

# **A Study of Protein Targeting Reveals Insights into Mitigating Protein Aggregation**

Thesis by

Thang Xuan Nguyen

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2013

(Defended May 31, 2013)

© 2013

Thang Xuan Nguyen

All Rights Reserve

**For my family:  
My mother, Vũ thị An;  
my sisters and brother, Chị Ti, Chị Vi & Anh Tèo;  
my nephews and nieces.  
and my entire extended family.  
It is my good karma to be connected to you all.**

**And in loving memory of my father,  
Nguyễn xuân Thọ (1946-2010),  
a warrior and an artist,  
who taught us to live with compassion, wisdom and courage.**



# Acknowledgements

I thank my advisor, Shu-ou Shan. Shu-ou is an intense boss who expects the best from her students and will not accept anything less. I appreciate all her efforts in guiding me as a graduate student and a scientist, to approach research with creativity, hard work and audacity. I will be forever grateful for the guidance that she has given me throughout my graduate career.

I thank my committee members – Professors Bil Clemons, Linda Hsieh-Wilson and David Chan – for their critiques, encouragements and intellectual conversations over the years.

I thank my UCI undergraduate advisor, Professor James Nowick, for his mentorship. James was the one who pushed me to apply to Caltech.

I thank my dear friends and colleagues in the Shan lab for making my experience at Caltech amazing. I am fortunate to have met all of you. Sowmya Chandrasekar, Peera Jarupornpan, Xin Zhang, Rumana Rashid, Vinh Lam, Shen Kuang, Ishu Saraogi, Aileen Ariosa, David Akopian, Dawei Zhang, Nathan Piece, Mike Rome, Meera Rao, George Liang, Cecilia Zurita-Lopez, Shaobin Guo, Camille McAvoy, Yu-Hsien Hwang Fu, Un-seng Chio, Samantha Piszkievicz and Sandhya Chandrasekaran. Thank you to my bay-mate Ishu Saraogi for sharing her perspectives in science and beyond. I must also add a special shout-out for Team Chloroplast – Sowmya, Peera, George, Sam and (now) Camille – because we are the awesome-est. I thank Aileen for being awesome too, and proofreading this manuscript. Thanks for the sweat, tears, hugs and laughter.

I thank the collaborators that have helped me in my research, both inside and outside of campus. I thank Alastair McDowall of the EM facility for his help in TEM imaging and trials with cryo-EM, Sonja Hess and her team at the Proteome Expression Laboratory for help with the

NEM alkylation and intact mass spectrometry work, Professor Betsy Komives of University of California San Diego for her help with H/D exchange trials and Professor Jeffery Kelly of the Scripps Research Institute for guidance with amyloid-beta experiments. I have enjoyed working with rotation students Va Si and Camille McAvoy and SURF student Samantha Piskiewicz, and wish them the best in their future endeavors. Collaboration within the Shan group has been a blessing and my research would not have gone far without work from the Chloroplast Team, especially from Sowmya Chandrasekar who taught me all the basics of biochemistry and shared the bulk of the work in Chapter 1 and Peera-Jaru-Ampornpan who has imparted to me knowledge and skills. Most importantly, I inherited the cpSRP43 project from her and she contributed significantly to Chapters 2 and 3 of this book.

I thank the support staff Margot Hoyt, Santiago Laparra, Manuel Lagang and Blanca Ortega – you make work in the lab much easier for us all. I thank Agnes Tong who deals with all matters related to the Chemistry Graduate Program. I thank the high school students who have volunteered in our lab throughout the years: Kevin Kim, Catherine Chan, Irene Chen, Albert Gianatan and Chris Chi.

I thank my parents for all they have given me. I thank my sisters and brothers for their support, emotionally and otherwise. I thank all my friends, especially the Tám Central Sangha – Ngân, Wil, Cat, PhuongDuy, Joslyn, Ethan, Andy, HồngNghi, Nikki, Vi, Thảo, Emily, Hậu-Nam, Triết, anh MinhAnh, and anh Tiên, for their friendship. Lastly, I thank my best friend, Jason Lộc, for the journey.

# Contents

<b>Dedication</b> . . . . .	i
<b>Acknowledgments</b> . . . . .	iii
<b>Table of Contents</b> . . . . .	v
<b>List of Figures</b> . . . . .	vii
<b>List of Tables</b> . . . . .	x

## Part One

### **Chapter 1:** Concerted Complex Assembly and GTPase Activation in the Chloroplast Signal

Recognition Particle . . . . .	1
--------------------------------	---

## Part Two

### **Chapter 2:** Mechanism of an ATP-independent Protein Disaggregase. Part I. Structure of a

Membrane Protein Aggregate Reveals a Mechanism of Recognition by its Chaperone . . . . .	47
--	----

### **Chapter 3:** Mechanism of an ATP-independent Protein Disaggregase. Part II. Distinct Molecular

Interactions Drive Multiple Steps During Aggregate Disassembly . . . . .	85
--	----

<b>Chapter 4:</b> Mitigating Protein Aggregation: Towards the Application of cpSRP43 as a Versatile Chaperone for Membrane Protein Systems . . . . .	131
--	-----



# List of Figures

- Figure 1.1** Fluorescence assays to report on complex assembly between cpSRP54 and cpFtsY
- Figure 1.2** Thermodynamic and kinetics for formation of the cpSRP54-cpFtsY complex
- Figure 1.3** Effects of cpFtsY mutations on its stimulated GTPase reaction with cpSRP54
- Figure 1.4** Effects of cpSRP54 mutations on its stimulated GTPase reaction with cpFtsY
- Figure 1.5** Effect of mutant GTPases on the targeting and integration of LHCP into thylakoid membranes
- Figure 1.6** Free energy profile for the GTP-dependent binding and activation cycles between the SRP and SR GTPases from bacteria or chloroplast
- Figure 1.S1** Cys-lite cpSRP54 and cys-less cpFtsY and their fluorescently labeled versions are enzymatically active
- Figure 1.S2** FRET and acrylodan fluorescence signals from cpSRP54•cpFtsY complex could be competed away by EDTA or unlabeled cpSRP54
- Figure 1.S3** Equilibrium titrations of the cpSRP54•cpFtsY complex in GTP using different fluorescence assays
- Figure 1.S4** Defects of additional cpSRP54 and cpFtsY IBD loop mutants in complex formation and GTPase activation, measured by the stimulated GTPase reaction
- Figure 1.S5** The effects of additional cpSRP54 and cpFtsY mutations on LHCP targeting and integration
- Figure 2.1** LHCP aggregates contain exposed hydrophobic surfaces as detected by small molecule dyes
- Figure 2.2** LHCP aggregates are resistant to many detergents

- Figure 2.3** LHCP forms aggregates after a critical concentration
- Figure 2.4** Transmission Electron Microscopy analysis of LHCP aggregates
- Figure 2.5** Atomic Force Microscopy analysis of LHCP aggregates
- Figure 2.6** Mapping the LHC aggregates reveals exposed motif
- Figure 2.7** L18 has a strong tendency to be exposed on the surface of the aggregate
- Figure 2.8** Analysis of pyrene excimer fluorescence reveals an amorphous aggregate core
- Figure 2.9** Model for the global organization of LHC proteins in the aggregate
- Figure 2.S1** Sequence Alignment of LHCP and Lhcb5
- Figure 3.1** cpSRP43 makes highly sequence-specific interactions with the FDPLGL motif in the L18 sequence
- Figure 3.2** cpSRP43 can interact with a variety of LHCP TM mutants
- Figure 3.3** Schematics depicting quantitative analysis of the cpSRP43-mediated disaggregation reaction. Concentration dependences of the kinetics
- Figure 3.4** L18-binding mutants uncouple initial recognition of the aggregate from its subsequent solubilization
- Figure 3.5** LHCP TM mutants exhibit a wide range of disaggregation efficiencies
- Figure 3.6** Time courses for the alkylation reactions of cysteine residues in the L18 shows accessibility of WT and mutant LHC proteins
- Figure 3.7** The irreversible TM mutants form ultrastable aggregates
- Figure 3.8** Linear free energy analysis of the cpSRP43-mediated disaggregation reaction
- Figure 3.9** Working model for cpSRP43-mediated disaggregation reaction
- Figure 4.1** cpSRP43 can rescue LHC aggregates with TM swapped from unrelated membrane proteins

- Figure 4.2** The minimal recognition motif within L18 can be reduced to 11 amino acid containing the essential FDPLGL motif
- Figure 4.3** Co-expression of cpSRP43 improves the yield of L11- and L18-fused SERP1
- Figure 4.4** cpSRP43 inhibits A $\beta$ <sub>1-40</sub> fibrillization

# List of Tables

<b>Table 1.1</b>	Summary of equilibrium and kinetic properties of cpSRP54 and cpFtsY IBD-loop mutants
<b>Table 1.S1</b>	Summary of different classes of mutational effects
<b>Table 1.S2</b>	Kinetic parameters of additional mutants outside of the IBD loop
<b>Table 1.S3</b>	Mant-GTP binding affinity to the individual mutant GTPases
<b>Table 1.S4</b>	Basal GTPase activity of cpFtsY mutants
<b>Table 2.S1</b>	Inverse central linewidth ( $\Delta H^{-1}$ ) and fraction alkylated by NEM of individual Lhcb5 cysteine mutants upon aggregate formation
<b>Table 3.1</b>	Description of the LHCP TM mutants
<b>Table 3.2</b>	Thermodynamic and kinetic parameters in the disaggregation reaction
<b>Table 3.3</b>	Summary of the thermodynamic and kinetic parameters of the L18-binding mutants.
<b>Table 3.4</b>	Summary of the thermodynamic and kinetic parameters of the LHCP TM mutants
<b>Table 4.1</b>	Sequence of TM swap mutants

# Chapter 1

## Concerted Complex Assembly and GTPase Activation in the Chloroplast Signal Recognition Particle

A version of this chapter has been published as:

Nguyen, T.X., Chandrasekar, S., Neher, S., Walter, P., and Shan, S. (2011) *Biochemistry*, **50** (33), 7208–7217.

**Abstract**

The universally conserved signal recognition particle (SRP) and SRP receptor (SR) mediate the cotranslational targeting of proteins to cellular membranes. In contrast, a unique chloroplast SRP in green plants is primarily dedicated to the post-translational targeting of light harvesting chlorophyll *a/b* binding (LHC) proteins. In both pathways, dimerization and activation between the SRP and SR GTPases mediate the delivery of cargo; whether and how the GTPase cycle in each system adapts to its distinct substrate proteins have been unclear. Here, we show that interactions at the active site essential for GTPase activation in the chloroplast SRP and SR play key roles in the assembly of the GTPase complex. In contrast to their cytosolic homologues, GTPase activation in the chloroplast SRP–SR complex contributes marginally to the targeting of LHC proteins. These results demonstrate that complex assembly and GTPase activation are highly coupled in the chloroplast SRP and SR and suggest that the chloroplast GTPases may forego the GTPase activation step as a key regulatory point. These features may reflect adaptations of the chloroplast SRP to the delivery of their unique substrate protein.

## Introduction

Cotranslational protein targeting by the signal recognition particle (SRP) and the SRP receptor (SR) is a universally conserved pathway essential for the proper structure and function of the cell. Cytosolic SRP recognizes ribosomes translating SRP substrates and, via interactions with SR, delivers its cargo—the ribosome-nascent chain complexes—to the eukaryotic endoplasmic reticulum or the prokaryotic plasma membrane [1,2]. The functional core of SRP consists of a universally conserved SRP54 subunit, or Ffh in bacteria, and an SRP RNA[3]. SRP54 is comprised of three domains: (i) a methionine-rich M-domain, which provides the binding site for the substrate protein and the SRP RNA [4]; (ii) a GTPase G-domain that shares homology with the Ras-fold[5]; and (iii) an N-terminal N-domain that interacts with the ribosome [6,7]. Together the N- and G-domains comprise a structural and functional unit called the NG-domain. The SR (FtsY in bacteria) also contains an NG-domain highly homologous to that in SRP54. The GTP-dependent interaction between the NG-domains of SRP and SR guides the delivery of cargo to protein translocation machineries on the target membrane, and subsequent GTP hydrolysis in the complex drives the dissociation of SRP and SR, recycling them for additional rounds of protein targeting [8].

A notable exception to this classic SRP pathway is provided by the chloroplast SRP (cpSRP) [9]. The cpSRP pathway still uses the conserved SRP54 and SR GTPases (called cpSRP54 and cpFtsY, respectively). However, cpSRP lacks the otherwise universally conserved SRP RNA and is instead a heterodimeric protein complex comprised of cpSRP54 and cpSRP43, a novel SRP subunit unique to the chloroplast of green plants [10-13]. The most significant difference between the cytosolic and chloroplast SRP pathways lies in the nature of their substrate proteins. The cytosolic SRP must recognize its cargos within a milieu of translating

ribosomes in the cytosol, based on signal sequences that differ widely in size, shape, and amino acid composition. In contrast, the cpSRP is dedicated primarily to the post-translational delivery of the light-harvesting chlorophyll a/b binding (LHC) family of proteins [14,15]. LHC proteins are synthesized in the cytosol and imported into the chloroplast stroma, where they are recognized and captured by the cpSRP [10]. Analogous to the cytosolic SRP pathway, the interaction of cpSRP with cpFtsY brings the LHC proteins to the Albino3 (Alb3) translocase on the thylakoid membrane, where the LHC proteins are integrated and assembled into light harvesting complexes [16].

The similarities and differences between the cytosolic and chloroplast SRP pathways raise intriguing questions: How do the targeting machineries in each pathway meet the unique challenges posed by their substrate proteins, and what are the roles of the SRP and SR GTPases in this adaptation? Extensive work on the cytosolic SRP showed that during the SRP-FtsY interaction a series of discrete conformational changes provide multiple opportunities to exert regulation [17-21]. Assembly of a stable SRP-FtsY complex requires the formation of a transient “early” intermediate, which subsequently rearranges to a stable, “closed” complex. GTPase activation in the complex requires yet another rearrangement, the movement of the highly conserved insertion box domain (IBD) loops, which positions multiple catalytic residues adjacent to the bound GTP molecules and activates GTP hydrolysis [17]. Importantly, each GTPase rearrangement allows the SRP and FtsY to sense and respond to their biological cues. A correct cargo can accelerate the assembly of the SRP-FtsY complex while delaying its GTPase activation [19]. Delayed GTP hydrolysis provides an important time window for the targeting complex to search for the translocation machinery before GTP hydrolysis drives its irreversible disassembly. Once at the target membrane, the movement of the IBD loops, which mediates GTPase

activation, is crucial for driving the initiation of protein translocation [22]. Finally, the timing of GTP hydrolysis provides an important fidelity checkpoint: incorrect cargos, which fail to delay GTPase activation, could be more promptly rejected through premature GTP hydrolysis [19]. Thus, the uncoupling of complex assembly and GTPase activation steps in the bacterial SRP and FtsY is crucial for ensuring the efficiency and fidelity of cotranslational protein targeting.

On the other hand, the cpSRP handles substrate proteins of a completely different nature. The LHC family of proteins comprises 30–50% of the protein content in the thylakoid membrane and are likely the most abundant membrane proteins on earth. The sheer abundance and rapid turnover of these proteins demand a highly robust and efficient pathway for their targeting and integration. Compared to its cytosolic homologue, specific substrate selection is much easier to achieve in the cpSRP, as members of the LHC protein family are highly homologous and share a conserved sequence motif, L18, that is specifically recognized by the cpSRP. Consequently, many features have evolved in the cpSRP pathway that may represent adaptations to its unique substrate proteins. For example, cpSRP uses cpSRP43 to efficiently capture the LHC proteins [23] as well as to help localize the targeting complex to Alb3 on the thylakoid membrane [24]. Here we address this issue from a different perspective: what are the similarities and differences in the GTPase cycles of the chloroplast versus cytosolic SRP and SR? Are there distinct features of the cpSRP54 and cpFtsY GTPases that may reflect their adaptation to the cpSRP pathway? Using a combination of fluorescence and mutational analyses, we dissected the molecular steps during the interaction of cpSRP54 and cpFtsY and probed the role of the GTPase cycle in the targeting of LHCP. The results showed that, despite many similarities with their bacterial homologues, cpSRP54 and cpFtsY undergo a much more streamlined GTPase cycle in which the

complex formation and GTPase activation processes are highly coupled. These differences may have evolved to maximize the efficiency of targeting for the highly abundant LHC proteins.

## Materials and Methods

**Mutagenesis, Protein Expression and Purification.** The bacterial expression plasmid for cpSRP54 was constructed by inserting the coding sequence of mature cpSRP54 from *Arabidopsis thaliana* between the NdeI and HindIII restriction sites in pET41(a) (Novagen). cpSRP54 was overexpressed in *Escherichia coli* Rosetta BL21 cells (Invitrogen) at 37 °C using 0.5 mM IPTG (EMD Biosciences). cpSRP54 was purified by cation exchange chromatography in buffer A (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5% glycerol), first using SP-Sepharose FF beads (GE Healthcare), followed by a MonoS HP column (GE Healthcare), both using a linear gradient of 150–600 mM NaCl.

The construct expressing mature cpFtsY fused to thioredoxin was a generous gift from R. Henry [25]. Thioredoxin-fused cpFtsY was overexpressed in *Escherichia coli* BL21-DE3\* cells (Invitrogen) at 37 °C using 0.5 mM IPTG (EMD Biosciences). cpFtsY was first purified over Talon resin (Clontech) in buffer B (50 mM KHEPES, pH 7.5, 150 mM NaCl, 1 mM PMSF) following manufacturer's instructions. Following thrombin digestion to remove the thioredoxin tag, cpFtsY was further purified by anion exchange chromatography over a MonoQ column (GE Healthcare) in buffer C (50 mM Tris, pH 7.5, 50 mM NaCl, 1 mM EDTA, 2 mM DTT) using a linear gradient of 50–300 mM NaCl, as previously described [26].

cpSRP54 and cpFtsY mutants were constructed using the QuikChange protocol (Stratagene) and were expressed and purified using the same procedures as those for the wild-type cpSRP54, with the following exceptions. Cys-less and single cysteine mutants of cpFtsY were expressed in *Escherichia coli* BL21-DE3\* cells (Invitrogen). Inclusion bodies containing mutant cpFtsY were solubilized using 8 M urea. Solubilized cpFtsY was refolded into the native structure by dialyzing in refolding buffer (100 mM Tris-HCl pH 8.0, 400 mM l-arginine, 5 mM

reduced glutathione, 0.5 mM oxidized glutathione, complete EDTA free protease inhibitor cocktail tablet). The refolded proteins were dialyzed in buffer A and purified by affinity chromatography using Ni-NTA (Qiagen) followed by anion exchange chromatography using MonoQ as described above for wild-type cpFtsY.

**Fluorescence Labeling.** For FRET experiments, single cysteine mutants were labeled with maleimide derivatives of coumarin N-(7-dimethylamino-4-methylcoumarin-3-yl)maleimide (DACM) and BODIPY-fluorescein-N-(2-aminoethyl)maleimide (BODIPY-FL) (Invitrogen). Proteins were dialyzed in labeling buffer (50 mM KHEPES, pH 7.5), 300 mM NaCl, 2 mM EDTA, and 10% glycerol) and treated with 2 mM TCEP at RT to reduce disulfide bonds. The labeling reaction was carried out with a 30-fold excess of dye over protein for 2 h at 4 °C and stopped by addition of 2 mM DTT. Acrylodan labeling was done similarly except that the labeling reaction was carried out for >12 h at 4 °C. The excess dye was removed by gel filtration using Sephadex G25 resin (Sigma-Aldrich). The absorbance of DACM, BODIPY-FL, and acrylodan ( $\epsilon_{363} = 27 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\epsilon_{504} = 79 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ , and  $\epsilon_{391} = 20 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ , respectively) was used to determine the concentration of labeled protein. The labeling efficiency was typically over 80% for all the probes, and the background labeling estimated from cys-less or cys-lite constructs was less than 10%.

**Fluorescence Measurement.** All measurements were carried out at 25 °C in assay buffer (50 mM KHEPES pH 7.5, 150 mM KOAc, 2 mM Mg(OAc)<sub>2</sub>, 0.01% Nikkol, 10% glycerol) on a Fluorolog 3-22 spectrofluorometer (Jobin Yvon). For formation of the GTP-bound cpSRP54-cpFtsY complex, 2 mM GTP (Sigma-Aldrich) was used to ensure that both proteins were predominantly GTP-bound. The amount of GDP generated during the course of the experiment was minimal, as estimated from the GTPase rate constants. For complex formation

with 5'-guanylylimidodiphosphate (GMPPNP), 200  $\mu\text{M}$  GMPPNP (Sigma-Aldrich) was used. For equilibrium or kinetic measurements using FRET, an excitation wavelength of 380 nm was used, and the donor fluorescence emission was monitored at 450 nm. The FRET efficiency was calculated as described [18]. For measurements using acrylodan-labeled cpSRP54, an excitation wavelength of 370 nm and an emission wavelength of 495 nm were used [19].

Equilibrium titrations were carried out using a constant concentration of labeled protein and varying concentrations of the binding partner. The data were fit to eqs 1 or 2

$$F_{\text{obs}} = F_1 \times \frac{[\text{cpSRP54}] + [\text{cpFtsY}] + K_d - \sqrt{([\text{cpSRP54}] + [\text{cpFtsY}] + K_d)^2 - 4 \times [\text{cpSRP54}][\text{cpFtsY}]}}{2 \times [\text{cpSRP54}]} \quad (1)$$

$$F_{\text{obs}} = F_1 \times \frac{[\text{cpSRP54}]}{K_d + [\text{cpSRP54}]} \quad (2)$$

where  $F_{\text{obs}}$  is the observed fluorescence at a particular protein concentration,  $F_1$  is the fluorescence with saturating protein, and  $K_d$  is the equilibrium dissociation constant of the complex.

The association rate constant ( $k_{\text{on}}$ ) for the cpSRP54-cpFtsY complex was measured using the stop-flow apparatus as described in ref [18]. For FRET, 0.5  $\mu\text{M}$  DACM-labeled cpFtsY was mixed with 1–50  $\mu\text{M}$  BODIPY-FL-labeled cpSRP54 in the presence of 2 mM GTP. For measurements based on acrylodan fluorescence, 0.5  $\mu\text{M}$  acrylodan-labeled cpSRP54 was mixed with 1–50  $\mu\text{M}$  wild-type cpFtsY. The observed rate constants ( $k_{\text{obs}}$ ) for each reaction were plotted against cpSRP54 or cpFtsY concentration, respectively, and fitted to a linear (eq 3) or hyperbolic function (eq 4)

$$k_{\text{obsd}} = k_{\text{on}} [\text{protein}] + k_{\text{off}} \quad (3)$$

$$k_{\text{obs}} = k_1 \times \frac{[\text{protein}]}{K_d + [\text{protein}]} \quad (4)$$

in which  $k_{\text{obs}}$  is observed rate of association at a particular protein concentration,  $k_{\text{on}}$  (slope) is the association rate constant and  $k_{\text{off,app}}$  (y-intercept) is the apparent dissociation rate constant, and  $k_1$  and  $K_d$  are defined in Figure 2E.

The dissociation rate constant ( $k_{\text{off}}$ ) was determined by a pulse-chase experiment. 2  $\mu\text{M}$  wild-type cpFtsY was incubated with 0.5  $\mu\text{M}$  acrylodan-labeled cpSRP54(234C, A142W) for 10 min to form the GTP-bound cpSRP54-cpFtsY complex and mixed with 200 mM EDTA or a 20-fold excess of unlabeled cpSRP54 to drive disassociation of the complex. The time course for decrease in acrylodan fluorescence was fit to a single-exponential function to obtain the dissociation rate constant. Both the complex association and dissociation kinetics were measured on a Kintek stopped-flow apparatus.

**GTPase Assays.** All GTPase assays were performed at 25 °C in assay buffer as described previously [26]. GTP hydrolysis reactions were followed and analyzed as described in ref [27]. The reciprocally stimulated GTPase reaction between cpFtsY and cpSRP was measured in multiple-turnover experiments ( $[\text{GTP}] > [\text{E}]$ ) with a small fixed amount of cpSRP54 (100 nM), varying concentrations of wild-type or mutant cpFtsY, and 100  $\mu\text{M}$  GTP. The cpFtsY concentration dependence of the observed rate ( $k_{\text{obs}}$ ) was fit to eq 5

$$k_{\text{obs}} = k_{\text{cat}} \times \frac{[\text{cpFtsY}]}{K_m + [\text{cpFtsY}]} \quad (5)$$

in which  $k_{\text{cat}}$  is the maximal rate constant with saturating cpFtsY and  $K_m$  is the concentration of cpFtsY required to reach half saturation. Analogous setups were used when cpSRP54 mutants were tested, with the concentration of cpSRP54 being varied instead that of cpFtsY.

The affinity of mutant cpFtsY for cpSRP54 was determined using an inhibition assay that measures the ability of mutant cpFtsY to compete with wild-type cpFtsY and inhibits its interaction with cpSRP54, as described in Shan *et al* [17]. The data were fit to eq 6

$$k_{\text{obs}} = k_0 \times \frac{K_i}{[\text{cpFtsY(mt)}] + K_i} + k_1 \times \frac{[\text{cpFtsY(mt)}]}{[\text{cpFtsY(mt)}] + K_i} \quad (6)$$

in which  $K_i$  is the inhibition constant,  $k_0$  is the rate constant of GTP hydrolysis in the absence of the inhibitor, and  $k_1$  is the rate constant of GTP hydrolysis from the cpSRP54-cpFtsY(mt) complex. At subsaturating concentrations of the wild-type cpFtsY ( $<K_m$ ), the value of  $K_i$  equals  $K_d$ , the dissociation constant of the cpSRP54-cpFtsY(mt) complex. Analogous setups and analyses were used when cpSRP54 mutants were tested.

**Gel Filtration.** Complex formation was carried out in column buffer [50 mM KHEPES, pH 7.5, 200 mM NaCl, 2 mM Mg(OAc)<sub>2</sub>, 2 mM DTT]. 10  $\mu$ M of cpSRP54 was mixed with 10  $\mu$ M wild-type or mutant cpFtsY in the presence of 450  $\mu$ M GMPPNP and incubated on ice for 10 min before loading onto Superdex 200 (GE Healthcare). Reference runs of the individual proteins confirmed the identities of the peaks.

**LHCP Integration Assay.** The thylakoids were collected from chloroplasts of 9–12 day old pea seedlings (Laxton Progressive 9) hypotonically lysed in lysis buffer (10 mM KHEPES, pH 8.0, 10 mM MgCl<sub>2</sub>) for 10 min as described by Yuan *et al.* [28]. The stromal extract was removed, and the thylakoid pellet was resuspended in lysis buffer and washed twice in import buffer (50 mM KHEPES, pH 8.0, 330 mM sorbitol) containing 1 M KOAc to remove residual cpFtsY associated with the membrane. Thylakoids were resuspended in import buffer to a concentration of 1 mg chlorophyll/mL (1 $\times$ ). Each 150  $\mu$ L light-harvesting chlorophyll *a/b* binding protein (LHCP) targeting/integration reaction contained 10  $\mu$ L of *in vitro* translated <sup>35</sup>S-methionine-labeled LHCP, 50  $\mu$ L of 1 $\times$  salt-washed thylakoid, 50 mM GTP, 50 mM ATP, 0.5  $\mu$ M cpSRP54, and varying concentrations of cpFtsY. Analogous setups were used when cpSRP54 mutants were tested. The reactions were incubated at 25 °C for 10 min before being quenched on ice. The reaction mixtures were thermolysin-treated for 40 min and centrifuged to

remove any nonintegrated LHCP in the supernatant. The resulting pellets were resuspended in 2× SDS and analyzed by SDS-PAGE. The two lower bands that represent the protease-protected fragments of the integrated LHCP (25 and 18.5 kDa) were quantified using a Molecular Dynamics Storm 840 and ImageQuant software (GE Healthcare).

## Results

**Fluorescence Assays to Monitor the cpSRP54–cpFtsY Interaction.** To directly visualize the interaction between cpSRP54 and cpFtsY in real time, we developed fluorescence-based assays, which have been used in the bacterial SRP and other systems to elucidate key features of protein interaction mechanisms. To this end, we constructed cys-lite and cys-less versions of cpSRP54 and cpFtsY, respectively. cpSRP54 has a solvent exposed cysteine 198 which can be mutated to serine to obtain cys-lite cpSRP54 without disrupting its interaction with cpFtsY (Figure 1.S1A); the remaining two cysteines in cpSRP54 are likely buried inside the folded protein, based on homology modeling with Ffh, and did not react significantly with fluorescent dyes in control experiments (Supplementary Figure 1.S1B). cpFtsY contains five native cysteines, all of which were replaced with serines. Cys-less cpFtsY was purified from inclusion bodies and refolded into the native structure. Refolded cys-less cpFtsY interacted with and stimulated cpSRP54's GTPase activity with efficiencies within two-fold of that of wild-type cpFtsY (Supplementary Figure 1.S1C).

As the crystal structure of cpSRP54 or its complex with cpFtsY is not available, we constructed a homology model of the cpSRP54-cpFtsY complex based on superposition of the crystal structure of apo-cpFtsY onto that of *T. aquaticus* FtsY in complex with Ffh (Figure 1.1A). On the basis of this model, single cysteines were introduced at solvent exposed positions and labeled with fluorescent dyes using thio-specific chemistry. In FRET experiments, a cysteine was engineered at residue 321 of cys-less cpFtsY and labeled with DACM as the donor fluorophore, and a cysteine was introduced at residue 234 of cys-lite cpSRP54 and labeled with BODIPY-FL as the acceptor dye (Figure 1.1A). Both probes are located at the N–G domain interface of the respective GTPases and are 30 Å apart as estimated from the homology model.

Significant FRET was observed upon assembly of the cpSRP54·cpFtsY complex in the presence of GTP (Figure 1.1B). At saturating protein concentrations and when complications from GTP hydrolysis were minimized (see below), the FRET efficiency in the cpSRP54·cpFtsY complex was 0.60 (Figure 1.2F). In addition, the cysteine at residue 234 of cys-lite cpSRP54 was labeled with an environmentally sensitive dye, acrylodan. Formation of the cpSRP54·cpFtsY complex with GTP induced a blue shift and a 30% increase in the fluorescence intensity of this dye (Figure 1.1C), providing an additional measurement of the cpSRP54–cpFtsY interaction. Fluorescently labeled cpSRP54 and cpFtsY interacted with and activated each other's GTPase activity with rate constants within 2-fold of the wild-type proteins (Supplementary Figure 1.S1D and E). Further, both the FRET and fluorescence change of cpSRP54(234C)-acrylodan upon complex formation could be competed away by EDTA or unlabeled protein (Supplementary Figures 1.S2A and B). Thus, these fluorescence assays faithfully report on the kinetics and stability of the cpSRP54–cpFtsY interaction.

**Two-Step Complex Assembly.** Using the fluorescence assays, we characterized the kinetics and stability of the interaction between cpSRP54 and cpFtsY. These analyses, however, were complicated by the hydrolysis of GTP, which occurs quickly in the cpSRP54·cpFtsY complex and drives rapid disassembly of the GTPase complex. In the bacterial SRP and FtsY GTPases, this problem can be overcome by using the nonhydrolyzable GTP analogue GMPPNP, which provides a good mimic for GTP to support efficient assembly of a stable SRP·FtsY complex [18]. However as shown below (Figure 1.2F), GMPPNP does not provide an adequate mimic of GTP to support stable complex assembly between cpSRP54 and cpFtsY. To overcome this problem, we used the mutant GTPases, cpSRP54(A142W) or cpFtsY(A168W). The corresponding mutations in bacterial SRP and FtsY, Ffh(A144W) and FtsY(A335W),

respectively, specifically disrupted GTPase activation in the Ffh-FtsY complex without affecting rapid and stable complex assembly [17]. Similarly, both mutants in cpSRP54 and cpFtsY allowed a stable cpSRP54-cpFtsY complex to be efficiently assembled (Figures 1.3B and 1.4B), but specifically blocked GTP hydrolysis in the complex, and thus provided a reasonable estimate for the kinetic and thermodynamic stabilities of the wild-type cpSRP54-cpFtsY complex.

We determined the kinetics of complex assembly in the presence of GTP by following either the gain of FRET (Figure 1.2A) or the increase in fluorescence of acrylodan-labeled cpSRP54 (Figure 1.2B). Surprisingly, the complex assembly rate constant ( $k_{on}$ ) measured using FRET was over three-fold faster than that determined using cpSRP54(234C)-acrylodan (Figure 1.2C). This difference was not caused by a larger deleterious effect of acrylodan labeling at cpSRP54(234C) on complex assembly, as cpSRP54(234C)-acrylodan exhibited comparable activity in the stimulated GTPase reaction to the cpSRP54 and cpFtsY labeled with the FRET dyes (compare Supplementary Figures 1.S1D and 1.S1E, ●). Instead, we reasoned that the difference in the observed complex assembly rates arises from the fact that the acrylodan probe reports on a local conformational change surrounding residue 234 that accompanies complex assembly, whereas FRET directly reports on approximation of distance between cpSRP54 and cpFtsY as soon as a complex is formed. This raised the possibility that assembly of the stable cpSRP54-cpFtsY complex occurs in two steps, with the initial formation of an intermediate detected by FRET followed by conformational rearrangement to form a more stable, final complex detected specifically by cpSRP54(234C)-acrylodan.

To provide additional evidence for this model, we analyzed the concentration dependence of the observed complex assembly rates using cpSRP54(234C)-acrylodan. If formation of a stable complex occurred in a single bimolecular association, then the observed complex

assembly rate constants should increase linearly with increasing protein concentration. In contrast, if additional steps were required for stable complex assembly, deviations from linearity would be expected. Indeed, the observed complex assembly rate constant exhibited a hyperbolic dependence on cpFtsY concentration and plateaued at  $6 \text{ s}^{-1}$  with saturating cpFtsY (Figure 1.2D). Control experiments showed that this plateau was unlikely to be caused by protein aggregation or inactivation at high concentrations (Supplementary Figure 1.S1E). These results are consistent with the formation of a transient intermediate with a  $K_d$  value of  $30 \text{ }\mu\text{M}$  during complex assembly (Figure 1.2E), such that complex formation is rate-limited by the bimolecular cpSRP54–cpFtsY association at low protein concentrations but becomes rate-limited by a unimolecular rearrangement from this intermediate at saturating protein concentrations. Together, these results strongly suggest that assembly of the cpSRP54·cpFtsY complex requires at least two steps.

We further determined the kinetic and thermodynamic stabilities of the cpSRP54·cpFtsY complex. The affinity of the cpSRP54·cpFtsY complex was measured by equilibrium titrations using mutant cpSRP54(A142W) or cpFtsY(A168W), as rapid GTP hydrolysis from the wild-type complex will artificially raise the observed equilibrium dissociation constant ( $K_d$ ) (Figure 1.2F vs Supplementary Figure 1.S3A, ●). These analyses yielded a  $K_d$  value of 300–500 nM using both the FRET assay and acrylodan-labeled cpSRP54 (Figure 1.2F and Supplementary Figure 1.S3B and C). In addition, pulse-chase experiments gave a dissociation rate constant of  $0.03 \text{ s}^{-1}$  for the cpSRP54(A142W)·cpFtsY complex (Figure 1.2G). In conjunction with the association rate constant measured above (Figure 1.2C), this yielded a  $K_d$  value of 200 nM for this complex, consistent with the value determined from equilibrium titrations.

**IBD Loops Play Essential Roles in Both Complex Assembly and GTPase Activation.**

To probe the molecular determinants essential for the interaction between cpSRP54 and cpFtsY, we generated a collection of site-directed mutant GTPases that map to the putative interaction surface of cpSRP54 and cpFtsY based on structural homology, with an emphasis on the universally conserved IBD loops (Figure 1.1A, magenta, and Supplementary Tables 1.S1 and 1.S2). Control experiments showed that the basal GTP binding and hydrolysis activity (Supplementary Tables 1.S3 and 1.S4, respectively) of the individual cpFtsY and cpSRP54 mutants were comparable to that of the wild-type proteins, ensuring that defects did not arise from disruption of the global structure of the mutant proteins. We then screened the mutants by monitoring the reciprocally stimulated GTPase reaction between cpSRP54 and cpFtsY (Figures 1.3A,B and 1.4A,B). As demonstrated above, the complex assembly rate constants measured directly using the fluorescence assays agreed well with the value of  $k_{\text{cat}}/K_m$  ( $3 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ ) in the stimulated GTPase reaction; further, dissociation of the  $^{\text{GTP}}\text{cpSRP54}\text{-cpFtsY}^{\text{GTP}}$  complex ( $0.030 \text{ s}^{-1}$ ) is at least 20-fold slower than GTP hydrolysis from this complex ( $0.7 \text{ s}^{-1}$ ). Both observations indicate that in the stimulated GTPase reaction the value of  $k_{\text{cat}}/K_m$  is rate-limited by, and hence reports on, the rate of assembly of a stable cpSRP54–cpFtsY complex, whereas the maximal rate constant  $k_{\text{cat}}$  reports on either GTP hydrolysis from the complex or a rate-limiting rearrangement that activates the chemical step.

