THE ZINC REGULATED ANTIVIRULENCE PATHWAY OF SALMONELLA IS A MULTI-PROTEIN IMMUNOGLOBULIN ADHESION SYSTEM

Gerd Prehna1,2, Yuling Li3,4, Nikolay Stoynov5, Mark Okon3,6, Marija Vukovic1,2, Lawrence P. McIntosh1,6, Leonard J. Foster5, B. Brett Finlay3,4, and Natalie C.J. Strynadka1,2.

1Department of Biochemistry, Life Sciences Center, 2Center for Blood Research, 3Michael Smith Laboratories, 4Department of Microbiology and Immunology, 5Centre for High-Throughput Biology and Department of Biochemistry & Molecular Biology, and 6Department of Chemistry. The University of British Columbia, Vancouver BC, Canada

*Running Title: Immunoglobulin Protein Adhesion in Salmonella antivirulence

Address correspondence to: N. C.J. Strynadka, Dept. Biochemistry, University of British Columbia, Vancouver BC, Fax: +1-604-822-0789, E-mail: natalie@byron.biochem.ubc.ca

Keywords: Antivirulence, Salmonella, ZirS, ZirT, NMR, SILAC

Background: Salmonella utilize ZirTS to decrease virulence for increased host transmission.

Results: ZirS is an immunoglobulin fold that binds the previously unknown zir protein ZirU; and both interact with the extracellular domains of ZirT.

Conclusions: The zir antivirulence pathway is an immunoglobulin adhesion system that assembles in the extracellular environment.

Significance: Understanding antivirulence completes the picture of host-microbe interactions for a full model of pathogenesis.

SUMMARY

The co-evolutionary relationship between pathogen and host has led to a regulatory cycle between virulence factors needed for survival and antivirulence factors required for host transmission. This is exemplified in Salmonella spp. by the zirTS antivirulence genes; a secretion pathway comprised of the outer membrane transporter ZirT, and its secreted partner, ZirS. ZirTS act within the gastrointestinal tract to function as a virulence modulator and during Salmonella shedding in anticipation of a new host. Together, ZirT and ZirS decrease virulence by lowering bacterial colonization at systemic sites through an unknown mechanism. To understand this mechanism, we have probed the zirTS pathway both structurally and biochemically. The NMR derived structural ensemble of the C-terminal domain of ZirS reveals an immunoglobulin super family fold (IgSF). Stable isotope labeling by amino acids in cell culture experiments show that the ZirS IgSF domain interacts with its transporter ZirT, and reveal a new protein interaction partner of the pathway, a protein encoded adjacent to zirTS that we have designated as ZirU. ZirU is secreted by ZirT and is also a predicted IgSF. Biochemical analysis delineates ZirT into an N-terminal porin domain and C-terminal extracellular soluble IgSF domains, while biophysical characterization suggests that the transporter undergoes self-association in a concentration dependent manner. We observe that ZirS and ZirU directly interact with each other and with the extracellular domains of ZirT. Here we show that the zir antivirulence pathway is a multi-protein immunoglobulin adhesion system consisting of a complex interplay between ZirS, ZirT, and ZirU.

Pathogenic microorganisms have the dual challenge of both adapting to survive within the environment of their host and mediating their transmission to new hosts. Such organisms employ a number of virulence genes, which are often encoded together on a localized genetic segment or pathogenicity island (1,2). These factors are readily identifiable as their disruption...
reduces or abolishes the severity of the disease phenotype of the infecting species. Prominent examples include the type III secretion system (T3SS), found in Gram-negative bacteria such as *Salmonella* and *Yersinia* (3), and the cytotoxin associated genes (cag) that constitute the type IV secretion system (T4SS) in the Gram-negative bacteria *Helicobacter pylori* (4). In both cases these genes encode proteins that interact directly within eukaryotic cells, gained through co-evolution with their mammalian hosts (4,5). These virulence factors frequently mimic the function of host proteins to take control of cellular signaling events (6,7) allowing the bacteria to evade the immune system and replicate.

