Mechanism of an ATP-independent protein disaggregase.
I. Structure of a membrane protein aggregate reveals a mechanism of recognition by its chaperone*

Thang X. Nguyen¹, Peera Jaru-Ampornpan¹, Vinh Q. Lam¹, Peigen Cao¹, Samantha Piszkiewicz¹, Sonja Hess², Shu-ou Shan¹

¹Division of Chemistry and Chemical Engineering and ²The Proteome Exploration Laboratory, Division of Biology, Beckman Institute, California Institute of Technology, Pasadena, CA 91125

*Running title: Structure of a membrane protein aggregate

To whom correspondence should be addressed: Shu-ou Shan, Division of Chemistry and Chemical Engineering, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA, USA; Tel.: (626) 395-3879; Fax: (626) 568-9430; E-mail: sshan@caltech.edu

Keywords: protein aggregation; LHC; disaggregate

Background: A novel chaperone, cpSRP43, recognizes and disassembles the aggregates formed by its client proteins.

Results: The client proteins of cpSRP43 form stable disc-shaped aggregates with the chaperone recognition motif displayed on the surface.

Conclusion: The surface-exposed motif on the aggregate allows it to be recognized by its chaperone.

Significance: Understanding the structure and energetics of protein aggregates provides insights into the mechanism of their disassembly.

SUMMARY

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

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

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

Previously, we identified an efficient disaggregase activity in the chloroplast Signal Recognition Particle 43 subunit (cpSRP43) (3). This provides an example in which a relatively small protein scaffold (38 kDa) can recognize and disrupt large protein aggregates in an ATP-independent mechanism (12), in contrast to the Clp/Hsp100 family of disaggregases. cpSRP43 is part of the protein targeting machinery, the cpSRP, that mediates the delivery of the light-harvesting chlorophyll a/b-binding (LHC) family of proteins to the thylakoid membrane (13-15). The most abundant member of the LHC family, LHCP, comprise ~30% of the proteins on the thylakoid membrane and is arguably the most abundant membrane protein on earth. The sheer abundance of these proteins and their highly hydrophobic nature demands highly effective chaperones that protect them from aggregation before arrival at the membrane. In the chloroplast stroma, this chaperone function is provided by cpSRP43, which effectively protects LHC proteins from aggregation and can even reverse preformed large LHC protein aggregates (12,16). cpSRP43 recognizes a highly-conserved 18-amino acid loop between the second and the third transmembrane (TM) domains of LHC proteins, termed L18 (17,18). In previous work, we showed that the specific interaction of cpSRP43 with the L18 motif is crucial for the chaperone and disaggregate activity of cpSRP43 (12). This and other observations led us to propose that, in the absence of external energy input, cpSRP43 uses specific binding energy with its substrate proteins to remodel and rescue LHC protein aggregation (12).

To gain insights into the molecular mechanism that underlies cpSRP43’s novel disaggregate activity, we need to first understand the nature of the LHC aggregate and identify the structural features that facilitate its disassembly by cpSRP43. To this end, we examined the nature and structure of the LHC aggregate using biophysical and biochemical techniques. We show that LHC proteins form disc-like particles with a relatively amorphous hydrophobic core, but exhibit a defined interior/exterior organization in which the L18 recognition motif is displayed on the solvent-exposed surface. This suggests an attractive mechanism for cpSRP43 to recognize the LHC aggregates and thus initiate their disassembly.

MATERIALS AND METHODS

Materials – LHCP, Lhcb5 and their mutants were purified under denaturing conditions as described (12), except that 6M GdmHCl was used instead of 8M urea for Lhcb5. αβ1-40 and re-crystallized thioflavin T (ThT) were generous gifts from Dr. J. W. Kelly. 1-anilino-8-naphthalene sulfonate (ANS) and bis-ANS were from Sigma and Invitrogen, respectively. n-dodecyl-N,N,N-dimethylamine-N-oxide (LDAO), n-dodecyl-β-D-maltopyranoside (DDM), n-octyl-β-D-glucopyranoside (β-OG) and n-nonyl-β-D-glucopyranoside (BNG) were from Anatrace. Triton X-100 (TX-100) was from Sigma and sodium dodecyl sulfate (SDS) was from BioRad. Urea and GdmHCl were molecular biology grade from MP and Sigma, respectively. 1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl methanethiosulfonate (MTSSL) was from Toronto Research Chemicals, N-ethyl-maleimide was from Sigma and N-(1-pyrene)-maleimide was from Invitrogen.

