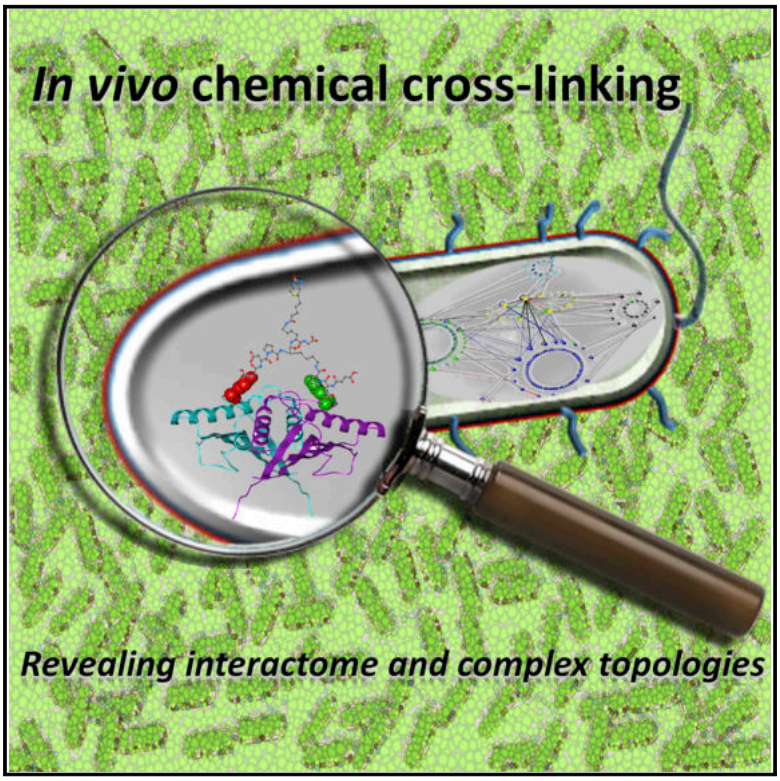
Article

Structure

Probing the Protein Interaction Network of *Pseudomonas aeruginosa* Cells by Chemical Cross-Linking Mass Spectrometry

Graphical Abstract



Authors

Arti T. Navare, Juan D. Chavez, ..., Colin Manoil, James E. Bruce

Correspondence

jimbruce@uw.edu

In Brief

Navare et al. apply in vivo chemical cross-linking mass spectrometry to reveal the first large-scale protein interaction network in *Pseudomonas aeruginosa* with over 600 cross-linked peptide pairs. Cross-linked sites provide constraints useful for complex structure prediction, even for membrane protein interactions.

Highlights

- 260 binary interactions of *P. aeruginosa* proteins were detected in live cells
- In vivo cross-linking enabled identification of membrane protein interactions
- Cross-link sites enabled complex structure predictions

Navare et al., 2015, Structure *23*, 762–773
April 7, 2015
http://dx.doi.org/10.1016/j.str.2015.01.022

CellPress

CellPress

Structure
Article

Probing the Protein Interaction Network of *Pseudomonas aeruginosa* Cells by Chemical Cross-Linking Mass Spectrometry

Arti T. Navare,[1] Juan D. Chavez,[1] Chunxiang Zheng,[1] Chad R. Weisbrod,[1] Jimmy K. Eng,[1] Richard Siehnel,[2] Pradeep K. Singh,[2] Colin Manoil,[1] and James E. Bruce[1,*]
[1]Department of Genome Sciences, University of Washington, 850 Republican Street, Seattle, WA 98109, USA
[2]Department of Microbiology, University of Washington Medical Center, 1959 N.E. Pacific Street, Seattle, WA 98195, USA
*Correspondence: jimbruce@uw.edu
http://dx.doi.org/10.1016/j.str.2015.01.022

SUMMARY

In pathogenic Gram-negative bacteria, interactions among membrane proteins are key mediators of host cell attachment, invasion, pathogenesis, and antibiotic resistance. Membrane protein interactions are highly dependent upon local properties and environment, warranting direct measurements on native protein complex structures as they exist in cells. Here we apply in vivo chemical cross-linking mass spectrometry, to reveal the first large-scale protein interaction network in *Pseudomonas aeruginosa*, an opportunistic human pathogen, by covalently linking interacting protein partners, thereby fixing protein complexes in vivo. A total of 626 cross-linked peptide pairs, including previously unknown interactions of many membrane proteins, are reported. These pairs not only define the existence of these interactions in cells but also provide linkage constraints for complex structure predictions. Structures of three membrane proteins, namely, SecD-SecF, OprF, and Oprl are predicted using in vivo cross-linked sites. These findings improve understanding of membrane protein interactions and structures in cells.

INTRODUCTION

Pseudomonas aeruginosa (*PA*) is a Gram-negative, opportunistic pathogen that causes serious and re-occurring infections in immunocompromised individuals, including patients with cystic fibrosis (CF) (Rosenstein and Zeitlin, 1998). The intrinsic virulence of *PA*, the emergence of multi-drug-resistant (MDR) strains, and the treatment-resistant phenotype of chronic *PA* infections (Oliver et al., 2000) underscore the need for the development of novel therapeutic options. Since nearly all biological functions in living cells are regulated by protein-protein interactions (PPIs), improved understanding of PPIs could yield new avenues for effective therapies. However, despite the significance that large-scale interaction data could hold in drug development, the current knowledge on *PA* PPIs is limited to only a few manually curated interactions(Goll et al., 2008) discovered using small-scale targeted experiments(Goure et al., 2004; Marquordt et al., 2003; Robert et al., 2005).

Recently, genome-scale prediction of PPIs in *PA* integrated various genomic features of *PA* genes and their protein products, such as essentiality, co-expression, co-localization, domain-domain interaction, co-operon or co-gene cluster information, etc., against a compiled reference dataset created from known PPIs in three closely related Gram-negative bacteria (Zhang et al., 2012). Such databases provide a key first step in defining relevant PPIs in bacterial pathogenicity and drug resistance, and can be greatly accelerated by experimental measurements. Furthermore, empirical measurements can reveal what interactions are actually present and can advance predictions, since two proteins that do not share a common operon and are thought to be present in different subcellular locations could, in reality, co-localize and interact in vivo, and thus, would not be predicted with high confidence to be interactors.

Mapping PPIs experimentally can be achieved by employing well-established methods such as the high-throughput yeast two-hybrid method of genome-level screening (Fields and Song, 1989), tandem affinity purification (Puig et al., 2001), and co-immunoprecipitation (Bauer and Kuster, 2003). However, many of these techniques require protein isolation or assay under non-native conditions that can increase both false-positive and false-negative results. Chemical cross-linking with mass spectrometry (XL-MS) has emerged as a powerful tool for mapping PPIs from protein complexes (Herzog et al., 2012; Singh et al., 2010; Sinz, 2003; Tang et al., 2005) and provides highly complementary information to conventional methods with several unique advantages. These include the ability to identify PPI networks that exist in vivo, formed by proteins that are in close proximity due to a direct interaction or due to their interactions with a third protein (Guerrero et al., 2006; Vasilescu et al., 2004). XL-MS data can also be utilized to provide structural information of functional protein complexes. For example, chemical cross-linking of intact infectious virus particles of the *Potato Leafroll Virus* revealed interactions sites among viral coat proteins that enabled prediction of interfacial surfaces based on cross-link site-directed docking (Chavez et al., 2012).

A third notable advantage of in vivo XL-MS is its ability to identify PPIs and structural features directly from intact cells (Zhang et al., 2008), including those of membrane proteins. Membrane proteins represent nearly 30% of the entire eukaryotic proteome (Wallin and von Heijne, 1998), include desirable drug targets

(Yildirim et al., 2007), and yet they are under-represented in the protein data bank with only a few hundred known crystal structures. Crystallization of intact membrane proteins is extremely challenging as these species are often structurally unstable once removed from their native environment. In vivo chemical cross-linking circumvents these technically challenging steps by linking membrane proteins in their native forms, capturing their complexes and interactions as they exist in cells (Chavez et al., 2013; Zhang et al., 2008). Due to these advantages, XL-MS has been applied to study protein complexes from a variety of samples including intact bacteria (Weisbrod et al., 2013a), virus particles (Chavez et al., 2012), mammalian cells (Vasilescu et al., 2004), and cell lysate (Rinner et al., 2008).

Given the value of protein interaction networks in identifying novel drug targets for human opportunistic bacteria (Cui et al., 2009), we focus on large-scale mapping of PPI networks in living *PA* cells using XL-MS. These efforts revealed several previously unknown interactions of structural outer membrane lipoproteins with periplasmic, cytoplasmic, and hypothetical proteins. The occurrence of such PPIs between proteins of different annotated subcellular locations highlights the dynamic nature of protein localizations. Protein functions are related to their network topology; hence, interactions detected between hypothetical proteins and lipoproteins provide insight on possible functional involvement of these hypothetical proteins. Furthermore, the experimentally observed peptide-peptide cross-links found within a protein and between two proteins allow prediction and visualization of structural features of proteins and complexes for which no structural or functional information is available. Therefore, results derived from proteome-wide cross-linking experiments not only extend the current understanding of the protein interaction network in *PA* but also provide the first in vivo topological details on these interactions using an unbiased approach. This comprehensive protein interaction network can serve as a valuable resource to the community; the cross-link site information from the present dataset can be utilized to derive structural information for proteins of interest.

