UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Mechanistic Analyses of Peroxisome-Related AAA+ Motors

Permalink
https://escholarship.org/uc/item/6t62x18z

Author
Castanzo, Dominic

Publication Date
2020

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Mechanistic Analyses of Peroxisome-Related AAA+ Motors

By

Dominic Castanzo

A dissertation in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Andreas Martin, Chair
Professor James Hurley
Professor Gloria Brar
Professor Arash Komeili

Summer 2020

ABSTRACT

Mechanistic analyses of peroxisome-related AAA+ motors

By

Dominic T Castanzo

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Professor Andreas Martin, Chair

AAA+ ATPases hydrolyze ATP to perform diverse cellular functions. While these complex machines have been the subject of biochemical studies for years, the details on what these motors contribute to cellular processes, how these motors perform work on their substrates, and the identify of their substrates remains unknown for many AAA+ ATPases. Specifically, it is unclear how the activity of two such enzymes, Pex1/Pex6 and Msp1, brings about their cellular functions. It is known that Pex1/Pex6 is essential for peroxisome biogenesis, but the identity of Pex1/Pex6's substrate is unknown. Furthermore, an understanding of how Pex1/Pex6 processes this substrate, and how this processing contributes to peroxisome biogenesis, is lacking. Msp1 is a poorly studied AAA+ ATPase embedded in the outer-mitochondrial membrane. It is responsible for extracting mislocalized tail-anchored proteins from this surface, but the mechanism of action it uses for extraction and how it recognizes mislocalized proteins remain unclear. Furthermore, although Msp1 has been shown to be present at peroxisomes, details regarding how its function may differ from its mitochondrial duties are unexplored. For my dissertation, I designed and established *in vitro* assays to study the biochemical mechanisms of both Pex1/Pex6 and Msp1.

To better understand Pex1/Pex6's role at the peroxisomal membrane, I focused on characterizing its primarily peroxisomal binding partner, Pex15. I solved the crystal structure of the soluble core domain of Pex15, and exploited its biochemical features to develop an assay that measures its mechanical unfolding. Furthermore, I demonstrated that the Pex1/Pex6 ATPase is capable of unfolding the soluble Pex15 domain, constituting the first direct evidence that Pex1/Pex6 unfolds substrates. I utilized these assays to define both, motor mechanism and substrate requirements to establish that the intrinsically disordered regions of Pex15 play crucial roles in Pex1/Pex6 interaction and engagement.

Although Pex15 functions as a peroxisomal tether for Pex1/Pex6, it occasionally mislocalizes to the outer-mitochondrial membrane, where it becomes a substrate for Msp1. Msp1 is believed to function as a hexamer; however, it fails to oligomerize when the N-terminal transmembrane domain is removed, presenting a challenge to *in vitro* characterization of this motor. I developed a system to study Msp1 activity *in vitro* by fusing it to a hexamerizing scaffold to encourage oligomerization. I used this system to

characterize Msp1's mechanism of action, provided direct evidence that it is capable of unfolding proteins by processive threading, and showed that it is promiscuous in substrate selection. Lastly, I demonstrated that Pex3 directly inhibits Msp1 unfoldase activity. Overall, my work helped elucidate AAA+ motor function and further our understanding of peroxisome biogenesis and proteome quality control at the peroxisome and mitochondria.

This work is dedicated to my wife, Caitlin. Without your constant love and support, this would not have been possible.

TABLE OF CONTENTS

TABLE OF FIGURES

Chapter 1: Introduction to protein localization and AAA+ ATPases

Chapter 2: Probing Pex15's domain architecture and structure

Chapter 3: Pex1/Pex6 unfolds Pex15 by processive threading

Chapter 4: The AAA+ ATPase Msp1 is a processive protein translocase with robust unfoldase activity

CHAPTER 1

Introduction to protein localization and AAA+ ATPases

1.1 The importance and benefits of subcellular compartments

Cells are not amorphous and disorganized bags of proteins, and the benefits of localizing, concentrating, and sequestering certain proteins or molecules is evidenced by the prolific evolution of prokaryotic and eukaryotic organelles. These organelles serve a variety of purposes to the organism. For one, localization within an organelle can increase the effective concentration of proteins and metabolites without the need to produce or import a greater number of such molecules into the cell. Concentrating proteins and metabolites can benefit cellular processes by increasing the efficiency of biochemical reactions to rates which would otherwise be unobtainable if the reactants were free to diffuse throughout the entire cell volume. Another benefit of organelles is the ability to establish unique chemical environments with properties distinct from the cytosol, varying in such characteristics as pH or oxidation/reduction potential. Compartmentalizing these unique microcosms is important both in maintaining the specific functions of that organelle and in protecting the rest of the cell from exposure to environments that are more acidic, oxidizing, etc. Beyond chemical environments, membrane-bound organelles can quarantine enzymatic and chemical reactions with toxic byproducts. For example, the ß-oxidation of very long chain fatty acids that occurs within human peroxisomes produces toxic hydrogen peroxide as a byproduct; not only do peroxisomes sequester the reactions in a compartment separate from other essential and delicate cellular processes (e.g. DNA maintenance, replication, and repair in the nucleus, or mRNA translation in the cytosol) but they also import and concentrate the enzymatic machinery necessary to break down hydrogen peroxide molecules (Frederick & Newcomb, 1969). As a result, peroxisomes are able to host chemical reactions that are both essential and potentially hazardous, without jeopardizing the health of the cell. Through these and other benefits, the evolution of subcellular organization and compartmentalization allows any single cell to perform the great diversity of enzymatic and chemical processes necessary for life.

1.2 An overview of transmembrane domain-containing protein targeting pathways

The function of each organelle derives from the architecture, luminal environment, and population of proteins localized to the surface and within the organelle. Thus, while subcellular compartmentalization affords cells the ability to create a variety of functionally distinct organelles, it also presents a new challenge: targeting thousands of structurally and biochemically diverse proteins to the proper organelle. Following translation by cytosolic ribosomes, soluble proteins can make their way to specific organelles by binding other organelle-localizing or -anchored proteins. However, post-translational release into the cytosol is not an option for many transmembrane-domain (TMD) containing proteins that, if not guided to the proper membrane, may form aggregates in the cytosol and become toxic to other cellular processes.

Perhaps one of the best characterized protein targeting mechanisms is the Signal Recognition Particle (SRP) pathway, whereby the SRP protein binds a signal sequence on a nascent polypeptide chain and stalls translation until the ribosome has been ushered to the ER membrane (Fig. 1.1). Once docked with the appropriate

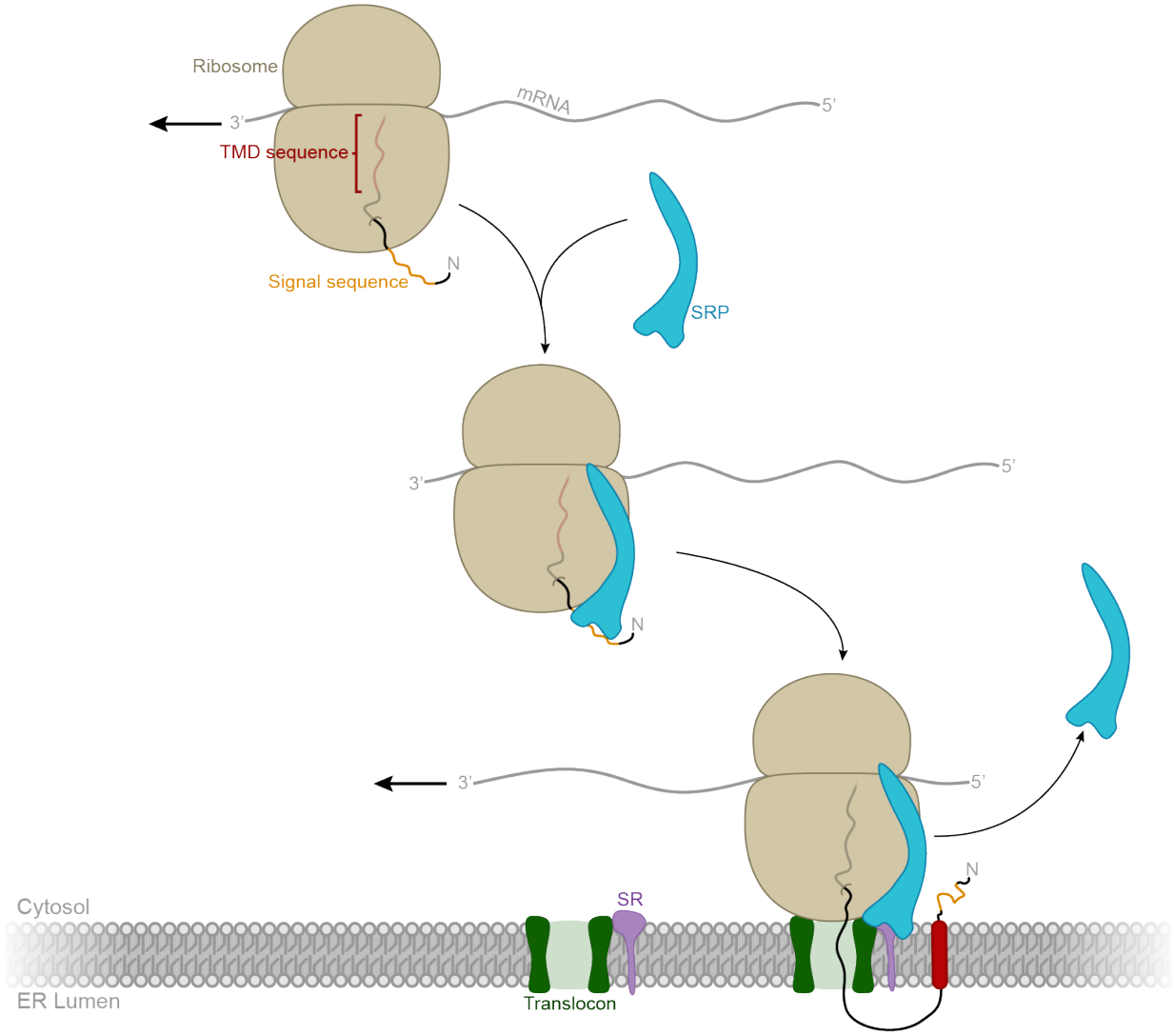


Figure 1.1 : The SRP-dependent co-translational protein targeting pathway. As the N-terminal signal sequence emerges from the exit tunnel it is detected and bound by SRP, which simultaneously acts to pause translation before the TMD-corresponding sequence has been translated or extruded. SRP guides the stalled ribosome to the ER, where the former binds the SRP receptor (SR) protein. SR associates with the translocon and, upon SRP release, translation is resumed. With the help of the translocon, the nascent protein's TMD is inserted into the ER membrane.

complexes at the ER, translation resumes and the newly-synthesized protein is delivered directly into the ER lumen or membrane (Egea et al., 2005). The SRP pathway is responsible for efficiently targeting proteins with N-terminal TMDs, while the burden of guiding proteins with TMDs in the middle of their sequence falls to the analogous SRP-independent (SND) pathway, which utilizes much of the same protein complexes at the ER membrane (Fig. 1.2; Noriega et al., 2014; Aviram et al., 2016).

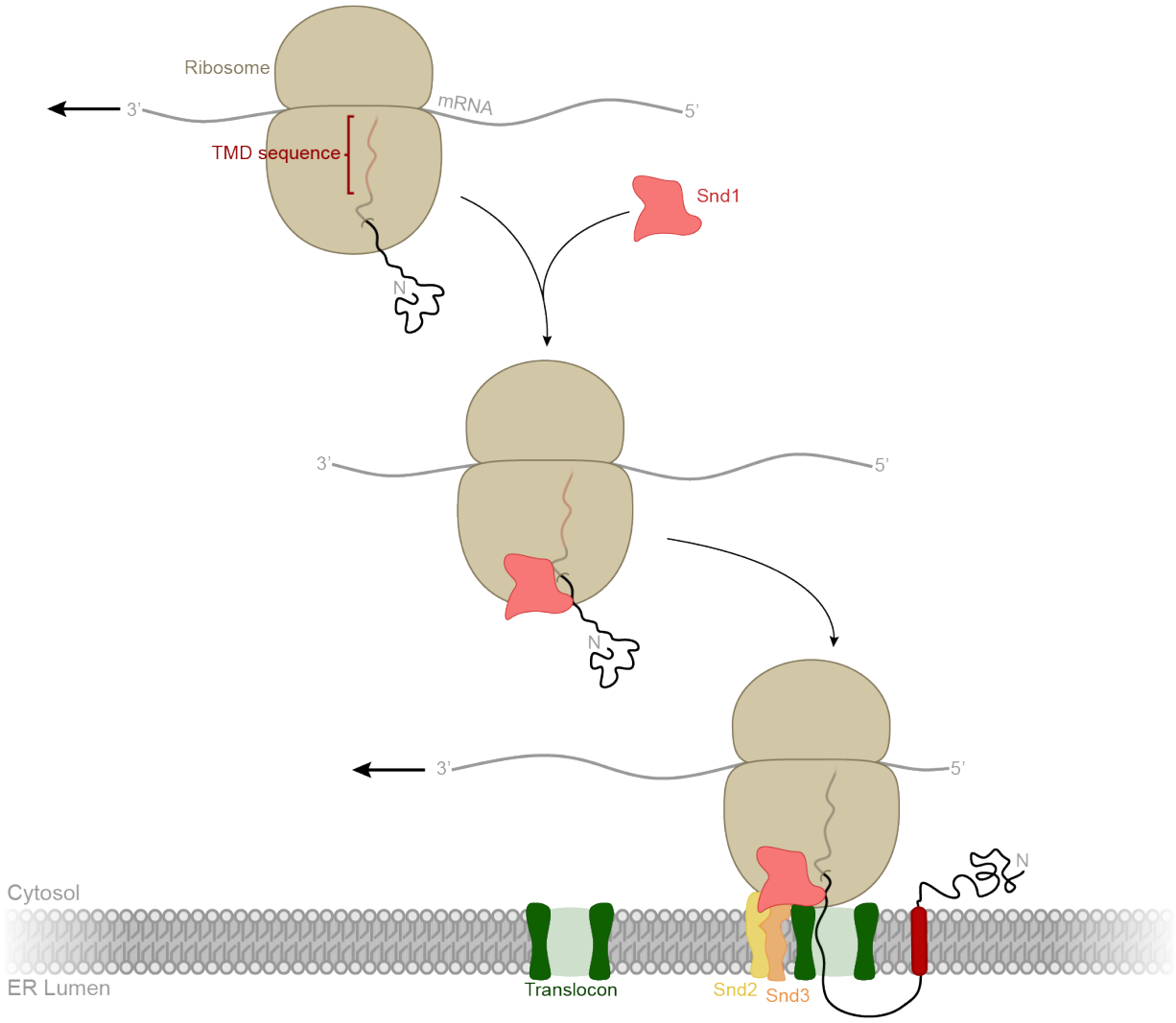


Figure 1.2 : The SND-dependent co-translational targeting pathway. During translation, Snd1 binds either the nascent chain, the ribosome, or both. Snd1 then brings the ribosome to the ER membrane by associating with ER-embedded Snd2 and Snd3. These two proteins also associate with the translocon, and translation proceeds as the translocon helps the nascent protein insert into the ER.

Proteins with a single TMD at their C-terminus, also called tail-anchored (TA) proteins, present a unique challenge to intracellular targeting pathways, since the ultimate destination of TA proteins is encoded in the physiochemical properties of the TMD itself. As a result, the entire protein is synthesized and released from the ribosome before it is recognized by various targeting pathways and shuttled to the appropriate membrane.

Much of the targeting information encoded in TA proteins comes from the length, hydrophobicity, alpha-helical content, and flanking residues of the TMDs, as these factors govern which chaperones are likely to bind the TMD segment upon release from the ribosome. TA proteins with highly-hydrophobic and alpha-helical TMDs are typically destined for the ER and are recognized by the Guided Entry of Tail-anchored proteins (GET) pathway (Schuldiner et al., 2008; Fig. 1.3), while TA proteins with moderately hydrophobic TMDs appear to instead bind the cytosolic protein calmodulin for transit to

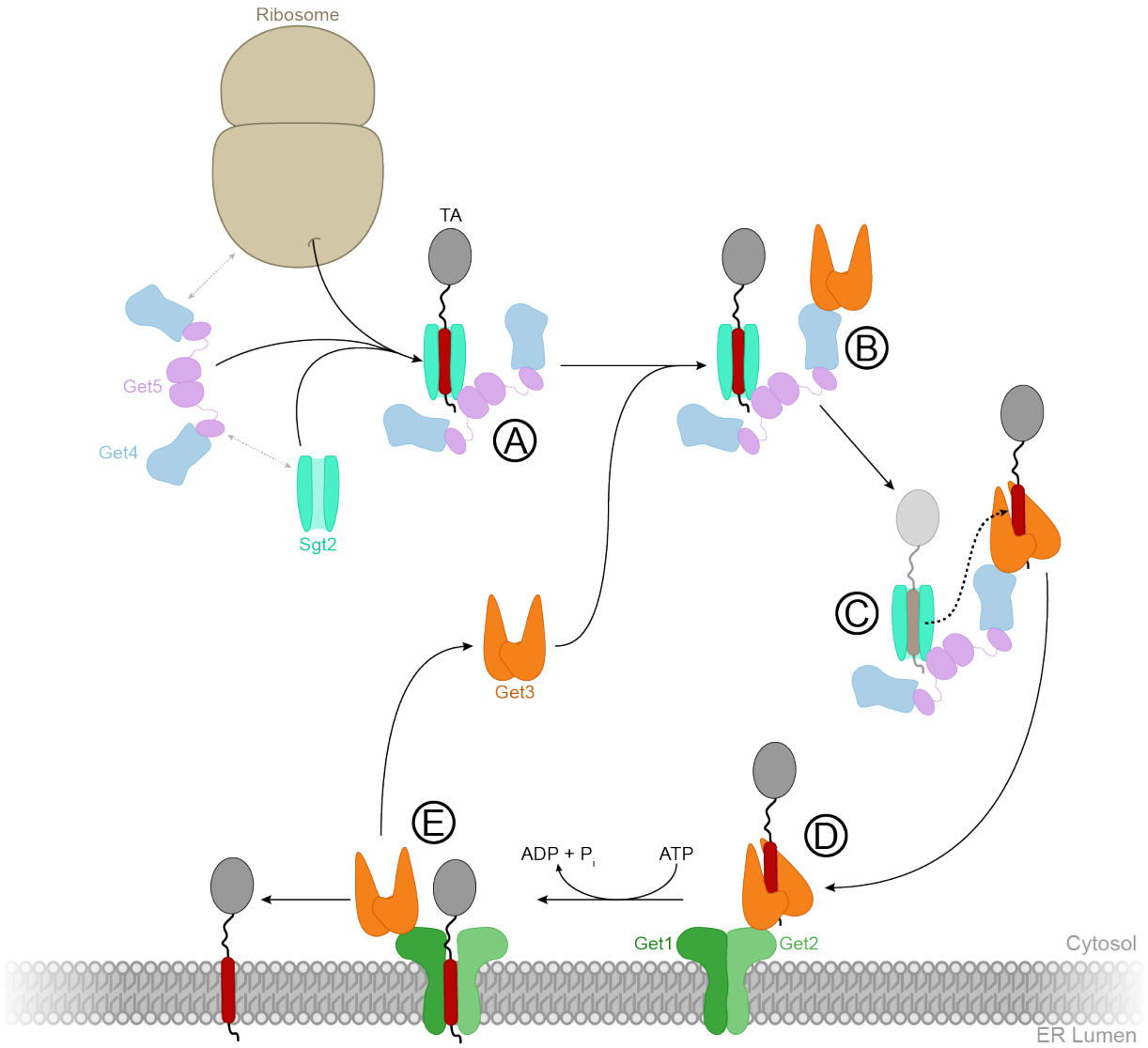


Figure 1.3 : An overview of the GET targeting pathway. (A) The chaperone Sgt2 binds the TMD of TA proteins after their release from the ribosome. The complex is then bound by the Get4/Get5 heterodimer, with Get5 directly binding Sgt2. (B) The heterodimer then recruits Get3 through an interaction with Get4. (C) The Get4/Get5 complex mediates a handoff of the TA protein from Sgt2 to Get3. (D) Get3 then brings the TA protein to the ER membrane. Through a progression of interactions with Get2, Get1, and varying affinities throughout the ATP-hydrolysis process, the TA protein is inserted into the ER membrane.

the ER (Guna et al., 2018). However, not all TA proteins must go to the ER, and cells have also had to devise mechanisms of post-translationally targeting such populations to other organelles: specifically, to mitochondria and peroxisomes.

While the proteinaceous machinery responsible for TA protein insertion into both the outer mitochondrial membrane (OMM) and the peroxisome are still unknown, many studies have attempted to profile what *kind* of TA proteins, in particular what kind of TMDs, end up in these organelles. TA proteins destined for the OMM membranes tend to have relatively short TMD helices and exhibit lower hydrophobicity than TMDs of TA protein in the ER (Kalbfleisch et al., 2007), while TA proteins in the peroxisomal membrane are similar but with moderate hydrophobicity. Additional information on a protein's intended destination may be conferred by a TMD's C-terminal element (CTE),

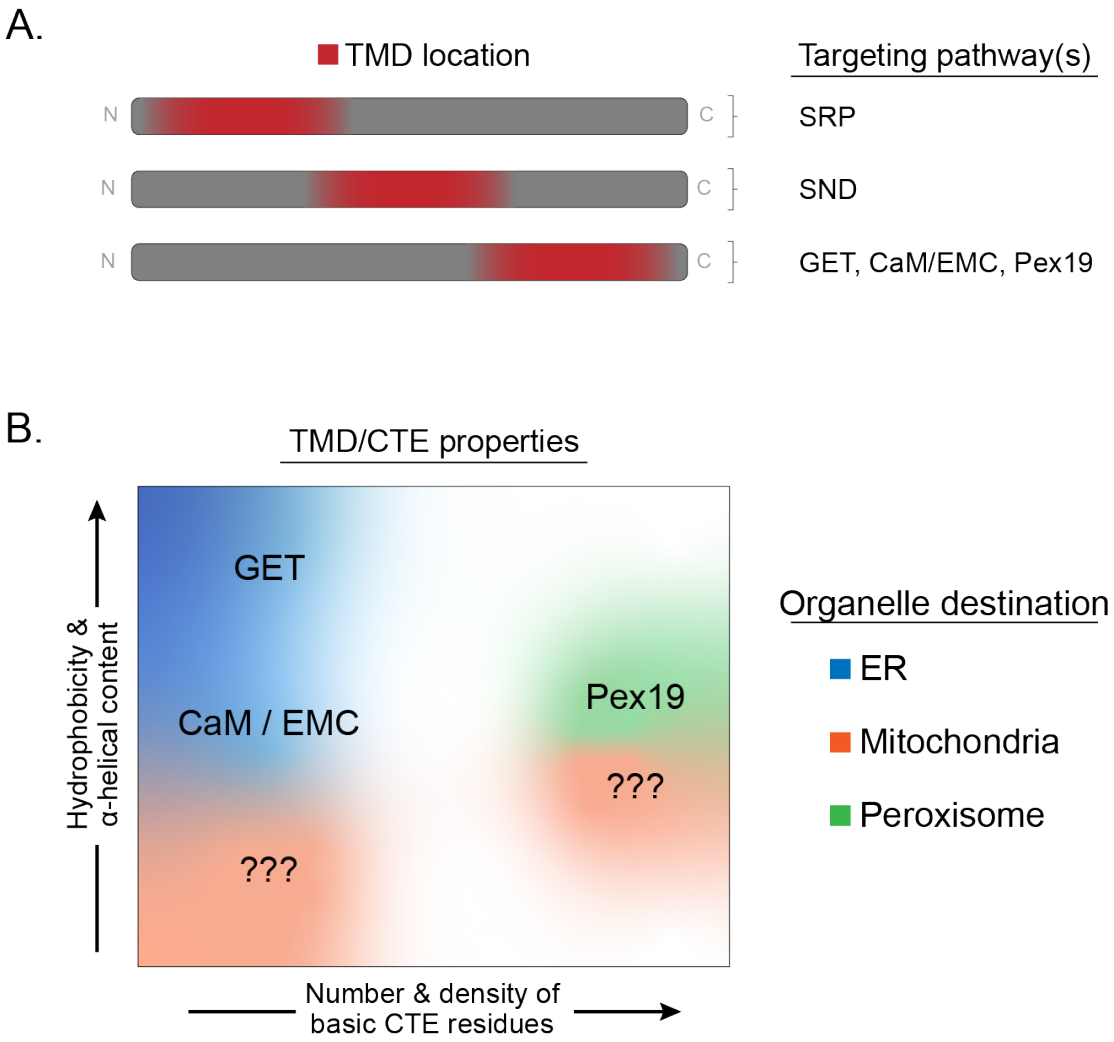


Figure 1.4 : An overview of targeting pathways for TMD-containing proteins. (A) The location of a proteins TMD helps determine which targeting pathway it may be directed by, with some, but not all, targeting pathways listed by TMD location. (B) For proteins with TMDs at or near their C-terminus (TA proteins), the hydrophobicity, alpha-helical content, and the number and density of basic residues in the C-terminal element (CTE) appear to be determining factors for varying targeting pathways and ultimate destinations. TMDs of high hydrophobicity and alpha-helical content are readily recognized by Sgt2 and are thus targeted by the GET pathway. TMDs of moderate hydrophobicity/alpha-helical nature are more likely to be recognized by CaM and inserted into the ER via the EMC. Proteins with TMDs of low hydrophobicity/alpha-helicity with a low density of basic residues at their CTE tend to be found in the mitochondria, though the molecular machinery involved is unclear. An overlapping category of proteins can be found at both peroxisomes and mitochondria: proteins whose TMDs are moderately hydrophobic but possess a high density of basic residues in their CTE. Some of these proteins are thought to be shuttled to the peroxisome by Pex19, but the exact molecular mechanism of insertion at either membrane is unclear.

the sequence just C-terminal of the TMD. Specifically, CTEs containing several basic residues are present in most peroxisomal TA proteins, but only in a portion of OMM TA proteins (Kalbfleisch et al., 2007). The hydrophobicity, alpha-helicity, and CTE charge of TA proteins can provide a rough prediction of where a TA protein will ultimately reside (depicted in Fig. 1.4); however, clear distinction between TA proteins destined for mitochondria and peroxisomes remains unclear (Chio et al., 2017).

The breadth of targeting pathways dedicated to localizing TA proteins with vaguely distinguishable or overlapping physiochemical properties means it is unlikely for

a TA protein to be left unattended in the cytosol following release by the ribosome. However, combined with the repeated observations that TA proteins can, if released by their chaperone, auto-insert into a membrane, this also produces the potential for proteins to mislocalize to the wrong cellular compartment. While cells have developed a variety of mechanisms to recognize unfolded proteins and target them for degradation, distinguishing an otherwise-healthy protein in the wrong membrane presents a different challenge.

1.3 The AAA+ ATPase Msp1 clears mislocalized TA-proteins from the OMM

Cell biologists have long been interested in identifying the proteins and pathways responsible for protein localization quality control, particularly with regards to mitochondrial proteins. In 1993, a group fused the OMM-localizing signal from the MAS70 protein to cytochrome c_1, thus rerouting the chimeric cytochrome c_1 from the mitochondrial intermembrane space to the OMM (Nakai et al., 1993). Since this mislocalization resulted in decreased cellular fermentation and thus poor growth on glycerol, the group then performed a genetic screen and selected for restored growth on glycerol in an attempt to identify genes responsible for mitochondrial protein sorting. They found that overexpression of one gene in particular restored viability on glycerol, and thus named it MSP1 (Mitochondrial Sorting Protein). The Msp1 protein was further characterized to be roughly 40 kDa, contain an N-terminal TMD, and localize to the OMM. Msp1's TMD anchors it in the OMM, while its soluble domain extends into the cytosol. Further analysis of sequence alignments suggested that MSP1 is a part of a then-novel family of putative ATPases. Interestingly, Msp1 function was suggested to be nonessential for cell growth and mitochondrial function in WT cells, which begs the question what Msp1 is contributing to cellular functions in WT yeast.

Further characterization of Msp1 came over 20 years later, when it was hypothesized that Msp1 could be responsible for the quality control of mislocalized TA proteins at mitochondria (Okreglak & Walter, 2014). The mislocalization of one protein in particular, Pex15, became the model system for Msp1 biochemical dissection. Pex15 has been suggested to be post-translationally targeted to peroxisomes, either directly via Pex19 or indirectly via the GET pathway (Elgersma et al., 1997, Okreglak & Walter, 2014). Interestingly, truncation of the last 30 amino acids of Pex15 ($Pex15^{354-383}$) leaves the protein with only a short stretch of basic residues after the TMD, which further hinders localization fidelity and increases mislocalization to mitochondria. How $Pex15^{354-383}$ mislocalizes specifically to mitochondria is unclear; however it is reasonable to speculate that the now-terminal basic residues, combined with Pex15's weakly-hydrophobic TMD, resemble the TMD/CTE physiochemical profile of TA proteins typically found in the OMM, and thus are adopted by whatever targeting pathways that typically guide TA proteins to the mitochondria. Alternatively, it may be that the lipid composition of the OMM is more amenable to spontaneous TMD incorporation than that of other organelle membranes. Pex15 appears to be prone to mislocalization, as impairment of the GET pathway in yeast can also lead to mislocalization of full-length Pex15 to the OMM. It is especially surprising that Pex15 is trafficked through the GET

pathway, given how weakly hydrophobic its TMD is, but deletion of various GET pathway components clearly causes mislocalization to mitochondria (Okreglak & Walter, 2014). Deletion of Msp1 (*msp1*Δ) independently leads to accumulation of Pex15 in the OMM, and an *msp1*Δ *get3*Δ double deletion further exacerbates this phenotype. The observations that Msp1 and Pex15 co-immunoprecipitate (co-IP), and that Pex15 is cleared from the OMM more slowly in *msp1*Δ cells, provide the framework for the model that Msp1's function is to clear mislocalized TA proteins from the OMM (Okreglak & Walter, 2014). Support of this model comes from observations that ATAD1/Thorase, the mammalian homolog of Msp1, performs a similar role in clearing mislocalized TA proteins from mitochondria in cultured human cells (Chen et al., 2014).

Msp1's classification as a AAA+ ATPase is independently very informative on the protein's structure and function. AAA+ ATPases commonly function as oligomers, like pentamers, hexamers, and heptamers, and possess a highly-conserved AAA-domain fold of 200-250 residues, forming an N-terminal large and a C-terminal small subdomain. ATP binding occurs between these subdomains and is mediated by Walker-A motif, while the Walker-B motif facilitates hydrolysis ((Walker et al., 1982, Fig. 1.5a). The Walker A motif is part of the highly conserved P loop and has the consensus sequence GXXXXGK[T/S], where X is any amino acid and the lysine directly interacts with phosphate groups of the bound ATP. Additionally, the threonine/serine of this motif helps coordinate a Mg^{2+} ion, which interacts with the ß-phosphate of the bound ATP as well as water molecules that also interact with residues of the Walker B motif (Fig. 1.5b, yellow). Mutations in the Walker A motif, typically replacement of the lysine, preclude ATP binding. The Walker B motif has the consensus sequence hhhhD[D/E], where h is any hydrophobic residue. The C-terminal aspartate/glutamate of this motif is the catalytic base that activates a water molecule to act as a nucleophile on the γ-phosphate of ATP during hydrolysis (Fig. 1.5b, salmon). Mutations in the Walker B motif, typically exchanging the C-terminal glutamate/aspartate for glutamine/asparagine, allow for ATP binding (because the Walker A motif is still intact), but abolish ATP hydrolysis. The ATP binding site of each AAA+ subunit is localized at the interface to the clockwise-next neighbor in the oligomeric assembly, e.g. the hexameric ring. While Walker A and Walker B motifs of each binding pocket are in the same subunit, thus making them cis-acting residues, the arginine finger motif is provided by the adjacent subunit, with a highly conserved arginine residue stabilizing ATP through interaction with the γ-phosphate and thus playing a critical role for hydrolysis (Fig. 1.5b, brick-red; Miller & Enemark, 2016). Protein translocases of the AAA+ family form hexameric rings with a central pore or channel, into which substrate-gripping pore loops protrude from every subunit. These pore loops interact with substrate polypeptides through bulky, hydrophobic residues such as phenylalanine, tyrosine, and tryptophan, and allow ATP-hydrolysis driven conformational changes of the motor hexamer to be transduced for mechanical substrate translocation (Martin et al., 2008).

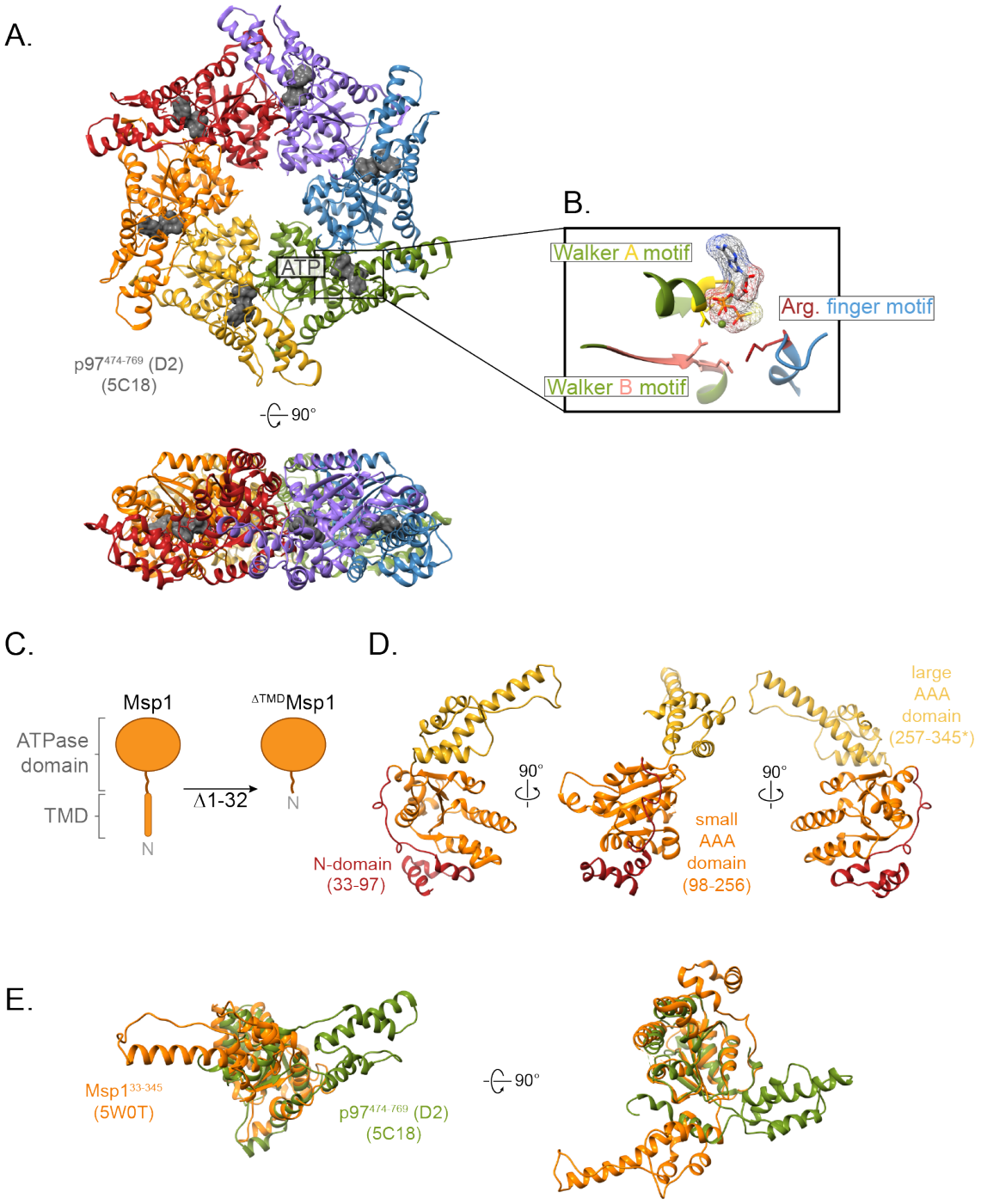


Figure 1.5 : An overview of AAA ATPases and Msp1 structure. (A) The D2 domain of p97 (5C18) exemplifies the conserved AAA ATPase hexameric architecture (individual protomers are differently colored). (B) Adjacent subunits form ATP-binding pockets, with one protomer contributing the Walker A and Walker B motifs, and the other contributing the arginine finger motif. (B) Cartoon schematic of full-length Msp1 and $^{\Delta TMD}$Msp1. (D) The Msp1 crystal structure (5W0T), colored by domain. (E) Msp1 crystallized in a conformation different to other conserved AAA domains.

Several recent studies have furthered our understanding of how Msp1 performs its function as a AAA+ motor in protein quality control. Wohlever et al. (2017) reconstituted heterologously-expressed full-length Msp1 into liposomes and observed *in vitro* extraction of a model TA substrate (Sumo fused to Sec22's TMD). Importantly, these data suggest that Msp1 can work independently of cofactors or binding partners, though they do not preclude the possibility that the presence of such interactions may either inhibit or enhance Msp1's activity. Perhaps unsurprisingly, this extraction activity is dependent on ATP hydrolysis, as Msp1 variants harboring the Walker B mutation (E193Q) failed to extract the TA substrate. Interestingly, extraction also depends on Msp1's pore loops. Msp1 has three pore-loops: loop 1 (on the N-terminal side of central pore), loop 2 (on the inside of the central pore), and loop 3 (on the C-terminal side of the central pore). Mutation of the loop 1 residues (W166A/Y167A) and the loop 3 residues (H206A/E207A) both abolish protein extraction from proteoliposomes, suggesting that these residues are likely involved in gripping the substrates during extraction. Low-resolution negative stain EM suggests that $Msp1^{E193Q}$ forms a hexamer, and the crystal structure of monomeric Msp1 without its TMD ($^{\Delta TMD}$Msp1) has been solved. This structure (Fig. 1.5c) shows three main modules: the N-domain, the small AAA+ subdomain, and the large AAA+ subdomain (Fig. 1.5d). Despite high sequence similarity with other AAA+ ATPases, such as the D2 domain of p97 (Fig. 1.5a), structural alignment of Msp1 with this domain reveals a striking disparity in conformation. Specifically, the long helix of Msp1's large AAA subdomain is rotated 180°, which is seemingly-incompatible with hexamerization (Fig. 1.5e, Wohlever et al., 2017). It is unclear whether this conformation represents a relevant biological state of monomeric Msp1 or the protein simply crystallized in this strange state. Interestingly, substitution of Msp1's TMD with TMDs from other proteins changes neither its *in vitro* nor its *in vivo* activity, suggesting the TMD is not specifically central to the hexamerization of Msp1 or its substrate recognition. A series of experiments with size-exclusion chromatography (SEC) suggest that $^{\Delta TMD}$Msp1 is unable to hexamerize both in the presence and absence of ATP; however, hexamerization does occur in the presence of a slowly-hydrolyzable ATP analog, ATPγS. Interestingly, however, $^{\Delta TMD}$Msp1 does not hexamerize in the presence of the non-hydrolyzable analog AMPPNP. Further studies found that the ATP-hydrolysis-incompetent Walker B mutant of Msp1 remains a monomer in the absence of nucleotide (apo), but forms stable hexamers in the presence of a variety of nucleotides (ATP, ATPγS, AMPPNP). In support of the conclusion that these constructs were indeed forming hexamers, low-resolution negative stain electron microscopy (EM) revealed hexameric particles of $^{\Delta TMD}$Msp1E193Q in the presence of ATP and $MgCl_2$, which were roughly 13 nm in diameter. It is noteworthy that this negative stain EM revealed only top-down or bottom-up views of the hexamers, which appeared to have a preferred orientation of lying flat on the grids. As a result of not seeing any side-views of the hexamers, only low-resolution top-down/bottom-up 2D class averages were obtained. Overall, this study by Wohlever et al. showed that Msp1 indeed forms a hexamer like other protein translocases of the AAA+ ATPase family, and that it can extract a model substrate from proteoliposomes without the help of binding partners or post-translational modifications to either itself or its substrate. Furthermore, this work demonstrated that extraction by Msp1 is dependent on both, its pore loops and its ability to hydrolyze ATP.

1.4 Pex15's role in peroxisome biogenesis

Peroxisomes are present in all nucleated cell types throughout Eukarya and are the site of various vital metabolic processes, most notably the β-oxidation of fatty acids (Islinger et al., 2018). In order to host these metabolic pathways, peroxisomes must import a wide array of proteins into their luminal matrix during peroxisome biogenesis. Failure to efficiently import matrix proteins and establish functional peroxisomes leads to a variety of genetic disorders collectively referred to as peroxisome biogenesis disorders. The most severe peroxisome biogenesis disorder, Zellweger Syndrome, is typically fatal before age 1 (Goldfischer et al., 1973; Klouwer et al., 2015). Unlike mitochondrial or ER import, the peroxisomal matrix protein import pathway involves transport of fully-folded proteins through a highly dynamic pore. This pore, commonly referred to as the importomer, is poorly understood, and the mechanism by which it facilitates the transport of large moieties into a densely-packed peroxisomal matrix remains unknown.

The peroxisomal protein import cycle of proteins tagged with a peroxisome targeting signal 1 (PTS1) can be broken down into four stages: first, the shuttle receptor, Pex5, binds cargo in the cytosol, recognizing PTS1 on the C-terminus of the cargo (Fig. 1.6a; Fodor et al., 2014; Freitas et al., 2011). Secondly, Pex5 associates with the transmembrane proteins of the docking complex, consisting of Pex13, Pex14 and Pex17, via di-aromatic pentapeptide motifs (WxxxF/Y) (Barnett et al., 2000; Bottger et al., 2000; Oliveira et al., 2002; Otera et al., 2002; Saidowsky et al., 2001; Schell-Steven et al., 2005) as well as the RING-domain containing E3 ubiquitin ligases Pex2, Pex10, and Pex12 (El Magraoui et al., 2012; Platte et al., 2009; Schwartzkopff et al., 2015). Many of these proteins appear to associate to form the transmembrane pore known as the importomer (Fig. 1.6b). Exactly when cargo is released, and the composition of the pore through which cargos pass, is unknown (Alencastre et al., 2009). Along with two E2 ubiquitin-conjugating enzymes (Pex4 and Ubc4; Tamura et al., 2014; Williams et al., 2007; Williams et al., 2008), the E3 ligases ubiquitinate Pex5, thus designating it for extraction. Mono-ubiquitination of an N-terminal cysteine destines Pex5 to be recycled into the cytosol to begin the matrix-protein import cycle anew, whereas polyubiquitination of N-terminal lysines targets Pex5 for proteasomal degradation (Platte et al., 2009). Extraction is catalyzed by the AAA+ ATPases Pex1 and Pex6, which form a complex that is recruited to the peroxisome by Pex15, a C-terminally TA protein with an N-terminal cytosolic domain (Fig. 1.6c; Birschmann et al., 2003; Elgersma et al., 1997). Previous work in the Martin Lab used negative stain electron microscopy to demonstrate that Pex1/Pex6 forms a heterohexamer of alternating Pex1 and Pex6 subunits (Gardner et al., 2015). It has long been suggested that Pex1/Pex6's ATP hydrolysis provides the physical force for Pex5 extraction from the membrane, that Pex1/Pex6 physically engages Pex5 and translocates it through its central pore, thereby severing its interactions with the docking complex components. Despite this being the predominant model of Pex1/Pex6 function, it has never been demonstrated that Pex1/Pex6 processes Pex5 in any way. In recent years, the structure

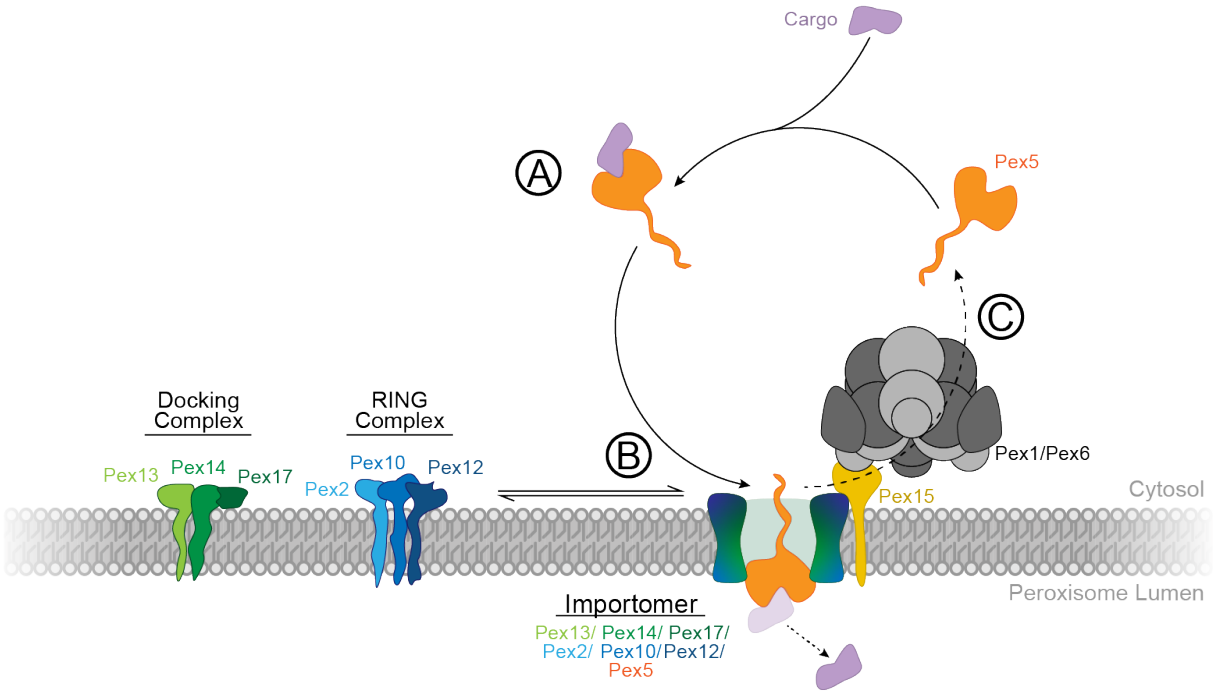


Figure 1.6 : An overview of the import cycle of peroxisomal matrix proteins. (A) Pex5 binds peroxisome matrix-bound cargo proteins in the cytosol, recognizing and binding a conserved C-terminal PTS1 motif. (B) Pex5 associates with the docking complex at the peroxisomal membrane and, along with a multitude of other peroxisomal membrane proteins, helps forms the importomer pore. Through an unclear mechanism, the fully-folded cargo traverses the membrane and is released into the lumen of the peroxisome. Pex5 is then ubiquitinated by members of the RING complex, which targets it for extraction. (C) Pex15 recruits the AAA ATPase heterohexameric complex Pex1/Pex6 to the peroxisomal membrane, and Pex1Pex6 catalyzes the release of Pex5 from the peroxisomal membrane through an unknown mechanism.

of Pex1/Pex6 has come into better focus with several studies publishing progressively-higher resolution structures obtained via cryo electron microscopy (Ciniawsky et al,. 2015; Tan et al., 2016). Despite these advances, a comprehensive model for the biochemical mechanism of Pex1/Pex6, as well as identification of an *in vivo* or model substrate, was lacking. While it had been recently theorized that the Pex1/Pex6 complex could thread substrates through its central pore (Ciniawsky et al., 2015), there was no direct evidence of Pex1/Pex6 unfoldase capabilities, much less demonstration of this phenomenon with any putative or model substrate.

