## Title

High-throughput Isolation and Characterization of Untagged Membrane Protein Complexes: Outer Membrane Complexes of Desulfovibrio vulgaris



## Journal

Journal of Proteome Research, 11(12)

## ISSN

1535-3893

## Authors

Walian, Peter J
Allen, Simon
Shatsky, Maxim
et al.

## Publication Date

2012-12-07

## DOI

10.1021/pr300548d

Peer reviewed

Journal of proteome research

Article

pubs.acs.org/jpr

# High-throughput Isolation and Characterization of Untagged Membrane Protein Complexes: Outer Membrane Complexes of *Desulfovibrio vulgaris*

Peter J. Walian,[*,†,‡] Simon Allen,[‡,§] Maxim Shatsky,[†,‡] Lucy Zeng,[†] Evelin D. Szakal,[§,∥] Haichuan Liu,[§] Steven C. Hall,[§] Susan J. Fisher,[†,§] Bonita R. Lam,[†] Mary E. Singer,[†] Jil T. Geller,[†] Steven E. Brenner,[⊥] John-Marc Chandonia,[†,⊥] Terry C. Hazen,[#] H. Ewa Witkowska,[§] Mark D. Biggin,[†] and Bing K. Jap[†]

[†]Lawrence Berkeley National Laboratory, Berkeley, California, United States

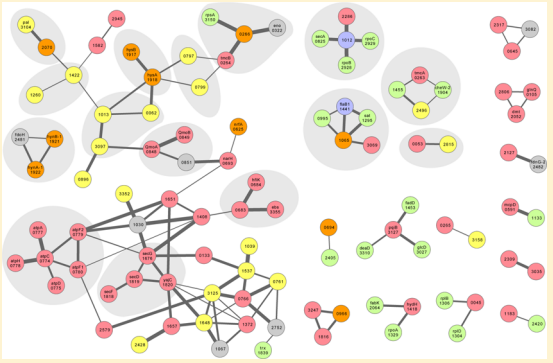
[§]University of California, San Francisco, California, United States

[⊥]University of California, Berkeley, California, United States

[#]University of Tennessee, Knoxville, Tennessee, United States

Supporting Information

**ABSTRACT:** Cell membranes represent the "front line" of cellular defense and the interface between a cell and its environment. To determine the range of proteins and protein complexes that are present in the cell membranes of a target organism, we have utilized a "tagless" process for the system-wide isolation and identification of native membrane protein complexes. As an initial subject for study, we have chosen the Gram-negative sulfate-reducing bacterium *Desulfovibrio vulgaris*. With this tagless methodology, we have identified about two-thirds of the outer membrane-associated proteins anticipated. Approximately three-fourths of these appear to form homomeric complexes. Statistical and machine-learning methods used to analyze data compiled over multiple experiments revealed networks of additional protein−protein interactions providing insight into heteromeric contacts made between proteins across this region of the cell. Taken together, these results establish a *D. vulgaris* outer membrane protein data set that will be essential for the detection and characterization of environment-driven changes in the outer membrane proteome and in the modeling of stress response pathways. The workflow utilized here should be effective for the global characterization of membrane protein complexes in a wide range of organisms.



**KEYWORDS:** *membrane proteins, protein complexes, protein−protein interactions, mass spectrometry, Gram-negative bacteria, interaction networks*

## INTRODUCTION

Cell membranes form the critical interface between a cell and its environment. Significant changes in response to environmental conditions are expected to take place through the proteins situated within these membranes. Membrane protein-associated changes may occur in the form of abundance level, protein−protein interactions, post-translational modifications and even mutations. To understand some of the earliest and perhaps most critical responses to stress, characterization of these changes on a molecular level is needed.

Numerous challenges arise when preparing membrane proteins in forms suitable for proteomics studies.[1−3] The preparation of stable, intact membrane protein complexes, in which native structures and protein−protein interactions have been preserved, poses a unique purification challenge largely due to the requirement for detergent solubilization.[4] It is widely recognized that membrane protein complex stability is sensitive to the method of solubilization and isolation and that the use of an unsuitable detergent or detergent-to-protein ratio, for example, can lead to the disruption of complexes or even the association of proteins into biologically irrelevant aggregates. Given the significant biochemical challenges, relatively few studies focusing on the membrane protein proteome of a target organism have been conducted. Studies of membrane protein complexes requiring the preservation of native-state interactions pose even greater difficulties. Significant progress in this area however has been reported by groups using native gel methods to preserve complexes prior to separation by SDS-PAGE.[5−8] Even finer control of the process has been gained through the coupling of chromatography with native gel methods facilitating the identification of 44 inner and outer membrane protein complexes from *Helicobacter pylori*[9] and 30 complexes of the *Escherichia coli* inner and outer membrane.[10]

**Received:** June 18, 2012
**Published:** October 25, 2012

In parallel with the aforementioned efforts, we have employed a "tagless" methodology to isolate, under conditions intended to maximize the probability of preserving native interactions, endogenously expressed membrane protein complexes for subsequent identification and characterization. This workflow involves successive mild but effective detergent solubilization optimized for specific membrane types, liquid chromatography, and both native and denaturing gel electrophoresis. Membrane proteins isolated through this process are subjected to in-gel digestion and identification by mass spectrometry (MS). In contrast to strategies employing affinity tags for the targeted purification of selected proteins, use of this "tagless" strategy is aimed at obtaining proteome-wide views of a target organism's membrane protein complexes.

As an initial subject for our studies we examined the membrane protein complexes present in the outer membrane of the bacterium, *Desulfovibrio vulgaris*. *D. vulgaris* (strain Hildenborough) is a Gram-negative sulfate-reducing bacterium recognized for its ability to reduce heavy metals and survive in physiologically demanding conditions.[11−13] Application of this microbe as a key component of large-scale bioremediation strategies appears promising.[14] The *D. vulgaris* genome[15] has been recently revised and found to contain 3403 protein-coding genes distributed across one genomic chromosome and one large plasmid.[16] The cataloging and characterization of protein complexes from this organism, grown under standard and stressed conditions, will provide data critical for modeling stress responses in *D. vulgaris* relevant to the efficient detoxification of heavy metal and radionuclide contaminated sites.

Here we report the results from our studies of *D. vulgaris* outer membrane preparations, derived from cultures grown to late exponential phase under standard conditions. MS analysis of proteins isolated through this workflow resulted in the identification of 296 proteins; of these, 70 are proposed to be outer membrane associated. We found that the bulk of highly stable *D. vulgaris* outer membrane protein complexes appear to be homomeric. To identify additional and potentially weaker interactions, statistical and machine learning-based methods were used. This analysis revealed a range of heteromeric protein−protein interactions taking place between proteins from different cellular compartments that included a number of homomeric complexes. These results are depicted here in interaction network format. The complex subunit identifications obtained from *D. vulgaris* cultures prepared under standard growth conditions have established a baseline outer membrane protein complex data set for this organism. As the processing methodology described here can be applied to cultures prepared under a range of growth conditions, in addition to providing a catalog of *D. vulgaris* outer membrane protein complexes, these data will serve as an essential reference for the detection and characterization of changes in the membrane protein complexes of cultures subjected to different environmental stressors. The results presented demonstrate the potential of this approach for managing the challenging task of globally processing and characterizing the membrane protein complexes of target organisms.

## MATERIALS AND METHODS

### Cell Growth

*D. vulgaris* Hildenborough (ATCC 29579) was obtained from the American Type Culture Collection (Manassas, VA). For this study five *D. vulgaris* culture sets were grown in defined lactate sulfate medium (LS4D medium) at 30 °C[17] and independently processed. LS4D medium consisted of 60 mM sodium lactate, 50 mM $Na_2SO_4$, 8.0 mM $MgCl_2$, 20 mM $NH_4Cl$, 2.2 mM $K_2HPO_4$, 0.6 mM $CaCl_2$, 30 mM PIPES [piperazine-N, N-bis (2-ethanesulfonic acid)], 12.5 mL of a trace mineral solution per liter, NaOH (to adjust the pH to 7.2), and 1.0 mL of a 10× vitamin solution per liter that was added after autoclaving.[18] The reductant used for LS4D medium was 5 mL per liter of an anaerobic titanium citrate solution. This solution contained 20% (wt/vol) titanium(III) chloride, 0.2 M sodium citrate, and 8.0% (wt/vol) $Na_2CO_3$. Cell growth was monitored using the optical density at 600 nm (OD600). Samples were harvested at the late exponential phase (OD600, ∼0.6) and stored at −80 °C.

### Cell Membrane Isolation

To reduce the presence of iron sulfide present as a consequence of growth in LS4D media cells were washed in cell wash buffer (20 mM HEPES, pH 7.4, 2 mM $NaN_3$, 100 mM KCl, 0.1 mM EDTA, 1 mM $MgCl_2$, 125 mM sucrose) prior to lysis. *D. vulgaris* cell pellets were resuspended in cell wash buffer and gently stirred at 4 °C until a uniform suspension was obtained. A broad-spectrum protease inhibitor (Complete, Roche) was added to the wash and lysis buffers. The suspension was transferred to 500 mL centrifuge bottles and spun at 10 000× *g* for 10 min. Washed cell pellets were resuspended in lysis buffer (25 mM HEPES, pH 7.6, 100 mM KCl, 12.5 mM $MgCl_2$, 0.1 mM EDTA, 20% glycerol) and processed through a gas-driven cell disruptor (EmulsiFlex-C5, Avestin) three times to break the cells open. To enhance the preservation of membrane protein complexes the cell disruptor was chilled with ice. The broken cell suspension was spun at low speed (10 000× *g* for 10 min) to remove unbroken cells; the supernatant from this step was spun at high speed (100 000× *g*, 1 h, 4 °C) to isolate membranes. To reduce the presence of high abundance soluble proteins, the membrane pellet was resuspended and washed several times in membrane wash buffer (20 mM HEPES, pH 7.4, 2 mM $NaN_3$, 100 mM NaCl, 1 mM $MgCl_2$, 0.1 mM EDTA). After each wash, membranes were pelleted by high speed centrifugation (100 000× *g*, 1 h, 4 °C). Membranes were either used immediately or quick-frozen and stored at −80 °C.

### Membrane Solubilization

The proteins of *D. vulgaris* cell membranes were extracted in a two-step process, generally based on the protocol of Baldermann et al.,[19] which targets solubilization of predominantly inner membrane proteins in the initial step and proteins of the outer membrane in the second. The cell membranes obtained from five separately prepared late exponential phase cultures were processed in this manner for our study. Four of the five experiments utilized the detergents $C_{12}E_9$ and octyl glucoside (OG) for solubilization of the inner and outer membranes respectively, while in one experiment the inner membrane was solubilized with Triton X-100 and the outer membrane with octyl POE (OP). Specifically, to extract proteins of the inner membrane, washed membranes were placed into a hand homogenizer along with solubilization buffer (20 mM HEPES pH7.4, 2 mM $NaN_3$, 100 mM NaCl, 1 mM $MgCl_2$ and 0.2 mg/mL lysozyme) containing 0.1% $C_{12}E_9$ or 1% Triton X-100 and solubilized on ice for 1 h. Samples were adjusted to obtain a final detergent-to-protein ratio of 1:1 while maintaining a protein concentration of 10 mg/mL. After solubilization, the detergent concentration was lowered to half the initial level by slow dilution with solubilization buffer only, and the sample centrifuged (100 000× *g* for 1.5 h at 4 °C) to pellet unsolubilized membranes enriched in outer membrane proteins. Outer


dx.doi.org/10.1021/pr300548d | J. Proteome Res. 2012, 11, 5720−5735

membrane pellets were processed in solubilization buffer containing a more aggressive detergent, either OG or OP, to extract proteins of the outer membrane. These membranes were placed into a hand homogenizer along with solubilization buffer containing 2% OP or 3% OG and solubilized on ice for 1 h. After solubilization, the detergent concentration was lowered by dilution with solubilization buffer (2× or 4× depending on the detergent used) and the sample centrifuged (100 000× *g* for 1.5 h at 4 °C). Isolated protein complexes were not frozen. The solubilized proteins of the outer membrane were then subjected to ion exchange chromatography.

### Ion Exchange Chromatography

Dependent on the amount of cell membrane processed, a 2−10 mL bed of anion exchange media (Q Sepharose HP; GE Healthcare) was equilibrated in buffer A (20 mM HEPES pH 7.4, 2 mM $NaN_3$) containing 2−3× the critical micelle concentration (CMC) of the solubilizing detergent. Detergent-solubilized samples were loaded onto the column and the column was washed to remove nonspecifically bound material using buffer A. Bound proteins were step eluted (with buffer A containing 1 M NaCl) in 50 mM increments over a range of 50−400 mM NaCl followed by a final step of 1 M NaCl. Eluate was collected in 2 or 8 mL fractions.