RESULTS

PA cells were cross-linked by a Protein Interaction Reporter (PIR) cross-linker called BDP-NHP (*n*-hydroxyphthalamide ester of biotin aspartate proline) and the cross-linked digest was analyzed using real-time analysis for cross-linking technology (ReACT), an in-house mass spectrometry-based method (Weisbrod et al., 2013a) (Figure S1A; Supplemental Experimental Procedures). A sample-specific stage 1 database comprising 9,122 unique peptides identified with $\leq$1% false discovery rate (FDR) corresponding to 1,282 unique proteins ($\geq$2 peptides/protein) (Table S1) was created and ReACT MS data were searched against the stage 1 database to identify cross-linked peptides. At this step, a higher FDR threshold of 5% was employed to filter the identified peptide matches since an additional requirement of peptide mass modification (BDP stump mass) was also employed to ensure confidence in the cross-linked peptide matches. It should be noted that in a given cross-linked peptide pair, either one or both the peptides may contain decoy sequences (target-decoy/decoy-decoy) instead of being assigned to target sequences (target-target). A global FDR of 3% was calculated to determine the percentage of cross-linked pairs containing one or both decoy sequences (target-decoy and decoy-decoy) rather than target sequences (target-target), as an estimate of the FDR of cross-linked peptide pair identification. The advantage of using a sample-specific stage 1 database containing target and decoy protein sequences for confident identification of cross-linked peptide pairs with reduced FDR has been demonstrated previously (Anderson et al., 2007; Chavez et al., 2013; Zhang et al., 2009). Representative proteins from all subcellular compartments were found cross-linked and their respective percentages in the stage 1 database (Figure S1B) were directly proportional to their constitutive numbers in the entire PAO1 database (http://www.pseudomonas.com/p_aerug.jsp), indicating that under the given cross-linking reaction conditions, the BDP-NHP cross-linker is not biased toward proteins of a particular cellular compartment. Identification of the cross-linked peptide pairs of cytoplasmic proteins indicate PIR molecules penetrate into *PA* cells, as has been demonstrated previously with mass spectrometry and immuno-nanogold electron microscopy imaging of other PIR cross-linked bacterial cells (Tang et al., 2007) and with confocal microscopy of PIR labeled human cells (Chavez et al., 2013).

After implementing the above data processing steps, we identified a total of 690 unique cross-linked peptides forming a total of 626 peptide pairs (Table S2) from a total of 211 proteins. All cross-linked peptide pairs identified in this work are presented in our online cross-linked peptide database XLink-DB (http://brucelab.gs.washington.edu/xlinkdb/specViewerPA.php). Of these, 437 (70%) were between peptides of a single protein with non-identical and non-overlapping sequences, and are referred to here as ''intra-cross-link pairs'', 53 (8%) corresponded to links between two peptides of a single protein with overlapping and identical sequences, and are termed ''unambiguous homodimer links.'' The remaining 136 (22%) cross-linked peptide pairs were between peptides of two different proteins and are referred as ''inter-cross-link pairs.'' It is worth mentioning that ReACT allows identification of hetero- and homodimer peptide pairs with equal probability. Several methods have been published to enable identification of non-cleavable cross-linked peptides (Rinner et al., 2008; Trnka et al., 2014; Yang et al., 2012). While it is conceivable that several of these strategies could show enhanced identification of homodimer peptide pairs since both peptides share common fragments in the complex MS^2 spectra, this is not the case with cleavable cross-linkers(Tang et al., 2005) and ReACT (Weisbrod et al., 2013a), and other similar methods(Kaake et al., 2014; Liu et al., 2012). ReACT calculates masses of the released cross-linked peptides as they are cleaved from the cross-linker and then fragments both the peptides separately for peptide identification. This is true for both homo- and heterodimeric peptide pairs, allowing identification of both types of cross-linked peptide pairs without a bias toward homodimeric links.

PIR XL-MS identify binary protein interactions because two linked peptides are used to establish connectivity. Thus, it is currently not straightforward to ascertain if cross-linking sites among several proteins indicate the presence of multi-protein complexes or simply several binary interactions. For intra-protein cross-links between peptides with non-identical or non-overlapping sequences, additional efforts are often required to determine if the two linked peptides belong to one protein

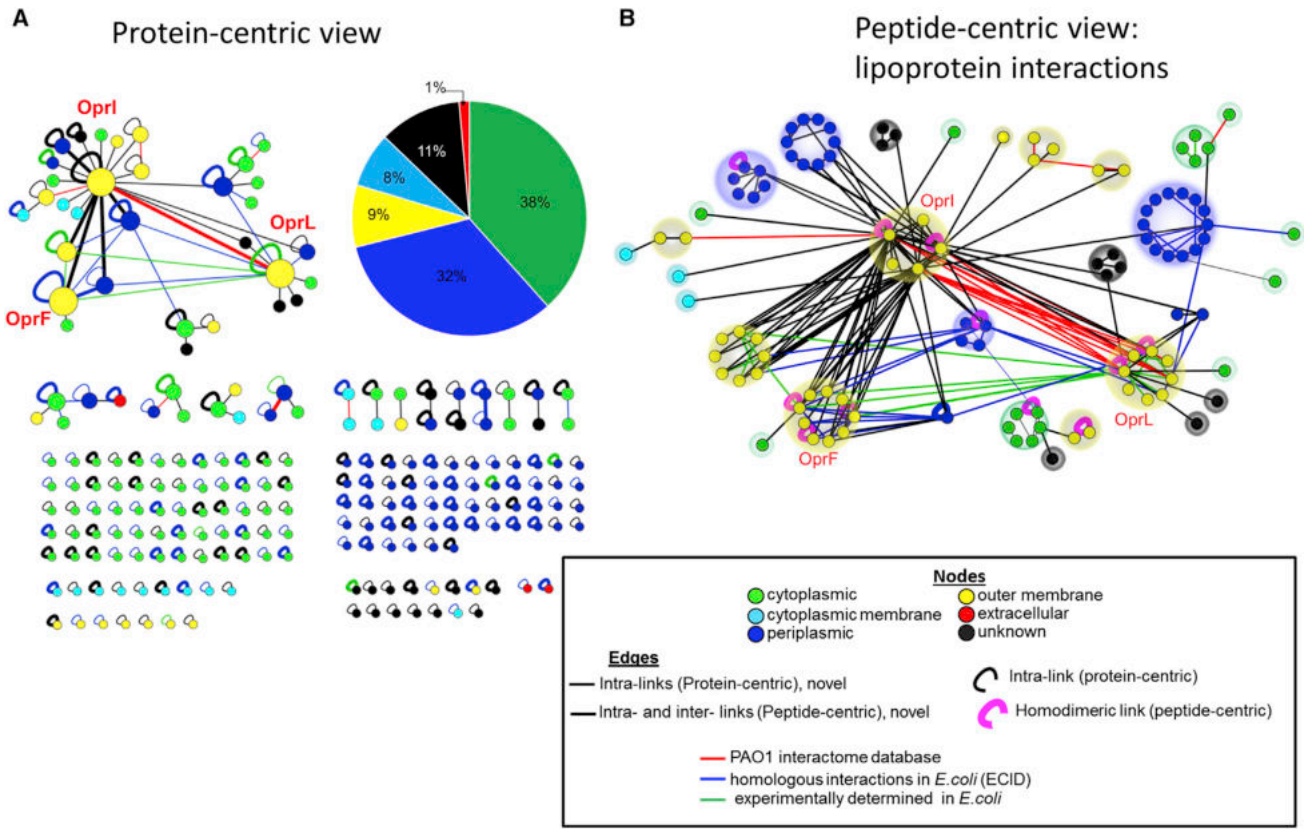


Figure 1. In Vivo Cross-Linking-Derived Protein-Protein Interaction Network of *P. aeruginosa*

PIR cross-linked derived large-scale PPI profile obtained from intact *PA* cells (A) and a peptide-centric view of the major lipoprotein interactions (B). Nodes in (A) represent proteins and connections denote cross-link-interactions. In the peptide-centric view of the lipoprotein interactions (B), multiple ($\geq$2) cross-linked peptides of a single protein are arranged in a circle layout. Edge colors in (A) and (B) represent computationally predicted *PA* protein interactions from PAO1 interactome database (red), homologous PPI in *E. coli* as reported in the EcID database (blue), homologous PPI from cross-linked *E. coli* cells (green), and novel interactions (black). Intra-links in (A) are shown as semicircular edges and the homodimeric links in (B) are indicated by colored semicircular (magenta) edges. Node colors from (A) and (B) denote protein subcellular locations. The PPI network layout was directed by connectivity of the nodes in the Cytoscape program and the three highly connected membrane proteins, namely OprI, OprL, and OprF, are highlighted by bigger node size. A pie chart from (A) shows the distribution of the subcellular locations of all the 211 identified cross-linked proteins.

monomer or two monomers of the same protein. In contrast, when both the cross-linked peptides share the same or overlapping amino acid sequences (homodimer link) and if the peptide sequence occurs just once within a protein, this unambiguously indicates the existence of a multimeric complex between two or more identical monomers of a protein (Anderson et al., 2007). We used the intra-cross-links, the unambiguous homodimers, and inter-protein cross-linked sites to build monomeric protein models and to filter homo- and heterodimeric protein models.