It is possible that a better understanding of Pex1/Pex6's recruitment to the peroxisome by Pex15 will help elucidate the details of its essential activity, as recent evidence suggests that Pex15's recruitment of Pex1/Pex6 to the peroxisome is only part of the story. Tamura et al. (2014) reported that the mammalian homologs of Pex15 and Pex14 directly interact both *in vivo* and *in vitro,* and postulated that the interaction between them is a key step in protein import, but this idea is yet to be corroborated for the yeast proteins. Tamura et al. also suggested that Pex14 binds to a small cytosol-exposed region in Pex15. However, due to low sequence conservation between the yeast and mammalian homologs, it is difficult to predict this equivalent region in the yeast Pex15. Understanding how Pex15 interacts with Pex14 and how abolishing this interaction affects the import cycle is critical to determining Pex15's role in peroxisomal protein import. Furthermore, it was demonstrated that the soluble cytosolic domain of Pex15 inhibits the rate of ATP hydrolysis by the heterohexamer (Gardner et al., 2015);

however, the mode of interaction between Pex15 and Pex1/Pex6, and the functional significance of Pex1/Pex6's ATP hydrolysis inhibition were unclear. More broadly, understanding more about the structure and domain architecture of Pex15 is paramount to understanding its ability to interact with a diverse array of binding partners, inhibit Pex1/Pex6 ATPase activity, and contribute to the import of peroxisomal matrix proteins.

Interestingly, Pex15 appears to be a substrate of Msp1 when it mislocalizes to the OMM. Since Msp1 is known to be present in the peroxisomal membrane as well, this means that Pex15 possesses the ability to interact with two different AAA+ ATPases in different capacities, providing us a unique opportunity to understand not only what Pex15 contributes to peroxisomes biogenesis, but also how two different motors recognize Pex15. A recent paper concludes that Msp1 does not extract Pex15 from the peroxisome due to the presence of Pex3, and it was proposed that direct binding to Pex3 shields Pex15 from recognition by Msp1 (Weir et al., 2017). A more detailed understanding of newly-revealed interactions will shed light on the mechanism by which Msp1 extracts Pex15 from the mitochondria in the absence of Pex3.

CHAPTER 2

Probing Pex15's domain architecture and structure

Our knowledge of Pex15's domain layout and architecture is limited to secondary structure predictions and the fact that it is a C-terminally tail-anchored transmembrane domain-containing protein. Structural information on Pex15, both about domain layout regarding structured/unstructured regions and about the structure of folded regions, will be paramount to understanding Pex15's role in peroxisome biogenesis. At the beginning of my PhD, I set out to learn more about Pex15 on a structural level, with the goal of solving the crystal structure of the soluble domain of Pex15.

2.1 Pex15 construct, expression, and purification optimization

Prior to this investigation, nothing was known about Pex15's structure except that it has a short TMD at its C-terminus, approximately encompassing amino acids 338-349. The Pex15$^{\Delta TMD}$ construct previously shown to bind Pex1/Pex6 and inhibit its ATPase activity included most of the soluble domain in Pex15$^{1-327}$ (Fig. 2.1a, Gardner et al., 2015). However, the low yields and poor solubility resulting from heterologous

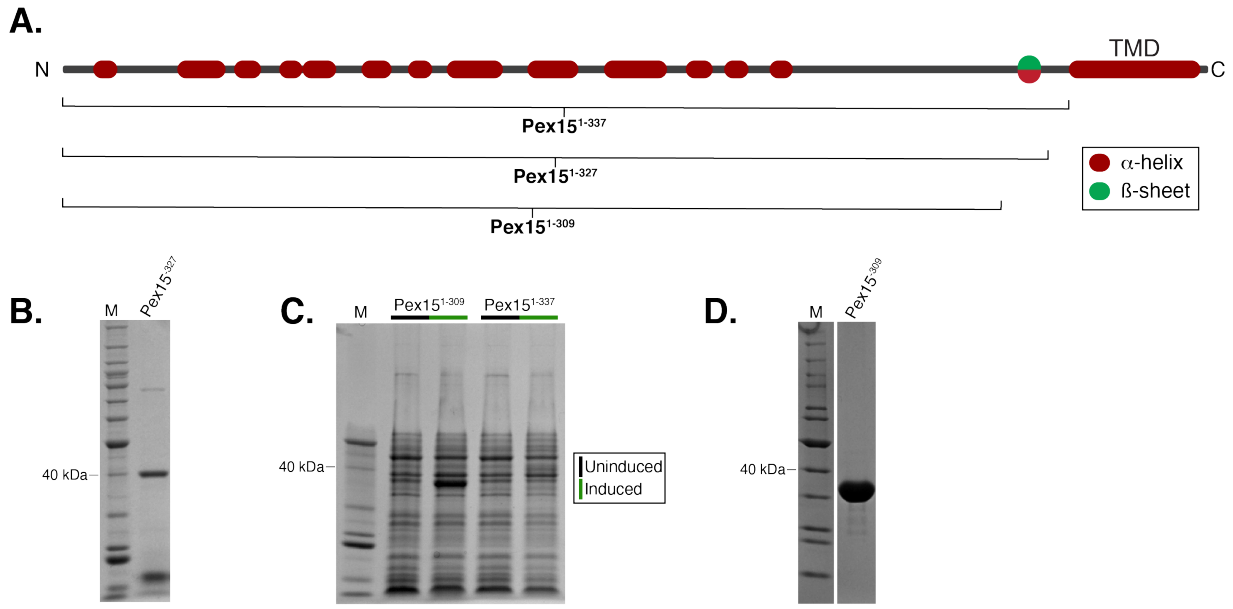


Figure 2.1 : Pex15 is predicted to be comprised of alpha-helices and unstructured regions. (A) A cartoon schematic of Pex15's predicted secondary structure, mapping relevant Pex15 truncation constructs of Pex15$^{1-309}$, Pex15$^{1-327}$, and Pex15$^{1-337}$. (B) Optimal purification of Pex15$^{1-327}$ yields pure protein, but in quantities insufficient for crystallization screening. (C) An induction test of differently-truncated Pex15 reveals robust expression of Pex15$^{1-309}$, but not Pex15$^{1-337}$. (D) Purification of Pex15$^{1-309}$ yields high quantities of pure Pex15$^{1-309}$.

expression of Pex15$^{1-327}$ in *E. coli* means this construct is poorly suited for crystallization (Fig. 2.1b). I conducted an expression screen of Pex15—varying such parameters as OD of induction, expression durations and temperatures, and the media—but was ultimately unable to determine conditions that yielded significantly more soluble protein in the cell lysate, and I therefore turned to construct optimization. Secondary structure prediction (JPredIV) suggested that Pex15 possesses a long unstructured region between its TMD and most of the other predicted helices in the middle of the soluble domain (Fig. 2.1a). The only exception to this observation is a 12-residue stretch of

ambiguous secondary structure identity (colored red and green in Fig. 2.1a). Since long unstructured regions tend to preclude protein crystallization, I evaluated different truncations of Pex15's C-terminus.

The expression of Pex15$^{1-337}$ mirrored that of the original Pex15$^{1-327}$ construct, showing no significant increase in band intensity on a gel between pre- and post-expression samples (Fig. 2.1c). In comparison to Pex15$^{1-327}$ and Pex15$^{1-337}$, the more aggressively truncated Pex15$^{1-309}$ demonstrated significantly better expression (Fig. 2.1c), and the purified protein was much more soluble (Fig. 2.1d). Notably, Pex15$^{1-309}$ required only a single-step His-tag affinity purification to reach approximately the same purity as longer versions after additional size-exclusion chromatography. Since screening crystallization conditions consumes a considerable quantity of protein, this was a great starting point.

Pex15$^{1-309}$ still showed some contaminants at low concentrations, whose presence may not have major effects on *in vitro* activity assays, but could interfere with crystallography. I therefore increased the stringency of washing steps during the affinity purification through the addition of higher salt concentrations, various detergents, or ATP to release potentially bound cellular chaperones, yet observed no further decrease in contaminants.

Although Pex15$^{1-309}$ expressed very well, it still contains a long unstructured region at its C-terminus. I hypothesized that this region may preclude crystallization and, using secondary structure prediction servers and the primary sequence as a guide, I developed another more truncated construct: Pex15$^{1-253}$. Like Pex15$^{1-309}$, Pex15$^{1-253}$ expressed and purified very well. Through these expression and purification tests, it is abundantly clear that any Pex15 construct including amino acids 310-327 demonstrates poor solubility, though it remains unclear whether this represents an important feature in Pex15's *in vivo* activity.

Initial attempts to crystallize Pex15$^{1-309}$ and Pex15$^{1-253}$ were unsuccessful. Hypothesizing that either the C-terminal unstructured region was poorly predicted or that there may be more unstructured regions in Pex15, I used limited proteolysis with trypsin to reveal well-folded regions of the protein. Trypsin is highly specific in cleaving only after basic residues (arginine and lysine). The technique of limited proteolysis employs a high concentration of protease for a short period of time (e.g. 10 min). During this time, the protease will cleave only the easily accessible sites (i.e. unstructured regions and exposed loops) and not residues participating in secondary structure.

I incubated 10 μM Pex15 with increasing concentrations of trypsin for 10 min at room temperature, then quenched the reaction by denaturing the protein in 2X SDS Sample Buffer (125 mM Tris pH 6.8, 4% SDS, 20% glycerol, 0.01% bromophenol blue, 10% ß-mercaptoethanol) with PMSF to specifically inhibit trypsin activity. I performed this assay with both Pex15$^{1-309}$ and Pex15$^{1-253}$ to investigate the contribution of the length of the C-terminal unstructured regions in core proteolysis. Pex15$^{1-309}$ has a predicted unstructured region of 70 amino acids (including uncleaved affinity purification tags), whereas Pex15$^{1-253}$ presumably lacks any unstructured domain at its C-terminus (assuming cleaved affinity tags). When incubated with trypsin, Pex15$^{1-309}$ readily formed two major degradation products (Fig. 2.2a). The larger of the two products ran on a gel identically to the non-proteolyzed Pex15$^{1-253}$ construct, suggesting that amino acid K253 is at or near the junction of structured and unstructured polypeptide, and confirming

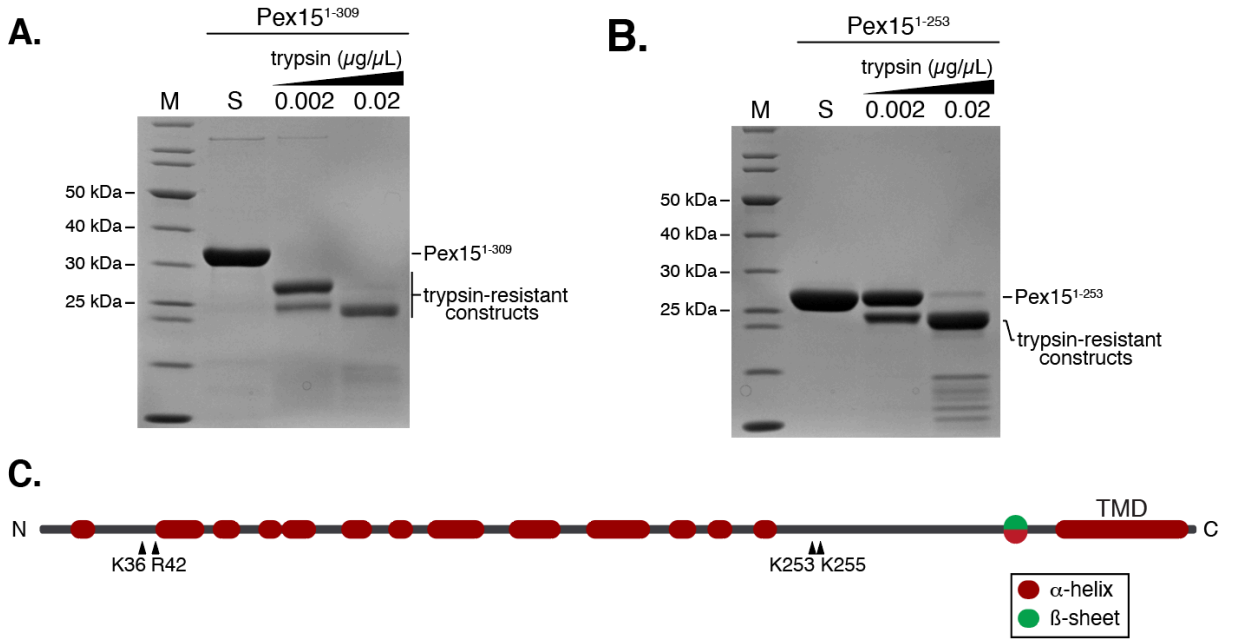


Figure 2.2 : Limited proteolysis suggests new boundaries of structured core domain. (A) Limited proteolysis of Pex15$^{1-309}$ with the proteasomal core domain reveals a degradation-resistant construct slightly smaller than the starting domain. (B) The same reaction as in (A) but with Pex15$^{1-253}$ shows no proteolysis by the proteasome core domain, suggesting the degradation-resistant domain seen in (A) likely included a long unstructured region that could not reach to the proteases in the core domain. (C) Limited proteolysis as in (A) but with trypsin reveals two protease-resistant stable domains. (D) Limited proteolysis as in (C) but with Pex15$^{1-253}$ suggests that the larger of the two stable domains formed in (C) is very similar to the starting Pex15$^{1-253}$ construct in (D). (E) Mass spectrometry of the products of (C) suggest four distinct trypsin cleavage sites after residues K36, R42, K253, and K255.

what was predicted by the secondary structure prediction algorithms. Pex15$^{1-309}$'s smaller degradation product, interestingly, was the same size as Pex15$^{1-253}$'s only major degradation product (Fig. 2.2b). Mass spectrometry revealed four major cut sites in Pex15$^{1-309}$: two in the C-terminal unstructured region at K253 and K255, and two near the N-terminus at K36 and R42 (Fig. 2.2c). From this we concluded that the first 35-45 amino acids of Pex15's N-terminus are largely unstructured, and we consequently produced two more crystallization candidates: Pex15$^{37-253}$, and Pex$^{43-253}$.

In conclusion, these data reveal that the soluble domain of Pex15 (Pex15$^{1-309}$) can be broken down into three subdomains: an N-terminal unstructured region (amino acids 1-42), a folded core domain (amino acids 43-253), and a C-terminal unstructured region (amino acids 254-309).

I next set out to investigate the basic secondary structure within the structured region of Pex15. Though the JPredIV secondary structure prediction server suggested that Pex15 was entirely alpha-helical, I sought to confirm this using circular dichroism (CD) spectroscopy. I analyzed the CD spectra of Pex15$^{1-309}$ and observed very strong signal matching the expected spectrum for a highly alpha-helical protein (Fig. 2.3).

2.2 Solving the crystal structure of Pex15

Early attempts to crystallize Pex15$^{1-309}$ were performed by setting crystal trays on sitting-well Intelli-Plate 96-3 LVR plates, with the three subwells containing different

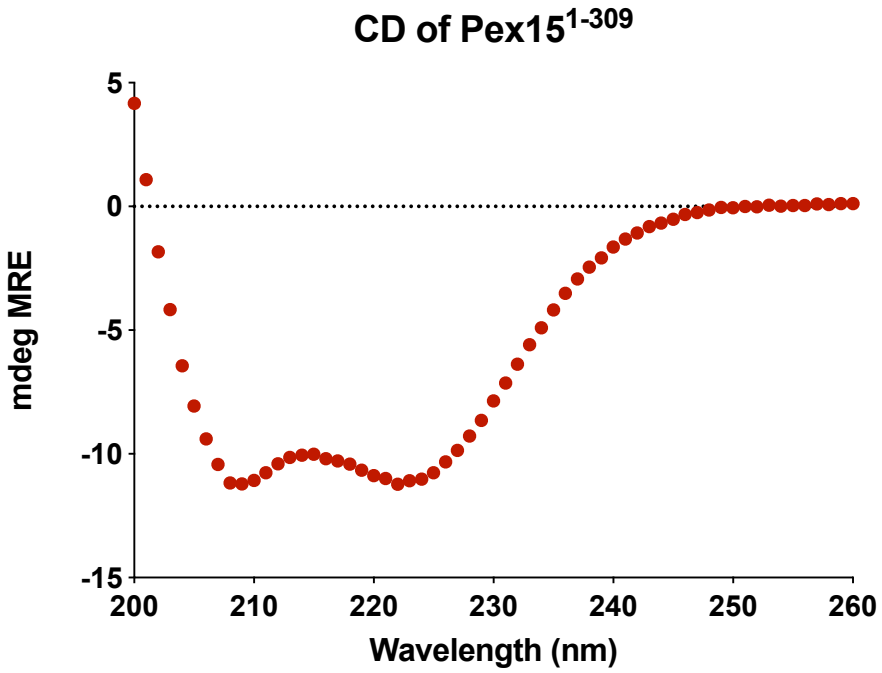


Figure 2.3 : CD spectroscopy confirms that Pex15 contains high alpha-helical content. CD spectroscopy of Pex15$^{1-309}$ suggests the high alpha-helical content. Samples were scanned at a constant concentration of 3 mg/mL.

Pex15$^{1-309}$ concentrations and a buffer-only control. I set trays of 5 mg/mL and 10 mg/mL Pex15$^{1-309}$ in minimal crystallography buffer (20 mM HEPES pH 7.6, 50 mM NaCl, 50 mM KCl, 0.5 mM TCEP) using four different commercial screens, JCSG I-IV, and incubated the trays at 18 °C. I manually screened the plates over 2-3 weeks, but observed nothing resembling orderly precipitation. I considered that the His and FLAG purification tags may interfere with crystallization, and therefore cloned a PreScission Protease (PSP, also called 3C protease) cleavage site (LEVLFQGP) in front of these C-terminal tags to remove them after purification. Since PSP cleaves after the glutamine residue, it leaving a 6 amino acid cleavage scar at the C-terminus of the protein. This cleavage was performed during a dialysis step to remove imidazole from the sample and allow a second pass over Ni-NTA for removal of all uncleaved, His-tag containing protein. Unfortunately, removal of the affinity tags did not improve crystallization of Pex15$^{1-309}$. Nevertheless, I continued to cleave off the tags from all Pex15 constructs for structural and biochemical characterization.

Secondary structure predictions of Pex15$^{1-309}$ suggest that the main folded domain, a series of alpha helices, ends around amino acid 253 (Fig. 2.2c). Hypothesizing that amino acids 254 through 309 are mostly unstructured and are therefore hindering crystallization, I truncated Pex15 even more aggressively to form a version without this C-terminal unstructured region: Pex15$^{1-253}$. I set crystal trays using the following screens: JCSG I-IV, Crystal Screen I-II, Wizard I-II, PEGs Suite I-II, and the PACT Suite. While the Pex15$^{1-253}$ in each of these plates displayed some semblance of orderly precipitation, ultimately none of these hits could be optimized in either small scale (400 nL droplets) or large scale (4 µL droplets) screens.

Since my initial attempts to crystallize Pex15$^{1-309}$ and Pex15$^{1-253}$ were unsuccessful, I turned my attention to more aggressively-truncated Pex15 constructs

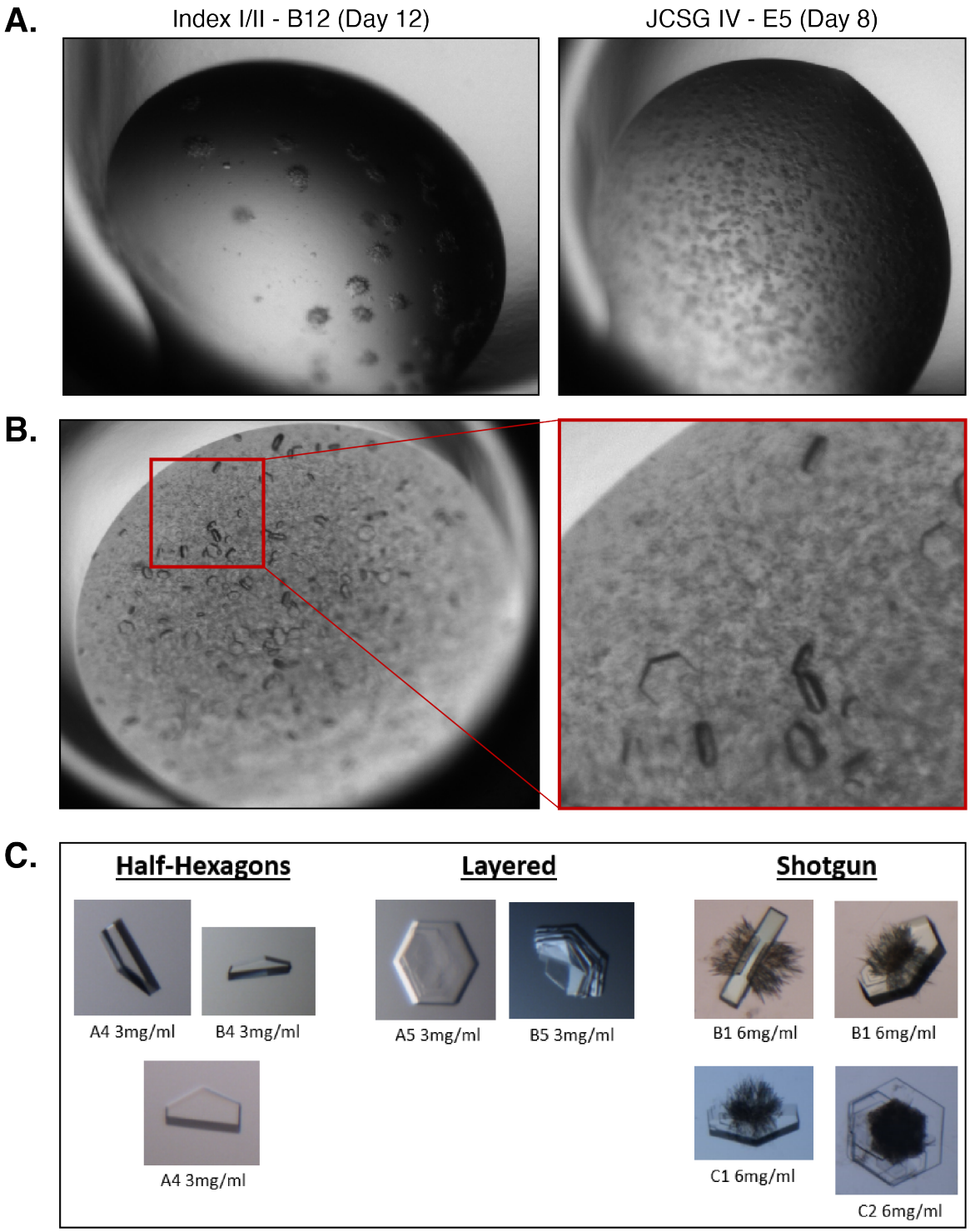


Figure 2.4 : Optimization of Pex15$^{37\text{-}253}$ crystallization. (A) Initial crystal hits of Pex15$^{37\text{-}253}$ were regular and proteinaceous, but lacked sharp edges associated with well-ordered crystals. (B) Optimization of conditions in (A) produced regular hexagonal crystals with sharp edges. (C) Further optimization of conditions in (B) ultimately produced a variety of crystal forms.

which were identified through limited proteolysis, starting with Pex15$^{37\text{-}253}$. The protein was purified by Ni-NTA affinity and gel-filtration chromatography, and peak fractions

were concentrated to ~9 mg/mL. I set trays of $Pex15^{37-253}$ (9 mg/mL) using the following screens: Index I/II, Wizard I-II, JCSG+, JCSG I-III, Crystal Screen I-II, and PACT. I observed two early hits, which yielded regular but poorly defined proteinaceous precipitations, and I developed small scale screens to optimize each condition (Fig. 2.4a). The first screen varied in sodium acetate trihydrate concentration in 0.2 M increments from 2.4 M to 3.4 M (all at pH 7.0), cross-screened with 0%, 1%, 2.5%, and 5% glycerol. The second screen was based off of a hit in 3.6 M sodium formate, 10% w/v glycerol. Through a combination of condition optimization, microseeding, and scaling up in volume, I ultimately obtained crystals from a matrix screening sodium formate (3.7 M to 4.1 M in 0.1 M increments) against glycerol (0%, 2.5%, 5%). These crystals presented in three different morphologies: half-hexagons, layered hexagons, and an unusual hexagon where the center nucleated spiky protrusions resembling a sea urchin's spines (Fig. 2.4b).

Although these crystals diffracted with a resolution up to 2.0 Å (Fig. 2.5a) and indexed to the P3 space group, they possessed unit cell dimensions of 55.8 Å x 55.8 Å x 539 Å. The extreme length of this unit cell presents a problem for data analysis: the diffractions we see represent the reciprocal spacing of points within a unit cell, thus an incredibly long unit cell (such as 539Å) will diffract with reflections very close to each other (Fig. 2.5a). We attempted to spread these reflections by tilting the detector and increasing the distance between the sample and the detector (Fig. 2.5b). Unfortunately, these efforts were insufficient to allow for easy distinction between the reflections, which remained too close together to be processed (Fig. 2.5c).

The long unit cell would not be an issue if we possessed a suitable model to use for molecular replacement to solve the structure, however no previous structures of Pex15 or its homologs (Pex26 in mammals, APEM9 in plants) are published. I attempted to predict Pex15's structure using structure prediction servers RaptorX, Sparks-X, and PHYRE2. While all three servers yielded entirely alpha-helical structures with slightly varying conformations, none were able to successfully serve as a model for molecular replacement. As a final effort we attempted to phase the data using a single artificial poly-alanine helix, but this too was unsuccessful.

Since I got only two hits with $Pex15^{37-253}$, we decided to pursue the other stable construct indicated by limited proteolysis: $Pex15^{43-253}$. Anticipating that after successful crystallization we may encounter the same issue with phasing the dataset, I elected to obtain phases experimentally by expressing and purifying selenomethionine-containing $Pex15^{43-253}$ (SeMet-$Pex15^{43-253}$).

This construct proved to be far more amenable to crystallization, as I was quickly able to identify 7 conditions promoting orderly precipitation across the JCSG I-IV screens (some examples are shown in Fig. 2.6a). The most promising hit contained 0.1 M MES pH 6.0 and 5% w/v PEG 6000, with the $Pex15^{43-253}$ being set at concentrations of 5, 10, and 15 mg/mL (Fig. 2.6a, right). In the scaled up focused screen for these conditions, I varied the PEG 600 concentration from 3-7% in 1% increments. These conditions yielded plentiful crystals, but some were very small. In an attempt to both vary the types of crystals and increase their size, I set another screen with the same conditions but by using traditional growth, microseeding, macroseeding, and streak seeding (using a cat whisker from Caesar the cat). From these screens I obtained large crystals in a variety of morphologies (Fig. 2.6b).

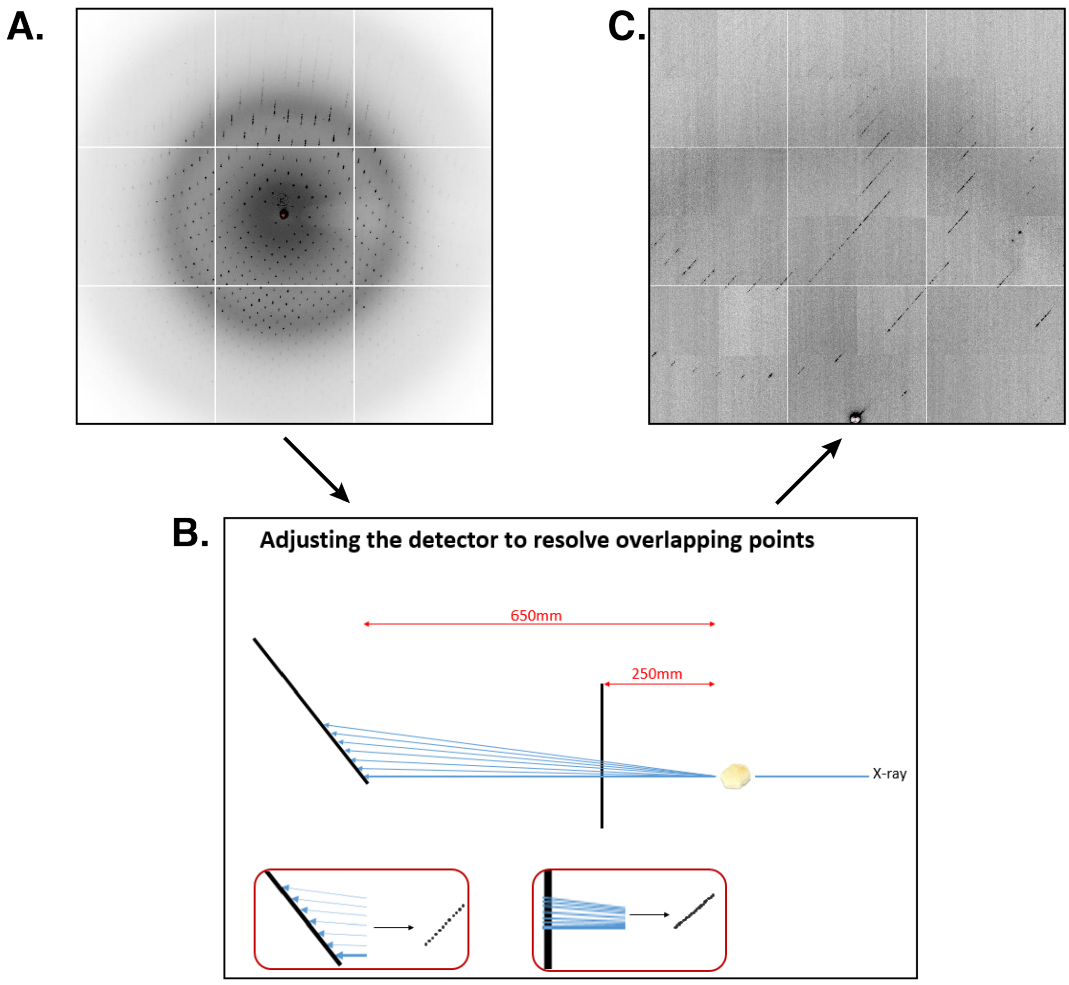


Figure 2.5 : Adjusting detector distance and angle did not improve data collection sufficiently. (A) Diffraction pattern from crystals depicted in Fig. 2.4c. Due to the calculated unit cell dimensions of 55.8 Å x 55.8 Å x 539 Å, reflections along the long axis were impossible to resolve, and they appeared as overlapping smears of reflections. (B) Cartoon schematic depicting our efforts to increase the distance of the detector and angle it such that the reflections might spread out. (C) Although the steps taken in part (B) did help spread the reflections, it was not improved to a degree to which an entire dataset could be collected and resolved.

Freezing the crystals in anticipation of collecting a dataset presented a challenge. High molecular weight PEGs are not particularly good cryoprotectants. Typically, for crystals grown in mother liquors without cryoprotectants, the crystals are looped and briefly bathed in a 1:1 mixture of 50% cryoprotectant and mother liquor, yielding a solution of 25% cryoprotectant. After a brief wash in these mixtures (as quickly as possible, typically 5-10 s) the crystals are re-looped and plunged into liquid nitrogen. The crystals formed from $Pex15^{43-253}$, however, rapidly dissolved during bathing in any cryoprotectant solution, including glycerol, 2-methyl-2,4-pentanediol (MPD), ethylene glycol, D-sorbitol, xylitol, PEG200, and PEG400. I hypothesized that the abruptness of the solution change was too much of a shock for the crystals. My solution to this problem was a gradual stepwise washing of the crystals: I transferred crystals from the mother liquor solution first to a solution of 5% cryoprotectant, allowed them to equilibrate for 5-10s, then looped them into the next solution of 10% cryoprotectant, and so on, until they were ultimately in 25% cryoprotectant and frozen in liquid nitrogen. I froze some crystals in 15%, 20%, and 25% cryoprotectant. This process is much gentler

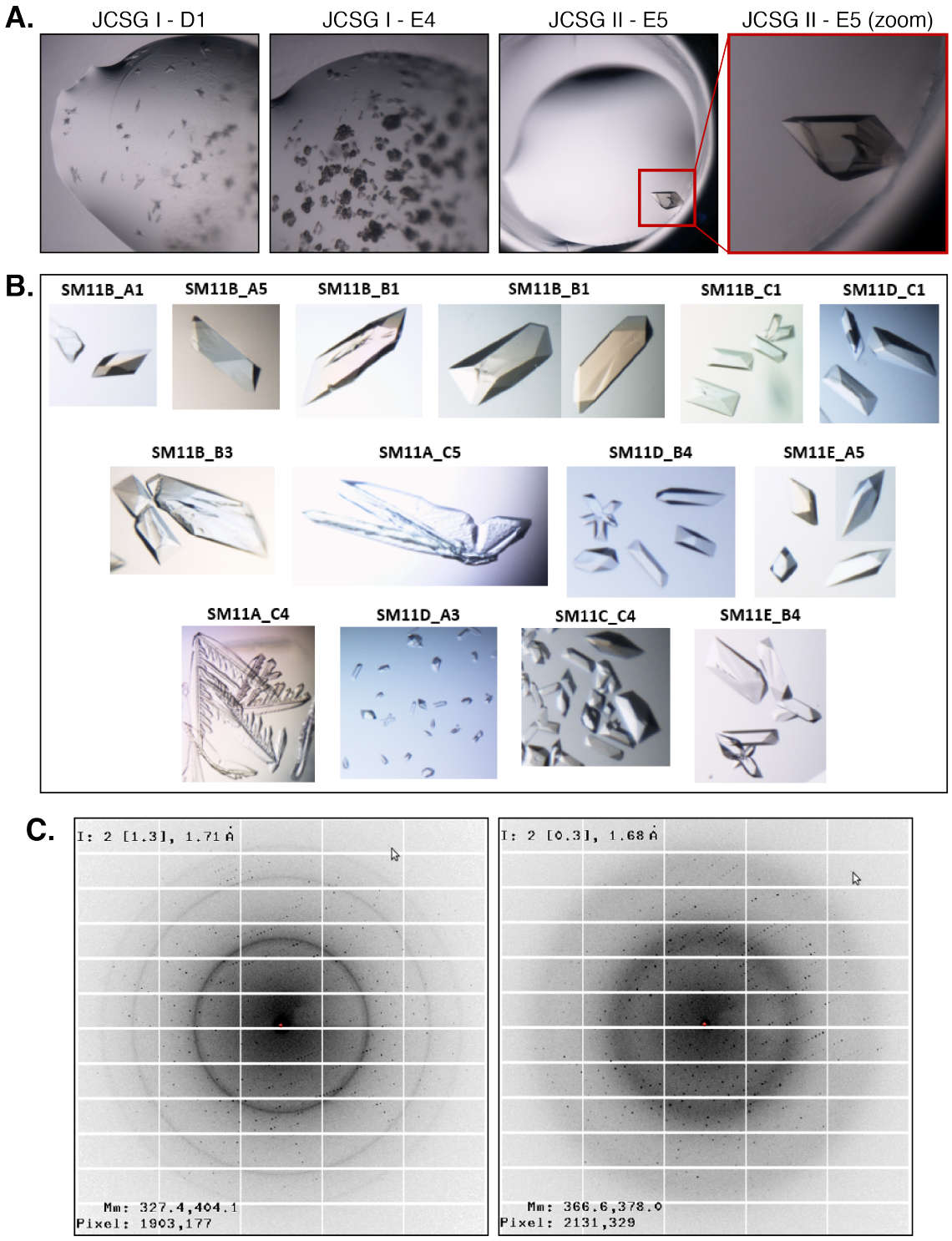


Figure 2.6 : Optimization of Pex15$^{43-253}$ crystallization. (A) Sparse matrix screen crystallization hits of SeMet-Pex15$^{43-253}$. (B) Various crystal forms produced from scaled-up and optimized crystal conditions based off some conditions identified in (A). (C) Example diffraction image of some crystals depicted in (B), showing high-resolution diffraction.

and proved to be far more agreeable with the Pex15$^{43-253}$ crystals, which could be better preserved through this method. Using this process, I looped crystals into a smaller array of cryoprotectants: D-sorbitol, ethylene glycol, PEG400, and MPD.

Many of the crystals diffracted to relatively high resolution (Fig. 2.6c), including crystals in different cryoprotectants across different concentrations of cryoprotectants. Although I collected many datasets, the best came from a prism crystal cryoprotected in 25% MPD. This was a SAD dataset collected at the selenium edge of absorbance at a wavelength of 0.979560Å, and it was submitted using XDS to be processed in the $P2_1$ 2_1 2_1 space group, with dimensions 47.7 Å x 58.8 Å x 85.8 Å (a much more reasonably

Table 2.1 : Data collection and refinement statistics (molecular replacement). The SeMet-Pex15$^{43\text{-}253}$ crystals diffracted to 1.39 Å, but throughout the process of refinement we used a 1.55 Å resolution cut-off, and the structure was better refined under these conditions.

	Pex15$^{43\text{-}253}$
Data collection	
Wavelength (Å)	0.979560
Space group	P 21 21 21
Cell dimensions	
a, *b*, *c* (Å)	47.7, 58.8, 85.8
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	48.52-1.55 (1.605-1.55)
R_{merge}	1.333 (1.486)
$I / \sigma I$	18.08 (1.80)
Completeness (%)	100 (100)
Redundancy	13.0 (12.4)
Refinement	
Resolution (Å)	48.52-1.55 (1.605-1.55)
No. reflections	35741 (3509)
R_{work} / R_{free}	0.1738 (0.2705) / 0.1890 (0.3022)
No. atoms	2035
Protein	1773
Water	262
B-factors	
Protein	20.71
Water	32.25
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.56

*Dataset collected from one crystal. *Values in parentheses are for highest-resolution shell.

[AU: Equations defining various *R*-values are standard and hence are no longer defined in the footnotes.]
[AU: Ramachandran statistics should be in Methods section at the end of Refinement subsection.]
Ramachandran favored (98%), allowed (1.8%), and outliers (0%).
[AU: Wavelength of data collection, temperature and beamline should all be in Methods section.]
Wavelength (0.979560Å), Temperature (100K), ALS BL8.3.1

sized unit cell). I submitted this dataset to the AutoSol program on Phenix, which used the experimentally obtained phases to solve the structure of Pex15$^{43\text{-}253}$ to a resolution of 1.39Å, with an R_{work}/R_{free} of 0.1932 / 0.2148 (Table 2.1).

I assumed the B-factors might be skewed by data in the higher-resolution shells/bins, so I decided to run four simultaneous refinements with increasingly conservative high-resolution cutoffs: 1.45 Å, 1.50 Å, 1.55 Å, and 1.60 Å. Without any resolution cutoff, there were around 50,000 reflections. At a 1.45 Å cutoff, the number of

reflections droped to 43,500, and this number continued to drop by ~3,500 reflections for each 0.05Å increase in resolution cutoff. While each refinement with a resolution cutoff succeeded in lowering the R_{work}/R_{free} values, I also lost reflections and resolution with these cutoffs. Ultimately, in an attempt to strike a balance between preserving the reflections while also cutting out skewed data in high-resolution shells, I elected to move forward with the 1.55 Å cutoff. After many rounds of refinement, the final structure of $Pex15^{43-253}$ was solved to a resolution of 1.55 Å with and R_{work}/R_{free} of 0.1738 / 0.1890.

An *ab initio* view of the final model built into the electron density for $Pex15^{43-253}$ demonstrates the high resolution (Fig. 2.7a). As predicted, the structure is entirely alpha-helical. It is composed of 12 alpha-helices that are packed together and form a roughly curved, or banana-like, shape (Fig. 2.7b). The remnant residues from PSP cleavage at the C-terminus, consisting of amino acids LEVLFQ, form an alpha-helix (colored orange in Fig. 2.7b) and seamlessly extend off the last residues of the 12^{th} alpha-helix. Throughout the center of the folded domain, numerous phenylalanines extend into the hydrophobic core. Another notable feature is that all seven cysteines of $Pex15^{1-309}$ are in this folded domain and buried inside hydrophobic core (Fig. 2.7c). Interestingly, the first six cysteines appear to be paired in the folded structure, with the side chain of cys_1 extending toward the side chain of cys_2, and so on. Given the high resolution of the structure, it is clear that none of these pairs are actually forming disulfide bridges. However, the protein was crystallized under reducing conditions (with 0.5 mM TCEP) and it is unclear whether these cysteine pairs *could* form disulfide bonds *in vivo* or under oxidizing conditions. While some of the cysteine pairs appear to be just a little too far to form disulfide bridges, it is reasonable to hypothesize that a small conformational adjustment could allow for formation of three different disulfide bridges in Pex15. It would be interesting to see if the proximity of these cysteine pairs is relevant to any Pex15 functions *in vivo*.

I subsequently used the structure of $Pex15^{43-253}$ as a model for molecular replacement to provide phases for the $Pex15^{37-253}$ crystal datasets (the hexagonal crystals with very long unit cells). Despite their distinct crystal forms, the structures and conformations of $Pex15^{37-253}$ and $Pex15^{43-253}$ turned out to be identical.