Light Scattering Assay – Light scattering experiments were performed as previously described (12). For formation of aggregates (Fig. 3, black), unfolded LHCP in 8 M urea was directly
diluted into Buffer D (50 mM KHEPES pH 7.5, 200 mM NaCl) to the desired final concentration; the final concentration of urea was equalized among different samples. The CMC is obtained as the x-intercept from the linear fit of the data (19). For serial dilution experiments (Fig. 3, red), the sample at 1 µM LHCP was serially diluted (by two-fold) into fresh buffer D and allowed 10–30 min to equilibrate before the measurement.

TEM – LHCP aggregates were formed by diluting unfolded LHCP in 8 M urea into Buffer D to the final concentration of 2 µM. After incubation at 25 °C for 5 minutes, the sample was diluted 5-fold and immediately deposited onto a glow-discharged 200-mesh Formvar grid (Ted Pella Inc., CA). After 45-second adsorption time, the grid was washed in water and then stained with 1% uranyl acetate for 45 seconds. TEM images were obtained on a 120 kV Tecnai T12 electron microscope coupled with a CCD camera. The diameters of the particles were measured using ImageJ (20).

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

Fluorescence – All fluorescence experiments were carried out in Buffer D using a Fluorolog 3-22 spectrofluorometer (Jobin Yvon). For bis-ANS experiments, 1 mM bis-ANS was added to Buffer D with or without 1 µM LHCP aggregate. The samples were excited at 395 nm and then scanned from 410 to 620 nm, with the excitation and emission band passes of 2 and 5 nm, respectively. For ThT experiments, 20 mM re-crystallized ThT was added to Buffer D containing no aggregate, aggregates from 1 or 5 µM LHCP, or 15 µM freshly sonicated Ap1-40 amyloid. The samples were excited at 440 nm and then scanned from 470 to 570 nm, with the excitation and emission band passes of 3 and 7 nm, respectively. For comparison, ThT fluorescence from 1 and 5 µM unfolded LHCP in 8 M urea were measured.

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

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

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

Sedimentation – Unfolded LHCP was diluted to 10 µM 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 µl of 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 boiled and visualized by SDS-PAGE.

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

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

**NEM Alkylation And MS Analysis** – Cysteine mutants of Lhcb5 in 6 M GdmHCl were reduced with 2.5 mM TCEP at RT for 2 hours. Each mutant was diluted into Buffer D to a final concentration of 3.3 µM and incubated on ice for 10 minutes to form the aggregate, followed by the addition of 100 µM NEM. The reaction was quenched with 50 mM DTT at various time points, concentrated under vacuum, redissolved in 0.2% formic acid, and ca. 25 pmol protein was analyzed on an LC-MSD SL 1100 series (Agilent). The samples were chromatographed on a 2.1 x 150 mm Zorbax 300SB-C3 column (Agilent) using a gradient consistent of 0.2% formic acid and 0.2% formic acid in acetonitrile (89.8%) and methanol (10%). Intact masses were measured in the single quadrupole and quantified using the software ChemStation software (Agilent). Control experiments where different ratios of un-alkylated and alkylated proteins were mixed and subjected to MS analysis shows the quantification of ratio of alkylated species to be reliable (Fig. 6E). The reported accessibilities were calculated as a ratio between the alkylation of each cysteine mutant under aggregation Buffer D versus denaturing 6 M GdmHCl.

**RESULTS**

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

We next used thioflavin T (ThT) to probe the structural organization of the LHCP aggregate. ThT is often used as a diagnostic for the formation of amyloid fibrils generated by amyloid-β (Aβ), α-synuclein, and other amyloidogenic proteins (25). Similar to bis-ANS, the fluorescence of ThT exhibited a significant increase in intensity and a blue shift in spectrum in the presence of the LHCP aggregate (Fig. 1B, blue lines). The extent of these fluorescence changes is comparable to that induced by mature amyloid fibrils generated by the Aβ1-40 peptide (Fig. 1B, red vs. blue and Fig. 1C). As microscopy analyses did not indicate fibril formation in the LHCP aggregate (see below), these results suggest that ThT is not highly specific for amyloid fibrils, consistent with recent work observing ThT fluorescence of non-fibrillar aggregates of β-lactoglobulin and transthyretin (19,26). Instead, this dye possibly binds to hydrophobic grooves that are often present in amyloid fibrils but can also be generated by other types of aggregates (27).