All 260 protein-level PPIs (Table S3) are depicted in Figure 1A in a protein-centric view, while a peptide-centric network of the highly cross-linked major lipoproteins is shown in Figure 1B. Each cross-linked protein (Figure 1A) or peptide (Figure 1B) is denoted as a node. In a peptide-centric view (Figure 1B), when more than one peptide of a protein is cross-linked, the corresponding peptide nodes are arranged in a circular cluster. Different edge colors indicate that the identified PPI is included in a computationally predicted PAO1 interactome database (edge color, red) and/or is a homologous PPI in *Escherichia coli*, previously reported in the EcID database (edge color, blue) and/or is identified from PIR cross-linked *E. coli* cells in our lab (edge color, green) (Weisbrod et al., 2013a; Zheng et al., 2011, and additional efforts http://brucelab.gs.washington.edu/xlinkdb/dataRetriever.php?tablename=ecoli_total&dataset=&privateFlag=1), or is a previously unknown PPI (edge color, black). An experimentally validated protein interaction database for *PA* does not currently exist; however, the cross-linking-derived binary protein interaction network was compared with a predicted PAO1 interactome database (Zhang et al., 2012) (red edges). Only seven PPIs between the cross-linked proteins that share an operon (SecD-SecF; SucC-SucD; LptE-LptF) or a subcellular location (i.e., membrane proteins, OprI-OprL, OprI-OprN; or cytoplasmic proteins, SucA-rpsG and rplB-fusA) were found in the predicted PAO1 interactome database. Empirical in vivo measurements of predicted interactions are important and can further advance predictions since two proteins that are not co-expressed or are thought to be present in different cellular compartments could, in reality, co-localize and interact in vivo, and thus would not be predicted as high confidence interactors. A majority of PPIs identified

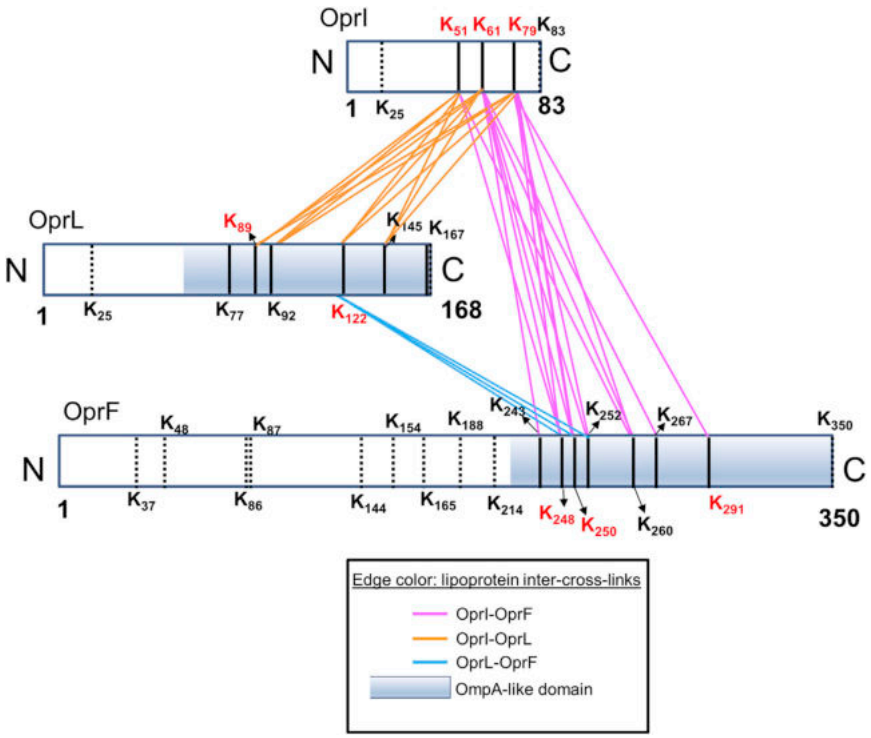


Figure 2. The Identified Inter-Cross-Linked Sites of the Major Lipoproteins OprI, OprF, and OprL

A block diagram shows relative protein lengths of OprI, OprF, and OprL with black lines denoting the positions of all Lys residues within the mature form of lipoprotein sequences. Dotted and solid black lines mark the positions of un-cross-linked and cross-linked Lys. Homodimeric sites are marked in red. Pink, orange, and light blue lines denote inter-cross-linking between OprI-OprF, OprI-OprL, and OprF-OprL, respectively. The OmpA-like region in OprL and OprF is shown as a shaded blue box. See Figure S2.

between PAO1 proteins were either reported for homologous *E. coli* proteins in the EcID database (Table S4) or were identified from the PIR cross-linked *E. coli* cells, suggesting that these interactions are conserved among the two Gram-negative bacteria.

Outer membrane lipoprotein, OprI, peptidoglycan (PG)-associated lipoprotein, OprL, and major porin, OprF, were among the highly connected proteins (larger nodes, Figure 1A), forming interactions with 22, 9, and 6 other PAO1 proteins, respectively. It is important to note that cross-linked outer membrane proteins contributed only a small proportion (9%) of the total detected cross-linked proteome, while the majority of cross-linked proteins were from soluble cytoplasmic (38%) and periplasmic (32%) regions (Figure 1A, pie chart). This suggests that while the PIR cross-linker molecules permeated through the cell membrane barrier into the inner cell compartments, the higher number of intra- and inter-cross-links detected for major lipoproteins is likely due to their high abundance in the cell envelope (Mizuno and Kageyama, 1979b). This compartment is first to come in contact with the incoming cross-linker so that during the short half-life of the PIR cross-linker in the timescale of tens of minutes, outer membrane proteins are likely to be readily cross-linked, provided that reactive lysine (Lys) sites within the protein sequence are accessible. The limitation of MS-based approaches for detecting low abundant species from complex biological systems is shared by in vivo cross-linking technology. As technologies improve with greater MS speeds, improved bioinformatics capabilities, and sample preparation workflow, a wider dynamic range of detection is achieved. It is also important to note that the cross-linking data enabled structural predictions of these abundant proteins and their complexes, for which no crystal structures are currently available. Most conventional methods of protein structure determination require highly purified isolated proteins and often fail even with highly abundant membrane proteins. In vivo cross-linking obviates the need to perturb native proteins to gain structural information and is complementary to other conventional structural methods.

A peptide-centric view of the lipoprotein network shows the number of unique cross-linked peptides of the lipoproteins and their interacting partners (Figure 1B). All three lipoproteins contain multiple residues that are accessible for forming intra-links (OprI, 21 intra-links; OprF, 9 intra-links; and OprL, 12 intra-links) and homodimeric intra-links (Figure 1B; Table S2). The existence of unambiguous homodimeric links between peptides with identical sequence confirms that the three lipoproteins form homodimeric complexes in vivo. The same sites of intra-cross-links were also involved in inter-protein cross-links with other *PA* proteins. These identified interactions involve proteins functionally linked to membrane processes such as secretion, transport, and several metabolic processes and structural details of these interactions are lacking. Interestingly, all the cross-linked sites of these PPIs were located near the carboxy termini of OprI, OprF, and OprL (Figure 2), suggesting that the C termini of these key membrane proteins are exposed, and are more likely to participate in PPIs compared with their membrane-bound (Khalid et al., 2006) or lipidated N termini (Mizuno and Kageyama, 1979b). A domain-domain interaction profile of all the cross-linked PAO1 proteins obtained from a Structural Classification of Proteins database (SCOP database, 1.75v) (Andreeva et al., 2008) revealed that the highly cross-linked carboxy termini of two of the major lipoproteins, OprF and OprL, share a common ''OmpA-like domain'' (Figure 2). In the case of OprI, all the associated PPIs were mapped to an unannotated domain due to the lack of a known structural or functional domain for this protein. This C-terminal unannotated region of OprI interacted with the highly conserved OmpA-like domains of five other membrane proteins, namely, lipotoxin (lptF), a probable membrane protein (Q9I4T3), and the two lipoproteins OprF and OprL.

The mature form of OprL has eight Lys residues, seven of which are located in the OmpA-like domain that spans from 56 to 168 amino acid residues. Of the seven Lys residues from the

OmpA-like domain, six were found inter- or intra-cross-linked, while the seventh Lys at the protein C terminus was not identified in any cross-linked relationship in the present study. When the protein sequences of OprL and its homolog PAL from *E. coli* were aligned, the OmpA-like domains of the two lipoproteins showed 40% sequence alignment (Figure S2A). In PAL, the OmpA-like region contains PG binding sites, PAL dimerization sites, and sites of known interactors of PAL, namely TolA, TolB, and OmpA (OprF homolog) as indicated in Figure S2 (Cascales and Lloubes, 2004). Interestingly, two OprL Lys residues, K89 and K122, involved in forming inter-links with the OmpA homolog, OprF, aligned with N93 and N127 of *E. coli* PAL, differ by single base substitutions at this codon. Furthermore, although these sites (N93 and N127) appear somewhat distant in sequence from the OmpA binding region in PAL, both flank the β-strand and loop region on the PAL C-terminal crystal structure that were shown by site-directed mutagenesis (Cascales and Lloubes, 2004) to be important for OmpA interaction (Figure S2B). The same Lys residues (K89 and K122 in OprL) were identified as the sites of unambiguous homodimeric links, confirming the existence of in vivo dimerization of OprL, as also shown previously for *E. coli* PAL (Cascales and Lloubes, 2004). The non-cross-linked Lys residue from the OprL N terminus aligned within the unstructured membrane-bound N-terminal of PAL (dotted line, Figure S2A) (Abergel et al., 2001) that was suggested to be a functionally silent (Cascales and Lloubes, 2004) flexible tail, only serving to anchor the protein into the outer membrane (Abergel et al., 2001). Therefore, the highly cross-linked C-terminal domain of OprL containing the OmpA-like region must also be functionally important and accessible to other *PA* proteins within the periplasmic region.