The vast majority of mutants exhibited defects in this reaction (Table 1.1 and Supplementary Tables 1.S1 and 1.S2). Among them, perturbations of the IBD loops produced the most deleterious effect on the reciprocally stimulated GTPase reaction between cpSRP54 and cpFtsY (Table 1.1), consistent with their high evolutionary conservation. Inspection of the concentration dependence of the stimulated GTPase reactions further suggested that the majority

of these mutants have defects in both the complex assembly and GTP hydrolysis steps. For example, the cpFtsY(A169W) and the corresponding cpSRP54(A143W) mutations not only reduced the GTPase rate from the complex by over 50-fold (Figures 1.3A and 1.4A and Table 1.1,  $k_{\text{cat}}$ ), but a significantly higher concentration of mutant proteins were required to reach saturation (Figures 1.3A and 1.4A, insets and Table 1.1,  $K_m$ ). Only two mutations, cpFtsY(A168W) and cpSRP54(A142W), were exceptions: both mutants reduced the maximal rate of GTP hydrolysis by 15–50-fold (Figures 1.3B and 1.4B and Table 1.1,  $k_{\text{cat}}$ ), but saturation in GTPase rate could be reached at low protein concentrations, suggesting that efficient complex assembly could occur in these mutants (Figures 1.3B and 1.4B, insets, and Table 1.1,  $K_m$ ).

To further dissect the contribution of each residue to complex assembly and/or GTPase activation, we used a well-established inhibition assay (Figures 1.3C and 1.4C) [17]. For example, if a mutant cpFtsY could bind cpSRP54 but failed to efficiently hydrolyze GTP, then it would compete with wild-type cpFtsY in binding and inhibit its stimulated GTPase reaction with cpSRP54. Under subsaturating concentrations of the wild-type cpSRP54 and cpFtsY, the inhibition constant  $K_i$  obtained from this assay represents the dissociation constant between the mutant and its partner GTPase. In agreement with their kinetic parameters from the stimulated GTPase reactions, cpFtsY(A168W) and cpSRP54(A142W) exhibited strong competitive binding to their respective partner GTPases, with inhibition constants below 0.2  $\mu\text{M}$  (Figures 1.3C and 1.4C, open circles, and Table 1.1,  $K_i$ ). In contrast, all the other deleterious mutations in the IBD loop severely impaired complex formation (Table 1.1). For example, cpFtsY(A169W) and cpSRP54(A143W) could not act as competitive inhibitors in the inhibition assay and exhibited  $K_i$  values over 50  $\mu\text{M}$  (Figures 1.3C and 1.4C, ●).

To independently corroborate the results from the inhibition assay, we used gel filtration and fluorescence analyses to independently evaluate the mutational effects on the stability and/or conformational changes of the complex. In gel filtration analyses, which qualitatively assess the ability of the mutant proteins to form a kinetically stable complex [29,30], cpSRP54(A142W) and cpFtsY(A168W) assembled complexes with efficiencies within two-fold of the wild-type proteins (Figures 1.3D and 1.4E, red vs black). On the other hand, all the other deleterious mutations in the IBD loops (residues D163, R166, A169 of cpFtsY and D137, R140, A143 of cpSRP54) showed no or little detectable complex formation (Figures 1.3D and 1.4E). Similarly, fluorescence assays showed that mutant cpSRP54(A142W) exhibited the same complex assembly rate constant as wild-type cpSRP54 (Figure 1.4D), and both cpSRP54(A142W) and cpFtsY(A168W) assembled stable complexes with their binding partners (Supplementary Figure S3B and C). In contrast, complex formation could not be detected for mutants cpSRP54(D137A) and cpSRP54(A169W) using the fluorescence assay (data not shown). Together, these results strongly suggest that the IBD loops, which provide key catalytic motifs for GTPase activation, are also intimately involved in the assembly of the cpSRP54–cpFtsY complex.

Two additional lines of evidence support this notion and showed that in the cpSRP54–cpFtsY complex interactions at the catalytic active site are tightly coupled to assembly of the GTPase complex. First, several mutations in the IBD loop of cpSRP54 caused extensive blue shift and increase in fluorescence intensity of the acrylodan labeled at cpSRP54(234C) compared to that of wild-type cpSRP54 (Figure 1.4F). This indicates that perturbation of the IBD loop effects a change in the local environment at the NG-domain interface of cpSRP54, a region critical for efficient complex assembly (Supplementary Table S1 and ref [30]). Second, replacement of the  $\beta,\gamma$ -bridging oxygen of GTP strongly reduced both the rate (data not shown)

and the stability of the complex (Figure 1.2F), and GMPPNP could not induce the change in the fluorescence of acrylodan-labeled cpSRP54(234C) upon complex formation (data not shown). Thus, interactions at sites critical for the chemical reaction are also integrally involved in the complex assembly process.

**Defects in Complex Formation and GTPase Activation Block LHCP Targeting.** To assess the contribution of the individual steps in cpSRP54 and cpFtsY's GTPase cycle to the protein targeting reaction, we tested the effects of the mutant cpFtsY and cpSRP54 GTPases on the targeting and translocation of LHCP. The overall efficiency of LHCP targeting and integration was analyzed based on protease protection of LHCP upon its proper integration into salt-washed thylakoid membranes (see Materials and Methods). *In vitro* translocation reached completion after 10 min and the reaction saturated at cpFtsY concentrations above 150 nM (Supplementary Figures S5A and S5B, respectively). On the basis of these observations, a concentration of 500 nM and a time point of 10 min were used to test the effect of mutant proteins on the efficiency of the targeting reaction.

In general, a significant defect in LHCP integration was observed only with a >10-fold reduction in the individual steps of cpSRP54 and cpFtsY's GTPase cycle (complex formation or GTPase activation). This is analogous to observations in the cotranslational protein targeting reaction carried out by bacterial SRP and FtsY [31] and suggests that the targeting of LHCP by cpSRP and cpFtsY is not the major rate-limiting step in the translocation/integration assay. Nevertheless, this assay revealed moderate to strong defects in LHCP integration for most of the mutant GTPases (Figures 1.5A,B and Supplementary Figure S5C). The two mutants that specifically block GTPase activation, cpSRP54(A142W) and cpFtsY(A168W), reduced translocation efficiency 2-fold (Figure 1.5A,B), suggesting that activated GTP hydrolysis in the

cpSRP54-cpFtsY complex is not crucial but does modulate the efficiency of the targeting reaction. In comparison, mutant GTPases that also impair complex assembly, such as cpFtsY(A169W) and cpSRP54(A143W), showed stronger defects in LHCP targeting and translocation (Figure 1.5A,B). The reduction in translocation efficiency of the various GTPase mutants correlated with their values of  $k_{cat}/K_m$  in the GTPase assay, an indicator for the efficiency of complex assembly (Figure 1.5C,D). Collectively, these results demonstrate that efficient assembly of the cpSRP54–cpFtsY complex is crucial for the targeting and integration of LHCP, whereas GTPase activation and/or GTP hydrolysis plays a modulatory role to help enhance the efficiency of targeting.

## Discussion

The interaction between the SRP and SR GTPases delivers cargo proteins to a target membrane and hence plays a crucial role in the proper localization of membrane proteins. During the interaction of the bacterial SRP and SR, formation of a stable complex is a two-step process that requires initial formation of a transient “early” intermediate, followed by a slow rearrangement of this intermediate to a stable complex (Figure 1.6, black line, steps 1 and 2) [18]. Here, real-time fluorescence analyses strongly suggest that a two-step assembly process also occurs during the interaction between cpSRP54 and cpFtsY. First, the complex assembly rate constant measured by acrylodan, which detects a local rearrangement at the NG-domain interface accompanying complex formation, is significantly slower than that reported by the FRET probes, which are less sensitive to the conformational state of the complex. Further, the observed complex assembly rate constant exhibits a hyperbolic, instead of linear, dependence on protein concentration. Both observations are indicative of the presence of an intermediate during complex assembly (Figures 1.2F and 1.6, red lines). Compared to the bacterial SRP and FtsY, the assembly intermediate formed by the chloroplast GTPases is less stable ( $K_d$  30  $\mu\text{M}$  compared to 4–8  $\mu\text{M}$  in the bacterial complex [18]) but rearranges to the stable complex much faster ( $k_1$  6  $\text{s}^{-1}$  compared to 0.6–1  $\text{s}^{-1}$  for the bacterial complex [18]) (Figure 1.6, red vs black lines, step 2). These observations suggest that the transient intermediate assembled by the chloroplast GTPases is more productive and possibly requires less extensive rearrangements to attain the final complex than their bacterial homologues. This is consistent with our previous observation that cpFtsY is preorganized into a conformation more conducive to stable complex assembly than the bacterial FtsY [26,32].

An important feature of the bacterial SRP system is that movement of the IBD loops, which activates GTP hydrolysis, can be conceptually and experimentally uncoupled from the rearrangements, in the rest of the protein, that mediate stable complex assembly [17,22]. Numerous mutations in the IBD loops result in specific inhibition of GTPase activation, without significantly disrupting formation of the complex [17]. Given these observations, it was surprising to find that the vast majority of mutations in the IBD loops of cpSRP54 and cpFtsY severely compromised assembly of the GTPase complex. This raises the possibility that in the cpSRP54·cpFtsY complex these two steps are more tightly coupled, as the catalytic active sites are also intimately involved in the assembly of the complex. Supporting this notion is the observation that conservative perturbations at the site of chemical transformation, such as replacement of the  $\beta$ ,  $\gamma$ -bridging oxygen of GTP with  $-\text{NH}-$ , severely disrupted complex stability and assembly rate, in contrast to the bacterial SRP with which GMPPNP specifically inhibits GTP hydrolysis [31]. Further, mutations of the IBD loops induced large changes in the fluorescence of an acrylodan dye over 30 Å away at the NG-domain interface, suggesting that the GTPase active sites are intimately linked to sites crucial for complex assembly. Taken together, these observations suggest that during the interaction between cpSRP54 and cpFtsY the complex formation and GTPase activation steps are highly coupled, in contrast to the cytosolic SRP·FtsY complex in which these processes occur in two distinct molecular steps (Figure 1.6, step 3, black vs red lines).

What gives rise to this difference? To address this question, one might begin by reflecting on the role of the multiple conformational steps during the assembly and activation of the cytosolic SRP·FtsY complex. Uncoupling complex formation and GTPase activation allows each of these steps to provide an independent fidelity checkpoint, thus providing the SRP multiple

opportunities to reject the incorrect cargos [19]. This is crucial for the bacterial SRP to achieve a high fidelity of substrate selection, as it has to recognize highly divergent signal sequences and to distinguish between the correct and incorrect cargos based on subtle variations [20]. We speculate that the absence of a similar challenge in attaining specific substrate selection may underlie the different behavior of the chloroplast GTPases. In contrast to the cytosolic SRP, the cpSRP is primarily dedicated to a highly conserved LHC family of proteins, and the cpSRP43 subunit can provide highly specific recognition of these substrates [14,15,23]. Although cpSRP54 also participates in the cotranslational targeting of several membrane proteins (such as D1 protein) [33], the number and diversity of these substrates are much more limited than those handled by the cytosolic SRP. It could therefore be envisioned that the chloroplast SRP system can afford to forego the GTPase activation step as an additional regulatory point.

Consistent with this notion, GTPase activation plays a less essential role in protein targeting by the cpSRP than the cytosolic SRP pathway. In the cytosolic SRP pathway, mutant GTPases that specifically block the activation of GTP hydrolysis severely inhibit protein targeting at late stages [31]. Thus, the molecular rearrangements that lead to GTPase activation, notably the movement of the IBD loops, play an essential role in the unloading of cargo from the SRP and the initiation of protein translocation. In contrast, mutations that specifically inhibit GTPase activation in the cpSRP54-cpFtsY complex resulted in only a 2-fold reduction in the targeting of LHCP. Although in previous reports, the observation that GMPPNP inhibited LHCP targeting has implicated a crucial role of GTP hydrolysis for LHCP targeting and integration [24,25], our findings here suggest that these defects could instead arise from the failure of GMPPNP to support efficient and stable cpSRP54-cpFtsY complex assembly. Indeed, mutant GTPases that impair complex assembly between cpSRP54 and cpFtsY led to much larger

deleterious effects on LHCP targeting, and the defects in their targeting efficiency correlated with their defects in complex assembly. Thus, GTPase activation or GTP hydrolysis plays a modulatory role in enhancing the targeting efficiency of LHCP but is not as crucial as is the case with the cytosolic SRP. The ability of cpSRP43 to directly interact with the Alb3 translocase and thus regulate substrate binding and release [24,34] might allow the cpSRP pathway to bypass the use of GTPase activation as a critical mechanism to drive the unloading of cargo from the cpSRP; this possibility remains to be tested.

Collectively, these results suggest a more streamlined cpSRP54–cpFtsY interaction cycle compared to their bacterial homologues (Figure 1.6), which might be a consequence of their adaptation to targeting a different set of substrate proteins. This pair of GTPases is primed to efficiently form a complex and to quickly turn over the complex (through rapid GTP hydrolysis), bypassing conformational steps that serve as important fidelity checkpoints in the bacterial SRP pathway. These features could allow the cpSRP pathway to cater to the LHC family of proteins, whose sequence conservation allows specific substrate selection to be more easily achieved, but whose high abundance demands a highly efficient targeting pathway with rapid turnover. In this light, one might further speculate that the complex series of dynamic conformational changes observed for the bacterial SRP·FtsY GTPase complex could be fine-tuned to allow efficient targeting only in response to the correct signal sequences while minimizing the targeting of empty ribosomes and incorrect cargo proteins. The divergent properties of the bacterial and chloroplast SRP and FtsY GTPases might reflect different mechanisms to achieve the balance between efficiency and selectivity as the two pathways adapt to distinct challenges posed by their substrate proteins.

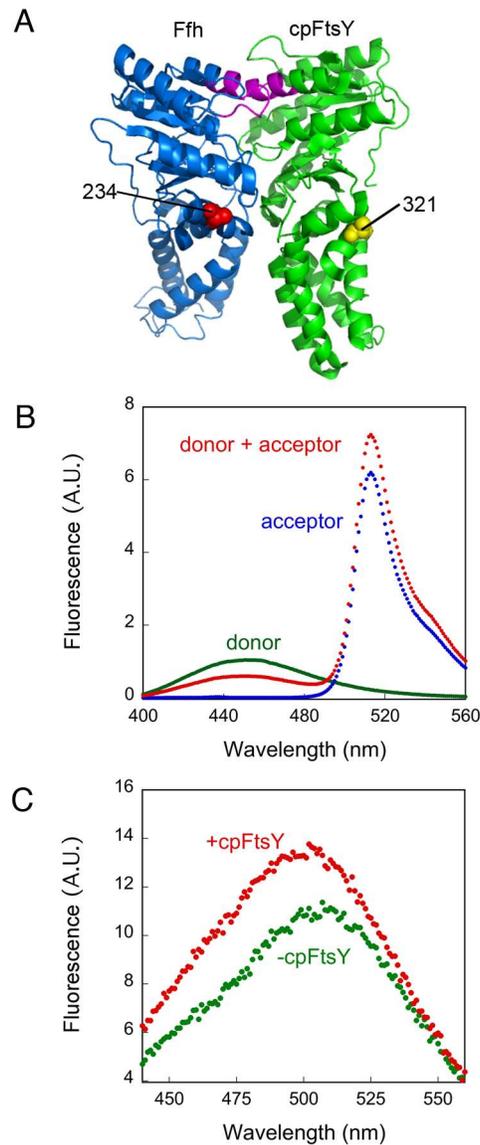
**Acknowledgment**

We thank X. Zhang, N. Pierce, and members of the Shan lab for helpful comments on the manuscript.

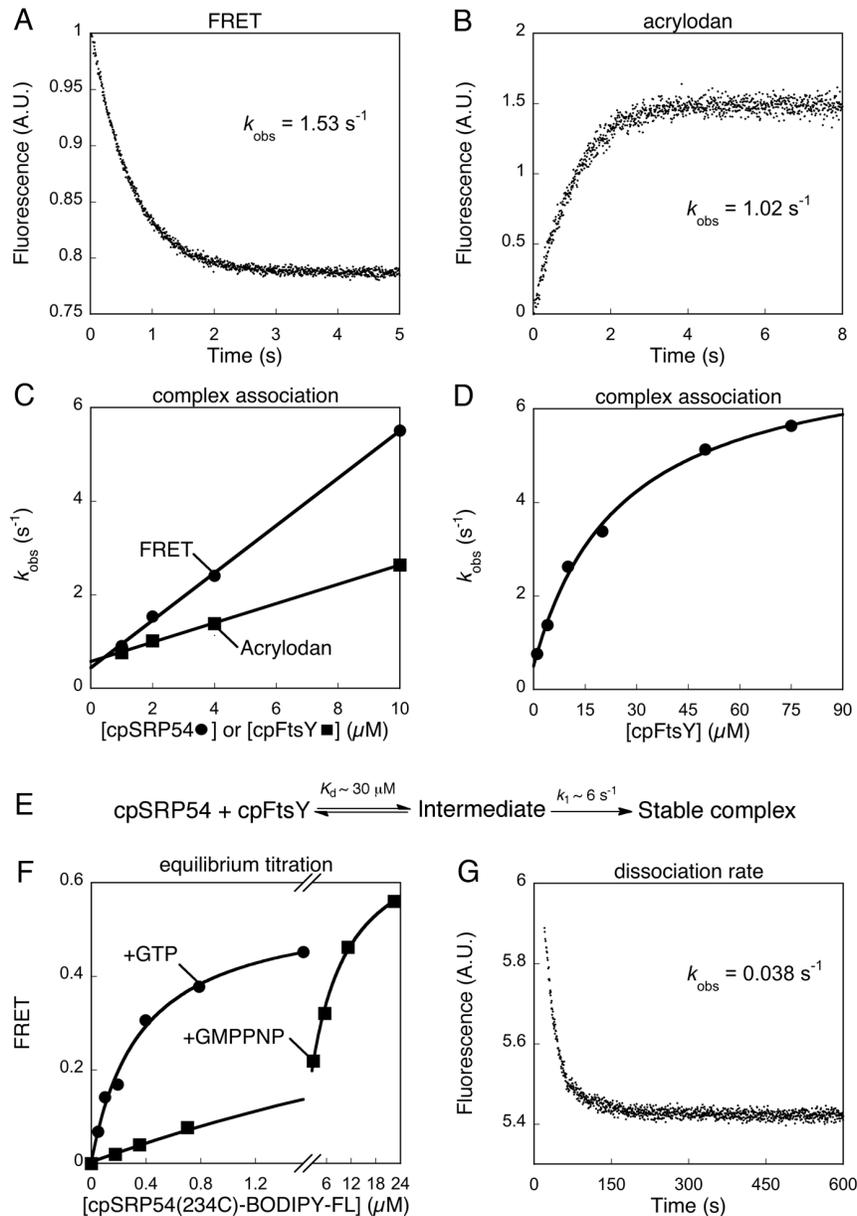
**Table 1.1** Summary of equilibrium and kinetic properties of cpSRP54 and cpFtsY IBD-loop mutants

mutations	$k_{\text{cat}}$ ( $\text{min}^{-1}$ )		$K_{\text{m}}$ ( $\mu\text{M}$ )		$k_{\text{cat}}/K_{\text{m}}$ ( $\mu\text{M}^{-1}\text{min}^{-1}$ )		$K_{\text{i}}$ ( $\mu\text{M}$ )	$K_{\text{d}}$ ( $\mu\text{M}$ )
WT	$27 \pm 3$	(1) <sup>a</sup>	$1.6 \pm 0.3$	(1)	$18 \pm 3$	(1)		$0.80^b$
A142W cpSRP54	$0.50 \pm 0.34$	(0.02)	$0.24 \pm 0.17$	(0.2)	$2.8 \pm 1.7$	(0.2)	$0.11 \pm 0.09$	$0.44^c, 0.16^d$
A168W cpFtsY	$1.8 \pm 0.1$	(0.07)	$0.27 \pm 0.02$	(0.2)	$6.8 \pm 0.6$	(0.4)	$0.23 \pm 0.15$	$0.52^3$
D137A cpSRP54	$1.0 \pm 0.5$	(0.04)	$5.1 \pm 0.1$	(3)	$0.19 \pm 0.09$	(0.01)	$2.0 \pm 1.3$	
D163A cpFtsY	$1.5 \pm 0.4$	(0.05)	$5.0 \pm 0.1$	(3)	$0.34 \pm 0.06$	(0.02)	ND	
R140A cpSRP54	$3.4 \pm 0.2$	(0.1)	$6.3 \pm 1.2$	(4)	$0.54 \pm 0.08$	(0.03)	$0.51 \pm 0.20$	
R166A cpFtsY	$4.3 \pm 0.8$	(0.2)	$5.1 \pm 1.7$	(3)	$0.89 \pm 0.18$	(0.05)	$3.2 \pm 1.4$	
A143L cpSRP54	$0.23 \pm 0.05$	(0.008)	$20 \pm 1$	(13)	$0.012 \pm 0.003$	(0.001)	> 40	
A143W cpSRP54	$0.52 \pm 0.07$	(0.02)	$15 \pm 7$	(10)	$0.037 \pm 0.012$	(0.002)	> 40	
A169L cpFtsY	$0.32 \pm 0.10$	(0.01)	$15 \pm 2$	(10)	$0.031 \pm 0.019$	(0.002)	> 40	
A169W cpFtsY	$0.36 \pm 0.15$	(0.01)	$13 \pm 5$	(8)	$0.028 \pm 0.001$	(0.002)	> 40	
F165A cpFtsY	$28 \pm 3$	(1)	$0.8 \pm 0.3$	(0.5)	$38 \pm 10$	(2)	ND	
A167W cpFtsY	11	(0.4)	7.4	(5)	1.5	(0.1)	ND	

<sup>a</sup>Values in parenthesis are relative to that of the wild-type proteins, which is normalized to 1. ND, not determined. <sup>b</sup>Apparent  $K_{\text{d}}$  values were obtained from equilibrium titrations using FRET in Figure 2F and Supplementary Figure S3. <sup>c</sup>Apparent  $K_{\text{d}}$  values were obtained from equilibrium titrations using acrylodan-labeled cpSRP54 in Supplementary Figures S3. <sup>d</sup> $K_{\text{d}}$  obtained from  $k_{\text{off}}/k_{\text{on}}$ .

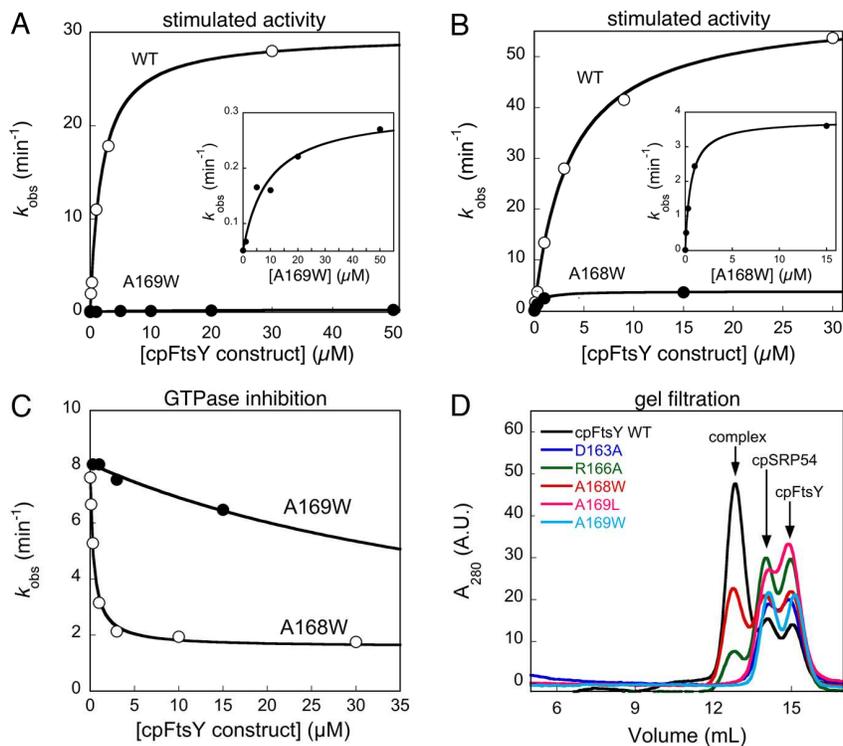


**Figure 1.1** Fluorescence assays to report on complex assembly between cpSRP54 and cpFtsY. (A) The positions of the FRET donor (yellow sphere) and acceptor (red sphere) probes in cpFtsY and cpSRP54, respectively, mapped onto a homology model of the complex generated by superimposing the crystal structure of cpFtsY (2OG2) onto that of the *T. aquaticus* Ffh·FtsY NG domain complex (1RJ9). The IBD loops in cpSRP54 and cpFtsY are highlighted in magenta. (B) Fluorescence emission spectra of donor-labeled cpFtsY (0.5  $\mu$ M, green), acceptor-labeled cpSRP54 (2  $\mu$ M, blue), and their complex formed with 2 mM GTP (red). (C) Fluorescence emission spectra of acrylodan-labeled cpSRP54(234C) in the absence (green) and presence (red) of cpFtsY (2  $\mu$ M).

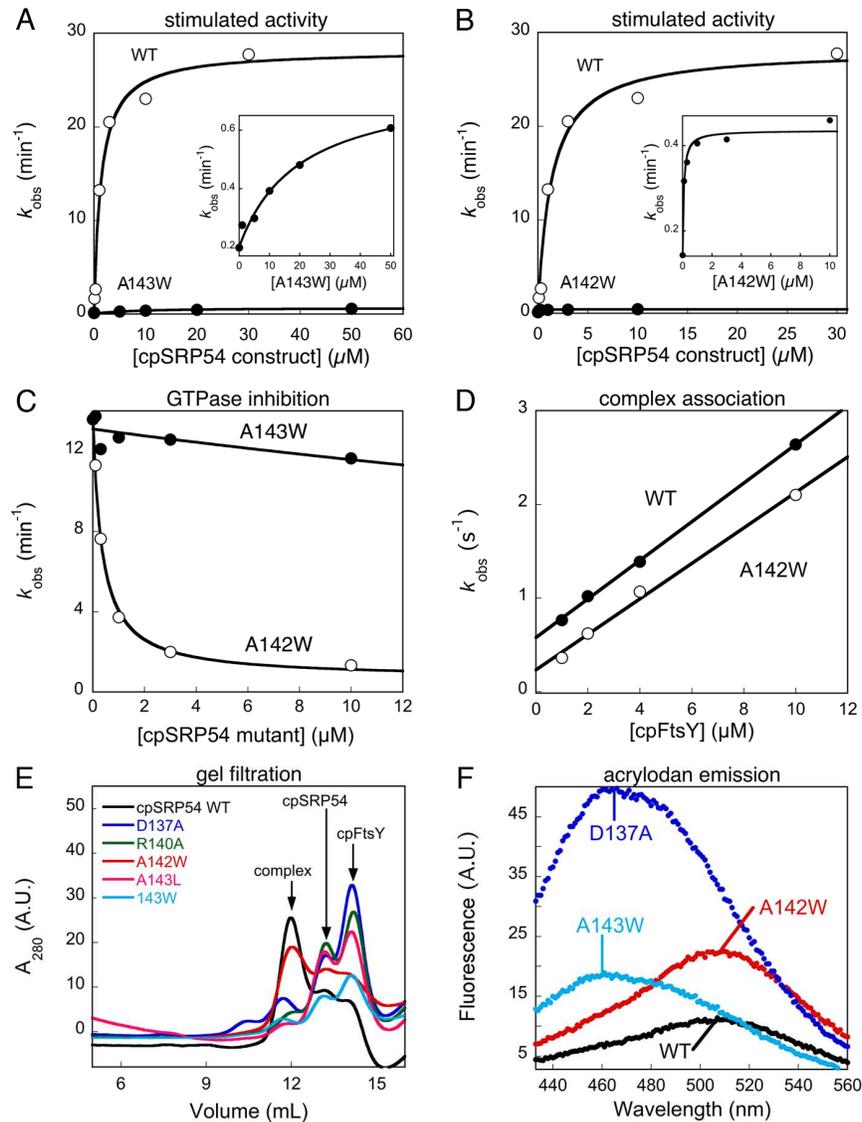


**Figure 1.2** Thermodynamic and kinetics for formation of the cpSRP54-cpFtsY complex. (A) Complex assembly between  $0.5 \mu\text{M}$  cpFtsY(321C)-DACM and  $2 \mu\text{M}$  cpSRP54(234C)-BODIPY-FL, measured in a stopped-flow apparatus as described in Materials and Methods. Single-exponential fit of the data gave a  $k_{\text{obs}}$  value of  $1.53 \text{ s}^{-1}$ . (B) Complex assembly between  $0.5 \mu\text{M}$  cpSRP54(234C)-acrylodan and  $2 \mu\text{M}$  cpFtsY, measured in a stopped-flow apparatus as described in Materials and Methods. Single-exponential fit of the data gave a  $k_{\text{obs}}$  value of  $1.02 \text{ s}^{-1}$ . (C) Association rate constants for cpSRP54–cpFtsY complex formation with GTP measured by FRET (●) and acrylodan fluorescence (■). Linear fits of the data gave complex assembly rate constants ( $k_{\text{on}}$ ) of  $5 \times 10^5$  and  $1.57 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$  with FRET and acrylodan fluorescence, respectively. (D) A hyperbolic dependence of complex assembly rate constants on cpFtsY concentration. The data were fit to eq 4 in the Materials and Methods, which gave a  $K_d$  value of  $30 \mu\text{M}$  and a rate constant of  $6 \text{ s}^{-1}$  at saturating cpFtsY. (E) A two-step schematic of cpSRP54–

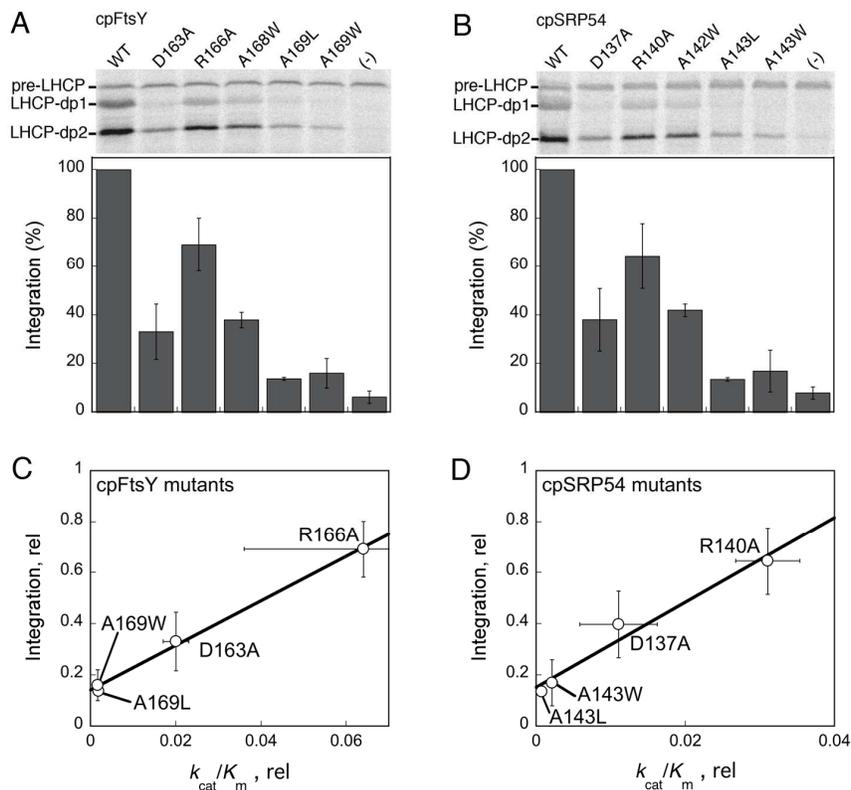
cpFtsY complex assembly. (F) Equilibrium titration of the cpSRP54·cpFtsY complex formed with GTP (●) or GMPPNP (■) measured by FRET. Complex formation with GTP was carried out using mutant cpFtsY(A168W) to minimize GTP hydrolysis. The data were fit to eq 2, which gave  $K_d$  values of 0.35  $\mu\text{M}$  with GTP and 7  $\mu\text{M}$  with GMPPNP. (G) Dissociation kinetics of the cpSRP54(234C, A142W)·cpFtsY complex, measured as described in the Materials and Methods. Single-exponential fit of the data gave an apparent dissociation rate constant of 0.038  $\text{s}^{-1}$ . After subtracting the GTP hydrolysis rate from this complex (0.008  $\text{s}^{-1}$ ), the corrected dissociation rate constant was 0.030  $\text{s}^{-1}$ .



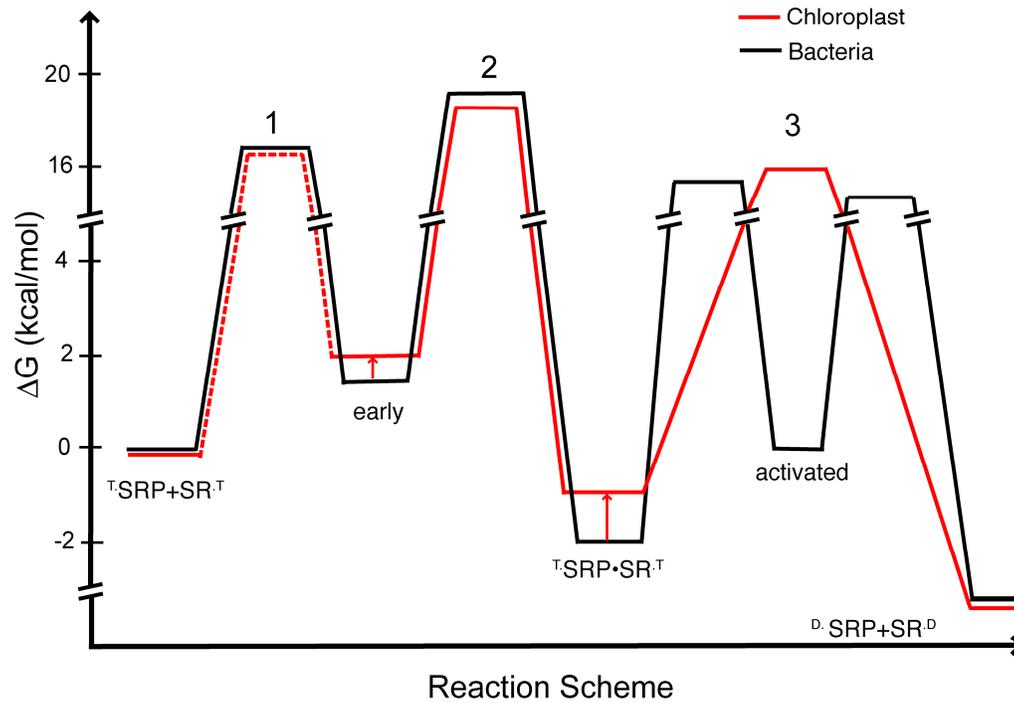
**Figure 1.3** Effects of cpFtsY mutations on its stimulated GTPase reaction with cpSRP54. (A, B) The stimulated GTPase reactions of wild-type cpFtsY (○) and mutants cpFtsY(A169W) (part A, ● and inset) and cpFtsY(A168W) (part B, ● and inset). (C) Inhibition assays for determining the affinities of mutants cpFtsY(A168W) (○) and cpFtsY(A169W) (●) for cpSRP54. The figures show representative data, and Table 1 summarizes the average values from two or more measurements. (D) Gel filtration analyses of stable complex formation of cpSRP54 with wild-type cpFtsY (black) and mutants cpFtsY D163A (blue), R166A (green), A168W (red), A169L (magenta), and A169W (cyan).