Contingent with virulence factors, pathogens also employ genes that encode for antivirulence function(s). Similar to virulence proteins, antivirulence proteins can be identified by their disruption, but instead result in a more severe disease or hypervirulence phenotype. These systems necessarily evolve with their virulence counterparts to facilitate the transmission to a new host by prolonging host survival (8-10), or are critical for survival in another part of the pathogen life-cycle (11). Although as a field antivirulence is relatively new, examples are readily appearing with the availability of pathogen genomes and bioinformatic techniques (8). Antivirulence proteins have been characterized in *Leishmania* spp., protozoan parasites that cause skin lesions and eventual death (10), and in *Salmonella* spp., a Gram-negative enteropathogenic bacteria that is a common cause of gastroenteritis and the causative agent of typhoid fever (12-15). In each case, the antivirulence proteins act to decrease population growth of the pathogen during specific points of infection, effectively increasing the levels of host survival required to allow for efficient transmission. In *Salmonella* spp., two characterized examples, *sciS* (12) and the *zirTS* pathway (13), highlight the pathogenic importance of antivirulence.

Upon ingestion of contaminated food or drinking water, *Salmonella enterica* serovar Typhimurium (*S. typhimurium*) localize to the intestinal mucosa and phagocytic cells of their human hosts where they internalize and replicate by virtue of their T3SS systems (3,14). Here the bacteria cause a self-limiting gastroenteritis, with diarrheal symptoms resulting when newly replicated *S. typhimurium* induce shedding of the intestinal mucosal lining. This ultimately leads to the release of the pathogen into the environment for transmission to new hosts (16). The gene *sciS* has been shown to be intimately involved in these replication events, where it regulates *Salmonella* division within macrophages (12). The deletion of *sciS* results in an increase of bacterium in macrophages infected *in vitro*, and an eight-fold increase of lethality in mice infected intragastrically (12). Together this suggests that *sciS* attenuates virulence at the late stages of infection (12).

The *zirTS* pathway also regulates virulence by attenuating population growth, but instead influences the early stages of infection (13). Deletion of *zirTS* genes results in a hypervirulence phenotype, with the median survival time of BALB/c mice dropping from 8.5 days to 6 days, and wild-type *Salmonella* outcompeting deletion strains by 4-fold as observed by competitive index (CI) experiments (13). Additionally, it was found that *zirTS* is expressed throughout the gastrointestinal tract rather than at systemic sites, and is highly expressed during bacterial shedding in fecal pellets at 8 weeks post infection (13). Furthermore, the system was shown to be negatively regulated by OxyR, a transcription factor that controls the expression of genes in response to oxidative stress (17), and also by the presence of zinc in the media. As *Salmonella* spp are primarily transmitted by the ingestion of contaminated fecal matter, *ZirTS* is expressed in the anticipation of infection of a new host to regulate bacterial virulence upon colonization of the intestinal tract (13).

The *zirTS* genes are located on a horizontally acquired genomic island that is conserved in pathogenic strains of *Salmonella* (13,18). *ZirS* (*Zinc regulated Secreted*) is a 31 kDa protein of unknown function that is secreted by *ZirT* (*Zinc regulated Transporter*), a 72 kDa outer-membrane (OM) β domain (19) with sequence homology to Invasin/Intimin proteins involved in bacterial-host adhesion (13). *ZirS* and *ZirT* are predicted to constitute a secretion system with aspects reminiscent of type V two-partner secretion systems (TPS) (20). However, *ZirS* lacks the N-terminal TPS domain typically
necessary for targeting of the secreted component to the cognate transporter in other type V systems. Additionally, ZirS has sequence homology (24\% identity) with a putative human zinc finger protein (NP_065798), and the open reading frame adjacent to ZirS, (STM1667) is a predicted thioperoxidase (13). In light of the negative regulation of the pathway by OxyR and zinc, these previous observations suggested a potential requirement for specific environmental regulatory cues for ZirTS function.

To elucidate the molecular details governing the mechanism of action and regulation of the zirTS antivirulence pathway in Salmonella, we have undertaken a structural approach using NMR spectroscopy and computational modeling, coupled with biochemical and biophysical analysis. We characterize features of the known components of the zirTS pathway, ZirS and ZirT, and identify an additional protein component that directly interacts with the system, named herein as ZirU. Our results show that this secretion pathway is completely divergent from type V or other secretion, describes the zir antivirulence pathway as a multi-protein immunoglobulin adhesion system, and offers mechanistic insight into an antivirulence mechanism with a new role for the Ig-fold in bacterial pathogenesis.

