

POST-TRANSLATIONAL MEMBRANE PROTEIN TARGETING BY  
THE CHLOROPLAST SIGNAL RECOGNITION PARTICLE

Thesis by

Peera Jaru-Ampornpan

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2011

(Defended April 29, 2011)

© 2011

Peera Jaru-Ampornpan

All Rights Reserved

## ACKNOWLEDGEMENT

First, I would like to thank my thesis advisor, Shu-ou Shan, for guidance during my graduate work. Shu-ou teaches by example of her rigor and depth in research, as well as her commendable work ethic. She sets high standards for her students, and I am fortunate to have an opportunity to try to meet those and learn many valuable lessons along the way. I am also grateful to my thesis committee, Professors Doug Rees, David Chan and Bil Clemons, for the insightful discussions about my projects we have enjoyed over the years.

I am privileged to have had opportunities to work with the brilliant scientists, who are also wonderful friends, of the Shan lab: Sowmya Chandrasekar, Xin Zhang, Rumana Rashid, Vinh Lam, Thang Nguyen, Kuang Shen, Ishu Saraogi, Aileen Ariosa, David Akopian, Dawei Zhang, Nathan Pierce, and Mike Rome. Special thanks to Sowmya and Thang who have weathered the Monday meetings together with me as Team Chloroplast for the last three years and have helped me proofread parts of this dissertation. Xin has taught me many useful skills, scientific and otherwise, and I thank him for those lessons and his friendship.

Collaboration has been an important part of my research experience at Caltech. I thank Zhenfeng Liu for his invaluable advice and discussion on LHCP and experiments, Peigen Cao for his help in AFM imaging, Alasdair McDowell for his help in TEM imaging, Alex Nisthal and his Robot for the LHCP and cpSRP43 scanning mutagenesis project, and Troy Walton for his help in the multi-angle light scattering trials. I have enjoyed working with two talented SURF students, Tony Jia and Jack Wang, and wish them the best for their future endeavors.

Many supporting staff members have helped smooth my Caltech education experience and become my good friends. I thank Margot Hoyt for being the world's best secretary and beyond; Alison Ross for all the administrative issues regarding the BMB option; Santiago Laparra for many liters of bacterial culture media; Tom Dunn, Jerry Becker, and Scott Smith for promptly fixing various instruments; Joe Drew, Ron Cohen, and Steve Gould of Chemistry purchasing who dealt with many odd purchasing requests for me; and Leslie Fox and Jill Brooks for keeping me sane with their exercise classes. In addition, we have had help from the high-school volunteers, Kevin Kim, Catherine Chan, Grace Qi, Joe Tung, and Irene Chen. I thank them for their hard work and the way their youth has brightened the lab, and wish them the best for their future endeavors.

I could not have done this without personal support from many friends and family. I thank my parents and my sister for being the coolest family I know. I thank all my friends, especially the Thai Scholar network, for the wonderful friendship they have shared with me. Last, I thank my best friend, Knot Pipatsrisawat, for the amazing journey we have together enjoyed.

## ABSTRACT

Post-translational transport of membrane proteins poses enormous challenges to the cells. The transport factors must accurately select and deliver the cargos to the appropriate target membranes. In addition, they have to provide chaperone for their hydrophobic cargos. To understand capacity and limitation of a post-translational transport factor, we studied one of the most efficient membrane protein transport pathways, the delivery of light-harvesting chlorophyll-binding (LHC) proteins to the thylakoid membrane. This targeting reaction is mediated by the chloroplast Signal Recognition Particle (cpSRP) and its receptor. Although the core SRP GTPases are close homologues of those in cytosolic SRP pathways, the unique features of cpSRP that might reflect its adaptation to the challenges in post-translational targeting include (i) the lack of the otherwise universally conserved SRP RNA, and (ii) the exclusive presence of a novel protein, cpSRP43. In the first part of this thesis, we define the thermodynamic and kinetic framework for the GTPase cycles of cpSRP and its receptor and uncover the molecular bases that enable their intrinsically fast interactions, such that they can bypass an SRP RNA, an essential accelerator for the cytosolic SRP–receptor interaction. The second part of the thesis is devoted to characterization of the chaperone function of cpSRP43. We show that cpSRP43 specifically and effectively prevents and reverses the aggregation of its cargo, LHC proteins. We further investigate the molecular mechanism of this novel disaggregase activity, using a combination of biochemical and structural approaches. In summary, this dissertation aims to understand how cpSRP and its receptor adapt to their unique requirements in efficiently transporting a family of highly abundant membrane proteins.

## TABLE OF CONTENTS

Acknowledgement.....	iii
Abstract.....	v
Table of Contents.....	vi
Summary.....	1
Chapter 1:	
Efficient Interaction Between Two GTPases Allows the Chloroplast SRP Pathway to Bypass the Requirement for an SRP RNA.....	9
Chapter 2:	
A Distinct Mechanism to Achieve Efficient SRP–SRP Receptor Interaction by the Chloroplast SRP Pathway.....	56
Chapter 3:	
ATP-Independent Reversal of a Membrane Protein Aggregate by a Chloroplast SRP Subunit.....	93
Chapter 4:	
Biochemical and Biophysical Characterization of the LHC Protein Aggregates...	140
Chapter 5:	
Probing the Mechanism of cpSRP43-Mediated Protein Disaggregation.....	166

# Summary

Compartmentalization offers cells the efficiency and the flexibility to simultaneously carry out vastly different chemical reactions in contained environments. Maintaining specific contents of cellular compartments requires precise control of protein transport from the cytosolic space of the cells, where most proteins are made, to the various organelles. Especially challenging is the transport of membrane proteins, which often are hydrophobic and hence are prone to aggregation during production and transport in aqueous environment. Therefore, cells devote their resources into membrane protein transport to ensure proper protein localization and prevent protein aggregation. Many pathways exist to transport different groups of membrane proteins with great efficiency and accuracy.

A major targeting pathway for membrane proteins is the Signal Recognition Particle (SRP) pathway. SRP handles about one-third of all the cellular proteins and targets secretory and membrane proteins to the endoplasmic reticulum membrane in eukaryotic cells or the plasma membrane in prokaryotic cells (1, 2). Although the size and the composition of the SRP machineries vary in different species, they share the core components. These include the SRP GTPases in the SRP and the SRP receptor (SR) and the SRP RNA that forms a complex with the SRP GTPase (1). The cytosolic SRP recognizes its cargos, the translating ribosomes, by binding to the signal sequences (3). Because of its co-translational nature of targeting, cytosolic SRP minimizes the exposure of hydrophobic regions of its membrane protein substrates. The first transmembrane

helix is taken as a signal sequence and is sequestered by direct binding to SRP. The rest of the protein is still either inside the ribosome tunnel or is not yet translated during the delivery by SRP (2, 4). Through its interactions with SR, SRP brings its cargo to the target membrane. These interactions involve an extensive series of conformational rearrangements that are modulated by unique environmental cues provided by the cargos and the membrane (5, 6, 7). Thereby, protein targeting by SRP achieves exquisite spatial and temporal regulations.

Recent discovery of an SRP-mediated targeting pathway in chloroplasts has brought some surprises (8, 9). As opposed to the cytosolic SRP that strictly mediates co-translational targeting, chloroplast SRP (cpSRP) post-translationally delivers its cargos, the nuclear-encoded light-harvesting chlorophyll a/b-binding (LHC) family of proteins, to the thylakoid membrane (8). The core SRP GTPases in the cpSRP pathway are highly homologous to those from the cytosolic SRP pathway, sharing over 60% sequence similarity (10, 11). Unexpectedly, however, the otherwise universally conserved SRP RNA has not been found in the cpSRP pathway. Instead, cpSRP is composed of two protein subunits: cpSRP54, which is an SRP GTPase that interacts with the reciprocal chloroplast SR GTPase (10), and cpSRP43, which is a novel protein unique to chloroplasts and functions in cargo recognition (12, 13, 14). Similar to cytosolic SRP, cpSRP brings its cargo to the thylakoid membrane via its interaction with the SR GTPase (15). In the thylakoid membrane, the LHC proteins bind photosynthetic pigments and form light-harvesting complexes that allow the photosynthetic reaction centers to capture light efficiently (16, 17). Present in all green plants, LHC proteins are arguably the most abundant membrane protein on earth (18). Due to the abundance and the essentiality of

the LHC proteins, the transport system that handles them must be highly effective. Moreover, the post-translational nature of targeting further complicates the situation, adding a requirement for a potent chaperone to prevent the aggregation of the hydrophobic LHC proteins during targeting. cpSRP must therefore adapt itself to accommodate these specialized needs. We believe that these “surprises”, the drastic deviations from the cytosolic SRP pathways, reflect such adaptations. This dissertation aims to understand the molecular bases of these adaptations that enable cpSRP to meet the challenges in LHC proteins transport.

First, the otherwise universally conserved SRP RNA is absent from cpSRP. In the cytosolic SRP systems, the SRP-SR interaction is extremely slow and involves multiple discrete conformational changes (5, 19). One of the important roles of the SRP RNA is to accelerate complex formation between the SRP and the SR GTPases by 400-fold, bringing their interaction rate to a physiological range for protein targeting (20, 21). How, then, might this seemingly crucial molecule be left out in the cpSRP system? More specifically, how do cpSRP GTPases that are highly homologous to the slowly-interacting cytosolic SRP GTPases manage the high-load protein targeting without the interaction accelerator? The first part of this dissertation aims to answer these questions.

Chapter 1 shows that the cpSRP pathway bypasses the SRP RNA but still sustains the fast rate of SRP-SR association required for efficient targeting. Indeed, the rate of SRP-SR complex assembly for the chloroplast proteins is 400-fold faster than that of the *E. coli* SRP proteins, matching the interaction rate of the bacterial proteins in the presence of the SRP RNA. The intrinsically fast complex assembly stems partly from the pre-organization of the free cpSRP receptor into a conformation that is conducive for

complex formation. This is not the case for the cytosolic SRP receptor, which undergoes this conformational change only upon binding to SRP (22). Therefore, an energy barrier in complex formation is readily reduced for cpSRP-cpSR interaction.

Chapter 2 reveals the molecular bases of this fast interaction. Cross-complementation study using the GTPases from both the chloroplast and *E. coli* pathways confirms the pre-organization of the cpSRP receptor and uncovers the stimulatory role of the cargo-binding “M” domain of cpSRP. The M-domain of *E. coli* SRP, which is the binding site for both the SRP RNA and the signal sequence, does not significantly stimulate the interactions of the cytosolic GTPases unless the SRP RNA is bound to it (23). On the contrary, the M-domain of cpSRP, which is the binding site for cargo-binding cpSRP43, raises the interaction rate by 50–100-fold. Therefore, the M-domain of cpSRP subsumes the function of the SRP RNA and eliminates the need for an external regulator for the chloroplast SRP-SR interaction.

Second, cpSRP43 is found exclusively in chloroplast SRPs and has no known homologues in other eukaryotic SRPs. *Chaos*, the cpSRP43-null mutant, has yellow leaves and sustains growth defects, implicating its vital role in LHC biogenesis (24). Later biochemical work detected its direct binding to the conserved 18-amino acid motif (called L18) on LHC proteins (12, 13, 14). Therefore, cpSRP43 is a cargo recognition module specific for cpSRP. Since the nature of post-translational targeting requires a potent chaperone for the membrane protein substrates, it is inviting to speculate that cpSRP43 assumes this chaperone function. The second part of this dissertation explores this idea and is devoted to characterizing the chaperone aspect of cpSRP43.

Chapter 3 discusses the role of cpSRP43 as a potent specific chaperone for the cargo. cpSRP43 can effectively prevent LHC proteins from aggregating in aqueous environment using the specific binding interactions that it establishes with its substrate. Despite its small size, cpSRP43 is an elongated molecule that can provide potential binding surfaces for LHC proteins. Serendipitously, we discovered that cpSRP43 also has a specific disaggregase activity toward the aggregates formed by its substrate protein. Using no external energy input, cpSRP43 can actively reverse aggregation of LHC proteins with speed and efficiency that rival the ATPase-based disaggregases. This is an exciting discovery because, to our knowledge, cpSRP43 represents a rare example of a class of novel disaggregases that utilize binding energy to dissolve insoluble protein aggregates.

Chapters 4 and 5 explore cpSRP43–LHC protein aggregates as a model system for studying a novel mechanism for protein disaggregation. Chapter 4 focuses on the biochemical and biophysical characterization of the aggregates formed by LHC proteins in an attempt to understand the basic nature of the aggregates handled by cpSRP43. We found that these aggregates are disc-shaped micelles with a  $\sim 12$  nm diameter, and they are thermodynamically and kinetically stable. Chapter 5 extends the mechanistic study of the cpSRP43-mediated disaggregation reaction and outlines the molecular requirements of the LHC protein aggregates for efficient disaggregation by cpSRP43. The disaggregation reaction can be dissected into at least two steps: initial binding to the aggregate and subsequent re-solubilization. For the first step, the aggregate must present the cpSRP43 recognition element, L18, on its surface. For the subsequent re-solubilization step, the overall binding energy provided by cpSRP43 must be enough to

overcome the internal packing interactions of the aggregates. Although future work is required, these studies allow us to start probing the capabilities and the limitations of this new disaggregase activity.

In summary, this dissertation aims to explain the “adaptations” made to satisfy special needs in the cpSRP pathway. At the center, the cytosolic and the chloroplast SRP pathways share highly homologous SRP GTPases. However, each system has evolved distinct mechanisms to enable them to fulfill their functions. The cytosolic SRP must select its substrates from a vast pool of translating ribosomes; the pathway has built-in multiple potential regulatory points, making use of the extensive conformational rearrangements of the SRP GTPases (7). The indispensable SRP RNA plays a regulatory role and serves as the molecular link between the correct cargos and the GTPases (25). On the other hand, the chloroplast SRP pathway, handling one conserved family of proteins, opts for efficiency by bypassing some conformational rearrangements and foregoing external regulators. To cope with the unique problem of protein aggregation during post-translational targeting, cpSRP43 has been evolved as an add-on cargo-binding module for cpSRP to provide effective chaperone during LHC transport. The cytosolic SRP and cpSRP represent only two examples of the numerous membrane protein targeting pathways. Other pathways, such as the GET pathway that delivers tail-anchored proteins (26, 27) or the bacterial SecB–SecA pathway that delivers bacterial outer membrane proteins (28), certainly have different requirements and have evolved distinct mechanisms to handle their own substrates. Lessons learned from this dissertation, as well as from numerous comparative studies, emphasize the versatility of nature in its capability to cater to different biological needs.

**References:**

1. Keenan, R. J., Freymann, D. M., Stroud, R. M., and Walter, P. (2001) The signal recognition particle, *Annu. Rev. Biochem.* 70, 755–775.
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. Walter, P., and Blobel, G. (1981) Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes, *J. Cell Biol.* 91, 557–561.
4. Zheng, N, and Gierasch, L. M. (1996) Signal sequences: the same yet different, *Cell* 86, 849–852.
5. Shan, S., Stroud, R., and Walter, P. (2004) Mechanism of association and reciprocal activation of two GTPases, *PLoS Biology* 2, e320.
6. Zhang, X., Schaffitzel, C., Ban, N., and Shan, S. (2009) Multiple conformational switches in a GTPase complex control co-translational protein targeting, *Proc. Natl. Acad. Sci. USA* 106, 1754–1759.
7. Zhang, X., Rashid, R., Wang, K., and Shan, S. (2010) Sequential checkpoints govern substrate selection during cotranslational protein targeting., *Science* 328, 757–760.
8. Schuenemann, 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. USA* 95, 10312–10316.
9. Schuenemann, D. (2004) Structure and function of the chloroplast signal recognition particle *Curr. Genet.* 44, 295–304.
10. Li, X., Henry, R., Yuan, J., Cline, K., and Hoffman, N.E. (1995) A chloroplast homologue of the signal recognition particle subunit SRP54 is involved in the posttranslational integration of a protein into thylakoid membranes, *Proc. Natl. Acad. Sci. USA* 92, 3789–3793.
11. 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.
12. 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.
13. 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.

14. 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.
15. Jaru-Ampornpan, P., Chandrasekar, S., 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.
16. Peter, G. F., and Thornber, J.P. (1991) Biochemical composition and organization of higher plant photosystem II light-harvesting pigment-proteins, *J. Biol. Chem.* 266, 16745–16754.
17. 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.
18. Jansson, S. (1999) A guide to the Lhc genes and their relatives in arabidopsis, *Trends Plant Sci.* 4, 236–240.
19. Zhang, X., Kung, S., and Shan, S. (2008) Demonstration of a multistep mechanism for assembly of the SRP- SRP receptor complex: implications for the catalytic role of SRP RNA, *J Mol Biol.* 381(3), 581–593.
20. Peluso, P., Herschlag, D., Nock, S., Freymann, D. M., Johnson, A. E., and Walter, P. (2000) Role of 4.5S RNA in assembly of the bacterial signal recognition particle with its receptor, *Science* 288, 1640–1643.
21. 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.
22. Shan, S., and Walter, P. (2003) Induced Nucleotide Specificity in a GTPase, *Proc. Natl. Acad. Sci. U.S.A.* 100, 4480–4485.
23. Bradshaw, N., and Walter, P. (2007) The signal recognition particle (SRP) RNA links conformational changes in the SRP to protein targeting., *Mol Biol Cell* 18, 2728–2734.
24. 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, *The Plant Cell* 11, 87–99.
25. Bradshaw, N., Neher, S.B., Booth, D.S., and Walter, P. (2009) Signal sequences activate the catalytic switch of SRP RNA, *Science* 323, 127–130.
26. Stefanovic, S., and Hegde, R.S. (2007) Identification of a targeting factor for posttranslational membrane protein insertion into the ER, *Cell* 128, 1147–1159.
27. Schuldiner, M., Metz, J., Schmid, V., Denic, V., Rakwalska, M., Schmitt, H.D., Schwappach, B., and Weissman, J.S. (2008) The GET complex mediates insertion of tail-anchored proteins into the ER membrane, *Cell* 134, 634–645.
28. 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.

# **Chapter 1:**

## **Efficient Interaction Between Two GTPases**

### **Allows the Chloroplast SRP Pathway to Bypass**

### **the Requirement for an SRP RNA**

A version of this chapter has been published as:

Jaru-Ampornpan, P., Chandrasekar, S., and Shan, S. (2007) *Mol. Biol. Cell*, **18** (7), 2636–2545.

**Abstract**

Co-translational protein targeting to membranes is regulated by two GTPases in the signal recognition particle (SRP) and the SRP receptor (SR); association between the two GTPases is slow and is accelerated 400-fold by the SRP RNA. Intriguingly, the otherwise universally conserved SRP RNA is missing in a novel chloroplast SRP pathway. We found that even in the absence of an SRP RNA, the chloroplast SRP and SR can interact efficiently with one another; the kinetics of interaction between the chloroplast GTPases is 400-fold faster than their bacterial homologues and matches the rate at which the bacterial SRP and SR interact with the help of SRP RNA. Biochemical analyses further suggest that the chloroplast FtsY is pre-organized in a conformation that allows optimal interaction with its binding partner, so that conformational changes during complex formation are minimized. Our results highlight intriguing differences between the classical and chloroplast SRP and SR GTPases, and help explain how the chloroplast SRP pathway can mediate efficient targeting of proteins to the thylakoid membrane in the absence of the SRP RNA, which plays an indispensable role in all the other SRP pathways.

## Introduction

SRP and SR comprise the major cellular machinery that delivers nascent proteins to the eukaryotic endoplasmic reticulum membrane or the bacterial plasma membrane (1, 2). The functional core of SRP is the SRP54 protein (called Ffh in bacteria) in complex with an SRP RNA, which recognizes the cargo protein and interacts with the SR (called FtsY in bacteria). The protein targeting reaction is regulated by the guanosine-5'-triphosphate (GTP)-binding domains in both SRP54 and SR. SRP recognizes the signal sequence on nascent polypeptides that emerge from a translating ribosome (3). The ribosome•nascent chain complex is delivered to the membrane via the interaction of SRP with SR when both proteins are bound with GTP (4, 5). Upon arrival at the membrane, SRP releases the cargo protein to a protein conducting channel embedded in the membrane (6, 7), where the nascent protein is either integrated into the membrane or translocated across the membrane to enter the secretory pathway. GTP hydrolysis is stimulated in the SRP•SR complex, which then drives disassembly and recycling of SRP and SR (8).

The SRP and SR GTPases comprise a unique subgroup in the GTPase superfamily (1). Both proteins have a GTPase, “G” domain that shares homology with the classical *Ras* GTPase fold (9, 10). In addition, the SRP-type GTPases contain an N-terminal four-helix bundle, the “N”-domain, that packs tightly against the G domain. The G- and N-domains form a structural and functional unit called the NG domain. Unlike classical signaling GTPases that undergo large conformational changes depending on whether GTP or guanosine 5-diphosphate (GDP) is bound, the structures of these GTPases are similar regardless of which nucleotide is bound (10, 11). Substantial

conformational changes occur only when the two GTPases form a complex with one another (12, 13). Most notably, the G- and N-domains readjust their relative positions such that the N-domains of both proteins move closer to the dimer interface and form additional interface contacts to stabilize the complex.

The importance of this N-G domain rearrangement is supported by biochemical analyses. Many mutations at this interface disrupt SRP–SR complex formation and protein targeting (14). Interestingly, unlike classical GTPases, free FtsY displays little discrimination between GTP and non-cognate nucleotides. In contrast, FtsY acquires substantial nucleotide specificity only when it binds SRP. These results have led to the proposal that during complex formation, FtsY changes from a non-discriminative, “open” state to a “closed” state in which specific interactions between GTP and active site residues are established (15). Consistent with these observations, the crystal structure showed that, upon complex formation, the rearrangement at the N-G domain interface brings the nucleotide specificity determinant, Asp449, closer to the bound GTP and within hydrogen bonding distance with the amino group of the guanine ring (12). Thus, the N-G domain rearrangement is primarily responsible for the open → closed conformational change that occurs during SRP–SR complex formation and precisely aligns active site residues with respect to the bound GTP.

The unique structural features of the SRP subgroup of GTPases confer upon them many characteristics that are distinct from canonical GTPases. Most importantly, SRP-type GTPases bind nucleotides much more weakly than signaling GTPases and release nucleotides quickly (16–19). Therefore, they do not employ nucleotide exchange factors to facilitate the conversion from the GDP- to the GTP-bound form. These GTPases also

do not utilize external GTPase-activating proteins; instead, SRP and SR reciprocally activate one another upon complex formation (20).

A third unique feature of the GTPases engaged in the SRP pathway is the requirement for a universally conserved SRP RNA. Mammalian SRP is a cytosolic ribonucleoprotein complex that consists of 6 polypeptides and a 7S SRP RNA molecule. Besides SRP54, the other protein components are not conserved, whereas the SRP RNA has been shown to play an indispensable role in protein targeting in all three kingdoms of life. In early biochemical studies on the mammalian SRP, the SRP RNA appeared to be nothing more than a scaffold that holds all the SRP proteins together in a complex (21, 22). The finding that bacteria contain a much simpler SRP, comprised solely of a complex of Ffh and the 4.5S SRP RNA, was therefore intriguing. This smaller RNA contains the most phylogenetically conserved region of the SRP RNA, domain IV, which is likely to have been maintained for functional purposes (23, 24). Subsequently, kinetic analyses of the role of the 4.5S SRP RNA on the GTPase cycles of Ffh and FtsY showed that a major role of this RNA is to accelerate complex formation between the two GTPases. In the absence of the SRP RNA, Ffh–FtsY association is extremely slow, with a rate constant of  $5 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$ . The SRP RNA accelerates their association kinetics by 400-fold, to a rate that can allow the SRP and SRP receptor to adequately carry out their biological functions, thus accounting for the indispensable role of the SRP RNA in the bacterial, archeal, and eukaryotic SRP pathways.

A novel SRP targeting pathway was discovered in the chloroplast (25). cpSRP54 and cpFtsY are the chloroplast homologues of SRP and SR GTPases, respectively (26–28). cpSRP54 recognizes its cargo, the light-harvesting chlorophyll-binding proteins, via

a protein adaptor cpSRP43 (29). Together, cpSRP54 and cpSRP43 deliver the cargo protein from the stroma to the thylakoid membrane via the GTP-dependent interaction between cpSRP54 and cpFtsY (28). Surprisingly, the otherwise universally conserved SRP RNA has not been found to date in the chloroplast SRP system. To rationalize the absence of the SRP RNA, we characterized the kinetic and thermodynamic features of the GTPase cycles of cpSRP54 and cpFtsY. We found that, unlike their bacterial and mammalian homologues, the chloroplast SRP and SR GTPases can efficiently interact with one another by themselves. This helps explain why the cpSRP pathway could bypass the requirement for an SRP RNA.

## Materials and Methods

**Protein expression and purification.** cpSRP54 from *A. thaliana* was expressed from baculovirus at the Protein Expression Facility of Caltech. Recombinant cpSRP54 is purified by affinity chromatography using Ni-NTA (Qiagen) and cation exchange over a MonoS column (GE Healthcare) using a linear gradient of 150–600 mM NaCl. cpFtsY from *A. thaliana* was expressed and purified as described (30). Two additional chromatographic steps [Superdex 75 and monoQ (GE Healthcare)] were added to remove contaminating GTPases. Mutant cpFtsY(D283N) was constructed using the QuickChange procedure (Stratagene) and was expressed and purified by the same procedure as that for wild-type cpFtsY.

**Kinetics.** All reactions were carried out at 25 °C in assay buffer [50 mM KHEPES (pH 7.5), 150 mM KOAc, 2mM Mg(OAc)<sub>2</sub>, 2mM DTT, 0.01% Nikkol]. GTP hydrolysis reactions were followed and analyzed as described (19). The general procedures for characterizing the basal and stimulated GTPase reactions between SRP and SR have been described in detail (15, 19, 31) and are summarized briefly here. The justification for how each microscopic rate constant was derived from these measurements is provided in Supplementary Material.

Basal GTPase or XTPase activities of cpSRP54, cpFtsY, and cpFtsY(D283N) were measured in single-turnover reactions as described ([GTP]  $\ll$  [E]; (19)). The dependence of the observed rate constant ( $k_{obsd}$ ) on protein concentration were fit to eq 1, in which  $k_{max}$  is the maximal rate constant at saturating protein concentrations, and  $K_{1/2}$  is the protein concentration required to reach half the maximal rate.

$$k_{obsd} = k_{max} \times \frac{[protein]}{K_{1/2} + [protein]} \quad (1)$$

The nucleotide affinities of the GTPases were determined using several independent methods. The GTP affinities for cpSRP54 and cpFtsY and the XTP (xanthosine-5'-triphosphate) affinity for cpFtsY(D283N) were obtained from the  $K_{1/2}$  values obtained in the fits of the basal GTPase or XTPase reactions to eq 1. Because the chemical step is rate-limiting for the basal GTPase and XTPase reactions,  $K_{1/2}$  is equal to  $K_d$ , the dissociation constant of the nucleotide. The affinities of GDP, GppNHp (5'-guanylylimido-diphosphate), XDP (xanthosine-5'-diphosphate) and XppNHp (5'-xanthyllylimido-diphosphate) were determined using these nucleotides as inhibitors of the basal GTPase or XTPase reactions (19). With sub-saturating protein, the inhibition constant  $K_i$  is equal to  $K_d$ . Finally, the binding of nucleotides to the GTPases was determined directly by using fluorescent N-methyl-anthraniloyl (mant) derivatives of GTP, GDP, and XTP, as described below.

The reciprocally stimulated GTPase reaction between cpSRP54 and cpFtsY was determined in multiple turnover reactions ( $[GTP] \gg [E]$ ) in the presence of a small, fixed amount of cpSRP54 and varying concentrations of cpFtsY, using a GTP concentration that saturates both GTPase sites. The concentration dependence of the observed rate constant ( $k_{obsd}$ ) is fit to eq 2, in which  $k_{cat}$  is the rate constant at saturating cpFtsY concentrations, and  $K_m$  is the

$$k_{obsd} = k_{cat} \times \frac{[cpFtsY]}{K_m + [cpFtsY]} \quad (2)$$

concentration of cpFtsY that gives half the maximal rate. The stimulated GTPase reaction between cpSRP54 and GTP-bound cpFtsY(D283N) was determined using the same experimental setup. The stimulated GTPase reaction of cpSRP54 by XTP-bound cpFtsY(D283N) was determined analogously, except that the concentration of GTP and XTP were adjusted such that cpSRP54 was predominantly occupied by GTP whereas cpFtsY(D283N) was predominantly occupied by XTP.

The cpFtsY(D283N)-stimulated GTP hydrolysis from cpSRP54 was also determined in single turnover experiments. The hydrolysis of trace GTP\* was monitored in the presence of sub-saturating cpSRP54 and varying amounts of cpFtsY(D283N), with 25  $\mu$ M XTP present to selectively occupy the active site of cpFtsY(D283N). Under these conditions, the third-order reaction:  $\text{GTP}^* + \text{cpSRP54} + \text{cpFtsY(D283N)} \xrightarrow{\text{XTP}}$  products was followed. The reciprocal reaction,  $\text{XTP}^* + \text{cpFtsY(D283N)} + \text{cpSRP54} \xrightarrow{\text{GTP}}$  products was determined using an analogous setup, except that the concentration of cpSRP54 was varied and 25  $\mu$ M GTP was present to selectively occupy the active site of cpSRP54. The data were fit to eq 1 above. Finally, first-order rate constants of the stimulated GTP and XTP hydrolysis reactions from the  $^*\text{GTP} \cdot \text{cpSRP54} \cdot \text{cpFtsY(D283N)} \cdot \text{XTP}^*$  complex were determined using high concentrations of both proteins (20–80  $\mu$ M) in the presence of stoichiometric amounts of their respective nucleotides. The reaction time courses were monitored in a Kintek quench flow apparatus and fit to a single-exponential rate equation to obtain the first-order rate constants.

The effect of XTP on the reaction  $^*\text{GTP} \cdot \text{cpSRP54} + \text{cpFtsY(D283N)} \xrightarrow{\text{GTP}^*}$  products was determined in the presence of sub-saturating concentrations of both proteins

and a high concentration of GTP (200  $\mu\text{M}$ ) to saturate both active sites. The XTP concentration dependence was fit to eq 3, in which  $k_0$  is the rate constant in the absence of any inhibitor,  $k_1$  is the rate

$$k_{obsd} = k_0 \times \frac{K_i^{app}}{K_i^{app} + [\text{XTP}]} + k_1 \times \frac{[\text{XTP}]}{[\text{XTP}] + K_i^{app}} \quad (3)$$

constant at infinite XTP concentrations, and  $K_i^{app}$  is the apparent inhibition constant of XTP determined from this experiment.  $K_i^{app}$  is related to the dissociation constant of XTP by eq 4,

$$K_i^{app} = K_d^{XTP} \times \left( 1 + \frac{[\text{GTP}]}{K_d^{GTP}} \right) \quad (4)$$

in which  $K_d^{XTP}$  and  $K_d^{GTP}$  are the dissociation constants of XTP and GTP for cpFtsY(D283N), respectively.

**Fluorescence.** All fluorescence measurements were conducted at 25 °C using the single-photon-counting Fluorolog 3-22 spectrofluorometer (Jobin Yvon). Fluorescence emission spectra of mant-derivatives of GTP, GDP, and XTP were acquired using an excitation wavelength of 356 nm. Nucleotide binding affinities were determined by recording the change in fluorescence intensity at 445 nm in the presence of 0.4–1  $\mu\text{M}$  mant-nucleotides and increasing concentrations of cpSRP54, cpFtsY, or cpFtsY(D283N).

The data were fit to eq 5,

$$F_{obsd} = (F_{max} - F_0) \times \frac{[\text{protein}]}{K_d + [\text{protein}]} \quad (5)$$

in which  $F_{max}$  is the fluorescence at saturating protein concentrations,  $F_0$  is the fluorescence in the absence of any protein, and  $K_d$  is the dissociation constant of the mant-nucleotide.

The rate constants for dissociation of mant-GTP and mant-GDP were determined using a pulse chase experiment as described (16). The time course for decay of fluorescence was followed in a stopped flow apparatus (Applied Photophysics) and fit to single exponential functions to obtain the dissociation rate constants.

## Results

To understand why and how the cpSRP pathway bypasses the requirement for an SRP RNA, which plays a critical role in facilitating the interaction between the SRP and SR GTPases in all the other SRP pathways, we characterized the rate and equilibrium of the individual steps in the GTPase cycles of cpSRP54 and cpFtsY and their GTP-dependent interaction with one another (Figure 1.1). Each protein can bind and hydrolyze GTP by itself (steps 1–3 for cpSRP54 and 1'–3' for cpFtsY). cpSRP54 form a stable complex with cpFtsY when both proteins are bound with GTP (step 4). Both GTP molecules are rapidly hydrolyzed from the complex (step 5). GTP hydrolysis destabilizes the complex and drives its dissociation (step 6). The rate and equilibrium constants for each step are summarized in Table 1.1. For simplicity, additional possibilities such as hydrolysis of one of the GTPs followed by complex disassembly are not shown; these possibilities are presented in the Discussion.

### *Basal GTPase cycles of cpSRP54 and cpFtsY.*

We first determined the basal GTPase activities of the individual GTPases. Both proteins hydrolyze GTP slowly, with maximal hydrolysis rates of 0.017 and 0.0045 min<sup>-1</sup> at saturating protein concentrations for cpSRP54 and cpFtsY, respectively (Figure 1.2). The protein concentration dependence of the hydrolysis rate gives the affinity of each protein for GTP. Both GTPases bind their substrates weakly, with dissociation constants of 2.8 and 2.1 μM for cpSRP54 and cpFtsY, respectively (Figure 1.2). We also determined the affinities of cpSRP54 and cpFtsY for GDP and the non-hydrolyzable GTP analogue GppNHp by using these nucleotides as competitive inhibitors of the basal

GTPase reactions. Both proteins bind GDP and GppNHp weakly, with inhibition constants in the micromolar range (Table 1.2).

We also directly measured the interaction of nucleotides with both GTPases using fluorescent mant-derivatives of GTP and GDP. Binding of both GTPases to mant-GTP or mant-GDP induces a 50–80% increase in fluorescence (Figure 1.3A and 1.3B, respectively). Titration of this fluorescence change as a function of protein concentration gave dissociation constants of 6.5 and 11  $\mu\text{M}$  for binding of mant-GTP and mant-GDP to cpSRP54, respectively, and 1.9 and 3.1  $\mu\text{M}$  for binding of mant-GTP and mant-GDP to cpFtsY, respectively (Figure 1.3C and 1.3D; Table 1.2). For cpFtsY, these affinities are the same, within error, as those of unmodified nucleotides determined using the GTPase assay. For cpSRP54, these affinities are only  $\sim$  twofold larger than those of unmodified GTP and GDP. Thus, the mant-group does not significantly perturb the binding of nucleotides.

A hallmark of the SRP subgroup of GTPases is the fast rate at which they release and exchange nucleotides. The weak nucleotide binding affinities of cpSRP54 and cpFtsY suggest that this is also the case for the chloroplast SRP GTPases. This was confirmed by directly measuring the dissociation rate constants of mant-GTP and mant-GDP. As expected, both cpSRP54 and cpFtsY release mant-GTP quickly, with dissociation rate constants of 10.4 and 5.4  $\text{s}^{-1}$ , respectively (Figure 1.3E and Table 1.1). Similarly, mant-GDP is released quickly by both cpSRP54 and cpFtsY, with dissociation rate constants of 32 and 8.1  $\text{s}^{-1}$ , respectively (Figure 1.3F and Table 1.1). Thus, analogous to their bacterial and mammalian homologues, the chloroplast SRP GTPases hydrolyze GTP slowly and can exchange nucleotides quickly, in contrast to classical

signaling GTPases that release nucleotide slowly (on the order of  $10^{-3} - 10^{-4} \text{ s}^{-1}$ ) and require external exchange factors to facilitate nucleotide release.

***Interaction between cpSRP54 and cpFtsY is much more efficient than classical SRP systems.***

In classical SRP systems, complex formation between the SRP and SR GTPases is very slow and is accelerated 400-fold by the SRP RNA (19, 32). Once a complex is formed, SRP and SR stimulate each other's GTPase activity and the rate of this stimulated GTPase reaction within the complex is also accelerated 5–10 fold by the SRP RNA (19, 32). As no SRP RNA has been found in the chloroplast SRP system, we asked whether and how efficiently cpSRP54 and cpFtsY can interact with and activate each other in the absence of an SRP RNA.

To this end, we determined the rate of stimulated GTP hydrolysis reaction in the presence of both cpSRP54 and cpFtsY; GTPase activation in the cpSRP54•cpFtsY complex provides a means to monitor complex formation between the two GTPases (19, 32). To our surprise, cpSRP54 and cpFtsY interact with each other efficiently even in the absence of an SRP RNA (Figure 1.4, ●). The slope of the initial linear portion of the protein concentration dependence, which represents the rate constant of the reaction:  $\text{GTP} \cdot \text{cpSRP54} + \text{cpFtsY} \cdot \text{GTP} \rightarrow \text{products} (k_{cat}/K_m)$ , is  $\sim 400$ -fold faster than that of the corresponding reaction between the *E. coli* GTPases in the absence of the SRP RNA (Figure 1.4, ▲). Indeed, this rate constant matches that of the *E. coli* GTPases in the presence of the 4.5S SRP RNA (Figure 1.4, ▼). The rate constant at saturating protein concentrations, which represents the rate of GTP hydrolysis within the

$\text{GTP}\cdot\text{cpSRP54}\cdot\text{cpFtsY}\cdot\text{GTP}$  complex, is also identical between the chloroplast and the *E. coli* GTPases in the presence of the SRP RNA (● vs. ▼), and eightfold faster than that of the *E. coli* GTPases without the RNA bound (▲).

In the *E. coli* SRP system, complex formation is rate-limiting for the reaction:  $\text{GTP}\cdot\text{SRP} + \text{FtsY}\cdot\text{GTP} \rightarrow \text{products}$  (both in the presence and absence of SRP RNA) (19). Therefore,  $k_{cat}/K_m$  is equal to the association rate constant between the two GTPases. If this were also true for the cpSRP54 and cpFtsY, then the association between cpSRP54 and cpFtsY would be 400-fold faster than their *E. coli* homologues. Alternatively,  $k_{cat}/K_m$  is limited by the chemical step instead of complex formation for the chloroplast GTPases. If this were true, then the difference in association rates between the chloroplast and *E. coli* GTPases would be even greater. Thus the results in Figure 1.4 demonstrate that complex formation between the chloroplast SRP and SR GTPases is much more efficient than that of their bacterial and mammalian homologues and thus do not need the help from an SRP RNA.

***cpFtsY exhibits high nucleotide specificity.***

Association between bacterial SRP and SR GTPases is slow presumably because significant domain rearrangements are required to form a stable complex, including a change from the open to the closed conformation that is manifested functionally as an increase in the nucleotide specificity of the *E. coli* FtsY ((15); see Introduction). We hypothesized that the chloroplast SRP GTPases are pre-organized in the closed

conformation even in the absence of their binding partner, thus reducing the cost for the open  $\rightarrow$  closed rearrangement and resulting in a faster rate of protein–protein interaction.

A prediction from this model is that cpFtsY can effectively discriminate between cognate and non-cognate nucleotides by itself without the help from cpSRP54. To test this idea, we mutated the conserved specificity determinant, Asp283, to an asparagine. This mutation converts many GTPases to XTP-specific proteins by swapping the hydrogen bond between the carboxylate oxygen of Asp and the exocyclic amino group of the guanine ring (33–36). As predicted, wild-type cpFtsY preferentially hydrolyzes GTP. The rate constant of the reaction:  $\text{GTP}^* + \text{FtsY} \rightarrow \text{GDP} + \text{P}_i^*$  is 37-fold faster than that of mutant cpFtsY(D283N) (Figure 1.5A). Similarly, mutant cpFtsY(D283N) hydrolyzes XTP much faster than wild-type cpFtsY (Figure 1.5B). In contrast, *E. coli* FtsY exhibits no more than a fourfold difference between wild-type and mutant GTPases in the hydrolysis rates of either nucleotide (15).

We next asked if cpFtsY can specifically bind its cognate nucleotide. Using both the GTPase assays (Figure 1.5C and 1.5D) and fluorescent mant-nucleotides (Figure 1.5E and 1.5F), we showed that wild-type cpFtsY preferentially binds guanine-based nucleotides, with affinities 40–70-fold higher than mutant cpFtsY(D283N) (Table 1.2). Analogously, mutant cpFtsY(D283N) preferentially binds xanthine-based nucleotides, with affinities 90–250-fold higher than wild type cpFtsY (Table 1.2). In contrast, *E. coli* FtsY exhibits no more than a twofold discrimination between wild-type and mutant GTPases for any nucleotides (15). Together, the results in this section show that, unlike

its bacterial homologue, the active site of cpFtsY can specifically recognize GTP even in the absence of cpSRP54. This is consistent with the notion that free cpFtsY is already in the closed conformation and pre-organized to interact with cpSRP54.

***GTPase activation between cpSRP54 and cpFtsY is reciprocal but asymmetric.***

The XTP-specific mutant cpFtsY(D283N) also allowed us to test whether cpSRP54 and cpFtsY reciprocally stimulate the GTPase activity of one another, as is the case for the bacterial system. If this were the case, XTP hydrolysis by cpFtsY(D283N) would be stimulated by cpSRP54 and, conversely, GTP hydrolysis by cpSRP54 would be stimulated by cpFtsY(D283N).

To examine the effect of cpFtsY(D283N) on GTP hydrolysis by cpSRP54, we measured the rate of GTP hydrolysis in the third-order reaction:  $\text{GTP}^* + \text{cpSRP54} + \text{D283N}^{\text{XTP}} \rightarrow \text{GDP} + \text{P}_i^*$ . As predicted, the rate of GTP hydrolysis is significantly stimulated by the presence of cpFtsY(D283N) (Figure 1.6A), consistent with the notion that cpFtsY acts as the activating protein for cpSRP54. Analogously, the reciprocal reaction, XTP hydrolysis by cpFtsY(D283N), is significantly stimulated by the presence of cpSRP54 (Figure 1.6B; the third-order reaction:  $\text{XTP}^* + \text{D283N} + \text{cpSRP54}^{\text{GTP}} \rightarrow \text{XDP} + \text{P}_i^*$  was followed).

Interestingly, the rate of stimulated GTP hydrolysis from cpSRP54 is ~ tenfold slower than that of XTP hydrolysis from cpFtsY(D283N) (see rates in Figure 1.6A and 1.6B), raising the possibility that nucleotide hydrolyses from the two GTPase sites in the complex are *not* symmetric. To test this possibility, we formed the

$\text{GTP}\cdot\text{cpSRP54}\cdot\text{cpFtsY(D283N)}\cdot\text{XTP}$  complex by using high concentrations of both proteins and stoichiometric amounts of GTP and XTP, and directly measured the rate constants for hydrolysis of both GTP and XTP from this complex. As shown in Figure 1.6C, the rate constant for XTP hydrolysis is  $3.7 \text{ min}^{-1}$  (squares), over fourfold faster than the rate constant of  $0.87 \text{ min}^{-1}$  for GTP hydrolysis (circles). This represents only a lower limit for the difference in hydrolysis rates between the two active sites, because cpFtsY(D283N) bound with GTP is much more active in binding and activating cpSRP54 (see the next section), even though it preferentially binds XTP by itself. Thus, part of the GTP hydrolysis rate observed in Figure 1.6C (circles) is contributed by an alternative  $\text{GTP}\cdot\text{cpSRP54}\cdot\text{cpFtsY(D283N)}\cdot\text{GTP}$  complex. The actual difference between the hydrolysis rates from the two active sites is larger than that observed in Figure 1.6C and is closer to the  $\sim$  tenfold difference observed in Figures 1.6A and 1.6B, which monitors the third-order reaction rates. Under these conditions, the observed reaction rates are determined by the affinity of free cpSRP54 and cpFtsY(D283N) for their respective nucleotides as well as the rate at which GTP and XTP are hydrolyzed from the respective active sites in the complex. Since cpSRP54 and cpFtsY(D283N) exhibit similar affinities for GTP and XTP, respectively (Table 1.2), the observed  $\sim$  tenfold difference in reaction rate (Figures 1.6A and 1.6B) primarily reflects the difference in hydrolysis rate from the two active sites. Thus, like the classical SRP systems, cpSRP54 and cpFtsY act as reciprocal activating proteins for one another, yet unlike their bacterial homologues, nucleotide hydrolyses from the two active sites are asymmetric.

***Mutant cpFtsY(D283N) prefers GTP over XTP upon complex formation with cpSRP54.***

Another intriguing observation from the results in Figure 1.6C is that the rate constants of the stimulated GTPase and XTPase reactions from the  $\text{GTP}\cdot\text{cpSRP54}\cdot\text{cpFtsY(D283N)}\cdot\text{XTP}$  complex (0.87 and  $3.7 \text{ min}^{-1}$ , respectively) are over tenfold slower than that from the wild-type  $\text{GTP}\cdot\text{cpSRP54}\cdot\text{cpFtsY}\cdot\text{GTP}$  complex (Figure 1.4), even accounting for the fact that two GTP molecules are hydrolyzed in the wild-type complex. Therefore, we suspected that the D283N mutation or the replacement of GTP with XTP renders cpFtsY less active in binding and activating cpSRP54. This is reminiscent of the behavior of an XTP-specific mutant of the *E. coli* SRP GTPase, SRP(D251N), which is deficient in binding and activating FtsY in its XTP-bound form. Instead, mutant SRP(D251N) can better bind and activate FtsY when bound to the non-cognate GTP (31).