### Molecular Sieve Chromatography

Molecular sieve chromatography was used to characterize and compare the properties of proteins solubilized in different detergents. Prior to molecular sieve chromatography, ion exchange (IEX) peak fractions were concentrated from 2 to 10-fold using a centrifugal filter device (Amicon Ultra, Millipore); 200 μL samples were injected into a Superdex 200 16/300 column (GE Healthcare) which had been equilibrated with running buffer B (20 mM HEPES pH7.4, 2 mM $NaN_3$, 100 mM NaCl) containing 2−3× CMC of the solubilizing detergent. Eluate was collected in 0.25 mL fractions.

### Blue Native and SDS Polyacrylamide Gel Electrophoresis

Blue native PAGE of the IEX and molecular sieve chromatography fractions was performed based on the protocol of Schägger et al.[20] with modifications. Briefly, 45 μL outer membrane preparation samples were mixed with 5 μL of glycerol and 3−4 μL of a stock of 5.0% Coomassie G-250 in 1 M amino-caproic acid. Precast acrylamide gels (4−12% Bis-Tris, 1.0 mm, Invitrogen) were equilibrated in a cathode buffer (50 mM Bis-Tris, pH 7.0) containing 0.05% dodecyl maltoside (DDM) and using an anode buffer of 20 mM Bis-Tris and 30 mM Tricine, pH 7.0. When running the gel with samples loaded, the cathode buffer was replaced with a solution containing 50 mM Bis-Tris, pH 7.0, 0.02% Coomassie G-250 and 0.05% DDM. The gels were run overnight at 70 V and 4 °C. Whole lanes containing multiple protein bands were cut from the native gels, 6 cm in length and 0.5 cm in width, and placed length-wise across the tops of gels containing one wide sample well in addition to the standards well (4−12% Bis-Tris, 1.5 mm, 2D, Invitrogen) for SDS-PAGE. Native gel lanes were incubated in Laemmli sample buffer (Sigma) for 20 min prior to running the SDS-PAGE gels with MOPS running buffer (Invitrogen) according to the manufacturer's directions.

### Preparation of Gel-based Samples for Mass Spectrometry

Second dimension SDS-PAGE gels were scanned for record keeping and visible spots were assigned an identification number. Visible gel spots were excised manually, transferred to 96-well digester plates (Digilab, Inc.) and stored at −20 °C until digestion. Gel pieces were subjected to an automated digestion using a ProGest robot (Digilab, Inc.). Briefly, the gel pieces were stripped of Coomassie blue stain and dehydrated with acetonitrile, the proteins were reduced with 10 mM DTT (60 °C, 30 min), and the reduced cysteine residues were then alkylated with 100 mM iodoacetamide (37 °C, 45 min). Prior to enzymatic digestion excess reagents were removed and the gel pieces were washed twice with 25 mM ammonium bicarbonate, dehydrated, and incubated with 250 ng sequencing grade trypsin (37 °C for 4 h). The resulting tryptic peptides were extracted from the gel with 10% formic acid.[21]

### Mass Spectrometry and Protein Identification in the Protein Complex Workflow

Digested samples were analyzed by MS using either LC−MALDI-MS/MS (AB Sciex 4800 MALDI-TOF/TOF) or LC−ESI−MS/MS (Thermo LTQ-XL) platforms. For further details see Supplementary Methods (Supporting Information).

The resulting data were searched against a custom *D. vulgaris* (Hildenborough) protein database (containing common contaminants) using ProteinPilot 3.0 (version 114732, AB Sciex). In the case of MALDI data t2d files were submitted directly to the search engine, for LTQ data the raw files were first converted to mgf files using Mascot Daemon software (Matrix Science) prior to their submission to ProteinPilot.

### Proteomics Survey of Outer Membrane Proteins

Fourteen IEX fractions representing the first step of the tagless workflow for the outer membrane protein complex fractionation were analyzed in parallel by two approaches: gel LC−MS and 2D LC−MS. For the gel LC−MS workflow, proteins from each IEX fraction were separated by SDS-PAGE (4−12% Bis-Tris, 1.0 mm, Invitrogen), 269 bands were cut out and proteins were digested robotically with trypsin as described above. For the 2D LC−MS workflow, proteins were digested in solution according to the published protocol[23] and mixtures of proteolytic peptides were submitted to off-line fractionation by reversed phase HPLC at alkaline pH[24,25] utilizing a Zorbax Extend column (4.6 × 100 mm, Agilent) on a Michrom Paradigm MS4 HPLC system equipped with a LEAP Technologies PAL autosampler. Samples were desalted using Oasis cartridges (Waters) prior to injection onto the column. The analytical column was developed at a flow rate of 0.7 mL/min using the following gradient: (i) 10 min isocratic at 6% B, (ii) linear increase from 6 to 38% B in 30 min, (iii) linear increase from 38 to 100% B in 6 min, (iv) isocratic at 100% B for 2 min, (v) re-equilibration at the initial conditions of 6% for 10 min, where solvent A was 0.1% ammonium hydroxide in water and solvent B was 0.1% ammonium hydroxide in 80% acetonitrile. Chromatography was monitored by UV absorption at 230 nm. Two-minute fractions were collected and then pooled to generate nine high pH samples per IEX fraction for further MS analysis. Solvents were removed by using vacuum centrifugation and peptide mixtures were reconstituted in 0.1% formic acid. Peptide mixtures derived from both the gel LC−MS and 2D LC−MS protocols were subjected to nanoLC−ESI−MS/MS analysis utilizing a LTQ XL mass spectrometer as described above. In total, 269 and 126 LC MS/MS analyses were performed for samples generated via the gel LC−MS and 2D LC−MS workflows, respectively.

### Assignment of Outer Membrane Localization

The list of 70 putative outer membrane-associated proteins (presented in Table 1) was culled from the larger list of 296


dx.doi.org/10.1021/pr300548d | *J. Proteome Res.* 2012, 11, 5720−5735

**Table 1. *D. vulgaris* Outer Membrane-associated Proteins Identified and the Oligomeric States Observed**

| Gene ID | annotation[a] | putative oligomeric state(s) |
|---|---|---|
| DVU0062 | RND efflux system, outer membrane protein, NodT family | Homotrimer, homotetramer |
| DVU0064 | Hypothetical protein | Homodimer |
| DVU0100 | TonB-dependent receptor | Monomer |
| DVU0133 | Hypothetical protein | Multimer |
| DVU0147 | Lipoprotein, putative | Homopentamer |
| DVU0243 | Lipoprotein, putative | Homodimer |
| DVU0249 | Lipoprotein, putative | Homodimer |
| DVU0255 | Hypothetical protein | Multimer |
| DVU0266 | Hypothetical protein | Homodimer |
| DVU0273 | Conserved hypothetical protein | Homotrimer |
| DVU0371 | Conserved hypothetical protein | Monomer, homodimer |
| DVU0397 | Rare lipoprotein A, putative | Homodimer, homotrimer |
| DVU0609 | Lipoprotein, putative | Homodimer |
| DVU0610 | Conserved hypothetical protein | Homotrimer |
| DVU0761 | Lipoprotein, putative | Homotrimer |
| DVU0766 | Transporter, putative | Multimer |
| DVU0797 | Conserved hypothetical protein | Homotrimer |
| DVU0799 | Conserved hypothetical protein | Homotrimer |
| DVU0851 | Hypothetical protein | Monomer, heteromer with 0848, 0693 |
| DVU0896 | Lipoprotein, NLP/P60 family | Homotrimer |
| DVU0954 | Organic solvent tolerance protein, putative | Monomer |
| DVU1008 | Hypothetical protein | Homodimer |
| DVU1012 | Hemolysin-type calcium-binding repeat protein | Monomer |
| DVU1013 | Type I secretion outer membrane protein, TolC family | Homotrimer |
| DVU1039 | Lipoprotein, putative | Homodimer, homotrimer |
| DVU1045 | Hypothetical protein | Homotrimer |
| DVU1065 | Peptidyl-prolyl cis−trans isomerase domain protein | Homodimer |
| DVU1067 | Membrane protein, Bmp family | Homodimer, homotrimer |
| DVU1195 | Lipoprotein, putative | Multimer |
| DVU1260 | Outer membrane protein P1, putative | Monomer, homodimer, homotrimer |
| DVU1273 | Bacterial type II/III secretion system protein | Homotrimer |
| DVU1408 | Hypothetical protein | Multimer |
| DVU1422 | OmpA family protein | Homodimer, homotrimer, homotetramer |
| DVU1455 | Conserved hypothetical protein | Homodimer |
| DVU1537 | Lipoprotein, putative | Multimer |
| DVU1548 | Outer membrane protein, OmpP1/FadL/TodX family | Monomer, homohexamer |
| DVU1581 | Hypothetical protein | Multimer |
| DVU1648 | Lipoprotein, putative | Homotetramer |
| DVU1657 | Hypothetical protein | Multimer |
| DVU1758 | Lipoprotein, putative | Homotetramer |
| DVU1837 | Competence protein, putative | Homodimer |
| DVU1842 | Lipoprotein, putative | Homodimer, homopentamer |
| DVU1887 | Hypothetical protein | Monomer |
| DVU1902 | Conserved hypothetical protein | Homodimer |
| DVU1917 | Periplasmic [NiFeSe] hydrogenase, small subunit | Heteromer with 1918 |
| DVU1918 | Periplasmic [NiFeSe] hydrogenase, large subunit | Heteromer with 1917 |
| DVU1921 | Periplasmic [NiFe] hydrogenase small subunit | Heteromer with 1922 |
| DVU1922 | Periplasmic [NiFe] hydrogenase large subunit | Heteromer with 1921 |
| DVU1952 | Hypothetical Protein | Homodimer, homotrimer |
| DVU2070 | TPR domain protein | Heteromer with 3104 |
| DVU2373 | Outer membrane protein, OMP85 family | Monomer, homodimer |
| DVU2428 | Lipoprotein, putative | Multimer |
| DVU2496 | Lipoprotein, putative | Homotetramer, homooctamer |
| DVU2497 | Lipoprotein, putative | Homooctamer |
| DVU2523 | Lipoprotein, putative | Multimer |
| DVU2579 | TPR domain protein | Homodimer |
| DVU2614 | Hypothetical protein | Multimer |
| DVU2628 | TPR domain protein | Homotrimer, homohexamer |
| DVU2630 | Lipoprotein, putative | Homodimer, homotrimer |
| DVU2815 | Outer membrane efflux protein | Homotetramer |
| DVU3090 | Outer membrane protein, OmpP1/FadL/TodX family | Monomer, homodimer |
| DVU3097 | Outer membrane efflux protein | Homotrimer |
| DVU3104 | Peptidoglycan-associated lipoprotein, putative | Heteromer with 2070, multimer |
| DVU3125 | Lipoprotein, putative | Homotetramer, multimer |
| DVU3141 | Lipoprotein, putative | Homotetramer |
| DVU3158 | VacJ lipoprotein, putative | Homodimer, homotetramer |
| DVU3344 | Hypothetical protein | Multimer |
| DVU3352 | Lipoprotein, putative | Homotrimer |
| DVUA0117 | Type III secretion lipoprotein | Homodimer |
| DVUA0147 | Conserved hypothetical protein | Homotetramer |

[a]Annotations listed as provided in MicrobesOnline (www.microbesonline.com).

proteins identified in the outer membrane preparations (Supplementary Tables 3 and 4, Supporting Information), removing those indicated to be derived from other cellular compartments such as the inner membrane or cytoplasmic space with the assistance of annotations from databases such as MicrobesOnline (http://www.microbesonline.org/). Proteins for which no definitive annotation was available (e.g., "hypothetical protein" or "conserved hypothetical protein") but which were consistently detected across the outer membrane preparations and found to have biophysical characteristics consistent with, and in some cases significant sequence homologies to, known outer membrane structures were kept on the list.

## Estimation of Homo-oligomeric State

For putative complexes that displayed only one constituent subunit in second dimension SDS-PAGE gels, assignment of a homo-oligomeric state was made by dividing the complex molecular weight estimated from native gel migration (adjusted for an average detergent-Coomassie blue dye micelle weight contribution of 20 kDa for complex weights greater than 160 kDa, 10 kDa for total weights in the range 100−160 kDa, and 5 kDa for total weights below 100 kDa) by the subunit molecular weight calculated from its amino acid sequence. In cases where the stoichiometry could not be discerned, due to a broad range in


dx.doi.org/10.1021/pr300548d | J. Proteome Res. 2012, 11, 5720−5735

native molecular weights observed, complexes were defined only as multimeric.

## Identification of Heteromeric Membrane Protein Complexes and Interactions

Although in many cases complexes, both homo- and heteromeric, could be identified through direct inspection of second dimension SDS-PAGE gels, the identification of weaker interactions and reduction of false positives required detailed analysis of the data across multiple experiments. Statistical and machine learning methods were therefore applied to extract additional biologically significant signals from the data. We derived a number of scoring functions that were predictive of whether two proteins were in fact present in a single heteromeric protein complex. These functions were applied to each pair of proteins that were observed to be from the same region of a blue native polyacrylamide gel electrophoresis (BN-PAGE) gel and generated an array of scores. A machine-learning method trained on gold standard sets was used to integrate these scores and predict whether each pair of proteins were members of a true heteromeric complex. Details of this procedure are described below.

