Published in final edited form as:

J Proteome Res. 2009 March ; 8(3): 1504–1514. doi:10.1021/pr800865d.

A method for investigating protein-protein interactions related to *Salmonella* Typhimurium pathogenesis

Saiful M. Chowdhury¹, Liang Shi¹, Hyunjin Yoon², Charles Ansong¹, Leah M. Rommereim³, Angela D. Norbeck¹, Kenneth J. Auberry¹, Ronald J. Moore¹, Joshua N. Adkins¹, Fred Heffron², and Richard D. Smith^{1,*}

¹Pacific Northwest National Laboratory, Richland, WA-99352

²Oregon Health and Science University, Portland, OR-97239

³Dartmouth College, Hanover, NH

Abstract

We successfully modified an existing method to investigate protein-protein interactions in the pathogenic bacterium *Salmonella enterica* serovar Typhimurium (*S. Typhimurium*). This method includes i) addition of a histidine-biotin-histidine tag to the bait proteins via recombinant DNA techniques; ii) *in vivo* cross-linking with formaldehyde; iii) tandem affinity purification of bait proteins under fully denaturing conditions; and iv) identification of the proteins cross-linked to the bait proteins by liquid-chromatography in conjunction with tandem mass-spectrometry. *In vivo* cross-linking stabilized protein interactions and permitted the subsequent two-step purification step conducted under denaturing conditions. The two-step purification greatly reduced nonspecific binding of non-cross-linked proteins to bait proteins. Two different negative controls were employed to eliminate the possibility of identifying background and non-specific proteins as interacting partners, especially those caused by non-specific binding to the stationary phase used for protein purification. In an initial demonstration of this approach, we tagged three *Salmonella* proteins—HimD, PduB and PhoP— with known binding partners that ranged from stable (e.g., HimD) to transient (i.e., PhoP). Distinct sets of interacting proteins were identified for each bait protein, including the known binding partners such as HimA for HimD, as well as unexpected binding partners. Our results suggest that novel protein-protein interactions may be critical to pathogenesis by *Salmonella*.

Keywords

HBH tag; formaldehyde; cross-linking; mass spectrometry; *in vivo* interactions

Introduction

A number of methods are currently used for studying the dynamic interactions among proteins, even though most are not carried out under physiologically relevant conditions¹. One of the more broadly applied techniques for studying cellular protein interactions *in vitro* is Yeast two-hybrid (Y2H); however, this technique suffers from significant false positive and false negative

*To whom correspondence should be addressed: Pacific Northwest National Laboratory, P.O Box 999, MSIN: K8-98, Richland, WA 99352, E-mail: E-mail: rds@pnl.gov.

Supporting Information **Available**: Bacterial strains, plasmid and primer sequences used in this study are available in the supplementary tables. Complete protein heat map data for all controls and cross-link experiments are also available in excel file format. This information is available free of charge via the Internet at <http://pubs.acs.org>.

identification rates, in part because of incomplete and/or inappropriate protein modifications^{2, 3}. Co-immunoprecipitation is another widely applied *in vivo* technique that uses antibodies to isolate proteins of interest together with interacting partners from their cellular environment⁴. Limitations of this method include the need to empirically determine appropriate interacting conditions, such as sensitivity to detergents present during the immunoprecipitation, and co-immunoprecipitation of many irrelevant proteins. In addition, proteases that are present in the cell may degrade the immunoprecipitate during purification, particularly when the purification step involves complexes such as the proteasome. Cross-reactivity between antibody and homologues of the antigen also generates false-positive partners. Several approaches have been developed to overcome these shortfalls^{5, 6}; for example, QUICK (quantitative immunoprecipitation combined with knockdown) utilizes SILAC (stable isotope labeling by amino acids in cell culture), co-immunoprecipitation and RNAi knock down to distinguish between background and interacting partners⁶. While promising, this method depends on successfully accomplishing several sequential molecular biology techniques. In addition, the method is currently not high-throughput, and requires antibody and knockdown constructs for each targeted protein.

A new approach that at least partially addresses these problems has recently been reported⁷. In this approach a histidine-biotin-histidine (HBH) tag is added to specific proteins via recombinant DNA techniques, and the resultant construct is transformed and expressed in the target organism. After cross-linking *in vivo* with formaldehyde to stabilize protein-protein interactions from the cell context, the bait proteins are isolated under denaturing conditions with the relevant interacting proteins attached via the cross-link and identified by mass spectrometry (MS). In a demonstration of the approach, tandem affinity purification under fully denaturing conditions significantly reduced and even eliminated non-specific binding, and consequently increased the likelihood of correct identifications between the bait and its binding partners^{7, 8}. Application of this method resulted in identification of particularly unstable proteins in yeast, such as those that have been ubiquitinated and deubiquitinated by enzyme activity during cell lysis and purification process⁷. This work arguably yielded the first complete interaction map of the 26 S proteasome in yeast⁸.

In the study reported herein, we refined the HBH tag approach for rapid label-free application to investigate protein-protein interactions in *Salmonella enterica* serovar Typhimurium (a.k.a. *S. Typhimurium*). This pathogenic bacterium causes acute gastroenteritis and diarrhea in humans and systemic infections in mice that lack functional Slc11a1 (formerly called Nramp1), whose symptoms are similar to that of human typhoid fever caused by *Salmonella enterica* serovar Typhi^{9, 10}. This similarity is the reason that *S. Typhimurium*-mediated systemic infection in mice serves as a model system for studying the pathogenesis and immunology of typhoid fever in humans. To establish systemic infection in mice, *S. Typhimurium* specifically uses a group of proteins (i.e., virulence factors) to coordinate *Salmonella* activity for evading the host immune systems¹¹. To investigate the protein-protein interactions that may enable *S. Typhimurium* to evade host immune responses, three protein virulence factors were chosen to serve as models to study the dynamic interactions of these pathogenic markers and other proteins during the course of infection.