After solving the Pex15 structure, I was interested in gaining further insight into the interacting moiety on Pex1/Pex6. Pulldowns with various constructs of each component lead us to the conclusion that Pex6's N1 domain is, at least in part, responsible for binding Pex15. I attempted to solve the crystal structure of Pex6 by expressing and purifying various truncations ($Pex6^{1-179}$, $Pex6^{1-184}$, and $Pex6^{1-215}$) and setting crystal screens. I was able to obtain crystals of $Pex6^{1-184}$ in 1 M LiCl, 0.1 M citric acid pH 5.0, and 20% PEG6000. These crystals were stable in a wide variety of cryoprotectants, and diffracted to a high resolution (~1.6 Å). However, the absence of phase information presented a challenge, and attempts to express SeMet-$Pex6^{1-184}$ (or any other SeMet-Pex6 variant) failed to either purify or crystallize. Ultimately, time constraints forced me to move on from this project. Given the structural information we later determined regarding Pex15 binding the Pex1/Pex6 complex (discussed in Chapter 3), it would be interesting to obtain a higher-resolution understanding of Pex6's N domains.

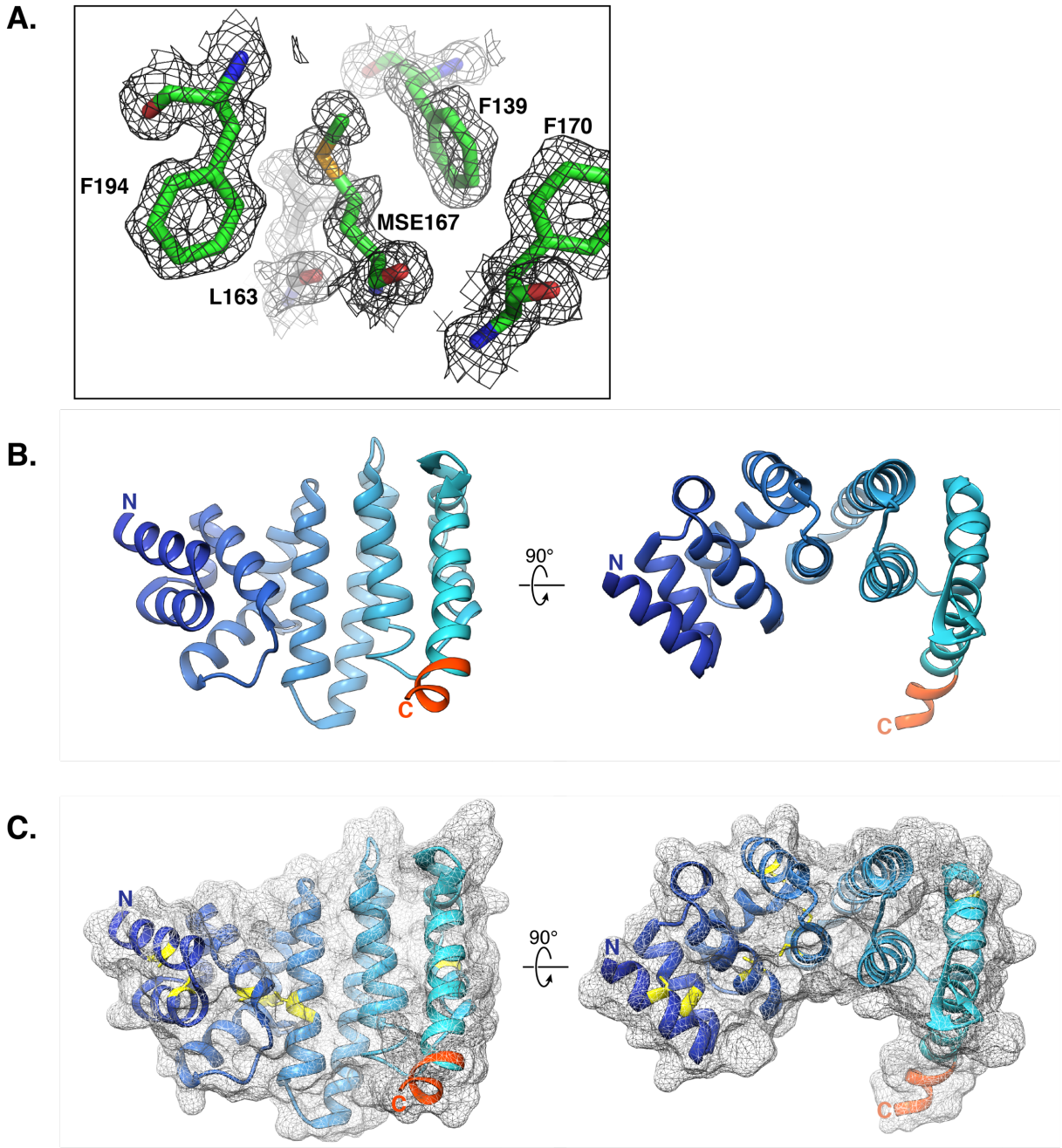


Figure 2.7 : The first glimpse of Pex15's crystal structure. (A) Example ab initio rendering of Pex15's crystal structure, depicted high-resolution density around several of its many internal phenylalanine residues, as well as one of the SeMet residues, which were used to determine phases. (B) Pex15's crystal structure depicts an entirely alpha-helical structure, with the C-terminal PSP tag (sequence LEVLFQ) ordered as a helical continuation of Pex15's final core helix (orange). (C) Surface-mesh rendering of Pex15's crystal structure, emphasizing its curved shape. Cysteines are colored yellow.

2.3 Attempted structure determination for Pex15's homologs: Pex26 and APEM9

Pex15's mammalian homolog is Pex26, a ~34 kDa TA protein in the peroxisome membrane. Although Pex15 and Pex26 have low sequence homology (13.1% identity),

they are believed to perform the same cellular role in peroxisome biogenesis: recruiting Pex1/Pex6 to the peroxisomal membrane. Pex1 (38% identity, 56% homology) and Pex6 (38% identity, 56% homology) are much more conserved from yeast to humans (NCBI database). Given these observations, I hypothesized that Pex15 and Pex26 would share structural homology. Since no structural information exists on Pex26, I set out to solve its crystal structure and, in the process, define the structured and unstructured domains of the soluble part of Pex26.

Table 2.2 : Point mutations in Pex26 reported in patients with peroxisome biogenesis disorders. These are only the missense point mutations reported in Pex26 that have appeared in patients with peroxisome biogenesis disorders. Other types of mutations (frameshifts and nonsense mutations) are not reported.

Mutation	Source
G89R	*Matsumoto et al., 2003*
M1T	*Matsumoto et al., 2003*
L45P	*Matsumoto et al., 2003*
L44P	*Furuki et al., 2006*
G98W	*Matsumoto et al., 2003*
P117L	*Furuki et al., 2006*
L153V	*Furuki et al., 2006*

There are numerous disease-related mutations identified for Pex26 (Tabel 2.2; Furuki et al., 2006; Matsumoto et al., 2003), and mapping them to a structure and analyzing their effects on Pex26's ability to bind or inhibit Pex1/Pex6 may provide a deeper understanding of the nature of this interaction and how it may relate to the varying severity of peroxisome biogenesis disorders. It is possible that several of these mutations simply disrupt the global structure of Pex26—especially for position in the hydrophobic core—but it would nevertheless be informative to describe which mutations disrupt structure and which, if any, are instead biochemically defective. Of particular interest are mutations characterized as producing "mild" phenotypes in humans: L153V and R98W (Furuki et al., 2006).

In order to be able to express and purify large quantities of Pex26, I first aimed to remove its TMD. Pex26 is 305 amino acids long, and prediction servers indicate that its TMD begins around amino acid 255 (Fig. 2.8a). Unlike Pex15, however, the boundaries of any unstructured and structured domains in Pex26 are poorly defined, and it was much more difficult to estimate where to truncate the protein. I began by cloning four different truncations of Pex26, each construct more strongly truncated from the C-terminus than the last: these constructs spanned amino acids 1-254, 1-248, 1-223, and 1-208. I expressed and purified these constructs similarly to $Pex15^{43-253}$, and only $Pex26^{1-223}$ expressed and purified well enough for crystallography screening (Fig. 2.8b). Unfortunately, I was unable to obtain or optimize crystals of this Pex26 construct. Pex15 and Pex26 share poor sequence homology, but it is possible that Pex15's unstructured N-terminal domain is an essential physical feature in the protein's function. If this were true, Pex26 may also have an unstructured N-terminal region. We elected again to employ limited proteolysis to identify the boundary between the unstructured region and the structured core. Limited proteolysis using trypsin revealed two cleavage sites in Pex26's N-terminus: after R20 and R26 (Fig. 2.8a, Fig. 2.8c). However, I was unsuccessful in crystallizing both the native and SeMet versions of $Pex26^{21-223}$ and $Pex26^{27-223}$, and time constraints forced me to move on from this effort.

If this were to be pursued further in the future, I believe a reasonable starting point would be limited proteolysis with different proteases. It is possible that the Pex26

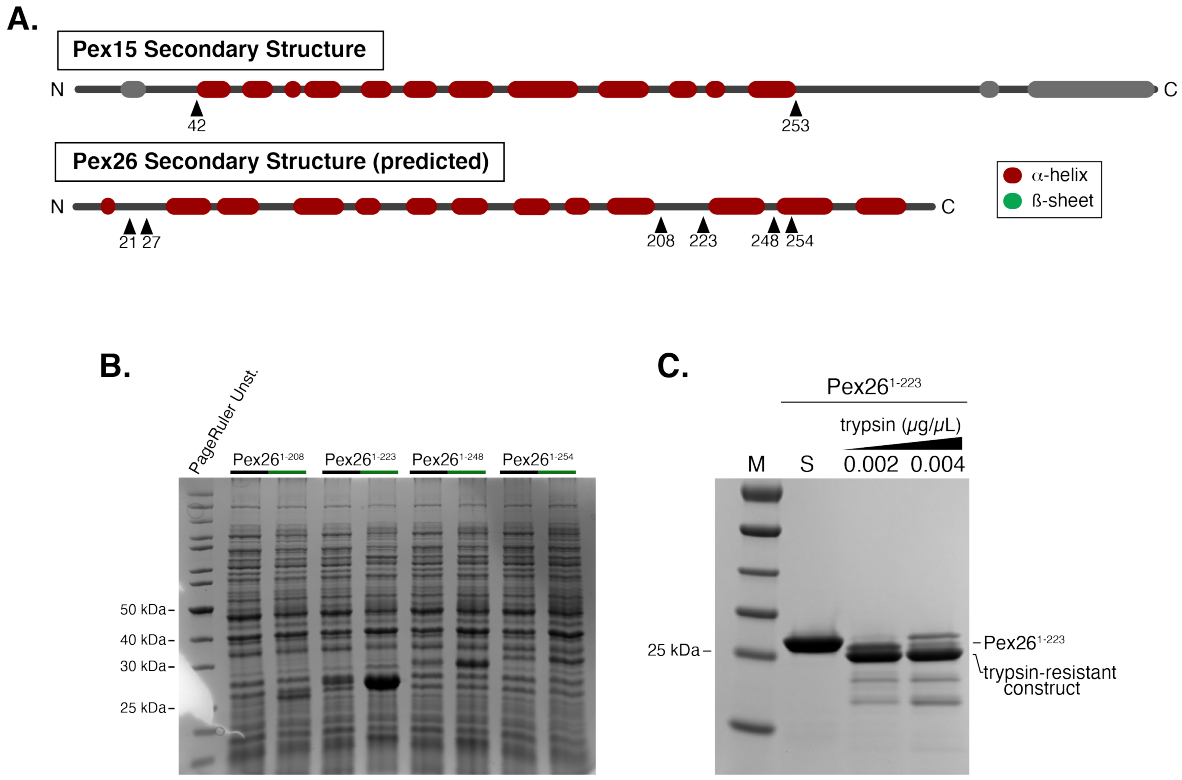


Figure 2.8 : Construct design and optimization for Pex26 crystallization. (A) Comparison of Pex15's now-known secondary structure (red) and remaining predicted secondary structure (grey) with the predicted secondary structure of the mammalian Pex26. The relevant cleavage sites discussed in Section 2.6 are marked along Pex26's primary structure depiction. (B) Pre- and post-induction gel of various Pex26 constructs (coomassie stained). $Pex26^{1-223}$ appears to express more than the other constructs. (C) Limited proteolysis of $Pex26^{1-223}$ using trypsin reveals a stable constructs only slightly smaller than the starting reactant. Mass spectrometry subsequently revealed these stable constructs are either $Pex26^{21-223}$ or $Pex26^{27-223}$.

construct suggested by trypsin still possesses unstructured regions that just happen to lack lysines and arginines, and that limited proteolysis with, for example, chymotrypsin or elastase might reveal a construct more amenable to crystallization. Furthermore, I only attempted crystallization using the JCSG I-IV screens, and there are many more crystallization screens that I did not utilize in trying to get Pex26 constructs to crystallize. A broader screening process may yield crystals without further construct optimization.

My interest in whether Pex15's structure, rather than its sequence, has been evolutionarily conserved is not restricted to its human homolog. Pex15's plant homolog, APEM9 (from *A. thaliana*), is thus defined by its homologous role in recruiting Pex1/Pex6 to the peroxisome membrane; however, like Pex26, APEM9 has very low sequence identity to Pex15 (13.5% identity). Interestingly, unlike Pex26, APEM9 is not predicted to be entirely alpha-helical (Fig. 2.9a). I approached crystal construct optimization similarly to my efforts with Pex26, first by testing expression of four different truncations from the C-terminus, each one more aggressive than the previous: $APEM9^{1-262}$, $APEM9^{1-243}$, $APEM9^{1-234}$, and $APEM9^{1-204}$. These truncations were specifically chosen based on the primary sequence and predicted secondary structure of APEM9—each is truncated just before or just after predicted secondary structure, taking into account the nature of nearby amino acids.

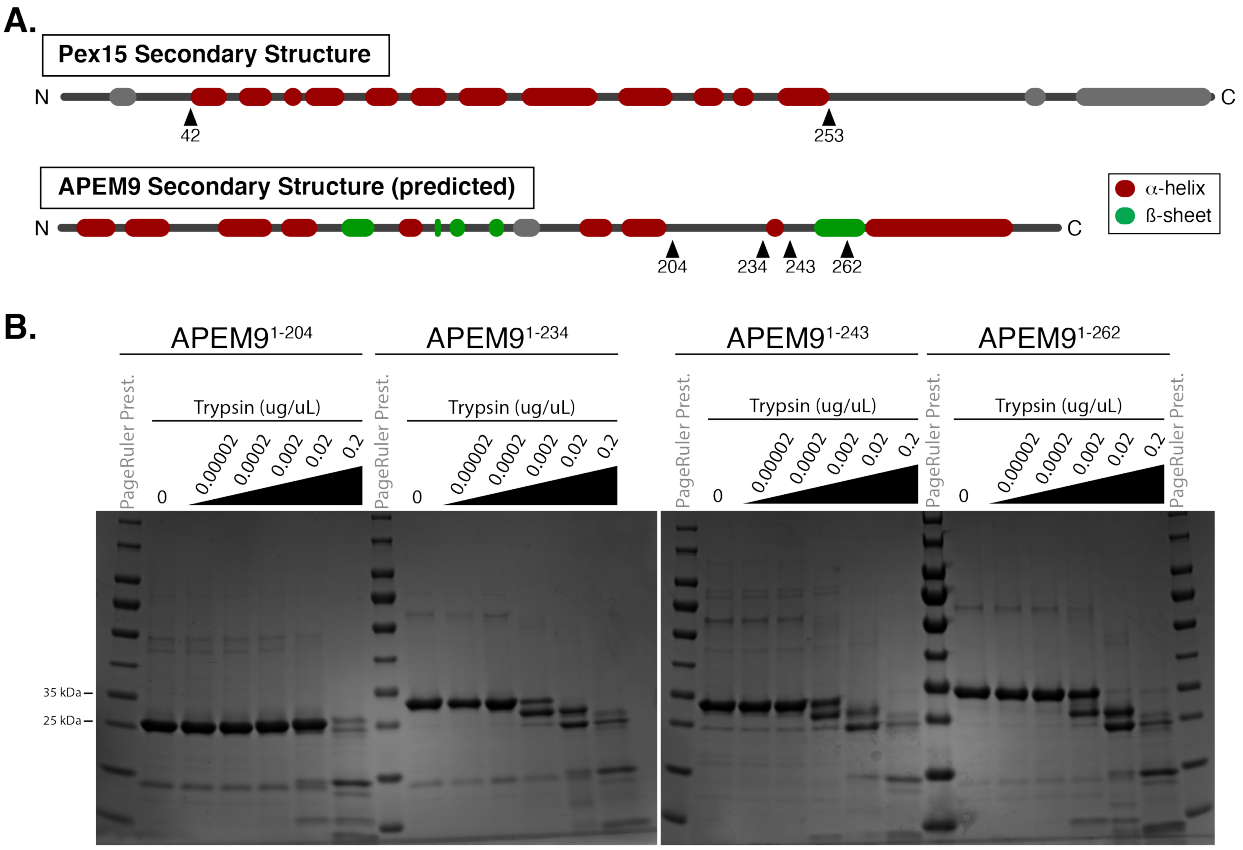


Figure 2.9 : Construct design and optimization for APEM9 crystallization. (A) Comparison of Pex15's now-known secondary structure (red) and remaining predicted secondary structure (grey) with the predicted secondary structure of the plant homolog APEM9. The relevant cleavage sites discussed in Section 2.7 are marked along APEM9's primary structure depiction. All constructs expressed and purified similarly (data not shown). (B) Limited proteolysis of all APEM9 constructs shows that, except for the smallest construct, all APEM9 variants subjected to trypsin proteolysis are reduced to a slightly smaller stable construct. We believe the smallest construct of APEM9$^{1-204}$ represents a construct very similar to the stable domains produced from trypsin digestion.

I expressed and purified all four constructs similarly to Pex15$^{43-253}$. While each expressed and purified quite well, the two shortest constructs appeared to be better behaved than the two longer constructs. Additionally, it does not appear that overnight cleavage with PSP to remove the affinity tag destabilized the protein or reduced its solubility. To determine the structural stability of these various constructs, I performed limited proteolysis. The longest constructs all appeared to be readily cleaved to form slightly smaller truncations of APEM9 (Fig 2.9b). Interestingly, however, APEM9$^{1-204}$ appeared not to be cleaved at all. I interpreted this finding as APEM9$^{1-204}$ already being a stable truncation of the protein, possessing few or small unstructured regions that contain lysines or arginines. I therefore set out to set crystallization trays of this construct at 5, 10, and 15 mg/mL protein.

One screen in particular—the PACT screen—yielded numerous hits in similar conditions, which included 0.1 M Bis Tris Propane and 20% w/v PEG3350 at various pH's (pH 6.5, 7.5, and 8.5) in addition to 0.2 M of various sodium halide salts. Interestingly, pH appeared to have a strong influence on crystal growth. Although there were clearly protein crystals forming in all three wells, a solution at pH 6.5 formed large, chunky prisms ideal for crystallography. The crystals in the pH 7.5 solution were smaller

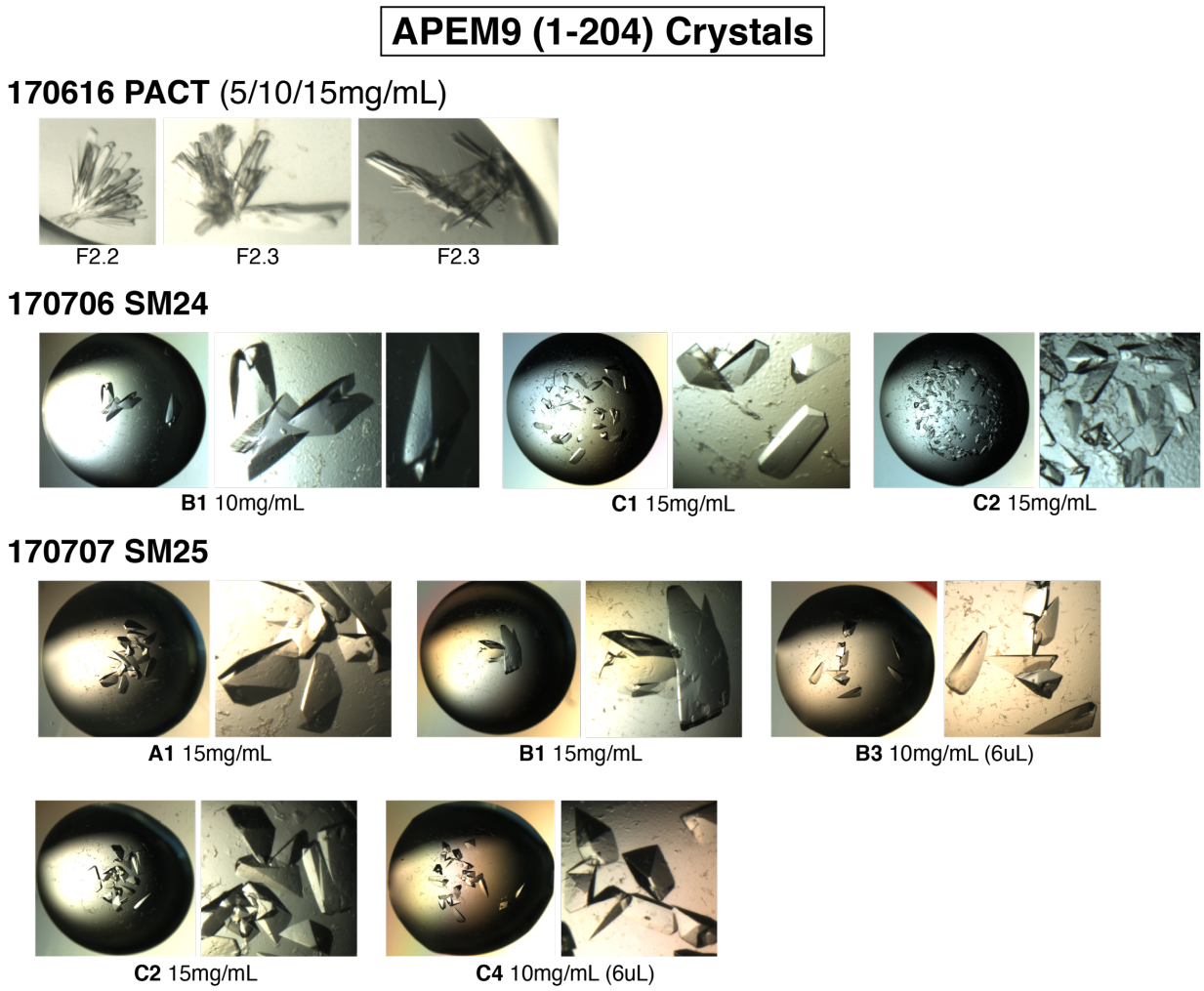


Figure 2.10 : APEM9 crystal forms. APEM9 crystallized readily in a series of related conditions described in Section 2.7.

and flatter, and the crystals in the pH 8.5 solution were almost microcrystalline. The best hit was in 0.1 M Bis Tris Propane pH 6.5, 0.2 M NaBr, 20% w/v PEG3350, so we set crystal trays of a focused screen surrounding these conditions. The first screen held 0.1 M Bis Tris Propane pH 6.5 constant and cross screened PEG3350 (15% to 25%, in 2.5% intervals) against NaBr (0.175 M to 0.225 M, in 0.125 M intervals). After identifying the best crystals at high NaBr concentrations and lower PEG3350 concentrations, we concocted a second screen which again held 0.1 M Bis Tris Propane pH 6.5 constant and cross screened PEG3350 (12% to 18%, in 1.5% intervals) against NaBr (0.210 M to 0.240 M, in 0.015 M intervals). This screen produced numerous large crystals ideal for looping and X-ray diffraction (Fig. 2.10).

While it was promising that APEM9$^{1-204}$ so readily crystallized, its predicted secondary-structure differences made molecular replacement with Pex15 for structure determination not feasible, and I decided to purify and crystallize SeMet-APEM9$^{1-204}$. Fortunately, this version of the protein expressed and purified well, and we were able to obtain crystals in the same conditions as the native protein. All proteins were purified into a minimal crystallography buffer (20 mM HEPES, 50 mM NaCl, 50 mM KCl, 0.5 mM TCEP, pH 7.6) prior to setting trays.

Looping the APEM9$^{1-204}$ and SeMet-APEM9$^{1-204}$ crystals proved challenging. Similar to Pex15, these crystals appear to disintegrate upon looping into a variety of cryoprotectants, including D-sorbitol, xylitol, MPD, glycerol, PEG400, ethylene glycol, and Bushmills whiskey (perhaps unsurprisingly, this one was the worst). I employed the same technique here as I did with Pex15 crystals, washing crystals in increasingly-higher concentrations of cryoprotectants from 5% to 25% in 5% increments. This process certainly was gentler on the proteins, but still I observed crystals fracturing and dissolving in the process. Ultimately, crystals were frozen in a variety of cryoprotectants at 20% or 25%. I also adjusted the crystallization conditions to include anywhere from 0.5% to 6% of the two most promising cryoprotectants based on previous looping efforts: glycerol and ethylene glycol. Additionally, I tried different combinations of these cryoprotectants, screening 1-3% glycerol against 1-3% ethylene glycol. I tried this technique with both native and SeMet crystals, but the crystals either did not grow or they continued to disintegrate when shifted to a higher cryo concentration. Unfortunately, time constraints forced me to move on from this effort without solving the structure.

If this were to be further pursued in the future, I believe there are a multitude of optimization avenues. First, it is possible that growing the crystals at a colder temperature would encourage slower precipitation and, hopefully, more stable and ordered unit cells. Secondly, there were many crystallization conditions from the PACT screen that we did not optimize. In particular, optimization of crystallization in sodium iodide, sodium thiocyanate, and sodium acetate are promising starting points, as they were the best-looking precipitates after the sodium bromide condition. One last optimization opportunity is, again, from limited proteolysis. Although APEM9$^{1-204}$ barely changed at all through exposure to trypsin, it is possible that APEM9$^{1-204}$ does indeed possess unstructured regions that just happen to lack lysines and arginines. Limited proteolysis with a different protease, such as chymotrypsin or elastase, may reveal other constructs amenable to crystallization.

2.4 Concluding remarks on Pex15's domain architecture and structure

Prior to this work, structural information for Pex15 or any of its known homologs was lacking. I optimized expression of a TMD-less Pex15 construct and defined the basic domains of Pex15: an unstructured N-terminal region (aa 1-42), a folded and highly alpha-helical core domain (aa 43-253), and an unstructured C-terminal region (aa 254-309). Beyond these residues lies a region of undefined order: a patch of hydrophobic residues spanning amino acids 313-324 that may form an alpha-helix. Generally speaking, inclusion of this region in any Pex15 construct correlates with weaker expression and poor solubility (especially when the affinity tags were removed via PSP cleavage). I solved the crystal structure of Pex15's core domain to a resolution of 1.55 Å, providing the first glimpse of this understudied component of peroxisome biogenesis. While I was unsuccessful in solving the crystal structures of the mammalian (Pex26) and plant (APEM9) homologs of Pex15, I made several important steps toward that goal, including identifying unstructured regions and even defining a host of crystallization conditions in the case of APEM9. In the future, it would be very interesting

to see if Pex15 homologs, which generally have poor sequence homology, are structurally conserved. Successful structure determination in the future will reveal whether natural selection helped preserve structural homology for these homologs despite their little sequence homology.

CHAPTER 3

Pex1/Pex6 unfolds Pex15 by processive threading

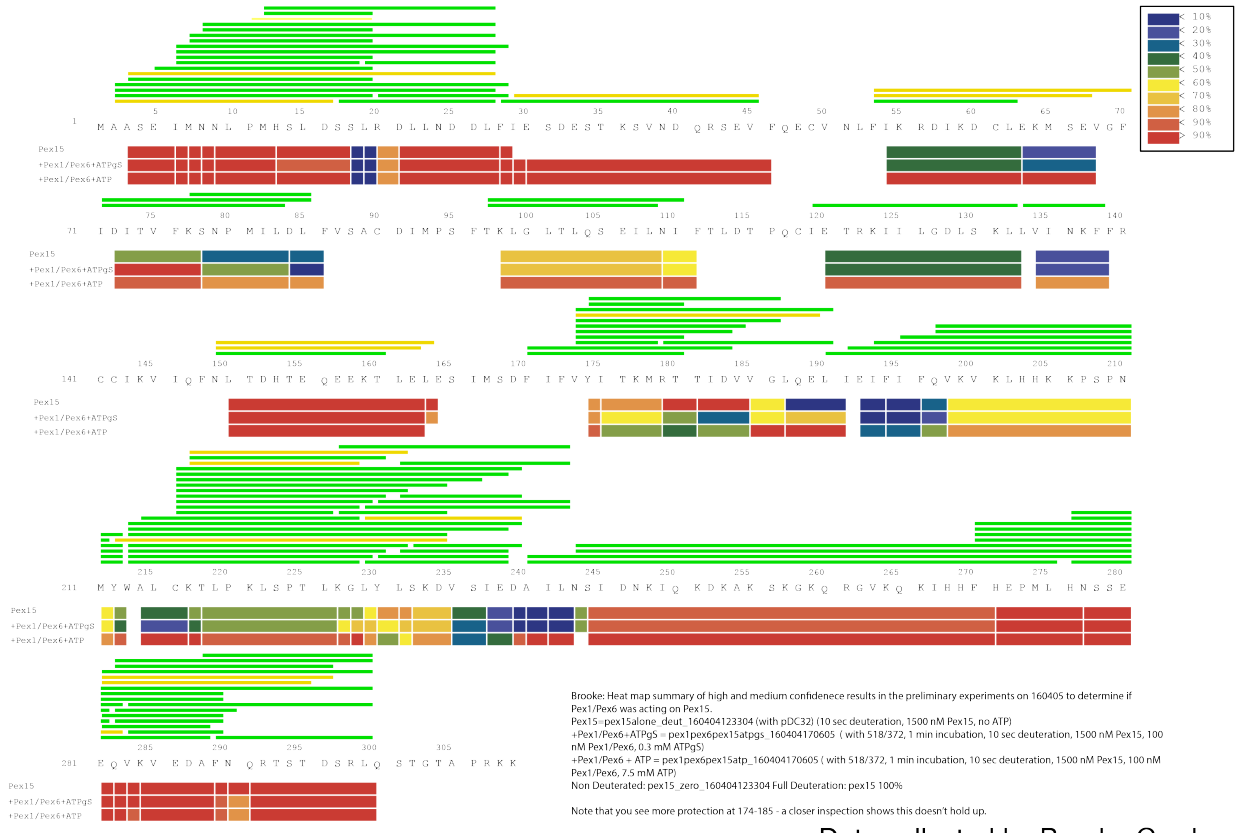
While it is relatively trivial to demonstrate that a member of the AAA+ ATPase family hydrolyzes ATP and that this activity is essential to its function *in vivo*, identifying the subject of that biochemical work is far more challenging. Some ATPases have a multitude of diverse substrates, while others are far more selective. In any case, although identifying substrates can be challenging, it provides valuable insight into the role of a AAA+ ATPase in the cell. It has long been known that Pex1 and Pex6 are AAA+ ATPases associated with peroxisome biogenesis. An earlier publication from our lab demonstrated that these two proteins form a heterohexameric motor with alternating Pex1 and Pex6 subunits that work together to hydrolyze ATP (Gardner et al., 2015). Furthermore, it was shown in this publication that $Pex15^{1-327}$ both binds Pex1/Pex6 and inhibits the motor's ATPase activity down to 20% of the motor's activity in isolation. While it has long been assumed that the shuttle receptor Pex5 is Pex1/Pex6's substrate, no one has been able to measure a physiologically-relevant affinity, much less so demonstrated that Pex5 is a Pex1/Pex6's substrate. We serendipitously discovered that Pex1/Pex6 unfolds soluble Pex15, and we characterized the mechanics of this process through extensive investigation of Pex1/Pex6 as an unfoldase and Pex15 as a substrate.

3.1 Pex1/Pex6 unfolds Pex15

Although it has long been established through *in vivo* experiments that Pex15 functions to recruit Pex1/Pex6 to the peroxisomal membrane, the nature of this interaction remained unclear. Specifically, little was known about the binding interface among these proteins. All that was known came from an *in vivo* mutagenesis screen of Pex15 that identified one residue in particular, L22, that, when mutated, weakened Pex15's interaction with Pex6, as determined by a yeast two-hybrid assay (Birschmann et al,. 2003). While limited proteolysis previously informed us that the first 42 amino acids of Pex15 are largely unstructured (Fig. 2.2c), secondary structure predictions suggested this region to form a short helix (Fig. 2.1a). Furthermore, it is possible this region is unstructured in isolation, but forms a helix when interacting with the Pex1/Pex6 complex.

To gain a higher resolution of Pex15's binding interface with the motor, a postdoc in the lab, Brooke Gardner, performed hydrogen-deuterium exchange (HDX) on $Pex15^{1-309}$ in both the absence and presence of Pex1/Pex6. Since amino acids 1-42 in Pex15 are generally unstructured and therefore solvent-exposed—and since mutations of L22 within that region appear to disrupt interaction with Pex1/Pex6 in *in vivo* assays—we hypothesized that this region would be protected from solvent in the presence of Pex1/Pex6 and thus show a slower rate of hydrogen-deuterium exchange. Furthermore, if other areas of Pex15 contribute to recruiting Pex1/Pex6 to the peroxisomal membrane, HDX would be helpful in mapping those on the recently solved structure of $Pex15^{43-253}$.

As expected, HDX of $Pex15^{1-309}$ revealed a quick rate of amide hydrogen exchange for amino acids 1-46 and 244-309, in general agreement with the boundaries between unstructured and structured segments determined by limited proteolysis (Fig.

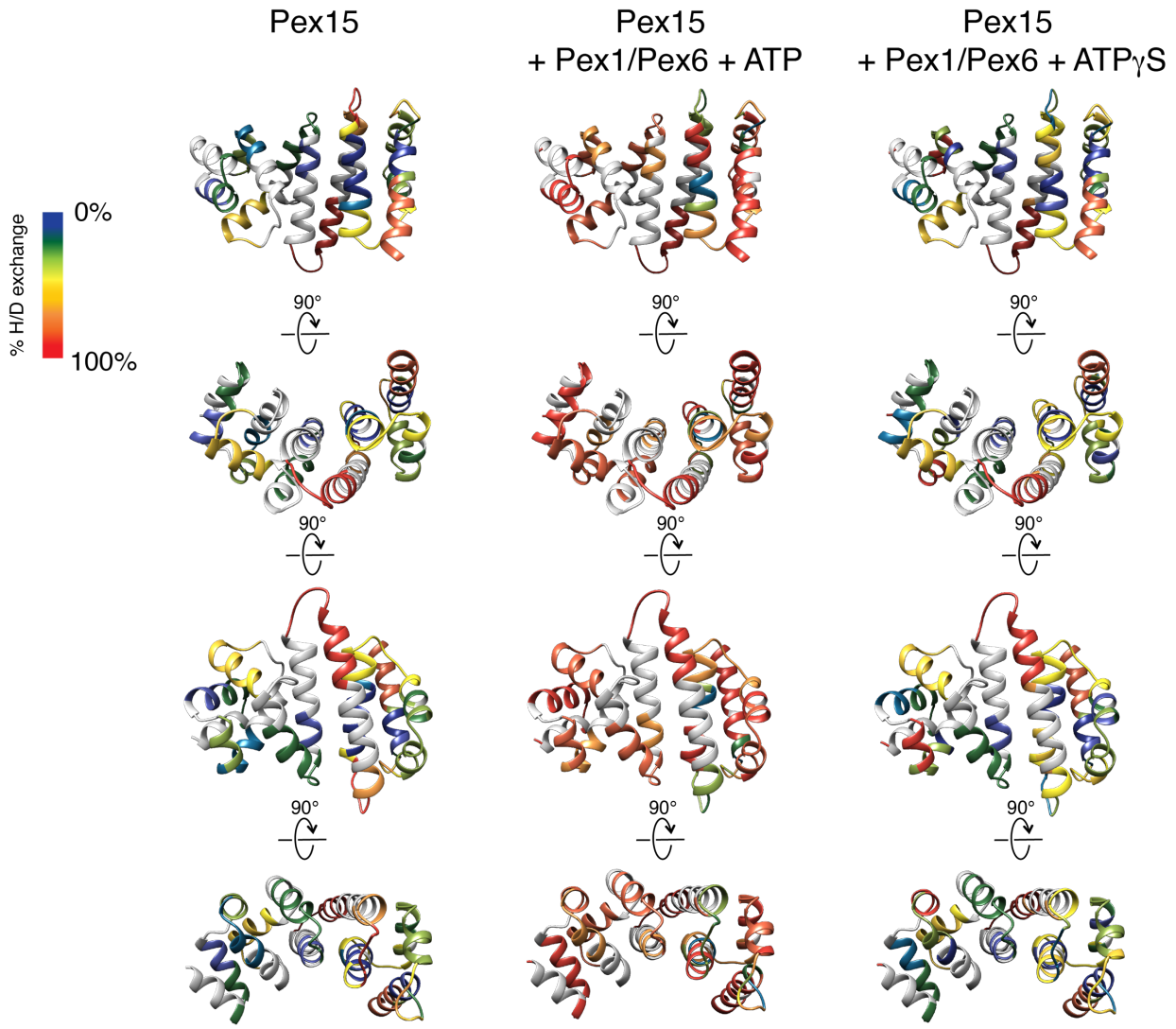


Data collected by Brooke Gardner

Fig. 3.1 : HDX peptide coverage of Pex15 alone, in the presence of Pex1/Pex6 with ATPγS, and in the presence of Pex1/Pex6 with ATP. Data collected by Dr. Brooke Gardner depicting peptide coverage throughout Pex15$^{1\text{-}309}$ after pepsin digestion (thin green and yellow bars above the primary sequence). The colored blocks below Pex15's primary sequence reflects the different conditions of Pex15 alone (top), Pex15 with Pex1/Pex6 preincubated in ATPγS (middle), and Pex15 with Pex1/Pex6 preincubated in ATP (bottom). The blocks are color-coordinated to match the rate of amide-hydrogen exchange, where red represents quickly-exchanging regions and blue represents slowly-exchanging regions.

3.1, Fig. 2.2c). Moreover, amide protons of the core domain (amino acids 47-243) generally exchanged much more slowly, which we would expect from residues within secondary and tertiary structures of a protein. Brooke then performed the same reaction in the presence of Pex1/Pex6 and ATP. However, rather than observing a decrease in rate of HDX anywhere on Pex15, she unexpectedly saw a global increase in rate throughout the protein, most notably within the core domain (Fig. 3.1). These data suggested that Pex1/Pex6 is completely unfolding Pex15. To confirm this hypothesis, Brooke performed the same reaction in the presence of ATPγS instead of ATP and observed exchange rates throughout Pex15 that resembled those for Pex15 alone, suggesting that Pex1/Pex6 is unfolding Pex15 in an ATPase-hydrolysis-dependent manner (Fig. 3.2).

Since this was the first demonstration of Pex1/Pex6 interacting with any other protein as a substrate, we decided to investigate its unfoldase capabilities further in order to understand more about the underlying mechanism. While HDX is a fine-tuned method able to provide extensive information on the tertiary structure of a protein—especially when the structure has already been solved—it is a labor-intensive and time-consuming process and is thus poorly suited for screening many motor or substrate mutants. I therefore set out to develop a more tractable unfoldase assay using



Data collected by Brooke Gardner

Fig. 3.2 : Peptide rate of HDX (Pex1/Pex6 + ATP) mapped onto Pex15's structure. Rates of amide hydrogen exchange determined in Fig. 3.1 mapped onto Pex15's recently-solved structure. Notably, many regions of $Pex15^{1\text{-}309}$ which are slowly-exchanging in isolation become quickly-exchanging in the presence of Pex1/Pex6 + ATP, but not in the presence of Pex1/Pex6 + ATPγS.

Pex1/Pex6 and Pex15. Unless otherwise stated, the standard substrate in these optimization assays is $Pex15^{1\text{-}309}$ with its affinity purification tags cleaved off.

We ultimately designed an assay that took advantage of $Pex15^{1\text{-}309}$'s exclusively-internal cysteine residues (Fig. 2.7c). In a fully folded protein, these cysteines should be inaccessible to even small reactive molecules, such as maleimide; however, upon unfolding, they would be free to react, for instance with fluorescein-5-maleimide (F5M). By incubating Pex15 with Pex1/Pex6 and F5M, and then running the reaction on a SDS PAGE gel, we should be able to visualize the extent to which Pex15 has been unfolded by imaging in the fluorescein channel—the more Pex15 got unfolded by Pex1/Pex6, the brighter the Pex15-corresponding band will be on the gel. Since Pex1/Pex6 contains Cys residues itself, we were initially not sure whether it would remain functional in the

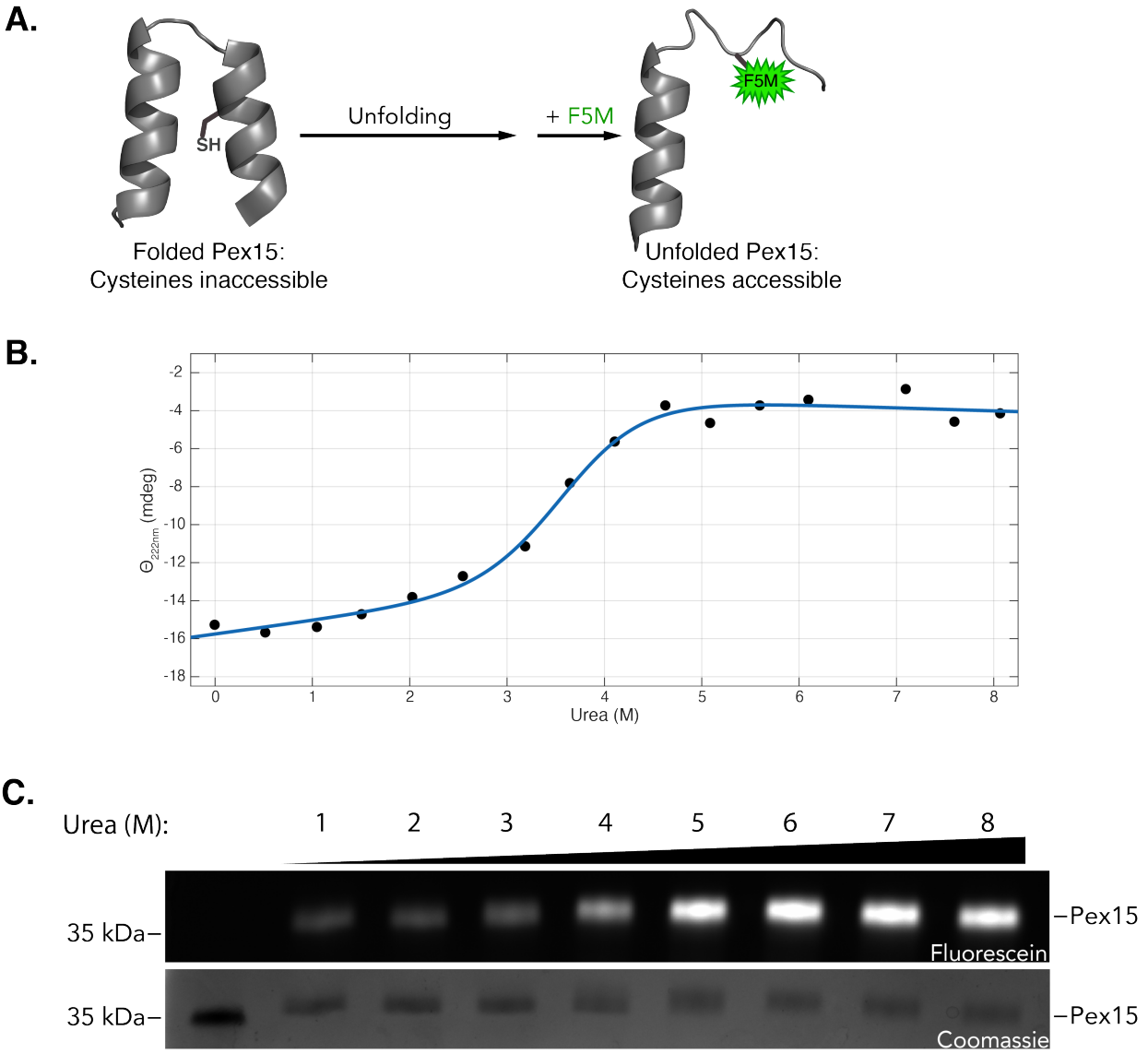


Fig. 3.3 : Establishing a positive control for maleimide unfoldase assays. (A) Cartoon depiction of the desired unfoldase assay, whereby internal cysteines are protected in the folded state, but become exposed and accessible for fluorescein-5-maleimide labeling upon unfolding by either urea or a protein motor. (B) CD spectroscopy of Pex15$^{1\text{-}309}$ incubated overnight in increasing urea concentrations, suggesting a Cm of ≈ 3.5 M. (C). SDS-PAGE gel imaged in the fluorescein channel (and subsequently coomassie stained) of Pex15 incubated in the presence of both F5M and increasing urea concentrations. Pex15 appears to be highly labeled by F5M at urea concentrations which CD indicated were sufficient to unfold Pex15. In all future maleimide-labeling reactions, 6 M urea was used as a positive control of unfolding and F5M labeling.

presence of F5M, so we elected to perform the reaction by first incubating Pex1/Pex6 with Pex15 for a defined duration before pulse labeling with F5M (Fig. 3.3a).