**Gold Standards.** Computational analysis was performed using curated gold standard sets of interacting and noninteracting pairs of proteins. Because the majority of these pairs have not been previously experimentally validated for *D. vulgaris*, they should be considered an "imperfect" gold standard. The positive gold standard set of interacting proteins includes pairs of *D. vulgaris* proteins that interact in stable complexes previously identified using low throughput experiments, as well as pairs of *D. vulgaris* proteins that were mapped to *E. coli* proteins annotated as interacting either in EcoCyc version 12.0[26] or in a recent set of reciprocal tandem affinity purification (TAP) experiments in *E. coli*.[27] This data set was then curated to account for known differences between *E. coli* and *D. vulgaris* complexes (e.g., the degradosome complex is not present in *D. vulgaris* due to the truncation of a scaffold protein[28]). We also excluded all interactions with ribosomal proteins, as this complex is atypical due to the RNA component as well as highly abundant, leading to many potential false positives. This resulted in an initial set of 12 pairs of proteins that were also observed in our data (i.e., comigrated in BN-PAGE gels). We supplemented this data set with an additional five pairs of proteins obtained by manual curation of a set of predicted protein complexes that were identified in a high throughput survey of *D. vulgaris* complexes (M. Biggin, unpublished data) that were also observed in our data; the complete set of pairs is listed in Supplementary Table 1 (Supporting Information). The probability of observing these 17 pairs solely due to chance was calculated (i.e., we estimated how many pairs would overlap between the gold standard set and a randomly shuffled *D. vulgaris* data set). On the basis of 10 000 shuffles, the probability of sharing 17 or more protein pairs was only 0.0127, implying that the magnitude of overlap between the two data sets being due to chance was highly unlikely.

A negative gold standard set of noninteracting protein pairs was prepared by randomizing pairs of proteins from the positive gold standard set. We included all pairs of *D. vulgaris* proteins mapping to *E. coli* proteins that (a) were present in a heteromeric complex in EcoCyc, but not observed to interact with each other in either EcoCyc or TAP experiments (both reciprocal and nonreciprocal interactions[27]), and (b) for which an interaction should have been possible to detect via TAP because both bait and prey were identified in other TAP pulldowns. We excluded pairs made between ribosomal proteins and other proteins, as well as pairs in which one partner was annotated as a protein chaperone or protease, since the latter functional categories are expected to form nonspecific complexes with a variety of partners. We observed a total of 146 pairs from our negative gold standards in the data set, which is what would be expected from randomly shuffled data ($152 \pm 37$). In contrast to the gold positive set, we do not observe enrichment in gold negative interactions.

Our mapping of *E. coli* proteins to *D. vulgaris* proteins was done using bidirectional best BLAST searches.[29] All predicted protein sequences encoded by the *D. vulgaris* Hildenborough genome[15] were queried against a database of protein sequences encoded by the *E. coli* K-12 MG1655 genome[30] using BLASTP 2.2.9 with default options, and all *E. coli* proteins were queried against a database of *D. vulgaris* proteins using the same method. Pairs in which each protein was the most significant hit for a query from the other genome, and for which both E-values were at least as significant as $10^{-4}$, were mapped to each other. Even when limiting results to those of high significance, it cannot be stated with certainty that the predicted interactions between these pairs of putative homologues will be conserved; however, many of these pairs are components of well-studied complexes that have been observed in multiple species (Supplementary Table 1, Supporting Information) in addition to *E. coli*.

**Scoring Functions.** For each pair of proteins that was observed to comigrate in BN-PAGE, we calculated a variety of scores that we expected would help distinguish true members of a heteromeric protein complex from false positives. Each function is described below.

*Ranking within SDS-PAGE Bands.* In cases where multiple proteins were observed in a single SDS-PAGE band, we expected that the proteins in which more unique peptides were observed would be more likely to be true members of a complex. We therefore ranked all proteins observed in a single band in numerical order from most peptides observed to least. The rank of a protein within a native gel band is defined as the maximal rank among its SDS-PAGE band rankings, while the rank of a pair of proteins is defined as the minimal value of the two rankings. The final protein pair rank over all native bands where two proteins comigrated is defined as the maximal value among its BN-PAGE rankings. Specifically, the global rank of proteins $i$ and $j$ is defined as $Rank(i,j) = \max_{n \in \text{Native}} (\min( \max_{s \in \text{SDS}_n} (r^i_{n,s}), \max_{s \in \text{SDS}_n} (r^j_{n,s})))$, where $r^i_{n,s}$ is the rank of protein $i$ in native band $n$ and SDS band $s$.

*Comigration of Two Proteins over All Experiments.* This scoring function is defined as the number of times two proteins comigrated in a native band divided by the sum of individual appearances of the two proteins in all the native bands. This number is also referred to as Dice's coefficient. This feature helps to resolve the problem of "frequent fliers"—"sticky" proteins that tend to bind nonspecifically to many other proteins. For sticky proteins, this value is close to zero, while for proteins that form specific interactions the value is higher.

*Probability for Two Proteins to Be Observed in a Native Band.* Given the distribution of proteins in experiment $E$ we approximate the probability that the observation of pair $(i,j)$ in native band $n$ is not due to chance: $p(i, n, E)^* p(j, n, E)$, where $p(i, n, E) = (1 - |n|/|E|)^{i^E}$, $|E|$ is the number of proteins in $E$, $|n|$ is number of proteins in band $n$, and $i^E$ is the number of times $i$


dx.doi.org/10.1021/pr300548d | J. Proteome Res. 2012, 11, 5720−5735

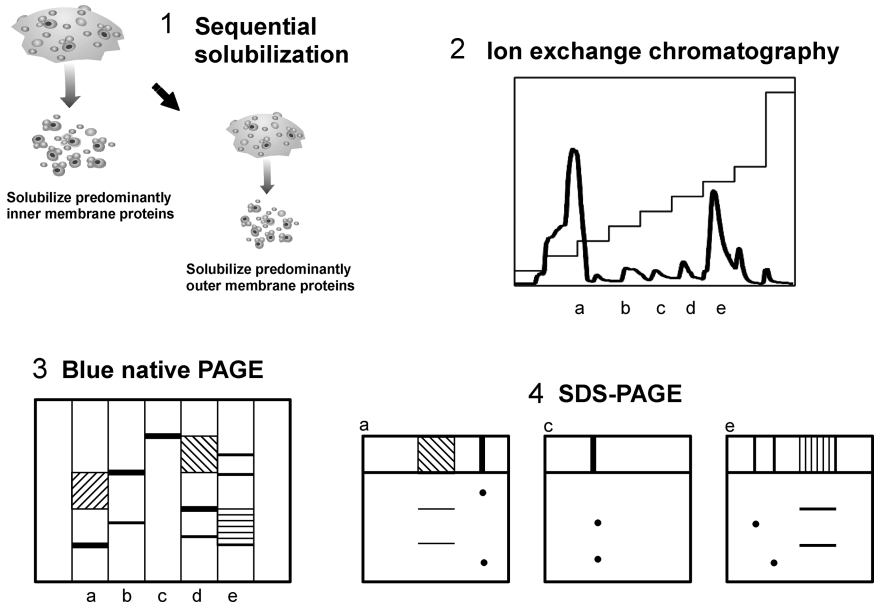



**Figure 1.** Isolation of untagged membrane protein complexes. In the first step of the process, Gram-negative cell membranes are initially treated with a mild detergent to solubilize protein predominantly of the inner membrane. Residual membranes are treated with a second detergent to solubilize protein predominantly of the outer membrane. In step two, solubilized membrane proteins are separated by IEX. Elution peak fractions are subjected to BN-PAGE in the third stage of the process to further separate complexes. Lastly, in step four, lanes are cut from the native gels and placed along the top of a new gel to run a second dimension of SDS-PAGE. In this way, putative complexes are separated into their subunits for subsequent excision and identification by MS.

is observed in $E$. The final score for each pair of proteins $(i,j)$, is the maximum value of $p(i, n, E)^* p(j, n, E)$ over all native bands and experiments. This function aims to address a potential concern not captured by the previous feature. Detergents can differ significantly in their ability to solubilize the full spectrum of membrane proteins. Therefore, in experiments where certain proteins are solubilized to a higher degree or are rendered less stable in solution, the probability of biologically irrelevant protein comigration can increase, in turn, creating an opportunity for false-positive associations. Native gel bands that contained relatively large numbers of proteins also represented cases with a higher probability for producing false-positives. Therefore, proteins detected across many native gel bands were down-weighted, as well as all proteins in very populated bands.

*Mismatch between a Protein's Molecular Weight and That Estimated for a Native Gel Band.* Even though the previous function down-weights instances when many proteins are present in a native band, it still assigns a relatively high score to a protein that appears only once in an experiment. Therefore, in order to further down-weight highly populated bands we take into consideration the native gel band molecular weight estimate and compare it to the sum of estimated molecular weights of the potential subunits found in that band. This weighting function is depicted in Supplementary Figure 1 (Supporting Information). When the total calculated molecular weight exceeds the experimental value by more than 1.25-fold, the function value starts to drop following a Gaussian distribution.

We confirmed that each feature produced statistically significant different distributions when computed for all pairs of proteins in our gold standard positive and negative data sets. According to Mann–Whitney U test the $p$-values for the four scores are 0.00027, 0.0096, 0.0028, and 0.0003. Distributions of gold positive and negative values for these features are depicted in Supplementary Figures 2–5 (Supporting Information).

**Classification of Protein Interactions.** The scores described above were computed for every pair of comigrating proteins seen in a native gel band. Pairs that matched with those in the gold standard set were used to train the Random Forest classifier as implemented in the WEKA package.[31] The classifier was then used to compute a score between zero (not likely to interact) and one (likely to interact) for all comigrating pairs. To estimate the effectiveness of the classifier on our data set we performed a 10-fold cross-validation. Supplementary Figure 6 (Supporting Information) shows a plot of the true positive rate versus false positive rate. For example, we can identify 47% of positive interactions from the gold standard with just 1% of false positives. All scored pairs are reported in Supplementary Table 2 (Supporting Information). Figure 6 shows the interaction network derived from the scored pair set drawn at a threshold of 0.6.

## RESULTS AND DISCUSSION

To stably extract untagged inner and outer membrane protein complexes, retaining native structures while maximizing yield, we used a procedure in which bacterial membranes were sequentially solubilized (Figure 1). Membranes were initially treated with a relatively mild detergent to extract proteins predominantly from the inner membrane. Residual membrane pellets were subsequently solubilized using a more aggressive detergent to extract proteins of the outer membrane. Following solubilization, proteins were chromatographically processed using ion exchange media. Chromatographically separated membrane protein complexes were subjected to blue native polyacrylamide gel electrophoresis (BN-PAGE) to further isolate putative complexes, and obtain estimates of their native molecular weight. In a final step, proteins were extracted from native gel lanes by a second dimension of SDS-PAGE revealing putative complex subunits and their molecular weights, and yielding samples suitable for in-gel digestion and MS analysis. In this manner,


dx.doi.org/10.1021/pr300548d | J. Proteome Res. 2012, 11, 5720–5735

*D. vulgaris* membrane protein complexes of the outer membrane were identified and networks of heteromeric protein−protein interactions involving these and other proteins determined.

## Purification and Identification of Membrane Protein Complexes

Isolation of endogenous *D. vulgaris* membrane protein complexes in quantities sufficient for subunit identification required substantial amounts of cell membranes. The typical yield of *D. vulgaris* cells per liter of cell culture is quite low in comparison to *E. coli*, for example, averaging less than 1 g. To maximize protein extraction, preserve native complexes and avoid the potential losses associated with protocols used to separate inner from outer membranes,[32,33] we elected to process whole *D. vulgaris* membranes directly using a two-step solubilization protocol with detergents matched to lipid bilayer type.[19] During the course of these studies we processed membranes derived from *D. vulgaris* cultures up to 50 L in volume, and typically obtained about 0.1 g of wet membrane per liter of cell culture. We did not observe any significant differences in complexes obtained based on whether membranes were used directly or the proteins were isolated from flash-frozen membranes. This was not entirely unexpected as membrane protein complexes obtain a significant degree of protection from destabilization while in the protective environment of the lipid bilayer.

In this process, cell membranes were first treated with a relatively mild detergent, either $C_{12}E_9$ or Triton X-100, effective in solubilizing proteins predominantly of the inner membrane. Unsolubilized membranes, enriched in proteins of the outer membrane, were separated from the solubilized material by ultracentrifugation. The residual membrane pellet, on average about two-thirds the weight of the starting pellet, was subsequently treated with a second, more aggressive, detergent—either octyl glucoside (OG) or octyl POE (OP). In both solubilization steps, detergent levels were adjusted so that a 1:1 (w/w) ratio of detergent-to-protein was applied; higher levels typically resulted in increased protein instability and ultimately lower yield of suitable proteins. Proteins solubilized from this residual membrane fraction were enriched with those of the outer membrane.

