SUPPLEMENTARY INFORMATION

Bacterial interactomes: interacting protein partners share similar function and are validated in independent assays more frequently than previously reported

Maxim Shatsky[1,#], Simon Allen[2,#], Barbara L Gold[3], Nancy L Liu[3], Thomas R Juba[4], Sonia A Reveco[1], Dwayne A Elias[5], Ramadevi Prathapam[3], Jennifer He[3], Wenhong Yang[3], Evelin D Szakal[2], Haichuan Liu[2], Mary E Singer[6], Jil T Geller[6], Bonita R Lam[6], Avneesh Saini[3], Valentine V Trotter[3], Steven C Hall[2], Susan J Fisher[2], Steven E Brenner[1,7], Swapnil R Chhabra[1], Terry C Hazen[8], Judy D Wall[4], H Ewa Witkowska[2], Mark D Biggin[9], John-Marc Chandonia[1,*] and Gareth Butland[3,*].

[1]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, United States.

[2]Department of Obstetrics, Gynecology and Reproductive Sciences, University of California, San Francisco, California, United States.

[3]Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, United States.

[4]Departments of Biochemistry and of Molecular Microbiology & Immunology, University of Missouri, Columbia, Missouri, United States.

[5]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, United States.

[6]Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, United States.

[7]Department of Plant and Microbial Biology, University of California, Berkeley, California, United States.

[8]Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, Tennessee, United States.

[9]Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, California, United States.

[#]These authors contributed equally to this work.

[*]Corresponding authors: gpbutland@lbl.gov and JMChandonia@lbl.gov.

Supplemental Text

Discussion of individual protein complexes from *D. vulgaris*

Complexes associated with sulfate reduction and energy conservation

Desulfovibrio vulgaris, a Deltaproteobacteria, is model for sulfate reducing bacterium (SRB). The biological reduction of sulfate (SO_4^{2-}) to sulfide (S^{2-}) by SRBs is an ancient process, dating back ~3.5 billion years (1). *D. vulgaris*, like most SRBs, is an obligate anaerobe that couples energy conservation to the reduction of sulfate using small carbon compounds, hydrogen or inorganic ions as electron donors (2). SRBs can degrade environmental contaminants, such as polyaromatic hydrocarbons, and reduce heavy metals, making them useful in environmental remediation efforts (3, 4), or they can have negative economic impacts, such as metal and concrete corrosion or oil souring (5, 6). The protein complexes associated with sulfate respiration and energy conservation have not surprisingly therefore been a focal point for many of the previous studies characterizing protein complexes from *D. vulgaris* and other SRMs (reviewed in (7, 8)).

Dissimilatory sulfite reductase (Dsr). Dsr is central to sulfate respiration. In *D. vulgaris* and other SRMs Dsr α and β subunits are encoded by *dsrA* and *dsrB* respectively as part of a *dsrABD* operon (7, 9, 10). The function of the *dsrD* gene has been unclear, with some reports suggesting it is not associated with DsrAB but is instead a transcriptional regulator while others suggest that it functions with DsrAB (11-13). An additional partner of the Dsr complex, DsrC or the gamma subunit, is encoded at a separate locus on the *D. vulgaris* chromosome (8, 14). Previous characterization has shown that Dsr can be isolated in several multimeric states, including $\alpha_2\beta_2$, $\alpha_2\beta_2\gamma_1$ and $\alpha_2\beta_2\gamma_2$ (15). Our affinity purifications of both DsrA and DsrB support these observations as DsrA and DsrB co-purified with each other; and DsrC was detected in purifications using tagged DsrA. However, tagged DsrD co-purified DsrA and DsrB, contrary to previous work (11) (Fig. 3: a_1 and supplemental Fig. S11 a). In addition, we identified protein DVU3273, previously annotated as a hypothetical protein of unknown function, in both DsrA and DsrB purifications and subsequent tagging of DVU3273 co-purified DsrA and DsrB and no other proteins. (For clarity, we note that all proteins previously annotated as hypothetical are still referred to as hypothetical even after identification in our MS data.) When compared to our purifications using tagged DsrA or DsrB, the yields in our DVU3273 and DsrD purifications were lower, suggesting that these two proteins bind substoichiometrically to the DsrAB complex. Unfortunately tagged DsrC failed to co-purify any proteins possibly due to the STF tag at the C-terminus disrupting folding or an interacting surface (9, 16). Thus, our analysis has identified a new Dsr interacting protein and strengthened the likelihood that DsrD is in fact involved, via direct association with the catalytic subunits, in Dsr function.

Quinone-interacting Membrane bound Oxidoreductase (QmoABC). The Qmo protein complex is essential for the reduction of sulfate but not sulfite and physically interacts with adenosine 5′-phosphosulfate (APS) reductase, which is encoded in the same operon (17, 18). In our study, tagged QmoB purifies QmoA and QmoC and also DVU0851/QmoD, which is encoded at the end of the *apsBA-qmoABC-DVU0851* operon (Fig. 3: a_2 and supplemental Fig. 11 b; (19)). In addition, tagged QmoD purifies QmoA, QmoB and QmoC. QmoD was previously

purified from *D. vulgaris* as part of a large set of proteins that included adenosine phosphosulfate reductase (ApsAB), but neither QmoC or QmoD were found associated with QmoA or QmoB in this earlier study (19). Recent work shows that a *D. vulgaris qmoD* mutant displays a slow growth phenotype when sulfate is the electron acceptor, consistent with impaired but not absent activity of Qmo (18). Together with our results, these genetic data suggest a role for QmoD in the assembly or stability of the Qmo complex rather than catalysis. Finally, we note that ApsA was also detected by MS in a tagged QmoD purification. This interaction is not present, however, in our high-confidence dataset because ApsA was removed by our computational filtering steps due to its presence in a wide array of purifications, likely because of its high natural abundance.

Flavin Oxidoreductase (FlxABCD). The *flxABCD-hdrABC* gene cluster in *D. vulgaris* encodes a flavin oxidoreductase (FlxABCD) and heterodisulfide reductase (HdrABC), both of which have recently been functionally linked to a role in ethanol metabolism, working with alcohol dehydrogenase, Adh1 (DVU2405) encoded upstream of *hdrC* (20). Based on sequence analysis, the Flx complex has been proposed to oxidize NADH and provide electrons to Hdr, which may in turn provide electrons to dissimilatory sulfite reductase via DsrC (8). We observed several interactions among members of the Flx and Hdr complexes (Fig. 3: a_3), notably FlxA reciprocally co-purified FlxB, and FlxB co-purified FlxCD and HdrA, thereby providing experimental evidence for a physical link between the Flx and Hdr complexes for the first time. Additionally, FlxA reciprocally co-purified with DVU2398, a predicted rubrerythrin encoded immediately adjacent to *flxA*, but previously not functionally linked with the *flx* gene cluster. Functional interaction of DVU2398 with the *flx* system is further supported by the finding that tagged DVU2398 co-purified FlxB and FlxCD in addition to FlxA.

Metalloprotein cofactor biosynthesis

Many sulfate respiration and energy conservation complexes contain ubiquitous metalloprotein cofactors such as heme, siroheme, iron sulfur (FeS) clusters and the more specialized NiFe(Se) cofactors found in hydrogenase enzymes (7, 8).

Iron sulfur clusters. Little is known about FeS cluster biosynthesis in obligate anaerobes such as *D. vulgaris*. However, we have observed several interactions between putative biosynthesis proteins, including that between cysteine desulfurase (NifS) and its gene neighbor NifU and the uncharacterized protein DVU2427 (Fig. 3: b_1), and also that between SufB and SufC (Fig. 3: b_2). The NifSU and SufBC complexes have been characterized in other species (21). SufBC likely has a role as a scaffold protein in FeS cluster synthesis. However, NifS is the sole annotated cysteine desulfurase in *D. vulgaris,* and it is not currently known if NifS and NifU function in nitrogen fixation only or have a more generic role in FeS cluster biosynthesis (22).

Heme biosynthesis. Heme biosynthesis in *Desulfovibrio* species and methanogenic Archaea is facilitated by a novel pathway not found in other organisms (23, 24). In this alternative heme biosynthesis (*ahb*) pathway, siroheme, a prosthetic group found in *D. vulgaris* dissimilatory and assimilatory sulfite reductases, is modified to heme *b* via a multienzyme pathway that includes AhbA, AhbB, AhbC and AhbD (24). Interestingly, we observed AhbA and AhbB to repeatedly co-purify with tagged DVU1000, an uncharacterized peptidase (Fig. 3: b_3). These data are consistent with the proposed role of AhbA and AhbB in catalyzing the first committed step in *D.*

vulgaris heme *b* synthesis, namely the decarboxylation of siroheme (23). The interaction of these proteins with a peptidase suggests a potential mechanism of post translational regulation of siroheme to heme *b* conversion.

