Classification: Biological Sciences: Biophysics and Computational Biology

SurA is a "Groove-y" Chaperone That Expands Unfolded Outer Membrane Proteins

Dagan C. Marx[1], Ashlee M. Plummer[1‡], Anneliese M. Faustino[2], Michaela A. Roskopf[1], Mathis J. Leblanc[1], Taylor Devlin[1], Henry J. Lessen[1], Ananya Majumdar[3], Barbara T. Amann[1], Patrick J. Fleming[1], Susan Krueger[4], Stephen D. Fried[2], Karen G. Fleming[1]*

[1]Thomas C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, MD

[2]Department of Chemistry, Johns Hopkins University, Baltimore, MD

[3]Biomolecular NMR Center, Johns Hopkins University, Baltimore, MD

[4]National Institute of Standards and Technology, Gaithersburg, MD

[‡] Current Address: Department of Cell Biology, Harvard Medical School, Boston, MA

*Corresponding Author: Karen.Fleming@jhu.edu

Keywords: Chaperones, Outer Membrane Protein Biogenesis, Periplasm, Circular Dichroism, Crosslinking, Crosslinking Mass Spectrometry, Small-Angle Neutron Scattering, Molecular Modeling

Abstract: Chaperone proteins play a critical role in the biogenesis of many nascent polypeptides *in vivo.* In the periplasm of *E. coli* this role is partially fulfilled by SurA, which promotes the efficient assembly of unfolded outer membrane proteins (uOMPs) into the bacterial outer membrane, though the mechanism by which SurA interacts with uOMPs is not well understood. Here we identify multiple conformations of SurA in solution, one of which contains a cradle-like groove in which client uOMPs bind. Access to this binding groove by clients is gated by the intrinsic conformational dynamics of SurA. Crosslinking mass spectrometry experiments identify multiple regions of native client uOMPs that bind to SurA, providing insight into the molecular determinants of SurA-uOMP interactions. In contrast to other periplasmic chaperones that encapsulate uOMPs, small angle neutron scattering data demonstrate that SurA binding greatly expands client uOMPs. These data can explain the dual roles of SurA as both a holdase and a foldase. Using an integrative modeling approach that combines crosslinking, mass spectrometry, small angle neutron scattering, and simulation, we propose structural models of SurA in complex with an unfolded protein client. We further find that multiple SurA monomers are able to bind discrete sites on a single uOMP. The structural arrangement of SurA and uOMPs provides the basis for a possible mechanism by which SurA binds and expands clients in a manner that facilitates their folding into the outer membrane.

Significance Statement: Outer membrane proteins play critical roles in bacterial physiology and increasingly are being exploited as antibiotic targets. Their biogenesis requires chaperones in the bacterial periplasm to safely ferry them to their destination membrane. We used crosslinking, mass spectrometry, and small angle neutron scattering to propose an ensemble of structural models that explain how one chaperone, SurA, stabilizes client outer membrane proteins through expansion of their overall size, which positions them for delivery to the BAM complex. This study highlights the use of an hybrid integrative approach and emerging methods in structural biology to map highly heterogeneous structural ensembles like that of an unfolded protein bound to a chaperone.

Introduction

Chaperone proteins are present in all cellular compartments across all domains of life and promote efficient folding of specific proteins, termed clients(1–4). In the absence of chaperone proteins, some clients are prone to populate misfolded states or form aggregates that can be toxic to the cell(5–7). Chaperones interact with these unfolded clients to suppress these off-pathway reactions and to ensure the formation of the functionally active conformation of their clients(8).

SurA is the most important chaperone in the *E. coli* outer membrane protein (OMP) biogenesis network, as it handles the majority of the flux of unfolded outer membrane proteins (uOMPs) through the periplasm(9–16). Eight OMPs have been identified as clients for SurA, as their expression is notably decreased in a strain of *E. coli* lacking SurA(15). SurA has been shown to act as both a holdase, preventing uOMP aggregation, and a foldase that catalyzes the folding of uOMPs into the outer membrane through an interaction with the β-barrel assembly machine (BAM)(14, 17–21). The driving forces for these dual functions must be accomplished entirely derived from SurA binding with clients, as there is no ATP in the periplasm.

The mechanism by which SurA binds client uOMPs is currently unknown. The binding constants for SurA binding to multiple OMPs as well as various peptides that mimic uOMP sequences are in the low micromolar range; however, the stoichiometry and structural arrangement of these interactions with native clients remains unclear(20, 22–28). Unlike other chaperones in the OMP biogenesis network that oligomerize to form cages around uOMPs, SurA has a modular structure with three distinct structural domains connected by flexible linkers: a “core” chaperone domain comprised of the N- and C-terminal helices and two peptidyl prolyl isomerase (PPIase) domains (P1 and P2)(29–31). Figure 1A shows the crystal structure of full length, monomeric SurA (PDB: 1M5Y) in which P1 is associated with the core domain, while P2 is structurally isolated.(32) The orientation of the P1 and P2 domains of SurA relative to the core domain has been hypothesized to be dynamic, though the details of the different conformations have not been previously described(20, 33). The core domain of SurA has been shown to be responsible for the majority of the binding energy to small uOMPs, with the P1 and P2 domains hypothesized to aid in binding larger clients(20, 34, 35). The great

diversity of SurA client OMPs, which can have 8-26 transmembrane (TM) beta-strands, suggests that SurA must recognize a common binding motif that is present in all clients(26, 36–38).

In this study, we use an integrative hybrid approach to elucidate the structural features of the complexes formed between SurA and two cognate uOMP clients, uOmpA and uOmpX(39–41). Combining experiment and simulation, we find that monomeric SurA adopts at least 4 distinct conformations in solution. Photo-crosslinking and mass spectrometry indicate that a cradle-like groove formed between the P1 and core domains is a common binding site of the transmembrane domain of uOmpA ($uOmpA_{171}$) and uOmpX. This cradle is only accessible when P1 is released from the core domain, forming an "open" state of SurA. Mass spectrometry further revealed specific regions on $uOmpA_{171}$ and uOmpX that bind within the SurA groove. Finally, small-angle neutron scattering (SANS) experiments reveal $uOmpA_{171}$ is greatly expanded in solution when bound to SurA compared to when it is it bound to other chaperones. Using this orthogonal structural information, a sparse ensemble of models of the SurA-$uOmpA_{171}$ complex was created that are consistent with all known experimental data. The sparse ensemble contains $uOmpA_{171}$ structures that can be occupied by one or more SurA and populate a wide distribution of conformational states that all present the β-signal on the exterior surface, which is known to be recognized by the BAM complex(42–45). Overall, our findings reveal how SurA binds and solubilizes uOMP clients in the periplasm and provide insight into a potential β-signal mediated hand-off mechanism of a uOMP from SurA to the BAM complex or the adjacent membrane(46).

Results

Apo SurA adopts multiple conformations in solution

The crystal structure shown in Figure 1A does not capture the conformational heterogeneity of the arrangements of the P1 and P2 domains that are inferred from genetic studies, which suggest that they adopt multiple states relative to the core domain(33). To determine the population distribution of SurA conformers, we measured the energetics of the interactions between the P1 or P2 with the core domain using chemical denaturation titrations. Due to the highly α-helical nature and size of the core domain relative to the smaller PPIase domains, we reasoned that the circular dichroism (CD) signal at 222nm reports primarily on the structure of the core domain. We measured the stabilities of SurA constructs lacking P1 ($SurA_{\Delta P1}$), P2 ($SurA_{\Delta P2}$), or both P1 and P2 ($SurA_{\Delta P1P2}$), as well as full length SurA (Figure S1, Table S1). Using thermodynamic analysis, we found that the P1 and P2 domains compete for binding to the core domain (Figure S2, Table S1). The interaction of P1 domain with the core is three-fold more favorable than the P2-core interaction. P1 and P2 undergo independent competition with each other for binding to the core domain.

Thus, in addition to the crystallized, "P1 closed" structure of monomeric SurA, two additional conformations of SurA must exist: an "open" SurA where both P1 and P2 are unbound from the core domain, and a "P2 closed" SurA where P2 is bound to the core domain with P1 unbound. Molecular dynamics simulations suggested an additional state of monomeric SurA, termed "collapsed", where P1 and P2 are bound to the core domain (Figure S3). Figure S4 and Table S2 summarize the four possible conformations of monomeric SurA in solution with "P1 closed" being the dominant conformation, followed by "P2 closed," then "open."

The open conformation of SurA exposes a cradle-like groove that binds client uOMPs

To identify the regions of SurA involved in binding uOMPs, we utilized chemical crosslinking by incorporating a photoactivatable unnatural amino acid (*para*-azido-Phenylalanine, pAF) at 36 individual surface-exposed positions on SurA. We mixed each of these $SurA_{pAF}$ variants either with $uOmpA_{171}$ or with uOmpX and measured the efficiency of crosslinking by quantitative SDS-PAGE (Figure S5, Table S3). Every $SurA_{pAF}$ variant crosslinks to both $uOmpA_{171}$ and uOmpX, though the crosslinking efficiency varies dramatically with position. The crosslinked complexes are primarily composed of one uOMP and one SurA (a one-to-one complex), though

in some cases, the formation of a complex corresponding to two SurAs and one uOMP is observed (two-to-one). These results indicate multiple segments of uOMP can be simultaneously occupied by SurA, in agreement with binding studies that found a cooperative Hill coefficient(20). In contrast, the non-client OmpLA only crosslinked to half of the $SurA_{pAF}$ variants and the variant with the highest crosslinking efficiency was less than half of the maximum raw crosslinking efficiencies for the cognate client uOMPs (34% vs 72%; Figure S6, Table S3).