**Figure 1.4** Effects of cpSRP54 mutations on its stimulated GTPase reaction with cpFtsY. (A, B) The stimulated GTPase reactions of cpFtsY with wild-type cpSRP54 (○) and mutants cpSRP54(A143W) (part A, ● and inset) and cpSRP54(A142W) (part B, ● and inset). (C) Inhibition assays for determining the affinities of mutants cpSRP54(A142W) (○) and cpSRP54(A143W) (●) for cpFtsY. The figures show representative data, and Table 1 summarizes the average values from two or more measurements. (D) Mutant cpSRP54(A142W) (○) exhibits the same GTP-dependent complex assembly kinetics as wild-type cpSRP54 (●), measured using acrylodan-labeled cpSRP54(234C) as described in the Materials and Methods. Linear fits of data gave complex formation rate constants of  $1.9 \times 10^5$  and  $2.3 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$  for mutant and wild-type cpSRP54, respectively. (E) Gel filtration analyses of stable complex formation of cpFtsY with wild-type cpSRP54 (black) and mutants cpSRP54 D137A (blue), R140A (green), A142W (red), A143L (magenta), and A143W (cyan). (F) Fluorescence emission spectra of the acrylodan labeled at cpSRP54(234C) in the wild-type protein (black) compared with mutants cpSRP54 D137A (blue), A142W (red), and A143W (cyan).



**Figure 1.5** Effect of mutant GTPases on the targeting and integration of LHCP into thylakoid membranes. LHCP-dp1 and -dp2 denote the two 18.5 and 21 kDa protease-protected fragments that represent LHCP successfully targeted and integrated into the thylakoid membrane. Pre-LHCP was added to the reaction after the protease treatment and served as a loading control. (A, B) LHCP integration efficiency by the individual cpFtsY and cpSRP54 mutants, respectively. The top panels show representative data, and the lower panels show quantification of two or more measurements. All the data were normalized to that of the wild-type protein, which was set to 100%. (C, D) Correlation of the translocation defect of cpFtsY (part C) and cpSRP54 (part D) mutants with their  $k_{cat}/K_m$  values.



**Figure 1.6** Free energy profile for the GTP-dependent binding and activation cycles between the SRP and SR GTPases from bacteria (black) or chloroplast (red). The values for the *E. coli* GTPases were obtained from refs [19] and [27]. A standard state of 1  $\mu$ M was used. The activation free energies were calculated from the observed association and dissociation rate constants using  $\Delta G^\ddagger = -RT \ln(kh/k_B T)$ , where  $R = 1.987 \text{ cal K}^{-1} \text{ mol}^{-1}$ , Planck constant  $h = 1.58 \times 10^{-37} \text{ kcal s}^{-1}$ , Boltzmann constant  $k_B = 3.3 \times 10^{-27} \text{ kcal s}$ , and  $T = 298 \text{ K}$ . Equilibrium stabilities of complexes were calculated using  $\Delta G = \Delta G^\circ - RT \ln(K/K^\circ)$ .

## Supplementary Tables and Figures

**Supplementary Table 1.S1** Summary of different classes of mutational effects<sup>a</sup>

Mutational Effect	cpFtsY Mutants	cpSRP54 Mutants	Location
I – Complex formation defective	G288W <sup>b</sup>	G255W	NG interface
II – GTPase activation defective	A168W	A142W	IBD loop
III – Both steps defective	N135A		Dimer interface
	N135W	Q109W	Dimer interface
	D163A	D137A	IBD loop
	R166A	R140A	IBD loop
	A169L	A143L	IBD loop
	A169W	A143W	IBD loop
		G256W	NG interface
Neutral Mutations	F165A		IBD loop
	A167W		IBD loop
	R220A	R193W	Dimer interface
	G288A		NG interface
	G289A		NG interface
	G289W		NG interface
		Q109A	Dimer interface

<sup>a</sup>Mutations of homologous residues in cpSRP54 and cpFtsY are presented in the same row, except for cpSRP54(Q109A).

<sup>b</sup>kinetic parameters from Jaru-Ampornpan *et al.* [30]

**Supplementary Table 1.S2** Kinetic parameters of additional mutants outside of the IBD loop

Mutants	$k_{\text{cat}}$ ( $\text{min}^{-1}$ )	$K_{\text{M}}$ ( $\mu\text{M}$ )	$k_{\text{cat}}/K_{\text{M}}$ ( $\mu\text{M}^{-1}\text{min}^{-1}$ )
WT	$27 \pm 3$ (1)	$1.6 \pm 0.3$ (1)	$18 \pm 3$ (1)
<u>cpFtsY</u>			
N135A	$5.9 \pm 2.1$ (0.2)	$4.3 \pm 0.9$ (3)	$1.4 \pm 0.2$ (0.1)
N135W	$8.0 \pm 1.2$ (0.3)	$2.7 \pm 0.9$ (2)	$3.3 \pm 1.2$ (0.2)
R220A	26 (1)	2.2 (1)	12 (0.6)
G288A	23 (0.9)	1.1 (0.7)	22 (1)
G289A	$24 \pm 1$ (0.9)	$1.1 \pm 0.2$ (0.7)	$23 \pm 3$ (1)
G289W	$29 \pm 9$ (1)	$2.5 \pm 1.2$ (2)	$14 \pm 10$ (0.8)
<u>cpSRP54</u>			
Q109A	16 (0.6)	2.0 (1)	8.2 (0.5)
Q109W	7.7 (0.3)	4.2 (3)	1.9 (0.1)
R193W	18 (0.7)	2.8 (2)	6.3 (0.4)
G255W <sup>a</sup>	$38 \pm 3$ (1)	$41 \pm 4$ (26)	$0.93 \pm 0.15$ (0.05)
G256W <sup>a</sup>	$6.5 \pm 0.2$ (0.2)	$33 \pm 6$ (21)	$0.20 \pm 0.03$ (0.01)

<sup>a</sup> $k_{\text{cat}}$  and  $K_{\text{m}}$  values for mutants that are severely compromised in complex formation are determined by extrapolation, as saturation of reaction was not reached up to 50  $\mu\text{M}$  of mutant protein.

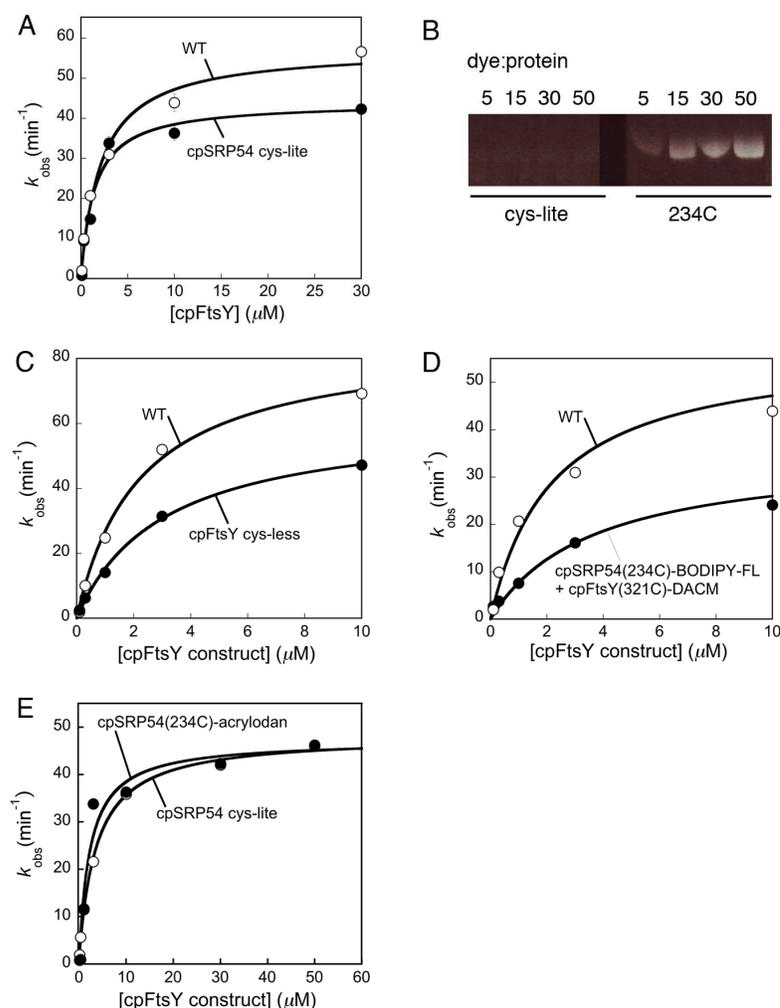
**Supplementary Table 1.S3** Mant-GTP binding affinity to the individual mutant GTPases<sup>a</sup>

GTPase constructs	$K_d$ ( $\mu$ M)
<b>cpFtsY</b>	
WT	1.8
D163A	1.2
R166A	3.1
A168W	3.1
A169L	ND
A169W	3.3
G288W	0.92
<b>cpSRP54</b>	
WT	6.7
D137A	2.6
R140A	58
A142W	2.6
A143L	68
A143W	3.5
G255W	8.2
G256W	35

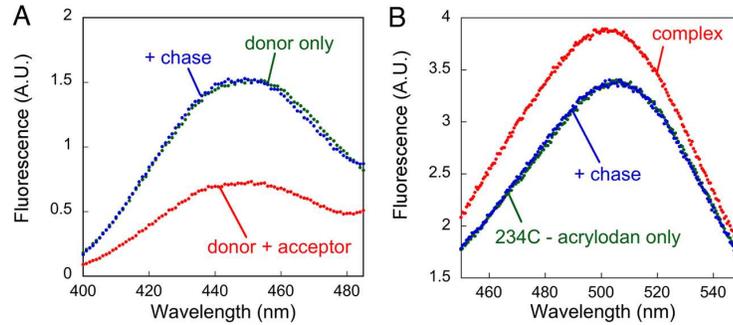
<sup>a</sup>Mant-GTP (Jena Bioscience) binding assay was performed as described in Jaru-Ampornpan *et al.* [26].

**Supplementary Table 1.S4** Basal GTPase activity of cpFtsY mutants

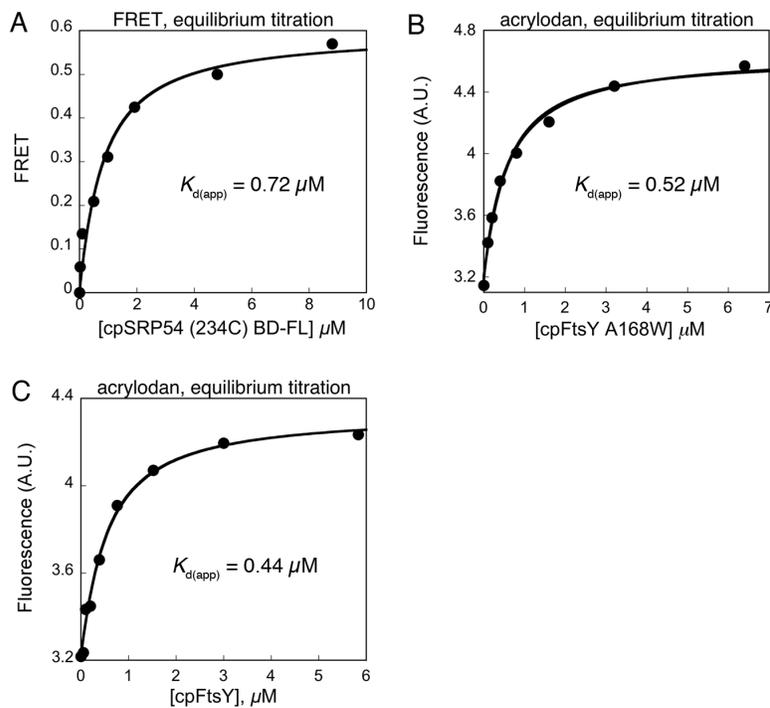
cpFtsY construct	$k_{\text{cat}}$ ( $\text{min}^{-1}$ )	$K_m$ ( $\mu\text{M}$ )
WT	$0.0039 \pm 0.0009$	$1.1 \pm 0.4$
N135A	$0.0077 \pm 0.0022$	$0.34 \pm 0.04$
N135W	$0.0020 \pm 0.0002$	$1.3 \pm 1.1$
R166A	$0.0028 \pm 0.0003$	$1.9 \pm 0.1$
D163A	0.011	0.44
F165A	$0.0124 \pm 0.0096$	$1.2 \pm 0.3$
A167W	0.0032	1.9
A168W	$0.0023 \pm 0.0002$	$1.1 \pm 0.6$
A169L	0.0038	2.1
A169W	$0.0088 \pm 0.0029$	$1.9 \pm 0.9$
R220A	$0.0050 \pm 0.0008$	$1.6 \pm 0.3$
G288A	$0.0031 \pm 0.0002$	$0.58 \pm 0.18$
G288W	$0.0039 \pm 0.0002$	$0.34 \pm 0.16$
G289A	$0.0029 \pm 0.0006$	$0.71 \pm 0.41$
G289W	$0.0030 \pm 0.0006$	$0.40 \pm 0.25$



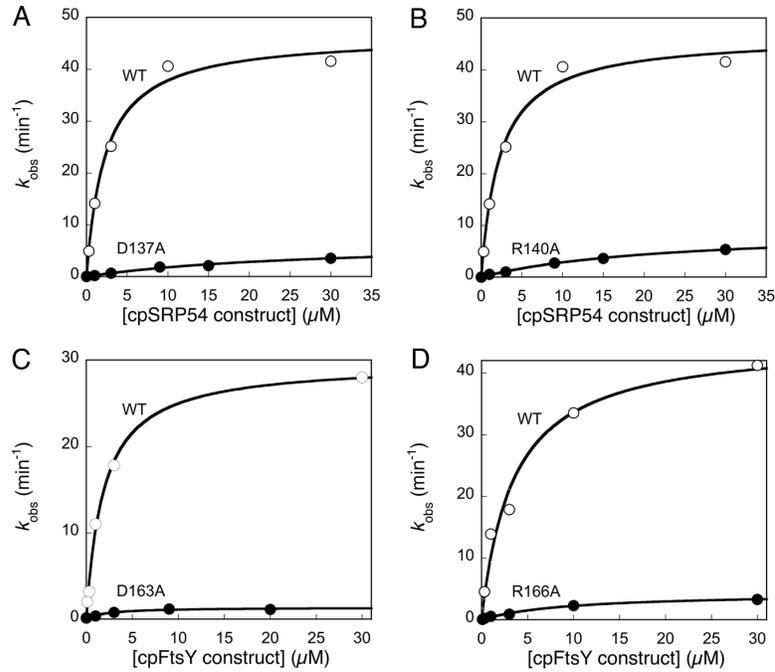
**Supplementary Figure 1.S1** Cys-lite cpSRP54 and cys-less cpFtsY and their fluorescently labeled versions are enzymatically active. All GTPase assays were performed as described in the *Methods*. cpSRP54(234C) refers to the cys-lite construct with a cysteine mutation at position 234 on cpSRP54. cpFtsY(321C) refers to the cys-less construct with a cysteine mutation at position 321 on cpFtsY. (A) The reciprocally stimulated GTPase reaction of cpSRP54 WT (○) or cpSRP54 cys-lite (●) with cpFtsY. Fits of the data to Eq 5 gave  $k_{\text{cat}}$  values of 56 and 54  $\text{min}^{-1}$ , respectively. (B) cpSRP54 cys-lite and cpSRP54(C234) were labeled with BODIPY-FL for 2hrs and loaded on a 10% SDS-PAGE gel. (C) The stimulated GTPase reaction of cpSRP54 with cpFtsY (○) or cys-less cpFtsY (●). Fits of the data to Eq 5 gave  $k_{\text{cat}}$  values of 56 and 39  $\text{min}^{-1}$ , respectively. (D) The stimulated GTPase reaction of cpFtsY with cpSRP54 (○), or of cpFtsY(321C)-DACM with cpSRP54(234C)-BODIPY-FL (●). Fits of the data to Eq 5 gave  $k_{\text{cat}}$  values of 39 and 20  $\text{min}^{-1}$ , respectively. (E) The reciprocally stimulated GTPase reaction of cys-lite cpSRP54 (○) or cpSRP54(C234)-acrylodan (●) with cpFtsY, which gave  $k_{\text{cat}}$  values of 51 and 43  $\text{min}^{-1}$ , respectively.



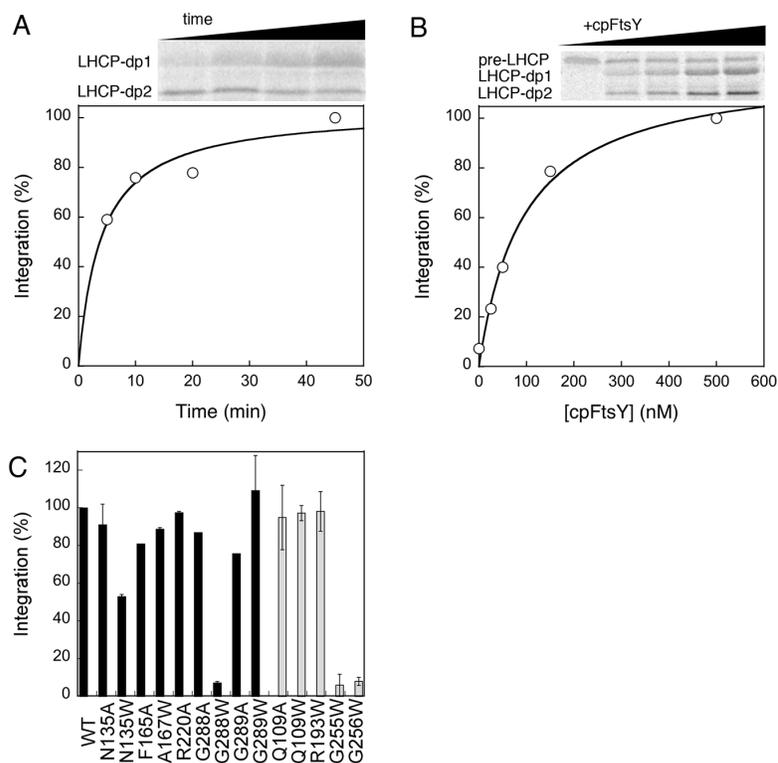
**Supplementary Figure 1.S2** FRET and acrylodan fluorescence signals from cpSRP54•cpFtsY complex could be competed away by EDTA or unlabeled cpSRP54. (A) Fluorescence emission spectra of cpFtsY(321C)-DACM (0.5 μM) in the absence (green) and presence of 2 μM cpSRP54(234C)-BODIPY-FL, with (blue) or without (red) 10 mM EDTA as the chase. Complex assembly was carried out in 2 mM GTP. (B) Acrylodan fluorescence increase upon formation of the cpSRP54•cpFtsY complex (green vs. red) could be chased by a 20-fold excess of unlabeled cpSRP54 (blue). Complex formation was carried out using 0.5 μM cpSRP54(234C)-acrylodan, 2 μM cpFtsY, and 2 mM GTP.



**Supplementary Figure 1.S3** Equilibrium titrations of the cpSRP54•cpFtsY complex in GTP using different fluorescence assays. (A) Complex assembly measured by FRET, using wildtype cpSRP54 and cpFtsY. (B) Complex assembly measured using cpSRP54(234C)-acrylodan and cpFtsY(A168W). (C) Complex assembly measured using cpSRP54(234C, A142W)-acrylodan and cpFtsY. The data were fit to Eqs. 1 or 2 in the Methods, and the  $K_d$  values are reported in Table 1.1.



**Supplementary Figure 1.S4** Defects of additional cpSRP54 and cpFtsY IBD loop mutants in complex formation and GTPase activation, measured by the stimulated GTPase reaction. (A, B) The stimulated GTPase reactions of wildtype cpSRP54 ( $\circ$ ), and mutants cpSRP54(D137A) (part A,  $\bullet$ ) and cpSRP54(R140A) (part B,  $\bullet$ ). (C, D) The stimulated GTPase reactions of wildtype cpFtsY ( $\circ$ ), and mutants cpFtsY(D163A) (part C,  $\bullet$ ) and cpFtsY(R166A) (part D,  $\bullet$ ). The figures show representative data, and Table 1.1 summarizes the average values from two or more measurements.



**Supplementary Figure 1.S5** The effects of additional cpSRP54 and cpFtsY mutations on LHCP targeting and integration. (A,B) Time- (part A) and cpFtsY concentration- (part B) dependent targeting and translocation of LHCP, carried out as described in *Methods*. (C) Integration efficiencies of additional cpFtsY (black bars) and cpSRP54 (grey bars) mutants.

## References

1. Koch, H. G., Moser, M., and Muller, M. (2003) Signal recognition particle-dependent protein targeting, universal to all kingdoms of life. *Re. Physiol. Biochem. Pharmacol.* **146**, 55–94
2. Walter, P., and Johnson, A. E. (1994) Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Ann. Rev. Cell Biol.* **10**, 87–119
3. Keenan, R. J., Freymann, D. M., Stroud, R. M., and Walter, P. (2001) The signal recognition particle. *Annu. Rev. Biochem.* **70**, 755–775
4. Zopf, D., Bernstein, H. D., Johnson, A. E., and Walter, P. (1990) The methionine-rich domain of the 54 kd protein subunit of the signal recognition particle contains an RNA binding site and can be crosslinked to a signal sequence. *EMBO J.* **9**, 4511–4517
5. Zopf, D., Bernstein, H. D., and Walter, P. (1993) GTPase Domain of the 54-kD subunit of the mammalian signal recognition particle is required for protein translocation but not for signal sequence binding. *J. Cell Biol.* **120**, 1113–1121
6. Halic, M., Blau, M., Becker, T., Mielke, T., Poll, M.R., Wild, K., Sinning, I., and Beckmann, R. (2006) Following the signal sequence from ribosomal tunnel exit to signal recognition particle. *Nature* **444**, 507–511
7. Schaffitzel, C., Oswald, M., Berger, I., Ishikawa, T., Abrahams, J.P., Koerten, H.K., Koning, R.I., and Ban, N. (2006) Structure of the E. coli signal recognition particle bound to a translating ribosome. *Nature* **444**, 503–506
8. Bourne, H. R., Sanders, D.A., and McCormick, F. (1991) The GTPase superfamily: Conserved structure and molecular mechanism. *Nature* **349**, 117–127
9. Eichacker, L. A., and Henry, R. (2001) Function of a chloroplast SRP in thylakoid protein export. *Biochim. et Biophys. Acta* **1541**, 120–134
10. Schunemann, D., Gupta, S., Persello-Cartieaux, F., Klimyuk, V. I., Jones, J. D. G., Nussaume, L., and Hoffman, N. E. (1998) A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 10312–10316
11. Funke, S., Knechten, T., Ollesch, J., and Schunemann, 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**, 8912–8917
12. Richter, C. V., Trager, C., and Schunemann, D. (2008) Evolutionary substitution of two amino acids in chloroplast SRP54 of higher plants cause its inability to bind SRP RNA. *FEBS Lett.* **582**, 3223–3229
13. Franklin, K. E., and Hoffman, N. E. (1993) Characterization of a chloroplast homologue of the 54-kDa subunit of the signal recognition particle. *J. Biol. Chem.* **268**, 22175–22180
14. Delille, J., Peterson, E. C., Johnson, T., Morre, M., Kight, A., and Henry, R. (2000) A novel precursor recognition element facilitates posttranslational binding to the signal recognition particle in chloroplasts. *Proc. Natl. Acad. Sci.* **97**, 1926–1931
15. Tu, C. J., Peterson, E. C., Henry, R., and Hoffman, N. E. (2000) The L18 domain of light-harvesting chlorophyll proteins binds to chloroplast signal recognition particle 43. *J. Biol. Chem.* **275**, 13187–13190
16. Tu, C.-J., Schuenemann, D., and Hoffman, N. E. (1999) 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.* **274**, 27219–27224
17. Shan, S., Stroud, R., and Walter, P. (2004) Mechanism of association and reciprocal activation of two GTPases. *Plos Biology* **2**, e320
  18. Zhang, X., Kung, S., and Shan, S. (2008) Demonstration of a two-step mechanism for assembly of the SRP-SRP receptor complex: implications for the catalytic role of SRP RNA. *J. Mol. Biol.* **381** 581–593
  19. Zhang, X., Schaffitzel, C., Ban, N., and Shan, S. (2009) Multiple conformational changes in a GTPase complex regulate protein targeting. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1754–1759
  20. Zhang, X., Rashid, R., Kai, W., and Shan, S. (2010) Sequential checkpoints govern fidelity during co-translational protein targeting. *Science* **328**, 757–760
  21. Egea, P. F., Shan, S., Napetschnig, J., Savage, D.F., Walter, P., and Stroud, R.M. (2004) Substrate twinning activates the signal recognition particle and its receptor. *Nature* **427**, 215–221
  22. Shan, S., Chandrasekar, S., and Walter, P. (2007) Conformational changes in the GTPase modules of SRP and its receptor drive initiation of protein translocation. *J. Cell Biol.* **178**, 611–620
  23. Jaru-Ampornpan, P., Shen, K., Lam, V.Q., Ali, M., Doniach, S., Jia, T.Z., and Shan, S. (2010) ATP-independent reversal of a membrane protein aggregate by a chloroplast SRP subunit. *Nature Struct. Molec. Biol.* **17**, 696–702
  24. Falk, S., Ravaud, S., Koch, J., and Sinning, I. (2010) The C-terminus of the Alb3 membrane insertase recruits cpSRP43 to the thylakoid membrane. *J. Biol. Chem.* **285**, 5954–5962
  25. Yuan, J., Kight, A., Goforth, R.L., Moore, M., Peterson, E.C., Sakon, J., and Henry, R. (2002) ATP stimulates signal recognition particle (SRP)/FtsY-supported protein integration in chloroplasts. *J. Biol. Chem.* **277**, 32400–32404
  26. Jaru-Ampornpan, P., Chandrasekar, S., and Shan, S. (2007) Efficient interaction between two GTPases allows the chloroplast SRP pathway to bypass the requirement for an SRP RNA. *Mol. Biol. Cell* **18**, 2636–2645
  27. Peluso, P., Shan, S., Nock, S., Herschlag, D., and Walter, P. (2001) Role of SRP RNA in the GTPase cycles of Ffh and FtsY. *Biochemistry* **40**, 15224–15233
  28. Yuan, J., Henry, R., and Cline, K. (1993) Stromal factor plays an essential role in protein integration into thylakoids that cannot be replaced by unfolding or by heat shock protein Hsp70. *Proc. Natl. Acad. Sci.* **90**, 8552–8556
  29. Shepotinovskaya, I. V., and Freymann, D.M. (2001) Conformational change of the N-domain on formation of the complex between the GTPase domains of *Thermus aquaticus* Ffh and FtsY. *Biochem. Biophys. Acta* **1597**, 107–114
  30. Jaru-Ampornpan, P., Nguyen, T.X., and Shan, S. (2009) A distinct mechanism to achieve efficient SRP-SRP receptor interaction by the chloroplast SRP. *Mol. Biol. Cell* **20**, 3965–3973
  31. Shan, S., and Walter, P. (2005) Co-translational protein targeting by the signal recognition particle. *FEBS Lett.* **579**, 921–926
  32. Chandrasekar, S., Chartron, S., Jaru-Ampornpan, P., and Shan, S. (2008) Crystal structure of the chloroplast signal recognition particle (SRP) receptor: domain arrangement modulates SRP-receptor interaction. *J. Mol. Biol.* **375**, 425–436

33. Nilsson, R., and van Wikj, K.L. . (2002) Transient interaction of cpSRP54 with elongating nascent chains of the chloroplast-encoded D1 protein; 'cpSRP54 caught in the act'. *FEBS Letters* **524**, 127–133
34. Lewis NE, M. N., Kathir KM, Rajalingam D, Kight AD, Daily A, Kumar TK, Henry RL, Goforth RL. (2010) A dynamic cpSRP43-Albino3 interaction mediates translocase regulation of chloroplast signal recognition particle (cpSRP)-targeting components. *J Biol Chem* **285**, 34220-34230

## **Chapter 2**

### **Mechanism of an ATP-Independent Protein**

### **Disaggregase.**

#### **Part I. Structure of a Membrane Protein Aggregate**

#### **Reveals a Mechanism of Recognition by Its Chaperone**

A version of this chapter has been published as

Nguyen, T.X., Jaru-Ampornpan, P., Lam, V.Q., Cao, P., Piszkiwicz, S., Hess, S., and Shan, S.

(2013) *J. Biol. Chem.*, **288** (19), 13420-30.

**Abstract**

Protein aggregation is detrimental to the maintenance of proper protein homeostasis in all cells. To overcome this problem, cells have evolved a network of molecular chaperones to prevent protein aggregation and even reverse existing protein aggregates. The most extensively studied ‘disaggregase’ systems are ATP-driven macromolecular machines. Recently, we reported an alternative ‘disaggregase’ system, in which the 38-kDa subunit of chloroplast Signal Recognition Particle (cpSRP43) efficiently reverses the aggregation of its substrates, the light-harvesting chlorophyll *a/b*-binding (LHC) proteins, in the absence of external energy input. To understand the molecular mechanism of this novel activity, here we used biophysical and biochemical methods to characterize the structure and nature of LHC protein aggregates. We show that LHC proteins form micellar, disc-shaped aggregates that are kinetically stable and detergent-resistant. Despite their non-amyloid nature, the LHC aggregates have a defined global organization, displaying the chaperone recognition motif on their solvent-accessible surfaces. These findings suggest an attractive mechanism for recognition of the LHC aggregate by cpSRP43 and provide important constraints to define the capability of cpSRP43’s disaggregase activity.

## Introduction

The proper folding of proteins into their native structures is essential for the function and survival of cells. However, environmental stress, molecular crowding, and potential exposure of hydrophobic regions of proteins during their biogenesis [1-3] pose challenges to protein folding *in vivo*. In this setting, improper intra- or intermolecular interactions can lead to the aggregation of proteins. Aggregate formation is detrimental to cells as it removes functional proteins [4]. Moreover, some aggregates, both amorphous ones and those that lead to highly ordered amyloid fibrils, are toxic to cells and have been implicated in a variety of protein folding diseases [5-7].

Cells have evolved elaborate mechanisms to overcome the problems associated with protein aggregation. A specialized class of molecular chaperones, the ‘disaggregases’, can perform the energetically uphill process of reversing protein aggregation. Thus far, studies of disaggregases have been dominated by the Clp/Hsp100 family of AAA<sup>+</sup> ATPases (ATPases associated with various cellular activities), such as ClpB in prokaryotes and Hsp104 in yeasts [3]. Both are large hexameric rings (>500 kDa) powered by mechanical forces from ATP hydrolysis and require additional co-chaperones to efficiently disassemble a variety of protein aggregates [8,9]. The complexity of these disaggregase systems and the promiscuity in their substrate selection has made it difficult to pinpoint their molecular mechanisms of action. Further, AAA<sup>+</sup>-disaggregase machines were only found in prokaryote and yeast, and no homologues have been identified in higher eukaryotes outside of plastids and mitochondria. It is conceivable that alternative mechanisms of disaggregation, such as the recently described Hsp110-70-40 system [10,11] could be used in higher eukaryotes. An understanding of alternative disaggregase systems can shed light on novel principles and mechanisms by which cellular chaperones overcome protein aggregates.

Previously, we identified an efficient disaggregase activity in the chloroplast Signal Recognition Particle 43 subunit (cpSRP43). This provides an example in which a relatively small protein scaffold (38 kDa) can recognize and disrupt large protein aggregates in an ATP-independent mechanism [12], in contrast to the Clp/Hsp100 family of disaggregases. cpSRP43 is part of the protein targeting machinery, the cpSRP, that mediates the delivery of the light-harvesting chlorophyll *a/b*-binding (LHC) family of proteins to the thylakoid membrane [13-15]. The most abundant member of the LHC family, LHCP, comprises ~30% of the proteins on the thylakoid membrane and is arguably the most abundant membrane protein on earth. The sheer abundance of these proteins and their highly hydrophobic nature demands highly effective chaperones that protect them from aggregation before arrival at the membrane. In the chloroplast stroma, this chaperone function is provided by cpSRP43, which effectively protects LHC proteins from aggregation and can even reverse preformed large LHC protein aggregates [12,16]. cpSRP43 recognizes a highly-conserved 18-amino acid loop between the second and the third transmembrane (TM) domains of LHC proteins, termed L18 [17,18]. In previous work, we showed that the specific interaction of cpSRP43 with the L18 motif is crucial for the chaperone and disaggregase activity of cpSRP43 [12]. This and other observations led us to propose that, in the absence of external energy input, cpSRP43 uses specific binding energy with its substrate proteins to remodel and rescue LHC protein aggregation [12].

To gain insights into the molecular mechanism that underlies cpSRP43's novel disaggregase activity, we need to first understand the nature of the LHC aggregate and identify the structural features that facilitate its disassembly by cpSRP43. To this end, we examined the nature and structure of the LHC aggregate using biophysical and biochemical techniques. We show that LHC proteins form disc-like particles with a relatively amorphous hydrophobic core,

but exhibit a defined interior/exterior organization in which the L18 recognition motif is displayed on the solvent-exposed surface. This suggests an attractive mechanism for cpSRP43 to recognize the LHC aggregates and thus initiate their disassembly.

## Materials and Methods

**Materials.** LHCP, Lhcb5 and their mutants were purified under denaturing conditions as described [12], except that 6M GdmHCl was used instead of 8M urea for Lhcb5.  $A\beta^{1-42}$  and recrystallized thioflavin T (ThT) were generous gifts from Dr. J. W. Kelly. 1-anilino-8-naphthalene sulfonate (ANS) and bis-ANS were from Sigma and Invitrogen, respectively. n-dodecyl-N,N,-dimethylamine-N-oxide (LDAO), n-dodecyl- $\beta$ -D-maltopyranoside (DDM), n-octyl- $\beta$ -D-glucopyranoside ( $\beta$ -OG) and n-nonyl- $\beta$ -D-glucopyranoside (BNG) were from Anatrace. Triton X-100 (TX-100) was from Sigma and sodium dodecyl sulfate (SDS) was from BioRad. Urea and GdmHCl were molecular biology grade from MP and Sigma, respectively. 1-oxy-2,2,5,5-tetramethylpyrrolidine-3-methyl methanethiosulfonate (MTSSL) was from Toronto Research Chemicals, N-ethyl-maleimide was from Sigma and N-(1-pyrene)-maleimide was from Invitrogen.

**Light Scattering Assay.** Light scattering experiments were performed as previously described [12]. For formation of aggregates (Figure 2.3, black), unfolded LHCP in 8 M urea was directly diluted into Buffer D (50 mM KHEPES pH 7.5, 200 mM NaCl) to the desired final concentration; the final concentration of urea was equalized among different samples. The CMC is obtained as the x-intercept from the linear fit of the data [19]. For serial dilution experiments (Figure 2.3, red), the sample at 1  $\mu$ M LHCP was serially diluted (by two-fold) into fresh buffer D and allowed 10 min to equilibrate before taking measurement.

**Transmission Electron Microscopy.** LHCP aggregates were formed by diluting unfolded LHCP in 8 M urea into Buffer D to the final concentration of 2  $\mu$ M. After incubation at 25 °C for 5 minutes, the sample was diluted five-fold and immediately deposited onto a glow-discharged 200-mesh Formvar grid (Ted Pella Inc., CA). After 45-second adsorption time, the

grid was washed in water and then stained with 1% uranyl acetate for 45 seconds. TEM images were obtained on a 120 kV Tecnai T12 electron microscope coupled with a CCD camera. The diameters of the particles were measured using ImageJ [20].

**Atomic Force Microscopy.** 1  $\mu\text{M}$  LHCP aggregate in Buffer D was deposited onto a freshly cleaved mica and incubated for 5 minutes at 25 °C to allow equilibration. The wafer was rinsed with Millipore water and dried under the weak flux of nitrogen. AFM images were taken immediately after the sample was prepared. A Digital Instrument Nanoscope IIIA AFM system in tapping mode was used throughout at ambient conditions. A sharp TESP tip (Veeco, CA) was used in the experiment. Typical values for the force constant, resonance frequency and tip radius were 42 N/m, 320 kHz, and 8 nm, respectively. Particle sizes were obtained by calculating the projected area of each particle at half maximal height onto the surface. This is because the apparent *lateral* size of surface features is usually overestimated due to the broadening effect of the AFM tip. The cross sectional area at half the maximum height provides a more realistic distribution of sizes of the particles.

**Fluorescence.** All fluorescence experiments were carried out in Buffer D using a Fluorolog 3-22 spectrofluorometer (Jobin Yvon). For bis-ANS experiments, 1 mM bis-ANS was added to Buffer D with or without 1  $\mu\text{M}$  LHCP aggregate. The samples were excited at 395 nm and then scanned from 410 to 620 nm, with the excitation and emission band passes of 2 and 5 nm, respectively. For ThT experiments, 20 mM re-crystallized ThT was added to Buffer D containing no aggregate, aggregates from 1 or 5  $\mu\text{M}$  LHCP, or 15  $\mu\text{M}$  freshly sonicated  $\text{A}\beta^{1-42}$  amyloid. The samples were excited at 440 nm and then scanned from 470 to 570 nm, with the excitation and emission band passes of 3 and 7 nm, respectively. For comparison, ThT fluorescence from 1 and 5  $\mu\text{M}$  unfolded LHCP in 8 M urea were measured.