We also sought to include a positive control for Pex15 unfolding in order to provide a maximum-unfolded reference for each experiment and normalize potential variations in protein concentration. Our first attempt included incubating a separate sample of Pex15 in guanidinium chloride (GdmCl) to chemically denature Pex15, thus exposing all cysteines. While GdmCl undoubtedly denatures Pex15, its ionic character also significantly changes the way proteins migrate on an SDS PAGE gel, thus confounding not only the analysis of protein in that lane, but also the adjacent lanes.

Instead of GdmCl, we ultimately settled on using urea to chemically denature Pex15, which does not distort protein bands on SDS PAGE gels.

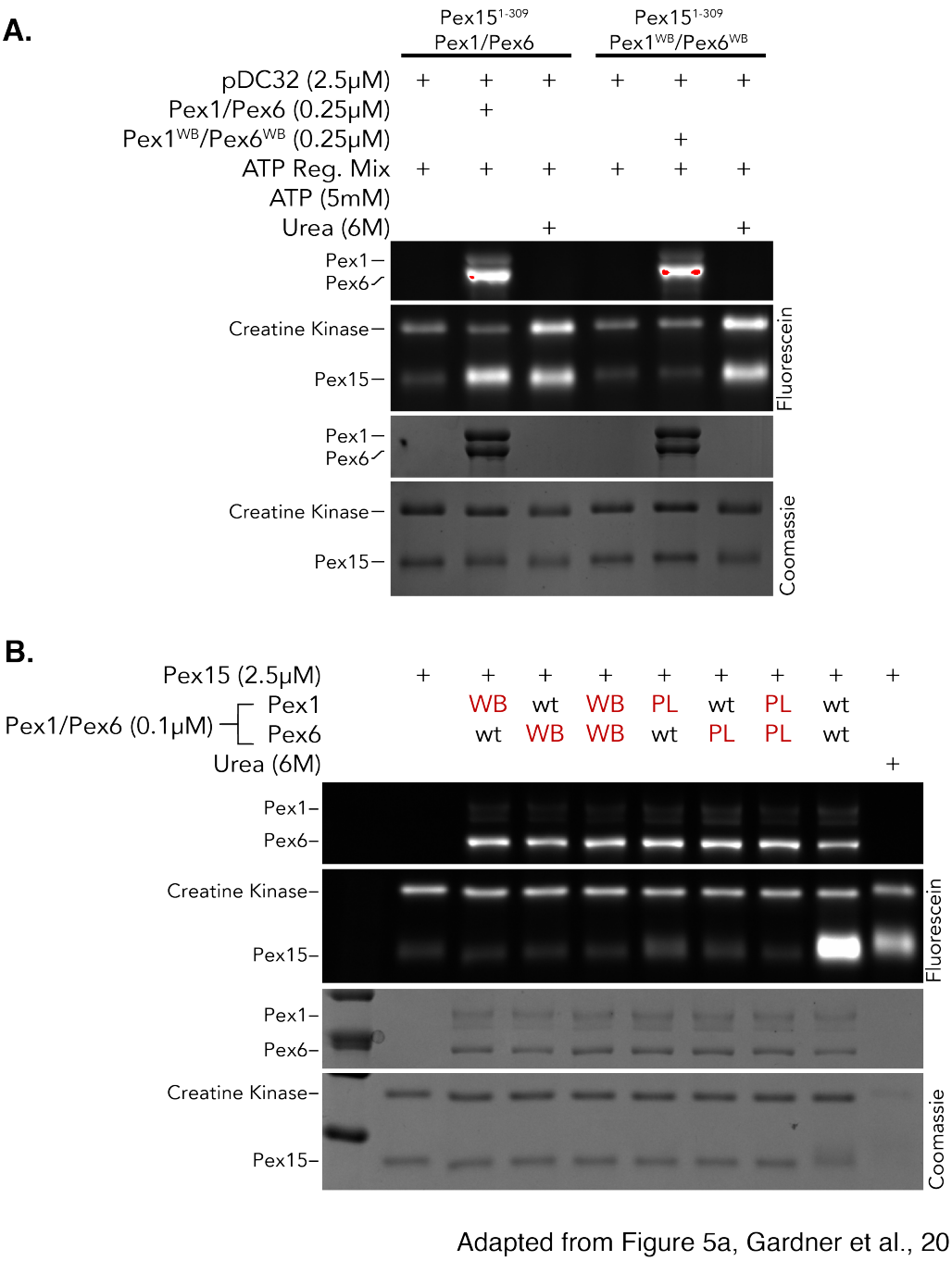
To determine the concentration of urea required for complete unfolding of Pex15, I incubated $Pex15^{1-309}$ with various urea concentrations at 24°C (RT) and monitored its secondary-structure content by CD spectroscopy. The midpoint of this urea-induced Pex15 unfolding was observed at C_m ~3.5 M (Fig. 3.3b). I elected to use 6M urea in the unfoldase assays, which were all performed at RT unless otherwise stated. As confirmation that this concentration of urea would serve as a positive control for unfolded Pex15, I incubated Pex15 in different concentrations of urea, added F5M, and analyzed it on a gel. The maleimide labeling agreed with my CD data and confirmed that Pex15's cysteines are maximally exposed at 6M urea (Fig. 3.3c).

Proteins are not rigid structures, and even well-folded domains such as Pex15 "breathe" as parts of them transiently sample the unfolded state. To account for this spontaneous unfolding and labeling with F5M, I included a negative control of Pex15 alone, without Pex1/Pex6. Additionally, while maleimide is highly reactive toward cysteines, it can to a lower extent also react with exposed lysines, and the Pex15-only control accounted for this as well.

Incubation of $Pex15^{1-309}$ with Pex1/Pex6 leads to a dramatic increase in labeling by F5M (Fig. 3.4a). Since the experimental conditions here were similar to those for HDX, we interpreted this increase in fluorescent labeling as Pex15 being unfolded by Pex1/Pex6. As further evidence that Pex1/Pex6 is responsible for this unfolding, I demonstrated that ATP-hydrolysis incompetent Walker B (WB) mutants of Pex1 and Pex6 are unable to unfold Pex15, as determined by low F5M labeling (Fig. 3.4a). Another notable control is the inclusion of an ATP regeneration mix. This mixture, included in all samples, contains creatine kinase, which also gets labeled with F5M could serve as a fluorescently labeled loading control in each experiment.

We first applied this maleimide labeling unfoldase assay to further our understanding of how Pex1/Pex6 acts on substrates. We screened a panel of Pex1/Pex6 mutants and compared their abilities to unfold Pex15. Having already established that a double $Pex1^{WB}/Pex6^{WB}$ mutant is unable to unfold Pex15, we tested whether motors with only one type of subunit mutated—either $Pex1^{WB}$ or $Pex6^{WB}$—were partially active. Unsurprisingly, nether $Pex1^{WB}$/Pex6 nor Pex1/$Pex6^{WB}$ were capable of unfolding Pex15. Since Pex1 and Pex6 subunits alternate in the Pex1/Pex6 heterohexamer, single-WB-mutant motors contain hydrolysis-incompetent ATPase pockets, and it is not surprising that they were unable to unfold Pex15 (Fig. 3.4b). Along these lines, we also evaluated single and double arginine finger mutants in the Pex1/Pex6 complex and found all mutants to be completely unfoldase-incompetent (data not shown).

Next, we sought to understand more about the mechanism of how Pex1/Pex6 unfolds Pex15. Hexameric AAA+ ATPases are capable of unfolding substrates by different approaches. One of them is a tug-and-release mechanism, utilized for instance by the N-ethylmaleimide-sensitive factor (NSF). NSF binds substrates via its N-terminal domain, which, upon ATP hydrolysis in the AAA domain, undergoes a large conformational change, leading to substrate unfolding. Not all motors show such significant conformational changes of their N-domains, and instead rely on substrate interactions with their ATPase domains to perform mechanical work. For these motors,



Adapted from Figure 5a, Gardner et al., 2018

Fig. 3.4 : Pex1/Pex6 unfolds Pex15, as detected by maleimide labeling unfoldase assays. (A) The Pex15-corresponding band is highly labeled with F5M in the presence of wild-type Pex1/Pex6, but not Walker B mutant Pex1/Pex6. Pex6 is also readily labeled by F5M. SDS PAGE gels were imaged first in the fluorescein channel, then coomassie stained. (B) Single- and double-mutants of Walker B motifs and pore loops all abrogate unfoldase activity by Pex1/Pex6, suggesting that the deficiencies of one subunit affects an adjacent, wild-type subunit. Pex15$^{1-309}$ was used for all assays in this figure.

pore loops projecting into the central pore of the hexamer engage with and pull on substrates, mechanically unfolding them in a hand-over-hand-like motion of individual AAA domains within a spiral staircase arrangement. Pore loops typically interact with substrate polypeptide chains through large aromatic residues (tryptophans, tyrosines,

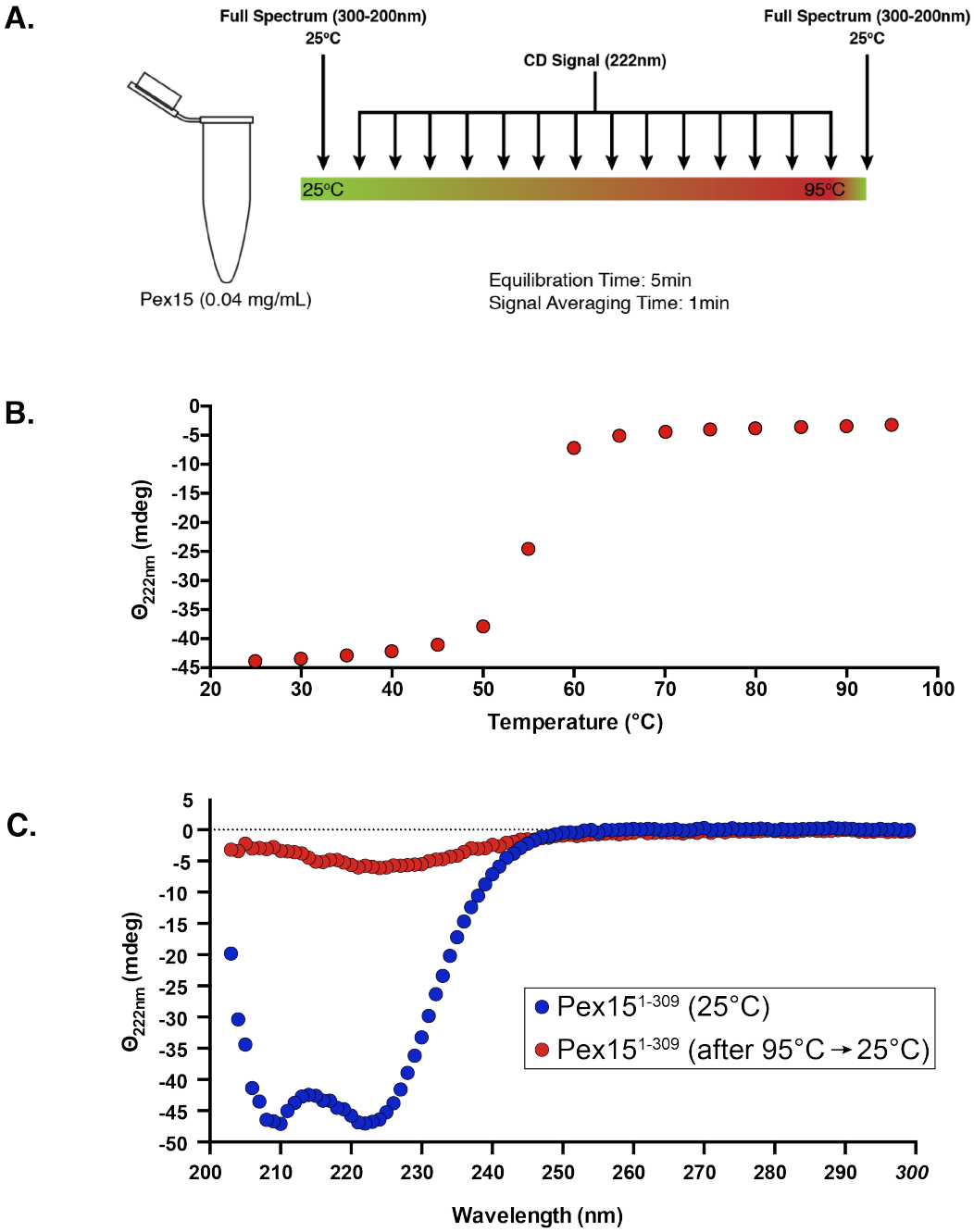


Fig. 3.5 : Toward measuring a refolding rate for Pex15: thermal denaturation. (A) Schematic depicting the proposed method of measuring Pex15 refolding, wherein CD spectra of $Pex15^{1-309}$ are measured before heating, throughout the heating process, then after the protein has been cooled to RT again and had a chance to refold. Full spectra are measured at the beginning and end, whereas measurements at Θ = 222nm are collected throughout the heating. (B) Θ_{222nm} measurements throughout heating indicate a T_m of ≈ 55°C. (C) Superimposed full CD spectra of Pex15 at RT before (blue) and after (red) heating to 95°C. The lack of alpha-helical signal after heating/cooling indicates that Pex15 has not refolded.

phenylalanines), and thus fail to grip their substrates when these residues are mutated to smaller amino acids.

Brooke had previously used negative stain EM to determine that the N-terminal domains of Pex1/Pex6 do not undergo significant conformational changes when comparing the motor in hydrolyzable ATP vs ATPγS, so we hypothesized that Pex1/Pex6 might rely on its pore loops and vertical AAA-domain movements to unfold substrates. We therefore compared the substrate-unfolding activity of wild-type Pex1/Pex6 to that of the pore-loop (PL) mutant variants Pex1^{PL}/Pex6, Pex1/Pex6^{PL}, and Pex1^{PL}/Pex6^{PL}. We found the double mutant and Pex1/Pex6^{PL} to be completely devoid of unfoldase activity (Fig. 3.4b). Surprisingly, however, the Pex1^{PL}/Pex6 mutant retained a low level of unfoldase activity (Fig. 3.4b). These data might suggest that Pex6 contributes more to substrate threading than Pex1, but we are unclear why this may be the case. Regardless, the observation that the double mutant is completely unable to unfold Pex15 supports the model that Pex1/Pex6 unfolds its substrates via processive threading through its central pore.

Demonstrating Pex1/Pex6's ability to unfold Pex15 was both novel and unexpected, as it is unknown what this may accomplish *in vivo* in the context of peroxisome biogenesis. While it is not impossible, no *in vivo* data suggest that Pex1/Pex6 extracts Pex15 from the peroxisome membrane. We therefore hypothesized that unfolding Pex15 may disrupt quaternary structure of protein complexes that Pex15 participates in. Furthermore, since persistence of unfolded Pex15 in the membrane would likely be inadvertently disruptive to other proteins and complexes, we hypothesized that Pex15 should be able to refold after being unfolded. I thus attempted to determine whether Pex15$^{1\text{-}309}$ could refold.

As proof of concept that Pex15$^{1\text{-}309}$ unfolded reversibly, I tested its ability to refold after thermal denaturation. Using CD spectroscopy, I first collected a full spectrum (200-300 nm) at RT, then gradually heating the sample to 95°C in 5°C increments, allowing the sample to equilibrate for 5 min at each temperature before measuring the CD signal at 222 nm (Fig. 3.5a). This measurements clearly indicated cooperative Pex15 unfolding, with a T_m around 55°C (Fig. 3.5b).Once at 95°C, the sample was cooled and another full spectrum measured. If protein unfolding were fully reversible, the post-denaturing spectrum should match the pre-denaturing spectrum, but this was not the case for Pex15. Even after being cooled, Pex15 remained unfolded (Fig. 3.5c). While these data might suggest that Pex15 cannot refold, many reversibly unfolding proteins aggregate if slowly heated up and held at high temperatures for an extended time.

Another way to determine whether a protein can refold is dilution out of a chemical denaturant. I incubated two samples of Pex15$^{1\text{-}309}$ in 8 M urea overnight (~19 h), then diluted one of them ten-fold to 0.8 M urea (a concentration at which Pex15$^{1\text{-}309}$ should be fully folded, judging by earlier CD data in Fig. 3.3b), and diluted the other sample into 8 M urea. Following dilution to 0.8 M or 8 M urea, the samples were again allowed to equilibrate at RT overnight (~20 h; Fig. 3.6a). The sample remaining in 8 M urea should be completely unfolded, whereas the sample diluted out of urea should resemble the native protein if it refolds. The data show that the CD spectra of both samples is identical and exhibit no evidence of secondary structure (Fig. 3.6b). From this, we conclude that Pex15$^{1\text{-}309}$ is unable to refold following rapid dilution out of chemically-denaturing conditions.

Although Pex15 was unable to refold following thermal or chemical denaturation, these shifts in physical and solvent parameters may poorly resemble the mechanical

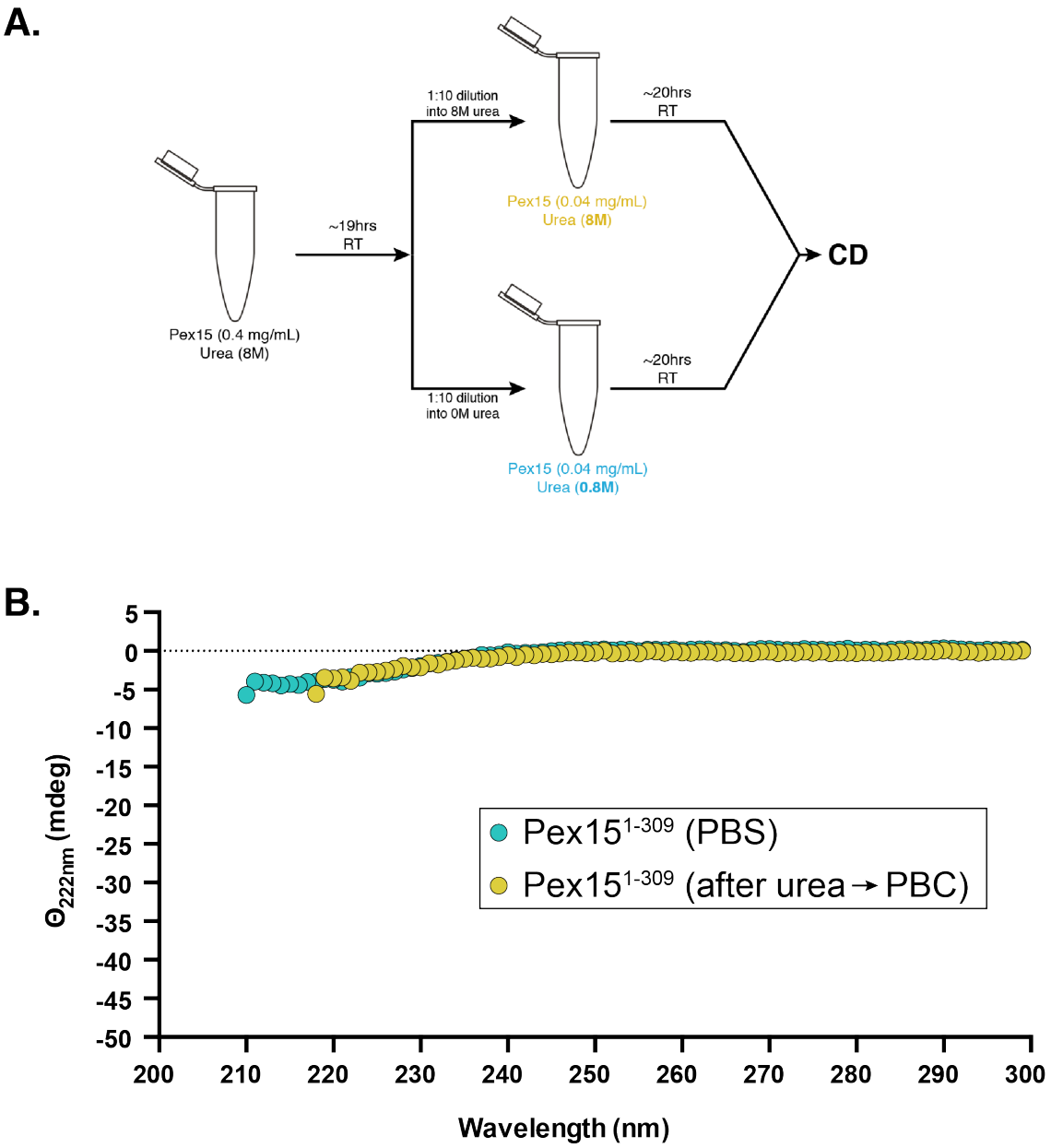


Fig. 3.6 : Toward measuring a refolding rate for Pex15: chemical denaturation. (A) Schematic depicting the proposed method of measuring Pex15 refolding, wherein Pex15$^{1\text{-}309}$ is unfolded chemically in 8 M urea, then diluted to ten-fold to 0.8 M urea using PBS (a concentration at which Pex15 can remain fully-folded, as seen in Fig. 3.3b). This is compared to Pex15 that has been diluted using 8 M urea, in which the protein should certainly remain unfolded. (B) Full CD spectra of Pex15 diluted in PBS (cyan) and 8 M urea (yellow) bear no semblance of its natural secondary structure, indicating Pex15 is unable to refold when diluted out of urea in this experimental setup.

unfolding that occurs when Pex1/Pex6 pulls on Pex15 and a potential co-translocational refolding when Pex15 exits the ATPase ring on the other side. Thus, to assess refolding in a scenario more similar to what may happen in the cell, I sought to establish a refolding assay using the maleimide labeling strategy previously employed. The idea was to allow Pex1/Pex6 to unfold Pex15$^{1\text{-}309}$ and then halt unfolding by deactivating Pex1/Pex6, allowing Pex15 to refold prior to addition of F5M.

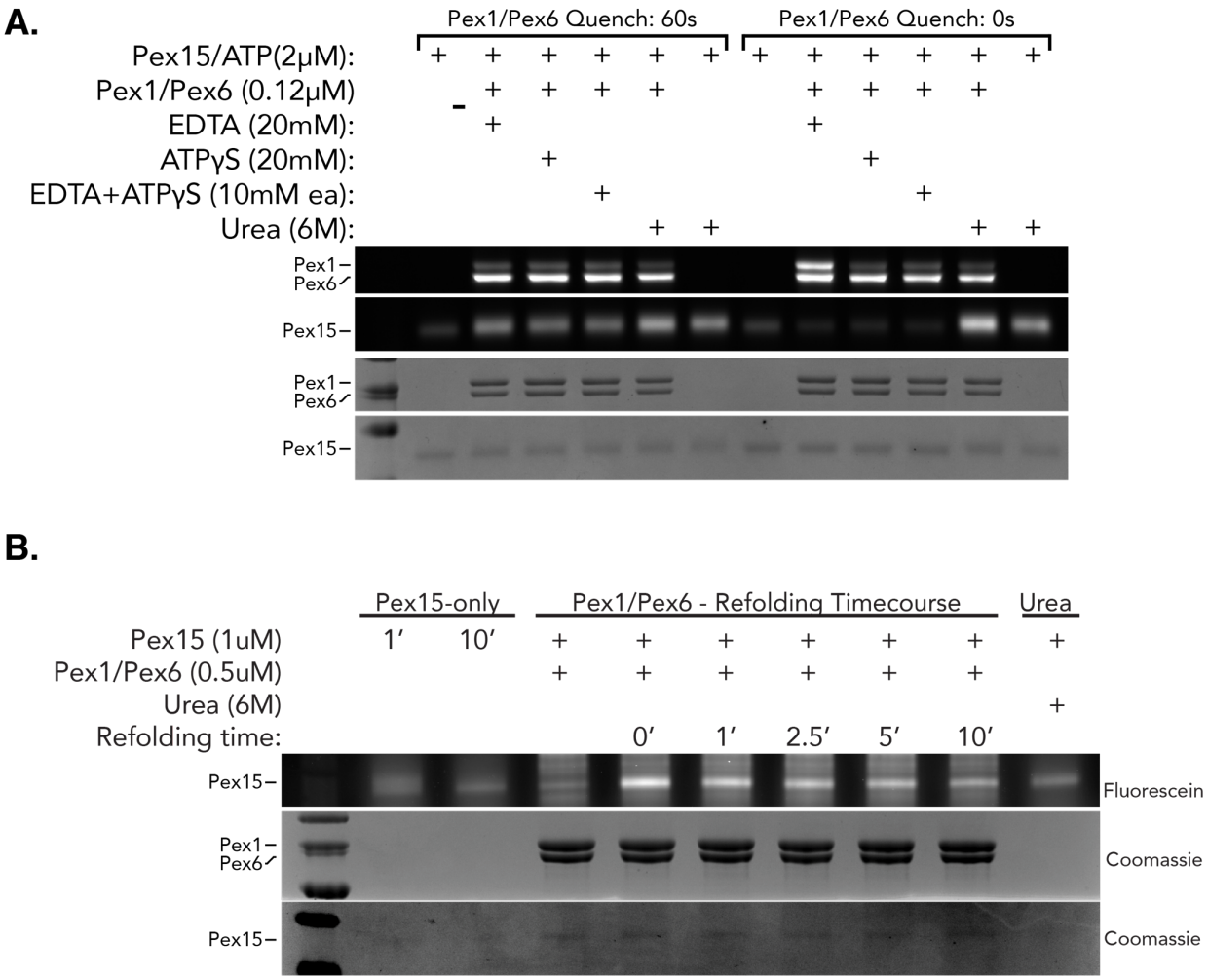


Fig. 3.7 : Toward measuring a refolding rate for Pex15: mechanical unfolding. (A) Pex1/Pex6-mediated unfolding of Pex15$^{1-309}$ can be quenched by pre-incubation with EDTA (20 mM), ATPγS (20 mM), or both (10 mM each). When pre-incubated with EDTA, Pex1 is readily labeled by F5M. We interpret this as dissociation of the Pex1/Pex6 complex, which subsequently exposes cysteines that were formerly at the Pex1/Pex6 binding interface and therefore shielded from F5M labeling. (B) Maleimide-labeling unfoldase assay wherein Pex1/Pex6 was allowed to unfold Pex15$^{1-309}$ before being quenched. Pex15 was then allowed time to refold and re-bury its cysteines before labeling with F5M. The Pex15-corresponding band in the fluorescein channel appears to be less-bright as the sample is given more time to refold; however, it is impossible to determine whether this reflects true refolding, unfolded protein aggregation, or the formation of disulfide bonds.

Before this could be done, I needed to identify conditions that would sufficiently deactivate Pex1/Pex6, and devised two ways to quench Pex1/Pex6 activity. The first method of quenching Pex1/Pex6 activity is through addition of EDTA, which chelates the Mg^{2+} that is essential for ATP hydrolysis. Maleimide labeling reactions contained 10 mM $MgCl_2$, and I therefore quenched the ATPase reaction through addition of 20 mM EDTA. The second method of quenching ATPase activity is through the addition of ATPγS that is non- or only slowly hydrolyzable. We knew from the HDX data that addition of ATPγS is sufficient to inhibit Pex1/Pex6-mediated unfolding of Pex15$^{1-309}$. Unfoldase assays typically included an ATP regeneration mix to maintain ATP levels, however, to ensure a rapid Pex1/Pex6 inhibition by ATPγS and avoid the conversion of ATPγS to ATP, I omitted the regeneration mix. Given Pex1/Pex6's hydrolysis rate of

9000 ATP per hexamer per minute, a motor concentration of 0.12 μM in the sample led to an ATP consumption of 1080 μM per minute. The 5 mM ATP present at the beginning of the reaction was thus well sufficient to support the ~ 1 min unfolding of Pex15$^{1-309}$, while allowing a rapid quenching through the addition of 20 mM ATPγS.

To test whether EDTA, ATPγS, or both could inhibit Pex1/Pex6 unfoldase activity, we used a maleimide labeling unfoldase assay. Preincubation of Pex1/Pex6 with EDTA (20 mM), ATPγS (20 mM), or the combination of both (10 mM of each) eliminated Pex15$^{1-309}$ unfolding, while addition of these inhibitors after a 1-minute unfolding period allowed for some unfolding of Pex15$^{1-309}$ (Fig. 3.7a). I elected to continue with EDTA since it was the cheapest and easiest to use in excess.

To evaluate whether Pex15$^{1-309}$ could refold following mechanical unfolding by Pex1/Pex6, I allowed Pex1/Pex6 to unfold Pex15$^{1-309}$ in ATP for 1 minute prior to the addition of EDTA. Following EDTA-mediated inhibition of Pex1/Pex6, Pex15$^{1-309}$ was allowed to refold for 10 minutes, during which samples were taken at various timepoints. Samples were incubated with F5M for a set duration and then quenched with reducing agent. Interestingly, Pex15$^{1-309}$ labeling after Pex1/Pex6 unfolding and inhibition of the motor declines over time (Fig. 3.7b), indicating that cysteines become increasingly inaccessible, either due to complete or partial refolding, or aggregation of unfolded Pex15$^{1-309}$. Based on these data, it was not possible to conclude whether Pex15 is indeed capable of properly refolding after mechanical unfolding by Pex1/Pex6.

3.2 Investigating Pex1/Pex6's substrate requirements

So far, we had used the maleimide unfoldase assay to learn more about the mechanism by which Pex1/Pex6 unfolds Pex15$^{1-309}$. To further our understanding of this

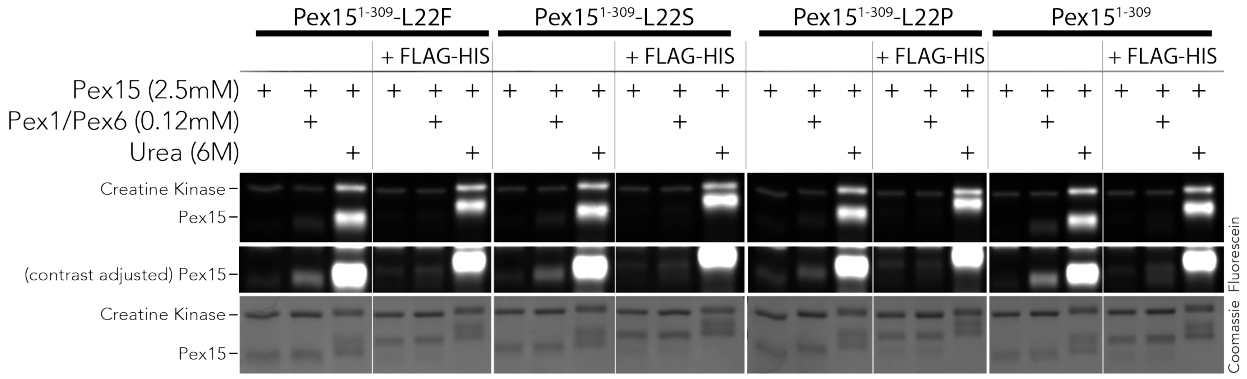


Fig. 3.8 : Pex1/Pex6 is sensitive to the C-terminal affinity purification tags on Pex15 substrates. Maleimide-labeling unfoldase assays demonstrating that various Pex15 constructs are unfolded more by Pex1/Pex6 when the C-terminal affinity purification tags are cleaved off using PSP.

interaction, we set out to define what makes an ideal substrate for Pex1/Pex6.

An *in vivo* study had previously shown that mutation to Pex15's L22 residue weakens the interaction with Pex1/Pex6 *in vivo* (Birschmann et al., 2003). I began by exchanging this residue for the amino acids they were mutated to in the aforementioned screen: L22F and L22S. As an additional parameter, I took note that this region

putatively forms an alpha-helix; to see if the helical nature of this region is paramount to its interaction with Pex1/Pex6, I also generated an L22P mutant. The L22F mutant was unfolded to the same extent as the native sequence protein, but the L22S mutant showed a slight defect in unfolding, and the L22P mutant must unfolded to even lower extent (Fig. 3.8). The most surprising result from this assay originated from comparing the unfolding of $Pex15^{1-309}$ with and without its C-terminal purification tags cleaved off. In the wild-type protein and all three mutants, Pex15 without its affinity purification tags ($Pex15^{1-309}$) was unfolded far more than Pex15 retaining the purification tags ($Pex15^{1-309}$-FLAG-His; Fig. 3.8). It is curious that making the construct smaller would make it a better substrate for Pex1/Pex6, so we concluded that the high charge density of consecutive 6xHis and FLAG epitopes presented a challenge for Pex1/Pex6 engagement or processing. We had previously not observed unfolding of $Pex15^{1-327}$-FLAG-His in maleimide labelings (despite this construct being the most potent inhibitor of Pex1/Pex6) and wondered if removing the affinity tags would make it a better substrate for Pex1/Pex6. Unfortunately, removal of these affinity tags causes the protein to precipitate (data not shown). Perhaps these tags were hitherto performing an unappreciated task in helping to solubilize this construct. Unless otherwise stated, all unfolding reactions hereafter were performed with tagless substrates.

Since no other mutants of Pex15 had been identified in the literature, we set out to characterize the importance of the rest of Pex15's unstructured N-terminal region by performing a series of truncations and testing Pex1/Pex6's ability to unfold these various constructs: $Pex15^{1-309}$, $Pex15^{12-309}$, $Pex15^{30-309}$, $Pex15^{43-309}$, and $Pex15^{57-309}$. The first three mutants had progressively longer truncations of Pex15's N-terminal unstructured region until it is completely gone ($Pex15^{43-309}$). Maleimide labeling unfoldase assays revealed that truncation of the first 11 residues has little effect on Pex1/Pex6's ability to recognize and unfold Pex15 (Fig. 3.9a). Interestingly, both $Pex15^{30-309}$ and $Pex15^{43-309}$ are unfolded similarly, but to a lesser extent than the full-length construct (Fig. 3.9a). From this we conclude that while Pex1/Pex6 does not require an N-terminal unstructured region to engage and unfold Pex15, the presence of this region facilitates unfolding. Whether it is the specific amino acid sequence from residues 12-29 that makes the difference or simply the length of this region that increases unfolding efficiency by Pex1/Pex6 was unclear. To distinguish between these two hypotheses, I mutated two highly-conserved residues (L22A and L23A) to help define the importance of sequence and length of Pex15's N-terminal unstructured region. As previously discussed, one of these residues, L22, was identified in a mutagenesis screen as a contributor to Pex15's interaction with Pex1/Pex6. Interestingly, mutation of these residues reduced Pex1/Pex6's unfoldase activity to the same extent as truncating the entire N-terminal domain ($Pex15^{43-309}$), and we concluded that these residues in the N-terminal unstructured domain are not essential but contribute to Pex15 recognition by Pex1/Pex6. The last mutant I evaluated was $Pex15^{53-309}$, which lacked the first alpha helix of Pex15's structured core domain. This first alpha helix contains the core's first cysteine, which is situated near the second cysteine (residing on the second helix). As a result, truncation of this helix exposes the second cysteine to the solvent and leads to increased F5M labeling of even the folded protein. However, the remaining five buried cysteines present enough of a dynamic range to distinguish folded from unfolded Pex15. Surprisingly, truncation of this first helix (along

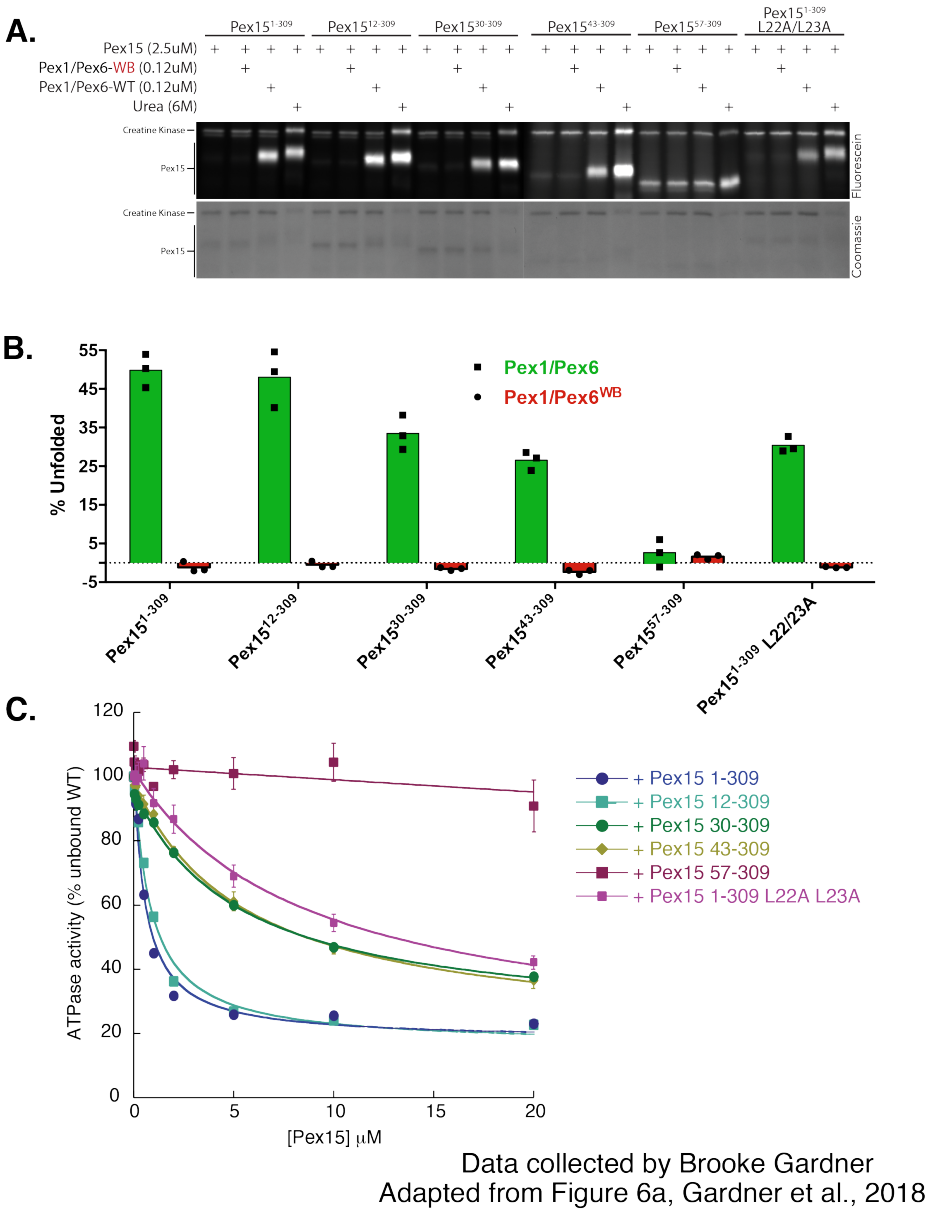


Fig. 3.9 : Truncations throughout Pex15's N-terminal unstructured region weaken ATPase inhibition, Pex15 unfoldability. (A) Maleimide-labeling unfoldase assays of progressively-aggressive truncations throughout Pex15's N-terminal unstructured region. While Pex15$^{1\text{-}309}$ and Pex15$^{12\text{-}309}$ are unfolded similarly, unfoldability diminishes for Pex15$^{30\text{-}309}$ and Pex15$^{43\text{-}309}$, the latter of which lacks any unstructured region altogether. Mutation of two highly-conserved leucines at positions 22 and 23 also reduces unfoldability. Surprisingly, Pex15$^{57\text{-}309}$, which is truncated through to Pex15's first core-domain alpha helix, is not unfolded at all, which may suggest this region is important for the Pex15:Pex1/Pex6 interaction. (B) Quantification of maleimide-labeling unfoldase assays depicted in (A) (n = 3, technical replicates). (C) ATPase assays showing inhibition of Pex1/Pex6 ATPase activity. Inhibition by Pex15 appears to correlate to substrate unfoldability by Pex1/Pex6.

with the rest of the unstructured N-terminal domain) completely abrogates Pex15 unfoldase by Pex1/Pex6 (Fig. 3.9a), and we conclude that residues in the unstructured region (L22, L23) and in the first helix of the structured core form the basis of Pex15's interaction with Pex1/Pex6.

All maleimide labeling reactions were analyzed by SDS-PAGE, and I used ImageQuant or the BioRad ChemiDoc imaging software (Image Lab) to accurately

quantify the band intensity and extent of Pex15 unfolding. For each set of experiments, I set the basal labeling intensity (Pex15 without urea or enzyme) as 0% unfolded, and the urea sample as 100% unfolded (after back-calculating for any sample dilution prior to loading the gel). I then subtracted the fluorescence intensity of the basal labeling reaction from the Pex15 intensity observed in the presence of Pex1/Pex6, and calculated the extent of unfolding as a percent of the maximum unfolding seen in the urea sample. All experiments were performed in triplicates. Plotting the labeling efficiency clearly demonstrates how $Pex15^{30-309}$ and $Pex15^{43-309}$ experience markedly less unfolding than the two longer constructs, and that $Pex15^{L22A/L23A}$ is also unfolded to a lesser extent (Fig. 3.9b). Interestingly, Pex15 in the presence of the ATPase-deficient Walker-B mutant $Pex1^{WB}/Pex6^{WB}$ appears to be F5M labeled to a lower extent than in the absence of any motor (see negative unfoldase values in Fig. 3.9b), suggesting that binding to inactive Pex1/Pex6 may stabilize Pex15 or further shield its largely buried cysteines.

These data prompted us to consider whether there is a relationship between substrate unfolding and substrate-mediated reduction of Pex1/Pex6 ATPase activity. Brooke tested each of these N-terminal truncations in ATPase assays and indeed found a clear correlation: $Pex15^{1-309}$ and $Pex15^{12-309}$ reduced the ATP hydrolysis rate of Pex1/Pex6 with a similar K_i that was lower than the ones observed for $Pex15^{30-309}$ and $Pex15^{43-309}$ (Fig. 3.9c). This result is expected if the N-terminal unstructured region is important for Pex15 binding to Pex1/Pex6 and if substrate processing decelerates ATP hydrolysis.

Brooke also collaborated with Dr. Saikat Chowdhury at The Scripps Research Institute to perform negative stain EM on Pex1/Pex6 and $Pex15^{1-327}$ in hopes of observing a bound complex. By comparing the obtained density to a previously-published negative stain EM model of Pex1/Pex6 alone, an extra blob of density was observed near the N-domains of every other subunit in the heterohexamer (Fig. 3.10a, gray). This extra density showed a roughly curved morphology, resembling the structure of $Pex15^{43-253}$. Indeed, the crystal structure of Pex15 could be fit unambiguously into this density (Fig. 3.10b, blue spectrum ribbon), in an orientation that puts the first, most N-terminal alpha helix of Pex15's structured core very close to the ATPase domain of Pex6. These structural data agree with our biochemistry results in suggesting that the N-terminal unstructured region and first helix of Pex15's core mediate Pex1/Pex6 binding.

During these efforts, Brooke was also working on using HDX to map Pex15's binding site on Pex1/Pex6. Her data identified a short stretch of amino acids between Pex6's N1 and N2 domains (amino acids 241-249) that experienced slower HDX in the presence of Pex15, suggesting it may be at the interface of the interaction. While neither a crystal structure nor a high-resolution cryo-EM structure exist for Pex1/Pex6, other labs have produced a structural model based off homology to other ATPases and their own understanding of protein folding and dynamics. When we fit this model into our negative stain EM density obtained in the presence of Pex15, we found that the $Pex6^{241-249}$ motif resides near the interface of Pex6 and the docked Pex15 core domain (Fig. 3.10b, magenta). In the structural model of Pex1/Pex6, these nine amino acids form a loop between two alpha-helices.