RNA synthesis and degradation

The above interactions associated with sulfate respiration and energy conservation were not the only source of novel and unexpected interactions that we uncovered. This is exemplified by our analysis of RNA synthesis and degradation complexes.

RNA polymerase. The expected RNA polymerase subunits were isolated using tagged RpoB, RpoC, sigma factor RpoN, and accessory RpoZ subunits (Fig. 3: c_1). In addition, two further proteins, DVU2459 and DVU2460, reproducibly co-purified with tagged RpoC; and tagged DVU2459 and DVU2460 co-purified each other, the RpoA, RpoB and RpoC subunits of RNA polymerase, and a subunit of DNA gyrase (supplemental Fig. S12). DVU2460 is uncharacterized, but DVU2459 is annotated as RdgC-like, a protein that binds to single and double stranded DNA, functions as a negative regulator of RecA, and may act at replication forks (25). Thus DVU2459/2460 quite likely act in transcription, DNA recombination and DNA repair.

RNA Polymerase subunits also reproducibly associated with other tagged proteins. For example, tagged Asd, an aspartate semi-aldehyde dehydrogenase, co-purified RpoC. In addition, the RNA Polymerase sigma-70 protein RpoD co-purified with three additional baits: DVU1368, a predicted NAD(FAD)-dependent dehydrogenase containing a C-terminal rhodanese domain; DVU1850, a cystathionine beta synthase domain protein; and DVU2360, a predicted FAD-binding ferredoxin-NADPH reductase (supplemental Fig. S12). DVU1368 reciprocally co-purified with DVU1850, further supporting a link between these proteins. While the functional significance of all of these interactions is not fully clear, the results illustrate that even the best characterized complexes may still be found to engage in additional interactions.

RNA degradosome. *D. vulgaris* does not encode a full length Ribonuclease E (Rne) ortholog, but it does encode a smaller Ribonuclease family G/E protein (Rne), which is homologous to parts of both *E. coli* Rng and Rne RNA endonuclease proteins (26). While our tagging of Rne was unsuccessful, this protein reproducibly co-purified as a prey with two other proteins (Fig. 3: c_2): DVU0838, a small conserved hypothetical protein containing an RNA binding domain and DVU1078, a conserved hypothetical protein containing a single stranded nucleic acid binding domain. DVU1078 protein also co-purified with tagged DVU0838, further supporting the interaction between these proteins.

Interactions were also observed for another RNA degradosome component, enolase, which in addition to its well function in glycolysis is a degradosome component in many bacteria (27). Enolase was reproducibly associated with two different groups of proteins (Fig. 3: c_3). The first group are three known or likely subunits of Acetyl-CoA carboxylase that are encoded adjacently on the genome: carboxyl transferase (AccA/D); biotin carboxylase (AccC); and a hypothetical protein DVU2224 that is weakly homologous to biotin carboxyl carrier proteins (28). To confirm that DVU2224 binds biotin, we purified the protein and showed that it is detected in a Western blot by a biotin binding-HRP conjugate (supplemental Fig. S13). The second group of proteins physically associated with enolase in our experiments included the ATP-dependent RNA

helicases DeaD and RhlE as well as hypothetical protein DVU0983, which contains a predicted bifunctional nuclease domain. In addition, while polynucleotide phosphorylase (Pnp) was not observed purifying with tagged enolase, we suspect that the two proteins are functionally linked because they co-purify together when multiple other proteins were used as bait. Taken together, these data support the idea that enolase maintains additional functions, possibly in RNA metabolism, in addition to its glycolytic activity in *D. vulgaris*.

Complexes associated with motility

Flagellar assembly. Our analysis supports and extends previous analyses of flagellar assembly complexes. For example, we find interactions between the hypothetical protein DVU0410; the carbon storage regulator, CsrA; the flagellar assembly factor, FliW; as well as flagellin proteins FlaB3 and FlaD (Fig. 3: d_1). CsrA and FliW are encoded by the *flgF* operon, which also encodes multiple flagellar structural components (26). These observations support recent work which has linked both FliW and CsrA to flagellin homeostasis in *Bacillus subtilis* and suggests that *D. vulgaris* employs a similar mechanism of regulation that additionally involves DVU0410 (29). Consistent with data from other species (30-32), we also observed an interaction between FlgM, which is annotated as an anti-sigma-28 factor and a negative regulator of flagellin synthesis, and FliA (sigma-28), the sigma factor for the flagellar operons (Fig. 3: d_2).

Chemotaxis receptors. In prokaryotes, transmembrane chemoreceptor proteins can spatially organize into multimeric clusters at one or both cell poles (33). We identified a network of interactions between nine proteins from the chemotaxis system: CheV, a CheW-like protein, and eight methyl-accepting chemotaxis receptors: DVU0608, DVU1169, DVU1857, DVU1869, DVU2295, DVU2309, DVU3035, and DVU3082 (Fig. 3: d_3). Among these proteins only two were tagged, CheV and DVU3082. The purification of tagged CheV identified the other eight chemoreceptors as prey. This is expected as CheV serves as a scaffold or an adaptor protein during formation of a multimeric cluster of chemoreceptor proteins in other organisms (33). The purification of DVU3082, by contrast, only detected the bait protein itself as well as two frequent fliers that were excluded by our subsequent filtering.

Interactions of proteins encoded by megaplasmid pDV1

Some metabolic functions, such as nitrogen fixation, are encoded on the *D. vulgaris* 200 kb megaplasmid pDV1, though many of its predicted 156 genes remain experimentally uncharacterized (26). Out of 11 protein-protein interactions between megaplasmid encoded proteins in our high-confidence dataset, only one is with a chromosomally encoded *D. vulgaris* protein. This supports the idea that the megaplasmid provides modular additional functions to the cell, and indeed the megaplasmid can be cured from the cell in a relatively facile manner under standard conditions, with limited but quantifiable impact on the cell (34).

Overall we predict at least five distinct protein complexes encoded by the megaplasmid, many of which appear to have a role in protein secretion (Fig. 3: $e_{1\text{-}5}$). DVUA0114, which contains a DUF1895 domain, co-purified DVUA0115, a YscF family protein, and DVUA0116, a HrpB1/HrpK ortholog, both of which are associated with type III secretion (Fig. 3: e_1) (35, 36). Proteins orthologous to the YopN/SycN/YscB/TyeA protein secretion complex of *Y. pestis* (DVUA0105-

DVUA0106) co-purified each other and a protein of unknown function, DVUA0111.1 (Fig. 3: e_2) (37). DVUA0021, a P-loop ATPase component of a transport system, interacts with its gene neighbor DVUA0022, a predicted multidrug efflux pump (Fig. 3: e_3). DVUA0126, a conserved protein annotated as a secretion system chaperone, co-purifies its gene neighbor DORF2161 (Fig. 3: e_4). Finally, the proteins of adjacent genes DVUA0071 and DVUA0072 interact (Fig. 3: e_5), both of which are predicted UDP-glycosyltransferases (26). The coherence of these data further supports the low false discovery rate for our survey that our regression analysis implies.

Additional novel complexes

Two further intriguing groups of interactions have been detected. One set of interactions not expected based on literature annotation are those made by AphA, which is homologous to histone deacetylases and AcuC-type acetoin utilization proteins (26). AphA reciprocally co-purified with DVU2240, a protein annotated as an N-methylhydantoinase A/acetone carboxylase, and DVU2969 an uncharacterized protein annotated as an acetoacetyl-CoA synthase (Fig. 3: f_1). While there is no obvious functional link to a cellular process based on the current annotation for these two proteins, their orthologs are gene neighbors in many species, supporting the idea of a functional link (38).

The second series of intriguing interactions were those of the hypothetical protein DVU0656 (Fig. 3: f_2). This protein contains a Pfam PF03966 domain and is a member of the Uncharacterized Protein Family UPF0434/Trm112p (39). No characterization of DVU0656, the complexes it forms, or its cellular role has been reported for any prokaryote. However, its yeast ortholog Trm112p binds and activates at least four *S*-adenosyl-methionine (SAM) dependent methyltransferases (MTs) linked to protein translation (40, 41). Strikingly, in *D. vulgaris* we also observe that DVU0656 co-purifies three MTs: HemK, a protein methyltransferase; UbiE, a ubiquinone/menaquinone MT; and Sun, a tRNA and rRNA cytosine-C5 MT. The interacting protein partners between yeast and *D. vulgaris* Trm112 orthologs display an incredible conservation of function in translation associated processes. The notable exception, yeast Bud23, does not have a prokaryotic homolog. However the closest Bud23 homologs in prokaryotes are UbiE family MTs (42). It has been suggested that Trm112p may bind to all partner MTs, can stimulate SAM binding and turnover, and aid substrate recognition by the associated MT (43). Given the high degree of functional coherence between the yeast and *D. vulgaris* interacting partners it is probable that some of the Trm112 functions are conserved in the *D. vulgaris* DVU0656 ortholog.