**LHCP Forms Stable Aggregates** – To probe the stability of the LHCP aggregate, we tested its solubility in various detergents, including LDAO, DDM, β-OG, BNG, and TX-100. By analyzing the amount of proteins in the soluble and insoluble fractions after medium-speed sedimentation (see...
Methods), we showed that none of these detergents was able to solubilize the LHCP aggregate at or above their respective concentrations typically used for membrane protein solubilization (Fig. 2A).

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

LHCP Forms Micellar, Disc-shaped Aggregates – Formation of large LHC aggregates can be monitored based on light scattering at 360 nm (12). The scattering intensity increases linearly with LHCP concentration above ~100 nM (Fig. 3, black), suggesting that aggregate formation was complete under these conditions. However, the linearity broke down at lower LHCP concentrations (Fig. 3 and inset, black). This was not due to limitations in instrument sensitivity: when pre-formed LHCP aggregates were diluted, linearity in light scattering intensity was observed at all concentrations and extrapolated through zero (Fig. 3 and inset, red). These observations show that: (i) the LHCP aggregate is kinetically stable and virtually irreversible once it has formed; and (ii) formation of the LHCP aggregate requires a critical protein concentration, reminiscent of the Critical Micellar Concentration during micelle formation. An analogous, ‘critical aggregate concentration’ of 125 nM was obtained for the LHCP aggregates from these data (see Methods). This micelle-like characteristic begins to suggest a globular morphology of the LHC aggregates.

To directly observe the global structure of LHC aggregates, we examined them using transmission electron microscopy (TEM) and atomic force microscopy (AFM). Negatively stained TEM images revealed LHCP aggregates to be circular particles (Fig. 4A,B). Analysis of the size of these particles resulted in a distribution that fits well to a Gaussian function, with diameters of 12±2 nm (Fig. 4C). Consistent with the EM images, AFM analysis also showed LHCP aggregates to be disc-shaped particles (Fig. 5A,B) with mean area of 214±94 nm² (Fig. 5B), or mean diameter of 16±5 nm, in good agreement with the EM measurements. Strikingly, the heights of the aggregates measured by AFM are ‘quantized’ and peaked at integrals of 0.7-0.8 nm (Fig. 5D and inset). These results suggest that LHC proteins form disc-shaped aggregates with a height of 0.7-0.8 nm, and these discs can further stack upon one another.

The L18 Recognition Motif is Displayed on the Aggregate Surface – To probe the global structure of LHC aggregates at higher resolution, we engineered 30 single-cysteine mutations, which span every 5-10 residues throughout the entire protein sequence of Lhcb5 (Fig. 6A, blue). Lhcb5 is a close homologue of LHCP (supplemental Fig. S1) that strongly depends on the cpSRP pathway for its biogenesis. Its aggregation is efficiently prevented and disassembled by cpSRP43 analogously to LHCP (12,28). All single-cysteine mutants were able to form light-scattering aggregates with the same extent and kinetics as wildtype Lhcb5, and thus can be used to probe the assembly of the wildtype aggregate (Fig. 6B).

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

Upon formation of the aggregate, the spin probes in all the TMs, loop 1 (between TM1 and TM2) and the C-terminus of Lhc5 displayed high \( \Delta H \) values and broad EPR spectra, suggesting that they are highly immobile and likely engaged in strong inter- or intramolecular interactions (Fig. 6D, supplemental Table 1 \( \Delta H^{-1} \) values are plotted). In contrast, spin probes placed in the L18 motif and the N-terminus of TM3 are highly mobile, indicating that these regions are free from any extensive interactions in the aggregate and are likely solvent-exposed (Fig. 6D). In addition, spin labels at the N-terminus of Lhc5 also displayed highly mobile spectra.

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

In agreement with the results of EPR measurements, the residues within the TMs exhibit low efficiency of alkylation, ranging from 20-40%, in contrast to the almost complete alkylation of the respective cysteines under denaturing conditions (Fig. 6G and supplemental Table 1). Residues in loop 1 and the C-terminus of the LHC protein exhibit slightly higher alkylation efficiency, 40-60%, indicating that these regions are partially buried in the aggregate but to a lesser extent than the TMs. In contrast, residues on the L18 loop and the N-terminal end of TM3 proximal to L18 are almost completely alkylated (90-100%), suggesting that these sites are highly solvent accessible and presented on the exterior of the aggregate. Finally, residues in the N-terminus of Lhc5 showed almost 100% reactivity, again demonstrating the exposure of this region on the aggregate surface.