Figure 1B shows the strong linear correlation of the normalized crosslinking efficiencies of the 36 $SurA_{pAF}$ variants to $uOmpA_{171}$ and uOmpX. The data reveal a conserved SurA region that crosslinks efficiently to both client uOMPs, demarcated by dotted lines in top right corner. These SurA residues are located on the core (gray) and P1 (blue) domains, indicating that P2 is not as important for binding these client uOMPs. The lack of high crosslinking to P2 agrees with published binding studies of SurA and $SurA_{\Delta P2}$ to $uOmpA_{171}$ that have shown that P2 does not thermodynamically contribute to $uOmpA_{171}$ binding(20).

Mapping the eight high-efficiency crosslinking sites (pink residues in Figure 1C-F) onto each of the four SurA conformations identifies that only the "open" conformation presents a uOMP binding groove in which crosslinking residues are co-localized to a single region of the protein (Figure 1D). The "open" conformation creates a 20-25Å chasm between the P1 and core domains, which is large enough to accommodate all types of secondary structure. The bottom of the putative OMP binding cradle is formed by the C-terminal helix of SurA with the walls formed by the N-terminal region of the core domain and P1 (Figure 1G). The bottom of the cradle is 25Å long, while the core and P1 walls are approximately 50Å long, which is long enough to sequester the entire length of a hydrophobic TM segment of an OMP from water. Figure 1G shows hydrophobic patches on the core and P1 walls of the cradle surrounded by polar regions. The C-terminal helix of the core domain has a 30Å hydrophobic stretch containing multiple aromatic residues which may also serve as a binding site for TM segments of uOMPs (Figure S7). Thus the occluded, cradle-like groove exposed in "open" SurA contains structural features that could interact favorably with client uOMPs.

SurA interacts with flexible, aromatic rich regions of $uOmpA_{171}$ and uOmpX

Crosslinking mass spectrometry (XL-MS) was used to locate where SurA crosslinks to the client uOMPs, $uOmpA_{171}$ and uOmpX. Seven high efficiency $SurA_{pAF}$ variants (four on the core domain and three on P1) were crosslinked to $uOmpA_{171}$ and four (two on the core and two on P1) were crosslinked to uOmpX. Photo-crosslinked samples were subject to proteolysis and the resulting crosslinked peptides were analyzed by LC-MS. Crosslinked peptides were identified (to a false discovery rate < 0.01) and mapped using the MeroX v2.0 software package (Figure S8 and S9)(47). Representative mass spectra of identified crosslinked peptides with their fragment ion peak assignments are provided in Figure S10A-I, and summary data of all peptide spectrum matches are provided as Supplementary Data 1 and 2 (see also SI Methods).

We considered a uOMP segment to specifically bind to SurA if it spanned residues that crosslinked to three or more SurA variants. Using this criterion, we identified six segments of $uOmpA_{171}$ that selectively bind to SurA. These segments are residues: 1-21, 51-73, 84-91, 95-104, 112-113, and 130-131 (Figures 2A, S11). For uOmpX we identified two segments: residues 41-47 and 68-72 (Figures 2B, S12). Importantly, when we challenged $uOmpA_{171}$ with a $SurA_{pAF}$ whose label was outside the "open" SurA groove ($SurA_{26,pAF}$), only a single spectrum matched to a crosslinked peptide, implying that the rich network of crosslinks between $uOmpA_{171}$ and SurA is dependent on pAF being placed in the cradle-like groove (Figure 2A). The presence of multiple SurA binding segments on uOMPs is consistent with the higher-order stoichiometries observed in our gel-based crosslinking experiments.

The segments on uOMPs that bind the "open" SurA groove vary in length and location and crosslink to residues on both the core and P1 domains, indicating clients can interact with both sides of the cradle. The sequences of identified cognate uOMP segments are unusually enriched in tyrosine residues (10 of the 13 tyrosines appear in segments; P=0.003 by Chi-square test). Glycines are also very common in binding segments, though here the enrichment is not significant because uOMPs have many glycines in general(48, 49). In the case of $OmpA_{171}$, many of these tyrosines and glycines in binding segments are highly conserved according to the Pfam family for the OmpA transmembrane domain (ID: PF01389)(50). Thus, the data support a model in which the "open" SurA groove preferentially binds highly flexible uOMP regions containing at least one tyrosine residue.

uOmpA exists in an expanded conformation when bound to SurA

SurA has no obvious cavity that is large enough to encapsulate an entire uOMP, unlike the binding modes used by other chaperones (Skp and DegP) in the OMP biogenesis network(30, 51). While the "open" SurA groove could protect small segments of the uOMP, it is not obvious how the remainder of a uOMP would be protected from aggregation. To address these questions, small angle neutron scattering (SANS) was used to measure the size of the crosslinked complex formed between the $SurA_{105,pAF}$ variant and $uOmpA_{171}$. This complex was chosen because it forms a single well defined, one-to-one crosslinking band as visualized by SDS-PAGE (Figure S5). SANS reports directly on the radius of gyration (R_G) and the distance of maximum dimension (D_{Max}). Additionally, the scattering contrast of each component in a complex can be selectively visualized by altering its deuteration state(52).

We collected SANS profiles of $(SurA_{105,pAF}\text{-}uOmpA_{171})_{XL}$ in multiple buffer conditions and different deuteration states to take advantage of this selective contrast feature (Figure 3A and Figure S13). Guinier analysis of data collected in 30% D_2O conditions on a sample composed of hydrogenated SurA and deuterated uOmpA was particularly insightful (Figure 3B). In this experiment, the scattering contrast is dominated by the $uOmpA_{171}$ component, and these data yielded an R_G value of 45 Å ± 3 Å (Tables S4 and S5). This value far exceeds the R_G observed for apo SurA in control experiments (35 Å) (Figure S14, Tables S4 and S5), indicating that $uOmpA_{171}$ is greatly expanded relative to a SurA monomer. Figure 3C shows the $P(r)$ distribution for the 30% D_2O condition where the D_{Max} of the complex is approximately 150 Å. Given the molar mass of $uOmpA_{171}$ (18 kDa), this D_{Max} value indicates that $uOmpA_{171}$ adopts an elongated conformation when bound to SurA.

Solution structures of SurA-uOmpA complexes reveal multiplicity

We created a total of 40 models of varying conformations and binding stoichiometries by docking the XL-MS based $uOmpA_{171}$ binding segments into the cradle-like groove of SurA as described in detail in the Supplemental Methods. Models were built with both unexpanded $uOmpA_{171}$, and expanded $uOmpA_{171}$ conformations consistent with the 30% D_2O SANS data (Figure S15). Motivated by SDS-PAGE analyses of crosslinked complexes and the multiplicity of binding sites on $uOmpA_{171}$, we included several structural models comprised of multiple SurA

protomers bound to a single $uOmpA_{171}$ (two, three, or four). We next tested these 40 structural models by calculating the theoretical SANS spectrum one would expect if various combinations of the structural models were present at different fractional populations. One of the SANS datasets (collected in 0% D_2O) was consistent with only 6 combinations that span across 6 structural models; the other SANS dataset (collected in 98% D_2O) was consistent with 81 combinations that span across 36 structural models. Overlays of these structural models are depicted in Figure 4.

The structural models and combinations thereof we have tested are not exhaustive; importantly however, we found that the SANS data could not be satisfactorily fit (with a reduced chi-square cut-off of 1) unless structural models with elongated $uOmpA_{171}$ with two (or more) bound SurA were included. Four of the six structural models consistent with the 0% D_2O SANS dataset contain an identical $uOmpA_{171}$ conformation with one and two SurA protomers bound, highlighting the existence of multiple binding sites. Figure 5 shows structural models of the one-to-one complexes with SurA bound to segments 51–73 and 130–131 on $uOmpA_{171}$, and a two-to-one complex with SurA occupying both of those sites.

Discussion

Based on the results described, we propose the existence of multiple SurA conformations including three novel SurA domain arrangements not previously observed in crystal structures of monomeric SurA(32). These new conformations suggest a dynamic arrangement of the P1 and P2 domains relative to the core domain. The evolution of structural plasticity in SurA points to functional roles for these conformers to aid in client recognition while at the same time protecting the occluded OMP-binding cradle from occupancy by non-cognate clients. This autoinhibition of SurA chaperone function prevents depletion of the availability of this important biogenesis factor that serves as both a foldase and a holdase in outer membrane protein maturation(12, 20). The competitive binding of P1 and P2 for the same region of the core domain likely regulates formation of the open conformation, as the cradle can be occluded by either the P1 or the P2 domains closing.

Comparison of the "P1 closed" and "open" conformations of SurA reveals a chemically heterogeneous cradle-like groove where uOMP segments bind. Unlike many chaperone-client interactions, there are no obvious uOMP binding sites on the P1 or core domain. This suggests uOMPs bind to SurA through a mixture of hydrophobic and polar interactions that could be mediated by the alternating chemical nature of side chains in a TM β-sheet. Indeed, mapping the uOMP segments that crosslink to SurA onto structures of $OmpA_{171}$ and OmpX reveals that every identified segment contains sequences that become β-sheet upon folding (Figure S16). The recognition and sequestration of TM β-sheet segments would reduce non-native β-sheet formation, which has been observed to happen in kinetic studies of the intrinsic folding pathway of $uOmpA_{171}$(53).