References

[1] Bontognali, T. R., Sessions, A. L., Allwood, A. C., Fischer, W. W., Grotzinger, J. P., Summons, R. E., and Eiler, J. M. (2012) Sulfur isotopes of organic matter preserved in 3.45-billion-year-old stromatolites reveal microbial metabolism. *Proc Natl Acad Sci U S A* **109**, 15146-15151.

[2] Zhou, J., He, Q., Hemme, C. L., Mukhopadhyay, A., Hillesland, K., Zhou, A., He, Z., Van Nostrand, J. D., Hazen, T. C., Stahl, D. A., Wall, J. D., and Arkin, A. P. (2011) How sulphate-reducing microorganisms cope with stress: lessons from systems biology. *Nat Rev Microbiol* **9**, 452-466.

[3] Lu, X. Y., Zhang, T., and Fang, H. H. (2011) Bacteria-mediated PAH degradation in soil and sediment. *Appl Microbiol Biotechnol* **89**, 1357-1371.

[4] Wall, J. D., and Krumholz, L. R. (2006) Uranium reduction. *Annu Rev Microbiol* **60**, 149-166.
[5] Enning, D., and Garrelfs, J. (2014) Corrosion of iron by sulfate-reducing bacteria: new views of an old problem. *Appl Environ Microbiol* **80**, 1226-1236.
[6] Gieg, L. M., Jack, T. R., and Foght, J. M. (2011) Biological souring and mitigation in oil reservoirs. *Appl Microbiol Biotechnol* **92**, 263-282.
[7] Grein, F., Ramos, A. R., Venceslau, S. S., and Pereira, I. A. (2013) Unifying concepts in anaerobic respiration: insights from dissimilatory sulfur metabolism. *Biochim Biophys Acta* **1827**, 145-160.
[8] Pereira, I. A., Ramos, A. R., Grein, F., Marques, M. C., da Silva, S. M., and Venceslau, S. S. (2011) A comparative genomic analysis of energy metabolism in sulfate reducing bacteria and archaea. *Front Microbiol* **2**, 69.
[9] Oliveira, T. F., Vonrhein, C., Matias, P. M., Venceslau, S. S., Pereira, I. A., and Archer, M. (2008) The crystal structure of Desulfovibrio vulgaris dissimilatory sulfite reductase bound to DsrC provides novel insights into the mechanism of sulfate respiration. *J Biol Chem* **283**, 34141-34149.
[10] Schiffer, A., Parey, K., Warkentin, E., Diederichs, K., Huber, H., Stetter, K. O., Kroneck, P. M., and Ermler, U. (2008) Structure of the dissimilatory sulfite reductase from the hyperthermophilic archaeon Archaeoglobus fulgidus. *J Mol Biol* **379**, 1063-1074.
[11] Hittel, D. S., and Voordouw, G. (2000) Overexpression, purification and immunodetection of DsrD from Desulfovibrio vulgaris Hildenborough. *Antonie Van Leeuwenhoek* **77**, 271-280.
[12] Laue, H., Friedrich, M., Ruff, J., and Cook, A. M. (2001) Dissimilatory sulfite reductase (desulfoviridin) of the taurine-degrading, non-sulfate-reducing bacterium Bilophila wadsworthia RZATAU contains a fused DsrB-DsrD subunit. *J Bacteriol* **183**, 1727-1733.
[13] Mizuno, N., Voordouw, G., Miki, K., Sarai, A., and Higuchi, Y. (2003) Crystal structure of dissimilatory sulfite reductase D (DsrD) protein--possible interaction with B- and Z-DNA by its winged-helix motif. *Structure* **11**, 1133-1140.
[14] Venceslau, S. S., Stockdreher, Y., Dahl, C., and Pereira, I. A. (2014) The "bacterial heterodisulfide" DsrC is a key protein in dissimilatory sulfur metabolism. *Biochim Biophys Acta*.
[15] Oliveira, T. F., Franklin, E., Afonso, J. P., Khan, A. R., Oldham, N. J., Pereira, I. A., and Archer, M. (2011) Structural insights into dissimilatory sulfite reductases: structure of desulforubidin from desulfomicrobium norvegicum. *Front Microbiol* **2**, 71.
[16] Oliveira, T. F., Vonrhein, C., Matias, P. M., Venceslau, S. S., Pereira, I. A., and Archer, M. (2008) Purification, crystallization and preliminary crystallographic analysis of a dissimilatory DsrAB sulfite reductase in complex with DsrC. *J Struct Biol* **164**, 236-239.
[17] Ramos, A. R., Keller, K. L., Wall, J. D., and Pereira, I. A. (2012) The Membrane QmoABC Complex Interacts Directly with the Dissimilatory Adenosine 5'-Phosphosulfate Reductase in Sulfate Reducing Bacteria. *Front Microbiol* **3**, 137.
[18] Zane, G. M., Yen, H. C., and Wall, J. D. (2010) 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* **76**, 5500-5509.
[19] Krumholz, L. R., Wang, L., Beck, D. A., Wang, T., Hackett, M., Mooney, B., Juba, T. R., McInerney, M. J., Meyer, B., Wall, J. D., and Stahl, D. A. (2013) Membrane protein complex of APS reductase and Qmo is present in Desulfovibrio vulgaris and Desulfovibrio alaskensis. *Microbiology* **159**, 2162-2168.
[20] Ramos, A. R., Grein, F., Oliveira, G. P., Venceslau, S. S., Keller, K. L., Wall, J. D., and Pereira, I. A. (2014) The FlxABCD-HdrABC proteins correspond to a novel NADH dehydrogenase/heterodisulfide reductase widespread in anaerobic bacteria and involved in ethanol metabolism in Desulfovibrio vulgaris Hildenborough. *Environ Microbiol*.

[21] Ayala-Castro, C., Saini, A., and Outten, F. W. (2008) Fe-S cluster assembly pathways in bacteria. *Microbiol Mol Biol Rev* **72**, 110-125, table of contents.
[22] Dos Santos, P. C., Johnson, D. C., Ragle, B. E., Unciuleac, M. C., and Dean, D. R. (2007) Controlled expression of nif and isc iron-sulfur protein maturation components reveals target specificity and limited functional replacement between the two systems. *J Bacteriol* **189**, 2854-2862.
[23] Bali, S., Lawrence, A. D., Lobo, S. A., Saraiva, L. M., Golding, B. T., Palmer, D. J., Howard, M. J., Ferguson, S. J., and Warren, M. J. (2011) Molecular hijacking of siroheme for the synthesis of heme and d1 heme. *Proc Natl Acad Sci U S A* **108**, 18260-18265.
[24] Lobo, S. A., Lawrence, A. D., Romao, C. V., Warren, M. J., Teixeira, M., and Saraiva, L. M. (2014) Characterisation of Desulfovibrio vulgaris haem b synthase, a radical SAM family member. *Biochim Biophys Acta*.
[25] Drees, J. C., Chitteni-Pattu, S., McCaslin, D. R., Inman, R. B., and Cox, M. M. (2006) Inhibition of RecA protein function by the RdgC protein from Escherichia coli. *J Biol Chem* **281**, 4708-4717.
[26] 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., and Arkin, A. P. (2010) MicrobesOnline: an integrated portal for comparative and functional genomics. *Nucleic Acids Res* **38**, D396-400.
[27] Bandyra, K. J., Bouvier, M., Carpousis, A. J., and Luisi, B. F. (2013) The social fabric of the RNA degradosome. *Biochim Biophys Acta* **1829**, 514-522.
[28] Broussard, T. C., Kobe, M. J., Pakhomova, S., Neau, D. B., Price, A. E., Champion, T. S., and Waldrop, G. L. (2013) The three-dimensional structure of the biotin carboxylase-biotin carboxyl carrier protein complex of E. coli acetyl-CoA carboxylase. *Structure* **21**, 650-657.
[29] Mukherjee, S., Yakhnin, H., Kysela, D., Sokoloski, J., Babitzke, P., and Kearns, D. B. (2011) CsrA-FliW interaction governs flagellin homeostasis and a checkpoint on flagellar morphogenesis in Bacillus subtilis. *Mol Microbiol* **82**, 447-461.
[30] Chadsey, M. S., and Hughes, K. T. (2001) A multipartite interaction between Salmonella transcription factor sigma28 and its anti-sigma factor FlgM: implications for sigma28 holoenzyme destabilization through stepwise binding. *J Mol Biol* **306**, 915-929.
[31] Sorenson, M. K., Ray, S. S., and Darst, S. A. (2004) Crystal structure of the flagellar sigma/anti-sigma complex sigma(28)/FlgM reveals an intact sigma factor in an inactive conformation. *Mol Cell* **14**, 127-138.
[32] Wosten, M. M., van Dijk, L., Veenendaal, A. K., de Zoete, M. R., Bleumink-Pluijm, N. M., and van Putten, J. P. (2010) Temperature-dependent FlgM/FliA complex formation regulates Campylobacter jejuni flagella length. *Mol Microbiol* **75**, 1577-1591.
[33] Baker, M. D., Wolanin, P. M., and Stock, J. B. (2006) Signal transduction in bacterial chemotaxis. *Bioessays* **28**, 9-22.
[34] Clark, M. E., Edelmann, R. E., Duley, M. L., Wall, J. D., and Fields, M. W. (2007) Biofilm formation in Desulfovibrio vulgaris Hildenborough is dependent upon protein filaments. *Environ Microbiol* **9**, 2844-2854.
[35] Hausner, J., Hartmann, N., Lorenz, C., and Buttner, D. (2013) The periplasmic HrpB1 protein from Xanthomonas spp. binds to peptidoglycan and to components of the type III secretion system. *Appl Environ Microbiol* **79**, 6312-6324.
[36] Sun, P., Tropea, J. E., Austin, B. P., Cherry, S., and Waugh, D. S. (2008) Structural characterization of the Yersinia pestis type III secretion system needle protein YscF in complex with its heterodimeric chaperone YscE/YscG. *J Mol Biol* **377**, 819-830.
[37] Plano, G. V., and Schesser, K. (2013) The Yersinia pestis type III secretion system: expression, assembly and role in the evasion of host defenses. *Immunol Res* **57**, 237-245.

