Harnessing Biological Tools of Protein Transport and Catalysis

Thesis by

Camille Z. McAvoy

In Partial Fulfillment of the Requirements for

the Degree of

Doctor of Philosophy in Chemistry

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2018

Defended May 4, 2018

Camille Z. McAvoy
ORCID: 0000-0002-9828-1538

ACKNOWLEDGEMENTS

ABSTRACT

This work covers two projects related to protein structure and function. The first focuses on studies of chloroplast signal recognition particle 43 (cpSRP43), its interaction with substrate (the light-harvesting, chlorophyll-binding proteins, or LHCP), the role of conformational change in its activity, and the use of cpSRP43 as a tool for handling nonnative proteins. This work utilizes a variety of biochemical and biophysical approaches including light scattering and electron paramagnetic resonance to probe the structure-function relationship of cpSRP43. The second project entails the study of the C-C bond formation mechanism of nitrogenase, a biological nitrogen fixer found in soil microorganisms. Together these projects make for an interesting story of the medicinal and agricultural applications of basic biochemistry.

PUBLISHED CONTENT AND CONTRIBUTIONS

Camille McAvoy, Alex Siegel , Samantha Piszkiewicz, Emily Miaou, Mansen Yu, Thang Nguyen, Annie Moradian, Michael J. Sweredoski, Sonja Hess, Shu-ou Shan. (2018) Two Distinct Sites of client protein interaction with the chaperone cpSRP43. JBC. doi:10.1074/jbc.RA118.002215

Camille Z. McAvoy participated in the conception of the project, performed light scattering and fluorescence anisotropy experiments, prepared the data, made a portion of the figures, wrote and edited the manuscript.

Fu-Cheng Liang, Gerard Kroon, **Camille Z. McAvoy**, Chris Chi, Peter E. Wright, and Shu-ou Shan. (2016) Conformational dynamics of a membrane protein chaperone enable spatially regulated substrate capture and release. Proc Natl Acad Sci USA 13(12): E1615-E1624.
doi:10.1073/pnas.1524777113

Camille Z. McAvoy participated in the conception of the project, performed linker mutant experiments, prepared the data, made a portion of the figures, and contributed to the editing of the manuscript.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Chapter IV: Increasing membrane protein expression through co-expression with an ATP-independent chaperone

Chapter V: Characterizing biocatalyzed carbon-carbon bond formation mechanisms for industrial synthesis applications

Chapter VI: Nitrogenase Substrate Scope

NOMENCLATURE

Protein Terminology
Chaperone. Protein that assists in folding/unfolding or assembly/disassembly of other proteins (e.g. prevention of aggregation)

Enzyme. Protein that lowers the activation energy barrier for a chemical reaction.

Conformational change. Change in structure of protein.

cpSRP43. Chloroplast signal recognition particle 43 (chaperone).

cpSRP54. Chloroplast signal recognition particle 54 (conformational regulator).

LHCP. Light-harvesting, chlorophyll-binding proteins.

Alb3. Translocase in thylakoid membrane.

Nitrogenase. Nitrogen-fixing enzyme found in soil microorganisms.

Technique Terminology
NEM-Alkylation. N-ethylmaleimide (NEM) alkylation involves studying protein complexes by adding an alkylation site on a protein and then incubating with an alkylating agent such as NEM.

Light scattering assay. Measures turbidity (cloudiness) of solution to detect aggregates.

Fluorescence Anisotropy. Measures binding of fluorescently-tagged peptide to a protein of interest.

HDX-MS. Hydrogen-deuterium exchange mass spectrometry measures local mobility through exchange of hydrogen (water) for deuterium (D2O).

EPR. Electron Paramagnetic Resonance, uses a tagged protein (typically with an MTSSL tag) with unpaired electron to study local mobility.

Western Blot. A tagged protein is recognized by an antibody (such as anti-His or anti-FLAG) that is then recognized by a secondary antibody (such as anti-mouse or anti-rabbit) which is detected by developing a film.

ThT Fluorescence. Measures increase in fluorescence in beta-sheet rich aggregates.

I n t r o d u c t i o n

UNDERSTANDING PROTEIN TRANSPORT AND CATALYSIS IN THE BIOLOGICAL WORLD

This thesis is broken down into two sections, each with a unique, interdisciplinary approach at understanding processes in the biological world as a platform for engineering useful tools. The first section of this thesis focuses on chaperones and protein transport and is broken down into four chapters: the first chapter focuses on substrate-chaperone interaction with a membrane protein, the second on the cycle of substrate capture and release, the third on conformational transitions that drive this cycle, and the fourth on applying this small, ATP-independent chaperone to the problem of membrane protein expression. The second section goes on to take a combined organic and biophysical approach to characterize a carbon-carbon bond formation reaction between nitrogenase and methyl isonitrile for industrial synthesis applications. This section then goes on to review nitrogenase substrate scope and assess the potential role of hydrazine as an intermediate in nitrogen reduction. Together these works showcase the usefulness and variety of macromolecular machines for treating disease and greening industry through sustainable synthesis.

Chapter 1

A MEMBRANE PROTEIN CHAPERONE BINDS TO SUBSTRATE THROUGH TWO UNIQUE MODES

C.Z.M. Performed light scattering experiments, sedimentation experiments, and anisotropy experiments.

A.S. Performed cross-linking experiments.

S.P. Performed alkylation-protection experiments.

E.M. Was supervised by C.Z.M. and contributed to light scattering and anisotropy experiments.

T.X.N. Performed translocation experiments

A.M., M.J.S., and S.H. assisted with analysis of mass spec data.

ABSTRACT

Integral membrane proteins are prone to aggregation and misfolding in aqueous environments and therefore require binding by molecular chaperones during their biogenesis, targeting, and assembly. Chloroplast signal recognition particle 43 (cpSRP43) is an ATP-independent chaperone required for the biogenesis of the most abundant class of membrane proteins, the light-harvesting chlorophyll a/b-binding proteins (LHCPs). Previous work has shown that cpSRP43 specifically recognizes an L18 loop sequence, which is conserved among LHCP paralogs. However, how cpSRP43 protects the transmembrane domains (TMDs) of LHCP from aggregation has been unclear. In this work, alkylation-protection experiments identified the first two TMDs of Lhcb$_5$, a member of the LHCP family, and their intervening loop as major sites of protection by cpSRP43. Site-directed mutagenesis identified a class of cpSRP43 mutants that bind tightly to the L18 sequence but are defective in chaperoning full-length LHCP. These mutations mapped to hydrophobic surfaces on the bridging helix and the β-hairpins lining the ankyrin repeat motif (ARM) of cpSRP43, implying that these regions are potential sites for interaction with the client TMDs. Our results suggest a working model for client protein interactions involving this membrane protein chaperone.

Proper protein folding and localization are critical for cellular protein homeostasis, which is acutely challenged by the post-translational targeting of integral membrane proteins. Before arrival at the target membrane, nascent membrane proteins are highly prone to aggregation in the cytosol and other aqueous cellular compartments. Thus, effective molecular chaperones or chaperone networks are required to minimize improper exposure of the transmembrane domains (TMDs) on newly synthesized membrane proteins and to maintain them in soluble, translocation-competent conformations. Many chaperone proteins are linked to membrane protein biogenesis, including SecB, Skp, and SurA that protect bacterial outer membrane proteins, and Hsp70 homologues implicated in the import of mitochondrial and chloroplast proteins (1-6).

The light-harvesting chlorophyll a/b-binding proteins (LHCP) comprise over 50% of the protein content on the thylakoid membrane of green plants (7) and form the most bundant family of membrane proteins on earth. LHCPs are nuclear-encoded, initially synthesized in the cytosol, and imported across the chloroplast envelope in a largely unfolded state (8). In the chloroplast stroma, LHCPs are protected in a soluble 'transit complex' by the chloroplast signal recognition particle (cpSRP), comprised of the cpSRP43 and cpSRP54 protein subunits (9-12). LHCPs are then delivered to the Alb3 translocase and inserted into the thylakoid membrane via interactions between the GTPase domains of cpSRP54 and its receptor cpFtsY (9, 13-16). Previous work showed that the cpSRP43 subunit binds tightly to and quantitatively prevents the aggregation of multiple members of the LHCP family, and that it is necessary and sufficient to chaperone LHCPs (17). Although the chaperone activity of cpSRP43 is allosterically regulated by additional components in the cpSRP pathway, such as cpSRP54 and Alb3 (18-23), the simple composition of the cpSRP43-LHCP chaperone-client pair and the robustness of cpSRP43's chaperone activity make this pair an excellent system to understand the interaction and regulation of membrane protein chaperones.

A long-standing question about the cpSRP43-LHCP system is the mechanism by which the TMDs on the substrate proteins are protected by cpSRP43. The substrate-binding domain (SBD) of cpSRP43 is comprised of Ankyrin repeat motifs (ARMs), capped at the N-terminus by a chromodomain (CD1) and at the C-terminus by a bridging helix (BH) (24-26). Biochemical and crystallographic analyses showed that a conserved Tyr204 in the third ARM recognizes an FDPLGL motif in L18, a conserved 18-amino acid sequence between the second and third TMDs of LHCP (17, 24, 27-28). However, interaction with a soluble loop sequence is unlikely to be sufficient to protect LHCPs, which contain three TMDs, from aggregation. The ability of cpSRP43 to quantitatively prevent full-length LHCPs from aggregation is highly suggestive of additional interactions between cpSRP43 and the substrate TMDs. Nevertheless, deletion of individual TMDs in LHCP or replacement with the TMDs from unrelated membrane proteins did not severely disrupt the cpSRP43-LHCP interaction (28); this lack of specificity rendered the putative TMD interactions of cpSRP43 particularly challenging to demonstrate and identify. Although a crosslinking study identified three additional residues in TM3 of LHCP that can crosslink to cpSRP43 (29), the study was limited to TM3 and did not examine additional possible interaction with the remainder of LHCP, nor the sites on cpSRP43 that interact with the substrate protein. Further, conformational

rearrangements occur in the cpSRP43 SBD upon substrate binding (25), making it particularly challenging to define the interaction of this chaperone with the substrate TMDs.

In this work, we used a combination of alkylation/protection and site-directed mutagenesis studies to understand the interaction between cpSRP43 and its client protein. The results completely defined the sites of LHCP that are protected by cpSRP43 upon their interaction, and identified a set of mutant cpSRP43s that are specifically disrupted in their ability to chaperone LHCP without affecting binding of the L18 motif. These observations suggest potential TMD interaction sites on this membrane protein chaperone.

RESULTS

Mapping the interaction sites of cpSRP43 on Lhcb$_5$ through alkylation/protection patterns. To define the sites on Lhcb$_5$ involved in complex formation with cpSRP43, we tested the ability of cpSRP43 to protect individual residues in Lhcb$_5$ from alkylation by N-ethylmaleimide (NEM). To this end, we purified a set of Lhcb$_5$ variants in which single cysteines were engineered at every 5-10 residues across the entire sequence of Lhcb$_5$ (30). Lhcb$_5$ is a close homologue of LHCP that strongly depends on the cpSRP pathway for its biogenesis, and previous work showed that Lhcb$_5$ forms a tight complex with cpSRP43 with an apparent K_d value of ~10 nM; further, because Lhcb$_5$ has only one native cysteine it was the ideal substrate to use for making single cysteine mutants (28). To ensure that all Lhcb$_5$'s are bound by the chaperone, we tested the efficiency of complex formation between each single cysteine variant of Lhcb$_5$ and cpSRP43 (Table 1), and cpSRP43•Lhcb$_5$ complexes were assembled using concentrations of cpSRP43 that are saturating for each Lhcb$_5$ mutant. Further, the correlation of % soluble substrate in light scattering and % translocation in membrane targeting (Figure 2G) indicates that real client-chaperone complexes are being observed through these experiments. The previously identified FDPLGL interaction motif in the L18 sequence was not tested, as point mutations at any of these residues severely impaired complex formation with cpSRP43 (24, 28). The efficiency of NEM alkylation was quantified by intact mass spectrometry and provides a direct measure for the solvent accessibility of individual cysteine residues in Lhcb$_5$ (30). Previous work showed that LHCPs are imported into the chloroplast and maintained by cpSRP in a largely unfolded state (13). All the engineered single cysteines in Lhcb$_5$ denatured in 6M GdmCl were alkylated by NEM to >90% completion (30). Thus, comparison of the alkylation efficiency in the cpSRP43•Lhcb$_5$ complex to that of chemically denatured Lhcb$_5$ provides a measure for the degree to which individual residues in Lhcb$_5$ are protected by interaction with cpSRP43. The observed alkylation-protection pattern differs significantly from studies of Lhcb$_5$ aggregation in aqueous buffer (30), particularly in TM1, the end of TM2, and the C-terminus, indicating that the alkylation-protection patterns reported here reflect complex formation and not client aggregation. Representative data for the complexes of two Lhcb$_5$ variants, E156C and V135C, are shown in Figures 1A and 1B, respectively. Deconvolution and quantification of the m/z spectrum showed that for the E156C mutant of Lhcb$_5$, a single alkylated species was present after a 10 minute alkylation reaction, indicating that this site was fully alkylated and thus relatively solvent exposed in the complex (Figure 1A). By contrast, the m/z spectrum of the V135C mutant of Lhcb$_5$ contained both the unalkylated and alkylated species (Figure 1B), indicating that this site was protected by cpSRP43 and likely involved in Lhcb$_5$-cpSRP43 interaction.

The results of the alkylation-protection experiments for all the Lhcb$_5$ variants are summarized in Figure 1C. When the alkylation efficiencies at 10 minutes were mapped onto the sequence of Lhcb$_5$ (Figure 1D), several patterns were observed. Residues 70–143, which span the first two TMDs of Lhcb$_5$ and their intervening loop, were modestly to heavily protected, suggesting that they interact with cpSRP43 in the cpSRP43•Lhcb$_5$ complex. Residues 190–200, which form the C-terminal part of TM3, were also protected, consistent with the results of a previous crosslinking study (30). In contrast, residues in the N- and C-terminal portions of Lhcb$_5$ were either highly accessible or modestly protected. In addition, residues in the loop connecting TM2 and TM3 of Lhcb$_5$ were relatively accessible, consistent with crystallographic analysis showing that the L18 peptide is bound at a solvent accessible site on the surface of cpSRP43 (29). Together, these results show that cpSRP43 contacts and protects its substrates at all three TMDs, with extensive interactions formed with TM1, TM2, their intervening loop, and the C-terminus of TM3.

cpSRP43's substrate binding domain is highly sensitive to point mutations. Previous work established that CD1, the ARMs, and the BH together form a structural and functional unit that comprises the SBD of cpSRP43 (25). To establish which sites of cpSRP43 are involved in complex formation with LHCP, we mutated all

solvent-exposed hydrophobic residues (Leu, Ile, Val, and Trp) in the SBD, as well as additional residues on the β-hairpins of the ARMs and on the BH (highlighted in blue in Figure 2A). Each residue was mutated to cysteine in an otherwise cysteine-less cpSRP43 (C118A, C240S). Cysteine-less cpSRP43 is 5-fold reduced in binding and chaperoning LHCP compared to wildtype cpSRP43 because it is shifted to a less active conformation (25), but otherwise behaves analogously to wildtype cpSRP43.

We tested each single cysteine mutant of cpSRP43 for
its ability to bind and protect LHCP from aggregation using a well-established light scattering assay (Figure 2B and 2C). In this assay, LHCP denatured and solubilized in 8M urea was added to a solution containing either buffer, cysteine-less cpSRP43 (referred to as WT), or the mutant cpSRP43 of interest, and the turbidity of the solution was monitored in real time. In the absence of cpSRP43, LHCP aggregated extensively in aqueous solution (Figures 2B and 2C, *green* lines). The presence of 2.5 µM cysteine-less cpSRP43 prevented the aggregation of ~55% of LHCP (Figs. 2B and 2C, *black* lines); this cpSRP43 concentration thus provides the most sensitive condition to screen for mutant cpSRP43s defective in chaperone activity. Further, the results observed through this light scattering assay were mirrored by sedimentation experiments (Figure 2E) in which mutant cpSRP43s were incubated with LHCP in aqueous buffer and separated into pellet (aggregated) and soluble fractions.

We found that single point mutations of a surprisingly large number of residues in the cpSRP43 SBD compromised its chaperone activity. Of the 33 single cysteine mutants tested, only 10 mutants exhibited chaperone activities within three-fold of that of cysteine-less cpSRP43 (Fig. 2B). Six mutants exhibited 3-5 fold reductions in the solubilization of LHCP compared to cysteine-less cpSRP43, and chaperone activity was undetectable for 17 mutants (Fig. 2C and 2D). The residues whose mutation induced modest or severe defects in chaperone activity span almost an entire surface of the cpSRP43 SBD (Figure 4). Thus, cpSRP43 is highly sensitive to conservative perturbations in its SBD.

Two distinct classes of defective cpSRP43 SBD. mutants. The large number of surface residues that exhibit a mutational defect in substrate binding could arise from an extensive interaction surface of cpSRP43 with LHCP, or from perturbation of the global conformation of the SBD by these mutations. Recent NMR studies showed that the SBD of apo-cpSRP43 intrinsically samples active and inactive conformations with equal probability (25), supporting the possibility that the conformation of cpSRP43 could be particularly susceptible to mutations. To control for mutational effects on the global conformation of the SBD, we tested the ability of cpSRP43 mutants to bind the L18 recognition motif of LHCP. All of the chaperone-defective mutations examined here are located away from the crystallographically identified L18 binding site of cpSRP43 (Y204 highlighted in Figures 3D & E; 24); thus, a defect in L18 binding caused by these mutations most likely arises from a global structural defect of the SBD, rather than disruption of a direct interaction with L18.

The binding affinity of cpSRP43 for the L18 motif was measured based on the cpSRP43-induced increase in the fluorescence anisotropy of a HiLyte-Fluor488-conjugated L11 peptide, which represents the minimal sequence in L18 required for high affinity binding to cpSRP43. LHCP's L11 minimal binding motif was identified via a series of deletions in the L18 motif centered around the DPLG sequence, which has been shown to directly interact with cpSRP43 (24). We tested the ability of cpSRP43 to prevent aggregation and disassemble aggregates of each L18 truncation construct and the results are summarized in Supplementary Table 1. GSFDPLGLADD, or the L11 motif, was the minimal motif required for cpSRP43 to recognize and chaperone LHCP with comparable efficiency to wildtype LHCP. Representative equilibrium titrations for L11-cpSRP43 binding are shown in Figures 3A and B. The equilibrium dissociation constants (K_d) for L11 binding to WT and mutant cpSRP43's, derived from the equilibrium titrations, are summarized in Table 2. The anisotropy change of L11 induced by a subsaturating concentration (0.19 µM) of each mutant cpSRP43 relative to that of WT cpSRP43 are summarized in Figure 3C.

We found that mutation of a large number of residues affected the interaction of cpSRP43 with the L18 motif. Eight mutants bound the L11 peptide an order of magnitude more weakly than WT cpSRP43 (K_d ~ 0.8–3.5 µM; Figure 3B and Table 2, *yellow*), and three mutants exhibited ~100-fold weakened binding to L11 (K_d > 10 µM; Table 2, *red*). In contrast, twelve mutant cpSRP43s bound the L11 peptide with K_d values within 5-fold of that of WT cpSRP43 (K_d < 0.6 µM; Figure 3A and Table 2, *green*). We defined these twelve mutants as Class I. Mutants in this class either bind L11 with similar affinity compared to WT cpSRP43 but are defective in chaperoning LHCP, or the modest reductions in L11 binding observed with these mutants were insufficient to

account for their complete loss of chaperone activity towards LHCP. Thus, Class I mutants specifically disrupt the ability of cpSRP43 to protect the TMDs of LHCP from aggregation. The remainder of the chaperone-defective mutants are designated as Class II. These mutants were substantially disrupted in interaction with the L18 motif, although the mutations are located at sites away from the vicinity of the L18 binding site; thus, these mutations disrupt L18 binding through an allosteric effect, by altering the conformation of the SBD.

DISCUSSION

cpSRP43 is a small, ATP-independent chaperone with a SBD comprised mostly of an ankyrin repeat motif. At a size of 25 kDa, the cpSRP43 SBD is able to effectively chaperone multi-pass membrane protein substrates comparable to its own size, providing an intriguing system to understand how a small protein scaffold interacts with and provides protection for large client proteins. Our previous understanding of the cpSRP43-LHCP interaction was limited to recognition of the L18 loop sequence in LHCP by cpSRP43-Tyr204 (17, 24, 27). In this work, the results of alkylation-protection experiments showed that LHCP is extensively protected in its complex with cpSRP43; the regions of protection span all three TMDs of LHCP as well as the loop between TM1 and TM2, suggesting that these sites are either directly contacted by cpSRP43, or protected via intramolecular interactions induced by cpSRP43. Furthermore, site-directed mutagenesis studies of the cpSRP43 SBD identified two classes of mutant cpSRP43s: Class I, which disrupts cpSRP43's ability to protect LHCP from aggregation without affecting high-affinity recognition of L18; and Class II, which allosterically disrupts binding of the L18 motif. Together, these results provide evidence for much more extensive cpSRP43-client interactions, and suggest potential sites of cpSRP43 that interact with the TMDs of LHCP.

When mapped onto the crystal structure of the cpSRP43 SBD, the two classes of mutants are enriched in different regions of the cpSRP43 SBD, suggesting that different surfaces in the SBD mediate distinct functions. The residues that give rise to Class II mutants are primarily located on the helices in the ARM (Figure 4B, *purple*). As the sites of Class II mutations are away from the previously identified L18 binding site (Tyr204; highlighted in *blue* in Figure 4), we attribute their defects to disruption of the active conformation of the cpSRP43 SBD. In contrast, residues that give rise to Class I mutants are enriched in the bridging helix, the β-hairpins in the ARM, and a hydrophobic surface in CD1 (Figure 4A, *orange*), suggesting that these regions may either form or are in close vicinity to the TMD binding sites in cpSRP43. In support of this notion, Class I mutations cluster on or near major hydrophobic surfaces on the cpSRP43 SBD (Figure 4C). This model is also consistent with the general structural and functional features of ankyrin repeat proteins, which are formed by individual repeats of helix-loop-helix folds connected by β-hairpins. Structural, computational, and protein engineering studies showed that intra- and inter-repeat interactions between the helices allow ankyrin repeat proteins to cooperatively fold into concave L-shaped structures; in contrast, the loops and β-hairpins, which project outward from the helices, often form the recognition sites for interaction partners (31-32). We therefore propose that client recognition by cpSRP43 may occur analogously, with the L18 sequence specifically recognized by the loop in Ank3, while the TMDs in LHCP are bound to and protected by the β-hairpins and the hydrophobic surface on Ank4 and BH.

The large number of residues in the cpSRP43 SBD, at which a single conservative mutation away from the direct interaction site severely disrupts substrate binding and chaperone activity, is extraordinary. This behavior is characteristic of molecular systems that do not robustly attain the active conformation, such that small perturbations of any interaction are sufficient to drive the molecule or complex into the inactive state (33-34). Likewise, the sensitivity of cpSRP43 to point mutations strongly suggests that its SBD is at the tipping point of a cooperative conformational change required to attain a chaperone-active conformation. This model is consistent with recent NMR data that detected distinct conformational states in the cpSRP43 SBD and showed a conformational equilibrium close to unity between the active and inactive states (25). As observed previously, this property of cpSRP43 may be particularly useful in enabling regulation, allowing cpSRP43 to be readily switched 'on' and 'off' by its regulators in the stroma and at the target membrane, respectively (25). The precise nature of the conformational changes in cpSRP43 remains an important question for future investigations.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification. Single cysteine mutants of Lhcb$_5$ and cpSRP43 were constructed using the QuikChange Mutagenesis procedure (Stratagene) according to manufacturer's instructions. WT and mutant cpSRP43, LHCP, and Lhcb$_5$ were overexpressed and purified as previously described (35).

Alkylation. Purified cpSRP43 and single-cysteine mutants of Lhcb$_5$ were incubated in Buffer D (200 mM NaCl, 50 mM KHEPES, pH 7.5) and treated with 4 mM TCEP overnight. Four µM Lhcb$_5$ and cpSRP43 were incubated for 10 minutes to allow complex formation. Alkylation reactions were performed and analyzed as described previously (30). Briefly, the pre-formed Lhcb$_5$-cpSRP43 complex was treated with 100 µM N-ethylmaleimide (NEM) for 2 or 10 minutes and quenched with 50 mM DTT. Quenched samples were treated with 0.2% formic acid and analyzed on an LC-MSD SL 1100 series (Agilent) using a 2.1 x 150 mm Zorbax 300SB-C3 column (Agilent) and a gradient consisting of 0.2% formic acid as solvent A and 0.2% formic acid in acetonitrile (89.8%) and methanol (10%) as solvent B. Intact masses were determined in the single quadrupole. Chemstation software (Agilent) was used to deconvolute the masses and quantify the proteins. The reported accessibilities were calculated as a ratio of the alkylation efficiency of each cysteine mutant observed in the cpSRP43-Lhcb$_5$ complex relative to that of the same mutant denatured in 6M GdmCl. As a control, chaperone activity of cpSRP43 with each Lhcb$_5$ mutant was monitored (Table 1) to ensure efficient complex formation for all Lhcb$_5$ mutants with chaperone.

Chaperone Activity of cpSRP43. The ability of cpSRP43 to prevent LHCP aggregation was measured as described (17, 28). Aggregates were removed via ultracentrifugation in a TLA-100 rotor (Beckman Coulter) at 100,000 rpm for 30 min at 4°C prior to the experiment. Light scattering experiments were performed by addition of 3 µL of 50 µM LHCP denatured in 8M urea to 150 µL buffer D (50 mM KHEPES, pH 7.5, 200 mM NaCl) or 2.5 µM cpSRP43 in buffer D. Light scattering was monitored at 360 nm on a UV-Vis spectrometer (Beckman Coulter) over time until equilibrium was reached. The percentage of soluble LHCP (% soluble) at equilibrium was plotted for each single-cysteine cpSRP43 mutant.

Sedimentation. Aggregates of purified LHCP and cpSRP43 WT and mutants were removed via ultracentrifugation in a TLA-100 rotor (Beckman Coulter) at 100,000 rpm for 30 min at 4°C prior to the experiment. Sedimentation was performed by adding LHCP to either buffer D (50 mM KHEPES, pH 7.5, 200 mM NaCl) alone or buffer D pre-incubated with cpSRP43 to a final concentration of 2.5 µM LHCP and 9.375 µM cpSRP43. Chaperone was left to prevent LHCP aggregation for 10 minutes at room temperature, and then samples were spun down at 18,000 g for 30 minutes at 4°C in an Eppendorf microcentrifuge. The soluble fractions were removed from each tube after sedimentation and pellet fractions were resuspended in an equal volume of 8M urea. Equal volumes of pellet and soluble were run on gels that were Coomassie-stained and quantified using ImageJ.

Translocation. Translocation was performed by incubating in vitro translated ^{35}S-labeled LHCP that had been alone in aqueous buffer or prevented from aggregation or disaggregated with WT cpSRP43 with purified thylakoid membranes, 300 µM cpFtsY, 300 µM cpSRP54, 75 mM ATP, and 75 mM GTP at room temperature for 30 minutes with light. Gels were visualized via autoradiography and quantified via ImageJ.

Measurement of L18 binding. The L18 binding affinity of cpSRP43 was measured using L11 (GSFDPLGLADD), the minimal binding motif in L18, conjugated to HiLyte-Fluor488. Anisotropy measurements were conducted in buffer D (50 mM KHEPES, pH 7.5, 200 mM NaCl) on a Fluorolog 3-22 spectrofluorometer (Jobin Yvon), using 100 nM HiLyte-Fluor488-labeled L11 and varying concentrations of cpSRP43. Samples were excited at 500 nm and fluorescence anisotropy was recorded at 527 nm, as previously described (25). The data were fit to Eq 1,

$$A_{obs} = A_0 + \Delta A \frac{[L11]+[pro]+K_d-\sqrt{([L11]+[pro]+K_d)^2-4[L11][pro]}}{2[L11]} \qquad (1)$$

in which [pro] is cpSRP43 concentration, A_{obsd} is the observed anisotropy value, A_0 is the anisotropy value without cpSRP43, ΔA is the change in anisotropy at saturating cpSRP43 concentrations, and K_d is the equilibrium dissociation constant for the interaction of cpSRP43 with L11-HiLyte-Fluor488.

Conflict of Interest: The authors declare no competing financial interests.

Author Contributions: S.-o.S. conceived and coordinated the study and wrote the paper. C.Z.M. designed, performed, and analyzed the experiments in Fig 2-4 and Table 2, carried out analyses in Fig 1 and Table 1, and wrote the paper. S.P. and S.H. designed, performed, and analyzed the experiments in Fig 1 and Table 1. E.M. performed experiments in Fig 2-4 and Table 2. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Hardy, S. J.; Randall, L. L., A kinetic partitioning model of selective binding of nonnative proteins by the bacterial chaperone SecB. *Science* **1991,** *251* (4992), 439-43.
2. Randall, L. L.; Hardy, S. J., SecB, one small chaperone in the complex milieu of the cell. *Cell Mol Life Sci* **2002,** *59* (10), 1617-23.
3. Walton, T. A.; Sandoval, C. M.; Fowler, C. A.; Pardi, A.; Sousa, M. C., The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** *106* (6), 1772-7.
4. Hennecke, G.; Nolte, J.; Volkmer-Engert, R.; Schneider-Mergener, J.; Behrens, S., The periplasmic chaperone SurA exploits two features characteristic of integral outer membrane proteins for selective substrate recognition. *J Biol Chem* **2005,** *280* (25), 23540-8.
5. Thoma, J.; Burmann, B. M.; Hiller, S.; Muller, D. J., Impact of holdase chaperones Skp and SurA on the folding of beta-barrel outer-membrane proteins. *Nat Struct Mol Biol* **2015,** *22* (10), 795-802.
6. Mihara, K.; Omura, T., Cytoplasmic chaperones in precursor targeting to mitochondria: the role of MSF and hsp 70. *Trends in cell biology* **1996,** *6* (3), 104-8.
7. Green BR, S. A., Light regulation of nuclear-encoded thylakoid proteins. In *Molecular Genetics of Photosynthesis*, B Andersson, A. S., J Barber, Ed. IRL Press: Oxford, 1996; pp 75–103.
8. Ouyang, M.; Li, X.; Ma, J.; Chi, W.; Xiao, J.; Zou, M.; Chen, F.; Lu, C.; Zhang, L., LTD is a protein required for sorting light-harvesting chlorophyll-binding proteins to the chloroplast SRP pathway. *Nature communications* **2011,** *2*, 277.
9. Schuenemann, D.; Gupta, S.; Persello-Cartieaux, F.; Klimyuk, V. I.; Jones, J. D.; Nussaume, L.; Hoffman, N. E., A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. *Proceedings of the National Academy of Sciences of the United States of America* **1998,** *95* (17), 10312-6.
10. Eichacker, L. A.; Henry, R., Function of a chloroplast SRP in thylakoid protein export. *Biochimica et biophysica acta* **2001,** *1541* (1-2), 120-34.

11. Groves, M. R.; Mant, A.; Kuhn, A.; Koch, J.; Dubel, S.; Robinson, C.; Sinning, I., Functional characterization of recombinant chloroplast signal recognition particle. *J Biol Chem* **2001,** *276* (30), 27778-86.

12. Klimyuk, V. I.; Persello-Cartieaux, F.; Havaux, M.; Contard-David, P.; Schuenemann, D.; Meiherhoff, K.; Gouet, P.; Jones, J. D.; Hoffman, N. E.; Nussaume, L., A chromodomain protein encoded by the arabidopsis CAO gene is a plant-specific component of the chloroplast signal recognition particle pathway that is involved in LHCP targeting. *The Plant cell* **1999,** *11* (1), 87-99.

13. Schunemann, D., Structure and function of the chloroplast signal recognition particle. *Current genetics* **2004,** *44* (6), 295-304.

14. Franklin, A. E.; Hoffman, N. E., Characterization of a chloroplast homologue of the 54-kDa subunit of the signal recognition particle. *J Biol Chem* **1993,** *268* (29), 22175-80.

15. Li, X.; Henry, R.; Yuan, J.; Cline, K.; Hoffman, N. E., A chloroplast homologue of the signal recognition particle subunit SRP54 is involved in the posttranslational integration of a protein into thylakoid membranes. *Proceedings of the National Academy of Sciences of the United States of America* **1995,** *92* (9), 3789-93.

16. Tu, C. J.; Schuenemann, D.; Hoffman, N. E., Chloroplast FtsY, chloroplast signal recognition particle, and GTP are required to reconstitute the soluble phase of light-harvesting chlorophyll protein transport into thylakoid membranes. *J Biol Chem* **1999,** *274* (38), 27219-24.

17. Jaru-Ampornpan, P.; Shen, K.; Lam, V. Q.; Ali, M.; Doniach, S.; Jia, T. Z.; Shan, S. O., ATP-independent reversal of a membrane protein aggregate by a chloroplast SRP subunit. *Nat Struct Mol Biol* **2010,** *17* (6), 696-702.

18. Gao, F.; Kight, A. D.; Henderson, R.; Jayanthi, S.; Patel, P.; Murchison, M.; Sharma, P.; Goforth, R. L.; Kumar, T. K.; Henry, R. L.; Heyes, C. D., Regulation of Structural Dynamics within a Signal Recognition Particle Promotes Binding of Protein Targeting Substrates. *J Biol Chem* **2015,** *290* (25), 15462-74.

19. Falk, S.; Ravaud, S.; Koch, J.; Sinning, I., The C terminus of the Alb3 membrane insertase recruits cpSRP43 to the thylakoid membrane. *J Biol Chem* **2010,** *285* (8), 5954-62.

20. Falk, S.; Sinning, I., The C terminus of Alb3 interacts with the chromodomains 2 and 3 of cpSRP43. *J Biol Chem* **2010,** *285* (53), le25-6; author reply le26-8.

21. Lewis, N. E.; Marty, N. J.; Kathir, K. M.; Rajalingam, D.; Kight, A. D.; Daily, A.; Kumar, T. K.; Henry, R. L.; Goforth, R. L., A dynamic cpSRP43-Albino3 interaction mediates translocase regulation of chloroplast signal recognition particle (cpSRP)-targeting components. *J Biol Chem* **2010,** *285* (44), 34220-30.

22. Dunschede, B.; Bals, T.; Funke, S.; Schunemann, D., Interaction studies between the chloroplast signal recognition particle subunit cpSRP43 and the full-length translocase Alb3 reveal a membrane-embedded binding region in Alb3 protein. *J Biol Chem* **2011,** *286* (40), 35187-95.

23. Horn, A.; Hennig, J.; Ahmed, Y. L.; Stier, G.; Wild, K.; Sattler, M.; Sinning, I., Structural basis for cpSRP43 chromodomain selectivity and dynamics in Alb3 insertase interaction. *Nature communications* **2015,** *6*, 8875.

24. Stengel, K. F.; Holdermann, I.; Cain, P.; Robinson, C.; Wild, K.; Sinning, I., Structural basis for specific substrate recognition by the chloroplast signal recognition particle protein cpSRP43. *Science* **2008,** *321* (5886), 253-6.

25. Liang, F. C.; Kroon, G.; McAvoy, C. Z.; Chi, C.; Wright, P. E.; Shan, S. O., Conformational dynamics of a membrane protein chaperone enables spatially regulated substrate capture and

release. *Proceedings of the National Academy of Sciences of the United States of America* **2016,** *113* (12), E1615-24.

26. Sivaraja, V.; Kumar, T. K.; Leena, P. S.; Chang, A. N.; Vidya, C.; Goforth, R. L.; Rajalingam, D.; Arvind, K.; Ye, J. L.; Chou, J.; Henry, R.; Yu, C., Three-dimensional solution structures of the chromodomains of cpSRP43. *J Biol Chem* **2005,** *280* (50), 41465-71.

27. Jonas-Straube, E.; Hutin, C.; Hoffman, N. E.; Schunemann, D., Functional analysis of the protein-interacting domains of chloroplast SRP43. *J Biol Chem* **2001,** *276* (27), 24654-60.

28. Jaru-Ampornpan, P.; Liang, F. C.; Nisthal, A.; Nguyen, T. X.; Wang, P.; Shen, K.; Mayo, S. L.; Shan, S. O., Mechanism of an ATP-independent protein disaggregase: II. distinct molecular interactions drive multiple steps during aggregate disassembly. *J Biol Chem* **2013,** *288* (19), 13431-45.

29. Cain, P.; Holdermann, I.; Sinning, I.; Johnson, A. E.; Robinson, C., Binding of chloroplast signal recognition particle to a thylakoid membrane protein substrate in aqueous solution and delineation of the cpSRP43-substrate interaction domain. *The Biochemical journal* **2011,** *437* (1), 149-55.

30. Nguyen, T. X.; Jaru-Ampornpan, P.; Lam, V. Q.; Cao, P.; Piszkiewicz, S.; Hess, S.; Shan, S. O., Mechanism of an ATP-independent protein disaggregase: I. structure of a membrane protein aggregate reveals a mechanism of recognition by its chaperone. *J Biol Chem* **2013,** *288* (19), 13420-30.

