Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of β-barrel proteins

Anna Konovalova, David H. Perlman, Charles E. Cowles, and Thomas J. Silhavy

Department of Molecular Biology, Lewis Thomas Laboratory, Princeton University, Princeton, NJ 08544

Contributed by Thomas J. Silhavy, September 5, 2014 (sent for review August 20, 2014)

RcsF (regulator of capsule synthesis) is an outer membrane (OM) lipoprotein that functions to sense defects such as changes in LPS. However, LPS is found in the outer leaflet, and RcsF was thought to be tethered to the inner leaflet by its lipidated N terminus, raising the question of how it monitors LPS. We show that RcsF has a transmembrane topology with the lipidated N terminus on the cell surface and the C-terminal signaling domain in the periplasm. Strikingly, the short, unstructured, charged transmembrane domain is threaded through the lumen of β-barrel OM proteins where it is protected from the hydrophobic membrane interior. We present evidence that these unusual complexes, which contain one protein inside another, are formed by the Bam complex that assembles all β-barrel proteins in the OM. The ability of the Bam complex to expose lipoproteins at the cell surface underscores the mechanistic versatility of the β-barrel assembly machine.

In addition to the cytoplasmic or inner membrane (IM), Gram-negative bacteria such as Escherichia coli have an additional outer membrane (OM) (1). The OM is essential for survival; it serves as a selective permeability barrier and protects the cell from toxic hydrophobic substances such as deterrents, antibiotics, and other harmful chemicals (2). In E. coli, the OM is an asymmetric bilayer with phospholipids in the inner leaflet and LPS in the outer leaflet. LPS molecules have strong lateral interactions and are responsible for the barrier function of the OM. There are two main classes of proteins in the OM: the first class is integral transmembrane proteins often referred to simply as outer membrane proteins (OMPs). These proteins assume a conformation known as a hydrophobic β barrel. β barrels are assembled by the essential BamABCDE complex in which BamA is itself a β barrel (3). In contrast, the second class of proteins is lipoproteins, most of which have a soluble protein domain that is anchored to the OM by N-terminal lipid moieties.

The lipoprotein export pathway to the OM is well characterized (4). Lipoproteins are synthesized in the cytoplasm and are targeted to the inner membrane Sec translocon by a lipoprotein signal sequence (SS), which contains a “lipobox” with a highly conserved cysteine (Cys) residue to which a diglyceride is attached by a thioether bond (5). After lipid modification and SS processing, a third acyl chain is added to the Cys residue at the newly formed N terminus. Lipoproteins are sorted to either the IM or the OM based on a few amino acid residues, which directly follow the lipidated Cys by a process known in E. coli as the “α2 rule” (6). OM lipoproteins are targeted to the Lol machine, which consists of an inner membrane ATP-driven LolCDE complex that extracts lipoproteins from the IM (7); a periplasmic chaperone LolA, which escorts lipoproteins across the aqueous periplasmic space (8); and an OM acceptor lipoprotein LolB, which anchors lipoproteins to the inner leaflet of the OM (9). The assembly pathway for most lipoproteins is assumed to end here, leaving the molecule anchored in inner leaflet of the OM by lipid moieties with the protein domain facing the periplasm (4).

There are only a few reports about surface-exposed lipoproteins (10). In most of the cases, how these lipoproteins are transported across OM is not understood. Recently we have shown that a fraction of Braun’s lipoprotein or Lpp, the most abundant protein in E. coli, is exposed on the cell surface (11). Lpp exists in two forms: a bound form in which Lpp is covalently linked to peptidoglycan (PG) and a free form, the function of which is not understood (12, 13). We discovered that the free form of Lpp adopts an unusual conformation with its C terminus exposed on the cell surface (11). We wish to understand the mechanism by which lipoproteins are transported across the OM. However, because the function of the free form of Lpp is unknown, it is not a useful model for this analysis.

RcsF (regulator of capsule synthesis) is an OM lipoprotein that functions as the sensory component of the Rcs stress response (14). Rcs regulates the production of the colanic acid capsule, which stabilizes the OM and helps combat envelope stress (15). Capsule production results in a mucoid colony phenotype, which, together with a few available transcriptional lacZ fusions, serves as a functional read out of the system (16). Although some envelope-associated signals can be sensed by periplasmically expressed RcsF (17), the strongest inducers of the Rcs system are LPS defects, caused either by mutations in LPS biosynthesis (rfa) (18) or drugs that target LPS directly, such as polymixin B (19). Sensing these LPS defects by RcsF strongly depends on RcsF localization to the OM (17). Because LPS is present in the outer leaflet of the OM, we therefore

Significance

In Escherichia coli, most outer membrane (OM) lipoproteins are thought to be soluble proteins that are simply tethered to the inner leaflet of this membrane by lipid moieties attached to the N terminus. Here we show that lipoprotein RcsF (regulator of capsule synthesis) adopts a transmembrane orientation with the lipidated N terminus on the cell surface and the folded C-terminal domain in the periplasm. The short, unstructured, polar linker domain spans the hydrophobic membrane by passing through the lumen of several different OM β-barrel proteins. This remarkable, interlocked structure is formed by the Bam complex, which folds and inserts all β-barrel proteins in the OM, suggesting that this assembly machine translocates the lipid moieties and then folds the β barrel around the RcsF linker.

Author contributions: A.K. and T.J.S. designed research; A.K., D.H.P., and C.E.C. performed research; A.K., D.H.P., and T.J.S. analyzed data; and A.K., D.H.P., and T.J.S. wrote the paper.

The authors declare no conflict of interest.