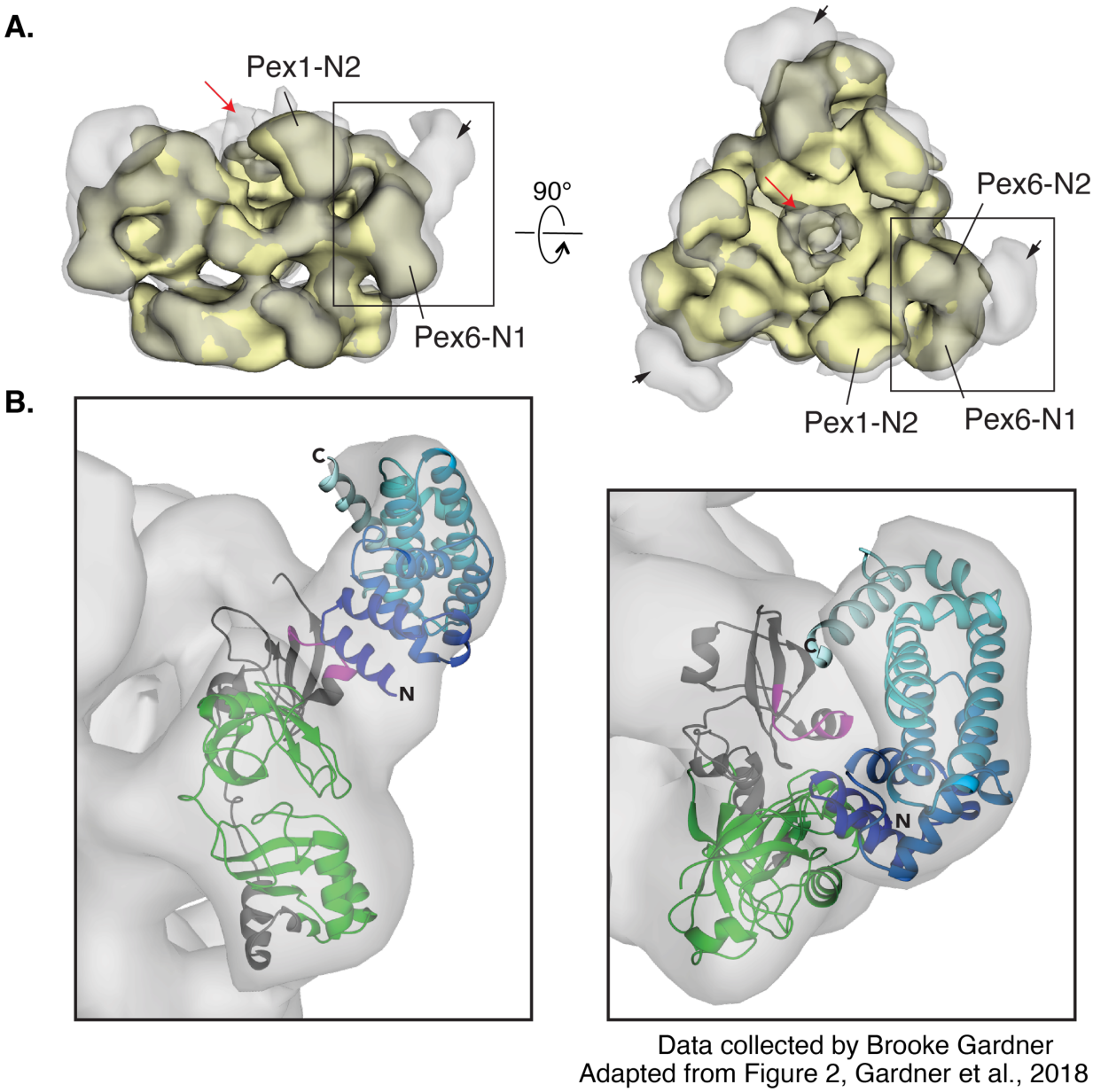


Fig. 3.10 : Negative-stain EM provides insight into Pex15 and Pex1/Pex6's interaction. (A) Negative-stain EM density of Pex1/Pex6 alone (yellow) superimposed with Pex1/Pex6 with Pex15 (grey), showing additional density both in the central pore and as an extension off of what we believe to be the Pex6 N1 domain. Pex15's curved shape allows for unambiguous docking of the Pex15 crystal structure into the differential density. (B) This docking positions Pex15's N-terminus, specifically its first structural alpha-helix, at the Pex15:Pex1/Pex6 binding interface, in agreement with the data presented in Fig. 3.11 describing the defects of omitting this region. Furthermore, this orientation implicates a small region (purple) of Pex6's N1/N2 domains (green) in interacting with Pex15 (blue). The Pex1/Pex6 structure is not definitively known, and these models reflect structural predictions based on homology to other ATPases.

I consequently exchanged the Pex6 residues 241-249 for serine or glycine (LRLDVTSGS → GSSGSSGSS) to biochemically determine whether this region mediates interaction with Pex15, but this mutant demonstrated poor solubility and did not yield enough protein. I therefore placed only two pairs of mutations, R242A/D244A and V245A/T246A, in the Pex6 of double-Walker-B mutant $Pex1^{WB}/Pex6^{WB}$ and

assessed the contributions of these residues to MBP-Pex15$^{1-309}$ binding in amylose-pulldown experiments (data not shown). The data did not suggest any significant differences in affinity (data not shown), however, pulldowns are not a very sensitive assay for more subtle changes in binding. I therefore performed ATPase inhibition assays comparing Pex15$^{1-327}$'s ability to inhibit R242A/D244A or V245A/T246A mutant Pex1/Pex6 in comparison to the WT motor. Interestingly, despite having a similar basal rate of ATPase activity, Pex1/Pex6$^{VT:AA}$ appeared less sensitive to inhibition by Pex15$^{1-327}$, while Pex1/Pex6$^{RD:AA}$'s sensitivity to Pex15's inhibition was similar to that of the WT motor's (Fig. 3.11). Through all these experiments, we have learned more about how Pex15 may recruit Pex1/Pex6 to the peroxisomal membrane *in vivo*.

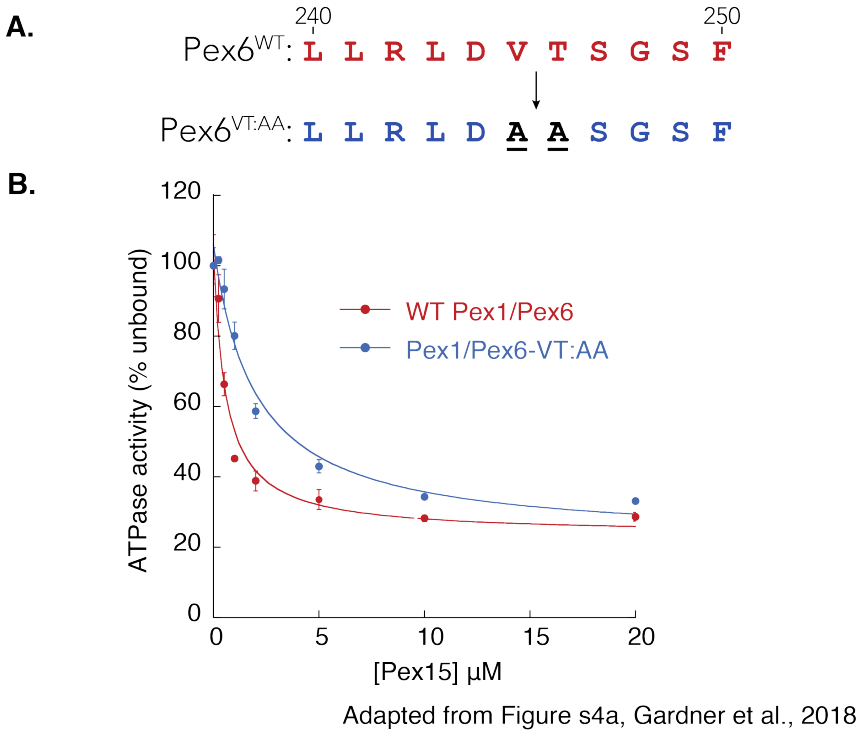


Fig. 3.11 : Mutation of Pex6 residues at suspected Pex6-Pex15 interface weakens ATPase inhibition. (A) Wild-type (red) and mutant (blue) sequences of Pex6 region inferred as the Pex15 binding site by negative-stain EM in Fig. 3.12b (purple), wherein two consecutive residues at positions 245 and 246, valine and threonine, have been mutated to alanines. (B) ATPase assays show that Pex15-mediated inhibition of this construct is diminished in the mutant Pex1/Pex6 construct, compared to the wild-type motor. These data implicate this region as part of the binding interface between Pex1/Pex6 and Pex15.

While it was clear from previous experiments that Pex15's N-terminus plays a crucial role in recruiting Pex1/Pex6 to the membrane, we still found it curious that the presence or absence of the His and FLAG affinity tags at the C-terminus appeared to also affect Pex15's unfoldability by Pex1/Pex6. From the perspective of Pex1/Pex6, substrate binding and substrate engagement can be two separate events—binding is the initial interaction with the complex, whereas engagement is the stable contact with the motor's pore loops for subsequent mechanical pulling and translocation. We have established that the initial interaction with the complex occurs between Pex15's N-terminal region and Pex6's N1 and N2 domains, but where Pex1/Pex6 engages Pex15 remained unclear. Since the C-terminus of Pex15 appeared to dictate unfoldability, we hypothesized that Pex1/Pex6 might engage the isolated Pex15 from its free C-terminus. To test this hypothesis, I cloned Pex15$^{1-309}$-MBP, where the free C-terminus was blocked by a folded MBP domain. This construct was not unfoldable by Pex1/Pex6, suggesting that Pex1/Pex6 indeed engages Pex15$^{1-309}$ from the C-terminus (Fig. 3.12). However, the engagement of Pex15's free C-terminus would not be a possible scenario in vivo, where Pex15 is anchored in the peroxisomal membrane. Thus, although we have extensively proven

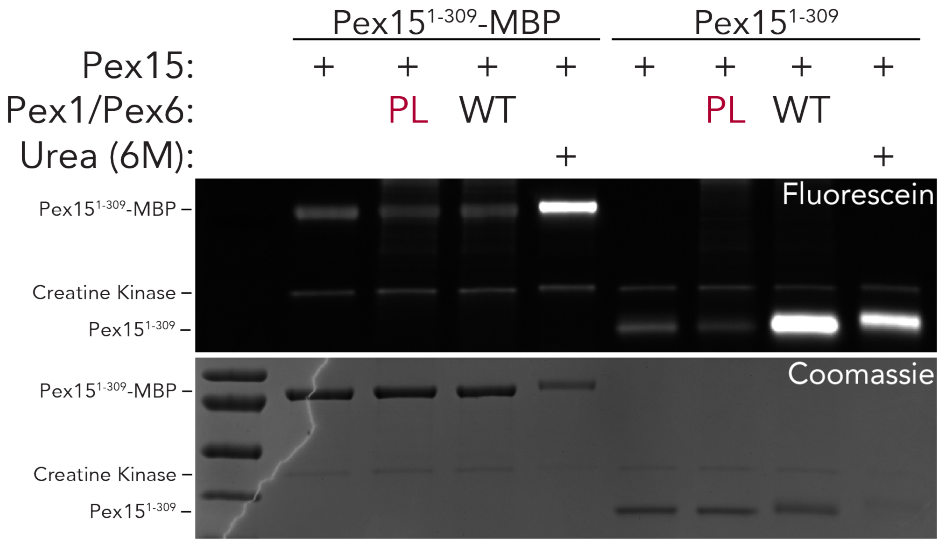


Fig. 3.12 : Pex1/Pex6 unfolds soluble Pex15 from its free C-terminus. Maleimide-labeling unfoldase assay demonstrating that Pex1/Pex6 is able to unfold Pex15, but not Pex15-MBP. These data suggest Pex1/Pex6 needs a free C-terminus in order to engage $Pex15^{1-309}$.

Pex1/Pex6's unfoldase capabilities, and it is conceivable that the motor engages the C-terminal unstructured region of Pex15 as a loop, the identity of an endogenous, bona fide substrate still remains unclear. Nevertheless, we continued using Pex15 as a model substrate combined with the maleimide labeling unfoldase assay to learn more about the Pex1/Pex6 motor.

First, I investigated the basic requirement for Pex1/Pex6 substrates, the length of an unstructured region. It is clear from previous experiments that Pex1/Pex6 unfolds Pex15 from the C-terminus of its ~56 amino acid unstructured tail (62 amino acids when including the PSP cleavage scar), whereas unfolding did not occur from the N-terminus. We hypothesized that this was because the N-terminus binds the Pex6 N1/N2 domains, allowing the C-terminus to reach into the central pore and contact the pore loops. It remained unclear how long this tail needs to be and whether specific sequences therein make it a substrate for Pex1/Pex6. I therefore made truncations throughout Pex15's C-terminal tail to elucidate these length and sequence requirements for Pex1/Pex6 substrates. I compared Pex1/Pex6's ability to unfold $Pex15^{1-253}$, $Pex15^{1-266}$, $Pex15^{1-277}$, $Pex15^{1-288}$, $Pex15^{1-299}$, and $Pex15^{1-309}$, and found that Pex1/Pex6 was only able to unfold the two longest constructs (Fig. 3.13a, b). To understand if this difference in unfolding is due to specific sequences present between amino acids 289 and 299 or simply due to the difference in length, I compared Pex1/Pex6's ability to unfold $Pex15^{1-299}$ with $Pex15^{1-288+GS}$ and $Pex15^{1-288\text{^}GS}$. The latter two constructs have the same length unstructured C-terminal tail as $Pex15^{1-299}$, but contain a 11-amino acid glycine-serine stretch either at the end (+GS) or as an insertion in the unstructured region (^GS). Addition or insertion of a glycine-serine linker allowed for some unfolding by Pex1/Pex6, but not as much as for $Pex15^{1-299}$ (Fig. 3.13c). These data thus implicate tail length as the predominant parameter in determining a substrate's unfoldability; however, tail sequence certainly plays a role. This is not surprising, given that many AAA+ ATPases appear to struggle with glycine-serine rich, "slippery" sequences due to lack of more

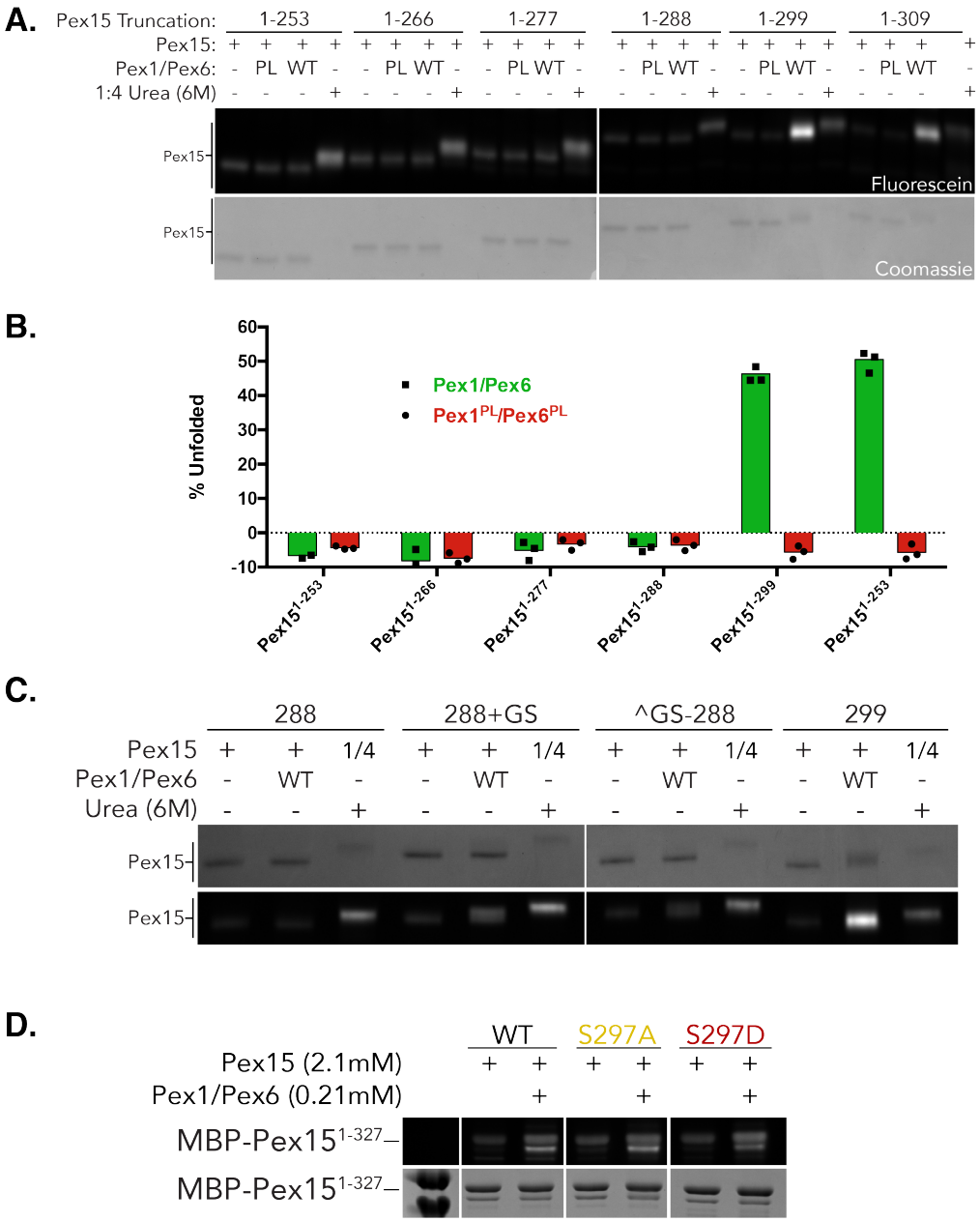


Fig. 3.13 : Truncations, mutations throughout Pex15's unstructured C-terminal region reveal a minimum length requirement for engagement. (A) Maleimide-labeling unfoldase assays of progressively-aggressive truncations throughout Pex15's C-terminal unstructured region demonstrate that Pex15 is not unfoldable when the C-terminal unstructured region is shorter that 45 amino acids, as is quantified in (B). (C) Extension of $Pex15^{1-288}$'s unstructured region using a glycine/serine sequence (288+GS) or insertion of this sequence 11aa from the natural end of $Pex15^{1-288}$ (^GS-288) recover some unfoldase activity by Pex1/Pex6. It is possible that both length and sequence contribute to Pex1/Pex6 substrate unfoldability. (D) Phosphomimetic mutation of Pex15's S297, which was thought to be phosphorylated in vivo, does not alter unfoldability by Pex1/Pex6.

bulky side chains and the consequently reduced grip of pore loops during mechanical pulling.

It is possible that Pex1/Pex6's interaction with Pex15 *in vivo* is regulated by a post-translational modification of the latter. A screen of phosphorylated proteins in *S. cerevisiae* found that Pex15 can be phosphorylated on a single residue: S297. While it

is unclear from this screen whether this phosphorylation event is related to Pex15's role in peroxisome biogenesis or rather part of an unrelated role in the cell, it is plausible that phosphorylation affects Pex15's affinity for Pex1/Pex6. To test this, I made two mutations in MBP-Pex15$^{1-327}$: replacing S297 with an aspartate, which is a phosphomimetic residue approximately reproducing the addition of a negatively-charged group to a serine, and an S297A mutation, which represents a non-phosphorylatable residue. Both variants purified similarly to the non-mutated protein, and both were unfolded to the same extent by Pex1/Pex6 (Fig. 3.13d). A phosphomimetic mutation may not fully mimic a phosphorylated serine; however, we still believe the best conclusion from these data is that phosphorylation at S297 neither weakens nor enhances Pex15's interaction with Pex1/Pex6. Furthermore, if Pex15 is unfolded by Pex1/Pex6 *in vivo*, S297 phosphorylation appears to neither inhibits nor enhance unfolding, and may therefore play a role in another cellular process.

3.3 Temperature dependence of Pex1/Pex6 pore-loop mutants

In 2015, a study published from Dr. Petra Wendler's lab suggested that Pex1/Pex6 variants with pore loops mutated in either Pex1 (Pex1^{PL}/Pex6) or Pex6 (Pex1/Pex6^{PL}) cause growth defect in yeast only when cells are grown on oleic acid. Oleic acid growth assays are the classic test for functioning peroxisomes, since *S. cerevisiae* depend of peroxisomal metabolism to utilize oleic acid as a carbon source. These results appeared to be in direct contradiction to my earlier experiments, which indicated that Pex1/Pex6^{PL} possessed no unfoldase activity. Furthermore, while Pex1^{PL}/Pex6 was still minimally active, it is surprising that a strain containing this strongly defective mutant is viable. Given that Dr. Wendler's lab performed *in vivo* experiments at 30°C (whereas our unfoldase assays occurred at 24°C), we hypothesized that the pore loop mutations in Pex1/Pex6 could cause a temperature-sensitive phenotype.

We therefore compared the ability of Pex1^{PL}/Pex6, Pex1/Pex6^{PL}, and Pex1^{PL}/Pex6^{PL} complexes to unfold Pex15 at 24°C and 30°C, using a maleimide labeling unfoldase assay. At both temperatures, the double mutant Pex1^{PL}/Pex6^{PL} was entirely unable to unfold Pex15. As before, Pex1/Pex6 ^{PL} was completely unable to unfold substrates at 24°C, and the Pex1^{PL}/Pex6 mutant retained a low level of unfoldase capability at this temperature. Substrate unfolding by both single mutants increased at 30°C (Fig. 3.14a), likely due to the temperature-correlated lower stability of Pex15 and/or the higher ATP-hydrolysis rate of the motor. To test whether the stability of Pex15$^{1-309}$ differs substantially for 24°C versus 30°C, I again utilized CD spectroscopy and urea-induced equilibrium unfolding. The midpoint of the Pex15 unfolding transition decreased from Cm ~ 3.6 M at 24°C to Cm ~ 2.6 M at 30°C (Fig. 3.14b), suggesting that Pex1/Pex6's ability to unfold Pex15, and therefore other substrates in the cell as well, may rely on the thermodynamic stability of those proteins, and perhaps exists in a delicate balance at physiologically optimal temperatures.

While it makes sense that a substrate's thermodynamic stability and its rate of unfolding by a AAA+ ATPase are inversely correlated, we could not rule out that

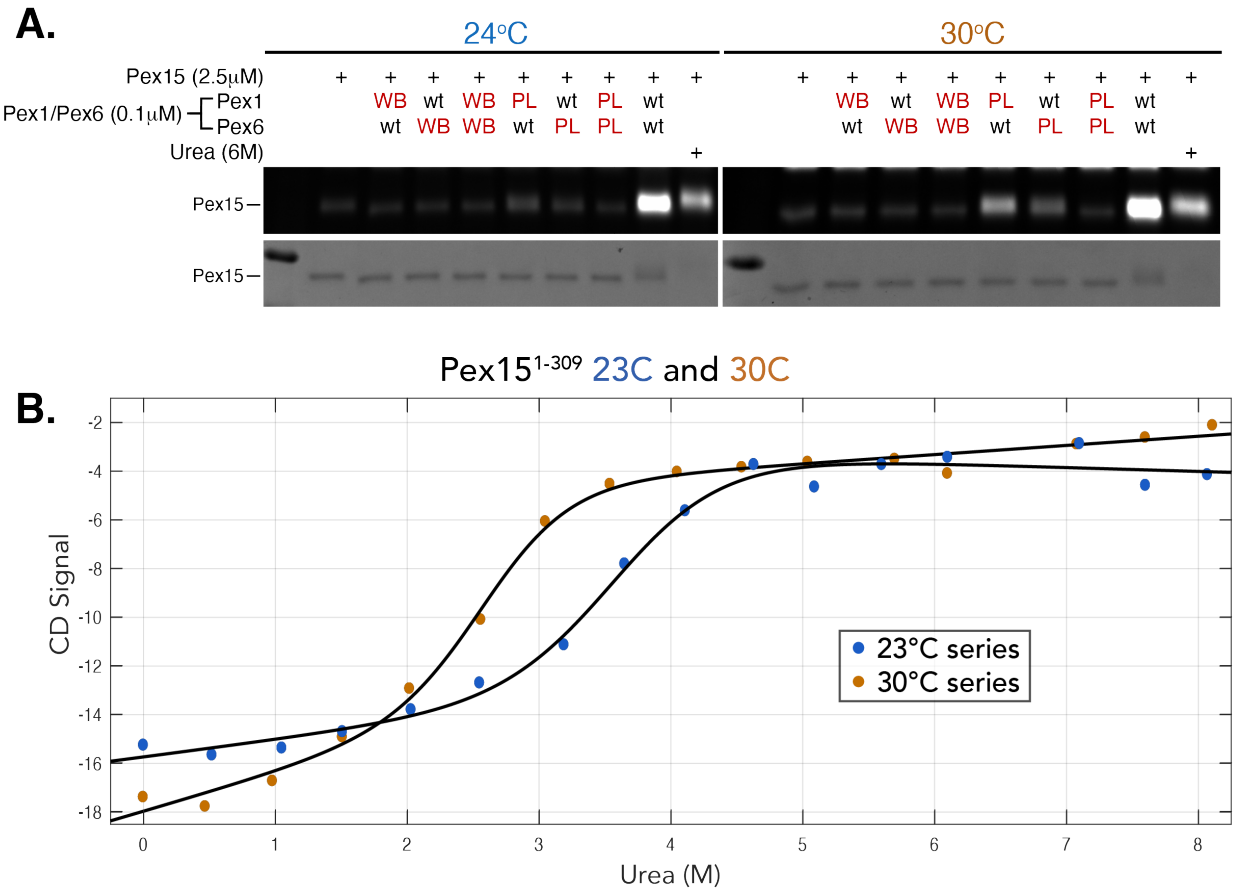


Fig. 3.14 : Pex1^{PL}/Pex6 exhibits temperature-dependent unfolding of Pex15. (A) Maleimide-labeling unfoldase assays of Pex15 using various single- and double-walker b and -pore loop mutants of Pex1/Pex6 at both RT (24°C) and 30°C. Pex1^{PL}/Pex6 appears more able to unfold Pex15 at 30°C than at RT, a temperature-depending change not observed in another other single or double mutant. (B) CD spectroscopy of Pex15 in increasing concentrations of urea at both RT and 30°C, showing a lower C_m at 30°C (~ 2.75 M) as compared to RT (~ 3.5 M). CD signal reflects Θ_{222nm} at each point.

differences in Pex1/Pex6^{PL}'s ATPase rates determined that a substrate like Pex15 could get unfolded at 30°C, but not at 24°C. To distinguish between these two models, I attempted to analyze Pex1/Pex6^{PL}'s ability to unfold a structurally-weakened Pex15 mutant. Such a destabilized Pex15 mutant may allow unfolding by Pex1/Pex6^{PL} and rescue *in vivo* growth defects at 24°C. Unsurprisingly, there were no documented Pex15 mutations explicitly categorized as "destabilizing," so I first set out to identify destabilizing mutations.

Pex15's hydrophobic core contains a series of phenylalanines projecting into the center of the folded domain (Fig. 3.15a). I hypothesized that mutation of one or more of these residues to alanines would destabilize Pex15 by creating a cavity and disrupting the hydrophobic packing. I identified four phenylalanines evenly distributed throughout Pex15 to mutate to alanine: F86, F111, F139, and F194 (Fig. 3.1bb). All four mutants expressed and purified normally, although F139A eluted off the SEC column earlier than expected with an uncharacteristically large void peak, a possible indication of aggregation.

First, I sought to establish that these mutants were indeed destabilized. Rather than performing urea-induced unfolding transitions measured by CD spectroscopy, I

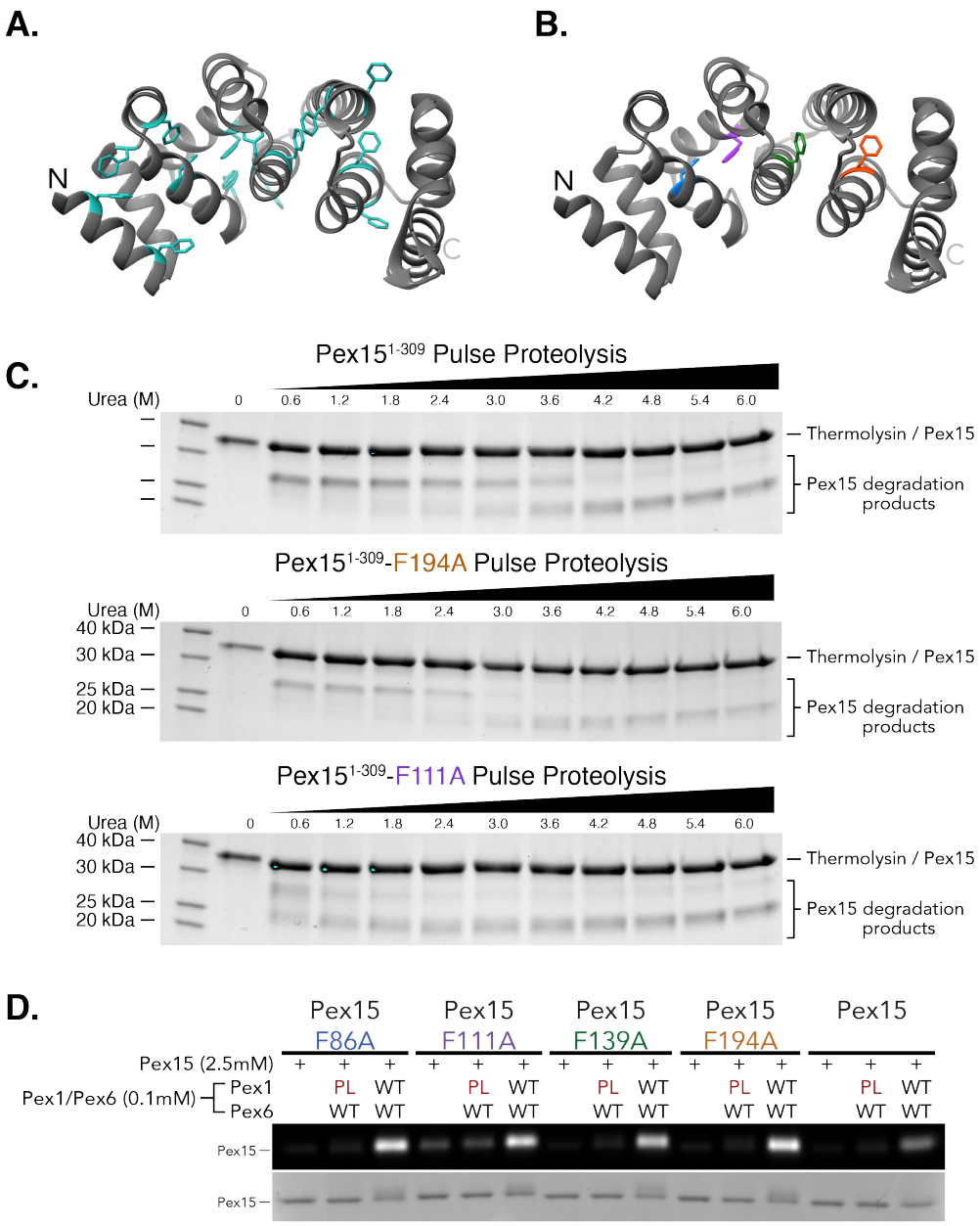


Fig. 3.15 : Mutation of internal phenylalanines destabilizes the Pex15 core domain. (A) Pex15's phenylalanines (cyan) extend into the hydrophobic core of the protein throughout the structure. (B) Four evenly-distributed phenylalanines identified for mutations in an attempt to destabilize Pex15: F86, F111, F139, and F194. (C) Pulse proteolysis on Pex15, Pex15^{F194A}, and Pex15^{F111A} suggest both mutants are destabilized, as compared to the wild-type sequence. Full-length Pex15$^{1-309}$ is the same size as thermolysin, but the appearance and disappearance of a slightly smaller band, likely Pex15's core domain with part of its naturally-unstructured region cleaved off, serves as a reference for protein stability. (D) Maleimide-labeling reactions do not indicated that any of the mutants are unfolded more than the wild-type sequence, for either wild-type Pex1/Pex6 or the pore loop mutant.

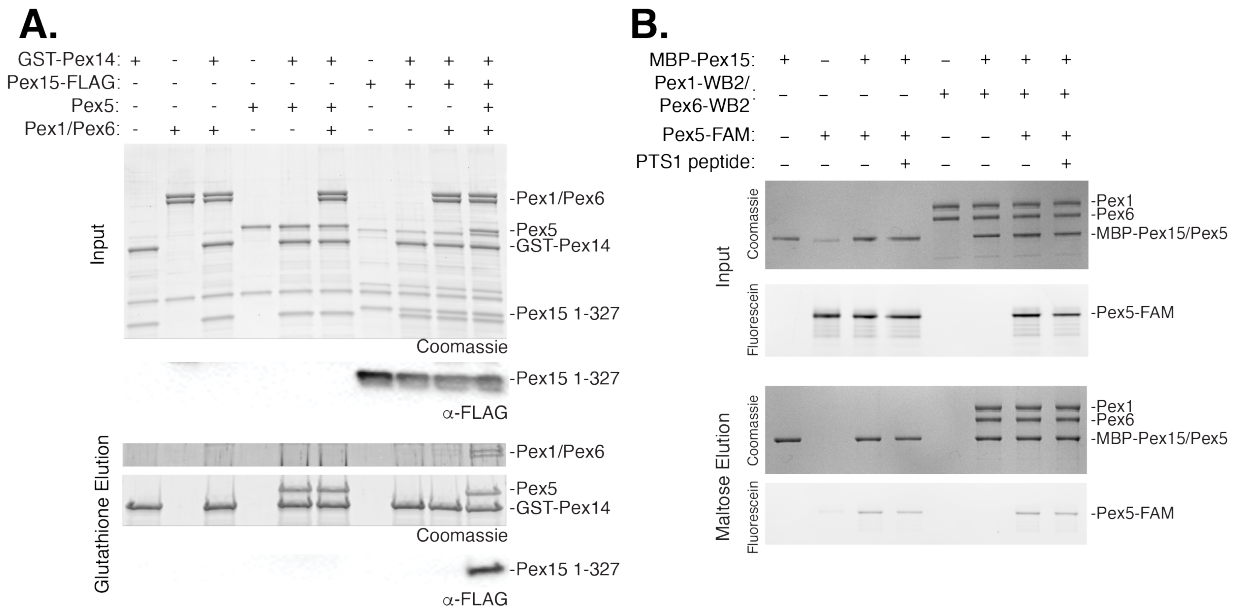
utilized the technique of pulse proteolysis. Similar to limited proteolysis, pulse proteolysis involves exposing a protein of interest to a promiscuous protease in different concentrations of urea. As the protein becomes more and more loosened and destabilized under increasingly denaturing conditions, it is increasingly proteolyzed to smaller fragments by the protease. Comparing the appearance of lower molecular-weight bands on the gel among different mutants is an effective way to assess relative

stability. I compared Pex15$^{1-309}$, Pex15$^{1-309_F111A}$, and Pex15$^{1-309_F194A}$ using pulse proteolysis with thermolysin and observed the appearance of proteolysis products significantly earlier in the F111A mutant, suggesting a more substantial destabilization, followed by the F194A mutant (Fig. 3.15c,d).

Having established that at least some of the mutations affected the global stability of Pex15, I proceeded to compare all mutants in maleimide labeling unfoldase assays using both wild-type Pex1/Pex6 and Pex1^{PL}/Pex6. Wild-type Pex1/Pex6 unfolded all destabilized Pex15 mutants to a greater extent than the unmutated Pex15$^{1-309}$; however, Pex1^{PL}/Pex6 showed no difference in unfolding for these Pex15 variants at 24°C (Fig. 3.15d). Pex15^{F111A} is clearly destabilized, as also indicated by a higher level of basal maleimide labeling in the absence of Pex1/Pex6 that suggests a more frequent sampling of partially unfolded states. Assuming that this Pex15 mutant was simply not destabilized enough to allow unfolding by Pex1^{PL}/Pex6, I tested unfolding of a double mutant: F86A/F194A. This mutant also did not exhibit greater unfolding by Pex1^{PL}/Pex6 at 24°C (data not shown). Since the Pex1^{PL}/Pex6 mutant cannot unfold Pex15^{F111A}, even though pulse-proteolysis and the maleimide labeling data clearly indicate the reduced stability of this substrate variants, we conclude that thermodynamic stability of Pex15 is not the predominant factor responsible for the temperature sensitivity observed for Pex1^{PL}/Pex6-containing cells. Given our uncertainty about whether Pex15 is Pex1/Pex6's endogenous substrate, however, it remains to be seen how pore-loop mutations affect substrate processing in vivo, e.g. the potential extraction of Pex5 from the peroxisomal membrane that may not require mechanical unfolding of a globular domain.

3.4 Reevaluating Pex15's role in peroxisome biogenesis

Prior to our investigations, Pex15's only reported role in *S. cerevisiae* was in recruiting Pex1/Pex6 to the peroxisome membrane. While providing evidence that Pex1/Pex6 could unfold Pex15 *in vivo* remains difficult, we began to explore whether Pex15 performed other roles in the cell. In human cells, Pex26 is reported to interact with Pex14, a member of the docking complex that initially binds the cargo-charged Pex5 at the peroxisomal membrane. It is possible that Pex15 plays a similar role in yeast, linking Pex1/Pex6 to the docking complex, although this interaction has not yet been explicitly reported in *S. cerevisiae*. Using a GST-tagged, TMD-less Pex14 construct, Brooke performed coimmunoprecipitations and observed that Pex1/Pex6 inetracts with GST-Pex14 only when both Pex15 and Pex5 were present (Fig. 3.16a). It is known that Pex5 interacts with Pex14, and Pex15 interacts with Pex1/Pex6. While a direct interaction between Pex5 and Pex1/Pex6 has long been postulated, direct evidence substantiating this model is lacking. Furthermore, direct interactions between Pex14 and Pex1/Pex6 or Pex5 and Pex15 have been neither hypothesized nor demonstrated. Thus, Brooke performed a subsequent coimmunoprecipitation demonstrating Pex5's ability to interact with MBP-Pex15$^{1-327}$ (Fig. 3.16b). In these immunoprecipitations, MBP-Pex15$^{1-327}$ is shown to be capable of binding both



Data collected by Brooke Gardner
Adapted from Figure 7, Gardner et al., 2018

Fig. 3.16 : Pex15 mediates interaction between Pex5/Pex14 and Pex1/Pex6. (A) GST-Pex14 pulldowns to determine direct or indirect interaction with Pex15, Pex5, and Pex1/Pex6. Pex5 appears to pull-down with Pex14 alone, implicating a direct interaction. Neither Pex15 nor Pex1/Pex6 interact with Pex14 directly, but both pull down with Pex14 in the presence of Pex5. (B) MBP-Pex15 pulldowns to determine direct or indirect interaction with $Pex1^{WB}$/$Pex6^{WB}$ mutants, Pex5, and the PTS1 peptide. Pex5 pulls down with Pex15 equally-well in both the absence and presence of the PTS1 peptide. Predictably, Pex1/Pex6 also pulls down with Pex15, and Pex5 both with and without the PTS1 peptide can also bind this complex.

Pex1/Pex6 and Pex5 simultaneously. These data suggest that Pex15 functions to link Pex1/Pex6 to the Pex5-bound docking complex on the peroxisomal membrane.

We were unable to provide evidence that Pex1/Pex6 is capable of unfolding a Pex15 construct with a blocked C-terminus, which is our best *in vitro* approximation of what full-length Pex15 anchored in a membrane would look like to Pex1/Pex6. Nonetheless, using Pex15 as a model substrate allowed us to produce strong evidence that Pex1/Pex6 is a capable unfoldase and able to unfold substrates via processive threading through its central pore. Furthermore, we used a variety of structural and biochemical techniques to identify Pex15's N-terminal unstructured region and the first helix of its globular core as key components in the interaction with Pex1/Pex6. Conversely, HDX and point mutagenesis have shed light on the nature of Pex1/Pex6's interaction with Pex15 through a disordered loop between Pex6's N1 and N2 domains. Our attempts to solve the crystal structure of $Pex6^{N1}$ yielded crystals that diffracted with high resolution, but could not be phased. Additionally, F→A point mutants in Pex15's core demonstrated that creating a cavity in Pex15's hydrophobic core could destabilize the folded domain and accelerate unfolding by wild-type Pex1/Pex6, yet have no effect on unfolding by the $Pex1^{PL}$/Pex6 mutant. It still remains unclear why $Pex1^{PL}$/Pex6 appears to be better unfoldase than Pex1/$Pex6^{PL}$. Finally, we demonstrated that Pex15's interaction with Pex1/Pex6, as well as its newly-documented interaction with Pex5, facilitates colocalization of Pex1/Pex6 with the docking complex. Since our study was published (Gardner et al., 2018), a new report demonstrated that Pex1/Pex6 can unfold monoubiquitinated Pex5 (Pedrosa et al., 2018). The authors demonstrated this

utilizing a modified form of our maleimide-labeling assay. Pex5 possesses several buried cysteines throughout its folded C-terminal TPR motifs. We had previously attempted to perform maleimide-labeling unfoldase assays with Pex5, but observed only a small change in attached F5M fluorescence between folded Pex5 and unfolded Pex5 (in 6M urea). Such a small dynamic range would make it difficult to draw conclusions about whether the domain is folded or not. Pedrosa et al., however, smartly used maleimide-PEG5000 to label Ubn-Pex5's cysteines. As a result, labeled Ubn-Pex5 runs as larger species on a gel. These data are among the first to show that Pex1/Pex6 processes Pex5 and that ubiquitination of Pex5 is a requirement for such activity.

CHAPTER 4

The AAA+ ATPase Msp1 is a processive protein translocase with robust unfoldase activity

Msp1 is a conserved AAA+ ATPase in the meiotic clade of the AAA+ protein superfamily. It resides in the outer mitochondrial membrane (OMM) and is proposed to function as a quality control agent of mislocalized tail-anchored (TA) proteins. Recent *in vivo* and *in vitro* approaches support a model where Msp1 functions by extracting mislocalized TA substrates from the OMM, but fail to define a mechanism by which Msp1 acts on its substrates during extraction. I developed a system for studying Msp1 motor function and mechanism by reconstituting the cytosolic portion of Msp1 through the use of a hexamerization scaffold. Based on data for various model substrates, we can conclude that Msp1 is a robust unfoldase, capable of unfolding a diversity of substrates. Furthermore, our approach establishes a framework for biochemically characterizing the molecular mechanisms of other membrane-anchored motors as soluble hexamers.

4.1 Reconstituting Msp1 activity in vitro

Previous work utilized size exclusion chromatography to establish that $^{\Delta TMD}$Msp1 does not form stable hexamers unless it is bound to ATPγS or harbors a Walker B mutation to prevent the hydrolysis of bound ATP (Wohlever et al., 2017), which presents a challenge to studying motor function *in vitro*. I therefore set out to facilitate $^{\Delta TMD}$Msp1 hexamerization by fusing it to another moiety that hexamerizes with higher affinity. I modeled these pursuits after two other publications: Shi et al., 2016, which utilized the *de novo*-synthesized hexameric coiled-coil (ccHex) domain to hexamerize $^{\Delta TMD}$Yme1 in lieu of its N-terminal TMDs; and Monroe et al., 2017, who utilized a C-terminal fusion of the hexameric Hcp1 protein (*P. aeruginosa*) to hexamerize Vps4. I elected to fuse the hexamerization domain to Msp1's N-terminus, replacing the TMD. In addition to ccHex and Hcp1, we also explored the use of a new hexamerization scaffold: the N-domain of the archaeal proteasome PAN (PANN).

After fusing PANN (Fig. 4.1a) and ccHex (Fig. 4.1b) to $^{\Delta TMD}$Msp1's N-terminus, I expressed and purified each in the presence of ATP. Size exclusion chromatography showed that while $^{\Delta TMD}$Msp1 elutes from a Sup6i SEC column as a monomer/dimer around 16.5mL, scaffolded PANN-Msp1 and ccHex-Msp1 clearly form higher order oligomers and elute much earlier (Fig. 4.1c). Both chimeras elute as two larger-molecular weight peaks: Peak 1 (~12 mL elution volume), and Peak 2 (~14 mL elution volume) (Fig. 4.1c). Sending Peak 1 or Peak 2 once again over a Sup6i column reveals that both populations are relatively static: Peak 1 eluted at 12.5 mL, and Peak 2 eluted a bit later at 14.5 mL (Fig. 4.1d).