31. Binz, H. K.; Amstutz, P.; Kohl, A.; Stumpp, M. T.; Briand, C.; Forrer, P.; Grutter, M. G.; Pluckthun, A., High-affinity binders selected from designed ankyrin repeat protein libraries. *Nature biotechnology* **2004,** *22* (5), 575-82.

32. Binz, H. K.; Stumpp, M. T.; Forrer, P.; Amstutz, P.; Pluckthun, A., Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. *Journal of molecular biology* **2003,** *332* (2), 489-503.

33. Marcos, E.; Crehuet, R.; Bahar, I., On the conservation of the slow conformational dynamics within the amino acid kinase family: NAGK the paradigm. *PLoS computational biology* **2010,** *6* (4), e1000738.

34. Kikani, C. K.; Antonysamy, S. A.; Bonanno, J. B.; Romero, R.; Zhang, F. F.; Russell, M.; Gheyi, T.; Iizuka, M.; Emtage, S.; Sauder, J. M.; Turk, B. E.; Burley, S. K.; Rutter, J., Structural bases of PAS domain-regulated kinase (PASK) activation in the absence of activation loop phosphorylation. *J Biol Chem* **2010,** *285* (52), 41034-43.

35. Jaru-Ampornpan, P.; Chandrasekar, S.; Shan, S. O., Efficient interaction between two GTPases allows the chloroplast SRP pathway to bypass the requirement for an SRP RNA. *Molecular biology of the cell* **2007,** *18* (7), 2636-45.

Table 1. Chaperone activity of cpSRP43 towards individual Lhcb$_5$ single-cysteine mutants (columns 2 and 3), and NEM alkylation efficiency at each cysteine in the presence of WT cpSRP43 (columns 4 and 5). [a]cpSRP43 chaperone activity was measured using 1 μM Lhcb$_5$ in the light scattering assay. % soluble Lhcb$_5$ at indicated cpSRP43 concentrations are reported. [b]Fraction of NEM modified Lhcb$_5$ in the cpSRP43•Lhcb$_5$ complex after indicated times of the alkylation reaction. All values represent mean ±S.D., with n = 2. N.D., not done.

Lhcb5 Construct	1 μM cpSRP43 Activity[a]	5 μM cpSRP43 Activity[a]	Alkylation[b] at 2'	Alkylation[b] at 10'
I40	ND	ND	1.0	1.0
G50	ND	ND	1.0	1.0
Q70	ND	ND	0.55 ± 0.02	0.67 ± 0.10
I75	84.48	ND	0.53 ± 0.05	0.66 ± 0.10
A85	102.15 ± 0.66	ND	0.33 ± 0.00	0.39 ± 0.02
P90	57.06 ± 31.69	ND	0.19	0.37 ± 0.14
C100 (WT)	95.59	ND	0.62 ± 0.16	0.66 ± 0.17
G110	ND	ND	0.29 ± 0.05	0.35 ± 0.03
N120	ND	ND	0.47 ± 0.04	0.52 ± 0.09
N125	74.62	ND	0.47 ± 0.00	0.56 ± 0.11
L130	68.18 ± 0.24	ND	0.28 ± 0.00	0.32 ± 0.01
V135	89.89 ± 14.78	ND	0.13	0.18 ± 0.01
G143	89.20	ND	0.34	0.45 ± 0.01
T150	95.84	ND	0.80	0.88 ± 0.03
E156	80.74	ND	1.0	1.0
D157	86.34	ND	1.0	1.0
G162	95.39	ND	1.0	1.0
A171	86.33	ND	1.0	1.0
L180	89.85	ND	0.91 ± 0.12	0.91 ± 0.13
I185	82.86 ± 5.02	ND	0.91 ± 0.12	0.86 ± 0.02
L190	59.79	95.04	0.44 ± 0.10	0.48 ± 0.11
M195	74.87 ± 23.58	95.50	0.17 ± 0.02	0.24 ± 0.05
I200	77.24	ND	ND	0.25 ± 0.10
V210	92.27	ND	0.81	0.91 ± 0.13
P220	88.82	ND	0.44	0.50 ± 0.17
A230	87.60	ND	0.56	0.64 ± 0.14

Table 2. K_d values for binding of HiLyte-conjugated L11 to individual cpSRP43 mutants. All cpSRP43 mutants shown in this table are derived from cysteine-less cpSRP43 (denoted as WT). Green highlights mutants that exhibit K_d values within 3-fold of WT cpSRP43; yellow highlights mutants exhibiting 3-5 fold defects in L11 binding; and red highlights mutants that are severely defective in L11 binding. * indicates that saturation could not be reached with the mutant during equilibrium titrations, and their K_d values for L11 were estimated assuming the same end point in the titration curve as cysteine-less cpSRP43.

Construct	Average K_d (µM)
WT	0.12 ± 0.029
L103C	0.14 ± 0.052
W106C	0.19 ± 0.027
W114C	3.03 ± 0.61
V124C	0.18 ± 0.058
V125C	0.78 ± 0.058
W133C	17.5* ± 3.65
V156C	0.55 ± 0.17
T162C	12.2* ± 0.44
F166C	3.53 ± 0.44
G193C	0.53 ± 0.16
G194C	2.38 ± 0.61
L195C	14.8* ± 5.26
T196C	0.87 ± 0.14
V222C	0.83 ± 0.22
L228C	0.083 ± 0.066
L231C	0.41 ± 0.029
I237C	0.17 ± 0.078
L238C	2.05 ± 0.19
R252C	0.42 ± 0.18
E256C	0.64 ± 0.017
I259C	0.94 ± 0.26
N260C	0.062 ± 0.0035
E263C	0.22 ± 0.063
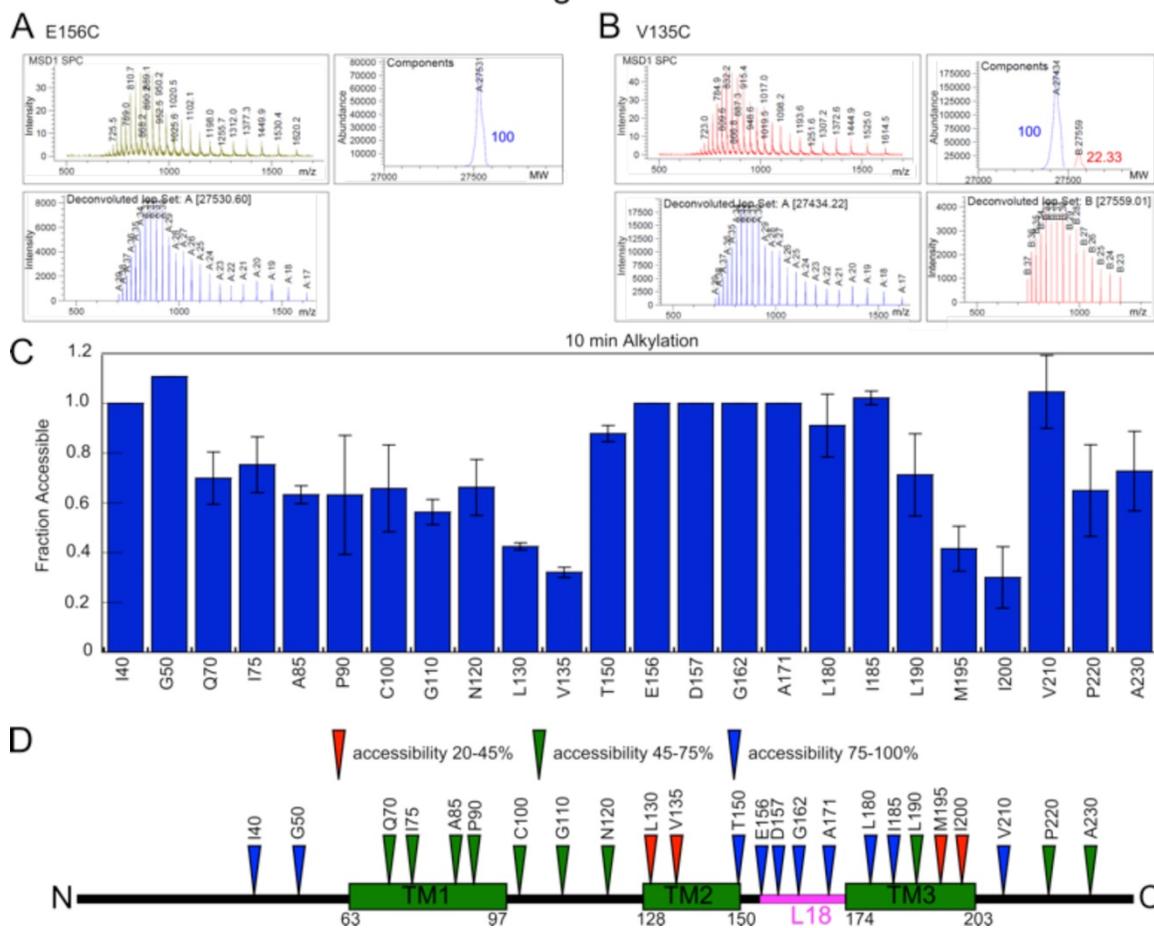
Figure 1. Alkylation pattern of Lhcb5 in the cpSRP43•Lhcb5 complex suggests potential sites of chaperone interaction on the substrate protein. (A, B) Mass spectrum (upper left), deconvolution (lower left), and component analysis (upper right) for a completely alkylated Lhcb5 residue, E156C (part A) and a partially alkylated Lhcb5 residue, V135C (part B). (C) Summary of the NEM alkylation efficiencies at individual sites in Lhcb5. Alkylation reactions were carried out for 10 minutes. For each engineered cysteine, 'Fraction Accessible' was calculated from the ratio of the fraction of alkylation in the cpSRP43•Lhcb5 complex relative to that of Lhcb5 dissolved in 6M GdmHCl. Error bars indicate S.E.M., with n = 2. (D) The alkylation pattern of Lhcb5 in complex with cpSRP43 is mapped onto the sequence of Lhcb5. Red, green, and blue triangles denote extensive protection (<45% alkylation), intermediate levels of protection (45-75% alkylation), and least protection (>75% alkylation), respectively.

Figure 2
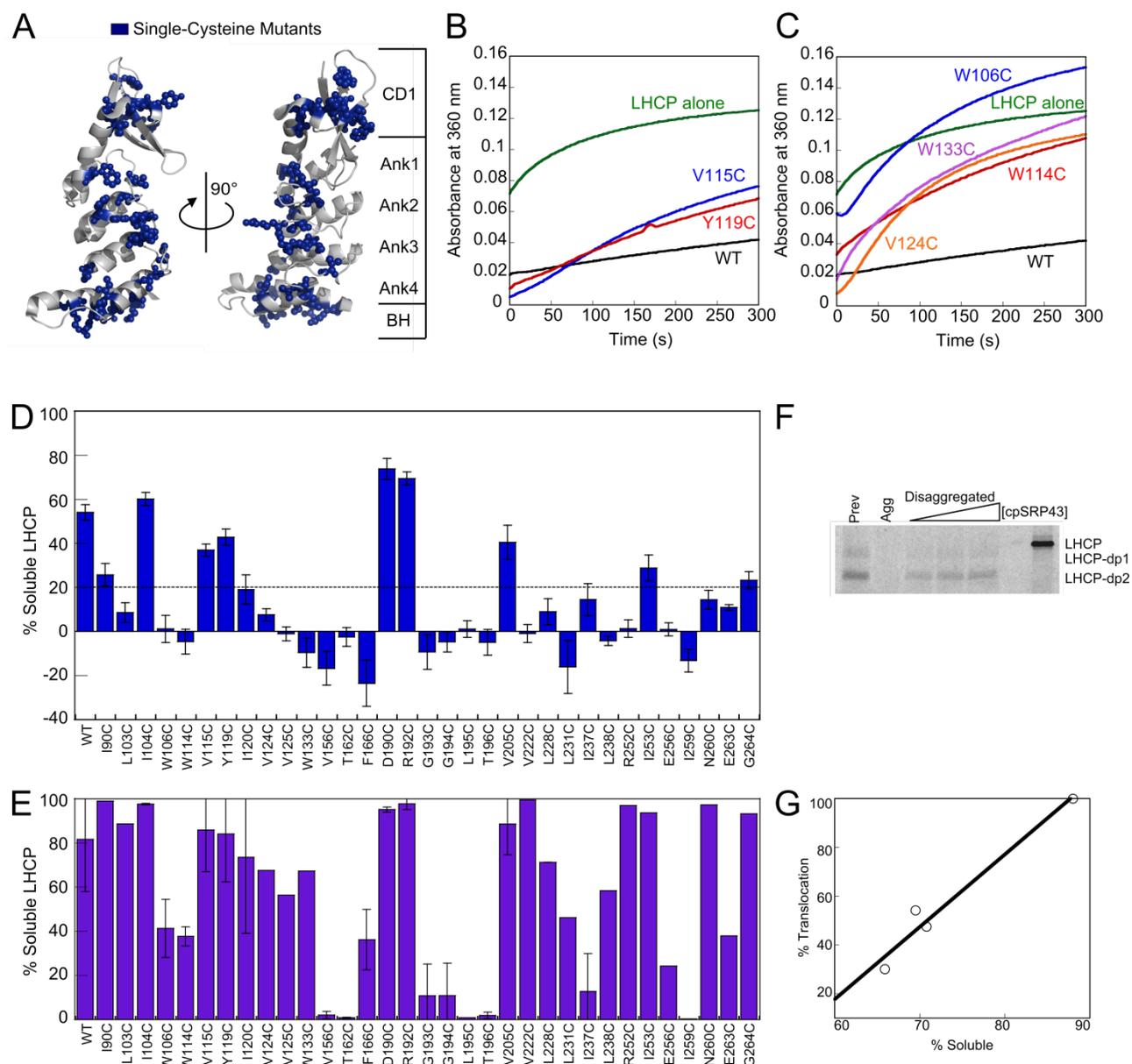
Figure 2. Single-cysteine mutants across the cpSRP43 SBD exhibit defects in chaperone activity. (A) Structure of cpSRP43 indicating all sites where cysteine mutations were made (blue). (B, C) Representative data showing the chaperone activity of neutral (part B) and defective (part C) cpSRP43 mutants. Light scattering traces are shown for LHCP diluted into aqueous buffer (green), into a solution containing cysteine-less WT (black), and into solutions containing the indicated cpSRP43 mutants. (D) Summary of the chaperone activity for all the single-cysteine mutants in the cpSRP43 SBD. Mutants exhibiting chaperone activity within three-fold that of cysteine-less cpSRP43 are considered neutral (above dashed line), whereas mutants with lower activity are considered defective (below dashed line). Error bars indicate S.E.M, with n = 3-13. (E) Sedimentation experimental results mirror trends observed through light scattering, supporting the classification of defective cpSRP43 single-cysteine mutants. Error bars indicate S.D., with n = 2. (F) Representative example of translocation with LHCP that has been prevented from aggregation or disaggregated by WT cpSRP43. (G) Correlation of % soluble LHCP with % translocation into thylakoid membrane as measured for WT cpSRP43.

Figure 3
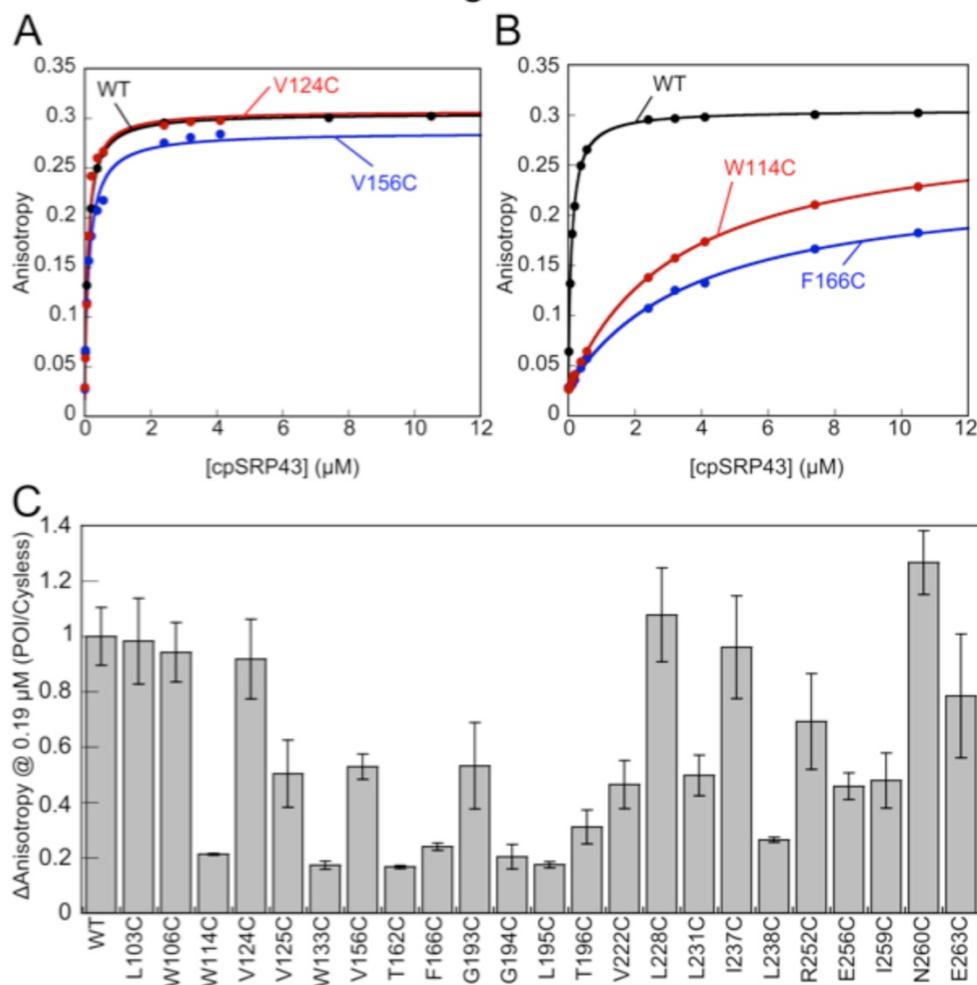
Figure 3. Characterization of the interaction of mutant cpSRP43s with the L18 motif. (A, B) Representative equilibrium titrations for the binding of WT and mutant cpSRP43s to Hylite-Fluor488 labeled L11. Representative data for cpSRP43 mutants that can bind L11 with high affinity are shown in part A, and those for mutants exhibiting weakened L11 binding are shown in part B. (C) Summary of the cpSRP43-induced changes in the fluorescence anisotropy of L11 at 0.19 μM, which is subsaturating for the binding of cysteine-less cpSRP43 to L11. The data for all mutants are normalized to that of cysteine-less cpSRP43 (denoted as WT).
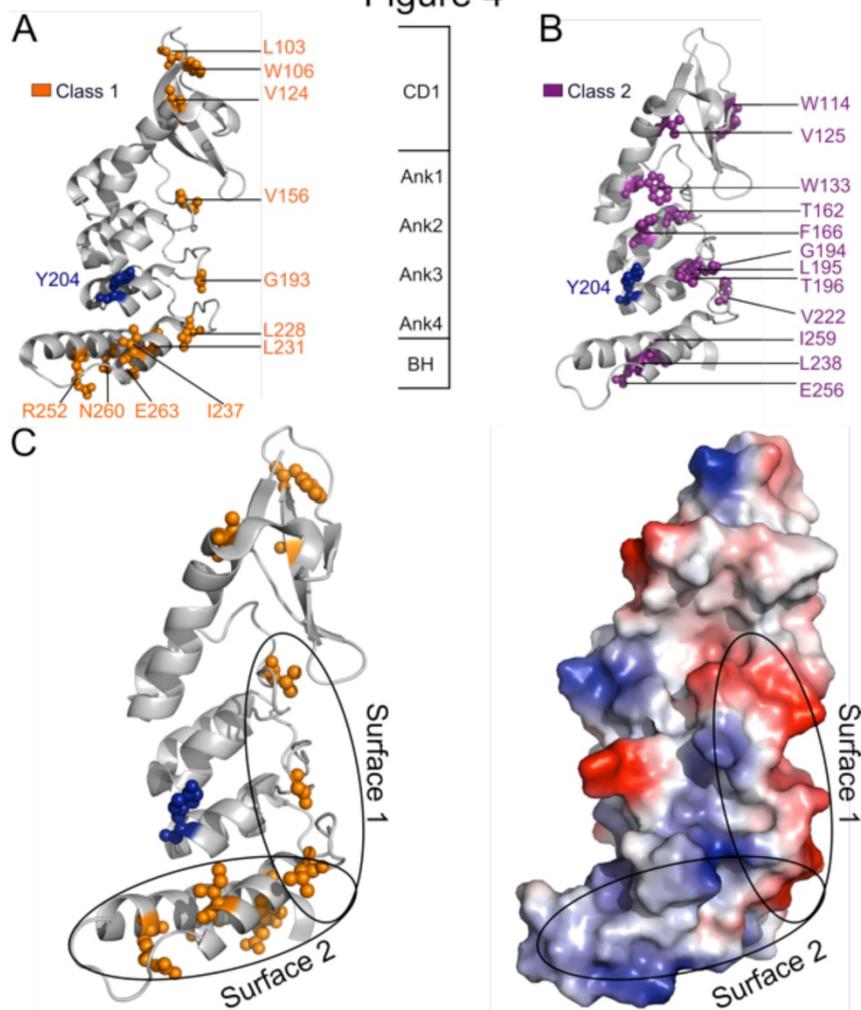
Figure 4. (A) Residues whose mutations led to defective chaperone activity for LHCP (measured in Figure 2D) but did not disrupt L18 binding (measured in Figure 3) are mapped onto the crystal structure of the cpSRP43's SBD (PDB ID 3DEP). These mutants are categorized as Class I and colored in orange. (B) Residues whose mutations disrupted both cpSRP43's chaperone activity and its interaction with the L18 motif are mapped onto the crystal structure of the cpSRP43 SBD. These mutants are categorized as Class II and colored in purple. (C) A putative model for the interaction surfaces of cpSRP43 with LHCP, with Tyr204 (blue) interacting with the L18 sequence, and the hydrophobic surfaces formed by Ank4, BH and the β-hairpins along the ankyrin repeat domain protecting the TMDs of LHCP.
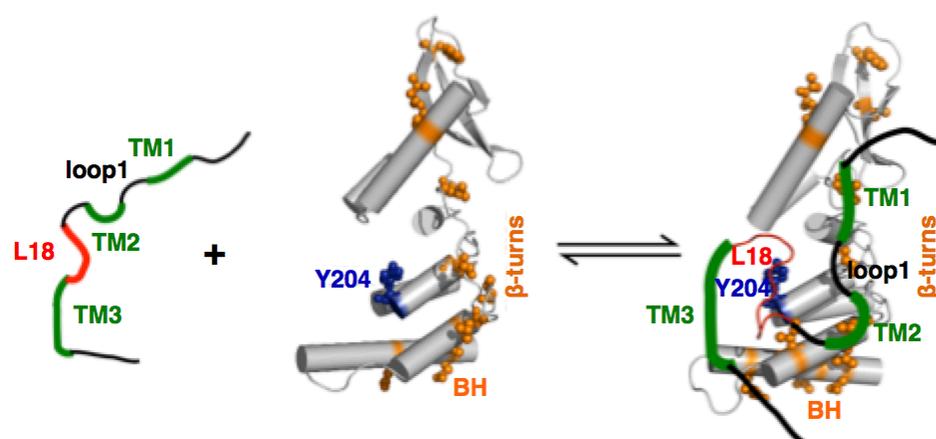
Figure 5. Final model of cpSRP43-Lhcb5 complex. This model shows sites of TM interaction based on the findings of this work.

Supplemental Data

Supplemental Table 1. Summary of L18 truncation experiments that identified L11 as the minimal motif of L18 necessary and sufficient for binding to cpSRP43.

LHCP Construct	L18 Motif	Prevention	Disaggregation
Wild type	VDPLYPGGSFDPLGLADD	Full	Full
Construct 1	YPGGSFDPLGLADD	Full	Full
L11	GSFDPLGLADD	Full	Full
Construct 3	GGSFDPLGL	None	None
Construct 4	DPLGLADD	None	None
Construct 5	FDPLGL	None	None
Construct 6	DPLG	None	None
Construct 7	GSFDPLGLAD	Partial	Partial
Construct 8	GSFDPLGLA	Partial	Partial
Construct 9	SFDPLGLADD	Partial	Partial

C h a p t e r   2

CONFORMATIONAL DYNAMICS OF A MEMBRANE PROTEIN CHAPERONE ENABLE SPATIALLY REGULATED SUBSTRATE CAPTURE AND RELEASE

Work published in:

Fu-Cheng Liang, Gerard Kroon, **Camille Z. McAvoy**, Chris Chi, Peter E. Wright, and Shu-ou Shan. (2016) Conformational dynamics of a membrane protein chaperone enable spatially regulated substrate capture and release. *PNAS*.

F-C. L. Performed light scattering, anisotropy, and NMR experiments.

G.K. Assisted with NMR experiments.

C.Z.M. Performed light scattering experiments.

C.C. Assisted with analysis of NMR data.

Abstract

Membrane protein biogenesis poses enormous challenges to cellular protein homeostasis and requires effective molecular chaperones. Compared with chaperones that promote soluble protein folding, membrane protein chaperones require tight spatiotemporal coordination of their substrate binding and release cycles. Here we define the chaperone cycle for cpSRP43, which protects the largest family of membrane proteins, the light harvesting chlorophyll a/b-binding proteins (LHCPs), during their delivery. Biochemical and NMR analyses demonstrate that cpSRP43 samples three distinct conformations. The stromal factor cpSRP54 drives cpSRP43 to the active state, allowing it to tightly bind substrate in the aqueous compartment. Bidentate interactions with the Alb3 translocase drive cpSRP43 to a partially inactive state, triggering selective release of LHCP's transmembrane domains in a productive unloading complex at the membrane. Our work demonstrates how the intrinsic conformational dynamics of a chaperone enables spatially coordinated substrate capture and release, which may be general to other ATP-independent chaperone systems.

Introduction

Protein homeostasis is essential for all cells and requires proper control of the folding, localization, and interactions of proteins. The biogenesis of membrane proteins poses a particular challenge to protein homeostasis. Before arrival at the membrane, newly synthesized membrane proteins need to traverse aqueous cellular compartments where they are highly prone to aggregation. Thus, the posttranslational targeting of membrane proteins relies critically on effective molecular chaperones that maintain nascent membrane proteins in translocation competent states. Many examples illustrate the intimate link between chaperone function and membrane protein biogenesis: SecB, Skp, and SurA protect bacterial outer membrane proteins (1–5), and Hsp70 homologs assist the import of mitochondrial or chloroplast proteins (6). Our understanding of membrane protein chaperones lags far behind that for soluble proteins, such as DnaK and GroEL. All chaperones need to switch between "open" and "closed" conformations to allow substrate release and binding, respectively. For many chaperones that promote the folding of soluble proteins, these switches can be driven either by ATPase cycles, such as Hsp70 (7) and GroEL (8), or by changes in environmental conditions, such as the acid-induced HdeA (9, 10) and oxidation-induced Hsp33 (11). In contrast, membrane protein chaperones must regulate their action spatially: they must effectively capture substrate proteins in the aqueous phase, and then facilely and productively release them at the target membrane. With few exceptions (1, 2), how membrane protein chaperones achieve spatiotemporal coordination of their chaperone cycle is not well understood.

The light harvesting chlorophyll a/b-binding proteins (LHCPs) provide an excellent model system to address these questions. Like >95% of organellar proteins, LHCPs are initially synthesized in the cytosol and imported across the chloroplast envelope in a largely unfolded state with the assistance of the LHCP translocation defect protein (12). In the stroma, LHCPs are protected in a soluble "transit complex" by the chloroplast signal recognition particle (cpSRP), comprised of cpSRP43 and cpSRP54 (13). Via interactions between the GTPase domains of cpSRP54 and its receptor cpFtsY, LHCPs are delivered to the Alb3 translocase and inserted into the thylakoid membrane (13–17). LHCPs comprise more than 50% of the proteins in the thylakoid membrane and are the most abundant membrane protein family on earth. Their sheer abundance, high aggregation propensity, and crucial roles in energy generation of green plants demand highly effective chaperone(s) during their biogenesis, making this a robust system to understand the function and mechanism of membrane protein chaperones.

Previous work showed that the cpSRP43 subunit in cpSRP binds tightly to and quantitatively prevents the aggregation of multiple members of the LHCP family, demonstrating that cpSRP43 is responsible for chaperone function (18, 19). cpSRP43 is comprised of multiple protein-interaction domains: three chromodomains (CDs) and an ankyrin repeat domain (A1–A4) between CD1 and CD2 (Fig. 1A) (14). Biochemical and crystallographic analyses showed that a conserved Tyr204 in the third ankyrin repeat recognizes a FDPLGL motif in L18, a conserved 18-amino acid sequence between TM2 and TM3 of LHCP (20– 22). In addition, aromatic cages in CD2 provide binding sites for a conserved RRKR motif in the C terminus of cpSRP54 (23). A recent study found that

cpSRP54 can induce compaction of cpSRP43 and enhance L18 peptide binding threefold, suggesting that cpSRP54 could positively regulate cpSRP43 (24). Finally, cpSRP43 also interacts directly with the C-terminal stromal domain of the Alb3 translocase (termed Alb3CT) that mediates the membrane insertion of LHCP (25–29), suggesting a potential mechanism to couple substrate release to the correct localization of LHCP and its imminent membrane insertion (30). The ability of cpSRP43 to directly bind Alb3 may also explain findings in earlier genetic studies that, when both cpSRP54 and cpFtsY are deleted, cpSRP43 by itself can mediate the targeting and insertion of some LHCP family members, albeit less efficiently (30).

Nevertheless, the molecular mechanism of cpSRP43's chaperone function remains elusive. Where is the substrate binding domain of this chaperone located? How does it interact with the targeting (cpSRP54) and translocation (Alb3) machineries to achieve accurate spatiotemporal regulation of its activity? More fundamentally, in the absence of an ATPase module, what propels the substrate binding and release cycle for this chaperone? In this work, a combination of biochemical and solution NMR studies addresses these questions and for the first time, to our knowledge, defines the complete chaperone cycle for a chaperone dedicated to integral membrane proteins. Our results show that cpSRP43 inherently exchanges between three distinct conformations; this allows it to be readily turned "on" by cpSRP54 in the aqueous stroma to enable tight substrate binding and to be readily switched to less active conformations by Alb3 at the membrane to enable facile substrate unloading. Furthermore, we show that Alb3 specifically induces the release of substrate TMDs, but not the L18 motif, from cpSRP43, suggesting a highly productive, stepwise mechanism for handover of the membrane protein substrates to the translocation machinery.

Results
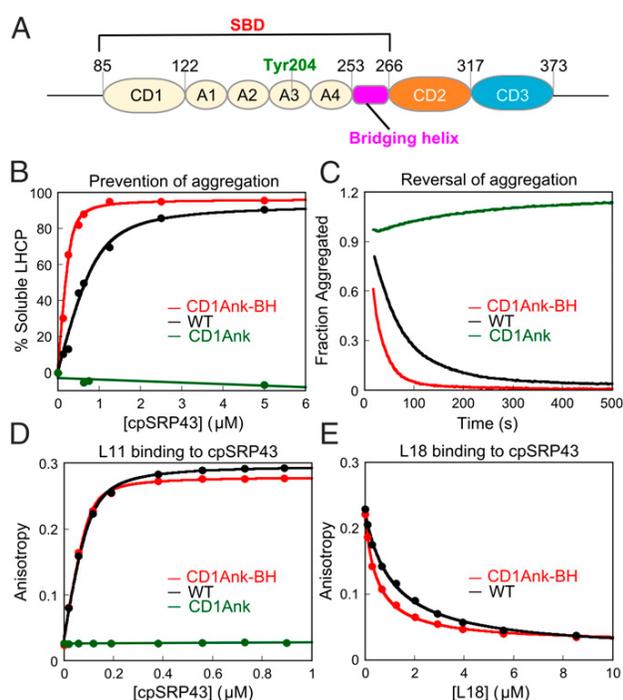
Figure 9. The CD1Ank-BH fragment is necessary and sufficient for the chaperone activity of cpSRP43. (A) Schematic of cpSRP43. CD, chromodomain; A1–A4, ankyrin repeats 1–4; BH, bridging helix; SBD, substrate binding domain. **(B)** Binding of LHCP to wildtype cpSRP43 (black) and to the CD1Ank (green) and CD1Ank-BH (red) fragments of cpSRP43 were measured by the ability of cpSRP43 to prevent LHCP aggregation (Materials and Methods). The data were fit to Eq. 1 and gave apparent Kd values of 170 and 32 nM for LHCP binding to cpSRP43 and to the CD1Ank-BH fragment, respectively. **(C)** CD1Ank-BH (red) can reverse LHCP aggregation more efficiently than WT cpSRP43 (black), but the CD1Ank (green) fragment cannot. **(D)** Binding of HiLyte-Fluor 488–labeled L11 peptide to WT and mutant cpSRP43, detected by fluorescence anisotropy. The data were fit to Eq. 2 and gave Kd values of 25 and 15 nM for the binding of dye-labeled L11 peptide to cpSRP43 (black) and to the CD1Ank-BH (red) fragment, respectively. **(E)** Binding of the L18 peptide to WT and mutant cpSRP43 was measured using L18 as a competitor of dye-labeled L11. The data were fit to Eq. 3 and gave K_i^{app} values of 1.1 and 0.5 μM for cpSRP43 (black) and the CD1Ank-BH (red) fragment, respectively. Errors of Kd and K_i^{app} values were estimated to be $\pm 10\%$ (SD) based on at least two measurements (technical replicates).

Defining the Substrate-Binding Domain for cpSRP43

In previous work, a fragment comprising CD1 and the ankyrin repeat domain of cpSRP43 (CD1Ank) could not chaperone LHCP (19). Nevertheless, crystallographic analysis (20) revealed an additional 20-amino acid-long helix, not present in other ankyrin repeat proteins, that bridges the ankyrin repeat domain and CD2 (here termed the bridging helix or BH; Fig. 1A). Intriguingly, addition of this BH to CD1Ank was necessary and sufficient to restore chaperone activity. Using a light scattering assay (19), which directly monitors the formation of large protein aggregates, we found that a protein fragment containing CD1, the ankyrin repeat domain, and BH (CD1Ank-BH) prevented the aggregation of LHCP fivefold more efficiently than full-length cpSRP43 (Fig. 1B). In a more stringent disaggregation assay (19), which monitors the ability of cpSRP43 to reverse preformed large LHCP aggregates, CD1Ank-BH was also more efficient than full-length cpSRP43 (Fig. 1C). Consistent with previous results (19), removal of BH (i.e., CD1Ank) abolished chaperone activity in both assays (Fig. 1 B and C). A key substrate recognition motif of cpSRP43 is the L18 sequence between TM2 and TM3 of LHCP (21, 31). To test whether the BH is important for this recognition, we measured cpSRP43 binding with L18 based on the increase in fluorescence anisotropy of a HiLyte-Fluor488–labeled L11 peptide, which comprises the minimal binding motif in L18 (20, 22). The CD1Ank-BH fragment binds tightly to the dye-labeled L11 peptide, with a Kd value (15 nM) twofold lower than that of WT cpSRP43 (25 nM) (Fig. 1D). Consistent with the results of the light scattering assay, removal of the BH abolished this binding (Fig. 1D). To rule out potential artifacts introduced by dye labeling, we also measured the binding of the WT L18 sequence to cpSRP43 by using it as a competitor of dye-labeled L11 in the

anisotropy assay. These experiments showed that L18 effectively bound to both WT cpSRP43 and CD1Ank-BH and competed with dye-labeled L11 (Fig. 1E). Together with previous mutational analyses (19), these results establish that the CD1Ank-BH fragment comprises the minimal substrate binding domain (SBD) of cpSRP43 and that the bridging helix connecting the ankyrin repeat domain to CD2 is crucial for substrate binding.

Chromodomain 2 Regulates Chaperone Activity
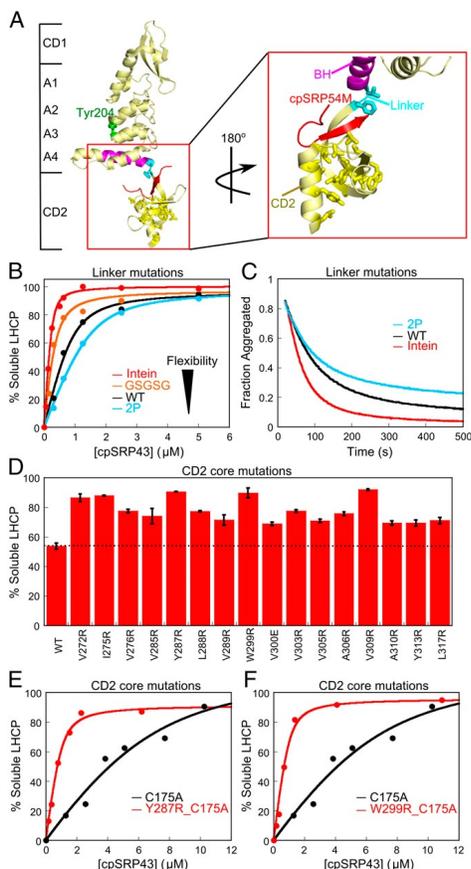
Intriguingly, we isolated many mutations in the neighboring CD2 that affect chaperone activity of the SBD. One class of these mutants resides in the linker connecting BH to CD2 (Fig. 2A, cyan). The chaperone activity of cpSRP43 is strongly correlated with the flexibility of the linker sequence, with longer and more flexible linkers [GSCFNGT (the Intein construct) or GSGSG insertion], leading to higher activities of cpSRP43 in preventing (Fig. 2B) and reversing (Fig. 2C) LHCP aggregation, whereas a conformationally more restricted linker 2P (two Pro replacing the natural QV) leads to lower activity (Fig. 2 B and C). A second class of mutants resides in the conserved hydrophobic core of CD2 (Fig. 2A, yellow). Virtually every single mutation introduced into this core hyperactivates the chaperone (Fig. 2D). Quantitative analysis of some of the chaperone (Fig. 2D). Quantitative analysis of some of the mutants further showed that cpSRP43's substrate binding is enhanced 10- to 20-fold in each mutant (Fig. 2 E and F). Thus, although CD2 does not directly bind the substrate protein, it regulates the conformation and activity of the SBD.