Although the burial of TMs in the Lhc5 aggregate is expected due to their hydrophobic nature, the low mobility and inaccessibility of loop 1 and the C-terminus of Lhc5 were surprising. We therefore asked if the burial of these loop regions results from topological constraints imposed by the neighboring TMs, or from the inherent physicochemical property of the looping sequence. To address this question, we altered the location of L18 in Lhc5 by either swapping it with loop 1 (TM1-L18-TM2-TM3) to construct the LoopSwap (LS) mutant, or with the protein sequence C-terminal to TM3 (TM1-TM2-TM3-L18) to construct the Cterm mutant. The aggregate formed by both constructs can be rescued by cpSRP43 (Fig. 7A), suggesting that cpSRP43, despite its specific interaction with L18, can tolerate variations in the remainder of its substrate’s sequence.

To probe the accessibility of individual motifs in the L18-swapped mutants, we probed the accessibility of engineered single cysteines in each domain by NEM alkylation. The alkylation efficiency of each motif in both L18-swapped mutants is similar to that of wildtype Lhc5: the L18 motif is highly accessible and almost completely alkylated, whereas loop 1 and the C-terminus regions exhibit medium levels of alkylation (Fig. 7B-D). These results indicate that the intrinsic properties of these sequences/domains determine their accessibility in the aggregate, and demonstrate that the L18 motif has a strong tendency to be displayed on the surface of protein aggregates.

LHC Aggregates Contain an Amorphous Hydrophobic Core—To determine whether the TMs of the LHC protein make specific intermolecular contacts in the buried core of the aggregate, we exploited the ability of pyrene labels to form excited-state dimers (excimers) when they are within 4-10 Å of each other. High pyrene excimer fluorescence reports on close proximity between specific sites within the aggregate. To this end, we mixed two proteins, each labeled with pyrene at a single cysteine residue, in all pair-wise combinations, allowed them to form the aggregates, and monitored for pyrene excimer fluorescence at 455 nm relative to the monomer fluorescence at 375 nm (see Methods). As a
positive control, we used a double-cysteine mutant V139C-L140C and labeled both positions with pyrene probes. The excimer ratio of this construct was high, ~0.8 (Fig. 8A). As the negative control, pyrene labeled L180C was used (excimer ratio = 0.08; Fig. 8A).

Two important lessons were learned from the results of these measurements. First, many pyrene pairs exhibit excimer fluorescence intensities substantially above the background and above the other pyrene pairs, with excimer ratios of ≥0.30 for homo-pyrene pairs at multiple residues in TM1 (67, 72, 88, 92) and for multiple hetero-pyrene pairs in all three TMs (Fig. 8B). Second, these excimer fluorescence intensities were still modest, up to 0.36 (Fig. 8B). These values are far below the values of 0.6-0.8 expected for specifically interacting pairs that are always in close proximity. These data indicate that in the LHC aggregate, the TMs form extensive intermolecular contacts in its hydrophobic core, but these interactions are much less specific than those observed in amyloid fibrils (29).

Taken together, the results demonstrate that: (i) LHC proteins form highly stable, disc-shaped aggregates; (ii) despite the possibly amorphous nature of the LHC aggregate core, it contains a defined global organization that can be reliably probed; and (iii) the L18 motif, among other regions of the LHC protein, is displayed on the surface of the aggregate and thus poised for interactions with cpSRP43 (Fig. 9). These results suggest an attractive model in which cpSRP43 could recognize the L18 motif presented on the surface of the aggregate, providing a starting point for its action as a disaggregate. Further, the exposure of the N-terminus and the N-terminal end of TM3 suggests additional potential interaction sites with cpSRP43 during aggregate recognition.