Similar to OprL, all the cross-linked Lys residues of OprF are located near the C-terminal region in the OmpA-like domain (Figure 2). The PG-associated major porin protein, OprF, has been studied extensively, first due to its proposed multifunctional roles, as a non-specific porin (Woodruff and Hancock, 1988), and its involvement in cell wall structural stability (Rawling et al., 1998), yet controversy still exists surrounding the structure. Early studies involving surface epitope mapping (Rawling et al., 1995) predicted a one-domain secondary structure for OprF with 15 transmembrane strands traversing the entire protein length (Figure S2Ci). A later report (Sugawara et al., 2006) suggested that OprF may exist in two forms: a highly abundant two-domain form containing membrane-bound N-terminal β-barrel with a periplasmic globular C-terminal domain and a low abundance one-domain form similar to that mentioned above, which contains only a β-barrel structure spanning the entire length of the protein (Figure S2Cii). Since all cross-link sites of OprF are located near the C-terminal region, the cross-linking results identified here are most consistent with the abundant two-domain conformer of OprF where the solvent-accessible C-terminal domain resides in the periplasmic space. Second, the cross-linked Lys residues of OprF were also identified in inter-protein links with Lys residues of PAL and OprI within domains of these proteins that are also located in the periplasmic region. This further suggests that these C-terminal cross-linked regions of all three outer membrane proteins exist close to one another (within 35 Å) within the solvent-accessible periplasmic space. Third, OprF formed homodimeric intra-links (Figure 1B) in the C-terminal region similar to its *E. coli* homolog OmpA where previous in vivo cross-linking data enabled OmpA C-terminal dimer structure prediction (Zheng et al., 2011). Subsequent site-directed mutagenesis experiments (Marcoux et al., 2014) confirmed that the dimerization interface of OmpA is present in the soluble periplasmic C terminus.

DISCUSSION

Membrane Protein Interactions of OprI

Unique insight that can be gained from the present in vivo cross-linking data relates to the intricate network of membrane proteins, such as the outer membrane lipoproteins OprI, OprF, and OprL (Figure 1). *PA* outer membrane protein OprI was first isolated and characterized as a homolog of the free-form of *E. coli* lipoprotein LPP by Mizuno and co-workers (Mizuno and Kageyama, 1979b). The same group also confirmed the association of OprF and OprL with the PG layer (Mizuno and Kageyama, 1979a). OprI was suggested to be present largely as an unbound form in *PA* strain PAO1 and the existence of the PG-bound form can vary among strains of the bacteria (Mizuno and Kageyama, 1979a). In the present study, OprI was identified as the highest connected protein, linked to 17 other proteins, including OprF and OprL (Table S2), in addition to the homodimeric intra-links (Table 1; Figures 1B and 2). Interestingly, all these interactions of OprI were near the C terminus at Lys residues 51, 61, and 79 (Figure 2), suggesting that this region of OprI must reside in the solvent-accessible periplasmic space and is likely to contain functionally important residues.

OprI-OprF-OprL Interaction: Potential Role in Cell Envelope Structure Stability

Our in vivo cross-linking efforts resulted in identification of several inter-protein interactions between the C-terminal of OprI and C-terminal OmpA-like domains of OprF and OprL, involving 13 and 11 inter-cross-link peptide pairs, respectively (Figure 2). In addition, OprL and OprF C-terminal inter-protein linkages were identified, suggesting the existence of an intracellular ternary lipoprotein complex in cells, consistent with that in *E. coli* (Cascales et al., 2002). The sequence homologs of OprI, OprF, and OprL in *E. coli*, namely, LPP, OmpA, and Pal were shown to form multi-protein complexes during in vivo cross-linking and immunoprecipitation experiments and it was suggested that interaction between PAL and OmpA provides cell membrane integrity in *E. coli* (Cascales et al., 2002). The presence of conserved PPIs between the homologous lipoproteins within *E. coli* and *PA* support similar functional roles of their interactions across the two closely related Gram-negative bacteria. In additional, OprF was proposed to be important for structural integrity (Rawling et al., 1998) where the possible association of the two-domain OprF conformer discussed above with OprL and OprI was suggested to be relevant for cell shape and outer membrane stability in *PA*. The observed inter-links between OprF and the two lipoproteins confirm the existence of these associations in vivo.

OprI Interactions with Bacterial Lipotoxins

Inter-protein links between OprI and two of the *PA* lipotoxins, namely LptF and LptE, were identified and a single inter-link

Table 1. A List of the Homodimeric Intra-Cross-Link Peptide Pairs of Protein OprI

Peptide 1	Peptide 2	Protein Uniport ID	Gene Name	Cross-Linked Lysine Site 1	Cross-Linked Lysine Site 2
MLEKASRK	MLEKASR	P11221	oprI	79	79
MLEKASR	MLEKASRK			79	79
KADEALGAAQKAQQTADEANER	KADEALGAAQK			51	51
KADEALGAAQKAQQTADEANER	ADEALGAAQKAQQTADEANER			61	61
MLEKASR	MLEKASR			79	79
MLEKASR	MLEKASR			79	79
MLEKASR	MLEKASR			79	79
KADEALGAAQKAQQTADEANER	KADEALGAAQKAQQTADEANER			61	61

Unique peptide pairs showing the site of PIR crosslinking (underlined) is listed, along with the residue number of the cross-linked Lys from each peptide. Numbering of cross-linked lysine residues within a protein sequence follows the Uniprot sequence annotation, starting with methionine as the first amino acid residue. Post-transnationally modified (oxidation) methionine residues in the cross-linked peptides are double underlined.

was also detected between the two lipotoxins (Table S2). LptF and LptE are the two outer membrane lipoproteins, with multifunctional roles as pro-inflammatory factors (Firoved et al., 2004) and in the survival of *PA* under harsher environments (Damron et al., 2009), that share an operon (Database) and were detected in the outer membrane vesicle (OMV) proteome of the bacteria (Choi et al., 2011). Based on their common operon, LptF and LptE were predicted to interact (Zhang et al., 2012). Our unbiased in vivo cross-linking data provide the first direct evidence of their interaction in living bacterial cells, along with interaction with major lipoprotein OprI. The positions of the inter-protein cross-linked sites between the two lipotoxins and OprI suggest that the C-terminal region of LptE and the OmpA-like domain of LptF are in proximity of the OprI periplasmic domain. The association between LptF and LptE proteins with OprI and their presence in OMVs raise the possibility that these interactions may be important for extracellular trafficking of these multifunctional lipoproteins via OMVs.

Interaction of Transmembrane Proteins SecD-SecF: Prediction of a Functional Protein Complex

The Sec protein machinery is a well-conserved secretion system in Gram-negative bacteria that enables export of newly synthesized membrane proteins and other secreted proteins across the inner membrane to their final destination. Two components of the Sec apparatus SecD and SecF, each contain six transmembrane helices and are known to form a complex and interact with the SecYEG translocon to facilitate protein transport from cytoplasmic to periplasmic space via proton motif force (Arkowitz and Wickner, 1994). The large periplasmic loop of SecD appears important to maintain proper function of the SecDF protein complex (Nouwen et al., 2005). Although reports of successful co-crystallization of SecD and SecF have not yet appeared, the crystal structure of a fused form of SecD and SecF (TsecDF) from *Thermus thermophilius* has been published (Tsukazaki et al., 2011) (Figure 3A). TsecDF can assume two forms: a substrate capturing F-form and substrate releasing I-form (Tsukazaki et al., 2011). The periplasmic head region (P1) of the SecD domain of TsecDF moves toward (I-form) or away from (F-form) the periplasmic side (P4) of the SecF domain, aided by a small hinge region in the base of P1, and as a result, a substrate protein is captured from the translocon and subsequently released into the periplasm (Tsukazaki et al., 2011). Critical residues identified in TsecDF include Leu106 and Leu243, which are required for proton motion and flexibility of the hinge region (Tsukazaki et al., 2011).

The PIR XL-MS dataset contain one cross-linked peptide pair of PAO1 SecD (ILGKTANLEFR) and SecF (MPSEDPELGKK), along with one intra-link peptide pair within protein SecD (Figure S1C). When the sequences of SecD and SecF were aligned with TsecDF, the position of these peptides was aligned with cross-linked peptide from PAO1 SecD (ILGKTANLEFR) situated on the functionally important flexible linker or hinge portion of the TsecD domain while the SecF peptide (MPSEDPELGKK) aligned with the loop region of the TsecDF domain (Figure 3A). Since loop regions are often involved in protein-protein recognition (Planas-Iglesias et al., 2013), it is possible that these loop regions of SecD and SecF are crucial for functional complex formation. Sequence alignment efforts also revealed that critical residue Leu106 in the TsecDF hinge region is aligned to Ile of the cross-linked peptide ILGKTANLEFR from SecD. Furthermore, the second Leu243 critical for proton transport is conserved in SecD along with other transmembrane amino acid residues of TsecDF (Figure S3A).

Due to the sequence similarities shared by SecD and SecF proteins with the fused TsecDF form, the protein sequences of PAO1 Sec proteins were concatenated to form a fused version and structural model prediction was performed with I-TASSER (Roy et al., 2010). The five predicted structures (Figure S3B) closely resembled the TsecDF structure and on the top model (TM score of 0.53 ± 0.15) as shown in Figure 3B (http://www.modelarchive.org/ DOI: ma-a4bdy), the observed cross-linked sites were positioned approximately 13 Å apart (C_α-C_α atoms of cross-linked Lys), in excellent agreement with previous PIR cross-linked site distributions (Chavez et al., 2013). These results illustrate that cross-linking data can identify locations of protein interaction interfaces and can shed light on relevant structural features of complexes that have not yet been successfully studied with crystallography.