Figure 10. Mutations in CD2 hyperactivate the chaperone. (A) Crystal structure of cpSRP43 CD1Ank-BH-CD2 fragment bound to 54M peptide [Protein Data Bank (PDB) ID code 3UI2]. Cyan shows the linker (V266, F267) between BH and CD2, yellow shows residues mutated in the hydrophobic core of CD2, magenta highlights the BH, and red shows the 54M peptide. (B and C) Chaperone activity of cpSRP43 linker mutants, measured by prevention (B) and reversal (C) of LHCP aggregation. The data in B were fitted to Eq. 1 and gave apparent Kd values of 33, 100, 177, and 256 nM for LHCP binding to intein-cpSRP43 (red), GSGSG-cpSRP43 (orange), WT cpSRP43 (black), and 2P-cpSRP43 (cyan), respectively. (D) Many point mutations in the conserved hydrophobic core of CD2 activate chaperone activity. Chaperone activities were measured by prevention of LHCP aggregation using 1 μM LHCP and 4 μM WT or mutant cpSRP43. Error bars represent SD, with n = 2. To rule out potential involvements of CD3, all measurements were made with the ΔCD3 construct. A soft mutation, C175A, was present in all constructs to increase the sensitivity of the mutational screen under these conditions. (E and F) Mutants Y287R (E) and W299R (F) enhance LHCP binding to cpSRP43 >10-fold. Chaperone activities were measured and analyzed as in B and gave apparent Kd values of 1.9 μM, 168 nM, and 141 nM for cpSRP43(C175A), cpSRP43(Y287R_C175A), and cpSRP43(W299_C175A), respectively. Errors of Kd values were estimated to be ±10% (SD) based on at least two measurements (technical replicates).

## The cpSRP54 M Domain Drives cpSRP43 to the Active Conformation

In the stroma, cpSRP43 is bound to cpSRP54, the other subunit of cpSRP. The C terminus of the cpSRP54 M-domain (termed 54M) binds at CD2 in the vicinity of the BH (Fig. 2A) (23, 32), placing it in an optimal position for regulating interdomain interactions of cpSRP43. In support of this notion, 54M enhances the binding affinity of LHCP to WT cpSRP43 sixfold under stringent low salt conditions (100 mM NaCl; Fig. 3A). This result is consistent with a recent study, which also reported that cpSRP43 becomes more compact and binds the L18 peptide threefold more strongly in the presence of cpSRP54 (24). Importantly, the superactive mutant, intein-cpSRP43, is not further stimulated by 54M (Fig. 3B). In contrast, a soft mutation C175A reduced cpSRP43's chaperone activity ~18-fold; this partially crippled mutant is much more strongly activated by 54M (Fig. 3C). Together, these results strongly suggest that 54M and superactive mutations in CD2 drive the same activating conformational change in cpSRP43 and that interdomain dynamics control cpSRP43's chaperone activity (Fig. 3D). CpSRP43 has been bound to shift between monomer and higher oligomeric forms depending on solution ionic strength (19). To test whether the activating conformational change of cpSRP43 is related to changes in its oligmeric state, we carried out gel filtration chromatography and analytical ultracentrifugation analyses. The results showed that, under high salt conditions (≥200 mM NaCl), WT or superactive intein-cpSRP43 is predominantly

Figure 11. The cpSPR54 M-domain activates cpSRP43 for substrate binding. The abilities of chaperone to prevent LHCP aggregation were measured for WT cpSRP43 (A), superactive intein-cpSRP43 (B), and partially defective mutant cpSRP43(C175A) (C) in the absence (●) and presence (○) of 54M. The data were fit to a Michaelis–Menton equation and gave apparent Kd values of 0.26 and 1.5 μM for WT cpSRP43 with and without 54M (A), 0.20 and 0.41 μM for intein-cpSRP43 with and without 54M (B), and 0.08 and 3.0 μM for cpSRP43(C175A) with and without 54M (C). In A and B, activities were measured under 100 mM NaCl, a stringent condition under which cpSRP43 exhibits slightly reduced activity, to overcome the saturation effects with highly active chaperone constructs and better reveal the stimulatory effects of 54M. (D) Scheme depicting the conformational change of cpSRP43 on substrate binding. TMD, transmembrane domain.

a monomer. cpSRP43 also binds the L11 peptide and 54M in monomeric form. Under low salt conditions (50 mM NaCl), cpSRP43 predominantly runs on Superdex 200 as a dimer as previously observed (19). Importantly, binding of L11 shifts dimeric cpSRP43 toward the monomeric form. These results, together with the previous observation that the shift of cpSRP43 to dimeric complexes at lower ion strength is correlated with loss of chaperone activity (19), strongly suggest that monomeric cpSRP43 is chaperone active. Thus, the substrate- and 54M induced rearrangement of cpSRP43 reflects an intramolecular conformational change, rather than changes in its oligomeric state.

NMR Spectroscopy Directly Reveals Chaperone Conformational Dynamics

To directly probe for conformational dynamics in cpSRP43, we used transverse relaxation optimized (TROSY) NMR spectroscopy. We assigned the cpSRP43 backbone resonances using 3D triple resonance TROSY spectra of 2H,13C,15N-labeled cpSRP43 and the CD1Ank-BH fragment; assignments for several CD2 cross-peaks in full-length cpSRP43 were obtained by transferring published assignments of the isolated CD2CD3 fragment (33). We were able to assign the amide 15N, 1HN, 13Cα, and 13Cβ resonances for 80% of residues. Intriguingly, the number of NMR cross-peaks in the 1H-15N TROSY spectrum of cpSRP43 far exceeded the number of residues in the protein. This puzzle was resolved during NMR assignments: we found that at least 12 residues in CD1 (highlighted in Fig. 4G in blue) give rise to two 1H-15N cross-peaks in the TROSY spectrum that have identical 13Cα and 13Cβ chemical shifts (examples for S92 and A95). This observation indicates that the component cross-peaks in the TROSY spectra arise from the same amino acids and represent two distinct conformations of cpSRP43 in slow exchange on the chemical shift timescale. Importantly, the relative intensities of the component crosspeaks change in different constructs and in response to different ligands. An example for S92 is shown in Fig. 4. In full-length cpSRP43, the intensities of the twocomponent cross-peaks are comparable (Fig. 4B). In the inactive CD1Ank construct, the upfield peak is dominant (Fig. 4A). In contrast, in the CD1Ank-BH construct or the superactive mutant intein-cpSRP43, which binds LHCP two- to fivefold more tightly than full-length cpSRP43 (Figs. 1B and 2B), the downfield peak becomes more intense (Fig. 4 C and D). When cpSRP43 is bound to the L18 peptide or to the activator, 54M, or the C-terminal peptide of 54M that binds cpSRP43 (54M peptide), the downfield peak is dominant (Fig. 4 E and F). Additional titration experiments with L18 and 54M peptide corroborated that the downfield peak can be assigned to the active conformation of cpSRP43 conducive to substrate binding. Further, the relative intensity of this peak is strongly correlated with chaperone activity (Fig. 4H). Collectively, the biochemical and NMR data demonstrate that the SBD of cpSRP43 exchanges between active and inactive conformations that are regulated by CD2, by the substrate, and by interaction with the cpSRP54 M-domain.
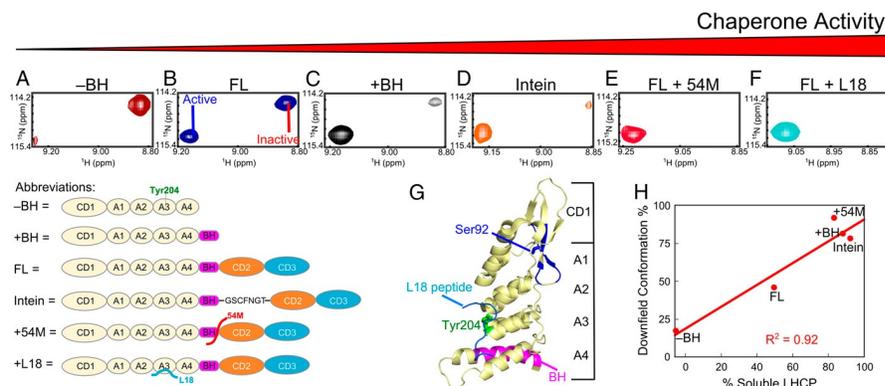
Figure 12. Conformational dynamics of cpSRP43 correlates with its chaperone activity. (A–F) Component cross-peaks for Ser92 in the TROSY spectrum of 2H,15Nlabeled CD1Ank fragment (A), full-length cpSRP43 (B), CD1Ank-BH fragment (C), intein-cpSRP43 (D), cpSRP43 bound to 54M peptide (E), and cpSRP43 bound to the L18 peptide (F). **(G)** Crystal structure of CD1Ank-BH fragment (PDB ID code 3DEP) highlighting the residues in CD1 (blue) for which component cross peaks reflecting the conformational dynamics of cpSRP43 have been unambiguously assigned. Green shows Tyr204 in Ank3, which binds L18 (cyan). Magenta highlights the bridging helix. **(H)** Correlation between the chaperone activity (% soluble LHCP observed with 0.625 µM chaperone) of each construct or complex and the relative intensity of the downfield component cross-peak.

Bidentate Interactions of Alb3 with cpSRP43 Drive Substrate Release
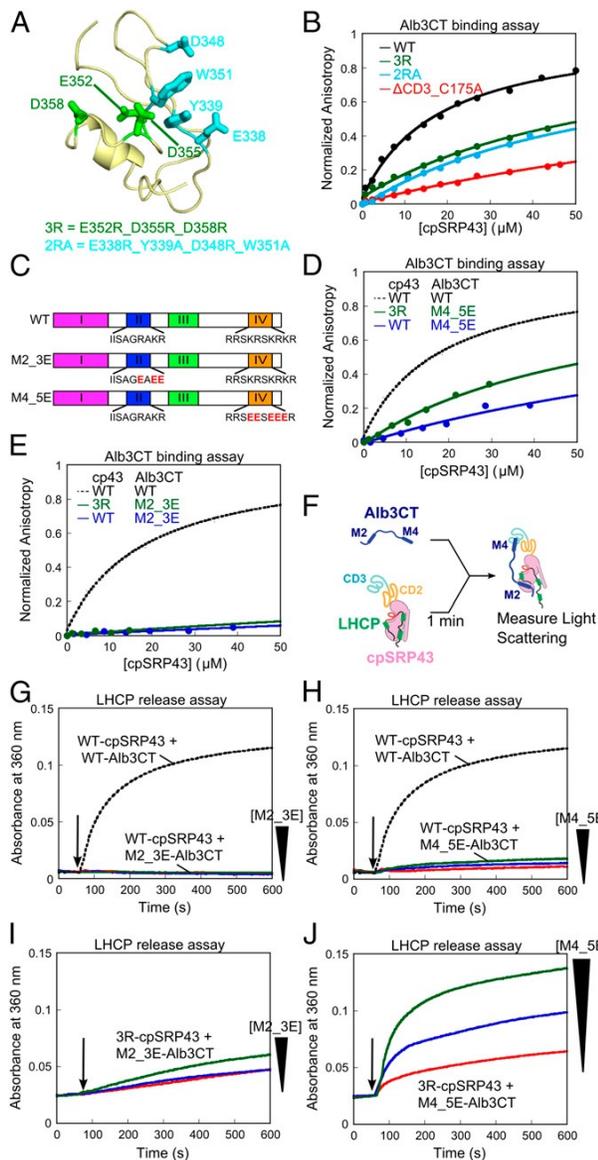
Recent data showed that cpSRP43 directly interacts with the C-terminal stromal domain of Alb3 (Alb3CT). However, which site(s) in cpSRP43 interact with Alb3CT is extensively debated (25–27, 29). To resolve this question, we titrated Alb3CT into cpSRP43 and monitored changes in the TROSY spectra. We found that Alb3CT substantially (>90%) broadened a specific set of cross-peaks in CD3, in contrast to the absence of any substrate- or 54M-induced perturbations in this domain. Many perturbed cross-peaks map to an extensive surface displaying conserved acidic residues and two aromatic residues (Fig. 5A), suggesting this to be a binding site for Alb3CT. This observation is consistent with a recent study in which a peptide corresponding to motif IV in Alb3CT is found to bind at an acidic surface in CD3 (29).

Figure 13. Chromodomain 3 of cpSRP43 binds motif IV of Alb3CT. (A) A structural model of CD3 (PDB ID code 1X32) highlighting the mutated residues. **(B)** Binding of fluorescein-labeled Alb3CT(S371C) to WT and mutant cpSRP43 were measured by changes in fluorescence anisotropy. The data were fit to Eq. 2 and gave Kd values of 18, >60, >63, and >160 µM for WT cpSRP43 and mutants 2RA, 3R, and ΔCD3, respectively. The anisotropy value of mutants was normalized to the same end point as WT cpSRP43. **(C)** Schematics of WT and charge reversal mutants of Alb3CT. **(D and E)** Binding of mutant Alb3CT(M4_5E) **(D)** and Alb3CT(M2_3E) **(E)** to WT cpSRP43 and mutant cpSRP43(3R), measured and analyzed as in B. The dashed lines indicate binding between WT Alb3CT and cpSRP43 from B and are shown for comparison. **(F)** Scheme for the LHCP TMD release assay. M2 and M4 denote Alb3CT motifs II and IV, respectively. **(G–J)** Alb3CTinduced TMD release from cpSRP43 for the reaction of WT cpSRP43 with mutant Alb3CT(M2_3E) **(G)**, WT cpSRP43 with Alb3CT(M4_5E) **(H)**, mutant cpSRP43(3R) with mutant Alb3CT(M2_3E) **(I)**, and mutant cpSRP43(3R) with mutant Alb3CT(M4_5E) **(J)**. The arrows indicate time of Alb3CT addition. Red, blue, and green indicate addition of 5, 10, and 20 µM Alb3CT, respectively. The dashed lines denote data of WT cpSRP43 with 5 µM Alb3CT and are shown for comparison.

To biochemically test this model, we generated two charge reversal mutants in CD3, E352R/D355R/D358R (3R; Fig. 5A, green), and E338R/Y339A/D348R/W351A (2RA; Fig. 5A, cyan). We measured the binding of Alb3CT to WT and mutant cpSRP43 based on the anisotropy change of fluorescein labeled at Alb3CT(S371C). Consistent with previous measurements (25, 26), WT cpSRP43 bound Alb3CT with an equilibrium dissociation constant (Kd) of 18 μM in this assay. Binding was weakened ninefold on deletion of CD3 (Fig. 5B, red), consistent with an important role of CD3 in cpSRP43-Alb3 binding (25). Both charge reversal mutants in CD3, cpSRP43(3R) and cpSRP43 (2RA), weakened the binding of Alb3CT fourfold (Fig. 5B).Alb3CT contains two conserved motifs, II and IV, that could interact with cpSRP43 (Fig. 5C) (25). Both motifs, especially motif IV, are enriched in basic residues, providing strong candidates for interacting with the acidic patch in CD3. We therefore introduced charge reversal mutations into motif II (M2_3E) or motif IV (M4_5E) of Alb3CT (Fig. 5C). Both mutations significantly weakened the binding of Alb3CT to cpSRP43 (Fig. 5 D and E, blue). Importantly, when the charge reversal mutants cpSRP43(3R) and Alb3CT(M4_5E) were combined to restore electrostatic complementarity, binding was partially rescued (Fig. 5D, green). In contrast, rescue was not observed when cpSRP43(3R) was combined with mutant Alb3CT(M2_3E) (Fig. 5E, green). These results suggest a specific electrostatic interaction between motif IV in Alb3CT and the acidic surface in CD3 of cpSRP43. To provide additional evidence for this model and to probe the function of Alb3CT, we developed an independent assay in which a preformed, soluble cpSRP43•LHCP complex is challenged with Alb3CT. If interaction with Alb3CT releases substrate from cpSRP43, this would lead to aggregation of LHCP that can be monitored in real time using light scattering (Fig. 5F). Indeed, addition of Alb3CT led to the reappearance of large LHCP aggregates in a dose-

Figure 14. Bidentate interaction of Alb3CT with cpSRP43. (A) Alb3CT-induced perturbation of cross-peaks in the TROSY spectra of 2H,15N-labeled CD1Ank-BH fragment. Spectra in the absence and presence of Alb3CT are shown in black and red, respectively. **(B)** Structure of CD1Ank-BH (PDB ID code 3DEO) highlighting residues for which NMR cross-peaks are perturbed by Alb3CT. Red denotes residues for which cross-peaks are broadened >70%; blue denotes residues whose cross-peaks are shifted by Alb3CT. Green shows Tyr204 in Ank3 where L18 binds. **(C)** Schematic of the mutant Alb3CT constructs used in this study. ΔM2 and ΔM4 denote deletion of motifs II and IV, respectively, ΔL denotes deletion of linker sequence. **(D)** Binding of WT Alb3CT (black) and mutant ΔM2 (red) and ΔM4 (blue) to cpSRP43, measured by their ability to compete with fluorescein-labeled Alb3CT(S371C) for binding to cpSRP43 and detected by fluorescence anisotropy. The data were fit to Eq. 3 and gave Kappi values of 41, 99, and 237 μM for WT Alb3CT and mutants ΔM2 and ΔM4, respectively. **(E and F)** LHCP release assays were carried out as outlined in Fig. 5F for WT Alb3CT (black), mutants ΔM2 (red) and ΔM4 (blue) (E), and mutants ΔL20 (blue), ΔL40 (orange) and ΔL47 (green) (F).

dependent manner. This result indicates that the hydrophobic TMDs of LHCP are no longer protected by cpSRP43 on Alb3CT binding and provides a robust assay to analyze the interaction and activity of Alb3CT. Both charge reversal mutants of Alb3CT, M2_3E and M4_5E, display severely compromised activities in this release assay (Fig. 5 G and H), supporting the importance of both motifs in Alb3CT activity. Importantly, combining the charge reversal mutant cpSRP43(3R) with Alb3CT(M4_5E) restored release activity (Fig. 5J), whereas the combination with Alb3CT(M2_3E) did not (Fig. 5I). Thus, in agreement with the recent structural analysis (29), cpSRP43 CD3 provides a platform to specifically bind motif IV in Alb3CT. Our results further show that this interaction is electrostatic in nature and important for the ability of Alb3CT to induce substrate release from cpSRP43. In addition to CD3, Alb3CT also induced substantial broadening of 17 cross-peaks in the TROSY spectrum of cpSRP43 that map to the SBD. This observation was corroborated when the TROSY experiment was repeated with the CD1Ank- BH construct: even in the absence of CD2 and CD3, addition of Alb3CT-induced broadening or shifts of a specific set of cross-peaks in the TROSY spectrum of CD1Ank-BH (Fig. 6A). Seven of the Alb3CT-perturbed cross-peaks map to CD1; another three map to Ank3 (Fig. 6B). Thus, Alb3CT also contacts the SBD. The simplest model to explain these results is that Alb3CT makes bidentate interactions with cpSRP43 in both the SBD and CD3. To test this model, we deleted motif II or IV of Alb3CT (Fig. 6C, ΔM2 and ΔM4, respectively). We measured the binding of WT and mutant Alb3CT to cpSRP43 by using them as competitors of fluorescein-labeled Alb3CT in the anisotropy assay (Fig. 6D). ΔM2 and ΔM4 weakened the binding of Alb3CT to cpSRP43 two- and sixfold, respectively (Fig. 6D). Both mutations also abolished the ability of Alb3CT to trigger LHCP release from cpSRP43 (Fig. 6E). The same effects were observed with charge reversal mutations in Alb3CT motifs II and IV interaction between Alb3CT and cpSRP43 that are essential to release LHCP from cpSRP43 at the membrane translocase (Fig. 5 D, E, G, and H). Together, these data provide strong evidence that Alb3CT uses motifs II and IV to make bidentate interactions with cpSRP43. The much larger effects of each mutant in the LHCP release assay than on cpSRP43-Alb3 binding (Fig. 6 D vs. E) further suggests that, although each motif could bind cpSRP43 at sufficiently high concentrations, the ability of Alb3CT to trigger substrate release requires the interactions mediated by both motifs.

If a bidentate interaction with Alb3CT was required to release LHCP from cpSRP43, then the spacing between motifs II and IV would also be important for this activity. In WT Alb3, these two motifs are bridged by a ~72-amino acid unstructured linker sequence, rendering it plausible that the two motifs can span the distance between the SBD and CD3 of cpSRP43. To test the importance of this spacing, we shortened the linker sequence (Fig. 6C, mutants ΔL20, ΔL40, and ΔL47). Although mutant ΔL20 did not exhibit a significant defect, further deletion of the linker sequence in mutants ΔL40 and ΔL47 compromised the ability of Alb3CT to induce LHCP release from cpSRP43 (Fig. 6F). Collectively, these data provide strong evidence for a bidentate interaction between Alb3CT and cpSRP43 that are essential to release LHCP from cpSRP43 at the membrane translocase.

Alb3CT Drives cpSRP43 to a Less Active Conformation

How does Alb3CT trigger substrate release from cpSRP43? Given that cpSRP43 samples between active and inactive conformations (Figs. 2–4), an attractive hypothesis is that Alb3CT drives cpSRP43 to a less active form. To test this model, we used two independent approaches. First, we took advantage of cpSRP43 variants, CD1Ank-BH and CD1Ank, which reside predominantly in the active and inactive conformations, respectively, and tested their binding with Alb3CT using the fluorescence anisotropy assay. We found that Alb3CT exhibits significant interactions with the CD1Ank fragment, although the interaction is 18-fold weaker than with full-length cpSRP43, consistent with the model that the SBD provides another binding site for Alb3CT (Fig. 7A). Importantly, Alb3CT bound to the CD1Ank fragment two- to threefold more strongly than to the CD1Ank-BH fragment, suggesting that it preferentially interacts with a cpSRP43 in a less active conformation. The stronger interaction of CD1Ank with Alb3CT than CD1Ank-BH also ruled out the possibility that irreversible misfolding of CD1Ank is responsible for its lack of chaperone activity. In the second approach, we took advantage of the TROSY spectra of cpSRP43, in which the relative intensity of component cross-peaks from individual amide groups directly detects cpSRP43 subpopulations in active and inactive conformations. Of the 12 residues in cpSRP43 for which component cross-peaks can be unambiguously assigned, residues S92, R93, Y102, L103 exhibited an Alb3CT-induced increase in the relative intensity of the peak that arise from the inactive conformation (Fig. 7B). These results provide independent support for the model that Alb3CT biases cpSRP43 to a less active conformation. In the targeting pathway, Alb3CT must interact with and release substrates from the LHCP•cpSRP43•cpSRP54 complex. To test whether this is possible, we sequentially added the 54M peptide, followed by Alb3CT, during 2D-TROSY experiments with the superactive intein-cpSRP43 (to mimic the conformation of substrate-bound cpSRP43). For five residues (S92, R93, T94, Y102, and L103), the intensity of the component cross-peaks representing the inactive conformation was significantly reduced on the addition of 54M, but these inactive peaks regained

Figure 15. Alb3CT preferentially binds and induces a less active conformation of cpSRP43. (A) Binding of Alb3CT to full-length cpSRP43 (black) and to the CD1Ank (green) and CD1Ank-BH (red) fragments, measured as in Fig. 5B. The data were fit to Eq. 2 and gave Kd values of 12, 216, and >530 μM for WT cpSRP43, CD1Ank, and CD1ANk-BH, respectively. Errors for Kd were estimated to be ±10% (SD) based on at least two technical replicates. **(B)** Region of the TROSY spectrum showing the component cross-peaks for Ser92 in 2H, 15N-labeled cpSRP43 in complex with Alb3CT. **(C)** The relative intensities of the cross-peaks corresponding to the inactive chaperone conformation are reduced by 54M (red bars) and enhanced by subsequent addition of Alb3CT (green bars). Data are shown for indicated residues in 2H, 15N-labeled inteincpSRP43. Relative intensities were determined from cross-peak heights. **(D)** 54M and Alb3CT bind to intein-cpSRP43 simultaneously. The cross-peaks for Tyr287 in the TROSY spectra of 2H, 15N-labeled intein-cpSRP43 are shown. The peak was shifted on 54M binding (red), and this perturbation persists on subsequent addition of Alb3CT (green). **(E)** Scheme depicting the cobinding of 54M and Alb3CT to cpSRP43. M2, Alb3CT motif II; M4, Alb3CT motif IV.

significant intensity on subsequent addition of Alb3CT (Fig. 7C). The effect of Alb3CT did not arise from dissociation of 54M from cpSRP43, as the binding of 54M to cpSRP43 could be detected by shifts of specific cross-peaks corresponding to residues in and near CD2 (Fig. 7D), and the 54M-induced perturbation of these cross-peaks remained even after the addition of Alb3CT (Fig. 7D). Collectively, the biochemical and NMR data strongly suggest that Alb3CT induces cpSRP43 to a less active conformation (Fig. 7E). The fact that this transition was observed with only a subset of residues further suggests that Alb3CT does not drive cpSRP43 into the completely inactive form but rather induces structural transitions in part of the cpSRP43 molecule. This notion is further supported by results in the following section.

Alb3CT Specifically Releases Substrate TMDs from cpSRP43

Our previous work showed that cpSRP43 binds LHCP via two sets of interactions: recognition of the L18 motif in loop 2 and promiscuous hydrophobic interactions with the TMDs in LHCP (22). When the cpSRP43•LHCP complex was challenged by Alb3CT in the release assay, the light scattering data indicated that Alb3CT antagonizes the interaction of substrate TMDs with cpSRP43 such that they are no longer protected from aggregation (Fig. 5F). To test whether Alb3CT also antagonizes the interaction of cpSRP43 with the L18 motif, we measured the binding of HiLyte-Fluor488–labeled L11 peptide to cpSRP43 based on fluorescence anisotropy (Fig. 8A). Intriguingly, the binding of L11 to cpSRP43 was unaffected, if not slightly stronger, in the presence of Alb3CT (Fig. 8A). To rule out artifacts from dye labeling, we also measured the binding of unlabeled L18 peptide to cpSRP43 by using it as a competitor of dye-labeled L11 (Fig. 8B). This experiment yielded the same result: L18 binds to cpSRP43 with comparable affinity with or without Alb3CT present. These data indicate that Alb3CT specifically antagonizes the interaction of cpSRP43 with the substrate TMDs, but does not affect the interaction of this chaperone with the L18 motif.

To test this model in the context of full-length LHCP, we reperformed the LHCP release assay except that, instead of monitoring the release reaction by light scattering, we monitored FRET between a donor dye (Atto488) labeled in the L18 motif of LHCP (G158C) and an acceptor dye [tetramethylrhodamine (TMR)] labeled at the native Cys297 of cpSRP43. In contrast to the release of TMDs detected by light scattering, the FRET assay showed no changes in, if not slightly higher, FRET efficiency between LHCP-L18 and cpSRP43-CD2 when the cpSRP43•LHCP complex was challenged with Alb3CT (Fig. 8C), suggesting that Alb3CT did not induce the release of L18 from cpSRP43. Together, these results support a model in which Alb3CT induces cpSRP43 into a distinct conformation in which the substrate TMDs are released, whereas the L18 motif of LHCP remains bound to the chaperone.

Figure 16. Alb3CT uncouples the interaction of cpSRP43 with the L18 motif and TMD of LHCP. (A) Binding of HiLyte-Fluor488–labeled L11 peptide to cpSRP43 with (blue) and without (red) Alb3CT present, measured using fluorescence anisotropy. The data were fit to Eq. 2 and gave Kd values of 22 and 11 nM in the absence and presence of 20 μMAlb3CT, respectively. (B) Binding of the L18 peptide to cpSRP43 in the absence (red) and presence (blue) of Alb3CT, measured by using L18 as a competitor of HiLyte-Fluor488–labeled L11 peptide in binding cpSRP43. The data were fit to Eq. 3 and gave Kappi values of 1.1 and 0.68 μM in the presence and absence of Alb3CT, respectively. (C) (Left) Scheme of the FRET-based LHCP release assay (Materials and Methods). (Right) FRET efficiency between Atto488-labeled LHCP (at L18) and TMR labeled cpSRP43 before and after challenge by Alb3CT. Values reported are mean ± SD, with n = 2. (D) Model for the chaperone cycle of cpSRP43 during LHCP targeting and insertion, as described in the text. TMD, transmembrane domain; SBD, substrate binding domain; M2, Alb3CT motif II; M4, Alb3CT motif IV.
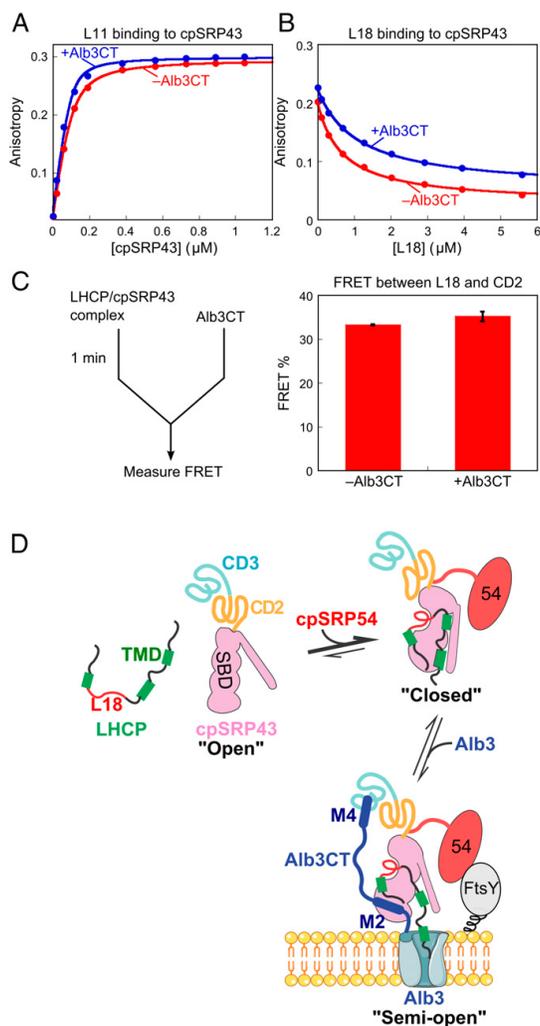
Discussion

Membrane proteins pose special challenges to protein homeostasis during their posttranslational targeting and require highly effective chaperones. Compared with chaperones that facilitate the folding of soluble proteins, membrane protein chaperones not only handle much more aggregation-prone client proteins, but must also regulate their substrate binding and release in response to spatial cues. Here, we define the complete chaperone cycle for cpSRP43, an ATP-independent chaperone dedicated to integral membrane proteins. Our results reveal a remarkably modular nature of this chaperone, wherein substrate binding and spatial regulations are mediated by distinct domains. Most importantly, the SBD of cpSRP43 intrinsically samples at least three distinct conformations. This conformational sampling enables this chaperone to be readily switched on by activators in the stroma and switched off by a negative regulator at the target membrane, driving highly coordinated substrate capture and release despite the lack of ATPase cycles.

Substrate Binding Domain Samples Multiple Conformations

Remarkably, a 25-kDa fragment in cpSRP43 comprised of CD1, the ankyrin repeats, and the bridging helix was sufficient for stoichiometric binding and chaperoning of LHCP by cpSRP43. Combined with previous mutational work that tested the effects of deleting CD1 and individual ankyrin repeats (19), this defines the CD1Ank-BH fragment as the minimal substrate binding domain for cpSRP43. How this small chaperone domain protects client proteins, which match its own size and contain three TMDs, will be the next challenging question. conformations. This property is most directly visualized in the 1H-15N TROSY spectra of cpSRP43, in which at least two distinct conformations in slow exchange gave rise to pairs of component cross-peaks for the same backbone NH. Although component cross-peaks are unambiguously assigned for 12 residues reported here, the number of cross-peaks in the TROSY spectrum is ~50% greater than the number of residues in cpSRP43, indicating that many more residues undergo analogous conformational sampling. A recent single molecule study also revealed a high degree of conformational fluctuation in cpSRP43 (24). Our ability to isolate a large set of superactive and defective mutations that lock this chaperone into distinct conformations provides further evidence for this model. Importantly, the relative intensity of the component cross-peaks in the TROSY spectra strongly correlate with chaperone activity in different cpSRP43 variants and in the presence of different regulators (more discussions below). Based on these observations, we propose that the SBD samples between an open conformation unable to bind substrates and a closed conformation conducive to tight substrate binding (Fig. 8D, Upper). SBD Is Activated by cpSRP54 in the Soluble Phase. Although CD2 does not directly bind substrate proteins, molecular events in this domain regulate substrate binding in the neighboring SBD. Intriguingly, CD2 by itself reduces substrate binding in the SBD and biases cpSRP43 toward the less active open conformation. Combined with the ability of a large number of point mutations in CD2 to hyperactivate cpSRP43 and drive the SBD to the closed conformation, these observations suggest the presence of strong evolutionary pressure to maintain a substantial population of apo-cpSRP43 in the inactive conformation, i.e., to keep this chaperone at the tipping point of conformational transitions. This conformational property enables cpSRP43

to be readily turned on by cpSRP54 in the stroma. The C-terminal tail of the cpSRP54 M-domain intercalates between the bridging helix and CD2, placing it at an optimal position for regulating interdomain interactions (23). In support of this model, we found here that 54M or the 54M peptide stimulates the chaperone activity of SBD. A recent study also found that cpSRP54 enhanced binding of L18 to cpSRP43 threefold (24). Our observation that the stimulatory effect of 54M is largely bypassed in superactive chaperone mutants and becomes more pronounced in defective chaperone mutants further indicates that 54M and superactive mutations in CD2 drive the same activating conformational change in the SBD. This 54M-driven rearrangement allows cpSRP43 to efficiently capture and tightly bind its substrates in the stroma, effectively protecting LHCPs from aggregation in the aqueous environment (Fig. 8D, Upper Right). The structural basis of the communication between CD2, 54M, and SBD remains a challenging question for future studies. Nevertheless, the available data provide intriguing clues. As mutations in the linker bridging the SBD and CD2 can lead to gain or loss of function, communication likely involves reorientation of CD2 relative to the SBD. This model is consistent with previous small angle X-ray scattering (SAXS) analysis (23) and with the observation of enhanced FRET efficiency between dye pairs in Ank3 and CD2 on cpSRP54 binding (24). The BH, which connects the ankyrin repeat domain to CD2 and whose deletion drives cpSRP43 into the inactive state, provides a prime candidate for mediating interdomain communication. Nevertheless, multiple residues in CD1 can sense the binding of L18 and 54M, which binds >50 Å away, indicating that substrate and 54M induces long-range communications that propagates through the entire length of the SBD.

Bidentate Interactions of Alb3CT Drive Coordinated Substrate Release

At the target membrane, cpSRP43 must facilely release its substrate to the Alb3 translocase. The finding that cpSRP43 directly interacts with the stromal domain of Alb3 (25, 27, 30) provides an attractive mechanism to couple the release of substrate to its correct localization and imminent insertion into membrane. Nevertheless, where and how cpSRP43 interacts with Alb3 has been highly controversial, with the ankyrin repeat domain or the chromodomains alternatively proposed as Alb3 binding sites (25–27, 29). Here, high-resolution NMR combined with biochemical analyses resolved this issue and showed that Alb3CT uses motifs II and IV to make bidentate interactions with both the SBD and CD3 of cpSRP43, respectively. As these sites are distinct from the 54M binding site (located in CD2), this rationalizes the observation that cpSRP43 can form a ternary complex with both cpSRP54 and Alb3CT in the NMR studies here and in previous pull-down experiments (25). The fact that Alb3CT and cpSRP54 biases cpSRP43 to different conformations also explains the anticooperative binding between these two factors (24, 25). Given the predominantly electrostatic nature of the interaction between CD3 and Alb3 motif IV, it is likely that the acidic patch on CD3 provides the site for initial recruitment of Alb3CT (29), which then enables Alb3 motif II to further contact the SBD to induce substrate release. Intriguingly, Alb3CT specifically induces the release of substrate TMDs from cpSRP43 without disrupting the interaction of the L18 motif with the chaperone. This observation has several important implications. First, Alb3CT does not simply reverse the effect of 54M and drive cpSRP43 to the completely inactive open conformation. Instead, Alb3CT induces a third, semiopen state of cpSRP43 in which its interactions with the L18 and TMDs on substrate protein are uncoupled (Fig. 8D, Lower). Second, this suggests a highly coordinated mechanism of substrate release at the

membrane, in which the TMDs are first released and could initiate their membrane insertion via Alb3, whereas the L18 loop of LHCP remains bound to cpSRP43. Formation of such a LHCP•cpSRP•Alb3 release complex at the membrane is consistent with the previous observation that cpSRP43 was trapped into high molecular weight species with LHCP on addition of Alb3CT (27). In contrast to the current mechanisms in which LHCP is completely released from cpSRP43, such a stepwise mechanism would minimize irreversible aggregation of LHCP and abortive targeting reactions, providing a more productive route for unloading the membrane protein substrate onto the translocase.