To test whether this is also the case for mutant cpFtsY(D283N), we measured the rate constant for GTP hydrolysis from the  $\text{GTP}\cdot\text{cpSRP54}\cdot\text{cpFtsY(D283N)}\cdot\text{GTP}$  complex when cpFtsY(D283N) is forced to bind its non-cognate nucleotide by using a high GTP concentration. When mutant cpFtsY(D283N) is bound with the non-cognate GTP, the rate of stimulated GTP hydrolysis is much faster than when it is bound with the cognate XTP (Figure 1.7A, diamonds vs. squares). The rate constant at saturating protein concentration, which represents the rate constant for GTP hydrolysis from the  $\text{GTP}\cdot\text{cpSRP54}\cdot\text{cpFtsY(D283N)}\cdot\text{GTP}$  complex, is comparable to that of the wild-type complex (Figure 1.7A; diamonds vs. circles), suggesting that the

$\text{GTP}\cdot\text{cpSRP54}\cdot\text{cpFtsY(D283N)}\cdot\text{GTP}$  complex achieves the same active conformation as the complex formed by the wild-type proteins. As a  $\sim$  fivefold higher concentration of mutant cpFtsY(D283N) than wild-type cpFtsY is required to reach saturation, complex formation is modestly compromised for GTP-bound cpFtsY(D283N) (Figure 1.7A; diamonds vs. circles). In contrast, no saturation is observed in the reaction with XTP-bound cpFtsY(D283N) up to 30  $\mu\text{M}$  (squares), indicating that complex formation is significantly compromised when the mutant is bound with its cognate nucleotide. Thus, mutant cpFtsY(D283N) prefers the non-cognate GTP over cognate XTP when it forms a complex with cpSRP54.

To provide independent evidence on this switch in nucleotide preference upon complex formation, we explored the effect of XTP on the rate of the reaction:  $\text{GTP}\cdot\text{cpSRP54} + \text{cpFtsY(D283N)}\cdot\text{GTP} \rightarrow \text{products}$ . If cpFtsY(D283N) is less active in binding and activating the GTPase reaction of cpSRP54 when it is bound with cognate XTP than with non-cognate GTP, then addition of XTP, which competes off the GTP bound at the active site of cpFtsY(D283N), should inhibit the stimulated GTPase reaction. As predicted, addition of XTP inhibits this stimulated reaction (Figure 1.7B). The observed inhibition constant for XTP is 9.0  $\mu\text{M}$ , consistent with the expected value of  $8.9\pm 0.9$   $\mu\text{M}$  given the affinities of mutant cpFtsY(D283N) for GTP and XTP and the GTP concentration used in this experiment (eq 4 in the Methods). This strongly suggests that the binding of XTP to cpFtsY(D283N) is responsible for the observed inhibitory effect. Taken together, the results in this section show that although cpFtsY(D283N) by itself exhibits a specificity for XTP, this mutant prefers the non-cognate GTP for interacting with and stimulating GTP hydrolysis from cpSRP54. Thus Asp283 and/or the

bound GTP play a much more important role than specifying the nucleotide preference of cpFtsY and likely participate in critical interface interactions with cpSRP54 in the cpSRP54•cpFtsY complex.

## Discussion

***The chloroplast SRP and SR GTPases are pre-organized to efficiently interact with each other.***

cpSRP54 and cpFtsY share 69.5% and 65.4% similarity with their *E. coli* homologues. All the essential motifs in the GTP binding pocket are highly conserved. As expected from the high sequence conservation, both proteins share many biochemical features characteristic of the SRP subfamily of GTPases, including weak nucleotide affinities, fast nucleotide exchange rates, and the ability to reciprocally stimulate each other's GTPase reaction after they form a complex.

Given these similarities, it is surprising that the otherwise universally conserved SRP RNA, which plays a crucial role in eukaryotic and prokaryotic SRP protein targeting pathways, is missing in the chloroplast SRP pathway. In *E. coli*, association between the SRP and SR GTPases is extremely slow, with a rate constant of  $5 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$  (19). This slow association rate does not appear to be caused by the extended N-terminal A-domain of *E. coli* FtsY, as truncating the N-terminal 46 amino acids of the A-domain results in identical kinetics of interaction with the SRP GTPase (S.S., unpublished results). Further, *T. aquaticus* FtsY, which lacks an extended N-terminal A-domain, also interacts with its binding partner very slowly in the absence of the SRP RNA (S.S., unpublished results). At this rate and the *in vivo* concentration of these GTPases (nanomolar range), the association between the two GTPases will take hours to complete, whereas protein targeting occurs on the time scale of seconds. One of the important contributions of the SRP RNA is to accelerate complex formation between the two GTPases by 400-fold (19,

32). Another contribution of the SRP RNA, albeit minor in comparison, is to increase the rate at which GTP is hydrolyzed from the SRP•SR complex by ~ five–tenfold (19, 32). GTP hydrolysis is known to drive disassembly and recycling of the SRP and SR after each round of targeting (8). Here, we showed that cpSRP54 and cpFtsY can interact efficiently with each other even in the absence of an SRP RNA: their association rate is at least as fast as that of their *E. coli* homologues that contain the SRP RNA, and GTP hydrolysis from the cpSRP54•cpFtsY complex also occurs at the same rate as the *E. coli* GTPase complex in the presence of the RNA. This helps explain how the chloroplast SRP system can bypass the requirement for the SRP RNA.

Why is the protein–protein interaction so efficient between the chloroplast GTPases? Interaction between the bacterial SRP and SR GTPases is slow presumably due to the requirement for extensive conformational changes during complex formation. One of the important rearrangements is a repositioning of the N-G domain interface, which led to a change of the GTPase site from a floppy, non-specific open state to a closed state in which active site interactions with the bound nucleotide are established (15). Thus, one possibility is that cpSRP54 and cpFtsY are pre-organized into the closed conformation that is ready to interact with each other. The results herein strongly suggest that this is the case at least for cpFtsY. Free cpFtsY can specifically recognize its cognate nucleotide, in contrast to *E. coli* FtsY which acquires nucleotide specificity only when it forms a complex with SRP. Further, cpFtsY exhibits higher affinities for GTP and GDP than its bacterial homologues, with dissociation constants of 2–3  $\mu\text{M}$  instead of 19–30  $\mu\text{M}$  for *E. coli* FtsY. Finally, the crystal structure of cpFtsY shows that among all the structures of free FtsY from various species, the conformation of apo-cpFtsY is closest to

that observed in the Ffh•FtsY complex, especially with regard to the relative orientation of the N- and G-domains (Chandrasekar et al., manuscript in preparation). These observations strongly support the notion that free cpFtsY is pre-organized in a closed conformation, and thus can interact with cpSRP54 without paying further energetic penalty to rearrange the relative position of the N- and G-domains. It remains to be seen whether cpSRP54 is similarly pre-organized into the closed conformation prior to interaction with cpFtsY.

It appears that the SRP RNA has been evolved to accelerate the very inefficient interaction between the SRP and SR GTPases in classical systems. Although models are abundant (32, 37, 38), the molecular mechanism by which the SRP RNA acts as a catalyst to accelerate both the formation and disassembly of the SRP•SR complex is still poorly understood. It is possible that in the transition state for complex assembly, the SRP RNA may provide a transient tether that facilitates the rearrangement of one or both GTPases into the closed conformation; alternatively, the RNA and the chloroplast GTPases may employ completely different mechanisms to attain faster association kinetics.

***Asymmetric nucleotide hydrolysis from the cpSRP54•cpFtsY complex.***

The crystal structure of the *T. aquaticus* Ffh•FtsY complex shows that the two GMPPCP molecules are bound at a composite active site formed at the dimer interface (12, 13). Consistent with the composite nature of the active site and the extensive degree of crosstalk between the two GTPase sites, the two nucleotides are hydrolyzed at the same rate from the *E. coli* SRP•FtsY complex. These observations have led to earlier

proposals of concerted GTP hydrolyses in the SRP•SR complex (20). In contrast to this notion, we showed here that nucleotide hydrolysis in the cpSRP54•cpFtsY complex can be asymmetric, with the nucleotide hydrolyzing ~ tenfold faster from the cpFtsY than the cpSRP54 active site. This observation argues against a concerted mechanism. Even in the *E. coli* system, mutant GTPases have been identified in which GTP is hydrolyzed much faster from the SRP than the FtsY active site (39). Further, when either one of the GTPases is bound with a non-hydrolyzable GTP analogue, it can still activate efficient GTP hydrolysis on its binding partner (unpublished results). Together, these results strongly suggest that hydrolyses of the two GTPs in the SRP•SR complex do not proceed through a concerted mechanism or an ordered pathway (i.e., one GTP must be hydrolyzed first before hydrolysis of the second GTP can occur). Rather, each active site can hydrolyze its bound GTP independently.

Even though the nucleotide is hydrolyzed ~ tenfold slower from cpSRP54 than from cpFtsY(D283N), multiple rounds of XTP hydrolysis from the cpSRP54•cpFtsY(D283N) complex is not blocked and occur as efficiently as single turnover reactions (data not shown). Thus, disassembly of the complex must occur on a faster time scale than the second hydrolysis event, implying that SRP and SR can dissociate from one another even when only one of the nucleotides is hydrolyzed. A similar observation was made for the *E. coli* SRP•SR complex (39). Together, the data from the *E. coli* and chloroplast systems suggest that only one GTP hydrolysis event is required to drive disassembly of the SRP•SR complex. It remains to be clarified how many GTPs need to be hydrolyzed during each round of protein targeting, and what the precise role of each GTP hydrolysis event is.

***The nucleotide specificity determinant of cpFtsY, Asp283, mediates molecular crosstalk between the two GTPases.***

Given the high specificity of cpFtsY(D283N) for XTP, it is surprising to find that this mutant prefers GTP over XTP when it forms a complex with cpSRP54. This strongly suggests that Asp283, in addition to conferring nucleotide specificity to cpFtsY, also contributes to interactions at the dimer interface. The behavior of cpFtsY(D283N) is reminiscent of an XTP-specific mutant of the *E. coli* SRP GTPase, Ffh(D251N), which also prefers the non-cognate GTP over cognate XTP when it forms a complex with FtsY (31). The crystal structure confirms that Asp251 makes an important interface contact with Lys390 from FtsY (12, 13). A similar interaction could be formed by Asp283 of cpFtsY with a hydrogen bond donor (–AH) at the interface of the cpSRP54•cpFtsY complex (Figure 1.8A). When cpFtsY(D283N) is bound to XTP, mutation of Asp283 to Asn destroys this interface contact and compromises the interaction between the two GTPases (Figure 1.8B). In contrast, replacement of XTP with GTP no longer constrains Asn283 in this particular configuration; a rotation around the C<sup>β</sup>–C<sup>γ</sup> bond can reposition the carbonyl oxygen of Asn283 close to the hydrogen bond donor from cpSRP54, thus restoring this interface contact (Figure 1.8C). Alternatively, the exocyclic amino group of GTP could directly interact with a hydrogen bond acceptor from cpSRP54 (–B; Figure 1.8D), therefore replacement of GTP with XTP compromises the cpSRP54–cpFtsY(D283N) interaction. In either scenario, our results map the G-IV motif of cpFtsY and its bound nucleotide to the dimer interface between the two GTPases, and demonstrate the presence of extensive crosstalk between the two GTPase sites.

## Perspective

The results here help rationalize why the chloroplast SRP targeting pathway bypasses the requirement for the SRP RNA, as the SRP and SR GTPases from chloroplast can interact efficiently with one another without the help from the SRP RNA. The novel cpSRP43 protein, which together with cpSRP54 forms the chloroplast SRP, has often been viewed as a functional replacement for the SRP RNA. Our results show that the chloroplast GTPases have evolved to efficiently interact with one another, and no additional acceleration has been observed in the presence of cpSRP43 (Supplementary Figure 1.S2). Therefore, cpSRP43 does not replace all of the functions of the SRP RNA. This novel chloroplast protein may have evolved to mediate other important roles of the SRP RNA in the protein targeting reaction, such as recognition of the cargo protein (25, 29). Analogously, the SRP RNA may have been evolved to interact with ribosomal RNAs during co-translational protein targeting in the classical SRP pathways (40, 41).

It is fascinating to speculate on the evolutionary origin of the vast difference in the kinetics of interaction between the SRP and SR GTPases from chloroplast vs. classical SRP pathways, and why cells have evolved the SRP RNA to deal with the inefficient interaction between SRP and SR GTPases during targeting of proteins from the cytosol to membrane compartments. An intriguing possibility is that the slow interaction kinetics between the SRP and SR GTPases in classical pathways allow for additional opportunities for regulation and for improving fidelity. The interaction kinetics between SRP and SR are still relatively slow even with the RNA present; *in vivo*, however, the presence of the ribosome, the cargo protein, or the membrane translocon could further accelerate the interaction between the two GTPases. The SRP RNA, bound

in close vicinity to the signal sequence binding site, could respond to cues such as cargo binding and mediate the additional stimulation of SRP–SR interaction by the cargo. In this way, the SRP RNA could mediate the communication between the cargo binding M-domain and the GTPase domain of SRP, and potentially provide a checkpoint to improve the fidelity of the classical SRP pathway, which needs to sort a vast number of cellular proteins to subcellular compartments. In contrast, a much smaller number of proteins need to be handled by the cpSRP pathway, thus alleviating the need for this regulatory RNA molecule.

**Acknowledgment**

We thank Dr. R. Henry for the bacterial expression vectors for cpSRP54 and cpFtsY, Dr. J. Campbell and the rest of the Shan lab for comments on the manuscript, and Dr. P. Walter for insightful discussions and intellectual support. S.S. was supported by career awards from the Burroughs Wellcome Fund and the Camille and Henry Dreyfus Foundation, P.J.A. is supported by the Bray fellowship.

**Table 1.1** Rate and equilibrium constants for the GTPase cycle of cpSRP54 and cpFtsY<sup>a</sup>

<b>Equilibrium or Rate Constant</b>	<b>Values determined</b>
$k_1$	$3.7 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$
$k_{-1}$	$10.4 \pm 0.8 \text{ s}^{-1}$
$K_1$	$2.8 \pm 0.4 \text{ }\mu\text{M}$
$k_2$	$0.017 \pm 0.002 \text{ min}^{-1}$
$k_3$	$2.9 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$
$k_{-3}$	$32 \pm 2 \text{ s}^{-1}$
$K_3$	$5.1 \pm 0.3 \text{ }\mu\text{M}$
$k_1'$	$2.6 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$
$k_{-1}'$	$5.4 \pm 0.3 \text{ s}^{-1}$
$K_1'$	$2.1 \pm 0.2 \text{ }\mu\text{M}$
$k_2'$	$0.0045 \pm 0.002 \text{ min}^{-1}$
$k_3'$	$2.6 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$
$k_{-3}'$	$8.1 \pm 0.2 \text{ s}^{-1}$
$K_3'$	$3.1 \pm 0.2 \text{ }\mu\text{M}$
$k_4$	$\geq 8.6 (\pm 0.3) \times 10^5 \text{ M}^{-1}\text{s}^{-1}$
$k_{-4}$	N. D.
$k_5$	$0.83 \pm 0.04 \text{ s}^{-1}$
$k_6$	$\geq 0.83 \text{ s}^{-1}$

<sup>a</sup> The rate and equilibrium constants are defined in Figure 1.1. The details and justifications of how each microscopic rate constant is determined from the experimental data are provided in Supplementary Material. N. D., not determined

**Table 1.2** Nucleotide affinities of cpSRP54, wild-type cpFtsY, and mutant cpFtsY(D283N)

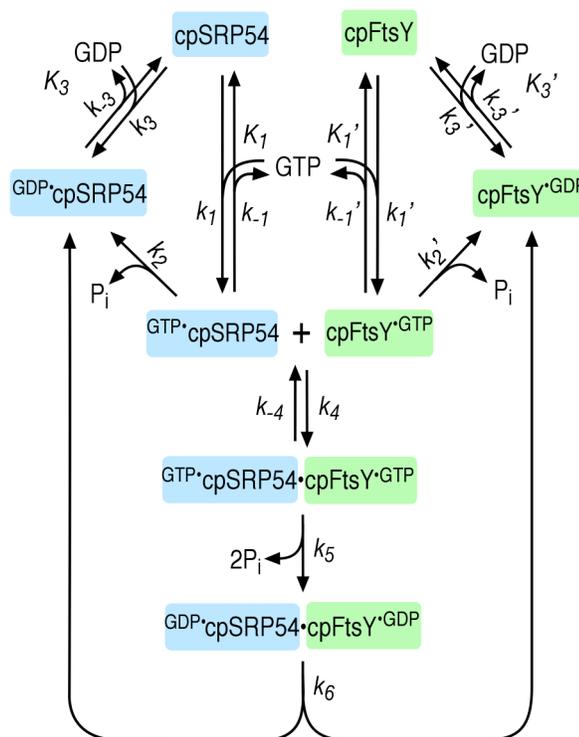
Protein	$K_d$ ( $\mu\text{M}$ )					
	GTP	GDP	GppNHp	XTP	XDP	XppNHp
cpSRP54	2.8 <sup>a</sup> /6.5 <sup>c</sup>	5.1 <sup>b</sup> /11 <sup>c</sup>	26 <sup>b</sup>	50 <sup>b</sup>	N.D. <sup>d</sup>	N.D. <sup>d</sup>
cpFtsY wild-type	2.1 <sup>a</sup> /1.9 <sup>c</sup>	3.1 <sup>c</sup>	4.6 <sup>b</sup>	510 <sup>b</sup>	557 <sup>b</sup>	970 <sup>b</sup>
cpFtsY(D283N)	76 <sup>b</sup>	180 <sup>b</sup>	360 <sup>b</sup>	2.2 <sup>a</sup> /2.7 <sup>c</sup>	6.5 <sup>c</sup>	34 <sup>b</sup>

<sup>a</sup>Determined by cpFtsY concentration dependences as described in Methods

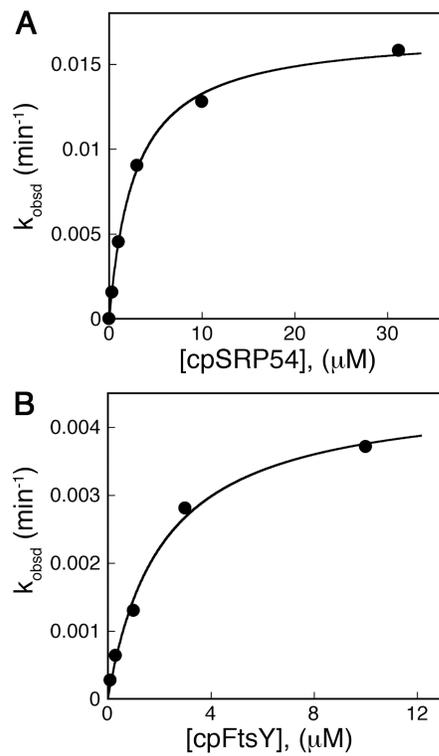
<sup>b</sup>Determined by inhibition methods as described in Methods

<sup>c</sup>Determined by fluorescence using mant-nucleotides as described in Methods

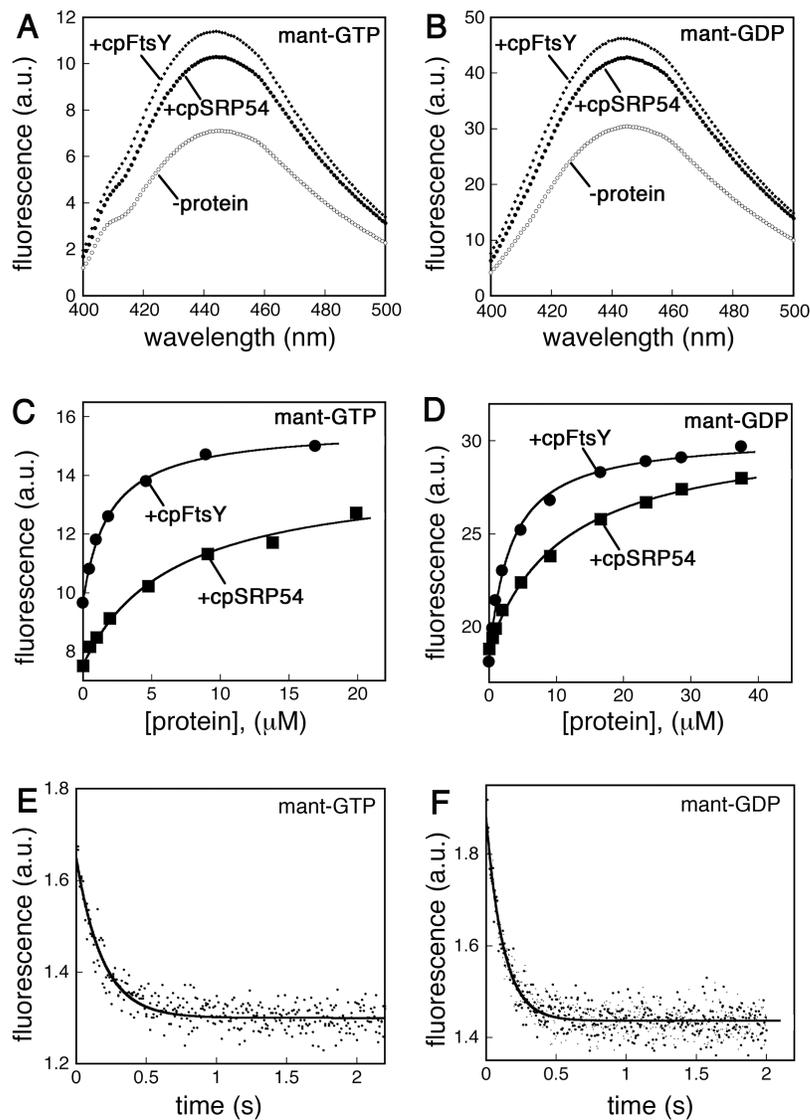
<sup>d</sup>N.D., not determined



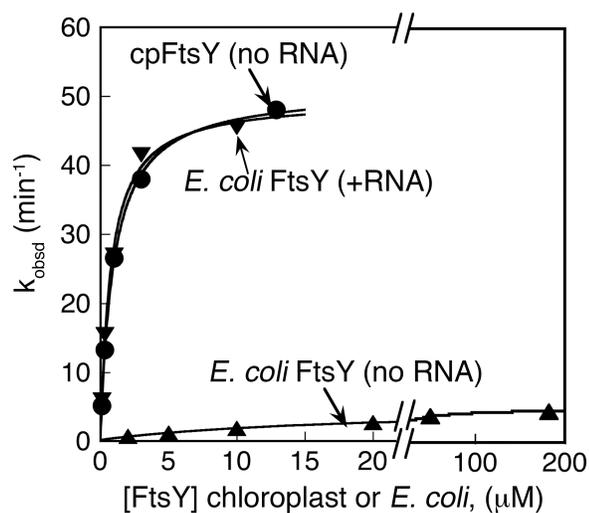
**Figure 1.1** Schematic depiction of the GTPase cycles of cpSRP54 (*blue*) and cpFtsY (*green*). Superscripts depict the nucleotide bound to each protein. The triangular cycles on the top left and right depict the basal GTPase cycles of cpSRP54 and cpFtsY, respectively. Binding of GTP to cpSRP54 (or cpFtsY) is characterized by the association rate constant  $k_1$  (or  $k_1'$ ), dissociation rate constant  $k_{-1}$  (or  $k_{-1}'$ ), and equilibrium dissociation constant  $K_1$  (or  $K_1'$ ). Rate constants for GTP hydrolysis from cpSRP54 and cpFtsY are denoted by  $k_2$  and  $k_2'$ , respectively. Binding of GDP to cpSRP54 (or cpFtsY) is characterized by the association rate constant  $k_3$  (or  $k_3'$ ), dissociation rate constant  $k_{-3}$  (or  $k_{-3}'$ ), and equilibrium dissociation constant  $K_3$  (or  $K_3'$ ). Complex formation between cpSRP54 and cpFtsY is characterized by the association rate constant  $k_4$  and dissociation rate constant  $k_{-4}$ . The two bound GTPs are hydrolyzed from the  $\text{GTP}\cdot\text{cpSRP54}\cdot\text{cpFtsY}\cdot\text{GTP}$  complex, represented collectively by the rate constant  $k_5$ . The  $\text{GDP}\cdot\text{cpSRP54}\cdot\text{cpFtsY}\cdot\text{GDP}$  complex then dissociates with a rate constant  $k_6$ .



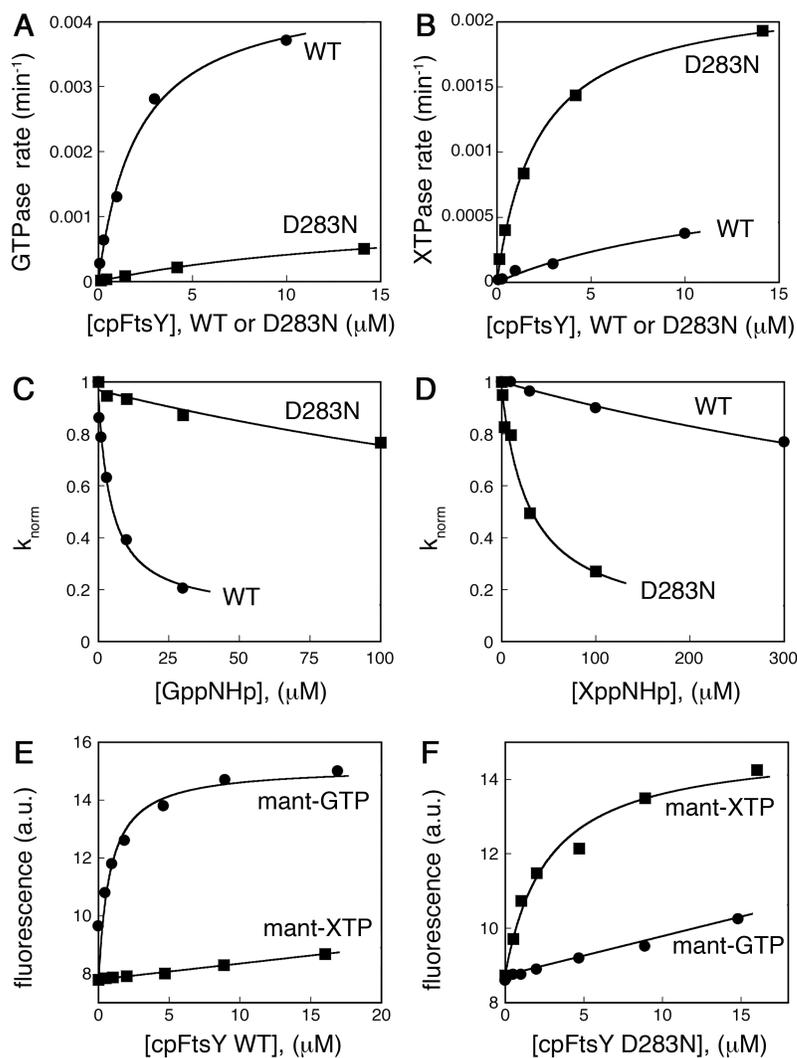
**Figure 1.2** Basal GTPase reactions of cpSRP54 (A) and cpFtsY (B). The data were fit to eq 1 in Methods and gave a  $k_{\text{max}}$  of  $0.017 \text{ min}^{-1}$  and  $K_{1/2}$  of  $2.8 \mu\text{M}$  for cpSRP54, and a  $k_{\text{max}}$  of  $0.0045 \text{ min}^{-1}$  and  $K_{1/2}$  of  $2.1 \mu\text{M}$  for cpFtsY.



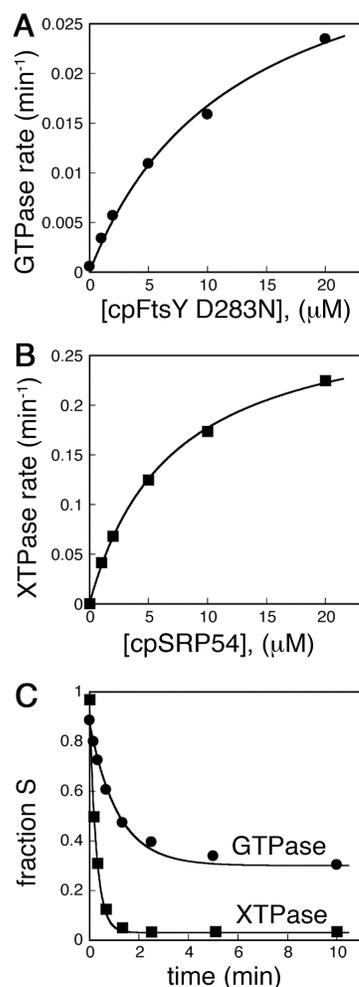
**Figure 1.3** Interaction of nucleotides with cpSRP54 and cpFtsY. (A–B) Fluorescence emission spectra of mant-GTP (A) or mant-GDP (B) in the absence of protein (O) and in the presence of 5  $\mu$ M cpSRP54 (●) or cpFtsY (◆). (C–D) Titration of the fluorescence changes of mant-GTP (C) and mant-GDP (C) in the presence of cpSRP54 (■) or cpFtsY (●). The data were fit to eq 5 in Methods and the  $K_d$  values are summarized in Table 1.2. (E–F) Dissociation of mant-GTP (E) and mant-GDP (F) from cpFtsY. The data were fit to single exponential rate equations and gave dissociation rate constants of 5.4 and 8.1  $s^{-1}$  for mant-GTP and mant-GDP, respectively.



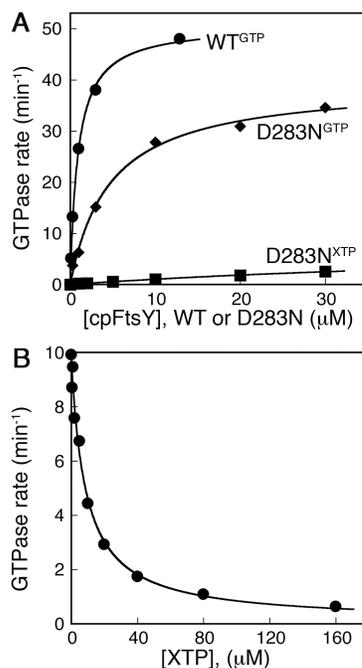
**Figure 1.4** Interaction of cpSRP54 and cpFtsY is much more efficient than that of their *E. coli* homologues. Rates of the stimulated GTPase reaction were determined for cpSRP54 (100 nM) and cpFtsY (●), or for *E. coli* Ffh (100 nM) and *E. coli* FtsY with (▼) and without 4.5S SRP RNA (▲). The data were fit to eq 2 in the Methods, and gave a  $k_{cat}$  value of  $50 \text{ min}^{-1}$  and a  $K_m$  value of  $0.97 \text{ μM}$  for the chloroplast GTPases, and  $k_{cat}$  values of 49 and  $4.8 \text{ min}^{-1}$  and  $K_m$  values of 0.76 and  $18 \text{ μM}$  for the *E. coli* GTPases with and without the SRP RNA, respectively.



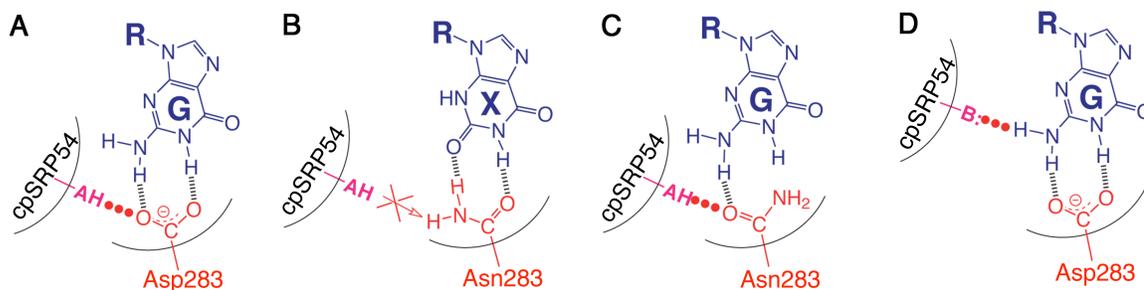
**Figure 1.5** cpFtsY preferentially binds and hydrolyzes its cognate nucleotide. (A–B) Basal GTPase (A) and XTPase (B) reactions of wild-type cpFtsY (●) and mutant cpFtsY(D283N) (■). The data were fit to eq 1 and gave a  $k_{\text{max}}$  value of  $0.0045 \text{ min}^{-1}$  and a  $K_{1/2}$  value of  $2.1 \text{ }\mu\text{M}$  for GTP hydrolysis by wild-type cpFtsY, and a  $k_{\text{max}}$  value of  $0.0022 \text{ min}^{-1}$  and a  $K_{1/2}$  value of  $2.2 \text{ }\mu\text{M}$  for XTP hydrolysis by mutant cpFtsY(D283N). (C) GppNHp binds more strongly to wild-type cpFtsY (●) than to mutant cpFtsY(D283N) (■). (D) XppNHp binds more strongly to mutant cpFtsY(D283N) (■) than to wild-type cpFtsY (●). The  $K_i$  values are reported in Table 1.2. (E–F) Titration of the change in fluorescence of mant-GTP (●) and mant-XTP (■) upon binding to wild-type cpFtsY (E) and mutant cpFtsY(D283N) (F). The data were fit to eq 5 and the  $K_d$  values are summarized in Table 1.2.



**Figure 1.6** Nucleotide hydrolyses from the cpSRP54•cpFtsY(D283N) complex are asymmetric. (A) Stimulation of the GTPase reaction of cpSRP54 by cpFtsY(D283N), determined as described in Methods using 0.2 μM cpSRP54 and 20 μM XTP. The data were fit to eq 1 and gave a maximal rate constant of 0.037 min<sup>-1</sup>. (B) Stimulation of the XTPase reaction of cpFtsY(D283N) by cpSRP54, determined as described in Methods using 0.2 μM cpFtsY(D283N) and 20 μM GTP. The data were fit to eq 1 and gave a maximal rate constant of 0.30 min<sup>-1</sup>. (C) Time courses for GTP and XTP hydrolyses from the <sup>GTP</sup>•cpSRP54•cpFtsY(D283N)•<sup>XTP</sup> complex, determined as described in Methods. The data were fit to single-exponential rate equations and gave rate constants of 0.86 and 3.7 min<sup>-1</sup> for the GTPase reaction (●) and XTPase reactions (■), respectively.



**Figure 1.7** cpFtsY(D283N) prefers GTP over XTP when it forms a complex with cpSRP54. (A) GTP hydrolysis rates when cpSRP54 (100–500 nM) interacts with wild-type cpFtsY (●), cpFtsY(D283N) bound to GTP (◆) and cpFtsY(D283N) bound to XTP (■). The following nucleotide concentrations were used: 100  $\mu\text{M}$  GTP for reaction with wild-type cpFtsY, 200  $\mu\text{M}$  GTP for reaction with cpFtsY(D283N) bound to GTP, and 20  $\mu\text{M}$  GTP and 50  $\mu\text{M}$  XTP for reaction with cpFtsY(D283N) bound to XTP. The data were fit to eq 2, which gave  $k_{cat}$  values of 50 (●) and 39  $\text{min}^{-1}$  (◆). (B) XTP inhibits the ability of GTP-bound cpFtsY(D283N) to stimulate GTP hydrolysis by cpSRP54. Reactions were carried out in the presence of 500 nM cpSRP54, 2 mM cpFtsY(D283N), and 200  $\mu\text{M}$  GTP, as described in the Methods. The data were fit to eq 3 and gave an apparent inhibition constant of 9.0  $\mu\text{M}$ .



**Figure 1.8** Model for the interactions of cpSRP54 with the side chain of cpFtsY Asp283 or with GTP. (A–C) Proposed interactions between the side chain of residue 283 with a hydrogen bond donor from cpSRP54 (–AH) for the wild-type cpSRP54•cpFtsY complex (A) and the cpSRP54•cpFtsY(D283N) complex with XTP (B) or GTP (C) bound to cpFtsY(D283N). (D) The GTP bound to cpFtsY interacts with a hydrogen bond acceptor (–B:) from cpSRP54.

## Supplementary Materials

### Determination of the microscopic rate constants:

*Basal GTP binding and hydrolysis ( $K_1$ ,  $k_2$  and  $K_1'$ ,  $k_2'$ ).* The chemical step is rate-limiting for the basal GTPase reaction of cpSRP54, as the maximal rate constant of GTP hydrolysis ( $0.017 \text{ min}^{-1}$ ; Figure 1.2A) is  $4 \times 10^4$ -fold slower than the rate at which GTP dissociates from the enzyme active site ( $10.4 \text{ s}^{-1}$ ; Figure 1.3E). Therefore, the  $K_{1/2}$  value obtained from the data in Figure 1.2A is equal to  $K_1$ , the equilibrium dissociation constant for GTP, and the  $k_{max}$  value from the same figure is equal to  $k_2$ , the rate constant for GTP hydrolysis from the  $\text{GTP} \cdot \text{cpSRP54}$  complex. For the same reason, the chemical step is rate-limiting for the basal GTPase reaction of cpFtsY. Therefore, the  $K_{1/2}$  value obtained from the data in Figure 1.2B is equal to  $K_1'$ , the dissociation constant for GTP, and the  $k_{max}$  value obtained from the same figure is equal to  $k_2'$ , the rate constant for GTP hydrolysis from the  $\text{cpFtsY} \cdot \text{GTP}$  complex. The values of  $K_1$  and  $K_1'$  were also determined independently by fluorescence assays (Figure 1.3C) as described in the text.

*GDP binding to cpSRP54 and cpFtsY ( $K_3$  and  $K_3'$ ).* The binding affinities of GDP for both proteins were determined by using GDP as a competitive inhibitor of the basal GTPase reaction, as described previously, and by fluorescence assays (Figure 1.3D) as described in the text.

*Nucleotide dissociation rate constants ( $k_{-1}$ ,  $k_{-1}'$  and  $k_{-3}$ ,  $k_{-3}'$ ).* The rate constants for nucleotide dissociation from each protein were measured using fluorescent mant-GTP and mant-GDP in pulse-chase experiments (Figure 1.3E and 1.3F and data not shown) as described in the text.

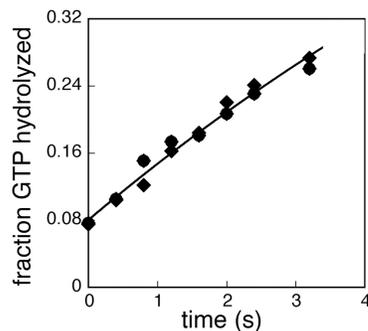
*Nucleotide association rate constants ( $k_1$ ,  $k_1'$  and  $k_3$ ,  $k_3'$ ).* The rate constants for binding of GTP and GDP to both proteins were obtained from the equilibrium dissociation constant and the dissociation rate constant for each nucleotide, determined as described above, using  $k_{on} = k_{off} / K_d$ .

*Rate constant for complex formation ( $k_4$ ).* The association rate constant between cpSRP54 and cpFtsY was not determined directly due to the lack of a direct protein–protein binding assay, and was estimated from the value of  $k_{cat}/K_m$  for the stimulated GTPase reaction; this value provides a lower limit for  $k_4$ , as explained in the Results section.

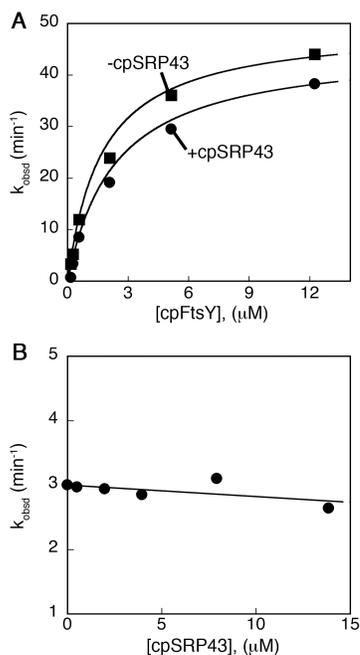
*Rate constant for GTP hydrolysis in the  $^{GTP}\bullet\text{cpSRP54}\bullet\text{cpFtsY}\bullet^{GTP}$  complex ( $k_5$ ).* This rate constant was derived from the value of  $k_{cat}$  determined from the stimulated GTPase reaction between cpSRP54 and cpFtsY (Figure 1.4, circles). Several observations suggest that product release is not rate-limiting for  $k_{cat}$ . First, the value of  $k_{cat}$  is the same, within experimental error, as the rate constant of GTP hydrolysis from the  $^{GTP}\bullet\text{cpSRP54}\bullet\text{cpFtsY}\bullet^{GTP}$  complex determined under single turnover conditions (data not shown). Second, the time course for the reaction:  $^{GTP}\bullet\text{cpSRP54}\bullet\text{cpFtsY}\bullet^{GTP} \rightarrow$  products is consistent with a single exponential rate without exhibiting a burst phase (Supplementary Figure 1.S1). Thus, steps prior to GTP hydrolysis, rather than product release, is rate-limiting for the stimulated GTPase reaction from the  $^{GTP}\bullet\text{cpSRP54}\bullet\text{cpFtsY}\bullet^{GTP}$  complex. Therefore,  $k_{cat}$  represents the sum of rate constants for hydrolysis of the two GTPs from the  $^{GTP}\bullet\text{cpSRP54}\bullet\text{cpFtsY}\bullet^{GTP}$  complex ( $k_5$ ) and may

be limited by either the chemical step itself, or a conformational change prior to GTP hydrolysis.

*Rate constant for complex dissociation ( $k_6$ ).* For the same reasons stated in the previous paragraph, the value of  $k_{cat}$  sets a lower limit for the rate of product release ( $k_6$ ), which has not been directly measured in this study.



**Supplementary Figure 1.S1** The time course for GTP hydrolysis from the cpSRP54•cpFtsY complex shows no obvious burst phase. The reaction was carried out in the presence of 12.5  $\mu\text{M}$  cpSRP54, 15  $\mu\text{M}$  cpFtsY, and 100  $\mu\text{M}$  GTP doped with trace amounts of GTP\*; the high concentration of protein relative to GTP is used to maximize the chance of observing the presence of a burst phase. The different symbols represent data from two independent measurements. The line is a fit of the initial part of the time course to a single-exponential rate equation.



**Supplementary Figure 1.S2** cpSRP43 shows no significant effect on GTPase activity of cpSRP54 and cpFtsY. (A) Rates of the stimulated GTPase reactions were determined for cpSRP54 (100 nM) and increasing concentration of cpFtsY in the presence of 100  $\mu$ M GTP. In the presence (●) of 1  $\mu$ M cpSRP43, the fit of the data to eq 2 in the Methods gave a  $k_{cat}$  value of 49.5  $\text{min}^{-1}$  and a  $K_m$  value of 1.7  $\mu$ M. In the absence (■) of cpSRP43, a  $k_{cat}$  value was 45.5  $\text{min}^{-1}$  and a  $K_m$  value was 2.3  $\mu$ M. (B) Rates of the stimulated GTPase reactions were determined for cpSRP54•cpFtsY complex (100 nM) in the presence of increasing concentration of cpSRP43. No significant stimulation or inhibition was observed compared to the reaction rate in the absence of cpSRP43.

**References:**

1. Keenan, R. J., Freymann, D. M., Stroud, R. M., and Walter, P. (2001) The signal recognition particle, *Annu. Rev. Biochem.* 70, 755–775.
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. Walter, P., Ibrahimi, I., and Blobel, G. (1981) Translocation of proteins across the endoplasmic reticulum I. Signal Recognition Protein (SRP) binds to *in vitro* assembled polysomes synthesizing secretory protein, *J. Cell. Biol.* 91, 545–550.
4. Gilmore, R., Blobel, G., and Walter, P. (1982a) Protein translocation across the endoplasmic reticulum: 1. Detection in the microsomal membrane of a receptor for the signal recognition particle, *J. Cell Biol.* 95, 463–469.
5. Gilmore, R., Walter, P., Blobel, G. (1982b) Protein translocation across the endoplasmic reticulum. II. Isolation and characterization of the signal recognition particle receptor, *J. Cell Biol.* 95, 470–477.
6. Gorlich, D., Prehn, S., Hartmann, E., Kalies, K. U., and Rapoport, T. A. (1992) §A mammalian homolog of Sec61p and SecYp is associated with ribosomes and nascent polypeptides during translocation, *Cell* 71, 489–503.
7. Simon, S. M., and Blobel, G. (1991) A protein-conducting channel in the endoplasmic reticulum, *Cell* 65, 371–380.
8. Connolly, T., Rapiejko, P. J., Gilmore, R. (1991) Requirement of GTP hydrolysis for dissociation of the signal recognition particle from its receptor, *Science* 252, 1171–1173.
9. Freymann, D. M., Keenan, R. J., Stroud, R. M., and Walter, P. (1997) Structure of the conserved GTPase domain of the signal recognition particle, *Nature* 385, 361–364.
10. Montoya, G., Svensson, C., Luirink, J., and Sinning, I. (1997) Crystal structure of the NG domain from the signal recognition particle receptor FtsY, *Nature* 385, 365–368.
11. Reyes, C. L., Rutenber, E., Walter, P., and Stroud, R.M. (2007) X-ray structures of the signal recognition particle receptor reveal targeting cycle intermediates, *PLoS ONE* 2, e607.
12. 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.
13. Focia, P. J., Shepotinovskaya, I.V., Seidler, J.A., and Freymann, D.M. (2004) Heterodimeric GTPase Core of the SRP Targeting Complex, *Science* 303, 373–377.
14. Lu, Y., Qi, H.-Y., Hyndman, J. B., Ulbrandt, N. D., Teplyakov, A., Tomasevic, N., and Bernstein, H. D. (2001) Evidence for a novel GTPase priming step in the SRP protein targeting pathway, *EMBO J.* 20, 6724–6734.
15. Shan, S., and Walter, P. (2003) Induced Nucleotide Specificity in a GTPase, *Proc. Natl. Acad. Sci. U.S.A.* 100, 4480–4485.