Homodimer Interactions of Membrane Lipoproteins

Protein dimerization/ oligomerization is ubiquitous in biology (Marianayagam et al., 2004). Dimerization of membrane proteins can facilitate ion channel formation (Zheng et al., 2011), signal

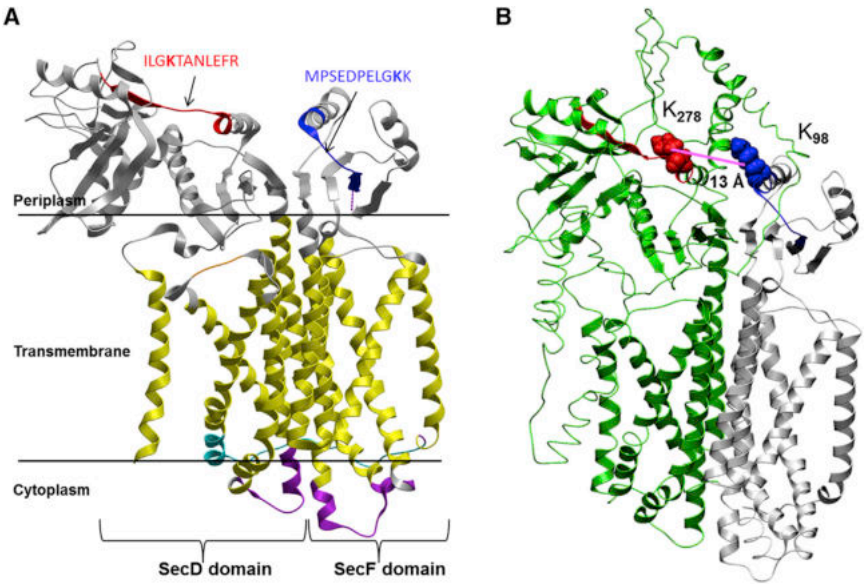


Figure 3. Known and Cross-Linking-Derived Predicted Structures of Fused Form of SecD and SecF Proteins from Bacteria *Thermus thermophilius* and *P. aeruginosa*, Respectively

(A) Crystal structure of the fused SecD-SecF protein (TsecDF) from *Thermus thermophilius* showing the SecD and SecF domains containing six transmembrane domains (yellow) each, and periplasmic (gray) and cytoplasmic (purple) domains. Location of the cross-linked peptides from PAO1 SecD (red) and SecF (blue) was mapped onto the crystal structure.

(B) The top scoring predicted structure of concatenated *PA* SecD-secF using I-TASSER. The green and gray portions denote the SecD and SecF portions of the concatenated complex. See Figure S3.

transduction (Yu et al., 2002), and can provide structural stability. Cross-linked products that include two linked identical or overlapping peptide sequences can yield unambiguous identification of homomultimeric interactions. Direct in vivo cross-linking measurements of the lipoprotein complexes revealed several unambiguous homodimer linkages as described below, and can help extend the current knowledge about their structures and possible functions as they exist in the cellular environment.

Homodimeric Links of Major Outer Membrane Lipoprotein, OprI

Unambiguous homodimer links of OprI involving three cross-linked sites were identified (Table 1) and the same sites were also identified in inter-linked relationships. Oligomeric forms of native OprI were detected by western blot in a study by Lin et al. (2010) and its homolog protein LPP from *E. coli* is known to exist as a homotrimer (Shu et al., 2000). LPP and OprI show similar properties such as their molecular weight, unique solubility in 10% trichloroacetic acid/SDS, composition of fatty acids, and both proteins are rich in α-helix (Mizuno and Kageyama, 1979b). Lpp and OprI share 25% sequence similarities, including the N-terminal lipidated cysteine, carboxyl-terminal Lys, which can be free or covalently bound to the PG layer, as shown in the sequence alignment of the proteins (Figure 4A). One Lys of LPP (Lys75), corresponding to Lys79 of OprI (highlighted in red) was also found cross-linked in a homodimeric intra-link (unpublished results). Due to the similarity of the two proteins, the crystal structure of LPP-56 mutant was used to predict the monomeric model of OprI using homology modeling software I-TASSER(Roy et al., 2010) (Figure 4B). Of the four predicted models, only a single model had a high confidence(C) score of 0.33 (C score range, −5 to 2), while each of the remaining three models had a C score of −5, signifying lower quality models (Figure S4A). The top model (TM score 0.76 ± 0.10) of OprI was then used with SymmDock to generate a dimer of OprI (Figures 4C and 4D). A cross-linking distance constraint of 35 Å was used as an additional criterion to filter the resulting dimeric models. As can be seen from Figures 4C and 4D, when no distance constraint information was used to generate the dimer structure, the resulting top 20 models showed an antiparallel orientation of the two monomers (Figure 4C) with a distance of 75 Å between one of the cross-linked Lys residues (K79) of the two monomers. Since the N-terminal cysteine residue in LPP is lipid modified and serves as an anchor in the outer membrane to increase cell integrity, this dimeric complex with antiparallel orientation would appears less favorable due to the fewer lipoprotein insertion sites in the outer membrane. On the contrary, the top 20 dimer models of OprI (Figure S4B) generated with the cross-linking distance constraints showed parallel orientation of the OprI monomers, with their N and C termini oriented in the same direction as shown in the case of the top most representative model in Figure 4D (the Model Archive DOI: ma-a89gj), and all the linear and non-linear distances (Table S5) measured between the cross-linked Lys residues of the monomers were well within the maximum distance expected for the PIR cross-linker (Chavez et al., 2013).

Homodimeric Links of Major Porin, OprF

XL-MS analysis identified two unambiguous homodimeric links of OprF, one at K291 and another at K248 (Table 2). An additional homodimer link was observed at K250 with both cross-linked peptides observed with the sequence SKVK, although this short sequence could have originated in several *PA* proteins. On the other hand, the peptide SKVK was also identified cross-linked to FDFDKFK which would be expected given the proximity of these two lysine residues in OprF, as well as three other peptides unique to OprF. (Figure 5A) Thus it is likely that SKVK indeed belongs to the major porin, although it was not used for cross-link site-directed docking below. The adjacent position of the homodimeric cross-link sites 248 and 250 suggest that the periplasmic region of OprF plays a role in the formation of dimeric or homomultimeric complexes, similar to that of its homolog in *E. coli*, OmpA. The interaction among OmpA C-terminal domains was first demonstrated with in vivo cross-linking and a model structure of the OmpA C-terminal dimer was proposed based on cross-link site-directed docking (Zheng et al., 2011). Recent in vitro studies confirmed OmpA dimer existence and

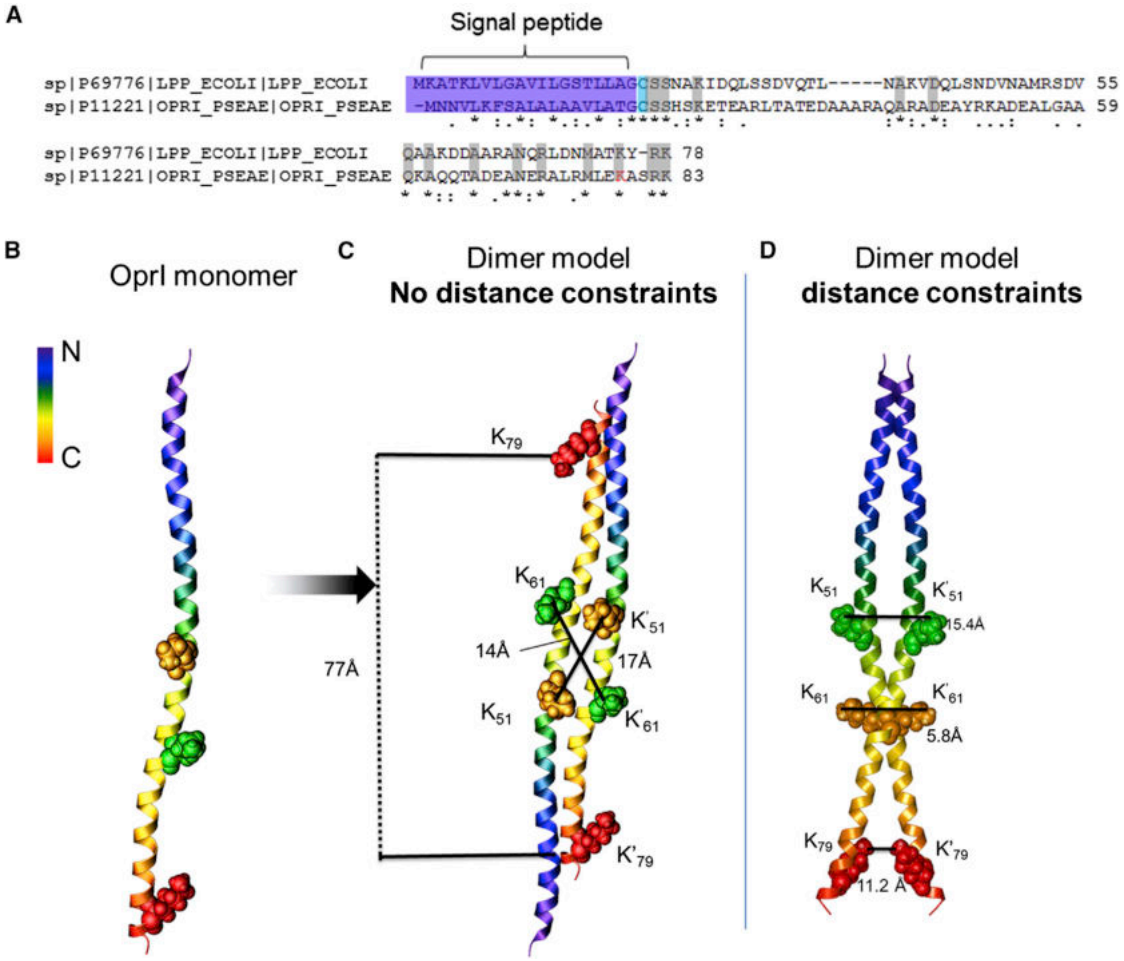


Figure 4. Prediction of OprI Monomer Structure Using the Template Structure of Homolog Protein LPP and Cross-Linking-Aided Selection of OprI Homodimer Complex Model

(A) Sequence alignment of OprI and its *E. coli* homolog LPP showed 25% sequence alignment. Conserved amino acid residues (gray), lipid-bound cysteines (cyan), the signal peptide sequence at the N-terminal (purple) are highlighted. Lys79 from OprI found inter- and intra-cross-linked is highlighted in red.