Summary

Our work here defines a rigorous framework for the chaperone cycle of cpSRP43 (Fig. 8D). The SBD of cpSRP43 samples at least three conformational states: an open state (Upper Left), a closed state (Upper Right), and semiopen state that binds tightly to L18 but not the TMDs of LHCP (Lower). CD2 biases the SBD to the open state, whereas interaction with cpSRP54 drives cpSRP43 to the closed state that binds LHCP tightly. When the transit complex is targeted to the thylakoid membrane via the interaction of cpSRP54 with cpFtsY, Alb3CT uses motifs II and IV (M2 and M4) to make bidentate interactions with SBD and CD3 of cpSRP43, respectively. These interactions induce cpSRP43 to a semiopen conformation, triggering the release of the substrate TMDs from cpSRP43 and initiating their membrane insertion, whereas the L18 motif remains bound to cpSRP43. Ultimate release of L18 from cpSRP43 might be driven by folding of LHCP in the membrane and binding of photosynthetic pigments. Although cpSRP43 is dedicated to the LHCP family of proteins, it embodies multiple new concepts that have emerged from the recent discovery of a wide array of chaperones: (i) the use of protein interaction energy instead of ATPase cycles to regulate substrate binding/release, which is found in cyclophilins (34), small heat-shock proteins (35–38), and all of the chaperones in the bacterial periplasmic or eukaryotic extracellular space; and (ii) the use of conformational flexibility for activation, which was also found in HdeA, Hsp33, and Hsp26 (39, 40), and may be a general feature of ATP-independent chaperones. The mechanism we describe here for cpSRP43 could facilitate understanding of these conceptually analogous chaperone systems, as well as methods for their investigation.

Materials and Methods

Protein Expression and Purification. cpSRP43 and Alb3CT mutants were constructed using the QuikChange procedure (Stratagene). CD1Ank constructs were deleted residues from 253 to 266 compared with CD1Ank-BH. In the linker mutants Intein (GSCFNGT) and GSGSG, the indicated sequences were inserted between V266 and F267. In mutant 2P, two prolines replaced the original linker residues (Q265 and V266). WT and mutant cpSRP43 and LHCP were overexpressed and purified as previous described (41). Alb3CT was overexpressed in BL21(DE3) cells and purified as previously described (25). 54M peptide (QKAPPGTARRKRKAC) was from Eton Bioscience (99% purity). Peptides L11 (GSFDPLGLADD), L18 (VDPLYPGGSFDPLGLADD), and L11 labeled with HiLyte-Fluor488 were purchased from AnaSpec (>95% purity). Single cysteine mutants of cpSRP43 and LHCP were labeled with fluorescent dyes via maleimide chemistry with 80–90% efficiency. Labeled proteins were purified through Sephadex G25 to remove free dye.

Chaperone Activity of cpSRP43. Two types of experiments were used. First, we measured the ability of cpSRP43 to prevent LHCP aggregation, as described previously (19, 22). Briefly, urea denatured LHCP was diluted into buffer containing varying concentrations of cpSRP43, and light scattering at 360 nm was measured over time until equilibrium was reached. The percentage of soluble LHCP (% soluble) at equilibrium was plotted as a function of cpSRP43 concentration. The data were fit to

$$\% \, soluble = \, \Delta A \frac{[LHCP] + pro + K_d - \sqrt{([LHCP] + [pro] + K_d)^2 - 4[LHCP][pro]}}{2[LHCP]}$$

in which [pro] is cpSRP43 concentration, ΔA is the soluble% at saturating cpSRP43 concentrations, and Kd is the apparent dissociation constant of LHCP interaction with cpSRP43. Second, we measured the ability of cpSRP43 to reverse preformed LHCP aggregates. LHCP (1 μM) was aggregated in buffer for 1 min, followed by addition of 4–5 μM WT or mutant cpSRP43, and the clearance of large LHCP aggregates was followed in real time by light scattering at 360 nm.

Measurement of L18 Binding. The interaction of L18 with cpSRP43 was assessed by two methods. First, binding of cpSRP43 to HiLyte-Fluor488–labeled L11 peptide, which contains the minimal interaction sequence in L18, was detected by changes in fluorescence anisotropy of the dye. Anisotropy measurements were conducted in buffer D (50 mM KHepes, pH 7.5, 200 mM NaCl) on Fluorolog 3–22 (Yobin Yvon), using 100 nM HiLyte-Fluor488–labeled L11 and varying concentrations of cpSRP43. The samples were excited at 500 nm, and the fluorescence anisotropy was recorded at 527 nm. The data were fitted to

$$A_{obsd} = A_0 + \, \Delta A \frac{[L11] + pro + K_d - \sqrt{([L11] + [pro] + K_d)^2 - 4[L11][pro]}}{2[L11]}$$

in which Aobsd is the observed anisotropy value, A0 is the anisotropy value without cpSRP43, ΔA is the change in anisotropy at saturating cpSRP43 concentrations, and Kd is the equilibrium dissociation constant for cpSRP43 interaction with L11-HiLyte-Fluor488. To independently measure the L18-cpSRP43 interaction without perturbations from the dye, unlabeled L18 peptide was used as a competitor for the binding of L11-HiLyte-Fluor488 to cpSRP43; 100 nM L11-HiLyte-Fluor488 was preincubated with 120 nM cpSRP43 for 5 min, and the complex was challenged with an increasing concentration of L18 peptide. Anisotropy values were recorded at equilibrium and plotted as a function of [L18]. The data were fit to

$$A_{obsd} = A_0 - \Delta A \frac{[L18]}{[L18] + K_i^{app\prime}}$$

in which Aobsd is the observed anisotropy value, A0 is the anisotropy value without L18 present, ΔA is the change in anisotropy at saturating L18 concentrations, and Kapp is the apparent inhibition constant. Alb3CT Binding to cpSRP43. Binding was detected by two methods. First, Alb3CT(S371C) was labeled with fluorescein-5′-maleimide (Invitrogen), and labeled protein was purified through Sephadex G25 (Sigma) to remove free dye. Binding of fluorescein-labeled Alb3CT(S371C) to cpSRP43 was detected by changes in fluorescence anisotropy, measured as described for L11 binding except that 200 nM Alb3CT(S371C)-fluorescein was used instead of L11. Samples were excited at 495 nm, and the fluorescence anisotropy was recorded at 512 nm. The data were fit to Eq. 2, with the exception that [Alb3CT] replaces [L11]. Second, unlabeled Alb3CT was used as competitors in the interaction of cpSRP43 with Alb3CT(S371C)-fluorescein; 200 nM Alb3CT-S371C-fluorescein was preincubated with 10 μM cpSRP43 for 5 min, and the complex was chased with increasing concentrations of WT or mutant Alb3CT. Fluorescence anisotropy values were recorded at equilibrium and plotted as a function of competitor concentration. The data were fit to Eq. 3, with the exception that [Alb3CT] replaces [L18].

Alb3-Induced Substrate Release. LHCP release was evaluated by two methods. First, a soluble LHCP•cpSRP43 complex was performed in buffer D using 1 μM LHCP and 5 μM cpSRP43. After 1-min incubation, the complex was chased with Alb3CT, and the reaction was monitored in real time by light scattering at 360 nm. Second, a single cysteine LHCP mutant (C80A,G158C) was labeled with Atto488 maleimide in the middle of L18. A single cysteine mutant of cpSRP43 (C175A-C297) was labeled with TMR-5-maleimide (Invitrogen). Labeled LHCP (50 nM) and cpSRP43 (2 μM) were preincubated for 5 min in buffer D with 0.1 mg/mL BSA, followed by addition of 5 μM Alb3CT. Fluorescence was monitored on a Fluorolog 3–22 spectrofluorometer (JobinYvon) using an excitation wavelength of 505 nm and emission wavelength of 525 nm. A control release reaction was performed using unlabeled cpSRP43 to obtain the intensity for the donor sample (ID). FRET efficiency was calculated as

$$FRET\ efficiency = \frac{I_D - I_{DA}}{I_D}$$

in which IDA is the fluorescence intensity of donor in the presence of acceptor-labeled cpSRP43.

NMR Spectroscopy. Multidimensional NMR spectra were recorded on Bruker Advance spectrometers operating at 800 and 900 MHz. All NMR spectra were acquired at 17 °C using 2H, 15N-labeled or 2H, 15N, 13C-labeled protein (~0.2 mM) in NMR buffer containing 10% (vol/vol) D2O at pH 6.5. Titration experiments were performed by serial addition of unlabeled ligands into the NMR sample containing 2H, 15N-labeled cpSRP43. NMR data were processed with NMRPipe (42) and analyzed with NMRView Java (43). Details on NMR sample preparation and assignments are described in SI Materials and Methods.

1. Hardy SJ, Randall LL (1991) A kinetic partitioning model of selective binding of nonnative proteins by the bacterial chaperone SecB. Science 251(4992):439–443.

2. Randall LL, Hardy SJ (2002) SecB, one small chaperone in the complex milieu of the cell. Cell Mol Life Sci 59(10):1617–1623.

3. Walton TA, Sandoval CM, Fowler CA, Pardi A, Sousa MC (2009) The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. Proc Natl Acad Sci USA 106(6):1772–1777.

4. Hennecke G, Nolte J, Volkmer-Engert R, Schneider-Mergener J, Behrens S (2005) The periplasmic chaperone SurA exploits two features characteristic of integral outer membrane proteins for selective substrate recognition. J Biol Chem 280(25):23540–23548.

5. Thoma J, Burmann BM, Hiller S, Müller DJ (2015) Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. Nat Struct Mol Biol 22(10):795–802.

6. Mihara K, Omura T (1996) Cytoplasmic chaperones in precursor targeting to mitochondria: The role of MSF and hsp 70. Trends Cell Biol 6(3):104–108.

7. Zhuravleva A, Clerico EM, Gierasch LM (2012) An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones. Cell 151(6): 1296–1307.

8. Fenton WA, Horwich AL (1997) GroEL-mediated protein folding. Protein Sci 6(4):743–760.

9. Tapley TL, et al. (2009) Structural plasticity of an acid-activated chaperone allows promiscuous substrate binding. Proc Natl Acad Sci USA 106(14):5557–5562.

10. Tapley TL, Franzmann TM, Chakraborty S, Jakob U, Bardwell JC (2010) Protein refolding by pH-triggered chaperone binding and release. Proc Natl Acad Sci USA 107(3):1071–1076.

11. Reichmann D, et al. (2012) Order out of disorder: Working cycle of an intrinsically unfolded chaperone. Cell 148(5):947–957.

12. Ouyang M, et al. (2011) LTD is a protein required for sorting light-harvesting chlorophyll-binding proteins to the chloroplast SRP pathway. Nat Commun 2:277.

13. Schuenemann D, et al. (1998) A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. Proc Natl Acad Sci USA 95(17): 10312–10316.

14. Schünemann D (2004) Structure and function of the chloroplast signal recognition particle. Curr Genet 44(6):295–304.

15. Franklin AE, Hoffman NE (1993) Characterization of a chloroplast homologue of the 54-kDa subunit of the signal recognition particle. J Biol Chem 268(29):22175–22180.

16. Li X, Henry R, Yuan J, Cline K, Hoffman NE (1995) A chloroplast homologue of the signal recognition particle subunit SRP54 is involved in the posttranslational integration of a protein into thylakoid membranes. Proc Natl Acad Sci USA 92(9):3789–3793.

17. Tu CJ, Schuenemann D, Hoffman NE (1999) Chloroplast FtsY, chloroplast signal recognition particle, and GTP are required to reconstitute the soluble phase of lightharvesting chlorophyll protein transport into thylakoid membranes. J Biol Chem 274(38):27219–27224.

18. Falk S, Sinning I (2010) cpSRP43 is a novel chaperone specific for light-harvesting chlorophyll a,b-binding proteins. J Biol Chem 285(28):21655–21661.

19. Jaru-Ampornpan P, et al. (2010) ATP-independent reversal of a membrane protein aggregate by a chloroplast SRP subunit. Nat Struct Mol Biol 17(6):696–702.

20. Stengel KF, et al. (2008) Structural basis for specific substrate recognition by the chloroplast signal recognition particle protein cpSRP43. Science 321(5886):253–256.

21. DeLille J, et al. (2000) A novel precursor recognition element facilitates posttranslational binding to the signal recognition particle in chloroplasts. Proc Natl Acad Sci USA 97(4):1926–1931.

22. Jaru-Ampornpan P, et al. (2013) Mechanism of an ATP-independent protein disaggregase:II. distinct molecular interactions drive multiple steps during aggregate disassembly. J Biol Chem 288(19):13431–13445.

23. Holdermann I, et al. (2012) Chromodomains read the arginine code of post-translational targeting. Nat Struct Mol Biol 19(2):260–263.

24. Gao F, et al. (2015) Regulation of structural dynamics within a signal recognition particle promotes binding of protein targeting substrates. J Biol Chem 290(25): 15462–15474.

25. Falk S, Ravaud S, Koch J, Sinning I (2010) The C terminus of the Alb3 membrane insertase recruits cpSRP43 to the thylakoid membrane. J Biol Chem 285(8):5954–5962.

26. Falk S, Sinning I (2010) The C terminus of Alb3 interacts with the chromodomains 2 and 3 of cpSRP43. J Biol Chem 285(53):le25–le26.

27. Lewis NE, et al. (2010) A dynamic cpSRP43-Albino3 interaction mediates translocase regulation of chloroplast signal recognition particle (cpSRP)-targeting components. J Biol Chem 285(44):34220–34230.

28. Dünschede B, Bals T, Funke S, Schünemann D (2011) Interaction studies between the chloroplast signal recognition particle subunit cpSRP43 and the full-length translocase Alb3 reveal a membrane-embedded binding region in Alb3 protein. J Biol Chem 286(40):35187–35195.

29. Horn A, et al. (2015) Structural basis for cpSRP43 chromodomain selectivity and dynamics in Alb3 insertase interaction. Nat Commun 6:8875.

30. Tzvetkova-Chevolleau T, et al. (2007) Canonical signal recognition particle components can be bypassed for posttranslational protein targeting in chloroplasts. Plant Cell 19(5):1635–1648.

31. Tu CJ, Peterson EC, Henry R, Hoffman NE (2000) The L18 domain of light-harvesting chlorophyll proteins binds to chloroplast signal recognition particle 43. J Biol Chem 275(18):13187–13190.

32. Funke S, Knechten T, Ollesch J, Schünemann D (2005) A unique sequence motif in the 54-kDa subunit of the chloroplast signal recognition particle mediates binding to the 43-kDa subunit. J Biol Chem 280(10):8912–8917.

33. Sivaraja V, et al. (2005) Three-dimensional solution structures of the chromodomains of cpSRP43. J Biol Chem 280(50):41465–41471.

34. Chakraborty A, et al. (2002) A single-domain cyclophilin from Leishmania donovani reactivates soluble aggregates of adenosine kinase by isomerase-independent chaperone function. J Biol Chem 277(49):47451–47460.

35. Kargul J, Laurent GJ (2012) Small heat shock proteins: Molecular protectors against the disease. Int J Biochem Cell Biol 44(10):1587.

36. Laskowska E, Matuszewska E, Kuczy_nska-Wi_snik D (2010) Small heat shock proteins and protein-misfolding diseases. Curr Pharm Biotechnol 11(2):146–157.

37. Mogk A, Deuerling E, Vorderwülbecke S, Vierling E, Bukau B (2003) Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. Mol Microbiol 50(2):585–595.

38. Veinger L, Diamant S, Buchner J, Goloubinoff P (1998) The small heat-shock protein IbpB from Escherichia coli stabilizes stress-denatured proteins for subsequent refolding by a multichaperone network. J Biol Chem 273(18):11032–11037.

39. Bardwell JC, Jakob U (2012) Conditional disorder in chaperone action. Trends Biochem Sci 37(12):517–525.

40. Foit L, George JS, Zhang BW, Brooks CL, 3rd, Bardwell JC (2013) Chaperone activation by unfolding. Proc Natl Acad Sci USA 110(14):E1254–E1262.

41. Jaru-Ampornpan P, Chandrasekar S, Shan SO (2007) Efficient interaction between two GTPases allows the chloroplast SRP pathway to bypass the requirement for an SRP RNA. Mol Biol Cell 18(7):2636–2645.

42. Delaglio F, et al. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6(3):277–293.

43. Johnson BA, Blevins RA (1994) NMR View: A computer program for the visualization and analysis of NMR data. J Biomol NMR 4(5):603–614.

44. Salzmann M, Pervushin K, Wider G, Senn H, Wüthrich K (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. Proc Natl Acad Sci USA 95(23):13585–13590.

45. Salzmann M, Wider G, Pervushin K, Senn H, Wuthrich K (1999) TROSY-type tripleresonance experiments for sequential NMR assignments of large proteins. J Am Chem Soc 121(4):844–848.

46. Lebowitz J, Lewis MS, Schuck P (2002) Modern analytical ultracentrifugation in protein science: A tutorial review. Protein Sci 11(9):2067–2079.

47. Schuck P, Perugini MA, Gonzales NR, Howlett GJ, Schubert D (2002) Size-distribution analysis of proteins by analytical ultracentrifugation: Strategies and application to model systems. Biophys J 82(2):1096–1111.

Supporting Information

SI Materials and Methods

Protein Production and Purification for NMR. Isotope-labeled full-length cpSRP43 and CD1Ank-BH were overexpressed in BL21(DE3) cells at 37 °C in M9 D2O media supplemented with 15NH4SO4 and 12C or 13C-glucose. Protein expression was induced at OD600 = 0.7 by addition of 1 mM IPTG for 15 h. Isotope-labeled cpSRP43 and CD1Ank-BH were purified using the same protocol as unlabeled cpSRP43. Proteins were exchanged into NMR buffer (50 mM phosphate, 150 mM NaCl, pH 6.5) using a PD MidiTrap G-25 column (GE Healthcare).

Assignment of cpSRP43 Backbone. Triple labeled (2H, 15N, 13C) cpSRP43 and CD1Ank-BH (~0.8 mM) were prepared in NMR buffer containing 10% (vol/vol) D2O. Backbone resonance assignments were made using 3D TROSY-HNCA, TROSY-HN(CO)CA, TROSY-HNCACB, and TROSY-HN(CO)CACB spectra (44, 45). A total of 252 of 323 nonproline backbone residues were assigned using RunAbout in NMRview java. Many of the cross-peaks associated with CD2 are weak, and only a subset (residues 278–287) could be assigned from the triple resonance spectra. Assignments for additional CD2 residues (F267, D273, I275, E277, L288, V289, W291, D293, G294, W299, V300, G302, D308, V309, and K311) were made by transferring published assignments for the isolated CD2 domain (33).

Characterization of the Oligomeric State of cpSRP43. The oligomeric state of cpSRP43 was assessed by two methods. First, cpSRP43 in the presence or absence of equal molar HiLyte-Fluor488–labeled L11 peptide were analyzed by size exclusion chromatography using Superdex 200 column (GE Healthcare). The column was equilibrated in 50 mM KHepes, pH 7.5, with either 200 mM or 50 mM NaCl. The protein elution profile was confirmed by SDS/PAGE, and the molecular mass on the column was further calibrated using the LMW kit (GE Healthcare). Second, velocity sedimentation-type analytical ultracentrifugation was performed using Optima XL-I (Beckman Coulter) with an absorbance optical detection system (280 nm). The sample was spun at 50,000 rpm at 20 °C with 7 μM of cpSRP43 alone and in complex with equal molar of HiLyte-Fluor488–labeled L11 peptide or cpSRP54M. Buffer viscosity, protein partial specific volumes, and density were calculated using the SEDNTERP (46). The observed sedimentation data were fitted to a single component system by using the SEDFIT software (47), and the sedimentation coefficient distribution was extracted from the fitting.
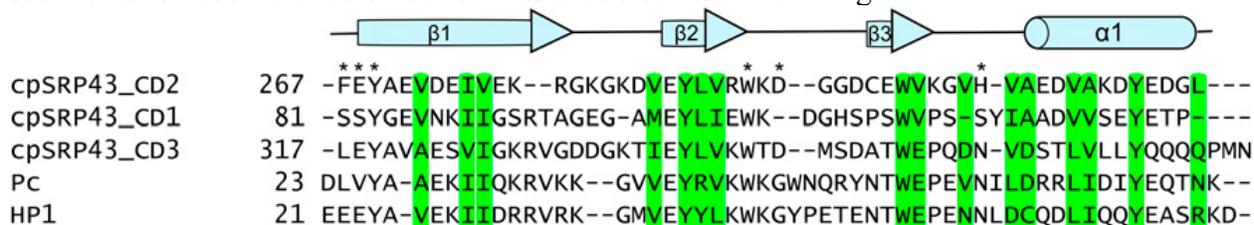

Fig. S1. Sequence alignment of cpSRP43 CD1, CD2, CD3, Polycomb (Pc), and heterochromatin protein 1 (HP1). Green highlights conserved hydrophobic residues in CD2 that were mutated in this study. Red asterisks denote aromatic cage residues in canonical

chromodomains. Black asterisks denote aromatic cage residues involved in 54M binding (23). Secondary structure information is labeled above the sequence highlighting β-sheet (β1, β2, β3) and α-helix (α1).
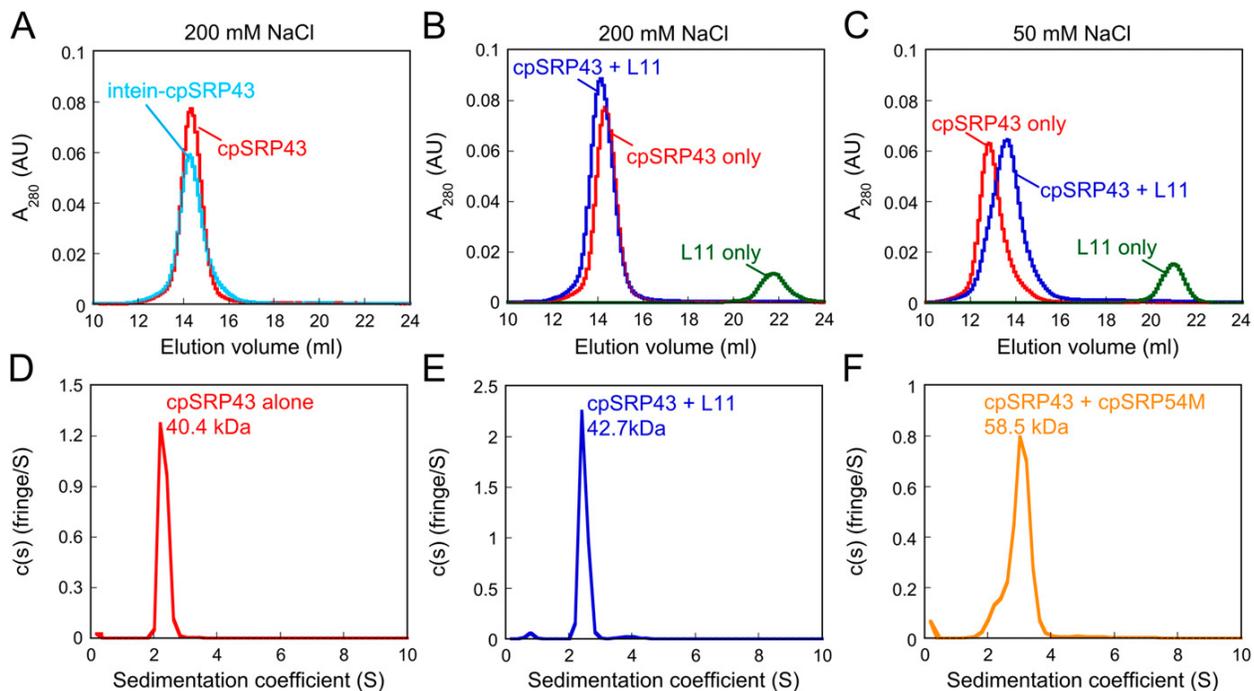
Fig. S2. CpSRP43 is active as a monomer. (A) WT (red) and superactive intein-cpSRP43 (cyan) runs as a monomer on Superdex 200 column with buffer containing 200 mM NaCl. (B) The complex of WT cpSRP43 and HiLyte-Fluor488–labeled L11 peptide (blue) was eluted as a 1:1 complex. (C) At lower ionic strength (50 mM NaCl), cpSRP43 exhibits oligomeric forms (red), but L11 binding shifts cpSRP43 to a lower molecular weight complex (blue). (D–F) Sedimentation coefficient distributions calculated from a velocity sedimentation experiment of cpSRP43 alone (D), cpSRP43 with HiLyte-Fluor488–labeled L11 peptide (E), and cpSRP43 with cpSRP54M (F) using buffer containing 200 mMNaCl. The experimental molecular mass is close to the predicted values of cpSRP43 and cpSRP54M (36 and 22 kDa, respectively), suggesting that the active form of cpSRP43 is monomer.
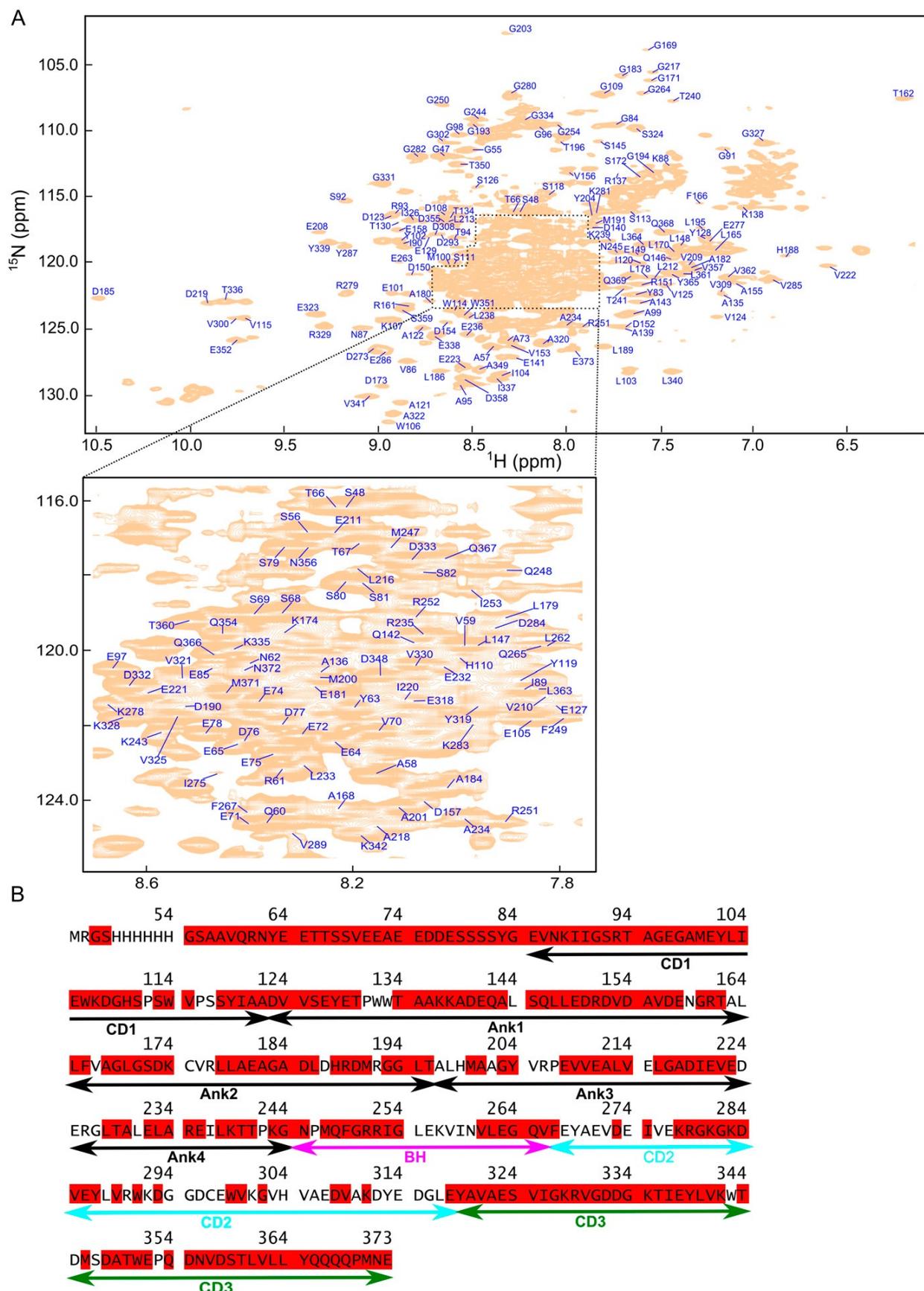
Fig. S3. Assignments of the cpSRP43 spectra. (A) 800 MHz 1H-15N TROSY spectrum of full-length 15N, 13C, 2H-labeled cpSRP43 is shown with the assigned residues indicated. The central area of the spectrum is enlarged for clarity. (B) Assigned residues are highlighted in red in the cpSRP43 sequence.
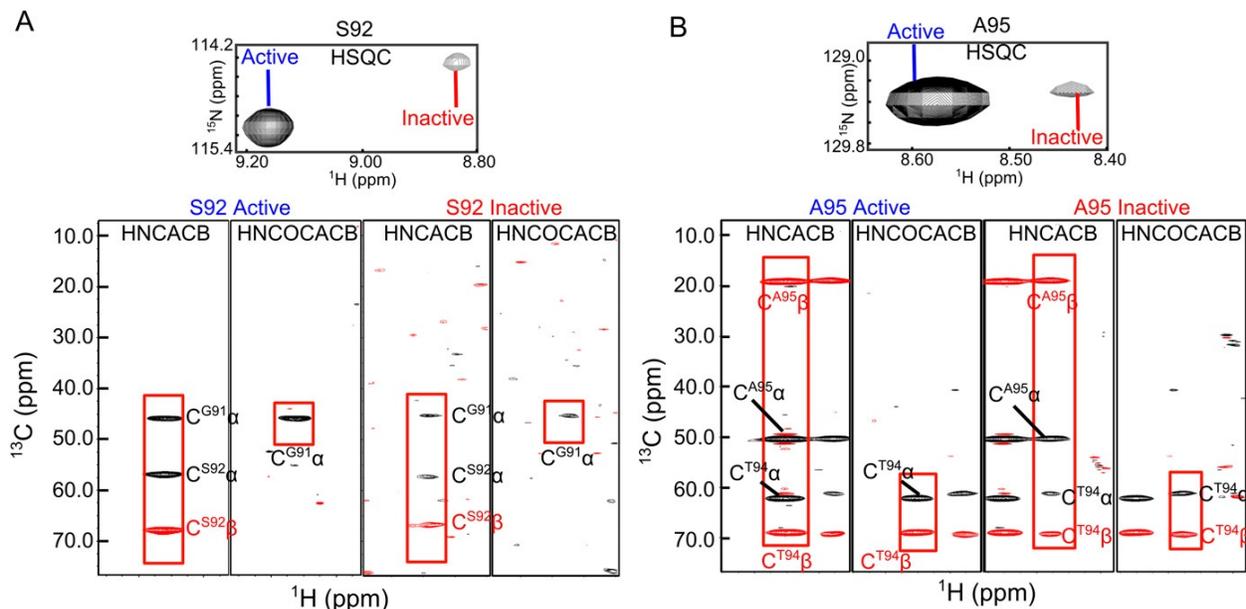
Fig. S4. Pairs of component cross-peaks in CD1 arise from the same amino acids. (Upper) Component cross-peaks of residues Ser92 (A) and Ala95 (B) in the TROSY spectrum of 2H, 15N, 13C-labeled CD1Ank-BH. (Lower) Strips from the HNCACB and HN(CO)CACB spectra of 15N, 13C, 2H-labeled CD1Ank-BH for the two component peaks of S92 and A95. For each residue, the two component peaks have connectivities to identical Cα and Cβ chemical shifts for residues i and i-1, indicating that they arise from the same amino acid. The red rectangles highlight the connectivities for the active state (Left) and inactive state (Right).
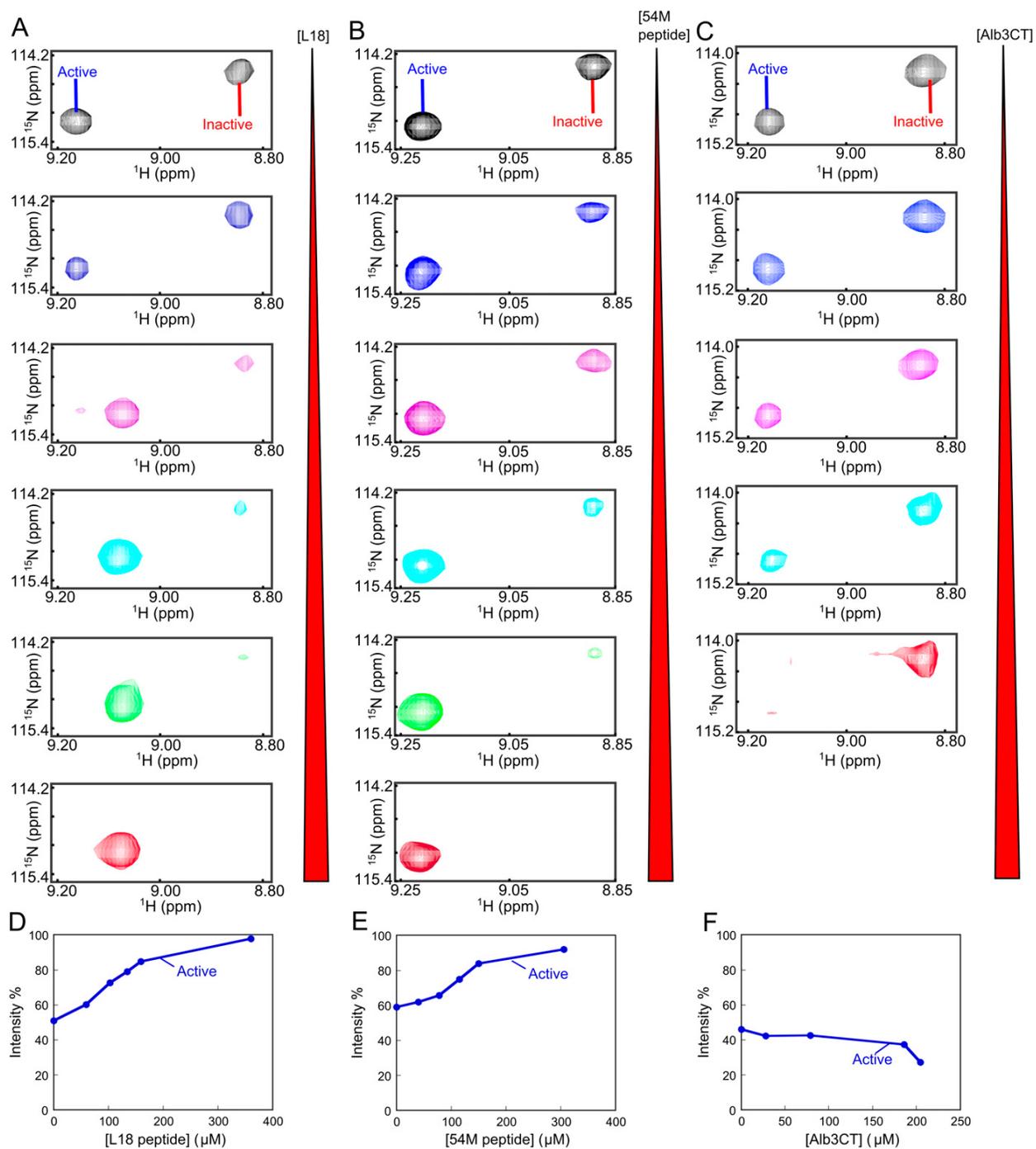
Fig. S5. The conformation of cpSRP43 is regulated by binding partners. (A–C) Titration of the L18 peptide (A), 54M peptide (B), and Alb3CT (C) into the 2H,15N-labeled cpSRP43 during the TROSY experiment. Component cross-peaks are shown for Ser92. (D–F) Quantification of the relative intensity of the component cross-peaks from the data in A–C.
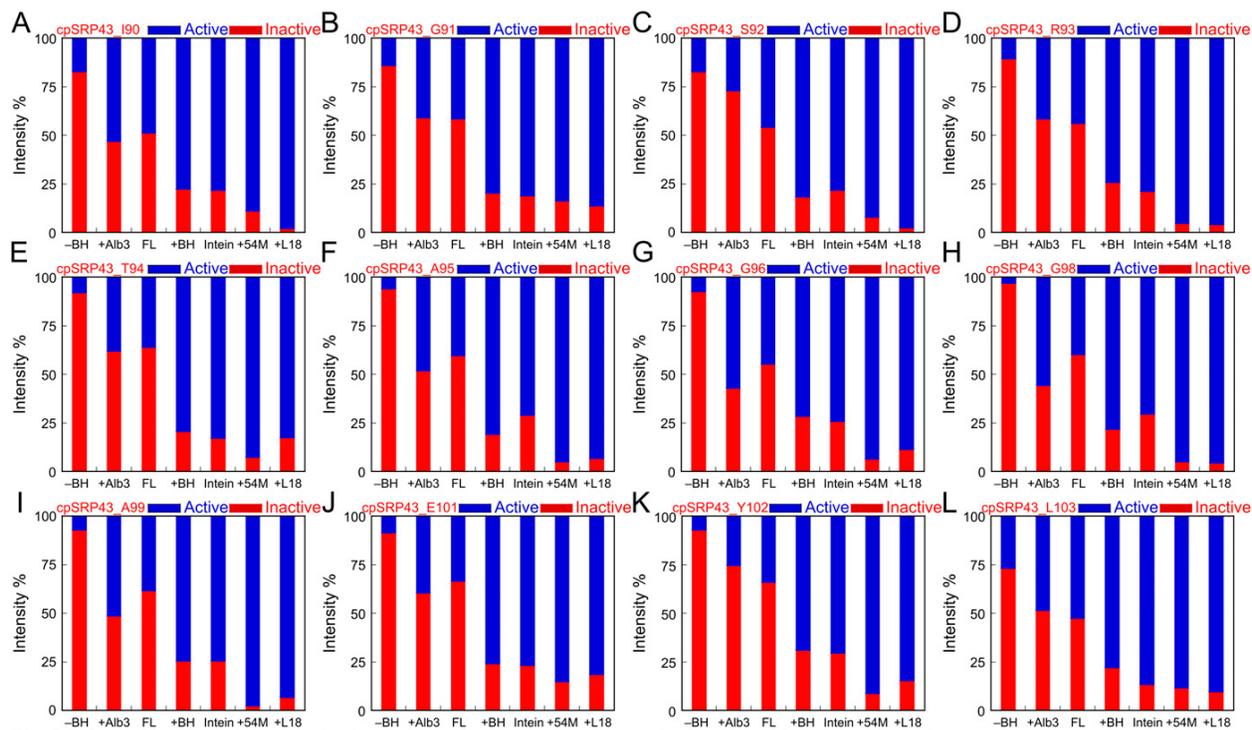
Fig. S6. Quantification of the relative intensities of component cross peaks for 12 residues in various chaperone constructs and with various ligands bound.–BH, CD1Ank fragment; +Alb3, full-length cpSRP43 + Alb3CT; FL, full-length cpSRP43; +BH, CD1Ank-BH fragment; Intein, Intein-cpSRP43; +54M, full-length cpSRP43 + 54M peptide; +L18, full-length cpSRP43 + L18 peptide.
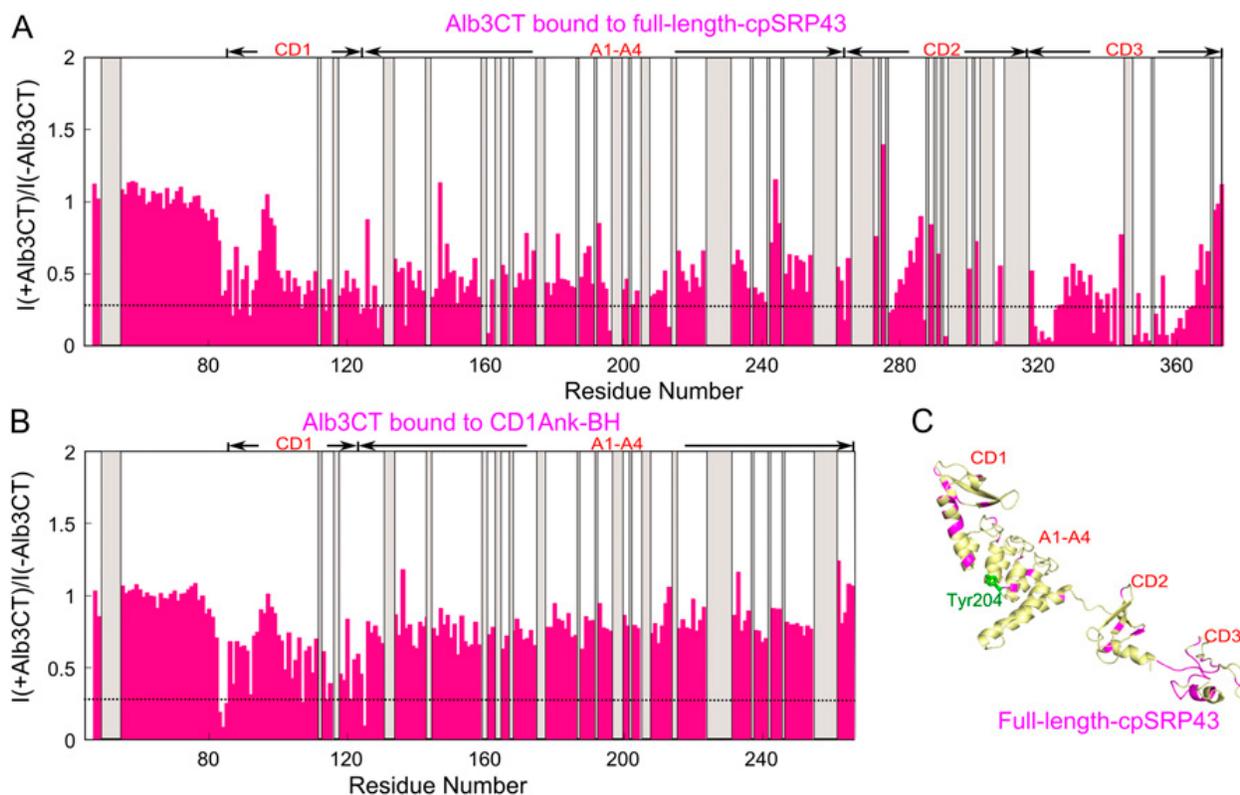
Fig. S7. Summary of the effect of Alb3CT on cross-peak intensities in the 1H-15N TROSY spectra of full-length cpSRP43 (A) and the CD1Ank-BH fragment (B). The intensities of each cross-peak in the presence and absence of Alb3CT [I(+Alb3CT) and I(−Alb3CT), respectively] were quantified and normalized to those of residues 1–63, a highly unstructured region whose intensities were unaffected by any binding partners, and their ratios were plotted. The gray bars denote unassigned residues (including all of the prolines), and the dashed lines show the cutoff where the peak is broadened ≥70%. (C) Structural model for full-length cpSRP43 (19). Dark red highlights residues whose cross-peaks are broadened >70% (from A) on binding to Alb3CT. Green highlights Tyr204 that binds L18.
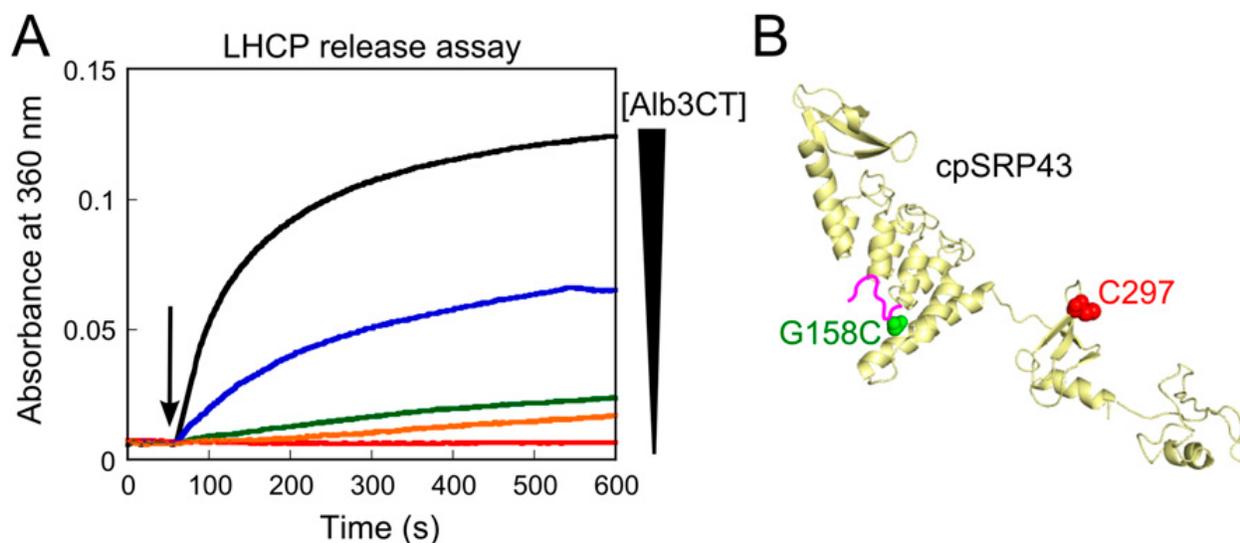
Fig. S8. (A) Alb3CT induces dose-dependent release of LHCP TMDs from cpSRP43, performed as outlined in Fig. 5F and described in Materials and Methods. Red, orange, green, blue, and black denote data obtained with 0, 0.625, 1.25, 2.5, and 5.0 μM Alb3CT, respectively. Errors were estimated to be ±10% (SD) based on three technical replicates. (B) Structural model of cpSRP43 highlighting