To achieve maximal separation of proteins within these samples a strong anion exchange resin was used for ion exchange chromatography (IEX). Typically, 10−50 mg of solubilized outer membrane proteins were loaded onto a column containing up to 10 mL of media. This was often more media than what would typically be used for processing similar amounts of protein from other sources. Additional resin was needed for *D. vulgaris* samples due to the competition between residual metals (from the culture media) and protein for binding to the column. Proteins were eluted from the column using a step gradient (in 50 mM NaCl increments, from 0 to 400 mM with a final step at 1 M) (Figure 2). The largest elution peak of outer membrane proteins was consistently centered at around 350 mM NaCl. Use of elution increments finer than 50 mM NaCl did not improve the separation of proteins and generally resulted in significant protein overlap between elution peaks. The fractions of each IEX elution peak were of a sufficient protein concentration that they could be used directly for BN-PAGE without the need for a concentration step prior to sample loading. This offered a significant advantage in maintaining protein stability as even incremental changes in protein concentration can increase the probability of aggregation.

Samples of the IEX elution peaks were surveyed by molecular sieve chromatography to assess molecular size distributions that could provide information on potential problems such as detergent effectiveness. As anticipated, it was found that extraction efficiency and chromatographic separation of the membrane proteins of *D. vulgaris* were sensitive to the choice of solubilizing detergent. For example, for the same amount of starting cell membranes, we found the detergent OG to be, on average, more effective in solubilizing proteins of the outer membrane than OP (Figure 3). Additionally, significant differences in the shape of the molecular sieve elution profiles reflecting dissimilarities in the average molecular weights and mobilities of the solubilized proteins were noted. Potential factors leading to these differences include the extent of protein solubilization and delipidation, and the physical and chemical properties of the detergents including micelle size. Based on these results we elected to use OG for the bulk of our experiments involving outer membrane solubilization.

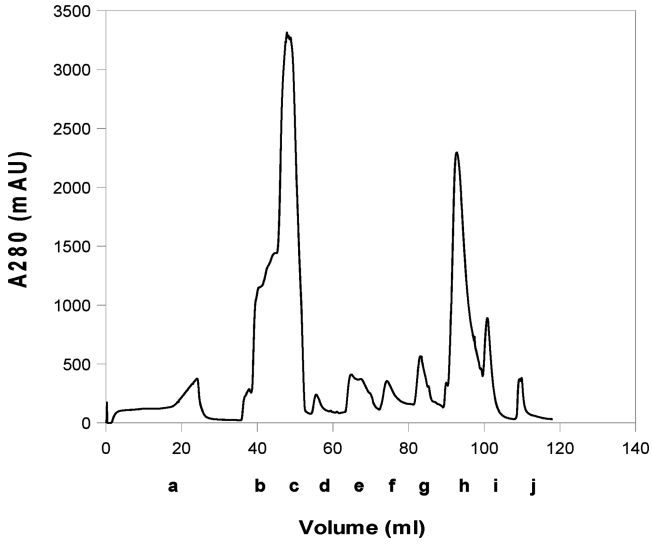



**Figure 2.** Anion exchange chromatography of *D. vulgaris* membrane proteins. Outer membrane proteins solubilized and eluted in the detergent OG. Proteins were eluted from these columns using a NaCl step gradient (0, 50, 100, 150, 200, 250, 300, 350, 400, and 1000 mM NaCl steps). The elution peaks associated with these values are indicated by the letters a−j, respectively.

As the level of outer membrane sample complexity in terms of the different types of constituent proteins is relatively low, about 1/10th that of the inner membrane samples, the application of molecular sieve chromatography to fractions of the IEX elution peaks prior to BN-PAGE did not notably improve separation of outer membrane samples. Molecular sieve chromatography was therefore not included as part of the regular processing of outer membrane preparations.

Outer membrane samples taken directly from IEX elution peak fractions were subjected to BN-PAGE to further separate candidate complexes, obtain native molecular weight estimates and prepare samples for a second dimension of SDS-PAGE. While the native gel process used here generally followed those previously described,[20,34,35] we found that band resolution and sensitivity could be improved by adjusting the Coomassie blue G-250 levels in the sample and cathode buffers to 0.5% and 0.02% respectively, and with the addition of 0.05% dodecyl maltoside to the cathode buffer (Figure 4). Sample volumes were adjusted so as to obtain the maximum signal possible from low abundance proteins while minimizing gel overloading from the most abundant ones. Running the gels slowly at lower voltages appeared to be beneficial for preserving complexes as


dx.doi.org/10.1021/pr300548d | *J. Proteome Res.* 2012, 11, 5720−5735

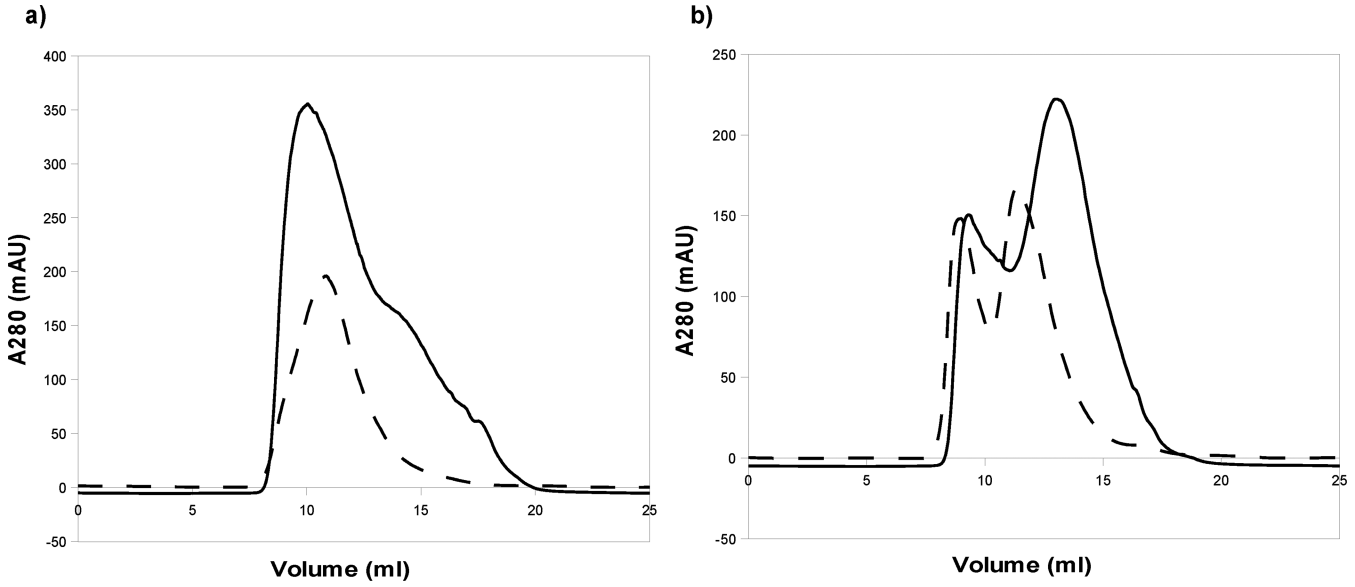



**Figure 3.** Detergent-based differences in *D. vulgaris* outer membrane protein isolation. (a) The continuous curve is a molecular sieve elution profile from an OG solubilized sample that eluted from an IEX column at 250 mM NaCl. The broken-line curve is from a similar experiment but where the membranes were solubilized in OP. (b) The continuous curve is again a molecular sieve elution profile from an OG solubilized sample that in this case eluted from an IEX column at 350 mM NaCl. The broken-line curve is from a similar experiment involving membranes solubilized in OP.

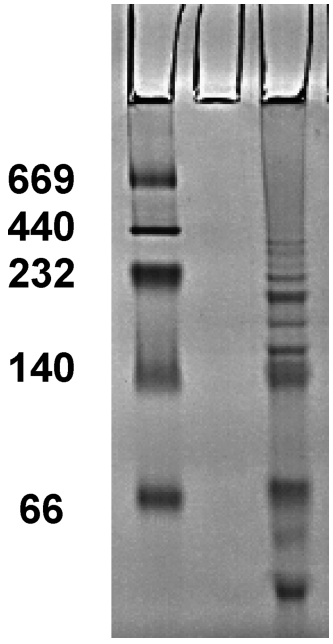



**Figure 4.** BN-PAGE of outer membrane proteins. Sample derived from a *D. vulgaris* culture grown to midlog phase under standard conditions. Lane 1 (from the left), molecular weight standards (669, 440, 232, 140, and 66 kDa); lane 3, protein from an IEX elution peak (250 mM NaCl) of OG solubilized membranes.

they migrated through the gels. Our best results were obtained running the gels overnight at 4 °C using a relatively low voltage (70 V).

Following BN-PAGE, SDS-PAGE of the putative membrane protein complexes sequestered in the native gels was performed. For these gels, lane strips (6.0 × 0.5 cm) were cut from the BN-PAGE gels using a tool especially prepared for this task and placed lengthwise along the top of single-well format SDS-PAGE gels. Laemmli sample buffer was added to the strips while situated in the wells avoiding excessive manipulation of the gels following treatment. Surveying wait times from 2 min to 1 h it was found that incubation with sample buffer for 20 min prior to running a gel was sufficient. Reliable transfer of protein from the native gel strips into the SDS-PAGE gels was accomplished without the need for applying an overlay of agarose around the gel strips due to the quality of the fit between a strip and the SDS-PAGE gel well. In this manner, second dimension SDS-PAGE gels were produced in high-throughput fashion. Gels for this study were prepared with samples isolated from a series of five *D. vulgaris* outer membrane preparations (Supplementary Figure 7, Supporting Information).

Images of the native sample and molecular weight standard lanes were combined with images of the corresponding SDS-PAGE gels to assist in spot processing (see for example, Figure 5). Protein spots were selected for MS analysis, given a control number, and assigned native complex and subunit molecular weight estimates. Spots appearing to originate from a common native gel band, as indicated by migration through the SDS-PAGE gel along the same trajectory, were classified as putative subunits of the same complex. When the spots of protein subunits from more than one complex were in close proximity, assignment of spots to other members of the same complex was aided by comparison of spot profile; spots emanating from the same complex often displayed similar shapes. To complete the characterization, this information was merged with the protein identifications obtained from MS of the processed spots.

Spots excised from the second dimension SDS-PAGE gels were subjected to in-gel digestion and analysis by liquid chromatography electrospray ionization tandem MS (LC−ESI−MS/MS) or liquid chromatography and matrix-assisted laser desorption ionization time-of-flight MS (LC−MALDI TOF MS/MS). The results of the MS protein identifications are provided in Supplementary Tables 3 and 4 (Supporting Information) that contain a complete set of the acquired MS data and a summary of the best results, respectively, for each of the 296 proteins that were identified by this workflow.


dx.doi.org/10.1021/pr300548d | J. Proteome Res. 2012, 11, 5720−5735

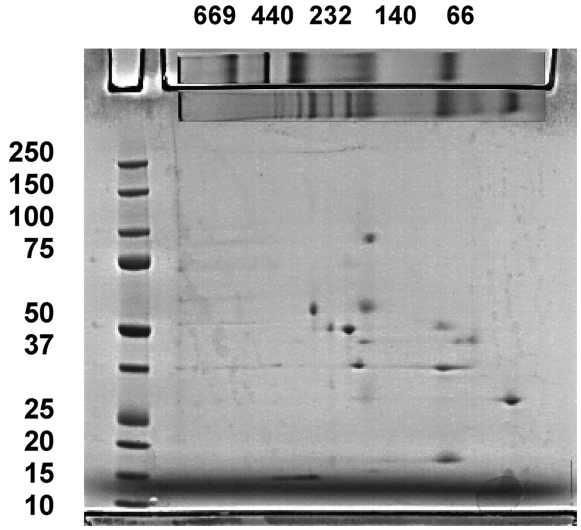



**Figure 5.** Second dimension SDS-PAGE of *D. vulgaris* outer-membrane proteins. The sample of outer membrane proteins used to prepare this gel came from an IEX peak eluting at 250 mM NaCl. Top of gel images—upper image, lane of the native gel molecular weight standards used (669, 440, 232, 140, and 66 kDa); lower image, native gel lane used to prepare the 2D gel. Left side, molecular weight standards (250, 150,100, 75, 50, 37, 25, 20, 15, and 10 kDa).