DISCUSSION

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

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

The morphology of the LHC aggregates bears resemblance to those of the soluble oligomeric intermediates that often precede amyloid fibril formation, which are disc-shaped, 9–25 nm in diameter, and 2–3 nm in height (32,33). Although earlier work tends to categorically describe these protein aggregates as ‘amorphous’, accumulating data suggest that there are nonetheless degrees of organization in some of these aggregates (34,35). For instance, the folding intermediates of bovine growth hormone, phosphoglycerate kinase, P22 tailspike and coat proteins participate in specific intermolecular interactions in their aggregation pathways (36-38). Likewise, although highly specific intermolecular interactions have not been detected in the LHC aggregates, more detailed analyses at the individual residue level provide convincing evidence that LHC aggregates have a defined ‘interior’ and ‘exterior’ that can be reliably probed, arguing against complete disorder in these aggregates.
What features of the LHC aggregate allow cpSRP43 to recognize it and initiate the disaggregation process? Answers to this question are central to understanding the mechanism of cpSRP43’s disaggregase activity. The results here strongly suggest that the formation of LHC aggregates is driven largely by hydrophobic collapse to bury its three TM domains. Importantly, we showed that the N-terminus of the LHC protein, the L18 motif, and the N-terminal segment of TM3 are displayed on the solvent-accessible surface when LHC proteins form aggregates. As the L18 motif is the primary recognition element for cpSRP43, its presentation on the exterior of the aggregate provides a very attractive mechanism by which cpSRP43 could recognize and anchor onto the aggregate to initiate the disassembly process (Fig. 9). Conceivably, the N-terminal fragment of TM3 proximal to the L18 motif could also contribute to this initial recognition, as previous work has detected crosslinks between cpSRP43 and residues at the N-terminus of TM3 (39). This and additional mutational studies suggest that TM3 is a likely candidate for cpSRP43 to initiate disruptions of the internal packing within the LHC aggregate (Jaru-Ampornpan et al., accompanying manuscript).

Further, the L18-swap experiments show that the L18 motif is a dominant sequence element that has a strong tendency to be displayed on the surface of a protein aggregate, likely due to its relatively polar amino acid composition and high propensity for disorder (Fig. 7). Considering that cpSRP43 has co-evolved with and is dedicated to the chaperoning of the LHC family of proteins, it is intriguing that the latter evolved a polar L18 recognition motif and made it accessible even when they form aggregates, which would enable cpSRP43 to readily recognize the aggregated LHC proteins. Although the physiological significance of cpSRP43’s disaggregase activity remains to be directly established, this observation is consistent with the possibility that this activity is beneficial, as it would enable cpSRP43 to rescue aggregated, off-pathway intermediates during the targeting or insertion of its substrate proteins (40).

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

On the other hand, the L18-swap experiments here and additional mutagenesis studies (Jaru-Ampornpan et al., accompanying manuscript) strongly suggest that the interaction of cpSRP43 with the remainder of the LHC protein, apart from L18, is highly adaptable, as a wide range of unnatural substrates can be effectively bound, chaperoned, and rescued by cpSRP43.

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

Structure of a membrane protein aggregate

20. Abramoff, M. D., Magelhaes, P.J. and Ram, S.J. (2004) Image processing with ImageJ. Biophotonics Int. 11, 36-42


Structure of a membrane protein aggregate

Acknowledgements—We thank A.N. Murray and Dr. J.W. Kelly for re-crystallized ThT and fibrilized Aβ1-40, A. McDowall for EM analyses, members of the Proteome Exploration Laboratory for mass spectrometry assistance, and members of the Shan group for helpful comments on the manuscript.

FOOTNOTES

* S.S. was supported by the David and Lucile Packard Fellowship in science and engineering, the Henry Dreyfus teacher-scholar award, and the BIG award from the American Federation for Aging Research. T.X.N. was supported by a grant from the Betty and Gordon Moore Foundation. AFM studies were supported by the Department of Energy grant DE-FG03-01ER46175 to J. Heath. The Proteome Exploration Laboratory is supported by the Betty and Gordon Moore Foundation and the Beckman Institute.

The abbreviations used are: cpSRP, chloroplast Signal Recognition Particle; LHC, light-harvesting chlorophyll a/b-binding; AAA+, ATPases associated with various cellular activities; TM, transmembrane; Aβ, amyloid beta; TEM, transmission electron microscopy; AFM, atomic force microscopy; EPR, electron paramagnetic resonance; and LC-MSD, tandem liquid chromatography and mass spectrometry.

FIGURE LEGENDS

FIGURE 1. LHCP aggregates contain exposed hydrophobic surfaces as detected by small molecule dyes. A, Fluorescence emission spectra of 1 mM bis-ANS with (blue) or without (black) 1 µM LHCP aggregate. B, Fluorescence spectra of 20 mM ThT in the absence (black) and presence of 1 (light blue) or 5 (dark blue) µM LHCP aggregate, or 15 µM Aβ1-40 (red). C, Quantification of the ThT fluorescence change at 484 nm per µM of protein (aggregate).