the positions of the FRET probes. Residue G158 in the L18 motif of LHCP (green) was mutated to cysteine and labeled with Atto488 (donor), and the native C297 in cpSRP43 (red) was labeled with TMR.

A

[54M peptide]

Peak1   Peak2

B

cpSRP43_G264   ▬ Peak2
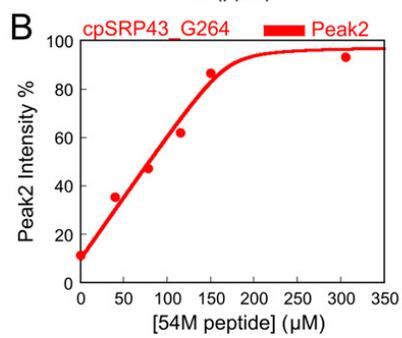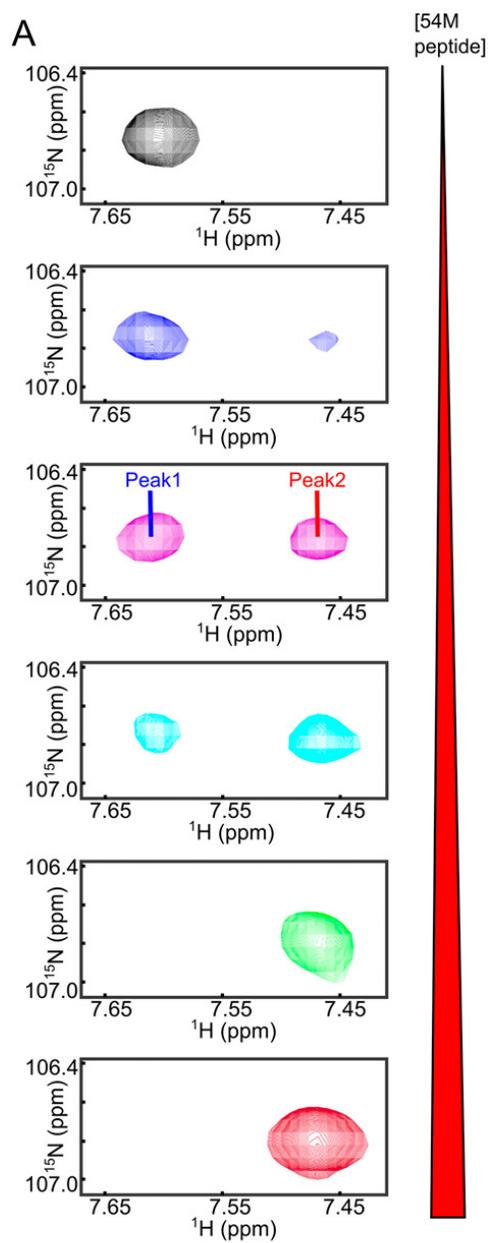
Fig. S9. Effects of 54M binding on the intensity of G264 (in BH). (A) Region of the TROSY spectrum of 2H, 15N-labeled cpSRP43 showing the effects of increasing concentrations of the 54M peptide on the G264 cross-peaks corresponding to unbound and 54M-bound cpSRP43 (labeled as peak 1 and peak 2, respectively). (B) Quantification of the relative intensities for the 54M-bound peak (peak 2) for G264. The line is a fit of the data to Eq. 1.

Additional Supplementary Material (Unpublished, performed by C.Z.M.)

Linker Mutant Prevention Data. The full linker mutant binding data is shown in Supplementary Figure 1. As this figure shows, changes in binding between the linker mutants are subtle. From this data, wild-type cpSRP43 has a Kd of 177 nM, the most flexible mutant, (GS)5, has a Kd of 100 nM, the least flexible mutant, 2P, has a Kd of 414 nM, the deltalinker mutant has a Kd of 211 nM, and 2G has a Kd of 186 nM.
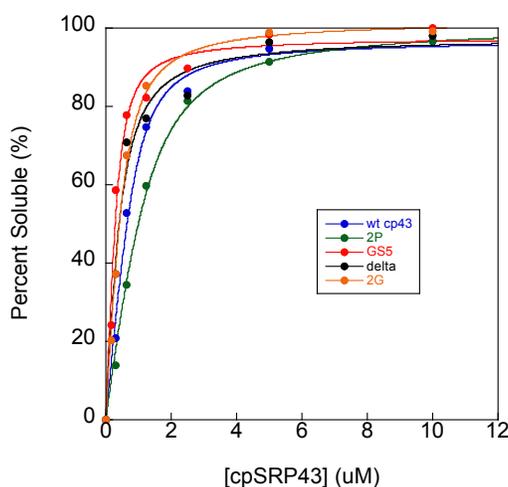
Additional Supplementary Figure 1. Full linker mutant characterization from LHCP aggregation/prevention experiments.

Linker Mutant Disaggregation Data. The full linker mutant disaggregation data is shown in Supplementary Figure 2. From this data, a correlation can be observed between linker mutant flexibility and disaggregation rate, namely that the more flexible linker, (GS)5, has a faster rate and the more rigid linker, 2P, has a slower rate. Similarly, end point analysis shows this same relationship between liker mutant flexibility and percent disaggregated at the end of the disaggregation experiment. This analysis is summarized in Supplementary Figure 3.

(a) (b)
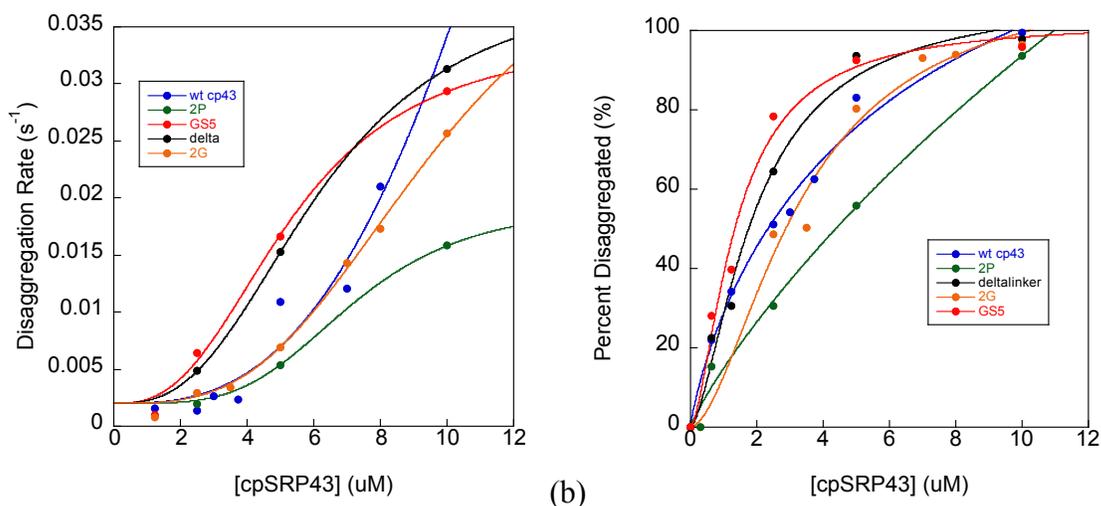
Additional Supplementary Figure 2. Full linker mutant characterization from LHCP disaggregation experiments. (a) A correlation can be observed between linker mutant flexibility and disaggregation rate. (b) Likewise, end point analysis shows a relationship between linker mutant flexibility and percent disaggregated at the end of the disaggregation experiment.
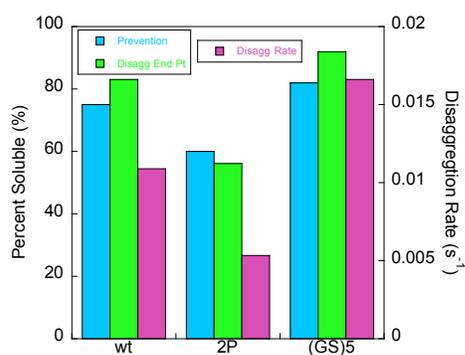
Supplementary Figure 3. Summary of prevention and disaggregation linker mutant data for defective 2P mutant and superactive (GS)5 mutant showing prevention, disaggregation endpoint, and disaggregation rate data.

C h a p t e r   3

AN UNFOLDED-FOLDED TRANSITION DRIVES CHAPERONE-MEMBRANE PROTEIN INTERACTION

An ATP-Independent Chaperone Undergoes an Unfolded-Folded Transition Upon Interaction with Substrate

Camille Z. McAvoy*, Fu-Cheng Liang*, Vinh Lam, Emily Miaou, Gerard Kroon, Peter Wright, Shu-ou Shan
*Authors contributed equally to this work

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125

Correspondence: sshan@caltech.edu

C.Z.M. Performed light scattering and EPR.

F-C.L. Performed NMR.

V.L. Performed HDX.

E.M. Assisted with light scattering under the supervision of C.Z.M.

G.K. Assisted with NMR.

Abstract

The chloroplast signal recognition particle component 43, cpSRP43, acts as an ATP-independent chaperone to the light harvesting chlorophyll-binding proteins (LHCP). We have found that cpSRP43 interacts with substrate via an unfolded to folded transition. The substrate binding domain, and particularly the bridging helix region, of cpSRP43 is partially unfolded in the apo protein. Upon interaction with co-chaperone cpSRP54, cpSRP43 shows mild structuring, and upon interaction with both cpSRP54 and LHCP, cpSRP43 shows significant helix formation from an otherwise unstructured state. These results are consistent with NMR data that suggests structuring of cpSRP43 upon interaction with cpSRP54 and LHCP. Overall this work represents a novel characterization of the cpSRP43 structural state. Further this unstructured to structured transition could account for the binding energy of cpSRP43 with LHCP and explain cpSRP43's ATP independence.

Introduction

Proper protein folding and localization are critical for cellular protein homeostasis which is acutely challenged by the post-translational targeting of integral membrane proteins. Before arrival at the target membrane, nascent membrane proteins are highly prone to aggregation in the cytosol and other aqueous cellular compartments. Thus, effective molecular chaperones or chaperone networks are required to minimize improper exposure of the transmembrane domains (TMDs) on newly synthesized membrane proteins and to maintain them in soluble, translocation-competent conformation.

Results

A bridging helix mediates conformational change in cpSRP43. Using a light scattering assay and single cysteine mutants along the solvent-exposed surface of cpSRP43's bridging helix (BH), we found that the BH is very sensitive to mutation and further that defective or even dead mutants can be rescued with cpSRP54M-domain. The mutations made along the BH are highlighted in Figure 1A with neutral mutants in green, defective mutants in orange, and the single dead mutant I259C in red. This dead mutant was further studied using the light scattering assay (Figure 1B). It was found that in the presence of 54M-domain, aggregation was not observed in the presence of I259C cpSRP43 mutant with LHCP. This could have been due either to 54M directly binding substrate (although previous work has shown that cpSRP54 alone has no chaperone activity) or through 54M exerting a conformational change in I259C cpSRP43, allowing this otherwise dead mutant to become active. Using 54M peptide, a short fragment of cpSRP54M known to bind to cpSRP43, we observed similar rescue of LHCP as was seen in the 54M-domain case, suggesting that 54M acts through conformational change in cpSRP43 rather than directly binding substrate. As a final control, we utilized a cpSRP54M-domain mutant in which only the portion of 54M that binds cpSRP43 is mutated (deltaRRKRK 54M); there was no prevention of aggregation observed with cpSRP43 I259C and the deltaRRKRK 54M mutant suggesting that 54M binding to cpSRP43 is necessary for chaperone activity, further corroborating cpSRP43's role as chaperone and 54M's role as a modulator of conformational change in cpSRP43. The same rescue of activity with 54M-domain was observed for defective cpSRP43 mutant E256C as seen in Figure 1C. Further, for all defective BH mutants, 54M was able to rescue chaperone activity (Fig 1D and 1E, respectively). One possibility could have been that the BH was directly interacting with CD2 and that these mutations interfered with those reactions and 54M was then able to reestablish them. However, deleting CD2CD3 from cpSRP43 to arrive at the minimal substrate binding domain (SBD, see Liang et al. PNAS 2016) and inserting each of the defective mutants into this construct did not result in reversal of the defective phenotype, indicating that BH-CD2 interactions alone do not account for the conformational changes observed in the bridging helix mutants.

EPR indicates that the nature of conformational change in cpSRP43 is likely an unfolded-folded transition. EPR was performed using neutral mutants throughout cpSRP43's substrate binding domain (SBD) in the presence and absence of cpSRP54M-domain and LHCP substrate. The neutral mutants are indicated in Figure 2A (purple) and their chaperone activity is given in Figure 2B; all mutants used for EPR had chaperone activity within threefold of cysteineless WT. EPR spectra of all mutants in the apo form (black), plus 54M-domain form (blue), and plus 54M-domain and LHCP (green) are given in Figure 2C. From these spectra, central linewidth (1/deltaH, Figure 2D) and mobility (intensity of mobile/intensity of immobile peak, Figure 2E) analyses were performed to compare relative mobility of each of the residues studied. These data, particularly in the bridging helix region (mutants I253C and G264C), indicate a potential shift from a partially unstructured state in the apo form to a somewhat more structured state with the addition of a 54M-domain and a fully structured state in the presence of 54M-domain and LHCP. Based on previously established central linewidth standards, the bridging helix region, which we would

expect to be a well-structured alpha-helix, goes from an unstructured state to an alpha-helical one going from apo to with 54M-domain and LHCP forms. This same trend is observed in the mobility analysis in Figure 2E.

HDX further corroborates the unfolded-folded model of cpSRP43 conformational change.
Protection patterns obtained through hydrogen-deuterium exchange further indicate the role of conformational change in cpSRP43 binding. cpSRP43 in various backbones and with various ligands was subjected to HDX and perturbation maps were generated to study the protection patterns of cpSRP43 under these various conditions. The addition of either substrate (Lhcb5) or cpSRP54M-domain causes structuring in Ank2, Ank3, and the BH of cpSRP43 (Figure 3A and 3C, respectively); this result agrees with the EPR results, suggesting that binding of 54M-domain and substrate cause cpSRP43 to go an unstructured-to-structured transition.

Discussion

Through EPR, HDX, and NMR we have demonstrated that cpSRP43 undergoes a structural transition upon substrate binding and moreover that this transition likely involves going from a partially unstructured to structured state. This represents a novel characterization of cpSRP43 structural state that accounts for its ATP-independence. That is, a structural transition is coupled to binding substrate, forgoing the need for energy from ATP.

Materials and Methods

Protein Expression and Purification. Single cysteine mutants of Lhcb$_5$ and cpSRP43 were constructed using the QuikChange Mutagenesis procedure (Stratagene) according to manufacturer's instructions. WT and mutant cpSRP43, LHCP, and wildtype and mutant Lhcb5 were overexpressed and purified as previously described (Jaru-Ampornpan Mol Biol Cell 2007).

Chaperone Activity of cpSRP43. The ability of cpSRP43 to prevent LHCP aggregation was measured as described previously (Jaru-Ampornpan et al. NSMB 2010, Jaru-Ampornpan et al. JBC 2013). Aggregates were removed via ultracentrifugation in a TLA-100 rotor (Beckman Coulter) at 100,000 rpm for 30 min at 4 °C prior to the experiment. Light scattering experiments were performed by addition of 3 µL of 50 µM LHCP denatured 8M urea to 150 µL buffer D (50 mM KHEPES, pH 7.5 and 200 mM NaCl) or 2.5 µM cpSRP43 in buffer D. Light scattering was monitored at 360 nm on a UV-Vis spectrometer (Beckman Coulter) over time until equilibrium was reached. The percentage of soluble LHCP (% soluble) at equilibrium was plotted for each single-cysteine cpSRP43 mutant.

Site-directed spin labeling and EPR measurements. Spin labeling reactions were performed in 50 mM Hepes, pH 7.5, 200 mM NaCl. Reduced and degasses single cysteine mutants of cpSRP43 were labeled with five- to tenfold molar excess of MTSSL (Toronto Research Chemicals) at room temperature in the dark for 2-3 h. Excess MTSSL was removed by gel filtration. EPR spectra were acquired using a 9.4-GHz (X-band) EMX EPR spectrometer (Bruker) equipped with an ER 4119HS cavity at room temperature. The concentrations of spin-labeled samples were 30-100 uM cpSRP43. 5-30 scans were accumulated and averaged using microwave power of 5 mW with modulation amplitude set at 1 gauss and a magnetic field sweep width of 100 gauss. Less than 2% background labeling was observed; background subtraction was therefore not necessary.

Hydrogen-deuterium exchange mass spectrometry. Approximately 10 uM protein used for each injection. For each differential HDX experiment, 120 uL of apo and 120 uL of bound sample used. Each HDX experiment consisted of 7 time points (0, 10, 30, 60, 300, 900, and 3600 seconds) in 3 replicates. 5 uL of each protein sample is diluted with 20 uL H_2O/D_2O HDX buffer for each injection (time point). This 25 uL protein sample is quenched with 25 uL buffer and then injected for LC-MS. The spectra (apo versus bound) are then analyzed via HDX work bench software to produce perturbation maps.

Tables

Table 1. $K_{1/2}$ and K_d values for binding of LHCP to individual cpSRP43 mutants obtained from light scattering data with and without cpSRP54M. All cpSRP43 mutants shown in this table are derived from cysteine-less cpSRP43 (denoted as WT). 1 µM LHCP was used for all light scattering experiments and 1:1 cpSRP43:54M domain was used for +54M experiments.

Construct	-54M $K_{1/2}$ (nM)	% Soluble at Saturation	Hill Coefficient	R value	+54M K_d (nM)
WT	2100	85%	2.1	1.0	50
R252C	6800	88%	2.3	0.99	840
E256C	8400	80%	4.6	1.0	180
I259C	No binding	No binding	No binding	No binding	980
E263C	3500	100%	2.0	0.98	460

Figures
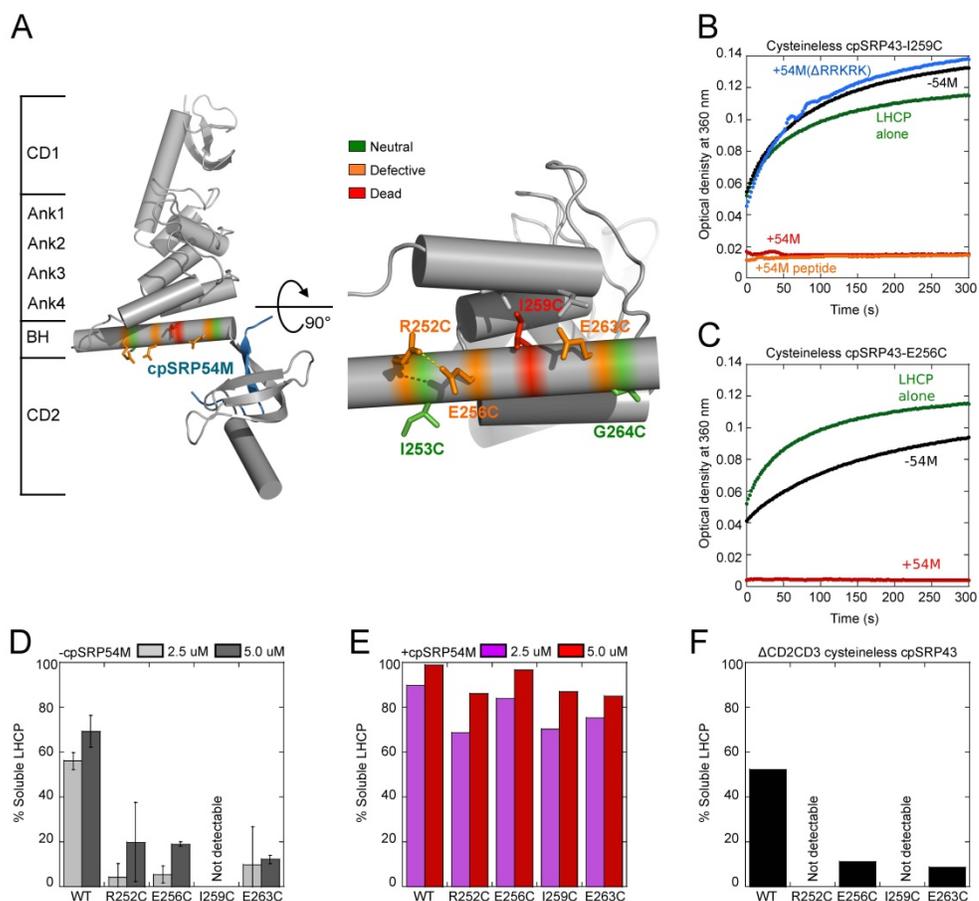
Figure 1

Figure 1. The bridging helix mediates the unfolded-folded conformational transition in cpSRP43. (A) Structure of cpSRP43 highlighting the residues in the bridging helix that were studied. **(B)** Chaperone activity spectra in the absence (green) and presence of cpSRP43-I259C mutant (black), and mutant plus 54M domain (light blue), 54M peptide (dark blue), and deltaRRKRK mutant (orange). This figure demonstrates that 54M induces a conformational change that allows cpSRP43 to perform its chaperone activity rather than 54M directly chaperoning LHCP. If 54M were acting as a chaperone we would not expect 54M peptide to have chaperone activity yet we see here that aggregation is prevented in the presence of the I259C cpSRP43 mutant and 54M peptide, implying that 54M induces a conformational change in cpSRP43. The deltaRRKRK 54M mutant does not bind to cpSRP43 and the observation that addition of this mutant does not restore chaperone activity is an indication that 54M must bind to cpSRP43 to induce chaperone activity. **(C)** Chaperone activity spectra in the absence (green) and presence of cysteineless cpSRP43-E256C mutant (black), and mutant plus 54M domain (light blue). **(D)** Activity of defective BH mutants in the absence of cpSRP54M-domain at 2.5 uM (light gray) and 5.0 uM (dark gray) concentrations of cpSRP43. **(E)** Activity of defective BH mutants in the presence of cpSRP54M-domain at 2.5 uM (purple) and 5.0 uM (red) concentrations of cpSRP43. **(F)** Activity of BH mutants in a deltaCD2CD3 cysteineless cpSRP43 backbone.
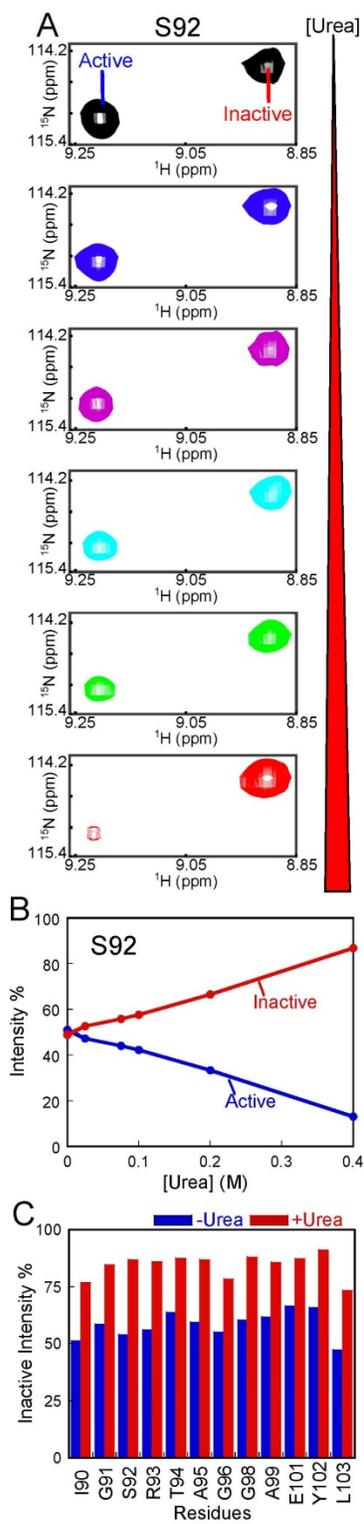
Figure 2

Figure 2. The component cross peaks of Ser92 in cpSRP43 are sensitive to urea unfolding. (A) Titration of the urea into the ^{2}H, ^{15}N-labeled full-length cpSRP43 during the TROSY experiment. Component cross peaks are shown for Ser92. (B) Quantification of the relative intensity of the component cross peaks, inactive (*red*) and active (*blue*), from the data in parts (*A*) are shown. (C) The relative intensities of the cross peaks corresponding to the inactive chaperone conformation are enhanced by urea (red bars) as compare to without urea (blue bar). Data are shown for indicated residues in ^{2}H, ^{15}N-labeled full length cpSRP43. Relative intensities were determined from cross peak heights.
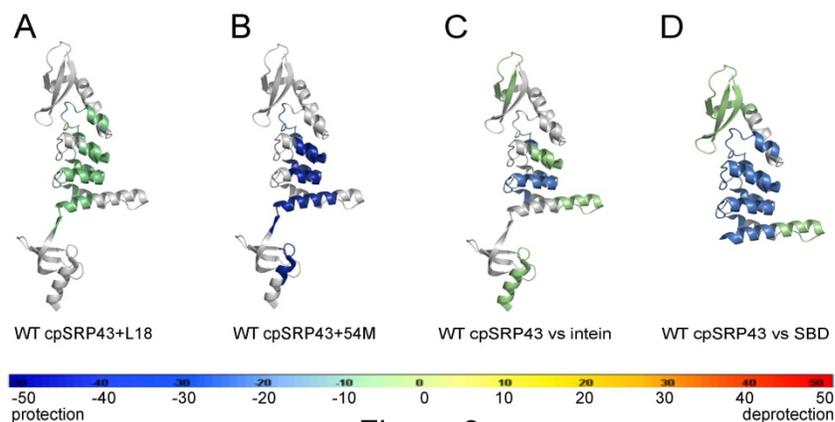
Figure 3

Figure 3. HDX perturbation results mapped onto cpSRP43 structure. (A) Full length cpSRP43 with L18. As indicated in the perturbation key, darker blue indicates more protection, whereas redder indicates deprotection, and gray indicates no change from the apo to the plus L18 state. (B) Protection pattern change from apo full-length cpSRP43 to full-length cpSRP43 with 54M (left). (C) Protection pattern change from WT cpSRP43 to superactive intein cpSRP43 (left). (D) Protection pattern change from WT cpSRP43 to superactive SBD cpSRP43 (left).
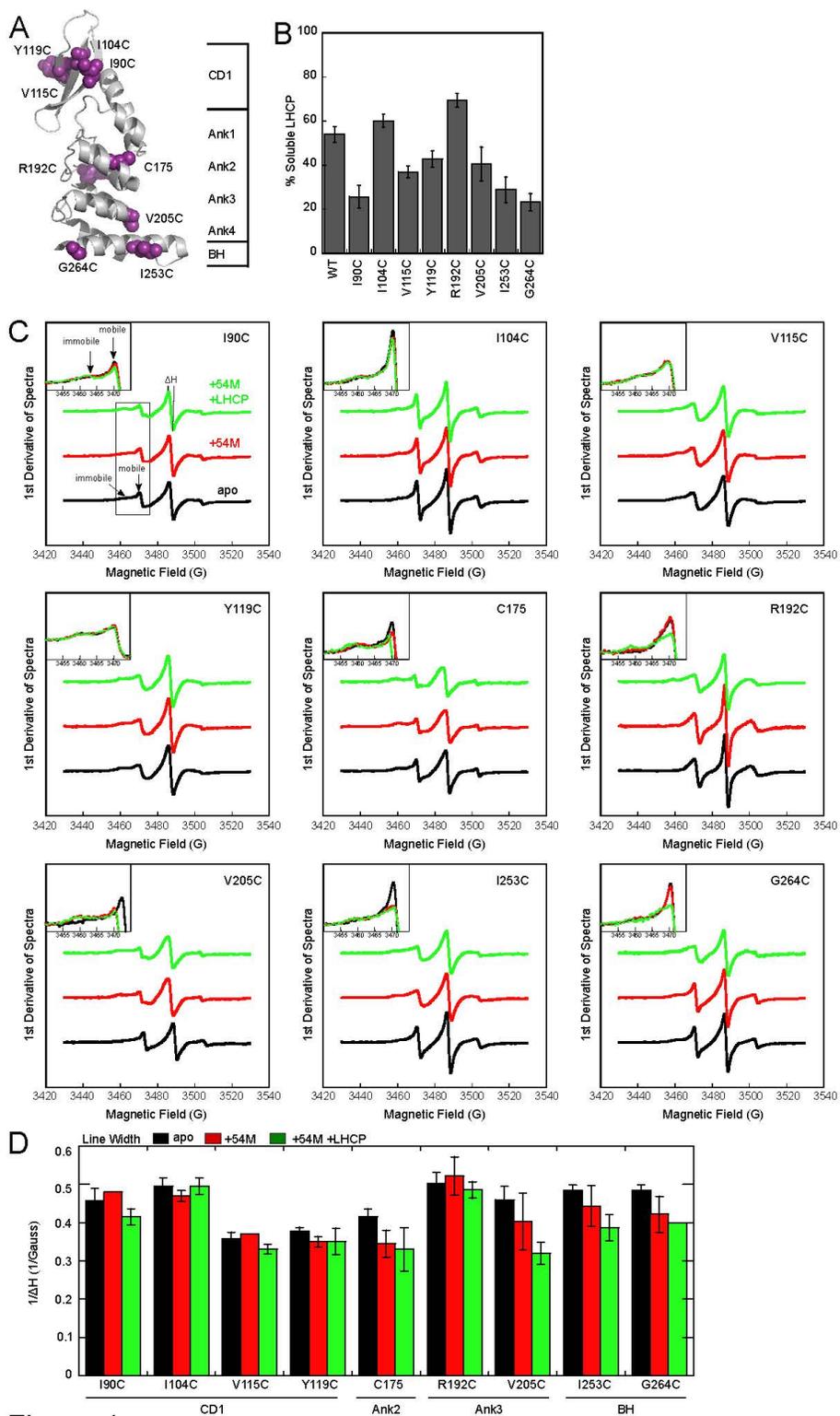
Figure 4