Our chemical crosslinking studies show that the "open" conformation of SurA, which is the least intrinsically stable conformation, is the active chaperone state. Residues on the surface of SurA that crosslink the most efficiently to uOMPs co-localize around and inside of a large groove in this "open" conformation that is occluded in all other conformations. We speculate that this binding site is specific for client-OMP segments because highly efficient crosslinking is only observed for cognate uOMP clients. Mass spectrometry identified multiple binding segments on each client uOMP that have very low sequence similarity but are each

highly enriched in tyrosines (Figures S11, S12). It is notable that the C-terminal helix of SurA (which forms the floor of the groove) contains a hydrophobic patch with multiple aromatic residues that may specifically interact with the tyrosines in uOMPs through a pi-stacking mediated mechanism (Figure S7).

Binding segments of $uOmpA_{171}$ were also found on loops and turns of the folded structure of $OmpA_{171}$ (Figure S16). A particularly intriguing possibility that is created by the cradle confinement is that SurA reduces the chain entropy in the bound region of a uOMP. This would stabilize flexible loops and turns in uOMPs that could allow for native beta-hairpin formation when bound to SurA. The BAM complex has been proposed to template individual beta-hairpins that are inserted sequentially into the outer membrane, but it is possible that SurA induces uOMPs to form beta hairpins prior to interactions with BAM or the adjacent, disrupted membrane(54–57). Due to the apparent flexibility of the linkers connecting the P1 and core domains of SurA, the "open" conformation may be able to accommodate larger structures than beta-hairpins, such as folded OMPs, as the current model of the cradle in "open" SurA is approximately the same dimensions as the TM domain of an eight-stranded OMP. We cannot exclude the possible stabilization of an OMP in its folded conformation by two or more wide "open" SurA monomers, similar to a proposed mechanism of chaperone function of DegP(30).

Additionally, we can identify differences in how SurA recognizes the clients $uOmpA_{171}$ and uOmpX with respect to the N-terminus of these clients. uOMPs are post translationally secreted N-terminus first by the Sec translocon into the periplasm, and it is likely that the N-terminus must be immediately recognized by chaperones to avoid aggregation and/or interaction with the inner membrane. Recently, low resolution cryo-EM structures have identified a potential interaction between SurA and the translocon in the presence and absence of OmpA(58). We observe a high density of crosslinked peptides in the first twenty residues of $uOmpA_{171}$, showing the SurA intrinsically recognizes the $uOmpA_{171}$ N-terminus. In contrast, the N-terminus of uOmpX does not appear to crosslink to SurA using the variants tested. This differential recognition of client uOMP N-termini by SurA hints at the possibility of multiple

uOMP biogenesis pathways that can diverge as soon as the uOMP is introduced into the periplasm.

The C-terminus of uOMPs is also very important, as it contains the β-signal, which is essential for *in vivo* and *in vitro* folding of OMPs(36, 42, 43, 59). The absence of robust crosslinking to the β-signal in both client uOMPs and the distal location of the β-signal in our structural models of SurA-bound $uOmpA_{171}$ is telling (Figures 2 and 4). The BAM complex is known to recognize the β-signal, hence it would be conducive for membrane insertion if SurA leaves this segment unobstructed and therefore presentable to the BAM complex(42, 59). Additionally, the expansion of uOMPs when bound to SurA could allow for uOMPs to reduce their "search space" when trying to find the BAM complex or facilitate passage through pores in the peptidoglycan network(60).

Other chaperones in the OMP biogenesis pathway (e.g., Skp and DegP) are holdases that form cages around uOMPs, which effectively inhibits uOMPs from accessing the BAM complex(29, 34). $uOmpA_{171}$ has an intrinsic s-value of 1.65 Svedbergs (R_G = 24Å), which is equivalent to the estimated R_G of 8-stranded uOMPs when bound to Skp. Our SANS experiments reveal $uOmpA_{171}$ has a R_G of 45Å and a D_{Max} of 150Å when bound to SurA, which is just shy of the width of the periplasm(51–53). We therefore propose that SurA binds and expands uOMPs in a way that increases the likelihood that the β-signal finds the BAM complex, or the adjacent disrupted membrane region, to initiate folding into the outer membrane.

However, the expansion of uOMPs also increases the chance for non-native intra- or inter-protein contacts to be formed which can lead to toxic misfolded proteins in the periplasm. Additionally, uOMPs must be shielded from interacting with the inner membrane, which can mediate aggregation(53). The presence of multiple segments on uOMPs that interact with SurA enables several SurA protomers to interact with a single uOMP, which could prevent these undesired outcomes. We identify two and six SurA binding segments on the smallest possible SurA client uOMPs (8 TM beta-strands), and hypothesize that the number of binding segments will increase with size. Thus we expect multiple SurA monomers will be necessary to effectively chaperone uOMP clients, in agreement with binding studies of uOMPs to SurA, where the Hill coefficient has been shown to be greater than 1(20). Sequential binding of uOMPs to multiple

SurA monomers could facilitate diffusion across the periplasm down the uOMP concentration gradient in a similar fashion to a proposed mechanism for transport through the nuclear pore complex(61). The presence of heterocomplexes of SurA and other chaperones (Skp, FkpA, and DegP) bound to a single uOMP could also facilitate uOMP diffusion, and could be necessary to solubilize larger, aggregation-prone clients(29).

In summary, our results provide structural insight into the conformational plasticity of proteins involved in uOMP biogenesis, as both SurA, and uOMPs bound to SurA are shown to be highly dynamic. We provide further evidence that SurA may be involved in the specific recognition of some uOMPs as they enter the periplasm, and solubilizes uOMPs in a manner that may induce productive interactions with the BAM complex (Figure 6). The mechanism that SurA utilizes is distinct from that of other chaperones described to date to the best of our knowledge, and likely reflects the unique challenges associated with maintaining a membrane protein in an aqueous compartment without ATP. Because SurA's mechanism requires it to access conformations with low fractional population that greatly differ from its lowest-energy conformation, much of the mechanism we propose would have been invisible to traditional structural biology techniques. This study therefore highlights the use of emerging methods in structural biology to map highly heterogeneous structural ensembles such as that of an unfolded protein to a chaperone.

Materials and Methods

Circular Dichroism- CD titrations were collected on an Aviv 62A DS spectropolarimeter using a 0.1 cm quartz cuvette. SurA and SurA domain deletion constructs were diluted to 1µM (SurA, $SurA_{\Delta P1}$, $SurA_{\Delta P2}$) or 1.5µM ($SurA_{\Delta P1P2}$) in buffer containing 20mM Tris (Fisher Scientific), pH 8.0. Equilibrium unfolding titrations were conducted by titrating a solution containing equivalent protein concentration to the analyte and 8M urea (ThermoFisher), 20mM Tris, pH8.0 using a computer controlled titrator (Hamilton) to maintain constant protein concentration. Each urea step was between 0.1 and 0.2M urea until a final concentration of 7M urea was reached, with 5 min of equilibration time with stirring between each reading. Signal at 222nm was averaged at each data point 30 seconds with stirring off. Thermodynamic stabilities were determined by fitting the unfolding titration curves to a two-state linear extrapolation model (Figure S1).

Molecular Dynamics of apo-SurA- All-atom molecular dynamics (MD) simulations were used to probe the conformational heterogeneity of apo-SurA. Because the crystallized conformation lacks several loops, we utilized the SWISS-MODEL homology-modeling server to build in these missing residues, along with a C-terminal 6x Histidine tag(62–64). Using CHARMM-GUI, we constructed an MD simulation system containing SurA solvated with 69,671 explicit water molecules and 200 mM NaCl in a 132 x 132 x 132 $Å^3$ box (65, 66). The CHARMM36 force field was utilized for these simulations which were run at 25 °C. These simulations were completed with a time step of 2 femtoseconds/step and 500,000 steps per run; consecutive 10 nanosecond runs were combined for the data shown in Figure S3. For each panel in this figure, data points are shown in 10 ps increments.

We monitored the backbone alpha-carbon root-mean-square deviation (Cα-RMSD) of each SurA conformer relative to the starting structure using the VMD (v. 1.9.3) RMSD Trajectory Tool(67) (Figure S3A). We quantified this parameter for both the entire protein and the individual structural units of SurA defined as follows: the N/PPIase-1/C domain as residues 1 to 255 and 371 to 417 and the PPIase-2 domain as residues 262 to 367 (Figure S3B). The protein R_G was also calculated as a function of simulation time using a previously published TCL script in VMD(68) (Figure S3C). After 85 ns of simulation, we observe a closing of SurA to form the collapsed conformation illustrated as a cartoon in Figure S4.