For pyrene excimer experiments, DTT-reduced single cysteine mutants of Lhcb5 in 6 M GdmHCl were labeled with a 30-fold molar excess of pyrene maleimide at room temperature in the dark for 2 hours. Excess pyrene was removed by gel filtration, and the efficiency of spin-labeling (90-100%) was determined by LC-MSD 1100 series (Agilent Technologies, Santa Clara, CA). The samples were prepared by diluting pyrene-labeled Lhcb5 pairs into Buffer D for a final concentration of 1.5  $\mu$ M for each variant. Spectra were obtained from excitation at 317 nm and then scanned from 360 to 560 nm, with the excitation and emission band passes of 3 and 6 nm, respectively. The amount of excimer fluorescence, indicated by a red shift to 445 nm, is normalized against the non-excited fluorescence signal at 376 nm. Statistically, when two variants A and B are mixed there is a population distribution of homo-pairs (e.g. 25% A-A and 25% B-B) and hetero-pairs (50% A-B). The equation below corrects for the real hetero-pair excimer ( $F_{AB}$ ):

$$F_{AB} = 2 \times (F_{AB, \text{app}} - 0.25F_A - 0.25F_B)$$

where  $F_{AB, \text{app}}$  is the apparent ratio of excimer fluorescence ( $I_{445}/I_{376}$ ) between two pyrene-labeled variants, and  $F_A$  and  $F_B$  is the ratio of excimer fluorescence of each individual variants measured separately.

**Sedimentation.** Unfolded LHCP was diluted to 10  $\mu$ M in Buffer D and incubated at 25  $^{\circ}$ C for 5 minutes. Aggregation was complete, judged by the absence of LHCP in the supernatant after centrifugation at 13,000 rpm in a microfuge for 30 minutes. The pellet was dissolved with 50  $\mu$ l of detergent or chemical denaturants at different concentrations for 30 minutes at 25  $^{\circ}$ C. The mixtures were then spun at 13,000 rpm in a microfuge for 30 minutes, and soluble (S) and pellet (P) fractions were boiled and visualized by SDS-PAGE.

**SDS Solubility.** For Figure 2.2B, the assay was performed as described for amyloid fibrils [21]. Briefly, aggregation of 10  $\mu$ M LHCP in Buffer D preceded for 5 minutes at 25 °C. The mixture was then mixed with 2% SDS-PAGE loading buffer and either incubated at 25 °C or 100 °C for ten minutes prior to SDS-PAGE. Only the proteins that migrated into the resolving gel (e.g. solubilized portion) were visualized.

**Spin Labeling and Electron Paramagnetic Resonance Measurements.** Spin labeling reactions were performed in 6 M GdmHCl, 50 mM KHEPES, pH 7.5, and 2 mM EDTA. Reduced and degassed single cysteine mutants of Lhcb5 were labeled with a three- to five-fold molar excess of MTSSL at room temperature in the dark for 2-3 hours. Excess MTSSL was removed by gel filtration, and the efficiency of spin-labeling (80-100%) was determined by EPR using a TEMPO calibration curve according to manufacturer's instructions (Bruker). EPR spectra were acquired using a 9.4-GHz (X-band) EMX EPR spectrometer (Bruker) equipped with an ER 4119HS cavity at 20-23 °C. To form the aggregate, the individual spin-labeled proteins in GdmHCl were diluted into Buffer D. The concentrations of the aggregate samples were 30-100  $\mu$ M. Data acquisition was previously described [22].

**NEM Alkylation and MS Analysis.** Cysteine mutants of Lhcb5 in 6 M GdmHCl were reduced with 2.5 mM TCEP at RT for 2 hours. Each mutant was diluted into Buffer D to a final concentration of 3.3  $\mu$ M and incubated on ice for 10 minutes to form the aggregate, followed by the addition of 100  $\mu$ M NEM. The reaction was quenched with 50 mM DTT at various time points, concentrated under vacuum, redissolved in 0.2% formic acid, and ca. 25 pmol protein was analyzed on an LC-MSD SL 1100 series (Agilent). The samples were chromatographed on a 2.1 x 150 mm Zorbax 300SB-C3 column (Agilent) using a gradient consistent of 0.2% formic acid and 0.2% formic acid in acetonitrile (89.8%) and methanol (10%). Intact masses were

measured in the single quadrupole and quantified using the software ChemStation software (Agilent). Control experiments where different ratios of un-alkylated and alkylated proteins were mixed and subjected to MS analysis shows the quantification of the ratio of alkylated species to be reliable (Figure 2.6E). The reported accessibilities were calculated as a ratio of the alkylation of each cysteine mutant under aggregation Buffer D versus denaturing 6 M GdmHCl.

## Results

**LHCP Aggregates Contain Exposed Hydrophobic Grooves.** To characterize the surface features of LHC protein aggregates, we used an established collection of small molecule dyes. Exposure of hydrophobic patches or crevices within aggregates can be probed by extrinsic fluorescent molecular dyes like ANS and bis-ANS [23,24]. We tested whether the aggregates of LHCP, the most abundant member of the LHC protein family, share this feature. Indeed, the fluorescence of both ANS (data not shown) and bis-ANS (Figure 2.1A) increased significantly in the presence of 1  $\mu$ M LHCP aggregate, accompanied by a blue shift of the fluorescence emission spectra. These results strongly suggest that LHCP aggregates contain exposed hydrophobic micro-domains that allow the binding of these dyes, consistent with the highly hydrophobic nature of this protein.

We next used thioflavin T (ThT) to probe the structural organization of the LHCP aggregate. ThT is often used as a diagnostic for the formation of amyloid fibrils generated by amyloid- $\beta$ (A $\beta$ ),  $\alpha$ -synuclein, and other amyloidogenic proteins [25]. Similar to bis-ANS, the fluorescence of ThT exhibited a significant increase in intensity and a blue shift in spectrum in the presence of the LHCP aggregate (Figure 2.1B, blue lines). The extent of these fluorescence changes is comparable to that induced by mature amyloid fibrils generated by the A $\beta$ <sup>1-40</sup> peptide (Figure 2.1B, red vs. blue and Figure 2.1C). As microscopy analyses did not indicate fibril formation in the LHCP aggregate (see below), these results suggest that ThT is not highly specific for amyloid fibrils, consistent with recent work observing ThT fluorescence of non-fibrillar aggregates of  $\beta$ -lactoglobulin and transthyretin [19,26]. Instead, this dye possibly binds to hydrophobic grooves that are often present in amyloid fibrils but can also be generated by other types of aggregates [27].

**LHCP Forms Stable Aggregates.** To probe the stability of the LHCP aggregate, we tested its solubility in various detergents, including LDAO, DDM,  $\beta$ -OG, BNG, and TX-100. By analyzing the amount of proteins in the soluble and insoluble fractions after medium-speed sedimentation (see Methods), we showed that none of these detergents were able to solubilize the LHCP aggregate at or above respective concentrations typically used for membrane protein solubilization (Figure 2.2A).

In addition, we tested the solubility of the LHCP aggregate in SDS using an established protocol for amyloid fibrils [21]. This assay evaluates solubility of the aggregate based on the mobility of the protein in SDS-PAGE after incubation with SDS-containing buffer at room temperature (see Methods). ‘SDS-insoluble’ amyloid fibrils or oligomeric protein aggregates cannot enter the resolving gel unless boiled [21]. LHCP aggregate showed significant resistance to 2% SDS in this procedure, as only 24% of the aggregates could be solubilized and migrated into the gel without boiling (Figure 2.2B, right panel). SDS could solubilize large LHCP aggregates only after extensive incubation and boiling of the sample (Figure 2.2B, left panel). Taken together, the detergent-resistance of the LHC protein aggregate suggests the presence of highly stable packing interactions within the aggregate that must be overcome by cpSRP43.

**LHCP Forms Micellar, Disc-shaped Aggregates.** Formation of large LHC aggregates can be monitored based on light scattering at 360 nm [12]. The scattering intensity increases linearly with LHCP concentration above ~100 nM (Figure 2.3, black), suggesting that aggregate formation was complete under these conditions. However, the linearity broke down at lower LHCP concentrations (Figure 2.3 and inset, black). This was not due to limitations in instrument sensitivity: when pre-formed LHCP aggregates were diluted, linearity in light scattering intensity was observed at all concentrations and extrapolated through zero (Figure 2.3 and inset, red).

These observations show that: (i) the LHCP aggregate is kinetically stable and virtually irreversible once it has formed; and (ii) formation of the LHCP aggregate requires a critical protein concentration, reminiscent of the critical micellar concentration during micelle formation. An analogous, ‘critical aggregate concentration’ of 125 nM was obtained for the LHCP aggregates from these data (see Methods). This micelle-like characteristic begins to suggest a globular morphology of the LHC aggregates.

To directly observe the global structure of LHC aggregates, we examined them using transmission electron microscopy (TEM) and atomic force microscopy (AFM). Negatively stained TEM images revealed LHCP aggregates to be circular particles (Figure 2.4A,B). Analysis of the size of these particles resulted in a distribution that fits well to a Gaussian function, with diameters of  $12\pm 2$  nm (Figure 2.4C). Consistent with the EM images, AFM analysis also showed LHCP aggregates to be disc-shaped particles (Figure 2.5A,B) with mean areas of  $214\pm 94$  nm<sup>2</sup> (Figure 2.5C), or mean diameters of  $16\pm 5$  nm, in good agreement with the EM measurements. Strikingly, the heights of the aggregates measured by AFM are ‘quantized’ and peaked at integrals of 0.7-0.8 nm (Figure 2.5D and inset). These results suggest that LHC proteins form disc-shaped aggregates with a height of 0.7-0.8 nm, and these discs can further stack upon one another.

**The L18 Recognition Motif is Displayed on the Aggregate Surface.** To probe the global structure of LHC aggregates at higher resolution, we engineered 30 single-cysteine mutations, which span every 5-10 residues throughout the entire protein sequence of Lhcb5 (Figure 2.6A, blue). Lhcb5 is a close homologue of LHCP (Supplementary Figure 2.S1) that strongly depends on the cpSRP pathway for its biogenesis and whose aggregate is efficiently prevented and disassembled by cpSRP43 [12,28]. All single-cysteine mutants were able to form

light-scattering aggregates with the same extent and kinetics as wildtype Lhcb5, and thus can be used to probe the assembly of the wildtype aggregate (Figure 2.6B).

With each single-cysteine mutant, we used two independent methods to measure their relative positions on the LHC aggregate. In the first approach, we labeled each cysteine with the nitroxide spin probe MTSSL in 6M GdmHCl, allowed for aggregation in aqueous buffer, and used electron paramagnetic resonance (EPR) spectroscopy to investigate the local backbone mobility of each specific site in the aggregate. A probe buried inside the aggregate will engage in strong interactions and have more restricted motion than those on the solvent-exposed surface, and hence exhibit broader central linewidth ( $\Delta H$ ) and hyperfine splitting (hfs) in the overall spectral width (Figure 2.6C, green versus red spectra). As a control, EPR measurements were carried out for each protein variant solubilized in 6M GdmHCl; all spin labels displayed similar, low values of  $\Delta H$  under these conditions, indicating the high mobility of the residues in the unfolded protein (Figure 2.6C, black).

Upon formation of the aggregate, the spin probes in all the TMs, loop 1 (between TM1 and TM2) and the C-terminus of Lhcb5 displayed high  $\Delta H$  values and broad EPR spectra, suggesting that they are highly immobile and likely engaged in strong inter- or intramolecular interactions (Figure 2.6D, Supplementary Table 1  $\Delta H^{-1}$  values are plotted). In contrast, spin probes placed in the L18 motif and the N-terminus of TM3 are highly mobile, indicating that these regions are free from any extensive interactions in the aggregate and are likely solvent-exposed (Figure 2.6D). In addition, spin labels at the N-terminus of Lhcb5 also displayed highly mobile spectra.

To independently probe the global architecture of LHC protein aggregates, we examined the susceptibility of the individual cysteine residues to alkylation with *N*-ethyl-maleimide

(NEM). The cysteine residues on the solvent-accessible surface of the aggregate will react rapidly and efficiently with NEM, whereas those buried within the aggregate will be alkylated much less efficiently (Figure 2.6F, green versus red curves). The efficiency of alkylation can be quantified by intact mass spectrometry and provides a direct measure for the solvent accessibility of individual residues in the LHC aggregate (Figure 2.6E). As a control for the intrinsic bias in the reactivity of cysteines at different positions, each single-cysteine mutant was solubilized in 6M GdmHCl and tested in parallel experiments (Figure 2.6F, black).

In agreement with the results of EPR measurements, the residues within the TMs exhibit low efficiency of alkylation, ranging from 20-40%, in contrast to the almost complete alkylation of the respective cysteines under denaturing conditions (Figure 2.6G and Supplementary Table 1). Residues in loop 1 and the C-terminus of the LHC protein exhibit slightly higher alkylation efficiency, 40-60%, indicating that these regions are partially buried in the aggregate but to a lesser extent than the TMs. In contrast, residues on the L18 loop and the N-terminal end of TM3 proximal to L18 are almost completely alkylated (90-100%), suggesting that these sites are highly solvent accessible and presented on the exterior of the aggregate. Finally, residues in the N-terminus of Lhcb5 showed almost 100% reactivity, again demonstrating the exposure of this region on the aggregate surface.

Although the burial of TMs in the Lhcb5 aggregate is expected due to their hydrophobic nature, the low mobility and inaccessibility of loop 1 and the C-terminus of Lhcb5 were surprising. We therefore asked if the burial of these loop regions results from topological constraints imposed by the neighboring TMs, or from the inherent physicochemical property of the looping sequence. To address this question, we altered the location of L18 in Lhcb5 by either swapping it with loop 1 (TM1-L18-TM2-TM3) to construct the LoopSwap (LS) mutant, or with

the protein sequence C-terminal to TM3 (TM1-TM2-TM3-L18) to construct the Cterm mutant. The aggregate formed by both constructs can be rescued by cpSRP43 (Figure 2.7A), suggesting that cpSRP43, despite its specific interaction with L18, can tolerate variations in the remainder of its substrate's sequence.

To probe the accessibility of individual motifs in the L18-swapped mutants, we probed the accessibility of engineered single cysteines in each domain by NEM alkylation. The alkylation efficiency of each motif in both L18-swapped mutants is similar to that of wildtype Lhcb5: the L18 motif is highly accessible and almost completely alkylated, whereas loop 1 and the C-terminus regions exhibit medium levels of alkylation (Figure 2.7B-D). These results indicate that the intrinsic properties of these sequences/domains determine their accessibility in the aggregate, and demonstrate that the L18 motif has a strong tendency to be displayed on the surface of protein aggregates.

**LHC Aggregates Contain an Amorphous Hydrophobic Core.** To determine whether the TMs of the LHC protein make specific intermolecular contacts in the buried core of the aggregate, we exploited the ability of pyrene labels to form excited-state dimers (excimers) when they are within 4–10 Å of each other. High pyrene excimer fluorescence reports on close proximity between specific sites within the aggregate. To this end, we mixed two proteins, each labeled with pyrene at a single cysteine residue, in all pair-wise combinations, allowed them to form the aggregates, and monitored for pyrene excimer fluorescence at 455 nm relative to the monomer fluorescence at 375 nm (see Methods). As a positive control, we used a double-cysteine mutant V139C-L140C and labeled both positions with pyrene probes. The excimer ratio of this construct was high, ~0.8 (Figure 2.8A). As the negative control, pyrene labeled L180C was used (excimer ratio = 0.08; Figure 2.8A).

Two important lessons were learned from the results of these measurements. First, many pyrene pairs exhibit excimer fluorescence intensities substantially above the background and above the other pyrene pairs, with excimer ratios of  $\geq 0.30$  for homo-pyrene pairs at multiple residues in TM1 (67, 72, 88, and 92) and for multiple hetero-pyrene pairs in all three TMs (Figure 2.8B). Second, these excimer fluorescence intensities were still modest, up to 0.36 (Figure 2.8B). These values are far below the values of 0.6-0.8 expected for specifically interacting pairs that are always in close proximity. These data indicate that in the LHC aggregate, the TMs form extensive intermolecular contacts in its hydrophobic core, but these interactions are much less specific than those observed in amyloid fibrils [29].

Taken together, the results demonstrate that: (i) LHC proteins form highly stable, disc-shaped aggregates; (ii) despite the possibly amorphous nature of the LHC aggregate core, it contains a defined global organization that can be reliably probed; and (iii) the L18 motif, among other regions of the LHC protein, is displayed on the surface of the aggregate and thus poised for interactions with cpSRP43 (Figure 2.9). These results suggest an attractive model in which cpSRP43 could recognize the L18 motif presented on the surface of the aggregate, providing a starting point for its action as a disaggregase. Further, the exposure of the N-terminus and the N-terminal end of TM3 suggests additional potential interaction sites with cpSRP43 during aggregate recognition.

## Discussion

The ability of cpSRP43 to prevent and reverse LHC protein aggregation demonstrates the diversity and capability of cellular chaperones and highlights a disaggregation mechanism that relies on binding interactions instead of external energy input. The robustness and simplicity of the cpSRP43-LHC disaggregase system provides an opportunity to unravel the mechanism by which a relatively small, ATP-independent chaperone can rescue insoluble protein aggregates. In this work, biophysical and biochemical analyses of the structure and energetics of the LHCP aggregate help define the capability of cpSRP43 as a protein disaggregase and suggest an attractive mechanism for how this chaperone recognizes the LHC protein aggregates to initiate the disaggregation reaction.

Using kinetic analyses, we previously showed that cpSRP43 can actively remodel and disassemble LHC aggregates [12]. To gauge the amount of energy cpSRP43 must overcome during disaggregation, here we examined the stability of the LHC aggregates. The results indicate that LHC aggregates are stable both kinetically and thermodynamically. First, extensive dilution of the aggregate did not lead to re-solubilization, suggesting that LHC aggregates, once formed, are kinetically stable. This is in contrast to the ‘salting out’ effect, in which protein precipitates are reversibly produced when the protein concentration exceeds the solubility limit [30]. Second, LHC aggregates are resistant to a variety of detergents, even up to 2% SDS, akin to highly stable fibrils and insoluble amyloid oligomers [31]. The stability of the LHC aggregate further supports the notion that its reversal requires the active participation of cpSRP43 to engage and disrupt the aggregate and showcases the capacity of this chaperone as a disaggregase. The morphology of the LHC aggregates bears resemblance to those of the soluble oligomeric intermediates that often precede amyloid fibril formation, which are disc-shaped, 9–25 nm in

diameter, and 2–3 nm in height [32,33]. Although earlier work tends to categorically describe these protein aggregates as ‘amorphous’, accumulating data suggest that there are nonetheless degrees of organization in some of these aggregates [34,35]. For instance, the folding intermediates of bovine growth hormone, phosphoglycerate kinase, P22 tailspike and coat proteins participate in specific intermolecular interactions in their aggregation pathways [36-38]. Likewise, although highly specific intermolecular interactions have not been detected in the LHC aggregates, more detailed analyses at the individual residue level provide convincing evidence that LHC aggregates have a defined ‘interior’ and ‘exterior’ that can be reliably probed, arguing against complete disorder in these aggregates.

What features of the LHC aggregate allow cpSRP43 to recognize it and initiate the disaggregation process? Answers to this question are central to understanding the mechanism of cpSRP43’s disaggregase activity. The results here strongly suggest that the formation of LHC aggregates is driven largely by hydrophobic collapse to bury its three TM domains. Importantly, we showed that the N-terminus of the LHC protein, the L18 motif, and the N-terminal segment of TM3 are displayed on the solvent-accessible surface when LHC proteins form aggregates. As the L18 motif is the primary recognition element for cpSRP43, its presentation on the exterior of the aggregate provides a very attractive mechanism by which cpSRP43 could recognize and anchor onto the aggregate to initiate the disassembly process (Figure 2.9). Conceivably, the N-terminal fragment of TM3 proximal to the L18 motif could also contribute to this initial recognition, as previous work has detected crosslinks between cpSRP43 and residues at the N-terminus of TM3 [39]. This and additional mutational studies suggest that TM3 is a likely candidate for cpSRP43 to initiate disruptions of the internal packing within the LHC aggregate (Jaru-Ampornpan *et al.*, accompanying manuscript).

Further, the L18-swap experiments show that the L18 motif is a dominant sequence element that has a strong tendency to be displayed on the surface of a protein aggregate, likely due to its relatively polar amino acid composition and high propensity for disorder (Figure 2.7). Considering that cpSRP43 has co-evolved with and is dedicated to the chaperoning of the LHC family of proteins, it is intriguing that the latter evolved a polar L18 recognition motif and made it accessible even when they form aggregates, which would enable cpSRP43 to readily recognize the aggregated LHC proteins. Although the physiological significance of cpSRP43's disaggregase activity remains to be directly established, this observation is consistent with the possibility that this activity is beneficial, as it would enable cpSRP43 to rescue aggregated, off-pathway intermediates during the targeting or insertion of its substrate proteins [40].

The mechanism of aggregate recognition proposed here for cpSRP43 is distinct from those proposed for ClpB/Hsp104, where exposed patches enriched in charged and hydrophobic amino acids are recognized by the disaggregases [41,42]. It can be speculated that a generalized mode of substrate recognition is optimal for Hsp70 and/or the AAA+ disaggregases, which must handle a broad range of substrates. In contrast, dedication of cpSRP43 to the LHC family of proteins allows them to adopt a more specific and effective mechanism, in which an exposed polar motif is used for recognition and ultimately enables the chaperone to gain access to the hydrophobic core. This mechanism of aggregate recognition could explain analogous disaggregase systems reported previously, such as the Mitochondria Import Stimulation Factor (MSF), whose ability to rescue aggregated mitochondrial precursor proteins also depends on the basic mitochondrial signal sequence that is likely displayed on the aggregate surface [43,44].

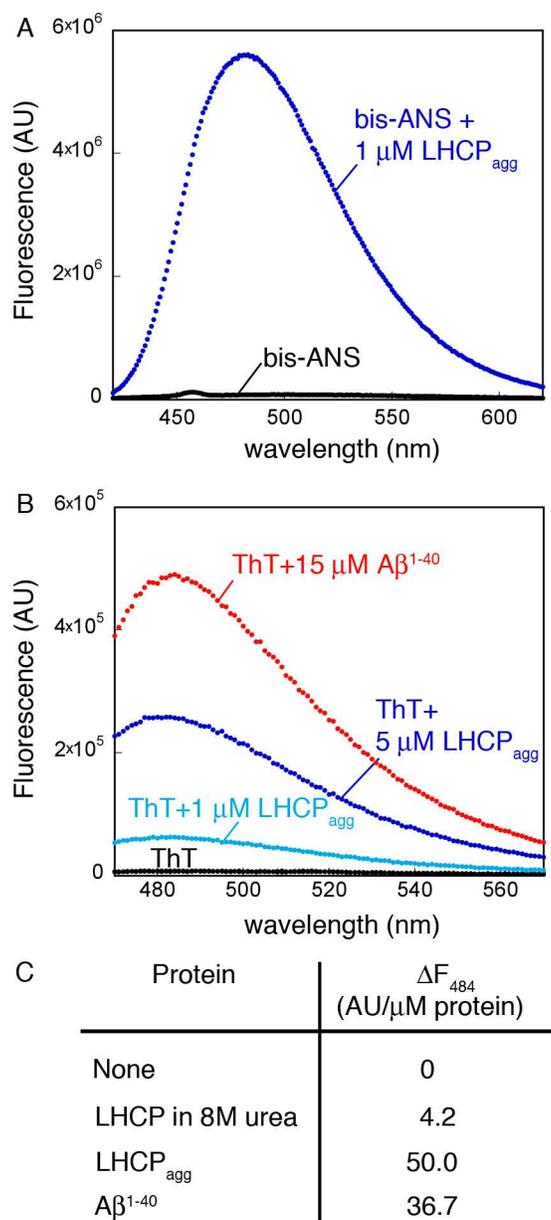
On the other hand, the L18-swap experiments here and additional mutagenesis studies (Jarupornpan *et al.*, accompanying manuscript) strongly suggest that the interaction of cpSRP43

with the remainder of the LHC protein, apart from L18, is highly adaptable, as a wide range of unnatural substrates can be effectively bound, chaperoned, and rescued by cpSRP43.

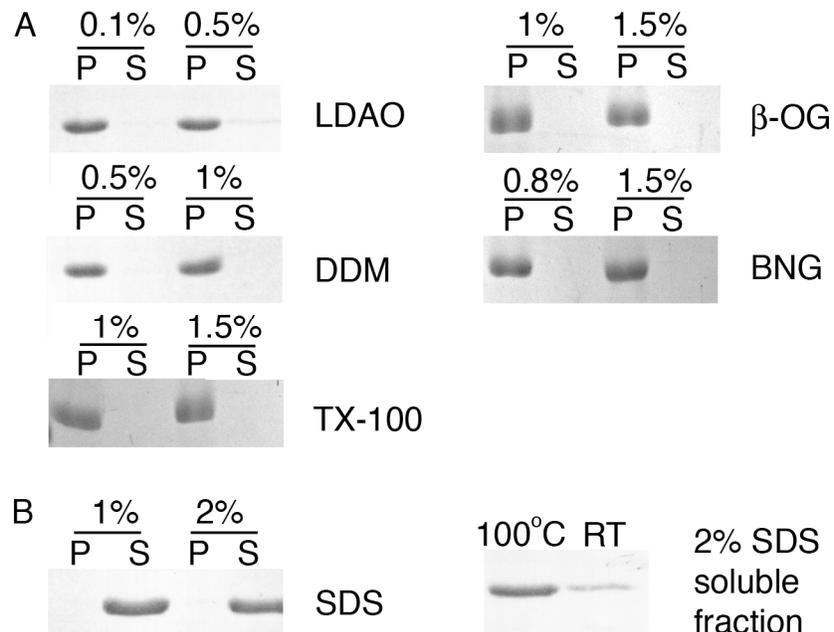
In summary, in-depth characterization of the nature and structure of the LHC protein aggregate suggest an attractive mechanism for its recognition by cpSRP43, and provide important constraints for the capability and limitation cpSRP43's disaggregase activity. In the accompanying paper (Jaru-Ampornpan *et al.*, accompanying manuscript), the lessons learned from this work are leveraged against structure-function analyses to propose a multi-step mechanism for the disaggregase reaction mediated by cpSRP43. These results provide a foundation for understanding the molecular basis of ATP-independent disaggregase systems, and guide the engineering of specific chaperone-substrate interactions for aggregates of similar nature.

**Acknowledgement**

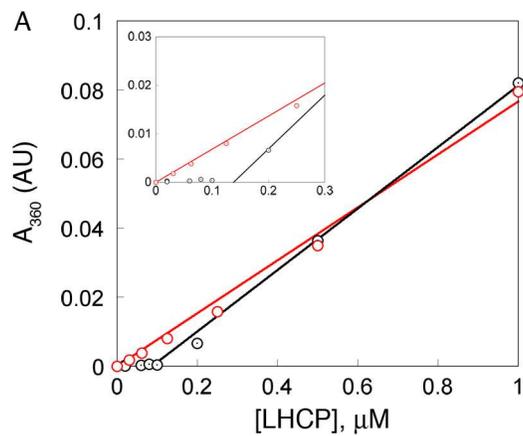
We thank A.N. Murray and Dr. J.W. Kelly for re-crystallized ThT and fibrilized A $\beta$ <sub>1-40</sub>, A. McDowall for TEM analyses, members of the Proteome Exploration Laboratory for mass spectrometry assistance, and members of the Shan group for helpful comments on the manuscript.



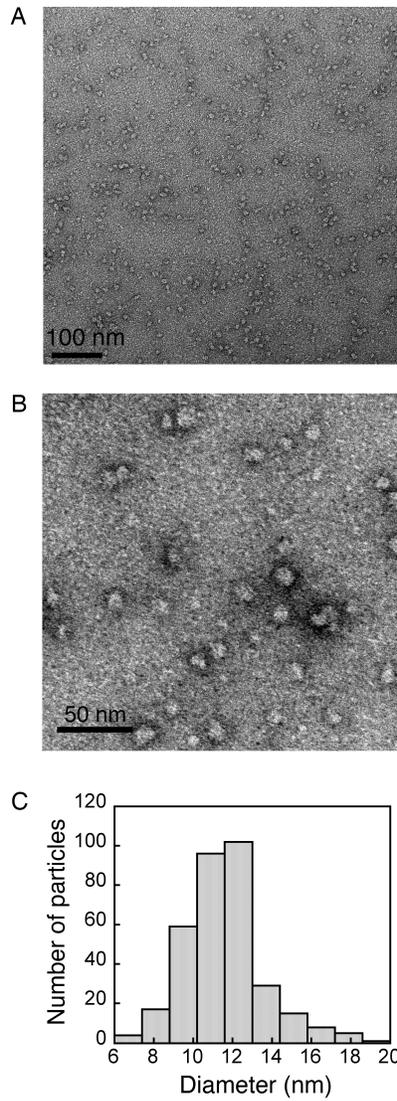
**Figure 2.1** LHCP aggregates contain exposed hydrophobic surfaces as detected by small molecule dyes. (A) Fluorescence emission spectra of 1 mM bis-ANS with (blue) or without (black) 1  $\mu\text{M}$  LHCP aggregate. (B) Fluorescence spectra of 20 mM ThT in the absence (black) and presence of 1 (light blue) or 5 (dark blue)  $\mu\text{M}$  LHCP aggregate, or 15  $\mu\text{M}$  A $\beta^{1-40}$  (red). (C) Quantification of the ThT fluorescence change at 484 nm per  $\mu\text{M}$  of protein (aggregate).



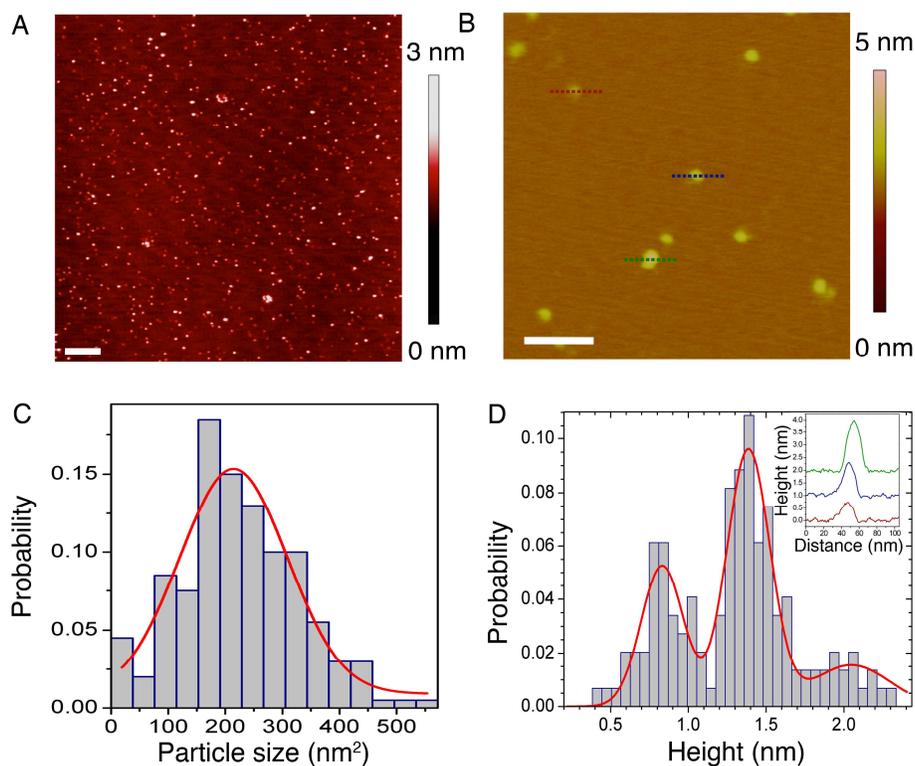
**Figure 2.2** LHCP aggregates are resistant to many detergents. (A) Sedimentation analysis of the ability of various detergents to resolubilize LHCP aggregates. CMC of LDAO, DDM, β-OG, BNG and TX-100 are 0.023%, 0.009%, 0.53%, 0.2% and 0.02%, respectively. P and S denote the pellet and soluble fractions, respectively. (B) SDS-solubility assay as described for amyloids [21] show partial solubility of LHCP aggregates in 2% SDS, right panel. The samples were directly loaded onto the gel and solubility was judged by the mobility of the protein into the resolving gel. Quantification using ImageJ revealed that 24% of the LHCP aggregates is soluble when the sample was not boiled ('RT'), compared to 87% for the boiled sample ('100 °C'). Left panel shows complete solubilization of LHCP by SDS (CMC 0.23%) after treatment as in (A).



**Figure 2.3** LHCP forms aggregates after a critical concentration. Light scattering intensities during formation of the aggregate (black) are compared with those from serial dilution of preformed aggregates (red). The inset highlights the lag phase at low concentrations during formation of the aggregate. AU, arbitrary units.



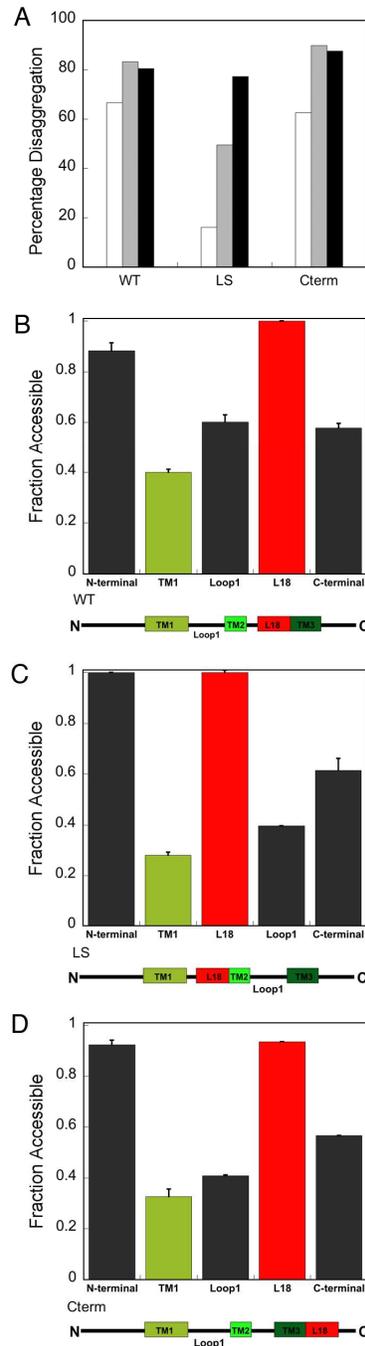
**Figure 2.4** TEM analysis of LHCP aggregates. (A) Large field view of a negatively stained TEM image of LHCP aggregates. (B) A zoomed-in image shows that LHCP aggregates are round particles. (C) Size distribution of the LHCP aggregates, measured from several independent experiments. The mean diameter is  $12 \pm 2$  nm.



**Figure 2.5** AFM analysis of LHCP aggregates. (A) Large field view of AFM topographic image showing well-separated LHCP aggregates. Large clusters are occasionally observed. The scale bar is 500 nm. (B) A zoomed-in region of the image reveals disc-shaped particles. The lines indicate particles whose heights were measured (red, blue, and green). The scale bar is 100 nm. (C) Size distribution of LHCP aggregates, measured from several regions on the surface. The red line is a Gaussian fit to the data, which gave a mean area of the particle of 214 nm<sup>2</sup>. (D) Height distribution of LHCP aggregates shows three populations of 0.8, 1.4, and 2.1 nm. The inset shows the height profiles for the representative particles indicated in B. Curves are vertically displaced for clarity.

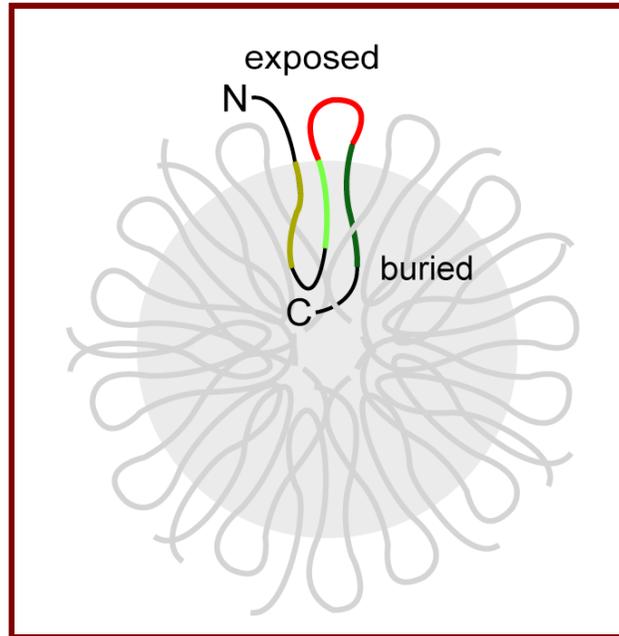


L130C and L170C in Lhcb5 aggregates and at residue L170C when Lhcb5 was dissolved in 6 M GdmHCl. (G) Summary of NEM accessibility of the single-site cysteines in the Lhcb5 aggregate. The color scheme is the same as in D.