Figure 4. EPR analysis indicates an unstructured-structured transition in cpSRP43 upon interaction with cpSRP54M-domain and LHCP. (A) Structure of cpSRP43 highlighting sites of structural transition. (B) Chaperone activity data showing that all mutants used for EPR were neutral mutants within threefold activity of the cysteineless WT cpSRP43. (C) Representative EPR spectra of cpSRP43 alone (black) and with 54M domain (red) and 54M domain plus LHCP (green). (D) Linewidth analysis of EPR spectra show that cpSRP43 in the apo form appears to be partially unstructured (black) but shows mild structuring upon addition of cpSRP54 M domain (red) and significant structuring upon addition of cpSRP54 M domain and LHCP (green).
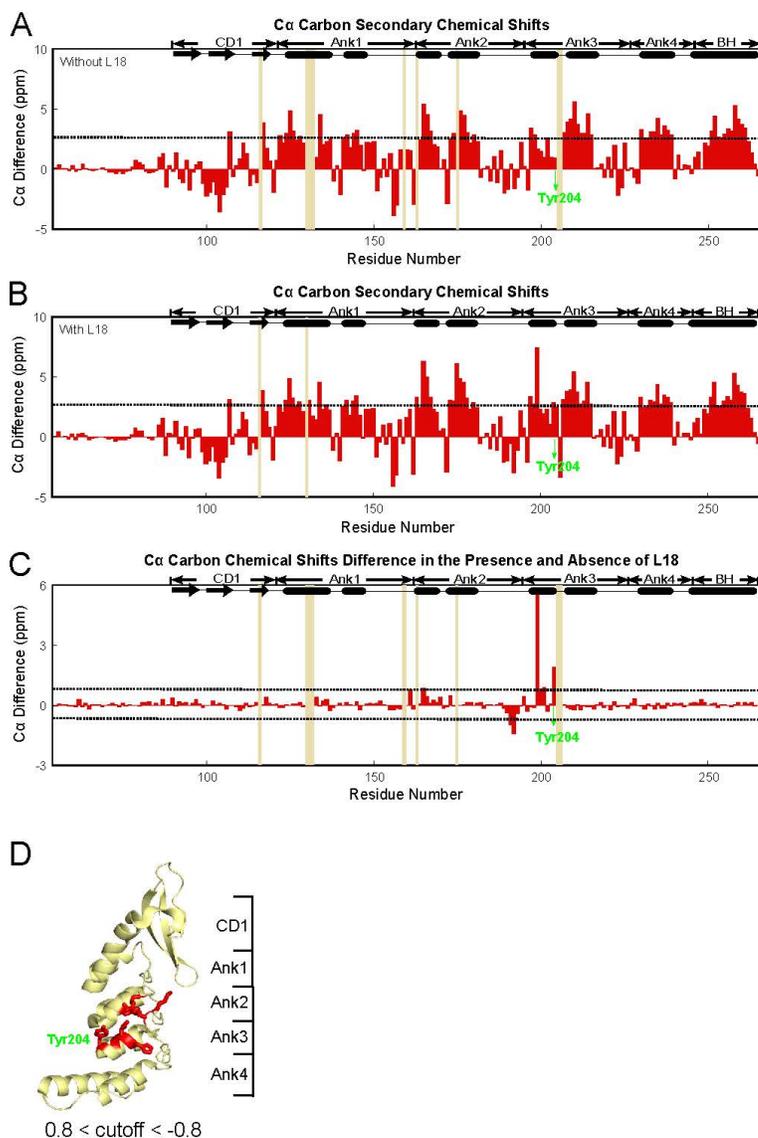
Figure 5. L18 binding induces secondary structure perturbation in ankyrin repeat region. (A and B) Secondary shift analysis of C$_{alpha}$ carbon shows the presence of L18 *(A)* induces more helical structure in the first helix of Ank3 as compared to the absence of L18 *(B)*. Dashed line is shown for the 2.8 ppm cutoff and yellow bar is shown for unassigned/ambiguous region. (C) C$_{alpha}$ carbon chemical shifts difference plot suggests L18 induced major structure perturbation in the first helix of Ank3 and some minor changes in Ank2. Dashed line is shown for the 0.8 ppm cutoff and yellow bar is shown for unassigned/ambiguous region. (D) NBD domain of cpSRP43 structure (PDB 3DEO) highlights the residues that have major C$_{alpha}$ chemical shift difference (*red*) upon the binding of L18. The cutoff 0.8ppm was used to show the major difference.
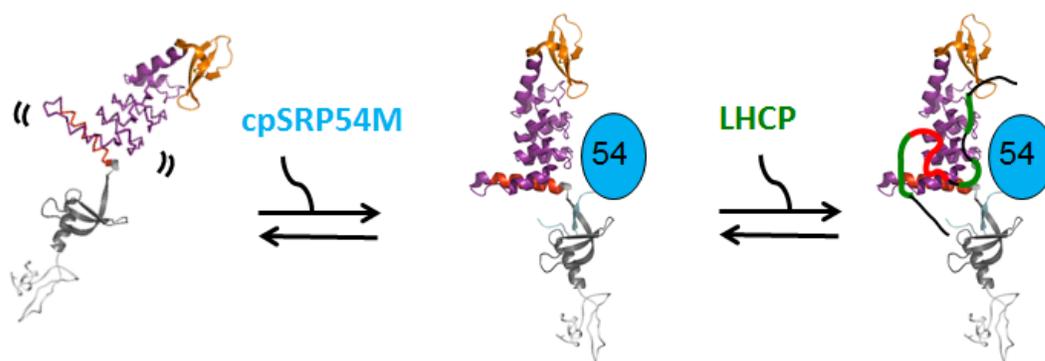
Figure 6. Final model of cpSRP43 structural transition. cpSRP43 exhibits spring-like motions going from a partially unstructured to structured state upon interaction with cpSRP54M-domain and substrate.

Supplementary Figures

A WT cpSRP43+L18
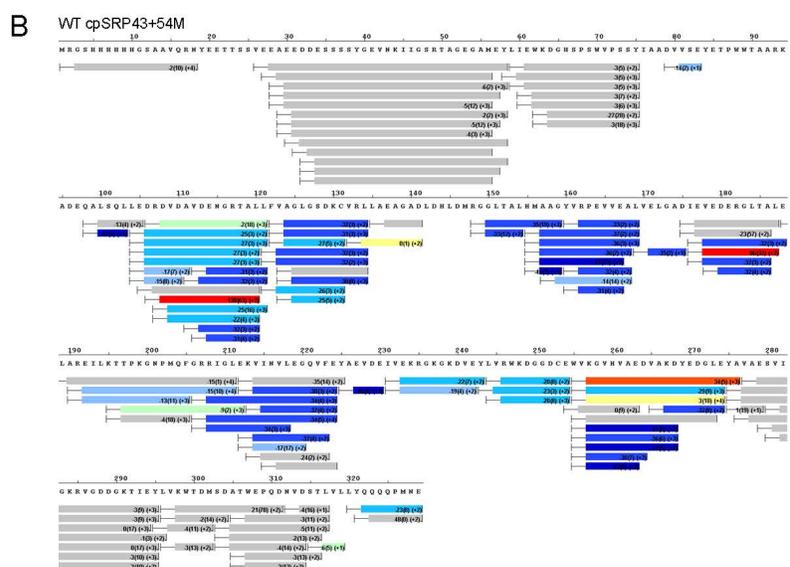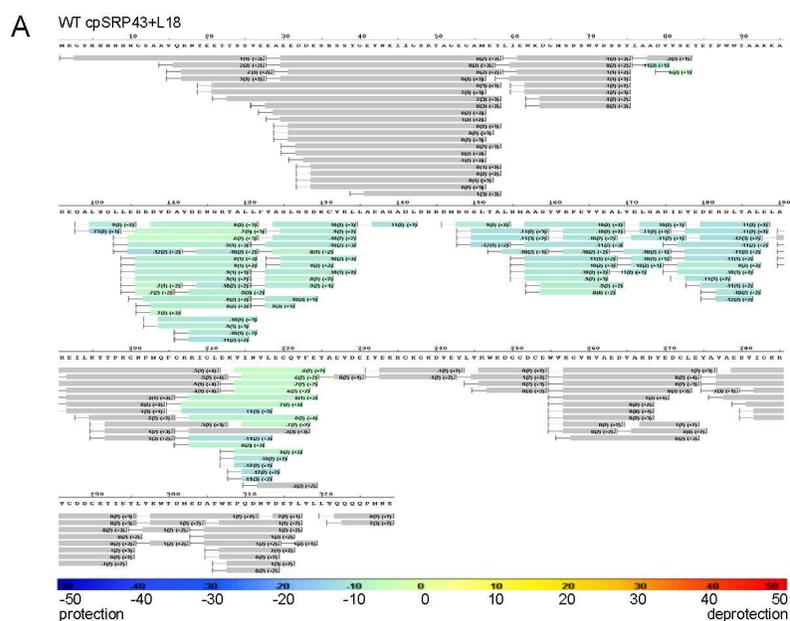
B WT cpSRP43+54M

Supplemental Figure 1

Supplementary Figure 1. Full HDX protection maps. (A) Perturbation map of full length WT cpSRP43 with L18 peptide. (B) Perturbation map of full length WT cpSRP43 with 54M domain.
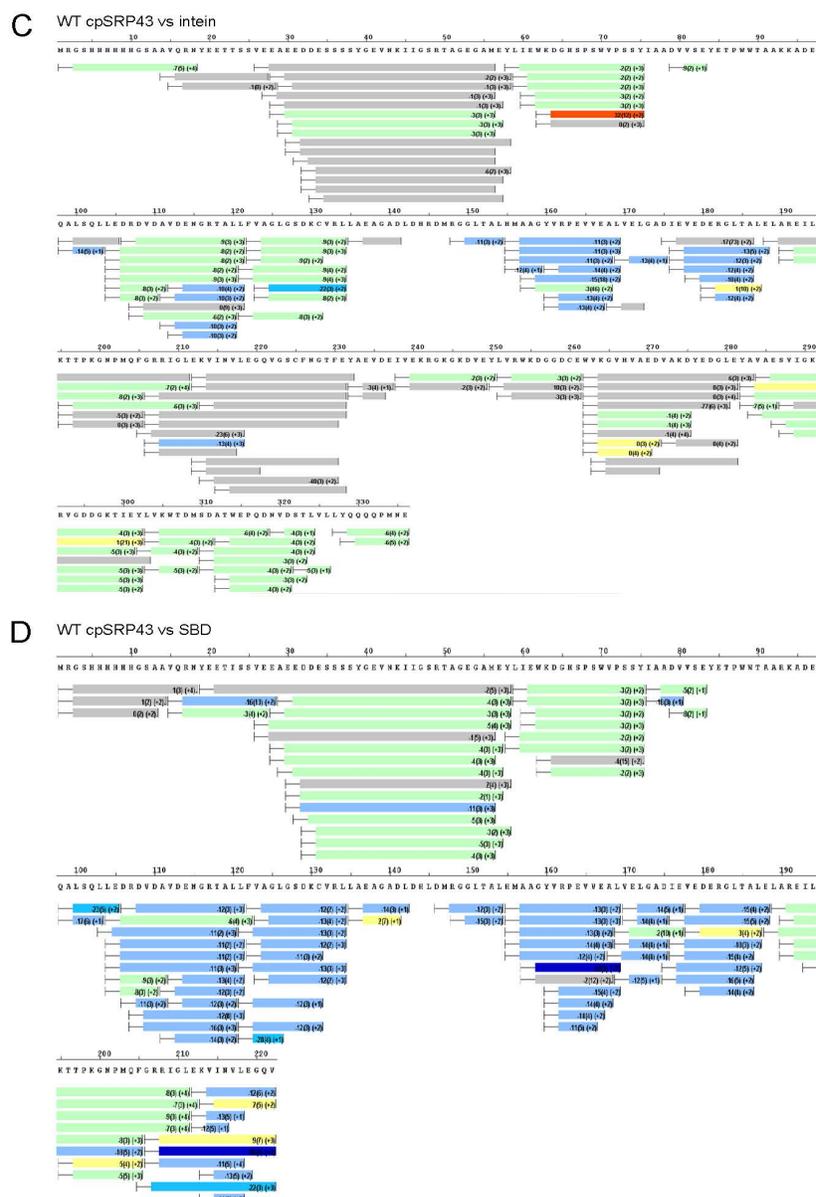
Supplemental Figure 2

Supplementary Figure 2. Full HDX protection maps. (C) Perturbation map of full length WT cpSRP43 vs superactive intein cpSRP43. **(D)** Perturbation map of full length WT cpSRP43 vs truncated superactive SBD alone.

C h a p t e r   4

INCREASING MEMBRANE PROTEIN EXPRESSION THROUGH CO-EXPRESSION WITH AN ATP-INDEPENDENT CHAPERONE

Camille Z. McAvoy, Thang X. Nguyen, Shu-ou Shan

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125

C.Z.M. Performed co-expression, Western blot, and analysis using ImageJ.

T.X.N. mentored C.Z.M.

Membrane protein and cpSRP43 co-expression

Membrane proteins are an important class of proteins whose study has been impeded by lacking overexpression yields. Approximately, 20-30% of genes encode membrane proteins and about 50% of drug targets are membrane proteins.[1,2,3] Although E. coli is the most widely used expression host, membrane protein yields are often insufficient for structural and functional studies.[1] Overexpression of both prokaryotic and eukaryotic membrane proteins leads to saturation of the Sec translocon which leads to: i) protein misfolding/aggregation in cytoplasm, ii) impaired respiration, and iii) activation of the Arc response, which leads to inefficient ATP production and the formation of acetate.[2,3]

Strategies to improve membrane protein expression have included engineering E coli strains, using organisms other than E. coli, engineering/selecting for membrane protein variants that show better expression, co-expression with chaperones, and expression using cell-free systems.[1] Although each of these strategies has shown some success, no strategy has yet been shown to be completely generalizable for broad membrane protein expression. Using our knowledge of the cpSRP43 chaperone, we seek to improve membrane protein yield by co-expressing membrane proteins with cpSRP43. Unlike other ATP-dependent disaggregases whose activity can be taxing on a cell's resources, cpSRP43 can act through recognition of an L18 fusion and promiscuous hydrophobic contacts with membrane proteins of interest. Our hypothesis is that cpSRP43 can help improve membrane protein expression by preventing the aggregation of membrane proteins as they are produced. Further, we theorize that cpSRP43 will help to keep membrane proteins in functional form by keeping them from going into inclusion bodies where helical bundle membrane proteins are rarely found to be functional.[18] Various pieces of evidence have hinted at cpSRP43's potential ability to chaperone non-native substrates, particularly data involving LHCP transmembrane domain (TM) swap mutants and preliminary studies of cpSRP43 with $A\beta_{40}$.

TM Swap Mutant Studies. LHCP consists of three transmembrane domains, an L18 motif, and N and C terminal regions, as shown in Figure 1.[4,5]

Figure 1. Light-harvesting, chlorophyll-binding protein (LHCP) organization (4,5). LHCP consists of three transmembrane domains (TM1, TM2, and TM3), an L18 motif, and N-and C-terminal portions (4,5).

To determine what interactions were important between cpSRP43 and substrate, the TMs of native LHCP were systematically exchanged with TMs from other membrane proteins, namely SERP1 and cytochrome B5. The SERP1 TM was swapped with LHCP TM2 and on another mutant it was swapped with LHCP TM3.[4] Likewise, a mutant was made in which cytochrome B5 was swapped with LHCP TM3.[4] Additionally, a truncation was made in which TM1 was removed from LHCP leaving only TM2, TM3, L18, and N- and C-terminal fragments.[4] From this truncated LHCP, cytochrome B5 was swapped with TM2 and SERP1 was swapped with TM3 to give a new

construct with none of the native LHCP TM's but with the same N-terminal, C-terminal, and L18 scaffold of LHCP as shown in Figure 2.[4,5]

Figure 2. LHCP mutant in which all native LHCP TM's have been replaced.[4,5] TM1 has been deleted; TM2 has been swapped for Cyb5, and TM3 has been swapped for Serp1. However, the L18 portion and scaffold (that is N-and C-terminal portions) are all the same as the native LHCP substrate.

Using this LHCP mutant with no native TM's, cpSRP43 was found to have comparable disaggregation activity and a modestly higher disaggregation rate than when using wild-type LHCP (Figure 3).[4,5] These results support the model that cpSRP43 interacts with its substrates in two ways: 1) through a specific L18 interaction and 2) through promiscuous hydrophobic interactions.
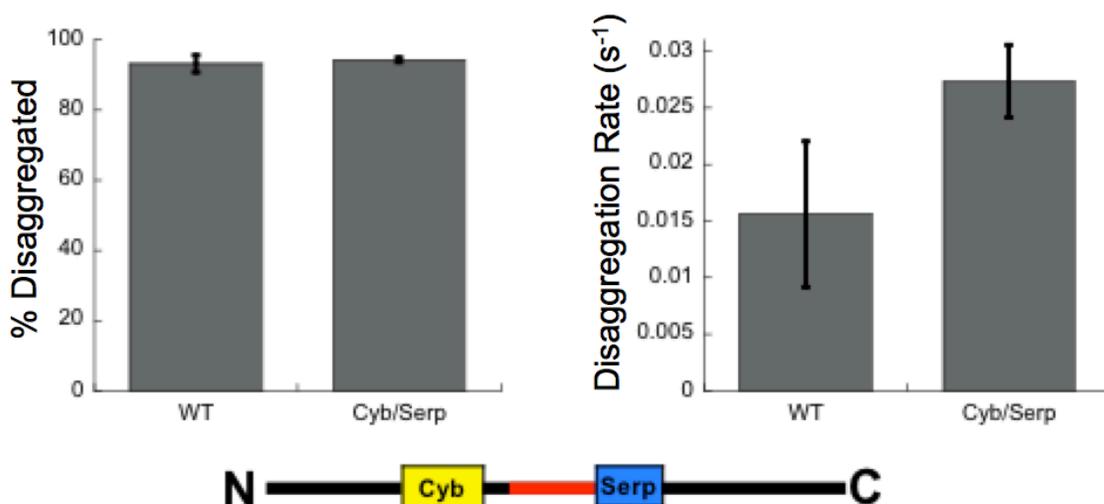
Figure 3. Disaggregation data for LHCP deltaTM1 in which TM2 has been swapped with Cyb5 and TM3 has been swapped with Serp1.[2,21] This mutant substrate shows comparable disaggregation properties to wild-type LHCP with a modestly faster disaggregation rate. cpSRP43 in these experiments is at a concentration of 5 uM.

To further test the capabilities of cpSRP43 to chaperone non-native substrates we attempted to express SERP1 fused to L18 in order to study this substrate and not just the SERP1 TM in the LHCP scaffold, as was studied for the TM swap mutants. However, the SERP1 protein could not be expressed in sufficient quantities to study.[6] To overcome this problem, we sought to utilize cpSRP43's dual interactions with substrates to improve membrane protein expression. We hypothesized that by fusing the L18 motif to membrane proteins, cpSRP43 could chaperone them in E. coli cells through specific L18 recognition and promiscuous hydrophobic interactions. This chaperone activity would be expected to be generalizable to various membrane proteins, given a positioning of the L18 motif in such a way that it is exposed on the aggregate surface for cpSRP43 to recognize.

With this knowledge in mind, we aimed to use cpSRP43 as a means of increasing membrane protein expression. We began these experiments first by optimizing cpSRP43 expression in E. coli. We then went on to test cpSRP43's activity by looking at co-expression with various membrane protein substrates. In doing so, we looked at factors such as temperature dependence of expression and L18 dependence of substrates. Because of cpSRP43's important interaction with L18 in LHCP, one would expect that this interaction is necessary for increased expression of membrane protein substrates. Finally, we looked at fractionation to determine the location of the expressed membrane protein substrates. It is known that membrane proteins that end up in inclusion bodies are rarely functional proteins.[1] Therefore, in our strategy to increase membrane protein co-expression we aim to improve localization to the membrane in order to get functional membrane proteins. Using our various expression, L18 dependence, and fractionation data we are able to assess the ability of cpSRP43 to increase membrane protein co-expression.

Results

Optimization of cpSRP43 expression for co-expression experiments. In trying to increase membrane protein expression, it was first necessary to optimize the expression of the cpSRP43 chaperone. For our initial studies, Serp1 was our model substrate. The results of this optimization are summarized in Figure 4. Our initial construct had both a His6 tag on cpSRP43 and on Serp1. However, having a His6 tag on both cpSRP43 and on the substrate made analysis ambiguous because it turned out that cpSRP43 had degradation products of similar molecular weight to the substrates of interest. Another major issue was that baseline cpSRP43 expression was low. We looked at different vectors, namely pACYCDuet, pQE and pET and ultimately used pQE for co-expression experiments because pQE was what was used in the past to express and purify cpSRP43 alone.[1] Despite this change, cpSRP43 expression was still relatively low.

We next turned to cell type as a means of optimizing expression. Comparing BL21-DE3*, Rosetta, and BL21-CodonPlus-DE3-RIL cells, we found dramatically increased expression with BL21-CodonPlus cells. To overcome the ambiguity of having both cpSRP43 and substrate being His6-tagged, we let the substrate remain His6-tagged (for easy detection by Western blot and ultimately for purification) and then FLAG-tagged cpSRP43. However, the switch from a His6-tag to a FLAG-tag on cpSRP43 dramatically diminished expression. Comparing cpSRP43 with no tag, FLAG, Strep, and His6 tags, we found that N-terminally His6-tagged cpSRP43 had the best expression. Because we found that the His6 tag was needed for high cpSRP43 yield, we ultimately switched the tag on the substrate to a FLAG tag (for detection via Western blot). Thus the conditions we ultimately settled upon after optimization were BL21-CodonPlus-DE3-RIL cells, pQE-80L vector, His6-cpSRP43, and FLAG-substrate, as summarized in Figure 4. The final construction of the co-expression plasmid is shown in Figure 5.

Cell Type	Vector	Tag on cpSRP43	Tag on substrate
BL21-DE3*	pACYCDuet	His6	His6
Rosetta	pQE-80L	FLAG	FLAG
BL21-CodonPlus-DE3-RIL	pET	Strep	

Figure 4. Optimization of cell type, vector, and tags for cpSRP43 expression. Within each portion of the chart is a gel showing pre-induction (-IPTG, left lane) and post-induction (+IPTG, right lane) conditions for cpSRP43 expression using various cell types, vectors, and tags on cpSRP43. The best conditions for cpSRP43 expression are circled in green: BL21-CodonPlus-DE3-RIL cells, a pQE-80L plasmid, and a His6 tag on cpSRP43.
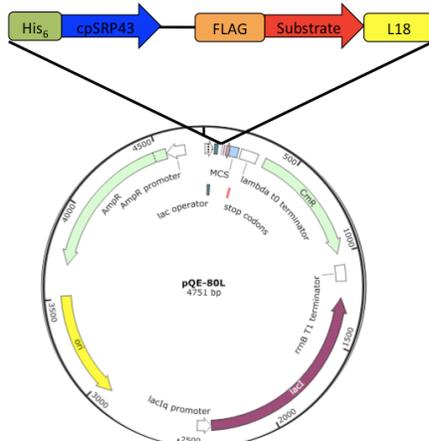
Figure 5. Schematic of pQE-80L plasmid containing the co-expression construction. cpSRP43 is N-terminally tagged with His6 and the substrate is N-terminally tagged with FLAG and C-terminally tagged with L18.

Substrate scope of co-expression of cpSRP43 with membrane proteins. The substrates tested were a combination of tail-anchored proteins and two transmembrane domain proteins. These substrates are summarized schematically in Figure 6. The initial substrates used for optimization of cpSRP43 expression conditions were Sec61B, a subunit of the protein transolocation machinery in the ER, Serp1, a stress-associated ER protein, and Lep1, an E. coli protein known as leader peptidase that is involved in removing amino-terminal peptides from exported proteins (Figure 6a). From these initial substrates, we expanded our scope to various tail-anchored membrane proteins (Figure 6b) including cytochrome b5 (Cyb5), an electron donor involved in lipid biosynthesis, SNARE proteins Sec22, Use1, and Bos1, and finally Scs2, an integral ER protein involved in phospholipid metabolism. Our last group of preliminary test substrates consists of two TM swap mutants of the large mechanosensitive channel (MscL) derived from S. aureus and E. coli (Figure 6c). These substrates, numbered 27 and 51, were C-terminally tagged with an L18 motif similarly to the tail-anchored substrates.

Parent Constructs
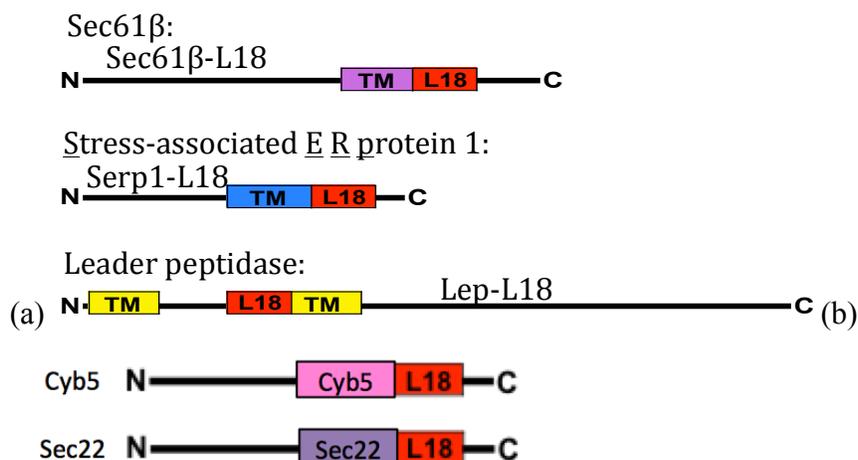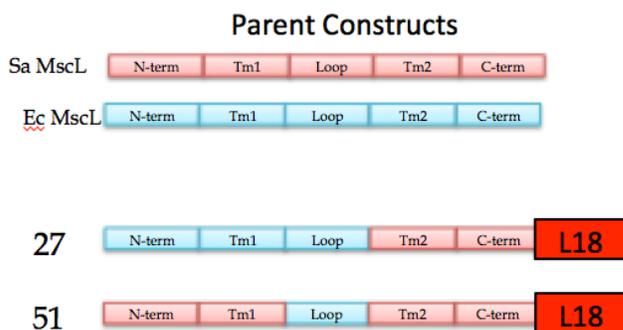
Figure 6. Preliminary membrane protein substrates. (a) First set of substrates developed while optimizing cpSRP43 expression conditions. These initial substrates include two tail-anchored proteins, Sec61B and Serp1, and a 2 TM bacterial substrate, Lep1. (b) Schematic of various tail-anchored proteins which include Cyb5 and Sec22 (pictured) as well as Use1, Scs2, and Bos1 (not pictured but following an analogous scheme). (c) Large mechanosensitive channel (MscL) constructs made by swapping components from S. aureus MscL and E. coli. The parent constructs are shown for S. aureus (peach) and E. coli (blue). The swapped mutants are labeled 27 and 51 and were tagged with the L18 motif.

Preliminary screening of these membrane protein substrates has shown promise for improved membrane protein expression with cpSRP43 co-expression. Promising substrates include Serp1, Sec61B, Sec22, Scs2, and Use1. The expression of each of these substrates with and without cpSRP43 co-expression is shown in Figures 7 and 8. All substrates were grown at 37˚C, 250 rpm and induced for four hours with 1.0mM IPTG at an $OD_{600} \sim 0.4-0.6$. In each figure, a Coomassie-stained gel confirms co-expression of cpSRP43 and an anti-FLAG western blot shows expression of the membrane protein of interest. Each western blot is then quantified to show the relative levels of membrane protein expression with and without cpSRP43. The quantifications are only a relative comparison and are not standardized to protein concentration. The relative protein expression levels are quantified such that 100% represents the highest protein expression level on the gel and other percentages represent protein expression relative to the 100% case. In this way protein expression can be compared directly between substrates with and without cpSRP43.

Unsuccessful substrates include Cyb5 and MscL27 which are not expressed within the level of western blot sensitivity either in the with or without cpSRP43 case. Intriguingly, in the case of MscL27, cpSRP43 expression is not seen on the Coomassie gel (Figure 7a). Unlike the other tail-anchored substrates, MscL27 forms a multimeric complex and it is possible that in trying to chaperone MscL27, cpSRP43 is degraded along with MscL27 if this substrate unable to form a complex as it is expressed. Lep1 likewise does not co-express with cpSRP43 and shows the same phenomenon as MscL27 in that cpSRP43 expression is not observed (data not shown). Future experiments such as pulse-chase analysis to determine both cpSRP43 and substrate degradation need to be performed in order to assess whether cpSRP43 is being degraded in its attempt to chaperone proteins like MscL27 and Lep1.
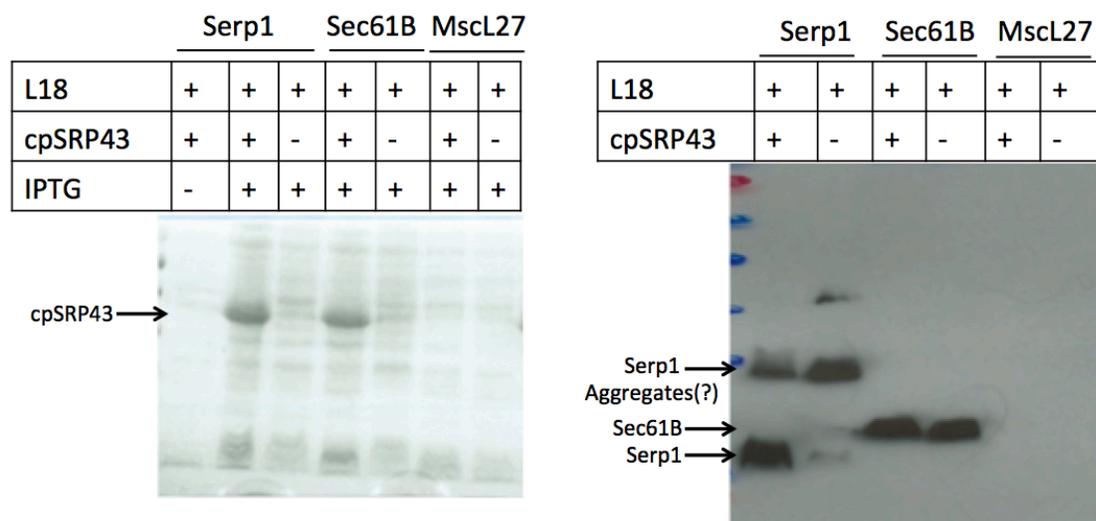
Figure 7. Co-expression of cpSRP43 with Serp1, Sec61B, and MscL27. The Coomassie-stained gel (a, left) shows cpSRP43 expression. The anti-FLAG western blot (b, right) shows the expression of Serp1, Sec61B, and MsCl27 with and without cpSRP43.
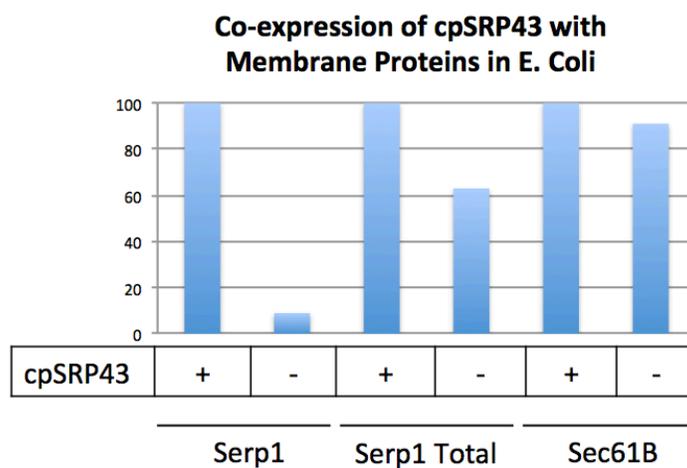
Figure 8. Quantification of co-expression anti-FLAG Western blot. Lanes 1 and 2 show quantification of only the Serp1 band found at the expected molecular weight of approximately 12 kDa. There is significantly more expression with cpSRP43 than with Serp1 alone. Lanes 3 and 4 show quantification of all bands combined in the Serp1 western blot, potentially representing dimers or aggregates of this protein. Including all bands, co-expression with cpSRP43 is still significantly improved relative to Serp1 alone. For Sec61B, shown in Lanes 5 and 6, expression is only modestly improved with cpSRP43 co-expression.

Figure 9. Coexpression of cpSRP43 with Cyb5, Sec22, MscL27, Scs2, Use1, and Bos1. cpSRP43 overexpression is seen for Cyb5, Sec22, Scs2, Use1, and Bos1 (prominent cpSRP43 band at ~38kDa) in all cases except for MscL27. MscL27 has the same vector construction and cpSRP43 gene as all of the other co-expression plasmids and underwent the same induction conditions.
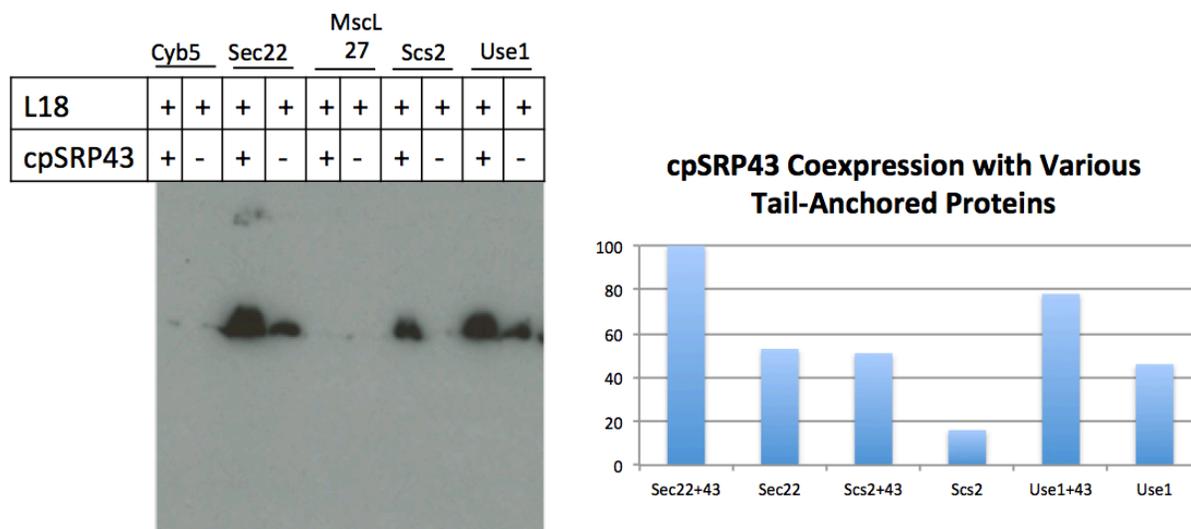
Figure 10. Anti-FLAG western blot and quantification of expression of various membrane proteins. Only substrates showing detectable band were quantified. Sec22, Scs2, and Use1 showed significantly improved expression with cpSRP43 co-expression. All bands were quantified relative the most prominent band, the Sec22 with cpSRP43 band, which has been set to 100% as a reference point.

Overall, these preliminary screens are promising for the use of cpSRP43 as a means of improving membrane protein co-expression. Additionally, as described in greater detail previously, structure-function cpSRP43 mutant studies have led to a superactive cpSRP43 mutant with improved binding and disaggregation activity that can potentially be used to improve the expression of substrates that did not show initial promise with wild-type cpSRP43 co-expression.