Biochemical Crosslinking

We use the library of SurA-*p*AF variants to understand which structural regions of the SurA chaperone are involved with binding to uOMP clients. 25 µM of each $SurA_{pAF}$ variant was mixed with 5 µM unfolded $OmpA_{171}$ ($uOmpA_{171}$) in 20 mM Tris (pH = 8.0) and 1 M Urea. Mixtures were then irradiated with UV light (wavelength, λ= 254 nm) for 5 minutes using a Spectroline MiniMax UV Lamp (Fisher #11-992-662). Aliquots were taken for SDS-PAGE analysis both pre- and post- exposure to UV light. These samples were subjected to electrophoresis using a 12 % precast gel (Mini-PROTEAN TGX, Bio-Rad) at a constant voltage of 150 V for 55 minutes at room-temperature. Using ImageJ, densitometry analysis on the loss of density of the $uOmpA_{171}$ band was utilized to quantitate crosslinking efficiency. Crosslinking efficiency values were corrected for the amount of $uOmpA_{171}$ band intensity lost when mixed with WT SurA (not containing pAF). A representative SDS-PAGE gel for each $SurA_{pAF}$ variant and $uOmpA_{171}$ is shown in Supplemental Figure 5. This same protocol was utilized to assess the crosslinking of $SurA_{pAF}$ variants to uOmpX and uOmpLA (Supplemental Figure 6).

Mass Spectrometry and Data Analysis-A Thermo Q-Exactive HF-X Orbitrap mass spectrometer was used to analyze protein digests. A full MS scan in positive ion mode was followed by ten data-dependent MS scans. The full MS scan was collected using a resolution of 30,000 (@ m/z 200), an AGC target of 3E6, a maximum injection time of 100 ms, and a scan range from 350 to 1500 m/z. The data-dependent scans were collected with a resolution of 15,000 (@ m/z 200), an AGC target of 2E5, a minimum AGC target of 8E3, a maximum injection time of 250 ms, and an isolation window of 2.0 m/z units. To dissociate precursors prior to their re-analysis by MS2, peptides were subjected to a stepped HCD with 22%, 25%, and 28% normalized collision energies. Fragments with charges of 1, 2, and >8 were excluded from analysis, and a dynamic exclusion window of 60.0 s was used for the data-dependent scans.

MS data were centroided and converted to the mzML file format using the msConvert application in the ProteoWizard Toolkit.(69) MS data were then analyzed for crosslinks using MeroX Version 2.0 (Götze, 2015).(47) p-azido-phenylalanine was added to the amino acid list with a mass of 188.06981084 Da. For tryptic digests, protease sites were allowed after arginine and lysine residues, with lysine blocked by proline as a cleavage site. For double digests,

protease sites were allowed at arginine, aspartic acid, lysine, and glutamic acid, with lysine blocked by proline as a cleavage site. For both tryptic digests and double digests, a maximum of three missed cleavages were allowed. For modifications, a maximum of two oxidations of methionine were allowed. p-azido-phenylalanine was input as a promiscuous crosslinking agent with a composition of $–N_2$, and a maximum Cα-Cα-distance of 30 Å. Default settings were used for the mass comparison, score, and false discovery rate (FDR). A MeroX score of 50 was selected as the acceptance cutoff for cross-linked peptide-spectrum matches (PSMs); all accepted PSMs had an FDR of <0.01 (Figures S8 and S9). In numerous situations, a crosslink site to p-azido-phenylalanine could not be pinpointed down to a specific residue with a given peptide-spectrum match because of insufficient fragment ion data. In this situation, limiting boundaries inferred from several peptide-spectrum matches were manually combined to determine a 'minimal' region of crosslink sites that were consistent with all the data. This manual data reduction process is explicitly shown in Supplementary Data 1 and 2, and results in the binding segments illustrated in Figure 2.

SANS Data Collection- All scattering experiments were collected at the National Institute of Standards and Technology Center for Neutron Research (Gaithersburg, MD) as previously described.(52) The scattering data presented here were collected on the NGB 30-m SANS Instrument (NIST). A neutron beam of wavelength $\lambda = 6$ Å (wavelength spread, $\Delta\lambda/\lambda = 0.15$) was utilized to collect scattering profiles from all samples described here on a 2D position-sensitive detector (64 cm x 64 cm) with 128 x 128 pixels at resolution of 0.5 cm pixel^{-1}. For data processing, raw counts were normalized to a common monitor count and then corrected for empty cell counts, ambient room background counts, and non-uniform detector response. Data were placed on an absolute scale by normalizing the scattering intensity to the incident beam flux. Radial averaging was utilized to produce scattering intensity profiles, I(q) versus q; $q = 4\pi\sin(\theta)/\lambda$, where 2θ is the scattering angle, λ is the neutron wavelength, and q is the magnitude of the scattering vector. Sample-to-detector distances of 5.0 m and 1.5 m were used to cover a range of 0.01 Å^{-1} < q < 0.4 Å^{-1}.

SANS on Fully Protonated SurA-uOmpA$_{171}$ Complex- SANS data were collected on a fully protonated SurA-uOmpA$_{171}$ complex to quantify the complex R_G. We scaled up the crosslinking reaction for uOmpA$_{171}$ and SurA$_{105,pAF}$, purified this complex by gel-filtration, and buffer exchanged this complex into GF buffer containing 98 % D_2O using the protocol described above for apo-SurA. This dataset is shown in Figure S13 in gold.

SANS on Protonated-SurA/Deuterated-uOmpA$_{171}$ Complex- The SurA $_{105,pAF}$ was crosslinked to deuterated-uOmpA$_{171}$ as described above for the fully protonated complex. This complex was purified as described above and buffer exchanged into either 0 % or 30 % D_2O GF buffer for SANS experiments (purple and green in Figures S13 and 3). We made three attempts to also collect scattering profiles in 80 % and 98 % D_2O of this complex but the I(0) values from Guinier fitting indicated that these samples contained aggregates. It is known that increased buffer concentrations of D_2O may promote self-association and aggregation of particularly hydrophobic proteins.(70)

Generation of uOmpA$_{171}$ Structures- We utilized a torsional angle Monte Carlo procedure combined with standard molecular dynamics (MD) to construct an ensemble of uOmpA conformations consistent with the experimentally measured sedimentation coefficient of 1.65 svedbergs for uOmpA in solution.(53) An extended polypeptide structure ($\phi = -78, \varphi = 149$) containing the mature TM domain sequence of E. coli OmpA (residues 2–172) was built using PyMOL.(71) This extended structure was partially collapsed using a coarse-grained torsion angle Monte Carlo procedure developed at Johns Hopkins University (REDUX, https://pages.jh.edu/fleming). The coarse-grained, expanded globule structures were converted back to all atom structures using PULCHRA(72) and subjected to MD simulations. These simulations were carried out with generalized Born implicit solvent electrostatics using NAMD(73) and the CHARMM22 force field (74) and experimental constraints incorporated through the collective variables module in NAMD. This procedure involved first collapsing the models to an effective anhydrous radius of gyration (R_G) of 20 Å and then re-expansion steps using the experimental target sedimentation coefficient = 1.65 s as calculated using HullRad.(75) Twelve independent uOmpA models were created.

Docking of uOmpA Structures to SurA **-**Four segments of the extended uOmpA structure described above (residues 2-21, 54-73, 84-104, 115-132) were independently submitted to the HADDOCK protein-protein docking web server.(76) These sequence segments were chosen to include those residues that were found to efficiently cross-link to $SurA_{105,PAF}$ using the photoactivatable unnatural amino acid (para-azido-Phenylalanine, pAF) as described above. Active and passive residues for HADDOCK were chosen from the surface of open SurA containing the most efficient cross-linking sites (Fig. S11).

In all cases the extended peptide was found to be docked to the open SurA surface in at least two major orientations by HADDOCK. Examples of these docked oligopeptides were inspected using molecular graphics to obtain target distances for docking the full length uOmpA models with s = 1.65 (described above) to the open form of SurA using targeted molecular dynamics. These simulations used the collective variables module of NAMD as described above, but with the target distances from HADDOCK peptide docking as distance restraints.

The uOmpA-open SurA models were further manipulated by targeted molecular dynamics to increase the uOmpA D_{max} to the target of 150 Å that was determined by P(r) analysis of the SANS data. A second open SurA model was then docked to exposed, known binding segments (Fig. S11) of the extended uOmpA. In three cases, a third open SurA was docked to remaining exposed known binding segments.

One extended polypeptide of uOmpA was generated with a D_{Max} ~250 Å and four open SurA models were docked to the four main segments on OmpA that displayed high efficiency cross-linking.

In all, twenty-three models containing one docked SurA, thirteen models containing two docked SurA, three models containing three SurA, and one model containing four SurA were built. Physical dimensions of these models are listed in Table S7. Values for R_G and D_{Max} were calculated using HullRad(75). All models contained CHARMM hydrogens and were used to calculate predicted SANS profiles using the SasCalc server.(77)

Acknowledgements

We thank the Johns Hopkins University Biomolecular NMR Center and the Center for Molecular Biophysics for providing facilities and resources. Access to NGB30 SANS was provided by the Center for High Resolution Neutron Scattering, a partnership between the National Institute of Standards and Technology and the National Science Foundation under Agreement No. DMR-1508249. We acknowledge the support of the National Institute of Standards and Technology, U.S. Department of Commerce, in providing the neutron research facilities used in this work. This work benefitted from CCP-SAS software developed through a joint EPSRC (EP/K039121/1) and NSF (CHE-1265821) grant. This work was supported by National Science Foundation (NSF) grant MCB1412108. D.C.M, A.M.P., T.D. and H.J.L. were supported by NIH training grant T32 GM008403. A.M.P. was supported by the NSF grant DGE 1232825. The authors thank lab members for helpful discussions.