$^{\Delta TMD}$Msp1, despite eluting off SEC late, possesses a low rate of ATP hydrolysis above background. At high concentrations (1.25 μM $^{\Delta TMD}$Msp1 'hexamer' ≈ 7.5 μM monomer) ATP hydrolysis can reach ~20 ATP hexamer^{-1} min-^{1} (Fig. 4.2a); however, dilution of $^{\Delta TMD}$Msp1 to lower concentrations (≤ 250nM) leads to a disproportional loss of ATPase activity, indicating hexamer dissociation (Fig. 4.2e). The scaffolded PANN-Msp1 and ccHex-Msp1 constructs, in contrast to $^{\Delta TMD}$Msp1, possess a much higher ATPase rate of ~180 min^{-1} and ~120 min^{-1}, respectively (Fig. 4.2a). Moreover, both constructs maintain a constant rate of ATP hydrolysis even upon dilution from 250 nM to

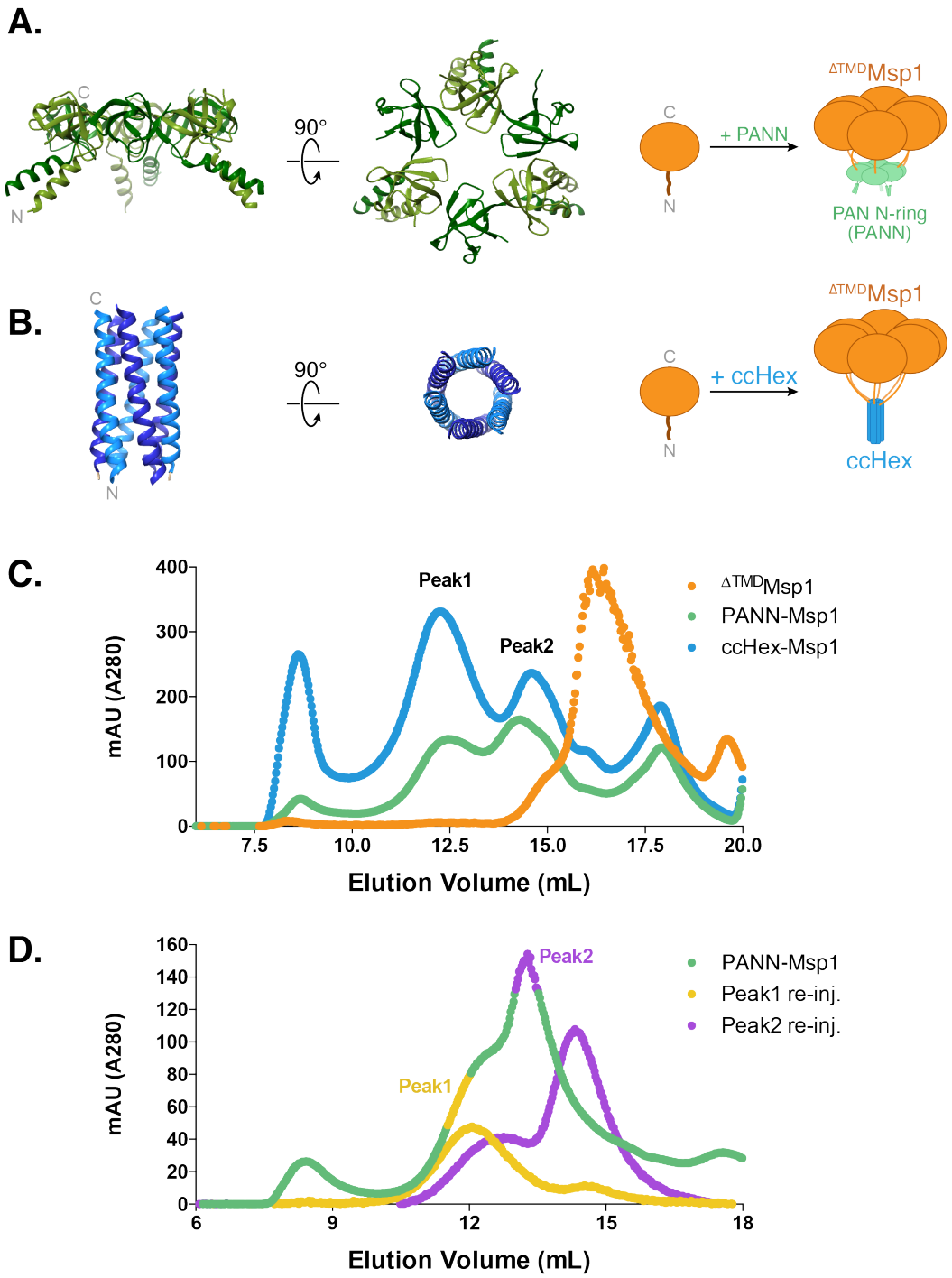


Fig. 4.1 : Fusing Msp1 to hexamerization scaffolds PANN and ccHex facilitate oligomerization of $^{\Delta TMD}$Msp1. (A) Crystal structure of the PAN N-ring (3H43), and cartoon rendering of PANN-Msp1. (B) Crystal structure of the ccHex domain (3R3K), and cartoon rendering of ccHex-Msp1. (C) Size exclusion chromatography elution profiles of $^{\Delta TMD}$Msp1, PANN-Msp1, and ccHex-Msp1 off of a Sup6i column. The latter two constructs elute much earlier as two peaks, p1 and p2, indicating oligomerization. Re-injection of p1 and p2 show that the populations are largely static, and elute again at their respective values.

25 nM (Fig. 4.2c, d, e). These data indicate that the Msp1 fusions are stable enough to remain hexameric at concentrations where $^{\Delta TMD}$Msp1 dissociates.

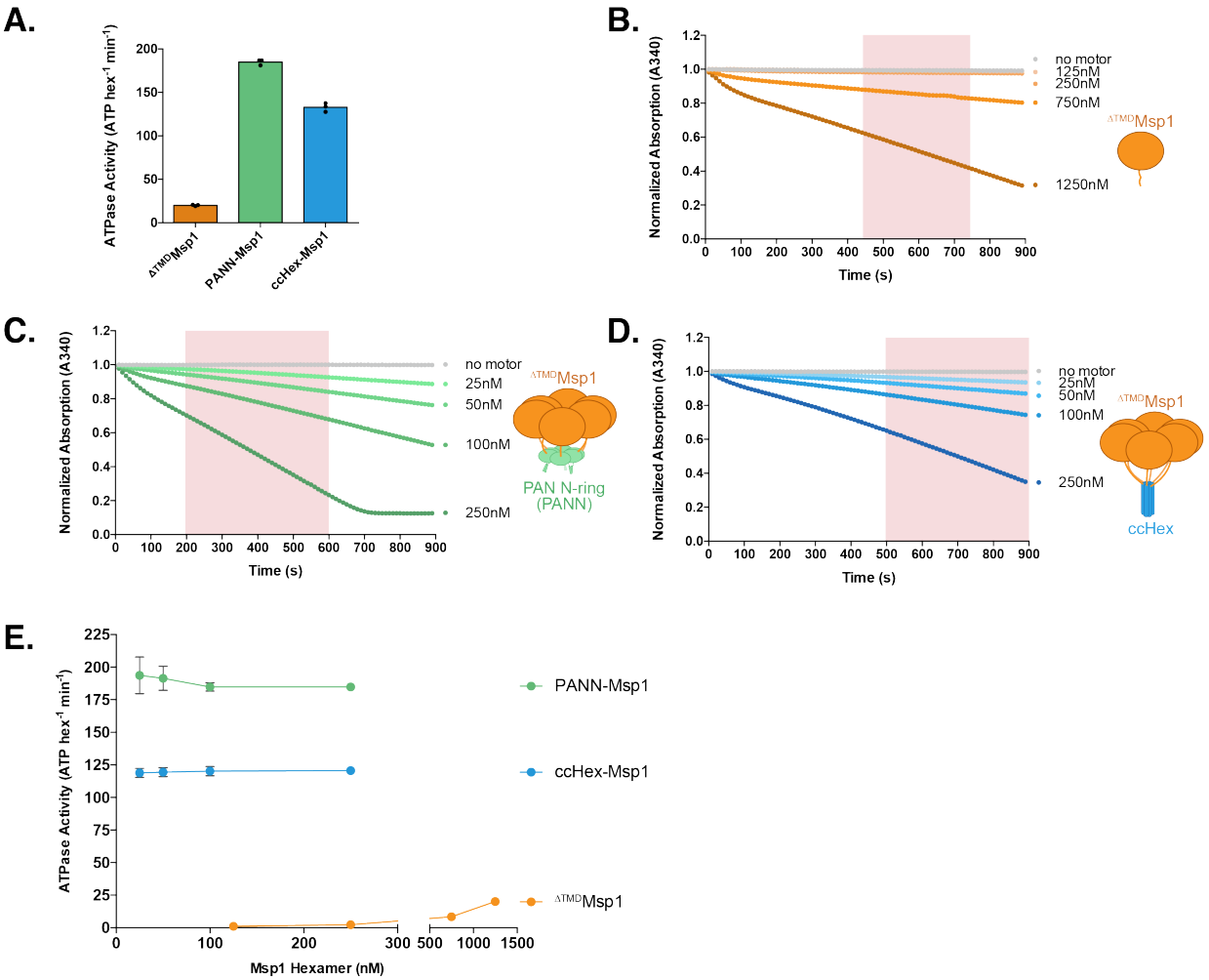


Fig. 4.2 : Hexamerized Msp1 possesses robust ATPase activity. (A) High concentrations of $^{\Delta TMD}$Msp1 (1250nM) produces a low rate of ATP hydrolysis, while both PANN-Msp1 and ccHex-Msp1 possess robust ATPase activity at a much lower concentration (100nM; n = 3, technical replicates). (B-D) Representative raw data of $^{\Delta TMD}$Msp1, PANN-Msp1, and ccHex-Msp1 ATPase activity at a range on concentrations, respectively. (E) ATPase data demonstrating the stability of PANN-Msp1 and ccHex-Msp1 hexamers, which retain robust ATPase activity even at very low concentrations. Conversely, $^{\Delta TMD}$Msp1's ATPase activity diminishes upon dilution.

The N-termini of adjacent PAN N-ring domains form paired helices (Fig. 4.1a). Since the PANN-Msp1 complex is purified using N-terminal His tags, we hypothesized that such a concentration of histidines could explain the strange oligomerization dynamics that yielded two higher-order peaks from SEC. Upon proteolytic removal of the tags, I observed of PANN-Msp1 populations from Peak1 toward Peak2 (Fig. 4.3a). Furthermore, the rate of ATP hydrolysis also increases in the absence of the tags (Fig. 4.3b). In order to learn more about these two different populations of PANN-Msp1, I performed negative stain electron microscopy (EM) with the help of Ben LaFrance from the Nogales Lab.

While the raw micrographs of the Peak1 population depicted regularly-sized yet amorphous proteinaceous masses, the 2D-class averages reveal a regularity in the density. Several of the classes appear to consist of starfish-like structures with three or four lobes projecting from a central density (Fig. 4.4a). Given the size of these masses, we wondered whether these are a collection of PANN-Msp1 hexamers conjoined by the

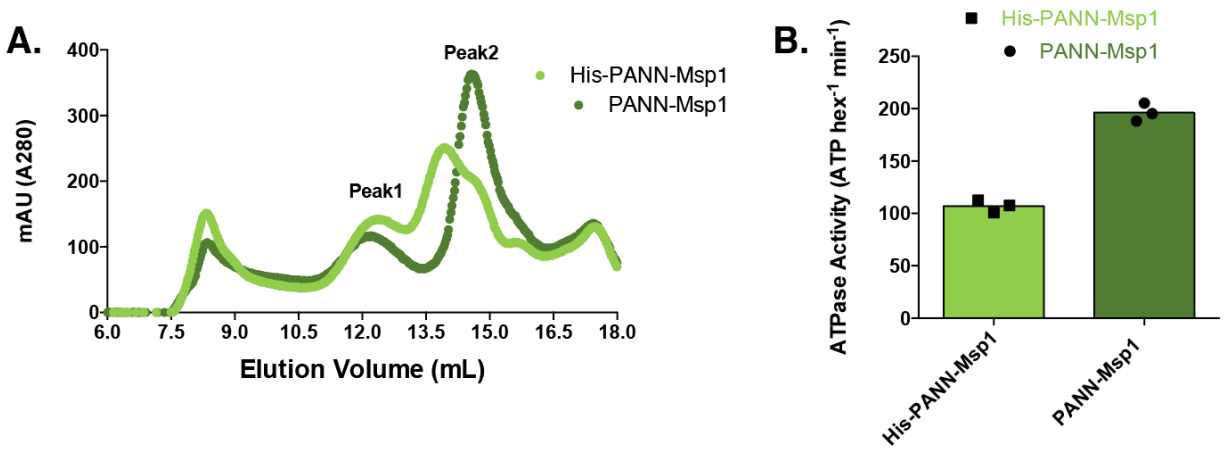


Fig. 4.3 : Construct and buffer optimization of PANN-Msp1 purification and ATPase activity. (A) Cleavage of PANN-Msp1's N-terminal 6xHis tags shifts the elution equilibrium toward the p2 population. (B) Tagless PANN-Msp1 has a higher basal ATPase rate than tagged PANN-Msp1 (n = 3, technical replicates).

PAN N-ring domains. Crude measurements of the PAN N-ring (solid wedge, ~65 Å across) and the Msp1-corresponding densities (hollow wedge, ~115 Å) show that the widths of these densities roughly match the diameter of the PAN N-ring (67 Å) and the $^{\Delta TMD}Msp1^{E193Q}$ hexamers (~135 Å) observed by Wohlever et al., 2017 (our data is shown in Fig. 4.4b). Given that the protein in this sample possesses ATPase activity, we surmise that PANN-Msp1 in Peak1 forms tri- or tetra-hexamers. Since removal of the His tags shifts the balance of PANN-Msp1 molecules from Peak1 to Peak2, we can further conclude that this higher-order oligomerization into tri- and tetra-hexamers is in part mediated by the His tags.

Negative stain EM of the Peak2 population revealed structures roughly corresponding to the shape and size we would expect for PANN-Msp1 hexamers. While the density does not possess perfect six-fold symmetry, it is evident that there is a radial

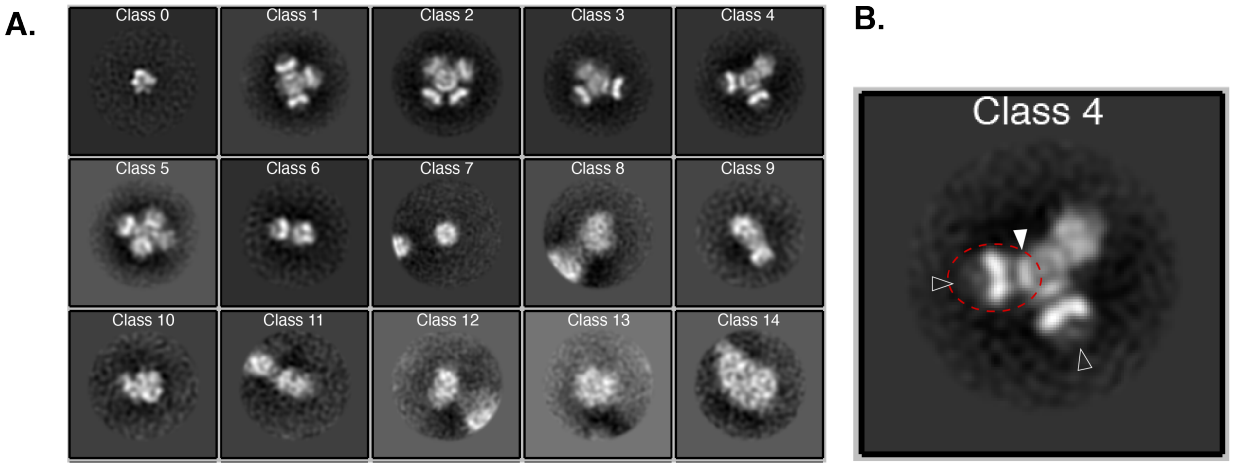


Fig. 4.4 : Negative-stain EM of the p1 population reveals PANN-Msp1 hexamer oligomers. (A) Various 2D class averages of the PANN-Msp1 p1 population, showing large three- and four-lobed complexes. (B) Enlarged image of Class 4, with what we believe is a single PANN-Msp1 hexamer circled in red. Density matching with width of the PANN domain is visible (solid wedge). An additional density on top of the Msp1 AAA domain is visible as well (hollow wedge), which was not visible in previous top-down or bottom-up views of the Msp1 hexamer reported in Wohlever et al., 2017. Data collected, processed, and refined by Ben LaFrance.

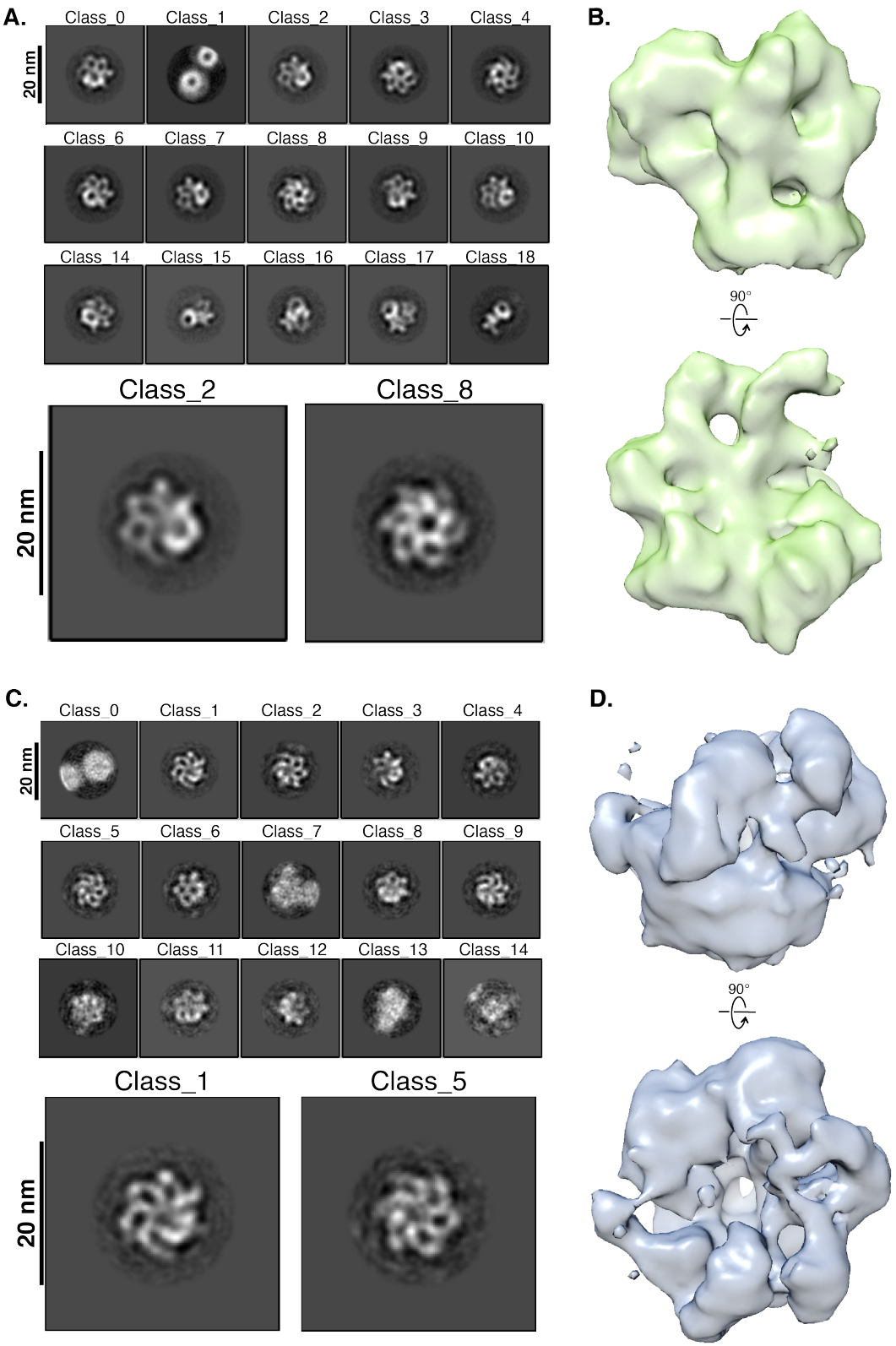


Fig. 4.5 : Negative-stain EM of the p2 population reveals distinct PANN-Msp1 hexamers. (A) 2D class averages of the PANN-Msp1 p2 population in ATPγS, which resemble distinct hexamers. (B) 3D reconstructed models generated from the class averages in (A). (C) 2D class averages of the PANN-Msp1 p2 population in ATP with Pex15$^{1-309}$. (D) 3D reconstructed models generated from the class averages in (C). Data collected, processed, and refined by Ben LaFrance.

regularity that resembles 5 of 6 protomers (Fig. 4.5a). Further processing of the 2D-

class averages into 3D reconstructions yields a density into which one can dock a model AAA hexamer (in this case, the D2 domain of p97), as well as an additional bottom density into which the published PAN N-ring structure easily fits (Fig. 4.5b). These data support the conclusion that the Peak2 population of PANN-Msp1 is composed of a species with approximately the expected shape and size of PANN-Msp1 hexamers.

It is suspected that Pex15 is a substrate of Msp1, and we therefore included Pex15 with PANN-Msp1 on grids to see if it induced any large-scale conformational changes in the motor. We mixed 10 μM $Pex15^{1-309}$ with PANN-Msp1 and incubated at RT for 5 min before applying it to a grid. Interestingly, we observed far fewer particles when Pex15 was included (~25,000 particles in the presence of Pex15, as opposed to ~46,000 particles without Pex15 from a similar number of micrographs). It is possible that such a high concentration of Pex15 coats the grid, making it more difficult for PANN-Msp1 to lay on the grid (Fig. 4.5c). From the particles we did observe, however, there were subtle differences from the substrate-free motor. First, the resolution in the substrate-containing sample was a bit lower, but whether this represents an increase in motor dynamics or is simply a result of averaging fewer particles is unclear. Furthermore, while only 5 of the 6 protomers are clearly visible in the substrate-free density, the motor density in the presence of substrate appears more complete, with the 6^{th} subunit slightly more visible (Fig. 4.5d). This may represent a difference in motor conformation when actively engaged to and processing a substrate.

Although Pex15 is a putative substrate for Msp1, this has never been directly demonstrated *in vitro*. Moreover, while it has been shown that Msp1 possesses the ability to extract TA proteins from membranes, how Msp1 mechanistically processes substrates is unknown. With a Pex15 unfoldase reaction already established, I sought to determine whether Msp1 possesses the ability to unfold its substrates. An extended 30 min maleimide labeling assay revealed modest unfoldase activity for PANN-Msp1, while no unfoldase activity was observed for the Walker-B mutant PANN-$Msp1^{E193Q}$ (Fig. 4.6a). These data demonstrate that hexameric Msp1 is capable of unfolding $Pex15^{1-309}$ in an ATP-hydrolysis-dependent manner.

It remained unclear whether Pex15 was threaded through the central pore of PANN to engage with Msp1's pore loops, or could enter Msp1's central channel through the gap between the PAN N-ring and Msp1's ATPase ring. Since both Peak1 and Peak2 appeared to be ATPase-competent, and since we can infer from the negative stain EM 2D class averages that Peak1 represents a mixture of tri- and tetra-hexamers conjoined by the N-terminal regions of the PAN N-ring coiled-coils, I decided to test whether both Peak1 and Peak2 possessed unfoldase activity. In maleimide labeling unfoldase assays, it appears that both Peak1 and Peak2 PANN-Msp1 complexes are capable of unfolding Pex15 to roughly the same extent (Fig. 4.6b). We cannot be certain that Pex15 is not making its way through the center of PANN before engaging with Msp1's pore loops, but given the steric obstacles present above the PAN N-ring in the tri- and tetra-hexamers, it seems more likely that Pex15 engages with the Msp1 pore loops by entering though the gap between the PAN N-ring and the ATPase ring. Previous studies on the PAN N-ring claim to identify a mutation (G113W) that blocks the central pore of the ring. I introduced this mutation to the PANN-Msp1 complex, but the motor exhibited significantly impaired ATPase activity, perhaps indicating improper

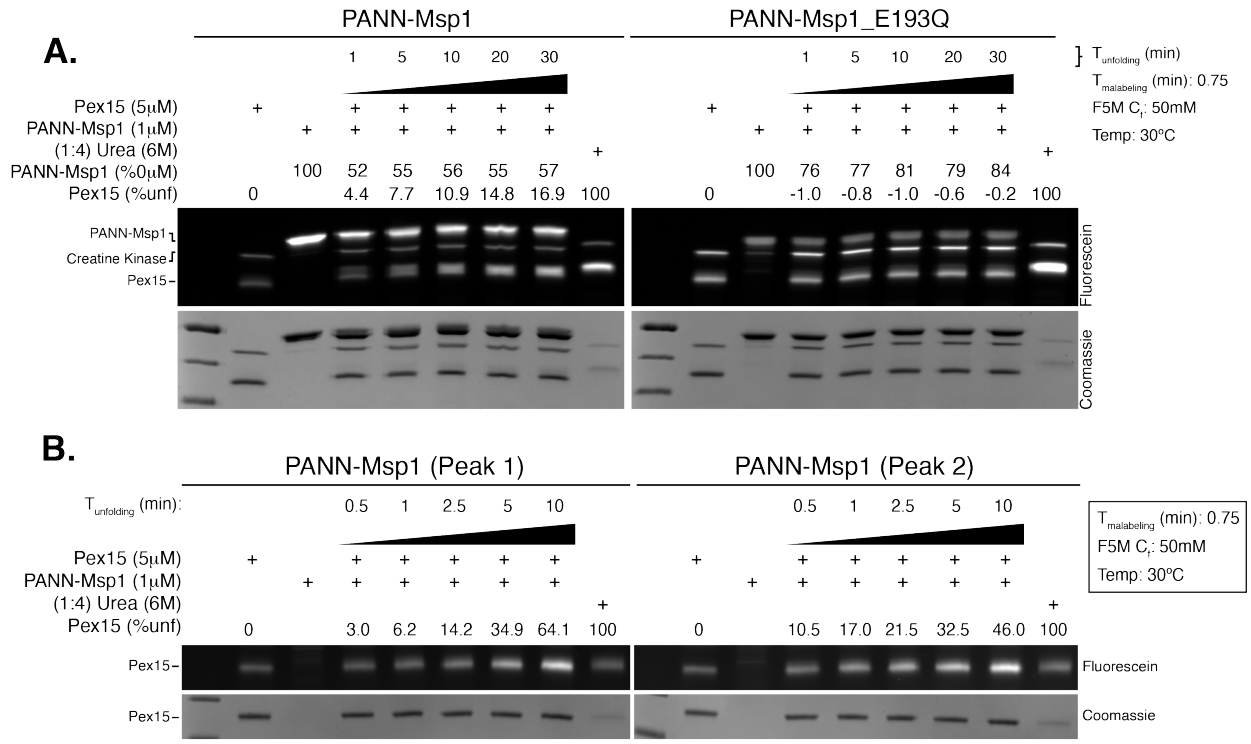


Fig. 4.6 : PANN-Msp1 is an unfoldase. (A) Maleimide labeling unfoldase reaction timecourse using wild-type PANN-Msp1 and Walker B mutant PANN-Msp1^{E193Q}, demonstrating that wild-type PANN-Msp1 unfolds Pex15 over time, while the hydrolysis-incompetent mutant is unable to. (B) Maleimide labeling unfoldase reaction comparing the unfoldase activities of the p1 and p2 populations of PANN-Msp1 reveal similar levels of unfoldase activity among both populations.

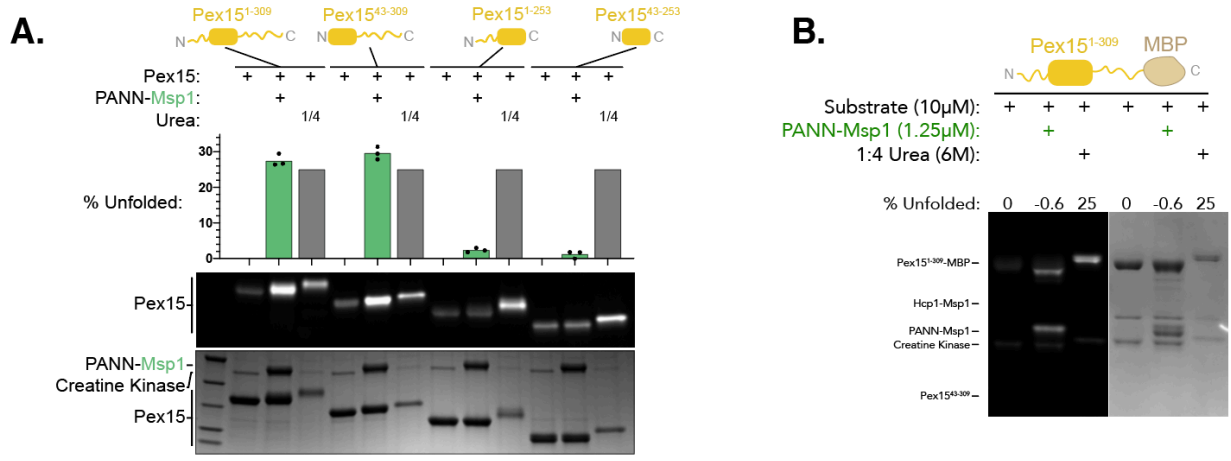


Fig. 4.7 : PANN-Msp1 requires Pex15's C-terminal unstructured region for unfolding. (A) Maleimide labeling unfoldase assays using the truncated Pex15$^{1-309}$, Pex15$^{34-309}$, Pex15$^{1-253}$, and Pex15$^{43-253}$. It appears that only Pex15$^{1-309}$ and Pex15$^{43-309}$ are unfoldable, while the shorter two constructs are not. (B) Maleimide labeling unfoldase assay demonstrating that PANN-Msp1 is unable to unfold Pex15-MBP.

oligomerization. It may be of value to repeat this experiment with a milder mutation to block the N-ring pore without impairing PANN hexamerization.

I initially observed that the presence of Pex15 stimulated the rate of ATP hydrolysis for His-PANN-Msp1. The earlier observation that removal of PANN-Msp1's N-terminal His tags leads to a higher basal rate of ATP hydrolysis suggests that PANN-

Msp1 complexes without the N-terminal His tags are “healthier”. If this were true in all cases, then cleaved PANN-Msp1 should unfold Pex15 to a greater extent than uncleaved complexes. In a side-by-side comparison this appeared to be true, as tagless complexes were able to unfold more Pex15 in the same amount of time (data not shown). Interestingly, although the basal ATPase rate of tagless PANN-Msp1 was higher than that of the tagged complex, it was not affected by the addition of Pex15. For all subsequent experiments, I used cleaved PANN-Msp1, lacking the N-terminal His tags.

Previous unfoldase experiments with Pex1/Pex6 highlighted the critical role that Pex15$^{1\text{-}309}$’s unstructured regions play in substrate recognition and engagement. To evaluate whether the same holds true for Msp1, I compared the unfoldability of Pex15$^{43\text{-}253}$, Pex15$^{1\text{-}253}$, Pex15$^{34\text{-}309}$, and Pex15$^{1\text{-}309}$, using the maleimide labeling unfoldase assay. While Pex15$^{1\text{-}309}$ and Pex15$^{43\text{-}309}$ were easily unfolded by PANN-Msp1, Pex15$^{1\text{-}253}$ and Pex15$^{43\text{-}253}$ were not (Fig. 4.7a). From these data, I concluded that soluble Pex15 is unfolded by Msp1 from the C-terminus. To confirm this, I tested the unfoldability of Pex15$^{1\text{-}309}$ with an MBP domain linearly fused to its C-terminus, and found that this construct could not be unraveled by PANN-Msp1, providing further evidence for a C-terminal engagement of Pex15$^{1\text{-}309}$ (Fig. 4.7b). Furthermore, these results suggest that soluble PANN-Msp1 is unable to engage an unstructured loop without a free terminus, as the Pex15$^{1\text{-}309}$-MBP construct still harbors a 56-residue unstructured region between the folded Pex15 and MBP domains. It remains unknown how native Msp1 engages and extracts mislocalized TA proteins from membranes *in vivo*—it is possible that native Msp1 possesses the ability to engage unstructured loops and that the presence of the PANN domain, although necessary for hexamerization of the soluble construct, inhibits this process *in vitro*. On the other hand, it is possible Msp1 naturally exists as monomers or dimers in the membrane, and then assembles around a substrate to extract it. While the soluble reconstituted *in vitro* system I have created lacks the ability to distinguish these models, we can still exploit it to learn more about the characteristics of Msp1 as an ATPase and its capabilities as an unfoldase.

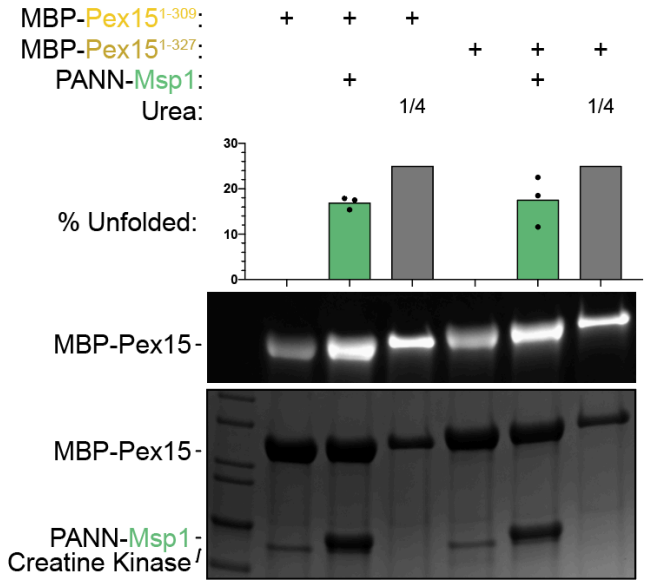


Fig. 4.8 : Inclusion of Pex15’s hydrophobic patch does not influence unfolding by PANN-Msp1. Maleimide labeling unfoldase assays comparing MBP-Pex15$^{1\text{-}309}$ (which is truncated before the hydrophobic patch) to MBP-Pex15$^{1\text{-}327}$ (including the hydrophobic patch) reveal that both substrates are unfolded to a similar extent (n = 3).

Toward the end of my PhD, a paper was published in EMBO (Li & Zheng et al., 2019) that used various *in vivo* assays to detail Msp1’s substrate selection process. Specifically, they suggested that a hydrophobic patch of residues just N-terminal of Pex15’s TMD is the primary recognition site for Msp1. More broadly, they posit that such a hydrophobic patch is displayed on all

mislocalized proteins, and that it is Msp1's signal for extraction. This sequence, corresponding to 12 residues (313-324, ITVLAGSFWAVL), is the sequence we learned to specifically avoid in our efforts to produce a Pex15 construct that expressed and purified well (Chapter 2, Section 2.1). Since I have developed an Msp1 unfoldase assay for Pex15, I sought to probe the effect of including this hydrophobic patch in my substrates to see if its presence changes the extent of unfolding in a maleimide labeling unfoldase assay.

N-terminally fusing a MBP increased the solubility of Pex15, and I was able to purify a small amount of MBP-Pex15$^{1-327}$, including the native hydrophobic patch sequence. This construct remained stable and soluble as long as the affinity purification tags were not cleaved off, and the protein was not concentrated after SEC. I was able to compare the unfolding of MBP-Pex15$^{1-309}$ and MBP-Pex15$^{1-327}$ in a maleimide labeling unfoldase assay, and observed no significant difference (Fig. 4.8). Ideally we would be able to compare a broad range of concentrations to determine whether the hydrophobic patch leads to changes in Km, but these efforts failed due to the low solubility of Pex15 including the hydrophobic patch. While it remains possible that this sequence plays an important role *in vivo*, perhaps by stimulating the oligomerization Msp1 monomers, our *in vitro* data suggest that it makes little to no difference in our soluble, reconstituted system.

4.2 Msp1 unfolds substrate via processive threading, and is inhibited by Pex3

Having reconstituted the two-component system of Msp1 processing Pex15, I sought to test the model postulated in Weir et al., 2017, wherein Pex3 inhibits Msp1's ability to process Pex15. Initial attempts to purify full-length Pex3 yielded very little protein that quickly precipitated even in the presence of detergents. Endogenous Pex3 is embedded in the peroxisomal membrane, and transmembrane domain prediction servers (TMHMM Server v. 2.0, DTU Health Tech) clearly delineate amino acids 20-39 as the transmembrane portion of the protein. Truncation of the ORF to Pex3$^{40-441}$ yields a more soluble species, which, however, co-purifies with a considerable amount of impurities. In an attempt to solubilize Pex3, I fused MBP to the N-terminus to create MBP-Pex3$^{40-441}$. This construct expresses and purifies very well, ultimately producing a clean sample of MBP-Pex3$^{40-441}$.

In agreement with the published model, addition of MBP-Pex3 into maleimide labeling reactions does indeed inhibit unfolding of Pex15 in a concentration-dependent manner (Fig. 4.9a). MBP-Pex3 appears to have at least one solvent-accessible cysteine, given how readily it is labeled by F5M, but there is a molar excess of F5M over all cysteines present in the reaction and we are confident that the reduced Pex15 fluorescence is due to inhibition of unfolding and not depletion of F5M.

To understand more about MBP-Pex3's inhibition of Pex15 unfolding by Msp1, I performed maleimide labeling unfoldase reactions in varying salt concentrations ranging from 10 mM to 150 mM of both NaCl and KCl. Interestingly, MBP-Pex3's ability to inhibit unfolding is inversely correlated to the ionic strength of the buffer: at the lowest salt

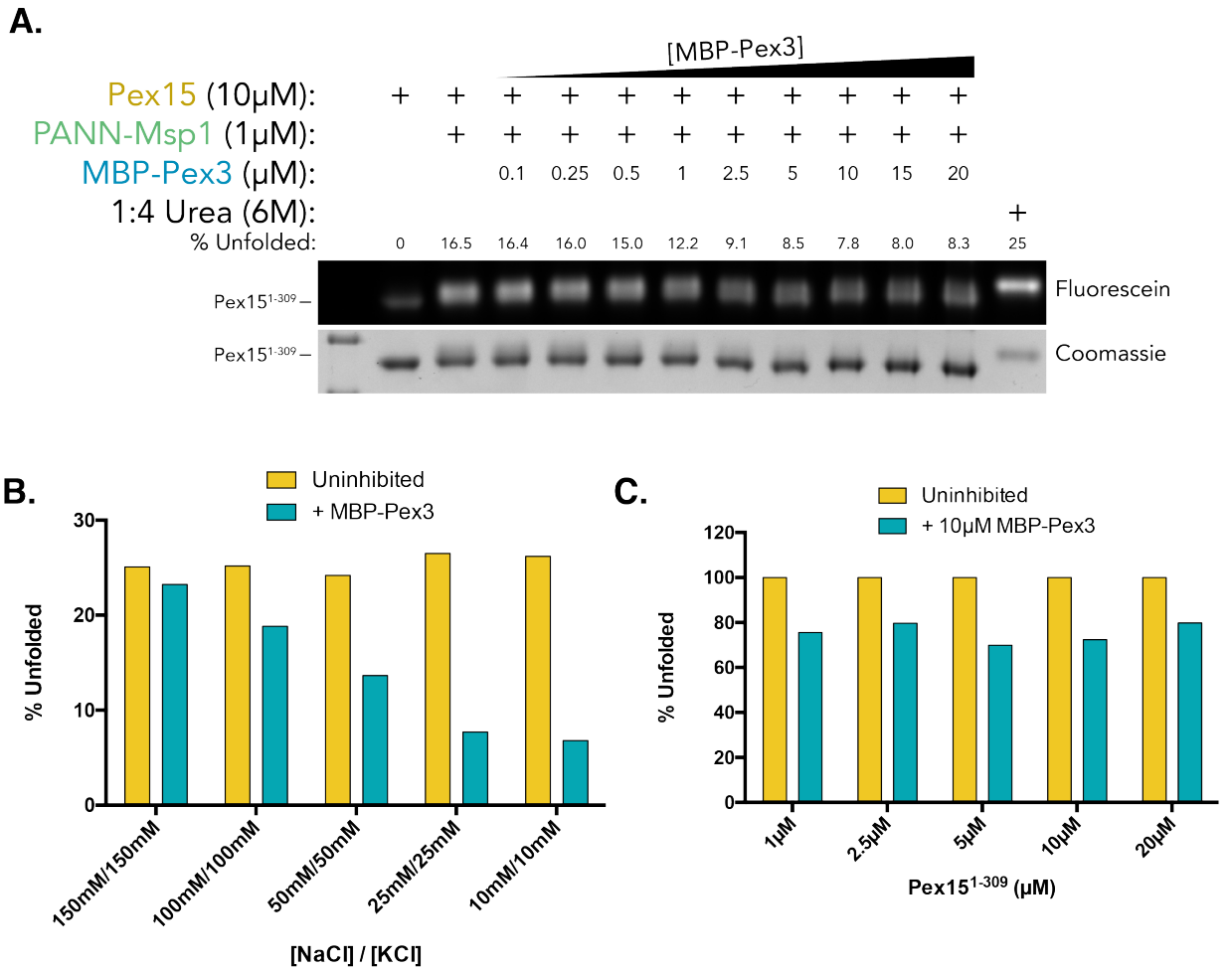


Fig. 4.9 : MBP-Pex3 inhibition of PANN-Msp1 unfoldase activity is concentration-dependent. (A) Maleimide labeling unfoldase reaction with titrated MBP-Pex3 reveals a concentration-dependent inhibition of Pex15 unfolding in a Pex3-dependent manner. (B) Quantified maleimide labeling unfoldase assays show that MBP-Pex3 inhibition of Pex15 unfolding is inversely related to salt concentration. (C) Quantified maleimide labeling unfoldase assays in which motor and inhibitor have been held at constant concentrations (as in (A)), but substrate concentration has been titrated. It appears that Pex3 is able to inhibit Pex15 unfolding to a similar extent, regardless of substrate concentration.

concentrations, MBP-Pex3 reduces Pex15 unfolding by nearly 75%, while at the highest salt concentration Pex3 barely inhibits unfolding at all (Fig. 4.9b).

The model postulated by Weir et al., 2017, suggests that Pex3 directly binds Pex15 to inhibit its unfolding. If this model were true, then titrating Pex15 from 1 μM to 20 μM while keeping MBP-Pex3 constant at 10μM in a maleimide labeling unfoldase reaction should change the extent of inhibition by MBP-Pex3, as the amount of shielded substrate is changing. Instead, I observed a constant unfoldase-activity inhibition of ~20% by 10μM MBP-Pex3 at all concentrations of Pex15 (Fig. 4.9c). While our data agree with the published findings that Pex3 inhibits Pex15 unfolding by Msp1, they indicate that Pex3 acts on Msp1, rather than shielding the substrate. If this were true, Pex3 should be able to inhibit unfolding of non-Pex15 substrates as well.