[38] Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C., and Jensen, L. J. (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* **41**, D808-815.

[39] Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R., and Bateman, A. (2010) The Pfam protein families database. *Nucleic Acids Res* **38**, D211-222.

[40] Figaro, S., Wacheul, L., Schillewaert, S., Graille, M., Huvelle, E., Mongeard, R., Zorbas, C., Lafontaine, D. L., and Heurgue-Hamard, V. (2012) Trm112 is required for Bud23-mediated methylation of the 18S rRNA at position G1575. *Mol Cell Biol* **32**, 2254-2267.

[41] Mazauric, M. H., Dirick, L., Purushothaman, S. K., Bjork, G. R., and Lapeyre, B. (2010) Trm112p is a 15-kDa zinc finger protein essential for the activity of two tRNA and one protein methyltransferases in yeast. *J Biol Chem* **285**, 18505-18515.

[42] White, J., Li, Z., Sardana, R., Bujnicki, J. M., Marcotte, E. M., and Johnson, A. W. (2008) Bud23 methylates G1575 of 18S rRNA and is required for efficient nuclear export of pre-40S subunits. *Mol Cell Biol* **28**, 3151-3161.

[43] Liger, D., Mora, L., Lazar, N., Figaro, S., Henri, J., Scrima, N., Buckingham, R. H., van Tilbeurgh, H., Heurgue-Hamard, V., and Graille, M. (2011) Mechanism of activation of methyltransferases involved in translation by the Trm112 'hub' protein. *Nucleic Acids Res* **39**, 6249-6259.

Species	*D. vulgaris*	*E. coli*	*M. pneumoniae*	*H. pylori*	*T. pallidum*	*C. jejuni*	*B. subtilis*	*Synechocystis* sp.
D. vulgaris	***3,530***	1089	242	651	434	774	932	851
E. coli	1089	***4,151***	262	683	431	837	1234	955
M. pneumoniae	242	262	***689***	203	199	212	310	242
H. pylori	651	683	203	***1,575***	329	889	592	548
T. pallidum	434	431	199	329	***1,036***	367	458	376
C. jejuni	774	837	212	889	367	***1,623***	737	635
B. subtilis	932	1234	310	592	458	737	***4,176***	896
Synechocystis sp.	851	955	242	548	376	635	896	***3,579***

Supplemental Table S1. Number of computed homologous proteins between the studied species and the number of genes (diagonal elements) for a given species.

Species	Positive Pairs	Negative Pairs
D. vulgaris	478	34,739
E. coli	1,535	24,3310
M. pneumoniae	56	1,823
H. pylori	375	12,818
T. pallidum	248	5,851
C. jejuni	386	18,699
B. subtilis	397	29,364
Synechocystis sp.	236	21,687

Supplemental Table S2. **Number of gold standard interologs as defined by the EcoCyc complexes.** The numbers are computed based on all protein pairs for a given species and not limited to reported PPIs. The small discrepancy between the number of pairs for *E. coli* in this table and S2 Table (number of PPIs in EcoCyc) is due to several genes annotated by MicrobesOnline as pseudo genes.

Study[1]	Species[2]	# protein pairs[3]	FDR based on EcoCyc[4]	% same operon pairs[5]	Fold same TIGR role[6]	% pairs in Y2H[7]	% pairs in AP-MS[8]
AP-MS Shatsky	*D. vulgaris*	82	20%	43%	8.2	33% (7/21)	47% (9/19)
AP-MS Hu	*E. coli*	248	31%	9.70%	5.2	19% (32/166)	16% (36/232)
AP-MS Arifuzzaman	*E. coli*	30	0%	27%	9.7	23% (5/22)	38% (9/24)
AP-MS Kuhner	*M. pneumoniae*	29	50%	28%	2.9	0% (0/17)	18% (2/11)
AP-MS all reciprocals	**multiple**	**389**	**27%**	**29%**	**6.4**	**22%**	**24%**
Y2H Titz	*T. pallidum*	4	-	0%	0	-	0% (0/3)
Y2H Parrish	*C. jejuni*	96	0%	9.4%	4.3	44% (14/32)	15% (4/26)
Y2H Rajagopala	*E. coli*	96	16%	18%	6.4	36% (8/22)	46% (44/95)
Y2H Hauser	*H. pylori*	42	0%	21%	7.1	73% (16/22)	42% (5/12)
Y2H Sato	*Synechocystis*	19	0%	26%	11	100% (1/1)	100% (1/1)
Y2H all reciprocals	**multiple**	**257**	**9.8%**	**18%**	**6.3**	**56%**	**44%**

Supplemental Table S3. PPI quality metrics for bait-prey prey-bait reciprocal protein pairs from bacterial interactomes. The PPI sets are taken from this study (supplemental Dataset S8) or from published studies described in the Materials and Methods and provided in supplemental Datasets S9-S17. For each set, the protein pairs analyzed are those reciprocally confirmed as bait-prey prey-bait pairs. Note that reciprocal PPIs were not identified in the Y2H PPIs for *B. subtilis*. The rows "all reciprocals" give the weighted mean quality metric scores for the set of reciprocal PPIs for the four AP-MS studies and, separately, for five Y2H surveys. The columns give from left to right: [1] the name and type of dataset; [2] the species; [3] the number of reciprocally confirmed protein pairs; [4] the FDR estimated using gold standard positive and negatives sets based only on complexes from the EcoCyc dataset or, in the case of the non *E. coli* studies, their orthologs; [5] the percent of protein pairs whose members are encoded in the same operon; [6] the fold enrichment of protein pairs for which both members have the same TIGR role over than expected among randomly chosen pairs of proteins. [7] The percent of protein pairs that are also found in the sum of all six Y2H datasets, excepting same study pairs in the case of the Y2H datasets. [8] The percent of protein pairs that are also found in the sum of all four AP-MS datasets, excepting same study pairs in the case of the AP-MS datasets.

Study[1]	Species[2]	% genes tagged and detected[3]	Preys detected per bait [4]	# pairs in high conf. set[5]	# proteins in high conf. set (percent of all proteins)[6]	High conf. pairs per bait[7]	% baits not in high conf. set[8]	% of high conf. set proteins in dimers[9]	Connected component size (std)[10]
Benchmark									
EcoCyc (no ribo.)	*E. coli*	NA	NA	1,549	710 (17%)	NA	NA	20%	3.9 (4.45)
Our high confidence									
Shatsky	*D. vulgaris*	27%	14.3	459	469 (14%)	0.48	64%	36%	3.39 (8.33)
Hu revised	*E. coli*	NA	NA	391	386 (9.3%)	0.45	63%	35%	3.6 (5.6)
Previous AP-MS interactomes									
Hu	*E. coli*	35%	22	5,993	1,757 (42%)	4.06	32%	4%	47 (271)
Hu (no ribo.)	*E. coli*	NA	NA	3,794	1,653 (40%)	2.57	36%	5%	34 (220)
Arifuzzaman	*E. coli*	64%	5.0	11,172	2,962 (71%)	4.19	14%	0.1%	740 (1,279)
Arifuzzaman (no ribo.)	*E. coli*	NA	NA	7,977	2,808 (68%)	3.48	23%	0.2%	702 (1,212)
Kuhner	*M. pneumoniae*	31%	14	1,058	410 (60%)	5.00	6%	0.1%	137 (190)
Kuhner (no ribo.)	*M. pneumoniae*	NA	NA	682	359 (52%)	3.21	23%	0.05%	120 (166)