References:

1. F. U. Hartl, M. Hayer-Hartl, Protein folding. Molecular chaperones in the cytosol: From nascent chain to folded protein. *Science (80-.).* **295**, 1852–1858 (2002).
2. Y. E. Kim, M. S. Hipp, A. Bracher, M. Hayer-Hartl, F. Ulrich Hartl, Molecular Chaperone Functions in Protein Folding and Proteostasis. *Annu. Rev. Biochem.* **82**, 323–355 (2013).
3. H. Saibil, Chaperone machines for protein folding, unfolding and disaggregation. *Nat. Rev. Mol. Cell Biol.* **14**, 630–642 (2013).
4. F. U. Hartl, Molecular chaperones in cellular protein folding. *Nature* **381**, 571–580 (1996).
5. F. Chiti, C. M. Dobson, Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu. Rev. Biochem.* **86**, 27–68 (2017).
6. C. M. Dobson, Principles of protein folding, misfolding and aggregation in *Seminars in Cell and Developmental Biology*, (Elsevier Ltd, 2004), pp. 3–16.
7. J. M. Barral, S. A. Broadley, G. Schaffar, F. U. Hartl, Roles of molecular chaperones in protein misfolding diseases. *Semin. Cell Dev. Biol.* **15**, 17–29 (2004).
8. F. U. Hartl, A. Bracher, M. Hayer-Hartl, Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324–332 (2011).
9. G. Mas, S. Hiller, Conformational plasticity of molecular chaperones involved in periplasmic and outer membrane protein folding. *FEMS Microbiol. Lett.* **365** (2018).
10. G. Mas, J. Thoma, S. Hiller, The Periplasmic Chaperones Skp and SurA. *Subcell. Biochem.* **92**, 169–186 (2019).
11. A. M. Plummer, K. G. Fleming, From Chaperones to the Membrane with a BAM! *Trends Biochem. Sci.* **41**, 872–882 (2016).
12. A. P. Chum, S. R. Shoemaker, P. J. Fleming, K. G. Fleming, Plasticity and transient binding are key ingredients of the periplasmic chaperone network. *Protein Sci.* **28**, 1340–1349 (2019).
13. F. Stull, J.-M. Betton, J. C. A. Bardwell, Periplasmic Chaperones and Prolyl Isomerases. *EcoSal Plus* **8** (2018).
14. S. M. Costello, A. M. Plummer, P. J. Fleming, K. G. Fleming, Dynamic periplasmic chaperone reservoir facilitates biogenesis of outer membrane proteins. *Proc. Natl. Acad.*

Sci. U. S. A. **113**, E4794–E4800 (2016).

15. K. Denoncin, J. Schwalm, D. Vertommen, T. J. Silhavy, J.-F. Collet, Dissecting the Escherichia coli periplasmic chaperone network using differential proteomics. *Proteomics* **12**, 1391–401 (2012).
16. S. W. Lazar, M. Almirón, A. Tormo, R. Kolter, Role of the Escherichia coli SurA protein in stationary-phase survival. *J. Bacteriol.* **180**, 5704–11 (1998).
17. J. G. Sklar, T. Wu, D. Kahne, T. J. Silhavy, Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli. *Genes Dev.* **21**, 2473–2484 (2007).
18. A. E. Rizzitello, J. R. Harper, T. J. Silhavy, Genetic evidence for parallel pathways of chaperone activity in the periplasm of Escherichia coli. *J. Bacteriol.* **183**, 6794–6800 (2001).
19. T. J. Silhavy, J. C. Malinverni, Assembly of Outer Membrane β-Barrel Proteins: the Bam Complex. *EcoSal Plus* **4** (2011).
20. J. R. Humes, *et al.*, The Role of SurA PPIase Domains in Preventing Aggregation of the Outer-Membrane Proteins tOmpA and OmpT. *J. Mol. Biol.* **431**, 1267–1283 (2019).
21. B. Schiffrin, *et al.*, Effects of Periplasmic Chaperones and Membrane Thickness on BamA-Catalyzed Outer-Membrane Protein Folding. *J. Mol. Biol.* **429**, 3776–3792 (2017).
22. E. W. Bell, E. J. Zheng, L. M. Ryno, Identification of inhibitors of the E. coli chaperone SurA using in silico and in vitro techniques. *Bioorg. Med. Chem. Lett.* **28**, 3540–3548 (2018).
23. E. Bitto, D. B. McKay, Binding of phage-display-selected peptides to the periplasmic chaperone protein SurA mimics binding of unfolded outer membrane proteins. *FEBS Lett.* **568**, 94–98 (2004).
24. K. H. Stymest, P. Klappa, The periplasmic peptidyl prolyl cis-trans isomerases PpiD and SurA have partially overlapping substrate specificities. *FEBS J.* **275**, 3470–3479 (2008).
25. E. Bitto, D. B. McKay, The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins. *J. Biol. Chem.* **278**, 49316–49322 (2003).
26. G. Hennecke, J. Nolte, R. Volkmer-Engert, J. Schneider-Mergener, S. Behrens, The periplasmic chaperone SurA exploits two features characteristic of integral outer

membrane proteins for selective substrate recognition. *J. Biol. Chem.* **280**, 23540–23548 (2005).

27. Q. Chai, *et al.*, Diverse sequences are functional at the C-terminus of the E. coli periplasmic chaperone SurA. *Protein Eng. Des. Sel.* **27**, 111–116 (2014).
28. H. M. Webb, L. W. Ruddock, R. J. Marchant, K. Jonas, P. Klappa, Interaction of the periplasmic peptidylprolyl cis-trans isomerase SurA with model peptides: The N-terminal region of SurA is essential and sufficient for peptide binding. *J. Biol. Chem.* **276**, 45622–45627 (2001).
29. B. Schiffrin, *et al.*, Skp is a multivalent chaperone of outer-membrane proteins. *Nat. Struct. Mol. Biol.* **23**, 786–793 (2016).
30. T. Krojer, *et al.*, Structural basis for the regulated protease and chaperone function of DegP. *Nature* **453**, 885–890 (2008).
31. C. W. Sandlin, N. R. Zaccai, K. G. Fleming, Skp Trimer Formation Is Insensitive to Salts in the Physiological Range. *Biochemistry* **54**, 7059–62 (2015).
32. E. Bitto, D. B. McKay, Crystallographic structure of SurA, a molecular chaperone that facilitates folding of outer membrane porins. *Structure* **10**, 1489–98 (2002).
33. G. R. Soltes, J. Schwalm, D. P. Ricci, T. J. Silhavy, The activity of Escherichia coli chaperone SurA is regulated by conformational changes involving a parvulin domain. *J. Bacteriol.* **198**, 921–929 (2016).
34. J. Thoma, B. M. Burmann, S. Hiller, D. J. Müller, Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. *Nat. Struct. Mol. Biol.* **advance on**, 795–802 (2015).
35. G. Li, *et al.*, Single-Molecule Detection Reveals Different Roles of Skp and SurA as Chaperones. *ACS Chem. Biol.* **13**, 1082–1089 (2018).
36. D. Vertommen, N. Ruiz, P. Leverrier, T. J. Silhavy, J. F. Collet, Characterization of the role of the escherichia coli periplasmic chaperone SurA using differential proteomics. *Proteomics* **9**, 2432–2443 (2009).
37. S. E. Rollauer, M. A. Sooreshjani, N. Noinaj, S. K. Buchanan, Outer membrane protein biogenesis in Gram-negative bacteria. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **370** (2015).

38. R. Koebnik, K. P. Locher, P. Van Gelder, Structure and function of bacterial outer membrane proteins: Barrels in a nutshell. *Mol. Microbiol.* **37**, 239–253 (2000).

39. A. Sali, *et al.*, Outcome of the First wwPDB Hybrid/Integrative Methods Task Force Workshop in *Structure*, (Cell Press, 2015), pp. 1156–1167.

40. M. P. Rout, A. Sali, Principles for Integrative Structural Biology Studies. *Cell* **177**, 1384–1403 (2019).

41. A. B. Ward, A. Sali, I. A. Wilson, Integrative structural biology. *Science (80-.).* **339**, 913–915 (2013).

42. M. Struyvé, M. Moons, J. Tommassen, Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *J. Mol. Biol.* **218**, 141–8 (1991).

43. D. R. Hendrixson, M. L. De La Morena, C. Stathopoulos, J. W. St Geme III, Structural determinants of processing and secretion of the Haemophilus influenzae Hap protein . *Mol. Microbiol.* **26**, 505–518 (1997).

44. V. Robert, *et al.*, Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol.* **4**, 1984–1995 (2006).

45. C. M. Sandoval, S. L. Baker, K. Jansen, S. I. Metzner, M. C. Sousa, Crystal structure of BamD: an essential component of the β-Barrel assembly machinery of gram-negative bacteria. *J. Mol. Biol.* **409**, 348–57 (2011).

46. T. J. Knowles, *et al.*, Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Mol. Microbiol.* **68**, 1216–27 (2008).

47. M. Götze, *et al.*, Automated assignment of MS/MS cleavable cross-links in protein 3d-structure analysis. *J. Am. Soc. Mass Spectrom.* **26**, 83–97 (2014).

48. M. B. Ulmschneider, M. S. P. Sansom, Amino acid distributions in integral membrane protein structures. *Biochim. Biophys. Acta - Biomembr.* **1512**, 1–14 (2001).

49. J. S. G. Slusky, R. L. Dunbrack, Charge asymmetry in the proteins of the outer membrane. *Bioinformatics* **29**, 2122–8 (2013).