1Present address: Department of Chemistry, Princeton University, Princeton, NJ 08544.
2Present address: Promega, Madison, WI 53711.
3To whom correspondence should be addressed. Email: tsilhavy@princeton.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1417138111/-/DCSupplemental.

PNAS | Published online September 29, 2014

www.pnas.org/cgi/doi/10.1073/pnas.1417138111
hypothesized that at least a portion of RcsF should be present on the cell surface.

RcsF consists of several domains (20, 21): a lipidated N-terminal Cys (residue 16), followed by a disordered proline-rich linker (residues 17–48), and a C-terminal folded domain (CTD, residues 49–134), the structure of which has been solved. The folded domain consists of a subdomain made up by short β-strand elements (residues 49–62) and a C-terminal RcsF core domain (residues 63–134) formed by a central three-stranded β sheet, one side of which is covered by a long α2 helix. The CTD is believed to function in signaling because it is sufficient to induce the Res system when expressed periplasmically (20). The linker domain therefore might be a sensory and/or a regulatory domain, as some linker mutations that result in moderate signaling phenotypes were reported (22). However, the exact role of the linker and how it is involved in sensing defects is unknown.

Here, we report that RcsF is indeed a surface-exposed lipoprotein, and it adopts a transmembrane conformation with the lipidated N terminus exposed at the cell surface and the core domain in the periplasm. Strikingly, the transmembrane domain crosses the OM through the lumen of β-barrel OMPs, which shield this highly charged sequence from the hydrophobic membrane interior. In addition to its ability to interact with several OMPs, RcsF also interacts with BamA, an essential component of the β-barrel assembly machine. We provide several lines of evidence that generation of RcsF/OMP complexes happens during OMP folding by the Bam complex. This demonstration that lipoprotein export across OM is tied to β-barrel assembly underscores the striking mechanistic versatility of the β-barrel assembly machine.

Results

**RcsF Is a Surface-Exposed Lipoprotein.** Several experimental approaches are commonly used to identify proteins exposed on the cell surface. Among them are selective surface biotinylation, surface proteolysis, and antibody-based whole cell assays (23). We used each of these methods to test whether RcsF can be detected on the cell surface (Fig. 1 and SI Appendix, Fig. S1). In a dot blot assay whole cells or a cell lysate of a WT or ΔrscF mutant strain were spotted onto a nitrocellulose membrane and probed with anti-RcsF polyclonal antibodies. We also used antibodies available in our laboratory against other lipoproteins (BamDCE and LptE) with known topology (Fig. 1A) as a control for cell envelope integrity. Because BamDCE and LptE are members of multiprotein complexes that might affect their detection, we also generated OM-targeted GFP (OM-GFP) by fusing the Lpp signal sequence to sGFP as an additional control. Anti-RcsF antibodies bind to the whole cells and cell lysates in an RcsF-dependent manner, whereas other control antibodies bind to the lysate but not whole cells (Fig. 1A). RcsF can also be labeled by an OM impermeable NHS-LC-LC biotin probe, and it is partially sensitive to an extracellularly added protease (SI Appendix, Fig. S1A and B). Taken together, these results suggest that at least a part of the RcsF molecule is exposed on the cell surface.

Next, we tested whether surface exposure of RcsF is dependent on an OM targeting signal. For that we generated two mutant RcsF proteins: one in which we introduced a Lol avoidance mutation (SM→DQ substitution at positions +2 and +3) that resulted in retention of RcsF in the inner membrane (RcsF<sup>ΔLol</sup>) (17) and one in which we replaced the lipoprotein signal sequence of RcsF with the well-characterized signal sequence of OmpA, which resulted in production of RcsF lacking an N-terminal lipid (RcsF<sup>SS-OmpA</sup>). As shown in Fig. 1B, RcsF<sup>ΔLol</sup> and RcsF<sup>SS-OmpA</sup> accumulated at WT levels but could not be detected on the cell surface, suggesting that OM targeting of RcsF by the Lol pathway is a prerequisite for surface localization.

To address the OM topology of RcsF, we used an epitope walking approach (23), which was successfully used to determine the topology of several *E. coli* OMPs (24–26). We introduced a FLAG epitope along the linker, as well as in the surface-exposed loop regions of the core domain. Of seventeen FLAG insertion variants made, nine were functional for signaling based on their ability to complement WT and activate the Res system in an rfa mutant background (SI Appendix, Fig. S1C). These functional FLAG insertion mutants were subjected to a dot blot assay with anti-FLAG and anti-RcsF antibodies (Fig. 1C). These FLAG insertions did not reduce RcsF levels or interfere with surface localization as anti-RcsF antibodies detected all mutant RcsFs on the cell surface. In contrast, anti-FLAG antibodies could not detect FLAG insertions in the core domain of RcsF. Interestingly, the first surface-detected FLAG insertion (R21) is only few amino acids away from the lipidated Cys residue, suggesting that the N-terminal lipid is anchored in the outer leaflet of the OM.