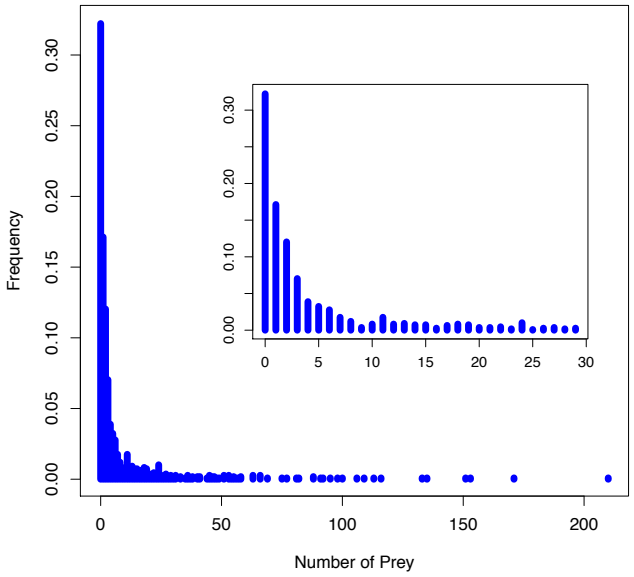
Supplemental Table S4. The connectivity of proposed bacterial AP-MS interactomes vs that of EcoCyc PPIs. PPI data are taken from the EcoCyc database (supplemental Dataset S6); our studies of *D. vulgaris* (supplemental Dataset S8); our reanalysis of Hu et al. 2009 (supplemental Dataset S18); Hu et al.'s original interactome (supplemental Dataset S9); Arifuzzaman et al, 2006 (supplemental Dataset S10); Kuhner et al, 2009 (supplemental Dataset S11). Because our high confidence *D. vulgaris* and *E. coli* interactomes excluded ribosomal proteins and chaperonins, for comparison protein pairs that included these proteins were also removed from the EcoCyc, Hu, Arifuzzaman and Kuhner interactomes in rows labeled "no ribo etc.". The columns give from left to right: [1] the name of the study; [2] the species; [3] the percent of genes in the genome successfully tagged and detected by MS as baits; [4] the mean number of prey proteins detected per bait prior to any computational filtering; [5] the number of protein pairs in the high confidence network; [6] the total number of proteins that are in high confidence pairs, i.e. excluding baits not part of any high confidence pair (the percent of high confidence protein pairs vs the total proteome); [7] the mean number of high confidence protein pairs per bait, including baits that are not part of a high confidence pair; [8] the percent of baits not included in a high confidence protein pair; [9] the percent of proteins that are part of a high confidence protein pair that are linked to only one other protein; [10] the mean size of connected components within the high confidence interactome, i.e. excluding baits not participating in a high confidence pair.

Study[1]	Species[2]	# protein pairs[3]	FDR based on EcoCyc[4]	% same operon pairs[5]	Reciprocal confrm. %[6]	Fold same TIGR role[7]	% pairs in Y2H[8]	% pairs in AP-MS[9]
Benchmarks								
EcoCyc	*E. coli*	1,549	0%	54%	NA	10	11% (77/670)	14% (130/906)
AP-MS reciprocals	multiple	389	27%	29%	NA	6.4	22%	24%
Y2H reciprocals	multiple	257	9.8%	18%	NA	6.3	56%	44%
Hu et al. AP-MS revised								
high confidence	*E. coli*	391	20%	15%	28%	6	25% (48/194)	12% (39/327)
low confidence	*E. coli*	2,961	80%	1.2%	5.5%	1.9	3.4 (57/1684)	4.5% (121/2698)

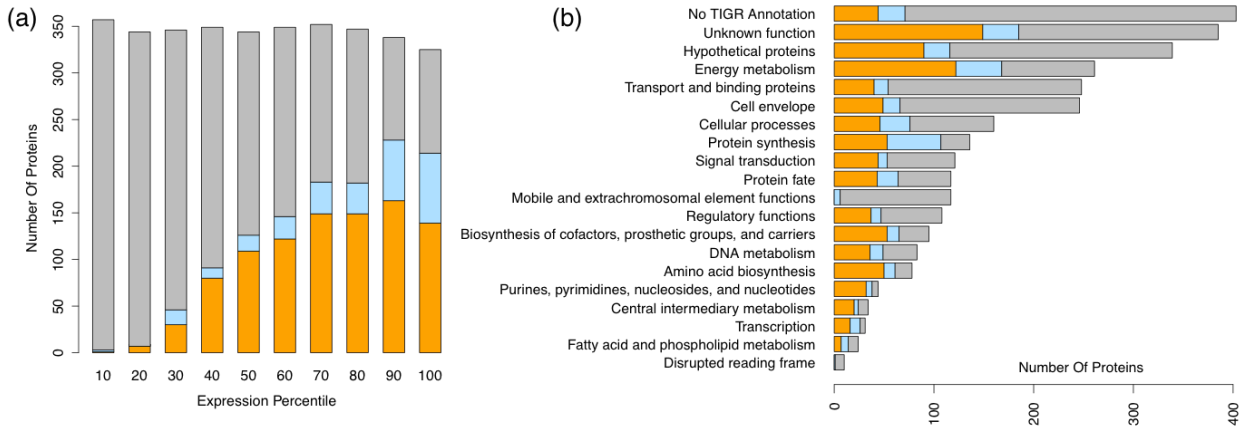
Supplemental Table S5. PPI quality metrics for benchmark datasets and for high and lower confidence sets from reanalysis of Hu et al.'s data. The PPI data for the three benchmark datasets from EcoCyc and the AP-MS and Y2H reciprocals are as described in Fig. 4 and the Experimental Procedures. The high confidence protein pairs from our reanalysis of Hu et al.'s data are shown as our those PPIs not selected in the rows "Hu et al. AP-MS revised". The columns give from left to right: [1] the name and type of dataset; [2] the species; [3] the number of protein pairs; [4] the FDR estimated using gold standard positive and negatives sets based only on complexes from the EcoCyc dataset or, in the case of most reciprocal PPIs, their orthologs; [5] the percent of protein pairs whose members are encoded in the same operon; [6] the percent of protein pairs that are reciprocally confirmed in the same study as both bait – prey and prey – bait pairs. [7] the fold enrichment of protein pairs for which both members have the same TIGR role over than expected among randomly chosen pairs of proteins. [8] The percent of protein pairs that are also found in the sum of all six Y2H datasets, excepting same study pairs in the case of the Y2H reciprocals. [9] The percent of protein pairs that are also found in the sum of all four AP-MS datasets, excepting same study pairs in the case of the AP-MS datasets.

Study[1]	Species[2]	# protein pairs[3]	# proteins in high conf. set (percent of all proteins)[4]	% high conf. set proteins in dimers[5]	Connected component size (std)[6]
Benchmark					
EcoCyc	*E. coli*	1,549	710 (17%)	20%	3.9 (4.45)
Our high confidence interactomes					
AP-MS Shatsky	*D. vulgaris*	459	469 (14%)	36%	3.39 (8.33)
AP-MS Hu revised	*E. coli*	391	386 (9.3%)	35%	3.6 (5.6)
Previous Y2H interactomes					
Y2H Titz	*T. pallidum*	979	578 (56%)	2%	72 (185)
Y2H Parrish	*C. jejuni*	2926	1108 (68%)	1%	48 (220)
Y2H Marchadier	*B. subtilis*	704	278 (7%)	3%	56 (107)
Y2H Rajagopala	*E. coli*	1776	1218 (29%)	10%	8.7 (81)
Y2H Hauser	*H. pylori*	728	526 (33%)	9%	18 (84)
Y2H Sato	*Synechocystis*	736	915 (26%)	25%	4.8 (22)

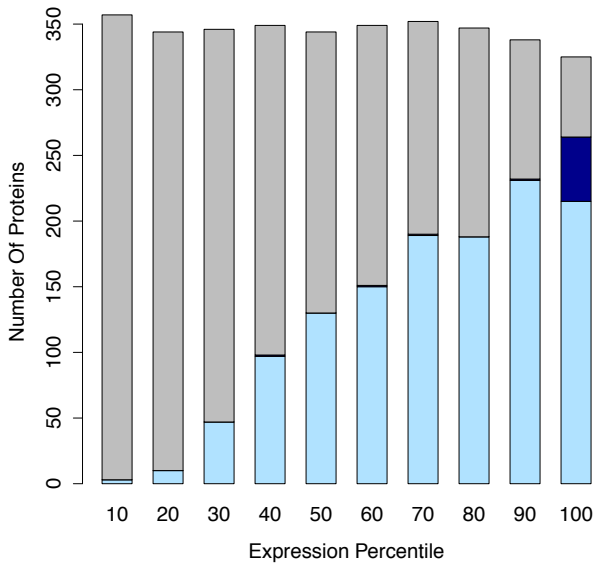
Supplemental Table S6. The connectivity of proposed bacterial Y2H interactomes vs that of EcoCyc and our high confidence interactomes. The PPI data are taken from the sets described Figs. 4 and 7 and Experimental Procedures. The columns give from left to right: [1] the name of the study; [2] the species; [3] the number of protein pairs in the high confidence network; [4] the total number of proteins that are in high confidence protein pairs (the percent of proteins in high confidence pairs vs the total proteome); [5] the percent of proteins that are part of a high confidence protein pair that are linked to only one other protein; [6] the mean size of connected components within the proposed interactomes, i.e. excluding baits not participating in a high confidence protein pair.