## Proteins of the *D. vulgaris* Outer Membrane

Processing and analysis of the *D. vulgaris* outer membrane sample sets yielded 296 protein identifications (Supplementary Tables 3 and 4, Supporting Information). This number represents not only the more abundant proteins of the outer membrane but the detection of, in general, lower levels of inner membrane and soluble proteins. The identification of these additional proteins was not surprising. Following the initial mild solubilization step to extract proteins of the inner membrane from the Gram-negative cell envelope, a small percentage of inner membrane proteins can remain with the nonsolubilized outer membrane. These residual inner membrane proteins therefore became part of the protein mixture obtained following solubilization of the outer membrane. Additionally, although cell membranes were washed prior to solubilization, small amounts of soluble proteins were retained with the membranes. Of the 296 proteins identified, 70 were designated as outer membrane-resident or associated. Assignment of these proteins as integral to the outer membrane or strongly associated with it was predominantly based on protein database annotations and the frequency of detection across the outer membrane preparations, with additional consideration given to biophysical properties and sequence analysis (such as the lack of predicted transmembrane helices and homology to known proteins of the outer membrane). Highly stable and abundant membrane protein complexes were identified by inspection of native and corresponding second dimension SDS-PAGE gels. Putative homomeric complexes were indicated by single, and heteromeric complexes by multiple, protein spots emanating from native gel bands of a molecular weight higher than those of the individual subunits. A majority of *D. vulgaris* outer membrane proteins identified (60) were detected as members of homomeric complexes with six proteins forming three heteromeric complexes. Table 1 contains a listing of these 70 proteins and the predominant oligomeric states observed for them.

The most prevalent category of proteins detected in these outer membrane preparations is the lipoproteins,[36,37] comprising over 35% of the proteins listed in Table 1. The fraction of proteins with nondefinitive annotations (e.g., hypothetical and conserved hypothetical) is also relatively large, representing more than 30% of the outer membrane proteins identified here. It should be noted that although these results suggest that these proteins participate in the formation of membrane protein complexes, some (e.g., DVU0266, 0273, 0371, 0851, 1887, 1902 and 2070) are not predicted to have transmembrane helices, $\beta$ barrel structures or lipid anchors, and instead may be membrane-associated through interactions with other proteins. A number of these proteins (DVU0266, 0273, 0371 and 0851) have been proposed to play roles in stress response processes.[38] Other proteins identified in these preparations that are not predicted to be integral membrane proteins (DVU2070, 2579 and 2628), contain tetratricopeptide repeat (TPR) motifs known to foster the formation of heteromeric complexes in organisms ranging from bacteria to humans.[39,40] It is important to stress that while relatively gentle biochemical methods were utilized for the outer membrane sample processing, certain peripheral−integral membrane protein interactions may have been disturbed through the isolation process particularly during BN-PAGE.

## Coverage of the *D. vulgaris* Outer Membrane Proteome

The degree to which the processing pipeline described here has covered the complete range of outer membrane proteins in *D. vulgaris* is difficult to accurately assess. Following a rule-of-thumb estimate that 2−3% of all Gram-negative bacteria protein coding genes are outer membrane associated,[41] approximately 85 outer membrane proteins would be expected for the *D. vulgaris* proteome. As we report detecting 70 putative outer membrane proteins, this metric would suggest that we have observed about 80% of the proteins anticipated. Annotations in the MicrobesOnline database[42] indicate at least 29 outer membrane proteins in *D. vulgaris*, and the PSORTdb[43] database of predictions of subcellular localizations annotates 32 *D. vulgaris* outer membrane proteins. Combining these annotations results in 44 distinct predicted outer membrane proteins, of which we observed 24 (55%). Using a shotgun MS approach, Brockman and colleagues identified over 2000 proteins of the *D. vulgaris* proteome[44] of which up to 68 may be affiliated with the outer membrane. Taken together, these findings would suggest that we have detected at least two-thirds of *D. vulgaris* outer membrane proteins; however the question remains as to whether the inability to detect certain membrane proteins is primarily due to losses incurred during the solubilization, chromatography and electrophoresis steps, naturally low abundance or a combination of these. In an effort to address this question, we performed a shotgun MS survey of proteins present in initial pipeline samples from each of the IEX elution peaks. This analysis utilized two parallel approaches: gel LC−MS, that is, a separation of proteins by SDS-PAGE followed by 1D LC−ESI−MS/MS of in-gel produced tryptic peptides, and 2D (high and low pH reversed phase) LC−ESI−MS/MS of peptides generated by in-solution digestion of proteins within each of the IEX elution peak samples. Of the 548 *D. vulgaris* proteins detected in these samples, 101 were identified as potential outer membrane proteins (Supplementary Table 5, Supporting Information). Based on this upper limit, the 70 identifications made following the processing of samples through the entire pipeline would represent about 69% coverage of detectable proteins. Of note, the great majority of outer membrane proteins detected in this proteomics survey, but not

observed following complete pipeline processing, were identified on the basis of single peptides. The observations of Zhang et al.[44] suggest that more than two-thirds of the proteins in this undetected subset are of low abundance, advancing the idea that the predominant factor in not detecting a given outer membrane protein with this methodology is the level of protein abundance. Coomassie blue staining of proteins in the second dimension SDS-PAGE gels prepared for this study may have been insufficient to enable visual detection of low-abundance proteins during sample preparation. Interestingly, in a side-by-side comparison very few additional spots could be visualized by silver staining these gels and in some cases spots visible on the Coomassie stained gels were not visible in the silver-stained gels. Factors, independent of abundance, also likely to have affected detection and identification of outer membrane proteins are the number of different fragments produced for a given protein, their lengths and hydrophobicities.

### Abundant Proteins of the Outer Membrane

Among the outer membrane proteins identified, three of the most abundant ones are a TolC-like protein (DVU1013) and two currently annotated as "conserved hypothetical" (DVU0797, 0799). The occurrence of a highly abundant TolC-like protein in *D. vulgaris* was not unanticipated as this class of channel-forming protein, a homotrimer of subunits forming one central channel and establishing a conduit between the inner and outer membrane, is essential for efflux processes and is also highly abundant in *E. coli*.[45] Sequence analysis of the two unannotated protein sequences revealed significant sequence similarity with each other (60% identity over the full sequence length) and with an annotated bacterial porin (48% identity to Dde_1011 from *Desulfovibrio desulfuricans* G20). Porins are channel-forming proteins found in the outer membranes of Gram-negative bacteria. As homotrimeric assemblies of large $\beta$ barrel subunits with each of the three subunits forming an independent channel, these complexes facilitate the entry and exit of a broad range of solutes.[46,47] Like the well-characterized bacterial porins of *E. coli*, both of these unannotated proteins are predicted to form homotrimers based on the 150−175 kDa native molecular weight estimate from observations of the complex in BN-PAGE, and the 50−52 kDa subunit molecular weight estimate obtained from SDS-PAGE. Taken together, these observations suggest that these two highly abundant *D. vulgaris* proteins are porins.

### Homologues of *E. coli* Outer Membrane Proteins in *D. vulgaris*

The *D. vulgaris* outer membrane proteins identified cover a broad range of functional categories. Under the category of general import systems, homologues of the *E. coli* FadL (specializing in the uptake of hydrophobic compounds;[48,49] DVU1260, 1548 and 3090) and TonB-dependent (sensing and uptake of specific solutes;[50] DVU0100) proteins have been detected, as well as a number of putative porins (general diffusion channels; DVU0799, 0797, and the lesser abundant DVU0273 and 0371). Proteins identified also include those proposed to be responsible for maintaining the structural integrity of the cell envelope such as the highly abundant OmpA family protein (DVU1422). Interestingly, this protein was observed here in dimeric, trimeric and tetrameric states. Even though annotated as an OmpA family protein, this protein does not contain the outer membrane $\beta$ barrel domain for which OmpA of *E. coli* is known;[51] of the two major domains in each protein only the C-terminal periplasmic domain is shared. In DVU1422 the $\beta$ barrel domain of *E. coli* OmpA has been replaced with a von Willebrand factor domain harboring a putative metal ion binding site. *E. coli* cross-linking studies found evidence of a dimer form of OmpA fostered through peptides of the C-terminal domain.[52] Together with the observation that expressed forms of the OmpA $\beta$ barrel domain alone do not oligomerize, these results suggest that the periplasmic domain plays an important role in the formation of oligomers. Other proteins potentially important in the structural maintenance of the cell include two proteins (DVU3104, 2070) associating to form a complex corresponding to one from the *E. coli* Tol-Pal system.[53] Proteins from this system have been found to participate in the network of interactions coupling regions of the outer membrane to peptidoglycan. The putative function of the complex found here was not at first apparent given the current *D. vulgaris* annotations. While the peptidoglycan-associated lipoprotein (Pal) homologue, DVU3104, was often detected as a homo-oligomer in our outer membrane preparations, it was also seen in a heteromeric complex with the protein DVU2070, which is annotated as containing a TPR domain. Sequence analysis revealed that DVU2070 has significant sequence similarity (31% identity, BLAST E = $10^{-18}$) with the *E. coli* protein YbgF. Notably, although both components of this complex share significant similarity with their counterparts in *E. coli*, the genes encoding the proteins of the *D. vulgaris* complex come from different operons while YbgF is encoded along with the pal and tol genes in a single operon in *E. coli*. A range of membrane proteins associated with *D. vulgaris* efflux systems was also observed. As expected, the largest number of such proteins were found participating in type I secretion systems (channel-forming structures spanning the outer membrane and periplasmic space); these included TolC-like (DVU1013, 2815, and 3097) and NodT-like (DVU0062) components. A member of a predicted type II secretion system, secretin DVU1273, was found as well. Interestingly, a member of a putative type III secretion system (DVUA0117), whose expression is expected to be subject to stress-based modulation, was detected in late exponential phase cultures grown under standard conditions. Also observed in the general identifications (but not within the set of confidently predicted heteromeric interactions) were proteins putatively associated with the transport of molecules destined for the outer membrane. Two *D. vulgaris* proteins (DVU2373 and DVU1837), homologues of subunits BamA and BamD, respectively, of the *E. coli* Bam outer membrane protein assembly complex[54] were detected but not as part of a complex. As with the potential Tol-Pal complex subunit (2070), the annotation of the potential Bam-like complex component DVU1837 did not suggest this as a potential function, being described only as a competence protein. Its homology to BamD of *E. coli* (73% coverage, 27% identity, BLAST E = $2 \times 10^{-19}$), however, suggests the possibility of a similar functional role in *D. vulgaris* although the nature of the interaction between the BamA and D-like proteins may be different from that in *E. coli*. With respect to another important transport system of *E. coli*, the Lpt lipopolysaccharide transport complex,[55,56] we did detect a homologue of the outer membrane based subunit LptD (DVU0954; currently annotated as an organic solvent tolerance protein); however, in contrast to the system in *E. coli*, no binding partners were observed. This is consistent with the observation that there is no apparent homologue of LptE (the outer membrane partner of LptD in the *E. coli* complex) in *D. vulgaris*.

### Periplasmic Hydrogenases

A particularly interesting finding in the group of identified proteins is the consistent observation of the periplasmic [NiFeSe]


dx.doi.org/10.1021/pr300548d | J. Proteome Res. 2012, 11, 5720−5735

(DVU1917, 1918), and to a somewhat lesser extent the [NiFe] (DVU1921, 1922), hydrogenase complexes in these preparations. These heteromeric complexes aid both sulfate reduction and energy production by catalyzing the breakdown of molecular hydrogen into protons and electrons.[57] We found these two complexes to be strongly associated with the outer membrane preparations. The potential membrane association of these proteins was not unanticipated as the lipobox sequence of [NiFeSe] hydrogenase has been shown, and that of an [NiFe] hydrogenase isoform projected, to facilitate acylation of their N-terminal cysteines.[58] Attachment of such hydrophobic groups would make these proteins lipoproteins and enable them to peripherally associate with cell membranes. Absent from the list of proteins identified is the [Fe] hydrogenase, which was not detected in these preparations. This observation is in agreement with research that found the expression of these hydrogenases to be tightly coupled to the available carbon source and hydrogen levels.[59] As the cell cultures processed for the work presented here were grown under standard conditions on defined LS4D media with mineral supplements (including nickel, iron and selenium salts), sodium lactate as a carbon source and the maintenance of low hydrogen levels, the relatively strong presence of the [NiFeSe] and undetectable level of [Fe] hydrogenase is as expected.

## Interaction Network Derived from *D. vulgaris* Outer Membrane Preparations

For the identification of additional and potentially weaker interactions, statistical and machine learning-based analysis of protein spots observed across multiple experiments was performed, as described in detail in the Methods section. Our models were trained on gold standard sets based primarily on well-studied protein complexes from *E. coli*, and integrate a variety of features derived from our MS data and experimental metadata (e.g., sizing standards). We identified interacting pairs based on a conservative interpretation of our predictions: in order to minimize false positives, we used a cutoff that corresponded to a 53% false negative rate on our training set, meaning that we also expect to miss a significant number of true interactions in our *D. vulgaris* data. In addition, since the threshold chosen corresponded to a 1% false positive rate on our training data, a small subset of our predicted interactions are expected to result from biologically irrelevant products of the isolation process; several cases are discussed below.