**Fig. 1.** RcsF is a surface-exposed lipoprotein. (A) RcsF is detected on the cell surface in a dot blot assay. Whole cells or cell lysates of a WT or ΔrscF strain were spotted on the nitrocellulose membrane and probed with corresponding antibodies. For OM-GFP, the expression plasmid was introduced into the indicated strains. (B) Surface exposure of RcsF depends on OM targeting signal sequence. Whole cells expressing the indicated RcsF variants were probed with anti-RcsF antibodies in a dot blot assay (Upper) and total RcsF protein levels were analyzed by an immunoblot assay with anti-RcsF antibodies. (C) The N terminus of RcsF is surface exposed. RcsF<sup>ΔLol</sup> and RcsF-FLAG variants were expressed from a plasmid and whole cells or cell lysates were probed with anti-FLAG, anti-RcsF and anti-BamC (as a control of cell envelope integrity) antibodies in a dot blot assay. Total RcsF protein levels were analyzed by an immunoblot assay with anti-RcsF antibodies. RcsF-FLAG variants are indicated by the amino acid number after which the FLAG tag was introduced.
In addition to Cys16, the RcsF sequence encodes four Cys residues, which form two nonconsecutive disulfide bonds to stabilize the structure of the core domain (20, 21). We performed cell surface labeling with Cys-specific cell-impermeable Mal-PEG label (20 kDa) to test whether the core domain of RcsF is present on the cell surface. For that we treated whole cells with a reducing agent to reduce disulfide bonds and then labeled cells with Mal-PEG. We observed labeling of BamA as expected (27) but not RcsF, even under conditions of RcsF overexpression (SI Appendix, Fig. S1D). Absence of RcsF Cys labeling further reinforced our epitope walking results, showing that the core domain of RcsF is not exposed on the cell surface.

These results suggest a transmembrane topology for RcsF with the lipided N terminus and the linker domain surface exposed and the core domain in the periplasm. We refer to this topology as “N out-C in.” This result is puzzling because the RcsF sequence does not encode a canonical transmembrane domain and the sequence that appears to span the membrane is highly charged. Therefore, to explain the transmembrane topology of RcsF, it seemed necessary to propose that surface-exposed RcsF must exist in a complex with a transmembrane OMP.

**RcsF Interacts with Several OMPs and BamA.** To identify the proteins interacting with RcsF, we used an in vivo chemical cross-linking approach in combination with high-resolution MS-based proteomics. For that, we generated a functional RcsF-Strep II tag fusion that allows affinity purification and used ethylene glycol bis(succinimidylsuccinate) (EGS) cross-linking to stabilize protein interactions. We purified RcsF-Strep using Strept-Tactin resin and subjected elution samples to SDS/PAGE, followed by in-gel digestion and high-resolution nano-flow ultra-high-performance LC MS (UPLC-MS). As a negative control for binding specificity, we used RcsF lacking a Strep tag, which was cross-linked and treated identically to the RcsF-Strep sample.

We identified three OMPs, OmpA, OmpC, and OmpF, as well as BamA as RcsF-interacting proteins, based on their prominent spectral count values (Table 1 and SI Appendix, Table S1). These proteins purified with RcsF-Strep but not RcsF (no tag) on the Strept-Tactin column, suggesting specific interaction. Because EGS can cross-link multiple members of protein complexes, we sought to confirm that RcsF interactions with OmpA, C, and F and BamA are direct. To confirm the interactions and to identify specific sites of RcsF interaction with the OMPs, we used site-specific photo-activated cross-linking with a genetically encoded BPA amino acid (28).

We generated 30 functional RcsF-Strep variants with BPA sites distributed along the linker as well as throughout the surface of the core domain and tested their ability to cross-link to OMPs by Western blot analysis (Fig. 2 and SI Appendix, Fig. S2). We observed that RcsF cross-links to OMPs at 10 sites of the linker domain and only a few surface sites in the core domain, many of which are in the linker facing region of the core domain (Fig. 2). Strikingly, the 10 cross-linking sites of the linker and β1α1 subdomain (RcsF 28–60) create a continuous interface for the RcsF/OMPs interaction. We also noticed that the most robust cross-linking to BamA occurs along β strands of the core domain, suggesting that RcsF interacts with BamA in a distinct fashion from other OMPs (Fig. 2).

Our site-specific cross-linking data and RcsF topology suggest that the linker/β1α1 subdomain mediates interaction with OMPs, raising the possibility that this region may traverse the barrel lumen of the OMPs to cross the OM. To identify residues of OMPs that interact with RcsF, we used several RcsF-Strep variants with pBPA at the linker sites followed by in vivo UV cross-linking, Strept-Tactin purification, SDS/PAGE with in-gel digestion, and high-resolution nano-flow UPLC-MS to sequence and map cross-linked peptides. We were able to identify four such cross-linked species: two cross-linked peptides of RcsF G60pBPA to OmpC and two to OmpF (Fig. 3A and B).

Both identified sites of cross-linking for OmpC and both for OmpF are derived from unique and independently characterized cross-linked peptides. However, they fall into the same specific region of the OmpC/F barrel based on the crystal structure (Fig. 3C) (29, 30). Strikingly, these residues face the lumen of the barrel (Fig. 3C), providing strong evidence that at least part of RcsF molecule resides inside the OmpC/F barrel. The placement of identified residues is consistent with our epitope walking results. The last epitope insertion that was detected on the cell surface follows amino acid V59. The FLAG epitope is eight amino acids long and would be exposed on top of the barrel based on the site of G60pBPA cross-linking.

While analyzing RcsF G60pBPA cross-linked samples, we also detected by MS a peptide with a N-terminal palmitoyl (C16:0) modification on a Cys residue (SI Appendix, Fig. S3). In addition, lipidated RcsF and nonlipidated RcsF migrate differently in SDS/PAGE, and we are unable to detect a band corresponding to nonlipidated RcsF in the WT strain (SI Appendix, Fig. S3). This result provides evidence that the lipid modifications are still present on RcsF in RcsF/OMP complexes.