Dependence of Membrane Protein Expression on Growth Temperature. Initial studies of temperature dependence for expression of the model membrane protein Serp1 showed little

difference between growth at 25˚C versus 37˚C as shown in Figure 11. However, it must be noted that these studies were conducted using an early vector construction of His6-cpSRP43 and His6-substrate in BL21DE3* cells, which was later optimized as described in the cpSRP43 expression optimization section. Therefore, although these experiments show little variation with growth at 25˚C versus 37˚C, the conditions must be repeated with the new optimization conditions (His6-cpSRP43 and FLAG-substrate, BL21-CodonPlus-DE3-RIL cells) and test substrates. It is interesting to note that under these pre-optimized conditions, there was very little expression of Serp1 in the absence of cpSRP43, whereas in the optimized conditions Serp1 alone expression is improved. This change in baseline Serp1 alone expression is discussed in more detail in the Supplementary Information section.
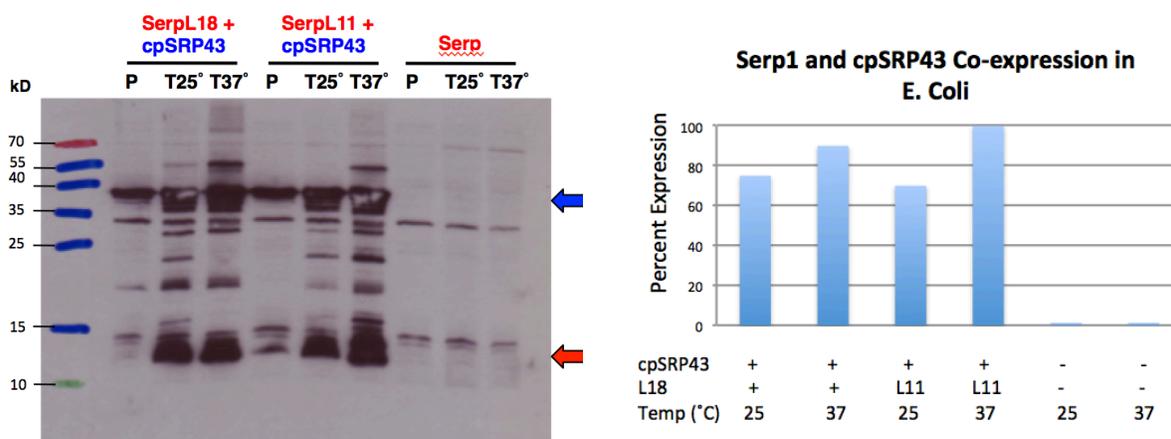

Figure 11. Expression of Serp with L18-tag, L11-tag, and no tag. (a, left) Anti-His6 western blot showing cpSRP43 (blue arrow, ~38 kDa) and Serp1 expression (red arrow, ~12 kDa). In this construction, it is obvious that cpSRP43 co-expression vastly improves Serp1 expression from the Serp1 alone case. Expression is improved with both the L18 and the truncated L11 tag. However, overall Serp1 expression with L18 or L11 is similar between the 25˚C and 37˚C cases, as shown in the quantification. **(b, right)** Quantification of Serp1 expression. All values are normalized to the highest expression band, Serp1-L11+cpSRP43. Expression without cpSRP43 is negligible and there is little quantitative variation between expression at 25˚C versus 37˚C.

Dependence of Membrane Protein Expression on L18-Tag. Because of the importance of the FDPLGL on the L18 region of LHCP for cpSRP43 aggregation prevention and disaggregation, we sought to study cpSRP43's mechanism of action for non-native substrates by comparing substrates with the L18 motif, the L11 motif (a truncated version of L18 that includes the critical FDPLGL residues), and no L18 tag. We initially conducted this study in the vector in which cpSRP43 was FLAG-tagged and Serp1 was His-tagged and BL21DE3* cells were used for expression. This preliminary experiment shows a lack of L18 dependence. Under these conditions, Serp1 alone (Figure 11a, Lane 7) does not show any expression in the absence of cpSRP43 (see Supplemental Figure 3 for comparison of Serp1 alone expression for different cell types and tags). However, Serp1 that has been tagged with L18, L11, and no tag and co-expressed with cpSRP43 do not show significant differences in expression. Surprisingly, the construct with no tag appears to have somewhat higher expression than the L11 or L18 tagged version. Although this experiment must be repeated in the optimized conditions, this provides interesting clues as to how cpSRP43 is acting to improve membrane protein co-expression. For instance, it is possible that cpSRP43 is able to utilize hydrophobic interactions with Serp1 to prevent aggregation before it occurs, analogously to the type of data observed for cpSRP43 and AB_{40}. Another substrate, Sec22, was tested for L18

dependence under the optimized conditions. Although increased expression was shown with co-expression of cpSRP43, differences in expression between L18, L11, and untagged substrates appear subtle. These L18 dependence experiments need to be repeated across the current pool of substrates to determine whether L18 dependence or lack of L18 dependence is a generalizable phenomenon. Such data will provide further insight into how cpSRP43 is able to improve expression of membrane protein substrates.

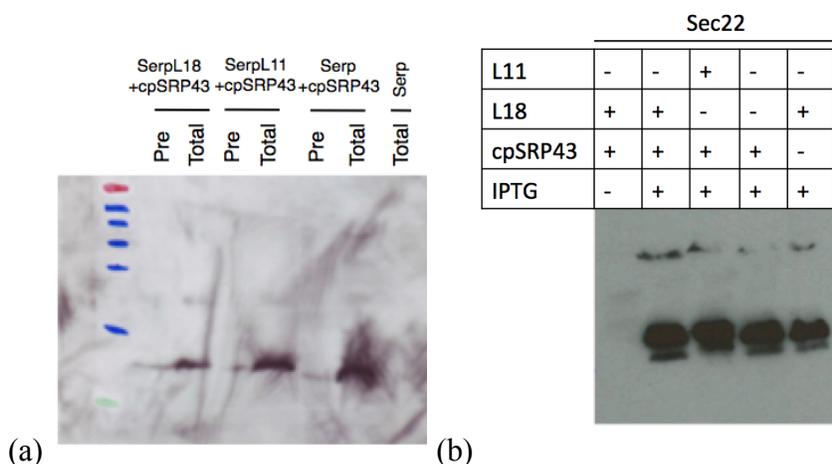|	Sec22				
L11	-	-	+	-	-
L18	+	+	-	-	+
cpSRP43	+	+	+	+	-
IPTG	-	+	+	+	+

(a)   (b)

Figure 12. L18 dependence study using Serp1 (a) and Sec22 (b) as model substrates. The Serp1 study was performed using His-tagged substrate, FLAG-tagged cpSRP43, and BL21DE3* cells whereas the Sec22 study was performed with the optimized cpSRP43 conditions. Preliminarily these substrates do not show a large increase in expression with the L18 motif, indicating that perhaps these substrates do not form aggregates with as strong packing interactions as LHCP or cpSRP43 is able to make promiscuous hydrophobic contacts with these substrates before they are able to aggregate.

Fractionation to Determine Cellular Location of Membrane Proteins During Co-Expression. Although the co-expression data has shown some promise, it is important to determine whether the membrane proteins that are being expressed are forming inclusion bodies, being transported to the membrane, or interacting with cpSRP43 or other chaperones and remaining soluble. This question is important because it is known that membrane proteins that go into inclusion bodies can often not be in a functional form.[18] To assess the distribution of membrane proteins between inclusion bodies, membrane, and soluble fractions, we used fractionation. Taking broken cells, we first pelleted them at 4000 g to remove any remaining unbroken cells. We then took the total fraction (that is, all of the broken cells) and centrifuged at 16,000 g to remove inclusion bodies, which are heavy relative to other cellular components. Finally, we ultracentrifuged for 2 hours at 100,000 g to pellet the membranes. The remaining supernatant after this ultracentrifugation was labeled as the soluble fraction. Thus far, we have performed fractionation on Serp1 and Sec61B. We also performed fractionation on LHCP, which is known to be found only in inclusion bodies, as a control for the procedure. Fractionation was performed on Serp1 in both low expression cpSRP43 conditions (where cpSRP43 was FLAG-tagged and Serp1 was His-tagged in BL21DE3* cells) and high expression cpSRP43 conditions from the final optimized conditions.

A. Serp1 Fractionation

Figure 13 shows the distribution of His6-tagged Serp1 co-expressed with FLAG-tagged cpSRP43 in BL21DE3* cells. Under these pre-optimized conditions where cpSRP43 has low expression, Serp1 is distributed between the inclusion body, membrane, and soluble fractions with the highest fraction in the membrane and lower fractions in the inclusion body and soluble lanes. Low cpSRP43 expression is shown in the accompanying Coomassie-stained gel where cpSRP43 appears to be predominantly in the soluble fraction. When Serp1 tagged with L11 was fractionated under these same pre-optimized, low cpSRP43 expressing conditions, it was found to be distributed between the inclusion body and the membrane with very little detectable in the soluble fraction (Figure 14). LHCP was used as a control to confirm the fractionation method. LHCP is known to be found in inclusion bodies, and these fractionation results, which show LHCP only in the inclusion body fraction and not in the membrane or soluble fractions, confirm that the inclusion bodies are being successfully separated from membrane and soluble fractions.
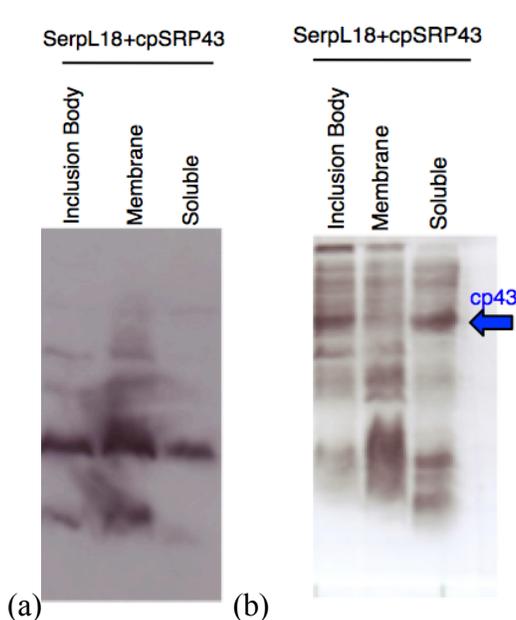
Figure 13. Fractionation of Serp1 tagged with L18 and co-expressed with cpSRP43 under pre-optimized conditions (low cpSRP43 expression). (a) anti-His western blot for Serp1 fractionation. Serp1 is distributed between the fractions with the highest amount in the membrane fraction. (b) Coomassie-stained gel showing cpSRP43 fractionation. cpSRP43 has very little in the membrane, slightly more in the inclusion body, and the highest fraction in the soluble portion.
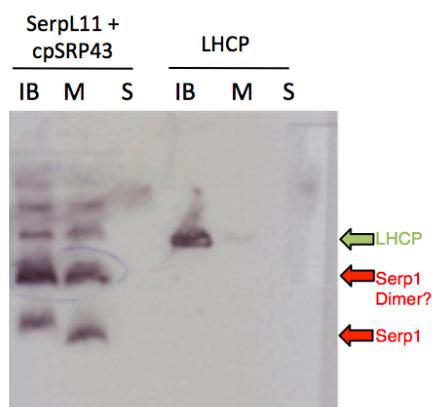
Figure 14. Anti-His western blot of the fractionation of Serp1 (with pre-optimized, low expressing cpSRP43 conditions) and LHCP. Serp1 is tagged with L11 and is co-expressed with cpSRP43. LHCP, which is known to be found in inclusion bodies, is expressed as a control. IB=Inclusion body, M=Membrane, S=Soluble.

These fractionation experiments were repeated for Serp1 and Sec61B in the optimized conditions (His-tagged cpSRP43, FLAG-tagged substrate, BL21-CodonPlus-DE3-RIL cells). Serp1 fractionation with and without cpSRP43 co-expression is shown in Figure 15. In the cpSRP43-Serp1 co-expression case, Serp1 is mainly found at the expected molecular weight of 12 kDa with the largest portion in the membrane fraction, some in the inclusion body fraction and little in the soluble fraction. In the Serp1 alone case most of Serp1 is found in the inclusion body fraction with little in the membrane or soluble fractions (note that there is some non-specific exposure on the far right side of the Serp1 alone western blot). Moreover, the Serp1 alone case shows many more higher molecular weight species (potentially oligomers) that could correspond to non-functional Serp1. cpSRP43 may therefore be able to keep Serp1 and other membrane proteins in a functional form by helping them to get to the membrane rather than allowing them to aggregate in inclusion bodies. This hypothesis must be further tested across a variety of substrates to more fully understand whether cpSRP43 is able to increase expression of functional membrane proteins by assisting them in getting to the membrane.
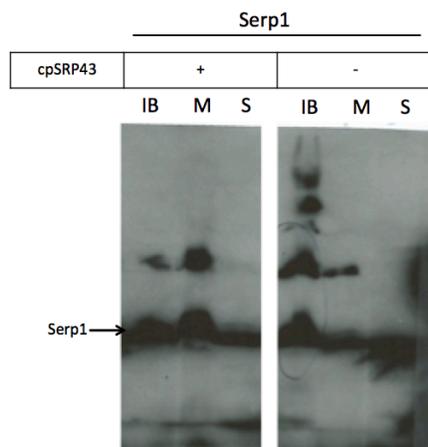
Figure 15. Anti-FLAG western blot of Serp1 fractionation. In the cpSRP43 co-expression case (left) the greatest fraction is found in the membrane at the expected 12 kDa molecular weight. By contrast, the Serp1 without cpSRP43 case (right) shows the largest fraction in the inclusion body and moreover this case shows many higher molecular weight bands, potentially indicating oligomerization of Serp1 in the inclusion body.

Sec61B Fractionation. Fractionation was also performed with Sec61B co-expressed with cpSRP43 under the optimized cpSRP43 expression conditions. In this fractionation, cpSRP43 was found to be distributed in the inclusion body, membrane, and soluble fractions, with the highest fraction in the soluble portion (Figure 16). An anit-FLAG western blot indicated that much of Sec61B was found in the membrane fraction and soluble fraction and little was found in the inclusion body (Figure 17). Although the high proportion in the membrane fraction is promising (approximately 40% by quantification) the similar amount in the soluble fraction is intriguing. It is possible that cpSRP43 is interacting with Sec61B, keeping it soluble. However, it is unclear what effect this will have on the functionality of the protein. This experiment must be repeated in both with and without cpSRP43 conditions in order to assess whether cpSRP43 co-expression alters the fractionation pattern.
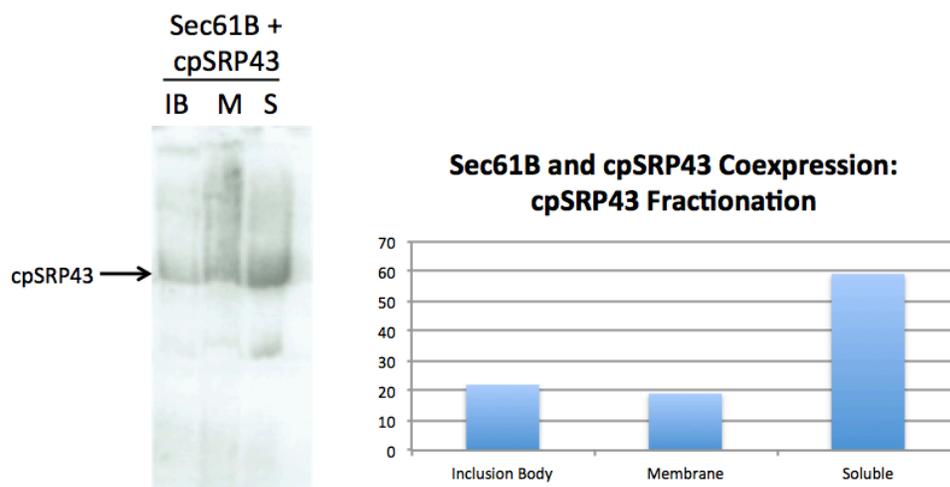
Figure 16. Coomassie-stained gel of fractionation of cpSRP43 co-expressed with Sec61B (left) and quantification of fractions (right). The greatest fraction is found in the soluble portion with smaller amounts found in the inclusion body and membrane fraction.
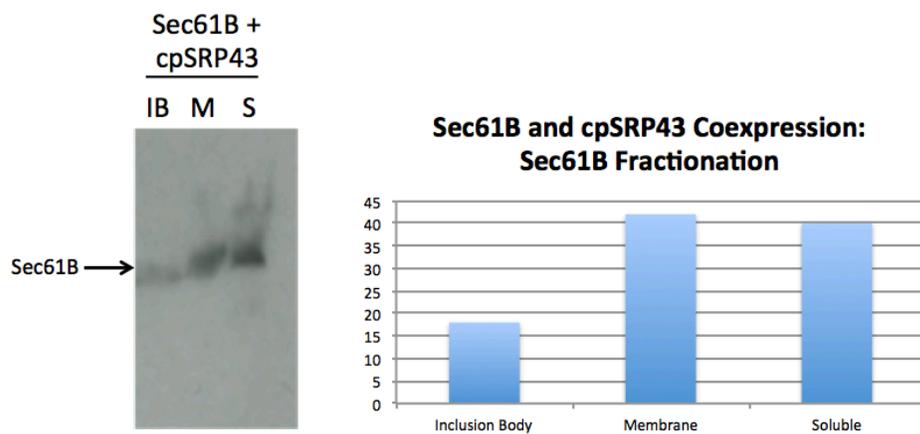
Figure 17. Anti-FLAG western blot of Sec61B co-expressed with cpSRP43 (left) and quantification of the fractions (right). Sec61B is found mostly in the membrane and soluble fractions with little in the inclusion body fraction.

Discussion

To summarize, we have tested the usefulness of cpSRP43 co-expression for improving membrane protein expression across a preliminary range of substrates including Serp1, Sec61B, MscL27, Cyb5, Sec22, Scs2, Bos1, and Use1. The most promising substrates thus far are Serp1, Sec61B, Sec22, Scs2, and Use1. Further we have assessed the necessity of the L18 motif for improving membrane protein co-expression and these initial results have indicated that this motif may not be necessary for some substrates, which could be promising as these substrates could potentially be purified without, for example, needing to have the L18 tag cleaved. To test for the localization of the expressed membrane proteins in the cell, we used fractionation and found that most cpSRP43 is found in the soluble fraction, as expected, and that a high proportion of Serp1 and Sec61B is found in the membrane fraction when co-expressed with cpSRP43. Our future work will include expanding our membrane protein substrate scope to further probe the usefulness of cpSRP43 co-expression. We will also determine the L18 dependence of our current and future substrates by making L18, L11, and deltaL18 versions of all substrates. Moreover, we will test the importance of the location of the L18 tag within the substrate. In the native LHCP substrate, the L18 tag is placed between TM2 and TM3. For multiple TM substrates such as MscL27 we will try placing the L18 tag in between its two TM's to assess whether this can increase expression. In tail-anchored substrates we have put the L18 tag on the C-terminus but we can likewise try putting the L18 tag on the N-terminus to determine if this is able to increase expression. By varying the location of the L18 tag and potentially obtaining substrates with more TMs, we can more fully assess the importance of the L18 tag in cpSRP43 chaperone activity of non-native substrates.

These cpSRP43 co-expression experiments have shown promise for cpSRP43 as a potential tool for increasing membrane protein expression. Because membrane proteins are found commonly in biological systems but are often not characterized due to their poor expression, developing tools for improving membrane protein expression is an important objective. Further, studying cpSRP43 co-expression can provide another approach for understanding cpSRP43 activity and mechanism for non-native substrates through L18 dependence studies and further through assessing whether cpSRP43 is able to deposit substrates on the membrane through fractionation. Although these experimental results provide a promising start for cpSRP43 as a tool for understanding and overcoming problems of protein aggregation, more work must be done to understand how cpSRP43 functions and what substrate scope this chaperone is applicable to for dealing with protein aggregation.

References

1. S. Schlegel, M. Klepsch, D. Gialama, D. Wickströn, D. Slotboom, and J. de Gier. Revolutionizing membrane protein overexpression in bacteria. *Microbial Biotechnology*. **2010**, *3*: 403-411.

2. M. Klepsch, J. Persson, and J. de Gier. Consequences of the Overexpression of a Eukaryotic Membrane Protein, the Human KDEL Receptor in Escherichia coli. *Journal of Molecular Biology*. **2011**: *407*, 532-542.

3. S. Wagner, L. Baars, A.J. Ytterberg, A. Klusmeier, C.S.Wagner, O. Nord, P.Å. Nygren, K.J. van Wijk and J.W. de Gier. Consequences of membrane protein overexpression in Escherichia coli. *Molecular and Cellular Proteomics*. **2007**, *9*: 1527-50

4. P. Jaru-Ampornpan, F. Liang, A. Nisthal, T. Nguyen, P. Wang, K. Shen, S. Mayo, and S. Shan. Mechanism of an ATP-independent Protein Disaggregase: II. Distinct Molecular Interactions Drive Multiple Steps During Aggregate Disassembly. *J. Biol. Chem.* **2013**, *288*:13431-13445.

5. Nguyen, T. A study of protein targeting reveals insights into mitigating protein aggregation. Dissertation (Ph.D.), California Institute of Technology, **2013**.

6. T. Nguyen, Shan Lab, unpublished data, personal communication **2013**.

7. P. Jaru-Ampornpan, K. Shen, V. Lam, M. Ali, S. Doniach, T. Jia, and S. Shan. ATP-independent reversal of a membrane protein aggregate by a chloroplast SRP subunit. *Nature Structural and Molecular Biology*. **2010**, *17*: 696-703.

8. Chiti, F; Dobson, C. Protein misfolding, functional amyloid and human disease. *Ann Rev of Biochem.* **2006**, *75*: 333-366.

9. Ahmed, M; Davis, J; Aucoin, D; Sato, T; Ahuja, S; Aimoto, S; Elliot, J; Nostrand, W; Smith, S. Structural conversion of neurotoxic amyloid-beta$_{1-42}$ oligomers to fibrils. *Nature Structural and Molecular Biology*. **2010**, *17*, 561-567.

10. Kodali, R; Wetzel, R. Polymorphism in the intermediates and products of amyloid assembly. *Current Opinion in Structural Biology*. **2007**, *17*, 48-57.

11. Luheshi, L; Dobson, C. Bridging the gap: From protein misfolding to protein misfolding diseases. *FEBS Letters*. **2009**, 583, 2581-2586.

12. Broadley, S; Hartl, F. The role of molecular chaperones in human misfolding diseases. *FEBS Letters*. **2009**, *583*, 2647-2653.

13. Wendler, P; Shorter, J; Snead, D; Plisson, C; Lindquist, S; Saibil, H. Motor Mechanism for Protein Threading through Hsp104. *Molecular Cell*. **2009**, *34*, 81-92.

14. Lindquist, S; Shorter, J. Hsp104 Catalyzes Formation and Elimination of Self-Replicating Sup35 Prion Conformers. *Science*. **2004**, *304*, 1793.

15. Duennwald, M. Echeverria, A; Shorter, J. Small Heat Shock Proteins Potentiate Amyloid Dissolution by Protein Disaggregases from Yeast and Humans. *PLoS Biology*. **2012**, *10*, 1-23.

16. T. Nguyen, P. Jaru-Ampornpan, V. Lam, P. Cao, S. Piszkiewicz, S. Hess, S. Shan. Mechanism of an ATP-independent protein disaggregase: I. Structure of a

membrane protein aggregate reveals a mechanism of recognition by its chaperone. *J. Biol. Chem.* **2013**, *288*: 13420-13430.

17. Schuenemann, D., Gupta, S., Persello-Cartieaux, F., Klimyuk, V. I., Jones, J. D., Nussaume, L., and Hoffman, N. E. A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. *Proc. Natl. Acad. Sci. U.S.A*. **1998**, *95*: 10312–10316.

18. Jason S. A guide to the Lhc genes and their relatives in Arabidopsis. *Trends in Plant Sci*. **1999**, *4*: 236-240.

19. Liu, Z., Yan, H., Wang, K., Kuang, T., Zhang, J., Gui, L., An, X., and Chang, W. Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature*. **2004**, *428*: 287–292.

20. Stengel, K. Holdermann, I. Cain, P. Robinson, C. Wild, K. Sinning, I. Structural basis for specific substrate recognition by chloroplast signal recognition particle protein cpSRP43. *Science*. **2008**, *321*: 253-256.

21. Tu, C. Peterson, E. Henry, R. Hoffman, N. The L18 domain of light-harvesting chlorophyll proteins binds to chloroplast signal recognition particle 43. *J. Biol. Chem*. **2000**, *275*: 13187-13190.

22. F. Liang, Shan Lab, unpublished data, personal communication **2014**.

23. Maji, S; Wang, L; Greenwald, J; Riek, R. Structure-activity relationship of amyloid fibrils. *FEBS Letters*. **2009**, *583*, 2610-2617.

24. Steiner, D; Forrer, P; Plückthun, A. Efficient Selection of DARPins with Sub-nanomolar Affinities using SRP Phage Display. *J. Mol. Bio*. **2008**, *382*, 1211-1227.

25. Zahnd, C; Wyler, E; Schwenk, J; Steiner, D; Lawrence, M; McKern, N; Pecorari, F; Ward, C; Joos, T; Plückthun, A. A Designed Ankyrin Repeat Protein Evolved to Picomolar Affinity to Her2. *J. Mol. Bio*. **2007**, *369*, 1015-1028.

26. Steiner, D; Forrer, P; Stumpp, M; Plückthun. Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display. *Nature Biotechnology*. **2006**, *24*, 823-831.

Supplementary Information

A result of switching from FLAG-cpSRP43 and His6-substrate in BL21DE3* cells to His6-cpSRP43 and FLAG-substrate in BL21-CodonPlus-DE3-RIL was seeing baseline expression of Serp1 without cpSRP43 increase. This change in Serp1 alone expression with tag and cell type is shown in Supplementary Figure 1.
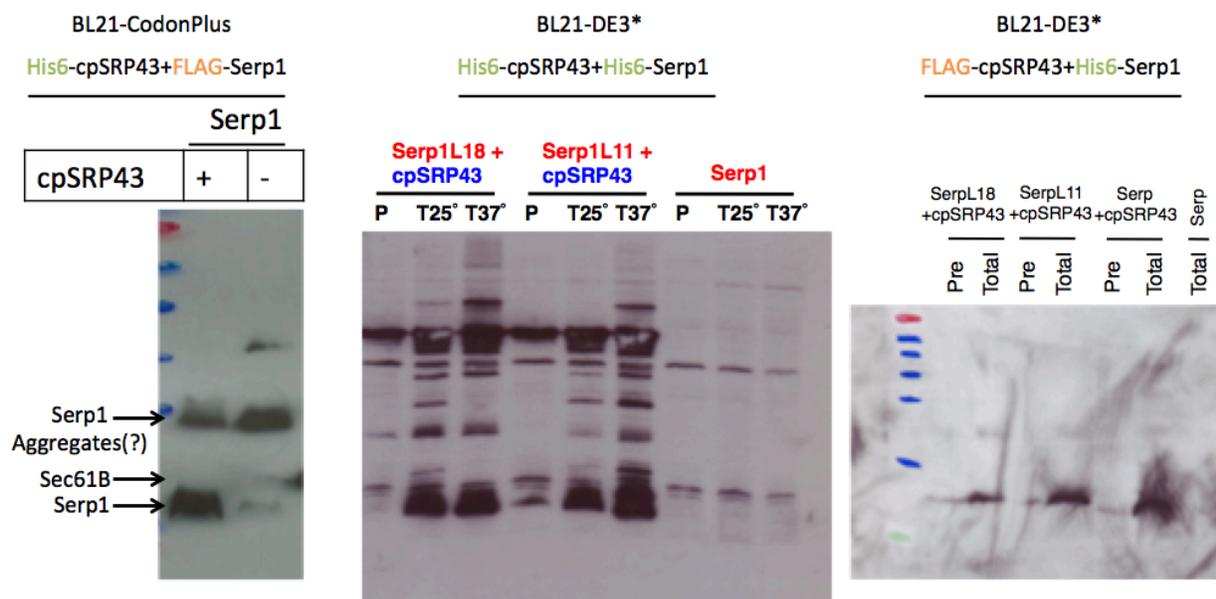
Supplementary Figure 1. Comparison of Serp1 alone expression between different tagging conditions and competent cell type.

C h a p t e r   5

CHARACTERIZING BIOCATALYZED CARBON-CARBON BOND FORMATION MECHANISMS FOR INDUSTRIAL SYNTHESIS APPLICATIONS

Camille Z. McAvoy and Douglas C. Rees

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125

Abstract

Carbon-carbon (C-C) bonds are fundamental components of a wide variety of industrial molecules including pharmaceuticals, fuel, and agricultural products; methods for forming C-C bonds are thus critical to industrial production. However, industrial syntheses often rely upon large quantities of toxic reagents and extreme conditions such as high temperature and pressure. Biologically-derived catalysts, which often function in aqueous and ambient conditions, have emerged as a powerful tool to help make industry more green and open up novel synthetic pathways. Here we use nitrogenase and a representative substrate, methyl isonitrile, to study a mechanistically uncharacterized C-C bond formation reaction. Nitrogenase is a two component metalloenzyme that catalyzes biological nitrogen fixation, the reduction of atmospheric nitrogen (N_2) to bioavailable ammonia (NH_3). Moreover, nitrogenase is capable of catalyzing other useful chemistries such as reducing methyl isonitrile to form methane, methylamine and intriguingly also ethane, ethylene, and higher hydrocarbons. We utilized an interdisciplinary kinetics-based approach to characterize the C-C bond formation mechanism between methyl isonitrile and nitrogenase. This work will serve not only as an important characterization of an applicable synthetic tool, but also as a foundation upon which future enzyme engineering work can be done to optimize this catalysis.

Introduction

Nitrogenase is the only known family of enzymes capable of biological nitrogen fixation, or the ATP-powered conversion of atmospheric nitrogen (N_2) to bioavailable ammonia (NH_3) (Figure 1). Nitrogen fixation is critical for making soil fertile, and as such nitrogenases are commonly found in microorganisms such as *Azotobacter vinelandii* and *Clostridium pasteurianum*. Significantly, this biocatalyzed process serves as an alternative to the Haber-Bosch process, a widely used nitrogen fixation method which requires conditions of high temperature and pressure.[1] Nitrogenase has long served as a system of interest to help green the industrial production of ammonia; here we seek to build upon this work by studying another important chemistry catalyzed by nitrogenase, carbon-carbon (C-C) bond formation.

Background and Significance

Although the complete mechanism of nitrogenase remains elusive, a variety of biophysical characterizations have provided key insights into the components of this complex biological system.[2-7] Nitrogenase is composed of two metalloprotein components: the iron (Fe) protein (called Av2 when purified from soil bacteria *Azotobacter vinelandii*) and the molybdenum-iron (MoFe) protein (Av1). Dinitrogen reduction is an ATP-dependent process initiated by electron transfer from the [4Fe:4S]-cluster of the Fe-protein to the MoFe-protein, which contains P-cluster [8Fe:7S] and FeMo-cofactor [7Fe:9S:C:Mo] metal centers.[2] These proteins and electron transfer reactions are described in Figure 2,[8,9] namely ATP-consumption drives electron transfer from the [4Fe:4S]-cluster to the P-cluster to the FeMo-cofactor.

Recent studies from the Rees Group have generated CO- and Se- bound forms of the cofactor generated under turnover conditions, providing further insight into the potential binding mechanism of substrates and inhibitors to the FeMo-cofactor.[4,7] The FeMo-cofactor is composed of a trigonal prism containing an interstitial carbide that is thought to stabilize the cofactor during rearrangements associated with substrate binding. Substrates bind only to the more highly reduced forms of the enzyme that are efficiently generated only with the Fe-protein and ATP consumption. These studies provide an excellent starting point to assess whether the structure containing methyl isonitrile bound to nitrogenase undergoing C-C bound formation, matches with the prediction from the two CO bound structure that this "high-CO" form may represent an intermediate relevant to the carbon-carbon (C-C) coupling reaction.[10]

Nitrogenase is a powerful enzyme capable of performing various reduction chemistries including the conversion of methyl isonitrile to methane and methylamine. Interestingly, additional products of nitrogenase-methyl isonitrile interactions include ethane, ethylene, and higher hydrocarbons, indicating that nitrogenase is capable of not only reduction but also C-C bond formation.[11-15] Understanding this mechanism could have significant applications for biocatalyzed syntheses of organic molecules involved in fuel, pharmaceuticals, and other chemical industries.

Specific Aims

We can organize this research endeavor into a few key biochemical questions: 1.What does the intermediate in ethane formation from methyl isonitrile look like while interacting with nitrogenase? 2.What are the kinetics of two carbon product formation from this interaction? 3.Will crystallizing methyl isonitrile in complex with nitrogenase provide insight into this mechanism?

To address these key research questions we have the following specific aims: Aim 1.) Form a complex between methyl isonitrile and nitrogenase under turnover conditions. Aim 2.) Perform gas chromatography under various conditions to study the mechanism and kinetics of this carbon-carbon bond formation. Aim 3.) Crystallize complex through optimization of crystallization conditions. Aim 4.) Determine structure and C-C bond formation mechanism through X-ray crystallography.

Experimental Approach

Our experimental approach involves an interdisciplinary series of experiments to extract useful organic tools from a biological system. The bacterial species *Azotobacter vinelandii* will be grown in a fermenter, and Av1 and Av2 will be purified from this cell lysate anaerobically using specialized tent and Schlenk line environments. Acetylene will be used as a standard substrate to confirm the activity of the Av1/Av2 complex via gas chromatography. Methyl isonitrile synthesis will be performed under vacuum as described in *Organic Synthesis*[11,12] and visualized in Figure 4. Gas chromatography using various concentrations of methyl isonitrile and nitrogenase components will be performed to quantify kinetics of product formation. Ultimately, methyl isonitrile in complex with nitrogenase will be generated under turnover conditions[4] and subsequently crystallized to directly study C-C bond formation. Solving the structure of this complex will provide insight into this important biocatalytic bond formation reaction.
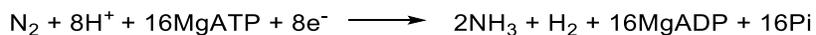
Experimental Approach Figures

$$N_2 + 8H^+ + 16MgATP + 8e^- \longrightarrow 2NH_3 + H_2 + 16MgADP + 16Pi$$

Figure 1. Nitrogen fixation performed by metallloenzyme nitrogenase. In this nitrogenase, dinitrogen is converted to ammonia through series of electron transfer reactions powered by ATP hydrolysis.[2]
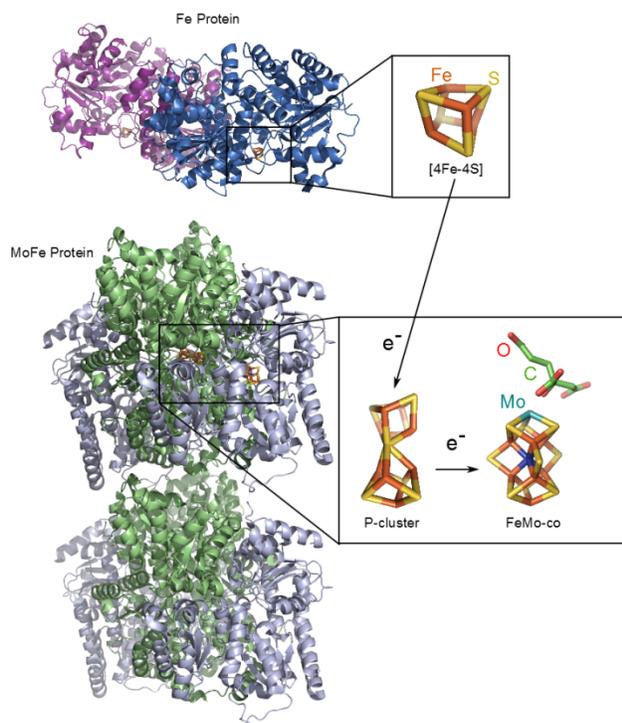

Figure 2. (left) Nitrogenase electron flow through Fe (composed of homodimeric gamma subunits in purple and blue)[16] and MoFe proteins (composed of alpha subunits in blue and beta subunits in green)[17] via 4Fe-4S cluster in Fe protein to P-cluster and FeMo-co in MoFe protein.
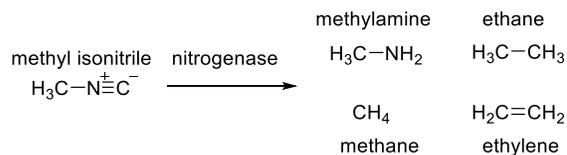

Figure 3. (above) Methyl isonitrile forms one and two carbon products in the presence of nitrogenase.
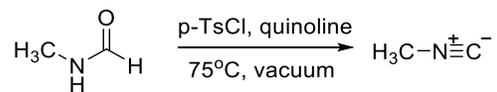

Figure 4. (above) Methyl isonitrile synthesis is performed under vacuum.[12]

Results

To begin this project, we first purified nitrogenase anaerobically (Figure 5). A fermenter was used to grow *Azotobacter vinelandii* (Av) which were then broken with an Emulsfilex and the resulting cell lysate was used to obtain Av1 (MoFe) and Av2 (Fe) from an ion exchange column (separating Av1 and Av2) followed by a size exclusion column for each protein. Purification was performed anaerobically using Schlenk lines.
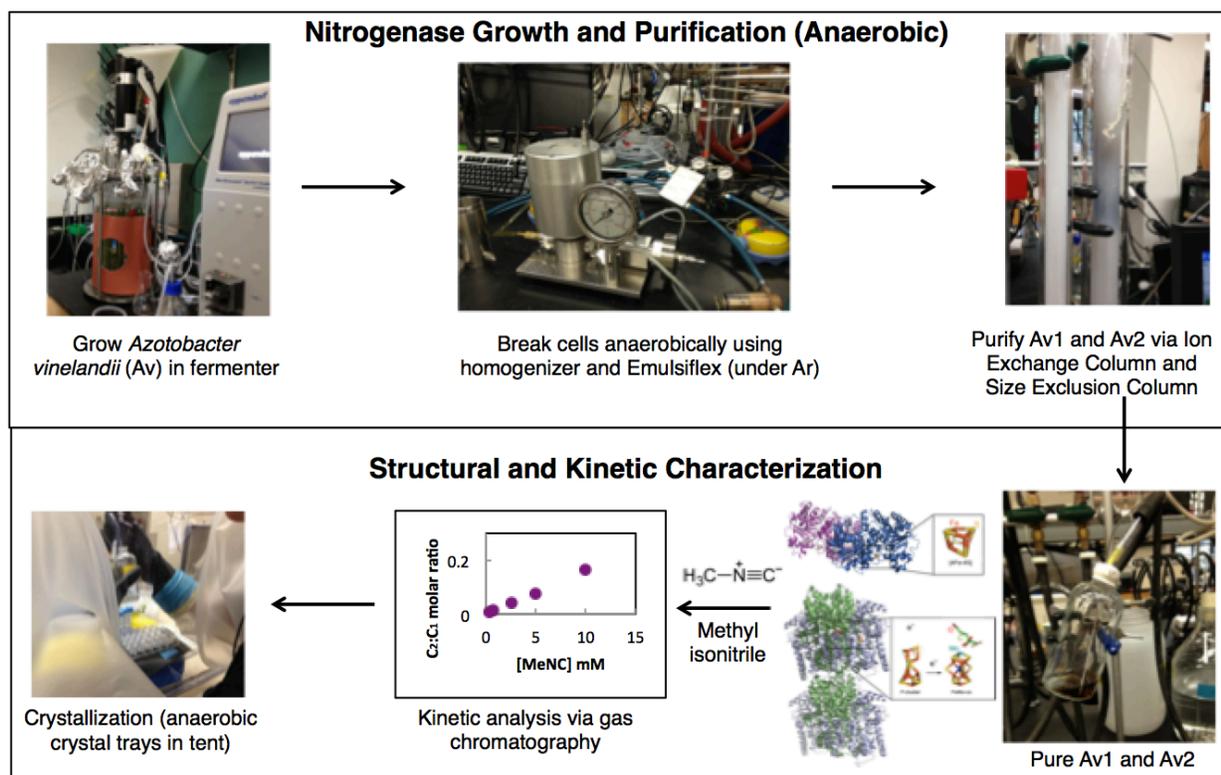
Figure 5. Purification of nitrogenase followed by gas chromatography and crystallization with methyl isonitrile.