To test this, I needed to develop a new unfoldase assay that does not rely on Pex15's buried cysteines. Further justification for establishing a new unfoldase assay is the desire to define Msp1's enzyme kinetics. To this end, I decided to track the unfolding of a fluorescent protein, mEOS3.2. mEOS3.2 is an ideal substrate, given its

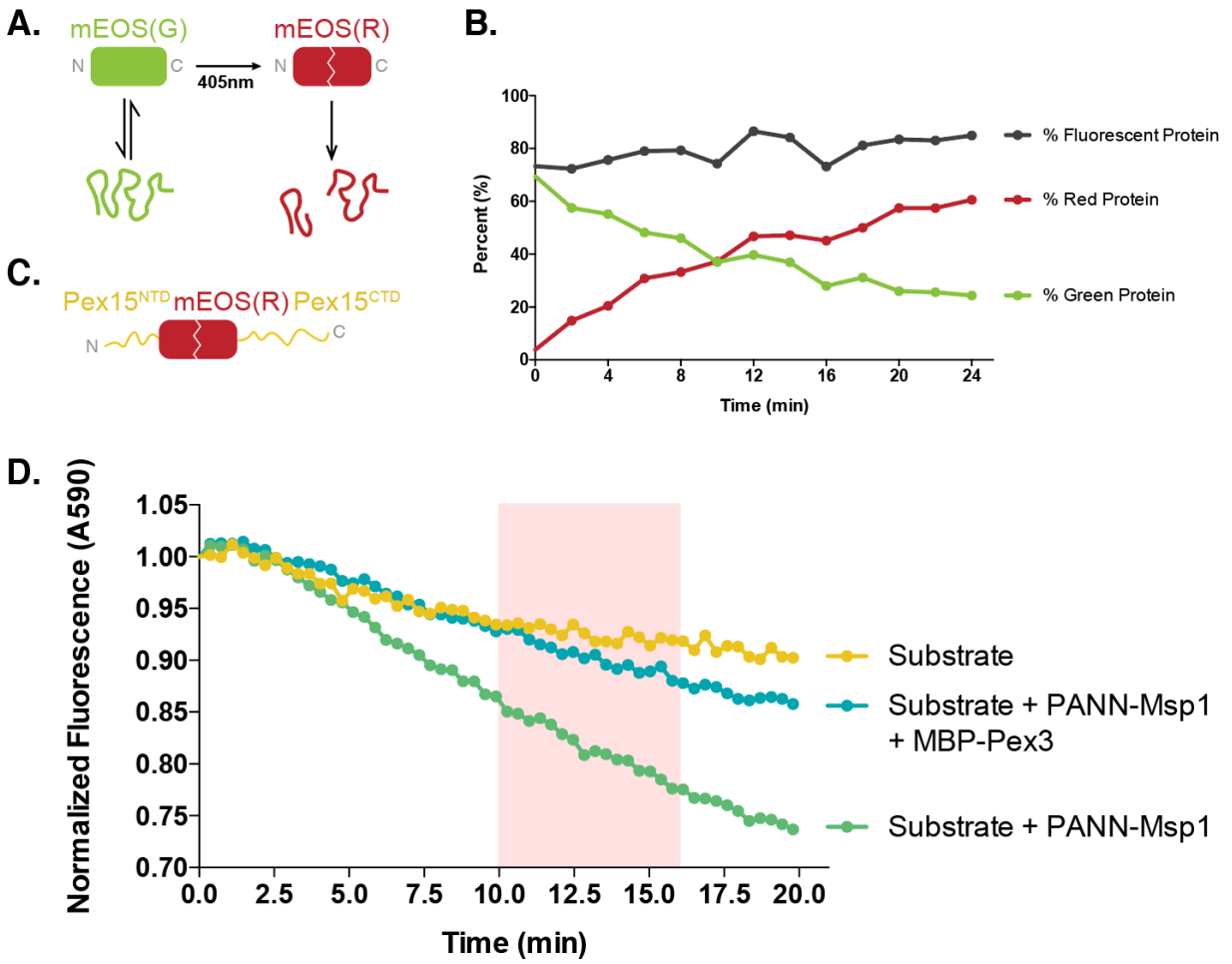


Fig. 4.10 : Developing a new substrate for kinetic unfoldase measurements. (A) Non-activated green mEOS (mEOS(G)) is able to refold after unfolding; however, after activation to red mEOS (mEOS(R)), the protein's backbone is cleaved and is therefore unable to refold once unfolded. (B) Cartoon depiction of the chimeric protein wherein Pex15's structured core domain has been replaced by mEOS. (C) Activation of the chimeric protein with 405nm lasers gradually converts mEOS(G) to mEOS(R). Activation time is given on the x axis. (D) Incubation of the chimeric protein with PANN-Msp1 in similar reaction conditions to maleimide labeling assays and tracking mEOS(R) fluorescence shows an accelerated loss of mEOS(R) in the presence of PANN-Msp1 (green), and that this loss of fluorescence is slowed in the presence of MBP-Pex3 (blue).

photoconvertible properties: when activated by a 405 nm laser, it auto-cleaves its backbone in one position and experiences a spectral shift from green (mEOS(G)) to red (mEOS(R)), while still remaining natively folded. However, once unfolded and separated, the two fragments of mEOS(R) cannot refold to a fluorescent protein (Fig. 4.10a), and the loss of mEOS(R) fluorescence can thus be used as a sensitive probe for protein unfolding. To activate mEOS, I exposed the protein to a 50 Watt, 405 nm laser for 2 minutes of activation, alternated with 2 minutes of cooling time on ice. I repeated this cycle 10 - 12 times. Even after 24 total minutes of activation, the protein continues photoconverting from green to red, which is depicted by the percent red increasing and the percent green decreasing throughout activation (Fig. 4.10b).

Not knowing Msp1's substrate specificity, I sought to evaluate and compare Msp1's ability to unfold two potential substrates: mEOS alone and $Pex15^{NTD}$-mEOS-$Pex15^{CTD}$, which is a mEOS domain with the unstructured regions of Pex15 (defined in Chapter 2, Section 2.2 and depicted in Fig. 4.10c). Whether or not PANN-Msp1 is able

to unfold this substrate will help elucidate how Msp1 recognizes or distinguishes substrates. I monitored the red protein fluorescence in a plate reader, exciting at 565 nm and reading emission at 585 nm. The data suggest that mEOS alone is not unfoldable by PANN-Msp1 (data not shown), which agrees with the above-described results that Msp1 requires a long unstructured region for engagement. The fluorescence of $Pex15^{NTD}$-mEOS-$Pex15^{CTD}$, however, steadily decreased in the presence of PANN-Msp1, providing evidence not only that PANN-Msp1 is strong enough to unfold the mEOS domain, but also that the core of Pex15 is not essential for substrate binding or engagement by PANN-Msp1 (Fig. 4.10d). Addition of MBP-Pex3 to the unfoldase reaction prompted a slower loss of red fluorescence, suggesting inhibition of mEOS(R) unfolding. These data further support our model that Pex3 likely inhibits the Msp1 motor directly, rather than interacting with the Pex15 substrate.

Having established unfoldase assays for both Pex15 and mEOS, we sought to determine whether PANN-Msp1 unfolds substrates via a tug-and-release mechanism or processive-threading through the central pore. We elected to test these two unfoldase mechanisms using mEOS-$Pex15^{1-309}$, a substrate developed by Kaili Carr during her time in the Martin Lab (Fig. 4.11a). Based on the observation that PANN-Msp1 cannot

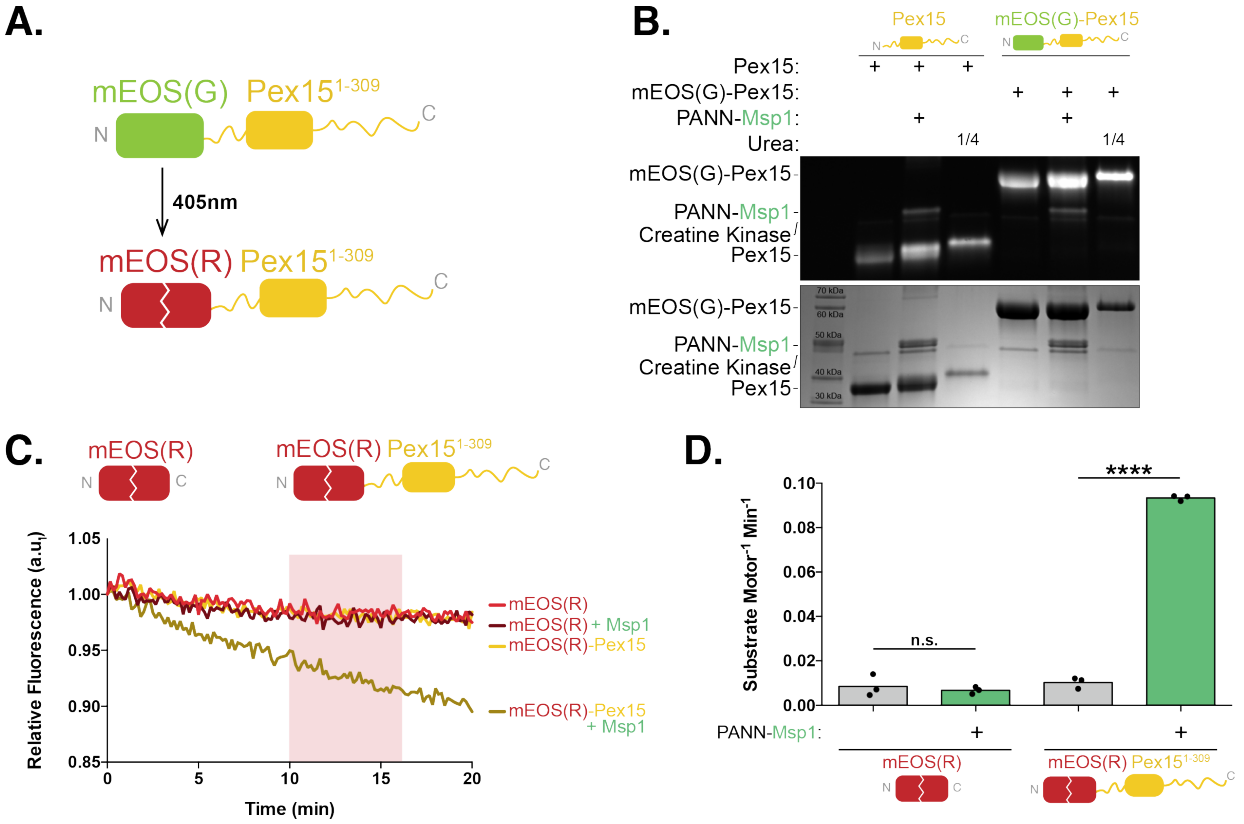


Fig. 4.11 : Determining rate of unfolding by tracking loss of mEOS fluorescence. (A) Cartoon depiction of model substrate with mEOS linearly fused to the N-terminus of Pex15, then activated using 405nm laser. (B) Maleimide labeling unfoldase assay demonstrating that the Pex15 domain of mEOS-Pex15 is still unfolded my PANN-Msp1 similarly to Pex15 alone. (C) A mEOS unfoldase assay, wherein mEOS(R) fluorescence is monitored, reveals that mEOS-Pex15 is unfolded by PANN-Msp1, but mEOS alone is not. (D) Quantification of data in (C), where the slope of the mEOS(R) signal loss and the % Red protein are used to calculate a rate of substrate $hexamer^{-1}$ $minute^{-1}$ unfoldase activity.

unfold isolated mEOS, the motor should also be unable to unfold mEOS-Pex15 from the N-terminus. Furthermore, since we believe PANN-Msp1 engages $Pex15^{1-309}$ from the C-terminus, we can be confident it will do the same in the mEOS-Pex15 construct.

Regardless of which unfolding method Msp1 employs, we should observe unfolding of Pex15. Using the mEOS(G) substrate was preferable to the red version in this assay, because the non-activated green protein is a single, contiguous polypeptide chain and runs as a single band on a gel, unlike the backbone-cleaved red mEOS. Indeed, when we evaluated mEOS(G)-Pex15 unfolding using a maleimide labeling unfoldase assay, we observe unfolding of the Pex15 domain (Fig. 4.11b). Next, we performed a mEOS unfoldase assay using mEOS(R) and observed loss of red fluorescence as well (Fig. 4.11c). We concluded that loss of mEOS fluorescence would occur only if PANN-Msp1 unfolded substrates via processive threading, starting at Pex15's C-terminus and consecutively unraveling the folded domains of Pex15 and mEOS. Thus, we have demonstrated that PANN-Msp1 is a processively-threading unfoldase, and that it can unfold multiple, sequential domains of a polypeptide. Furthermore, I was able to determine the rate of unfolding with ~0.1 substrate $hexamer^{-1}$ $minute^{-1}$ (Fig. 4.11d) at 20 μM substrate. The finding that PANN-Msp1 processively unfolds multiple domains presents an opportunity to assess the mechanical unfolding of any domain: if this domain is inserted between mEOS and Pex15, it can be inferred from loss of mEOS fluorescence that said domain was unfolded by PANN-Msp1. We therefore set out to test whether or not PANN-Msp1 can unfold three folded domains of particular interest: MBP, ubiquitin, and SUMO.

MBP was of interest, because it was used as a solubilizing moiety for Pex3 as well as certain particularly unstable variants of Pex15 (see Chapter 4, Section 4.1). Incubation of mEOS-MBP-Pex15 with PANN-Msp1 led to a loss of mEOS(R) fluorescence, suggesting that PANN-Msp1 is able to unfold MBP (Fig. 4.12a). PANN-Msp1 unfolded mEOS-Pex15 and mEOS-MBP-Pex15 at a similar rate, hinting that unfolding of MBP is not the rate-limiting step for processing the latter substrate. As a positive control, I performed a maleimide labeling reaction with the non-activated substrate mEOS(G)-MBP-Pex15 and observed unfolding of Pex15 (Fig. 4.12b); it is worth noting that MBP contains no cysteines, and thus is undetectable in maleimide-labeling reactions.

Many proteins embedded in or localized to the OMM become ubiquitinated and, given Msp1's natural role in TA protein extraction as well as in mitoCPR, it is possible that some substrates are ubiquitinated prior to recognition by Msp1. To assess whether Msp1 can unfold ubiquitin, I cloned mEOS-Ubn-Pex15 and tested it in the fluorescence-based unfolding assay. In the presence of motor, I observed a loss of mEOS(R) fluorescence, indicating that ubiquitin was unfolded. Interestingly, mEOS-Ubn-Pex15 was unfolded at nearly twice the rate (0.2 substrate / hexamer / minute) of mEOS-Pex15 (0.1 substrate / hexamer / minute of mEOS-Pex15; Fig. 4.12a). One potential explanation for this is that Pex15 unfolding is the rate-limiting step of substrate processing and the fusion of ubiquitin may destabilize Pex15. The Marqusee lab in collaboration with our group has recently shown that the attachment of ubiquitin to lysine side chains can significantly destabilize protein substrates and lead to an acceleration of mechanical unfolding by ATP-dependent motor proteins (Carroll & Green et al., 2020). Ubiquitin in the mEOS-Ubn-Pex15 fusion construct may act

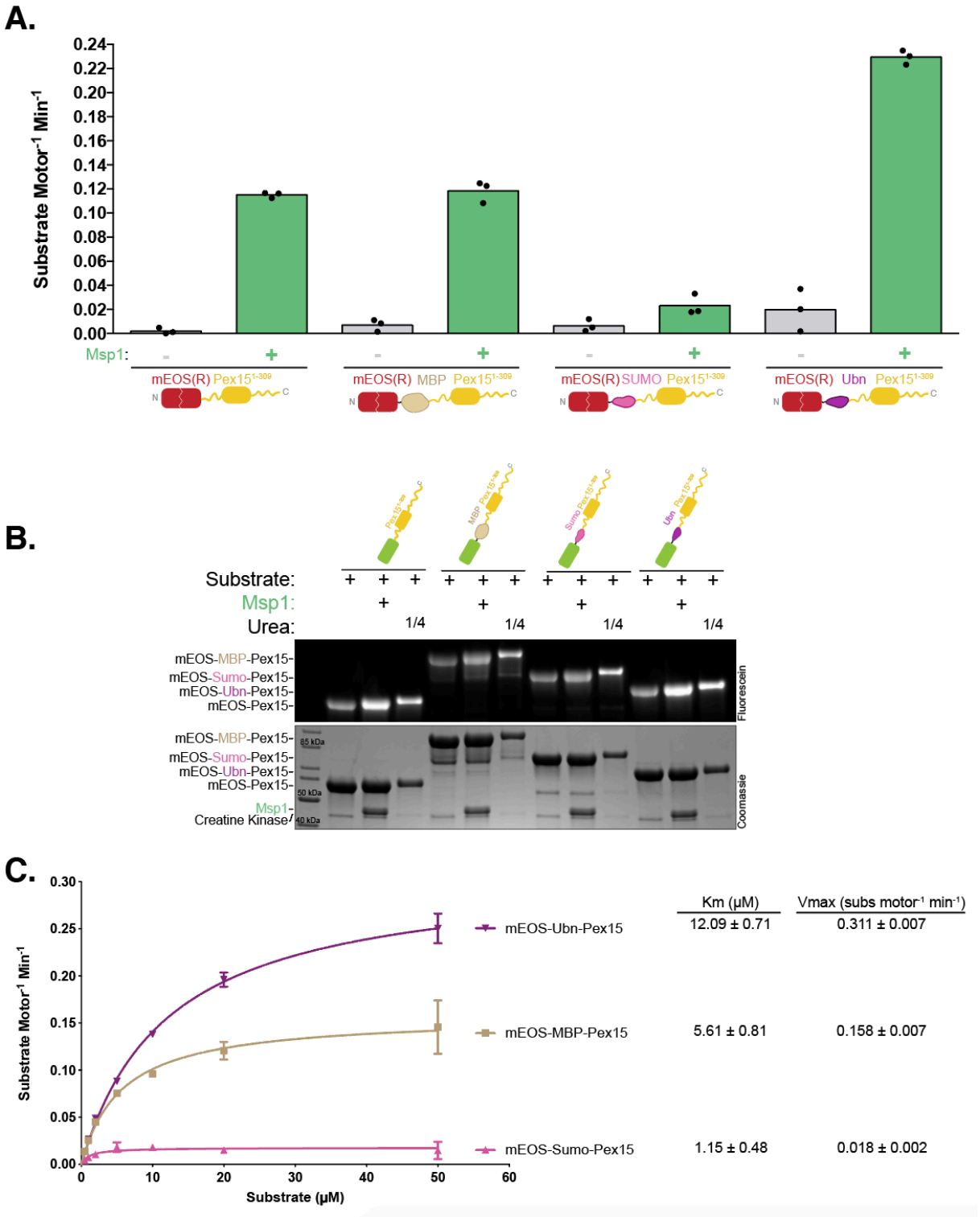


Fig. 4.12 : Michaelis-Menten kinetics reveals differences of K_m and V_{max} among varying substrate domains. (A) mEOS unfoldase assays of insert constructs suggest that MBP, SUMO, and ubiquitin are all unfoldable by PANN-Msp1, but to varying extent, with MBP and ubiquitin being readily unfolded while the motor struggles with SUMO. (B) Maleimide labeling unfoldase reactions showing the Pex15 domain of all substrates is unfolded in these reactions. (C) Michaelis-Menten kinetics of all substrates show differences in Km and Vmax depending on the insert (n = 3, technical replicates).

destabilizing on Pex15 due to proximity effects, although the details of this interaction

remain unclear, especially considering the separation by a 42-amino-acid linker (Pex15's unstructured N-terminal domain).

The Small Ubiquitin-like MOdifier (SUMO) protein is of interest due to its previous utilization in published Msp1 proteoliposome extraction assays as the folded domain component of the SUMO-TMD substrate (Wohlever et al., 2017). Interestingly, in contrast to mEOS-Ubn-Pex15, I observe a very slow rate of unfolding for SUMO (Fig. 4.12a). That SUMO itself represented a tough unfolding barrier in this case was indicated by maleimide labeling unfoldase assays, showing normal engagement and unfolding of the Pex15 moiety in mEOS-SUMO-Pex15 (Fig. 4.12b).

I next measured Michaelis-Menten (MM) kinetics of Msp1. Early attempts to perform MM measurements using $Pex15^{NTD}$-mEOS-$Pex15^{CTD}$ yielded a clear and obvious unfoldase-rate response to low and moderate substrate concentrations (0.5μM – 10μM); however, the unfoldase rate did not approach a maximum velocity but continued increasing with further increased substrate concentrations. At high concentrations (≥ 20 μM), mEOS-Pex15 exhibits unexpected unfolding kinetics. We believe that both mEOS and Pex15 are rather insoluble when unfolded and, at higher concentrations, potentially prompt other folded substrates to co-aggregate (and therefore lose fluorescence) in a motor-independent manner. The result of attempting a MM curve with this substrate is, again, a proportional and linear unfoldase rate response with protein concentration, thus precluding accurate determination of K_m and V_{max} values. I found that insertion of a solubilizing domain, MBP, between the mEOS and Pex15 domains adequately ameliorate the apparent insolubility and therefore the runaway unfolding the previous lead to a linear rate increase. I therefore was able to fit data from substrate titrations to MM curves for mEOS-MBP-Pex15, mEOS-Ubn-Pex15, and mEOS-SUMO-Pex15, and was able to calculate Km and Vmax values (Fig. 4.12c). It is curious that unfolding of SUMO appears to present such a challenge for Msp1, especially considering that the only other *in vitro* reconstitution of Msp1 activity utilized SUMO-TMD^{Sec22} as a model substrate (Wohlever et al., 2017). However, it is possible that in this assay for extraction of SUMO-TMD^{Sec22} from proteoliposomes, engagement of the substrate occurred such that only the TMD was pulled on and translocated, whereas the SUMO domain itself was spared from unfolding by Msp1.

The earlier observation that PANN-Msp1 can unfold $Pex15^{NTD}$-mEOS-$Pex15^{CTD}$ at roughly the same rate at the mEOS-Pex15 substrate provides evidence that Pex15's core folded domain is not critical for initial binding to the motor. It remained unclear, however, whether the N- and C-terminal unstructured regions of Pex15 played a role in recognition by Msp1. I therefore tested PANN-Msp1's ability to unfold a non-Pex15 model substrate with a different Intrinsically Disordered Region (IDR): mEOS-IDR. This substrate includes the folded mEOS3.2 domain, followed by an IDR derived from CyclinB, with the amino acid sequence: AHGGKHTFNNENVSARLGGACSIAVQAPAQHTFNNENVSARLGGALSIAVQAPAQ. The CyclinB tail is 55 amino acids long and extends to 68 amino acids with the inclusion of a glycine-serine linker and a PSP cleavage scar. In comparison, Pex15's unstructured C-terminal segment is 62 amino acids (including PSP cleavage scar). In a mEOS unfoldase assay, PANN-Msp1 was able to unfold mEOS-IDR at approximately the same rate as mEOS-Pex15, demonstrating PANN-Msp1's promiscuity as an unfoldase (Fig. 4.13a). This promiscuity aligns with Msp1's cellular function of

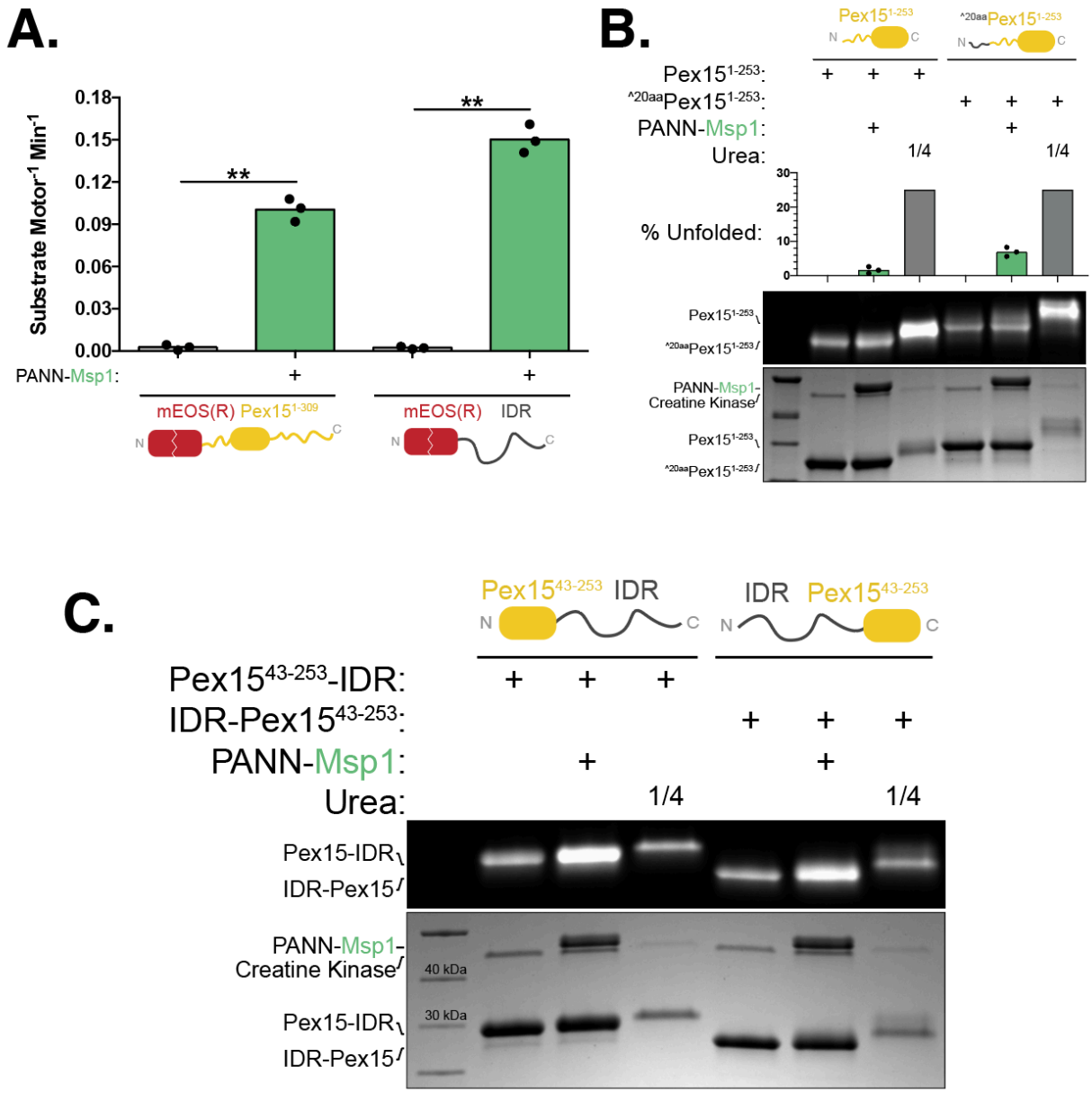


Fig. 4.13 : PANN-Msp1 is a promiscuous, bi-directional unfoldase. (A) mEOS unfoldase assay demonstrating that mEOS-Pex15 is unfolded at roughly the same rate as mEOS-IDR. (B) Extending Pex15$^{1\text{-}253}$'s N-terminus by 20 amino acids permits engagement and unfolding by PANN-Msp1. (C) Maleimide labeling unfoldase reactions reveal that PANN-Msp1 can unfold both Pex15$^{43\text{-}253}$-IDR and IDR-Pex15$^{43\text{-}253}$, indicating the motor does not have any inherent bias toward substrate processing directionality.

identifying and extracting a diverse variety of mislocalized proteins from the ER, peroxisome, and Golgi (Li & Zheng et al., 2019).

I have previously demonstrated that PANN-Msp1 is able to unfold Pex15$^{43\text{-}309}$ but not Pex15$^{1\text{-}253}$, and it remains unclear whether this difference is due to the specific sequence of Pex15's unstructured C-terminal region or simply the fact that it is longer. To determine if length was the underlying factor, I first extended Pex15$^{1\text{-}253}$'s N-terminus with 20 amino acids from the IDR sequence (to produce a substrate with ~62 amino acids of putative random coil at the N-terminus: $^{\text{20aa}}$Pex15$^{1\text{-}253}$). Notably, this

substrate's N-terminal unstructured region is now the same length as Pex15$^{43\text{-}309}$'s unstructured C-terminal region, which I have demonstrated to be sufficient for substrate engagement. This modest N-terminal extension was enough to be recognized by PANN-Msp1, and a small fraction of the protein was unfolded (Fig. 4.13b). The observation that only a small fraction of the protein was unfolded prompted reconsideration of whether Pex15's unstructured N-terminal region was indeed completely unstructured. To circumnavigate any undiscovered structural differences between the unstructured N- and C-terminal regions, I replaced Pex15's N-terminal unstructured region with the aforementioned full-length IDR (IDR-Pex15$^{43\text{-}253}$) and compared its unfolding with that of the reverse substrate (Pex15$^{43\text{-}253}$-IDR). PANN-Msp1 is able to unfold IDR-Pex15$^{43\text{-}253}$ to a similar extent as Pex15$^{43\text{-}253}$-IDR (Fig. 4.13c).

I sought to expand upon this investigation by analyzing the kinetics of N-terminal versus C-terminal unfolding, and created two similar model substrates, IDR-mEOS and mEOS-IDR, to assess potential differences in K_m or V_{max} by Michaelis-Menten analyses. While mEOS-IDR was easily unfolded by Msp1, IDR-mEOS was not (data not shown), pointing to important differences in mEOS unfoldability from the N- and C-termini. Studies of numerous other AAA+ protein unfoldases have shown that the local stability of a substrate near the engagement site, rather than its global stability, determine mechanical unfolding. Similarly here, Msp1 may experience a higher thermodynamic or kinetic unfolding barrier when pulling on mEOS from the N-terminus.

However, my data on Pex15 with N- or C-terminal extension clearly indicate that PANN-Msp1 can process its substrate in N-to-C or C-to-N direction, thus having no directional preference for translocation. While it is tempting to further analyze the exact length-requirements for Msp1 engagement, the fact that PANN resides at Msp1's N-terminus and may represent an artificial spacer for engagement with Msp1's pore loops means that any such length measurements would only reflect this *in vitro* system, and not be relevant *in vivo*.

4.3 Pex3 directly and specifically inhibits Msp1's unfoldase activity

Pex3's proposed mechanism of inhibition is that it directly interacts with Pex15, yet this interaction has never been described in the extensive existing peroxisome literature. My earlier experiments demonstrated MBP-Pex3's ability to inhibit PANN-Msp1 unfolding of Pex15^{NTD}-mEOS-Pex15^{CTD}, suggesting that Pex15's folded core domain is not responsible for any potential interaction with Pex3. Wondering whether Pex3 interacted with Pex15 at all, we tested its inhibition of mEOS-MBP-Pex15 (chosen over mEOS-Pex15 for its enhanced solubility and stability) and mEOS-IDR, and found that unfolding of both substrates was inhibited to a similar extent (Fig. 4.14a). To determine whether MBP-Pex3 inhibition was a general effect on AAA ATPases, I evaluated its ability to inhibit Pex1/Pex6 unfolding of Pex15, but observed no change in substrate unfolding (Fig. 4.14b). In summary, these data suggest not only a direct

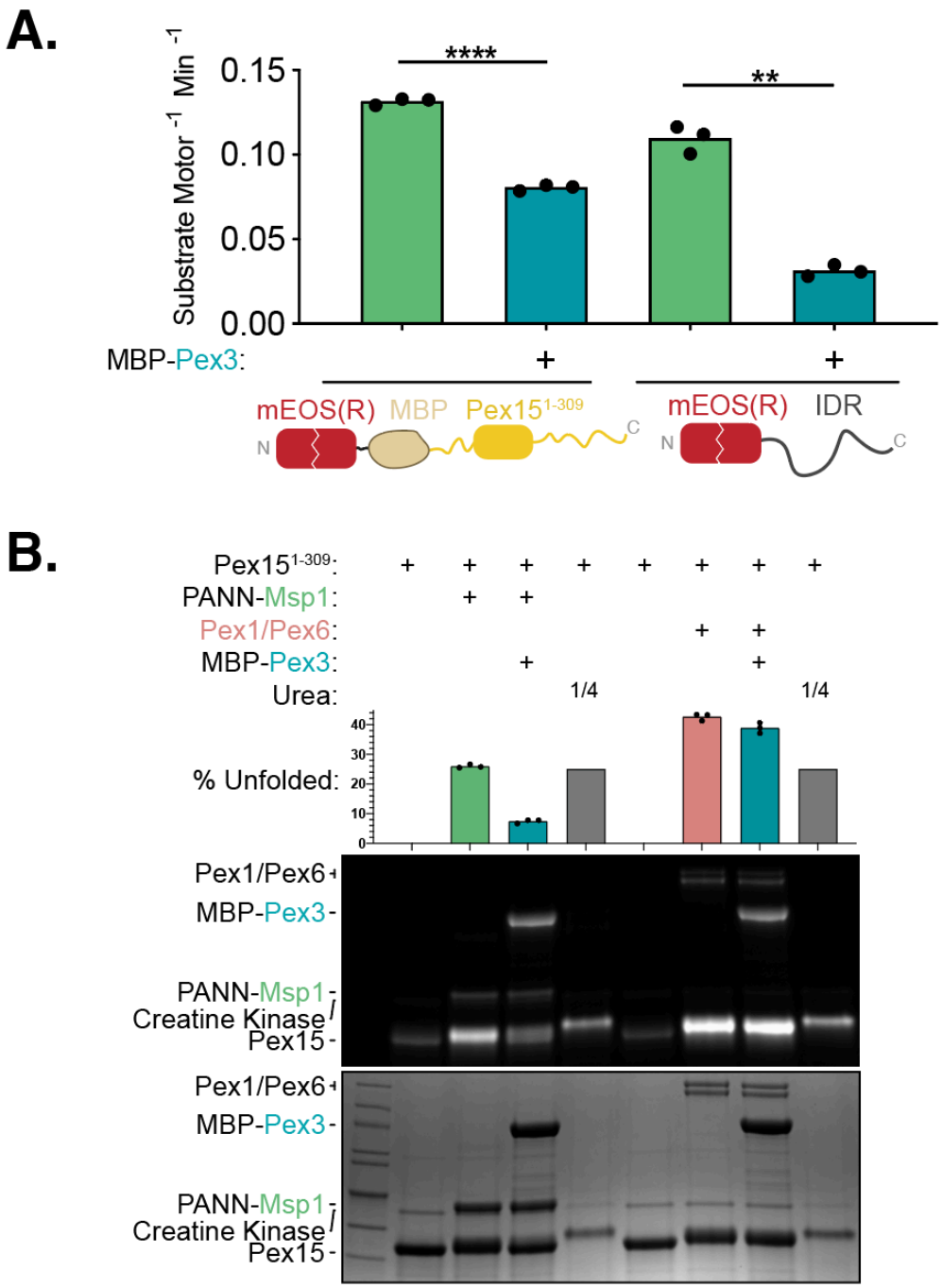


Fig. 4.14 : MBP-Pex3 inhibits unfolding of non-Pex15 substrates. (A) MBP-Pex3 inhibits PANN-Msp1 unfolding of both Pex15-based and non-Pex15 substrates, indicating that the motor is inhibited regardless of substrate. (B) MBP-Pex3 inhibits Msp1-mediated unfolding of Pex15$^{1-309}$, but not Pex1/Pex6-mediated unfolding of the same substrate, indicating inhibition specificity for Msp1.

interaction between MBP-Pex3 and PANN-Msp1, but also a degree of specificity by the inhibitor, given its potency against Msp1, but not Pex1/Pex6.

This mode of PANN-Msp1 inhibition suggests a direct interaction between MBP-Pex3 and the scaffolded motor. To test this further, I performed amylose pulldowns. Prior experience informed us that Pex15 tends to bind nonspecifically to many resin

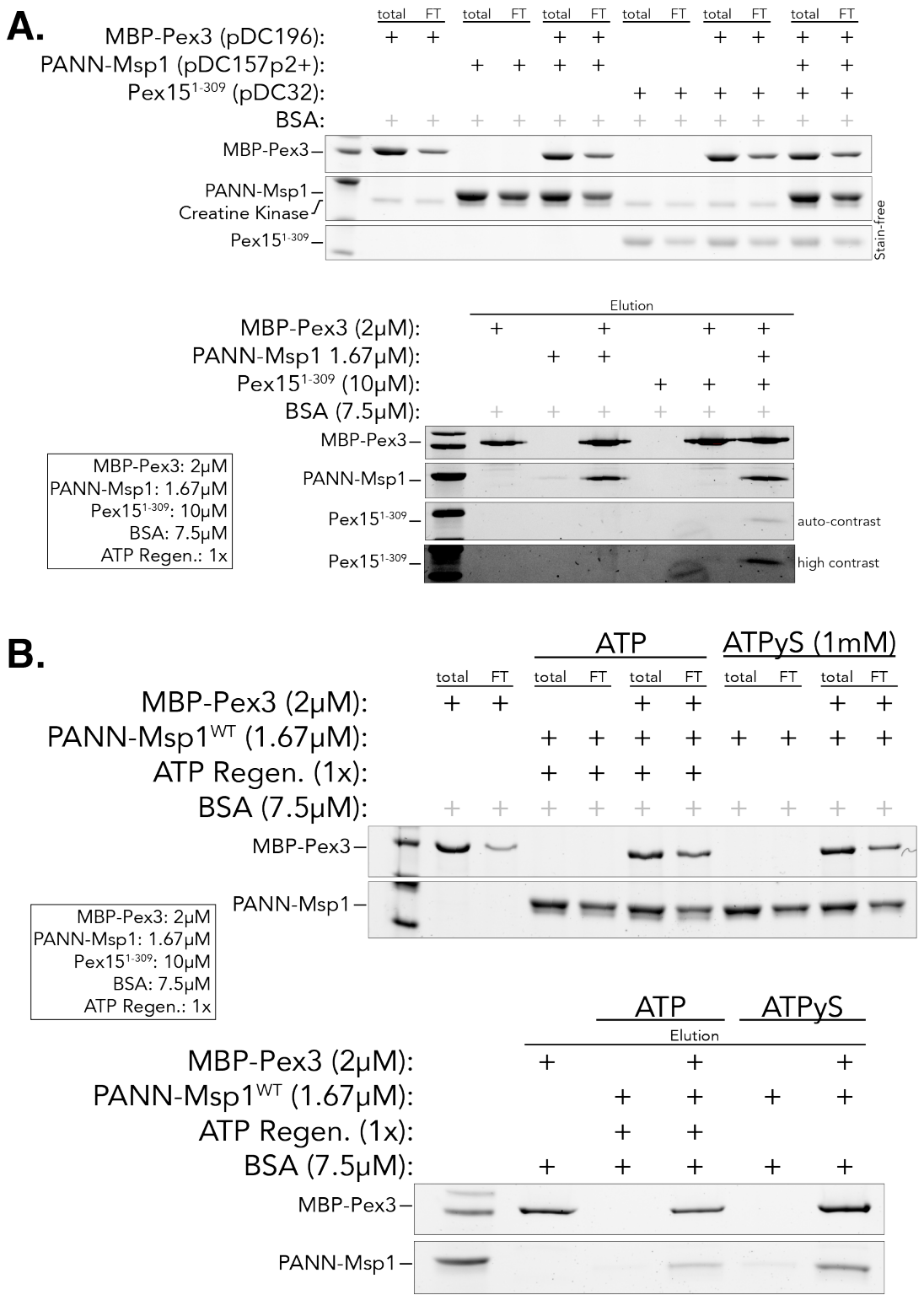


Fig. 4.15 : MBP-Pex3 interacts directly with PANN-Msp1, but not with Pex15. (A) MBP-Pex3 pull downs reveal no direct interaction between Pex3 and Pex15; however, PANN-Msp1 does directly interact with MBP-Pex3, and Pex15 can pull down with MBP-Pex3 in the presence of PANN-Msp1. Pull downs were performed in the presence of an ATP regeneration system. (B) MBP-Pex3 pull downs reveal that PANN-Msp1 interacts with MBP-Pex3 in the presence of both ATP and ATPyS.

types (Ni-NTA modified agarose beads, FLAG-antibody resin), but not amylose resin

(data not shown). I therefore performed pulldowns of MBP-Pex3 in the absence or presence of PANN-Msp1, Pex15$^{1-309}$, or both, and found that MBP-Pex3 interacts directly with PANN-Msp1 but not Pex15 (Fig. 4.15a). Importantly, PANN-Msp1 did not pull down with MBP alone, indicating a direct interaction mediated by the Pex3 domain. Furthermore, when MBP-Pex3 was incubated in the presence of PANN-Msp1 and Pex15, Pex15 could be found in the elution fraction, suggesting that PANN-Msp1 has the capacity to interact with both, MBP-Pex3 and Pex15 simultaneously (Fig. 4.15a). These experiments were performed in the presence of an ATP-regeneration system. In a separate experiment, it was determined that MBP-Pex3 can also bind PANN-Msp1 in the presence of 1 mM ATPγS instead of ATP (Fig. 4.15b).

My findings thus clearly point towards a direct interaction between MBP-Pex3 and PANN-Msp1, and contradict the previously proposed model of Pex3 binding and shielding the Pex15 substrate. The latter model was solely based on *in vivo* and cell-biological data, and it cannot be ruled out that a direct interaction between Pex3 and Pex15 is mediated by their transmembrane regions, which had been removed in our *in vitro* system.

Due to its natural positioning at the entrance to the Msp1 pore, it is possible that the PAN N-ring is involved in recruiting either the substrate, MBP-Pex3, or both to the PANN-Msp1 complex. The most obvious way to test this is to attempt unfoldase reactions without PANN, but this leaves the Msp1 as monomers that are unfolding incompetent. Given that monomeric $^{\Delta TMD}$Msp1 does not pull-down with MBP-Pex3 (data not shown), we can surmise that Pex3 binding requires oligomerization of Msp1. To assess the contribution of PANN to Pex15 or Pex3 binding, I performed pulldowns with PANN-Msp1 vs PANN alone at equimolar concentration as well as 5x higher concentration of PANN. Although PANN did appear to stick non-specifically to the

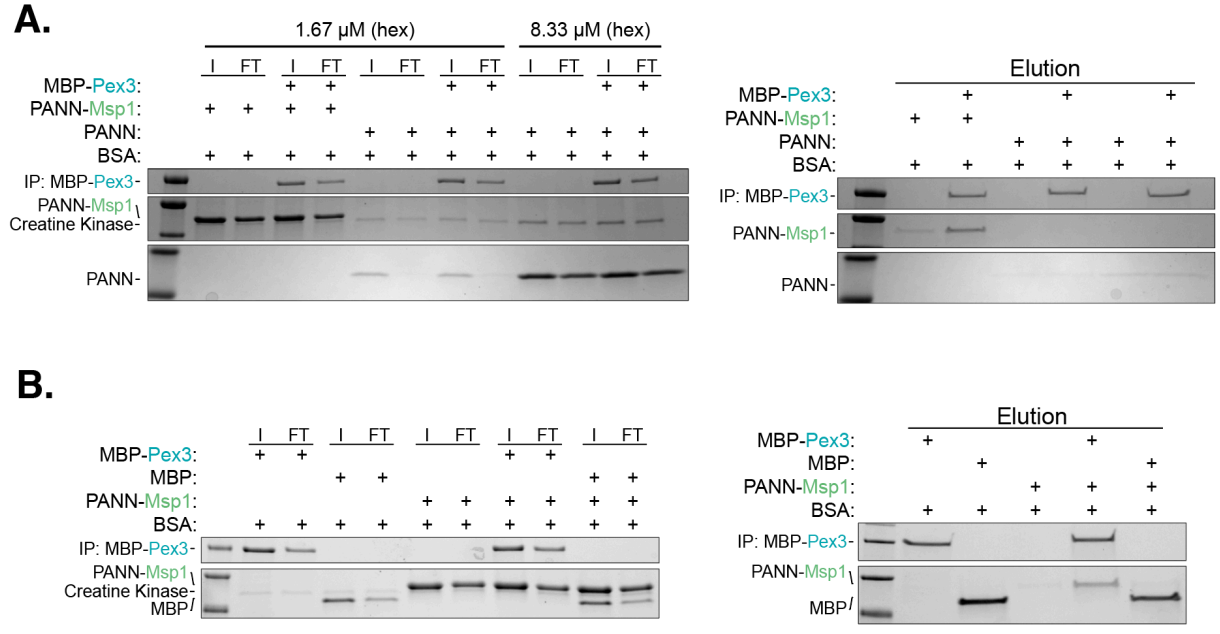


Fig. 4.16 : PANN does not mediate the interaction of MBP-Pex3 and PANN. (A) A pulldown of either PANN or PANN-Msp1 with MBP-Pex3 reveals only the Msp1-containing construct will interact with MBP-Pex3. (B) MBP alone is unable to pull down PANN-Msp1, implicating the Pex3 domain in interacting with the scaffolded motor.

amylose resin, I observed no increase in the presence of MBP-Pex3, suggesting no direction interaction (Fig. 4.16a). Conversely, MBP alone is incapable of pulling down PANN-Msp1, indicating that the interaction occurs via the Pex3 domain of the MBP-Pex3 construct (Fig. 4.16b).

The best way to assess whether the PANN domain is responsible for substrate or inhibitor recruitment to the PANN-Msp1 complex is to use an alternative domain for Msp1 hexamerization and test if the motor remains a potent unfoldase that is inhibited by MBP-Pex3. I fused ccHex in place of PANN to $^{\Delta TMD}$Msp1 and observed a similar higher-molecular weight elution off SEC, characteristic of hexamerization. I compared a series of variable linker lengths between the ccHex domain and Msp1, including glycine-serine linkers of 6-, 9-, 12-, 15-, and 18-amino acids. All proteins expressed and purified similarly well, but ccHex-Msp1 with a 9-amino acid GS linker possessed the best unfoldase activity, and this construct was used for all subsequent assays. Although ccHex-Msp1 eluted off SEC with a similar profile as PANN-Msp1, I sought to confirm its oligomeric states as individual hexamers by using negative stain EM. 2D-class averages and 3D reconstructions of ccHex-Msp1 in ATPγS indeed confirmed that the construct is hexameric and is the expected shape and size (Fig. 4.17a, with Ben LaFrance of the Nogales Lab). While the ccHex domain is poorly resolved (which is not unexpected given its small size), the Msp1 ATPase domain aligns well with the ATPase domain from the PANN-Msp1 3D reconstruction. Similar to PANN-Msp1, ccHex-Msp1 appeared to have an orientation bias on the grid, leading to predominantly top-down and bottom-up views of the hexamer, though ccHex-Msp1 was seen considerably more frequently at an angle 45° offset from a top-down view (Fig. 4.17b).