We tested the effect of OMP deletions on RcsF surface exposure. Mutants individually deleted for ompA, ompC, or ompF, as well as an ompR mutant defective in transcription of both ompC and ompF genes, did not show decreased levels of surface exposed RcsF (SI Appendix, Fig. S2) nor did they exhibit defects in RcsF signaling except for the ompR mutation, which leads to twofold activation of RcsF. We cannot analyze the triple mutant, ompACF, because its OM is quite unstable due to the loss of all three major OMPs. Because of this OMP redundancy for RcsF localization, RcsF surface exposure is not dependent on any single OMP.

**RcsF/OMP Complexes Are Formed During OMP Assembly by Bam Machine.** RcsF also interacts with BamA; however, the RcsF/BamA interaction is distinct from the interaction with other

---

**Table 1. Quantitative analysis of RcsF-interacting proteins identified in the EGS cross-link pull-down experiments**

| Protein | RcsF-Strep tag | RcsF no tag (negative control) |
|---------|----------------|-------------------------------|
|         | Replicate 1 | Replicate 2 | Replicate 3 | Replicate 1 | Replicate 2 | Replicate 3 |
| RcsF    | 71 7 | 88 7 | 86 7 | 0 0 | 2 2 | 3 3 |
| OmpA    | 81 18 | 91 18 | 78 20 | 0 0 | 0 0 | 2 2 |
| OmpC    | 49 16 | 79 20 | 49 16 | 0 0 | 0 0 | 0 0 |
| OmpF    | 49 14 | 50 15 | 38 13 | 0 0 | 0 0 | 0 0 |
| BamA    | 44 25 | 58 31 | 52 25 | 0 0 | 0 0 | 0 0 |

---

Konovalova et al.
OMP. Because of this, and because it is difficult to imagine how the lipidated linker region of RcsF could enter the lumen of a folded OMP, we hypothesized that RcsF/OMP complexes are formed during OMP folding by the Bam complex. To test this hypothesis, we used two different approaches.

In the first approach, we tested whether mutations in the Bam complex can affect levels of RcsF/OMP complexes in vivo. We focused on the RcsF/OmpA complex because, unlike OmpC and OmpF, OmpA levels are not dramatically reduced in the different of bam mutants, simplifying the interpretation of the results (SI Appendix, Fig. S4).

In vivo EGS cross-linking results in the formation of an RcsF/OmpA complex of molecular weight of ~50 kDa, which corresponds to a 1:1 ratio of RcsF to OmpA (Fig. 4A). As expected, RcsF/OmpA cross-linking requires the RcsF lipoprotein signal sequence (SI Appendix, Fig. S5). Two other prominent cross-linked bands of molecular weight of ~35 and 22 kDa were observed. The 22-kDa band is likely to be an RcsF cross-linked dimer. The 35-kDa band is the product of RcsF/Lpp cross-linking (SI Appendix, Fig. S6). We have not explored either of these interactions further. However, we note that Lpp is the most abundant protein in E. coli, and an lpp-null mutation did not affect either RscF signaling or its surface exposure.

To test whether mutations in the Bam complex can affect levels of RcsF/OMP complexes, we used two different mutations in bam that affect Bam function to different extents. bamA101 is a mutation in the promoter region of bamA that causes a reduction in BamA levels (31). Thus, in this mutant, fewer Bam complexes are formed, but those Bam complexes that do form are made of WT proteins. The bamA101 mutation does not confer significant OM assembly or permeability phenotypes. Strains carrying the double mutation bamA616 and bamB exhibit compromised Bam function (SI Appendix, Fig. S4). Importantly, levels of RcsF and OmpA are not affected in bamA101 and bamA616 bamB mutants under the conditions used (Fig. 4B). However, we observed a twofold decrease of RcsF/OmpA cross-linking in bamA101 and sevenfold decrease in bamA616 bamB. This decrease was specific to RcsF/OmpA because RcsF/Lpp cross-linking did not decrease in these mutants.

Both bamA101 and bamA616 bamB are Rsc-activating mutations judged by rprA-lacZ reporter fusions, and they also confer a mucoid phenotype. Therefore, decreased levels of RcsF/OmpA complexes in these strains could be a result of RscF activation. To test this, we performed EGS cross-linking in an Rsc-activating rfaP mutant. In the rfaP strain, levels of RcsF/OmpA were similar to that of the WT (Fig. 4B). Therefore, we attribute the decrease in RcsF/OmpA cross-linking in the bam mutants directly to the reduced activity of the Bam machine.

In the second approach, we tested whether RscF can interact with OmpA in vitro using purified proteins. We took advantage of a simplified system, in which OmpA can be folded in vitro, in the presence of detergent. Detergent-folded OmpA has the same properties as OmpA folded by the Bam machine in vivo, e.g., the barrel is heat modifiable and protease resistant (33). We set up our experiment under two different reaction conditions. In the first reaction, we refolded OmpA in detergent and then added RscF to the folded OmpA (Fig. 5A). In the second reaction, we folded OmpA in the presence of RscF (Fig. 5B). We used UV photo cross-linking of RscF-pBPA variants as measure of the RscF/OmpA interaction. We purified an RscF-pBPA variant at the linker site that interacts with OmpA in vivo (RscF K40pBPA), as well as an RscF-pBPA variant that does not interact with OmpA in vivo (RscF D72pBPA) to serve as a negative control. When RscF is added to previously refolded OmpA, no product corresponding to RscF x OmpA was formed after UV treatment (Fig. 5A). When OmpA was refolded in the
Fig. 3. RcsF G60pBPA cross-links inside the lumen of OmpC and OmpF. Culture expressing RcsF-Strep G60pBPA was subjected to in vivo UV cross-linking. RcsF-Strep G60pBPA complexes were affinity purified and resolved on SDS/PAGE. Cross-linked peptides were rendered by chymotryptic digestion and subjected to high-resolution MS-based proteomic analysis. A and B contain MS/MS spectra assigned to RcsF-OmpC and RcsF-OmpF cross-linked peptide species, with their b-and y-type fragment ions indicated in red and blue, respectively. Prominent ions are labeled with their empirical m/z values and b or y ion assignments. All matched fragment ions are additionally represented as flags on the sequence diagrams above the spectra. The pBPA residue is indicated by X. (C) Periplasmic and side view of the extracellular loop 3 in the lumen of the OmpC and OmpF barrel, based on the crystal structures (PDB ID codes 2J1N and 2ZFG, respectively). Sites of RcsF cross-linking identified in A and B are highlighted in red.
Mutations in the Bam complex affect levels of RcsF/OmpA complexes