16. Jagath, J. R., Rodnina, M. V., Lentzen, G., and Wintermeyer, W. (1998) Interaction of guanine nucleotides with the signal recognition particle from *Escherichia coli*, *Biochemistry* 37, 15408–15413.
17. Jagath, J. R., Rodnina, M. V., and Wintermeyer, W. (2000) Conformational changes in the bacterial SRP receptor FtsY upon binding of guanine nucleotides and SRP, *J. Mol. Biol.* 295, 745–753.
18. Moser, C., Mol, O., Goody, R. S., and Sinning, I. (1997) The signal recognition particle receptor of *Escherichia coli* (FTsY) has a nucleotide exchange factor built into the GTPase domain, *Proc. Natl. Acad. Sci. U. S. A.* 94, 11339–11344.
19. 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.
20. Powers, T., and Walter, P. (1995) Reciprocal stimulation of GTP hydrolysis by two directly interacting GTPases, *Science* 269, 1422–1424.
21. Walter, P. and Blobel, G. (1982) Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum, *Nature* 299, 691–698.
22. Walter, P. and Blobel, G. (1983) Disassembly and reconstitution of signal recognition particle, *Cell* 34, 525–533.
23. Poritz, M. A., Strub, K., and Walter, P. (1988) Human SRP RNA and *E. coli*. 4.5S RNA contain a highly homologous structural domain, *Cell* 55, 4–6.
24. Struck, C. R. J., Toschka, H. Y., Specht, T., Erdmann, V. A. (1988) Common structural features between eukaryotic 7SL RNAs, eubacterial 4.5S RNA and scRNA and archaeobacterial 7S RNA, *Nuc. Acids Res.* 16, 7740–7746.
25. Schuenemann, 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. USA* 95, 10312–10316.
26. 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.
27. Li, X., Henry, R., Yuan, J., Cline, K., and Hoffman, N.E. (1995) A chloroplast homologue of the signal recognition particle subunit SRP54 is involved in the posttranslational integration of a protein into thylakoid membranes, *Proc. Natl. Acad. Sci. USA* 92, 3789–3793.
28. 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.
29. 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.
30. Yuan, J., Kight, A., Goforth, R.L., Moore, M., Peterson, E.C., Sakons, J., and Henry, R. (2002) ATP stimulates signal recognition particle (SRP)/FtsY-supported protein integration in chloroplasts, *J. Biol. Chem.* 277, 32400–32404.

31. Shan, S., and Walter, P. (2005) Molecular crosstalk between the nucleotide specificity determinant of the SRP GTPases and the SRP receptor, *Biochemistry* 44, 6214–6222.
32. Peluso, P., Herschlag, D., Nock, S., Freymann, D. M., Johnson, A. E., and Walter, P. (2000) Role of 4.5S RNA in assembly of the bacterial signal recognition particle with its receptor, *Science* 288, 1640–1643.
33. Bishop, A., Buzko, O., Heyeck-Dumas, S., Jung, I., Kraybill, B., Liu, Y., Shah, K., Ulrich, S., Witucki, L., Yang, F., Zhang, C., and Shokat, K. M. (2000) Unnatural Ligands for engineered proteins: new tools for chemical genetics, *Annu. Rev. Biophys. Biomol. Struct.* 29, 577–606.
34. Hwang, Y. W., and Miller, D. L. (1987) A mutation that alters the nucleotide specificity of elongation factor Tu, a GTP regulatory protein, *J. Biol. Chem.* 262, 13081–13085.
35. Weijland, A., Parlato, G., and Parmeggiani, A. (1994) Elongation factor Tu D138N, a mutant with modified substrate specificity, as a tool to study energy consumption in protein biosynthesis, *Biochemistry* 33, 10711–10717.
36. Zhong, X.-M., Chen-Hwang, M.-C., Hwang, Y. W. (1995) Switching nucleotide specificity of Ha-Ras p21 by a single amino acid substitution at aspartate 119, *J. Biol. Chem.* 270, 10002–10007.
37. Spangord, R. J., Siu, F., Ke, A., and Doudna J.A. (2005) RNA-mediated interaction between the peptide-binding and GTPase domains of the signal recognition particle, *Nat. Struct. Mol. Biol.* 12, 1116–1122.
38. Buskiewicz, I., Kubarenko, A., Peske, F., Rodnina, M.V., and Wintermeyer, W. (2005) Domain rearrangement of SRP protein Ffh upon binding 4.5S RNA and the SRP receptor FtsY, *RNA* 11, 947–957.
39. Shan, S., Stroud, R., Walter, P. (2004) Mechanism of association and reciprocal activation of two GTPases, *PLoS Biology* 2, e320.
40. Halic, M., Blau, M., Becker, T., Mielke, T., Pool, 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.
41. 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.

## **Chapter 2:**

# **A Distinct Mechanism to Achieve Efficient SRP– SRP Receptor Interaction by the Chloroplast SRP Pathway**

A version of this chapter has been published as:

Jaru-Ampornpan, P., Nguyen, T. X., and Shan, S. (2009) *Mol. Biol. Cell*, **20** (7), 3965–3973.

**Abstract**

Co-translational protein targeting by the signal recognition particle (SRP) requires the SRP RNA, which accelerates the interaction between the SRP and SRP receptor 200-fold. This otherwise universally conserved SRP RNA is missing in the chloroplast SRP (cpSRP) pathway. Instead, the cpSRP and cpSRP receptor (cpFtsY) by themselves can interact 200-fold faster than their bacterial homologues. Here, cross-complementation analyses revealed the molecular origin underlying their efficient interaction. We found that cpFtsY is five–tenfold more efficient than *E. coli* FtsY at interacting with the GTPase domain of SRP from both chloroplast and bacteria, suggesting that cpFtsY is pre-organized into a conformation more conducive to complex formation. Further, the cargo-binding M-domain of cpSRP provides an additional 100-fold acceleration for the interaction between the chloroplast GTPases, functionally mimicking the effect of the SRP RNA in the co-translational targeting pathway. The stimulatory effect of the SRP RNA or the M-domain of cpSRP is unique to each pathway. These results strongly suggest that the M-domain of SRP actively communicates with the SRP and SR GTPases, and that the cytosolic and chloroplast SRP pathways have evolved distinct molecular mechanisms (RNA vs. protein) to mediate this communication.

## Introduction

The signal recognition particle (SRP) and the SRP receptor (SR) comprise the major cellular machineries that co-translationally deliver newly synthesized proteins from the cytosol to target membranes (1, 2). Co-translational protein targeting begins with recognition of the cargo—ribosomes translating nascent polypeptides containing signal sequences—by the SRP (3). The cargo is brought to the vicinity of the target membrane via the interaction between the SRP and SRP receptor (FtsY in bacteria) (4). Upon arrival at the membrane, SRP unloads its cargo to the protein-conducting channel, composed of the sec61p complex (or secYEG complex in bacteria) (5–7). The SRP and SRP receptor also reciprocally stimulate each other's GTPase activity (8). Thus after cargo unloading, GTP hydrolysis drives disassembly of the SRP•SR complex, returning the components into the cytosol for the next round of protein targeting (9).

The SRP pathway is conserved throughout all three kingdoms of life. Although the protein components of SRP and SR vary across species, the functional core of SRP is a highly conserved ribonucleoprotein complex, comprised of a 54-kD SRP GTPase (SRP54 in eukaryotes or Ffh in bacteria) and an SRP RNA (2). The SRP receptor also contains a conserved GTPase domain that is highly homologous to the GTPase domain in SRP54, and together the GTPase domains of SRP and SR form a unique subgroup in the GTPase superfamily (2). Both proteins contain a central GTPase “G” domain that adopts the classical *Ras*-type GTPase fold (10, 11). Unique to the SRP family of GTPases is an N-terminal extension, termed the “N” domain, that forms a four-helix bundle (10, 11). The N- and G-domains form a structural and functional unit called the NG-domain. In addition to the GTPase domains, the SRP and SR proteins contain unique effector

domains that allow them to carry out their biological functions. SRP has a C-terminal extension, a methionine-rich “M” domain, which interacts with the SRP RNA (12) and with the signal sequence of the cargo (13) SR has an N-terminal extension, an acidic “A” domain, which interacts with the target membrane (14) and potentially with the sec translocon (15).

SRP and SR form a complex with one another directly through their GTPase domains, and reciprocally activate each other’s GTPase activity within the complex (8). Both structural and biochemical analyses suggested that these GTPases undergo major structural rearrangements during complex formation (16, 17). One of the important conformational changes involves the intra-molecular rearrangement at the interface between the N- and the G-domains (16–18). Two conserved motifs at the N–G domain interface, “ALLEADV” on the N-domain and “DARGG” on the G-domain, act as a fulcrum that mediates the re-positioning of the N-domain relative to the G-domain in both SRP and SR (17). In addition, an inhibitory element from the first helix of the N-domain is removed (19). These structural rearrangements bring the two N-domains into close proximity with one another, allowing them to make additional interface contacts that stabilize the complex (18) (17). After a stable SRP•SR complex is formed, additional conformational rearrangements occur in both GTPase active sites to activate GTP hydrolysis within the complex (20).

A novel SRP-dependent protein targeting pathway has been found in chloroplast (21). A unique feature of the cpSRP pathway is that it utilizes a post-translational mode of targeting. Instead of recognizing ribosome•nascent chain complexes as cargo, the cpSRP recognizes light-harvesting chlorophyll-binding proteins (LHCPs) that are

imported into the chloroplast as fully synthesized proteins, and delivers LHCPs from the chloroplast stroma to the thylakoid membrane (22, 23). Analogous to the cytosolic SRP pathways, the cpSRP pathway is mediated by two GTPases, cpSRP54 and cpFtsY, that are close homologues of the cytosolic SRP54 and SR GTPases, respectively. Intriguingly, the other strictly conserved component of the cytosolic SRP pathway, the SRP RNA, has not been found in the cpSRP pathway. Instead, a novel 43-kD protein, cpSRP43, binds to a unique C-terminal extension in cpSRP54, and together the cpSRP43•cpSRP54 complex constitute the chloroplast SRP (24). Although early models suggested that cpSRP43 might act as a functional homologue of the SRP RNA to regulate the GTPase activity of the chloroplast SRP and SRP receptor (see below; (25)), kinetic analyses showed that cpSRP43 does not considerably affect either the complex formation or GTP hydrolysis rates of cpSRP54 and cpFtsY (26). Instead, cpSRP43 interacts specifically with the cargo, the LHCPs, to facilitate substrate recognition (23).

In cytosolic SRP pathways, complex formation between the SRP and SR GTPases is extremely slow, presumably because it is limited by the extensive conformational changes required to form a stable complex (27, 28). The SRP RNA overcomes this problem by enhancing the association rate between the two GTPases 200-fold, bringing the SRP–SR interaction rate to a range appropriate for their biological function (29). Moreover, the SRP RNA accelerates the rate at which the SRP•SR complex hydrolyzes GTP five–tenfold (27, 30). Many reports have suggested that the SRP RNA may play a regulatory role by bridging the communication between cargo binding and the GTPase cycle (12, 31, 32). The SRP RNA therefore plays a crucial role

in the SRP pathway, explaining why it is highly conserved from bacteria, archaea, to eukaryotes.

How does the chloroplast SRP bypass such a key component? Previous kinetic analyses revealed that in the absence of the SRP RNA, the association kinetics between cpSRP54 and cpFtsY is 200-fold faster than that of their *E. coli* homologues, and matches the rate of the RNA-stimulated interaction between bacterial SRP and SR (26). This provides a simple explanation for the absence of the SRP RNA in the cpSRP pathway, but also raises additional questions. What governs the kinetics of interaction between the SRP and SR GTPases? How can the chloroplast GTPases interact much more efficiently than their bacterial homologues despite their high sequence homology? The crystal structure of apo-cpFtsY shows that, compared to free bacterial FtsY, the conformation of apo-cpFtsY is closer to that observed in the Ffh•FtsY complex, suggesting that some of the N–G rearrangement is already in place in cpFtsY prior to complex formation (33). This and additional biochemical results led to a model in which cpFtsY is pre-organized in a conformation that is more conducive to interaction with its binding partner, and thus bypasses some of the conformational changes that limit the rate of association between the bacterial SRP and SR GTPases.

In this work, we present additional evidence for this model by showing that cpFtsY is intrinsically five–tenfold more efficient at interacting with the SRP GTPase. More importantly, we found that the cargo-binding M-domain of cpSRP54, without the help from the SRP RNA, provides an additional ~ 100-fold stimulation in complex formation between the cpSRP and cpFtsY GTPases. Both of these factors allow the chloroplast SRP and SR GTPases to achieve the same efficiency of interaction as the

RNA-catalyzed interaction between their bacterial homologues. The stimulatory effects of the SRP RNA and the M-domain of cpSRP54 are specific to their homologous binding partners and not interchangeable across species, suggesting that the classical and the cpSRP pathways have diverged to use different molecular mechanisms to mediate the communication between the M-domains and the GTPase modules.

## Materials and Methods

**Protein expression and purification.** *E. coli* Ffh and FtsY (47-497) were expressed and purified as described (27). The coding sequence of *E. coli* Ffh NG (1-295) was cloned into pET 28b (Novagen) between NcoI and XhoI restriction sites. The recombinant protein, with a His<sub>6</sub> tag at the C-terminus, was expressed in BL21 DE3\* (Invitrogen) and purified using Ni-NTA affinity column (Qiagen). *E. coli* Ffh NG (1-295) was further purified by cation exchange over a MonoS column (GE Healthcare) using a linear gradient of 150–600 mM NaCl. FtsY (47-497) interacts with Ffh with the same kinetics as either full-length FtsY or FtsY-NG (Supplementary Figure 2.S1), thus the large A-domain in *E. coli* FtsY does not affect the interaction between the SRP and SR GTPases.

cpSRP54 and cpFtsY were expressed and purified as described (26). Mutations of cpFtsY were introduced using the QuikChange Mutagenesis protocol (Stratagene). cpFtsY G288W was purified using the same procedure as that for the wild-type protein. cpFtsY F71V and F71A were purified from inclusion bodies as described (33). The coding sequence of cpSRP54 NG (1-294 of the mature protein) and a His<sub>6</sub> tag at the C-terminus was cloned into pAcUW51 (BD Biosciences) between BamHI and HindIII restriction sites. The resulting plasmid was then used for protein expression from baculovirus at the Protein Expression Center of Caltech. The recombinant cpSRP54 NG-His<sub>6</sub> was purified by affinity chromatography using Ni-NTA twice.

To construct the domain swap mutant proteins, pDMF6 encoding *E. coli* Ffh (10) was modified to contain an EcoRI site before the start of the Ffh M domain. The plasmid encoding FfhNG-cpSRP54M was constructed by replacing the sequence of FfhM (residues 296-453) with a PCR fragment encoding the cpSRP54M (residues 296-488)

using the EcoRI and BamHI restriction sites. The chimeric protein was expressed in Rosetta competent cells (Novagen) and purified using the same procedure as that for the wild-type Ffh protein (27).

**Kinetics.** All GTPase assays were performed at 25 °C in assay buffer [50 mM KHEPES (pH 7.5), 150 mM KOAc, 2mM Mg(OAc)<sub>2</sub>, 2mM DTT, 0.01% Nikkol, 10% glycerol]. GTP hydrolysis reactions were followed and analyzed as described (27). The reciprocally stimulated GTPase reaction between SRP and SR was determined in multiple turnover reactions ([GTP] >> [E]). The concentration dependence of the observed rate constant ( $k_{obsd}$ ) is fit to eq 1, in which  $k_{cat}$  is the rate constant at saturating SR concentrations, and  $K_m$  is the concentration of SR that gives half the maximal rate.

$$k_{obsd} = k_{cat} \times \frac{[SR]}{K_m + [SR]} \quad (1)$$

In these measurements, the basal GTPase rates from FtsY or cpFtsY were determined in side-by-side experiments (Supplementary Table 2.S1) and subtracted from the rates of the stimulated GTPase reactions prior to data analysis. The rate constants  $k_{cat}/K_m$  are listed in Table 2.1. The measurements that are directly compared were performed in side-by-side experiments. The figures show representative data, and Table 2.1 shows the average values from three or more measurements.

**Gel filtration chromatography.** 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]. For cpFtsY mutants (Figure 2.2C and 2.2D), 2 nmols of cpSRP54 and cpFtsY variants were mixed in

the presence of 450  $\mu\text{M}$  GppNHp and the mixture was incubated on ice for 5 minutes before being loaded onto Superdex 200 (GE Healthcare). For experiments in Figure 2.4 and 2.5 nmols of either cpSRP54 or cpSRP54 NG was mixed with equimolar cpFtsY in the presence of 450  $\mu\text{M}$  GppNHp. The mixture was incubated on ice for specified periods of time before being loaded onto Superdex 200. The identities of the peaks were confirmed by reference runs of the individual proteins.

## Results

To better understand the molecular mechanism by which the chloroplast SRP and SR GTPases achieve the same association kinetics as their bacterial homologues without the help from the SRP RNA, a series of cross-complementation experiments were carried out in which we tested the ability of the bacterial SRP receptor to interact with the cpSRP GTPase, and vice versa. The first goal of these experiments is to determine whether the core GTPase modules of SRP and SR, which comprise the heterodimer interface, are conserved across different species. The second goal is to identify unique molecular determinants in each pathway that allow the two different pairs of SRP and SR to efficiently interact with one another.

### *cpFtsY is intrinsically faster than E. coli FtsY at interacting with the SRP GTPase.*

We first asked how well the core GTPase domains from the *E. coli* and chloroplast pathways are conserved. To this end, we tested whether the SRP and SRP receptor GTPases can interact with one another across different species. The SRP and SRP receptor reciprocally stimulate the GTPase activity of each other, providing a convenient assay to monitor complex formation between the two GTPases (27). In this assay, the observed rate constant of GTP hydrolysis is monitored as a function of SR concentration. The slope of the initial linear portion of the concentration dependence represents the rate constant of the reaction:  $\text{GTP} \cdot \text{SRP} + \text{SR} \cdot \text{GTP} \rightarrow \text{products}$  ( $k_{cat}/K_m$ ), and the rate at saturating SR concentrations ( $k_{cat}$ ) represents the GTP hydrolysis rate once the complex is formed. For the *E. coli* GTPases,  $k_{cat}/K_m$  is equal to the association rate constant between SRP and SR during complex formation (27). For the chloroplast

GTPases, this rate constant provides a lower limit for the association rate constant between cpSRP54 and cpFtsY to form an active complex (26). In situations where the value of  $k_{cat}$  is comparable, the differences in  $k_{cat}/K_m$  reflect differences in either the rate or stability of complex formation. Therefore for the analyses below, we used the  $k_{cat}/K_m$  values as indices to compare the relative ability of the SRP and SR GTPases to form a complex with their binding partners.

The SRP and SRP receptors from both systems can cross-react with their heterologous binding partners. The chloroplast SRP receptor cpFtsY can interact with the *E. coli* SRP GTPase Ffh (Figure 2.1A, closed circles) and with the isolated NG-domain of Ffh (Ffh NG; Figure 2.1B, closed circles), with rate constants similar to those with its homologous partner, the NG-domain of cpSRP54 (cpSRP54 NG; Figure 2.1C, closed circles and Table 2.1). Analogously, in the absence of the SRP RNA, the *E. coli* SRP receptor FtsY can interact with its heterologous partner cpSRP54 NG (Figure 2.1C, open circles) with rates similar to those with its homologous partners, Ffh and Ffh NG (Figure 2.1 A and B, open circles). Therefore, the core GTPase modules of SRP and SRP receptor from the two pathways are largely conserved and interchangeable.

An interesting observation from the results in Figure 2.1 is that, in all three cases, cpFtsY is more efficient at interacting with the SRP GTPases than *E. coli* FtsY. When the binding partner is cpSRP54 NG, the  $k_{cat}/K_m$  value for cpFtsY is fivefold above that for *E. coli* FtsY (Figure 2.1C and Table 2.1). Even with the heterologous partners, *E. coli* Ffh and Ffh NG, cpFtsY exhibits about tenfold faster  $k_{cat}/K_m$  over that of *E. coli* FtsY (Figure 2.1 A and B, and Table 2.1). As the GTPase rates at saturating FtsY concentrations (i.e.,  $k_{cat}$ ) are within twofold of each other for FtsY compared to cpFtsY,

these differences in  $k_{cat}/K_m$  values stem primarily from differences in complex formation. Further, the basal GTPase rates of cpFtsY and FtsY are similar to one another and are at least 200-fold slower than the stimulated reaction rates (Supplementary Table 2.S1), indicating that the higher reactivity of cpFtsY over FtsY observed in Figure 2.1 reflects a higher efficiency of complex assembly with cpFtsY. These results provide independent evidence for the previously proposed model that cpFtsY is pre-organized in a conformation that is more conducive to interaction with the SRP GTPases than bacterial FtsY. This effect partly explains why cpSRP54 and cpFtsY can efficiently interact with one another in the absence of the SRP RNA (26, 33).

What are the molecular features in cpFtsY that allow it to interact more efficiently with the SRP GTPases? Complex formation requires the rearrangement of the N-domain relative to the G-domain. Previous crystallographic analyses suggest that, compared to bacterial FtsY, the relative position of the G- and N-domains in cpFtsY is more similar to that in the structure of the Ffh•FtsY complex (33). This may arise, in part, from the tighter packing interactions at the N–G domain interface, especially between the conserved ALLVSDF and SARGG motifs (highlighted in green and blue, respectively, in Figure 2.2A). In cpFtsY, the aromatic ring of Phe71 from the ALLVSDF motif inserts into the core of the N-domain and packs against the SARGG motif (Figure 2.2A). Phe71 is uniquely conserved among chloroplast FtsYs and is replaced by smaller residues in other species. We probed the importance of this packing by mutagenesis. Mutation of cpFtsY Phe71 to valine, its corresponding residue in *E. coli* FtsY, reduces the interaction rate of cpFtsY with cpSRP54 sixfold (Figure 2.2B, green circles). Mutating this residue to Ala reduces the rate even further (~ eightfold; Figure 2.2B, green squares). The

conserved SARGG motif also contributes significantly in the domain–domain packing interaction, as mutation of the universally conserved Gly288 to a bulky tryptophan is detrimental, reducing the value of  $k_{cat}/K_m$  76-fold (Figure 2.2B, blue). None of these mutations significantly reduce the basal GTPase activity of cpFtsY (Supplementary Table 2.S1), indicating that the observed defects are specific to the interaction of cpSRP54 with cpFtsY.

To provide additional evidence that these mutations impair complex formation between cpSRP54 and cpFtsY, we directly measured complex formation using gel filtration chromatography. SRP and SR GTPases form a stable complex in the presence of GppNHp, and the complex can be separated from the monomers by Superdex 200 (Figure 2.2C; (34)). With wild-type cpFtsY efficient complex formation with cpSRP54 was observed, whereas with mutant cpFtsY G288W no detectable complex formation could be found during gel filtration chromatography analysis (Figure 2.2C). Mutant cpFtsY F71V also exhibits a defect in complex formation (Figure 2.2D); the smaller defect of cpFtsY F71V than cpFtsY G288W in the gel filtration analysis is consistent with the less severe reduction of this mutant in  $k_{cat}/K_m$  in the GTPase assay. Together, these results demonstrate that the packing interaction at the N–G domain interface is important for the formation of the SRP•SR complex and possibly gives rise to the advantage of cpFtsY in interacting with the SRP GTPases.

***The M-domain of cpSRP54 accelerates cpSRP54–cpFtsY association.***

The results above demonstrate that the higher reactivity of cpFtsY over *E. coli* FtsY contributes five–tenfold to the 200-fold more efficient association between cpSRP

and cpFtsY in the absence of the SRP RNA (Figure 2.1). We hypothesized that the remaining 50–100-fold effect could arise from cpSRP54, in particular its unique M-domain that interacts with cpSRP43 instead of the SRP RNA.

To test this hypothesis, we compared the interaction rate of cpSRP54 with that of cpSRP54 NG. Remarkably, full-length cpSRP54 exhibits ~ 100-fold faster association kinetics ( $k_{cat}/K_m$ ) compared to the isolated NG-domain of cpSRP54 (Figure 2.3A, open squares vs. circles). Thus, the M-domain of cpSRP54 can act as a functional mimic of the SRP RNA and accelerates the interaction between the cpSRP54 and cpFtsY GTPase domains. The effect of the M-domain is specific to the interaction between the two chloroplast GTPases, as the basal GTP binding and hydrolysis activity of cpSRP54 NG is indistinguishable, within experimental errors, from that of full-length cpSRP54 (Supplementary Table 2.S1).

The faster  $k_{cat}/K_m$  value in the presence of cpSRP54 M-domain implies that the M-domain accelerates the kinetics of protein association between cpSRP54 and cpFtsY. This conclusion is confirmed independently by gel filtration chromatography. With full-length cpSRP54, complex formation is very fast, as the peak representing the cpSRP54•cpFtsY complex is clearly visible as soon as the two proteins are mixed together (Figure 2.4A, black). Complex formation is close to completion within 5 minutes, with less than 40% of cpFtsY remaining in the monomer form (Figure 2.4A, red). In contrast, complex formation is much slower in the case of cpSRP54 NG (Figure 2.4B). Only about 5% of cpFtsY went into the complex after an hour of incubation (Figure 2.4B, red). Qualitatively, these results provide additional evidence that the M-domain of cpSRP54 stimulates complex formation between cpSRP54 and cpFtsY.

The stimulatory effect of the cpSRP54 M-domain is most intriguing in light of the fact that *E. coli* Ffh exhibits similar interaction kinetics with FtsY regardless of whether its M-domain is present (Figure 2.3B; (31, 33)). The interaction between the *E. coli* GTPases is only stimulated when the M-domain binds the SRP RNA (Figure 2.3B, (27)). The SRP RNA, however, does not affect the kinetics of cpSRP54–cpFtsY association (Figure 2.3B). In summary, the results in this section demonstrate that in both the bacterial and chloroplast SRP pathways, the cargo-binding M-domain of SRP communicates with the GTPase domains and stimulates the interaction between the SRP and SR GTPases. These results also suggest that each pathway has evolved unique molecular mechanisms (RNA vs protein) to achieve this communication (see more below).

***The M-domains of SRP specifically communicate with their homologous receptors in each pathway.***

The SRP RNA stimulates the association kinetics between bacterial Ffh and FtsY ~ 200-fold. The results above showed that the M-domain of cpSRP54 stimulates complex formation between the cpSRP and cpFtsY GTPases. We next asked whether the effects of the SRP RNA and the M-domain of cpSRP54 are interchangeable between the two pathways, as the core NG-domains of these proteins can interact with the heterologous partners (Figure 2.1). We therefore tested whether the SRP RNA can exert its stimulatory effect in reactions containing cpFtsY, and analogously, whether the M-domain of cpSRP54 can exert its stimulatory effect in reactions containing *E. coli* FtsY.

Using the GTPase assay in this mix-and-match experiment, we systematically analyzed the effect of the SRP RNA and the M-domain of cpSRP54 on the two different SRP receptors. With *E. coli* Ffh, the association rate between Ffh and FtsY is stimulated 376-fold by the SRP RNA (Figure 2.5A, open vs. closed circles; (27)). In contrast, there is less than twofold difference when the binding partner is cpFtsY instead of *E. coli* FtsY (Figure 2.5A inset, open vs. closed squares and Table 2.1). These results suggest that cpFtsY, unlike *E. coli* FtsY, lacks the ability to respond to the SRP RNA bound to Ffh. Similarly, when cpFtsY interacts with its homologous partner cpSRP54, the SRP RNA does not provide any rate acceleration (Figure 2.5B, open vs. closed squares, and Figure 2.3B). The SRP RNA has no effect on the interaction of *E. coli* FtsY either when paired with cpSRP54 (Figure 2.5B). These results are expected in light of recent work that demonstrates that cpSRP54 does not bind the bacterial SRP RNA ((35); P. J.-A. and S.S., data not shown).

On the other hand, the cpSRP54 M-domain only exerts a stimulatory effect on reactions containing its homologous binding partner cpFtsY (Figure 2.3A). With *E. coli* FtsY as the binding partner, no difference in the association rate is observed for cpSRP54 compared to cpSRP54 NG (Figure 2.5C and Table 2.1). Thus, *E. coli* FtsY lacks the ability to communicate with and respond to the M-domain of cpSRP54.

If the M-domain of cpSRP54 can act as an independent structural unit to stimulate complex formation with cpFtsY, then fusion of the cpSRP54 M-domain to the NG domain of Ffh should stimulate the interaction of Ffh-NG with cpFtsY. To test this possibility, we constructed a chimeric protein, FfhNG-cpSRP54M, by replacing the M-domain of Ffh (including the linker between the G- and M-domains) with that of

cpSRP54. As predicted, the chimeric protein containing the M-domain from cpSRP54 forms an active complex with cpFtsY with a rate constant ( $k_{cat}/K_m$ ) that is ~ 15-fold faster than Ffh NG (Figure 2.6A, circles vs. squares). This stimulation is specific to the interaction between the two GTPases, as the basal GTPase activity of the fusion protein is similar to those of Ffh NG or Ffh (Supplementary Table 2.S1). This is in contrast to *E. coli* Ffh in which the Ffh M-domain does not appreciably affect the interaction of its NG-domain with cpFtsY (Table 2.1). Unfortunately, the effect of the SRP RNA could not be tested in the reciprocal fusion protein, cpSRP54 NG-Ffh M, as the RNA binding motif in the Ffh M-domain of this chimeric protein does not appear to be well formed and has lost the ability to bind the SRP RNA ( $K_d \geq 10 \mu\text{M}$ ; P. J.-A. and S.S., data not shown).

The stimulation induced by the cpSRP54 M-domain in the chimeric protein is specific to cpFtsY, as the fusion protein interacts with *E. coli* FtsY at the same rate as Ffh NG does (Figure 2.6B, circles vs. dashed line; Table 2.1). Further, the interaction of the chimeric protein with *E. coli* FtsY is 100-fold slower than its interaction with cpFtsY (Figure 2.6B, circles vs. dotted line). If no stimulation arises from the M-domain of cpSRP54, only a five–tenfold rate difference between the reactions of cpFtsY and *E. coli* FtsY would be expected (see Figure 2.1). Thus the M-domain of cpSRP54, even when fused to the GTPase domain from a cytosolic SRP, can provide a 10–20-fold stimulation of interactions with cpFtsY. The extent of stimulation by the cpSRP54 M-domain is ~ fivefold smaller in the fusion protein than in native cpSRP54, suggesting that there are additional inter-domain communications between the M- and NG-domains of cpSRP54 that helps position the M-domain for interacting with cpFtsY that cannot be perfectly captured in the fusion protein. Nevertheless, the results with the fusion protein provide

additional support for the model that the M-domain of cpSRP54 acts as a functional mimic of the SRP RNA and kinetically regulates the interaction between the cpSRP54 and cpFtsY GTPases.

## Discussion

Two major differences exist between the cytosolic and chloroplast SRP pathways. First, the cytosolic and chloroplast SRPs recognize significantly different forms of “cargo”. The cytosolic SRP interacts with ribosomes•nascent chain complexes bearing SRP signal sequences (3, 36), whereas the cpSRP binds to its substrates, LHCPs, as fully translated proteins (22, 23). Second, the cpSRP lacks the SRP RNA which is otherwise universally conserved in all the other SRP pathways. Instead, the cpSRP consists of the cpSRP54 GTPase and a novel protein only found in chloroplast, cpSRP43 (21). Previously, we showed that cpSRP54 and cpFtsY can form a complex with one another at rates 200-fold faster than that of their bacterial homologues, therefore bypassing the requirement for the SRP RNA (26). Here, we underscored the molecular mechanisms underlying the large difference in interaction rates between the bacterial and chloroplast SRP and SR GTPases.

Previous biochemical and structural works have suggested a model in which the conformational rearrangement at the N–G domain interface required for SRP–SR complex formation is partly achieved in free cpFtsY, thus allowing it to interact more efficiently with its binding partner cpSRP54 (26, 33). In this work, we provide independent biochemical support for this model by showing that cpFtsY is five–tenfold more efficient at interacting with the GTPase domain of SRP, even when the binding partner is the heterologous *E. coli* Ffh (Figure 2.1). Mutational analyses further supported the importance of the domain arrangement in cpFtsY, especially at the N–G domain interface, to the formation of the cpSRP54•cpFtsY complex ((33) and this work). These results, along with the previous work, support the model that cpFtsY is pre-

organized in a conformation that allows it to better interact with the GTPase domain of SRP.

Even with the higher reactivity of cpFtsY, the isolated GTPase domains of SRP and SR interact very slowly. For the *E. coli* SRP and SR GTPases, their interaction rate is accelerated 200-fold by the SRP RNA. Intriguingly, we found here that the M-domain of cpSRP54 acts as a functional mimic of the SRP RNA, stimulating the interaction between cpSRP54 and cpFtsY ~ 100-fold. This, together with the higher reactivity of cpFtsY, allows cpSRP54 and cpFtsY to achieve the same interaction rate as the RNA-catalyzed interaction between the bacterial SRP and FtsY, and alleviates the otherwise strict requirement for the SRP RNA in cytosolic SRP pathways. These results, together with previous work, provide strong evidence that the cargo-responding domains of the SRPs from both bacterial and chloroplast systems communicate with the GTPase domains and kinetically regulate complex formation between the SRP and SR GTPases (37).

It is interesting to note that, while the GTPase modules (the NG-domains) of SRP and SR can interact with their heterologous binding partners across species, the effects exerted by the M-domains or the SRP RNA are not interchangeable. The stimulatory effect of the SRP RNA or the M-domain of cpSRP54 during complex formation can only be attained when the homologous binding partners are paired together. The SRP RNA can only exert its stimulatory effect during the interaction of *E. coli* Ffh with *E. coli* FtsY. Analogously, the M-domain of cpSRP54 can only exert its stimulatory effect during the interaction of cpSRP54 or the chimeric protein (Ffh NG-cpSRP54 M) with cpFtsY. This specificity implies that the two pathways have evolved distinct mechanisms to mediate

communication between the M- and the GTPase domains. In cytosolic SRP pathways, the SRP receptor has evolved to establish a specific communication with the SRP RNA. Conversely, in the cpSRP pathway, the cpFtsY has evolved to establish a specific communication with the M-domain of cpSRP54.

How does the SRP RNA or the cpSRP54 M-domain stimulate complex formation between the SRP and SR GTPases? Although the detailed molecular mechanism remains unclear, three possible models can be envisioned based on previous and this work. First, the SRP RNA helps to pre-position the Ffh NG-domain such that it is more active at interacting with FtsY (37). By analogy, the M-domain of cpSRP54 might pre-position the NG-domain of cpSRP54. Second, the SRP RNA positions the M-domain of Ffh and allows it to transiently interact with the SRP or SR GTPase during complex formation (38), whereas in cpSRP54 the M-domain itself is properly positioned to establish these interactions. Third, the two pathways use distinct mechanisms to stimulate complex formation. The SRP RNA may provide a direct tether that holds the cytosolic SRP and SR GTPases together during complex formation (29), whereas cpSRP54 could use its M-domain to provide this tether. Our data appear to favor the third possibility. This is because the *E. coli* SRP, even though its M- and NG-domains would be pre-positioned by the SRP RNA, cannot efficiently interact with the chloroplast SRP receptor. Analogously cpSRP54, even though its M- and NG-domains would be pre-positioned, cannot efficiently interact with the *E. coli* FtsY. The stimulatory effect of the SRP RNA and the M-domain of cpSRP54 are highly specific to their homologous receptors, arguing against the first two models, in which the origin of the stimulatory effect would be more generic. These data also suggest that response elements must exist in the GTPases that specifically

interact with the SRP RNA in the case of cytosolic SRP or with the cpSRP54 M-domain in the case of the chloroplast SRP.

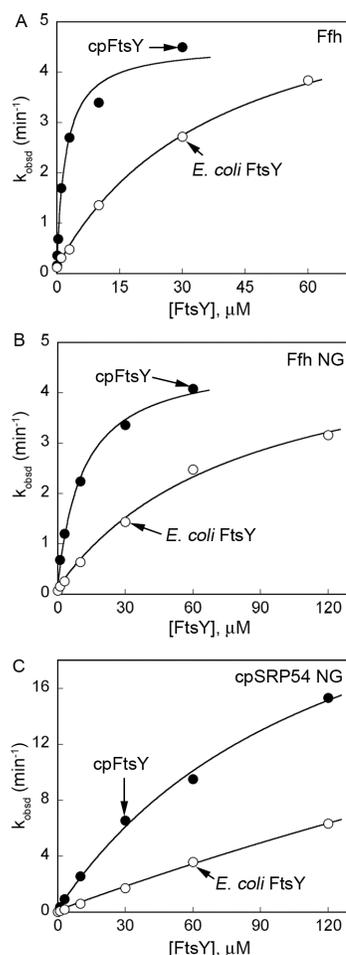
It was recently shown that in cytosolic SRP pathways, the SRP RNA exerts its stimulatory effect on SRP-SR complex assembly only in the presence of cargo or stimulatory detergents such as Nikkol that partially mimic the effect of the cargo (32, 39). This led to the proposal that the SRP RNA acts as a molecular linker that turns on the GTPase cycles of SRP and SR in response to signal sequence binding in the M-domain. Similarly, we found that the stimulatory effect of the cpSRP54 M-domain on the cpSRP54-cpFtsY interaction is also dependent on the presence of the stimulatory detergent Nikkol (Supplementary Figure 2.S2). This suggests that, analogous to the cytosolic SRP, the stimulatory effect of the cpSRP54 M-domain on complex formation between the chloroplast SRP and SR GTPases might occur only in response to binding of its cargo LHCP. Thus the M-domain of cpSRP54 might have also subsumed the function of the SRP RNA as a molecular linker that bridges the communication between cargo binding and SRP-SR complex formation.

**Acknowledgements**

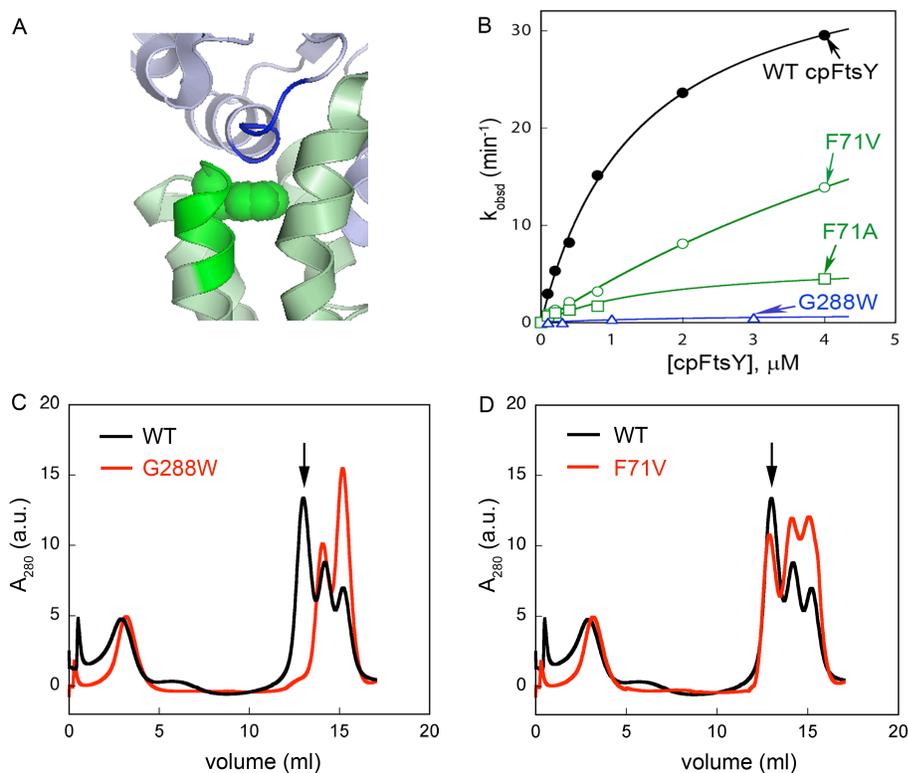
We thank the members of the Shan laboratory for helpful comments on the manuscript. This work was supported by NIH grant GM078024 to S.S. S.S. was supported by the Burroughs Wellcome Fund career award, the Beckman Young Investigator award, and the Packard and Lucile award in science and engineering. P.J.-A. was supported by a fellowship from the Brey Endowment foundation.

**Table 2.1** Summary of the  $k_{cat}/K_m$ ,  $k_{cat}$ , and  $K_m$  values

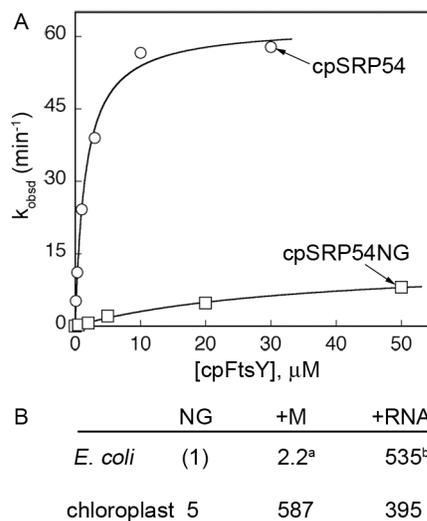
SRP GTPase construct	SRP receptor	SRP RNA	$k_{cat}/K_m$ ( $10^6 \text{ M}^{-1} \text{ min}^{-1}$ )	$(k_{cat}/K_m)^{\text{rel}}$	$k_{cat}$ ( $\text{min}^{-1}$ )	$K_m$ ( $\mu\text{M}$ )
Ffh NG	FtsY	–	$0.06 \pm 0.01$	0.4	$4.2 \pm 1.3$	$68 \pm 11$
	cpFtsY	–	$0.74 \pm 0.18$	5	$3.6 \pm 1.5$	$5.8 \pm 3.9$
Ffh	FtsY	–	$0.16 \pm 0.01$	(1)	$5.1 \pm 1.3$	$32 \pm 10$
	FtsY	+	$60.1 \pm 11.7$	376	$57.3 \pm 7.5$	$1.0 \pm 0.3$
	cpFtsY	–	$1.77 \pm 0.18$	11	$5.1 \pm 1.4$	$3.0 \pm 0.7$
	cpFtsY	+	$2.70 \pm 0.56$	17	$9.0 \pm 2.4$	$3.9 \pm 1.3$
cpSRP54 NG	FtsY	–	$0.06 \pm 0.01$	0.2	$29.5 \pm 0.7$	$494 \pm 62$
	cpFtsY	–	$0.31 \pm 0.09$	(1)	$25.1 \pm 6.4$	$85 \pm 28$
cpSRP54	FtsY	–	$0.06 \pm 0.01$	0.2	$17.6 \pm 0.6$	$256 \pm 0.7$
	FtsY	+	$0.06 \pm 0.01$	0.2	$19.7 \pm 0.5$	$260 \pm 14$
	cpFtsY	–	$35.2 \pm 9.60$	114	$55.5 \pm 16$	$1.9 \pm 0.5$
	cpFtsY	+	$23.5 \pm 7.78$	76	$31.7 \pm 2.9$	$1.4 \pm 0.3$



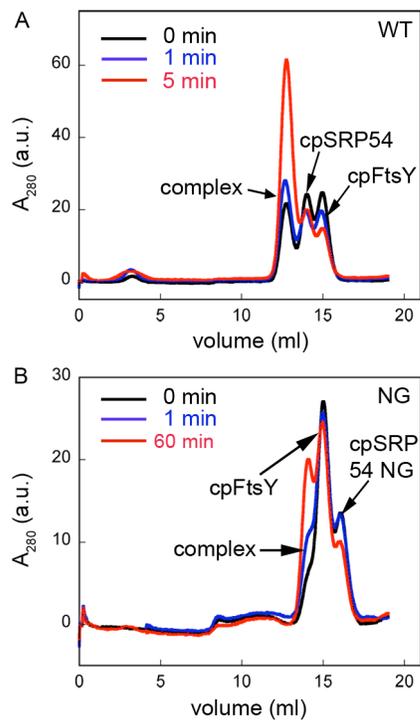
**Figure 2.1** cpFtsY is intrinsically faster than *E. coli* FtsY at interacting with the SRP GTPases. Rate constants for the stimulated GTPase reactions were determined with 500 nM Ffh, Ffh NG, or cpSRP54 NG, and with varying concentrations of cpFtsY or *E. coli* FtsY in the presence of 200  $\mu\text{M}$  GTP. (A) Reactions of *E. coli* Ffh with cpFtsY (●) or with *E. coli* FtsY (○). The data were fit to eq 1 and gave a  $k_{\text{cat}}$  value of 4.4  $\text{min}^{-1}$  and a  $K_m$  value of 2.2  $\mu\text{M}$  for cpFtsY, and a  $k_{\text{cat}}$  value of 6.0  $\text{min}^{-1}$  and a  $K_m$  value of 39  $\mu\text{M}$  for *E. coli* FtsY. (B) Reactions of *E. coli* Ffh NG with cpFtsY (●) or with *E. coli* FtsY (○). The data were fit to eq 1 and gave a  $k_{\text{cat}}$  value of 4.6  $\text{min}^{-1}$  and a  $K_m$  value of 10  $\mu\text{M}$  for cpFtsY, and a  $k_{\text{cat}}$  value of 5.1  $\text{min}^{-1}$  and a  $K_m$  value of 75  $\mu\text{M}$  for *E. coli* FtsY. (C) Reactions of cpSRP54 NG with cpFtsY (●) or with *E. coli* FtsY (○). The data were fit to eq 1 and gave a  $k_{\text{cat}}$  value of 30  $\text{min}^{-1}$  and a  $K_m$  value of 120  $\mu\text{M}$  for cpFtsY, and a  $k_{\text{cat}}$  value  $\geq 16$   $\text{min}^{-1}$  and a  $K_m$  value of 200  $\mu\text{M}$  for *E. coli* FtsY. The values of  $k_{\text{cat}}/K_m$  are listed for comparison in Table 2.1.