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

FIGURE 3. LHCP forms aggregates after a critical concentration. Light scattering intensities during formation of the aggregate (black) are compared to those from serial dilution of pre-formed aggregates (red). The inset highlights the lag phase at low concentrations during formation of the aggregate.

FIGURE 4. TEM analysis of LHCP aggregates. A, Large field view of a negatively stained TEM image of LHCP aggregates. B, A zoomed-in image shows that LHCP aggregates are round particles. C, Size distribution of the LHCP aggregates, measured from several independent experiments. The mean diameter is 12±2 nm.

FIGURE 5. AFM analysis of LHCP aggregates. A, Large field view of AFM topographic image showing well-separated LHCP aggregates. Large clusters are occasionally observed. The scale bar is 500 nm. B, A zoomed-in region of the image reveals disc-shaped particles. The lines
indicate particles whose heights were measured (red, blue, and green). The scale bar is 100 nm. C. Size distribution of LHCP aggregates, measured from several regions on the surface. The red line is a Gaussian fit to the data, which gave a mean area of the particle of 214 nm$^2$. D, Height distribution of LHCP aggregates shows three population of 0.8, 1.4 and 2.1 nm. The inset shows the height profiles for the representative particles indicated in (B). Curves are vertically displaced for clarity.

**FIGURE 6.** Mapping the LHC aggregates reveals exposed motif. A, Lhcb5 sequence, with residues mutated to cysteine in blue, the TM s underlined in green, and the L18 peptide underlined red. B, Light scattering from aggregates (A$_{360}$) 5 min after dilution of each single-cysteine mutants into Buffer D, final 1µM Lhcb5 proteins. Values are relative to that of wildtype protein. C, Representative EPR spectra of the spin probes placed at buried site L130C of TM2 (green) and exposed site 170C of L18 (red) upon Lhcb5 aggregation, and at 170C when Lhcb5 is solubilized in 6 M GdmHCl (black). D, Summary of the mobility of different residues in the Lhcb5 aggregate, reported in values of $\Delta H^{1}$. Residues in the TM s are in green, L18 in red, and the remainder of Lhcb5 in black. E, control experiment shows intact mass spectrometry can be used for the quantification of the efficiency of NEM alkylation. Wildtype Lhcb5, which contains one native cysteine (C100), was reacted with NEM to completion in 6 M GdmHCl. Different known ratios of the NEM-modified Lhcb5 were mixed with unreacted protein and submitted for MS analysis. F, Time courses for the alkylation reactions of representative cysteines at residues L130C, L170C in Lhcb5 aggregates, and at residue L170C when Lhcb5 is dissolved in 6 M GdmHCl. G, Summary of NEM accessibility of the single-site cysteines in the Lhcb5 aggregate. Color scheme is the same as in (C).

**FIGURE 7.** L18 has a strong tendency to be exposed on the surface of the aggregate. A, The extent of wildtype and mutant Lhcb5 aggregates (1 µM) resolubilized by 10 (white), 20 (gray), and 40 (black) µM cpSRP43. The aggregates formed by L18-swapped mutants TM1-L18-TM2-TM3 (LS) and TM1-TM2-TM3-L18 (Cterm) can be rescued by cpSRP43, although LS mutant required a higher concentration of cpSRP43. B, NEM accessibility analysis of residues in the wildtype aggregates. Regions probed include the N-terminus (G50C), TM1 (A80C), loop 1 (N120C), L18 (G162C) and C-terminus (A230C). NEM accessibility analysis of the same residues (C) in LS construct and (D) Cterm construct.

**FIGURE 8.** Analysis of pyrene excimer fluorescence reveals an amorphous aggregate core. A, representative pyrene spectra for the homo-pair L180C were recorded when it was solubilized in 6 M GdmHCl (black trace). Pyrene-labeled homo-pair A67C and L180C after aggregate formation in Buffer D shows excimer fluorescence (red and blue traces, respectively). Doubly pyrene-labeled Lhcb5 mutant V139C-L140C in Buffer D served as positive control for pyrene excimer fluorescence between residues in close proximity (orange trace). B, the chart lists pyrene excimer fluorescence ($I_{445}/I_{376}$) for all pair-wise combinations (see Methods); excimer fluorescence values between 0.21 and 0.29 are highlighted in yellow, and values $\geq 0.3$ are in orange.