Published online September 29, 2014
PNAS

E4355

RcsF/OmpA complexes are formed during OmpA folding in vitro. (PNAS PLUS)

∆EGS (linked RcsFxOmpA band has the expected size of 50 kDa. (with EGS and analyzed by immunoblot with anti-RcsF antibodies. The cross-

Konovalova et al.

RcsFxOmpA complex for RcsF K40

presence of RcsF, we observed UV-induced formation of the

RcsFxOmpA complex in vivo. (A) Cells of indicated strains were subjected to in vivo cross-linking with EGS and analyzed by immunoblot with anti-RcsF antibodies. The cross-linked RcsFxOmpA band has the expected size of 50 kDa. (B) In vivo cross-linking with EGS (Upper) performed as in A. Total OmpA and RcsF levels in each strain were analyzed in un–cross-linked samples (Lower) by anti-OmpA and anti-RcsF antibodies. Intensity of the OmpA, RcsF, and RcsF/OmpA bands were quantified by ImageJ and normalized to the WT.

Discussion

Based on the data presented, we propose a model (Fig. 6) in which the lipidated N terminus and linker domain of RcsF are surface exposed, whereas the core domain remains in the periplasm. The short, but highly charged sequences of RcsF that traverse the OM do so by passing through the lumen of OMPs, and the formation of these RcsF/OMP complexes occurs during OMP folding by the Bam machine. RcsF does not have barrel specificity; RcsF can interact with multiple OMPs. This lack of OMP specificity explains why RcsF surface exposure is not dependent on any one single OMP. The RcsF/OMP complex provides another example of an OM protein complex that forms during assembly from components targeted to the OM by different pathways. In addition, it suggests that the Bam complex can translocate lipid-containing protein sequences from one membrane face to the other.

OM lipoproteins have a variety of important functions: they participate in envelope biogenesis, stress responses, cell adhesion, virulence, etc. (10). Most of what we know about lipoprotein biogenesis comes from the Tokuda laboratory, where the Lol pathway was identified and characterized in great detail (4). For a long time, it has been assumed that anchoring of lipoproteins to the inner leaflet of OM by LolB is the last step in lipoprotein transport and that lipoproteins remain in this orientation with the protein domain facing the periplasm (4). Only recently have we started to appreciate that lipoproteins can be exposed on the cell surface (10). Most of these surface-exposed lipoproteins were identified in pathogenic bacteria because of their function in adhesion and virulence and because of their antigenic properties. However, there are only few examples where the signals and cellular factors required for surface exposure have been studied.

![Figure 4](https://example.com/fig4.png)

**Fig. 4.** Mutations in the Bam complex affect levels of RcsF/OmpA complexes in vivo. (A) Cells of indicated strains were subjected to in vivo cross-linking with EGS and analyzed by immunoblot with anti-RcsF antibodies. The cross-linked RcsFxOmpA band has the expected size of 50 kDa. (B) In vivo cross-linking with EGS (Upper) performed as in A. Total OmpA and RcsF levels in each strain were analyzed in un–cross-linked samples (Lower) by anti-OmpA and anti-RcsF antibodies. Intensity of the OmpA, RcsF, and RcsF/OmpA bands were quantified by ImageJ and normalized to the WT.

![Figure 5](https://example.com/fig5.png)

**Fig. 5.** RcsF/OmpA complexes are formed during OmpA folding in vitro. (A) Urea denatured FLAG-OmpA was refolded in TBS with 0.1% Triton X-100 for 2 h and then an equimolar amount of purified RcsF-Strep K40pBPA (positive for in vivo cross-linking), and RcsF-Strep D72pBPA (negative for in vivo cross-linking) was added; an equal volume of buffer was added to a control reaction (−). Reactions were incubated for additional 2 h and then subjected to UV cross-linking. Reactions were stopped by the addition of SDS-loading buffer and split into unboiled (−) and boiled (+) samples and analyzed by immunoblot for the presence of the RcsFxOmpA band with anti-FLAG antibodies. Note, folded OmpA (f) remains folded in −boil samples and migrates faster on SDS/PAGE than unfolded OmpA (u). (B) Urea denatured FLAG-OmpA was refolded in TBS with 0.1% Triton X-100 in the absence (−) or presence of an equimolar amount of the corresponding RcsF-Strep pBPA variant for 2 h. Samples were UV cross-linked and treated as described in A.
The first reported surface-exposed lipoprotein is pullulanase, PulA, in *Klebsiella oxytoca* (34, 35). Identification of PulA led to the discovery of its transport machinery (36), which later turned out to be a type II secretion system (T2SS) like that found in many different bacteria (37). Although Pul is a well-studied prototypical T2SS, lipoproteins are rather unusual substrates because most T2SS substrates are soluble proteins released into the media (38). Even though we know a lot about T2SS structure and function, it is still unclear how exactly PulA is anchored to the OM after it has been exported through the OM secretin PulD (37). Precise topology is known for only two surface-exposed lipoproteins in *E. coli*, CsgG and Wza, and they are very different from RcsF. The lipidated N terminus of both is anchored to the inner leaflet of the OM, and both proteins have a helical transmembrane helix that oligomerizes to form an octomeric pore in the OM (39, 40). Accordingly, the assembly of these proteins is likely to be Bam independent.