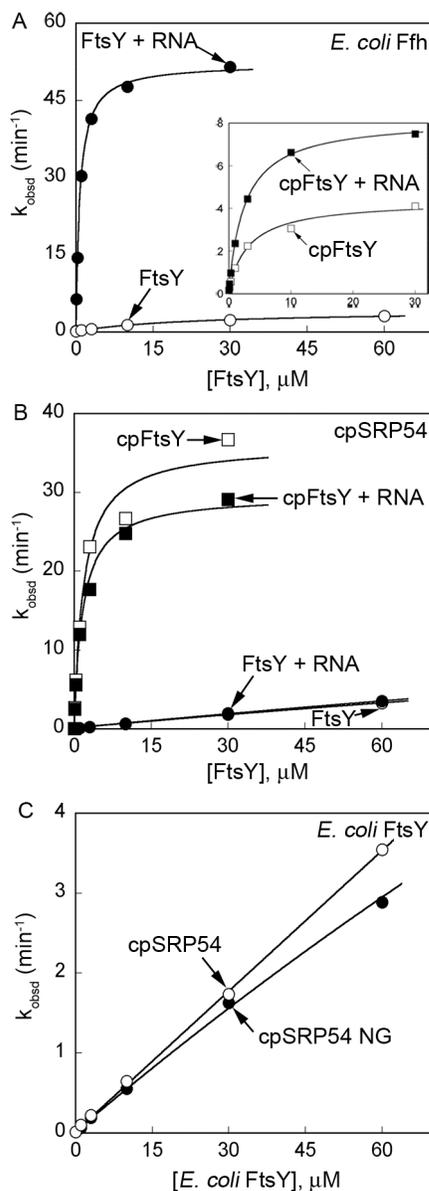
**Figure 2.2** Mutations at the N–G domain interface disrupt formation of the cpSRP54•cpFtsY complex. (A) The N–G domain interface of cpFtsY (PDB 2OG2). The G-domain is shown in pale blue, and the N-domain is shown in pale green. The conserved ALLVSDF and SARGG motifs are highlighted in darker shades of green and blue, respectively. F71 (green) is highlighted in space-filled representation. (B) Rate constants for the stimulated GTPase reactions of cpSRP54 with wild type cpFtsY (●), cpFtsY F71V (green circles), cpFtsY F71A (green squares), or cpFtsY G288W (blue triangles). The data were fit to eq 1 and gave  $k_{\text{cat}}/K_m$  values of  $2.9 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$  for wild-type cpFtsY,  $4.8 \times 10^6 \text{ M}^{-1} \text{ min}^{-1}$  for cpFtsY F71V,  $3.7 \times 10^6 \text{ M}^{-1} \text{ min}^{-1}$  for cpFtsY F71A, and  $3.8 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$  for cpFtsY G288W. Reactions contained 100 nM of cpSRP54 and varying concentrations of cpFtsY in the presence of 100  $\mu\text{M}$  GTP. (C) Complex formation between cpSRP54 and wild-type cpFtsY (black) or mutant cpFtsY G288W (red) was monitored on Superdex 200. An arrow marks the position where the cpSRP54-cpFtsY complex appears. (D) Complex formation between cpSRP54 and wild-type cpFtsY (black) or mutant cpFtsY F71V (red) was monitored on Superdex 200. An arrow marks the position where the cpSRP54–cpFtsY complex appears.



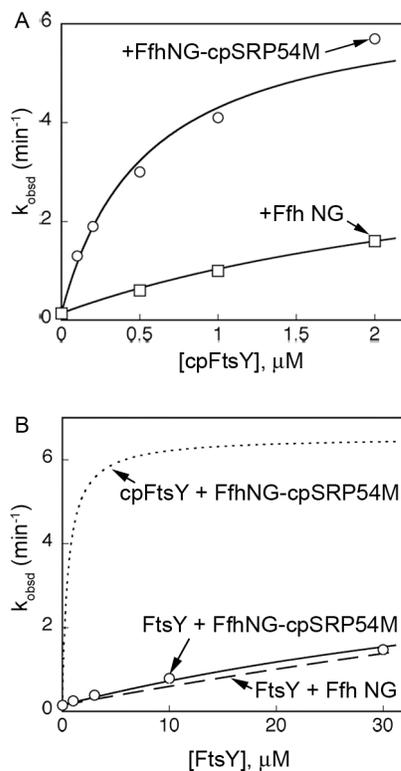
**Figure 2.3** The M-domain of cpSRP54 accelerates the interaction rate of cpSRP54 with cpFtsY. (A) Rate constants for the stimulated GTPase reactions of cpFtsY with cpSRP54 (○) or with cpSRP54 NG (□). The data were fit to eq 1, and the  $k_{cat}/K_m$  values are listed in Table 2.1. (B) Summary of the effect of M-domain and the SRP RNA on complex formation. The  $k_{cat}/K_m$  value of the reference reaction Ffh NG + FtsY → products was set to 1. The effect of the M-domain of Ffh<sup>a</sup> and of SRP RNA<sup>b</sup> has been previously reported (<sup>a</sup> from 31, 33; <sup>b</sup> from 26).



**Figure 2.4** cpSRP54 NG is defective in complex formation. Complex formation between cpFtsY and full-length cpSRP54 (A) or cpSRP54 NG (B) was monitored on Superdex 200 as in Figure 2.2. Reactions were incubated for specified lengths of time before the protein mixtures were loaded onto the column.



**Figure 2.5** The stimulatory effect of the M-domain or the SRP RNA is specific to the SRP receptor in each pathway. (A) Rate constants for the stimulated GTPase reaction of *E. coli* Ffh with *E. coli* FtsY in the presence (●) or absence (○) of 4.5S SRP RNA, or with cpFtsY in the presence (■) or absence (□) of 4.5S SRP RNA (inset). (B) Rate constants for the stimulated GTPase reaction of cpSRP54 with *E. coli* FtsY in the presence (●) or absence (○) of 4.5S SRP RNA, or with cpFtsY in the presence (■) or absence (□) of 4.5S SRP RNA. (C) Rate constants for the stimulated GTPase reaction of *E. coli* FtsY with cpSRP54 (○) or cpSRP54 NG (●). The  $k_{\text{cat}}/K_m$  values were reported in Table 2.1.



**Figure 2.6** The M-domain of cpSRP54 can stimulate interactions with cpFtsY when fused to Ffh-NG. (A) Rate constants for the stimulated GTPase reaction of cpFtsY with the chimeric protein Ffh NG-cpSRP54 M (○) or Ffh NG (□). The data were fit to eq 1 and gave  $k_{cat}/K_m$  values of  $1.2 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$  for Ffh NG-cpSRP54 M and  $7.4 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$  for Ffh NG. Reactions contained 100 nM of FfhNG-cpSRP54M or 500 nM of FfhNG and varying concentrations of cpFtsY in the presence of 100  $\mu\text{M}$  or 200  $\mu\text{M}$  GTP, respectively. (B) The stimulatory effect of the cpSRP54 M-domain in the chimeric protein is specific to cpFtsY. The stimulated GTPase reaction of Ffh NG-cpSRP54M with *E. coli* FtsY (○) was determined as in part A, and nonlinear fits of the data to eq 1 gave  $k_{cat}/K_m$  values of  $6.6 \times 10^4 \text{ M}^{-1} \text{ min}^{-1}$ . The dotted line represents the reaction of the fusion protein with cpFtsY (from part A) and was shown for comparison. The dashed line represents the reaction of Ffh NG with *E. coli* FtsY (from Figure 2.1B) and was shown for comparison.

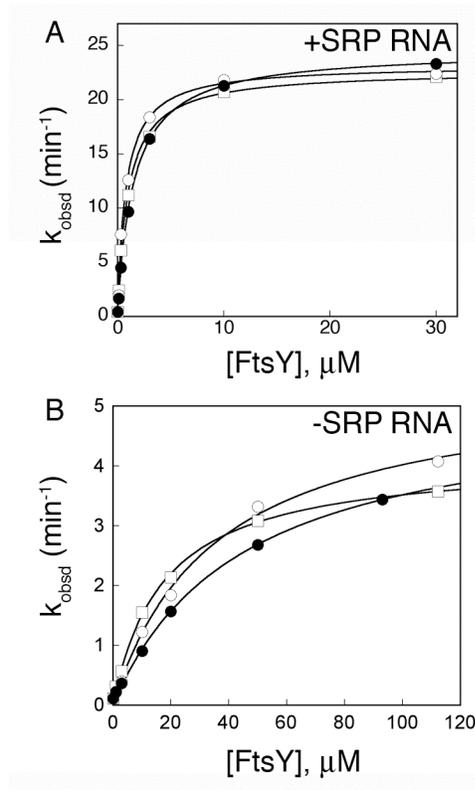
## Supplementary Materials

**Supplementary Table 2.S1** Kinetic constants for the basal GTPase activities of SRP and FtsY proteins used in this study

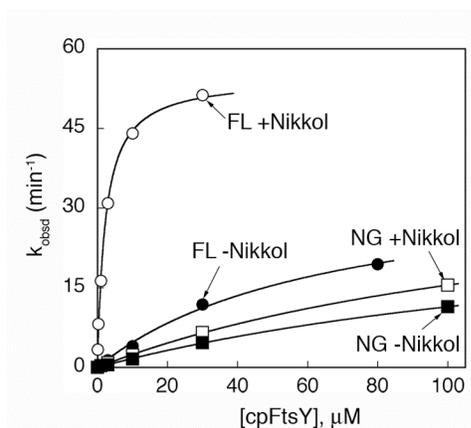
Protein construct	$K_m$ ( $\mu\text{M}$ )	$k_{max}$ ( $\text{min}^{-1}$ )
Ffh <sup>1</sup>	0.3±0.05	0.093±0.002
Ffh NG	1.1±0.2	0.11±0.002
Ffh NG-cpSRP54 M	1.4±0.2	0.12±0.002
<i>E. coli</i> FtsY <sup>1</sup>	14±2	0.012±0.002
<i>E. coli</i> FtsY NG	9.0±3.7	0.0097±0.002
cpSRP54 <sup>2</sup>	2.8±0.4	0.017±0.002
cpSRP54 NG	5.0±0.6	0.015±0.003
cpFtsY <sup>2</sup>	2.1±0.2	0.0045±0.002
cpFtsY F71V	0.3±0.07	0.004±0.0002
cpFtsY F71A	≤4.0	≤0.018
cpFtsY G288W	0.55±0.49	0.0068±0.0044

<sup>1</sup>These values have been previously reported in (27) and confirmed in this study.

<sup>2</sup>These values have been previously reported in (26) and confirmed in this study.



**Supplementary Figure 2.S1** *E. coli* FtsY constructs with different N-terminal extensions exhibit similar reaction rates in the stimulated GTPase reactions with Ffh in the presence (A) or absence (B) of SRP RNA. Rate constants for the stimulated GTPase reaction of *E. coli* SRP or Ffh with *E. coli* FtsY NG (●), *E. coli* FtsY (47-497) (○) and *E. coli* FtsY full-length (□) were determined as described in Methods. In the case of SRP (A), the fit of the data to eq 1 gave  $k_{\text{cat}}/K_m$  values of  $2.2 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$  for FtsY NG,  $2.8 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$  for FtsY (47-497) and  $1.5 \times 10^7 \text{ M}^{-1} \text{ min}^{-1}$  for full-length FtsY. In the case of Ffh (B), the fit of the data to eq 1 gave  $k_{\text{cat}}/K_m$  values of  $1.2 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$  for FtsY NG,  $1.5 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$  for FtsY (47-497) and  $2.3 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$  for full-length FtsY.



**Supplementary Figure 2.S2** The stimulatory effect of cpSRP54 M-domain is observed only in the presence of Nikkol. Rate constants for the stimulated GTPase reactions were determined in the presence (open symbols) and absence (closed symbols) of 0.01% Nikkol with full-length cpSRP54 and cpSRP54-NG, and gave  $k_{\text{cat}}/K_m$  values of  $2.4 \times 10^7$  and  $2.6 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$  for full-length cpSRP54 and cpSRP54-NG in the presence of Nikkol, respectively, and  $5.4 \times 10^5$  and  $1.9 \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$  for cpSRP54 and cpSRP54-NG in the absence of Nikkol, respectively.

**References:**

1. 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.
2. Keenan, R. J., Freymann, D. M., Stroud, R. M., and Walter, P. (2001) The signal recognition particle, *Annu. Rev. Biochem.* 70, 755–775.
3. Walter, P., Ibrahimi, I., and Blobel, G. (1981) Translocation of proteins across the endoplasmic reticulum I. Signal Recognition Protein (SRP) binds to *in vitro* assembled polysomes synthesizing secretory protein, *J. Cell. Biol.* 91, 545–550.
4. Gilmore, R., Blobel, G., and Walter, P. (1982a) Protein translocation across the endoplasmic reticulum: 1. Detection in the microsomal membrane of a receptor for the signal recognition particle, *J. Cell Biol.* 95, 463–469.
5. Simon, S. M., and Blobel, G. (1991) A protein-conducting channel in the endoplasmic reticulum, *Cell* 65, 371–380.
6. Gorlich, D., Prehn, S., Hartmann, E., Kalies, K. U., and Rapoport, T. A. (1992) A mammalian homolog of Sec61p and SecYp is associated with ribosomes and nascent polypeptides during translocation, *Cell* 71, 489–503.
7. Halic, M., Blau, M., Becker, T., Mielke, T., Pool, 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.
8. Powers, T., and Walter, P. (1995) Reciprocal stimulation of GTP hydrolysis by two directly interacting GTPases, *Science* 269, 1422–1424.
9. Connolly, T., Rapiejko, P. J., Gilmore, R. (1991) Requirement of GTP hydrolysis for dissociation of the signal recognition particle from its receptor, *Science* 252, 1171–1173.
10. Freymann, D. M., Keenan, R. J., Stroud, R. M., and Walter, P. (1997) Structure of the conserved GTPase domain of the signal recognition particle, *Nature* 385, 361–364.
11. Montoya, G., Svensson, C., Lührink, J., and Sinning, I. (1997) Crystal structure of the NG domain from the signal recognition particle receptor FtsY, *Nature* 385, 365–368.
12. Batey, R. T., Rambo, R. P., Lucast, L., Rha, B., and Doudna, J. A. (2000) Crystal structure of the ribonucleoprotein core of the signal recognition particle, *Science* 287, 1232–1239.
13. 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.
14. Parlitz, R., Eitan, A., Stjepanovic, G., Bahari, L., Bange, G., Bibi, E., and Sinning, I. (2007) *Escherichia coli* signal recognition particle receptor FtsY contains an essential and autonomous membrane-binding amphipathic helix, *J Biol Chem.* 282, 32176–32184.

15. Angelini, S., Deitermann, S., and Koch, H.G. (2005) FtsY, the bacterial signal-recognition particle receptor, interacts functionally and physically with the SecYEG translocon, *EMBO Rep.* 6, 476–481.
16. Shan, S., and Walter, P. (2003) Induced Nucleotide Specificity in a GTPase, *Proc. Natl. Acad. Sci. U.S.A.* 100, 4480–4485.
17. Focia, P. J., Shepotinovskaya, I.V., Seidler, J.A., and Freymann, D.M. (2004) Heterodimeric GTPase Core of the SRP Targeting Complex, *Science* 303, 373–377.
18. 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.
19. Neher, S. B., Bradshaw, N., Floor, S.N., Gross, J.D., Walter, P. (2008) SRP RNA controls a conformational switch regulating the SRP-SRP receptor interaction, *Nat Struct Mol Biol.* 15, 916–923.
20. Shan, S., Stroud, R., Walter, P. (2004) Mechanism of association and reciprocal activation of two GTPases, *PLoS Biology* 2, e320.
21. Schuenemann, 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. USA* 95, 10312–10316.
22. 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.
23. 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.
24. 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.
25. Goforth, R. L., Peterson, E.C., Yuan, J., Moore, M.J., Kight, A.D., Lohse, M.B., Sakon, J., and Henry, R.L. (2004) Regulation of the GTPase cycle in post-translational signal recognition particle-based protein targeting involves cpSRP43, *J. Biol. Chem.* 279, 43077–43084.
26. Jaru-Ampornpan, P., Chandrasekar, S., 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. Zhang, X., Kung, S., and Shan, S. (2008) Demonstration of a multistep mechanism for assembly of the SRP- SRP receptor complex: implications for the catalytic role of SRP RNA., *J Mol Biol.* 381(3), 581–593.
29. Peluso, P., Herschlag, D., Nock, S., Freymann, D. M., Johnson, A. E., and Walter, P. (2000) Role of 4.5S RNA in assembly of the bacterial signal recognition particle with its receptor, *Science* 288, 1640–1643.

30. Siu, F. Y., Spanggord, R.J., and Doudna, J.A. (2007) SRP RNA provides the physiologically essential GTPase activation function in cotranslational protein targeting, *RNA* 13, 240–250.
31. Bradshaw, N., and Walter, P. (2007) The signal recognition particle (SRP) RNA links conformational changes in the SRP to protein targeting, *Mol Biol Cell* 18, 2728–2734.
32. Bradshaw, N., Neher, S.B., Booth, D.S., and Walter, P. (2009) Signal sequences activate the catalytic switch of SRP RNA, *Science* 323, 127–130.
33. Chandrasekar, S., Chartron, J., Jaru-Ampornpan, P., and Shan, S. (2008) Structure of the chloroplast signal recognition particle (SRP) receptor: domain arrangements govern the SRP-receptor interaction, *J Mol Biol* 375, 425–436.
34. 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, *Biochemica et Biophysica Acta* 1597, 107–114.
35. Richter, C. V., Träger, C., and Schünemann, 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.
36. 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.
37. Jagath, J. R., Matassova, N. B., de Leeuw, E., Warnecke, J. M., Lentzen, G., Rodnina, M. V., Luirink, J., and Wintermeyer, W. (2001) Important role of the tetraloop region of 4.5S RNA in SRP binding to its receptor FtsY, *RNA* 7, 293–301.
38. Zheng, N., and Gierasch, L.M. (1997) Domain interactions in E. coli SRP: stabilization of M domain by RNA is required for effective signal sequence modulation of NG domain, *Mol Cell* 1, 79–87.
39. Zhang, X., Schaffitzel, C., Ban, N., and Shan, S. (2009) Multiple conformational switches in a GTPase complex control co-translational protein targeting, *Proc. Natl. Acad. Sci. USA* 106, 1754–1759

## **Chapter 3:**

### **ATP-Independent Reversal of a Membrane**

### **Protein Aggregate by a Chloroplast SRP Subunit**

A version of this chapter has been published as:

Jaru-Ampornpan, P., Shen, K., Lam, V.Q., Ali, M., Doniach, S., Jia, T.Z., and Shan, S. (2010) *Nat. Struct. Mol. Bio.*, **17** (6), 696–702.

K.S., V.Q.L, M.A., and S.D. contributed to the SAXS experiments and analyses in this paper.

**Abstract**

Membrane proteins impose enormous challenges to cellular protein homeostasis during their post-translational targeting, and require chaperones to keep them soluble and translocation-competent. Here we show that a novel targeting factor in the chloroplast Signal Recognition Particle (cpSRP), cpSRP43, is a highly specific molecular chaperone that efficiently reverses the aggregation of its substrate proteins. In contrast to AAA<sup>+</sup>-chaperones, cpSRP43 utilizes specific binding interactions with its substrate to mediate its disaggregase activity. This “disaggregase” capability can allow targeting machineries to more effectively capture their protein substrates, and emphasizes a close connection between protein folding and trafficking processes. Moreover, cpSRP43 provides the first example of an ATP-independent disaggregase, and demonstrates that efficient reversal of protein aggregation can be attained by specific binding interactions between a chaperone and its substrate.

## Introduction

Protein homeostasis, or proteostasis, is essential for all living cells. It requires precise control of the folding of proteins, their interactions, and their proper localization in a cell (1). Central to the proteostasis network is the cooperative action of an elaborate set of molecular chaperones, which ensures productive protein folding and effectively prevents the misfolding and aggregation of proteins (2–4). Once a protein aggregates, however, only a few chaperones have been identified that can reverse this detrimental process. The central players in these “disaggregase” systems are members of the Clp/Hsp104 family of proteins, which share an architecture of hexameric rings assembled from the ATPases associated with various cellular activities (AAA<sup>+</sup>) and use repetitive cycles of ATP binding and hydrolysis to drive disaggregation (5). These disaggregases often collaborate with the Hsp70/40 chaperones to achieve the efficient reversal of protein aggregation (6).

Membrane proteins pose enormous challenges to the maintenance of proper proteostasis during their post-translational transport. En route to their cellular destinations, membrane proteins must traverse aqueous environments in which they are prone to aggregation or misfolding. Therefore protein-targeting machineries, an essential part of the proteostasis network, must provide chaperones to protect their membrane protein substrates from aggregation and to keep them in a translocation-competent state. Examples include SecB that targets outer membrane proteins to the bacterial plasma membrane (7), Skp that chaperones bacterial outer membrane proteins in the periplasmic space (8), Get3/TRC40 that delivers tail-anchored proteins to the endoplasmic reticulum (9), and the mitochondrial import stimulation factor (MSF) or Hsp70 homologues that

deliver mitochondria- and chloroplast-resident membrane proteins (10). These examples underscore an essential link between chaperone function and protein trafficking.

The localization of light-harvesting complexes in chloroplasts represents a major membrane protein targeting pathway in nature. The substrates of this targeting reaction are the light-harvesting chlorophyll *a/b* binding (LHC) family of proteins, of which the most abundant member, LHCP, constitutes roughly 50% of the thylakoid proteins and is likely the most abundant membrane protein on earth (11). LHCP is synthesized in the cytosol and imported into the chloroplast stroma, where it is delivered to the thylakoid membrane by the chloroplast signal recognition particle (cpSRP) (12). cpSRP is a heterodimer (13) of cpSRP54, which interacts with the SRP receptor on the thylakoid membrane (14), and cpSRP43, a unique chloroplast protein. cpSRP43 is replete with domains and motifs that typically mediate protein–protein interactions: four ankyrin repeats (A1–A4) and three chromodomains (CD1–CD3) (15–17). The ankyrin repeats have been implicated in LHCP recognition (18,19), CD2 interacts with the cpSRP54 M-domain (20), whereas the functions of CD1 and CD3 remain elusive.

LHCP is a highly hydrophobic protein as it is comprised primarily of three transmembrane helices, and its proper folding and assembly require the hydrophobic environment provided by the thylakoid membrane and the binding of 18 photosynthetic pigments at its core (21–23) (Figure 3.1A). Thus in the aqueous environment of the stroma, it is essential to keep LHCP in a soluble, translocation-competent form. Based on the observation of a soluble “transit complex” between LHCP and cpSRP in native gels, cpSRP has been implied to provide a chaperone that maintains the solubility of

LHCP (12). It was further suggested that cpSRP54 binds to the third transmembrane helix of LHCP (13), whereas cpSRP43 binds a highly conserved stretch of 18 amino acids (L18, Figure 3.1A, pink) preceding the third transmembrane domain of LHCP (24, 25). Nevertheless, the potency of cpSRP as a molecular chaperone, the subunit(s) responsible for its chaperone activity, and its mechanism of action have remained unclear.

Here we show that cpSRP43 is a specific and highly effective molecular chaperone for the LHC family of proteins. Importantly, cpSRP43 not only prevents the aggregation of LHCP but also actively re-solubilizes existing LHCP aggregates. In contrast to chaperones built from AAA<sup>+</sup>-ATPases, cpSRP43 uses specific and extensive binding interactions with its substrate to propel the efficient reversal of protein aggregation. These findings demonstrate that a cellular targeting factor could be highly effective at overcoming protein aggregation problems, and that efficient protein disaggregation can be achieved by a small and simple protein fold.

## Materials and Methods

**Protein expression and purification.** Mature cpSRP43, cpSRP54, and cpFtsY were expressed and purified as described (18, 26). cpSRP43-R161A and cpSRP43-Y204A were constructed using the QuikChange procedure (Stratagene), and were expressed and purified as wild-type protein. All deletion mutants of cpSRP43 were cloned between the BamHI and XhoI restriction sites in pGEX-4T-3 (GE Healthcare). The protein segment corresponding to each deletion mutant has been described (17, 18). The GST fusion proteins were expressed in BL21 (DE3)\* cells and purified with Glutathione-S-sepharose (GE Healthcare) in PBS buffer. After Thrombin cleavage, the resulting cpSRP43 proteins were further purified using a MonoQ column (GE Healthcare). Mutant LHCPs,  $\Delta$ DPLG and L164K, were constructed using the QuikChange procedure. LHCA1 and LHCB5 were subcloned into pQE-80L (Qiagen). All LHCP variants were expressed as His-tagged proteins in BL21 (DE3)\* cells as described (27). Inclusion bodies containing recombinant LHCP were purified with Ni-NTA resins (Qiagen) under denaturing condition with 8 M urea.

**Sedimentation assay.** All measurements were performed in Buffer D [50 mM KHEPES (pH 7.5), 200 mM NaCl]. All reactions contained a final concentration of 5  $\mu$ M LHCP and 10  $\mu$ M chloroplast proteins. After incubation at room temperature for ten minutes, the mixtures were centrifuged at top speed in a microfuge, and soluble and pellet fractions were analyzed by SDS-PAGE.

**Light scattering assay.** Measurements were performed in Buffer D on a Beckman DU-640 spectrophotometer. A final concentration of 1  $\mu$ M LHCP was used except when otherwise specified. To analyze prevention of LHCP aggregation, urea-denatured LHCP

(or variants) was diluted into buffer with or without chaperone proteins and the absorbance at 360 nm was recorded for five minutes. The absorbance readings were normalized to that of the sample with no chaperone in side-by-side experiments. Time traces shown are representative of three or more side-by-side experiments. Error bars denote standard deviations from three or more experiments. To analyze reversal of LHCP aggregation, urea-denatured LHCP was allowed to aggregate in buffer for 30 seconds. Chaperone proteins were then added and the measurement continued for ten minutes. All absorbance readings were normalized to that at  $t=30$  s for the reaction that received no chaperone. The disaggregation time courses were fit to eq 1,

$$A = A_f + \Delta A e^{-k_{obsd}t} \quad (1)$$

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 to reach equilibrium. The forward disaggregation rate constants,  $k_f$  (Figure 3.4F), were obtained from the values of  $k_{obsd}$  and  $\Delta A$  as described in Supplementary Note. The concentration dependence of  $k_f$  was fit to eq 2,

$$k_f = k_0 \times \frac{K_d^n}{K_d^n + [cpSRP43]^n} + k_1 \times \frac{[cpSRP43]^n}{K_d^n + [cpSRP43]^n} \quad (2)$$

in which  $k_0$  is the rate of spontaneous LHCP disaggregation in the absence of the chaperone,  $K_d$  is an average equilibrium dissociation constant for binding of cpSRP43 to LHCP aggregates, and  $n$  is the Hill coefficient.

***Kinetic simulations.*** Kinetic simulations were performed with Berkeley-Madonna, version 8.2.3 (R. I. Macey, G. F. Oster, University of California at Berkeley). Details of the simulation are described in Supplementary Note.

**Fluorescence anisotropy.** LHCP were labeled with fluorescein-5'-maleimide (Invitrogen) in denaturing conditions [8 M urea, 50 mM K-HEPES (pH 7.0), 5 mM EDTA]. The labeling efficiency was typically 25–30%. Fluorescence measurements were conducted in SRP buffer [50 mM K-HEPES (pH 7.5), 150 mM KOAc, 2 mM Mg(OAc)<sub>2</sub>] using a Fluorolog 3-22 spectrofluorometer (Jobin Yvon). Labeled LHCP (100 nM) was diluted into buffer containing different concentrations of chloroplast proteins. The samples were excited at 450 nm and the fluorescence anisotropy was recorded at 524 nm. The data were fit to eq 3,

$$A_{obsd} = A_0 + \Delta A \times \frac{[\text{LHCP}] + [\text{pro}] + K_d - \sqrt{([\text{LHCP}] + [\text{pro}] + K_d)^2 - 4 \times [\text{LHCP}][\text{pro}]}}{2 \times [\text{LHCP}]} \quad (3)$$

in which [pro] is the chaperone concentration,  $A_{obsd}$  is the observed anisotropy value,  $A_0$  is the anisotropy value at [pro] = 0,  $\Delta A$  is the total change in anisotropy, and  $K_d$  is the equilibrium dissociation constant.

**LHCP translocation assay.** Translocation assay is based on protease protection of <sup>35</sup>S-labeled LHCP when it is properly integrated into the thylakoid membrane and was performed as described (27). The reactions contained 1 μM cpSRP43 (or its variants), 1 μM cpSRP54, 1 μM cpFtsY (except in the last lane), 1 mM GTP, 1 mM ATP, and salt-washed thylakoid membrane. The integration efficiency was quantified from the intensity of radioactive bands using ImageQuant and normalized to the reaction of wild-type cpSRP43.

**SAXS.** SAXS measurements were performed at beamline 12-ID at the Advanced Photon Source, Argonne National Laboratory (Argonne, IL) with the X-ray energy set at 12 keV. Data were averaged from five exposures (0.2 seconds) at 25 °C using a sample-detector

distance of 2 meters. Background from buffer [20 mM K-HEPES (pH 7.5), 150 mM KOAc, 150 mM NaCl, 1.5 mM Mg(OAc)<sub>2</sub>, 2 mM DTT] was subtracted, and no radiation damage was observed (data not shown). The SAXS data acquired at three different cpSRP43 concentrations showed overlapping Kratky plots (Supplementary Figure 3.S7A), and Guinier analysis of the scattering profile yielded a radius of gyration of 33±1 Å at all protein concentrations, indicating the high quality of data and the absence of aggregation or inter-particle interference.

For the SAXS data obtained at 200 µM cpSRP43 (Figure 3.8A, blue), the program GNOM was used to calculate the intramolecular distance distribution  $P(r)$ . This provided the input for molecular dynamics simulations using DAMMIN and GASBOR (28, 29) to reconstruct a dummy atom model. Ten independent simulations were performed using each software, and all the runs generated the same overall shape. The results were filtered and averaged using SUPCOMB and DAMAVER (30, 31). The filtered models from different simulation software converged on the same shape. The surface map was obtained with Situs (32, 33) and visualized with Chimera (34). Rigid-body docking of the structures of individual fragments of cpSRP43 into the surface map was performed manually based on their shape and the connections between the C- and N-termini of adjacent fragments. The surface map calculated from the molecular model was close to that from the dummy atom model (Supplementary Figure 3.S7B).

## Results

### *cpSRP43 prevents LHCP aggregation.*

Previous work suggested that cpSRP can maintain the solubility of LHCP, based on the ability of a ~ 200 kDa complex, comprised of both cpSRP subunits and LHCP, to migrate into native gels. Nevertheless, the majority of LHCP still deposited as insoluble aggregates, indicating a low efficiency of reconstitution (12). We optimized the reconstitution by: (i) using buffer conditions under which cpSRP43 is most active at interacting with LHCP (Supplementary Note and Supplementary Figure 3.S1); and (ii) presenting urea-denatured LHCP (see Methods) as a defined substrate for cpSRP. The latter strategy was based on the consideration that LHCP enters the chloroplast via two translocases on the chloroplast envelope, both of which translocate unfolded polypeptides (35), and that the proper folding and assembly of LHCP require photosynthetic pigments in the thylakoid membrane (22, 23). Thus, the substrate for cpSRP is most likely a largely unfolded LHCP molecule.

Using a sedimentation assay, we found that virtually all LHCP aggregated in aqueous buffer (Figure 3.1B, lane 1). A twofold molar excess of cpSRP allowed almost all the LHCP to be retained in the soluble fraction (Figure 3.1B, lane 4), demonstrating robust reconstitution of cpSRP's chaperone activity. Consistent with previous observations (25), cpSRP54 alone could not prevent the aggregation of LHCP (Figure 3.1B, compare lanes 6 and 7). To our surprise, cpSRP43 alone could retain LHCP in the soluble fraction, suggesting that cpSRP43 has the ability to chaperone LHCP by itself.

To independently test this conclusion, we used light scattering to monitor formation of high-molecular-weight LHCP aggregates in real time. LHCP aggregated

extensively when it was diluted from urea into aqueous buffer; the aggregation was rapid and close to completion during manual mixing (Figure 3.1C). The light-scattering intensity at equilibrium correlated linearly with LHCP concentration (Figure 3.1C and D), indicating that this assay quantitatively measures the amount of aggregates in our experimental range. LHCP aggregation was reduced ~ 80% when urea-solubilized LHCP was diluted into a solution containing equimolar cpSRP (Figure 3.1E, blue vs. black). Higher concentrations of cpSRP completely suppressed LHCP aggregation (Figure 3.1F, blue). Consistent with results from the sedimentation assay, cpSRP43 prevented LHCP aggregation as efficiently as cpSRP (Figure 3.1E and F, green vs. blue), whereas cpSRP54 did not (Figure 3.1E and F, red). Neither cpFtsY, the chloroplast SRP receptor, nor BSA suppressed LHCP aggregation (Figure 3.1E), further suggesting that the chaperone activity stemmed specifically from cpSRP43. These results indicate that cpSRP43 is primarily responsible for maintaining the solubility of LHCP, whereas cpSRP54 exhibits no significant chaperone activity by itself.

***cpSRP43 binds LHCP with high affinity.***

To quantitatively characterize the binding interactions between cpSRP and its substrate, we labeled a native cysteine (C79) in LHCP with fluorescein-5'-maleimide. Binding of cpSRP or cpSRP43 was detected as an increase in the fluorescence anisotropy of fluorescein-labeled LHCP (Figure 3.2A); this anisotropy change was competed by the L18 peptide (Figure 3.2B), suggesting that it is specific to the LHCP–cpSRP complex. Equilibrium titrations based on this anisotropy change were consistent with 1:1 binding between cpSRP and LHCP, and showed that LHCP bound to cpSRP with an apparent

dissociation constant ( $K_d$ ) of 97 nM (Figure 3.2A, blue). This is likely an upper limit for the true  $K_d$  value of cpSRP–LHCP binding, as a small fraction of LHCP possibly aggregated prior to the addition of cpSRP. Importantly, cpSRP43 bound LHCP with an apparent  $K_d$  of 138 nM, close to that observed with cpSRP (Figure 3.2A, green). In contrast, neither cpSRP54 nor cpFtsY by themselves induced significant anisotropy changes (Figure 3.2A, red and gold). Together, the results of this and previous sections demonstrate that cpSRP43 is sufficient for high affinity binding between cpSRP and its substrate. As suggested previously, cpSRP54 may contribute additional binding interactions for LHCP (36); these interactions could be transient in nature, or did not result in a net increase in overall binding affinity to LHCP.

***cpSRP43 provides a chaperone for the LHC protein family.***

We next tested the ability of cpSRP43 to chaperone other members of the LHCP family including LHCA1 and LHCB5, two close homologues of LHCP (Lhcb1 gene product) (37). Both of these proteins aggregated upon dilution from urea into aqueous buffer, although LHCB5 aggregated more slowly and to a lesser extent than LHCP or LHCA1 (Figure 3.3A and B). In the case of LHCA1, equimolar cpSRP43 could partially help prevent its aggregation whereas equimolar cpSRP suppressed aggregation more efficiently (Figure 3.3A). In the case of LHCB5, equimolar cpSRP43 or cpSRP completely prevented aggregate formation (Figure 3.3B). Thus cpSRP43 can chaperone different members of the LHC protein family and, with more challenging substrates such as LHCA1, cpSRP54 could enhance the chaperone activity of cpSRP43, although cpSRP54 by itself could not chaperone these proteins (data not shown).

***cpSRP43 actively reverses LHCP aggregation.***

AAA<sup>+</sup>-chaperones such as ClpB and Hsp104 exhibit the ability to re-solubilize protein aggregates (5). To test if cpSRP43 can reverse the aggregation of LHCP, we changed the order of addition and allowed LHCP to aggregate upon dilution from 8M urea into aqueous buffer. cpSRP or cpSRP43 was then added when the aggregation was close to completion (Figure 3.4A). Surprisingly, a twofold excess of cpSRP allowed LHCP to partition back into the soluble fraction even after LHCP had already aggregated (Figure 3.4B, compare lanes 4 and 1), and cpSRP43 was sufficient for re-solubilizing the LHCP aggregates (Figure 3.4B, lane 6). As expected, neither cpSRP54 nor BSA reversed LHCP aggregation (Figure 3.4B and data not shown). Thus cpSRP43 not only prevents, but also readily reverses the aggregation of LHCP.

What mechanism underlies this disaggregase activity? Two alternative models could be envisioned. In a passive mechanism, cpSRP43 binds free LHCP molecules that have transiently dissociated from the LHCP aggregate, and prevents them from re-aggregating (Figure 3.4C). Since the aggregation of LHCP ( $k_{-1}$ ) and the binding between cpSRP and soluble LHCP molecules ( $k_2$ ) are fast (Figure 3.1C and data not shown), the rate of disaggregation via this mechanism would be rate-limited by the slow dissociation of LHCP from the aggregates ( $k_1$ ). Therefore, this model predicts that increasing cpSRP43 concentrations would only drive the equilibrium, but would *not* affect the kinetics of LHCP disaggregation (Figure 3.4C, right, and Supplementary Figure 3.S2). Alternatively, cpSRP43 could interact with and remodel the LHCP aggregates, displacing individual LHCP molecules from the aggregate and converting them to soluble

cpSRP43•LHCP complexes (Figure 3.4D). This model predicts that both the equilibrium and rate constants of LHCP re-solubilization will be highly dependent on cpSRP43 concentration (Figure 3.4D, right).

To distinguish between these two possibilities, we followed the disaggregation reaction in real time using the light scattering assay. Addition of increasing amounts of cpSRP43 resulted in increasingly more efficient reversal of LHCP aggregation (Figure 3.4E), and at sufficiently high cpSRP43 concentrations, the disaggregation of LHCP was complete within 200 seconds or less (Figure 3.4E and F). cpSRP43 was able to dissolve the LHCP aggregates with efficiencies that are within twofold of those observed with cpSRP (Supplementary Figure 3.S3). Quantitative analysis of the rate and equilibrium of LHCP disaggregation (Supplementary Note) led to several important conclusions. First, the equilibrium for the disaggregation reaction became more favorable with increasing concentrations of cpSRP43 (Figure 3.4E) and cpSRP (Supplementary Figure 3.S3), consistent with the notion that binding of cpSRP43 prevented LHCP from re-aggregating. Second, the rate constants of disaggregation increased significantly with increasing concentrations of cpSRP43 (Figure 3.4E and F) or cpSRP (Supplementary Figure 3.S3). Third, the disaggregation rate constants exhibit a cooperative dependence on cpSRP43 concentration, with a Hill coefficient of 1.8 (Figure 3.4F and Supplementary Table 3.S1). This suggests that, although each cpSRP43 binds one soluble LHCP molecule (19) (Figure 3.2A), disaggregation requires the cooperative action of more than one cpSRP43 molecule to dislodge LHCP from the aggregates. These results are consistent with predictions from the active mechanism (Figure 3.4D) but could not be accounted for by the passive mechanism (Figure 3.4C and Supplementary Figure 3.S2), and strongly

suggest that cpSRP43 is an effective molecular chaperone that actively dissolves the aggregates formed by its substrate protein.

***Specific binding interactions drive chaperone activity.***

Most of the known chaperones that reverse protein aggregation are large macromolecular assemblies built from AAA<sup>+</sup>-ATPases and rely on mechanical forces powered by ATP hydrolysis. How does cpSRP43, a small protein with no ATPase sites, efficiently reverse protein aggregation? We reasoned that the AAA<sup>+</sup>-chaperones need to act on a variety of substrates via highly promiscuous interactions, and have not evolved specific and extensive interactions with their substrates during disaggregation (38–40). cpSRP43, on the other hand, is dedicated to the LHC family of proteins. We therefore hypothesized that cpSRP43, instead of being driven by ATP hydrolysis, utilizes specific binding interactions with its substrate to drive its chaperone/disaggregase activity.

Previous work showed that cpSRP43 specifically binds to the L18 motif of LHCP (Figure 3.1A, pink), a sequence highly conserved throughout the LHC protein family (24, 25). The crystal structure of an L18 peptide bound to the CD1–Ank4 fragment of cpSRP43 identified a DPLG motif in L18 as an important binding site for cpSRP43 (19) (Supplementary Figure 3.S4A). We tested the importance of these binding interactions by deleting this motif ( $\Delta$ DPLG) or introducing a single mutation, L164K, into DPLG. cpSRP43 or cpSRP, even at a tenfold molar excess, could not suppress the aggregation of  $\Delta$ DPLG (Figure 3.5A and Supplementary Figure 3.S4B). Similarly, the L164K mutation severely disrupted the binding LHCP to cpSRP (Figure 3.5B), and abolished the ability of cpSRP43 or cpSRP to prevent LHCP aggregation (Figure 3.5A and Supplementary

Figure 3.S4B). These results are consistent with previous work that showed that LHCP-L164K failed to integrate into the thylakoid membrane by the cpSRP pathway (19).

Reciprocally, we mutated residues in cpSRP43 that make important contacts to the L18 peptide (19) (R161A and Y204A). cpSRP43-R161A exhibited significantly reduced chaperone activity, requiring a tenfold molar excess to attain the same solubilization of LHCP as equimolar wild-type cpSRP43 (Figure 3.5C). cpSRP43-Y204A completely abolished the ability of cpSRP43 to suppress LHCP aggregation (Figure 3.5C). Similar results were obtained with cpSRP complexes assembled from these cpSRP43 mutants (Supplementary Figure 3.S4C). The defects of these mutant proteins in chaperoning LHCP correlated with their defects in binding LHCP: cpSRP R161A bound LHCP with a  $K_d$  value an order of magnitude higher than that of wild-type cpSRP (Figure 3.5D, squares vs. circles), whereas the cpSRP-Y204A mutation more severely disrupted LHCP binding (Figure 3.5D, diamonds). Together, these mutational results demonstrate that cpSRP43 exhibits high specificity for the LHC family of proteins, and that these specific interactions are essential for the chaperone activity of cpSRP43.

### ***Essential roles of chromodomains.***

Previous work revealed a highly modular domain structure of cpSRP43, with three CDs and an ankyrin repeat domain between the first and second chromodomains (15–17) (Figure 3.6A). As cpSRP43 is a protein targeting factor, some of these motifs could be used for functions other than chaperoning LHCP. We therefore defined the

minimal domain requirement for the chaperone activity of cpSRP43 by testing cpSRP43 mutants in which the individual structural motifs were systematically deleted.

Deletion of any of the ankyrin repeats in cpSRP43 abolished its ability to prevent LHCP aggregation, indicating that all the ankyrin repeats are required for chaperone activity (Figure 3.6B). Surprisingly, the CD1–Ank4 fragment ( $\Delta$ CD2 $\Delta$ CD3), despite its ability to bind the L18 peptide as well as wild-type cpSRP43 (19), failed to prevent the aggregation of LHCP (Figure 3.6C), suggesting that additional interactions between the chromodomains of cpSRP43 and the remainder of LHCP are essential for the chaperone activity of cpSRP43. While deletion of the first chromodomain abolished the ability of cpSRP43 to suppress LHCP aggregation (Figure 3.6C), mutants in which either the second or the third chromodomain was deleted could prevent and reverse the aggregation of LHCP almost as efficiently as wild-type cpSRP43 (Figure 3.6C and D, and Supplementary Figure 3.S5). Thus, all of the ankyrin repeats and at least one chromodomain on both the N- and C-termini of the ankyrin repeat domain are required to support cpSRP43's chaperone activity.

Consistent with these results, the cpSRP43 deletion mutants that can prevent and reverse LHCP aggregation,  $\Delta$ CD2 and  $\Delta$ CD3, exhibited high affinity binding to LHCP, with  $K_d$  values within two- to threefold of that of wild-type cpSRP43 (Figure 3.6E; green and gold vs. black). In contrast,  $\Delta$ CD1 and  $\Delta$ CD2 $\Delta$ CD3 bound to LHCP with much weaker affinities (Figure 3.6E, blue and red). The results in Figures 3.4 and 3.5 demonstrate a strong correlation between the strength of the cpSRP43–LHCP binding interactions and the ability of cpSRP43 to chaperone LHCP, supporting the notion that these binding interactions and chaperone activity are highly coupled. Together with the

observation that full-length LHCP binds cpSRP43 at least 10–20-fold stronger than the L18 peptide, these results further indicate that the interaction of LHCP with cpSRP43 is extensive and involves not only the previously identified contacts between the L18 motif and ankyrin repeats, but also interactions of the transmembrane domains of LHCP and the chromodomains of cpSRP43. Finally, the mutants that failed to efficiently bind and chaperone LHCP exhibited strong defects in the targeting and integration of LHCP to the thylakoid membrane, whereas  $\Delta$ CD3, which showed no appreciable defect in the chaperone activity, only mildly affected the integration efficiency (Figure 3.7). Although  $\Delta$ CD2 showed no significant defect in chaperone activity, this deletion mutant could not support translocation because CD2 is required to interact with cpSRP54 (20). These results highlight the essential role of cpSRP43's chaperone activity in maintaining the translocation competence of LHCP.

### ***Structural reconstruction of cpSRP43.***

To address how cpSRP43 could provide sufficient surface to bind a substrate of almost its own size, we reconstructed the global structure of cpSRP43 using small-angle X-ray scattering (SAXS). SAXS reports on the global size and shape of macromolecules in solution, and, in combination with molecular dynamics simulations, can generate a global structural model at resolutions of 10–15 Å (41). High quality SAXS data were acquired for cpSRP43 (Figure 3.8A, blue). Based on this SAXS profile, multiple independent molecular dynamics simulations using different software converged on the dummy atom model shown in Figure 3.8B. The reconstruction was further validated by

calculating a theoretical Kratky curve from this model, which overlapped well with the experimental profile (Figure 3.8A, red vs. blue).

The reconstruction revealed cpSRP43 to be an elongated, curved molecule  $\sim 120$  Å in length and  $\sim 40$  Å in sectional diameter (Figure 3.8B), consistent with an earlier suggestion based on analytical ultracentrifugation results (13). This narrow shape allowed us to dock the previously obtained high-resolution structures of the individual fragments (19, 42) successively into the SAXS reconstruction to generate a molecular model for cpSRP43 (Figure 3.8C). The structure of the CD1–Ank4 fragment fit well into the longer arm of the SAXS reconstructed shape; the small curvature in the crystal structure (19) was independently observed in the SAXS model, increasing our confidence in the position and orientation of this fragment. The good fit of the crystal structure of this fragment into the SAXS reconstruction also suggested that no major structural changes in the CD1–Ank4 fragment were induced by CD2 and CD3; therefore, it seems unlikely that the defect of this fragment in binding and chaperoning LHCP (Figure 3.8C and E) stems from an inactive conformation of CD1–Ank4 without the additional chromodomains. CD2 and CD3 were fit into the middle and other end of the reconstructed shape, respectively; their precise orientations could not be assigned at this resolution without further structural or biochemical constraints (Figure 3.8C). Nevertheless, the elongated shape of cpSRP43 revealed by this model suggests that this chaperone could provide extensive surface area for binding its substrate protein despite its small size, and might be well suited to keep the LHCP molecules in an extended, translocation competent conformation.

## Discussion

Post-translational targeting of membrane proteins poses enormous challenges to cellular proteostasis and mandates intimate coupling between the transport and chaperone functions of protein-targeting machineries. The high abundance and highly hydrophobic nature of the LHC family of proteins necessitate a highly effective molecular chaperone during their transport. Here we demonstrated that cpSRP43 efficiently fulfills these requirements. Thus, cpSRP provides a robust model system to test the limits of the chaperone capacity of targeting machineries and to understand their mechanism of action.