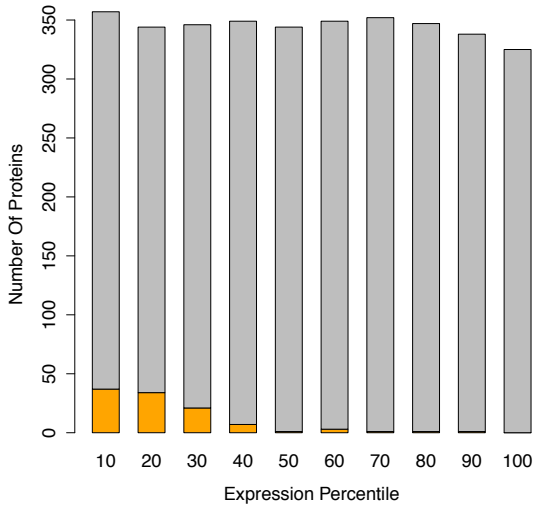
Supplemental Figure S1. Histogram of the number of prey identified per purification. Y-axis shows frequency of purifications with a given number of prey proteins (x-axis) identified by MS by two or more peptides, mean = 7.7, median = 2. The inset shows a zoom-in region up to 30 prey from the main plot.



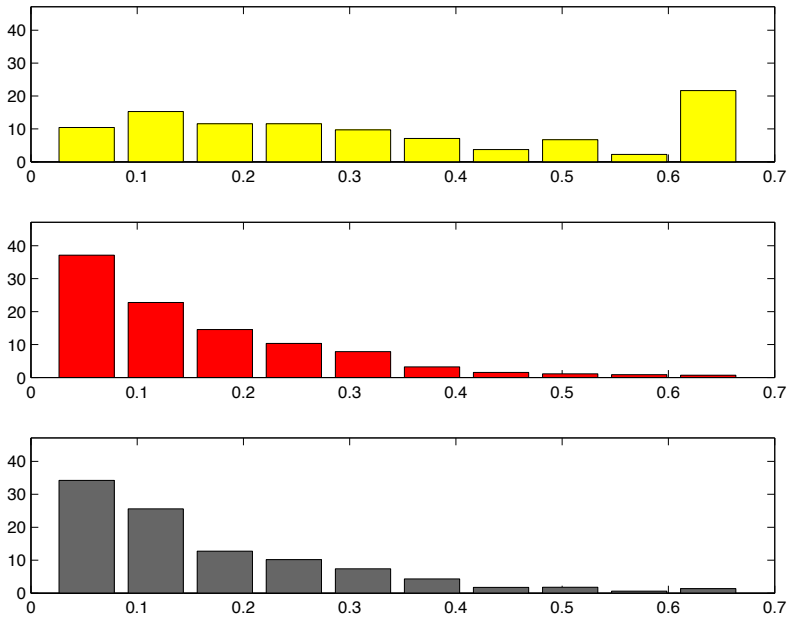
Supplemental Figure S2. Expression levels and TIGR role annotation of MS identified proteins. (a) The shaded sections of each bar show the number of *D. vulgaris* genes encoding tagged bait proteins detected by MS (orange), encoding proteins only detected as prey proteins by MS (cyan), and not detected in by MS (grey). The data are plotted as a function of mRNA microarray expression percentile, with highly expressed genes to the right. Ribosomal genes have been excluded. (b) TIGR roles of MS detected tagged bait proteins (orange), prey only proteins detected by MS (cyan), and all other genes except ribosomal proteins (grey).



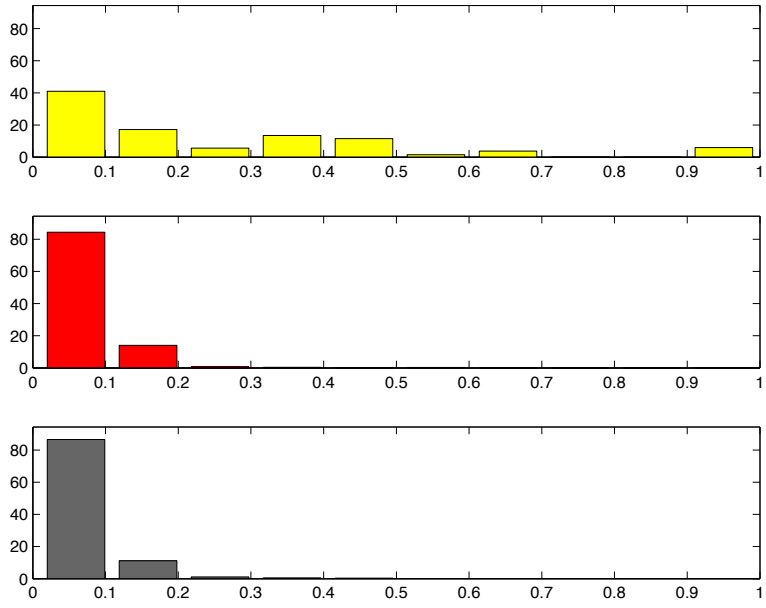
Supplemental Figure S3. mRNA expression levels of proteins detected by MS by two or more peptides. mRNA expression levels are shown when ribosomal proteins are included (dark blue) or excluded (light blue). All other *D. vulgaris* protein-encoding genes are in grey.



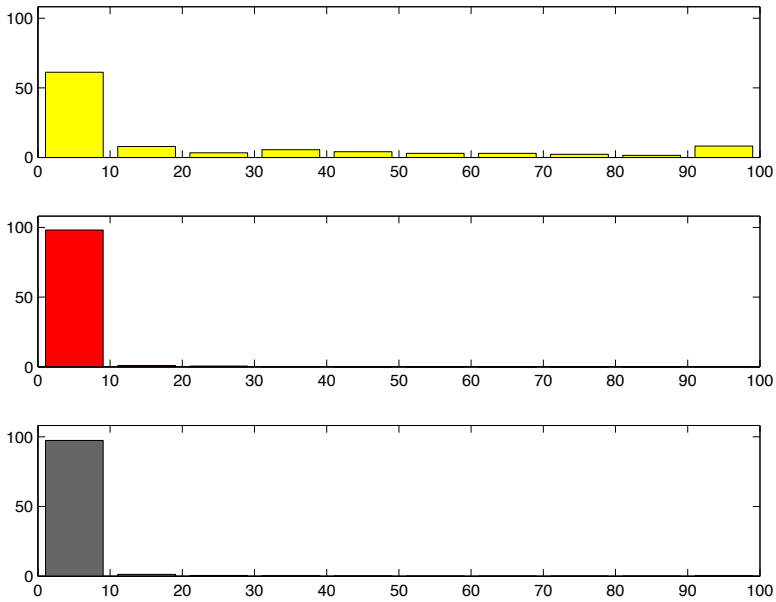
Supplemental Figure S4. mRNA expression levels of genes with the assigned TIGR role "*mobile and extrachromosomal element function*". mRNA expression levels are shown for "*mobile and extrachromosomal element function*" TIGR role genes (orange) and all other protein-encoding genes (grey).



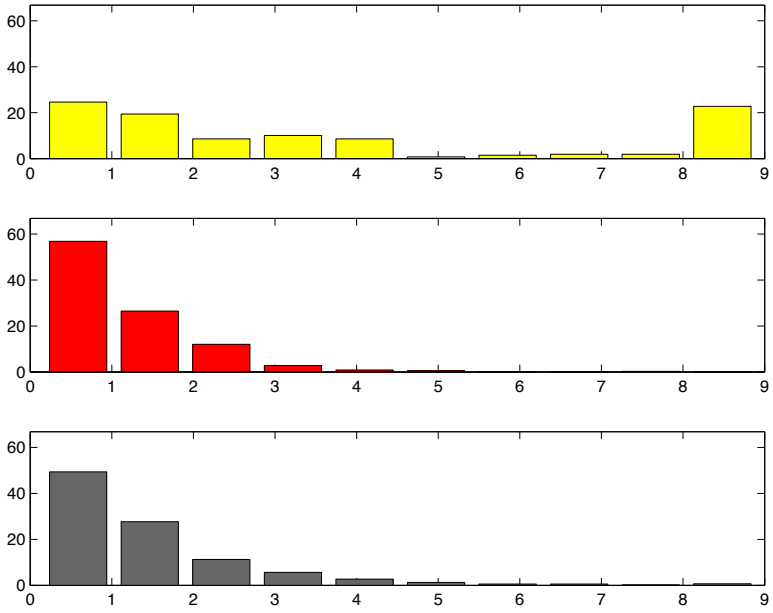
Supplemental Figure S5. Distribution of dice coefficients (feature (1)). Distributions are shown for the Gold Positive set (yellow), Gold Negative set (red), and for all pairs (grey). X-axis – dice coefficient. Y-axis – percentage.