50. S. El-Gebali, *et al.*, The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, D427–D432 (2019).

51. B. M. Burmann, C. Wang, S. Hiller, Conformation and dynamics of the periplasmic membrane-protein-chaperone complexes OmpX-Skp and tOmpA-Skp. *Nat. Struct. Mol. Biol.* **20**, 1265–1272 (2013).

52. N. R. Zaccai, *et al.*, "Deuterium Labeling Together with Contrast Variation Small-Angle Neutron Scattering Suggests How Skp Captures and Releases Unfolded Outer Membrane Proteins" in *Methods in Enzymology*, (Academic Press Inc., 2016), pp. 159–210.

53. E. J. Danoff, K. G. Fleming, The soluble, periplasmic domain of OmpA folds as an independent unit and displays chaperone activity by reducing the self-association propensity of the unfolded OmpA transmembrane β-barrel. *Biophys. Chem.* **159**, 194–204 (2011).

54. M. T. Doyle, H. D. Bernstein, Bacterial outer membrane proteins assemble via asymmetric interactions with the BamA β-barrel. *Nat. Commun.* **10** (2019).

55. A. I. C. Höhr, *et al.*, Membrane protein insertion through a mitochondrial β-barrel gate. *Science (80-.).* **359** (2018).

56. N. Noinaj, *et al.*, Structural insight into the biogenesis of β-barrel membrane proteins. *Nature* **501**, 385–390 (2013).

57. J. Lee, *et al.*, Formation of a β-barrel membrane protein is catalyzed by the interior surface of the assembly machine protein BamA. *Elife* **8** (2019).

58. S. Alvira, *et al.*, Trans-membrane association of the Sec and BAM complexes for bacterial outer-membrane biogenesis. *bioRxiv* https:/doi.org/10.1101/589077 (December 9, 2019).

59. D. Gessmann, *et al.*, Outer membrane β-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 5878–5883 (2014).

60. M. Pazos, K. Peters, Peptidoglycan. *Subcell. Biochem.* **92**, 127–168 (2019).

61. L. Maguire, M. Stefferson, M. D. Betterton, L. E. Hough, Design principles of selective transport through biopolymer barriers. *Phys. Rev. E* **100** (2019).

62. L. Bordoli, *et al.*, Protein structure homology modeling using SWISS-MODEL workspace. *Nat. Protoc.* **4**, 1–13 (2009).

63. K. Arnold, L. Bordoli, J. Kopp, T. Schwede, The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195–201 (2006).

64. M. Biasini, *et al.*, SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42** (2014).

65. J. Lee, *et al.*, CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).

66. S. Jo, T. Kim, V. G. Iyer, W. Im, CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).

67. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33- (1996).

68. A. Minoia, rgyr.tcl. *http://chembytes.wdfiles.com/local--files/yasc-vmd/rgyr.tcl* (2010).

69. M. C. Chambers, *et al.*, A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* (2012) https:/doi.org/10.1038/nbt.2377.

70. J. J. Lee, D. S. Berns, Protein aggregation. The effect of deuterium oxide on large protein aggregates of C-phycocyanin. *Biochem. J.* **110**, 465–470 (1968).

71. , PyMOL. *PyMOL Mol. Graph. Syst. Version 1.3 Schrödinger, LLC.*

72. P. Rotkiewicz, J. Skolnick, Fast procedure for reconstruction of full-atom protein models from reduced representations. *J. Comput. Chem.* (2008) https:/doi.org/10.1002/jcc.20906.

73. J. C. Phillips, *et al.*, Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).

74. A. D. MacKerell, *et al.*, All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).

75. P. J. Fleming, K. G. Fleming, HullRad: Fast Calculations of Folded and Disordered Protein and Nucleic Acid Hydrodynamic Properties. *Biophys. J.* **In Press** (2018).

76. G. C. P. Van Zundert, *et al.*, The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* (2016)

https:/doi.org/10.1016/j.jmb.2015.09.014.

77. J. E. Curtis, S. Raghunandan, H. Nanda, S. Krueger, SASSIE: A program to study intrinsically disordered biological molecules and macromolecular ensembles using experimental scattering restraints. *Comput. Phys. Commun.* **183**, 382–389 (2012).

78. D. Eisenberg, E. Schwarz, M. Komaromy, R. Wall, Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **179**, 125–142 (1984).

Figures

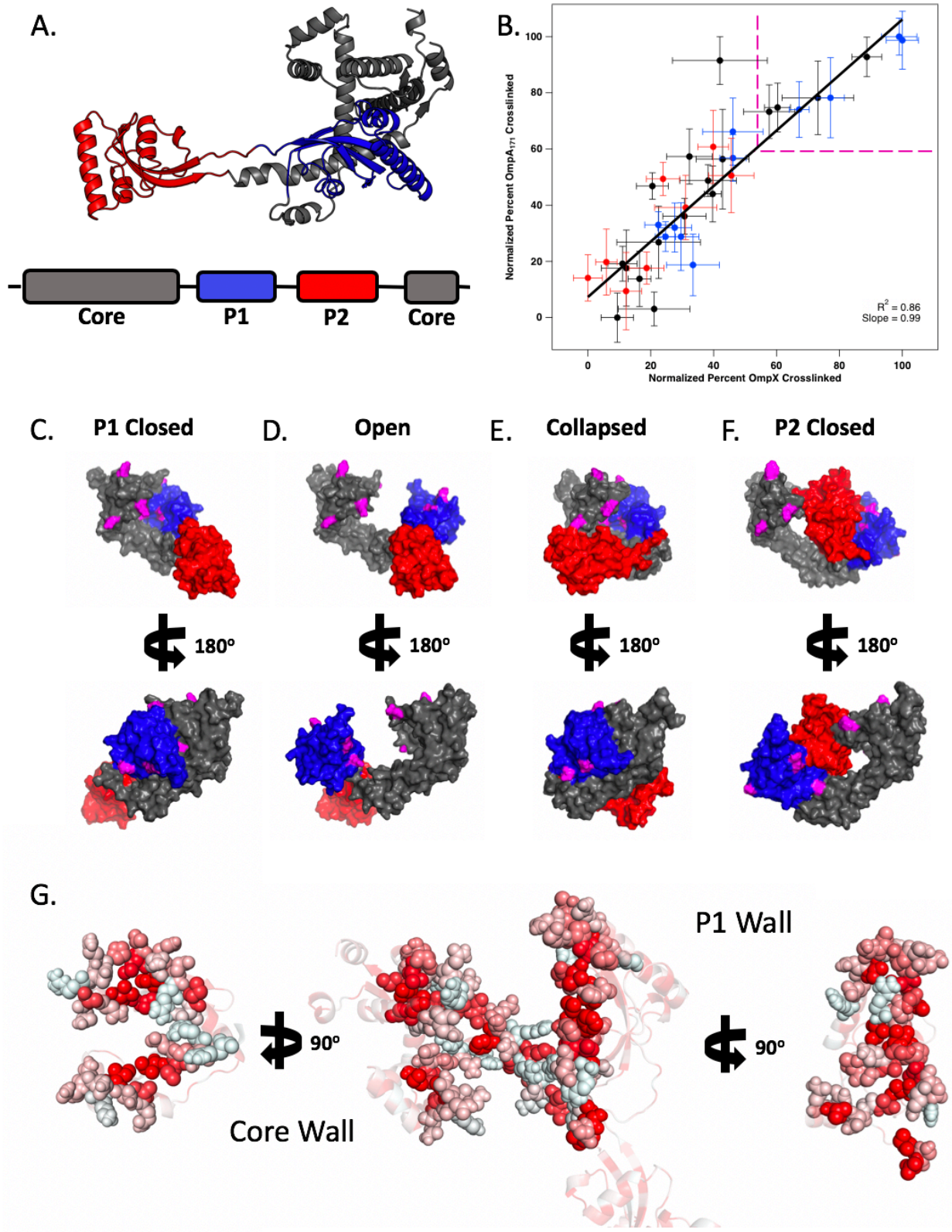
A.
Core
P1
P2
Core
B.
Normalized Percent OmpA171 Crosslinked
Normalized Percent OmpX Crosslinked
R2 = 0.86
Slope = 0.99
C. P1 Closed
D. Open
E. Collapsed
F. P2 Closed
180°
180°
180°
180°
G.
P1 Wall
90°
90°
Core Wall

Figure 1. Photo-crosslinking identifies a uOmp-binding cradle-like groove in the "open" conformation of SurA. (**A**) The x-ray crystal structure of "P1 closed" SurA is shown(PDB: 1M5Y) with the core, P1, and P2 domains colored grey, blue, and red; respectively. (**B**) A linear correlation of the normalized crosslinking percentages of the 36 $SurA_{pAF}$ variants to $uOmpA_{171}$ (y-axis) and uOmpX (x-axis). Each point represents an individual $SurA_{pAF}$ variant (colored by the domain) and error bars correspond to the standard deviation (n = 3-5). The slope of the correlation crosslinking efficiencies to both SurA client uOMPs is 0.99 (R^2 = 0.86). The dotted lines delineate a set of residues (found in the core and P1 domains) with high crosslinking efficiencies to both uOmpA (> 70%) and uOmpX (>55%). (**C-F**) Structural models of four possible SurA conformations are depicted, keeping the color scheme from Panel A. The positions of the eight high-efficiency crosslinking sites are shown in pink on each model. **Panel C** shows "P1 closed" SurA, highlighting the fact that pink residues found on the core and P1 domains are on opposite sides of the protein (shown by 180° rotation). **Panel D** shows "open" SurA, with all eight pink residues co-localized around the putative uOMP binding groove. To achieve this conformation, the P1 domain not only undocks from the core domain, but also rotates to allow residues 245 and 260 to localize to the binding groove. **Panels E and F** show the "collapsed" and "P2 closed" conformations, which are also incompatible with the proposed uOMP binding model. **(G)** A structural model of the uOMP binding, cradle-like groove is lined by both hydrophobic and polar residues. Residues are shown with space-filling representation colored by hydrophobicity (Eisenberg scale; red = most hydrophobic, white = least hydrophobic).(78) The middle panel looks down into the cradle, showing the C-terminal helix of SurA that lines the bottom of the cradle, with the rest of SurA show as a semi-transparent cartoon representation. Images to the left and right show the walls of the cradle rotated 90°, which contain small hydrophobic patches (the rest of SurA is omitted in these images for clarity).