EXPERIMENTAL PROCEDURES

**Expression and Purification of ZirS-IgSF, ZirT-IgSF, and ZirU.** ZirS (136-276) was expressed as an N-terminal GST-tagged protein in pGEX6p1 (GE Healthcare). Isotopically labeled samples were prepared with E. coli BL21(DE3) grown at 37\°C in M9 minimal medium containing 1 g/L 15NH4Cl and/or 3 g/L 13C6-glucose. At OD600 >0.6 the cells were induced with 1mM IPTG and the temperature lowered to 20\°C. After 20 hrs the cells were harvested, lysed by an Avestin homogenizer and the lysate cleared by centrifugation. GST-ZirS was purified by passing the lysate over Q-sepharose followed by Glutathione-sepharose at 4\°C. The material was then digested O/N during dialysis with 3C protease and GST removed using a mono-Q column. The ZirS material was concentrated and further purified by gel filtration using an SD75 column (GE Healthcare) equilibrated in 50 mM Tris pH 7.1, 50 mM NaCl, 0.1 mM TCEP and 0.1 mM PMSF. All steps were carried out at 4\°C. ZirT (430-660) and ZirU (38-281) were purified as N-terminally GST tagged constructs as per ZirS, but GST was not removed. The final buffer was 50mM Tris pH 7.5, 100 mM NaCl, 0.1 mM TCEP.

**Expression and Purification of ZirT.** ZirT (1-660) was first cloned into pET52b (Novagen) then subcloned with the his-tag and thrombin site into pBAD. The vector was transformed into S. typhimurium lacking InvG and ZirT expression induced with 0.1\% Arabinose O/N at 20\°C. Cells were harvested by centrifugation, lysed with an Avestin homogenizer, and inclusion bodies removed at 7000 rpm for 20 minutes. Membranes were harvested by ultracentrifugation at 100,000xg for 1hr and resuspended in 50mM HEPES pH 7.5 200 mM NaCl 10\% glycerol 0.1 mM PMSF 0.1 mM TCEP. ZirT was extracted from membranes in the same buffer plus 1\% LDAO (Anatrace) and unextracted material removed by ultracentrifugation. ZirT was purified by Nickel sepharose in buffer with 0.2\% LDAO followed by gel-filtration with Superose-6 in buffer plus 0.8\% βOG (Anatrace).

**NMR Spectroscopy** NMR spectra were recorded at 25\°C or 35\°C on Varian Unity 500 and cryoprobe-equipped Inova 600 MHz spectrometers. Final ZirS samples were composed of 0.5 mM protein in 50 mM Tris pH 7.1, 50 mM NaCl, 0.1 mM TCEP and 0.1 mM PMSF with the addition of 5\% D2O. Signals from the 1H, 13C, and 15N nuclei of ZirS were detected and assigned by standard heteronuclear NMR experiments (21). 15N NOE relaxation data for ZirS were recorded using a 600 MHz spectrometer at 20\°C following established methods (22). The NMR data were processed with NMRpipe (23) and analyzed with Sparky (Goddard, T.D. and Kneller, D.G., Sparky 3, University of San Francisco). The chemical shift assignments of ZirS have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu/) under accession code XXXX).

**Structure Calculation** The ZirS structural ensemble was calculated using CYANA, as summarized in Table 1 (24,25). Backbone dihedral angles were predicted from chemical shifts by TALOS+ (26). Distance restraints were derived from a 100 msec mixing time simultaneous 3D 13C/15N-edited NOESY-HSQC recorded with a 600
MHz spectrometer at 25°C (27) and assigned automatically. Hydrogen bond restraints were based on diagnostic NOE patterns. The calculation consisted of 7 cycles of refinement with 100 iterations per cycle (25). The 10 structures with the lowest calculated target function were chosen as a representative ensemble and refined in water using CNS (28,29). The coordinates of the ZirS ensemble have been deposited in the Research Collaboratory for Structural Bioinformatics Protein Data Bank (XXXX).

**Computational Modeling** Structural models for the ZirS N-terminal domain were calculated ab initio using the server Robetta (30) and manually with Rosetta (31) imposing a disulfide restraint. Residues 44-146 of ZirS were chosen to model due to secondary structure predictions with Jpred (32). The forced disulfide constraint resulted in poorly-packed folds consisting of isolated β-hairpins, whereas the Robetta server yielded folded domains which were analyzed by the Dali-server (33). ZirT (430-660) was modeled by the server Phyre2 (34) and with the program Modeller (35). Both constructed highly similar homology models with Invasin (pdb 1cwv) as the template.