The three proteins were selected as bait proteins not only because of their suspected roles in virulence, but also because they have either known or suspected interacting proteins. The first of these virulence factors, integration host factor beta subunit (a.k.a. HimD or STM0982) is an ~11 kDa DNA binding protein¹¹ that produces a stable interaction with protein Ihf α (HimA)¹². The role of this complex is to affect DNA structural organization, replication, and site specific recombination by binding to a conserved DNA sequence resulting in a conformational change. The second virulence factor, PhoP, is a membrane associated two-component ~25 kDa regulatory protein¹³ that has a transient interaction with the protein PhoQ. This two component

protein complex has been reported to have a role in sensing external divalent cations, such as Mg^{2+} . Cell growth under MgM shock condition (depletion of magnesium) results in activation of this virulent gene expression^{14, 15}. The third virulence factor, PduB is a ~24 kDa protein, which is a component of polyhedral bodies in *Salmonella enterica*. A recent study showed that the Pdu operon, of which PduB is a member, enhances virulence¹⁶. The Pdu operon encoded proteins for propanediol degradation^{17,18, 19}.

We applied the refined HBH tag approach to confirm and discover new protein interactions with the three selected virulence factors. The use of different controls to help screen different sources of background allowed us to confidently study the interactions of these proteins. Experimentally determined interactions not only served to demonstrate the general capability of this HBH-tag-based method towards bacterial species, but also provided novel insights into the pathogenesis of *S. Typhimurium*.

Experimental procedures

Chemicals and standard protocols

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western immunoblot assays were conducted according to the instructions from Invitrogen (Carlsbad, CA). Ni-nitrilotriacetic acid (NTA) agarose and anti-RGSHis antibody were purchased from Qiagen (Valencia, CA) and ImmunoPure Immobilized Streptavidin, from Pierce (Rockford, IL).

Bacterial strains and gene cloning

The wild type strain *S. Typhimurium* 14028 and four other derivatives of wild type strains were used in this study. These derivative strains are labeled LS1001, LS1073, LS1084 and LS1085. LS1001 contains the plasmid pKD46 that possesses the phage λ -Red recombinant system used for inserting the coding sequence for HBH into the chromosome of the wild type strain^{7, 11, 20, 21}. LS1073 and LS1085 were generated, fusing the coding sequence for the HBH tag into the HimD (STM0982) and PduB (STM2039) encoding genes, respectively. These mutant strains were validated by PCR for gene sequence and Western blot analysis for tagged protein expression. Derivative strain LS1084 was constructed to express the HBH tagged version of the PhoP (STM1231) gene and the PhoQ (STM1230) encoding gene. This was performed using a plasmid expressing PhoP tagged with HBH at the C-terminus, the coding sequence of PhoP as well as its promoter were PCR amplified and inserted into the plasmid pWKS30HBH as an AvrII/SmaI fragment²². Plasmid pWKS30HBH, a derivative of the low copy plasmid pWKS30, was constructed by inserting 530 bp nucleotide sequences encoding HBH tag between SmaI and BamHI in pWKS30, which is designed to place an HBH tag on target proteins at C-termini²³. For two component regulatory systems, the partner sensor proteins were also placed on the same plasmid using SacI and XbaI sites. Genes cloned on pWKS30HBH were verified by sequencing. All bacterial strains and plasmids used in this study are listed in Supplementary Table 1, and primers used in this study are listed in Supplementary Table 2.

Culturing conditions

All media used for cross-linking contained 4 μ M biotin and all bacterial cells were grown aerobically with agitation (150 rpm). For HimD-HBH, LS1073 was grown in 40 mL of LB medium (10 g/L of tryptone, 5 g/L of yeast extract and 10 g/L of NaCl) supplemented with 50 μ g/mL of kanamycin at 37°C for 16 h. For HBH-PhoP and PduB-HBH, LS1084-5 was cultured in 40 mL of N-minimal medium (100 mM Tris, pH 7.7, 0.5 mM KCl, 7.5 mM NH_4SO_4 , 0.5 mM K_2SO_4 , 1 mM KH_2PO_4) supplemented with 8 μ M of $MgCl_2$, 0.1% Casamino acids, 38 mM glycerol, and 15 μ g/mL of ampicillin at 37°C for 16 h. Cells were harvested (6000 $\times$ g,

15 min, room temperature), resuspended in 40 mL of prewarmed MgM medium (pH 5) with 50 μ M MgCl₂, and grown for 2 h at 37°C. Wild type cells were grown under identical conditions except that antibiotics were omitted.

Cross-linking

After harvesting, cells were washed once with 40 mL of PBS and resuspended in 1 mL of PBS. Different concentrations of formaldehyde (0.5–3%) were directly added to the LS1073 cells to determine optimal cross-linking conditions, and the cells were allowed to incubate for 10 min at 30°C. The cross-linking reaction was quenched by adding 2.5 M glycine to a final concentration of 0.125 M. After cross-linking, cells were collected and washed with dH₂O (4°C). The optimal condition for subsequent tandem affinity purification of cross-linked proteins was determined to be 1% formaldehyde for 10 min at 30°C, which was used in all subsequent experiments (Figure 1).

Tandem affinity purification of cross-linked and non-cross-linked proteins

After cross-linking, *Salmonella* cells were lysed by bead beating in 1 mL lysis buffer (8 M urea, 300 mM NaCl, 50 mM NaH₂PO₄, 0.5% Nonidet P-40, 1 mM PMSF, pH 8). The cell lysates were then sonicated for 20 min to fragment DNA. After centrifugation at 15,000 $\times$ g for 1 min, the supernatants were incubated with Ni-NTA agarose beads for 16 h at room temperature, after which, the beads were washed twice with 20 bed volumes of lysis buffer, pH 8; twice with 20 bed volumes of lysis buffer, pH 6.3; and once with 20 bed volumes of lysis buffer, pH 6.3, and 10 mM imidazole to remove nonspecifically bound proteins. The proteins were eluted from the Ni-NTA agarose beads using 10 bed volumes of lysis buffer (pH 4.3). The eluate was raised to pH 8 with 1 M Tris (pH 8) and then applied directly to immobilized Streptavidin beads and incubated at 4°C for 16 h. The beads were washed stringently using 10 bed volumes each of buffer A (8 M urea, 0.2 M NaCl, 2% SDS, 100 mM Tris, 5 mM β -mercaptoethanol, pH 7.5) and buffer B (buffer A plus 0.2% SDS) to remove any remaining contaminants. Residual urea and SDS were removed with buffer C (0.2 M NaCl, 100 mM Tris, pH 7.5). A portion (20 μ L) of the beads was resuspended in SDS-PAGE loading buffer for Western blot analysis with the anti-RGSHis antibody. The remaining beads were used for subsequent MS analyses. Controls included HBH-tagged strains without cross-linking and wild type cross-linked strains with 1% formaldehyde.