**FIGURE 9.** Model for the global organization of LHC proteins in the aggregate. L18 is in red, the TM s are in different shades of green, and the other looping sequences are in black. The shaded region depicts the buried core of the aggregate. For clarity, only one LHCP molecule in the aggregate is highlighted.
Structure of a membrane protein aggregate

Figures 1

<table>
<thead>
<tr>
<th>Protein</th>
<th>Relative F_rel (per µM protein)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>1</td>
</tr>
<tr>
<td>LHCP in 8M urea</td>
<td>1.2</td>
</tr>
<tr>
<td>LHCP <em>reg</em></td>
<td>6.8</td>
</tr>
<tr>
<td>Aβ1-40</td>
<td>5.2</td>
</tr>
</tbody>
</table>
### Figure 2

<table>
<thead>
<tr>
<th></th>
<th>0.1%</th>
<th>0.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>P</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>P</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1%</th>
<th>1.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>P</td>
<td>S</td>
</tr>
</tbody>
</table>

**LDAO**

<table>
<thead>
<tr>
<th></th>
<th>0.8%</th>
<th>1.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDM</td>
<td>P</td>
<td>S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1%</th>
<th>1.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>TX-100</td>
<td>P</td>
<td>S</td>
</tr>
</tbody>
</table>

**β-OG**

<table>
<thead>
<tr>
<th></th>
<th>0.8%</th>
<th>1.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNG</td>
<td>P</td>
<td>S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1%</th>
<th>2%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDS</td>
<td>P</td>
<td>S</td>
</tr>
</tbody>
</table>

**100°C RT**

**2% SDS soluble fraction**

### Figure 3

#### A

<table>
<thead>
<tr>
<th>$A_{590}(AU)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>[LHCP], μM</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
</tbody>
</table>

| 0.2         |
| 0.4         |
| 0.6         |
| 0.8         |
| 1           |

Inset graph: 
- Red line
- Green line
- Blue line

Point data: 
- Red circle
- Green square
- Blue triangle
Figure 4

A

B

C

Structure of a membrane protein aggregate
Structure of a membrane protein aggregate

Figure 5

A

B

C

D
Structure of a membrane protein aggregate

Figure 6

A

KAPAPAKSA VSETSDELAK WYGFDRIIFL PGGILDASAIE PEYLNGEVAQ
DGYDPFGILG KKRENFAKVR AFELHARAGA MLGAAGFIIP RALNKCANC
GPEAWFRTG ALLOGOQIN YPHKNIPILN VLAHAPVVL LNGAEYPRF
NGLLFFEKRH PGGFDNIGL EKDEFCHLRL RRKKEANGRL AMGRMLTFT
QAYTVGEQFV ENLAKHLSDE FQNNLITYA GTRERTPL

B

Extent of Aggregation (mol A)

Lhcb5 Mutants

C

Magnetic Field (Gauss)

3475 3500 3525 3550

D

HMH

3.3

4.2

3.02

AE L170C

AE = 1.74

AE = 1.74

AE = 3.02

AE L170C

AE L180C

E

M/S Reactiv (Reactions % Reacted)

Standard Input (% Reacted)

F

Fraction Accessible in GdmCl

0.2

0.4

0.6

0.8

1

0

5

10

G

Fraction Accessible

0.1

0.2

0.3

0.4

0.5

Residue Number

N

C

Downloaded from http://www.jbc.org/ by guest on September 1, 2017
Structure of a membrane protein aggregate

Figure 7

A

Percentage Disaggregation

W1  LS  Clein

B

Fraction Accessible

N-terminaL  TM1  Loop1  L18  C-terminaL

C

Fraction Accessible

N-terminaL  TM1  Loop1  L18  C-terminaL

D

Fraction Accessible

N-terminaL  TM1  Loop1  L18  C-terminaL
Mechanism of an ATP-independent protein disaggregate. I. Structure of a membrane protein aggregate reveals a mechanism of recognition by its chaperone
Thang X. Nguyen, Peera Jaru-Ampornpan, Vinh Q. Lam, Peigen Cao, Samantha Piszkiewicz, Sonja Hess and Shu-ou Shan

*J. Biol. Chem.* published online March 22, 2013

Access the most updated version of this article at doi: 10.1074/jbc.M113.462812

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2013/03/22/M113.462812.DC1

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2013/03/22/jbc.M113.462812.full.html#ref-list-1