In other bacteria, such as spirochetes, lipoproteins are believed to be anchored in the outer leaflet of the OM and surface exposed by default (41, 42). Both model lipoproteins OspA and OspC have a common feature: the presence of the N-terminal disordered linker domain, which serves as a cis-acting determinant for surface exposure (43, 44). It is unclear how these lipoproteins are transported across the OM, but the presence of an OMP that serves as a lipoprotein “flippase” has been suggested (43).

We previously reported that a fraction of Lpp, the most abundant lipoprotein in *E. coli*, is exposed on the cell surface (11), whereas the other fraction remains in the periplasm where it is covalently bound to PG (12, 13). It is hard to determine whether RcsF, like Lpp, exists in two different cellular locations. The labeling and cross-linking techniques used in this study are known to be low efficiency and hence cannot be used for direct quantification, and there is no functional assay to distinguish between surface exposed and periplasmic RcsF. However, we raise the possibility that RcsF might exist in different locations to sense diverse envelope defects and that perhaps the RcsF/OMP complex is used to detect certain alterations to the outer leaflet of the OM.

We used site specific cross-linking to identify residues of RcsF that interact with the OMPS and BamA. We showed that the linker together with the β1α1 subdomain create a continuous interaction interface with the OMPS. We also observed cross-linking along one face of the core domain. Two topology models could explain these cross-linking results. In the first model, the C-terminal end of the RcsF linker and the β1α1 subdomain would lie inside the barrel, whereas the core domain would interact with OMPS on the periplasmic side of the membrane. In an alternative model, the core domain also would reside inside the barrel and one side of its surface would be occluded from OMP interaction by the linker domain. Both of these models suggest that the core domain would not be detected on the cell surface and could explain our epitope walking results. However, we favor the first model, because the barrels of the RcsF-interacting OMPS simply are not big enough to accommodate the folded core domain.

In both reported structures of the soluble domain of RcsF, β1α1 creates a separate subdomain on the side of RcsF core (20, 21). However, because this subdomain is somewhat distant from the core and lacks extensive contacts with the core, it might exist in a different orientation in the full length RcsF. Because sites of cross-linking to OMPS follow the linear sequences of the linker and β1α1 subdomain, we suggest this subdomain exists in an extended conformation in the RcsF/OMP complexes.

We mapped OmpC and OmpF peptides that cross-link to RcsF G60pBPA. Two peptides were found for each protein, and both peptides mapped to the same region of the corresponding OMP based on their crystal structure (29, 30). OmpC and OmpF are structural homologs, and the fact that the RcsF interaction sites were similar for OmpC and OmpF suggests that this dependence is highly specific. Strikingly, the interaction site mapped to the extracellular loop 3 (L3), which folds back and occludes part of the lumen of both OmpC and OmpF. L3 has been previously implicated in the interaction with colicins, which use the lumen of OmpF as an entry way into the cell. Mutations in L3 are known to confer colicin resistance (45), and OmpF was successfully crystalized with the T83 peptide of colicin E9 and E3 inside the barrel lumen (30, 46). The T83 peptide is an intrinsically unstructured part of the N-terminal transport domain of E colicins, and it is the first domain to enter OmpF lumen (47). Indeed, recent studies show that the unstructured colicin E9 N-terminal domain threads through two adjacent OmpF trimer subunits in opposite orientation (48). Our results suggest the nature of the RcsF unstructured linker interaction with OmpC/F might be similar to that of unstructured colicin linker.

We do not know how RcsF can be accommodated in the OmpA barrel. The available OmpA structure consists of two domains: an eight-stranded, N-terminal β-barrel domain and C-terminal periplasmic, peptidoglycan-binding domain (49). This small β-barrel domain of OmpA was subjected to extensive structural studies (50–52). These studies suggest that the OmpA lumen does not form a channel because of the presence of a salt bridge. However, recent biochemical studies have proposed an electro-chemical gate opening model in which the salt bridge can “open up,” thus allowing the β barrel to function as a pore (53). Regardless of the pore state of OmpA, it is highly unlikely that the eight-stranded β-barrel domain of OmpA could accommodate the RcsF linker. However, there is also some biochemical, as well as physiological evidence that OmpA can adopt a so-called large barrel conformation in which the C-terminal domain can join the narrow barrel to form a large pore (54). The presence of the large pore conformation in vivo has remained exceedingly difficult to prove or disprove experimentally, and for this reason, this OmpA structure has been a matter of debate for decades. It is possible to delete the periplasmic domain, leaving the eight-stranded OmpA β barrel intact (55). Attempts to cross-link RcsF to this truncated OmpA have failed, perhaps indicating that RcsF interacts with the larger OmpA β barrel. However, because the structure of the large barrel remains unknown, it is difficult to test whether or not RcsF is inside the OmpA barrel. More work is required to clarify the nature of the RcsF-OmpA interaction.