The ability of cpSRP43 to reverse LHCP aggregation is intriguing; only a few other chaperones, all of which based on AAA<sup>+</sup>-ATPase assemblies, have been demonstrated to effectively reverse protein aggregation (5). The efficiency with which cpSRP43 re-solubilizes LHCP aggregates is on par with those exhibited by the AAA<sup>+</sup>-ATPase machines. In a similar experimental setup, ClpB, with the help of DnaK/DnaJ/GrpE, dissolves protein aggregates with half-times varying from several minutes to a few hours (43, 44). Hsp104, the eukaryotic homologue of ClpB, dissolves heat-aggregated GFP and facilitates its refolding on a time scale of minutes to hours (45). With the help of an adaptor protein MecA, ClpC reverses protein aggregation on a time scale similar to ClpB (46). Here, cpSRP43 at a concentration of 8  $\mu$ M or higher completed the disaggregation process within 100–200 seconds. This efficiency is remarkable given that all the other chaperones are massive macromolecular machines of over 600 kDa and rely on mechanical forces powered by ATP hydrolysis to effect their disaggregase activity, whereas cpSRP43 contains no ATPase sites and the minimal functional unit required to support its chaperone activity is  $\sim$  35 kDa, only slightly larger

than its substrate protein. Although a strict comparison of the disaggregase activity between cpSRP43 and other chaperones could not be made due to the largely unknown and possibly different nature of LHCP aggregates compared to those of previously used model proteins, it is evident that cpSRP43 can function as an effective disaggregase for its substrate proteins without any co-chaperone or ATP consumption. Indeed, cpSRP43 could also reverse heat-aggregated LHCP, albeit with less efficiency (Supplementary Figure 3.S6); this suggests that cpSRP43 can re-solubilize LHCP aggregates generated under different conditions, but the physical or chemical nature of the aggregate affects the efficiency with which this chaperone works.

How does a relatively small chaperone such as cpSRP43 efficiently reverse protein aggregation without ATP hydrolysis? Although the precise molecular mechanism remains to be defined, our results here provided several important clues. First, cpSRP43 has established highly specific and extensive interactions with its substrate, using not only its ankyrin repeats to contact the L18 motif of LHCP, but also additional interactions involving its chromodomains and the transmembrane domains of LHCP. The extended structure of cpSRP43 is consistent with the notion that this chaperone could provide extensive binding surfaces for its substrate protein despite its small size. These binding interactions are crucial for supporting the chaperone activity of cpSRP43. Second, the cooperative dependence of the protein disaggregation rate on cpSRP43 concentration strongly suggests that binding of the first cpSRP43 molecule induces conformational changes in the aggregated LHCP so that the second cpSRP43 molecule can bind more strongly. This supports an active role of cpSRP43 in remodeling the LHCP aggregates, and suggests that the binding interactions of cpSRP43 with LHCP induce changes in the

aggregated LHCP molecules that may disrupt their contacts within the aggregate, thereby dislodging LHCP molecules from the aggregate. As a protein-targeting factor, cpSRP43 most likely keeps solubilized LHCP molecules in a largely unfolded, translocation competent state. The thylakoid membrane environment and the binding of chlorophylls eventually drive the proper folding and assembly of LHCP.

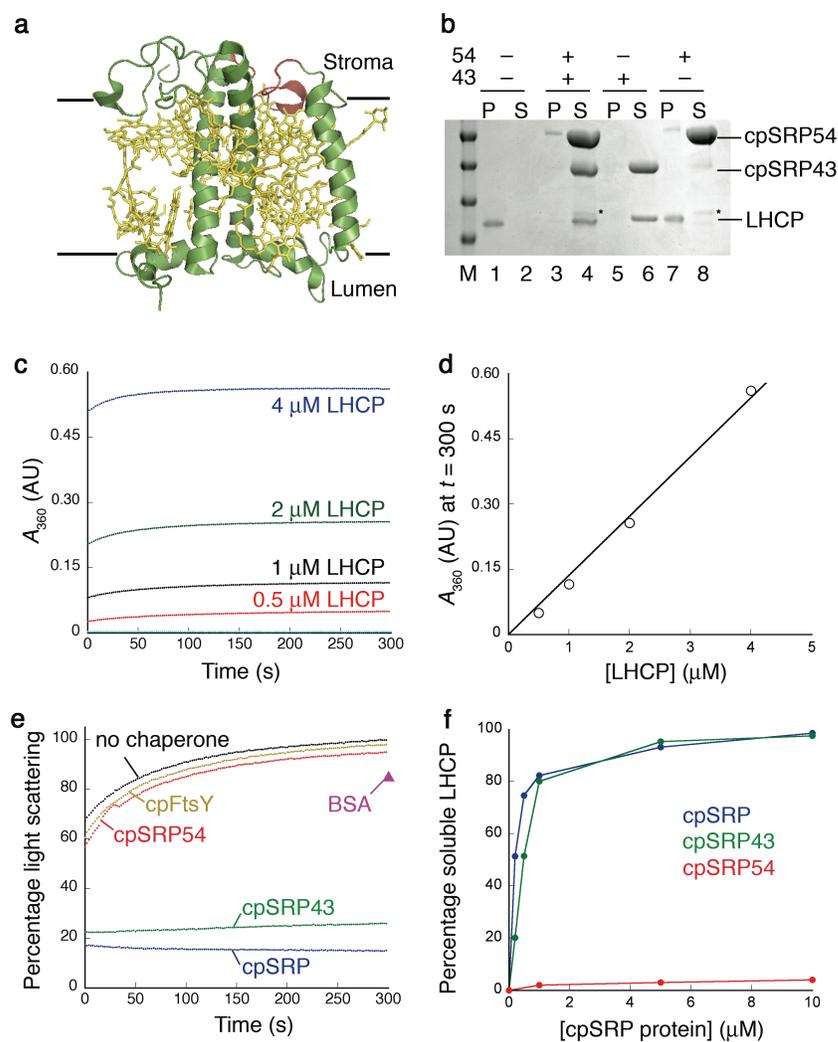
The ability of targeting factors to reverse protein aggregation would allow them to more efficiently capture their substrate proteins, and may reflect a more general feature of chaperones involved in post-translational protein targeting, such as SecB, MSF, and Hsp70. Indeed, SecB has been suggested to passively disaggregate proteins by binding to polypeptides that have dissociated from the aggregate (47). MSF likely provides another example in which a protein targeting factor can efficiently reverse aggregation (48). This chaperone restores the import competence of aggregated precursor proteins in an ATP-dependent manner (48), and the observation that aggregated protein substrates stimulate MSF's ATPase activity (49) strongly supports an active mechanism of disaggregation. Analogous to cpSRP43, MSF specifically recognizes the presequence of mitochondrial precursor proteins, and these specific binding interactions play a crucial role in its chaperone activity (49).

To our knowledge, cpSRP43 provides the first example of a simple solution to overcome protein aggregation problems without energy input from ATP. The key difference between the action of cpSRP43 and ATP-driven disaggregases may arise from their different substrate specificity. Most chaperones from the Clp and Hsp family have evolved to bind a variety of substrates via generic hydrophobic interactions, such that they can rescue proteins from aggregation regardless of sequence identity (38–40).

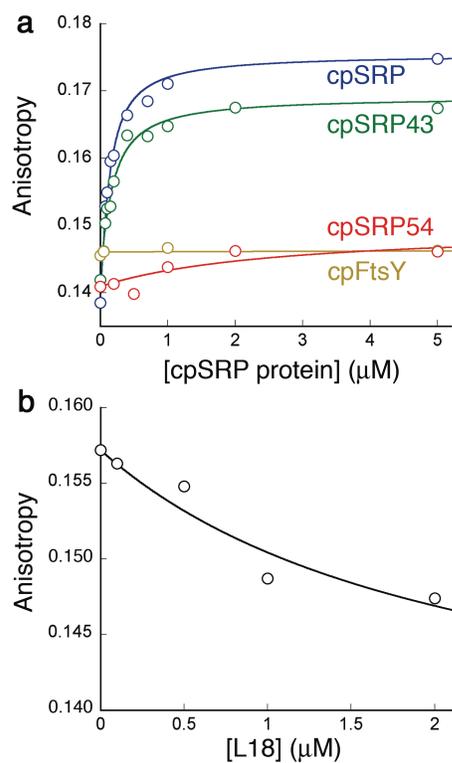
Sacrificing specificity for variety, these chaperones may resort to larger and more elaborate architectures, cooperative action of multiple chaperones, as well as mechanical forces powered by ATP hydrolysis to exert their action. In contrast, cpSRP43 is found only in the chloroplasts of green plants, and its evolution likely coincided with that of its substrates, the LHCPs (11). Thus, cpSRP43 is dedicated to the LHC family of proteins, and hence has established extensive and highly specific binding interactions with its substrates. With adequate binding interactions, cpSRP43 can bypass the massive architecture, the elaborate chaperone network, and the dependence on ATP. Thus, cpSRP43 illustrates a simple principle that efficient reversal of protein aggregation can be attained with a small protein fold and without external energy input, as long as adequate binding interactions are established between a chaperone and its substrate.

**Acknowledgements**

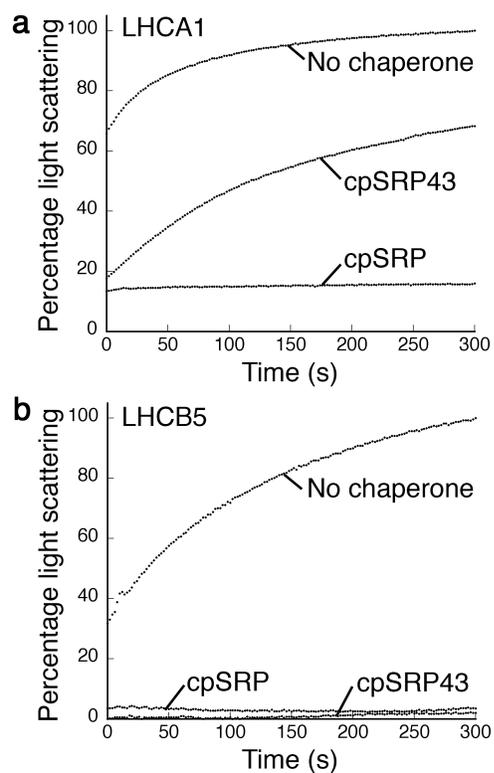
We thank T.X. Nguyen for help with the LHCP translocation assay, Drs. R. Henry for expression plasmids of cpSRP43 and LHCP, C. Robinson for plasmids encoding LHCA1 and LHCB5, Z. Liu for help and advice with LHCP reconstitution, A. Sim and V. Chu of the Doniach group and D. Herschlag for help and advice with the SAXS experiments, and D.C. Rees, W.M. Clemons, A. Varshavsky, N. Pierce, and members of the Shan laboratory for comments on the manuscript. This work was supported by NIH grant GM078024, and career awards from the Burroughs Wellcome Foundation, the Henry and Camille Dreyfus foundation, the Beckman foundation, and the Packard foundation to S.S., and by NIH program project grant P01-GM-66275 to S.D. and D.H.



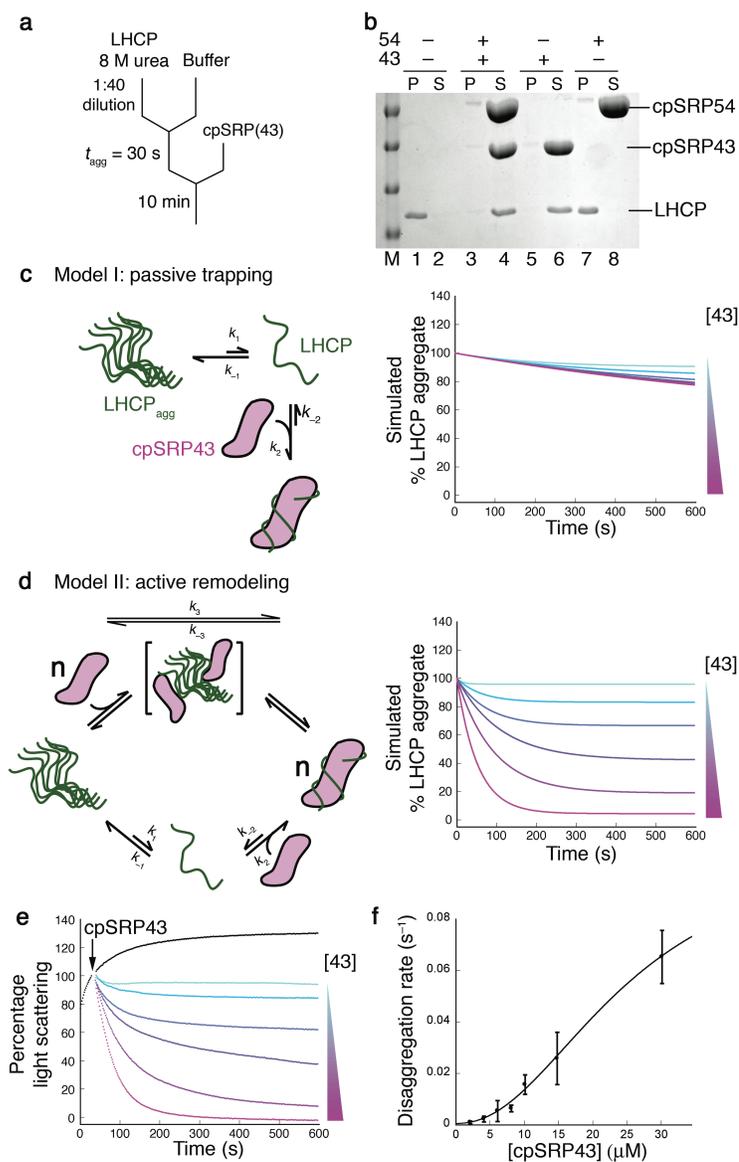
**Figure 3.1** cpSRP43 is sufficient for preventing aggregation of LHCP. (A) The crystal structure of an LHCP monomer (21) (PDB ID: 1RWT) in complex with photosynthetic pigments (gold sticks). The black lines depict the thylakoid membrane. Pink highlights the L18 motif. (B) Sedimentation analysis of the ability of cpSRP or its individual subunits to prevent LHCP aggregation. P and S denote the pellet and soluble fractions, respectively. The asterisks mark a small contamination during the preparation of cpSRP54. (C) Time courses for aggregation of LHCP at different starting LHCP concentrations. (D) The light scattering from aggregates is proportional to LHCP concentration. (E) Time courses for LHCP aggregation in the absence (black) or presence of cpSRP (blue), cpSRP43 (green), cpSRP54 (red), or cpFtsY (gold). The magenta triangle represents LHCP aggregation in the presence of BSA. (F) Concentration dependence of LHCP solubilization by cpSRP (blue), cpSRP43 (green), and cpSRP54 (red).



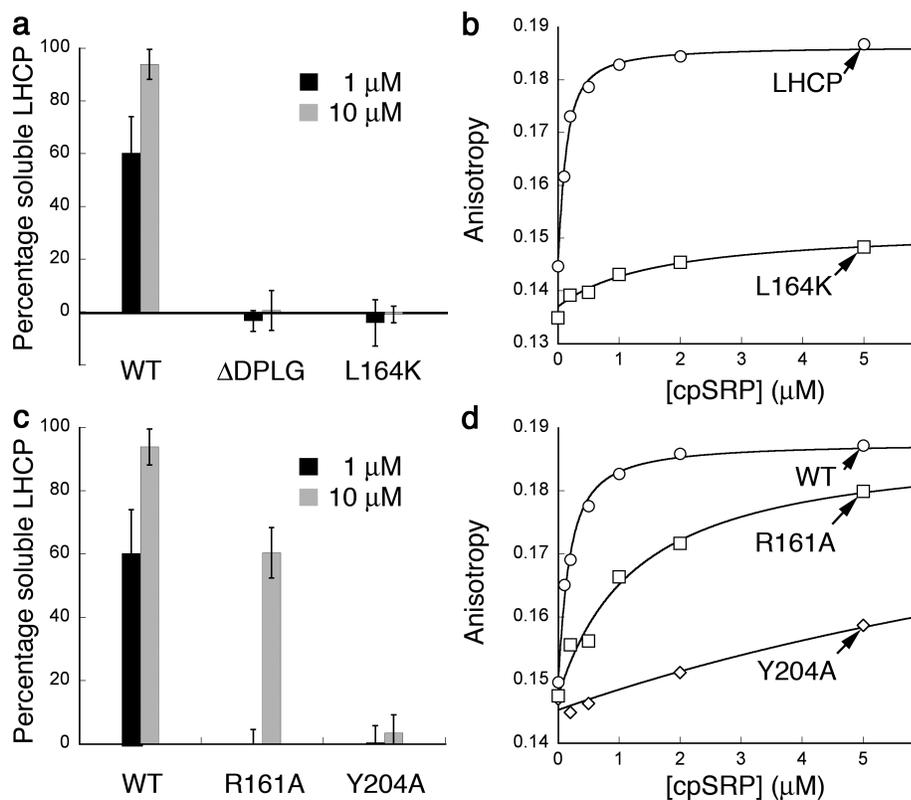
**Figure 3.2** LHCP binds with high affinity to cpSRP43. (A) Binding of LHCP to cpSRP components measured by fluorescence anisotropy. The data were fit to eq 3 and gave  $K_d$  values of 97 nM for cpSRP (blue) and 138 nM for cpSRP43 (green). cpSRP54 (red) and cpFtsY (gold) showed no significant binding to LHCP. (B) The L18 peptide competes with fluorescein-labeled LHCP in binding to cpSRP43. Nonlinear fit of the data gave an apparent  $K_i$  of 2.2  $\mu\text{M}$ , close to the  $K_d$  value of the L18–cpSRP43 interaction observed previously (19).



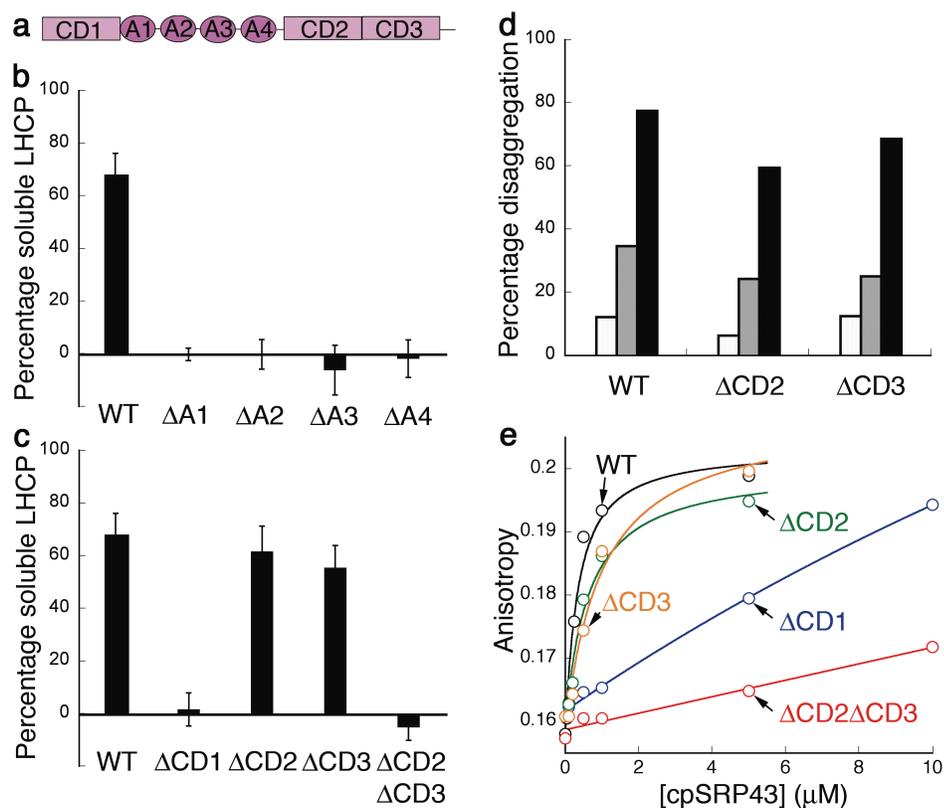
**Figure 3.3** cpSRP and cpSRP43 chaperone various members of the LHC family. Time courses for aggregation of LHCA1 (A) and LHCB5 (B) in the absence or presence of equimolar cpSRP or cpSRP43.



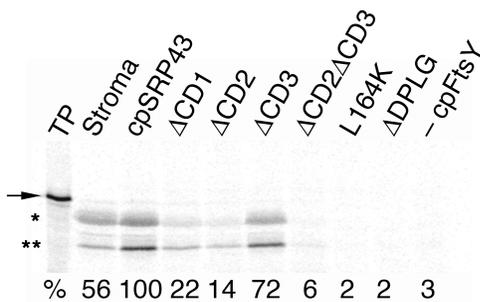
**Figure 3.4** cpSRP43 actively reverses LHCP aggregation. (A) Reaction scheme of the disaggregation assay. (B) Sedimentation analysis of LHCP disaggregation by cpSRP or its individual subunits. P and S denote the pellet and soluble fractions, respectively. (C–D) Models for LHCP disaggregation via a passive (C) or an active mechanism (D), as described in text. The right panels show kinetic simulations for each model at varying concentrations of cpSRP43 (see also Supplementary Figure 3.S2) as described in Supplementary Note. (E) Time courses for disaggregation of LHCP (1  $\mu\text{M}$ ) at varying concentrations (2–10  $\mu\text{M}$ ) of cpSRP43. The black arrow marks the time of cpSRP43 addition. (F) Concentration dependence of the forward rate constants of disaggregation reactions ( $k_j$ ; see Methods). Fits to eq 2 gave a Hill coefficient of 1.8.



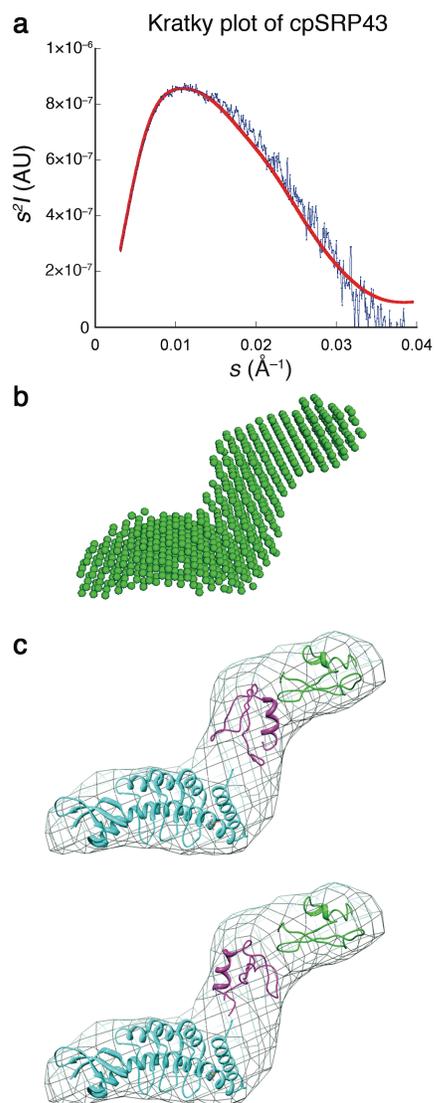
**Figure 3.5** Specific binding interactions between LHCP and cpSRP43 are essential for the chaperone activity. (A) The amount of soluble LHCP (1  $\mu\text{M}$ ) or its variants,  $\Delta\text{DPLG}$  and L164K, at equilibrium in the presence of 1  $\mu\text{M}$  (black) or 10  $\mu\text{M}$  (gray) cpSRP43. (B) Binding of LHCP (○) and LHCP-L164K (□) to cpSRP. Fits of data to eq 3 gave  $K_d$  values of 76 nM for LHCP and  $>5 \mu\text{M}$  for L164K. (C) The amount of soluble LHCP (1  $\mu\text{M}$ ) at equilibrium in the presence of 1  $\mu\text{M}$  (black) or 10  $\mu\text{M}$  (gray) of cpSRP43 or its mutants, R161A and Y204A. (D) Binding of LHCP to cpSRP (○), cpSRP-R161A (□) and cpSRP-Y204A (◇). Fits of data to eq 3 gave  $K_d$  values of 128 nM for wild-type cpSRP, 1.2  $\mu\text{M}$  for cpSRP-R161A, and  $>5 \mu\text{M}$  for cpSRP-Y204A.



**Figure 3.6** Chromodomains are essential for cpSRP43's chaperone activity. (A) Domain composition of cpSRP43. CD denotes the chromodomain, and A1–A4 denotes ankyrin repeats 1–4. (B–C) The amount of soluble LHCP (1  $\mu M$ ) at equilibrium in the presence of 1  $\mu M$  ankyrin (B) or chromodomain (C) deletion mutants of cpSRP43. (D) The amount of LHCP re-solubilized by 2 (white), 4 (gray), and 8  $\mu M$  cpSRP43,  $\Delta CD2$ , or  $\Delta CD3$  at equilibrium, obtained from the time courses in Supplementary Figure 3.S5. (E) Binding of LHCP to cpSRP43 (black;  $K_d = 296$  nM),  $\Delta CD1$  (blue;  $K_d > 10$   $\mu M$ ),  $\Delta CD2$  (green;  $K_d = 530$  nM),  $\Delta CD3$  (gold;  $K_d = 886$  nM), and  $\Delta CD2 \Delta CD3$  (red;  $K_d > 10$   $\mu M$ ).  $K_d$  values were from fits of data to eq 3.



**Figure 3.7** cpSRP43 or LHCP mutants defective in chaperone activity could not support LHCP targeting and translocation. The arrow marks full-length LHCP (TP, translation product); the single and double asterisks mark the protected fragments of LHCP that represent products of proper LHCP integration into the thylakoid membrane. The integration efficiency was quantified relative to that of wild-type cpSRP43 and is shown at the bottom.



**Figure 3.8** SAXS reconstruction of full-length cpSRP43 reveals an elongated shape. (A) Experimental (blue) and theoretical (red) SAXS profiles of cpSRP43 in Kratky's representation.  $s$  denotes momentum transfer, and  $I$  denotes scattering intensity in arbitrary units. The theoretical curve was calculated from the dummy atom model in B. (B) Dummy atom model of full-length cpSRP43, reconstructed from the SAXS profile as described in Methods. (C) Molecular models generated from rigid-body docking of the structures of individual cpSRP43 fragments into the SAXS reconstructed shape in B. CD1–Ank4 (PDB ID: 3DEO)(19) is in cyan. CD2 (PDB ID: 1X3Q)(42) is in pink. CD3 (PDB ID: 1X3P)(42) is in green. Note that multiple orientations of CD2 and CD3 are possible and cannot be resolved at this resolution; two possible orientations of CD2 are shown here.

## Supplementary Materials

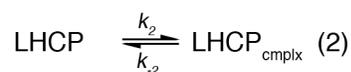
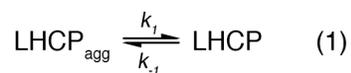
### Supplementary Note

#### *Optimization of conditions for reconstituting the LHCP-cpSRP interaction.*

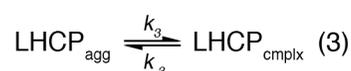
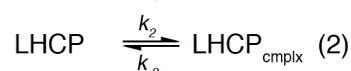
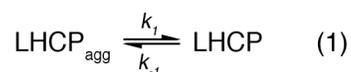
Previous work suggested that both the cpSRP43 and cpSRP54 subunits are necessary for the formation of a soluble transit complex with LHCP (12). Our results here strongly suggested that cpSRP43 is sufficient to bind and solubilize LHCP. The disparity between this and the previous studies could stem, in part, from the differences in experimental conditions. We have found that the activity and oligomeric state of cpSRP43 is sensitive to ionic strength. In buffers with low ionic strength (< 100 mM NaCl), which were typically used in previous studies, cpSRP43 exists as higher molecular weight complexes (Supplementary Figure 3.S1A). Under these low-salt conditions, cpSRP43 was sub-optimal for interacting with LHCP (Supplementary Figure 3.S1B and C). Presumably, the presence of cpSRP54 shifted the conformational equilibrium of cpSRP43 toward the more active monomeric state, and this likely contributed to the apparent requirement of cpSRP54 for interacting with LHCP under the previous assay conditions.

#### *Kinetic simulations.*

The Berkeley Madonna software was used to perform kinetic simulations for the different models of LHCP disaggregation (Figure 3.4C and D). For the passive model in Figure 3.4C, the following reactions were modeled (eq 1–2):



For the active mechanism in Figure 3.4D, the following reactions were modeled (eq 1–3):



in which  $\text{LHCP}_{\text{agg}}$  denotes the aggregated LHCP molecules, LHCP denotes a soluble LHCP monomer,  $\text{LHCP}_{\text{cmplx}}$  denotes the cpSRP43-LHCP complex, and the rate constants are defined in Figure 3.3C and D. As this software considers only first-order or pseudo-first-order reactions, the concentration of cpSRP43 was varied by varying the values of  $k_2$  and  $k_3$  during the simulation. For the aggregation reaction (eq 1), the rate constant for LHCP aggregation ( $k_{-1}$ ) was estimated to be  $0.05 \text{ s}^{-1}$ , based on the observation that LHCP aggregation is 70–80% complete within the first 15–20 seconds of mixing (Figure 3.1C, 3.4E, and Supplementary Figures 3.S3 and 3.S5). As the light scattering from aggregates was linear with LHCP concentration starting from  $0.5 \text{ }\mu\text{M}$  (Figure 3.1C) indicating complete aggregate formation above this concentration, we assumed that the equilibrium favors aggregation by 100-fold under the experimental conditions. This gave the apparent rate of spontaneous disaggregation ( $k_1$ ) of  $0.0005 \text{ s}^{-1}$ ; changing the value of  $k_1$  from  $0.00005$  to  $0.05 \text{ s}^{-1}$  altered the time courses, but did not affect the conclusion that

the disaggregation rates are independent of cpSRP43 concentrations in the passive mechanism (Supplementary Figure 3.S2).

For the binding reaction between LHCP monomer and cpSRP43 (eq 2), we assumed an association rate constant ( $k_{on}$ ) of  $1 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ , which is typical for bimolecular association between proteins. Varying the values of  $k_{on}$  from  $10^6$  to  $10^8 \text{ M}^{-1}\text{s}^{-1}$  had no effect on the result of simulation (not shown). The value of  $k_{on}$  and the experimentally determined  $K_d$  value of  $\sim 100 \text{ nM}$  (Figure 3.2A) were used to calculate the dissociation rate constant ( $k_{-2}$ ). The apparent associating rate constants,  $k_2$ , were calculated as  $k_{on} \times [\text{cpSRP43}]$ .

For the active disaggregation reaction (eq 3), we allowed the simulation program to fit the  $k_3$  and  $k_{-3}$  values using experimental data. The values obtained from the simulation are comparable to the rate constants of disaggregation obtained from manual fitting of data (Supplementary Table 3.S1; see the next section).

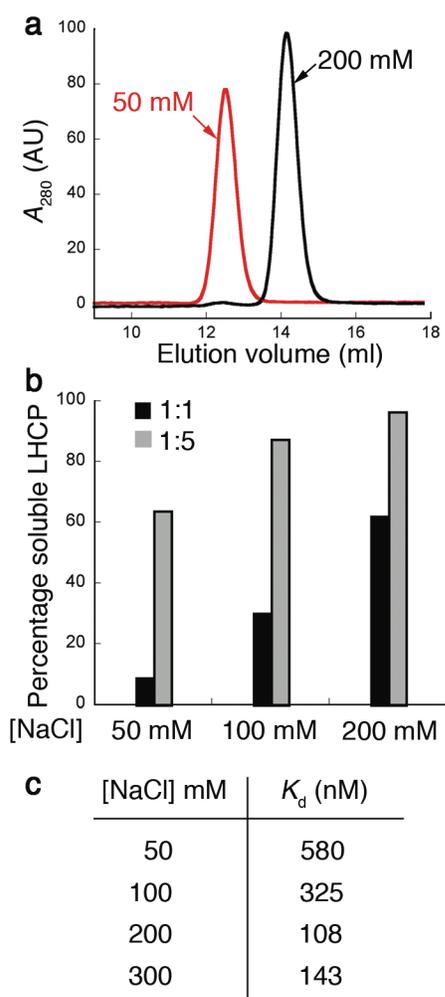
### ***Analysis of rate constants for protein disaggregation.***

The rate constants of protein disaggregation were obtained using several independent approaches. (i) The time courses of LHCP disaggregation in Figure 3.4E were fit to eq 1 in Methods to obtain the observed rate constants to reach equilibrium ( $k_{\text{obsd}}$ ), which is the sum of forward disaggregation and reverse re-aggregation processes ( $k_{\text{obsd}} = k_f + k_r$ ), and the equilibrium of each disaggregation reaction, which is determined by the relative magnitude of the disaggregation and re-aggregation processes [ $K_{\text{obsd}} = k_f/(k_f + k_r)$ ]. Using these relationships, we calculated the net rate constants for the disaggregation process ( $k_f$ ). (ii) In cases where a substantial amount of disaggregation

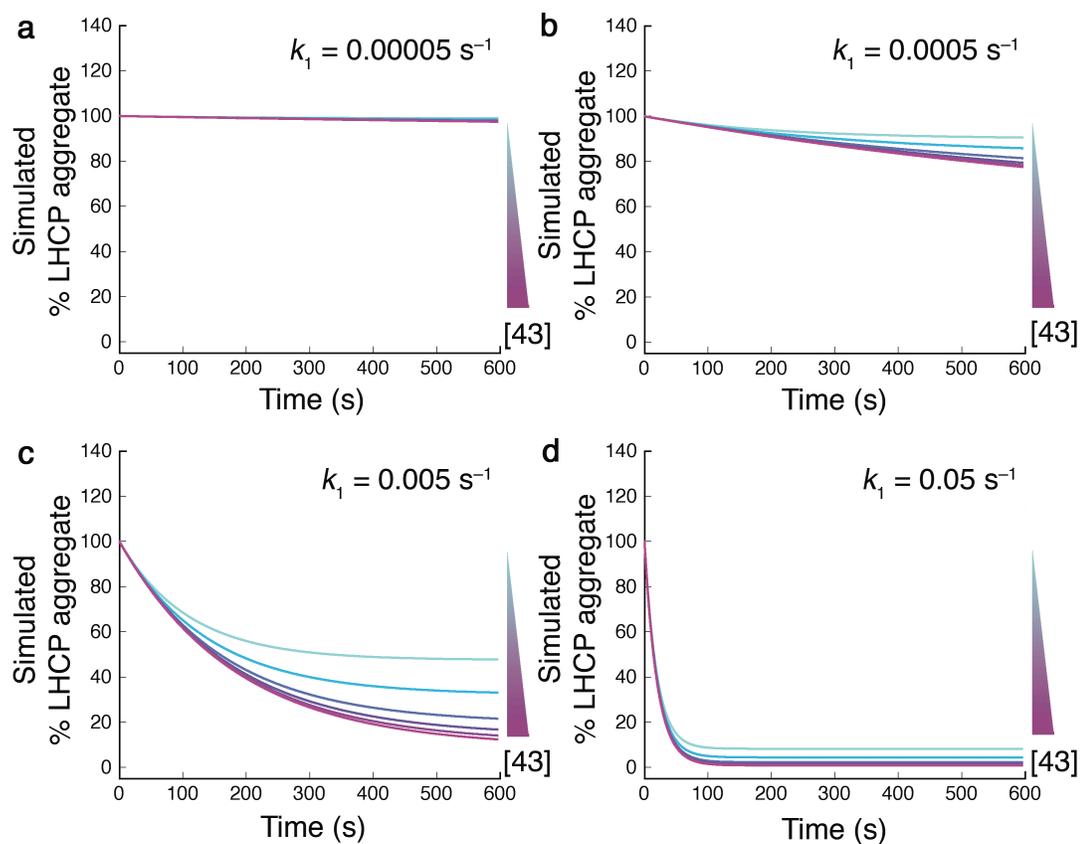
was achieved, the rate constants for the forward disaggregation reactions can also be estimated from the initial rates, since at earlier times the contribution from the reverse re-aggregation process is negligible. (iii) The Berkeley-Madonna program was used to fit the data in Figure 3.4E to the active remodeling mechanism in Figure 3.4D. The rate constants obtained from these different approaches were the same, within experimental error (Supplementary Table 3.S1).

**Supplementary Table 3.S1** Comparison of disaggregation rate constants obtained from fits of data and simulation

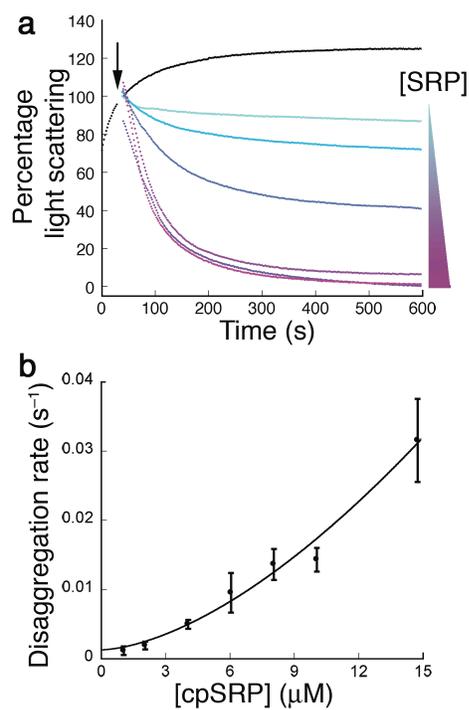
[cpSRP43] ( $\mu\text{M}$ )	$k_f$ from manual fits of data ( $\text{s}^{-1}$ )	$k_3$ from simulation ( $\text{s}^{-1}$ )
2	0.0018	0.0023
4	0.0034	0.0041
6	0.0036	0.0050
8	0.0084	0.0084
10	0.0187	0.0180



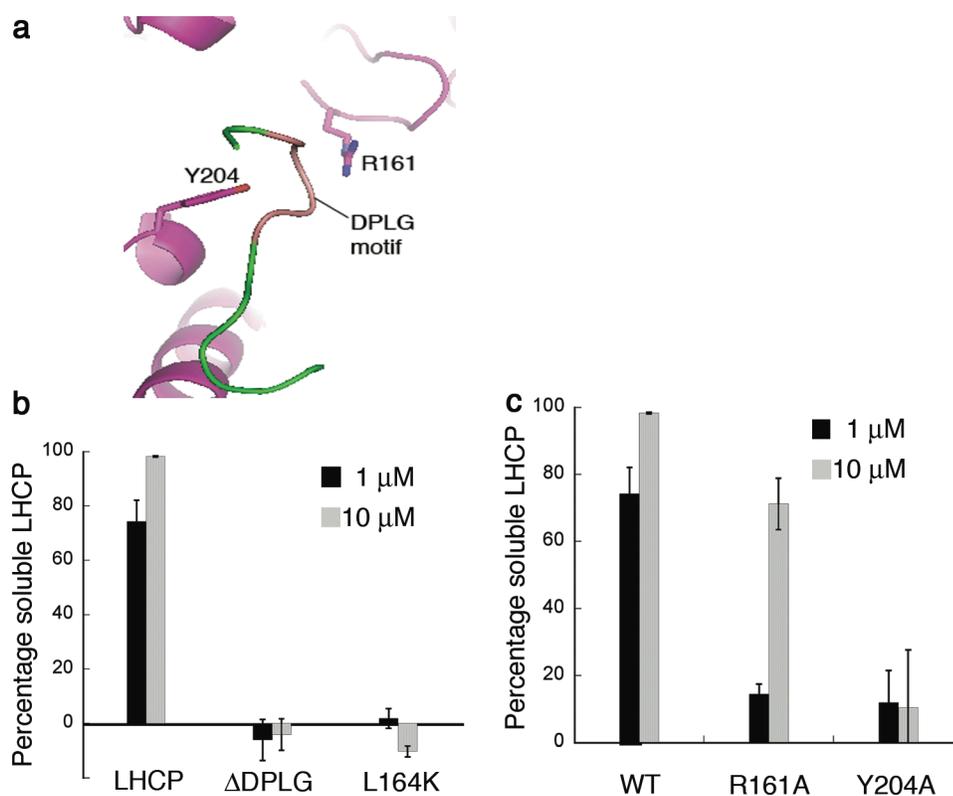
**Supplementary Figure 3.S1** cpSRP43 is active at higher ionic strength, but is present as inactive multimeric forms at lower ionic strength. (A) cpSRP43 runs as a monomer on Superdex 200 in buffer containing 200 mM NaCl (black trace), but exhibits aberrant mobility in buffer containing 50 mM NaCl (red trace). (B) The amount of solubilized LHCP (1  $\mu$ M) at equilibrium in the presence of 1  $\mu$ M (black) or 5  $\mu$ M (gray) cpSRP43 in buffer with different salt concentrations. (C) List of the  $K_d$  values for the LHCP-cpSRP43 interaction, determined by fluorescence anisotropy, at different salt concentrations.



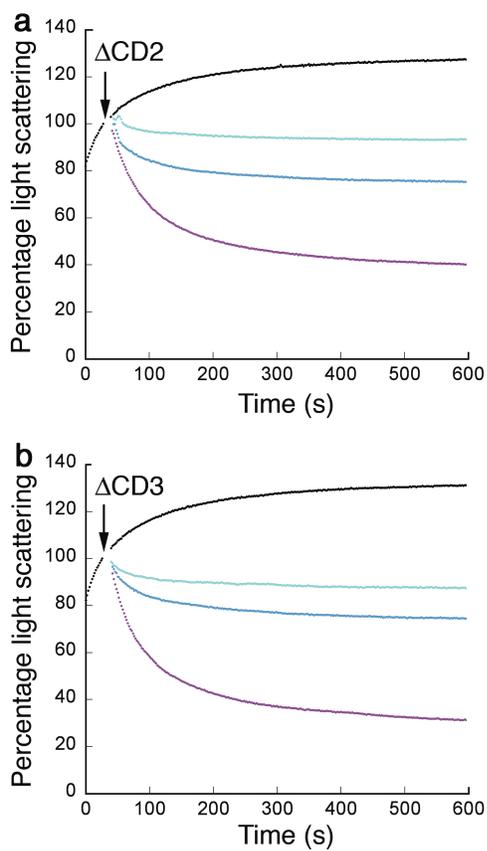
**Supplementary Figure 3.S2** Kinetic simulations based on the passive mechanism in Figure 3.4C, which showed that the rates of disaggregation are independent of cpSRP43 concentration when the rates of spontaneous disaggregation ( $k_1$ ) were varied over the range of  $0.00005 - 0.05 \text{ s}^{-1}$ .



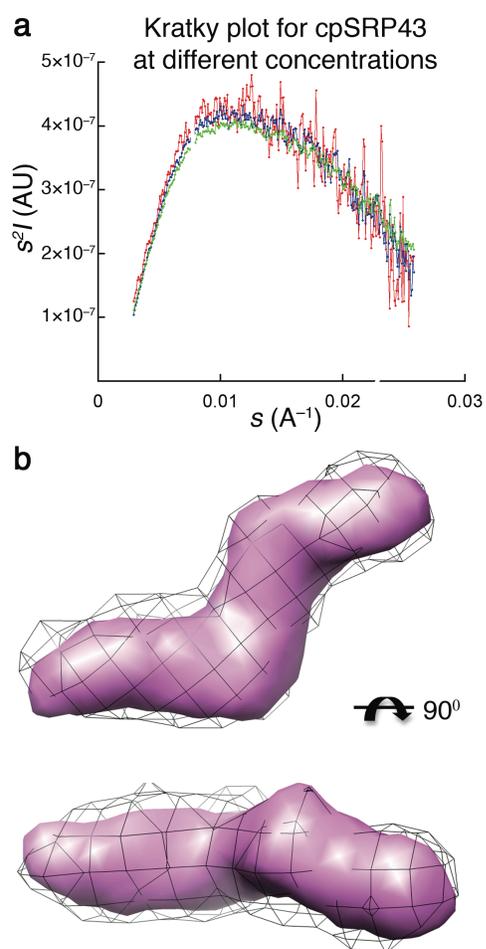
**Supplementary Figure 3.S3** cpSRP actively reversed LHCP aggregation. (A) Time courses for disaggregation of LHCP (1  $\mu M$ ) at varying concentrations (2–10  $\mu M$ ) of cpSRP. The black arrow marks the time of cpSRP addition. (B) Concentration dependence of the forward rate constants of disaggregation reactions ( $k_f$ ; see text for derivation). The fits to eq 3 in Methods gave a Hill coefficient of 1.7.



**Supplementary Figure 3.S4** Specific binding interactions between the L18 motif of LHCP and cpSRP are essential for chaperone activity. (A) The crystal structure of the L18 peptide (green) bound to the CD1–Ank4 fragment of cpSRP43 (magenta)(19). The DPLG motif is highlighted in pink. The two residues from cpSRP43 that make important contacts to this motif, R161 and Y204, are shown in sticks. (B) The amount of soluble LHCP (1  $\mu\text{M}$ ) or its mutants,  $\Delta\text{DPLG}$  and L164K, at equilibrium in the presence of 1  $\mu\text{M}$  (black) or 10  $\mu\text{M}$  (gray) cpSRP. (C) The amount of soluble LHCP (1  $\mu\text{M}$ ) at equilibrium in the presence of 1  $\mu\text{M}$  (black) or 10  $\mu\text{M}$  (gray) cpSRP or its mutants, cpSRP-R161A and cpSRP-Y204A.



**Supplementary Figure 3.S5** cpSRP43 deletion mutants,  $\Delta CD2$  and  $\Delta CD3$ , can efficiently reverse LHCP aggregation. Time courses for disaggregation reactions by 2  $\mu M$  (cyan), 4  $\mu M$  (blue), and 8  $\mu M$  (magenta) of cpSRP43  $\Delta CD2$  (A) and cpSRP43  $\Delta CD3$  (B).



**Supplementary Figure 3.S6** SAXS analysis of full-length cpSRP43. (A) Experimental SAXS profiles in Kratky's representation at cpSRP43 concentrations of 50  $\mu\text{M}$  (red), 100  $\mu\text{M}$  (blue), and 200  $\mu\text{M}$  (green). The agreement of the three curves indicated that no aggregation or interparticle interference occurred during the experiment. (B) Comparison of surface maps calculated from the dummy atom model in Figure 3.8B (pink shell) and from the docking model in Figure 3.8C (gray mesh).