Trypsin digestion

Washed beads were treated with 40 μ L (10mM) DTT for 1 h at 56°C. After that, 20 μ L of 55 mM iodoacetamide was added to block free thiols. The reaction was allowed to proceed for 30 min in the dark with constant stirring. The beads were then dissolved in 700 μ L of 50 mM NH₃HCO₃, which raised the pH to 7-8, and the proteins digested for 8 h at 37°C with 5 μ g of trypsin (Promega). The digestion was stopped by adding 10 μ L of water/1% TFA, and the supernatant was collected after centrifugation. The beads were then washed twice with 200 μ L of 25% acetonitrile and 0.1% TFA solution, and the washings were combined with the supernatant. The solution was desalted using a C-18 SPE column prior to liquid chromatography tandem mass spectrometric (LC-MS/MS) analysis, and the final peptide concentration was determined using a standard BCA protein/peptide assay method (Pierce).

LC-MS/MS and data analysis

LC-MS/MS was performed with two technical replicates for each sample, utilizing an LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA) equipped with a nano-electrospray ion source. The Orbitrap was utilized to obtain high resolution MS and the low resolution ion-trap was utilized to obtain MS/MS spectra. Peptides were identified by using SEQUEST™ to search the MS/MS spectra from each replicate LC-MS/MS analysis. The

searches were performed using the annotated *Salmonella* fasta data file provided by the Institute of Genomic Research (www.jcvi.org/, September 19, 2004) that contains 4550 protein sequences. A standard parameter file with fixed carbamidomethyl modification on cysteines and variable oxidative modification on methionine residues was considered.

Data were filtered utilizing confident SEQUEST™ filtering criteria, and only fully tryptic peptides were considered for data analysis. Fully tryptic peptides of all charge states with XCorr ≥ 1.5 and $\Delta Cn \geq 0.1$ were considered initially for peptide screening²⁴. The same criteria were utilized to stringently filter the background and non-specific interactions. After that all peptides from the interactor proteins were reevaluated by charge states and XCorrs, utilizing filtering criteria commonly used for ion trap data, i.e., CS1 ≥ 1.5 , CS2 ≥ 2.0 , CS3 ≥ 3.0 . A protein that was exclusively identified in replicate LC-MS/MS analyses of the cross-linked samples, but not in bead and bait control samples was considered as an interacting partner of HimD, PduB, or PhoP. To estimate the FDR for these experiments, the experimental dataset of PhoP (controls and cross-linking experiments) was searched using a decoy database approach²⁵ (forward and reverse sequences of *Salmonella* Typhimurium) and the False Discovery Rate (FDR) was estimated to be ~2%.

As formaldehyde can form non-specific modification in proteins and SEQUEST searches were performed only for unmodified peptides, we also considered every peptide identification in our analysis and did not require multiple peptides per protein per LC-MS/MS analysis. To provide added confidence in identifications, we independently utilized the accurate mass and time (AMT) tag approach²⁶, where cross-linked LC-MS/MS datasets were peak matched with our AMT tag database of *S. Typhimurium*. In addition, mass spectra for single peptide identifications were manually validated and assessed for mass measurement accuracy.

Results

Optimal cross-linking conditions

The HimD protein was used as a model to determine the optimal formaldehyde concentrations for cross-linking. Following cross-linking with 0%, 0.5%, 1%, 2%, and 3% formaldehyde, LS1073 cells that contained the tagged HimD protein were subjected to SDS-PAGE and Western blot analysis. Western analysis showed that IHF β (HimD) migrated at higher molecular weight with increasing concentrations of formaldehyde (Figure 1). Cross-linking was successful at 0.5%, 1%, and 2% formaldehyde; however, cross-linking at the highest concentration (i.e., 3%) was so complete that the proteins did not migrate into the gel. These results are consistent with previously reported observations^{7, 8, 27}.

Formaldehyde is a non-specific cross-linker and at higher concentrations this reagent can produce several non-specific modifications in proteins^{28, 29}. In addition, higher concentrations of formaldehyde reportedly tend to cause more proteins to cross-link with DNA, which reduces the amount of total soluble proteins in the lysate⁷. Therefore, to attenuate more non-specific protein modifications, we opted to use 1% formaldehyde with a relatively short incubation time for *in vivo* cross-linking of all target proteins utilized in this study.

Overall method

In order to generate a list of possible *in vivo* interacting partners for three selected virulent proteins in *S. Typhimurium*, HimD, PhoP and PduB, three different experimental schemes were performed simultaneously for each targeted protein (Figure 2). Both wild type and strains containing HBH tagged proteins were grown under identical conditions.

Two control experiments were employed to distinguish bead and non-specific interacting proteins from interacting partners in the cross-linked sample. The first control was used to

identify the highly specific bead interacting proteins under cross-linking conditions. This control was prepared by treating wild type cells containing no HBH tagged bait with 1% formaldehyde. After the two-step tandem affinity purification (His-Ni and Biotin-Streptavidin) and on bead tryptic digestion, the peptides consisted largely of abundant non-specific bead interacting proteins; this sample was referred to as “Bead control” (Figure 2). The second control sample was used to identify proteins that bound non-specifically to the bait-bead combination. First, cells containing the HBH tagged version of HimD, PhoP and PduB were separated into two equal fractions. In one of these fractions, no formaldehyde cross-linking was performed, and the sample was referred to as “Bait control” (Figure 2). After two step tandem-affinity purification, the peptide list generated for this control sample consisted of peptides from bait proteins, as well as peptides from proteins that bind non-specifically to beads and bait.