Site-specific cross-linking demonstrates that the nature of the interaction of RcsF with BamA is distinct from its interaction with other OMPS. We suggest that RcsF interacts with BamA
during RcsF/OMP assembly and that RcsF/OMP complexes are the product of this assembly reaction. Our in vitro experiments support this model by showing that RcsF/OmpA complexes must form during OmpA folding. Our in vivo data further support our model because mutations, which decrease Bam machine activity, lead to a reduction of RcsF/OmpA complexes, and the level of reduction correlates with the severity of the mutations. Whether or not the RcsF/BamA cross-linked complexes represent true assembly intermediates will require additional analysis.

It is not unprecedented that the RcsF/Omp complex requires β-barrel folding and the Bam machine to form. The LptD/E complex is the first example of a β-barrel/lipoprotein complex that requires Bam for assembly. LptD/E is the OM component of the LPS transport machine (56, 57). LptE is an OM lipoprotein that resides inside the lumen of a β-barrel protein LptD (58). LptE has several functions: it is required for LptD assembly (59, 60), plugging the LptD barrel (61), and LPS export across the OM (56, 62). However, the RcsF/Omp complexes are fundamentally different from LptD/LptE. LptE resides almost completely inside the lumen of a large 26-stranded barrel of LptD (63, 64). Interaction interface between two proteins is extensive, with a surface area >3,000 Å² (63, 64). Moreover, folding of the LptD barrel absolutely requires not only Bam but also the presence of LptE. No LptD exists without its lipoprotein partner (59, 60). In contrast, RcsF interacts with OMPs via a short sequence that exists in an extended conformation; RcsF does not have strict barrel specificity, and it can interact with multiple OMPs demonstrating a lack of extensive sequence-specific contacts. Unlike LptD, any of RcsF interacting OMPs can fold and function independent of RcsF. Based on ribosomal profiling, RcsF abundance was estimated to be around 3,000 molecules per cell. OmpA, OmpC, and OmpF are the most abundant OMPs, and they exist in at least hundred-fold excess to RcsF (65). Therefore, RcsF is unlikely to interfere with OMP functions, and RcsF simply uses these OMPs to adopt its OM topology and reach the cell surface.

There is yet another fundamental difference between LptD/E and RcsF/Omp. Our results indicate that the lipid is still attached to RcsF, and they show that a FLAG tag located just four amino acids from the N-terminal Cys residue is exposed on the cell surface. This result suggests that the N-terminal lipids of RcsF are exposed in the inner leaflet as they are in LptE, but rather in the outer leaflet of the OM. How the lipids are flipped across the membrane to assume this outer leaflet localization remains unknown. Recently, it has been suggested that BamA causes a local distortion of the OM bilayer that allows BamA to release folded OMPs into the membrane (66, 67). Such a distortion might also create a favorable environment for the RcsF lipids to flip, and thus BamA might serve the second function in RcsF/Omp biogenesis by catalyzing lipid translocation. Whether Bam activity is sufficient for lipid translocation and other mechanistic aspects of RcsF/Omp assembly by Bam will require additional research.

Bam is a β-barrel assembly machine that is also known to be able to assemble very complex substrates such as multimeric OMPs, autotransporters, which are β-barrels with secreted domains, and LptD/E. These substrates are fundamentally different in their structures, and their assembly pathways must require distinct activities of Bam. RcsF/Omp complexes represent yet another class of substrates assembled by this truly remarkable machine that operates in the absence of any obvious energy source.

### Experimental Procedures

**Dot Blot Assay.** Strains were grown to midlog and concentrated to OD₆₀₀ = 1 in PBS (Na₂HPO₄, 10 mM; KH₂PO₄, 1.8 mM; KCl, 2.7 mM; NaCl, 137 mM; pH 7.4). Cell lysates were prepare by sonication of the cell suspension in PBS supplemented with 10 mM EDTA. Two microliters of cell suspension or cell lysate was spotted on a nitrocellulose membrane and air dried. Membranes were blocked with 2% (w/v) nonfat dried milk (NFDM) in PBS for 30 min at room temperature and probed with indicated antibodies (SI Appendix, Table S4) for 1 h at room temperature. Membranes were washed five times for 3 min with PBS and probed with secondary antibodies for 1 h at room temperature.

**In Vivo EGS Cross-Linking.** Strains were grown to midlog, washed twice in PBS, and concentrated to OD₆₀₀ = 10 in PBS. Cells were incubated for 10 min at 37 °C, and EGS (Pierce) was added to a final concentration of 1 mM for 20 min at 37 °C and then quenched with Tris-Cl (100 mM final). Cells were pelleted and resuspended in SDS loading buffer and analyzed by immunoblotting.

**Purification of EGS Cross-Linked RcsF and MS.** Strains expressing RcsF-Strep tag or RcsFmut were grown in 1.5 L of Luria-Bertani (LB) medium with 0.02% PO₃, 0.02% BPA or RcsF-Strep D72 was grown in 1.5 L of LB with 0.16 mM pBPA (Bachem) until midlog. Cells were concentrated in ice-cold Tris-buffered saline (TBS) to OD₆₀₀ = 1. The cell suspension was subjected to UV irradiation at 365 nm for 15 min at 4 °C, 2x SDS sample loading buffer was added, and samples were analyzed by immunoblotting.