In a maleimide labeling unfoldase assay, MBP-Pex3 failed to inhibit ccHex-Msp1 unfolding of Pex15$^{1\text{-}309}$ (data not shown). While it is tempting to conclude that PANN is therefore responsible for MBP-Pex3's recruitment to the motor, it is noteworthy to reiterate the substantial architectural differences between PANN and ccHex. The narrow coiled-coil of ccHex right above the Msp1 ring is sufficient to drive hexamerization and allow ATP-hydrolysis-dependent substrate unfolding, but may preclude effectors, like Pex3, from binding or influence particular conformations of the motor. It may be that PANN fortuitously offers a suitable geometric arrangement of anchor points for $^{\Delta TMD}$Msp1's N-terminus, and, importantly, we found no positive evidence to suggest a direct interaction between PANN and MBP-Pex3.

Another hexamerization method we pursued was that of Hcp1, a large homohexameric protein complex from *P. aeruginosa* (Mougous et al., 2006). A recent study used this protein to hexamerize Vps4—another AAA ATPase in the meiotic clade—via fusion to the motor's C-terminus. Unlike ccHex, Hcp1's geometry is similar to PANN's, with the hexamer forming a donut-shaped structure. Notably, however, the Hcp1 hexamer is twice the mass of PANN: whereas each PANN monomer is only 8.7 kDa (52.2 kDa hexamer), a Hcp1 monomer is around 17.5 kDa (105 kDa hexamer).

Similar to ccHex-Msp1, I screened various linker lengths of 6-, 12-, and 18-amino acid glycine-serine linkers. The shortest linker length yielded the most protein with the highest ATPase activity, though this is only relative the other Hcp1 constructs, which all demonstrated poor solubility (data not shown). Using negative stain EM, I confirmed that Hcp1-Msp1 forms stable and discrete hexamers (Figure 4.18a, with Ben LaFrance in the Nogales Lab). Unlike PANN-Msp1 and ccHex-Msp1, Hcp1-Msp1 hexamers

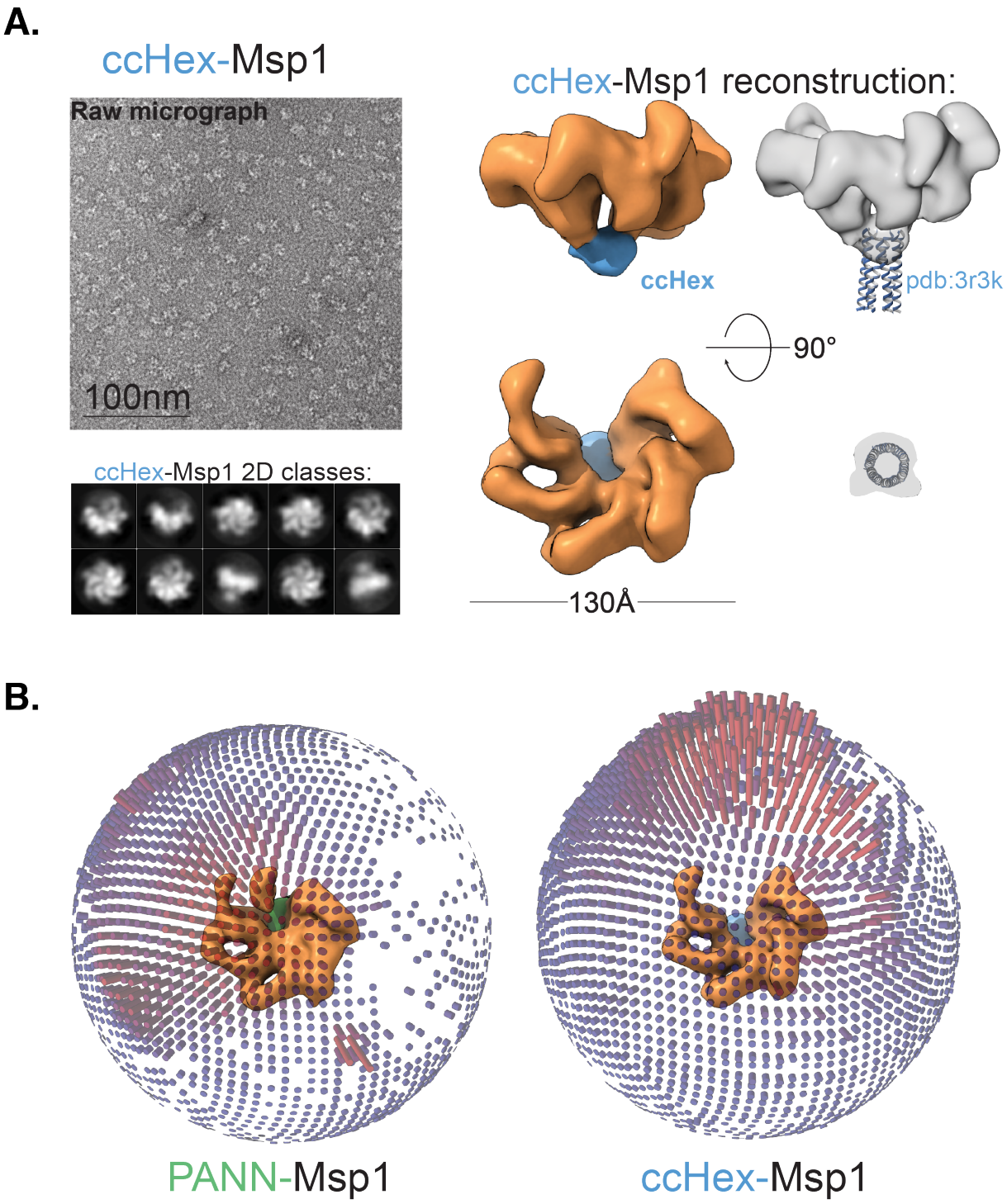


Fig. 4.17 : Negative-stain EM of ccHex-Msp1. (A) Negative stain EM sample micrograph, select 2D class averages, and 3D reconstructions of ccHex-Msp1. ccHex's crystal structure is docked into its corresponding density (3R3K). (B) Orientation bias of ccHex-Msp1 reveal a predominantly top-down view of the motor on grids. Data collected, processed, and refined by Ben LaFrance.

displayed a dramatically different orientation bias on negative stain grids, as they were seen almost exclusively from the side (Fig. 4.18b), yet the AAA ATPase domains still aligned very well across all three scaffolded complexes. In maleimide labeling reactions,

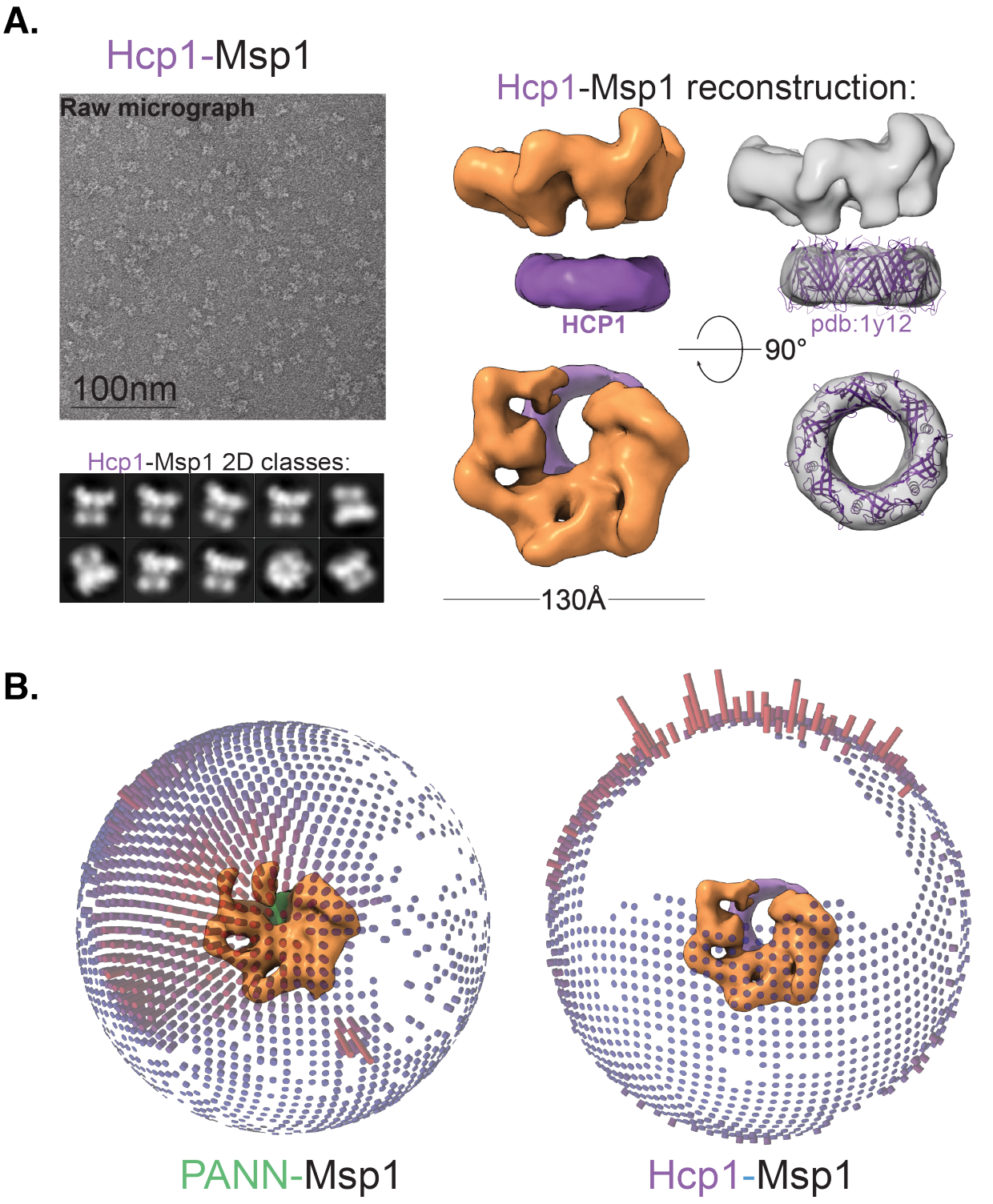


Fig. 4.18: Negative-stain EM of Hcp1-Msp1. (A) Negative stain EM sample micrograph, select 2D class averages, and 3D reconstructions of Hcp1-Msp1. Hcp1's structure is docked into its corresponding density (1Y12). (B) Orientation bias of Hcp1-Msp1 reveal a predominantly side view of the motor on grids. Data collected, processed, and refined by Ben LaFrance.

Hcp1-Msp1 was able to unfold Pex15$^{1\text{-}309}$ to a similar extent as PANN-Msp1 yet was not inhibited by the presence of MBP-Pex3 (data not shown). Again, it is possible that the

large structure presents a steric inhibition to Pex3 interaction, but we do not have aby data supporting or contradicting this.

Regarding the differential orientation on EM grids, it may be of interest to utilize the various hexamerization domains for future structural studies of other homohexameric complexes in order to obtain different views of the hexamers. Though we did not pursue this further, the strong agreement of AAA-domain conformations among the three hexamerization constructs would make it possible to combine all three datasets and, by masking out the scaffolds itself, obtain a 3D reconstruction of said complex with considerable orientation distribution.

4.4 Challenges of working with full-length Pex15

Although the soluble, hexamerized $^{\Delta TMD}$Msp1 constructs allowed me to get important insights into basic motor function and substrate recognition, the physiological function and mechanisms of Msp1 would be best studied in the context of a membrane, for instance using proteoliposomes. A major step in developing a proteoliposome-based extraction assay is the selection of an appropriate model substrate. While Wohlever et al. (2017) provided evidence that their poorly-described SUMO-TMD^{Sec22} substrate is extracted by Msp1, I have struggled to reproduce this assay, particularly the substrate-trapping step involving its interaction with GST-tagged SGTA, a TMD-binding eukaryotic chaperone. Furthermore, this substrate is not suited to answer the question whether unfolding is part of Msp1-mediated extraction. I sought to use Pex15 since it is already a putative *in vivo* substrate. Unfortunately, all attempts to purify detergent-solubilized full length Pex15 yielded a series of very high molecular-weight bands at the top of the gel and elution very early off SEC (data not shown), indicative of oligomerization or aggregation. Even fusions (MBP) and a solubilizing mutation (A317P) failed to coerce the protein into a well-behaved, monomeric state. I hypothesized that Pex15's folded core domain is sensitive to the presence of detergents, and that the laddering pattern on PAGE gels represented aggregates of unfolded protein.

In order to solubilize TMD-containing proteins for purification, sonicated and pre-cleared cell lysates were rocked in 1% DDM (~20mM) for at least 30 min at 4°C, and then maintained in 0.05% DDM throughout all washes and elutions (0.05% DDM, or 1 mM DDM, is roughly 5 times higher than DDM's CMC). I tested whether Pex15's structured core is sensitive to DDM by titrating the detergent before performing a maleimide labeling reaction (to assess the degree to which internal residues are exposed by the detergent). At a concentration corresponding to DDM's CMC (0.2 mM), Pex15 does not appear to be significantly destabilized; however, upon incubation in higher concentrations—such as the solubilization concentration of 20 mM or even the lower wash buffer concentration of 1 mM DDM—Pex15 displayed increased labeling by fluorescein-5-maleimide, indicating exposure of internal cysteines and destabilization of Pex15's core domain (Fig. 4.19c). Furthermore, incubation of Pex15 in 20 mM DDM and then dilution to 1 mM DDM (mimicking the shift from solubilizing concentrations of DDM

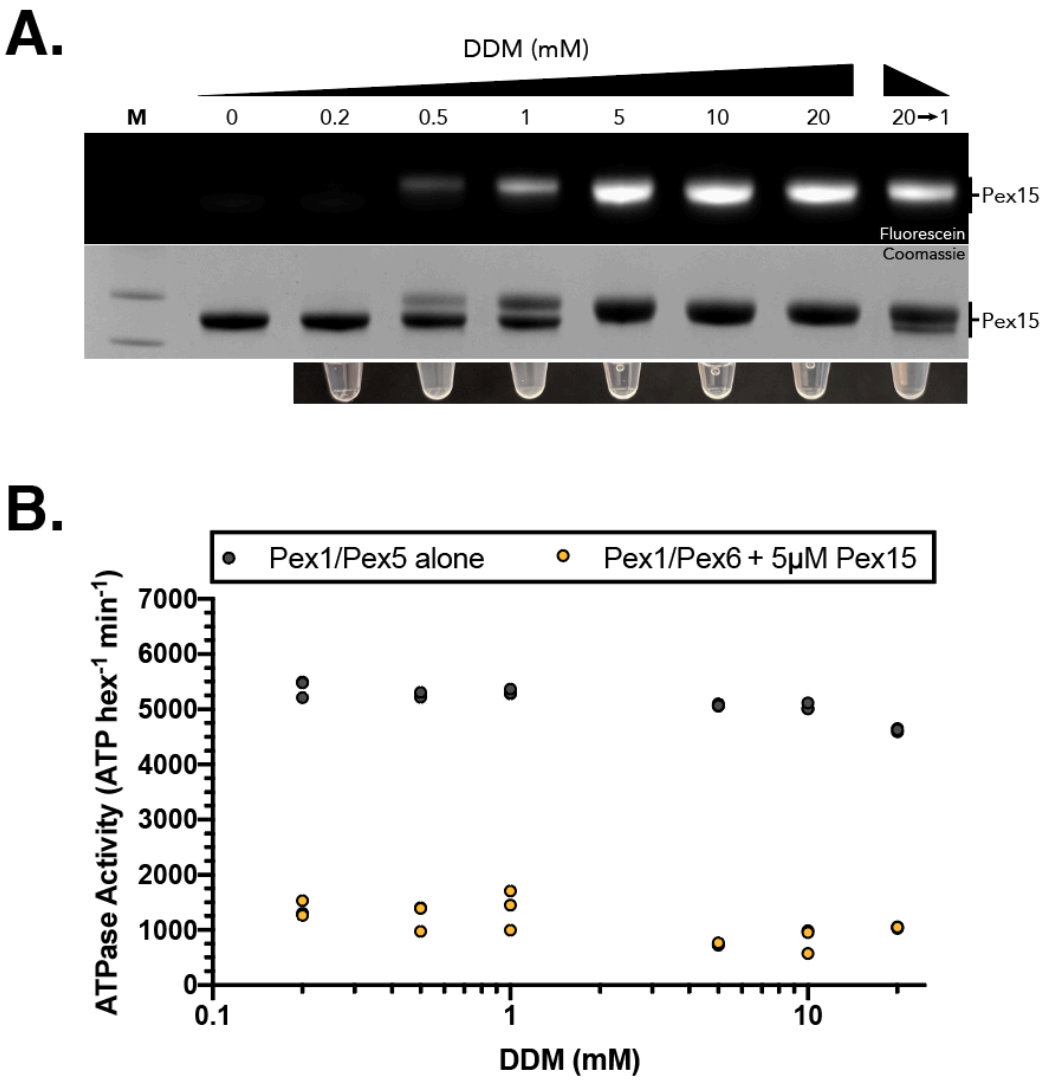


Fig. 4.19 : Pex15's folded domain is sensitive to detergent. (A) Pex15$^{1\text{-}309}$ incubated in increasing concentrations of DDM is labeled with F5M, suggesting gradual exposure of internal cysteines. (B) ATPase assays reveal that, even at high DDM concentrations where maleimide labelings suggest Pex15 is poorly structured (A), Pex15 is still capable of inhibiting Pex1/Pex6 ATPase activity (n = 3, technical replicates).

to washing concentrations of DDM), led to substantial labeling, suggesting that Pex15 struggles to properly refold in these conditions.

I decided to test whether these destabilized Pex15 constructs still generally maintained their ability to inhibit Pex1/Pex6 ATPase activity. Surprisingly, Pex1/Pex6 appears to have robust ATPase activity in various concentrations of DDM ranging from 0.2 mM to 20 mM. Furthermore, the addition of detergent-incubated Pex15 still inhibited Pex1/Pex6 ATPase activity, even at concentrations of DDM that appeared to unfold Pex15 (Fig. 4.19d). It is possible that DDM loosens Pex15's core structure without completely denaturing it; still, however, this may be responsible for the high-MW laddering we consistently see in our detergent-solubilized Pex15-TMD constructs.

I sought to develop a new method of producing TMD-anchored Pex15 proteoliposomes without exposing Pex15 to detergents at all. The idea was to purify Pex15$^{1\text{-}309}$-LPETGG as a soluble construct and append it to a TMD already

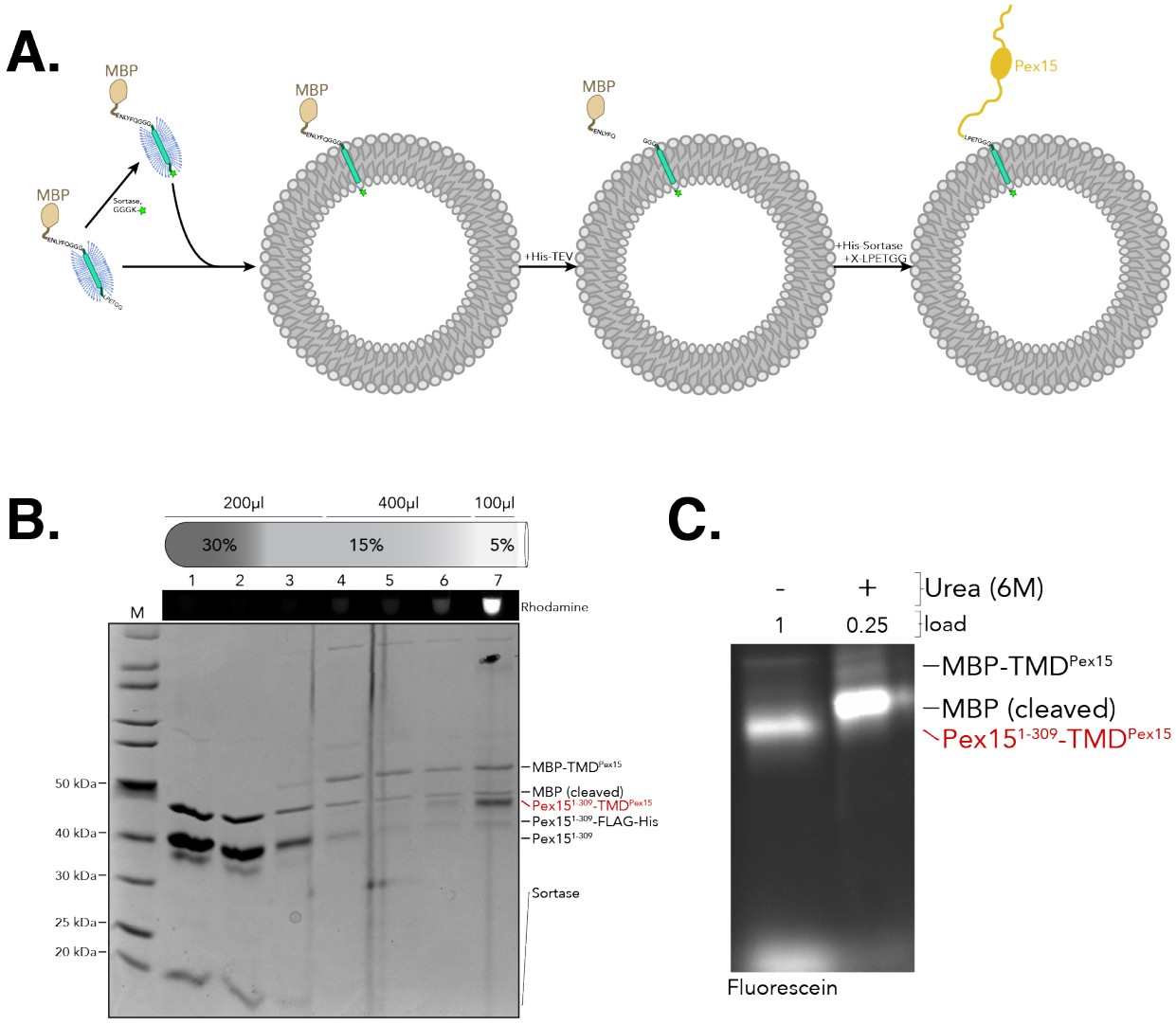


Fig. 4.20 : Using sortase to append soluble Pex15 to liposome-integrated TMDs. (A) Designed proteoliposome sortasing assay, wherein detergent-solubilized MBP-TMD^{Pex15} is incorporated into proteoliposomes before the MBP is cleaved using TEV, which leaves a GGG motif at the N-terminus of the liposome-incorporated GGG-TMD^{Pex15} peptide. Sortase is then used to append Pex15$^{1-309}$-LPETGG to the liposome-incorporated TMD^{Pex15} peptide. Following the sortase reaction, proteoliposomes are then floated to the top of a sucrose gradient to separate it from cleaved MBP and unreacted sortase and Pex15$^{1-309}$. (B) Liposome flotation assay following the steps depicted in (A) show that unreacted Pex15$^{1-309}$, sortase, and cleaved MBP all remain at the bottom of the sucrose gradient. Uncleaved full-length MBP-TMD^{Pex15} floats with the liposomes (band around 52 kDa). A novel species also floats with the proteoliposomes, appearing just below the cleaved MBP band on the gel. The molecular weight of this species corresponds to the desired reaction product Pex15$^{1-309}$-TMD^{Pex15}. (C) Incubation of this fraction in urea leads to an increase in F5M labeling, suggesting that the initial construct is properly-folded Pex15.

reconstituted in the liposome. I expressed and purified MBP-TEV-GG-TMD^{Sec22} and incorporated it into liposomes, then used TEV to cleave MBP off the liposome-anchored TMD, exposing a GGG-sortase-competent motif at the N-terminus of the GGG-TMD^{Sec22} peptide. I then used sortase to append the C-terminus of Pex15$^{1-309}$-LPETGG to the N-terminus of the TMD in the liposomes to form Pex15$^{1-309}$-TMD^{Sec22}. The liposomes contained 3% Ni-NTA lipids, and the Pex15$^{1-309}$-LPETGG, sortase, and TEV protease all had a His tag. His-tag interaction with Ni-NTA was expected to bring these components to the membrane and increase their local concentrations, thus potentially increasing reaction efficiency and yield. Conversely, coordination by the Ni-NTA groups

could restrict movements of each components and prevent their interaction in the proper orientation to allow the reaction to progress. I thus performed the sortase reaction twice, in the absence and presence of imidazole, but observed no difference, leaving unclear which conditions may be better suited. After the sortase reaction, I floated the liposomes to the top of a high-imidazole sucrose gradient to separate them from the other reactants (cleaved MBP, unreacted Pex15$^{1-309}$-LPETGG, and sortase; Fig. 4.20a). SDS-PAGE analysis of the ultracentrifugation fractions revealed that the reaction worked to some extent, as I could identify a band corresponding to Pex15$^{1-309}$-TMD that floats to the top of the gradient (Fig. 4.20b). That the liposomes were in the same fraction as Pex15$^{1-309}$-TMD was confirmed by the presence of 0.05% rhodamine-conjugated lipids, which were specifically included for tracking purposes (Fig. 4.20b, above the gel). Importantly, the final product band appears to float with the liposomes, while the other, non-incorporated reactants remain at the bottom of the gradient, unable to float with the proteoliposomes. To assess the folding state of the liposome-appended Pex15, I performed a maleimide labeling assay by pulse-labeling with F5M in the absence or presence of 6 M urea. The resulting gel showed a huge increase in fluorescence in the presence of urea, suggesting unfolding of previously structured Pex15 (Fig. 4.20c).

Although these proof-of-principle experiments confirmed that it is a possible to reconstitute Pex15-containing proteoliposomes without ever exposing Pex15 to detergent, the yield was low, and this method was unlikely to provide me with enough liposomes to reasonably conduct experiments. There are limits to boosting efficiency, as well: increasing sortase concentration without increasing the amount of the limiting component (the liposome-anchored TMD's) encourages sortase to use water in the hydrolysis reaction, which cleaves off the affinity tags and renders the resulting Pex15$^{1-309}$ incapable of participating in further reactions (the band corresponding to the product of this hydrolysis can be seen just below the Pex15$^{1-309}$-LPETGG-FLAG-His band in Fig. 4.20b). Increasing the concentration of Pex15$^{1-309}$-LPETGG-FLAG-His does not significantly increase product yield, but it does make it more difficult to separate the unreacted species from the proteoliposomes in the subsequence flotation assay. For potential future experiments using this approach, significant yield optimization will be necessary for procuring enough material for bulk biochemical analyses.

4.5 Concluding remarks on the functional reconstitution of Msp1

Until now, detailed analyses of Msp1 structure and mechanism have been hindered by the inability of the unanchored motor to form soluble hexamers *in vitro*. In my PhD thesis, I developed a strategy to form soluble, stable, and functional Msp1 hexamers by fusing $^{\Delta TMD}$Msp1 to several hexameric scaffolds. I have demonstrated not only that each hexameric Msp1 construct possesses ATPase activity, but that this activity remains stable with decreasing Msp1 concentrations, indicative of a stable hexamer assembly. Many AAA+ ATPases are unfoldases (Gardner et al., 2018; Gates et al., 2017; Ripstein et al., 2017; Olszewski et al., 2019; Martin et al., 2008), but Msp1's capabilities in this respect were unexplored. I showed for the first time that hexameric Msp1 is a robust bidirectional protein translocase and unfoldase. Msp1 requires a long

unstructured region for substrate engagement and is capable of unfolding a variety of folded domains from either N- or C-terminus. Msp1's ability to process substrates in both directions may be particularly relevant for its *in vivo* role in extracting TA proteins from membranes, and its promiscuity in engaging diverse IDRs is consistent with the multitude of mislocalized proteins it likely encounters in the peroxisome, ER, Golgi, and plasma membrane (Okreglak & Walter, 2014; Chen et al., 2014; Li & Zheng et al., 2019).

The importance of Msp1's pore loops had previously been established (Wohlever et al., 2017), but it remained unclear whether Msp1 unfolds substrates via processive threading or a tug-and-release mechanism. I showed that hexameric Msp1 can consecutively unfold linearly-fused folded domains by processive threading through the central pore. Based on my acquired data, I am unable to conclude whether the protein unfolding activity is essential for substrate extraction in either of Msp1's proposed roles as a quality control agent in TA mislocalization and the mitoCPR pathway. One possible mechanism of extraction involves hexameric Msp1 engaging an unstructured loop of a substrate near the TMD and initially translocating both strands, thus pulling the TMD out of the membrane. Thereafter, Msp1 can either continue translocation to unfold the structured domain from C- to N-terminus, or laterally release the substrate from the ring without further translocation. Alternatively, monomeric or dimeric Msp1 can bind and assemble around a substrate, and subsequently translocate in N-to-C-terminal direction to extract the TMD. While our reconstituted system allowed detailed mechanistic and enzymatic analyses of Msp1's unfoldase activity, a deeper understanding of Msp1's oligomerization dynamics may be necessary to explore the motor's extractase activity.

Lastly, *in vivo* studies suggested that Pex3 binds Pex15 at the peroxisome and shields it from extraction by Msp1. Our data agree that soluble MBP-Pex3$^{40\text{-}441}$ inhibits Pex15 unfolding by hexameric Msp1, but indicated a direct inhibition of Msp1 rather than substrate-shielding. We confirmed this model by demonstrating that MBP-Pex3 inhibits unfolding of a non-Pex15 substrate, mEOS-IDR. Importantly, MBP-Pex3 has no effect on the unfoldase activity of Pex1/Pex6, which functions in close proximity to Pex3 at the peroxisome membrane. Given this information, an outstanding question in the field remains about the function of Msp1 in the peroxisomal membrane, considering its general inhibition by Pex3. It is possible, however, that Pex15 and Pex3 are localized near the importomer, where Pex15 performs an important function in recruiting Pex1/Pex6 to the peroxisomal membrane, while Msp1 functions as a mislocalization quality control agent elsewhere in the peroxisome. Extensive studies will be necessary to further elucidate Msp1's role at the peroxisome, as well as the factors that govern the selection of substrates destined for extraction.

REFERENCES

Alencastre IS, Rodrigues TA, Grou CP, Fransen M, Sá-Miranda C, Azevedo JE. Mapping the cargo protein membrane translocation step into the PEX5 cycling pathway. *J Biol Chem* **284:40**, 27243-51 (2009)

Aviram N, Ast T, Costa EA, Arakel EC, Chuartzman SG, Jan CH, Haßdenteufel S, Dudek J, Jung M, Schorr S, Zimmermann R, Schwappach B, Weissman JS, Schuldiner M. The SND proteins constitute an alternative targeting route to the endoplasmic reticulum. *Nature* **540:7631**, 134-138 (2016)

Barnett P, Bottger G, Klein ATJ, Tabak HF, Distel B. The peroxisomal membrane protein Pex13p shows a novel mode of SH3 interaction. *EMBO J* **19:23**, 6382-6391 (2000)

Birschmann I, Stroobants AK, van den Berg M, Schäfer A, Rosenkrantz K, Kunau W-H, Tabak HF. Pex15p of *Saccharomyces cerevisiae* provides a molecular basis for recruitment of the AAA peroxin Pex6p to peroxisomal membranes. *Mol Biol Cell* **14:6**, 2226-2236 (2003)

Borgese N, Colombo S, Pedrazzini. The tale of tail-anchored proteins. *J Cell Biol* **161:6**, 1013-1019 (2003)

Bottger G, Barnett P, Klein ATJ, Kragt A, Tabak HF, Distel B. Saccharomyces cerevisiae PTS1 receptor Pex5p interacts with the SH3 domain of the peroxisomal membrane protein Pex13p in an unconventional, non-PXXP-related manner. *Mol Biol Cell* **11:11**, 3963-76 (2000)

Brambillasca S, Yabal M, Soffientini P, Stefanovic S, Makarow M, Hegde RS, Borgese N. Transmembrane topogenesis of a tail-anchored protein is modulated by membrane lipid composition. *EMBO J* **24:14**, 2533-42 (2005)

Brambillasca S, Yabal M, Makarow M, Borgese N. Unassisted translocation of large polypeptide domains across phospholipid bilayers. *J Cell Biol* **175:5**, 767-777 (2006)

Carroll E, Greene E, Martin A, Marqusee S. Ubiquitination modulates a protein energy landscape site-specifically with consequences for proteasomal degradation. *Nat Chem Biol*; doi: https://doi.org/10.1038/s41589-020-0556-3

Chang Y-W, Chuang Y-C, Ho Y-C, Cheng M-Y, Sun Y-J, Hsaio C-D, Wang C. Crystal structure of Get4-Get5 complex and its interactions with Sgt2, Get3, and Ydj1. *J Biol Chem* **285:13**, 9962-9970 (2010)

Chio US, Cho H, Shan S-O. Mechanisms of tail-anchored membrane protein targeting and insertion. *Annu Rev Cell Dev Biol* **33**, 417-438 (2017)

Chen Y, Pieuchot L, Loh RA, Yang J, Kari TM, Wong JY, Jedd G. Hydrophobic handoff for direct delivery of peroxisome tail-anchored proteins. *Nat Commun* **5**, 5790 (2014)

Chen YC, Umanah GK, Dephoure N, Andrabi SA, Gygi SP, Dawson TM, Dawson VL, Rutter J. Msp1/ATAD1 maintains mitochondrial function by facilitating the degradation of mislocalized tail-anchored proteins. *EMBO J* **33:14**, 1548-64 (2014)

Cichocki BA, Krumpe K, Vitali DG, Rapaport D. Pex19 is involved in importing dually targeted tail-anchored proteins to both mitochondria and peroxisomes. *Traffic* **19:10**, 770-785 (2018)

Ciniawsky S, Grimm I, Saffian D, Girzalsky W, Erdmann R, Wendler P. Molecular snapshots of the Pex1/6 AAA+ complex in action. *Nat Commun* **6:7331** (2015)

Denic V. A portrait of the GET pathway as a surprisingly complicated young man. *Trends Biochem Sci* **37:10**, 411-417 (2012)

Egea PF, Stroud RM, Walter P. Targeting proteins to membranes: structure of the signal recognition particle. *Curr Opin Struct Biol* **15:2**, 213-20 (2005)

El Magraoui F, Bäumer BE, Platta HW, Baumann JS, Girzalsky W, Erdmann R. The RING-type ubiquitin ligases Pex2p, Pex10p, and Pex12p form a heteromeric complex that displays enhanced activity in an ubiquitin conjugating enzyme-selective manner. *FEBS J* **279:11**, 2060-70 (2012)

Elgersma Y, Kwast L, van den Berg M, Snyder WB, Distel B, Subramani S, Tabak HF. Overexpression of Pex15p, a phosphorylated peroxisomal integral membrane protein required for peroxisome assembly in S. cerevisiae, causes proliferation of the endoplasmic reticulum membrane. *EMBO J* **16:24**, 7326-41 (1997)

Fodor K, Wolf J, Reglinski K, Passon DM, Lou Y, Schliebs W, Erdmann R, Wilmanns M. Ligand-induced compaction of the PEX5 receptor-binding cavity impacts protein import efficiency into peroxisomes. *Traffic* **16:1**, 85-98 (2015)

Frederick SE & Newcomb EH. Cytochemical localization of catalase in lead microbodies (peroxisomes). *J Cell Biol* **43:2**, 343-53 (1969)

Freitas MO, Francisco T, Rodrigues TA, Alencastre IS, Pinto MP, Grou CP, Carvalho AF, Fransen M, Sá-Miranda C, Azevedo JE. PEX5 protein binds monomeric catalase blocking its tetramerization and releases it upon binding the N-terminal domain of PEX14. *J Biol Chem* **286:47**, 40509-19 (2011)

Furuki S, Tamura S, Matsumoto N, Miyana N, Moser A, Moser HW, Fujiki Y. Mutations in the peroxin Pex26p responsible for peroxisome biogenesis disorders of complementation group 8 impair its stability, peroxisomal localization, and interaction with the Pex1p x Pex6p complex. *J Biol Chem* **281:3**, 1317-23 (2006)

Gardner BM, Castanzo DT, Chowdhury S, Stjepanovic G, Stefely MS, Hurley JH, Lander GC. The peroxisomal AAA-ATPase Pex1/Pex6 unfolds substrates by processive threading. *Nat Commun* **9:1**, 135 (2018)

Gardner BM, Chowdhury S, Lander GC, Martin A. The Pex1/Pex6 complex is a heterohexameric AAA+ motor with alternating and highly coordinated subunits. *J Mol Biol* **427:6PtB**, 1375-1388 (2015)

Gates SN, Yokom AL, Lin J, Jackrel ME, Rizo AN, Kendsersky NM, Buell CE, Sweeny EA, Mack KL, Chuang E, Torrente MP, Su M, Shorter J, Southworth DR. Ratchet-like polypeptide translocation mechanism of the AAA+ disaggregase Hsp104. *Science* **357(6348)**, 273-279 (2017)

Guna A & Hegde RS. Transmembrane domain recognition during membrane protein biogenesis and quality control. *Curr Biol* **28:8**, R498-R511 (2018)

Islinger M, Voelkl A, Fahimi HD, Schrader M. The peroxisome: an update on mysteries 2.0. *Histochem Cell Biol* **150:5**, 443-471 (2018)

Kalbfleisch T, Cambon A, Wattenberg BW. A bioinformatics approach to identifying tail-anchored proteins in the human genome. *Traffic* **8:12**, 1687-94 (2007)

Klouwer FCC, Berendse K, Ferdinandusse S, Wanders RJA, Engelen M, Poll-The BT. Zellweger spectrum disorders: clinical overview and management approach. *Orphanet J Rare Dis* **10:151** (2015)

Li L, Zheng J, Wu X, Jiang H. Mitochondrial AAA-ATPase Msp1 detects mislocalized tail-anchored proteins through a dual-recognition mechanism. *EMBO Rep* **20:4**, e46989 (2019)

Mariappan M, Li X, Stefanovic S, Sharma A, Mateja A, Keenan RJ, Hegde RS. A ribosome-associating factor chaperones tail-anchored membrane proteins. *Nature* **466:7310**, 1120-4 (2010)

Martin A, Baker TA, Sauer RT. Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. *Nat Struct Mol Biol* **15:11**, 1147-51 (2008)

Mateja A & Keenan RJ. A structural perspective on tail-anchored protein biogenesis by the GET pathway. *Curr Opin Struct Biol* **51**, 195-202 (2018)

Matsumoto N, Tamura S, Furuki S, Miyata N, Moser A, Shimozawa N, Moser HW, Suzuki Y, Kondo N, Fujiki Y. Mutations in the novel peroxin gene PEX26 that cause peroxisome-biogenesis disorders of complementation group 8 provide a genotype-phenotype correlation. *Am J Hum Genet* **73:2**, 233-36 (2003)

Miller JM & Enemark EJ. Fundamental characteristics of AAA+ protein family structure and function. *Archaea* **9294307** (2016)

Mougous JD, Cuff ME, Raunser S, Shen A, Zhou M, Gifford CA, Goodman AL, Joachimiak G, Ordoñez CL, Lory S, Walz T, Joachimiak A, Mekalonos JJ. A virulence locus of Pseudomonas aeruginosa encodes a protein secretion apparatus. *Science* **312(5779)**, 1526-30 (2006)

Nakai M, Endo T, Hase T, Matsubara H. Intramitochondrial protein sorting. Isolation and characterization of the yeast MSP1 gene which belongs to a novel family of putative ATPases. *J Biol Chem* **268:32**, 24262-9 (1993)

Noriega TR, Tsai A, Elvekrog MM, Petrov A, Neher SB, Chen J, Bradshaw N, Puglisi JD, Walter P. Signal recognition particle-ribosome binding is sensitive to nascent-chain length. *J Biol Chem* **289:28**, 19294-305 (2014)

Okreglak V & Walter P. The conserved AAA-ATPase Msp1 confers organelle specificity to tail-anchored proteins. *Proc Natl Acad Sci USA* **111:22**, 8019-24 (2014)

Oliveira ME, Reguenga C, Gouveia AM, Guimarães CP, Schliebs W, Kunau WH, Silva MT, Sá-Miranda C, Azevedo JE. Mammalian Pex14p: membrane topology and characterization of the Pex14p-Pex14p interaction. *Biochim Biophys Acta* **1567:1-2**, 13-22 (2002)

Olszewski MM, Williams C, Dong KC, Martin A. The Cdc48 unfoldase prepares well-folded protein substrates for degradation by the 26S proteasome. *Commun Biol* **2:29** (2019).

Otera H, Setoguchi K, Hamasaki M, Kumashiro T, Shimizu N, Fujiki Y. Peroxisomal targeting signal receptor Pex5p interacts with cargoes and import machinery components in a spatiotemporally differentiated manner: conserved Pex5p WXXXF/Y motifs are critical for matrix protein import. *Mol Cell Biol* **22:6**, 1639-55 (2002)

Platta HW, El Magraoui F, Bäumer BE, Schlee D, Girzalsky W, Erdmann R. Pex2 and Pex12 function as protein-ubiquitin ligases in peroxisomal protein import. *Mol Cell Biol* **29:20**, 5505-5516 (2009)

Rao M, Okreglak V, Chio US, Cho H, Walter P, Shan S-O. Multiple selection filters ensure accurate tail-anchored membrane protein targeting. *eLife* **5**, e21301 (2016)

Ripstein ZA, Huang R, Augustyniak R, Kay LE, Rubinstein JL. Structure of a AAA+ unfoldase in the process of unfolding substrate. *eLife* **6**, e25754 (2017)

Saidowsky J, Dodt G, Kirchberg K, Wegner A, Nastainczyk W, Kunau WH, Schliebs W. The di-aromatic pentapeptide repeats of the human peroxisome import receptor PEX5 are separate high affinity binding sites for the peroxisomal membrane protein PEX14. *J Biol Chem* **276:37**, 34524-9 (2001)

Schell-Steven A, Stein K, Amoros M, Landgraf C, Volkmer-Engert R, Rottensteiner H, Erdmann R. Identification of a novel, intraperoxisomal pex14-binding site in pex13: association of pex13 with the docking complex is essential for peroxisomal matrix protein import. *Mol Cell Biol* **25:8**, 3007-18 (2005)

Schuldiner M, Metz J, Schmid V, Denic V, Rakwalska M, Schmitt HD, Schwappach B, Weissman JS. The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* **134:4**, 634-645 (2008)

Schwartzkopff B, Platta HW, Hasan S, Girzalsky W, Erdmann R. Cysteine-specific ubiquitination protects the peroxisomal import receptor Pex5p against proteasomal degradation. *Biosci Rep* **35:3**, e00215 (2015)

Siegel V & Walter P. Each of the activities of signal recognition particle (SRP) is contained with a distinct domain: analysis of biochemical mutants of SRP. *Cell* **52:1**, 39-49 (1988)

Tamura S, Matsumoto N, Takeba R, Fujiki Y. AAA peroxins and their recruiter Pex26p modulate the interactions of peroxins involved in peroxisomal protein import. *J Biol Chem* **289:35**, 24336-46

Tan D, Blok NB, Rapoport TA, Walz T. Structures of the double-ring AAA ATPase Pex1-Pex6 involved in peroxisome biogenesis. *FEBS J* **283:6**, 986-92 (2016)

Walker JE, Saraste M, Runswick MJ, Gay NJ. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* **1:8**, 945-951

Wang F, Brown EC, Mak G, Zhuang J, Denic V. A chaperone cascade sorts proteins for posttranslational membrane insertion into the endoplasmic reticulum. *Mol Cell* **40:1**, 159-71 (2010)

Weidberg H & Amon A. MitoCPR-A surveillance pathway that protects mitochondria in response to protein import stress. *Science* **360:6385**, eaan4146 (2018)

Weir NR, Kamber RA, Martenson JS, Denic V. The AAA protein Msp1 mediates clearance of excess tail-anchored proteins from the peroxisomal membrane. *eLife* **6**, e28507 (2017)

Williams C, van den Berg M, Geers E, Distel B. Pex10p functions as an E3 ligase for the Ubc4-dependent ubiquitination of Pex5p. *Biochim Biophys Res Commun* **374:3**, 620-4 (2008)

Williams C, van den Berg M, Sprenger RR, Distel B. A conserved cysteine is essential for Pex4p-dependent ubiquitination of the peroxisomal receptor Pex5p. *J Biol Chem* **282:31**, 22534-43 (2007)

Wohlever ML, Mateja A, McGilvray PT, Day KJ, Keenan RJ. Msp1 is a membrane protein dislocase for tail-anchored proteins. *Mol Cell* **67:2**, 194-202 (2017)