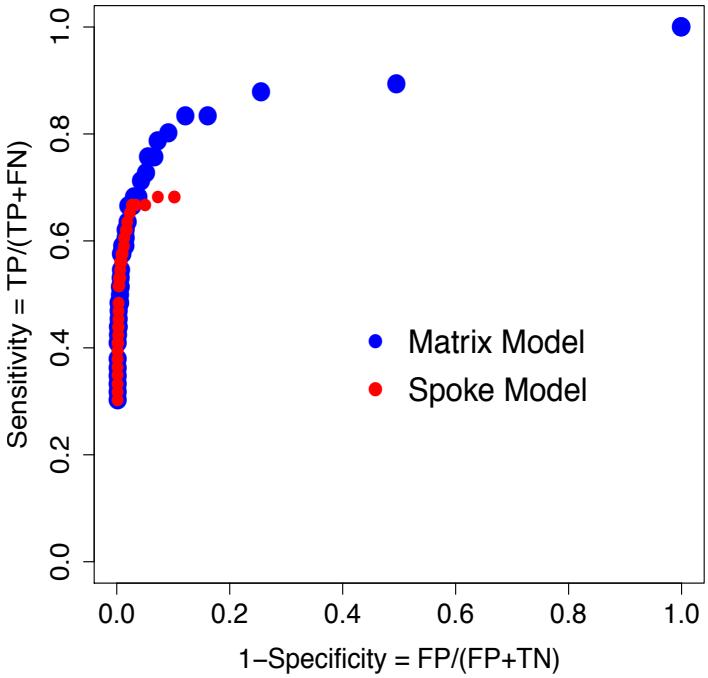
Supplemental Figure S6. Distribution of completeness scores (feature (2)). Distributions are shown for the Gold Positive set (yellow), Gold Negative set (red), and for all pairs (grey). X-axis – completeness score. Y-axis – percentage.



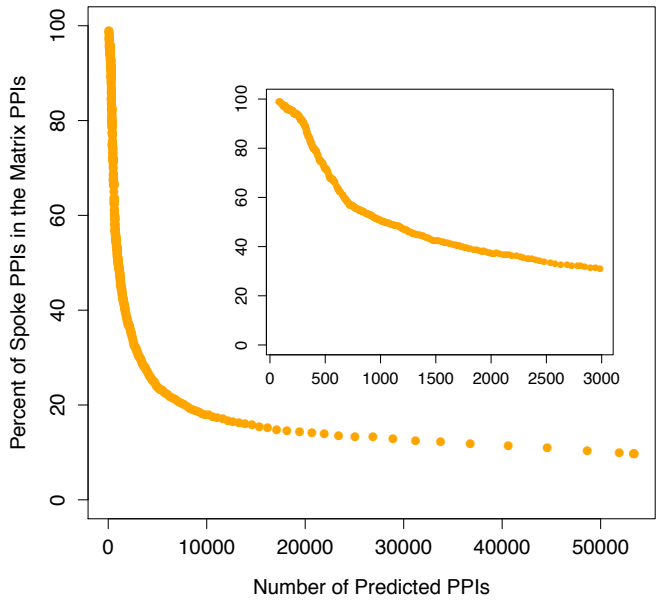
Supplemental Figure S7. Distribution of matrix model NSAF scores (feature (3)). Distributions are shown for the Gold Positive set (yellow), Gold Negative set (red), and for all pairs (grey). X-axis – matrix model NSAF scores. Y-axis – percentage.



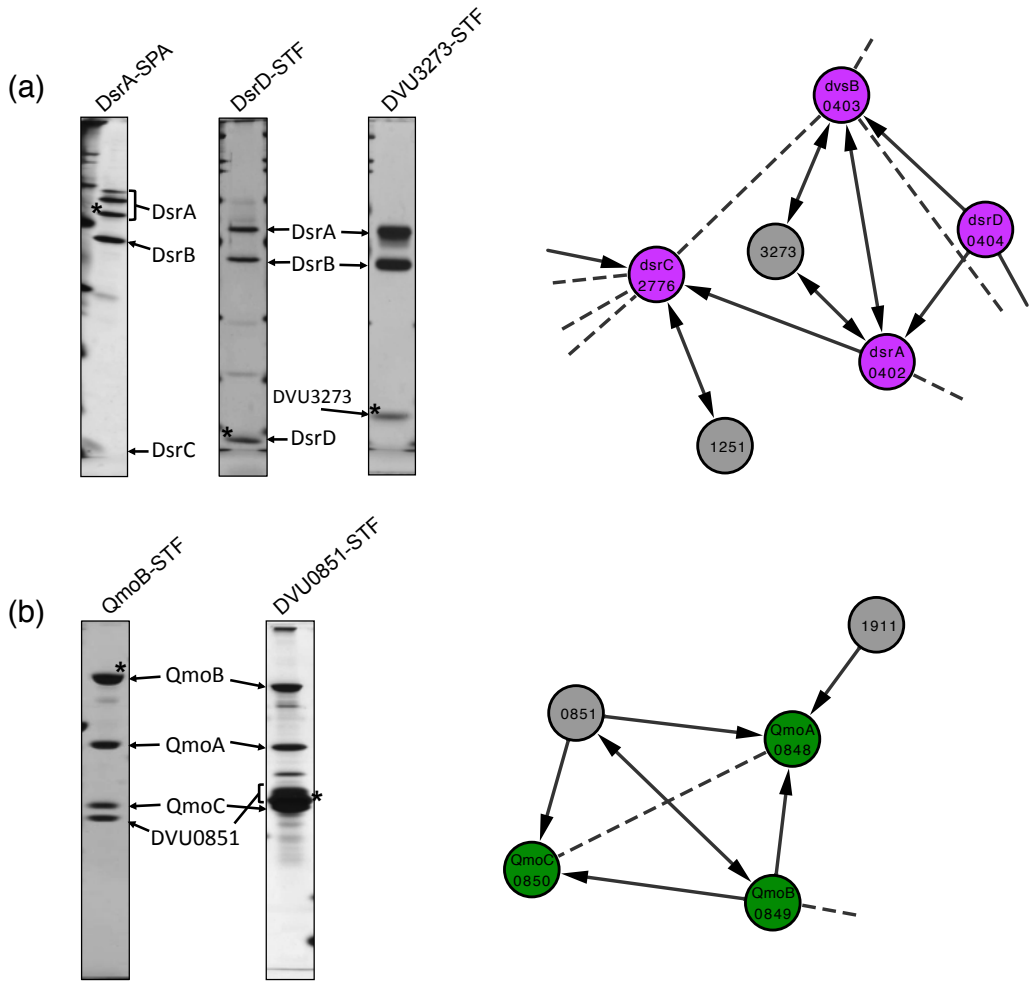
Supplemental Figure S8. Distribution of matrix model ANA scores (feature (4)). Distributions are shown for the Gold Positive set (yellow), Gold Negative set (red), and for all pairs (grey). X-axis – matrix model ANA scores. Y-axis – percentage.



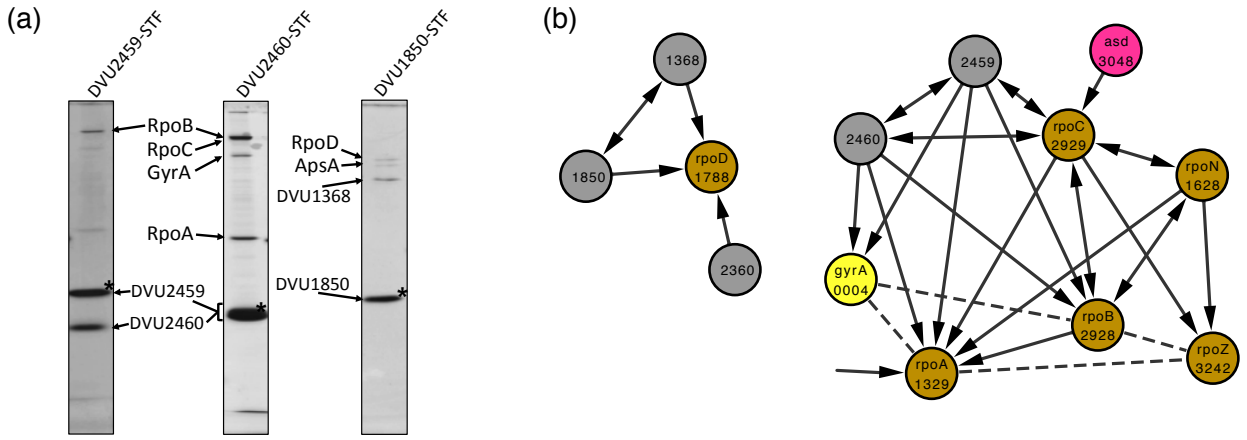
Supplemental Figure S9. Sensitivity and specificity of predicted PPIs as measured from the cross-validation procedure.