**Identification of RcsF G60pBPA Cross-Linked Peptides.** Culture expressing RcsF-Strep G60pBPA was grown in 1.5 L LB with 0.16 mM pBPA and subjected to in vivo UV cross-linking as described above. RcsF-Strep G60pBPA was subjected to purification and UPLC-MS as described in SI Appendix, Experimental Procedures. pBPA cross-linked peptides were mapped with the StavroX 3.2.2 software (69). To validate cross-linked peptide matches made by StavroX, we determined the most abundant G60pBPA cross-linking peptides and matched them with high confidence by the SEQUEST HT search engine within Proteome Discoverer from a purified, non-cross-linked RcsF G60pBPA protein sample. We then programmed the mass of this peptide into SEQUEST HT as a variable modification specified at any residue. Identical RcsF-OmpC and RcsF-OmpF cross-linked peptides were found by SEQUEST HT as were mapped by the StavroX software. Details of the StavroX and SEQUEST HT searches are described in SI Appendix, Experimental Procedures.

**In Vivo Detergent Folding of OmpA in the Presence of RcsF.** FLAG-OmpA was expressed and purified as previously described (70). RcsF-Strep K40pBPA and RcsF-Strep D72pBPA were purified as described in SI Appendix, Experimental Procedures. In vitro refolding of OmpA was performed under two different conditions (1): 5 μM of FLAG-OmpA in 8 M urea was 10-fold diluted in TBS with 0.1% Triton X-100 and incubated at room temperature for 2 h. Purified RcsF-Strep K40pBPA and RcsF-Strep D72pBPA in TBS with 0.1% Triton X-100 was then added to a final concentration of 0.5 μM. For the control, an equal volume of buffer was added instead of RcsF. Reactions were incubated for 2 h at room temperature and then were split in half, and one sample was subjected to UV cross-linking at 365 nm for 15 min at 4 °C, 4x SDS sample loading buffer was added to each sample, which was split in half again, and one half was boiled for 10 min. Samples were analyzed by immunoblotting (2). FLAG-OmpA (5 μM) in 8 M urea was 10-fold diluted in TBS with 0.1% Triton X-100 containing 0.5 μM of purified RcsF-Strep K40pBPA or RcsF-Strep D72pBPA and incubated at room temperature for 2 h. Samples were subjected to UV cross-linking and analyzed as described above.

**ACKNOWLEDGMENTS.** We thank Natacha Ruiz, Daniel Kahne, and members of the T.J.S. laboratory for critically reading the manuscript. We thank Marcin Grabowicz for a gift of pBT263-55-lpp-gfp plasmid and Dante Ricci for the DPR576 strain. We thank David Westwood and Daniel Kahne for providing purified FLAG-OmpA and for consulting on in vitro work. We thank Jean-François Collet and his laboratory for a gift of anti-RcsF antibodies and for helpful discussions. This work was supported by National Institute of General Medical Sciences Grant GM34821 (to T.J.S.).
3. Ricci DP, Silhavy TJ (2012) The Bam machine: A molecular cooperator. Biochim Biophys Acta 1826(10):1566–1577.

4. Tokuda H, Matsuyama S (2004) Sorting of lipoproteins to the outer membrane in Escherichia coli. Biochim Biophys Acta 1683(2):121–216.

5. Hayashi S, Wu HC (1990) Lipoproteins in bacteria. J Bioenerg Biomembr 22(3):451–471.

6. Matsuoka S, Yokota T, Tokuda H (1997) A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of Escherichia coli. EMBO J 16(23):6947–6955.

7. Zückert WR (1992) Secretion of bacterial lipoproteins: Through the cytoplasmic membrane, the periplasm and beyond. Biochim Biophys Acta 1183(1):159–156.

8. Nishikawa M, Franchi C, Gatti R, La Fleur, M, Sorice, A (2008) A magic beanstalk or a membrane escalator. Biochim Biophys Acta 1843(8):1568–1577.

9. Dong C, et al. (2006) Wza the translocon for E. coli capsular polysaccharides defines membrane and surface properties of Escherichia coli. Proc Natl Acad Sci USA 103(29):10675–10679.

10. Zuckert WR (1992) Probing the Borrelia burgdorferi surface lipoprotein secretion pathway using a conditionally folding protein domain. J Bacteriol 175(12):6724–6732.

11. Chen S, Zuckert WR (2011) Surface localization determinants of Borrelia OspC/Vp family lipoproteins. J Bacteriol 193(11):2814–2825.

12. Jeantier D, et al. (1994) Structural and functional alterations of a colicin-resistant mutant of Escherichia coli. Proc Natl Acad Sci USA 91(2):10675–10679.

13. Konovalova et al. | [Image 0x1 to 19x816]

14. Tokuda H, Matsuyama S (2004) Sorting of lipoproteins to the outer membrane in Escherichia coli. Biochim Biophys Acta 1683(2):121–216.

15. Hayashi S, Wu HC (1990) Lipoproteins in bacteria. J Bioenerg Biomembr 22(3):451–471.

16. Matsuoka S, Yokota T, Tokuda H (1997) A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of Escherichia coli. EMBO J 16(23):6947–6955.

17. Zückert WR (1992) Secretion of bacterial lipoproteins: Through the cytoplasmic membrane, the periplasm and beyond. Biochim Biophys Acta 1183(1):159–156.

18. Nishikawa M, Franchi C, Gatti R, La Fleur, M, Sorice, A (2008) A magic beanstalk or a membrane escalator. Biochim Biophys Acta 1843(8):1568–1577.

19. Dong C, et al. (2006) Wza the translocon for E. coli capsular polysaccharides defines membrane and surface properties of Escherichia coli. Proc Natl Acad Sci USA 103(29):10675–10679.

20. Zuckert WR (1992) Probing the Borrelia burgdorferi surface lipoprotein secretion pathway using a conditionally folding protein domain. J Bacteriol 175(12):6724–6732.

21. Chen S, Zuckert WR (2011) Surface localization determinants of Borrelia OspC/Vp family lipoproteins. J Bacteriol 193(11):2814–2825.