The other fraction of the HBH tagged cell sample was used to prepare the cross-linked sample. This fraction was treated with 1% formaldehyde to fix interactions that were constrained by $\sim 2\text{\AA}^{30, 31}$. Referred to as “cross-link”, this tandem affinity purified and digested sample consisted of peptides from 1) bait proteins, 2) proteins that bind non-specifically to beads and bait, and, 3) interacting partners. Proteins that were observed exclusively in the cross-link sample, but not in the bead and bait control samples were identified as potential interacting partners of each HBH tagged targeted protein.

Method refinements using HimD

HimD, a global transcriptional factor that usually forms a heterodimer with HimA, was used for initial method development. This heterodimer plays an important role in regulating gene expressions of many *Salmonella* virulence factors¹².

Three independent experiments, each consisting of replicate LC-MS/MS analyses of cross-link, bead control, and bait control samples (for a total of 18 analyses) were performed to identify the interacting partners of HimD and also to evaluate the reproducibility of the refined method. In one experiment, the cross-link and bead control samples were prepared with 0.5% formaldehyde. In the other two experiments, the cross-link and bead control samples were prepared with 1.0% formaldehyde and served as biological replicates. After purifications and on bead trypsin digestion, replicate LC-MS/MS analyses were performed for all samples, i.e., cross-link, bead and bait controls. Peptides were identified utilizing SEQUESTTM and filtered on the basis of XCorr (≥ 1.5) and ΔCn (≥ 0.1) values, as well as on the basis of whether the N- and C-termini were consistent with tryptic cleavage. For initial peptide screenings, a low XCorr cut off value was used to stringently filter background and non-specific proteins. The MS/MS spectra of interacting partners identified in the cross-link samples were manually investigated and XCorr values were reevaluated (see data analysis section) to remove false positive identifications. A combined protein heat map and heat map of identified interacting partners for HimD experiments were constructed using number of peptides observed for each protein (Figure 3).

Bait HimD was observed in 5 out of 6 cross-link sample analyses, as well as in all bait control analyses. HimA was the only protein observed exclusively in all cross-linked samples, and not in the controls (Figure 3, bottom panel). The interacting partner HimA was observed in low abundances in the cross-linked sample prepared with 0.5% formaldehyde and one of the cross-linked samples prepared with 1% formaldehyde. HimA was also found in high abundances in the other cross-linked sample prepared with 1% formaldehyde. A representative MS/MS spectrum of one peptide from protein HimA (+2, XCorr = 3.14, ΔCn = 0.58), as well as a peptide from bait HimD is shown in Figure 4. The LC-MS/MS datasets of cross-linked HimD samples were peak matched with the AMT tag approach, which independently showed the presence of interacting partner HimA (STM1339) along with the bait HimD (STM0982) (Table

1). The repeated *in vivo* identification of HimA, a known physiological binding partner of HimD from different biological replicates confirms our general approach, and demonstrates the reproducibility of this method.

Method evaluation using PduB and PhoP

To test the broad application of our method we also tagged two other proteins (PduB and PhoP) that are known to be involved in the pathogenic lifestyle of *S. Typhimurium*. The first of these, PduB is involved in 1, 2-propanediol utilization (pdu) and is proposed to have a role in formation of polyhedral organelles (or polyhedral bodies) where degradation of 1,2-propanediol occurs¹⁹. After culturing cells under conditions that increase PduB expression¹⁶, the cells were cross-linked as described above. LC-MS/MS analysis of tandem affinity purified PduB-HBH identified a number of peptides from additional proteins that are also thought to be part of polyhedral bodies. A protein heat map was generated based on the number of peptides identified in each LC-MS/MS dataset of bead, bait and cross-link samples of PduB (Figure 5 A). Note that 10 additional proteins that were not observed in control samples were identified in two technical LC-MS/MS replicate analyses of cross-link samples. Peptide abundance of these interacting partners along with the bait PduB is represented in another heat map (Figure 5 A, bottom panel). All of these proteins were also confidently identified in our AMT tag analysis of PduB cross-linked datasets.

The interacting partners identified for PduB are summarized in Table 2. The list contains two proteins from the Pdu family, PduA and PduJ. An MS/MS of one peptide observed for bait PduB and one interaction partner PduA is shown in Figure 6. While the precise function of PduJ has not been characterized, PduA proteins are the structural proteins that form the shell of polyhedral bodies³². The cross-linking of PduB with PduJ and PduA suggests a structural role for both PduB and PduJ in the formation of polyhedral bodies. Additionally, the cross-linking of PduB with proteins involved in protein synthesis and folding, such as RpsK, FusaA, and MopB, suggests that synthesis of polyhedral bodies, which were very abundant, was likely in progress *in vivo* at the time proteins were cross-linked with formaldehyde. Interactions of PduB with NuoG, GuaB and enolase, which are all involved in respiration, are consistent with the respiratory function of polyhedral bodies.

As part of bacterial two-component signal transduction systems, PhoP is a transcriptional regulator whose regulatory activity is controlled by histidine phosphorylation from the cognate signal sensor. Both PhoP/PhoQ signal transduction systems play critical roles in *Salmonella* pathogenesis^{14, 15, 33, 34}. Figure 5 B shows a protein heat map utilizing the number of peptides identified for each protein in replicate LC-MS/MS analyses of cross-link and control samples of PhoP. Twenty one proteins were unique between the replicate cross-link samples (Figure 5B, bottom panel). After AMT tag analysis, thirteen proteins were identified as potential interactors of PhoP. The interacting partners of PhoP are summarized in Table 3. PhoQ, a known binding partner, was among the identified proteins that were cross-linked to HBH-PhoP. An MS/MS spectrum of one peptide from PhoQ protein, as well as a peptide from the bait PhoP is shown in Figure 7. In addition, Fur and YqjL (two potential binding partners)³⁵ were also found to be cross-linked with HBH-PhoP. Fur is one of the four major regulators that control relevant stress responses in *Salmonella*; the other three are RpoS, PhoPQ, and OmpR/EnvZ³⁶. In addition to Fur, this method also revealed a putative universal stress protein STM1652 as an interacting partner of PhoP. One statistically significant interaction with Hfq, a translational regulator that controls both message stability and translational efficiency via its interaction with small regulatory RNAs³⁷, was also revealed.