(B) OprI monomer structure was predicted using I-TASSER, using LPP-56 mutant protein as a template. The cross-linked Lys molecules are shown on the ribbon structure of OprI.

(C and D) Two monomers of OprI were docked onto each other using SymmDock to generate a dimer, with (C) and without (D) using the cross-linking site distance constraint information. The resulting dimer models were evaluated for the distance between the cross-linked Lys. See Figure S4A and S4B.

demonstrated the requirement of the OmpA C-terminal domain for dimer formation (Marcoux et al., 2014). Sequence alignment of OmpA and OprF C termini showed that the homodimeric cross-linked region in OprF aligns with the dimerization interface of OmpA (Figure 5A). Moreover, Marcoux et al. (2014) also showed that the identified inter-linked OmpA site K213 we reported (Zheng et al., 2011) (or K192 numbering with removal of leader sequence) was critical to in vitro dimer formation. Although this Lys is not conserved in OprF, Lys residues K248 and K250, which exist nearby in the aligned OprF sequence, were observed in homodimer links. Homodimeric links in this region of OprF that align with the OmpA dimerization interface suggest that this region in OprF may serve a similar role in oligomerization.

To further gain information on the structural orientation of the dimeric complex of OprF, a model structure of the C-terminal domain of the OprF monomer was predicted using I-TASSER, using a distance constraint of 35 Å between the intra-cross-linked Lys residues. The resulting top scoring model (C score, 0.79; TM score, 0.82 ± 0.08) satisfied the distance constraints for all the observed intra-links (Figure S4C). The predicted structure of the OprF C-terminal domain was superimposed onto the corresponding region of the OmpA crystal structure (PDB ID, 2MQE) (Ishida et al., 2014) to identify regions of variability and the two proteins showed high structural similarity, also reflected in the high TM score of the predicted model. Importantly, the region of the OmpA dimerization interface and the corresponding superimposed region of the OprF monomer containing cross-link sites showed identical secondary structures. This high-quality monomeric model of the OprF C-terminal domain was submitted to SymmDock to generate a symmetrical dimeric model, where homodimeric links at Lys 248, 250, and 291 were used as distance constraints between the dimerization interfaces. The resulting top 20 dimer models (Figure S4D) were

Table 2. A List of the Intra-Links of Protein OprF

Peptide 1	Peptide 2	Protein Uniport ID	Gene Name	Cross-Linked Lysine Site 1	Cross-Linked Lysine Site 2
FDFDKSK	VKENSYADIK	P13794	oprF	248	252
FDFDKSK	ENSYADIKNLADFMK			248	260
SKVK	FDFDKSK			250	248
SKVK	VQLDVKFDFDK			250	243
SKVK	ENSYADIKNLADFMK			250	260
NLADFMKQYPSTSTTVEGHTDSVGTDAYNQK	SKVK			267	250
QYPSTSTTVEGHTDSVGTDAYNQKLSER	SKVK			291	250
FDFDKSK*	FDFDKSK*			248*	248*
QYPSTSTTVEGHTDSVGTDAYNQKLSER*	QYPSTSTTVEGHTDSVGTDAYNQKLSER*			291*	291*
SKVK*	SKVK*			250*	250*

The homodimeric intra-cross-linked peptide pairs are marked with an asterisk (*). Unique peptide pairs and the cross-linked lysine residue number (underlined) from each peptide are listed. Numbering of cross-linked lysine residues within a protein sequence follows the Uniprot sequence annotation, starting with methionine as the first amino acid residue.

evaluated for the shortest non-linear distance between C_β atoms of at least two of the three lysine residues (K248, K250, and K291) from each monomer (Table S6). A dimer model of OprF with the shortest non-linear distances (C_β-C_β) of 9.8 Å, 11.6 Å, between residues 248, 250, respectively, and 39.6 Å between Lys 291 was retained (the Model Archive DOI: ma-axj2r; Figures 5B and 5C). As shown in Figure 5C, both monomeric units of the dimer are oriented parallel to one another with their N-terminal ends, which would connect to the membrane-bound β-barrel region extended toward the outer membrane (Brinkman et al., 2000), while their C-terminal globular regions reside in the periplasmic space. The putative PG binding sites (highlighted in blue, Figure 5C) of OprF are located within the periplasmic region in the top dimer model. This region between residues 284 and 297 is highly conserved within the C-terminal region of various OmpA-related outer membrane proteins including MotB from *E. coli*, and was proposed to be a PG-interacting site (Abergel et al., 2001). All homodimeric cross-linked sites are located in the periplasmic region and are solvent accessible, as inferred from their surface-exposed positions (Figure 5C). Moreover, cross-linked Lys residues K248 and K250 are situated at the dimer interface, and their respective solvent-accessible surface distances 9.8 Å and 11.6 Å are within the expected range of cross-linkable distances for the BDP-NHP cross-linker.

The results presented here reveal the first view of the high confidence PPI network in *P. aeruginosa* with more than 600 cross-linked peptide pairs. A large-scale interaction network such as presented here is a resource for microbiologists to examine the protein-protein interactions in greater detail in this human pathogen. For example, these results reveal previously unknown interactions among many membrane proteins, such as OprI with lipotoxins LptE and LptF. These results offer new insight on lipotoxin interactions, the roles these proteins hold in inflammation in lung epithelia, which may be involved in CF disease progression and bacterial survival in harsh environments as encountered in the colonized CF lung (Damron et al., 2009). The cross-linking sites were used to define interfacial regions of protein complexes and enable cross-link site-directed complex structure predictions. The examples presented here have focused on the selected membrane proteins for which little other structural information currently exists. However, the network of identified cross-linked peptides can be utilized to predict structures of other proteins and protein complexes of interest using the steps described in Figure S5. Knowledge of protein secondary structures, PPIs, and interaction interfaces can also be used to develop synthetic peptides and peptidomimetics targeting specific PPIs essential for antibiotic resistance functions.

All cross-linked peptide pairs identified in this work are presented in our online cross-linked peptide database XLink-DB (http://brucelab.gs.washington.edu/xlinkdb/) as a resource for the research community. The identified cross-linked sites and linkage constraints can be used with protein modeling and docking tools such as I-TASSR, PatchDock, and SymmDock to predict and filter putative monomeric and dimeric protein structures. As we show for OprI, even one or two identified cross-linked peptide pairs can serve to greatly reduce the possible relative orientations of protein subunits in a multimeric complex. Recent work published by Marcoux et al. (2014) illustrates the value of PIR cross-linking data in structural work in which the previously published PIR-cross-linking results in *E. coli* (Zheng et al., 2011) were utilized to enable prediction of outer membrane protein A (OmpA) dimer structures. We believe that the present results may hold additional general utility to help guide future measurements using conventional structural biology techniques.

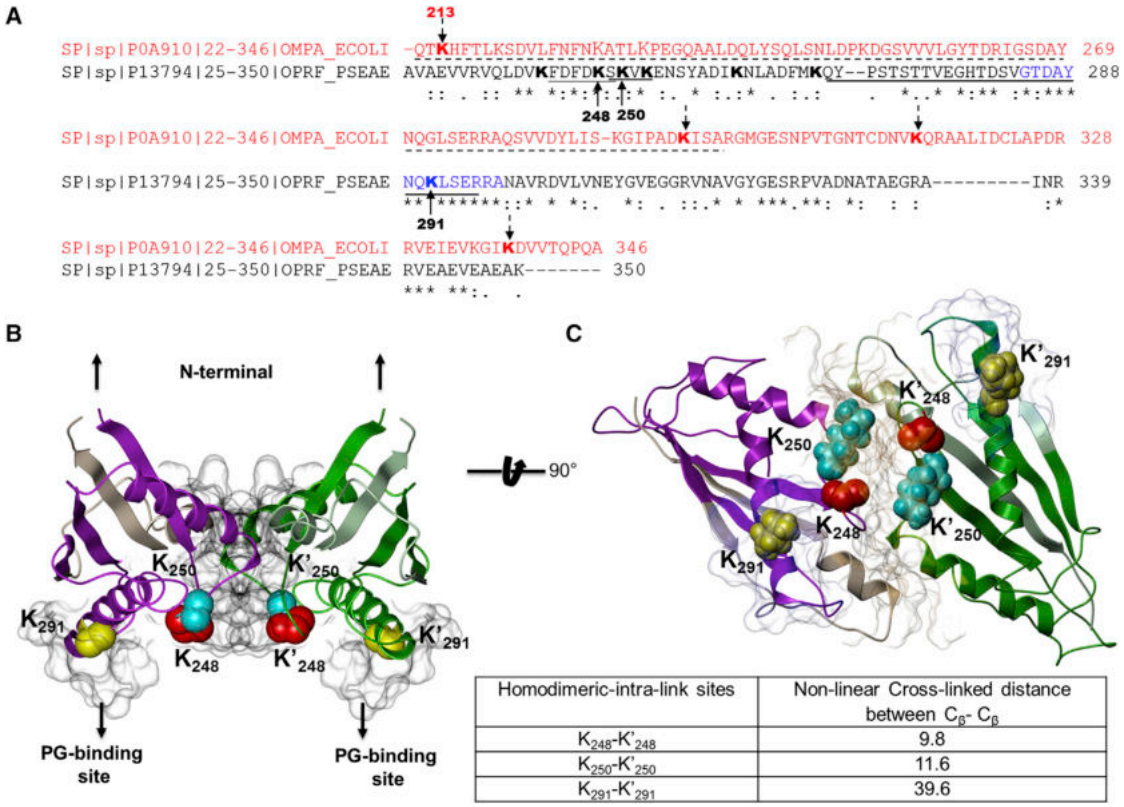
EXPERIMENTAL PROCEDURES

Synthesis of Protein Interaction Reporter Cross-Linker

The PIR cross-linker BDP was synthesized with Fmoc chemistry and esterified to produce the activated *n*-hydroxyphthalamide (NHP) ester of BDP as described previously (Chavez et al., 2013).