**References:**

1. Balch, W. E., Morimoto, R.I., Dillin, A., and Kelly, J.F. (2008) Adapting proteostasis for disease intervention, *Science* 319, 916–919.
2. Hayer-Hartl, M., and Hartl, F.U. (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. Doyle, S. M., Hoskins, J.R., and Wickner, S. (2007) Collaboration between the ClpB AAA+ remodeling protein and the DnaK chaperone system, *Proc. Natl. Acad. Sci. USA* 104, 11138–11144.
7. 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.
8. Walton, T. A., Sandoval, C.M., Fowler, C.A., Pardi, A., and Sousa, M.C. (2009) The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains, *Proc. Natl. Acad. Sci. USA* 106, 1772–1777.
9. Stefanovic, S., and Hegde, R.S. (2007) Identification of a targeting factor for posttranslational membrane protein insertion into the ER, *Cell* 128, 1147–1159.
10. Mihara, K., and Omura, T. (1996) Cytoplasmic chaperones in precursor targeting to mitochondria: the role of MSF and hsp70, *Trends Cell Biol.* 6, 104–108.
11. Schuenemann, D. (2004) Structure and function of the chloroplast signal recognition particle, *Curr. Genet.* 44, 295–304.
12. Schuenemann, 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. USA* 95, 10312–10316.
13. 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.
14. 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.
15. 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, *The Plant Cell* 11, 87–99.
16. Eichacker, L. A., and Henry, R. (2001) Function of a chloroplast SRP in thylakoid protein export, *Biochim. Biophys. Acta* 1541, 120–134.

17. Jonas-Straube, E., Hutin, C., Hoffman, N.E., and Schuenemann D. (2001) Functional analysis of the protein-interacting domains of chloroplast SRP43, *J. Biol. Chem.* 276, 24654–24660.
18. Goforth, R. L., Peterson, E.C., Yuan, J., Moore, M.J., Kight, A.D., Lohse, M.B., Sakon, J., and Henry, R.L. (2004) Regulation of the GTPase cycle in post-translational signal recognition particle-based protein targeting involves cpSRP43, *J. Biol. Chem.* 279, 43077–43084.
19. 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.
20. Hermkes, R., Funke, S., Richter, C., Kuhlmann, J., and Schünemann, D. (2006) The  $\alpha$ -helix of the second chromodomain of the 43 kDa subunit of the chloroplast signal recognition particle facilitates binding to the 54 kDa subunit, *FEBS Lett.* 580, 3107–3111.
21. 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.
22. Cammarata, K., and Schmidt, G.W. (1992) In vitro reconstitution of a light-harvesting gene product: deletion mutagenesis and analyses of pigment binding, *Biochemistry* 31, 2779–2789.
23. Paulsen, H., Rumler, U., and Rudiger, W. (1990) Reconstitution of pigment-containing complexes from light-harvesting chlorophyll a/b-binding protein overexpressed in *Escherichia coli*, *Planta* 181, 204–211.
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.* 97, 1926–1931.
25. 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.
26. Jaru-Ampornpan, P., Chandrasekar, S., 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. Yuan, J., Kight, A., Goforth, R.L., Moore, M., Peterson, E.C., Sakons, J., and Henry, R. (2002) ATP stimulates signal recognition particle (SRP)/FtsY-supported protein integration in chloroplasts, *J. Biol. Chem.* 277, 32400–32404.
28. Svergun, D. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing, *Biophys. J.* 76, 2879–2886.
29. Svergun, D. I., Petoukhov, M.V. and Koch, M.H.J. . (2001) Determination of domain structure of proteins from X-ray solution scattering, *Biophys. J.* 80, 2946–2953.
30. Kozin, M., and Svergun, D.I. (2001) Automated matching of high and low resolution structural models, *J. Appl. Crystallogr.* 34, 33–41.

31. Volkov, V., and Svergun, D.I. (2003) Uniqueness of ab initio shape determination in small-angle scattering, *J. Appl. Crystallogr.* *36*, 860–864.
32. Wriggers, W., Milligan, R. A., and McCammon, J. A. (1999) Situs: a package for docking crystal structures into low-resolution maps from electron microscopy, *J. Struct. Biol.* *125*, 185–195.
33. Wriggers, W., and Chacon, P. (2001) Using Situs for the registration of protein structures with low-resolution bead models from X-ray solution scattering, *J. Appl. Crystallogr.* *34*, 773–776.
34. Pettersen, E., Goddard, T., Huang, C., Couch, G., Greenblatt, M., Meng, E., and Ferrin, E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis, *J. Comput. Chem.* *13*, 1605–1612.
35. Oreb, M., Tews, I., and Schleiff, E. (2008) Policing Tic 'n' Toc, the doorway to chloroplasts, *Trends Cell Biol.* *18*, 19–27.
36. Li, X., Henry, R., Yuan, J., Cline, K., and Hoffman, N.E. (1995) A chloroplast homologue of the signal recognition particle subunit SRP54 is involved in the posttranslational integration of a protein into thylakoid membranes, *Proc. Natl. Acad. Sci. USA* *92*, 3789–3793.
37. Jansson, S. (1999) A guide to the Lhc genes and their relatives in arabidopsis, *Trends Plant Sci.* *4*, 236–240.
38. Barkow, S. R., Levchenko, I., Baker, T.A., and Sauer, R.T. (2009) Polypeptide translocation by the AAA+ ClpXP protease machine, *Chem Biol.* *16*, 605–612.
39. Tomoyasu, T., Mogk, A., Langen, H., Goloubinoff, P., and Bukau, B. (2001) Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the *Escherichia coli* cytosol, *Mol. Microbiol.* *40*, 397–413.
40. Mogk, A., Dougan, D., Weibezahn, J., Schlieker, C., Turgay, K., and Bukau, B. (2004) Broad yet high substrate specificity: the challenge of AAA+ proteins, *J. Struct. Biol.* *146*, 90–98.
41. Ali, M., Lipfert, J., Seifert, S., Herschlag, D., and Doniach, S. (2010) The ligand-free state of the TPP riboswitch: a partially folded RNA structure, *J. Mol. Biol.* *396*, 153–165.
42. Sivaraja, V., Kumar, T.K., Leena, P.S., Chang, A.N., Vidya, C., Goforth, R.L., Rajalingam, D., Arvind, K., Ye, J.L., Chou, J., Henry, R., and Yu, C. (2006) Three-dimensional solution structures of the chromodomains of cpSRP43, *J Biol Chem.* *280*, 41465–41471.
43. Goloubinoff, P., Mogk, A., Ben Zvi, A.P., Tomoyasu, T., Bukau, B. (1999) Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network, *Proc. Natl. Acad. Sci. USA* *96*, 12732–12737.
44. Mogk, A., Schlieker, C., Friedrich, K.L., Schoenfeld, H.-J., Vierling, E., and Bukau, B. (2003) Refolding of substrates bound to small Hsps relies on a disaggregation reaction mediated most efficiently by ClpB/DnaK, *J. Biol. Chem.* *278*, 31033–31042.
45. 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.

46. Schlothauer, T., Mogk, A., Dougan, D.A., Bukau, B., and Turgay, K. (2003) MecA, an adaptor protein necessary for ClpC chaperone activity, *Proc. Natl. Acad. Sci. USA* 100, 2306–2311.
47. Panse, V. G., Vogel, P., Trommer, W.E., and Varadarajan, R. (2000) A thermodynamic coupling mechanism for the disaggregation of a model peptide substrate by chaperone SecB, *J. Biol. Chem.* 275, 18698–18703.
48. 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.
49. Komiya, T., Hachiya, N., Sakaguchi, M., Omura, T., Mihara, K. (1994) Recognition of mitochondria-targeting signals by a cytosolic import stimulation factor, MSF, *J Biol Chem.* 269, 30893–30897.

**Chapter 4:**  
**Biochemical and Biophysical**  
**Characterization of the LHC Protein**  
**Aggregates**

**Abstract**

Protein aggregation is detrimental to cells. To overcome this problem, cells have evolved specialized chaperone systems to reverse existing protein aggregates; these “disaggregases” are often ATP-dependent macromolecular machines. Recently, we found that the 43-kDa subunit of chloroplast Signal Recognition Particle (cpSRP43) can efficiently reverse the aggregation of its substrate proteins, the light-harvesting chlorophyll a/b-binding (LHC) proteins, in the absence of external energy input. To understand the molecular mechanism of this novel disaggregase activity, it is imperative to first know the nature of the protein aggregates handled by cpSRP43. Using biophysical and biochemical methods, we found that LHC proteins form disc-shaped aggregates  $\sim 12$  nm in diameter and 1–2 nm in height. They contain exposed hydrophobic grooves that could be probed by a variety of small fluorescent dyes. LHC protein aggregates are kinetically stable and resistant to detergents, but can be resolubilized by strong chemical denaturants. These observations provide important clues to understanding the capability and specificity of cpSRP43’s disaggregase activity.

## Introduction

The proper folding of proteins into their three-dimensional structures is essential for their function. *In vivo*, many factors pose challenges to protein folding, including environmental stress and molecular crowding (1). Within this setting, improper intra- or intermolecular interactions can occur and lead to the aggregation of proteins. Protein aggregation is detrimental to cells as it deprives the cells of functional proteins. Moreover, some aggregates, most notoriously those that form amyloid fibrils, are toxic to cells and cause numerous protein-folding diseases (2).

Despite early suggestions that protein aggregates are amorphous (3), emerging evidence has argued for some degrees of order and specificity in aggregate formation (4). For instance, folding intermediates of bovine growth hormone, phosphoglycerate kinase, P22 tailspike, and coat proteins have been shown to participate in specific intermolecular interactions in their aggregation pathways (5, 6). The extreme case of ordered protein aggregates is the amyloid fibrils, which form extensive cross- $\beta$  strand structures (7). Many techniques have been developed to map out the structures of protein aggregates, especially the various intermediates during amyloid fibril formation. For example, transmission electron microscopy (TEM) and high-resolution atomic force microscopy (AFM) reveal global size, shape, and morphology of the aggregates (8, 9). Scanning transmission electron microscopy (STEM) can provide an estimate of the mass and/or density of the aggregates (10). Fluorescent small molecule dyes probe the presence of  $\beta$ -sheet structures or exposed hydrophobic patches in the aggregates (11, 12). Because of their unconventional order, the amyloid fibrils can be crystallized, and their X-ray crystal structures provide unprecedented resolution in visualization of protein aggregates (7, 13).

These powerful techniques reveal that protein aggregates can form highly-ordered structures that are thermodynamically and kinetically stable and contain specific intra- or intermolecular interactions that are not trivial to break.

Nevertheless, cells have evolved ways to handle protein aggregates. A specialized class of molecular chaperones, the disaggregases, can perform an energetically uphill process of reversing protein aggregation. The most well-characterized disaggregases belong to the Clp/Hsp100 family of AAA+ ATPases (ATPases associated with various cellular activities), such as ClpB in prokaryotes and Hsp104 in yeasts (*14*). Both are large hexameric rings that are powered by mechanical forces from ATP hydrolysis, and both require additional co-chaperones to efficiently dismantle protein aggregates (*15, 16*). These disaggregases are capable of reversing the aggregation of many model proteins. The complexity of these disaggregase systems and the promiscuity in their substrate selection have made it difficult to pinpoint their molecular mechanisms of action.

Recently, we found an efficient disaggregase activity in the 43-kDa subunit of the chloroplast Signal Recognition Particle (cpSRP43) (*17*). Unlike the Clp/Hsp100 family of disaggregases, cpSRP43 is a highly specific chaperone/disaggregase for its substrates, the light-harvesting chlorophyll a/b-binding (LHC) proteins (*18*). LHC proteins are synthesized in the cytosol and are likely unfolded when they enter the chloroplast, where they are recognized and delivered to the thylakoid membrane by the cpSRP (*18*). LHC proteins are highly hydrophobic proteins containing three transmembrane (TM) helices (*19*). A well-conserved 18-amino acid loop between the second and the third TM helices, termed L18, provides a specific recognition element for high affinity binding by

cpSRP43 (17, 20, 21). In previous work we showed that the specific interaction of cpSRP43 with the L18 motif is crucial for the disaggregation activity of cpSRP43. This and other observations led us to propose that, in the absence of external energy input, cpSRP43 uses specific binding energy with its substrate protein to drive its disaggregase activity (17).

To understand the molecular mechanism that underlies cpSRP43's novel disaggregase activity, we need to understand the nature of the LHC protein aggregate that cpSRP43 recognizes and subsequently remodels. What does it look like? What is presented outside of the aggregate to enable its recognition by cpSRP43? How strong are the internal packing interactions of the aggregate, which need to be overcome by the disaggregase? This chapter aims to answer some of these questions via several biochemical and biophysical methods. We found that the LHCP aggregates are disc-like particles that are ~ 12 nm wide and 1–2 nm thick. They have exposed hydrophobic patches that could bind to environmentally sensitive fluorescent dyes. Additionally, we investigated the stability of the aggregates using various detergents and chemical denaturants.

## Results

### *LHCP forms stable, micellar aggregates.*

Previously, we developed an assay for the formation of large LHCP aggregates based on light scattering (17). At LHCP concentrations above  $\sim 100$  nM, there was a linear relationship between the light scattering intensity and the concentration of LHCP (Figure 4.1A, black), suggesting that aggregation was complete under these conditions. However, the linearity broke down at LHCP concentrations below 100 nM (Figure 4.1A and inset, black). This was not because of limitations in instrument sensitivity; when pre-formed LHCP aggregates were serially diluted, linearity in light scattering intensity was observed throughout the entire concentration range and extrapolated through the zero point (Figure 4.1A inset, red). These observations showed that: (i) the LHCP aggregate is kinetically stable and virtually irreversible once it has formed; and (ii) formation of the LHCP aggregate involves a lag phase below a critical protein concentration, which is reminiscent of the critical micellar concentration (CMC) during the formation of micelles. Based on this information, we estimated the CMC values for aggregate formation by LHCP and a number of LHCP variants as the X-intercepts from linear fits to data (Figure 4.1A and B; (22)). These analyses yielded CMC values of 100–200 nM for the different LHCP aggregates (Figure 4.1B).

### *LHCPs form disc-shaped aggregates with diameters of $\sim 12$ nm.*

To investigate the global structure of the aggregate formed by LHCP, we employed TEM and AFM. The negatively stained TEM images revealed LHCP aggregates to be circular particles with diameters of  $12 \pm 2$  nm (Figure 4.2). Consistent

with the EM images, AFM analysis also showed LHCP aggregates to be disc-shaped particles (Figure 4.3A). Analysis of the area of these particles resulted in a distribution that could be fit with a Gaussian function, giving a peak at 214 nm<sup>2</sup> (Figure 4.3B). Assuming that LHCP aggregates were circular in shape, the diameter estimated from the mean area was 16±5 nm, in good agreement with the EM measurements. The heights of the LHCP aggregates measured by AFM were in the range of 1–2 nm (Figures 4.3C and D).

***LHCP aggregates contain exposed hydrophobic grooves.***

Many protein aggregates contain exposed hydrophobic microdomains that could be probed by small-molecule fluorescent dyes such as 1-anilino-8-naphthalene sulfonate (ANS) and bis-ANS, as the fluorescence intensity of these dyes increases upon exposure to a non-polar environment (12). We asked whether the LHCP aggregates share this feature with other aggregates. Indeed, the fluorescence of both ANS (not shown) and bis-ANS (Figure 4.4A) increased significantly upon the addition of 1µM LHCP aggregate, accompanied by a blue shift of the fluorescence emission spectra (Figure 4.4A). These results strongly suggested that the LHCP aggregate contains exposed hydrophobic microdomains that allow the binding of these dyes, consistent with the highly hydrophobic nature of this protein.

We also used Thioflavin T (ThT) to probe the structural organization of the LHCP aggregate. ThT is an environmentally sensitive fluorescent dye that is often used as a diagnostic for the formation of amyloid fibrils generated by amyloid-β (Aβ), α-synuclein, and other amyloidogenic proteins (11). Similar to bis-ANS, the fluorescence of ThT also

exhibited a significant increase in intensity and a blue shift in its spectrum in the presence of the LHCP aggregate (Figure 4.4B, blue lines). The extent of these fluorescence changes are comparable to that induced by mature amyloid fibrils generated from the amyloid- $\beta$  peptide (Figure 4.4B, red vs. blue, and Figure 4.4C). As both the TEM and AFM analyses did not indicate amyloid formation in the LHCP aggregate (see above), these results support suggestions from recent work that ThT is not highly specific for amyloid fibrils. 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 (23).

***LHCP aggregates are detergent-resistant but can be solubilized by strong chemical denaturants.***

In order to reverse aggregation, cpSRP43 must overcome the packing interactions that stabilize the LHCP aggregate. To probe the stability of this aggregate, we tested whether it could be solubilized by various detergents, including n-Dodecyl-N,N,-Dimethylamine-N-Oxide (LDAO), n-Dodecyl- $\beta$ -D-Maltopyranoside (DDM), Triton X-100 (TX-100), n-Nonyl- $\beta$ -D-Glucopyranoside (BNG), and n-Octyl- $\beta$ -D-Glucopyranoside ( $\beta$ -OG). The results of the sedimentation assay (see Methods) showed that none of these detergents were able to solubilize the LHCP aggregate at or above the concentrations used for membrane protein solubilization (Figure 4.5A). Sodium dodecyl sulfate (SDS) could completely solubilize the LHCP aggregate, as LHCP is present in the soluble fraction after incubation with SDS for 30 minutes (Figure 4.5B, left panel).

We also tested SDS solubility of the LHCP aggregate using an established protocol for amyloid fibrils (24). Unlike the sedimentation assay in which centrifugation

is used to separate pellet and soluble fractions after incubation with the detergents, this assay judges solubility of the aggregate by the mobility of the proteins into the separating gel after incubation with the SDS-containing buffer at room temperature ((24); see Methods). The solubilized amyloid fibrils can migrate into the separating gel without boiling, whereas the aggregates cannot enter into the separating gel unless boiled (24). Based on this standard, LHCP aggregate showed partial resistance to 2% SDS as only 24% of the aggregates could migrate into the gel without boiling (Figure 4.5B).

Inspired by the insights gained from quantitative analyses of protein unfolding by chemical denaturants, we also probed the stability of the LHCP aggregate by analyzing its solubility in chemical denaturants. Using the sedimentation assay, we showed that both guanidinium hydrochloride (GdmHCl) and urea could effectively solubilize the LHCP aggregates in a concentration-dependent manner (Figure 4.6A). Quantification of the amount of solubilized LHCP as a function of urea concentration gave an aggregate solubilization curve analogous to the protein denaturation curves (Figure 4.6B). 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 ( $U_{50}$ ; see Methods). These parameters provide quantitative empirical measures of the energetics of the internal packing interactions that stabilize the LHCP aggregate. In summary, the results in this section demonstrate that LHCP forms stable aggregates that are resistant to various detergents, but could be dissolved by strong chemical denaturants such as SDS, GdmHCl, and urea.

## Discussion

To understand the molecular mechanism of cpSRP43's disaggregase activity, it is essential to understand the nature of its substrate, the LHCP aggregate. Many techniques have been employed to characterize the LHCP aggregates. We found that LHCP forms disc-shaped aggregates ~ 12 nm in diameter and 1–2 nm in thickness. The aggregates contain exposed hydrophobic surfaces as probed by several environmentally sensitive small molecule dyes. Further, analyses of aggregate formation and solubilization suggest that LHCP forms micelle-like aggregates that are thermodynamically and kinetically stable.

It is interesting to note that the morphology of the LHCP aggregates bears some resemblance to the soluble oligomeric intermediates that precede the amyloid fibril formation, which are often disc-shaped, 9–25 nm in diameter and 2–3 nm in height (8, 9). Like the mature amyloid fibrils, these intermediate aggregates have been suggested to cause cytotoxicity that leads to diseases (25–27). Although this type of protein aggregates has been termed “amorphous”, only small heterogeneity was observed in TEM and AFM (see Figure 4.2 and 4.3), arguing against complete disorder in the structure of these aggregates. Nevertheless, more experiments are required to probe if there is a specific pattern in structural organization of the LHCP aggregate.

How stable are the LHCP aggregates? This is an important question because it directly relates to the amount of energy cpSRP43 has to overcome during disaggregation, and hence provides a measure of the capacity of this chaperone as a disaggregase. Some protein aggregates, such as protein precipitates produced when the protein concentration exceeds the solubility limit (e.g., “salting out” effect) or the “soluble aggregates” formed

by oligomeric protein complexes, can be readily re-dissolved by dilution in fresh buffer and retain native conformations (28). This is not the case for highly stable amyloid fibrils, which are detergent-insoluble and could only be dissolved in strong chemical denaturants (29). The results here indicate that LHCP aggregates are stable both kinetically and thermodynamically. First, extensive dilution of the aggregate does not lead to re-solubilization (Figure 4.1A, red), suggesting that LHCP aggregates, once formed, are kinetically stable. Second, LHCP aggregates are resistant to a variety of detergents, even up to 2% SDS, and are solubilized only by strong chemical denaturants. Finally, both the protein concentration dependence of aggregate formation and the urea concentration dependence of re-solubilization of this aggregate are cooperative, suggesting cooperativity in the interactions that stabilize the aggregate. Importantly, quantitative analyses based on the urea solubilization curves provide an empirical method to assess the stability of the aggregates. This will allow us to systematically correlate perturbations in the stability of the aggregate with changes in the efficiency of the disaggregation reaction and can teach us about the roles that these internal packing interactions play in the disaggregation process mediated by cpSRP43.

Many questions remain to be addressed regarding the nature and structure of the LHCP aggregate. For instance, how many LHCP monomers reside in each aggregate particle? This question could be explored by STEM, in which the electron density from unstained samples is compared to that from the standard specimen with known oligomeric composition, and the mass-per-area can be calculated based on the electron density (10). This technique has been successful in calculating the number of A $\beta$  peptides per length in the amyloid fibrils (8, 30). Another interesting question is whether

any secondary structures are present in the aggregate, and if so, what these structures are. The strong staining of ThT suggests the possibility that the aggregate could contain  $\beta$ -sheet rich areas since ThT prefers binding to cross- $\beta$  strand structures (11). However, the native LHCP is a helical protein (19). This question can be explored further by Fourier-transformed infrared spectroscopy (FT-IR) or 2D-IR (31).

Yet the key question that is central to understanding the mechanism of cpSRP43's disaggregase activity is: what features of the LHCP aggregate allow cpSRP43 to recognize it and initiate the disaggregation process? Which parts of the LHCP are presented outside the aggregate, and which parts are buried inside? Although additional experimental evidence is required, the results here strongly suggest that the formation of LHCP aggregates involves hydrophobic collapse driven by the sequestering of its three TM helices. As a result, they might leave their more hydrophilic segments exposed to solvent, similar to the formation of detergent micelles. These parts would likely include the loops in between the TMs, especially the L18 motif. Indeed, preliminary results from chemical modification and electron paramagnetic resonance experiments have indicated that the L18 motif is solvent-exposed when LHCP forms aggregates (T.X. Nguyen, V. Q. Lam, unpublished data). As the L18 is the primary recognition element for cpSRP43, its exposure on the exterior of the aggregate provides a very attractive mechanism for how cpSRP43 could recognize and anchor onto the aggregate to start the disassembly process.

In summary, characterization of the LHCP aggregate is vital to understanding the molecular mechanism by which cpSRP43 reverses LHCP aggregation. For a disaggregase that is highly specific to its substrate, the nature of LHCP aggregates is key

to understanding the capacity and limitations of cpSRP43. Further experiments are necessary for elucidating higher-resolution architecture of this aggregate.

## Materials and Methods

**Materials.** LHCP and LHCP mutants were purified under a denaturing condition as described (17). A $\beta^{1-40}$  and re-crystallized ThT were generous gifts from Dr. J. W. Kelly. ANS and bis-ANS were purchased from Sigma and Invitrogen, respectively. LDAO, DDM,  $\beta$ -OG, and BNG were obtained from Anatrace. SDS was from BioRad. Urea and guanidinium chloride were molecular biology grade from MP and Sigma, respectively.

**Light scattering.** Light scattering experiments were performed as previously described (17). For formation of aggregates (Figure 4.1A, black), unfolded LHCP in 8M urea was directly diluted into Buffer D (50 mM K-HEPES pH 7.5, 200 mM NaCl) to the final concentration in each data point, controlling for equal final concentration of urea. For serial dilution experiments (Figure 4.1A, red), the sample at 1  $\mu$ M LHCP was serially diluted (by twofold) into fresh buffer D. The CMCs are obtained as the X-intercepts from linear fits of data from the aggregate formation concentration series.

**TEM.** LHCP aggregate was formed by diluting unfolded LHCP in 8M urea into Buffer D to the final concentration of 2  $\mu$ M. After incubation at 25 °C for 5 minutes, the sample was diluted fivefold 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 (32).

**AFM.** 1  $\mu$ 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 then rinsed with Millipore water and dried under the weak flux of nitrogen. AFM images were

immediately taken 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. The distribution of particle sizes was obtained by calculating the projected area of each particle at half maximum 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. By taking the cross section area at half the maximum height, one can obtain 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  $\mu\text{M}$  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  $\mu\text{M}$  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  $\text{A}\beta^{1-40}$ . 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 8M urea were measured.

**Sedimentation assay.** Unfolded LHCP was diluted to 10  $\mu\text{M}$  in Buffer D and incubated at 25 °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  $\mu\text{l}$  of various detergent or chemical denaturants at different

concentrations for 30 minutes at 25 °C. The mixtures were then spun at 13,000 rpm in a microfuge for 30 minutes, and soluble (S) and pellet (P) fractions were visualized by SDS-PAGE.

For Figure 4.6B, the intensity of the Coomassie-stained bands were quantified using ImageJ (32). The data were fit with a function derived from the two-state model for protein folding (33), shown below:

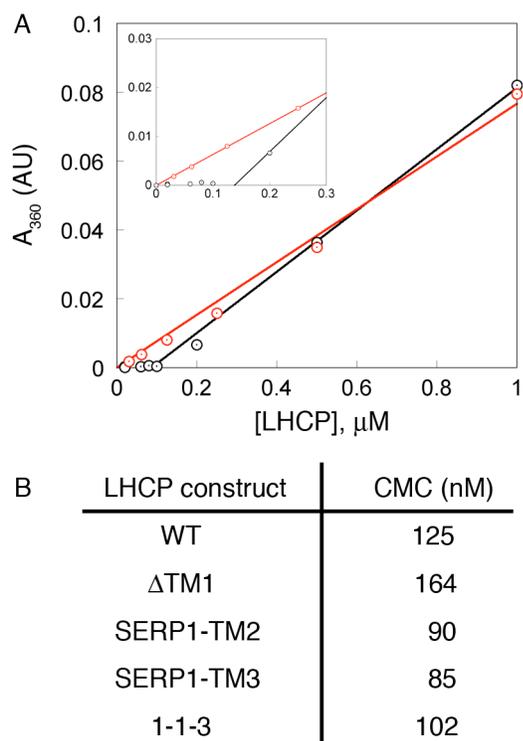
$$\theta = \frac{1}{1 + e^{-m(U_{50} + [\text{urea}])/RT}} \quad (1)$$

where  $\theta$  is the fraction soluble [ $s/(s+p)$ ],  $R$  is the gas constant, and  $T$  is temperature. The fit gave  $U_{50}$ , which is the urea concentration at which 50% of LHCP aggregate was solubilized, and  $m$ , which represents a constant of proportionality.  $\Delta G^\circ$ , which represents free energy of transfer of the aggregate from water to urea, could be calculated from fits to eq 5 ( $= -mU_{50}$ ).

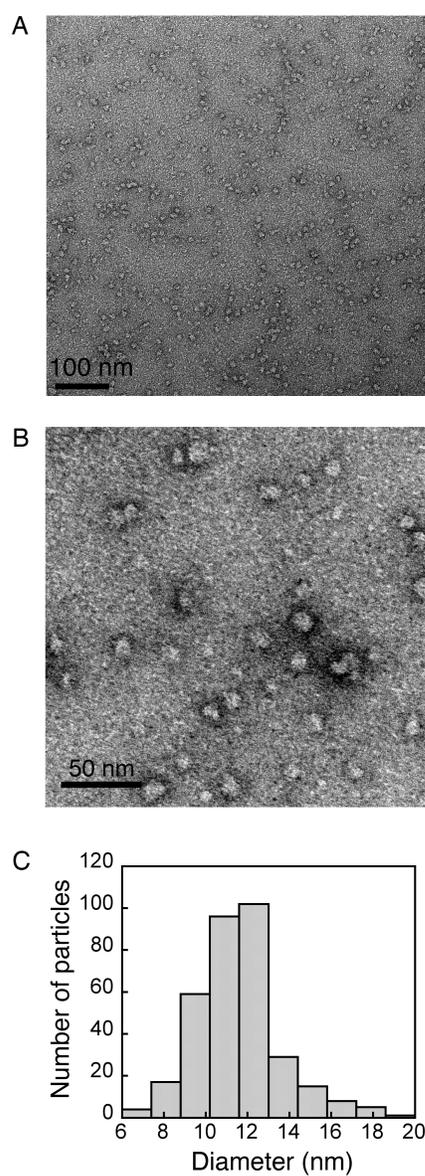
**SDS solubility assay.** For Figure 4.5B, the assay was performed as described for amyloid fibrils (24). Briefly, aggregation of 10  $\mu$ M LHCP in Buffer D proceeded for 5 minutes at 25 °C. The mixture was then mixed with 2% SDS-PAGE loading buffer and either incubated at 25 °C or boiled at 100 °C for ten minutes prior to gel loading. Total mixture was loaded without centrifugal separation, and only the proteins that migrated into the separating gel (e.g., solubilized portion) were visualized.

**Acknowledgement.**

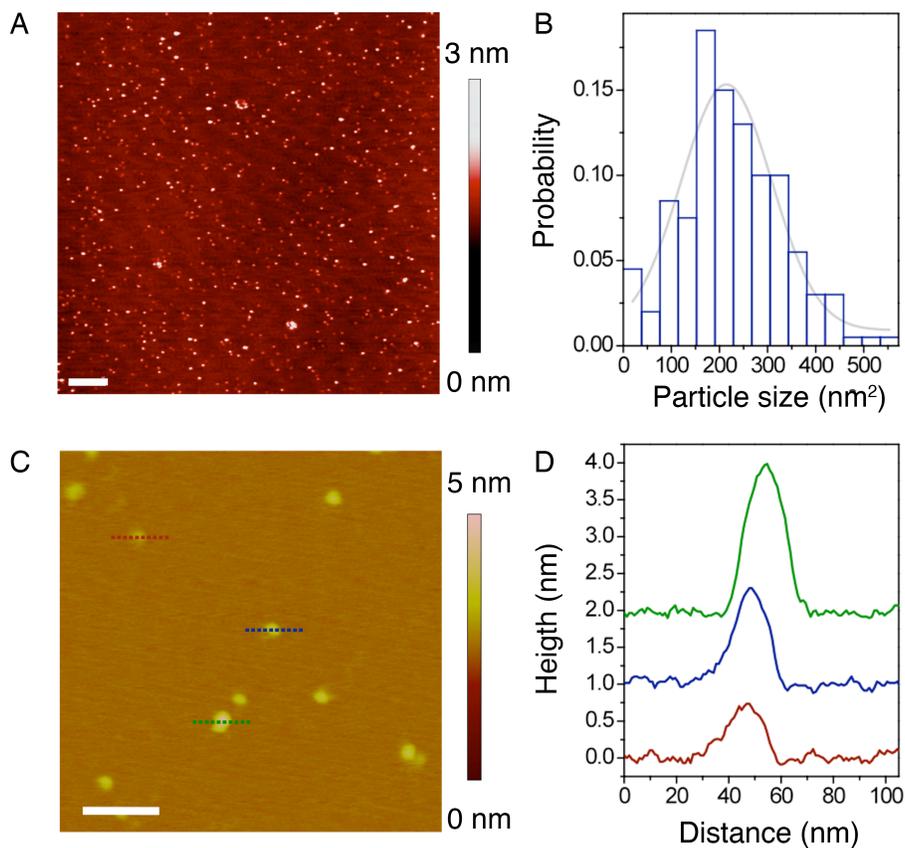
We thank A.N. Murray and Dr. J.W. Kelly for re-crystallized ThT and A $\beta^{1-40}$ , and Dr. A. McDowell and P. Cao for extensive help with TEM and AFM analyses, respectively.



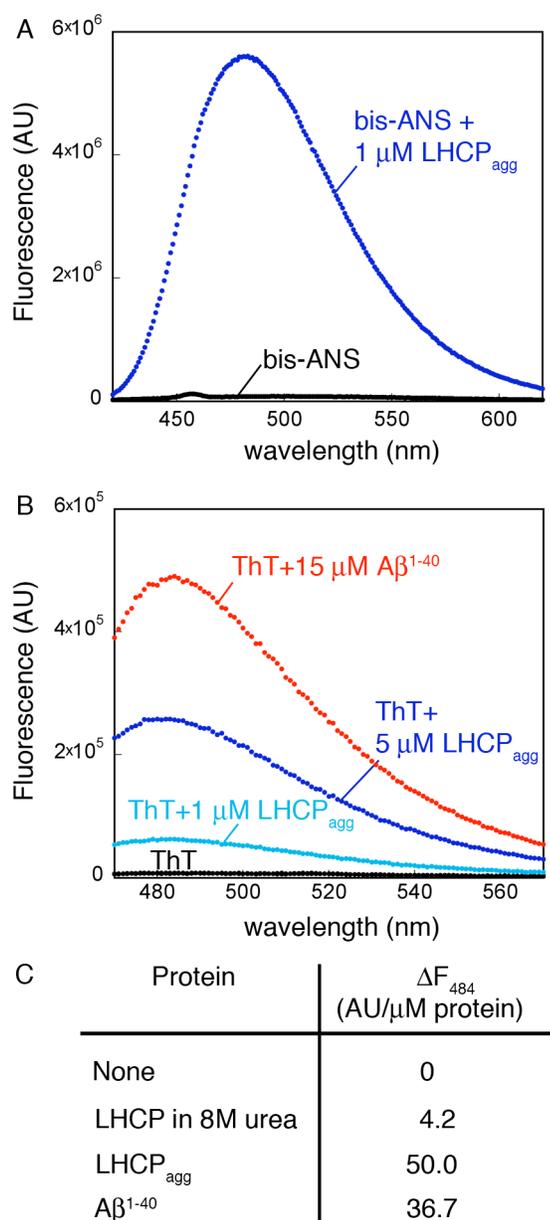
**Figure 4.1** LHCP forms aggregates after a critical concentration. (A) Formation of the LHCP involves a lag phase. Light scattering intensities during formation of the aggregate (black) are compared to those obtained from serial dilution of pre-formed aggregates (red). The inset highlights the lag phase at low concentrations that is only present during formation of the aggregate. (B) Summary of the CMC values of aggregate formation by different LHCP mutants.



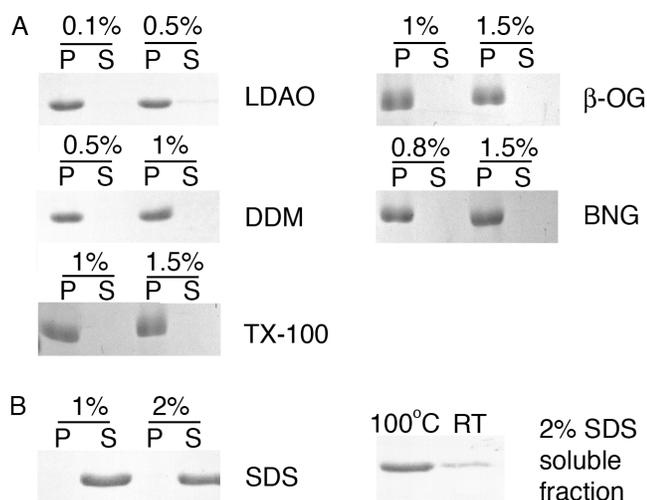
**Figure 4.2** TEM analysis of LHCP aggregates. (A) Large-field view of a negatively stained TEM image of LHCP aggregates shows well-separated and similarly sized particles. (B) A zoomed-in image shows that LHCP aggregates are round-shaped particles. (C) Distribution of the diameter of LHCP aggregates, measured from several independent experiments. The mean diameter is  $12 \pm 2$  nm.



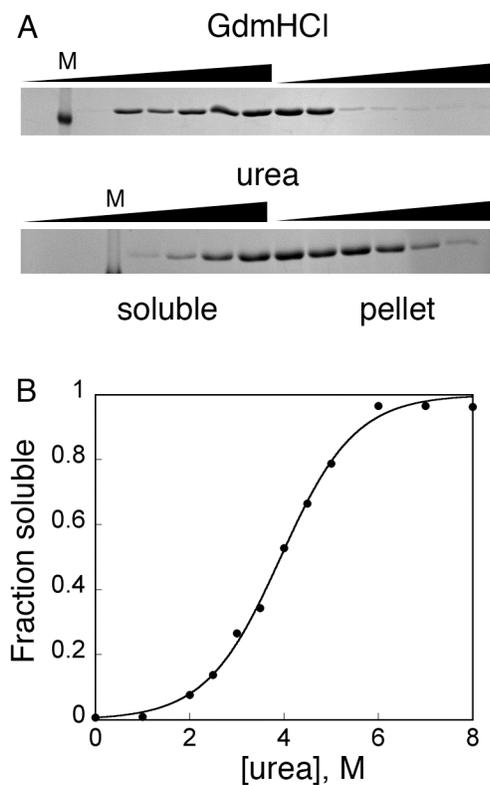
**Figure 4.3** AFM analysis of LHCP aggregates. (A) Large-field view of AFM topographic image shows well-separated LHCP aggregates. Large clusters are occasionally observed. The scale bar is 500 nm. (B) Distribution of the area of LHCP aggregates, measured from several regions on the surface. The gray line is a Gaussian fit to the data, which gave a mean area of the particle of 214 nm<sup>2</sup>. (C) 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. (D) The height profiles for the particles indicated in C. Curves are vertically displaced for clarity.



**Figure 4.4** LHCP aggregates contain exposed hydrophobic surfaces as probed by small molecule dyes. (A) Fluorescence emission spectra of 1  $\mu\text{M}$  bis-ANS with (blue) or without (black) 1  $\mu\text{M}$  LHCP aggregate. (B) Fluorescence spectra of 20  $\mu\text{M}$  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) ThT fluorescence change at 484 nm per  $\mu\text{M}$  of protein.



**Figure 4.5** LHCP aggregates are resistant to many detergents. (A) Sedimentation analysis of the ability of various detergents to resolubilize LHCP aggregates. P and S denote the pellet and soluble fractions, respectively. (B) Sedimentation analysis (left) and SDS-solubility assay as described for amyloids ((24); right) show partial solubility of LHCP aggregates in 2% SDS. In the left panel, the samples were centrifuged to separate pellet and soluble fractions and boiled prior to gel loading. In the right panel, the samples were not separated by centrifugation, and solubility was judged by the mobility of the protein band into the separating gel. Quantification using ImageJ revealed that 24% of the LHCP aggregate is soluble when the sample was not boiled (RT), compared to 87% for the boiled sample (100 °C).



**Figure 4.6** LHCP aggregates can be re-solubilized by chemical denaturants. (A) Sedimentation analysis of the ability of guanidinium chloride (GdmHCl) and urea to re-solubilize LHCP aggregates. ‘M’ denotes the protein marker lane. (B) Quantification of the solubilized fraction revealed a cooperative dependence on urea concentration. The fit to eq 1 gave a  $\Delta G^0$  of 2.9 kcal/mol and a  $U_{50}$  of 3.93 M.

**References:**

1. Hayer-Hartl, M., and Hartl, F.U. (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein, *Science* 295, 1852–1858.
2. Luheshi, L. M., and Dobson, C.M. (2009) Bridging the gap: from protein misfolding to protein misfolding diseases, *FEBS Lett.* 583, 2581–2586.
3. 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.
4. Kopito, R. R. (2000) Aggresomes, inclusion bodies and protein aggregation, *Trends in Cell Biology* 10, 524–530.
5. Brems, D. N., Plaisted, S.M., Havel, H.A., and Tomich, C.S.C. (1988) Stabilization of an associated folding intermediate of bovine growth hormone by site-directed mutagenesis, *Proc. Natl. Acad. Sci. USA* 85, 3367–3371.
6. Speed, M. A., Wang, D.I.C., and King, J. (1996) Specific aggregation of partially folded polypeptide chains: The molecular basis of inclusion body composition, *Nature Biotechnology* 14, 1283–1287.
7. Nelson, R., Sawaya, M.R., Balbirnie, M., Madsen, A.Ø., Riek, C., Grothe, R., and Eisenberg, D. (2005) Structure of the cross-beta spine of amyloid-like fibrils, *Nature* 435, 773–778.
8. Goldsbury, C., Wirtz, S., Muller, S.A., Sunderji, S., Wicki, P., Aepli, 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.
9. Mastrangelo, I., Ahmed, M., Sato, T., Liu, W., Wang, C., Hough, P., and Smith, S.O. (2006) High-resolution atomic force microscopy of soluble A $\beta$  42 oligomers, *J. Mol. Biol.* 358, 106–119.
10. Wall, J. S., and Hainfeld, J.F. (1986) Mass mapping with the scanning transmission electron microscope, *Annu. Rev. Biophys. Biophys. Chem.* 15, 355–376.
11. LeVine, H. I. (1999) Quantification of beta-sheet amyloid fibril structures with thioflavin T., *Methods Enzymol.* 309, 274–284.
12. 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.
13. Ivanova, M. I., Sievers, S.A., Sawaya, M.R., Wall, J.S., and Eisenberg, D. (2009) Molecular basis for insulin fibril assembly, *Proc. Natl. Acad. Sci. USA* 106, 18990–18995.
14. Doyle, S. M., and Wickner, S. (2008) Hsp104 and ClpB: protein disaggregating machines, *Trends Biochem. Sci.* 34, 40–48.
15. Glover, J. R., and Lindquist, S. (1998) Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins, *Cell* 94, 73–82.
16. Doyle, S. M., Hoskins, J.R., and Wickner, S. (2007) Collaboration between the ClpB AAA+ remodeling protein and the DnaK chaperone system, *Proc. Natl. Acad. Sci. USA* 104, 11138–11144.

17. 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.
18. Schuenemann, 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. USA* 95, 10312–10316.
19. 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.
20. 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.
21. 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.
22. 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.
23. Groenning, M., Olsen, L., van de Weert, M., Flink, J.M., Frokjaer, S., and Jorgensen, 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.
24. Chernoff, Y. O., Uptain, S.M., and Lindquist, S.L. (2002) Analysis of prion factors in yeast, *Methods Enzymol.* 351, 499–538.
25. McLean, C. A., Cherny, R.A., Fraser, F.W., Fuller, S.J., Smith, M.J., Beyreuther, K., Bush, A.I., and Masters, C.L. (1999) Soluble pool of A $\beta$  amyloid as a determinant of severity of neurodegeneration in Alzheimer's disease, *Ann. Neurol.* 46, 860–866.
26. Walsh, D. M., Klyubin, I., Fadeeva, J.V., Cullen, W.K., Anwyl, R., Wolfe, M.S., Rowan, M.J., and Selkoe, D.J. (2002) Naturally secreted oligomers of amyloid- $\beta$  protein potently inhibit hippocampal long-term potentiation *in vivo*, *Nature* 416, 535–539.
27. Dahlgren, K. N., Manelli, A.M., Stine, W.B. Jr, Baker, L.K., Krafft, G.A., and LaDu, M.J. (2002) Oligomeric and fibrillar species of amyloid- $\beta$  peptides differentially affect neuronal viability, *J. Biol. Chem.* 277, 32046–32053.
28. Fink, A. L. (1998) Protein aggregation: folding aggregates, inclusion bodies and amyloid, *Folding and Design* 3, 9–23.
29. 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.
30. Sen, A., Baxa, U., Simon, M.N., Wall, J.S., Sabate, R., Saupe, S.J., and Steven, A.C. (2007) Mass analysis by scanning transmission electron microscopy and

electron diffraction validate predictions of stacked  $\beta$ -solenoid model of HET-s prion fibrils, *J. Biol. Chem.* 282, 5545–5550.

31. Strasfeld, D. B., Ling, Y.L., Gupta, R., Raleigh, D.P., and Zanni, M.T. (2009) Strategies for Extracting Structural Information from 2D IR Spectroscopy of Amyloid: Application to Islet Amyloid Polypeptide, *J. Phys. Chem. B* 113, 15679–15691.
32. Abramoff, M. D., Magelhaes, P.J., and Ram, S.J. (2004) Image Processing with ImageJ, *Biophotonics International* 11, 36–42.
33. Fersht, A. (1998) *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*, W.H. Freeman and Company, New York.

## **Chapter 5:**

### **Probing the Mechanism of cpSRP43-Mediated Protein Disaggregation.**

**Abstract**

Clearing protein aggregates is a daunting task for cells. Thus far, the best characterized “disaggregase” systems belong to the Clp/Hsp100 family of AAA+ ATPases, which use mechanical forces powered by ATP hydrolysis to remodel protein aggregates. Recently, we described an alternative system that can disassemble protein aggregates: the 43-kDa subunit of chloroplast Signal Recognition Particle (cpSRP43). With no co-chaperones or ATPase activity, cpSRP43 utilizes specific binding energy established with its substrate, light-harvesting chlorophyll a/b-binding proteins (LHCP), to power disaggregation. The molecular mechanism of cpSRP43-mediated disaggregation remains to be elucidated. In this chapter, molecular genetics analysis showed that the mechanism of disaggregation by cpSRP43 can be dissected into two steps with distinct molecular requirements: (i) initial recognition, which depends on the binding of cpSRP43 to important recognition motifs exposed on the surface of the aggregate; and (ii) subsequent remodeling/solubilization of the aggregate, which the binding interaction of LHCP to cpSRP43 competes with the internal packing interactions within the aggregate. This work establishes a useful framework to understand the mechanism of action of ATP-independent protein disaggregases.