Validation of Interactions of PhoP and Hfq

We next tried to validate the novel interaction of PhoP with Hfq. Hfq was tagged with HBH and analyzed using formaldehyde. We were unable to observe tagged Hfq in our bait control experiment, which suggests the tagged version may express poorly. As a result, we performed a global proteomics analysis of *Salmonella* with Hfq deleted from the genome to assess the effect of Hfq on PhoPQ and/or PhoPQ-regulated genes. From our global analysis we observed 22 members of the PhoPQ regulon under the MgM Shock growth condition (Supplementary data). Hfq appeared to negatively regulate the expression of 11 proteins from this set, including PhoP, PhoQ, and the PhoP-activated gene PagC (Figure 8, A, B and Supplementary data). A recent study suggested that Hfq positively regulates the expression of the PhoP-repressed gene PrgI³⁷. Consistent with this report we also observed down regulation of PrgI, PrgH, and PrgK following loss of Hfq under the *Salmonella* pathogenicity island (SPI-1) that induced during LB Log growth condition (Supplementary data). These results suggest a role for Hfq in modulating expression of the PhoPQ regulon, which is consistent with and supports the cross-linking observation that Hfq interacts with PhoP.

Discussion

In this study, we refined the HBH-tag method, which was initially developed to identify protein-protein interactions in *Saccharomyces cerevisiae*^{7, 8}. One of the difficulties typically encountered by this type of affinity purification/MS-based method is the high background caused by nonspecific binding of proteins to the affinity reagents (e.g. beads). This difficulty is partially addressed by using *in vivo* cross-linking to stabilize most interactions under physiological conditions, which permits subsequent purification steps under fully denaturing conditions. These vigorous purification steps significantly reduce nonspecific binding and consequently increase the chance of identifying proteins cross-linked to a tagged bait protein. The use of two controls, “bead” and “bait” for filtering bead and affinity tag induced non-specific interactions further distinguish background identifications, which in turn reduces the number of false-positive interactions. These controls do lead to a concern that some of the relevant interaction partners observed in the bait and wild-type bead controls may be discarded. However, the loss of ambiguous protein interactions may be preferable to identifying a multitude of false interactions over the course of a rapid discovery-based approach, such as used here. Depending on research goals this challenge could be partially mitigated with experimental design and the use of isotopic labeling cell culture methods, such as SILAC³⁸.

The refined method was applied to the pathogenic bacterium *S. Typhimurium* to identify the protein-protein interactions important to *S. Typhimurium* virulence. Each of the three tagged proteins described in this work are required for virulence. The effects on virulence of each knockout are expressed by their LD₅₀ values for each of the disrupted genes HimD, PhoP, and Pdu are >10⁸, >10⁸, and >10² respectively^{16, 39, 40}. The greatest challenge to this work was ascertaining that the fusion protein was still active and that the tag was not located in a region of the protein essential for its interaction with other proteins or with DNA. Although Tagwerker *et al.* reported no effect to cell morphology and/or cell viability when HBH tagged alleles were incorporated into eight different essential yeast genes and emphasized that HBH tag generally did not interfere with protein function,⁷ we could not demonstrate virulence in the case of PhoP construct. Nevertheless, an expected interaction between PhoP and PhoQ was observed. Interestingly, the fact that we observed an interaction between PhoP and PhoQ that is considered to be very transient¹⁴ may be indicative of how the PhoP tagged form is inactivated.

This general approach for identifying protein complexes has been little appreciated or attempted outside of a relatively narrow field^{7, 8, 41}, and is applicable for addressing many other biological questions, but shortcomings remain. Currently, the direct (bait to identified interaction protein) and indirect (bait to intermediate to identified interacting protein)

interactions cannot be distinguished by this method, although technological advances may allow the cross-linked peptides to be identified to provide information on the proximity of interactions. Using formaldehyde as a cross-linking reagent also generates non-specific chemical modifications in several amino acids residues that are not well characterized in the literature. This is one cause of the low yield of unmodified peptides used to identify interactions. Digestion performed on beads with denaturing conditions may also result in poor recoveries. Moreover, on bead digestion also limits our ability to reverse inter cross-linked proteins by boiling. Kast and co-worker recently reported a comprehensive review about the utility of formaldehyde for *in vivo* cross-linking and also mentioned limitations for a formaldehyde cross-linking study⁴².

In this work, due to unknown modification states introduced by formaldehyde, we searched for only fully tryptic unmodified peptides. Both biologically significant and known transient interactions, such as PhoP with PhoQ, were also observed with few peptides. A single peptide from Hfq (average MW 11,133 kDa) was observed in both cross-link samples with high confidence (charge state +2, XCorr = 4.0, $\Delta Cn = 0.4$) as well as PduJ (+2, XCorr = 3.6, $\Delta Cn = 0.5$). We also used the AMT tag approach to independently analyze the data to provide peptide identifications. In the PhoP cross-linked sample, the peptide from Hfq was peak matched with our *S. Typhimurium* AMT tag database. In addition, mass accuracy of the precursor peptides of these interactions fell within confident ranges (i.e., 0.5 to 3.9 ppm without the use of internal calibration standards; Figure 4, 6, and 7).

While expected interactions (e.g., HimA with HimD, PhoP with PhoQ) and several components of polyhedral bodies were observed, our analyses also revealed many new and unexpected protein interactions, such as PhoP with Hfq. PhoP has been previously shown to regulate translation of the gene *ssrA*, but the mechanism is not known. The observation that Hfq interacts with PhoP hints at involvement of Hfq in this regulation. In combination, our cross-linking and global proteomics observations demonstrate a functional role for Hfq in PhoPQ mediated processes. This interaction between PhoP and Hfq might provide a mechanistic explanation of how PhoP post-transcriptionally regulates the protein SsrA¹³ and provide insight into the regulation of *Salmonella* pathogenesis in general.