Two-Step In Vivo Cross-Linking

P. aeruginosa (strain PAO1) cells were cross-linked using BDP-NHP and processed as illustrated in Figure S1A. Cells were washed first in PBS and then in a 170 mM Na_2HPO_4 (pH 7.4) buffer. BDP-NHP was added to a final concentration of 10 mM and the first step of the cross-linking reaction was carried out for 1 hr followed by centrifugation of the cross-linked cells. The cell pellet was cross-linked as before by adding BDP-NHP (10 mM). After the reaction was over, the cells were washed extensively in PBS buffer. The cross-linking reaction was performed twice to allow for the addition of up to 20 mM BDP-NHP to increase the duration of cellular exposure to the soluble cross-linker. The



Homodimeric-intra-link sites	Non-linear Cross-linked distance between C_β- C_β
K_{248}-K'_{248}	9.8
K_{250}-K'_{250}	11.6
K_{291}-K'_{291}	39.6

Figure 5. Sequence Alignment of OmpA-Like Domains of OprF and *E. coli* Homolog OmpA, Showing Locations of Cross-Linked Sites and the Top Selected OprF Homodimeric Model, Based on the Cross-Link-Derived Distance Constraints

(A) Sequence alignment of the OmpA-like regions of *PA* OprF (black) and *E. coli* homolog OmpA (red). The sites of homodimeric intra-links in OmpA identified by in vivo cross-linking (Zheng et al., 2011; Weisbrod et al., 2013a) are indicated by dotted arrows, and the sequence region reported to be important for OmpA dimerization (Marcoux et al., 2014) is underlined (dotted line). All cross-linked Lys residues of OprF are shown in bold. The peptide sequences containing sites of homodimeric intra-links (K248, K250, and K291) are underlined by solid lines and the corresponding cross-linked Lys residues are marked by solid arrows. Putative PG-binding sites on OprF (residues 284–297) are shown in blue.

(B) Top OprF dimer model with the shortest Xwalk distance between K248 (red), K250 (cyan), and K291 (yellow). The space filled region highlights the predicted dimer interface and the aligned PG-binding site.

(C) Xwalk distances between Lys residues are shown in the table under the rotated view of the dimer model. See Figures S4C and S4D.

cross-linked sample was further processed and analyzed by liquid chromatography-mass spectrometry (LC-MS) as described in the Supplemental Experimental Procedures.

LC-MS Analysis

LC-MS analysis was performed using reverse-phase nano-ultra performance liquid chromatography (Waters) coupled to a Velos-FTICR-MS (Thermo Scientific) (Weisbrod et al., 2013b). The stage 1 samples were analyzed in data-dependent acquisition mode. The cross-linked samples were analyzed by operating the mass spectrometer in ReACT mode (Weisbrod et al., 2013a), which optimizes analysis efficiency by focusing on cross-linked peptide ions that satisfy expected PIR mass relationships, followed by subsequent fragmentation of each of the cross-linked peptide ions (Supplemental Experimental Procedures).

Identification of Cross-Linked Proteins

A cross-linked protein-enriched stage 1 database was constructed as described in the Supplemental Experimental Procedures. The ReACT acquisition data were searched against the stage 1 database using SEQUEST with a variable modification of BDP stump mass (197.032422 Da) on lysine. The resulting peptide matches were filtered at a static FDR of 5%. The retained peptide matches were mapped to their respective ReACT-identified PIR relationships and peptide pairs meeting the following criteria were retained: (1) both peptides must contain a residual BDP modification on the protein N terminus or on an internal lysine side chain, and (2) their respective mass measurement errors must be $\leq$5 ppm.

Protein Structure Prediction and Docking

Monomeric protein structures were predicted with I-TASSER (Zhang, 2008) using user-defined distance constraints ($\leq$35 Å) as are derived from the present cross-linking results. The resulting models were accessed for short linear distances between Cα atoms of intra-cross-linked lysines. Docking of the top monomeric models was performed using SymmDock to generate homodimeric models (Schneidman-Duhovny et al., 2005). All the above steps involved in the cross-linking-derived protein structure modeling are shown in Figure S5. The top 20 docked models were filtered based on the non-linear distances calculated using the Xwalk algorithm (Kahraman et al., 2011) and were deposited into the Model Archive site (http://www.modelarchive.org/). Numbering of cross-linked lysine residues within the protein sequence is consistent with Uniprot sequence annotation, starting with methionine as the first amino acid residue.

ACCESSION NUMBERS

DOI numbers for the predicted protein models have been deposited in the Model Archive Structural Database as follows: SecDF-concatenated-Pao1, ma-a4bdy; OprI-dimer-Pao1, ma-a89gj; OprF-dimer-PAO1, ma-axj2r.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, six tables, OprI dimer, OprF dimer, and SecD-SecF models and can be found with this article online at http://dx.doi.org/10.1016/j.str.2015.01.022.

ACKNOWLEDGMENTS

This work was supported by NIH grants 5R01AI101307, 1R01GM097112, 5R01GM086688, and 7S10RR025107.

Received: October 15, 2014
Revised: January 13, 2015
Accepted: January 23, 2015
Published: March 19, 2015

REFERENCES

Abergel, C., Walburger, A., Chenivesse, S., and Lazdunski, C. (2001). Crystallization and preliminary crystallographic study of the peptidoglycan-associated lipoprotein from *Escherichia coli*. Acta Crystallogr. D Biol. Crystallogr. *57*, 317–319.

Anderson, G.A., Tolic, N., Tang, X., Zheng, C., and Bruce, J.E. (2007). Informatics strategies for large-scale novel cross-linking analysis. J. Proteome Res. *6*, 3412–3421.

Andreeva, A., Howorth, D., Chandonia, J.M., Brenner, S.E., Hubbard, T.J., Chothia, C., and Murzin, A.G. (2008). Data growth and its impact on the SCOP database: new developments. Nucleic Acids Res. *36*, D419–D425.

Arkowitz, R.A., and Wickner, W. (1994). SecD and SecF are required for the proton electrochemical gradient stimulation of preprotein translocation. EMBO J. *13*, 954–963.

Bauer, A., and Kuster, B. (2003). Affinity purification-mass spectrometry. Powerful tools for the characterization of protein complexes. Eur. J. Biochem. *270*, 570–578.

Brinkman, F.S., Bains, M., and Hancock, R.E. (2000). The amino terminus of *Pseudomonas aeruginosa* outer membrane protein OprF forms channels in lipid bilayer membranes: correlation with a three-dimensional model. J. Bacteriol. *182*, 5251–5255.

Cascales, E., and Lloubes, R. (2004). Deletion analyses of the peptidoglycan-associated lipoprotein Pal reveals three independent binding sequences including a TolA box. Mol. Microbiol. *51*, 873–885.

Cascales, E., Bernadac, A., Gavioli, M., Lazzaroni, J.C., and Lloubes, R. (2002). Pal lipoprotein of *Escherichia coli* plays a major role in outer membrane integrity. J. Bacteriol. *184*, 754–759.

Chavez, J.D., Cilia, M., Weisbrod, C.R., Ju, H.J., Eng, J.K., Gray, S.M., and Bruce, J.E. (2012). Cross-linking measurements of the Potato leafroll virus reveal protein interaction topologies required for virion stability, aphid transmission, and virus-plant interactions. J. Proteome Res. *11*, 2968–2981.

Chavez, J.D., Weisbrod, C.R., Zheng, C., Eng, J.K., and Bruce, J.E. (2013). Protein interactions, post-translational modifications and topologies in human cells. Mol. Cell. Proteomics *12*, 1451–1467.

Choi, D.S., Kim, D.K., Choi, S.J., Lee, J., Choi, J.P., Rho, S., Park, S.H., Kim, Y.K., Hwang, D., and Gho, Y.S. (2011). Proteomic analysis of outer membrane vesicles derived from *Pseudomonas aeruginosa*. Proteomics *11*, 3424–3429.

Cui, T., Zhang, L., Wang, X., and He, Z.G. (2009). Uncovering new signaling proteins and potential drug targets through the interactome analysis of *Mycobacterium tuberculosis*. BMC Genomics *10*, 118.

Damron, F.H., Napper, J., Teter, M.A., and Yu, H.D. (2009). Lipotoxin F of *Pseudomonas aeruginosa* is an AlgU-dependent and alginate-independent outer membrane protein involved in resistance to oxidative stress and adhesion to A549 human lung epithelia. Microbiology *155*, 1028–1038.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. Nature *340*, 245–246.

Firoved, A.M., Ornatowski, W., and Deretic, V. (2004). Microarray analysis reveals induction of lipoprotein genes in mucoid *Pseudomonas aeruginosa*: implications for inflammation in cystic fibrosis. Infect. Immun. *72*, 5012–5018.

Goll, J., Rajagopala, S.V., Shiau, S.C., Wu, H., Lamb, B.T., and Uetz, P. (2008). MPIDB: the microbial protein interaction database. Bioinformatics *24*, 1743–1744.

Goure, J., Pastor, A., Faudry, E., Chabert, J., Dessen, A., and Attree, I. (2004). The V antigen of *Pseudomonas aeruginosa* is required for assembly of the functional PopB/PopD translocation pore in host cell membranes. Infect. Immun. *72*, 4741–4750.

Guerrero, C., Tagwerker, C., Kaiser, P., and Huang, L. (2006). An integrated mass spectrometry-based proteomic approach: quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26 S proteasome-interacting network. Mol. Cell. Proteomics *5*, 366–378.