**Figure 2.7** L18 has a strong tendency to be exposed on the surface of the aggregate. (A) The extent of wild-type and mutant Lhcb5 aggregates (1  $\mu$ M) resolubilized by 10 (white), 20 (gray), and 40 (black)  $\mu$ M cpSRP43. The aggregates formed by L18-swapped mutants TM1-L18-TM2-TM3 (LoopSwap (LS)) and TM1-TM2-TM3-L18 (C-terminal (Cterm)) can be rescued by cpSRP43, although the LoopSwap mutant required a higher concentration of cpSRP43. (B) NEM accessibility analysis of residues in the wild-type aggregates. Regions probed include the N terminus (G50C), TM1 (A80C), loop 1 (N120C), L18 (G162C), and C terminus (A230C). (C) and (D) NEM accessibility analysis of the same residues in the LoopSwap construct (C) and in the C-terminal construct (D). Error bars in all panels indicate S.D.





**Figure 2.9** Model for the global organization of LHC proteins in the aggregate. L18 is in red, the TMs are in different shades of green, and the other looping sequences are in black. The shaded region depicts the buried core of the aggregate. For clarity, only one LHCP molecule in the aggregate is highlighted.

**Supplementary Table 2.S1.** Inverse central linewidth ( $\Delta H^{-1}$ ) and fraction alkylated by NEM of individual Lhcb5 cysteine mutants upon aggregate formation in Buffer D as graphed in Figures 2.6C and 2.6E, respectively.

Residue	Fraction Alkylated	$\Delta H^{-1}$ (Gauss <sup>-1</sup> )
A10C	0.99 ± 0.02	0.342
K20C	0.90 ± 0.02	0.223
L30C	1.00 ± 0.00	0.250
I40C	1.00 ± 0.00	ND
G50C	0.88 ± 0.03	0.251
G60C	0.94 ± 0.01	0.139
Q70C	0.47 ± 0.02	0.123
I75C	0.49 ± 0.07	0.118
A80C	0.40 ± 0.01	ND
A85C	0.51 ± 0.04	0.112
P90C	0.29 ± 0.07	0.141
C100	0.74 ± 0.09	0.209
V105C	0.66 ± 0.00	ND
G110C	0.44 ± 0.02	0.125
N120C	0.60 ± 0.03	0.125
L130C	0.25 ± 0.02	0.147
V135C	0.22 ± 0.05	0.111
L140C	0.30 ± 0.04	0.112
T150C	0.66 ± 0.03	0.127
H160C	1.00 ± 0.00	0.272
D166C	0.87 ± 0.02	ND
P167C	0.84 ± 0.01	ND
L170C	0.96 ± 0.05	0.340
L180C	0.90 ± 0.01	0.241
A185C	0.95 ± 0.03	0.214
L190C	0.48 ± 0.08	0.116
M195C	0.19 ± 0.02	0.116
I200C	0.22 ± 0.02	0.112
V210C	0.81 ± 0.05	0.129
P220C	0.57 ± 0.05	0.132
A230C	0.58 ± 0.02	0.135



## References

1. Hartl, F. U., and Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **295**, 1852–1858
2. Balch, W. E., Morimoto, R.I., Dillin, A., and Kelly, J.W. (2008) Adapting proteostasis for disease intervention. *Science* **319**, 916–919
3. Doyle, S. M., and Wickner, S. (2008) Hsp104 and ClpB: protein disaggregating machines. *Trends Biochem. Sci.* **34**, 40–48
4. Weibezahn, J., Tessarz, P., Schlieker, C., Zahn, R., Maglica, Z., Lee, S., Zentgraf, H., Weber-Ban, E.U., Dougan, D.A., Tsai, F.T., Mogk, A., and Bukau, B. (2004) Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. *Cell* **119**, 653–665
5. Luheshi, L. M., and Dobson, C.M. (2009) Bridging the gap: from protein misfolding to protein misfolding diseases. *FEBS Lett.* **583**, 2581–2586
6. Broadley, S. A., and Hartl, F.U. (2009) The role of molecular chaperones in human misfolding diseases. *FEBS Lett.* **583**, 2647–2653
7. Powers, E. T., Morimoto, R.I., Dillin, A., Kelly, J.W., and Balch, W.E. (2009) Biological and chemical approaches to diseases of proteostasis deficiency. *Annu. Rev. Biochem.* **78**, 959–991
8. Doyle, S. M., Shorter, J., Zolkiewski, M., Hoskins, J.R., Lindquist, S., and Wickner, S. (2007) Asymmetric deceleration of ClpB or Hsp104 ATPase activity unleashes protein-remodeling activity. *Nat. Struct. Mol. Biol.* **14**, 114–122
9. Glover, J. R., and Lindquist, S. (1998) Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins. *Cell* **94**, 73–82
10. Duennwald, M. L., Echeverria, A., and Shorter, J. (2012) Small heat shock proteins potentiate amyloid dissolution by protein disaggregases from yeast and humans. *PLoS Biol.* **10**, e1001346
11. Rampelt, H., Kirstein-Miles, J., Nillegoda, N.B., Chi, K., Scholz, S.R., Morimoto, R.I., and Bukau, B. (2012) Metazoan Hsp70 machines use Hsp110 to power protein disaggregation. *EMBO J.* **31**, 4221–4235
12. Jaru-Ampornpan, P., Shen, K., Lam, V.Q., Ali, M., Doniach, S., Jia, T.Z., and Shan, S. (2010) ATP-independent reversal of a membrane protein aggregate by a chloroplast SRP subunit. *Nat. Struct. Mol. Biol.* **17**, 696–702
13. Schuenemann, D., Gupta, S., Persello-Cartieaux, F., Klimyuk, V.I., Jones, J.D., Nussaume, L., and Hoffman, N.E. (1998) A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10312–10316
14. Groves, M. R., Mant, A., Kuhn, A., Koch, J., Dubel, S., Robinson, C., and Sinning, I. (2001) Functional characterization of recombinant chloroplast signal recognition particle. *J. Biol. Chem.* **276**, 27778–27786
15. Tu, C.-J., Schuenemann, D., and Hoffman, N. E. (1999) 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.* **274**, 27219–27224
16. Falk, S., and Sinning, I. (2010) cpSRP43 is a novel chaperone specific for light-harvesting chlorophyll a,b-binding proteins. *J. Biol. Chem.* **285**, 21655–21661

17. Delille, J., Peterson, E.C., Johnson, T., Morre, M., Kight, A., and Henry, R. (2000) A novel precursor recognition element facilitates posttranslational binding to the signal recognition particle in chloroplasts. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1926–1931
18. Tu, C. J., Peterson, E.C., Henry, R., and Hoffman, N.E. (2000) The L18 domain of light-harvesting chlorophyll proteins binds to chloroplast signal recognition particle 43. *J. Biol. Chem.* **275**, 13187–13190
19. Hurshman, A.R., White, J.T., Powers, E.T., and Kelly J.W. (2004) Transthyretin aggregation under partially denaturing conditions is a downhill polymerization. *Biochemistry* **43**, 7365–7381
20. Abramoff, M. D., Magelhaes, P.J. and Ram, S.J. (2004) Image processing with ImageJ. *Biophotonics Int.* **11**, 36–42
21. Chernoff, Y. O., Uptain, S.M., and Lindquist, S.L. (2002) Analysis of prion factors in yeast. *Methods Enzymol.* **351**, 499–538
22. Lam, V. Q., Akopian, D., Rome, M., Henningsen, D., and Shan, S. (2010) Lipid activation of the signal recognition particle receptor provides spatial coordination of protein targeting. *J. Cel. Biol.* **190**, 623–635
23. Stryer, L. (1965) The interaction of a naphthalene dye with apomyoglobin and apohemoglobin. A fluorescent probe of non-polar binding sites. *J Mol Biol.* **13**, 482–495
24. Rosen, C. G., and Weber, G. (1969) Dimer formation from 1-anilino-8-naphthalenesulfonate catalyzed by bovine serum albumin. A new fluorescent molecule with exceptional binding properties. *Biochemistry* **8**, 3915–3920
25. LeVine, H. (1999) Quantification of beta-sheet amyloid fibril structures with thioflavin T. *Methods Enzymol.* **309**, 274–284
26. Carrotta, R., Bauer, R., Waninge, R., and Rischel, C. (2001) Conformational characterization of oligomeric intermediates and aggregates in  $\beta$ -lactoglobulin heat aggregation. *Protein Sci.* **10**, 1312–1318
27. Groenning, M., Olsen, L., van de Weert, M., Flink, J.M., Frokjaer, S., and Jørgensen, F.S. (2007) Study on the binding of Thioflavin T to beta-sheet-rich and non-beta-sheet cavities. *J. Struct. Biol.* **158**, 358–369
28. Jansson, S. (1999) A guide to the Lhc genes and their relatives in Arabidopsis. *Trends Plant Sci.* **4**, 236–240
29. Krishnan, R., and Lindquist, S. (2005) Structural insights into a yeast prion illuminate nucleation and strain diversity. *Nature* **435**, 765–772
30. Fink, A. L. (1998) Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Fold Des.* **3**, 9–23
31. Serio, T. R., Cashikar, A.G., Kowal, A.S., Sawicki, G.J., Moslehi, J.J., Serpell, L., Arnsdorf, M.F., and Lindquist, S.L. (2000) Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* **289**, 1317–1321
32. Goldsbury, C., Wirtz, S., Muller, S.A., Sunderji, S., Wicki, P., Aebi, U., and Frey, P. (2000) Studies on the in vitro assembly of Ab 1-40: Implications for the search for Ab fibril formation inhibitors. *J. Struct. Biol.* **130**, 217–231
33. Mastrangelo, I. A., Ahmed, M., Sato, T., Liu W., Wang, C., Hough, P., and Smith, S.O. (2006) High-resolution atomic force microscopy of Soluble Ab42 oligomers. *J. Mol. Biol.* **358**, 106–119

34. Zettlmeissl, G., Rudolph, R., and Jaenicke, R. (1979) Reconstitution of lactic dehydrogenase. Noncovalent aggregation vs. reactivation. 1. Physical properties and kinetics of aggregation. *Biochemistry* **18**, 5567–5571
35. Kopito, R. R., and Sitia, R. (2000) Aggresomes and Russell bodies. Symptoms of cellular indigestion? *EMBO Rep.* **1**, 225–231
36. Brems, D. N., Plaisted, S.M., Havel, H.A., and Tomich, C.S. (1988) Stabilization of an associated folding intermediate of bovine growth hormone by site-directed mutagenesis. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3367–3371
37. Mitraki, A., Betton, J.M., Desmadril, M., and Yon, J.M. (1987) Quasi-irreversibility in the unfolding-refolding transition of phosphoglycerate kinase induced by guanidine hydrochloride. *Eur. J. Biochem.* **163**, 29–34
38. Speed, M. A., Wang, D.I., and King, J. (1996) Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition. *Nat. Biotechnol.* **14**, 1283–1287
39. Cain, P., Hodlermann, I., Sinning, I., Johnson, A.E., and Robinson, C. (2011) Binding of chloroplast signal recognition particle to a thylakoid membrane protein substrate in aqueous solution and delineation of the cpSRP43-substrate interaction domain. *Biochem. J.* **437**, 149–155
40. Henry, R. (2010) SRP: adapting to life in the chloroplast. *Nat. Struct. Mol. Biol.* **17**, 676–678
41. Lum, R., Niggemann, M., and Glover, J.R. (2008) Peptide and protein binding in the axial channel of Hsp104. Insights into the mechanism of protein unfolding. *J. Biol. Chem.* **283**, 30139–30150
42. Schlieker, C., Weibezahn, J., Patzelt, H., Tessarz, P., Strub, C., Zeth, K., Erbse, A., Schneider-Mergener, J., Chin, J.W., Schultz, P.G., Bukau, B., and Mogk, A. (2004) Substrate recognition by the AAA+ chaperone ClpB. *Nat. Struct. Mol. Biol.* **11**, 607–615
43. Hachiya, N., Mihara, K., Suda, K., Horst, M., Schatz, G., and Lithgow, T. (1993) A mitochondrial import factor purified from rat liver cytosol is an ATP-dependent conformational modulator for precursor proteins. *EMBO J.* **12**, 1579–1586
44. Hachiya, N., Komiya, T., Alam, R., Iwahashi, J., Sakaguchi, M., Omura, T., and Mihara, K. (1994) MSF, a novel cytoplasmic chaperone which functions in precursor targeting to mitochondria. *EMBO J.* **13**, 5146–5154



# Chapter 3

## Mechanism of an ATP-Independent Protein

### Disaggregase.

## Part II. Distinct Molecular Interactions Drive Multiple Steps during Aggregate Disassembly

A version of this chapter has been published as

Jaru-Ampornpan, P., Liang, F.C., Nisthal, A., Nguyen, T.X., Wang, P., Shen, K., Mayo, S.L.,  
and Shan, S. (2013) *J. Biol. Chem.*, **288** (19), 13431-45.

**Abstract**

The ability of molecular chaperones to overcome the misfolding and aggregation of proteins is essential for the maintenance of proper protein homeostasis in all cells. Thus far, the best studied disaggregase systems are the Clp/Hsp100 family of “ATPases associated with various cellular activities” (AAA+) ATPases, which use mechanical forces powered by ATP hydrolysis to remodel protein aggregates. An alternative system to disassemble large protein aggregates is provided by the 38-kDa subunit of the chloroplast signal recognition particle (cpSRP43), which uses binding energy with its substrate proteins to drive disaggregation. The mechanism of this novel chaperone remains unclear. Here, molecular genetics and structure-activity analyses show that the action of cpSRP43 can be dissected into two steps with distinct molecular requirements: (i) initial recognition, during which cpSRP43 binds specifically to a recognition motif displayed on the surface of the aggregate; and (ii) aggregate remodeling, during which highly adaptable binding interactions of cpSRP43 with hydrophobic transmembrane domains of the substrate protein compete with the packing interactions within the aggregate. This establishes a useful framework to understand the molecular mechanism by which binding interactions from a molecular chaperone can be used to overcome protein aggregates in the absence of external energy input from ATP.

## Introduction

Protein homeostasis is vital to all cells and requires the proper production, folding, localization, assembly and degradation of all cellular proteins [1]. Crucial to the maintenance of protein homeostasis is an elaborate network of ‘molecular chaperones’ [2-4], which prevents the misfolding and aggregation of proteins by protecting exposed hydrophobic residues in non-native states or unstructured regions and, in some cases, actively promotes protein folding [2]. However, under stress conditions, the folding capacity of the chaperone network could be exceeded or impaired, leading to protein aggregation. A special set of chaperone machineries, the ‘disaggregases’, plays a crucial role in rescuing these detrimental processes. The best-studied disaggregases belong to the Hsp100 family: Hsp104 in yeast and ClpB in bacteria [5]. Both are members of the ‘ATPases associated with various cellular activities’ (AAA<sup>+</sup>)<sup>4</sup> superfamily that assemble into hexameric ring structures [5]. These disaggregases use repetitive ATPase cycles and, in collaboration with their co-chaperones, remodel large protein aggregates via translocation of the substrate protein through their central pores [6-8].

Despite the fascinating activity displayed by ClpB/Hsp104, their homologues have not been found beyond bacteria and yeast cells. Nevertheless, multiple lines of evidence indicate that maintenance of protein homeostasis in mammalian cells is critically dependent on cellular programs to overcome the deleterious effects of protein aggregation [9]. Recently, it was demonstrated that ATP-independent actions of the mammalian Hsp110 and small heat shock proteins can engage and facilitate the remodeling of protein aggregates in collaboration with Hsp70 and Hsp40 homologues [10,11]. These observations suggest that cells, especially higher eukaryotic cells, have evolved alternative strategies and mechanisms to rescue protein aggregates.

Recently, we described a novel disaggregase system that operates independently of ATP: the 38-kDa subunit of the chloroplast Signal Recognition Particle (cpSRP43) [12]. The substrates of this chaperone belong to the light-harvesting chlorophyll *a/b*-binding (LHC) family of proteins, which are delivered by the cpSRP from the chloroplast stroma to the thylakoid membrane [13]. The most abundant member of this family, LHCP, comprises up to 50% of the protein content in the thylakoid membrane and is likely the most abundant membrane protein on earth [13,14]. LHC proteins contain three hydrophobic transmembrane (TM) helices, making them highly prone to aggregation as they traverse aqueous compartments in the cell [14,15]. Recently, we and others showed that the cpSRP43 subunit of cpSRP acts as an effective molecular chaperone for the LHC proteins [12,16]. Intriguingly, cpSRP43 also efficiently reverses the aggregation of LHC proteins without the requirements for ATP hydrolysis or co-chaperones [12,16].

cpSRP43 provides a valuable example of a novel class of ATP-independent chaperones/disaggregases that operates with energy derived solely from binding interactions with its substrate proteins. Understanding its mechanism of action will provide valuable insights into alternative principles and approaches that can be used to overcome protein aggregation problems. An increasing number of examples speaks to the generality of this phenomenon. Mitochondrial import stimulation factor (MSF) reverses the aggregation of mitochondrial precursor proteins and restores their import in an ATP-independent mode [17,18]. Small heat shock proteins play crucial roles in remodeling protein aggregates and facilitate their resolubilization by Hsp70/100 [10,19]. Cyclophilins reactivate the aggregates of adenosine kinase [20]. ATP-independent disaggregase activities have also been reported in nematode and

mammalian tissue [9,21,22]. However, the mechanism by which protein aggregates can be disassembled based solely on a chaperone's substrate binding energy remains elusive.

Many questions arise in addressing these mechanisms. First, what are the precise binding interactions between cpSRP43 and its substrate proteins? Previous work has demonstrated a specific interaction of cpSRP43 with a highly conserved 18-amino acid loop, L18, between TM2 and TM3 of LHC proteins [23,24]. However, the ability of cpSRP43 to prevent LHC proteins from aggregation implies that it must also protect the hydrophobic TMs of the substrate protein. Consistent with this notion, the binding affinity between cpSRP43 and full-length LHCP is at least an order of magnitude higher than that for the L18 peptide [12,25]. Thus additional interactions most likely exist between LHCP and cpSRP43, but the nature of these interactions remains to be determined. Second, how does cpSRP43 use these binding interactions to effect the reversal of protein aggregation? Previous kinetic analyses revealed that disaggregation is a cooperative process during which multiple cpSRP43 molecules recognize and actively remodel the LHCP aggregate [12]. However, how the recognition and remodeling of the protein aggregate was accomplished by cpSRP43 has been elusive.

By combining molecular genetics with kinetic and thermodynamic analysis, here we present evidence that the interaction of cpSRP43 with its substrate proteins is comprised of two components: sequence-specific recognition of the L18 motif and highly promiscuous interactions with hydrophobic TMs. These interactions enable distinct steps in the cpSRP43-mediated disaggregation of LHC proteins: initial recognition and subsequent remodeling and disruption of the aggregate. The balance of these binding interactions with the energetics of packing interactions within the aggregate dictates the efficiency of the disaggregation reaction.

## Materials and Methods

**Materials.** To construct the LHCP TM mutants (Table 3.1), a pair of unique restriction sites was introduced into the expression plasmid encoding LHCP before and after the sequences encoding TM1, TM2, or TM3. The sequences coding for the TMs were replaced with PCR fragments encoding alternative TMs using the corresponding restriction sites. TM deletion mutants and Lhcb<sub>5</sub> cysteine mutants were constructed using the QuikChange mutagenesis procedure (Stratagene). cpSRP43, LHCP and its variants were purified as described [12].

**Determination of Binding Affinity between cpSRP43 and Soluble Protein Substrates.** Two independent methods were used to determine the apparent dissociation constant  $K_d^{app}$  for cpSRP43-substrate binding: (i) Prevention of LHCP aggregation by cpSRP43, monitored by light scattering at 360 nm after a 10-minute incubation of the substrate protein with varying concentrations of cpSRP43[12]. The light scattering is linearly proportional to the concentration of LHCP except at very low concentrations ([12] [26]). The percentage of soluble substrates (% soluble) was analyzed as a function of cpSRP43 concentration. The data were fit to eq 1,

$$\frac{[L]}{[L] + [43]} = \frac{[L]}{[L] + K_d^{app} \left( \frac{[L]}{[43]} \right)} \quad (1)$$

in which [L] is the LHC protein concentration and [43] is the cpSRP43 concentration. (ii) Fluorescence anisotropy, as described previously [12]. Briefly, all anisotropy measurements were conducted at room temperature using a Fluorolog 3-22 spectrofluorometer (Jobin Yvon). Fluorescein-labeled LHCP or its variants (50 nM) were diluted into buffer containing different concentrations of cpSRP43. The samples were excited at 450 nm and the fluorescence anisotropy was recorded at 524 nm. The data were fit to equation 2,

$$A_{obsd} = A_0 + \Delta A \frac{[cpSRP43]}{K_d + [cpSRP43]} \quad (2)$$

in which  $A_{obsd}$  is the observed anisotropy value,  $A_0$  is the anisotropy value without cpSRP43,  $\Delta A$  is the total change in anisotropy, and  $K_d$  is the equilibrium dissociation constant. The  $K_d$  values measured by these two methods produced consistent results for the substrates tested (Fig 2A).

### Thermodynamic and Kinetic Analyses of cpSRP43-Mediated Disaggregation.

Disaggregation reactions were performed as previously described [12], with the exception that aggregate formation was allowed to proceed for 1 minute before the addition of cpSRP43. The observed light scattering intensity was normalized to that prior to the addition of cpSRP43. The disaggregation time courses were fit to an exponential function 3,

$$A = A_f + \Delta A e^{-k_{obsd} t} \quad (3)$$

in which  $A$  is the observed light scattering,  $A_f$  is the amount of light scattering at  $t \rightarrow \infty$ ,  $\Delta A$  is the extent of light scattering change, and  $k_{obsd}$  is the observed rate constant. The fractions disaggregated ( $K$ ) were calculated as  $[\Delta A / (\Delta A + A_f)]$ . The cpSRP43 concentration dependences of the value of  $K$  were fit to eq 4,

$$K = K_{max} \frac{[cpSRP43]^h}{K_{1/2}^h + [cpSRP43]^h} \quad (4)$$

in which  $K_{max}$  is the extent of disaggregation at saturating cpSRP43 concentration,  $K_{1/2}$  is the concentration of cpSRP43 that enables 50% solubilization of the aggregates, and  $h$  is the Hill coefficient.

Kinetic analysis was performed and analyzed as described previously [12] to obtain the forward rate disaggregation rate constant,  $k_f$ , from the observed rate constants ( $k_{obsd}$ , Eq 3) and the extent of disaggregation ( $K$ ). The cpSRP43 concentration dependence of the  $k_f$  values was fit to eq 5,

$$k = k_0 + \frac{k_{\max} \langle K_m \rangle^h}{K_m + \langle K_m \rangle^h} \quad (5)$$

in which  $k_0$  is the rate of spontaneous LHCP disaggregation in the absence of the chaperone,  $\langle K_m \rangle$  is the concentration of cpSRP43 required to achieve half maximal disaggregation rate,  $h$  is the Hill coefficient, and  $k_{\max}$  is the disaggregation rate constant at saturating cpSRP43 concentration.

For some of the irreversible mutants (red, Table 3.4), neither the reaction equilibrium nor the kinetics showed detectable cooperative concentration dependences. Therefore, the data were fit to Michaelis-Menten equations (eqs 6 and 7),

$$K = K_{\max} \times \frac{[L]}{K_m + [L]} \quad (6)$$

$$k = k_{\max} \times \frac{[L]}{K_m + [L]} \quad (7)$$

Although direct evidence remains to be obtained, the following strongly suggests that the formation of an initial recognition complex between cpSRP43 and LHCP aggregates is fast compared to subsequent remodeling and disassembly of the aggregate. First, in all the binding experiments the cpSRP43-LHCP interaction is complete within the timescale of manual mixing ( $\leq 15$  s), much faster than the overall disaggregation rates. Second, given an affinity of  $\sim 2 \mu\text{M}$  for the cpSRP43-L18 motif interaction [12,25] and the typical range of macromolecular association rate constants ( $10^6$ – $10^8 \text{ M}^{-1}\text{s}^{-1}$ ), the dissociation rate constant of the recognition complex would be in the range of  $2 - 200 \text{ s}^{-1}$ , much faster than the overall disaggregation reaction. Together, these observations suggest that the remodeling and disassembly of the aggregate is the rate-limiting step in the disaggregation reaction. Therefore, the cpSRP43

concentration required to achieve half of the maximal rate of disaggregation,  $\langle K_m \rangle$ , provides an empirical measure for the average binding affinity of cpSRP43 to the aggregate.

**Determination of the Energetics of Aggregate Formation.** The energetics of packing interactions that drive aggregate formation were probed with sedimentation assay. Briefly, pre-formed LHCP aggregates (10  $\mu$ M) were re-solubilized by various concentrations of GdmHCl or urea for 30 minutes at 25 °C. The mixtures were centrifuged at 18,000 g for 30 minutes, and soluble (S) and pellet (P) fractions were visualized by SDS-PAGE. The intensity of the Coomassie Blue-stained bands for the pellet and soluble fractions were quantified using ImageJ [27]. The data were fit to a two-state model (eq 8) analogous to that for protein folding [28]

$$f(\text{soluble}) = \frac{1}{1 + e^{\frac{m(U_{50} - U)}{RT}}} \quad (8)$$

in which fraction soluble is calculated as  $[S/(S+P)]$ , R is the gas constant, and T is temperature,  $U_{50}$  is the urea concentration to achieve 50% solubilization, and  $m$  is a constant of proportionality.

**Mathematical Analyses.** Linear regression analysis was performed using Mathematica to identify a weighted linear combination of  $U_{50}$  and  $\ln K_d$  values that best reproduced the  $\ln k_{\max}$  values in Table 4. This was carried out by identifying the global minimum for the scoring function (eq 9):

$$f(\alpha, \beta, \gamma) = \sum (\ln k_{\max} - \alpha U_{50} - \beta \ln K_d + \gamma)^2 \quad (9)$$

*LHCP Scanning Mutagenesis* – All molecular biology manipulations, including site-directed mutagenesis, transformation, and plating, were performed on a Tecan Freedom EVO liquid-handling robot (Nisthal and Mayo, manuscript in preparation). Constructs were sequence verified and re-arrayed into master plates. These master plates served to inoculate 10 mL volumes of Instant TB autoinduction media in 24-well plates. After overnight expression at 37 °C, the hexahistidine-tagged proteins were purified under denaturing conditions. Cysteine scanning of

the L18 motif was carried out using QuikChange mutagenesis (Stratagene). Single cysteine mutants of LHCP were analyzed using the light scattering assay as described in the section:

*Determination of binding affinity between cpSRP43 and soluble protein substrates.*

**Plate-based Aggregation Prevention and Disaggregase Activity Assay.** All LHCP variants were normalized to 45  $\mu\text{M}$ , and then diluted to 5  $\mu\text{M}$  in 384-well plates by the liquid-handling robot. In the aggregation prevention assay, either no, equimolar (5  $\mu\text{M}$ ), or a 1:3 molar ratio (15  $\mu\text{M}$ ) of traditionally purified cpSRP43 was already present in each reaction well. The reaction was followed by absorbance at 360 nm and allowed to proceed for at least 20 minutes. When measuring disaggregase activity, the cpSRP43 concentration was raised to a 1:6 molar ratio (30  $\mu\text{M}$ ) and added  $\sim 1$  min after diluting the LHCP protein into aqueous buffer. Again, the reaction was followed for 20 minutes by measuring the absorbance at 360 nm. For both types of assays, the first time point was measured  $\sim 5$  min after mixing and the final time point was used for data analysis. Percent chaperone activity is defined as:

$$\frac{A_{360}(\text{mutant}) - A_{360}(\text{no cpSRP43})}{A_{360}(\text{wt}) - A_{360}(\text{no cpSRP43})} \times 100 \quad (10)$$

where the equimolar chaperone<sub>A360</sub> value can be substituted for the A<sub>360</sub> values of other chaperone conditions. Relative chaperone activity is then calculated by normalizing the percent chaperone activity to the wt LHCP value for each assay plate.

## Results

**Bipartite Interactions of cpSRP43 with Soluble LHCP.** To identify binding interactions of cpSRP43 with LHCP that are crucial to its chaperone and disaggregase activities, we performed exhaustive alanine scanning mutagenesis in LHCP and assayed the chaperone activity using automated protocols on a Tecan Freedom EVO liquid-handling robot (Nisthal & Mayo, manuscript in preparation). Residues in the conserved L18 sequence between TM2 and TM3 of LHCP were further mutated to glycine and lysine. We tested the mutational effects on cpSRP43's interaction with LHCP by measuring the ability of cpSRP43 to (i) bind and thus prevent the aggregation of LHCP (Figures 3.1A,B and 3.1D,E); and (ii) reverse existing LHCP aggregates (Figure 3.1C and 3.1F). Most mutations outside the L18 region, a conserved 18 amino acid sequence, result in modest to marginal effects on both the prevention and disaggregation activities of cpSRP43 (Figure 3.1A-C). On the other hand, single mutations of every residue in an FDPLGL motif in L18 had large deleterious effects (Figure 3.1D-F), indicating that this motif plays a crucial role in the ability of cpSRP43 to bind and chaperone LHCP. In contrast, mutations in the remainder of the LHCP had modest to marginal effects (Figure 3.1D-F). An independent cysteine mutagenesis scan of the L18 sequence yielded the same results (Figure 3.1G,H). These results extend previous studies [12,25] and together, they show that cpSRP43 makes highly sequence-specific interactions with the XDPLGX motif in the L18 sequence.

The absence of significant defects resulting from point mutations of the remainder of LHCP (Figure 3.1A-C) suggests that the interactions of cpSRP43 with the TMs of LHCP are likely promiscuous. To provide independent evidence for this notion and to further probe the nature of cpSRP43's interaction with the TMs of the substrate protein, we constructed LHCP variants in which the individual TMs are deleted or swapped. In addition, the TMs in LHCP were

replaced with those from unrelated membrane proteins, including the tail-anchored proteins SERP1, Sec61b, and cytochrome  $b_5$  (Table 3.1 for nomenclature and composition of all LHCP TM mutants used in this study). If the interactions of cpSRP43 with the TMs are sequence-specific, these mutations should significantly reduce the ability of cpSRP43 to bind and chaperone LHCP [12]. On the other hand, if these interactions arise from generic hydrophobic interactions or backbone contacts, these TM replacements should not substantially disrupt the chaperone activity. We quantitatively measured the binding interactions of the TM mutants with cpSRP43 using two independent approaches: (i) The ability of cpSRP43 to bind and thus prevent the aggregation of LHC proteins, which provides a convenient measure for the apparent binding affinity ( $K_d^{app}$ ) between this chaperone and the soluble LHCP; (ii) Equilibrium titrations based on changes in the fluorescence anisotropy of fluorescein-labeled LHCP upon its binding to cpSRP43 [12]. The values of  $K_d^{app}$  obtained from the two assays were comparable with one another (Figure 3.2A and [12]).

All the LHCP TM mutants tested could be efficiently bound and protected from aggregation by cpSRP43 (Figure 3.2B-F), with efficiencies that differ no more than five-fold from wild-type LHCP. Some mutants, such as  $\Delta$ TM3, SERP2 and Sec2, bound cpSRP43 with even higher affinity than wild-type LHCP and are hence more readily protected by this chaperone (Figure 3.2B-D, green). Collectively, all the TM replacement mutants exhibit moderate to high binding affinities for cpSRP43, which are 10–100 fold higher than that of cpSRP43 for the isolated L18 peptide [25]. This strongly suggests that the hydrophobic TMs contribute additional binding interactions with cpSRP43. Further, these interactions are fairly generic and highly adaptable, in contrast to the strictly sequence-specific interactions of the L18

motif. Finally, these results show that cpSRP43 can protect a variety of aggregation-prone proteins, as long as the L18 motif is present to provide specific recognition.

**A Quantitative Framework to Analyze cpSRP43-Mediated Disassembly of LHC Aggregates.** To understand how the binding interactions from cpSRP43 are used to drive the disassembly of LHC aggregates, it is crucial to establish a quantitative framework that describes the energetics of the individual steps of this reaction. The disaggregation reaction mediated by cpSRP43 can be studied under single turnover conditions [28], minimizing complications from multiple turnover and facilitating interpretation of data. Both the kinetics and equilibrium of this reaction exhibit saturable cooperative concentration dependences (Figure 3.3A,B; [12]), strongly suggesting that the reaction involves at least two steps: (i) a higher-order step dependent on cpSRP43 concentration, presumably the assembly of a ‘recognition complex’ between cpSRP43 and the aggregate (Figure 3.3C, step 1); followed by (ii) a unimolecular step independent of chaperone concentration, presumably involving the remodeling and disruption of the aggregate to generate resolubilized LHC•cpSRP43 complexes (Figure 3.3C, step 2). Important parameters can be extracted from these data to empirically report on the energetics of these steps (Figure 3.3 & Table 3.2). Assuming that the initial recognition step is fast compared to the subsequent remodeling steps (see Methods), the cpSRP43 concentration required to achieve half of the maximal disaggregation rate provides an empirical measure for the average binding affinity of cpSRP43 to the LHC aggregate (Figure 3.3A,C and Table 3.2,  $\langle K_m \rangle$ ). The Hill co-efficient,  $h$ , denotes the minimum number of cpSRP43 molecules that cooperatively act together to disrupt the aggregate (Figure 3.3C and Table 3.2). The maximal rate of disaggregation at saturating chaperone concentration,  $k_{\max}$ , measures the energetic barrier for remodeling and disrupting the aggregate once the initial recognition complex is formed (Figure 3.3A,C and Table 3.2). In

equilibrium measurements, the fraction of LHC proteins resolubilized at saturating cpSRP43 concentrations,  $K_{\max}$ , reports on the extent to which the interactions between LHC and cpSRP43 overcome the forces that stabilize the aggregate. Finally, at a sub-saturating cpSRP43 concentration, the observed kinetics and equilibrium of LHC resolubilization ( $k_{\text{app}}$  and  $K_{\text{app}}$ , respectively) measures the overall barrier to reach the transition state and the final cpSRP43-LHC complex, respectively.

To provide independent evidence that the disaggregation reaction can be experimentally dissected into distinct steps and to probe the molecular determinants that underlie each step, we characterized mutant cpSRP43 or LHC proteins that exhibit different defects in the disaggregation reaction. Below, we present evidence for two distinct classes of mutants that uncouple the initial recognition of the protein aggregate from its subsequent remodeling and solubilization, and for the distinct molecular determinants and interactions that underlie these steps.

**Interaction with the L18 motif is essential for initial recognition of the aggregate.** In the preceding paper [26], the results of both electron paramagnetic resonance (EPR) and chemical modification experiments showed that in LHC aggregates, the hydrophobic TMs are buried in the interior whereas the L18 motif is displayed on the exterior. These results suggest an attractive model in which cpSRP43 could recognize the L18 motif presented on the surface of the aggregate, initiating its action as a disaggregase. If this were the case, mutant LHC or cpSRP43s that specifically disrupt the L18-cpSRP43 interaction would impair the initial recognition of the aggregate, exhibiting defects in disaggregation at low chaperone concentrations. As binding is a higher-order process, the defects of these mutants could be overcome when a sufficiently high chaperone concentration is used to drive the initial binding.

To test this hypothesis, we examined how mutations in the L18 motif of LHC or in the L18-binding sites of cpSRP43 affect the efficiency of disaggregation.

We identified two mutations in the L18 motif of Lhcb<sub>5</sub> (a close homologue of LHCP), H160C and L170C, that weaken substrate binding with cpSRP43. Equilibrium binding assays showed that wild-type Lhcb<sub>5</sub> binds tightly to cpSRP43, with a  $K_d^{\text{app}}$  value of 10 nM, whereas mutants H160C and L170C exhibited weakened binding, with  $K_d^{\text{app}}$  values of 30 nM and 1.1 mM, respectively (Figure 3.4A and Table 3.3). Reciprocally, mutation of Arg161 in cpSRP43 (R161A), which provides an important hydrogen bond partner with L18 [25], significantly reduces the binding affinity of cpSRP43 to LHCP ( $K_d^{\text{app}} = 1.2$  mM, compared to 138 nM with wild-type cpSRP43; [12]).