Among all proteins identified in these preparations, we classified 104 of these as participating in heteromeric protein−protein interactions. This set of interactions is not limited to those proteins whose principle association is with the outer membrane and, as mentioned earlier, also includes proteins classified as periplasmic, extracellular, inner membrane and cytoplasmic. A significant number of the complexes identified in these preparations do not involve outer membrane proteins. The network of interactions among these proteins is depicted in Figure 6; a listing of these proteins with their functional annotations is provided in Table 2. It should be noted that monomers and homomeric complexes, although detected in the outer membrane, are not included in Figure 6 unless they were also found to participate in heteromeric interactions. The relative confidence in these interactions is indicated by the width of the lines connecting the protein nodes; the wider the line, the more likely the interaction is to be correct. A subset of the clusters appearing in this interaction network will be discussed.

Within this network of interactions, several heteromeric complexes of the *D. vulgaris* outer membrane, described in the previous sections, were identified as anticipated − a Tol-Pal (DVU3104, 2070) complex (Figure 6A), and the [NiFeSe] (DVU1917, 1918) and [NiFe] (DVU1921, 1922) hydrogenases (Figure 6B and C, respectively); these are located in the upper left region of Figure 6. The Tol-Pal complex homologue, is found to interact with an OmpA family protein (DVU1422) (Figure 6D). The two hydrogenases are found in differing interaction patterns—the [NiFeSe] hydrogenase forms contacts with two TolC-like proteins (DVU0062, 1013) (Figure 6E) and two putative porins as described earlier (DVU0797, 0799) (Figure 6M); the [NiFe] complex's sole observed interaction is via the putative inner membrane protein subunit of formate dehydrogenase (DVU2481). For the relatively abundant OmpA family (DVU1422) and P1 (DVU1260) outer membrane proteins, a few interactions with several outer membrane and periplasmic proteins (DVU1013, 2070, 1582) as well as between themselves were identified, albeit with lower confidence. Interestingly, the most abundant TolC family protein (DVU1013) displayed high-confidence interactions with two other outer membrane efflux proteins (DVU0062 and 3097) (Figure 6E). To the right of these clusters in the figure is a network formed around a protein annotated as hypothetical (DVU0266) (Figure 6O). This putative periplasmic protein participates in high-confidence interactions with ferredoxin (DVU0264), a membrane-based comember of its predicted operon, enolase (DVU0322) an enzyme involved in carbohydrate degradation, and a ribosomal protein (DVU3150). Interaction of enolase with DVU0266 may be associated with the movement of this protein to the cell surface, while a functional purpose for the binding of a ribosomal protein to DVU0266 is not clear.

In the middle region of Figure 6 are displayed interactions between the subunits of ATP synthase (DVU0774 - 0780) (Figure 6J) and a number of putative outer membrane proteins (such as DVU1648, 3125, and 1537) interacting with the Sec translocase components SecDFG (DVU1819, 1818 and 1676) and YajC (DVU1820) (Figure 6L). Given the relatively sparse annotation of these putative outer membrane proteins it is difficult to predict with confidence the purpose of these interactions. One possibility is that a number of them may be proteins in the process of transport via the Sec translocase. Others may be providing support to proteins in the process of transiting the periplasmic space for the outer membrane. Nearby, the two members of the HflKC complex (DVU0683, 0684), an inner membrane-based inhibitor of FtsH proteolytic activity in *E. coli*,[60] can be seen to be strongly linked (Figure 6N). Although the HflKC complex observed here is homologous to that found in *E. coli*, we did not detect forms of HflKC participating in an extended complex with the zinc metalloprotease FtsH. Instead, we identified a strong interaction between complex subunit HflC and a protein proposed to be a member of the Band 7 family (DVU3355). All three members of this expanded complex share homology with this membrane protein family where the representative member, Band 7, is thought to regulate cation conductance. Also in this general region, two of the quinone-interacting membrane-bound oxidoreductase (Qmo) subunits (DVU0848, 0849) are found strongly linked (Figure 6 K). The Qmo complex is an essential component of the *D. Vulgaris* electron pathway for sulfate reduction.[61] Interacting with the Qmo complex through the QmoA subunit (DVU0848) is a protein annotated as hypothetical (DVU0851) that, like

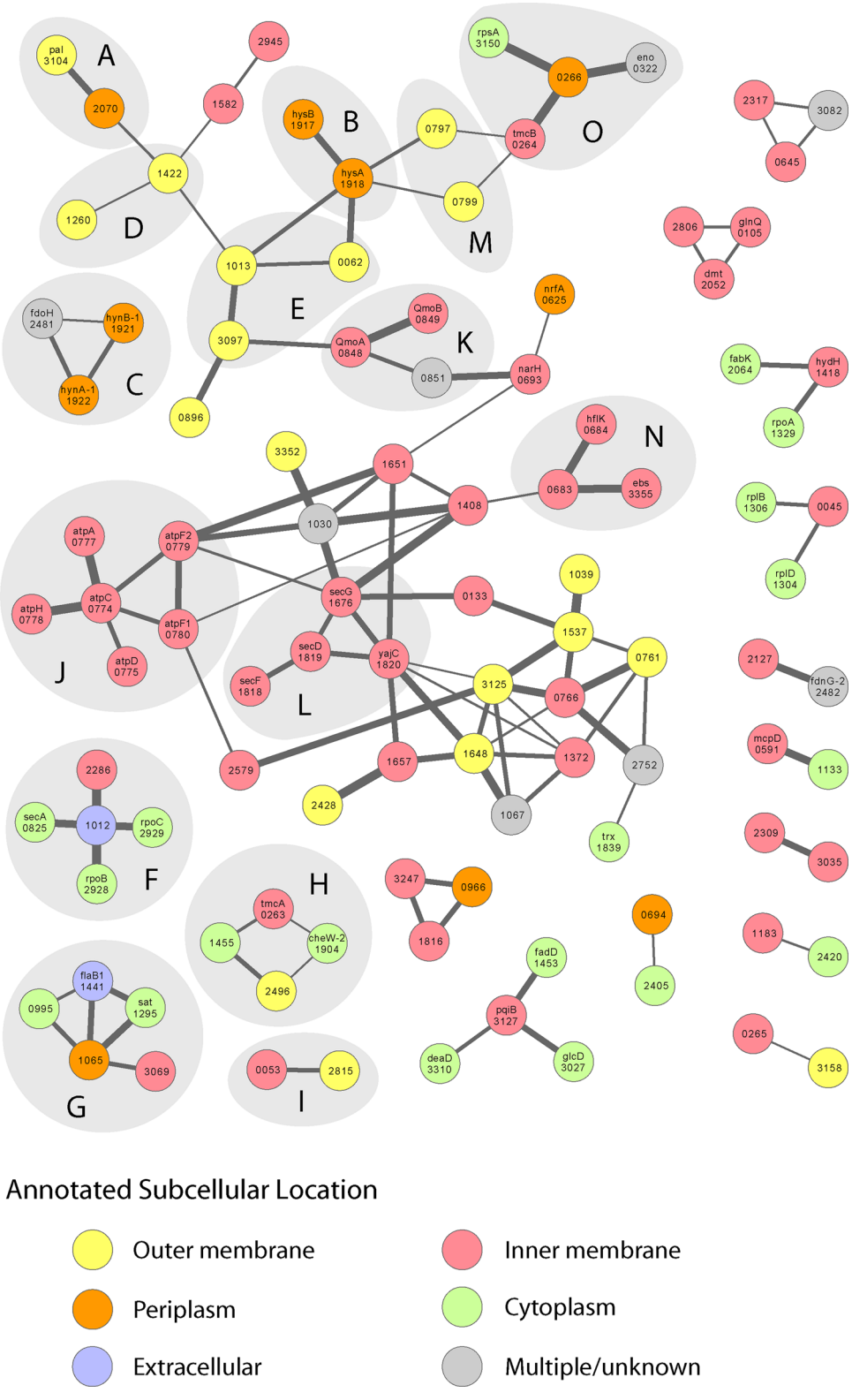



**Figure 6.** Interaction network of proteins detected in *D. vulgaris* outer membrane preparations. In addition to proteins of the outer membrane, the network of interactions detected involves putative inner membrane, periplasmic, cytoplasmic, and even extracellular proteins. Each node represents an observed protein; the gene number for each protein is situated upon its corresponding node. The color of the node indicates the putative cellular location (see legend). The lines connecting the nodes vary in thickness according to the likelihood of their interactions. Thicker lines indicate a relatively greater likelihood of interaction between proteins. Proteins detected solely as monomers or homomeric oligomers without evidence of heteromeric protein−protein interactions, although identified in the outer membrane preparations, are not depicted here.

subunit A, is a protein without a predicted transmembrane domain. Based on gene deletion studies, DVU0851 was found to be nonessential for sulfate reduction.[61] Its deletion however was observed to delay the onset of activity somewhat, suggesting a


dx.doi.org/10.1021/pr300548d | J. Proteome Res. 2012, 11, 5720−5735

**Table 2. *D. vulgaris* Proteins of the Interaction Network Shown in Figure 6**

| Gene ID | annotation[a] | Gene ID | annotation[a] |
|---|---|---|---|
| DVU0045 | flagellar biosynthesis protein, FliO, putative | DVU1441 | flaB1, flagellin |
| DVU0053 | sulfate permease, putative | DVU1453 | fadD, long-chain-fatty-acid–CoA ligase |
| DVU0062 | RND efflux system, outer membrane protein,NodT family | DVU1455 | conserved hypothetical protein |
| DVU0105 | glnQ, glutamine ABC transporter, ATP-binding protein | DVU1537 | lipoprotein, putative |
| DVU0133 | hypothetical protein | DVU1582 | hypothetical protein |
| DVU0263 | tmcA, Transmembrane complex, tetraheme cytochrome c3 | DVU1648 | lipoprotein, putative |
| DVU0264 | tmcB, Transmembrane complex, ferredoxin, 2 [4Fe-4S] | DVU1651 | conserved hypothetical protein |
| DVU0265 | membrane protein, putative | DVU1657 | hypothetical protein |
| DVU0266 | hypothetical protein | DVU1676 | secG, preprotein translocase, SecG subunit |
| DVU0322 | eno, enolase | DVU1816 | conserved hypothetical protein |
| DVU0591 | mcpD, methyl-accepting chemotaxis protein | DVU1818 | secF, protein-export membrane protein SecF |
| DVU0625 | nrfA, cytochrome c nitrite reductase, catalytic subunit NrfA, putative | DVU1819 | secD, protein-export membrane protein SecD |
| DVU0645 | methyl-accepting chemotaxis protein | DVU1820 | yajC, preprotein translocase, YajC subunit |
| DVU0683 | hflC protein, putative | DVU1839 | trx, thioredoxin |
| DVU0684 | hflK, hflK protein, putative | DVU1904 | cheW-2, chemotaxis protein CheW |
| DVU0693 | narH, molybdopterin oxidoreductase, iron−sulfur cluster-binding subunit, putative (TIGR), containing cytochrome c heme-binding site | DVU1917 | hysB, periplasmic [NiFeSe] hydrogenase, small subunit |
| DVU0694 | molybdopterin oxidoreductase, molybdopterin-binding subunit, putative | DVU1918 | hysA, periplasmic [NiFeSe] hydrogenase, large subunit, selenocysteine-containing |
| DVU0761 | lipoprotein, putative | DVU1921 | hynB-1, periplasmic [NiFe] hydrogenase, small subunit, isozyme 1 |
| DVU0766 | transporter, putative | DVU1922 | hynA-1, periplasmic [NiFe] hydrogenase, large subunit, isozyme 1 |
| DVU0774 | atpC, ATP synthase, F1 epsilon subunit | DVU2052 | dmt, glycosyl transferase, group 2 family protein |
| DVU0775 | atpD, ATP synthase, F1 beta subunit | DVU2064 | fabK, oxidoreductase, 2-nitropropane dioxygenase family |
| DVU0777 | atpA, ATP synthase, F1 alpha subunit | DVU2070 | TPR domain protein |
| DVU0778 | atpH, ATP synthase, F1 delta subunit | DVU2127 | von Willebrand factor type A domain protein |
| DVU0779 | atpF2, ATP synthase F0, B subunit, putative | DVU2286 | hydrogenase, CooM subunit, putative |
| DVU0780 | atpF1, ATP synthase F0, B subunit, putative | DVU2309 | methyl-accepting chemotaxis protein, putative |
| DVU0797 | conserved hypothetical protein | DVU2317 | methyl-accepting chemotaxis protein, putative |
| DVU0799 | conserved hypothetical protein | DVU2405 | alcohol dehydrogenase, iron-containing |
| DVU0825 | secA, preprotein translocase, SecA subunit | DVU2420 | conserved hypothetical protein |
| DVU0848 | QmoA, Quinone-interacting membrane-bound oxidoreductase | DVU2428 | lipoprotein, putative |
| DVU0849 | QmoB, Quinone-interacting membrane-bound oxidoreductase | DVU2481 | fdoH, formate dehydrogenase, beta subunit, putative |
| DVU0851 | hypothetical protein | DVU2482 | fdnG-2, formate dehydrogenase, alpha subunit, selenocysteine-containing |
| DVU0896 | lipoprotein, NLP/P60 family | DVU2496 | lipoprotein, putative |
| DVU0966 | amino acid ABC transporter, periplasmic amino acid-binding protein | DVU2579 | TPR domain protein |
| DVU0995 | ThiJ/PfpI family protein | DVU2752 | rhodanese-like domain protein |
| DVU1012 | hemolysin-type calcium-binding repeat protein | DVU2806 | MotA/TolQ/ExbB proton channel family protein |
| DVU1013 | type I secretion outer membrane protein, TolC family | DVU2815 | outer membrane efflux protein |
| DVU1030 | universal stress protein family | DVU2928 | rpoB, DNA-directed RNA polymerase, beta subunit |
| DVU1039 | lipoprotein, putative | DVU2929 | rpoC, DNA-directed RNA polymerase, beta prime subunit |
| DVU1065 | peptidyl-prolyl cis−trans isomerse domain protein | DVU2945 | conserved domain protein |
| DVU1067 | membrane protein, Bmp family | DVU3027 | glcD, glycolate oxidase, subunit GlcD |
| DVU1133 | hypothetical protein | DVU3035 | methyl-accepting chemotaxis protein, putative |
| DVU1183 | HD domain protein | DVU3069 | conserved hypothetical protein TIGR00247 |
| DVU1260 | outer membrane protein P1, putative | DVU3082 | methyl-accepting chemotaxis protein |
| DVU1295 | sat, sulfate adenylyltransferase | DVU3097 | outer membrane efflux protein |
| DVU1304 | rplD, ribosomal protein L4 | DVU3104 | pal, peptidoglycan-associated lipoprotein, putative |
| DVU1306 | rplB, ribosomal protein L2 | DVU3125 | lipoprotein, putative |
| DVU1329 | rpoA, DNA-directed RNA polymerase, alpha subunit | DVU3127 | pqiB, paraquat-inducible protein B |
| DVU1372 | membrane protein, putative | DVU3150 | rpsA, ribosomal protein S1 |
| DVU1408 | hypothetical protein | DVU3158 | vacJhomolog, vacJ lipoprotein, putative |
| DVU1418 | hydH, sensory box histidine kinase | DVU3247 | efflux transporter, RND family, MFP subunit |
| DVU1422 | OmpA family protein | DVU3310 | deaD, ATP-dependent RNA helicase, DEAD/DEAH family |
| | | DVU3352 | lipoprotein, putative |
| | | DVU3355 | ebs, SPFH domain/Band 7 family protein |