Conclusion

The reported tandem affinity purification and formaldehyde cross-linking method proved useful for identifying *in vivo* protein-protein interactions. Cross-linking followed by tandem affinity purification under denaturing conditions significantly reduced background proteins and in some cases eliminated nonspecific binding proteins to a given target protein. Additionally, the use of dual control strategies that mimicked cross-linking and non-cross-linking conditions allowed background and non-specific interactions to be distinguished from interacting partners. The reproducibility achieved for identifying known interactions *in vivo* is also indicative that this refined method holds promise for identifying *bona fide* interacting partners. Application of this method to the pathogen *S. Typhimurium* resulted in identification of *in vivo* protein interacting partners for three proteins associated with virulence, including both known and novel interactors. These known and novel interacting proteins provide new insights into the dynamic processes of *S. Typhimurium* virulence.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We gratefully acknowledge the contributions of Therese RW. Clauss, Brianne O. Petritis, Karl K. Weitz, Nicola Tolić, Samuel O. Purvine, Penny Colton, and Drs. Xiuxia Du, Joshua Turse, Ashoka D. Polpitiya, Matthew E. Monroe and Joseph N. Brown for discussions input and suggestions in preparing this publication. We also thank Dr. Peter Kaiser at University of California, Irvine for providing pFA6a-HBH-kanMX6.

Portions of this work were supported by the National Institute of Allergy and Infectious Diseases NIH/DHHS through interagency agreement Y1-AI-4894-01, National Center for Research Resources (NCRR) grant no. RR18522, and the Laboratory Directed Research and Development program at PNNL. Significant portions of this work were performed in the Environmental Molecular Science Laboratory, a United States Department of Energy (DOE) national scientific user facility at Pacific Northwest National Laboratory (PNNL) in Richland, WA. PNNL is operated for the DOE Battelle Memorial Institute under contract DE-AC05-76RLO01830.

References

1. Gingras AC, Aebersold R, Raught B. Advances in protein complex analysis using mass spectrometry. *J Physiol* 2005;563(Pt 1):11–21. [PubMed: 15611014]
2. Tucker CL, Gera JF, Uetz P. Towards an understanding of complex protein networks. *Trends Cell Biol* 2001;11(3):102–6. [PubMed: 11306254]
3. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, Qureshi-Emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadamar G, Yang M, Johnston M, Fields S, Rothberg JM. A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 2000;403(6770):623–7. [PubMed: 10688190]
4. Ransone LJ. Detection of protein-protein interactions by coimmunoprecipitation and dimerization. *Methods Enzymol* 1995;254:491–7. [PubMed: 8531709]
5. Chang IF. Mass spectrometry-based proteomic analysis of the epitope-tag affinity purified protein complexes in eukaryotes. *Proteomics* 2006;6(23):6158–66. [PubMed: 17072909]
6. Selbach M, Mann M. Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). *Nat Methods* 2006;3(12):981–3. [PubMed: 17072306]
7. Tagwerker C, Flick K, Cui M, Guerrero C, Dou Y, Auer B, Baldi P, Huang L, Kaiser P. A tandem affinity tag for two-step purification under fully denaturing conditions: application in ubiquitin profiling and protein complex identification combined with in vivocross-linking. *Mol Cell Proteomics* 2006;5(4):737–48. [PubMed: 16432255]
8. Guerrero C, Tagwerker C, Kaiser P, Huang L. An integrated mass spectrometry-based proteomic approach: quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26 S proteasome-interacting network. *Mol Cell Proteomics* 2006;5(2):366–78. [PubMed: 16284124]
9. Ansong C, Yoon H, Norbeck AD, Gustin JK, McDermott JE, Mottaz HM, Rue J, Adkins JN, Heffron F, Smith RD. Proteomics analysis of the causative agent of typhoid fever. *J Proteome Res* 2008;7(2):546–57. [PubMed: 18166006]
10. Chalker RB, Blaser MJ. A review of human salmonellosis: III. Magnitude of *Salmonella* infection in the United States. *Rev Infect Dis* 1988;10(1):111–24. [PubMed: 2832925]
11. Shi L, Adkins JN, Coleman JR, Schepmoes AA, Dohnkova A, Mottaz HM, Norbeck AD, Purvine SO, Manes NP, Smallwood HS, Wang H, Forbes J, Gros P, Uzzau S, Rodland KD, Heffron F, Smith RD, Squier TC. Proteomic analysis of *Salmonella enterica* serovar typhimurium isolated from RAW 264.7 macrophages: identification of a novel protein that contributes to the replication of serovar typhimurium inside macrophages. *J Biol Chem* 2006;281(39):29131–40. [PubMed: 16893888]
12. Mangan MW, Lucchini S, Danino V, Croinin TO, Hinton JC, Dorman CJ. The integration host factor (IHF) integrates stationary-phase and virulence gene expression in *Salmonella enterica* serovar Typhimurium. *Mol Microbiol* 2006;59(6):1831–47. [PubMed: 16553887]
13. Bijlsma JJ, Groisman EA. The PhoP/PhoQ system controls the intramacrophage type three secretion system of *Salmonella enterica*. *Mol Microbiol* 2005;57(1):85–96. [PubMed: 15948951]
14. Miller SI. PhoP/PhoQ: macrophage-specific modulators of *Salmonella* virulence? *Mol Microbiol* 1991;5(9):2073–8. [PubMed: 1766380]