Consistent with defects in recognition of the LHCP aggregate, mutant cpSRP43-R161A exhibited severe defects in the reversal of LHCP aggregates at low chaperone concentrations (Figure 3.4B, magenta vs. black; Table 3.3,  $K_{\text{app}}$  and  $k_{\text{app}}$ ). However, when the concentration of the mutant chaperone was raised to compensate for the binding defect, cpSRP43-R161A could reverse LHCP aggregation. At saturating chaperone concentrations, close to 50% solubilization of the aggregate could be attained (Figure 3.4B, magenta and Table 3.3). Analogously, the aggregates formed by the Lhcb<sub>5</sub> mutants, H160C and L170C, exhibited defects in the disaggregation reaction that can be rescued by higher cpSRP43 concentrations (Figure 3.4C,D and Table 3.3). At saturating chaperone concentrations, the equilibrium and kinetics of disaggregation with the mutant aggregates are within two-fold of those of wild-type Lhcb<sub>5</sub> (Figure 3.4C,D and Table 3.3). Finally, all three mutants exhibited much higher values of  $\langle K_m \rangle$  in the disaggregation reaction compared to the wild-type protein (Figure 3.4 and Table 3.3), which correlated with their reductions in substrate binding affinity (Figure 3.4A and [12]). Together,

these results showed that L18 binding is a key requirement for the initial recognition of the aggregate by cpSRP43; further, this recognition event can be uncoupled from the subsequent concentration-independent step(s) in the disaggregation reaction.

**A Class of LHCP TM Mutants Specifically Blocks the Disaggregation Process.** Since the LHCP TM mutants contain intact L18 motifs, they provide a collection of substrates to probe for additional molecular requirements that underlie cpSRP43's disaggregase activity. Surprisingly, although all the TM mutants can be efficiently bound and prevented from aggregation by cpSRP43 (Figure 3.2), they exhibit striking differences in the thermodynamics and kinetics of the disaggregation reaction (Figure 3.5, and Table 3.4).

The aggregates formed by some of the TM mutants, notably those of  $\Delta$ TM3, SERP2, Sec2 and Cyb2, showed disaggregation kinetics and efficiencies that are comparable to or even higher than that of wild-type LHCP (Figure 3.5A-H and Table 3.4, green). Notably, the aggregates formed by a group of mutants, especially 1-1-3,  $\Delta$ TM2, 1-2-2, and 1-3-2, were virtually irreversible even when saturation in disaggregation rate constants has been reached at high cpSRP43 concentrations (Figure 3.5A-H and Table 3.4, red). To a lesser extent, mutants 1-2-1 and  $\Delta$ TM1 also exhibited significant reductions in the disaggregation rates even when saturation was reached at high cpSRP43 concentrations (Figure 3.5 E-H and Table 3.4). In the aggregate formed by all these mutants, the L18 motif is highly accessible and solvent-exposed (Figure 3.6A-C); this and the observation that saturation in disaggregation kinetics can be reached with these mutants indicate that their defects could not be accounted for by the inability of cpSRP43 to recognize the aggregate. In addition to substantial reductions in the maximal disaggregation rates, the disaggregation reaction of these mutants lost cooperative dependence on cpSRP43, further supporting a specific defect in the ability of cpSRP43 to remodel and

resolubilize the aggregate. Together, these results provide strong evidence for the presence of an additional ‘remodeling’ step in cpSRP43’s disaggregase mechanism, whose molecular requirements are distinct from the initial recognition step.

**The Irreversible LHCP TM Mutants Form Ultra-stable Aggregates.** Unlike the L18-binding mutants, the irreversible LHCP TM mutants can bind reasonably well to cpSRP43. What caused their defects in disaggregation? The results from chemical modification and EPR experiments showed that the TM segments are buried inside the aggregate and engage in strong interactions [26]. We hypothesized that the internal packing interactions within the aggregates are altered in these TM mutants, which could present higher barriers for cpSRP43 to remodel and disrupt the aggregate. To test this hypothesis, we probed the energetics of the packing interactions that stabilize the aggregate by quantitatively analyzing its solubility in chemical denaturants. Using the sedimentation assay, we showed that both guanidinium hydrochloride (GdmHCl) and urea could effectively solubilize the LHCP aggregate in a concentration-dependent manner (Figure 3.7A). Quantification of the amount of solubilized LHCP as a function of urea concentration gave an aggregate solubilization curve analogous to protein unfolding curves (Figure 3.7B-E; [28]). Based on a two-state model, quantitative analyses of these data yielded information about the energetics of transfer of LHCP from urea to water ( $\Delta G^\circ$ ) and the urea concentration required to achieve 50% solubilization (Table 3.4,  $U_{50}$ ; see Methods). These parameters provide quantitative empirical measures of the energetics of the internal packing interactions that drive aggregate formation.

The aggregates formed by the LHCP TM mutants exhibited a wide range of stabilities, with  $U_{50}$  values ranging from 2.5 to 5.7 M (Figure 3.7B-E and 3.Table 4). Notably, the four ‘irreversible’ mutants that could bind cpSRP43 but whose aggregates could not be efficiently

resolubilized exhibited the highest  $U_{50}$  values (4.7–5.7 M; Figure 3.7B-E and Table 3.4, red). In contrast, some of the mutant aggregates that are more readily re-solubilized by cpSRP43, such as  $\Delta$ TM3, displayed the lowest  $U_{50}$  values (2.5–3.3 M; Figure 3.7B and Table 3.4, green). These results strongly suggest that the internal packing interactions within the aggregate provide a crucial barrier to the efficiency with which cpSRP43 can disrupt protein aggregates.

**Linear Free Energy Analysis to Probe the Energetic Determinants of Disaggregation Efficiency.** The collection of LHCP TM mutants, which display a wide range of disaggregation efficiencies and kinetics (Table 3.4), further allowed us to systematically probe the contributions of different molecular features and the nature of the rate-limiting remodeling complex (Figure 3.3C, ‡) during the disaggregation reaction. To this end, we evaluated how the maximal rate of disaggregation ( $k_{\max}$ ) correlates with the two energetic parameters that were varied in this set of mutants: (i) the binding affinity between cpSRP43 and the solubilized LHCP ( $K_d^{\text{app}}$ ), which ultimately drives the disaggregation reaction; and (ii) the energetics of packing interactions that drive aggregate formation ( $U_{50}$ ), which must be overcome by cpSRP43 during disaggregation. A strong correlation was found between the maximal disaggregation rate constant and a weighted combination of the  $U_{50}$  and  $K_d^{\text{app}}$  values (Figure 3.8A,  $R^2 = 0.96$ ), but not with either of the parameters alone (data not shown). This correlation strongly suggests that once a recognition complex is formed, the competition between the packing interactions that stabilize the aggregate and additional binding interactions that cpSRP43 establishes with the TMs of LHCP dictates the resolubilization of the aggregate. Finally, two of the mutants,  $\Delta$ TM1 and  $\Delta$ TM3, exhibit significant deviations from this correlation (Figure 3.8A, blue), suggesting preferences in the disaggregation pathway by cpSRP43 that are not accounted for by these two parameters (see the Discussion).

Analysis of the relationship between the equilibrium and rate constants of disaggregation provided further insights into the nature of the rate-limiting remodeling complex (Figure 3.3C, ‡). At a subsaturating cpSRP43 concentration (chosen at 4  $\mu\text{M}$ ), the free energy barrier ( $\Delta\Delta G \sim \ln K_{\text{app}}$ ) and the activation barrier ( $\Delta\Delta G^{\ddagger} \sim \ln k_{\text{app}}$ ) of the disaggregation reaction showed an excellent linear correlation for the entire set of LHCP mutants (Figure 3.8B), giving a slope of  $\approx 0.73$ . Analogous to the  $\phi$ -value analysis of protein folding [28], this correlation could be used to infer the nature and structure of the transition state *relative to* the substrate (the LHCP aggregate) and the product (the solubilized LHCP•cpSRP43 complex) of the reaction. The observation of an  $\phi$ -value that approaches unity implies a fairly late transition state and strongly suggests that, in the rate-limiting remodeling complex, a substantial fraction of packing interactions within the LHCP aggregate are disrupted and those with cpSRP43 are formed, albeit not to the same extent as those in the resolubilized cpSRP43-LHCP complex.

## Discussion

cpSRP43 provides an example of a novel class of chaperones that can effect the reversal of insoluble protein aggregates based solely on ATP-independent binding interactions with its substrate protein. The simplicity of this system makes it an accessible model system to delineate the molecular mechanisms as well as the capability and limitations of ATP-independent disaggregases. Here, mutational analyses revealed distinct sets of binding interactions that this chaperone establishes with its substrate proteins. Further, molecular genetics combined with thermodynamic and kinetic analyses allowed us to dissect the molecular steps during the cpSRP43-mediated disaggregation reaction and revealed distinct molecular requirements and interactions that underlie each step. These results, combined with previous work [12,26], led us to propose a two-step working model for the action of cpSRP43 as a protein disaggregase (Figure 3.9A).

**Bipartite interactions of cpSRP43 with substrate protein.** Previous work has established a specific interaction of cpSRP43 with L18, a conserved and relatively hydrophilic segment between TM2 and TM3 of the LHC family of proteins [23-25]. The mutagenesis results here further demonstrated that this interaction is localized to the most conserved FDPLGL motif within L18, emphasizing the highly specific nature of this recognition. This is consistent with crystallographic observations in which DPLG was found to form a ‘turn’ that wraps around Tyr204 of cpSRP43, whereas the side chains in the remainder of the L18 peptide were not well resolved [25]. Nevertheless, the ability of cpSRP43 to protect LHC proteins from aggregation implies that additional interactions must exist between this chaperone and the hydrophobic TMs on its substrate proteins. Although the precise motif(s) that mediate these additional interactions remain to be determined, the results here demonstrate that cpSRP43’s interactions with the TMs

are highly promiscuous, enabling it to bind and chaperone a variety of substrate variants in which the TMs were removed or replaced. Some of the substrate variants could be bound and chaperoned by cpSRP43 even more effectively than wild-type LHCP. Together, these results establish two important components of cpSRP43's binding interaction with substrate protein: highly specific recognition of the FDPLGL motif in the L18 segment, and generic hydrophobic interactions with the TMs of the substrate protein that are highly adaptable. As discussed below, these two sets of binding interactions contribute to distinct stages in cpSRP43's action as a disaggregase.

**Different binding interactions drive distinct stages of cpSRP43's disaggregase activity.**

*1. Recognition of the aggregate (Figure 3.9A, step 1).* To initiate the disaggregation reaction, cpSRP43 must first recognize and engage the LHC aggregates. Although additional interactions cannot be excluded, an attractive mechanism to drive this initial recognition is the binding of cpSRP43 to the L18 motif, which is displayed on the solvent-accessible exterior of the LHC aggregate (Figure 3.9A, step 1). In support of this model, mutant proteins that disrupt the interaction of cpSRP43 with the L18 motif exhibit defects in disaggregation at low cpSRP43 concentrations and require much higher chaperone concentrations to reach saturation (Figure 3.9B,  $\Delta\Delta G_1$ ). Consistent with a specific defect of these mutants in a binding step, their defects could be overcome by increasing the chaperone concentration, such that the maximal rate and efficiency of the disaggregation reaction with these mutants are within two-fold of that of the wild-type protein (Figure 3.9B).

Intriguingly, the values of  $\langle K_m \rangle$ , which provide a proxy for the binding of cpSRP43 to the LHCP aggregate, are considerably weaker than the binding of cpSRP43 to the L18 peptide [25]

and, with the exception of  $\Delta$ TM1, vary from  $\sim 6$  to  $16 \mu\text{M}$  across the different TM mutants (Table 3.4, white and green). On the one hand, this variation is much smaller than the up to 50-fold changes in the binding affinity between cpSRP43 and the soluble substrate protein (Table 3.4,  $K_d^{\text{app}}$ ), supporting the notion that interaction with the L18 motif is a major driving force for the initial recognition step and is less sensitive to variations in the TM segments of LHCP. On the other hand, such variation, though modest, could not be explained by the simplest model in which the recognition step is equivalent to interaction of cpSRP43 with an isolated L18 peptide. It is possible that the L18 motif is presented in different configurations on the aggregates formed by the different TM mutants, which could cause the observed variations. Consistent with this possibility, EPR experiments showed that in the aggregate, spin probes in the DPLG motif exhibit much lower mobility than the remainder of the L18 sequence [26], suggesting that this motif might contact the remainder of the aggregate and needs to undergo a rearrangement in order to interact with cpSRP43. Additional structural or sequence elements presented on the aggregate surface could also be recognized by cpSRP43. Consistent with this possibility, cpSRP43 crosslinks to residues at the N-terminus of TM3 [29]; this segment is also exposed on the surface of the LHC aggregate [26] and available for recognition by cpSRP43.

2. *Remodeling and disruption of the aggregate (Figure 3.9A, step 2)*. The class of ‘irreversible’ TM mutants (Table 3.4, red), which exhibits severe defects in maximal disaggregation rate constants ( $k_{\text{max}}$ ), provides strong evidence for a distinct remodeling step in the disaggregation reaction (Figure 3.9A, step 2) that has different molecular and energetic requirements to the initial recognition step. As cpSRP43 effectively prevents the aggregation of these mutant LHCPs, the defects of these mutants in disaggregation are most likely kinetic, rather than thermodynamic in origin. Further, the observation that all the ‘irreversible’ mutants

form significantly more stable aggregates than wild-type LHCP (Figure 3.9C,  $\Delta\Delta G_{\text{agg}}$ ) strongly suggests that the packing interactions within the aggregate present a major barrier for disaggregation (Figure 3.9C,  $k_{\text{max}}' \gg k_{\text{max}}$ ), and that these packing interactions need to be substantially disrupted in the rate-limiting remodeling complex (Figure 3.8A, species in bracket).

Additional insights into the remodeling step are provided by analyses of the entire series of TM mutants, which display a wide range of binding interactions with cpSRP43, packing interactions within the aggregate, and kinetics of disaggregation. In this series of mutants, the best predictor for disaggregation kinetics is provided by a combination of two energetic parameters: the packing interactions within the aggregate ( $U_{50}$ ) compensated by the available binding interactions between cpSRP43 and solubilized LHCP ( $K_d^{\text{app}}$ ; Figure 3.8A). This correlation is striking, and implies that the transition state (or rate-limiting remodeling complex) for the disaggregation reaction involves substantial global disruption of the aggregate. Further, these disruptions are compensated by the establishment of additional binding interactions of cpSRP43 with the TMs of the dislodged LHCP molecules. A notable example of the latter is SERP2, which forms an aggregate with a  $U_{50}$  value comparable to wild-type LHCP, 1-3-3 or Cyb2, but displays the fastest disaggregation kinetics as cpSRP43 establishes the strongest binding interactions with this mutant. Together, these results strongly support the model that, once cpSRP43 recognizes and ‘latches’ onto the LHCP aggregate, the competition between its binding interactions with the TM segments of LHCP molecules and the packing interactions that stabilize the aggregate dictates the efficiency of the disaggregation process.

□ value analysis, which compares the extent to which mutations affect the barrier to reach the transition state vs. that to the resolubilized cpSRP43-LHCP complex, provided additional insights into the nature of the rate-limiting remodeling complex during the

disaggregation process. The  $\alpha$ -value of 0.73 observed here rules out early ( $\alpha \sim 0$ ) transition states and suggests a fairly late structure for the rate-limiting remodeling intermediate [28], in which a substantial portion of the packing interactions within the LHCP aggregate is disrupted and significant binding interactions with cpSRP43 have been established. A slightly alternative model, which takes into account potential heterogeneity in the action of cpSRP43, is that cpSRP43 disrupts the packing interactions at certain parts of the aggregate more extensively than at others, giving rise to an  $\alpha$ -value less than unity. As formation of the aggregate is a highly cooperative process [26], it is conceivable that extensive disruption at multiple parts of the LHC aggregate could lead to the collapse of the network of packing interactions that drive aggregate formation (Figure 3.9A), thus leading to its solubilization.

**Perspective.** The analyses here established two key requirements for how a chaperone can use binding interactions to reverse a protein aggregate. First, the chaperone must efficiently recognize and latch onto the target aggregate, through interactions with structural or sequence motifs displayed on the exterior of the aggregate. Second, the chaperone must effectively compete with and replace the internal packing interactions of the aggregate, by interacting with and protecting the segments of the substrate protein buried in the interior of the aggregate. Although the specifics of each system differ, these principles may be generalized to other ATP-independent chaperones that participate in or facilitate protein disaggregation processes.

Nevertheless, these do not represent all the molecular features in the cpSRP43-mediated disaggregation reaction. Most notably, mutants  $\Delta$ TM1 and  $\Delta$ TM3 are clear outliers in the correlation analysis (Figure 3.8A, blue squares). These deviations imply that factors in addition to the two parameters explored here ( $U_{50}$  and  $K_d$ ) contribute to the disaggregation reaction, and suggest preferred pathways in the action of cpSRP43. For example,  $\Delta$ TM3 forms the loosest

aggregate and can bind cpSRP43 tightly, but its rate of disaggregation is significantly slower than that expected from these considerations. Coupled with the observation that N-terminal residues of TM3 are also highly accessible on the aggregate [26] and can contact cpSRP43 [29], this raises the intriguing possibility that cpSRP43 preferentially exerts its action on TM3 during the remodeling process to most effectively disrupt the aggregate. In contrast,  $\Delta$ TM1 forms one of the tightest aggregates and has weaker binding interactions with cpSRP43, yet its maximal rate of disaggregation by cpSRP43 far exceeded what would be expected based on these parameters. This led us to speculate that TM1 is not a preferred site of action of cpSRP43 during the disaggregation process. Finally, despite these deviations, the aggregates formed by both of these mutants and by ‘hybrid’ substrates containing TMs from unrelated membrane proteins are efficiently reversed by cpSRP43, demonstrating the remarkable adaptability of this chaperone.

It is noteworthy to compare the mechanism of aggregate remodeling by cpSRP43 to that of the ClpB/Hsp104 disaggregases. Much of the insights into the action of ClpB/Hsp104 are based on analogy with the ClpAP/ClpXP proteases [8], which use cycles of ATP binding and hydrolysis to drive repetitive movements of the substrate binding loops, forcing the polypeptides through a constricted pore in the hexameric assembly and thus unfolding the substrate protein. By analogy, ClpB/Hsp104 could use ATPase cycles to drive translocation of a polypeptide, extracting it out of protein aggregates [30,31]. In this mechanism, each disaggregase machine can locally sever an aggregate without disrupting the remainder of the aggregate. This is consistent with the observation that local, rather than global, structure and stability near the recognition sites dictate the efficiency of ClpA/ClpX [32,33], and with the ability of Hsp104 to generate more amyloid fragments and thus promote amyloid propagation [34,35]. Although the precise molecular details remain to be elucidated, our results suggest that cpSRP43 acts globally,

rather than locally, on the protein aggregate. The rate-limiting step in the reaction pathway of cpSRP43 involves the generation of a late intermediate in which the packing interactions within the entire aggregate are extensively disrupted and which requires the cooperative action of multiple cpSRP43 molecules. Conceivably, in the absence of external energy input, individual cpSRP43 molecules cannot compete with the packing interactions inside the aggregate and extract a soluble LHC molecule from it. Instead, multiple chaperones collectively disrupt and collapse the entire aggregate (Figure 3.9A). The results here provide a valuable framework to probe the capability, effectiveness, and limitations of this alternative ATP-independent chaperone mechanism, and to understand the design principles by which binding energy can be used to overcome the problems of protein aggregation.

**Acknowledgement**

We would like to thank Drs. W.M. Clemons, J. Chartron, C. Suloway for the plasmids of SERP1, Sec61 $\beta$  and cytochrome b<sub>5</sub> and the Shan lab for helpful comments on the manuscript.

**TABLE 3.1** Description of the LHCP TM mutants.

MRKSATTKKV ASSGSPWYGP DRVKYLGPFSS GESPSYLTGE FPGDYGWDTA GLSADPETFS<sub>60</sub>  
 KNRELEVIHS RWAMLGALGC VFPELLSRNG VKFGEAVWFK AGSQIFSEGG LDYLGNP<sub>120</sub>  
 HAQSILAIWA TM1 TQVILMGAVE GYRIAGGPLG L18 EVVDPLYPGG SFDPLGLADD PEAFaelkV<sub>180</sub>  
 ELKNGRLAMF TM2 SMFGFFVQAI VTGKGPLENL ADHLADPVNN NAWSYATNFV PGK<sub>233</sub>  
TM3

Construct	LHCP TM Replaced	Replaced by TM from	Sequence of the TM Replacement
WT	N/A	N/A	-
ΔTM1	TM1	-	-
ΔTM2	TM2	-	-
ΔTM3	TM3	-	-
1-1-3	TM2	LHCP TM1	PETFSKNRELEVIHSRWAMLGALGCVFPELLSRNG
1-3-3	TM2	LHCP TM3	PEAFaelkVKELKNGRLAMFSGFFVQAI
SERP2	TM2	SERP1	ASVGPWLLALFIFVCGSAIF
Sec2	TM2	Sec61b	VPVLVMSLLFIASVFM
Cyb2	TM2	Cytochrome b5	NSSWWTNWVIPAISALIVALMY
1-2-1	TM3	LHCP TM1	PETFSKNRELEVIHSRWAMLGALGCVFPELLSRNG
1-2-2	TM3	LHCP TM2	SILAIWATQVILMGAVEGYRIA
SERP3	TM3	SERP1	ASVGPWLLALFIFVCGSAIF
1-3-2	TM2, TM3	LHCP TM3, LHCP TM2	PEAFaelkVKELKNGRLAMFSGFFVQAI SILAIWATQVILMGAVEGYRIA

**TABLE 3.2** Thermodynamic and kinetic parameters in the disaggregation reaction.

<b>Parameter</b>	<b>Definition</b>	<b>Assay</b>	<b>Equation<sup>1</sup></b>
$\langle K_m \rangle$	cpSRP43 concentration that achieves half maximal rate of disaggregation	Disaggregation	<b>5</b>
<b>h</b>	Hill coefficient	Disaggregation	<b>5</b>
$k_{\max}$	Maximal disaggregation rate constant at saturating cpSRP43 concentration	Disaggregation	<b>5</b>
$K_{\max}$	Maximal fraction disaggregated at saturating cpSRP43 concentration	Disaggregation	<b>4</b>
$k_{\text{app}}$	Rate constant of disaggregation at a sub-saturating cpSRP43 concentration	Disaggregation	<b>5</b>
$K_{\text{app}}$	Fraction disaggregated at a sub-saturating cpSRP43 concentration	Disaggregation	<b>4</b>
$K_d^{\text{app}}$	Apparent dissociation constant of the soluble cpSRP43•LHCP complex	1. Prevention of aggregation 2. Fluorescence anisotropy	<b>1, 2</b>
$U_{50}$	Urea concentration required for 50% re-solubilization of the aggregate	Sedimentation	<b>8</b>

<sup>1</sup>Reference to equations in Methods.

**TABLE 3.3** Summary of the thermodynamic and kinetic parameters of the L18-binding mutants.

<b>Construct</b>	$K_d^{\text{app}}$ (nM)	$\langle K_m \rangle$ ( $\mu\text{M}$ )	$K_{\text{max}}$	$K_{\text{app}}^{4\mu\text{M}}$	$k_{\text{max}}$ ( $\text{s}^{-1}$ )	$k_{\text{app}}^{4\mu\text{M}}$ ( $\text{s}^{-1}$ )
cpSRP43 R161A	1200 <sup>1</sup>	>50 <sup>2</sup>	>0.40 <sup>2</sup>	0.02	N.D.	N.D.
Lhcb <sub>5</sub>	10	8.8	1.06	0.51	0.029	0.0042
Lhcb <sub>5</sub> H160C	30	64	0.85	0.12	0.025	0.0021
Lhcb <sub>5</sub> L170C	1100	>90 <sup>2</sup>	>0.35 <sup>2</sup>	0.03	>0.014	N.D.

Values reported are from Figure 3.1. N.D. = not determined.

<sup>1</sup>previously determined by fluorescence anisotropy [12].

<sup>2</sup>denotes the values at the highest cpSRP43 concentration used.

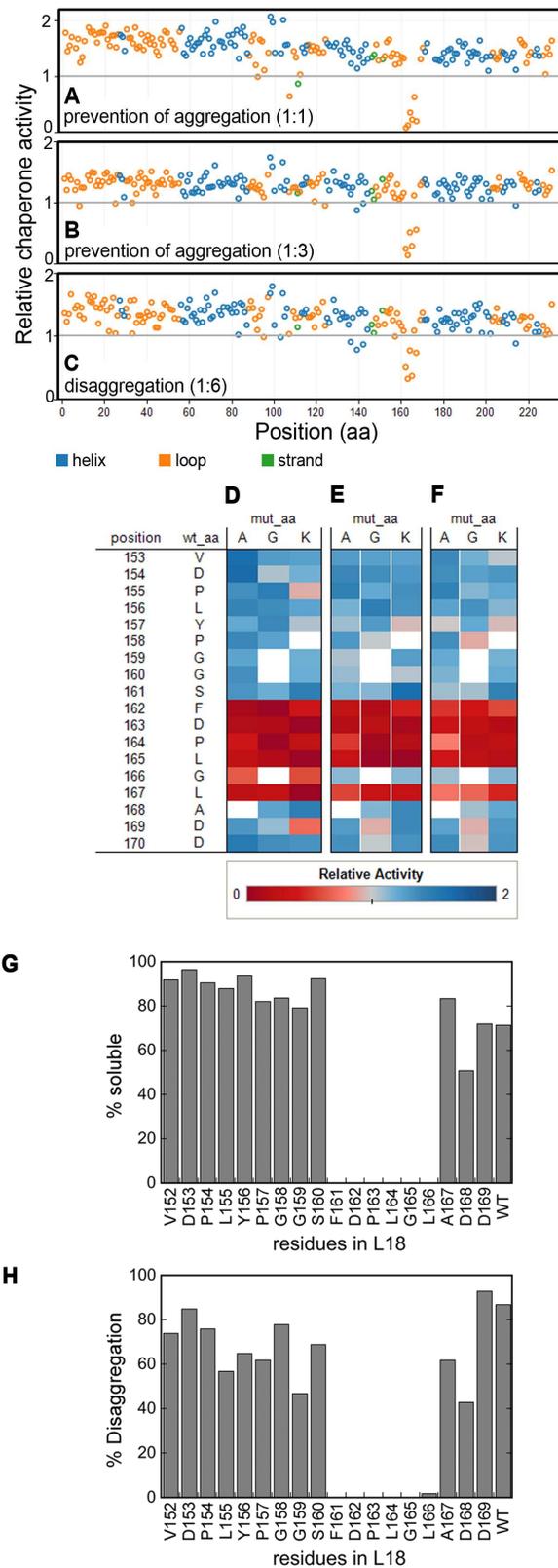
**TABLE 3.4** Summary of the thermodynamic and kinetic parameters of the LHCP TM mutants.

Substrate	$K_{\max}$	$K_{\text{app}}^{4\mu\text{M}}$	$\langle K_m \rangle (\mu\text{M})$	$h$	$k_{\max} (\text{s}^{-1})$	$k_{\text{app}}^{4\mu\text{M}} (\text{s}^{-1})$	$K_d^{\text{app}} (\text{nM})$	$U_{50} (\text{M})$
WT	0.98±0.02	0.70±0.05	8.8±4.1	2.9±0.5	0.049±0.005	0.0068±0.001	111±3	3.8±0.2
1-3-3	1.01±0.04	0.72±0.08	11.7±1.1	1.8±0.1	0.068±0.006	0.0090±0.001	144±34	3.7±0.1
SERP3	1.19±0.10	0.33±0.03	12.2±1.9	2.6±0.4	0.025±0.005	0.0021±0.0003	207±51	4.0±0.1
1-2-1	0.95±0.01	0.08±0.01	15.6±0.3	4.6±0.4	0.0099±0.002	0.0011±0.0001	234±16	4.4±0.3
$\Delta\text{TM1}$	0.97±0.02	0.13±0.04	42.5±15	1.6±0.5	0.053±0.021	0.0017±0.0002	413±76	4.7±0.0
SERP2	1.05±0.03	0.96±0.04	8.5±1.3	2.3±0.2	0.25 ±0.04	0.042±0.009	9±5	3.5±0.1
$\Delta\text{TM3}$	1.01±0.02	0.95±0.03	5.7±0.7	2.4±0.3	0.11 ±0.013	0.033±0.009	26±12	2.5±0.1
Sec2	1.09±0.02	0.81±0.01	6.6±0.5	2.9±0.3	0.088±0.008	0.017±0.005	36±17	3.3±0.1
Cyb2	1.12±0.11	0.66±0.04	6.8±0.4	3.6±0.5	0.056±0.004	0.0086±0.001	51±24	3.6±0.2
$\Delta\text{TM2}^{\#}$	0.62	0.04	N.D.	N/A	0.003	0.0004	216±88	5.7±0.1
1-2-2 <sup>#</sup>	N.D.	0.03	N.D.	N/A	0.006	0.0002	489±95	4.7±0.1
1-3-2 <sup>#</sup>	N.D.	0.03	N.D.	N/A	0.005	0.0003	456±206	4.8±0.1
1-1-3 <sup>#</sup>	0.56	0.02	N.D.	N/A	0.003	0.0003	490±57	5.7±0.1

N.D. =not determined. Values reported are average from two or more independent experiments  $\pm$  standard deviation.

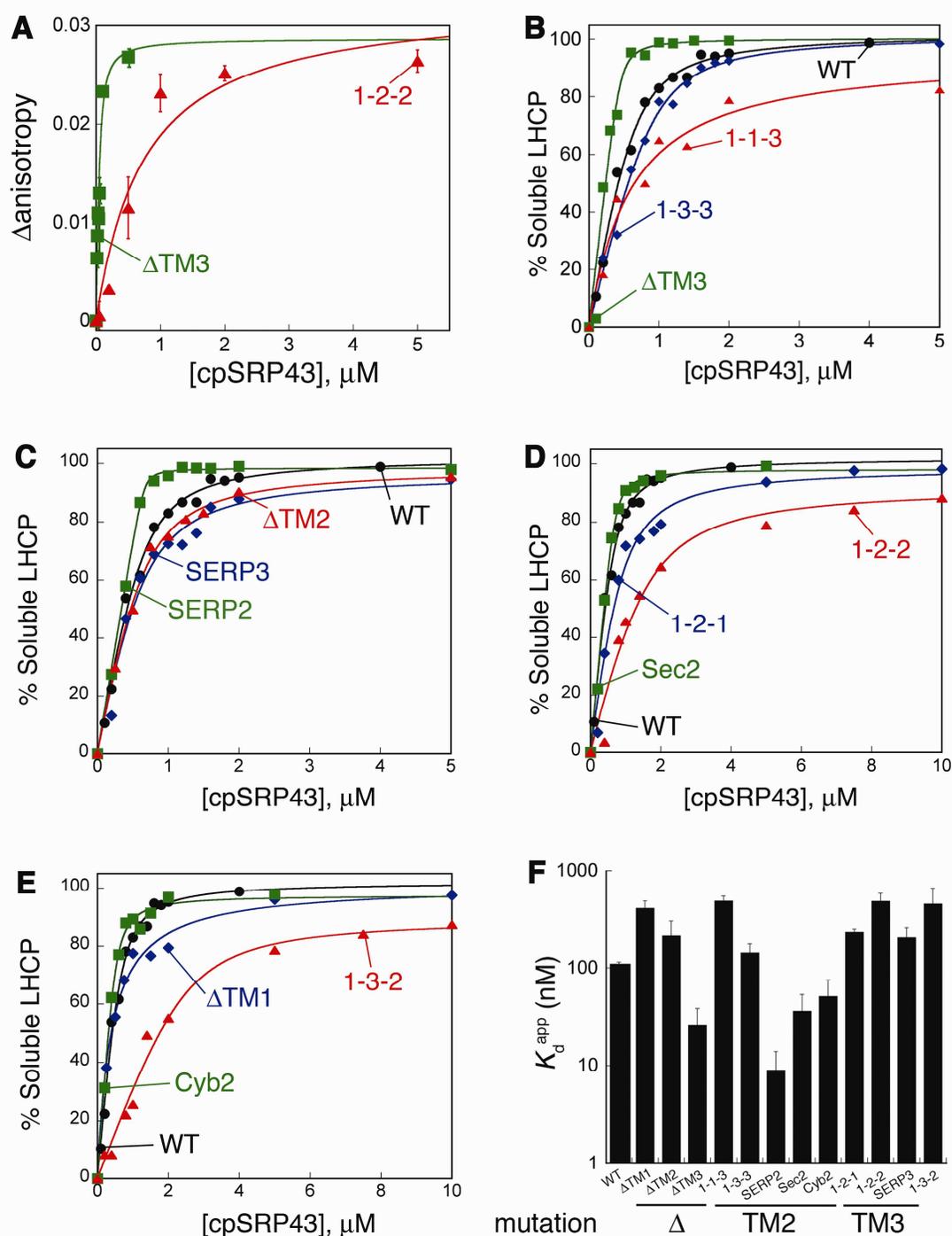
<sup>#</sup> denotes mutants that are fit with the Michaelis-Menten equation (eqs 6 & 7 in Methods).  $k_{\max}$  values were estimated for these mutants as the observed disaggregation rate constants reached saturation at the highest cpSRP43 concentrations. Accurate  $\langle K_m \rangle$  values could not be determined due to the extremely slow reaction of these mutants at low cpSRP43 concentrations, and are hence not reported.

N/A = not applicable.



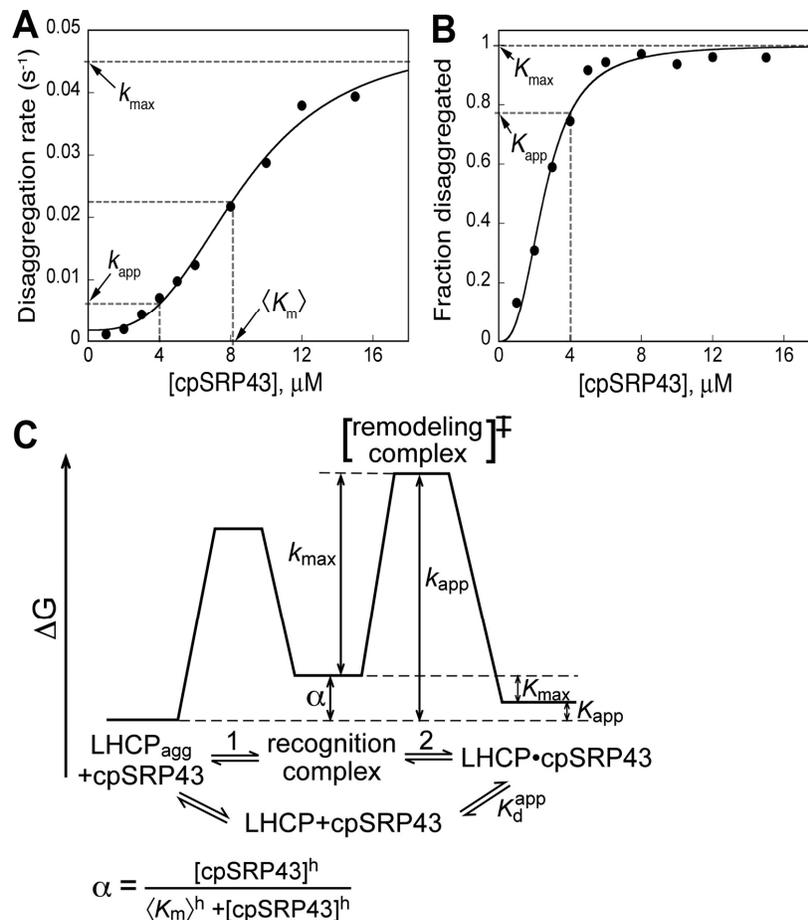
**Figure 3.1** cpSRP43 makes highly sequence-specific interactions with the FDPLGL motif in the L18 sequence. A–F, alanine-scanning mutagenesis of the entire LHCP (A–C), and alanine-,

glycine-, and lysine-scanning mutagenesis within the L18 sequence of LHCP (D–F). The aggregation prevention activity of cpSRP43 was assayed at 1:1 (A and D) and 1:3 (B and E) molar ratios of LHCP to cpSRP43. aa, amino acids; mut, mutant. In C and F, the disaggregase activity was measured at 1:6 molar ratio of LHCP to cpSRP43. All assays were performed in 384-well plates using a Tecan Freedom EVO liquid-handling robot, as described under “Experimental Procedures.” G and H, single-cysteine substitutions at individual residues in L18 were tested for their ability to prevent the aggregation of LHCP (G) and to resolubilize existing LHCP aggregates (H). In G, a 1:1 ratio of cpSRP43 and LHCP was used. In H, a 5:1 ratio of cpSRP43 relative to LHCP (in aggregates) was used.