[a]Annotations listed as provided in MicrobesOnline (www.microbesonline.com).


dx.doi.org/10.1021/pr300548d | J. Proteome Res. 2012, 11, 5720−5735

role in Qmo optimization. Strongly coupled to this protein is the iron−sulfur cluster-binding subunit of molybdopterin oxidoreductase (DVU0693). Notably, a number of the proposed inner membrane subunits of these complexes were not detected in these preparations (e.g., QmoC, $AtpF_0$ a, c and the Sec translocase transmembrane pore forming subunits SecE and SecY). Such an outcome was not unanticipated as the biochemical process employed in the present study was optimized specifically to maximize outer membrane protein extraction yield and maintain the stability of solubilized outer membrane protein complexes. As indicated earlier, inner membrane protein complexes are typically more sensitive to the choice of detergent and solubilization protocol; selection of an incompatible detergent can lead to complex destabilization, subunit denaturation and aggregation, and protein precipitation. An equivalent study on the inner membrane protein complexes of *D. vulgaris* will therefore be best approached selecting detergents and protocols optimized for this class of protein.

A number of relatively small networks can be seen in the lower portion of Figure 6. One of the more interesting of these involves the interaction of protein translocase subunit SecA (DVU0825) (Figure 6F) with a hemolysin-type calcium-binding repeat protein (DVU1012), a 316 kDa extracellular protein likely observed in a stage of the secretion process. Surprisingly, this interaction cluster also involves two RNA polymerase subunits (DVU2928, 2929) and the M subunit of the inner membrane protein Coo hydrogenase (DVU2286), all binding to the hemolysin-type calcium-binding repeat protein; the nature of the interaction pattern however does not suggest that these interactions take place simultaneously. Other clusters of interest involve a peptidyl-prolyl cis−trans isomerase domain protein (DVU1065), a putative periplasm-based chaperone, interacting with flagellin (DVU1441) (Figure 6G). While not defining a specialized activity for this protein, association of a chaperone with flagellin may be indicative of a support role in flagellar assembly. The additional strongly predicted interactions in this cluster are made by sulfate adenyltransferase (DVU1295) to both flagellin and the peptidyl-prolyl cis−trans isomerase domain protein, an interaction for which a functional role is not readily apparent. An additional possibility is that this may be an example of a biologically irrelevant interaction produced through the isolation process, particularly given the large number of proteins originating from several cellular compartments and preparation through a generalized process. In a different cluster involving a putative outer membrane protein, a lipoprotein (DVU2496) is seen to interact with an uncharacterized protein (DVU1455) that in turn interacts with the tetraheme cytochrome c3 (DVU0263) (Figure 6H). An additional low-confidence interaction with the chemotaxis protein CheW (DVU1904) is suggested to take place between the cytochrome and lipoprotein. However, this could again be reflective of a nonbiological interaction as CheW is expected to reside in the cytoplasm. Lastly, an outer membrane efflux protein (DVU2815) is seen here interacting with a putative sulfate permease (DVU0053) (Figure 6I). An inner membrane-anchored fusion protein, coupling transporters to efflux channels in type I secretion systems, was not detected in the complex possibly due to a low presence in the outer membrane preparations. This initial detection of interactions has likely identified components of an as of yet uncharacterized *D. vulgaris* sulfate transport complex that may play an important role in the function of this sulfate-reducing bacterium.

## CONCLUSION

The results of this outer membrane proteome study demonstrate that mild isolation and purification procedures can, even under high-throughput circumstances, produce a range of *D. vulgaris* membrane protein complexes suitable for identification and characterization. Not only do these results compare favorably to predictions and earlier studies pertaining to the *D. vulgaris* outer membrane proteome, but they also indicate significantly broader coverage of expected complexes in comparison to studies of the more thoroughly characterized *E. coli*. Based on the success of these efforts, work on the substantially larger arrays of proteins generated through the processing of *D. vulgaris* inner membrane preparations will begin. Processing of outer and inner membrane protein complexes can be repeated for cultures of *D. vulgaris* grown to stationary phase and biofilms, as well as in the presence of environmental stressors such as elevated nitrate levels. Through these efforts the data needed to characterize stress-induced changes, such as those pertaining to relative abundance and protein−protein interactions, can be assembled. These data, in turn, will facilitate the modeling of stress response pathways. The results presented here suggest that this processing pipeline should be an effective tool for the high-throughput isolation and identification of membrane protein complexes, supporting the global characterization of membrane proteins in a wide range of organisms.

## ASSOCIATED CONTENT

### Supporting Information

Supplementary Methods, Figures 1−8 and Tables 1−5. This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

### Corresponding Author

*Phone: 510-486-7469. Fax: 510-486-6488. E-mail: pjwalian@lbl.gov.

### Present Address

[‖]Therabiol Inc., San Francisco, CA 94107

### Author Contributions

[‡]These authors contributed equally to this work.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work conducted by ENIGMA - Ecosystems and Networks Integrated with Genes and Molecular Assemblies (http://enigma.lbl.gov), a Scientific Focus Area Program at Lawrence Berkeley National Laboratory, was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Mass spectrometry analysis was performed by the UCSF Sandler-Moore Mass Spectrometry Core Facility, which acknowledges support from the Sandler Family Foundation, the Gordon and Betty Moore Foundation, and NIH/NCI Cancer Center Support Grant P30 CA082103. We thank Chansu Park, Gerald Stampfel and Swan Lin for technical assistance. This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government

## ■ ABBREVIATIONS

MS, mass spectrometry; BN-PAGE, blue-native polyacrylamide gel electrophoresis; OG, octyl glucoside; OP, octyl POE; DDM, dodecyl maltoside; IEX, ion exchange chromatography; CMC, critical micelle concentration; TAP, tandem affinity purification.

## ■ REFERENCES

(1) Weiner, J. H.; Li, L. Proteome of the *Escherichia coli* envelope and technological challenges in membrane proteome analysis. *Biochim. Biophys. Acta* **2008**, *1778* (9), 1698−1713.

(2) Speers, A. E.; Wu, C. C. Proteomics of integral membrane proteins − theory and application. *Chem. Rev.* **2007**, *107* (8), 3687−3714.

(3) Tan, S.; Tan, H. T.; Chung, M. C. M. Membrane proteins and membrane proteomics. *Proteomics* **2008**, *8* (19), 3924−3932.

(4) Walian, P.; Cross, T. A.; Jap, B. K. Structural genomics of membrane proteins. *Genome Biol.* **2004**, *5* (4), 215.

(5) Marzoa, J.; Sánchez, S.; Ferreirós, C. M.; Criado, M. T. Identification of *Neisseria meningitidis* outer membrane vesicle complexes using 2-D high resolution clear native/SDS-PAGE. *J. Proteome Res.* **2010**, *9* (1), 611−619.

(6) Pan, J.-Y.; Li, H.; Ma, Y.; Chen, P.; Zhao, P.; Wang, S.-Y.; Peng, X.-X. Complexome of *Escherichia coli* envelope proteins under normal physiological conditions. *J. Proteome Res.* **2010**, *9* (7), 3730−3740.

(7) Yang, X.; Promnares, K.; Qin, J.; He, M.; Shroder, D. Y.; Kariu, T.; Wang, Y.; Pal, U. Characterization of multiprotein complexes of the *Borrelia burgdorferi* outer membrane vesicles. *J. Proteome Res.* **2011**, *10* (10), 4556−4566.

(8) Zheng, J.; Wei, C.; Zhao, L.; Liu, L.; Leng, W.; Li, W.; Jin, Q. Combining blue native polyacrylamide gel electrophoresis with liquid chromatography tandem mass spectrometry as an effective strategy for analyzing potential membrane protein complexes of *Mycobacterium bovis* bacillus Calmette-Guérin. *BMC Genomics* **2011**, *12*, 40.

(9) Bernarde, C.; Lehours, P.; Lasserre, J.-P.; Castroviejo, M.; Bonneu, M.; Mégraud, F.; Ménard, A. Complexomics study of two *Helicobacter pylori* strains of two pathological origins. *Mol. Cell. Proteomics* **2010**, *9* (12), 2796−2826.

(10) Maddalo, G.; Stenberg-Bruzell, F.; Goetzke, H.; Toddo, S.; Bjorkholm, P.; Eriksson, H.; Chovanec, P.; Genevaux, P.; Lehtio, J.; Ilag, L. L.; Daley, D. O. Systematic analysis of native membrane protein complexes in *Escherichia coli*. *J. Proteome Res.* **2011**, *10* (4), 1848−1859.

(11) Michel, C.; Brugna, M.; Aubert, C.; Bernadac, A.; Bruschi, M. Enzymatic reduction of chromate: comparative studies using sulfate-reducing bacteria. *Appl. Microbiol. Biotechnol.* **2001**, *55* (1), 95−100.

(12) Payne, R. B.; Gentry, D. M.; Rapp-Giles, B. J.; Casalot, L.; Wall, J. D. Uranium reduction by *Desulfovibrio desulfuricans* Strain G20 and a Cytochrome c3 mutant. *Appl. Environ. Microbiol.* **2002**, *68* (6), 3129−3132.

(13) Muyzer, G.; Stams, A. J. M. The ecology and biotechnology of sulphate-reducing bacteria. *Nat. Rev. Microbiol.* **2008**, *6* (6), 441−454.

(14) Wall, J. D.; Krumholz, L. R. Uranium reduction. *Annu. Rev. Microbiol.* **2006**, *60*, 149−166.

(15) Heidelberg, J. F.; Seshadri, R.; Haveman, S. A.; Hemme, C. L.; Paulsen, I. T.; Kolonay, J. F.; Eisen, J. A.; Ward, N.; Methe, B.; Brinkac, L. M.; Daugherty, S. C.; Deboy, R. T.; Dodson, R. J.; Durkin, A. S.; Madupu, R.; Nelson, W. C.; Sullivan, S. A.; Fouts, D.; Haft, D. H.; Selengut, J.; Peterson, J. D.; Davidsen, T. M.; Zafar, N.; Zhou, L.; Radune, D.; Dimitrov, G.; Hance, M.; Tran, K.; Khouri, H.; Gill, J.; Utterback, T. R.; Feldblyum, T. V.; Wall, J. D.; Voordouw, G.; Fraser, C. M. The genome sequence of the anaerobic sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough. *Nat. Biotechnol.* **2004**, *22* (5), 554−559.