15. van Velkinburgh JC, Gunn JS. PhoP-PhoQ-regulated loci are required for enhanced bile resistance in *Salmonella* spp. *Infect Immun* 1999;67(4):1614–22. [PubMed: 10084994]
16. Adkins JN, Mottaz HM, Norbeck AD, Gustin JK, Rue J, Clauss TR, Purvine SO, Rodland KD, Heffron F, Smith RD. Analysis of the *Salmonella typhimurium* proteome through environmental response toward infectious conditions. *Mol Cell Proteomics* 2006;5(8):1450–61. [PubMed: 16684765]
17. Walter D, Ailion M, Roth J. Genetic characterization of the pdu operon: use of 1,2-propanediol in *Salmonella typhimurium*. *J Bacteriol* 1997;179(4):1013–22. [PubMed: 9023178]
18. Conner CP, Heithoff DM, Julio SM, Sinsheimer RL, Mahan MJ. Differential patterns of acquired virulence genes distinguish *Salmonella* strains. *Proc Natl Acad Sci U S A* 1998;95(8):4641–5. [PubMed: 9539791]
19. Bobik TA, Havemann GD, Busch RJ, Williams DS, Aldrich HC. The propanediol utilization (pdu) operon of *Salmonella enterica* serovar Typhimurium LT2 includes genes necessary for formation of polyhedral organelles involved in coenzyme B(12)-dependent 1, 2-propanediol degradation. *J Bacteriol* 1999;181(19):5967–75. [PubMed: 10498708]
20. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 2000;97(12):6640–5. [PubMed: 10829079]
21. Uzzau S, Figueroa-Bossi N, Rubino S, Bossi L. Epitope tagging of chromosomal genes in *Salmonella*. *Proc Natl Acad Sci U S A* 2001;98(26):15264–9. [PubMed: 11742086]
22. Tagwerker C, Zhang H, Wang X, Larsen LS, Lathrop RH, Hatfield GW, Auer B, Huang L, Kaiser P. HB tag modules for PCR-based gene tagging and tandem affinity purification in *Saccharomyces cerevisiae*. *Yeast* 2006;23(8):623–32. [PubMed: 16823883]
23. Wang RF, Kushner SR. Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* 1991;100:195–9. [PubMed: 2055470]
24. Strittmatter EF, Kangas LJ, Petritis K, Mottaz HM, Anderson GA, Shen Y, Jacobs JM, Camp DG 2nd, Smith RD. Application of peptide LC retention time information in a discriminant function for peptide identification by tandem mass spectrometry. *J Proteome Res* 2004;3(4):760–9. [PubMed: 15359729]
25. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* 2007;4(3):207–14. [PubMed: 17327847]
26. Conrads TP, Anderson GA, Veenstra TD, Pasa-Tolic L, Smith RD. Utility of accurate mass tags for proteome-wide protein identification. *Anal Chem* 2000;72(14):3349–54. [PubMed: 10939410]
27. Schmitt-Ulms G, Legname G, Baldwin MA, Ball HL, Bradon N, Bosque PJ, Crossin KL, Edelman GM, DeArmond SJ, Cohen FE, Prusiner SB. Binding of neural cell adhesion molecules (N-CAMs) to the cellular prion protein. *J Mol Biol* 2001;314(5):1209–25. [PubMed: 11743735]
28. Metz B, Kersten GF, Baart GJ, de Jong A, Meiring H, ten Hove J, van Steenberg MJ, Hennink WE, Crommelin DJ, Jiskoot W. Identification of formaldehyde-induced modifications in proteins: reactions with insulin. *Bioconjug Chem* 2006;17(3):815–22. [PubMed: 16704222]
29. Metz B, Kersten GF, Hoogerhout P, Brugghe HF, Timmermans HA, de Jong A, Meiring H, ten Hove J, Hennink WE, Crommelin DJ, Jiskoot W. Identification of formaldehyde-induced modifications in proteins: reactions with model peptides. *J Biol Chem* 2004;279(8):6235–43. [PubMed: 14638685]
30. Orlando V, Strutt H, Paro R. Analysis of chromatin structure by in vivo formaldehyde cross-linking. *Methods* 1997;11(2):205–14. [PubMed: 8993033]
31. Vasilescu J, Figeys D. Mapping protein-protein interactions by mass spectrometry. *Curr Opin Biotechnol* 2006;17(4):394–9. [PubMed: 16822661]
32. Havemann GD, Sampson EM, Bobik TA. PduA is a shell protein of polyhedral organelles involved in coenzyme B(12)-dependent degradation of 1,2-propanediol in *Salmonella enterica* serovar typhimurium LT2. *J Bacteriol* 2002;184(5):1253–61. [PubMed: 11844753]
33. Groisman EA, Chiao E, Lipps CJ, Heffron F. *Salmonella typhimurium* phoP virulence gene is a transcriptional regulator. *Proc Natl Acad Sci U S A* 1989;86(18):7077–81. [PubMed: 2674945]
34. Groisman EA, Heffron F, Solomon F. Molecular genetic analysis of the *Escherichia coli* phoP locus. *J Bacteriol* 1992;174(2):486–91. [PubMed: 1530848]
35. Worley MJ, Nieman GS, Geddes K, Heffron F. *Salmonella typhimurium* disseminates within its host by manipulating the motility of infected cells. *Proc Natl Acad Sci U S A* 2006;103(47):17915–20. [PubMed: 17095609]

36. Rychlik I, Barrow PA. Salmonella stress management and its relevance to behaviour during intestinal colonisation and infection. *FEMS Microbiol Rev* 2005;29(5):1021–40. [PubMed: 16023758]
37. Sittka A, Pfeiffer V, Tedin K, Vogel J. The RNA chaperone Hfq is essential for the virulence of *Salmonella typhimurium*. *Mol Microbiol* 2007;63(1):193–217. [PubMed: 17163975]
38. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 2002;1(5):376–86. [PubMed: 12118079]
39. Galan JE, Curtiss R 3rd. Virulence and vaccine potential of phoP mutants of *Salmonella typhimurium*. *Microb Pathog* 1989;6(6):433–43. [PubMed: 2671582]
40. Marshall DG, Sheehan BJ, Dorman CJ. A role for the leucine-responsive regulatory protein and integration host factor in the regulation of the *Salmonella* plasmid virulence (spv) locus in *Salmonella typhimurium*. *Mol Microbiol* 1999;34(1):134–45. [PubMed: 10540292]
41. Wang X, Chen CF, Baker PR, Chen PL, Kaiser P, Huang L. Mass spectrometric characterization of the affinity-purified human 26S proteasome complex. *Biochemistry* 2007;46(11):3553–65. [PubMed: 17323924]
42. Sutherland BW, Toews J, Kast J. Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. *J Mass Spectrom* 2008;43(6):699–715. [PubMed: 18438963]

Figure 1. Cross-linking and purification of HimD-HBH. **I.** Western-blot analysis of HimD-HBH cross-linked with different (0.5-3%) formaldehyde concentrations. S; standard. **II.** Western-blot analysis of purified HimD-HBH. A: no bait/cross-linking control, B: bait/no cross-linking control; C: bait/cross-linking with 1% formaldehyde.

