eisenbraun, B., Key, J., Sanschagrin, P. C., Timony, M. A., Ottaviano, M., & Sliz, P. (2013). Collaboration gets the most out of software. Elife, 2, e01456. https://doi.org/10.7554/eLife.01456

Newman, J. R., & Fuqua, C. (1999). Broad-host-range expression vectors that carry the L-arabinose-inducible Escherichia coli araBAD promoter and the araC regulator. Gene, 227(2), 197-203. https://doi.org/10.1016/s0378-1119(98)00601-5

Nicholls, T. J., Spåhr, H., Jiang, S., Siira, S. J., Koolmeister, C., Sharma, S., Kauppila, J. H. K., Jiang, M., Kaever, V., Rackham, O., Chabes, A., Falkenberg, M., Filipovska, A., Larsson, N. G., & Gustafsson, C. M. (2019). Dinucleotide Degradation by REXO2 Maintains Promoter Specificity in Mammalian Mitochondria. Mol Cell, 76(5), 784-796.e786. https://doi.org/10.1016/j.molcel.2019.09.010

Niyogi, S. K., & Datta, A. K. (1975). A novel oligoribonuclease of Escherichia coli. I. Isolation and properties. J Biol Chem, 250(18), 7307-7312.

Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol, 302(1), 205-217. https://doi.org/10.1006/jmbi.2000.4042

Omelchenko, M. V., Galperin, M. Y., Wolf, Y. I., & Koonin, E. V. (2010). Non-homologous isofunctional enzymes: a systematic analysis of alternative solutions in enzyme evolution. Biol Direct, 5, 31. https://doi.org/10.1186/1745-6150-5-31

Orr, M. W., Donaldson, G. P., Severin, G. B., Wang, J., Sintim, H. O., Waters, C. M., & Lee, V. T. (2015). Oligoribonuclease is the primary degradative enzyme for pGpG in Pseudomonas aeruginosa that is required for cyclic-di-GMP turnover. Proc Natl Acad Sci U S A, 112(36), E5048-5057. https://doi.org/10.1073/pnas.1507245112

Orr, M. W., Weiss, C. A., Severin, G. B., Turdiev, H., Kim, S. K., Turdiev, A., Liu, K., Tu, B. P., Waters, C. M., Winkler, W. C., & Lee, V. T. (2018). A Subset of Exoribonucleases Serve as Degradative Enzymes for pGpG in c-di-GMP Signaling. J Bacteriol, 200(24). https://doi.org/10.1128/jb.00300-18

Palace, S. G., Proulx, M. K., Lu, S., Baker, R. E., & Goguen, J. D. (2014). Genome-wide mutant fitness profiling identifies nutritional requirements for optimal growth of Yersinia pestis in deep tissue. mBio, 5(4). https://doi.org/10.1128/mBio.01385-14

Pappas, G., Papadimitriou, P., Akritidis, N., Christou, L., & Tsianos, E. V. (2006). The new global map of human brucellosis. Lancet Infect Dis, 6(2), 91-99. https://doi.org/10.1016/s1473-3099(06)70382-6

Park, J., Lee, S. Y., Jeong, H., Kang, M. G., Van Haute, L., Minczuk, M., Seo, J. K., Jun, Y., Myung, K., Rhee, H. W., & Lee, C. (2019). The structure of human EXD2 reveals a chimeric 3' to 5' exonuclease domain that discriminates substrates via metal coordination. Nucleic Acids Res, 47(13), 7078-7093. https://doi.org/10.1093/nar/gkz454

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci, 30(1), 70-82. https://doi.org/10.1002/pro.3943

Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods, 14(3), 290-296. https://doi.org/10.1038/nmeth.4169

Roelofs, K. G., Wang, J., Sintim, H. O., & Lee, V. T. (2011). Differential radial capillary action of ligand assay for high-throughput detection of protein-metabolite interactions. Proc Natl Acad Sci U S A, 108(37), 15528-15533. https://doi.org/10.1073/pnas.1018949108