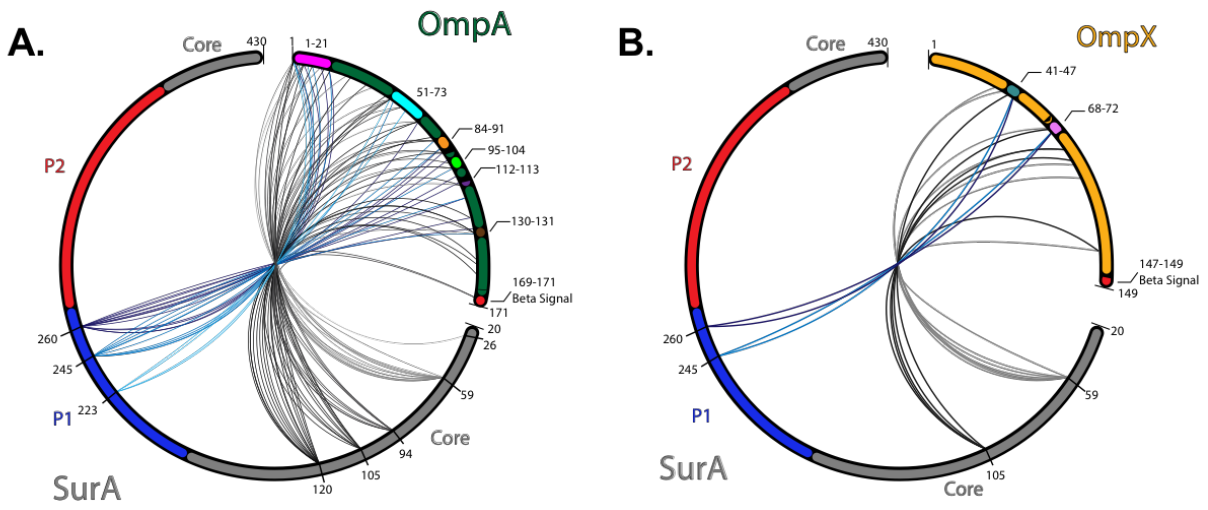

Figure 2. Cross-linking mass spectrometry identifies segments on two client uOmps that bind to SurA's groove. The SurA sequence is represented by the left side of both circles with structural domains colored as in Figure 1A. Each pAF variant used in XL-MS studies are labeled on the outside of the circle, according to their position in the sequence. The uOMP sequence that was crosslinked to SurA is shown on the right side of the circle, with $uOmpA_{171}$ colored green **(A)**, and uOmpX colored gold **(B)**. Each arc connecting the SurA and uOMP sequences corresponds to an identified crosslinked peptide. Specific SurA binding segments are highlighted with varying colors and are labeled for each uOMP. The β-signal is colored in red for both uOMPs, and has a very low density of crosslinked peptides identified.

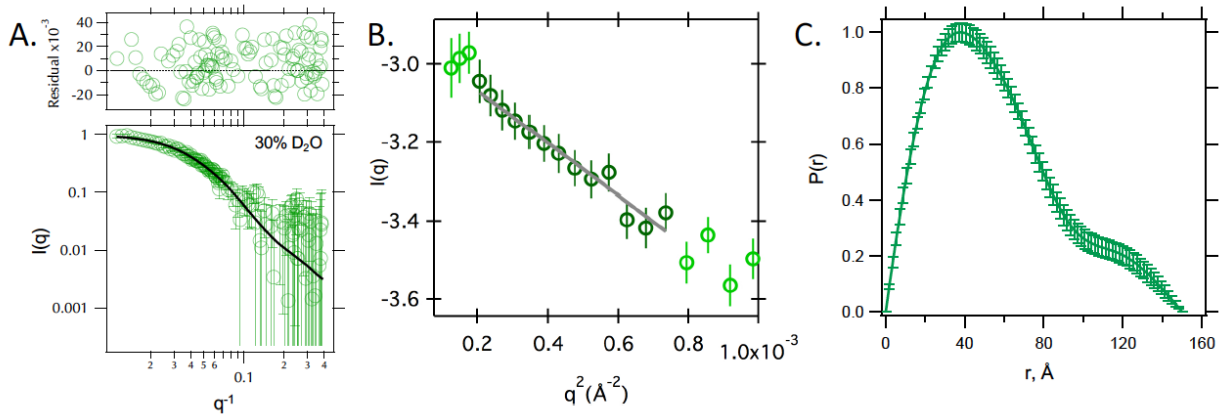

Figure 3. SANS of a SurA•uOmpA complex reveals an expanded uOmpA. (**A**) Raw scattering profile of proteo-SurA$_{105,pAF}$ chemically crosslinked to deutero-uOmpA$_{171}$ in 30% D_2O buffer is shown in green, along with a fit to the weight-average sparse ensemble of models created for this buffer condition(described further in Figure S13 and SI Methods). Error bars represent the standard error of the mean with respect to the number of pixels used in the data averaging. Randomly distributed residuals between data and model are shown above. (**B**) Linear fit of the Guinier region of the SANS profile determines the R_G of the complex is 45 Å. (**C**) *P*(*R*) distribution function inferred from the Guinier fit; D_{Max} is estimated to be 150 Å.