Methyl isonitrile was purchased from the abcd company (a German chemical company). With nitrogenase and methyl isonitrile in hand, we first sought to reproduce observation of C-C bond formation via gas chromatography.

Our first task was to use gas chromatography and PeakSimple software to optimize for C-C bond formation conditions. These reactions are known as activity assays as they are a measure of nitrogenase activity for a given substrate (Figure 6). Performing an activity requires turnover conditions, also referred to as an ATP regeneration system such a setup requires MgCl2, ATP, creatine phosphate, phosphocreatine kinase, and a buffered solution. Activity assays are then

performed with these conditions under Argon in the presence of dithionate, nitrogenase, and substrate of interest, in this case methyl isonitrile.
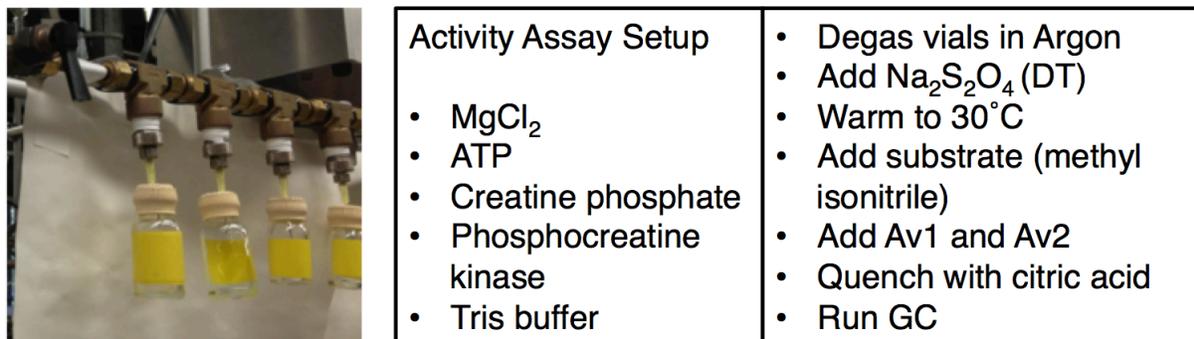
Activity Assay Setup	• Degas vials in Argon
	• Add $Na_2S_2O_4$ (DT)
• $MgCl_2$	• Warm to 30˚C
• ATP	• Add substrate (methyl
• Creatine phosphate	isonitrile)
• Phosphocreatine	• Add Av1 and Av2
kinase	• Quench with citric acid
• Tris buffer	• Run GC

Figure 6. Anaerobic setup (left), activity assay conditions (center), and activity assay procedure (right) for monitoring nitrogenase activity.

We first tested the activity of the purified nitrogenase using a standard acetylene reduction assay (Figure 7) and found that the activity fell on the low side of the expected range (usually ~2000 nmol ethylene/min/mg Av1), meaning we could then go on to use this batch of nitrogenase with our substrate of interest, methyl isonitrile.
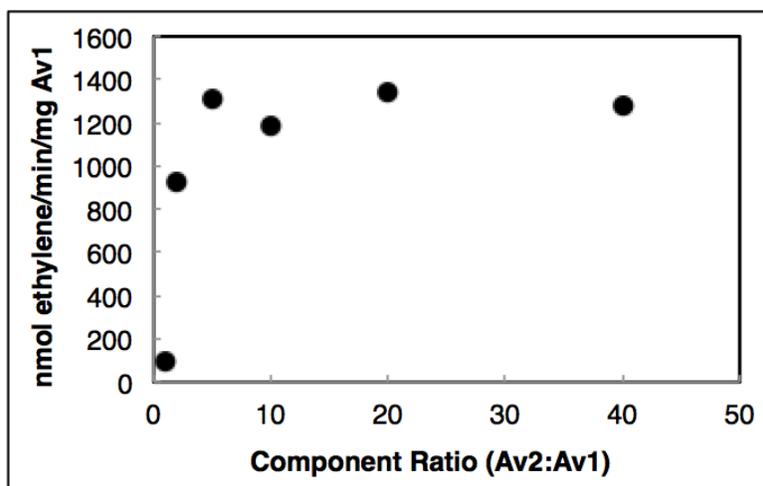
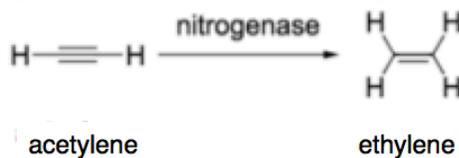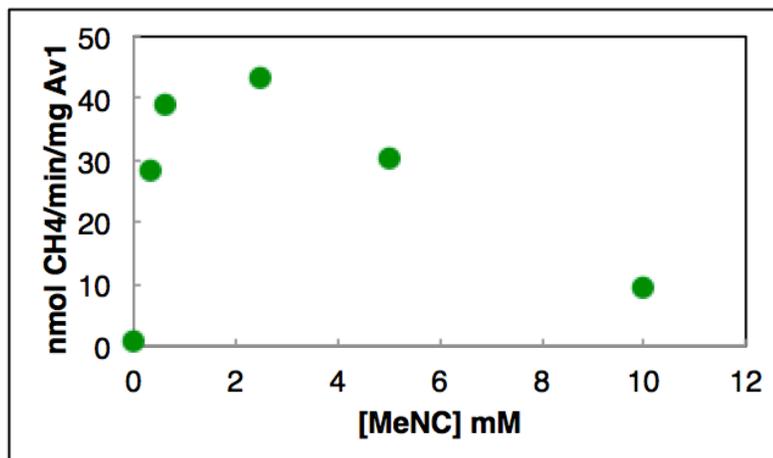Activity Assay Conditions: 10', 0.125 mg Av1, pH 7.5

Figure 7. Acetylene reduction assay as monitored by gas chromatography.

As expected,[15] when methyl isonitrile is incubated with nitrogenase it is reduced to methane (Figure 8). Methane formation peaks around 2.5 mM MeCN and becomes inhibited after this concentration as MeCN can act as both a substrate and an inhibitor.

Activity Assay Conditions: 10', 0.125 mg Av1, CR 10, pH 7.5
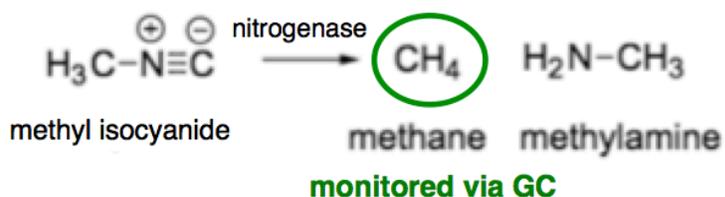
Figure 8. Methyl isonitrile is reduced to methane in the presence of nitrogenase.

When methyl isonitrile is incubated with nitrogenase it also forms ethane and ethylene (Figure 9). C2 bond formation appears to peak around 5mM methyl isonitrile. These results are similar to those found in the literature for C2 bond formation under methyl isocyanide titration conditions.[13,14] To further understand this C-C bond formation chemistry, we will attempt to optimize conditions for C2 product formation.
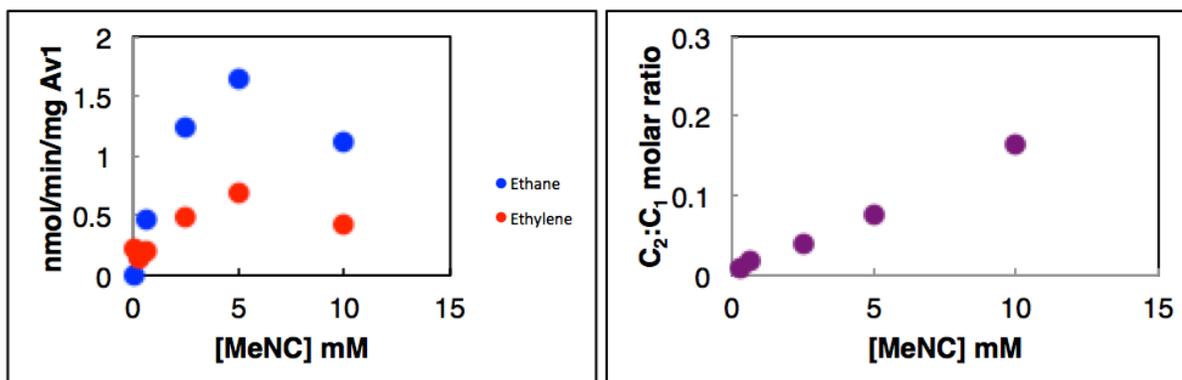
Activity Assay Conditions: 10', 0.125 mg Av1, CR 10, pH7.5
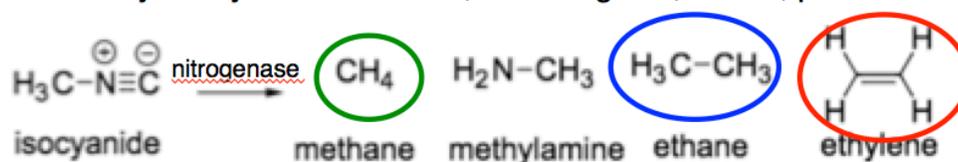
Figure 9. Methyl isonitrile forms C-C bonds in the presence of nitrogenase. Peak bond formation appears at about 5 mM methyl isonitrile.

Following the methyl isonitrile titration, favorable conditions for C2 bond formation were found at CR 10 and pH 7.5. Through this project we found optimized conditions for C2 production, namely a component ratio of 10, pH 7.5, [MeCN] = 5.0 mM. With these conditions in hand we sought to use crystallization to study the structure of nitrogenase interacting with methyl isonitrile. Although these conditions did not yield crystals, the optimization of C2 product formation is still very valuable for understanding C-C bond formation from nitrogenase.

References

1. Rees, D; Tezcan, FA; Haynes, C; Walton, M; Andrade, S; Einsle, O; and Howard, J (2005) Structural basis of biological nitrogen fixation. Phil. Trans. R. Soc. A., 363, 971-984.

2. Burgess, BK, and Lowe, DJ. (1996) Mechanism of Molybdenum Nitrogenase. Chem. Rev. 96. 2983-3011.

3. Seefeldt, L; Hoffman, B, and Dean, D. (2009) Mechanism of Mo-Dependent Nitrogenase. Annu Rev Biochem. 78: 701-729.

4. Spatzal, T; Perez, K; Einsle, O; Howard, J, and Rees, D (2014) Ligand binding to the FeMo-cofactor: Structures of CO-bound and reactivated nitrogenase. Science 345: 1620-1623.

5. Wolle, D; Kim, C; Dean, D and Howard, J. (1992) Ionic Interactions in the Nitrogenase Complex. JBC, 267, 6: 3667-3673.

6. Yang, K; Haynes, C; Spatzal, T; Rees, D, and Howard, J (2013) Turnover-Dependent Inactivation of Nitrogenase MoFe-Protein at High pH. Biochemistry, 53: 333-343.

7. Spatzal, T; Perez, K; Howard, J, and Rees, D. (2015) Catalysis-dependent selenium incorporation and migration in the nitrogenase active site iron-molybdenum cofactor. eLife; 10.7554/eLife.11620.

8. Jang, S.B.; Seefeldt, L.C.; Peters, J.W. (2000) Insighting into nucleotide signal transduction in nitrogenase: structure of an iron protein with MgADP bound. Biochemistry, 39: 14745-14752.

9. Einsle, O; Tezcan, F.A.; Andrade, S.L.; Schmid, B.; Yoshida, M.; Howard, J.B.; Rees, D.C. (2002) Nitrogenase MoFe-protein at 1.16 A resolution: a central ligand in the FeMo-cofactor. Science, 297: 1696-1700.

10. Spatzal, T et al. Unpublished.

11. Smith, JO. (2012) Candidacy Report. California Institute of Technology.

12. Schuster, RE; Scott, JE, and Casanova, J. Methyl isocyanide synthesis. Organic Synthesis, Coll. Vol. 5, p. 772 (1973); Vol. 46 p. 75 (1966).

13. Kelly, M; Postgate, J.R., and Richards, R.L. (1966) Reduction of Cyanide and Isocyanide by Nitrogenase of Azotobacter chroococcum. Biochem J. 102, 1: 1c-3c

14. Kelly, M. (1967) The Kinetics of the Reduction of Isocyanides, Acetylenes and the Cyanide Ion by Nitrogenase Preparation from *Azotobacter chroococcum* and the Effects of Inhibitors. Biochem. J. 107,1: 1-6.

15. Rubinson, JF; Corbin, JL; and Burgess, BK. (1983) Nitrogenase Reactivity: Methyl Isocyanide as Substrate and Inhibitor. Biochemistry. 22, 6260-6268.

16. Georgiadis, MM; Komiya, H; Chakrabarti, P; Woo, D; Kornuc, JJ; Rees, DC. (1992) Crystallographic structure of nitrogenase iron protein from Azotobacter vinelandii. Science. 257, 1653-1659.

17. Kim, J, and Rees, DC. (1992) Structural models for the metal centers in the nitrogenase molybdenum-iron protein. Science. 257, 1677-1682.

18. Research and Markets. "The Ethylene Technology Report 2016."

19. Ghanta, M; Fahey, D, and Subramaniam, B. (2013) Environmental impacts of ethylene production from diverse feedstocks and energy sources. Appl Petrochem Res, 4: 167-179.

Chapter 6

NITROGENASE SUBSTRATE SCOPE

Camille Z. McAvoy and Douglas C. Rees

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125

Abstract

To compliment nitrogenase experimental work on methyl isonitrile I performed a literature review on nitrogenase substrates. Nitrogenase is capable of handling a wide variety of substrates, which I have organized here in a substrate table. This review shows the breadth of reactions nitrogenase is capable of catalyzing. Further, I go on to assess the role of hydrazine in nitrogenase mechanism and present the cases for and against this potential intermediate.

Introduction

Understanding the nitrogenase substrate scope is valuable for further understanding the enzyme itself. In this review I identify various substrates through the literature and compile a thorough list of these substrates. Among this substrate scope is hydrazine, a potential intermediate in dinitrogen reduction. I explore the evidence for and against hydrazine as an intermediate in this enzymatic reaction.

Table of Nitrogenase Substrates

This table represents a literature survey of know substrates and inhibitors of nitrogenase presented with relevant binding information and sources.

Table 1. Table of nitrogenase substrates demonstrating the variety of structures nitrogenase is capable of accommodating.

Substrate	Product(s)	K_M
H^+	H_2	
N_2	$2NH_3$	0.1-0.2 atm
N_2H_4	$2NH_3$ (2e⁻)	20-30 mM
N_3^-	N_2+NH_3 (2e⁻) $N_2H_4+NH_3$ (6e⁻) $3NH_3$ (8e⁻)	1-15mM (HN_3)
N_2O	N_2 (2e⁻)	1 mM
NO_2^-	NH3 (6e-)	
C_2H_2	C_2H_4 cis $C_2H_2D_2$	0.003 – 0.02 atm
C_2H_4	no reduction?	
$CH_3C≡CH$	CH_3CHCH_2	0.5 atm
$C_2H_5C≡CH$	$C_2H_5CHCH_2$	too large to measure
$CH_3C≡CCH_3$	no reduction? $CH_3C≡CCH_3$	
$H_2C=C=CH_2$	$CH_3C≡CH$ 2,3 D_2-propene	0.02-0.08 atm
3,3-difluorocyclopropene	$CH_3CF=CH_2+HF$(4e⁻) $CH_2CH=CH_3+2HF$?	0.03 atm
cyclopropene	cyclopropane (2e⁻) cis D_2 cyclopropane $CH_3C≡CH$ (2e⁻)	0.012 atm
$N≡CNH_2$		
diazirine	$CH_3NH_2+NH_3$ (6e⁻) CH_4+2NH_3 (8e)	0.0012 atm

HCN (CN⁻?)	CH_3NH_2 (4e⁻) CH_4+NH_3 (8e⁻) CH_2NH (2e⁻)	4.5 mM HCN
CH_3CN	$C_2H_6+NH_3$ (6e⁻)	500-100 mM
C_2H_5CN	$C_3H_8+NH_3$ (6e⁻)	
C_3H_7CN	$C_4H_{10}+NH_3$ (6e⁻)	
$(CH_3)_2CHCN$		not a substrate
CH_3NC	CH_3NHCH_3 (4e⁻) $CH_2NH_2+CH_4$ (6e⁻)	0.7 mM
C_2H_5NC		
$H_2C=CHNC$	$C_3H_6+NH_3$ (6e⁻) $C_3H_8 + NH_3$ (8e⁻)	(0.8 mM)? 10-50 mM
H_2	inhibitor	
CO	inhibitor	
NO	inhibitor	
CH_2N_2		
CO_2		
CS_2		1.7 mM
COS		
$CH_3C_6H_4C\equiv CH$	$CH_3C_6H_4C=CH$	
$HC\equiv C-CH_2OH$	$HC=C-CH_2OH$	
NH_2OH	NH_3	30 mM
$HC\equiv CCH_2N_2$	$HC=CCH_2N_2$	
$H2C\equiv CCH2OH$	$H2C=CHCH2OH$	(alpha-70Ala)
HN=NCH3	HN=NCH3	
HN=NH	NH3	4.5 mM
HN=NCH3	inhibitor	
H+	HD	
$N\equiv CS^-$	HCN+H2S, (2e-) CH4+NH3(6e-)	0.9 mM
OCN⁻	CO+H2S (2e-)	6.1 mM (pH 6.0) 20 mM (pH 6.5)
Trans-dimethyldiazerene		
$HC\equiv CCHOH$	$HC=CCHOH$	
$HC\equiv CCH2NH2$	$HC=CCH2NH2$	
SeCN-		

References

1. B.K. Burgess in Molybdenum Enzymes, Cofactors and Model Systems, ACS Symp. 535, Stiefel, Newton and Coucouvanis, eds, pp. 144-169 (1993)

2. B.K. Burgess in Molybdenum Enzymes, Spiro, ed. pp. 161-219. (1985)

3. Richards. Reactions of msll molecules at transition metal sites: studies relevant to nitrogenase, an organometallic enzyme. Coordination Chem Rev. pp. 83-97 (1996)

4. Santos et al. Alkyne substrate interaction with Nitrogenase MoFe protein. J. Inorg. Biochemistry. pp. 1642-1648 (2007)

5. Shaw et al. Nitrite and Hydroxylamine as Nitrogenase Substrates: Mechanistic Implications for the Pathway of N2 Reduction. JACS pp. 12776-12783

6. Seedfelt et al. Mechanism of Mo-Dependent Nitrogenase. Annu. Rev. Biochem. 78 pp. 701-722 (2009) – Hydrazine as an intermediate in N2 reduction

7. Barney et al. Diazene Is a Substrate for Nitrogenase: Insights into the Pathway for N2 Reduction. Biochemistry, 46. pp. 6784-6794 (2007)

8. Barney et al. A methyldiazene-derived species bound to the nitrogenase active-site FeMo cofactor: Implications for mechanism. PNAS 103, pp. 17113-17118 (2006)

9. Sellman et al. The nitrogenase catalyzed N2 dependent HD formation: a model reaction and its significance for the FeMoco function. Coord. Chem. Rev. pp. 545-561 (2000)

10. Rasche et al. Reduction of Thiocyanate, Cyanate, and Carbon Disulfide by Nitrogenase: Kinetic Characterization and EPR Spectroscopic Analysis. Biochemistry. pp. 8574-8585. (1997)

11. Barney et al. (2006) Breaking the N2 triple bond: insights into the nitrogenase mechanism. Dalton Trans. 2277-2284.

Potential Role of Hydrazine as an Intermediate in Dinitrogen Reduction by Nitrogenase

Abstract

The metalloenzyme nitrogenase catalyzes the conversion of dinitrogen to ammonia. The mechanism of nitrogenase's conversion of dinitrogen to ammonia remains unclear despite decades of research. Here we will review evidence for and against hydrazine as an intermediate in this pathway. Further, our analysis will serve as a review of nitrogenase mechanistic literature.

Introduction

Nitrogenase is a metalloenzyme from soil bacteria that catalyzes the conversion of dintrogen (N_2) to bioavailable ammonia as shown in the below equation[1]:

$$N_2 + 8e^- + 16MgATP + 8H^+ \rightarrow 2NH_3 + H_2 + 16MgADP + 16P_i$$
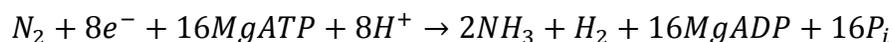
Nitrogenase is composed of two component proteins, the MoFe protein (also known as Av1 in *Azotobacter vinelandii*) and the Fe protein (or Av2) which work together to carry out dinitrogen reduction. MoFe protein contains the binding site for substrates composed of a 9S-Mo-X-homocitrate cofactor (FeMo-cofactor). MoFe also contains the P-cluster [8Fe-7S] which mediates electron transfer from the Fe protein. The Fe proton transfers electrons to MoFe coupled with stoichiometry of 1e⁻/2 ATP.

Although this enzyme has been extensively studied, its mechanism remains elusive. One possibility is a Chatt-type mechanism or alternating pathway mechanism in which nitrogenase is first converted to diazene which goes through a hydrazido state to arrive at hydrazine that is then converted to ammonia. An alternative mechanism, called the distal pathway, involves going through a diazenido state, releasing one ammonium to a nitrido state then to an imido state and finally the second ammonium. In this review we will assess evidence for and against hydrazine as an intermediate in nitrogenase's conversion of dinitrogen to ammonia.

Results

There are five general arguments that support hydrazine as an intermediate: 1) hydrazine can act as a substrate of nitrogenase,[1–9] 2) hydrazine can be detected when nitrogenase is acid-quenched,[10] 3) V-nitrogenase releases hydrazine as a minor product,[11–15] 4) the conversion of hydrazine to

ammonia can be catalyzed by various nitrogenase-like synthetic catalysts,[16–20] and 5) hydrazine can be observed as an intermediate via EPR.[8,9] These possibilities are summarized in the table below with their respective sources.

Table 2. Arguments for and against hydrazine as an intermediate with the respective sources.

For Hydrazine Intermediate	Against Hydrazine Intermediate	Source
1) Hydrazine acts as a substrate to nitrogenase	Low reactivity in WT nitrogenase; nitrogenase must be mutated to get higher activity	1-9
2) Hydrazine appears when nitrogenase is acid quenched	Could be an artifact of the presence of acid	10
3) Hydrazine is released by V-nitrogenase during N_2 reduction as a minor product; formation of small amounts of hydrazine during catalytic activity	Doesn't necessarily mean Mo-nitrogenase follows same mechanism; isn't necessarily on pathway to NH_3	11-15
4) Hydrazine to ammonia can be catalyzed by MFe_3S_4 catalysts, various nitrogenase metal catalytic mimics, molybdenum thiolate complexes, and mononuclear MCP*Me$_3$ cores.	How physiologically relevant to nitrogenase are these catalysts	16-21

1) Although these evidences support hydrazine as an intermediate for biological nitrogen fixation, there is also significant evidence against this intermediate. Although hydrazine can act as a substrate of nitrogenase, it is a very poor substrate in WT nitrogenase. To study hydrazine as a substrate, mutant forms of nitrogenase were made. To accommodate hydrazine, nitrogenase was mutated to α-70^{Ala}/α-195^{Gln}, hydrazine was used as a substrate, and this complex was freeze trapped and monitored by EPR.[1] Additionally $^{15}N_2H_4$ was used with pulsed ENDOR to confirm that the intermediate observed by EPR was derived from hydrazine. In another work studying hydrazine as a substrate, it was found that when Fe protein is substituted with an E^{II}-DPTA system and a β-98^{His} mutant of MoFe is made, hydrazine is reduced to ammonia but N_2 is not reduced.[4] Although hydrazine is able to act as a substrate in this system, it is not clear whether it is able to do

the same in the WT nitrogenase system. In the absence of mutations, hydrazine is a poor substrate for WT nitrogenase, constituting only about 10% of electron flow compared to protons. Further, hydrazine can act as a noncompetitive inhibitor of acetylene reduction. Changing α-70^{Val} to alanine greatly improves hydrazine reduction, as previously obsserved.[9] When α-70^{Val} is substituted with isoleucine, hydrazine reduction is eliminated. Lukoyanov et al. used cryoannealing to trap a diazene (2N2H) species and monitor it via EPR.[8] Hydrazine could then act as an intermediate after decay of this species. Thus Lukoyanov et al. do not specifically observe hydrazine as an intermediate but rather a diazene-level intermediate that could be on pathway to a hydrazine intermediate. Although hydrazine has been observed by EPR in the nitrogenase system this may not be an on pathway intermediate but rather an artifact of reaction conditions since WT nitrogenase is very unreactive with hydrazine. Overall it is unclear whether mutants that interact with hydrazine are relevant to WT nitrogenase and thus whether hydrazine acts as a substrate in the native nitrogenase remains unclear. It is possible that if hydrazine acts as an intermediate, these mutations mimic conditions for hydrazine already having access to active site upon reduction from dinitrogen.

2) The second argument involves seeing hydrazine upon acid or alkali quench. Although hydrazine could be a captured intermediate, it is also possible that acid or base produces hydrazine and that hydrazine would not be there otherwise. That is, it is possible another nitrogenous intermediate underwent maturation under these conditions to yield the observed hydrazine. Again, in straying from physiological conditions it is challenging to determine whether or not this evidence is biologically relevant. Hydrazine is a poor substrate of nitrogenase and is not found to accumulate

in solution, suggesting that if this species is relevant it is as an intermediate within nitrogenase and not as a free species.

3) It was found by Dilworth et al. that vanadium-based enzyme develops around 0.5% N_2H_2 per converted N_2 during catalytic turnover.[13] Thus in the Vanadium nitrogenase it appears that hydrazine is a product in V-nitrogenase dinitrogen reduction. Hydrazine is found to be free in solution and not enzyme-bound under these conditions and this suggests that hydrazine is a product for VFe and not a bound intermediate, in contrast to the Thorneley et al. quenching case. This does provide evidence for a four-electron reduced dinitrogen hydride species on V-nitrogenase capable of dissociating hydrazine.[13] Further, more hydrazine is produced from the vanadium nitrogenase with increasing temperature.[12] However, this doesn't necessarily mean that the molybdenum nitrogenase has the same mechanism and thus the same intermediate. Hydrazine has not been observed for the molybdenum nitrogenase as it has been for the Vanadium nitrogenase so this may represent a divergence between the two systems. Although V-nitrogenase and Mo-nitrogenase are similar in structure and function, V-nitrogenase evolves more H_2 during N_2 reduction.

4) Crossland et al. show η^2 coordination of hydrazine to iron in a small molecule comples, which may mimic nitrogenase intermediates during turnover.[16] Coucouvanis et al. demonstrate the reduction of hydrazine to ammonia using MFe_3S_4 clusters (M=Mo, V) demonstrating that nitrogenase could possibly perform similar hydrazine reduction chemistry. Although hydrazine can be transformed to ammonia by small molecule clusters, this again isn't for certain evidence that hydrazine acts as an intermediate. Rather it is indicative of hydrazine's potential for reactivity,

which demonstrates that hydrazine can be converted to ammonia but it is unclear whether this happens in the physiological case.

Rittle et al. show that N_2 can be converted to hydrazine and then to ammonia on a Fe catalyst via Fe=NNH$_2$, thus demonstrating a hybrid distal to alternating pathway for N_2 reduction. This indicates that hydrazine can appear in the reduction pathway even if its not a purely alternating model which opens the possibility of hydrazine acting as an intermediate in a variety of models. Again, although this is observed through synthetic catalysis it is uncertain how well this physiologically mimics nitrogenase. Further, hydrazine was observed after acid quenching and it is unclear whether this represents a biologically relevant condition.

Conclusion

Overall, there is strong evidence to support hydrazine as an intermediate in N_2 reduction to ammonia by nitrogenase. Hydrazine can act as a substrate, though it is a weak substrate of WT nitrogenase and works better with mutated forms of the enzyme.[1] Further, hydrazine has been observed upon acid quenching of nitrogenase indicating that it is a potential intermediate in N_2 to NH$_3$ reduction.[10] However, despite this evidence it is not certain whether hydrazine in on pathway to N_2 reduction as a hydrazine intermediate has not been directly detected as a product important in N_2 reduction. Its poor reactivity as a substrate for WT nitrogenase indicates that it may not be an on pathway intermediate; similarly the need to mutate nitrogenase to accommodate hydrazine to observe this intermediate also indicates the alternating mechanism may not necessarily be physiological. Further, acid quenching could cause the production of hydrazine which may not be in the active site otherwise. Thus it is unclear whether hydrazine is an intermediate in N_2 reduction.

References

1. Barney, B. M., Laryukhin, M., Igarashi, R. Y., Lee, H. I., Dos Santos, P. C., Yang, T. C., Seefeldt, L. C. Trapping a Hydrazine Reduction Intermediate on Nitrogenase Active Site. *Biochemistry* **44,** 8030–8037 (2005).
2. Yandulov, D. V., and Schrock, R. R. Catalytic reduction of dinitrogen to ammonia at a single molybdenum center. *Science (80-.).* **301,** 76–78 (2003).
3. Davis, L. Hydrazine as a substrate and inhibitor of Azotobacter vinelandii Nitrogenase. *Biochem Biophys.* **204,** 270–276 (1980).
4. Danyal K;Inglet B;Vincent K;Barney B;Hoffman B;Armstrong F;Dean D;Seefeldt L. Uncoupling Nitrogenase: Catalytic Reduction of Hydrazine to Ammonia by a MoFe Protein in the Absence of Fe Protein-ATP. *JACS* **132,** 13197–13199 (2010).
5. Barney B;Lee H;Dos Santos P;Hoffman B;Dean D;Seefeldt L. Breaking the N2 triple bond: insights into the nitrogenase mechanism. *Dalt. Trans* 2277–2284 (2006).
6. Hoffman B;Dean D;Seefeldt L. Climbing Nitrogenase: Toward a Mechanism of Enzymatic Nitrogen Fixation. *Acc. Chem. Res.* **42,** 609–619 (2009).
7. Crossland, J; Tyler, D. Iron-dinitrogen coordination chemistry: Dinitrogen activation and reactivity. *Coord. Chem. Rev.* **254,** 1883–1894 (2010).
8. Lukoyanov. Identification of Key Catalytic Intermediate Demonstrates That Nitrogenase Is Activated by Reversible Exchange of N2 for H2. *JACS* **137,** 3610–3615 (2015).
9. Barney, BM, Robert Y. Igarashi, Patricia C. Dos Santos, D. R. D. & Seefeldt, and L. C. Substrate Interaction at an Iron-Sulfur Face of the FeMo-cofactor during Nitrogenase Catalysis. *JBC* **279,** 53621–53624 (2004).
10. Thorneley, R. N. F., Eady, R. R., and Lowe, D. J. Biological nitrogen fixation by way of an enzyme-bound dinitrogen-hydride intermediate. *Nature* **272,** 557–558
11. Fisher K; Dilworth M;Newton W. Azotobacter vinelandii Vanadium Nitrogenase: Formaldehyde Is a Product of Catalyzed HCN Reduction, and Excess Ammonia Arises Directly from Catalyzed Azide Reduction. *Biochemistry* **45,** 4190–4198
12. Dilworth, M. J., Eldridge, M. E., and Eady, R. R. The molybdenum and vanadium nitrogenases of Azotobacter chroococcum: effect of elevated temperature on N2 reduction. *Biochem. J* **289,** 395–400 (1993).
13. Dilworth, M.J.; Eady, R. R. Hydrazine is a product of dinitrogen reduction by the vanadium-nitrogenase from Azotobacter chroococcum. *Biochem J* **277,** 465–468
14. Hoffman. Nitrogenase: A Draft Mechanism. *Acc. Chem. Res.* **46,** 587–595 (2013).
15. Tsagkalidis, C. W. and D. R. Hydrazine complexes of vanadium(I1 and III). *Inorganica Chim. Acta* **205,** 239–243 (1993).
16. Crossland, JL, Lev N. Zakharov, and D. R. T. Synthesis and Characterization of an Iron(II) η2-Hydrazine Complex. *Inorg. Chem* **46,** 10476−10478 (2007).
17. Coucouvanis. Catalytic and stoichiometric multielectron reduction of hydrazine to ammonia and acetylene to ethylene with clusters that contain the MFe3S4 cores (M- Mo, V). Relevance to the function of nitrogenase. *J. Mol. Catal. A Chem.* **107,** 123–135 (1996).
18. Chu, W, Chi-Chin Wu, and H.-F. H. Catalytic Reduction of Hydrazine to Ammonia by a Vanadium Thiolate Complex. *Inorg. Chem* **45,** 3164−3166 (2006).

19. Rehder, D. Vanadium nitrogenase. *J. Inorg. Biochem.* **80,** 133–136 (2000).
20. Vela, J, Stoian, S, Flaschenriem, CJ, Munck, E. & Holland, P. L. A Sulfido-Bridged Diiron(II) Compound and Its Reactions with Nitrogenase-Relevant Substrates. *J. AM. CHEM. SOC* **126,** 4522–4523 (2004).
21. Rittle, J and Peters, J. An Fe-N2 Complex That Generates Hydrazine and Ammonia via Fe=NNH2: Demonstrating a Hybrid Distal-to-Alternating Pathway for N2 Reduction. *J. AM. CHEM. SOC* **138**, 4243-4248 (2016).

Conclusion

OF CHAPERONES AND ENZYMES

Biological tools are useful case studies for organic chemists, biochemists, and beyond. My results from cpSRP43 and nitrogenase demonstrate that we should keep learning from biological systems to create pharmaceuticals and other industrial products.

Through studying cpSRP43 we demonstrated the interaction of cpSRP43 with its substrate, LHCP, and characterized this interaction surface to be the TM's, loop 1, and L18, from the point of view of the substrate and the beta-sheets, bridging helix, and Y204 from the point of view of the chaperone. Further we showed how cpSRP43 uses effectors to be finely spatially and temporally controlled. cpSRP43 was biophysically characterized to undergo an unfolded-folded transition upon binding substrate and cpSRP54M. cpSRP43 was then used as a tool to increase membrane protein expression. cpSRP43 can also potentially be used as a tool to prevent the aggregation of amyloid-beta found in Alzheimer's and other such disease-related aggregation-prone proteins. cpSRP43 is thus a well, characterized tool with great engineering potential to tackle protein aggregates in human disease.

In studying nitrogenase, I optimized conditions for C-C bond formation with methyl isonitrile. These conditions were then used to setup crystal trays to study the structure of this complex. Further, I studied the substrate scope of nitrogenase and the role of hydrazine in the nitrogen fixation pathway.

Overall, these works serve as fascinating case studies into using existing biological machinery for useful medical and industrial applications.