Rosta, E., Yang, W., & Hummer, G. (2014). Calcium inhibition of ribonuclease H1 two-metal ion catalysis. J Am Chem Soc, 136(8), 3137-3144. https://doi.org/10.1021/ja411408x

Rozewicki, J., Li, S., Amada, K. M., Standley, D. M., & Katoh, K. (2019). MAFFT-DASH: integrated protein sequence and structural alignment. Nucleic Acids Res, 47(W1), W5-w10. https://doi.org/10.1093/nar/gkz342

Schmier, B. J., Nelersa, C. M., & Malhotra, A. (2017). Structural Basis for the Bidirectional Activity of Bacillus nanoRNase NrnA. Sci Rep, 7(1), 11085. https://doi.org/10.1038/s41598-017-09403-x

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods, 9(7), 671-675. https://doi.org/10.1038/nmeth.2089

Steinegger, M., & Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol, 35(11), 1026-1028. https://doi.org/10.1038/nbt.3988

Sternon, J. F., Godessart, P., Gonçalves de Freitas, R., Van der Henst, M., Poncin, K., Francis, N., Willemart, K., Christen, M., Christen, B., Letesson, J. J., & De Bolle, X. (2018). Transposon Sequencing of Brucella abortus Uncovers Essential Genes for

Growth In Vitro and Inside Macrophages. Infect Immun, 86(8). https://doi.org/10.1128/iai.00312-18

Tareen, A., & Kinney, J. B. (2020). Logomaker: beautiful sequence logos in Python. Bioinformatics, 36(7), 2272-2274. https://doi.org/10.1093/bioinformatics/btz921

Vvedenskaya, I. O., Sharp, J. S., Goldman, S. R., Kanabar, P. N., Livny, J., Dove, S. L., & Nickels, B. E. (2012). Growth phase-dependent control of transcription start site selection and gene expression by nanoRNAs. Genes Dev, 26(13), 1498-1507. https://doi.org/10.1101/gad.192732.112

Wasmuth, E. V., Januszyk, K., & Lima, C. D. (2014). Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA. Nature, 511(7510), 435-439. https://doi.org/10.1038/nature13406

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr, 67(Pt 4), 235-242. https://doi.org/10.1107/s0907444910045749

Yu, D., & Deutscher, M. P. (1995). Oligoribonuclease is distinct from the other known exoribonucleases of Escherichia coli. J Bacteriol, 177(14), 4137-4139. https://doi.org/10.1128/jb.177.14.4137-4139.1995

Yuan, Z., Gao, F., Yin, K., & Gu, L. (2018). NrnC, an RNase D-Like Protein From Agrobacterium, Is a Novel Octameric Nuclease That Specifically Degrades dsDNA but Leaves dsRNA Intact. Front Microbiol, 9, 3230. https://doi.org/10.3389/fmicb.2018.03230

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J Struct Biol, 193(1), 1-12. https://doi.org/10.1016/j.jsb.2015.11.003

Zhang, X., Zhu, L., & Deutscher, M. P. (1998). Oligoribonuclease is encoded by a highly conserved gene in the 3'-5' exonuclease superfamily. J Bacteriol, 180(10), 2779-2781. https://doi.org/10.1128/jb.180.10.2779-2781.1998

Zheng, S. Q., Palovcak, E., Armache, J. P., Verba, K. A., Cheng, Y., & Agard, D. A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods, 14(4), 331-332. https://doi.org/10.1038/nmeth.4193

Zivanov, J., Nakane, T., Forsberg, B. O., Kimanius, D., Hagen, W. J., Lindahl, E., & Scheres, S. H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife, 7. https://doi.org/10.7554/eLife.42166

Zuo, Y., & Deutscher, M. P. (2001). Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res, 29(5), 1017-1026. https://doi.org/10.1093/nar/29.5.1017

Zuo, Y., Wang, Y., & Malhotra, A. (2005). Crystal structure of Escherichia coli RNase D, an exoribonuclease involved in structured RNA processing. Structure, 13(7), 973-984. https://doi.org/10.1016/j.str.2005.04.015