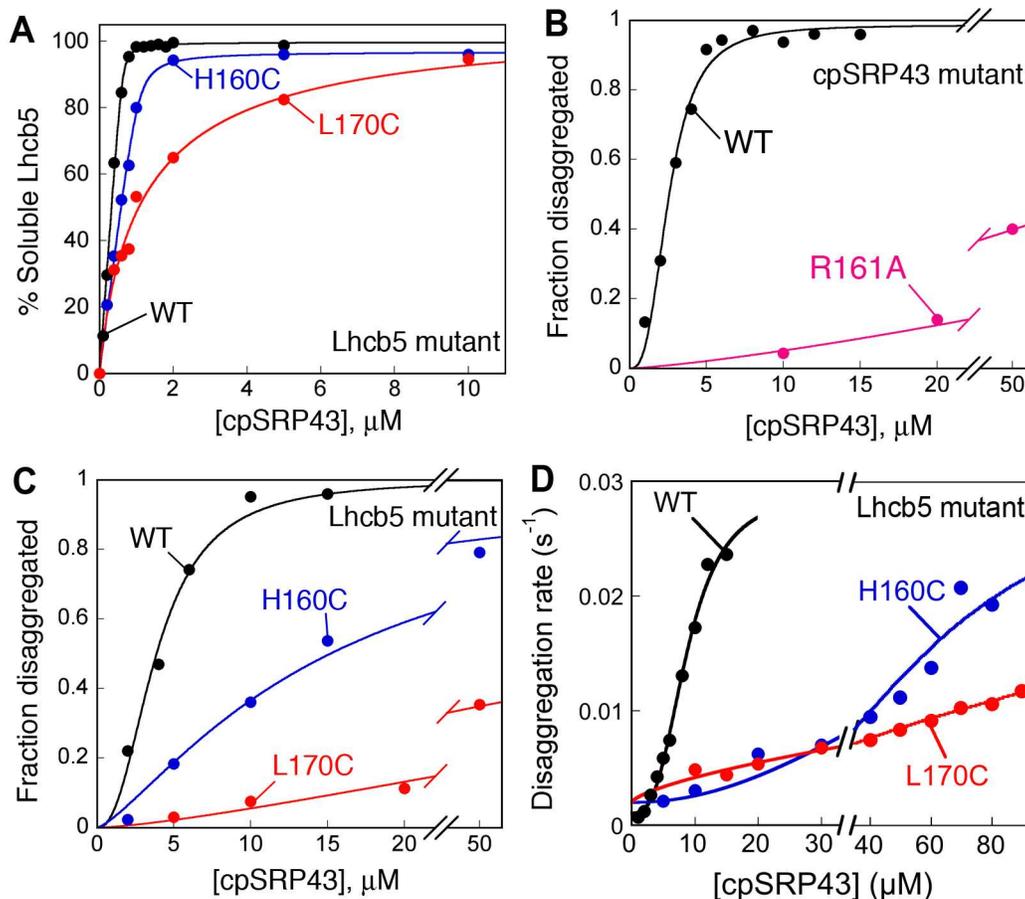


**Figure 3.2** cpSRP43 can interact with a variety of LHCP TM mutants. (A) Binding of cpSRP43 to TM mutants as measured by changes in anisotropy. Fits of data gave  $K_d$  values of 22 nM for  $\Delta$ TM3 and 713 nM for 1-2-2. For comparison, the  $K_d$  values measured by light scattering were 26 and 489 nM, respectively (Table 3.4). (B–E) Binding of cpSRP43 to LHCP and its TM mutants as measured by the ability of cpSRP43 to prevent the aggregation of substrate proteins (see “Experimental Procedures”). The data were fit to Equation 1 (see “Experimental Procedures”) and gave  $K_d^{app}$  values that are summarized in Table 3.4. (F) Summary of the  $K_d^{app}$

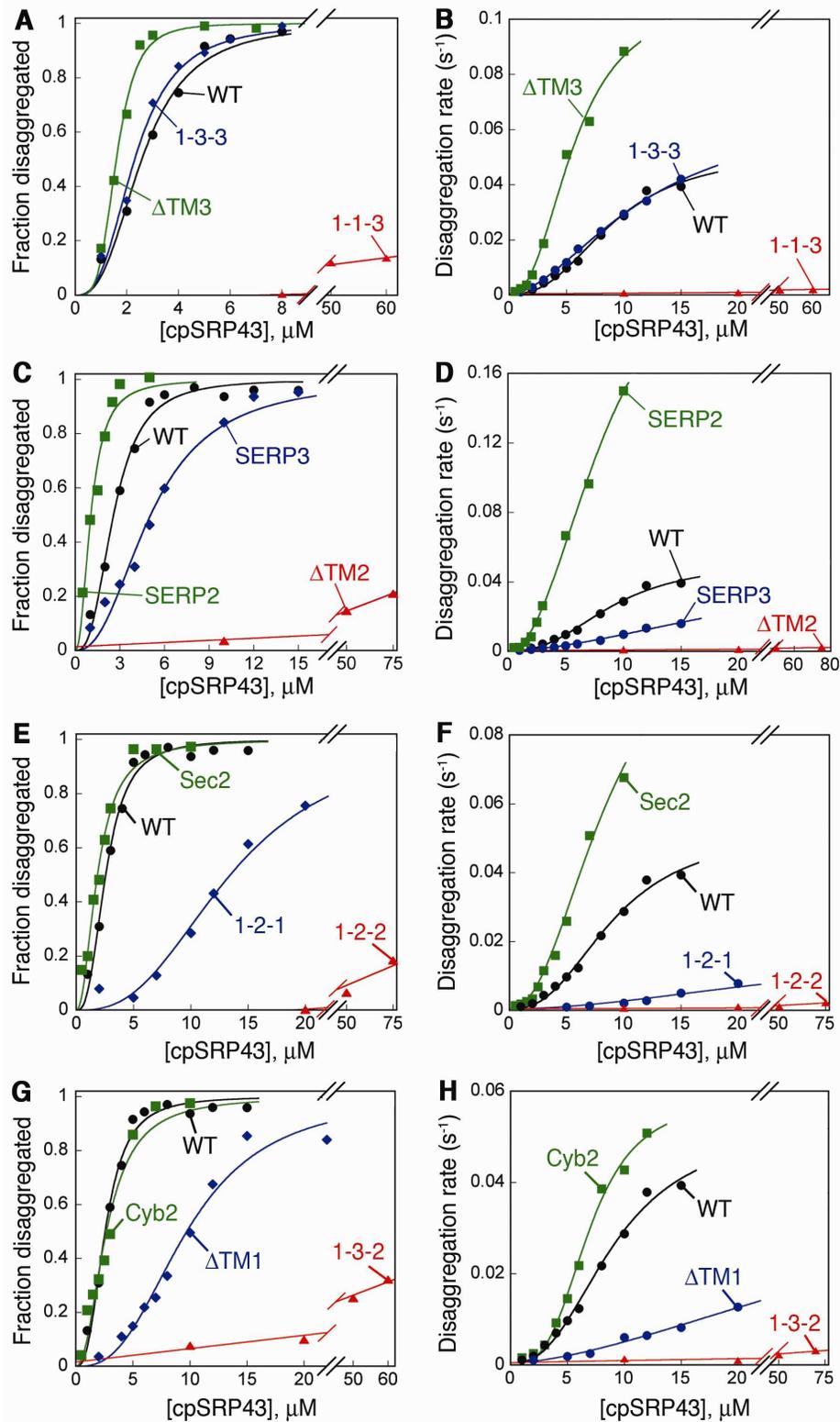
values of all the LHCP TM mutants characterized in this study. Values of  $K_d^{\text{app}}$  were determined by a combination of light scattering and fluorescence anisotropy assays.



**Figure 3.3** Schematics depicting quantitative analysis of the cpSRP43-mediated disaggregation reaction. Concentration dependences of the kinetics (A) and equilibrium (B) of the cpSRP43-mediated reversal of wild-type LHCP aggregate yield important parameters that report on the energetics of different steps of the disaggregation pathway (C).

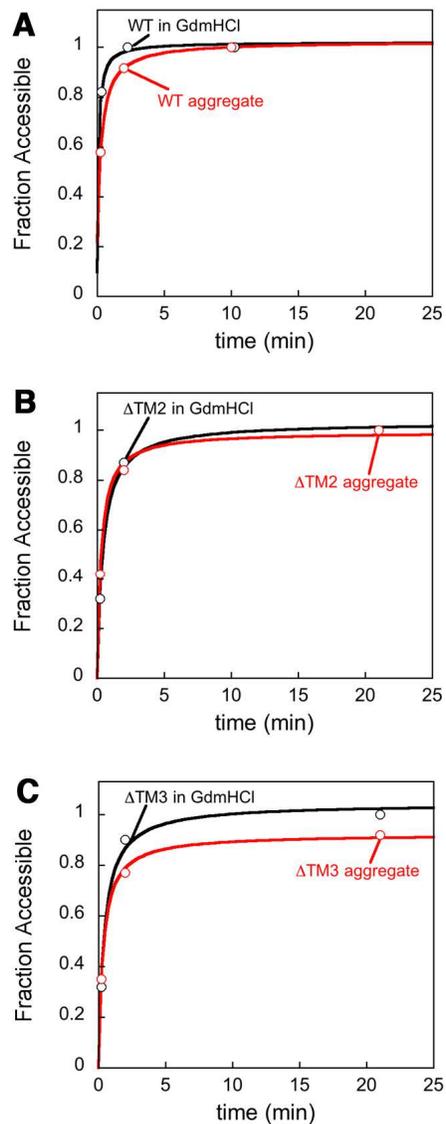


**Figure 3.4** L18-binding mutants uncouple initial recognition of the aggregate from its subsequent solubilization. (A) Binding of cpSRP43 to wild-type Lhcb<sub>5</sub> and L18 mutants H160C and H170C. The data were fit to Equation 1 (see “Experimental Procedures”) and gave  $K_d^{app}$  values of 10 nM for wild-type Lhcb<sub>5</sub> (black), 30 nM for Lhcb<sub>5</sub>-H160C (blue), and 1.1 μM for Lhcb<sub>5</sub>-L160C (red). (B) Concentration dependences for the equilibrium of disaggregation of LHCP by wild-type cpSRP43 (black) or mutant cpSRP43(R161A) (magenta). (C and D) Chaperone concentration dependences for the equilibrium (C) and kinetics (D) of disaggregation of Lhcb<sub>5</sub> (black), Lhcb<sub>5</sub>-H160C (blue), and Lhcb<sub>5</sub>-L170C (red) by wild-type cpSRP43.

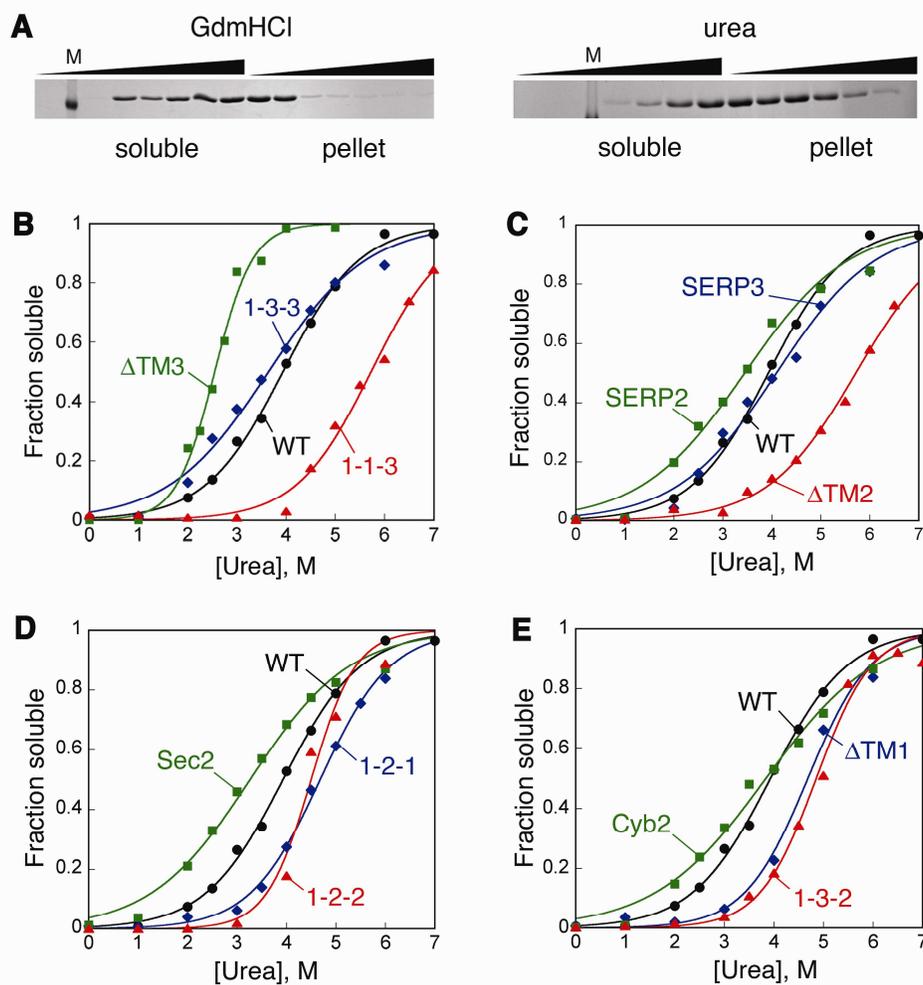


**Figure 3.5** LHCP TM mutants exhibit a wide range of disaggregation efficiencies. (A–H) Representative concentration dependences of the equilibrium (A, C, E, and G) and kinetics (B, D, F, and H) for disassembly of the aggregates formed by the LHCP TM mutants. The data for

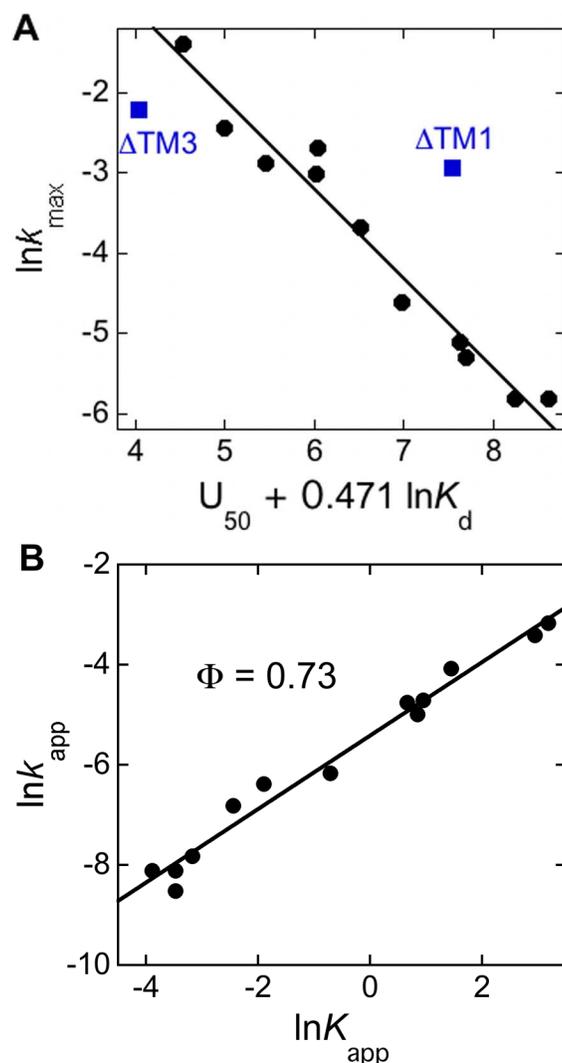
wild-type LHCP (black) were shown as a reference of comparison in all four sets. The data in A, C, E, and G were fit to Equation 4 (black, blue, and green) or Equation 6 (red) to obtain  $K_{\max}$  values and extract  $K_{\text{app}}^{4\mu\text{M}}$  values at 4  $\mu\text{M}$  cpSRP43. The data in B, D, F, and H were fit to Equation 5 (black, blue, and green) or Equation 7 (red) to obtain  $k_{\max}$ ,  $\langle K_{\text{m}} \rangle$ , and  $h$  values and to extract  $k_{\text{app}}^{4\mu\text{M}}$  values at 4  $\mu\text{M}$  cpSRP43. All the thermodynamic and kinetic parameters were reported in Table 3.4.



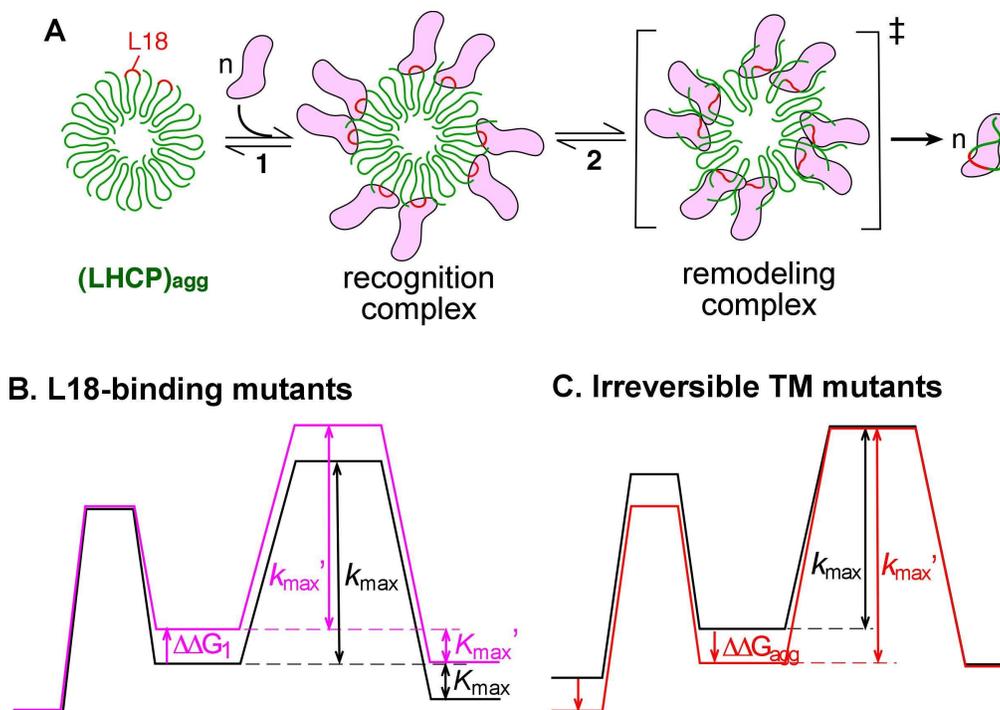
**Figure 3.6** Time courses for the alkylation reactions of cysteine residues in the L18 shows accessibility of WT (A) and mutant (B and C) LHC proteins. A–C, Lhcb5 L170C (A),  $\Delta$ TM2 G158C (B), and  $\Delta$ TM3 G158C (C) were labeled with 30-fold excess *N-ethyl-maleimide* in denaturant guanidinium hydrochloride (GdmHCl), pH 7.5 (black traces) and in aqueous buffer, pH 7.5 (red traces), and the reactions were quenched at different time points with DTT and subjected to intact protein mass spectrometry as in the accompanying manuscript [26].



**Figure 3.7** The irreversible TM mutants form ultrastable aggregates. (A) Sedimentation analysis of the ability of guanidinium chloride (GdmHCl) and urea to resolubilize LHCP aggregates. M denotes the protein marker lane. (B–E) Urea solubilization curves of LHCP and its TM mutants. The data were fit to Equation 8 (see “Experimental Procedures”) and gave  $U_{50}$  values (Table 3.4).



**Figure 3.8** Linear free energy analysis of the cpSRP43-mediated disaggregation reaction. (A) The maximal disaggregation rate constant strongly correlates with a weighted combination of the  $U_{50}$  and  $K_d^{\text{app}}$  values. Values for the analysis are from Table 3.4. The black line represents a linear fit to the data ( $R^2 = 0.96$ ).  $\Delta\text{TM1}$  and  $\Delta\text{TM3}$  (blue) were marked as outliers and were not included in the correlation. (B)  $\Phi$  value analysis of LHCP disaggregation. The values of  $K_{\text{app}}$  and  $k_{\text{app}}$  were calculated from fits of disaggregation equilibrium and kinetic data to Equations 4 and 5 (“Experimental Procedures”), respectively, and the concentration of cpSRP43 was chosen at 4 mM. Linear fit of the data (black line,  $R^2 = 0.98$ ) gave a slope ( $\Phi$  value) of 0.73.



**Figure 3.9** A, Working model for cpSRP43-mediated disaggregation reaction. Step 1 depicts initial binding of cpSRP43 (magenta) to the LHCP aggregate (green), which occurs via recognition of the solvent-exposed L18 motif (red). Step 2 depicts the cooperative action of cpSRP43 molecules to compete with and disrupt the packing interactions between the LHCP TM segments within the aggregate, leading to its resolubilization. B and C, qualitative free energy diagrams summarizing the effects of the L18-binding mutants that disrupt the initial binding step (B) and the irreversible TM mutants that disrupt the remodeling step (C), as described under “Results.” The figures are not drawn to scale.

## References

1. Balch, W. E., Morimoto, R.I., Dillin, A., and Kelly, J.W. (2008) Adapting proteostasis for disease intervention. *Science* **319**, 916–919
2. Hartl, F. U., and Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **295**, 1852–1858
3. Chang, H. C., Tang, Y.C., Hayer-Hartl, M, and Hartl, F.U. (2007) Snapshot: molecular chaperones, Part I. *Cell* **128**, 212-213
4. Tang, Y. C., Chang, H.C., Hayer-Hartl, M, and Hartl, F.U. (2007) Snapshot: molecular chaperones, Part II. *Cell* **128**, 412-413
5. Doyle, S. M., and Wickner, S. (2008) Hsp104 and ClpB: protein disaggregating machines. *Trends Biochem. Sci.* **34**, 40–48
6. Glover, J. R., and Lindquist, S. (1998) Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins. *Cell* **94**, 73–82
7. Goloubinoff, P., Mogk, A., Zvi, A.P.B., Tomoyasu, T., and Bukau, B. (1999) Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13732-13737
8. Haslberger, T., Bukau, B., and Mogk, A. (2010) Towards a unifying mechanism for ClpB/Hsp104-mediated protein disaggregation and prion propagation. *Biochem. Cell Biol.* **88**, 63-75
9. Cohen, E., Bieschke, J., Perciavalle, R.M., Kelly, H.W., and Dillin, A. (2006) Opposing activities protect against age-onset proteotoxicity. *Science* **313**, 1604-1610
10. Shorter, J. (2011) The mammalian disaggregase machinery: Hsp110 synergizes with Hsp70 and Hsp40 to catalyze protein disaggregation and reactivation in a cell-free system. *PLoS One* **6**, e26319
11. Rampelt, H., Kirstein-Miles, J., Nillegoda, N.B., Chi, K., Scholz, S.R., Morimoto, R.I., and Bukau, B. (2012) Metazoan Hsp70 machines use Hsp110 to power protein disaggregation. *EMBO J.* **31**, 4221-4235
12. Jaru-Ampornpan, P., Shen, K., Lam, V.Q., Ali, M., Doniach, S., Jia, T.Z., and Shan, S. (2010) ATP-independent reversal of a membrane protein aggregate by a chloroplast SRP subunit. *Nat. Struct. Mol. Biol.* **17**, 696–702
13. Schuenemann, D., Gupta, S., Persello-Cartieaux, F., Klimyuk, V.I., Jones, J.D., Nussaume, L., and Hoffman, N.E. (1998) A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10312–10316
14. Jansson, S. (1999) A guide to the Lhc genes and their relatives in Arabidopsis. *Trends Plant Sci.* **4**, 236–240
15. Liu, Z., Yan, H., Wang, K., Kuang, T., Zhang, J., Gui, L., An, X., and Chang, W. (2004) Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature* **428**, 287-292
16. Falk, S., Ravaud, S., Koch, J., and Sinning, I. (2010) The C-terminus of the Alb3 membrane insertase recruits cpSRP43 to the thylakoid membrane. *J. Biol. Chem.* **285**, 5954–5962
17. Hachiya, N., Mihara, K., Suda, K., Horst, M., Schatz, G., and Lithgow, T. (1993) A mitochondrial import factor purified from rat liver cytosol is an ATP-dependent conformational modulator for precursor proteins. *EMBO J.* **12**, 1579–1586

18. Hachiya, N., Komiya, T., Alam, R., Iwahashi, J., Sakaguchi, M., Omura, T., and Mihara, K. (1994) MSF, a novel cytoplasmic chaperone which functions in precursor targeting to mitochondria. *EMBO J.* **13**, 5146–5154
19. Rampelt, H., Kirstein-Miles, J., Nillegoda, N.B., Chi, K., Scholz, S.R., Morimoto, R.I., and Bukau, B. (2012) Metazoan Hsp70 machines use Hsp110 to power protein disaggregation. *EMBO J.* **31**, 4221–4235
20. Chakraborty, A., Das, I., Datta, R., Sen, B., Bhattacharyya, D., Mandal, C., and Datta, A.K. (2002) A single-domain cyclophilin from *Leishmania donovani* reactivates soluble aggregates of adenosine kinase by isomerization-independent chaperone function. *J. Biol. Chem.* **277**, 47451–47460
21. Bieschke, J., Cohen, E., Murray, A.N., Dillin, A., and Kelly, J.W. (2009) A kinetic assessment of the *C. elegans* amyloid disaggregase activity enables uncoupling of disassembly and proteolysis. *Protein Science* **18**, 2231–2241
22. Murray, A. N., Solomon, J.P., Wang, Y.-J., Balch, W.E., and Kelly, J.W. (2010) Discovery and characterization of a mammalian amyloid disaggregation activity. *Protein Science* **19**, 836–846
23. Tu, C. J., Peterson, E.C., Henry, R., and Hoffman, N.E. (2000) The L18 domain of light-harvesting chlorophyll proteins binds to chloroplast signal recognition particle 43. *J. Biol. Chem.* **275**, 13187–13190
24. Delille, J., Peterson, E.C., Johnson, T., Morre, M., Kight, A., and Henry, R. (2000) A novel precursor recognition element facilitates posttranslational binding to the signal recognition particle in chloroplasts. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1926–1931
25. Stengel, K. F., Holdermann, I., Cain, P., Robinson, C., Wild, K., and Sinning, I. (2008) Structural basis for specific substrate recognition by the chloroplast signal recognition particle protein cpSRP43. *Science* **321**, 253–256
26. Nguyen, T. X., Jaru-Ampornpan, P., Lam, V.Q., Cao, P., Piskiewicz, S., Hess, S., and Shan, S. (2013) Mechanism of an ATP-independent protein disaggregase. Part I. Structure of a membrane protein aggregate reveals a mechanism of recognition by its chaperone. *J. Biol. Chem.* **288**, 13420–13430
27. Abramoff, M. D., Magelhaes, P.J. and Ram, S.J. (2004) Image processing with ImageJ. *Biophotonics Int.* **11**, 36–42
28. Fersht, A. (1998) *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*, W.H. Freeman and Company, New York.
29. Cain, P., Holdermann, I., Sinning, I., Johnson, A.E., and Robinson, C. (2011) Binding of chloroplast signal recognition particle to a thylakoid membrane protein substrate in aqueous solution and delineation of the cpSRP43-substrate interaction domain. *Biochem. J.* **437**, 149–155
30. Weibezahn, J., Tessarz, P., Schlieker, C., Zahn, R., Maglica, Z., Lee, S., Zentgraf, H., Weber-Ban, E.U., Dougan, D.A., Tsai, F.T., Mogk, A., and Bukau, B. (2003) Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. *Cell* **119**, 653–665
31. Schlieker, C., Weibezahn, J., Patzelt, H., Tessarz, P., Strub, C., Zeth, K., Erbse, A., Schneider-Mergener, J., Chin, J.W., Schultz, P.G., Bukau, B., and Mogk, A. (2004) Substrate recognition by the AAA+ chaperone ClpB. *Nat. Struct. Mol. Biol.* **11**, 607–615

32. Lee, C., Schwartz, M.P., Prakash, S., Iwakura, M., and Matouschek, A. (2001) ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. *Mol. Cell* **7**, 627-637
33. Kenniston, J. A., Burton, R.E., Siddiqui, S.M., Baker, T.A., and Sauer, R.T. (2004) Effects of local protein stability and the geometric position of the substrate degradation tag on the efficiency of ClpXP denaturation and degradation. *J. Struct. Biol.* **146**, 130-140
34. Shorter, J., and Lindquist, S. (2004) Hsp104 catalyzes formation and elimination of self-replicating Sup35 prion conformers. *Science* **304**, 1793-1797
35. Shorter, J., and Lindquist, S. (2006) Destruction or potentiation of different prions catalyzed by similar Hsp104 remodeling activities. *Mol. Cell* **23**, 425-438

## **Chapter 4**

### **Mitigating Protein Aggregation:**

### **Towards the Application of cpSRP43 as a Versatile**

### **Chaperone for Membrane Protein Systems**

This unpublished work was done with the help of McAvoy, C., Si, V., and Piskiewicz, S.

**Abstract**

In recent biochemical studies, the chloroplast Signal Recognition Particle 43 (cpSRP43) has been shown to possess robust molecular chaperone activity in preventing and reversing the aggregation of its cognate substrate, the proteins belonging to a family of light-harvesting chlorophyll-*a/b* binding (LHC) proteins. Additionally, it has been hypothesized that cpSRP43's chaperone function is rooted in its tight binding affinity to the recognition motif (L18) in LHCs that is specific, and additional hydrophobic contacts with the substrate that is more generic. Here, we show preliminary data to suggest that cpSRP43 chaperone activity can be extended to alternative substrates unrelated to the natural interaction. First, we show cpSRP43 *in vitro* can reverse the aggregation of LHC swap mutants where transmembrane segments of unrelated proteins have been swapped to replace that of wild-type LHC. Second, we show the potential of implementing the co-expression of cpSRP43 and "L18-tagging" as a strategy to improve the low yield of membrane proteins in the bacterial expression system. Lastly, we demonstrate that cpSRP43 inhibits the fibrillization of amyloid beta peptide ( $A\beta_{1-40}$ ). These results attest to cpSRP43's potential as a molecular chaperone and provides the impetus for further engineering endeavors to address problems that stem from protein aggregation.

## Introduction

The maintenance of functional proteins is handled by the protein homeostasis network that ensures the proper folding of proteins, traffics nascent proteins to their cellular destinations and clears them at the end of their life cycles [1,2]. The working of this network is essential for the survival of the cell. Failure at any step along the network not only deprives the cell of functional proteins, as proteins are misfolded in inactive conformations or mis-localized, but may also lead to cellular toxic protein aggregation. Protein misfolding and aggregation are associated with a wide spectrum of human diseases, most notably those of the amyloidosis family such as Alzheimer's and Parkinson's disease [3].

Since its conceptualization in 1987 by R. John Ellis, the class of proteins known collectively as molecular chaperones has been shown to preempt and sometimes reverse the phenomenon of protein aggregation [1,4,5]. A particular subsection of this family has been found to traffic membrane proteins in their post-translational form [6-8]. These membrane proteins are often highly hydrophobic proteins that must traverse through an aqueous cytosol before arriving at their lipid environment. Throughout this process, the membrane proteins must be efficiently chaperoned to prevent the aberrant protein-protein interaction that leads to their aggregation. Examples of this class of multifunctional proteins that demonstrate the intimate link between molecular chaperone function and membrane protein biogenesis are found in the Hsp70 that facilitates the translocation of mitochondrial proteins [9,10], the components of the Guided Entry of Tail-anchored proteins (GET) pathway [11,12], and the chloroplast Signal Recognition Particle 43 (cpSRP43) that assists the import of thylakoid proteins [13,14].

Recently, we and others described a novel function of cpSRP43 as a chaperone that not only can prevent but also can reverse the aggregation of proteins without the external energy

input of ATP hydrolysis or co-chaperones [15,16]. Rather, the disaggregation mechanism of cpSRP43 relies on the intrinsic binding with its substrate, the light-harvesting chlorophyll *a/b*-binding (LHC) family of proteins [17,18]. The LHC proteins are composed of three highly hydrophobic transmembrane (TM) domains, making them highly prone to aggregation and requiring a robust chaperone system during their transit through the aqueous compartment of the stroma before reaching the thylakoid membrane [19,20]. We have proposed that cpSRP43 employs two modes of substrate binding to facilitate disaggregation reactions: (1) a sequence-specific binding of the recognition motif composed of an 18 amino acid loop between TM2 and TM3 of LHC proteins, called L18 and (2) a more generic hydrophobic interaction with the TMs.

Previous results suggest that the generic hydrophobic interactions of cpSRP43 with the TMs of its substrate allows this interaction to be more adaptable [18]. This more adaptable interaction allows cpSRP43 to still disassemble substrates of LHC mutants that have their TMs swapped out and replaced with TMs of unrelated membrane proteins, including mammalian tail-anchored proteins like the stress-associated endoplasmic reticulum protein 1 (SERP1), the protein transport protein Sec61 beta subunit (Sec61 $\beta$ ) and cytochrome *b*<sub>5</sub> [18]. With this insight, we wondered if we could harness the potential of this small chaperone to be engineered to handle a broader range of aggregation-prone protein substrates. By exploiting the simplicity of the single-component disaggregase and dualistic specificity and adaptability of the cpSRP43-LHC system, we can shed light on this model system that can guide the engineering of protein scaffolds that can target and remove a broad array of aggregates of interest [21,22]. Here we present two application cases that support the great bioengineering potential of cpSRP43 as a robust chaperone. In the first case, we show initial attempts at improving the yield of hard-to-express proteins, specifically SERP1, by employing the strategy of cpSRP43 co-expression

coupled with “L18-tagging” of the protein of interest as a way to prevent the degradation of these unstable proteins during synthesis due to their aggregation propensity. In the second case, we show the capability of cpSRP43 to suppress the fibrillization of amyloid beta peptide ( $A\beta_{1-40}$ ), which is intimately linked with Alzheimer’s disease.

## Materials and Methods

**Materials.** To construct the LHCP TM mutants (Table 1), a pair of unique restriction sites was introduced into the expression plasmid encoding LHCP  $\Delta$ TM1 before and after the sequences encoding TM2 or TM3. The sequences coding for the TMs were replaced with PCR fragments encoding alternative TMs using the corresponding restriction sites. Truncated L18 LHCP mutants, SERP1-L11 and SERP1-L18 were constructed using QuikChange PCR (Stratagene). To construct the co-expression plasmid, the PCR fragments encoding for SERP1, SERP1-L11, or SERP1-L18 were introduced downstream of the cpSRP43 gene within the multiple cloning site of the vector pQE (Qiagen). cpSRP43, LHCP and its variants were purified as described [15].  $A\beta_{1-40}$  and re-crystallized thioflavin T (ThT) were generous gifts from Dr. J. W. Kelly. Urea is from Sigma.

**Light Scattering Assay.** Light scattering experiments were performed as previously described [15]. For formation of aggregates (Figure 4.1, black), unfolded mutant substrate in 8 M urea was directly diluted into Buffer D (50 mM KHEPES pH 7.5, 200 mM NaCl) to 2  $\mu$ M final concentration. Disaggregation reaction was initiated 60 s after substrate aggregation by addition of various concentrations of cpSRP43. Kinetic analysis was performed as described previously to obtain the forward disaggregation rate (Figure 4.1D) [15].

**Preparation of Amyloid Beta Peptide (1-40).** Lyophilized synthetic  $A\beta_{1-40}$  (DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV) was a generous gift from the J. Kelly lab and monomerization of the peptide is as followed from protocol [23]. For each assay requiring  $A\beta_{1-40}$ , 1 mg of the peptide was dissolved in 600  $\mu$ L 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP, Sigma) and incubated at 25  $^{\circ}$ C for 2 h. HFIP was removed by blowing a gentle stream of argon over the solution, and the resulting film of peptide was dissolved in 5 mM NaOH

(500  $\mu\text{L}$ ) with 3.5 mM TCEP (Thermo Scientific) added and then sonicated for 2 h in ice-cold water bath. The resulting solution containing monomerized amyloid- $\beta$  was first passed through a 0.22- $\mu\text{m}$  filter (Millipore) and then through a 10-kDa-cutoff Centricon filter (Millipore).

**ThioflavinT (ThT) Fluorescence Assay.** To detect formation of fibrils using ThT fluorescence, the monomerized  $\text{A}\beta_{1-40}$  was diluted 10  $\mu\text{M}$  into 50 mM sodium phosphate buffer (150 mM NaCl, 3.5 mM TCEP, pH 7.4) containing 20  $\mu\text{M}$  re-crystallized ThT, and aggregation was initiated with automatic 5-second shaking every 10 minutes at 37  $^{\circ}\text{C}$  in a 96-well microplate reader Flexstation 3 (Molecular Devices). Fluorescence readings were taken every 10 minutes with excitation wavelength 440 and emission wavelength 485. For experiments testing the effect of cpSRP43 on fibrillization, cpSRP43 was pre-incubated in the buffer in a 1:1 molar ratio.