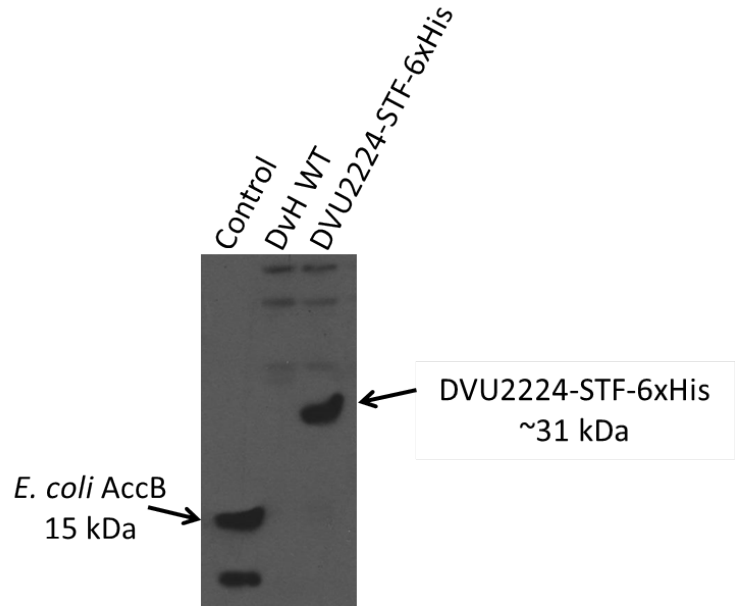
Supplemental Figure S10. Distribution of spoke (bait-prey) pairs within the whole set of matrix (prey-prey) pairs. The inset shows a zoom-in region up to the first 3000 pairs from the main plot.



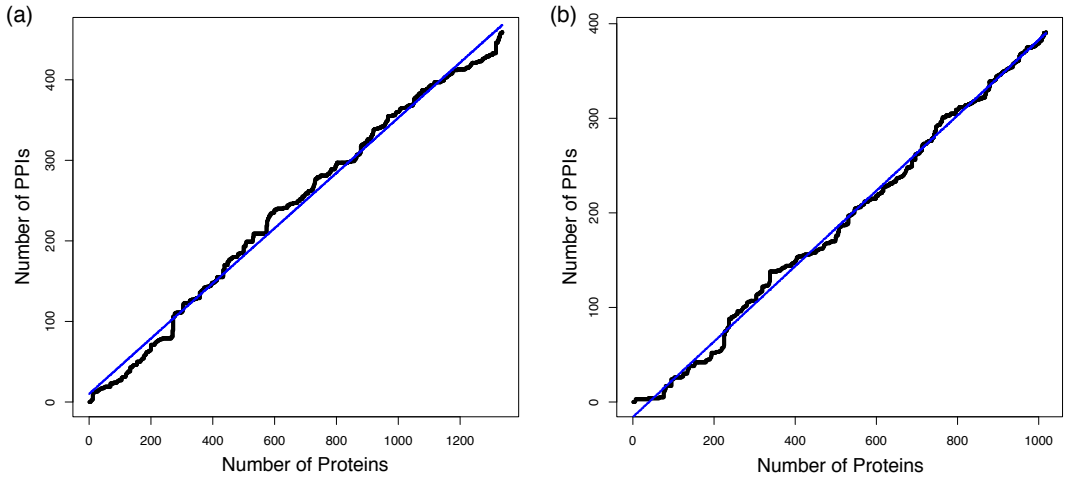
Supplemental Figure S11. Protein complexes involved in sulfate respiration. Silver stained SDS-PAGE gel images of TAP purified eluates from (a) DsrA-SPA, DsrD-STF and DVU3273-STF affinity tagged strains associated with dissimilatory sulfite reductase and (b) QmoB-STF and DVU0851-STF affinity tagged strains associated with the quinone-interacting membrane bound oxidoreductase (left) and corresponding interaction networks (right). Arrows indicate directional bait-to-prey associations. Colors are by annotated TIGR role as indicated in Figure 3. *denotes tagged protein.



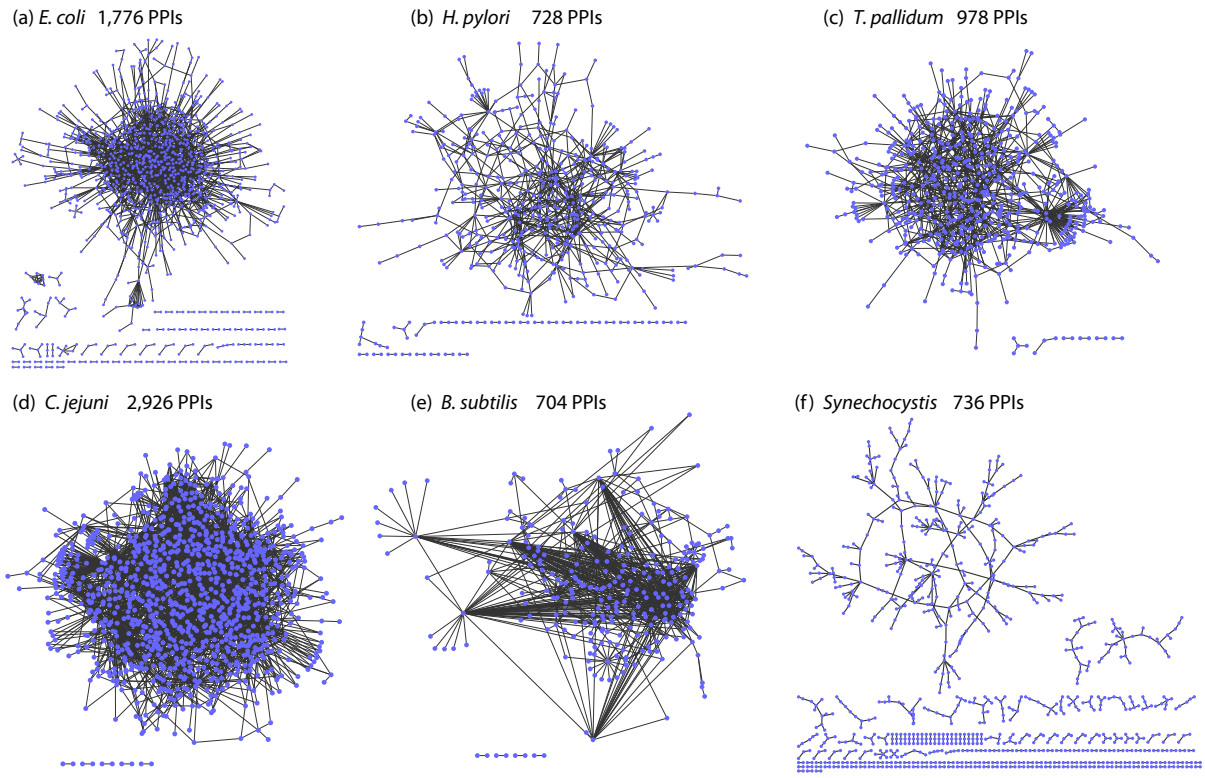
Supplemental Figure S12. RNA polymerase associated protein complexes. (a) Silver stained SDS-PAGE gel images of TAP purified eluates from DVU2459-STF, DVU2460-STF and DVU1850-STF affinity tagged strains associated with RNA polymerase subunits (b) Protein interaction networks for DVU2459, DVU2460 and DVU1850. Arrows indicate directional bait-to-prey associations. Colors are by annotated TIGR role as indicated in Figure 3. *denotes tagged protein.



Supplemental Figure S13. DVU2224 is a biotin binding protein. Western blot analysis of *D. vulgaris* wild type and CAT400742 (DVU2224-STF-6xHis) cell free extracts immunoprecipitated using Anti-FLAG agarose beads and probed using streptavidin-HRP conjugate (Sigma). Proteins were visualized using chemiluminescent substrate. *E. coli* Biotin Carboxyl Carrier Protein (AccB) present in *E. coli* BW25113 cell free extracts is visualized as a control.



Supplemental Figure S14. Linear correlation between the number of detected proteins and the number of inferred PPIs. The x-axis shows the accumulated number of distinct detected proteins in AP-MS studies in *D. vulgaris* (a) and *E. coli* (b) sorted by purification experiments, i.e. as more purifications are analyzed the more distinct proteins we observe in total. The y-axis shows the number of high confident PPIs that are present in at least one of the experiments accumulated on x-axis.



Supplemental Figure S15. Previously published Y2H interactomes for six bacterial species. The interactomes are defined in the Experimental Procedures.