## Introduction

Protein homeostasis is vital to all cells and requires the balance of protein production, folding, localization, assembly, and degradation (1). Crucial to the maintenance of protein homeostasis is an elaborate network of “molecular chaperones” (2–4), which ensure proper protein folding and prevent protein aggregation by interacting with exposed hydrophobic residues or unstructured backbone regions that are present in non-native proteins (2). However, due to environmental stress, the capacity of the chaperone network could be exceeded or impaired, and protein aggregation ensues. To overcome this, cells have evolved a special set of machineries called disaggregases to rescue these protein aggregates. The most studied disaggregases belong to the Clp/Hsp100 family: yeast heat shock protein 104 (Hsp104) and its bacterial homologue ClpB (5). Both are members of the ATPases associated with various cellular activities (AAA+) superfamily and assemble into hexameric ring structures (5). These disaggregases use repetitive ATPase cycles and, in collaboration with their co-chaperones, mostly Hsp70 and Hsp40, remodel large protein aggregates via unfolding and translocation of the substrate polypeptides through their central pores (6–8). Although present in plant and eukaryotic mitochondria, homologues of these disaggregases have not been found in higher eukaryotes (5). Instead, the recent discovery of ATP-independent disaggregase activities in homogenates derived from *C. elegans* and human tissue raised the possibility that alternative disaggregation mechanisms exist in higher eukaryotes (9, 10).

Recently, we described a novel disaggregase system that operates independently of ATP: the 43-kDa subunit of the chloroplast Signal Recognition Particle (cpSRP43)

(11). The substrates for 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 (12). The most abundant member of this protein family, LHCP, comprises up to 50% of the protein content in the thylakoid membrane (13). LHC proteins contain three highly hydrophobic transmembrane (TM) helices, making them highly prone to aggregation while traversing aqueous compartments in the cell (13, 14). Work from our and other laboratories have shown that cpSRP, in particular the cpSRP43 subunit, acts as a potent molecular chaperone for the LHC proteins (11). Intriguingly, cpSRP43 could also efficiently reverse the aggregation of its substrate proteins without the requirements for ATP hydrolysis or co-chaperones. Instead, cpSRP43 utilizes its specific binding energy to power its disaggregation activity. The basis for binding specificity is provided by a highly conserved recognition element, L18, a relatively hydrophilic 18-amino acid loop between TM2 and TM3 of the LHC proteins (15, 16). Nevertheless, the binding energy provided by L18 is moderate, with a dissociation constant ( $K_d$ ) of  $\sim 1\text{--}2\ \mu\text{M}$  for cpSRP43 (17), whereas the binding affinity between cpSRP43 and full-length LHCP is 138 nM (11). Moreover, the ability of cpSRP43 to prevent LHCP from aggregation implicitly indicates sequestration of hydrophobic regions of LHCP upon binding to cpSRP43. Therefore, we proposed that additional interactions must between LHCP and cpSRP43. This is supported by the fact that cpSRP43 is an elongated molecule replete with domains and motifs typically involved in protein–protein interactions: four ankyrin repeats (Ank1–Ank4) and three chromodomains (CDs) (11, 17–20). These domains can provide possible interaction interfaces for the LHCP molecule.

cpSRP43 represents an example of a novel class of disaggregases that operate with energy derived entirely from binding interactions with its substrates. Understanding its mechanism will provide valuable insights into alternative approaches that the cells use to handle protein aggregates. Previous kinetic analyses revealed that disaggregation is a cooperative process and suggested that cpSRP43 binds to and actively remodels the LHCP aggregate (11). We propose that initial recognition of the LHCP aggregate occurs via interactions with L18. This is substantiated by evidence from both electron paramagnetic resonance (EPR) and chemical modification experiments (V.Q. Lam and T.X. Nguyen, unpublished results), both of which showed that LHCP aggregates form by burying their hydrophobic TM segments in the interior, whereas the L18 motif is exposed on the exterior, corroborating an earlier speculation that LHCP aggregates are micellar (see Chapter 4). These results suggest an attractive starting point for disaggregation: the LHCP aggregates expose the L18 motif on the exterior, poised for recognition by cpSRP43.

After initial recognition of the aggregate, what happens next? The observed cooperativity suggested that binding of one cpSRP43 facilitates binding of subsequent cpSRP43 molecules, implying an aggregate remodeling step after cpSRP43 binding. This step could be independent of binding to the exposed L18. Using distinct classes of mutants, we present evidence that (i) the remodeling step could be distinguished from the initial binding step and (ii) the stabilities of both the starting LHCP aggregates and the final cpSRP43•LHCP soluble complex play large roles in determining the extent of aggregate remodeling and re-solubilization.

## Materials and Methods

**Materials.** To construct the LHCP TM mutant proteins, the expression plasmid encoding LHCP was modified to contain a pair of unique restriction sites before and after the regions coding TM1, TM2, or TM3. Then, the sequences coding for original TMs were replaced with PCR fragments encoding exogenous TMs using corresponding restriction sites. TM deletion mutants were constructed using QuikChange (Stratagene) with primers omitting the coding regions of TM1, TM2, or TM3. Description of each mutant was summarized in Supplementary Table 5.S1. Lhcb5 cystein mutants were constructed using QuikChange (Stratagene). LHCP and its variants were purified under a denaturing condition as described (11).

**Light scattering.** Light scattering experiments were performed as previously described (11). To determine  $K_d^{app}$  by light scattering, prevention of LHCP aggregation was used. Briefly, absorbance at 360 nm was measured after 10 minute incubation with varying concentrations of cpSRP43. The light scattering signals were normalized to the absorbance recorded from the sample with no chaperone, and percentage soluble (% soluble) was calculated by subtracting percentage light scattering at each cpSRP43 concentration from 100. The light scattering is linearly proportional to the concentration of LHCP except at very low concentrations. Data were fit with the quadratic eq 1

$$\% \text{ soluble} = 100 \times \frac{[\text{LHCP}] + [\text{cpSRP43}] + K_d^{app} - \sqrt{([\text{LHCP}] + [\text{cpSRP43}] + K_d^{app})^2 - 4 \times [\text{LHCP}][\text{cpSRP43}]}}{2 \times [\text{LHCP}]} \quad (1)$$

in which  $K_d^{app}$  is the apparent dissociation constant. Due to the inaccuracy of the measurement at very low unbound LHCP concentrations (see Chapter 4), the  $K_d^{app}$  values could be mildly underestimated. Nevertheless, comparison of the  $K_d$  values measured by

light scattering and fluorescence anisotropy for wild-type LHCP ( $111 \pm 3$  nM compared to 138 nM measured by anisotropy; (11)) and a few of the LHCP TM mutants showed that the two methods produced comparable numbers (Supplementary Figure 5.S1).

Disaggregation reactions were performed as previously described (11), with the exception that the aggregation period was 1 minute. The disaggregation time courses were fit to an exponential function (eq 2)

$$A = A_f + \Delta A e^{-k_{obsd}t} \quad (2)$$

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 observed light scattering is normalized to that prior to the addition of cpSRP43 for each disaggregation reaction. The fractions disaggregated ( $\theta$ ) were calculated by subtracting  $A_f$  from 1. The concentration dependences of fractions disaggregated were fit to eq 3

$$\theta = \theta_{max} \times \frac{[cpSRP43]^n}{(K_m^{dis})^n + [cpSRP43]^n} \quad (3)$$

in which  $\theta_{max}$  is the extent of disaggregation at saturating cpSRP43 concentration,  $K_m^{dis}$  is the average affinity of cpSRP43 to LHCP aggregates, and  $n$  is the Hill coefficient.

Kinetic analysis was performed and analyzed as described previously (11). Briefly, the concentration dependence of the disaggregation rate constants,  $k_f$ , was fit to eq 4,

$$k_f = k_0 \times \frac{K_d^n}{K_d^n + [cpSRP43]^n} + k_{max} \times \frac{[cpSRP43]^n}{K_d^n + [cpSRP43]^n} \quad (4)$$

in which  $k_0$  is the rate of spontaneous LHCP disaggregation in the absence of the chaperone,  $K_d$  is an average equilibrium dissociation constant for binding of cpSRP43 to LHCP aggregates,  $n$  is the Hill coefficient, and  $k_{\max}$  is the disaggregation rate constant at saturating cpSRP43 concentration.

**Fluorescence anisotropy.** Fluorescence labeling of LHCP and its variants and fluorescence anisotropy were carried out as described previously (11).

**Sedimentation assay.** Urea solubilization experiments and data analyses were carried out as described (Chapter 4). Briefly, re-solubilization of the LHCP aggregates by urea was monitored by SDS-PAGE. The intensity of the Coomassie-stained bands for the pellet and soluble fractions were quantified using ImageJ (21). The data were fit with a function derived from the two-state model for protein folding (22), shown below:

$$\theta = \frac{1}{1 + e^{-m(U_{50} + [\text{urea}])/RT}} \quad (5)$$

where  $\theta$  is the fraction soluble [ $s/(s+p)$ ],  $R$  is the gas constant, and  $T$  is temperature. The fit gave  $U_{50}$ , which is the urea concentration at which 50% of LHCP aggregate was solubilized, and  $m$ , which represents a constant of proportionality.  $\Delta G^\circ$ , which represents free energy of transfer of the aggregate from water to urea, could be calculated from fits to eq 5 ( $= -mU_{50}$ ).

**Mathematical analyses.** All correlations were obtained by regression fits of the obtained parameters in Table 5.3 using the “polyfitn” tool in Matlab (23).

## Results

The current working model of LHCP disaggregation by cpSRP43 involves two steps: (i) initial binding and (ii) subsequent remodeling of the aggregate. To provide supporting evidence for this model, we sought mutants that specifically block distinct stages of disaggregation. In this work, we present thermodynamic and kinetic analyses of disaggregation reactions of aggregates formed by various LHCP mutants and characterize intrinsic properties of these mutants to reveal the distinct molecular requirements for each step of disaggregation. Table 5.1 summarizes the parameters used for comparison in this study.

### *Binding to L18 on the LHCP aggregate is the first step to disaggregation.*

Previous results led us to propose that cpSRP43 recognizes and binds to the LHCP aggregates via interactions with the solvent-exposed L18 motif (T.X. Nguyen and V.Q. Lam, unpublished results). If this were the case, mutants of LHCP or cpSRP43 that specifically impair the L18–cpSRP43 interaction would exhibit defects in the disaggregation at low chaperone concentrations. However, as binding is a higher-order process, the defects of these mutants should be rescued when a sufficiently high concentration of the chaperone is used to drive binding. To test this hypothesis, we examined the disaggregation efficiencies of the L18 mutants on LHCP or the L18-binding mutants on cpSRP43.

The crystal structure of the cpSRP43–L18 complex identifies Arg161 in the ankyrin repeats of cpSRP43 as an important hydrogen bond partner with the L18 peptide (17). As expected, the cpSRP43-R161A mutant significantly reduces its binding affinity

to solubilized LHCP ( $K_d = 1.2 \mu\text{M}$ , compared to 138 nM for wild-type cpSRP43; Table 5.2; (11)). Reciprocally, we identified two mutants in the L18 region of Lhcb5, a close homologue of LHCP, that were defective in binding to cpSRP43: H160C and L170C. Using the light scattering assay for prevention of aggregation, we could conveniently measure apparent binding affinity of various LHC protein mutants (see Methods). The concentration dependence of the amount of the soluble LHCP gave a binding curve that could be fit with a quadratic equation to yield an apparent dissociation constant ( $K_d^{\text{app}}$ ), which is comparable to those measured by fluorescence anisotropy (See Methods and Supplementary Figure 5.S1; (11)). This assay yields  $K_d^{\text{app}}$  values of 30 nM and 1.1  $\mu\text{M}$  for Lhcb5-H160C and -L170C, respectively, compared to 10 nM for wild-type Lhcb5 (Table 5.2 and Supplementary Figure 5.S2).

Consistent with a defect in recognition of the LHCP aggregate, 10  $\mu\text{M}$  of cpSRP43-R161A could not reverse LHCP aggregation (Figure 5.1A, magenta), whereas almost 100% reversal of aggregation could be obtained for wild-type cpSRP43 at this concentration (Figure 5.1A, black; (11)). However, when the concentration of the mutant chaperone was raised to compensate for the binding defect, cpSRP43-R161A could efficiently reverse LHCP aggregation (Figure 5.1B, magenta, and Table 5.2). Analogously, the aggregates formed by the L18 mutants in Lhcb5 showed defects in the disaggregation reaction that can be rescued by higher cpSRP43 concentrations (Figure 5.1B and Table 5.2). The more severe mutant, Lhcb5-L170C, required a much higher concentration of cpSRP43 to achieve efficient solubilization of the aggregate than the less severe mutant (Figure 5.1B, red). Together, the results in this section showed that L18 binding is an important requirement for disaggregation, and suggest that this initial

recognition step can be uncoupled from the subsequent step(s) that are unimolecular in nature.

***cpSRP43 makes non-specific contacts with the hydrophobic TM segments.***

In order to remodel the aggregate in the subsequent step(s), cpSRP43 must compete with the internal packing of the aggregates and provide alternative binding interactions for the hydrophobic TMs that were buried inside the aggregate. However, it was unclear whether these interactions are highly sequence-specific, like that of the L18, or are more generic hydrophobic contacts, like those of other promiscuous chaperone–substrate interactions (2). To address this question, we constructed various LHCP mutants in which the individual TMs are deleted or replaced with other TM segments in LHCP or the TMs from unrelated membrane proteins (see Supplementary Table 5.S1 for nomenclature and description of all LHCP TM mutants used in this study). We reasoned that if these interactions were sequence-specific, deletion or replacement of the TM helices should result in significant reduction in cpSRP43’s chaperone activity, analogous to mutations in the L18 motif (11). On the other hand, if these interactions arose from generic hydrophobic interactions or backbone contacts, many of these TM replacements would not result in deleterious defects in the chaperone activity.

We purified several LHCP TM mutants and showed that cpSRP43 could still bind and protect them from aggregation (Figure 5.2, A and B). Deletion of TM helices did not result in significant defects in the chaperone activity of cpSRP43, although these LHCP deletion mutants still aggregated to similar extents as wild-type LHCP (Figure 5.2 and data not shown). Some mutants, such as  $\Delta$ TM3, SERP2, Sec2, interact even better and

are hence more readily protected by cpSRP43, whereas other mutants, such as 1-3-2, 1-2-2, and 1-1-3 exhibited moderately lower (two–fivefold) affinity to cpSRP43 (Figure 5.2B). Nevertheless, virtually all of the TM replacement mutants exhibit moderate-to-high binding affinities for cpSRP43 in their soluble form, and these binding affinities are significantly higher than that of cpSRP43 for the isolated L18 peptide (17). This strongly suggests that the hydrophobic TMs in LHCP contribute additional binding interactions with cpSRP43, and these interactions are fairly generic and highly adaptable, in contrast to the strictly sequence-specific interactions of cpSRP43 with 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.

***Some LHCP TM mutant aggregates are virtually irreversible.***

Since these TM mutants contain identical L18 motifs, they provide a good collection of substrates to probe for other molecular requirements for efficient disaggregation. Using the established light scattering assay (Supplementary Figure 5.S3), we analyzed both the equilibrium and the rate of the disaggregation reactions for these LHCP TM variants. As a group, they exhibited striking differences in both thermodynamics and kinetics of disaggregation (Figure 5.3 and Table 5.3).

Plots of the disaggregated fraction showed cooperative dependence on the concentrations of cpSRP43 (Figure 5.3A). Fits of these plots gave two important thermodynamic parameters: (i) the average affinities of cpSRP43 for binding to the various LHCP aggregates ( $K_m^{\text{dis}}$ ), which reflect the concentrations of cpSRP43 needed for efficient disaggregation; and (ii) the disaggregated fractions at saturating cpSRP43

concentration ( $\theta_{\max}$ ), which represent the extent of disaggregation reactions for these LHCP mutant aggregates once cpSRP43 has bound to the aggregate. Additionally, the  $\theta_{4\mu M}^{app}$  values, the extent of disaggregation at sub-saturating cpSRP43 concentration, provide important practical measures of disaggregation efficiency especially for mutants whose  $\theta_{\max}$  values cannot be accurately determined.

Surprisingly, even though these LHCP TM mutants showed significant binding interactions with cpSRP43 in their soluble forms, we found that some LHCP TM mutants, especially 1-1-3,  $\Delta$ TM2, 1-2-2, and 1-3-2, were virtually irreversible (Figure 5.3A, red, and Supplementary Figures 5.S4–5.S6). To a lesser extent,  $\Delta$ TM1 and 1-2-1 also show large defects in disaggregation equilibrium (Supplementary Figures 5.S7 and 5.S8). These mutants carry solvent-exposed wild-type L18 in their aggregates (T.X. Nguyen, unpublished results). Therefore, we hypothesized that this group of mutants did not block the recognition step but possibly disrupted the remodeling step. On the other hand, some TM mutants showed disaggregation efficiencies that are similar or even better than the wild-type LHCP aggregate (Figure 5.3, blue and green, and Supplementary Figures 5.S9–5.S12).

The disaggregation rate constants also showed cooperative dependence on the concentrations of cpSRP43 (Figure 5.3B). The  $k_{\max}$  values, which represent the disaggregation rate constants at saturating cpSRP43 concentrations, reveal significant differences in disaggregation kinetics among these LHCP TM mutants (Table 5.3). The irreversible mutants, if at all, were disaggregated at extremely slow rates (Figure 5.3B, red, and Supplementary Figures 5.S4–5.S6). Meanwhile, other mutant aggregates, notably those of  $\Delta$ TM3, SERP2, Sec2, and Cyb2, could be re-dissolved more easily than

the wild-type aggregate and exhibited much faster disaggregation rates (Figure 5.3 A and B, green, and Supplementary Figures 5.S10–5.S12).

Overall, the results in this section indicated that (i) the irreversible mutants specifically block the remodeling step with minimal perturbations in binding, providing additional support for the two-step disaggregation mechanism, and (ii) this group of LHCP TM mutants displays a wide range of disaggregation efficiencies and kinetics (Table 5.3). Further analyses of the molecular features underlying these differences may provide valuable insights into the molecular mechanism of protein disaggregation.

***The irreversible LHCP TM mutants form ultra-stable aggregates.***

Unlike the L18-binding mutants, the irreversible LHCP TM mutants bind reasonably well to cpSRP43. What then could cause the drastic defects in the equilibrium and kinetics of disaggregation? Previous results from chemical modification and EPR experiments showed that the TM helices engage in strong interactions and are buried inside the aggregate (T.X. Nguyen and V.Q. Lam, unpublished results). Since the TM helices are considerably altered in these LHCP mutants, we hypothesized that the internal packing interactions of their aggregates are different and may play a role in disaggregation. To test this hypothesis, we probed the strength of the packing interactions inside the aggregates using urea solubilization (Chapter 4). For comparison, we used the  $U_{50}$  values, the urea concentrations at which 50% of LHCP aggregates are solubilized (Table 5.1), which provide empirical measures for the strength of the packing interactions that stabilize the aggregates formed by the LHCP TM variants.

As expected, aggregates from LHCP TM mutants exhibited a wide range of stabilities, with the  $U_{50}$  values from 2.5 to 5.7 M (Table 5.3). Mutant aggregates that are easily disaggregated collectively displayed low  $U_{50}$  values (2.5–3.6 M; Figure 5.4, green, and Supplementary Figures 5.S10–5.S12), whereas wild-type LHCP and the neutral mutants exhibited intermediate  $U_{50}$  values (3.7–4.7 M; Figure 5.4, black and blue, and Supplementary Figures 5.S7–5.S9). Importantly, the four irreversible mutants that could bind but were extremely slow and inefficient in disaggregation had the highest  $U_{50}$  values (4.7–5.7 M; Figure 5.4, red, and Supplementary Figures 5.S4–5.S6). These results indicate that the internal packing interactions are crucial in determining the extent of a disaggregation reaction.

### ***Molecular determinants of disaggregation efficiency.***

To understand the contributions of the molecular features dictating the disaggregation more rigorously, we correlated the extent of disaggregation with the two thermodynamic parameters that we measured, the  $K_d^{\text{app}}$  values and the  $U_{50}$  values. First, we reasoned that binding between cpSRP43 and the LHCP monomer should provide the ultimate driving force for the disaggregation reaction, as the cpSRP43•LHCP soluble complex is the product of the reaction. However, we barely observed a correlation between the apparent binding affinities of these mutants in their soluble forms ( $K_d^{\text{app}}$ ) with the extent of disaggregation ( $\theta_{4\mu\text{M}}^{\text{app}}$ ) (Figure 5.5A;  $R^2 = 0.58$ ).

A qualitative observation that the four irreversible mutants did not have severe binding defects yet were very difficult to reverse (Table 5.3, red) prompted us to ask if other properties, especially the internal packing interactions, additionally contributed to

the disaggregation efficiency. When we included a dimension representing the stability of the LHCP aggregates in our mathematical analysis, we could observe an improvement in the correlation between the extent of disaggregation and the stabilities of both the final product (the cpSRP•LHCP soluble complex;  $K_d^{\text{app}}$ ) and the starting aggregate ( $U_{50}$ ) (Figure 5.5B;  $R^2 = 0.77$ ). The data points highlighted in blue in Figure 5.5 were derived from the set of Lhcb5 proteins (Figure 5.1B), which aggregated more slowly and presented another layer of complication for disaggregation analysis compared to the set of LHCP TM mutants. It is possible that these different LHC proteins were disaggregated in pathway-specific manners and that they had additional less well-understood molecular requirements for efficient disaggregation. Presently, we could not unambiguously pinpoint the origins of these outliers. Nevertheless, these correlational analyses show that both binding to LHCP and competing with aggregate packing interactions are crucial in determining the overall disaggregation efficiencies.

Kinetic analysis provides additional insights into the rate-limiting step of the mechanism. We observed a decent negative correlation between the  $U_{50}$  values and the maximal disaggregation rate constants,  $k_{\text{max}}$  (Figure 5.6A;  $R^2 = 0.78$ ). This observation strongly suggests that packing interactions inside the aggregates determine the activation barrier that cpSRP43 has to overcome during disaggregation. Further analyses of the differences in the equilibrium and the rates of disaggregation between the wild-type and mutant LHCPs yielded valuable information regarding the transition state of the disaggregation reaction. For each mutant, we calculated its perturbation on the free energy of the overall reaction ( $\Delta\Delta G \sim \ln K^{\text{app}}$ ) relative to its perturbation on the free energy of the transition state ( $\Delta\Delta G^\ddagger \sim \ln k^{\text{app}}$ ) and observed a linear relationship between

the two terms for the set of the LHCP mutants (Figure 5.6B). Analogous to the classical Bronsted  $\beta$ -value analysis or the  $\Phi$ -value analysis of the protein-folding mutants (22), this correlation could be used to infer information about the transition state of the reaction. The fractional slope of 0.5 (Figure 5.6B) could not be unambiguously interpreted. It possibly resulted from either the partially formed cpSRP43–LHCP interactions or the mixture of transition states from different kinetic pathways (22). However, the value significantly higher than zero implies that, in the transition state of the reaction, interactions between LHCP molecules and the disaggregase were already formed, albeit not as strong as those ultimately present in the final soluble product.

## Discussion

Thermodynamic and kinetic analyses of the disaggregation reactions of various LHCP mutants allow us to dissect the molecular steps during the disassembly of the LHCP aggregates by cpSRP43, and revealed several distinct and crucial molecular requirements for the disaggregation reaction mediated by cpSRP43. These data led us to propose a two-step working model for the action of cpSRP43 as a protein disaggregase (Figure 5.7A).

### *Step 1: Binding to the LHCP aggregate.*

To initiate the disaggregation reaction, cpSRP43 must recognize and engage the LHCP aggregates. This likely occurs via cpSRP43's binding to its primary recognition motif L18, which is presented on the exterior of the LHCP aggregate (Figure 5.7A, step 1). In support of this model, mutants of cpSRP43 or LHC proteins that disrupt the interactions between the L18 motif and cpSRP43 specifically affect this initial step mainly by destabilizing the intermediate cpSRP43•LHCP<sub>agg</sub> complex due to their compromised affinity of the L18 motif (Figure 5.7B, magenta line). This increases the free energy cost for the first binding step (Figure 5.7B,  $\Delta G_{\text{bind}}$  magenta vs. black). As binding is a higher-order process, this defect could be overcome by increasing the concentration of the disaggregase.

Markedly, there is a positive correlation between the overall apparent binding affinities between cpSRP43 and soluble LHCP ( $K_d^{\text{app}}$ ) and the binding affinities between cpSRP43 and LHCP aggregates ( $K_m^{\text{dis}}$ ) (Table 5.3). This is surprising because, if the aggregate recognition step involved exclusively the interactions between the L18 motif

and cpSRP43, the  $K_m^{\text{dis}}$  values should reflect only the affinity of cpSRP43 to the L18 motif and should be almost invariant for the group of LHCP TM mutants carrying wild-type L18 motifs. The strong correlation could imply that more than, in the recognition step, cpSRP43 also interacts with other regions in LHCP besides L18. In support of this speculation, recent study has proposed that cpSRP43 makes additional contacts with LHCP, particularly the beginning of the third TM helix of LHCP (24). Alternatively, this correlation could suggest different presentations of the L18 motifs by the aggregates formed by various LHCP TM mutants such that the average binding constants to the aggregates are different. Further studies to distinguish these possibilities are needed.

***Step 2: Remodeling of the LHCP aggregate.***

For the subsequent aggregate remodeling step, cpSRP43 must break the aggregate packing interactions and provide sufficient binding surfaces for the dislodged LHCP monomers, preventing re-aggregation (Figure 5.7A, step 2). Once cpSRP43 is bound to the aggregate, the extent of the disaggregation reaction (represented by  $\theta_{\text{max}}$ ) is determined by the free energy of the remodeling step, which is the difference in free energy between the cpSRP43•LHCP complex and the intermediate cpSRP43•LHCP<sub>agg</sub> complex (Figure 5.7B,  $\Delta G_{\text{rem}}$ ). Therefore, the modulations in the stability of either complex will induce changes in the equilibrium of disaggregation. For example, the irreversible mutants form ultra-stable aggregates (Figure 5.4, red, and Table 5.3, red), which likely result in the stabilization of the intermediate to the same extent (Figure 5.7C). This decreases the drive for the remodeling step. For these mutants, this stabilization of the aggregate and the intermediate state could render the subsequent

remodeling step energetically unfavorable (Figure 5.7C, red vs. black  $\Delta G_{\text{rem}}$ ). In this scenario, disaggregation of these mutants cannot efficiently occur even at high cpSRP43 concentrations, because the remodeling step is independent of chaperone concentration. On the contrary, the mutants that are easily reversible (Table 5.3, green) have an opposite free energy landscape. Their aggregates are destabilized, and their soluble cpSRP43•LHCP complexes are stabilized compared to wildtype. Hence, they exhibit greater disaggregation efficiency due to a larger thermodynamic drive.

The stability of the LHCP aggregates also influences the disaggregation kinetics, as suggested by the observed negative correlation between the disaggregation rates and the  $U_{50}$  values (Figure 5.6A). The mutants that form less stable aggregates and are easily reversible also have lower activation barrier for remodeling, resulting in a faster disaggregation rate. In contrast, the irreversible mutants with ultra-high stability of their aggregates (Figure 5.7C, red) also have a much larger activation barrier to proceed (Figure 5.7C, red  $\Delta G^\ddagger$ ) and hence exhibit slower disaggregation kinetics. The manifestation of the mutational effects on the disaggregation rate is remarkable, and strongly suggests that the rate-limiting transition state for the disaggregation process is fairly late and requires disruption of a substantial amount of internal structures within the aggregate.

### ***Comparison with the Clp/Hsp100 family of protein remodeling complexes.***

It is noteworthy to compare lessons learned from cpSRP43 to those from the force-generating protein-remodeling complexes such as the Clp/Hsp100 family of ATPases. First, how do the disaggregases select their substrates in the initial step? We

note that the two disaggregase systems differ greatly in their substrate selection.

cpSRP43 is a specific chaperone for the LHC family of proteins, whereas the Clp/Hsp100 ATPases handle a large fraction of the cellular proteome (25). The specificity in cpSRP43-mediated disaggregation is imparted by the recognition of the strictly-conserved L18 motif on the exterior of the aggregate. On the other hand, the AAA+ remodeling machines displays much more promiscuity in substrate selection. A binding study using a peptide library showed that ClpB prefers peptides that are enriched in aromatic and basic residues with no specific sequence patterns (26). These peptides comprise the hydrophobic cores of most native proteins and may be exposed as disordered loops when proteins aggregate (6). The less stringent requirements for their substrates allow the Clp/Hsp100 ATPases to be the central disaggregases for the cells.

Second, how do these disaggregase remodel their substrates in the subsequent step(s)? Although the precise molecular details remain to be elucidated, our results herein provide some clues to the mechanism of aggregate remodeling by cpSRP43. First, the cooperativity observed in disaggregation strongly argues for the changes in the structure of the LHCP aggregates brought upon by cpSRP43 binding and suggests possible destabilization of the aggregates. Since the aggregate formation is a cooperative process, as evident in both urea solubilization and disaggregation by cpSRP43, such destabilization likely collapses the whole structure. Second, the negative correlation between the rate of disaggregation and the global stability of the aggregates (Figure 5.6A) further supports disassembly of whole aggregates (see more below). Third, after disaggregation is complete, cpSRP43 remains in soluble complexes with dislodged LHCP monomers to prevent re-aggregation, and the extent of this equilibrium is

determined by the binding affinities between the LCHP monomer and cpSRP43. In contrast, Clp/Hsp100 ATPases use the threading mechanism for protein remodeling (6). Cycles of ATP binding and hydrolysis govern how the residues at the central pores of these hexameric rings bind or release substrates in a unidirectional manner, forcing polypeptides through small constrictions and possibly dislodging one aggregated molecule at a time (26, 27). Unfortunately, the relationship between the stability of the aggregates and the disaggregation rate has not been systematically investigated. Nevertheless, lessons learned from other members of the Clp/Hsp100 ATPases, such as the protease-associated ClpA and ClpX unfoldases, can provide valuable insights. For both unfoldases, it has been shown that the overall thermodynamic stability of the native protein substrates poorly correlates with the effectiveness of unfolding/degradation (28, 29). Rather, the local structure and the stability of the regions adjacent to the degradation tags have a dominant effect on unfolding (28–30). This observation corroborates with the proposed threading mechanism for the Clp/Hsp100 family, because these remodeling machines sequentially unravel their substrates from the recognition sites. This is in stark contrast with the case of cpSRP43, in which global stability of the aggregate correlates well with the rate of disaggregation (Figure 5.6A), suggesting that the two systems utilize distinct mechanisms in remodeling the protein/aggregate substrates. Instead of remodeling sequentially in parts like the Clp/Hsp100 proteins, cpSRP43 likely works on the whole aggregate at once.

### ***Perspective.***

The observation that cpSRP43 is only found in green plants suggests possible co-evolution with its substrates, the LHC proteins. Dedicated to the LHC protein transport,

cpSRP43 can establish extensive binding interactions with its substrate, sufficient to power its disaggregase activity. By the same token, it is intriguing that LHC proteins make the L18 accessible even when they form aggregates, allowing for binding to cpSRP43. Although the physiological significance of the disaggregase activity of cpSRP43 remains to be shown, it is conceivable that this activity is beneficial for cpSRP43, as it might be able to rescue unproductive membrane targeting or insertion of its substrates (31).

Recent works have reported protein-based, ATP-independent disaggregation activities in eukaryotic tissues (9, 10). Although the molecules responsible for these activities and their mechanism have not been identified, these findings suggest that cells have evolved methods to reverse protein aggregates that are alternative to the usage of ATP hydrolysis. As our work has shown, using binding energy to replace packing interactions inside the aggregates can be an efficient strategy for disaggregation. cpSRP43 now stands as a rare example of a class of disaggregases that can remodel and re-solubilize insoluble protein aggregates using only binding energy.

**Table 5.1** Description of the thermodynamic and kinetic parameter discussed in this study

Parameter	Definition	Assay	Reference <sup>1</sup>
$\theta_{\max}$	Maximal fraction disaggregated at saturating cpSRP43 concentration	Light scattering (Disaggregation)	3
$\theta_{4\mu\text{M}}^{\text{app}}$	Fraction disaggregated at sub-saturating cpSRP43 concentration (4 $\mu\text{M}$ )	Light scattering (Disaggregation)	3
$K_{\text{m}}^{\text{dis}}$	Average equilibrium dissociation constant for binding of cpSRP43 to LHCP aggregates	Light scattering (Disaggregation)	3
<b>n</b>	Hill coefficient	Light scattering (Disaggregation)	3
$k_{\max}$	Maximal disaggregation rate constant	Light scattering (Disaggregation)	4
$K_{\text{d}}^{\text{app}}$	Apparent dissociation constant for cpSRP43•LHCP complex	Light scattering (Prevention) and fluorescence anisotropy	1
$U_{50}$	Urea concentration for 50% re-solubilization of LHCP <sub>agg</sub>	Sedimentation	5

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

**Table 5.2** Summary of the thermodynamic parameters of the L18-binding mutants

<b>Construct</b>	$K_d^{app}$ (nM)	$K_m^{dis}$ ( $\mu$ M)	$\theta_{max}$	$\theta_{4\mu M}^{app}$	$\theta_{50\mu M}^{app}$
cpSRP43 R161A	1200 <sup>1</sup>	>50 <sup>2</sup>	N.D.	0.02	0.40
Lhcb5	10	4.1	1.06	0.51	1.06
Lhcb5 H160C	30	11.4	0.85	0.12	0.79
Lhcb5 L170C	1100	>50 <sup>2</sup>	N.D.	0.03	0.35

N.D. = not determined. Values reported are from Figure 1.

<sup>1</sup>previously determined by fluorescence anisotropy in (11)

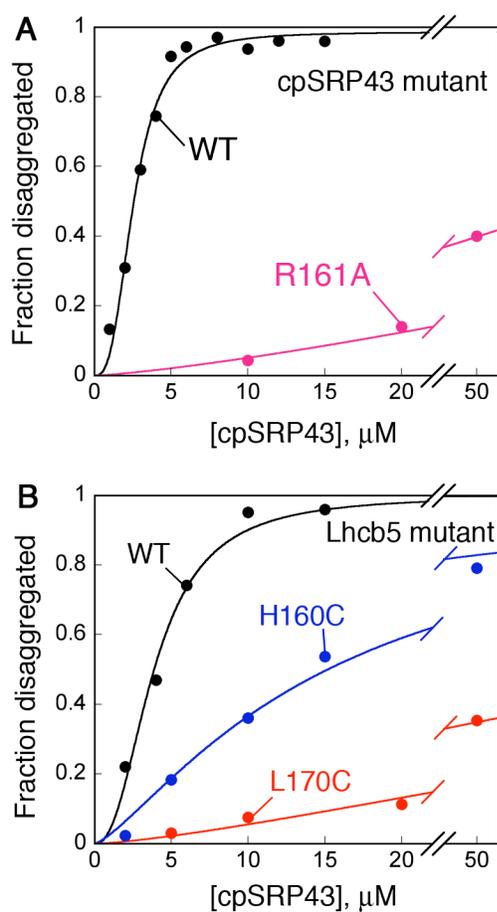
<sup>2</sup>denotes the values at the highest cpSRP43 concentration used

**Table 5.3** Summary of the thermodynamic and kinetic parameters of the LHCP TM mutants

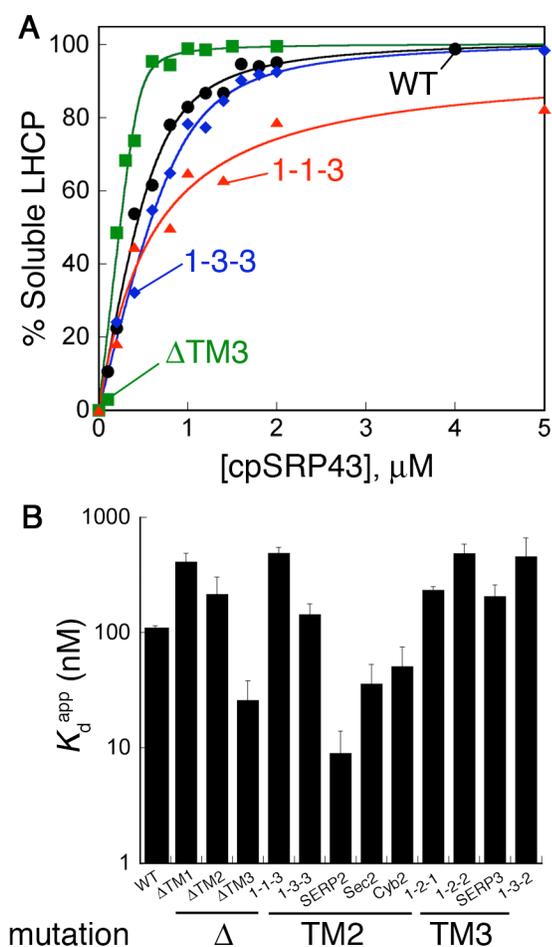
Construct	$\theta_{\max}$	$\theta_{4\mu\text{M}}^{\text{app}}$	$K_{\text{m}}^{\text{dis}}$ ( $\mu\text{M}$ )	n	$k_{\max}$ ( $\text{s}^{-1}$ )	$K_{\text{d}}^{\text{app}}$ (nM)	$U_{50}$ (M)
WT	0.99±0.03	0.70±0.06	2.6±0.2	2.5±0.1	0.048±0.011	111±3	3.8±0.2
1-3-3	1.00±0.05	0.72±0.10	2.8±0.4	2.5±0.6	0.046±0.000	144±34	3.7±0.1
SERP3	1.05±0.06	0.30±0.01	6.0±0.7	2.3±0.3	>0.030 <sup>1</sup>	207±51	4.0±0.1
1-2-1	0.88±0.01	0.02±0.00	12.8±0.9	3.4±0.0	0.015±0.000	234±16	4.4±0.3
$\Delta\text{TM1}$	0.91±0.00	0.09±0.04	8.4±0.8	3.1±0.3	>0.025 <sup>1</sup>	413±76	4.7±0.0
SERP2	1.04±0.04	0.95±0.03	1.3±0.1	2.2±0.4	0.109±0.006	9±5	3.5±0.1
$\Delta\text{TM3}$	0.98±0.02	0.95±0.03	1.7±0.1	4.1±0.0	0.065±0.003	26±12	2.5±0.1
Sec2	1.09±0.04	0.81±0.03	2.2±0.3	1.9±0.0	0.082±0.019	36±17	3.3±0.1
Cyb2	1.04±0.05	0.69±0.00	2.9±0.2	2.1±0.0	0.055	51±24	3.6±0.2
$\Delta\text{TM2}$	N.D.	0.01	>75 <sup>1</sup>	N.D.	0.003	216±88	5.7±0.1
1-2-2	N.D.	0.00	>75 <sup>1</sup>	N.D.	>0.002 <sup>1</sup>	489±95	4.7±0.1
1-3-2	N.D.	0.03	>60 <sup>1</sup>	N.D.	>0.004 <sup>1</sup>	456±206	4.8±0.1
1-1-3	N.D.	0.00	>60 <sup>1</sup>	N.D.	0.003	490±57	5.7±0.1

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

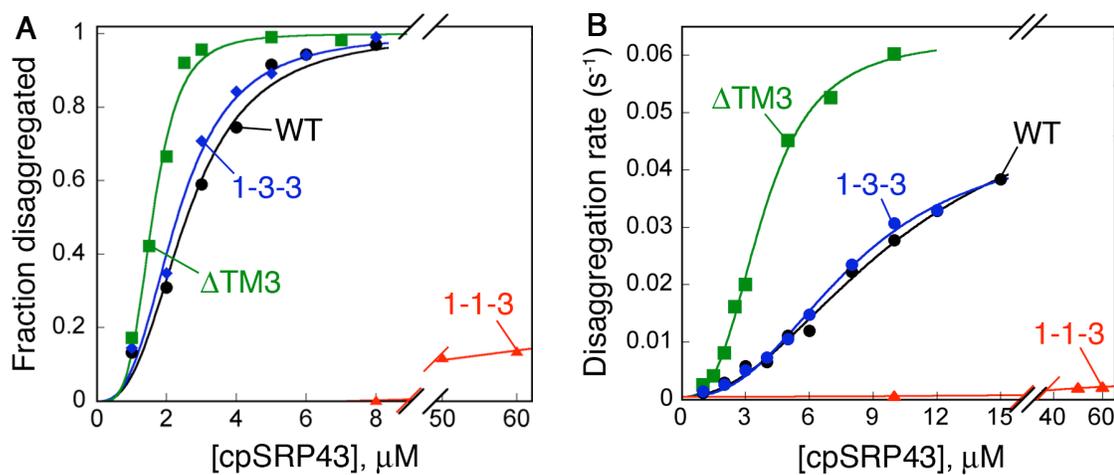
<sup>1</sup>denotes the values at the highest cpSRP43 concentration used



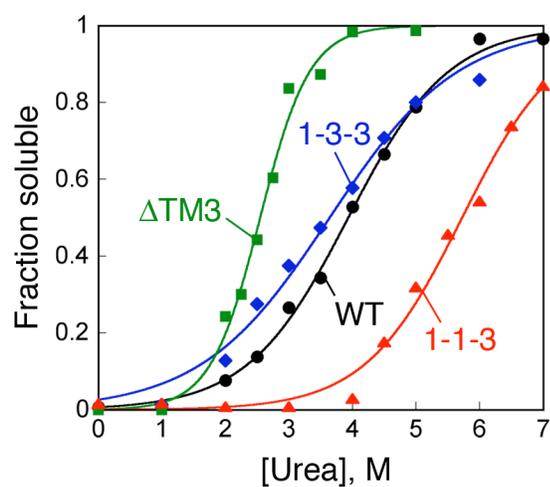
**Figure 5.1** L18 mutants uncouple initial binding from subsequent aggregate solubilization. (A) Concentration dependence for disaggregation of LHCP by wild-type (black) or R161A mutant (magenta) of cpSRP43. (B) Concentration dependence for disaggregation of Lhcb5 (black), Lhcb5-H160C (blue), and Lhcb5-L170C (red) by wild-type cpSRP43.



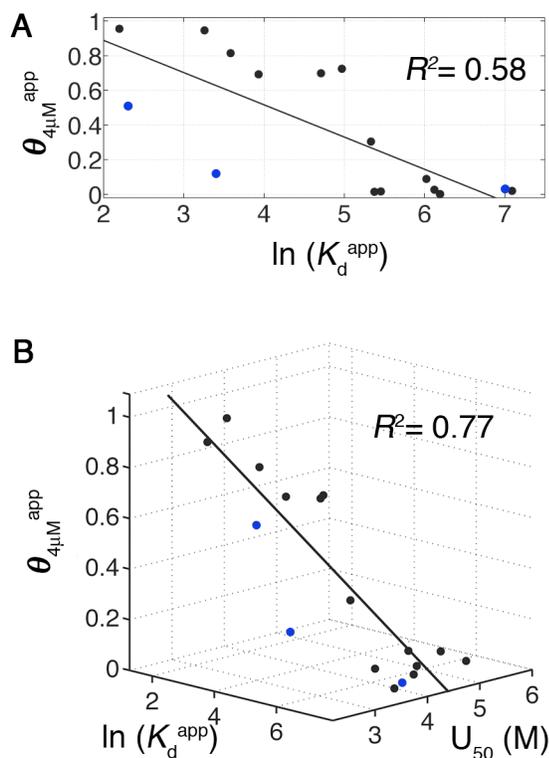
**Figure 5.2** cpSRP43 can interact with LHCP TM mutants. (A) Binding of cpSRP43 to LHCP and its TM mutants as measured by light scattering. The data were fit to eq 1 and gave  $K_d^{\text{app}}$  values of 109 nM for LHCP (black), 120 nM for 1-3-3 (blue), 13 nM for  $\Delta\text{TM3}$  (green), and 529 nM for 1-1-3 (red). (B)  $K_d^{\text{app}}$  values of the LHCP TM mutants in this study. Values are from Table 5.3. See Supplementary Table 5.S1 for nomenclature and description of the LHCP TM mutants.



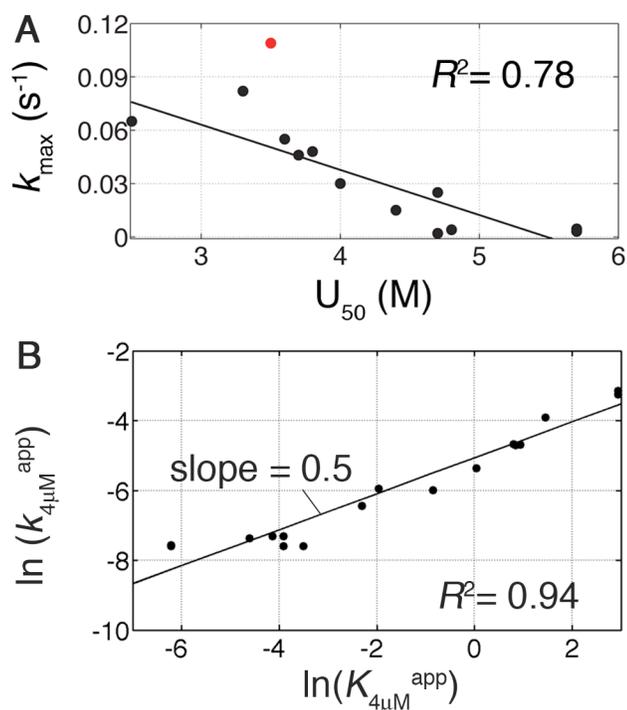
**Figure 5.3** LHCP TM mutants show different disintegration efficiencies. Concentration dependences of disintegration equilibrium (A) and rate (B) for LHCP (black), 1-3-3 (blue),  $\Delta\text{TM3}$  (green), and 1-1-3 (red). For A, the data were fit to eq 3 and gave  $\theta_{\text{max}}$ ,  $K_{\text{m}}^{\text{dis}}$  and  $n$  values.  $\theta_{4\mu\text{M}}^{\text{app}}$  values were also extracted from these plots. For B, the data were fit to eq 4 and gave  $k_{\text{max}}$  values. Thermodynamic and kinetic parameters were reported in Table 5.3.