**Transmission Electron Microscopy.**  $\text{A}\beta_{1-40}$  samples with and without cpSRP43 samples from the ThT assays were taken at  $t = 48$  h, when ThT fluorescence has saturated, indicative of the formation of fibrils. 4  $\mu\text{L}$  of samples of 10  $\mu\text{M}$   $\text{A}\beta_{1-40}$  were immediately deposited onto a glow-discharged 200-mesh Formvar grid (Ted Pella Inc., CA). After a 60-second adsorption time, the grid was washed in water twice and then stained with 1% uranyl acetate for 60 seconds. TEM images were obtained on a 120 kV Tecnai T12 electron microscope coupled with a CCD camera. The diameters of the particles were measured using ImageJ [24].

## Results

**cpSRP43 Can Recognize and Rescue Aggregates of LHCP Mutants with TMs Swapped with that of Un-related Membrane Proteins.** We first tested the capability of cpSRP43 to recognize and reverse the aggregation of LHCP mutants where the TMs have been swapped with those of unrelated membrane proteins. As a simple scaffold, we started with a mutant whose aggregates remain to be recognized and dismantled by cpSRP43, LHCP TM1 deletion mutant ( $\Delta$ TM1) [25]. From  $\Delta$ TM1, we replaced the remaining TM2 and TM3 with TMs of the tail-anchored proteins SERP1 (Serp), Sec61 $\beta$  (Sec) and cytochrome b5 (Cyb) to create a library of all nine possible permutations (See Table 4.1). Of the nine mutants cloned, eight were successfully expressed and purified from inclusion bodies. Of the eight mutants that we obtained, we subjected three to be extensively tested on their ability to be disaggregated by cpSRP43: Serp/Cyb, Cyb/Serp and Cyb/Sec, corresponding to TM2/TM3 swapping.

All three TM-swapped mutants aggregated upon dilution out of 8 M Urea and into aqueous buffer D and were, indeed, disaggregated by the addition of cpSRP43 in a concentration dependent manner (Figure 4.1 A-C). The extent of disaggregation is saturated with 1:5-molar excess of chaperone concentration, similar to that reported for wild-type LHCP disaggregation [15,18]. Interestingly, the rate of disaggregation at this cpSRP43 concentration for Cyb/Serp and Serp/Cyb is more than two-fold faster than that of wild-type LHCP, while the rate for Cyb/Sec aggregates is comparable to wild-type level (Figure 4.1D). This suggested that cpSRP43 is amendable to handling non-cognate protein aggregates by concurrently interacting with generic hydrophobic TMs and the sequence-specific recognition motif L18.

**Decipher the Minimal Recognition Motif within L18 for Chaperone Activity.** To further test the limit of cpSRP43 in handling alternative proteins, we asked if it would accept a

substrate that it has not encountered, mediated by the bipartite binding from the L18 motif engineered into the target protein and generic hydrophobic contacts with the TM available in the target. Given that only the FDPLGL motif in L18 is sensitive to mutations as assayed by cpSRP43 binding, we asked if we could “trim” down the recognition motif to minimize the effects of its insertion on the folding or function of the target protein [18]. To answer this question, we first deleted seven residues from the N-terminal end of L18 as it is farther away from the sensitive FDPLGL motif, constructing LHCP with only 11 residues left in the recognition motif (referred to as L11). The L11 LHCP mutant could be disaggregated by cpSRP43 to the same extent of that of wild-type (Figure 4.2). However, when one additional residue was trimmed from L11 either on the N- or C- terminal ends, the ability of cpSRP43 to rescue the aggregate was drastically reduced, re-solubilizing about 20% of the substrate proteins (Figure 4.2). With this result, we have demonstrated a “trimmed” L11 is the minimal recognition motif required for efficient cpSRP43 disaggregation.

**Test whether cpSRP43 improves the yield, solubility or localization of membrane proteins *in vivo*.** With the minimal recognition motif construct determined, we proceeded to construct SERP1-L11 and SERP1-L18, where the recognition motif L11 or L18 is engineered C-terminal to the TM of SERP1. However, these mutant proteins were not successfully obtained due to low expression in *E. coli*, as the target protein SERP1 itself has low expression (Figure 4.3, lanes 8 and 9). This is not surprising as the over-expression of membrane proteins often presents a daunting challenge due to their propensity for degradation by the cell during their expression. Degradation of these membrane proteins is attributed to their highly hydrophobic transmembrane domains that tend to misfold, mislocalize and/or aggregate [26-28]. One strategy to overcome the abysmal expression of membrane proteins has relied on co-expression with a

cognate chaperone that will stabilize the finicky membrane protein [29,30]. Using the same principle we wondered if we could engineer a non-cognate protein, like SERP1, that can be chaperoned by cpSRP43 and thus improve its yield by co-expression with cpSRP43. This strategy not only provides us with recombinant client proteins to test disaggregase activity of cpSRP43 towards alternative substrates, it also tests the ability of cpSRP43 to overcome the aggregation of target proteins in a cellular environment.

To test if the yield of the client proteins can be improved with co-expression with cpSRP43, we cloned in the genes for our client proteins down-stream of the genes encoding cpSRP43 in the pQE vector under the phage T5 promoter; previously, expression of cpSRP43 has shown to be robust in this bacterial expression system [15]. Our preliminary experiments showed that the expression of SERP1 (~ 9 kDa protein) is significantly enhanced upon fusion with L18 or L11 and co-expressed with cpSRP43. Expression of the ~ 11 kDa fusion constructs His<sub>6</sub>-SERP1-L11 and His<sub>6</sub>-SERP1-L18 was markedly improved at both temperature conditions (25 and 37 °C) as detected from whole cell analysis of post-IPTG induction samples by Western blot with anti-His antibody (Figure 4.3, lanes 2,3 and 5,6).

**cpSRP43 interferes with A $\beta$ <sub>1-40</sub> fibrillization *in vitro*.** Another candidate for an alternative substrate for cpSRP43's chaperone action is the 40 amino acid polypeptide fragment amyloid beta (A $\beta$ <sub>1-40</sub>) whose aggregation is associated with Alzheimer's disease. Although the question of which aggregate form is the toxic species in the pathology is still contested in the field, we wanted to see if cpSRP43 can interact with the hydrophobic polypeptide, as a preliminary test of the adaptability of this protein. First, we analyzed the effect of cpSRP43 on A $\beta$ <sub>1-40</sub> aggregation using a Thioflavin T (ThT) binding assay [31]. Purified A $\beta$ <sub>1-40</sub> peptide was incubated in the presence and absence of cpSRP43 and fibrillogenesis was monitored by

measuring ThT fluorescence emission at 485 nm (Figure 4.4A). In the absence of cpSRP43, we observed ThT-positive,  $\beta$ -sheet-rich  $A\beta_{1-40}$  aggregates after a lag phase of  $\sim 30$  h, consistent with a nucleated polymerization mechanism and with previously published studies [23,31,32]. In contrast, formation of  $A\beta_{1-40}$  aggregates was suppressed in the presence of cpSRP43 (Figure 4.4A). cpSRP43 at a molar ratio of 1:1 reduced the ThT fluorescence to  $\sim 10\%$  compared to that in solvent alone.

Next, we investigated the effect of cpSRP43 on  $A\beta_{1-40}$  fibrillogenesis by negative-stain Electron Microscopy (EM). In the absence of cpSRP43, we observed formation of predominantly fibrillar  $A\beta_{1-40}$  structures with a diameter of  $\sim 10$  nm and length of  $\sim 0.5-3$   $\mu\text{m}$  (Figure 4.4B), supporting the results of the ThT assays and congruent with previously published works [33,34]. In contrast, an equimolar concentration of cpSRP43 relative to  $A\beta_{1-40}$  markedly reduced fibril assembly in favor of spherical protein aggregates with diameters less than  $\sim 20$  nm (Figure 4.4C-D). It appears that cpSRP efficiently prevents the amyloidogenesis of  $A\beta_{1-40}$  but stimulates the assembly of compact, spherical oligomers. It remains to be tested whether these oligomers are the neural toxic species or off-pathway, non-toxic oligomers that have been isolated in other cases [35,36].

## Discussion

These studies reveal that cpSRP43's interaction with hydrophobic protein motifs is quite generic and can be used to disrupt a larger set of protein aggregates beyond LHC proteins, as long as the FDPLGL motif is present to allow this chaperone to recognize and engage the aggregate. In order to accommodate for this requirement we showed that the minimal motif for recognition could be trimmed down to 11 amino acids that still contained FDPLGL (Figure 4.2). Further truncation of this "L11" motif was not effective in the disaggregation mechanism, likely because a certain minimal length of peptide is needed to allow for the FDPLGL motif to escape the aggregate and be exposed for recognition by cpSRP43. On the other hand, we showed how cpSRP43 could accommodate mutant substrates by a more generic interaction with hydrophobic protein motifs, as TMs swapped from unrelated integral membrane proteins can still be disaggregated by cpSRP43 (Figure 4.1). This observation was the first implication for the high adaptability of cpSRP43 to be tailored as a molecular chaperone for alternative substrates.

We first tested the limits of cpSRP43's chaperone capability by asking if it could improve the low expression of integral membrane proteins, like SERP1. A major factor that could contribute to this problem is overload on the chaperones and protein targeting machinery in the bacterial expression systems, which leads to the aggregation, misfolding and/or degradation of the expressed protein, as manifested by the induction of heat shock proteins and proteases upon membrane protein expression [37]. As a naturally evolved chaperone that assists in membrane protein targeting, cpSRP43 is uniquely suited to help overcome the problem of poorly expressed proteins by preventing the aggregation and degradation of these proteins. In support of this notion, our preliminary experiments showed that the expression of the membrane protein SERP1 is significantly enhanced upon fusion with L18 or L11 and co-expression with cpSRP43 (Figure

4.3). It remains to be tested whether cpSRP43 helps maintain membrane proteins in translocation competent conformations and thus allows them to be more productively integrated into the membrane, but preliminary data from differential centrifugation shows SERP1-L18 and SERP1-L11 is split between the insoluble and the bacterial membrane fraction when co-expressed with cpSRP43 (C. McAvoy, unpublished data). As proteins that are stable in a proper environment are less likely to be flagged for degradation, these initial findings suggest the ability of cpSRP43 to prevent membrane proteins from aggregating and thus escape the degradation pathway.

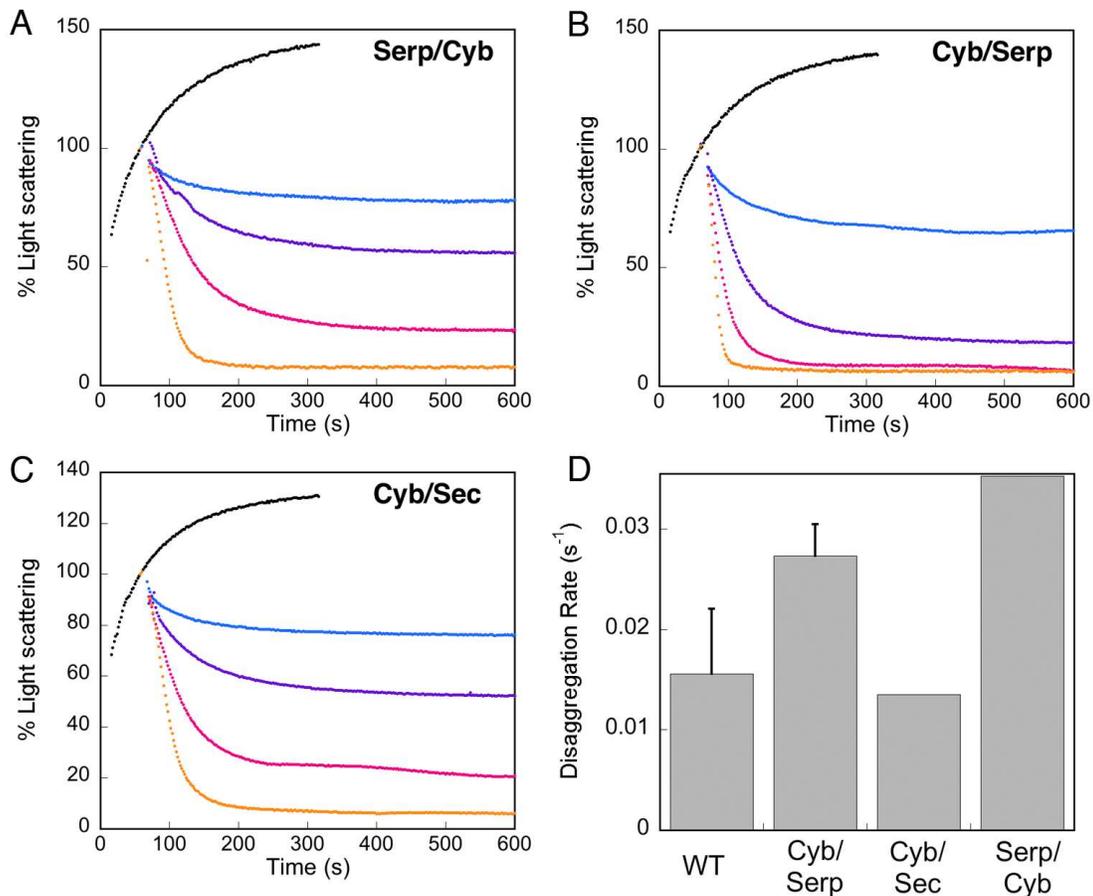
As another testament to the ability of cpSRP43 to bind generic hydrophobic protein segments, the interaction of the amyloid peptide  $A\beta_{1-40}$  with cpSRP43 was able to affect its aggregation trajectory. Strikingly, even without the L18 motif essential for disaggregation, cpSRP43 could interact non-specifically with the  $A\beta_{1-40}$  peptide to inhibit the formation of fibrils (Figure 4.4). Unlike the strategy of “tagging” proteins with the FDPLGL recognition motif to improve expression of membrane proteins, introduction of L11 or L18 into disease-related aggregation-prone proteins in biological systems is not a viable solution. However, this result highlights the adaptability of cpSRP43 for binding of generic hydrophobic motifs and demonstrates the feasibility of using engineering tools to evolve the highly module cpSRP43 to develop novel, high affinity binding scaffolds for disease-related targets. Collectively, these results show the great potential of the LHC-cpSRP43 system to be manipulated and adapted to address the problems associated with protein aggregation.

**Acknowledgement**

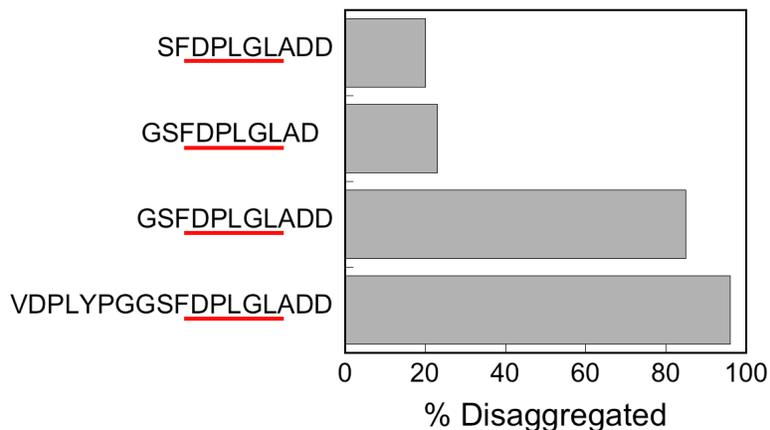
We would like to thank Drs. W.M. Clemons, J. Chartron, and C. Suloway for the plasmids of SERP1, Sec61 $\beta$  and cytochrome b<sub>5</sub>, Drs. J. Kelly, X. Li and X. Zhang for providing materials and protocols for the A $\beta$ <sub>1-40</sub> aggregation experiments, and A. McDowall for help with the EM work. Cloning and experiments were done with the help of rotation students Va Si and Camille  
McAvoy.

**TABLE 4.1** Sequence of TM swap mutants. TM domains of SERP1, Sec61 $\beta$ , and cytochrome b5 are highlighted in cyan, yellow and magenta, respectively. L18 motif is in red.

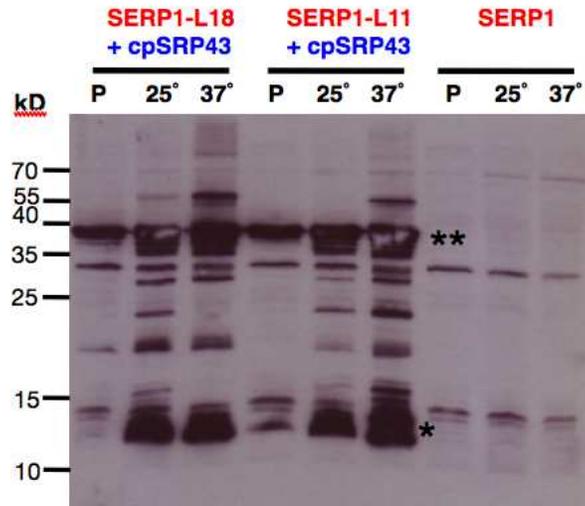
Replacement TM2/TM3	Construct Sequence
Serp/Serp	M R G S H H H H H H G S M R K S A T T K K V A S S G S P W Y G P D R V K Y L G P F S G E S P S Y L T G E F P G D Y G W D T A G L S A D V K F G E A V W F K A G S Q I F S E G G L D Y L G N P S L V H E L A S V G P W L L A L F I F V V C G S A I F G T P L G E V V D P L Y P G G S F D P L G L A D D L Q A S V G P W L L A L F I F V V C G S A I F V D G K G P L E N L A D H L A D P V N N N A W S Y A T N F V P G K
Serp/Sec	M R G S H H H H H H G S M R K S A T T K K V A S S G S P W Y G P D R V K Y L G P F S G E S P S Y L T G E F P G D Y G W D T A G L S A D V K F G E A V W F K A G S Q I F S E G G L D Y L G N P S L V H E L A S V G P W L L A L F I F V V C G S A I F G T P L G E V V D P L Y P G G S F D P L G L A D D L Q V P V L V M S L L F I A S V F M V D G K G P L E N L A D H L A D P V N N N A W S Y A T N F V P G K
Serp/Cyb	M R G S H H H H H H G S M R K S A T T K K V A S S G S P W Y G P D R V K Y L G P F S G E S P S Y L T G E F P G D Y G W D T A G L S A D V K F G E A V W F K A G S Q I F S E G G L D Y L G N P S L V H E L A S V G P W L L A L F I F V V C G S A I F G T P L G E V V D P L Y P G G S F D P L G L A D D L Q N S S W W T N W V I P A I S A L I V A L M Y R L Y V D G K G P L E N L A D H L A D P V N N N A W S Y A T N F V P G K
Sec/Serp	M R G S H H H H H H G S M R K S A T T K K V A S S G S P W Y G P D R V K Y L G P F S G E S P S Y L T G E F P G D Y G W D T A G L S A D V K F G E A V W F K A G S Q I F S E G G L D Y L G N P S L V H E L V P V L V M S L L F I A S V F M L G T P L G E V V D P L Y P G G S F D P L G L A D D L Q A S V G P W L L A L F I F V V C G S A I F V D G K G P L E N L A D H L A D P V N N N A W S Y A T N F V P G K
Sec/Sec	M R G S H H H H H H G S M R K S A T T K K V A S S G S P W Y G P D R V K Y L G P F S G E S P S Y L T G E F P G D Y G W D T A G L S A D V K F G E A V W F K A G S Q I F S E G G L D Y L G N P S L V H E L V P V L V M S L L F I A S V F M L G T P L G E V V D P L Y P G G S F D P L G L A D D L Q V P V L V M S L L F I A S V
Sec/Cyb	M R G S H H H H H H G S M R K S A T T K K V A S S G S P W Y G P D R V K Y L G P F S G E S P S Y L T G E F P G D Y G W D T A G L S A D V K F G E A V W F K A G S Q I F S E G G L D Y L G N P S L V H E L V P V L V M S L L F I A S V F M L G T P L G E V V D P L Y P G G S F D P L G L A D D L Q N S S W W T N W V I P A I S A L I V A L M Y R L Y V D G K G P L E N L A D H L A D P V N N N A W S Y A T N F V P G K
Cyb/Serp	M R G S H H H H H H G S M R K S A T T K K V A S S G S P W Y G P D R V K Y L G P F S G E S P S Y L T G E F P G D Y G W D T A G L S A D V K F G E A V W F K A G S Q I F S E G G L D Y L G N P S L V H E L N S S W W T N W V I P A I S A L I V A L M Y R L Y G T P L G E V V D P L Y P G G S F D P L G L A D D L Q A S V G P W L L A L F I F V V C G S A I F V D G K G P L E N L A D H L A D P V N N N A W S Y A T N F V P G K
Cyb/Sec	M R G S H H H H H H G S M R K S A T T K K V A S S G S P W Y G P D R V K Y L G P F S G E S P S Y L T G E F P G D Y G W D T A G L S A D V K F G E A V W F K A G S Q I F S E G G L D Y L G N P S L V H E L N S S W W T N W V I P A I S A L I V A L M Y R L Y G T P L G E V V D P L Y P G G S F D P L G L A D D L Q V P V L V M S L L F I A S V F M V D G K G P L E N L A D H L A D P V N N N A W S Y A T N F V P G K
Cyb/Cyb	M R G S H H H H H H G S M R K S A T T K K V A S S G S P W Y G P D R V K Y L G P F S G E S P S Y L T G E F P G D Y G W D T A G L S A D V K F G E A V W F K A G S Q I F S E G G L D Y L G N P S L V H E L N S S W W T N W V I P A I S A L I V A L M Y R L Y G T P L G E V V D P L Y P G G S F D P L G L A D D L Q N S S W W T N W V I P A I S A L I V A L M Y R L Y V D G K G P L E N L A D H L A D P V N N N A W S Y A T N F V P G K



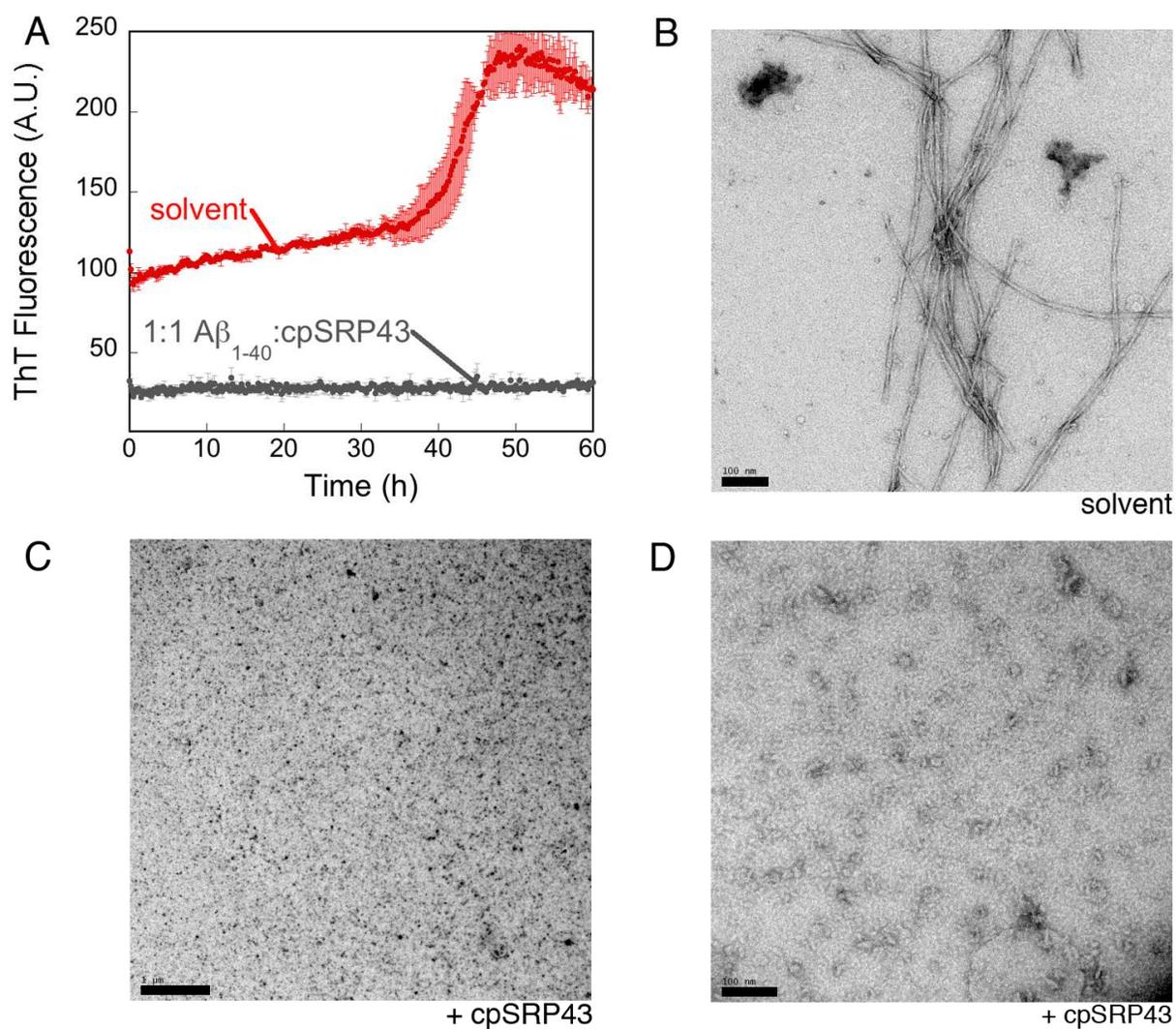
**Figure 4.1** cpSRP43 can rescue LHC aggregates with TM swapped from unrelated membrane proteins. Disaggregation time courses with varying concentrations of cpSRP43 (2, 4, 6 and 10  $\mu\text{M}$  correspond to blue, violet, magenta and orange traces, respectively) for 2  $\mu\text{M}$  substrates (A) Serp/Cyb (B) Cyb/Serp and (C) Cyb/Sec. Black trace is the aggregation kinetics of substrates without cpSRP43. (D) Comparison of TM mutant disaggregation rates at 10  $\mu\text{M}$  cpSRP43 and 2  $\mu\text{M}$  substrate; WT here is the LHCP with TM1 deleted, which served as the general platform for the swapped transmembrane.



**Figure 4.2** The minimal recognition motif within L18 can be reduced to 11 amino acid containing the essential FDPLGL motif. The amount of soluble wild-type LHCP (denoted VDPLYPGGSFDPLGLADD) is compared to the amount of soluble minimal recognition motif variants in the presence of 5  $\mu$ M cpSRP43 (five-fold excess to aggregate substrate) in disaggregation assays. The minimal recognition motif (GSFDPLGLADD) is referred to as “L11” in the text.



**Figure 4.3** Co-expression of cpSRP43 (\*\*\*) improves the yield of L11- and L18-fused SERP1 (\*). Cells grown at 25 °C and 37 °C for 4 hours after IPTG induction were harvested, lysed with BugBuster and separated by SDS-PAGE, then transferred to nitrocellulose and probed with an anti-His antibody. cpSRP43, SERP1, SERP1-L11 and SERP1-L18 are N-terminally His<sub>6</sub>-tagged. P is the pre-IPTG induction sample.



**Figure 4.4** cpSRP43 inhibits A $\beta_{1-40}$  fibrillization. (A) Effect of cpSRP43 on A $\beta_{1-40}$  (10  $\mu$ M) fibrillogenesis with agitation at 37  $^{\circ}$ C as measured by thioflavin (ThT) fluorescence. Results represent means  $\pm$  s.d. (n = 2). (B) Analysis of A $\beta_{1-40}$  aggregation reaction by EM of sample in A taken at t = 50 h. (B-C) Analysis of cpSRP43-treated A $\beta_{1-40}$  aggregation reaction where A $\beta_{1-40}$  was added into the solvent with cpSRP43 pre-incubated at molar ratio of 1:1 ratio. Samples taken at t = 50 h.

## References

1. Broadley, S. A., and Hartl, F.U. (2009) The role of molecular chaperones in human misfolding diseases. *FEBS Lett.* **583**, 2647-2653
2. Hartl, F. U., and Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **295**, 1852-1858
3. Luheshi, L. M., and Dobson, C.M. (2009) Bridging the gap: from protein misfolding to protein misfolding diseases. *FEBS Lett.* **583**, 2581-2586
4. Ellis, J. (1987) Proteins as molecular chaperones. *Nature* **328**, 378-379
5. Doyle, S. M., and Wickner, S. (2008) Hsp104 and ClpB: protein disaggregating machines. *Trends Biochem. Sci.* **34**, 40-48
6. Randall, L. L., and Hardy, S.J.S. (2002) SecB, one small chaperone in the complex milieu of the cell. *Cell. Mol. Life Sci.* **59**, 1617-1623
7. Walton, T. A., Sandoval, C.M., Fowler, C.A, Pardi, A., and Sousa, M.C. (2009) The cavity chaperone Skp protects its substrates from aggregation but allows independent folding of substrate domains. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1772-1777
8. Mariappan, M., Li, X. Stefanovic, S., Sharma, A., Mateja, A. Keenan, R.J., and Hegde, R.S. (2010) A ribosome-associating factor chaperones tail anchored membrane proteins. *Nature* **466**, 1120-1124
9. Deshaies, R. J., Koch, B.D., Werner-Washburne, M., Craig, E.A., and Schekman, R. (1988) A subfamily of stress proteins facilitates translocation of secretory and mitochondrial precursor polypeptides. *Nature* **332**, 800-805
10. Stuarda, R. A., Cyra, D.M., Craigh, E.A., and Neupert, W. (1994) Mitochondrial molecular chaperones: their role in protein translocation. *Trends Biochem. Sci.* **19**, 87-92
11. Mariappan, M., Li, X., Stefanovic, S., Sharma, A., Mateja, A., Keenan, R.J., and Hegde, R.S. (2010) A ribosome-associating factor chaperones tail-anchored membrane proteins. *Nature* **466**, 1120-1124
12. Wang, F., Brown, E.C., Mak, G., Zhuang, J., and Denic, V. (2010) A chaperone cascade sorts proteins for posttranslational membrane insertion into the endoplasmic reticulum. *Mol. Cell* **40**, 159-171
13. Schuenemann, D., Gupta, S., Persello-Cartieaux, F., Klimyuk, V.I., Jones, J.D., Nussaume, L., and Hoffman, N.E. (1998) A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10312-10316
14. Jaru-Ampornpan, P., Shen, K., Lam, V.Q., Ali, M., Doniach, S., Jia, T.Z., and Shan, S. (2010) ATP-independent reversal of a membrane protein aggregate by a chloroplast SRP subunit. *Nat. Struct. Mol. Biol.* **17**, 696-702
15. Jaru-Ampornpan, P., Shen, K., Lam, V.Q., Ali, M., Doniach, S., Jia, T.Z., and Shan, S. (2010) ATP-independent reversal of a membrane protein aggregate by a chloroplast SRP subunit. *Nat. Struct. Mol. Biol.* **17**, 696-702
16. Falk, S., and Sinning, I. (2010) cpSRP43 is a novel chaperone specific for light-harvesting chlorophyll a,b-binding proteins. *J. Biol. Chem.* **285**, 21655-21661
17. Nguyen, T. X., Jaru-Ampornpan, P., Lam, V.Q., Cao, P., Piszkiwicz, S., Hess, S., and Shan, S. (2013) Mechanism of an ATP-independent protein disaggregase. Part I. Structure of a membrane protein aggregate reveals a mechanism of recognition by its chaperone. *J. Biol. Chem.* **288**, 13420-13430

18. Jaru-Ampornpan, P., Liang, F.-C., Nisthal, A., Nguyen, T. X., Wang, P., Shen, K., Mayo, S. L., and Shan, S. (2013) Mechanism of an ATP-independent protein disaggregase: Part II. Distinct molecular interactions drive multiple steps during aggregate disassembly. *J. Biol. Chem.* **288**, 13431-13445
19. Jansson, S. (1999) A guide to the Lhc genes and their relatives in Arabidopsis. *Trends Plant Sci.* **4**, 236–240
20. Liu, Z., Yan, H., Wang, K., Kuang, T., Zhang, J., Gui, L., An, X., and Chang, W. (2004) Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature* **428**, 287-292
21. Amin, P., Sy, D.A.C., Pilgrim, M.L., Parry, D.H., Nussaume, L., and Hoffman, N.E. (1999) Arabidopsis mutants lacking the 43- and 54-kilodalton subunits of the chloroplast signal recognition particle have distinct phenotypes. *Plants Physiology* **121**, 61-70
22. Klimyuk, V. I., Persello-Cartieaux, F., Havaux, M., Contard-David, P., Schuenemann, D., Meierhoff, K., Gouet, P., Jones, J.D.G., Hoffman, N.E., and Nussaume, L. (1999) 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. *Plant Cell* **11**, 87-99
23. Lee, J., Culyba, E.K., Powers, E.T., and Kelly, J.W. (2011) Amyloid-β forms fibrils by nucleated conformational conversion of oligomers. *Nat. Chem. Biol.* **7**, 602-609
24. Abramoff, M. D., Magelhaes, P.J. and Ram, S.J. (2004) Image processing with ImageJ. *Biophotonics Int.* **11**, 36–42
25. Jaru-Ampornpan P., L. F.-C., Nisthal A., Nguyen T. X., Wang P., Shen K., Mayo S. L., and Shan S. (2013) Mechanism of an ATP-independent protein disaggregase: Part II. Distinct molecular interactions drive multiple steps during aggregate disassembly. *J. Biol. Chem.* **288**, 13431-13445
26. Laage, R., and Langosch, D. (2001) Strategies for prokaryotic expression of eukaryotic membrane proteins. *Traffic* **2**, 99-104
27. Tate, C. G. (2001) Overexpression of mammalian integral membrane proteins for structural studies. *FEBS Lett.* **31**, 94-98
28. Su, P. C., Si, W., Baker, D.L., and Berger, B.W. (2013) High-yield membrane protein expression from E. coli using an engineered outer membrane protein F fusion. *Protein Sci.* **22**, 434-443
29. Chartron, J. W., VanderVelde, D.G., and Clemons, W.M. (2012) Structures of the Sgt2/SGTA dimerization domain with the Get5/UBL4A UBL domain reveal a novel interaction that forms a conserved dynamic interface. *Cell Rep.* **2**, 1620-1632
30. Rome, M. E., Rao, M., Clemons, W.M., and Shan, S. (2013) Precise timing of ATPase activation drives targeting of tail-anchored proteins. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 7666-7671
31. LeVine, H. (1999) Quantification of beta-sheet amyloid fibril structures with thioflavin T. *Methods Enzymol.* **309**, 274–284
32. Bieschke, J., Zhang, Q., Powers, E.T., Lerner, R.A., and Kelly, J.W. (2005) Oxidative metabolites accelerate Alzheimer's amyloidogenesis by a two-step mechanism, eliminating the requirement for nucleation. *Biochemistry* **44**, 4977-4983
33. Goldsbury, C. S., Wirtz, S., Müller, S.A., Sunderji, S., Wicki, P., Aebi, U., and Frey, P. (2000) Studies on the *in vitro* assembly of a beta 1-40: implications for the search for a beta fibril formation inhibitors. *J. Struct. Biol.* **130**, 217-231

34. Schmidt, M., Sachse, C., Richter, W., Xu, C., Fändrich, M., and Grigorieff, N. (2009) Comparison of Alzheimer Abeta(1-40) and Abeta(1-42) amyloid fibrils reveals similar protofilament structures. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 19813-19818
35. Ehrnhoefer, D. E., Bieschke, J., Boeddrich, A., Herbst, M., Masino, L., Lurz, R., Engemann, S., Pastore, A., and Wanker, E.E. (2008) EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nat. Struct. Mol. Biol.* **15**, 558-566
36. Narayan, P., Orte, A., Clarke, R.W., Bolognesi, B., Hook, S., Ganzinger, K.A., Meehan, S., Wilson, M.R., Dobson, C.M., and Klenerman, D. (2011) The extracellular chaperone clusterin sequesters oligomeric forms of the amyloid- $\beta$ (1-40) peptide. *Nat. Struct. Mol. Biol.* **19**, 79-83
37. Wagner, S., Baars, L., Ytterberg, A.J., Klussmeier, A., Wagner, C.S., Nord, O., Nygren, P.A., van Wijk, K.J., de Gier, J.W. (2007) Consequences of membrane protein overexpression in *Escherichia coli*. *Mol. Cell Proteomics* **6**, 1527-1550