Herzog, F., Kahraman, A., Boehringer, D., Mak, R., Bracher, A., Walzthoeni, T., Leitner, A., Beck, M., Hartl, F.U., Ban, N., et al. (2012). Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. Science *337*, 1348–1352.

Ishida, H., Garcia-Herrero, A., and Vogel, H.J. (2014). The periplasmic domain of *Escherichia coli* outer membrane protein A can undergo a localized temperature dependent structural transition. Biochim. Biophys. Acta *1838*, 3014–3024.

Kaake, R.M., Wang, X., Burke, A., Yu, C., Kandur, W., Yang, Y., Novtisky, E.J., Second, T., Duan, J., Kao, A., et al. (2014). A new in vivo cross-linking mass spectrometry platform to define protein-protein interactions in living cells. Mol. Cell. Proteomics *13*, 3533–3543.

Kahraman, A., Malmstrom, L., and Aebersold, R. (2011). Xwalk: computing and visualizing distances in cross-linking experiments. Bioinformatics *27*, 2163–2164.

Khalid, S., Bond, P.J., Deol, S.S., and Sansom, M.S. (2006). Modeling and simulations of a bacterial outer membrane protein: OprF from *Pseudomonas aeruginosa*. Proteins *63*, 6–15.

Lin, Y.M., Wu, S.J., Chang, T.W., Wang, C.F., Suen, C.S., Hwang, M.J., Chang, M.D., Chen, Y.T., and Liao, Y.D. (2010). Outer membrane protein I of *Pseudomonas aeruginosa* is a target of cationic antimicrobial peptide/protein. J. Biol. Chem. *285*, 8985–8994.

Liu, F., Wu, C., Sweedler, J.V., and Goshe, M.B. (2012). An enhanced protein crosslink identification strategy using CID-cleavable chemical crosslinkers and LC/MS(n) analysis. Proteomics *12*, 401–405.

Marcoux, J., Politis, A., Rinehart, D., Marshall, D.P., Wallace, M.I., Tamm, L.K., and Robinson, C.V. (2014). Mass spectrometry defines the C-terminal dimerization domain and enables modeling of the structure of full-length OmpA. Structure *22*, 781–790.

Marianayagam, N.J., Sunde, M., and Matthews, J.M. (2004). The power of two: protein dimerization in biology. Trends Biochem. Sci. *29*, 618–625.

Marquordt, C., Fang, Q., Will, E., Peng, J., von Figura, K., and Dierks, T. (2003). Posttranslational modification of serine to formylglycine in bacterial sulfatases. Recognition of the modification motif by the iron-sulfur protein AtsB. J. Biol. Chem. *278*, 2212–2218.

Mizuno, T., and Kageyama, M. (1979a). Isolation and characterization of major outer membrane proteins of *Pseudomonas aeruginosa* strain PAO with special reference to peptidoglycan-associated protein. J. Biochem. *86*, 979–989.

Mizuno, T., and Kageyama, M. (1979b). Isolation of characterization of a major outer membrane protein of *Pseudomonas aeruginosa*. Evidence for the occurrence of a lipoprotein. J. Biochem. *85*, 115–122.

Nouwen, N., Piwowarek, M., Berrelkamp, G., and Driessen, A.J. (2005). The large first periplasmic loop of SecD and SecF plays an important role in SecDF functioning. J. Bacteriol. *187*, 5857–5860.

Oliver, A., Canton, R., Campo, P., Baquero, F., and Blazquez, J. (2000). High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. Science *288*, 1251–1254.

Planas-Iglesias, J., Bonet, J., Garcia-Garcia, J., Marin-Lopez, M.A., Feliu, E., and Oliva, B. (2013). Understanding protein-protein interactions using local structural features. J. Mol. Biol. *425*, 1210–1224.

Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Seraphin, B. (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods *24*, 218–229.

Rawling, E.G., Martin, N.L., and Hancock, R.E. (1995). Epitope mapping of the *Pseudomonas aeruginosa* major outer membrane porin protein OprF. Infect. Immun. *63*, 38–42.

Rawling, E.G., Brinkman, F.S., and Hancock, R.E. (1998). Roles of the carboxy-terminal half of *Pseudomonas aeruginosa* major outer membrane protein OprF in cell shape, growth in low-osmolarity medium, and peptidoglycan association. J. Bacteriol. *180*, 3556–3562.

Rinner, O., Seebacher, J., Walzthoeni, T., Mueller, L.N., Beck, M., Schmidt, A., Mueller, M., and Aebersold, R. (2008). Identification of cross-linked peptides from large sequence databases. Nat. Methods *5*, 315–318.

Robert, V., Filloux, A., and Michel, G.P. (2005). Subcomplexes from the Xcp secretion system of *Pseudomonas aeruginosa*. FEMS Microbiol. Lett. *252*, 43–50.

Rosenstein, B.J., and Zeitlin, P.L. (1998). Cystic fibrosis. Lancet *351*, 277–282.

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. *5*, 725–738.

Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., and Wolfson, H.J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res. *33*, W363–W367.

Shu, W., Liu, J., Ji, H., and Lu, M. (2000). Core structure of the outer membrane lipoprotein from *Escherichia coli* at 1.9 A resolution. J. Mol. Biol. *299*, 1101–1112.

Singh, P., Panchaud, A., and Goodlett, D.R. (2010). Chemical cross-linking and mass spectrometry as a low-resolution protein structure determination technique. Anal. Chem. *82*, 2636–2642.

Sinz, A. (2003). Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. J. Mass Spectrom. *38*, 1225–1237.

Sugawara, E., Nestorovich, E.M., Bezrukov, S.M., and Nikaido, H. (2006). *Pseudomonas aeruginosa* porin OprF exists in two different conformations. J. Biol. Chem. *281*, 16220–16229.

Tang, X., Munske, G.R., Siems, W.F., and Bruce, J.E. (2005). Mass spectrometry identifiable cross-linking strategy for studying protein-protein interactions. Anal. Chem. *77*, 311–318.

Tang, X., Yi, W., Munske, G.R., Adhikari, D.P., Zakharova, N.L., and Bruce, J.E. (2007). Profiling the membrane proteome of *Shewanella oneidensis* MR-1 with new affinity labeling probes. J. Proteome Res. *6*, 724–734.

Trnka, M.J., Baker, P.R., Robinson, P.J., Burlingame, A.L., and Chalkley, R.J. (2014). Matching cross-linked peptide spectra: only as good as the worse identification. Mol. Cell. Proteomics *13*, 420–434.

Tsukazaki, T., Mori, H., Echizen, Y., Ishitani, R., Fukai, S., Tanaka, T., Perederina, A., Vassylyev, D.G., Kohno, T., Maturana, A.D., et al. (2011). Structure and function of a membrane component SecDF that enhances protein export. Nature *474*, 235–238.

Vasilescu, J., Guo, X., and Kast, J. (2004). Identification of protein-protein interactions using in vivo cross-linking and mass spectrometry. Proteomics *4*, 3845–3854.

Wallin, E., and von Heijne, G. (1998). Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Sci. *7*, 1029–1038.

Weisbrod, C.R., Chavez, J.D., Eng, J.K., Yang, L., Zheng, C.X., and Bruce, J.E. (2013a). In vivo protein interaction network identified with a novel real-time cross-linked peptide identification strategy. J. Proteome Res. *12*, 1569–1579.

Weisbrod, C.R., Hoopmann, M.R., Senko, M.W., and Bruce, J.E. (2013b). Performance evaluation of a dual linear ion trap-Fourier transform ion cyclotron resonance mass spectrometer for proteomics research. J. Proteomics *88*, 109–119.

Woodruff, W.A., and Hancock, R.E. (1988). Construction and characterization of *Pseudomonas aeruginosa* protein F-deficient mutants after in vitro and in vivo insertion mutagenesis of the cloned gene. J. Bacteriol. *170*, 2592–2598.

Yang, B., Wu, Y.J., Zhu, M., Fan, S.B., Lin, J., Zhang, K., Li, S., Chi, H., Li, Y.X., Chen, H.F., et al. (2012). Identification of cross-linked peptides from complex samples. Nat. Methods *9*, 904–906.

Yildirim, M.A., Goh, K.I., Cusick, M.E., Barabasi, A.L., and Vidal, M. (2007). Drug-target network. Nat. Biotechnol. *25*, 1119–1126.

Yu, X., Sharma, K.D., Takahashi, T., Iwamoto, R., and Mekada, E. (2002). Ligand-independent dimer formation of epidermal growth factor receptor (EGFR) is a step separable from ligand-induced EGFR signaling. Mol. Biol. Cell *13*, 2547–2557.

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. BMC Bioinformatics *9*, 40.

Zhang, H., Tang, X., Munske, G.R., Zakharova, N., Yang, L., Zheng, C., Wolff, M.A., Tolic, N., Anderson, G.A., Shi, L., et al. (2008). In vivo identification of the outer membrane protein OmcA-MtrC interaction network in Shewanella oneidensis MR-1 cells using novel hydrophobic chemical cross-linkers. J. Proteome Res. *7*, 1712–1720.

Zhang, H., Tang, X., Munske, G.R., Tolic, N., Anderson, G.A., and Bruce, J.E. (2009). Identification of protein-protein interactions and topologies in living cells with chemical cross-linking and mass spectrometry. Mol. Cell. Proteomics *8*, 409–420.

Zhang, M., Su, S., Bhatnagar, R.K., Hassett, D.J., and Lu, L.J. (2012). Prediction and analysis of the protein interactome in *Pseudomonas aeruginosa* to enable network-based drug target selection. PLoS One *7*, e41202.

Zheng, C., Yang, L., Hoopmann, M.R., Eng, J.K., Tang, X., Weisbrod, C.R., and Bruce, J.E. (2011). Cross-linking measurements of in vivo protein complex topologies. Mol. Cell. Proteomics *10*, M110 006841.