(16) Price, M. N.; Deutschbauer, A. M.; Kuehl, J. V.; Liu, H.; Witkowska, H. E.; Arkin, A. P. Evidence-based annotation of transcripts and proteins in the sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough. *J. Bacteriol.* **2011**, *193* (20), 5716−5727.

(17) Redding, A. M.; Mukhopadhyay, A.; Joyner, D. C.; Hazen, T. C.; Keasling, J. D. Study of nitrate stress in *Desulfovibrio vulgaris* Hildenborough using iTRAQ proteomics. *Brief. Funct. Genomics Proteomics* **2006**, *5* (2), 133−143.

(18) Brandis, A.; Thauer, R. K. Growth of *Desulfovibrio* species on hydrogen and sulfate as sole energy source. *J. Gen. Microbiol.* **1981**, *126* (1), 249−252.

(19) Baldermann, C.; Lupas, A.; Lubieniecki, J.; Engelhardt, H. The regulated outer membrane protein Omp21 from *Comamonas acidovorans* is identified as a member of a new family of eight-stranded beta-sheet proteins by its sequence and properties. *J. Bacteriol.* **1998**, *180* (15), 3741−3749.

(20) Schägger, H.; Cramer, W. A.; von Jagow, G. Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal. Biochem.* **1994**, *217* (2), 220−230.

(21) Schilling, B.; Murray, J.; Yoo, C. B.; Row, R. H.; Cusack, M. P.; Capaldi, R. A.; Gibson, B. W. Proteomic analysis of succinate dehydrogenase and ubiquinol-cytochrome c reductase (Complex II and III) isolated by immunoprecipitation from bovine and mouse heart mitochondria. *Biochim. Biophys. Acta* **2006**, *1762* (2), 213−222.

(22) Shilov, I. V.; Seymour, S. L.; Patel, A. A.; Loboda, A.; Tang, W. H.; Keating, S. P.; Hunter, C. L.; Nuwaysir, L. M.; Schaeffer, D. A. The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol. Cell. Proteomics* **2007**, *6* (9), 1638−1655.

(23) Keshishian, H.; Addona, T.; Burgess, M.; Kuhn, E.; Carr, S. A. Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Mol. Cell. Proteomics* **2007**, *6* (12), 2212−2229.

(24) Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. Orthogonality of separation in two-dimensional liquid chromatography. *Anal. Chem.* **2005**, *77* (19), 6426−6434.

(25) Dowell, J. A.; Frost, D. C.; Zhang, J.; Li, L. Comparison of two-dimensional fractionation techniques for shotgun proteomics. *Anal. Chem.* **2008**, *80* (17), 6715−6723.

(26) Karp, P. D.; Riley, M.; Saier, M.; Paulsen, I. T.; Collado-Vides, J.; Paley, S. M.; Pellegrini-Toole, A.; Bonavides, C.; Gama-Castro, S. The EcoCyc database. *Nucleic Acids Res.* **2002**, *30* (1), 56−58.

(27) Hu, P.; Janga, S. C.; Babu, M.; Diaz-Mejia, J. J.; Butland, G.; Yang, W.; Pogoutse, O.; Guo, X.; Phanse, S.; Wong, P.; Chandran, S.; Christopoulos, C.; Nazarians-Armavil, A.; Nasseri, N. K.; Musso, G.; Ali, M.; Nazemof, N.; Eroukova, V.; Golshani, A.; Paccanaro, A.; Greenblatt, J. F.; Moreno-Hagelsieb, G.; Emili, A. Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins. *PLoS Biol.* **2009**, *7* (4), e1000096.

(28) Chhabra, S. R.; Joachimiak, M. P.; Petzold, C. J.; Zane, G. M.; Price, M. N.; Reveco, S. A.; Fok, V.; Johanson, A. R.; Batth, T. S.; Singer, M.; Chandonia, J.-M.; Joyner, D.; Hazen, T. C.; Arkin, A. P.; Wall, J. D.; Singh, A. K.; Keasling, J. D. Towards a rigorous network of

protein-protein interactions of the model sulfate reducer *Desulfovibrio vulgaris* Hildenborough. *PLoS One* **2011**, *6* (6), e21470.
(29) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215* (3), 403−410.
(30) Blattner, F. R.; Plunkett, G., III; Bloch, C. A.; Perna, N. T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J. D.; Rode, C. K.; Mayhew, G. F.; Gregor, J.; Davis, N. W.; Kirkpatrick, H. A.; Goeden, M. A.; Rose, D. J.; Mau, B.; Shao, Y. The complete genome sequence of *Escherichia coli* K-12. *Science* **1997**, *277* (5332), 1453−1474.
(31) Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I. H. The WEKA data mining software: an update. *SIGKDD Explor.* **2009**, *11* (1), 10−18.
(32) Osborn, M. J.; Gander, J. E.; Parisi, E; Carson, J. Mechanism of assembly of the outer membrane of *Salmonella typhimurium*. *J. Biol. Chem.* **1972**, *247* (12), 3962−3972.
(33) De Vrije, T.; Tommassen, J.; De Kruijff, B. Optimal posttranslational translocation of the precursor of PhoE protein across *Escherichia coli* membrane vesicles requires both ATP and the protonmotive force. *Biochim. Biophys. Acta* **1987**, *900* (1), 63−72.
(34) Wittig, I.; Braun, H.-P.; Schägger, H. Blue native PAGE. *Nat. Protoc.* **2006**, *1* (1), 418−428.
(35) Brookes, P. S.; Pinner, A.; Ramachandran, A.; Coward, L.; Barnes, S.; Kim, H.; Darley-Usmar, V. M. High throughput two-dimensional blue-native electrophoresis: a tool for functional proteomics of mitochondria and signaling complexes. *Proteomics* **2002**, *2* (8), 969−977.
(36) Yamaguchi, K.; Yu, F.; Inouye, M. A single amino acid determinant of the membrane localization of lipoproteins in *E. coli*. *Cell* **1988**, *53* (3), 423−432.
(37) Narita, S.; Matsuyama, S.; Tokuda, H. Lipoprotein trafficking in *Escherichia coli*. *Arch. Microbiol.* **2004**, *182* (1), 1−6.
(38) Elias, D. A.; Mukhopadhyay, A.; Joachimiak, M. P.; Drury, E. C.; Redding, A. M.; Yen, H.-C. B.; Fields, M. W.; Hazen, T. C.; Arkin, A. P.; Keasling, J. D.; Wall, J. D. Expression profiling of hypothetical genes in *Desulfovibrio vulgaris* leads to improved functional annotation. *Nucleic Acids Res.* **2009**, *37* (9), 2926−2939.
(39) Blatch, G. L.; Lässle, M. The tetratricopeptide repeat: a structural motif mediating protein−protein interactions. *Bioessays* **1999**, *21* (11), 932−939.
(40) D'Andrea, L. D.; Regan, L. TPR proteins: the versatile helix. *Trends Biochem. Sci.* **2003**, *28* (12), 655−662.
(41) Casadio, R.; Fariselli, P.; Finocchiaro, G.; Martelli, P. L. Fishing new proteins in the twilight zone of genomes: the test case of outer membrane proteins in *Escherichia coli* K12, *Escherichia coli* O157:H7, and outer Gram-negative bacteria. *Protein Sci.* **2003**, *12* (6), 1158−1168.
(42) Dehal, P. S.; Joachimiak, M. P.; Price, M. N.; Bates, J. T.; Baumohl, J. K.; Chivian, D.; Friedland, G. D.; Huang, K. H.; Keller, K.; Novichkov, P. S.; Dubchak, I. L.; Alm, E. J.; Arkin, A. P. MicrobesOnline: an integrated portal for comparative and functional genomics. *Nucleic Acids Res.* **2010**, *38* (suppl 1), D396−D400.
(43) Yu, N. Y.; Laird, M. R.; Spencer, C.; Brinkman, F. S. L. PSORTdb–an expanded, auto-updated, user-friendly protein subcellular localization database for Bacteria and Archaea. *Nucleic Acids Res.* **2011**, *39* (suppl 1), D241−D244.
(44) Zhang, W.; Gritsenko, M. A.; Moore, R. J.; Culley, D. E.; Nie, L.; Petritis, K.; Strittmatter, E. F.; Camp, D. G., II; Smith, R. D.; Brockman, F. J. A proteomic view of *Desulfovibrio vulgaris* metabolism as determined by liquid chromatography coupled with tandem mass spectrometry. *Proteomics* **2006**, *6* (15), 4286−4299.
(45) Koronakis, V.; Li, J.; Koronakis, E.; Stauffer, K. Structure of TolC, the outer membrane component of the bacterial type I efflux system, derived from two-dimensional crystals. *Mol. Microbiol.* **1997**, *23* (3), 617−626.
(46) Koebnik, R.; Locher, K. P.; Van Gelder, P. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol. Microbiol.* **2000**, *37* (2), 239−253.
(47) Jap, B. K.; Walian, P. J. Structure and functional mechanism of porins. *Physiol. Rev.* **1996**, *76* (4), 1073−1088.
(48) van den Berg, B.; Black, P. N.; Clemons, W. M., Jr.; Rapoport, T. A. Crystal structure of the long-chain fatty acid transporter FadL. *Science* **2004**, *304* (5676), 1506−1509.
(49) Hearn, E. M.; Patel, D. R.; Lepore, B. W.; Indic, M.; van den Berg, B. Transmembrane passage of hydrophobic compounds through a protein channel wall. *Nature* **2009**, *458* (7236), 367−371.
(50) Koebnik, R. TonB-dependent trans-envelope signaling: the exception or the rule? *Trends Microbiol.* **2005**, *13* (8), 343−347.
(51) Pautsch, A.; Schulz, G. E. High-resolution structure of the OmpA membrane domain. *J. Mol. Biol.* **2000**, *298* (2), 273−282.
(52) Zheng, C.; Yang, L.; Hoopmann, M. R.; Eng, J. K.; Tang, X.; Weisbrod, C. R.; Bruce, J. E. Cross-linking measurements of *in vivo* protein complex topologies. *Mol. Cell. Proteomics* **2011**, *10* (10), No. M110.006841.
(53) Parsons, L. M.; Lin, F.; Orban, J. Peptidoglycan recognition by Pal, an outer membrane lipoprotein. *Biochemistry* **2006**, *45* (7), 2122−2128.
(54) Hagan, C. L.; Kim, S.; Kahne, D. Reconstitution of outer membrane protein assembly from purified components. *Science* **2010**, *328* (5980), 890−892.
(55) Sperandeo, P.; Lau, F. K.; Carpentieri, A.; De Castro, C.; Molinaro, A.; Deho, G.; Silhavy, T. J.; Polissi, A. Functional analysis of the protein machinery required for transport of lipopolysaccharide to the outer membrane of *Escherichia coli*. *J. Bacteriol.* **2008**, *190* (13), 4460−4469.
(56) Wu, T.; McCandlish, A. C.; Gronenberg, L. S.; Chng, S.-S.; Silhavy, T. J.; Kahne, D. Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (31), 11754−11759.
(57) Valente, F. M. A.; Oliveira, A. S. F.; Gnadt, N.; Pacheco, I.; Coelho, A. V.; Xavier, A. V.; Teixeira, M.; Soares, C. M.; Pereira, I. A. C. Hydrogenases in *Desulfovibrio vulgaris* Hildenborough: structural and physiologic characterization of the membrane-bound [NiFeSe] hydrogenase. *J. Biol. Inorg. Chem.* **2005**, *10* (6), 667−682.
(58) Valente, F. M. A.; Pereira, P. M.; Venceslau, S. S.; Regalla, M.; Coelho, A. V.; Pereira, I. A.C. The [NiFeSe] hydrogenase from *Desulfovibrio vulgaris* Hildenborough is a bacterial lipoprotein lacking a typical lipoprotein signal peptide. *FEBS Lett.* **2007**, *581* (18), 3341−3344.
(59) Caffrey, S. M.; Park, H.-S.; Voordouw, J. K.; He, Z.; Zhou, J.; Voordouw, G. Function of periplasmic hydrogenases in the sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough. *J. Bacteriol.* **2007**, *189* (17), 6159−6167.
(60) Kihara, A.; Akiyama, Y.; Ito, K. A protease complex in the *Escherichia coli* plasma membrane: HflKC (HflA) forms a complex with FtsH (HflB), regulating its proteolytic activity against SecY. *EMBO J.* **1996**, *15* (22), 6122−6131.
(61) Zane, G. M.; Yen, H. B.; Wall, J. D. Effect of the deletion of qmoABC and the promoter-distal gene encoding a hypothetical protein on sulfate reduction in *Desulfovibrio vulgaris* Hildenborough. *Appl. Environ. Microbiol.* **2010**, *76* (16), 5500−5509.