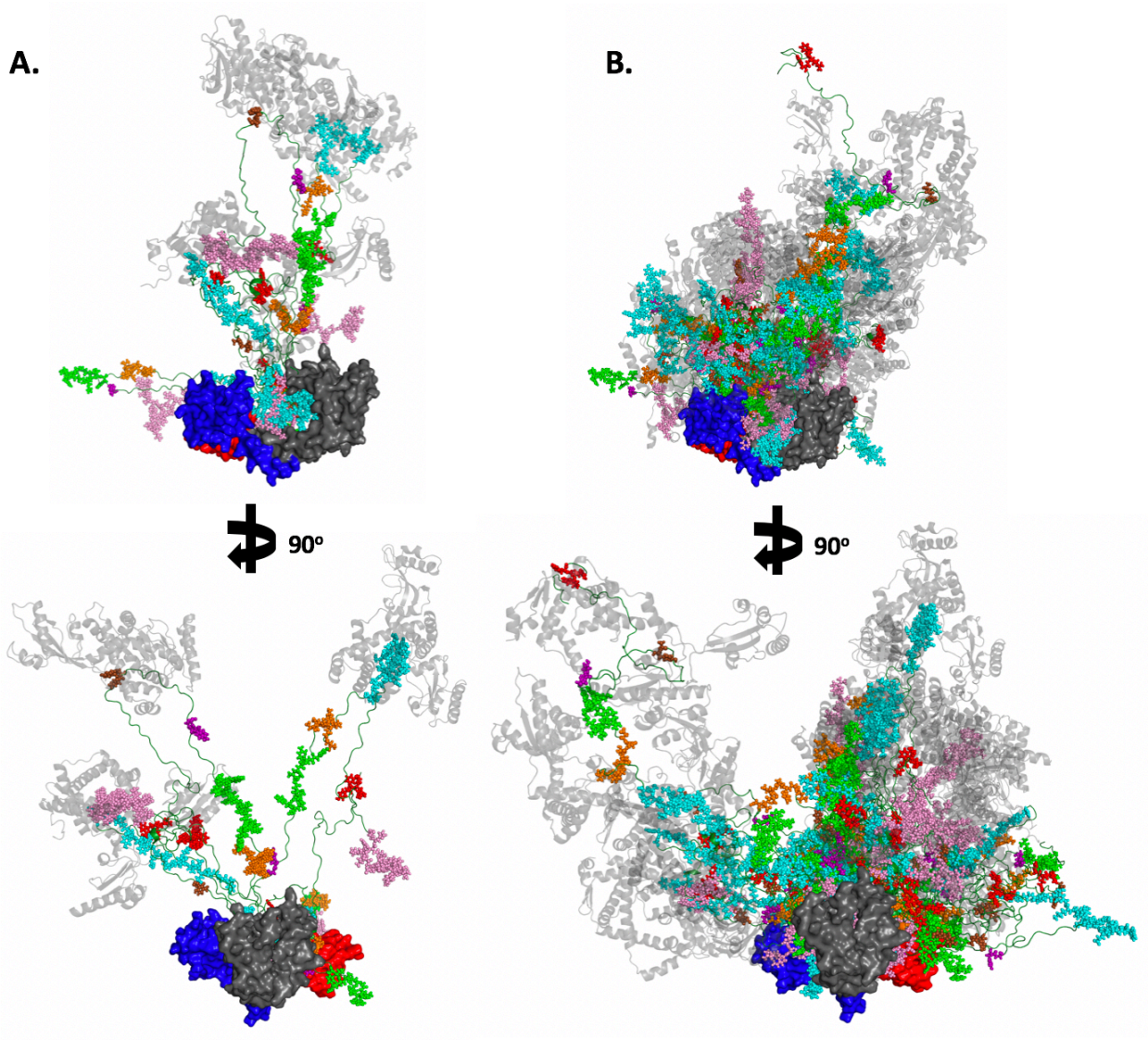

Figure 4. Structural models of sparse ensembles that reproduce SANS profiles by SASSIE analysis. (**A**) The six structural models included in the sparse ensemble that describe the 0% D_2O SANS dataset are shown aligned to SurA (shown in surface representation, colored as in Figure 1A). $uOmpA_{171}$ is shown in green, with identified SurA binding segments highlighted in space-filling representation (colored according to Figure 2A). Additional SurA monomers found to bind $uOmpA_{171}$ in this sparse ensemble (two-to-one and three-to-one complexes) are shown as semi-transparent cartoons. The β-signal for each $uOmpA_{171}$ conformation are colored red and shown in a space-filling representation and are extended away from SurA. **(B**) 36 structural models included in the sparse ensemble that describe the 98% D_2O SANS data-set are shown (representations and coloring are identical to Panel 4A).

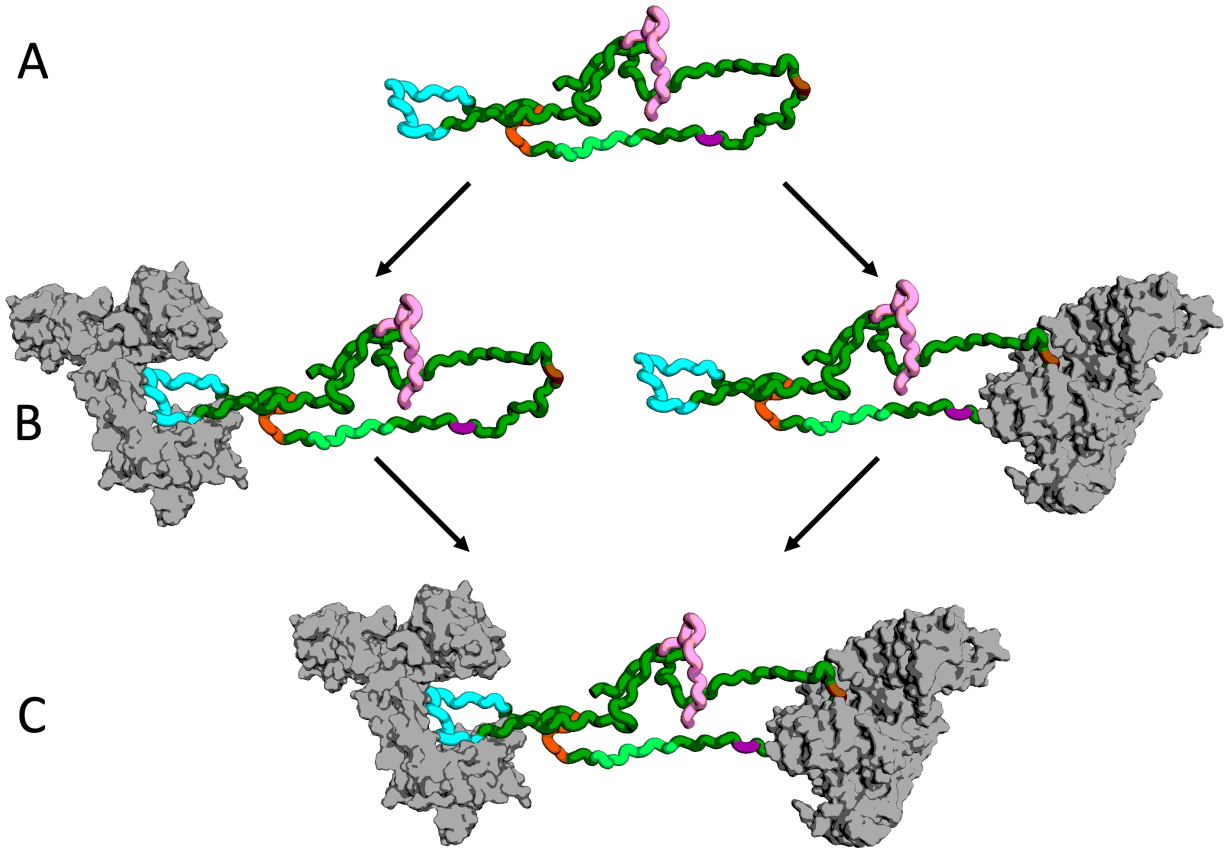

Figure 5. Sparse ensembles of SurA•uOmpA$_{171}$ complexes highlight dynamic occupancy of uOmpA$_{171}$ by SurA. (A) This conformation of uOmpA1$_{71}$ was represented multiple times in the sparse ensemble of structures which describe the 0% D_2O SANS dataset. SurA binding segments are colored as in Figure 2A. **(B)** Two one-to-one complexes were identified with SurA bound to different segments of uOmpA$_{171}$; residues: 51-73 (cyan) and 130-131 (brown). **(C)** The two-to-one complex with SurA bound to both binding segments on uOmpA$_{171}$ shown in Panel B, which is also represented in the 0% D_2O sparse ensemble. The presence of both one-to-one complexes and the two-to-one complex in our sparse ensemble highlights the ability of our modeling methods to capture possible states that together describe the dynamics that are used by SurA to solubilize uOmpA$_{171}$. Multiple SurA bound to a single uOMP allow for expansion of uOMPs while minimizing the chance for aggregation.

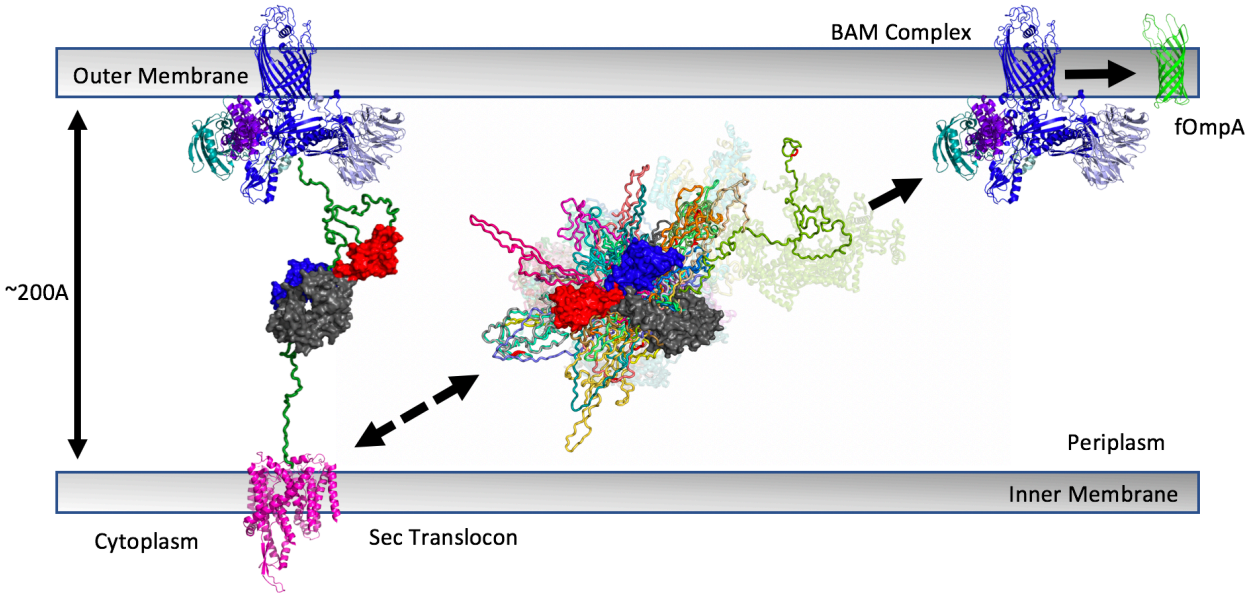

Figure 6. Proposed mechanism for SurA-dependent uOmp biogenesis. uOmps are post-translationally secreted through the Sec translocon, N-to-C terminally. The emerging uOMP N-terminus in the periplasm may be recognized by SurA. After complete translocation into the periplasm, one or more SurA protomers bind specific segments on uOMP clients, solubilizing the uOMP in highly expanded conformations roughly the size of the periplasm (shown on the left). Possible alignment of the translocon and BAM, which catalyzes folding into the outer membrane, could allow the cell to couple the expansion of uOMPs by SurA with chemical energy (ATP hydrolysis) in the cytoplasm to drive OMP biogenesis. The wide variety of conformations available to the SurA-bound uOMP (shown on the right) allow it to more easily encounter an unaligned BAM complex.. Our models suggest that SurA extends the C-terminal β-signal sequence on uOMPs (shown in red on the elongated, green $uOmpA_{171}$ model), which preferentially interacts with BAM in a mechanism reminiscent of fly fishing.