Figure 2.

Experimental scheme for confident identification of protein interaction partners. Bead control - wild type *Salmonella* cell after 1% formaldehyde cross-linking condition. Bait control - HBH tagged bait only control but no formaldehyde cross-linking. Cross-link - HBH tagged bait and formaldehyde cross-linking condition.

Figure 3.

Protein-protein interaction studies of HimD protein utilizing three independent biological experiments (one 0.5 and two 1% formaldehyde concentrations). Top panel- complete protein heat map generated utilizing number of peptides observed for each protein in 18 LC-MS/MS datasets (3 biological samples $\times$ 3 experiments [two controls and one cross-link experiment for each] $\times$ 2 technical LC-MS/MS replicates). Bottom panel - heat map of interacting partner of HimD (STM0982) from three biological experiments. The abundances of baits are shown in comparison in the bottom. A1, A2 - bead controls (0.5% formaldehyde), B1, B2, E1, E2, H1, H2- bait controls (0% formaldehyde), C1, C2 (Cross-link, 0.5% formaldehyde), D1, D2, G1, G2 (same bead controls, 1% formaldehyde), F1, F2, I1, I2 (cross-link, 1% formaldehyde).

Peptide abundances are represented with color scale limit (0-2, green to red). See the tables and supplementary data for original peptide counts.

Figure 4.

Mass spectra of HimD (STM0982) and its interacting partner HimA. A - Mass spectrum of one peptide observed from the bait protein HimD (STM0982). B - Mass spectrum of one peptide observed from its interacting partner HimA (STM1339).

Figure 5.

Heat map illustrating the number of peptides observed for each protein in all LC-MS/MS data sets for A) PduB (STM2039) and B) PhoP (STM1231) proteins. Top panels: All peptides that passed SEQUEST filtering criteria in 6 LC-MS/MS datasets for each bait [a biological sample $\times$ 3 experiments (two controls and one cross-link experiment for each) $\times$ 2 technical LC-MS/MS replicates]. Bottom panels: putative protein interactors following removal of proteins observed in control samples. Peptide abundances are represented with color scale limit (0-2, green to red). See the tables and supplementary data for original peptide counts.

Figure 6.

Mass spectra of PduB (STM2039) and one of its interacting partners. A - Mass spectrum of one peptide observed from the bait protein PduB (STM2039). B - Mass spectrum of one peptide observed from its interacting partner PduA (STM2038).

Figure 7.

Mass spectra of PhoP (STM1231) and one of its interacting partners. A - Mass spectrum of one peptide observed from the bait protein PhoP (STM1231). B - Mass spectrum of one peptide observed from its interacting partner PhoQ (STM1230). C - Mass spectrum of one peptide from Hfq (STM4361).

Figure 8. Global proteomics study used to demonstrate a functional role of PhoP (STM1231) with Hfq using AMT (accurate mass and time) tag analysis from 3 biological replicate samples. A) Heat map of abundances of peptide from 11 protein of PhoPQ regulon that regulated after deleting hfq gene. B) Peptide abundances from PhoP are shown exclusively in second heat map. Three biological replicate – in an infectious mimicking condition “Shock” with an hfq deletion is on the left and wild type on the right.

Table 1

List of protein identified as interacting partners of HimD. Peptides exclusively observed in cross-linked dataset are considered as interacting partners. Total number of peptide observed in SEQUEST and AMT tag analysis of bait HimD and interactor HimA is shown. Additional details including peptide sequences and their XCorr's are provided in the supplementary table 3.

Proteins	Count of peptides	Protein descriptions
STM0982 (Bait)	5	integration host factor (IHF), beta subunit site-specific (himD)
STM1339	5	integration host factor (IHF), alpha subunit (himA)

Table 2

List of proteins identified as interacting partners along with the bait PduB (STM2039). Number of peptides observed from each protein by SEQUEST and peak matching with AMT tag database of *S. Typhimurium* is presented. Additional details are provided in supplementary table 4.

Proteins	Count of peptides	Protein descriptions
STM1730	3	putative cytoplasmic protein (yciE)
STM2038	2	Propanediol utilization: polyhedral bodies (pduA)
STM2039 (Bait)	10	Propanediol utilization: polyhedral bodies (pudB)
STM2045	1	Propanediol utilization: polyhedral bodies (pduJ)
STM2323	3	NADH dehydrogenase I chain G (nuoG)
STM2511	1	IMP dehydrogenase (guaB)
STM2885	2	cell invasion protein (sipB)
STM2952	7	enolase (eno)
STM3417	5	30S ribosomal subunit protein S11 (rpsK)
STM3446	7	protein chain elongation factor EF-G, GTP-binding (fusA)
STM4329	2	chaperone Hsp10, affects cell division (mopB)

Table 3

List of protein identified as interacting partners of PhoP. Proteins exclusively observed in cross-linked dataset are considered as interacting partners. Total number of peptide observed in SEQUEST and AMT tag analysis of bait PhoP and its interactors are shown. Additional details are provides in supplementary table 5.

Protein Name	Count of Peptides	Protein Descriptions
STM0693	1	transcriptional repressor of iron-responsive genes (Fur family) (fur)
STM0730	2	citrate synthase (gltA)
STM1230	1	sensory kinase protein in two-component regulatory system (phoQ)
STM1231 Bait	22	response regulator in two-component regulatory system with (phoP)
STM2081	1	gluconate-6-phosphate dehydrogenase, decarboxylating (gnd)
STM2555	1	serine hydroxymethyltransferase (glyA)
STM2771	4	Flagellar synthesis: phase 2 flagellin (filament structural (fljB)
STM3068	1	fructose-bisphosphate aldolase (fba)
STM3215	2	putative transcriptional regulator (yqjI)
STM3417	3	30S ribosomal subunit protein S11 (rpsK)
STM3418	2	30S ribosomal subunit protein S13 (rpsM)
STM3441	2	30S ribosomal subunit protein S10 (rpsJ)
STM3983	1	3-hydroxyacyl-coA dehydrogenase (EC 1.1.1.35) of 4-enzyme FadB (fadB)
STM4361	1	host factor I for bacteriophage Q beta (hfq)