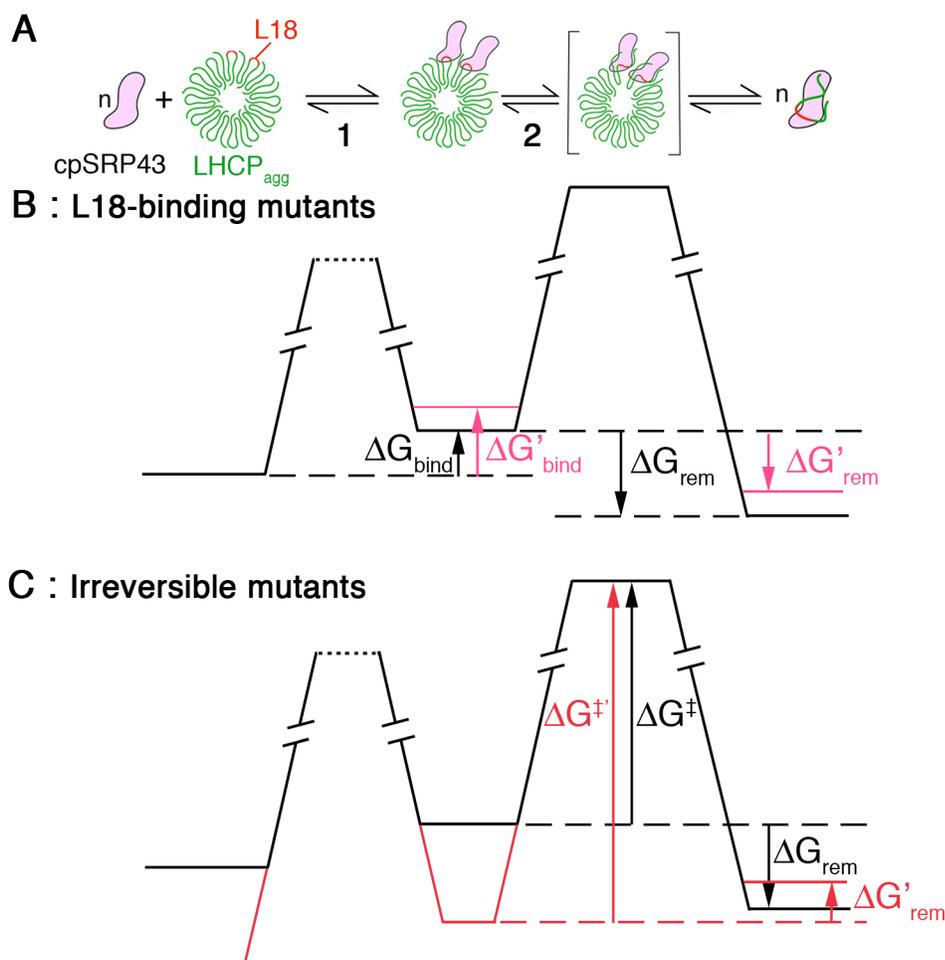
**Figure 5.4** The irreversible LHCP mutant aggregates are ultra-stable. Urea solubilization curves of LHCP and its TM mutants. The data were fit to eq 5 and gave  $U_{50}$  values of 3.9 M for LHCP (black), 3.6 M for 1-3-3 (blue), 2.5 M for  $\Delta$ TM3 (green), and 5.7 M for 1-1-3 (red).



**Figure 5.5** Stability of both the soluble cpSRP43•LHCP complex and the LHCP aggregates dictates disaggregation. (A) Correlation between the extent of disaggregation ( $\theta_{4\mu\text{M}}^{\text{app}}$ ) and the stability of the final soluble complex ( $K_d^{\text{app}}$ ). Values are from Table 5.3. The black line represents the regression fit ( $R^2=0.58$ ). (B) Correlation between the extent of disaggregation ( $\theta_{4\mu\text{M}}^{\text{app}}$ ) and the stability of the starting aggregate ( $U_{50}$ ) and the final soluble complex ( $K_d^{\text{app}}$ ). Values are from Table 5.3. The black line represents the regression fit ( $R^2=0.77$ ). Outliers highlighted in blue (Lhcb5, Lhcb5-H160C and Lhcb5-L170C) were included in the analyses.

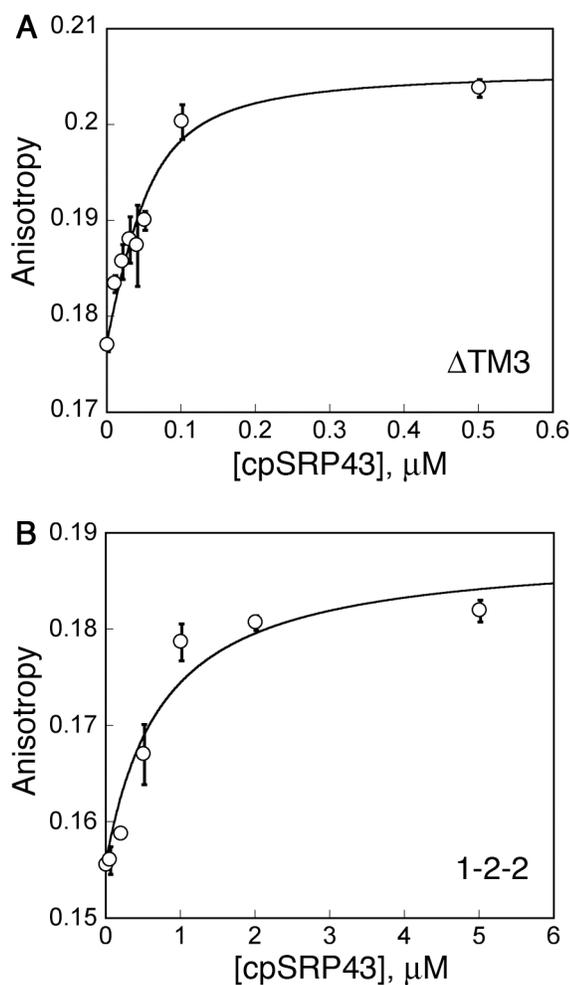


**Figure 5.6** The rate and transition state analyses of LHCP disaggregation. (A) Correlation between  $U_{50}$  and the maximal rate of disaggregation ( $k_{\max}$ ). Values are from Table 5.3. The black line represents the regression fit ( $R^2=0.78$ ). The red data point was excluded from the linear regression. (B) The  $\Phi$ -analysis of LHCP disaggregation. The values of  $K_{4\mu\text{M}}^{\text{app}}$  are calculated from  $\theta_{4\mu\text{M}}^{\text{app}}$  values. The  $k_{4\mu\text{M}}^{\text{app}}$  values are calculated from fits to eq 4. The black line represents a linear fit ( $R^2=0.94$ ) with a slope of 0.5.

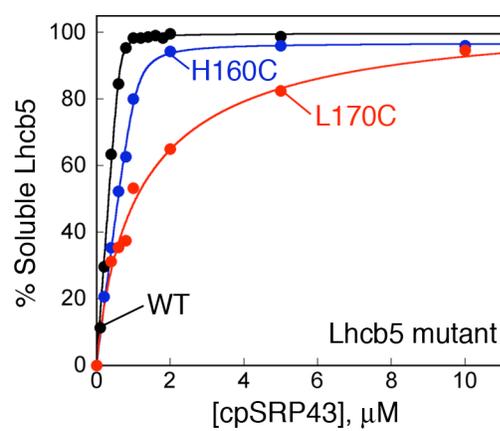


**Figure 5.7** (A) Working model for cpSRP43-mediated disaggregation. Step 1 depicts initial binding of cpSRP43 (magenta) to the LHCP aggregate (green) via the solvent-exposed L18 motif (red). Step 2 depicts subsequent remodeling and re-solubilization of the LHCP aggregate. Two molecules of cpSRP43 are depicted for simplicity. These cpSRP43 molecules cooperatively disrupt the aggregate packing and lead to complete solubilization. (B) and (C) Qualitative free-energy diagrams summarizing the effects of the mutants that disrupt each step. The L18-binding mutants (B) primarily disrupt the initial binding step. Due to its defect in binding to the aggregate via compromised L18 affinity, the free energy cost to proceed to the intermediate cpSRP43•LHCP<sub>agg</sub> complex ( $\Delta G_{\text{bind}}$ ) is enlarged compared to wildtype (magenta vs. black). However, this defect can be overcome by high concentration of cpSRP43 to drive binding. The irreversible mutants (C) mostly affect the remodeling step by creating a deep well at the intermediate step due to the high stability of their aggregates. This results in both a larger activation barrier for remodeling ( $\Delta G_{\text{rem}}$ , red vs. black) and a thermodynamically unfavorable reaction ( $\Delta G_{\text{rem}}$ , red vs. black). The figures are not drawn to scale.

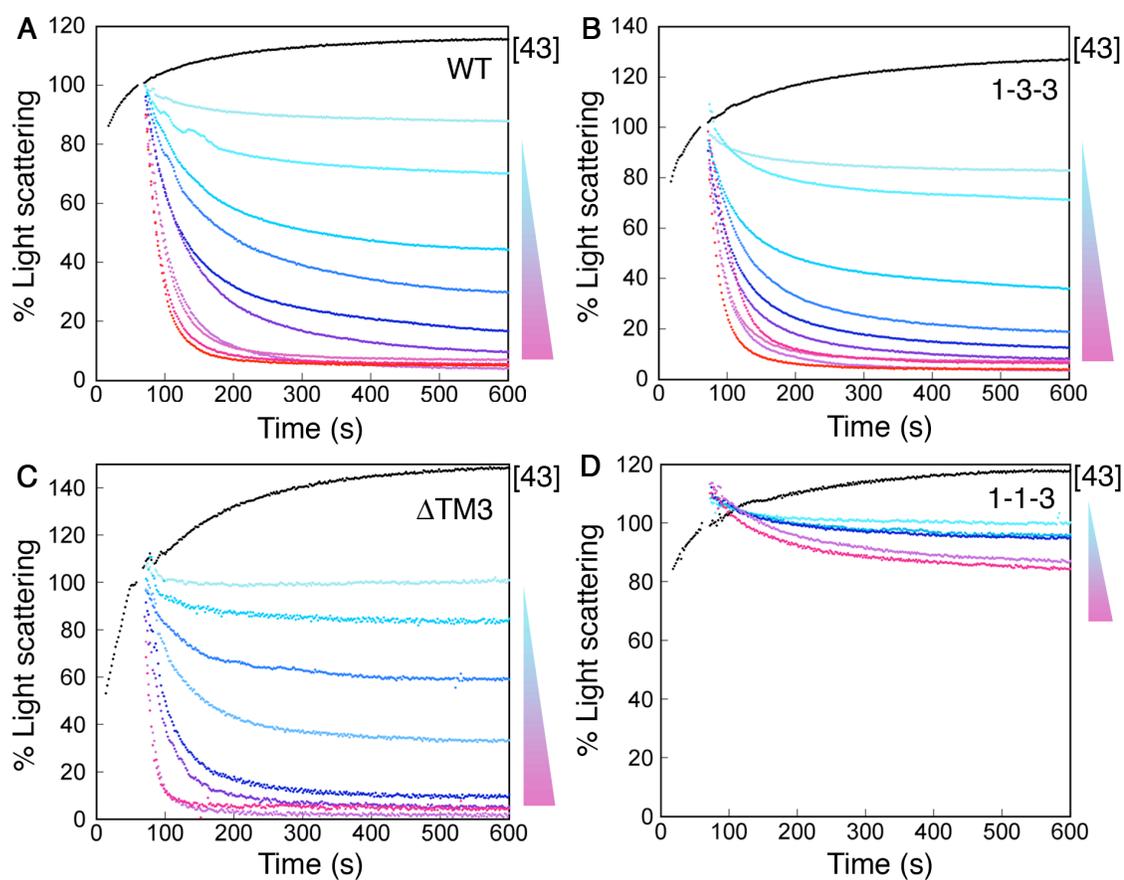




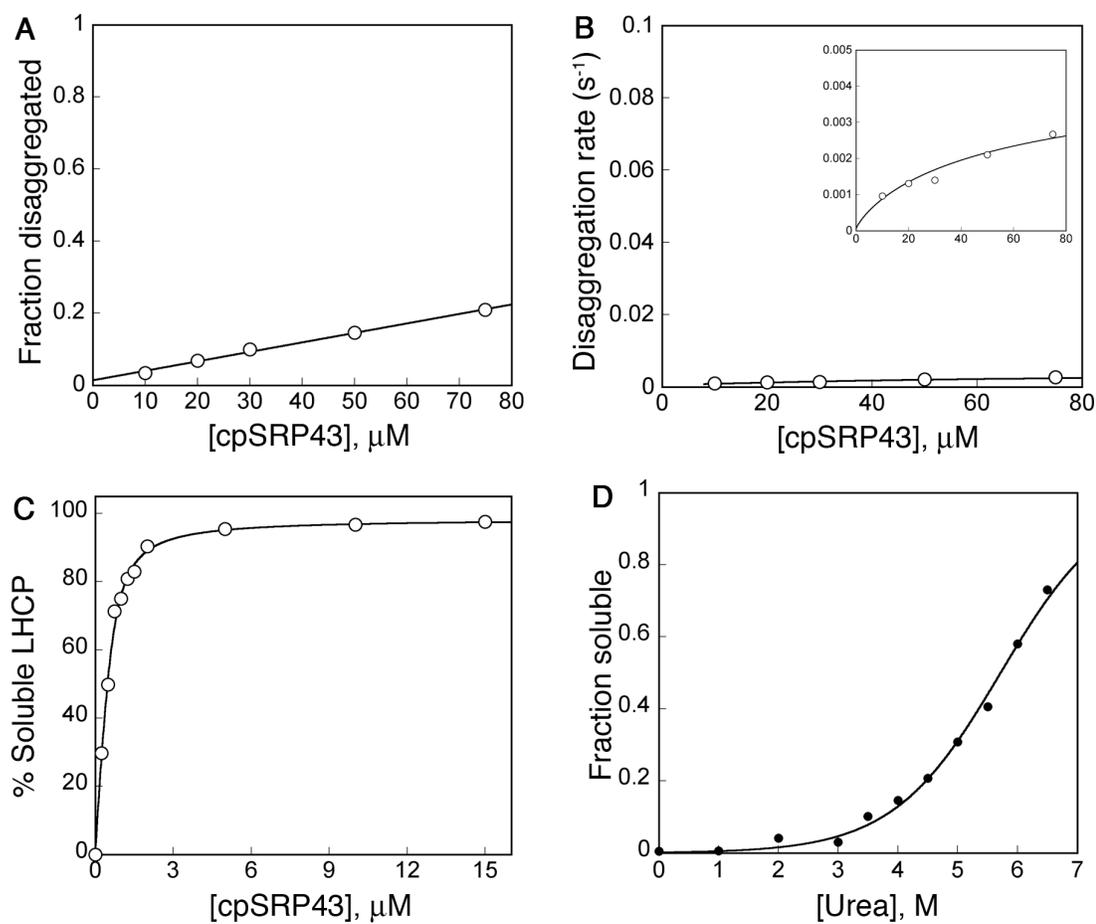
**Supplementary Figure 5.S1** The binding affinities from light scattering are comparable to those from fluorescent anisotropy. Binding of  $\Delta$ TM3 (A) and 1-2-2 (B) as measured by fluorescence anisotropy. The data were fit to a quadratic equation and gave  $K_d$  values of 22 nM for  $\Delta$ TM3 and 713 nM for 1-2-2. For comparison, the  $K_d^{\text{app}}$  values measured by light scattering were 26 nM and 489 nM, respectively (Table 5.3).



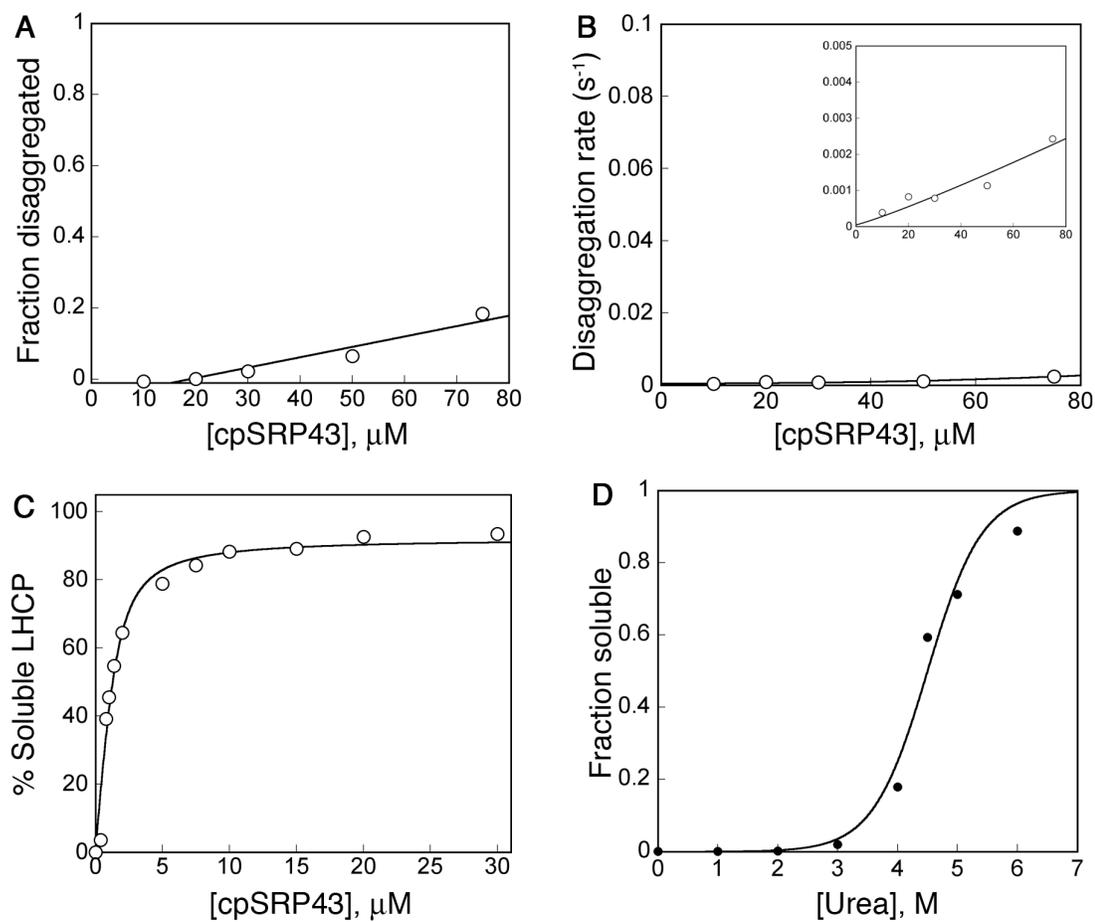
**Supplementary Figure 5.S2** Binding of cpSRP43 to Lhcb5 and its L18 mutants as measured by light scattering. The data were fit to eq 1 and gave  $K_d^{\text{app}}$  values of 10 nM for Lhcb5 (black), 30 nM for Lhcb5-H160C (blue), and 1.1  $\mu\text{M}$  for Lhcb5-L160C (red). The values are reported in Table 5.2.



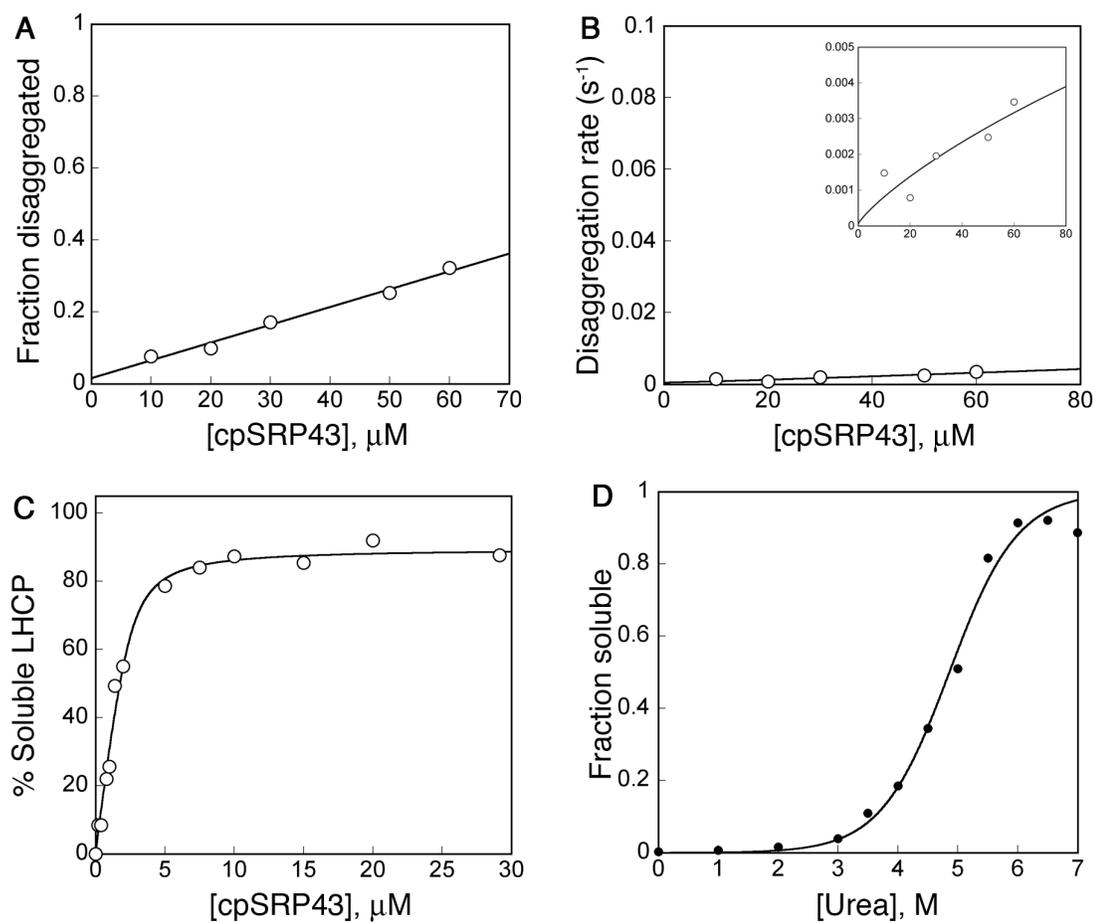
**Supplementary Figure 5.S3** Disaggregation time courses with varying concentrations of cpSRP43 for LHCP WT (1–15  $\mu$ M), 1-3-3 (1–15  $\mu$ M),  $\Delta$ TM3 (0.5–7  $\mu$ M), and 1-1-3 (10–60  $\mu$ M). These curves are representative of time courses derived from the disaggregation reactions of other LHCP TM mutants.



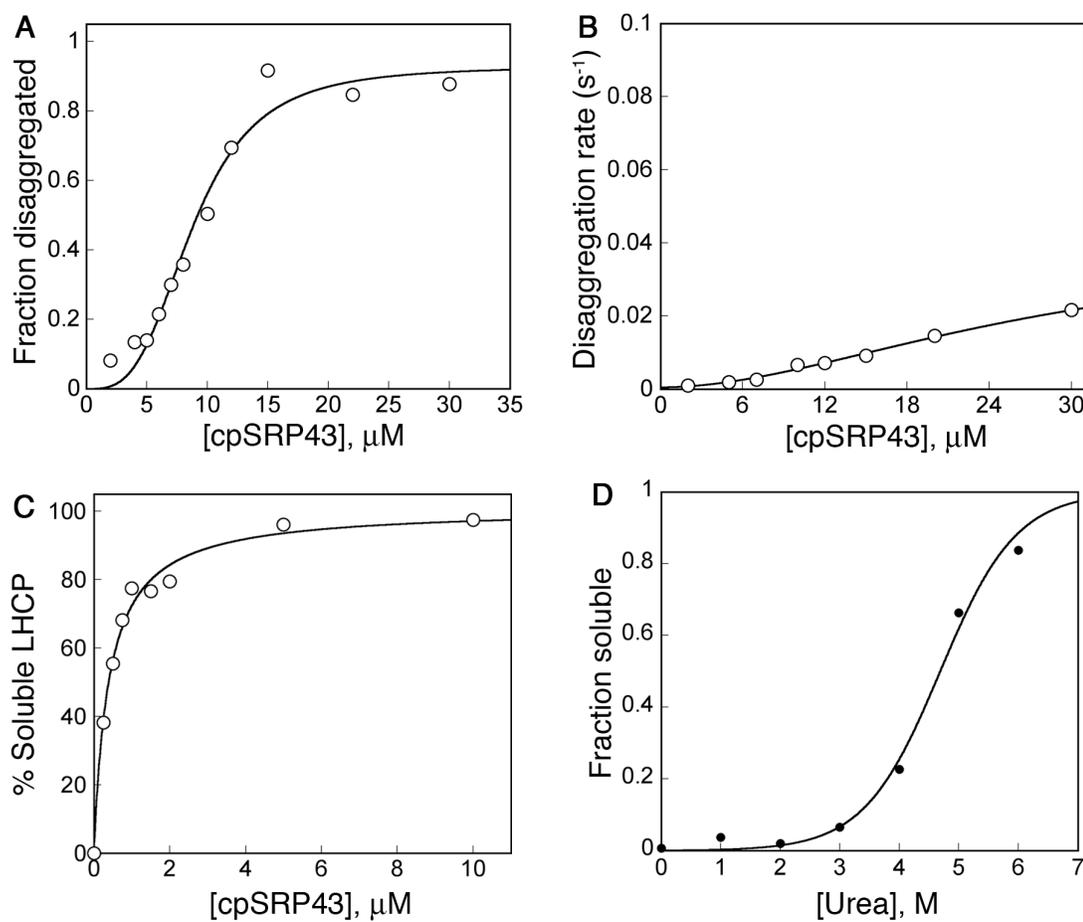
**Supplementary Figure 5.S4** Thermodynamic and kinetic analyses for  $\Delta$ TM2. (A) and (B) Concentration dependence of disaggregation equilibrium (A) and rate (B). (C) Binding of cpSRP43 to  $\Delta$ TM2. (D) Urea solubilization curve for  $\Delta$ TM2.



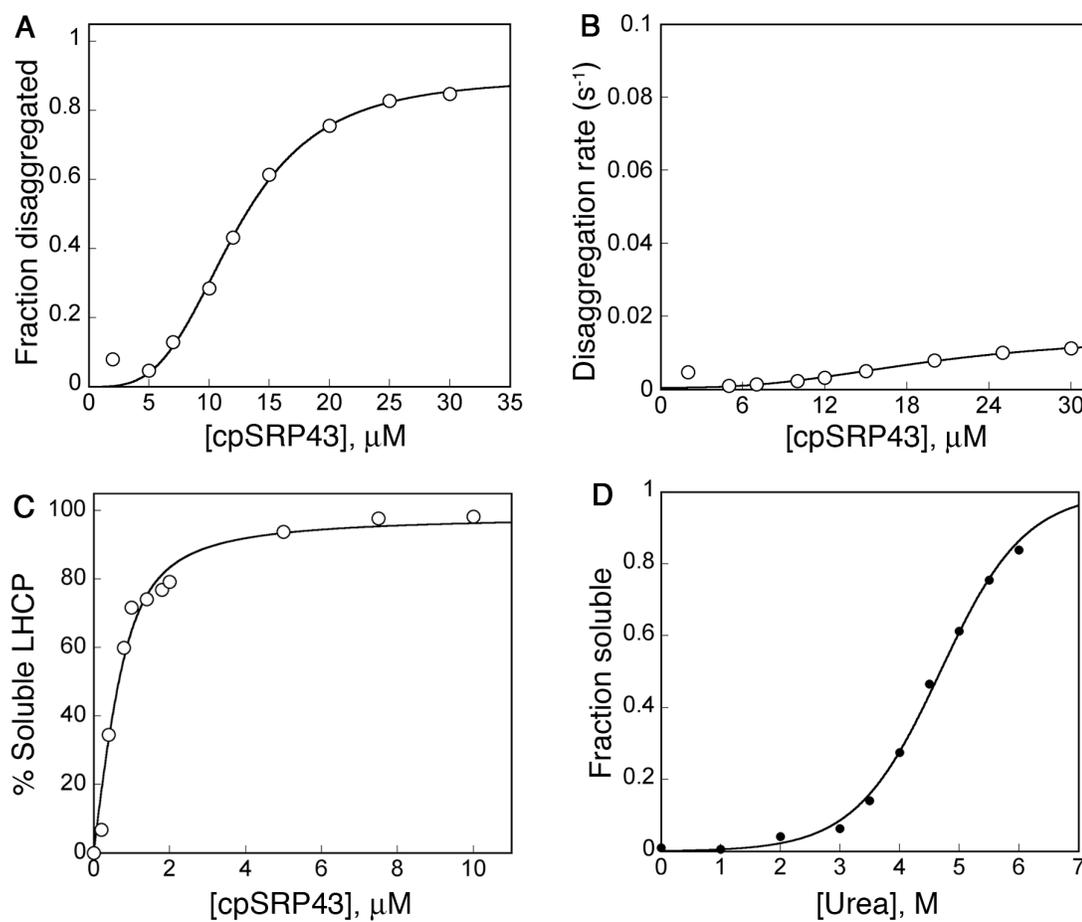
**Supplementary Figure 5.S5** Thermodynamic and kinetic analyses for 1-2-2. (A) and (B) Concentration dependence of disaggregation equilibrium (A) and rate (B). (C) Binding of cpSRP43 to 1-2-2. (D) Urea solubilization curve for 1-2-2.



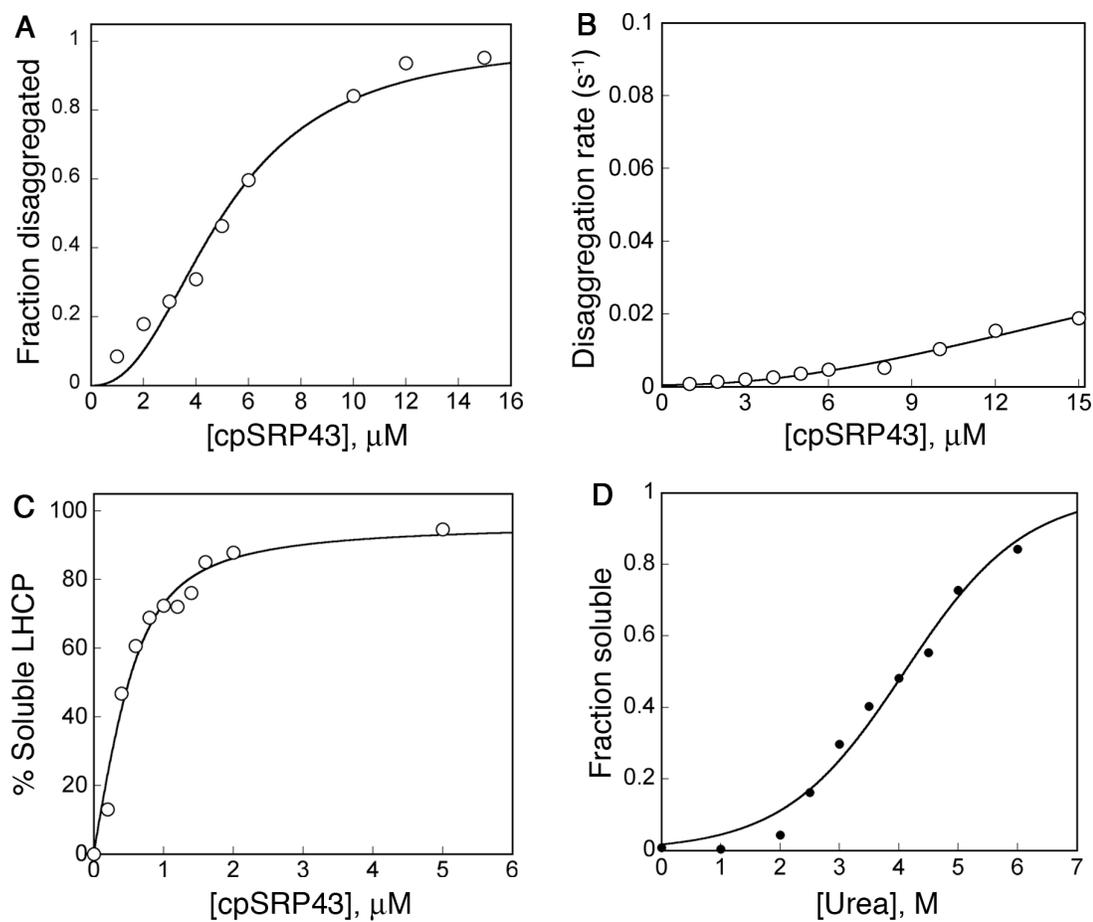
**Supplementary Figure 5.S6** Thermodynamic and kinetic analyses for 1-3-2. (A) and (B) Concentration dependence of disaggregation equilibrium (A) and rate (B). (C) Binding of cpSRP43 to 1-3-2. (D) Urea solubilization curve for 1-3-2.



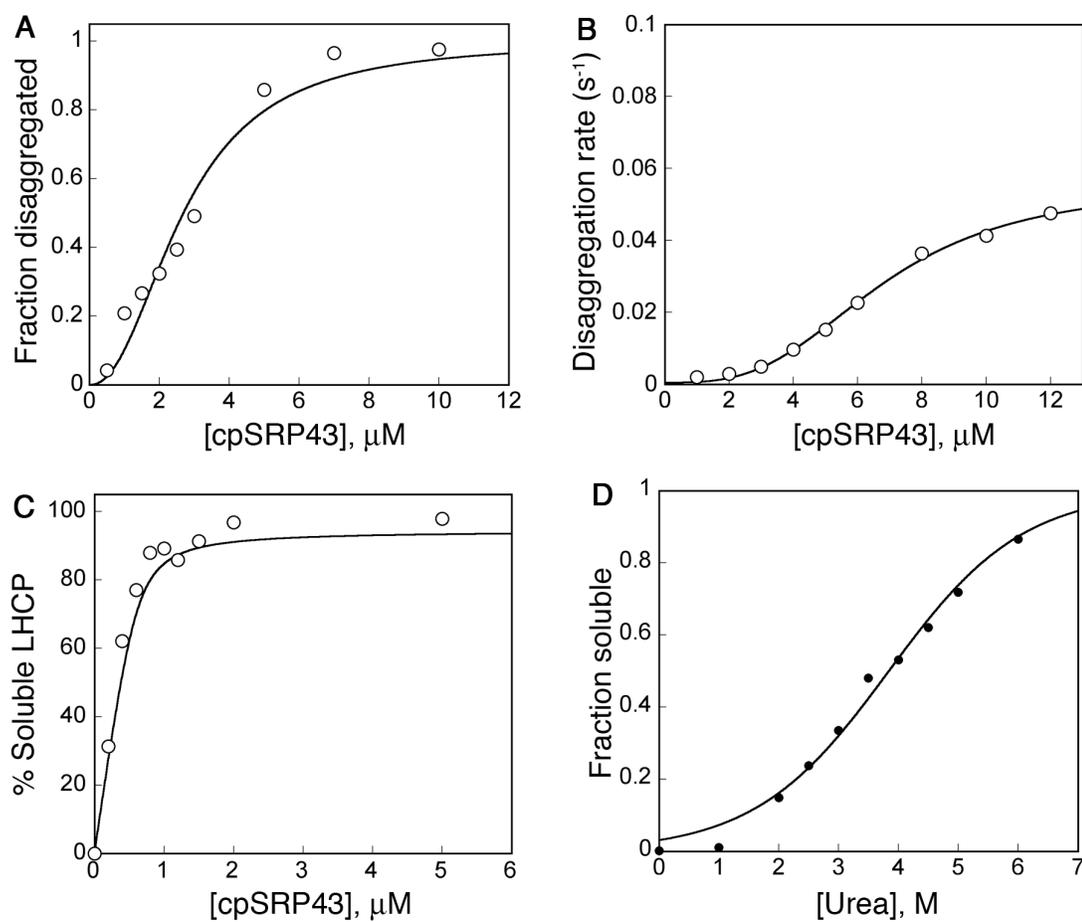
**Supplementary Figure 5.S7** Thermodynamic and kinetic analyses for  $\Delta$ TM1. (A) and (B) Concentration dependence of disaggregation equilibrium (A) and rate (B). (C) Binding of cpSRP43 to  $\Delta$ TM1. (D) Urea solubilization curve for  $\Delta$ TM1.



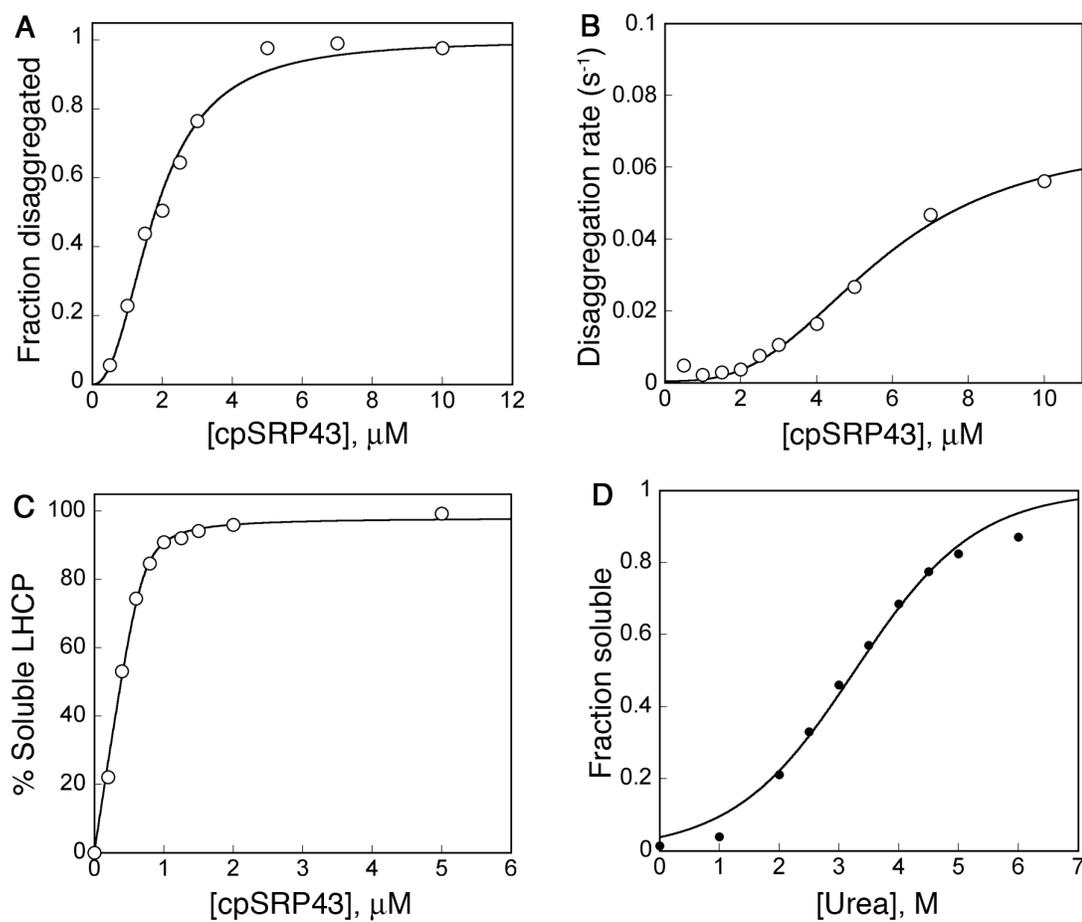
**Supplementary Figure 5.S8** Thermodynamic and kinetic analyses for 1-2-1. (A) and (B) Concentration dependence of disaggregation equilibrium (A) and rate (B). (C) Binding of cpSRP43 to 1-2-1. (D) Urea solubilization curve for 1-2-1.



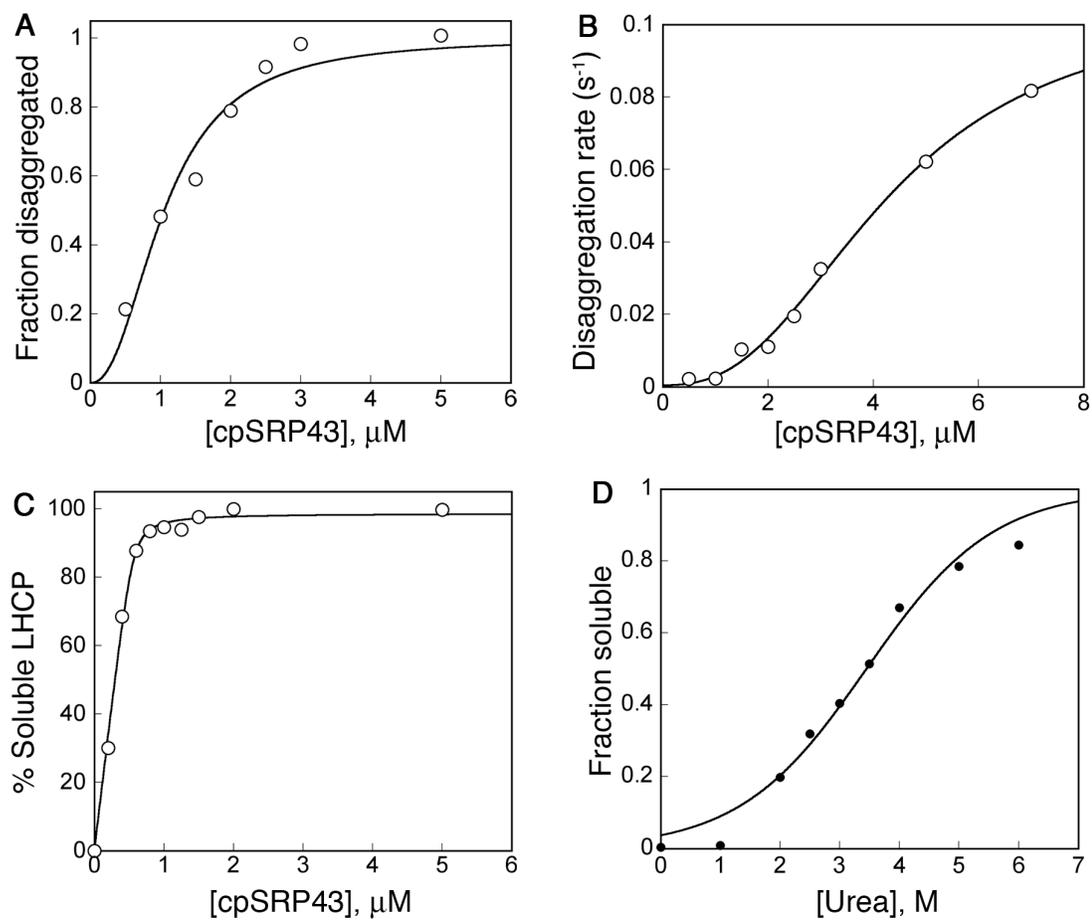
**Supplementary Figure 5.S9** Thermodynamic and kinetic analyses for SERP3. (A) and (B) Concentration dependence of disaggregation equilibrium (A) and rate (B). (C) Binding of cpSRP43 to SERP3. (D) Urea solubilization curve for SERP3.



**Supplementary Figure 5.S10** Thermodynamic and kinetic analyses for Cyb2. (A) and (B) Concentration dependence of disaggregation equilibrium (A) and rate (B and inset). (C) Binding of cpSRP43 to Cyb2. (D) Urea solubilization curve for Cyb2.



**Supplementary Figure 5.S11** Thermodynamic and kinetic analyses for Sec2. (A) and (B) Concentration dependence of disaggregation equilibrium (A) and rate (B and inset). (C) Binding of cpSRP43 to Sec2. (D) Urea solubilization curve for Sec2.



**Supplementary Figure 5.S12** Thermodynamic and kinetic analyses for SERP2. (A) and (B) Concentration dependence of disaggregation equilibrium (A) and rate (B and inset). (C) Binding of cpSRP43 to SERP2. (D) Urea solubilization curve for SERP2.

## References:

1. Balch, W. E., Morimoto, R.I., Dillin, A., and Kelly, J.F. (2008) Adapting proteostasis for disease intervention, *Science* 319, 916–919.
2. Hayer-Hartl, M., and Hartl, F.U. (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. 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.
7. Glover, J. R., and Lindquist, S. (1998) Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins, *Cell* 94, 73–82.
8. Goloubinoff, P., Mogk, A., Ben Zvi, A.P., Tomoyasu, T., and Bukau, B. (1999) Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network, *Proc. Natl. Acad. Sci. USA* 96, 13732–12737.
9. Bieschke, J., Cohen, E., Murray, A.N., Dillin, A., and Kelly, J.W. (2009) A kinetic assessment of the *C. elegans* amyloid disaggregation activity enables uncoupling of disassembly and proteolysis, *Protein Science* 18, 2231–2241.
10. 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.
11. 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.
12. Schuenemann, 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. USA* 95, 10312–10316.
13. Jansson, S. (1999) A guide to the Lhc genes and their relatives in arabidopsis, *Trends Plant Sci.* 4, 236–240.
14. 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.
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. 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.

17. 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.
18. 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, *The Plant Cell* 11, 87–99.
19. Eichacker, L. A., and Henry, R. (2001) Function of a chloroplast SRP in thylakoid protein export, *Biochim. Biophys. Acta* 1541, 120–134.
20. Jonas-Straube, E., Hutin, C., Hoffman, N.E., and Schuenemann D. (2001) Functional analysis of the protein-interacting domains of chloroplast SRP43, *J. Biol. Chem.* 276, 24654–24660.
21. Abramoff, M. D., Magelhaes, P.J., and Ram, S.J. (2004) Image Processing with ImageJ, *Biophotonics International* 11, 36–42.
22. Fersht, A. (1998) *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*, W.H. Freeman and Company, New York.
23. D'Errico, J. (2006) Polyfitn, N-d polynomial regression model.
24. 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.* Epub ahead of print (April 5, 2011).
25. Mogk, A., Dougan, D., Weibezahn, J., Schlieker, C., Turgay, K., and Bukau, B. (2004) Broad yet high substrate specificity: the challenge of AAA+ proteins, *J. Struct. Biol.* 146, 90–98.
26. 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.
27. 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.
28. 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.
29. Kenniston, J. A., Baker, T.A., Fernandez, J.M., and Sauer, R.T. (2003) Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine, *Cell* 114, 511–520.
30. 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.
31. Henry, R. L. (2010) SRP: Adapting to life in the chloroplast, *Nat Struct Mol Biol.* 17, 676–677.