**β Domain Refolding** ZirT porin domain constructs were overexpressed as C-terminal His-tagged constructs in the vector pET52b and purified as inclusion bodies. A concentration of 0.15 mg/mL purified material was resuspended in 20 mM potassium phosphate pH 8 and 2% SDS at room temperature. Refolding was induced with the addition of 1.5 M 2-methyl-2,4-pentanediol (MPD) and monitored by SDS-page gel. Secondary structure was measured by circular dichroism spectroscopy after concentration to 4.2 mg/mL (Jasco J-810). The CD data was de-convoluted using Dichroweb (36) and compared to Jpred (32) secondary structure predictions.

**ZirT Topology Assay** ZirT was expressed as described above. After 1 hr of induction at 37°C, 1 mL of cells were harvested washed 3x in PBS to remove Arabinose, and then resuspended in 500 µL. Thrombin was added at a total of 0 µg, 8 µg, or 80 µg and incubated for 1 hr or 2 hrs. The reaction was stopped with the addition of PBS and 1 mM PMSF and washed 3 times in the same buffer. The cells were resuspended in a final volume of 300 µL lysed by sonication and membranes harvested. Samples were analyzed by western-blot with anti-histag (Cedarlane) and developed by Immobilon Western Chemiluminescent HRP substrate (Millipore). Quantification was performed with Genesnap Chemiluminescent HRP substrate (Millipore). For samples containing ZirT, ZirT was added in trans on the vector pBR322. Expression and secretion were monitored in the pellet and in the media by western blot with anti-HA and anti-DnaK antibody and goat-anti-mouse-HRP secondary antibody for detection. All media fractions were sterile filtered before visualization.

**Metal Binding Analysis** The presence of metals was tested for by inductively coupled plasma mass spectrometry (ICP-MS) with a NexION 3000 ICP mass spectrometer (Perkin Elmer) and analyzed by the NexION software program. GST was used as a negative control and the zinc binding New Delhi β-lactamase (kindly supplied by Dustin King) as a positive control. Purified GST-ZirS (136-276), GST-ZirU (38-281) and ZirT-histag (430-660) were assayed. All proteins were diluted to 5 µM in 20 mM HEPES pH 7.5 100mM NaCl then diluted 1/10 in an internal standard (10 µg/L Sc⁴⁺, 1% nitric acid, Inorganic Ventures). Each sample was measured three times and the average abundance recorded.

**SILAC Salmonella** and HeLa cells were grown in minimal media supplemented with isotopically labeled Arginine (¹³C₆) and Lysine (D₈) for “heavy” extracts and unlabeled amino acids for “light” extracts. For bait, 40pmol of
GST-ZirS IgSF was bound to glutathione beads and an equivalent molar ratio of GST alone as the control. The co-precipitation was performed according to standard methods (37) and proteins were digested with endopeptidase LysC and trypsin (mass-spec grade, Promega) as described (PMID 12724530). Digested peptides were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) on a LTQ-OrbitrapXL (ThermoFisher, Bremen, Germany) on-line coupled to an Agilent 1100 Series nanoflow HPLC using a nanospray ionization source (Proxeon Biosystems) holding columns packed into 15-cm-long, 75-μm-inner diameter fused silica emitters (8-μm-diameter opening, pulled on a P-2000 laser puller from Sutter Instruments) using 3-μm-diameter ReproSil Pur C18 beads. Buffer A consisted of 0.5% acetic acid, and buffer B consisted of 0.5% acetic acid and 80% acetonitrile. Gradients were run from 6% B to 30% B over 60 min, then 30% B to 80% B in the next 10 min, held at 80% B for 5 min, and then dropped to 6% B for another 15 min to recondition the column. The LTQ-OrbitrapXL was set to acquire a full-range scan at 60,000 resolution from 350 to 1500 Th in the Orbitrap and to simultaneously fragment the top five peptide ions in each cycle in the LTQ.

Protein identification and quantification were done using Proteome Discover (v.1.2, ThermoFisher, Bremen, Germany) and Mascot (v2.3, Matrix Science) to search against the International Protein Index (IPI) Human (v3.69, 149,771 sequences - common serum contaminants added and all reversed sequences were concatenated) database with the following criteria: electrospray ionization-ion trap fragmentation characteristics, tryptic specificity with up to one missed cleavages; ±10 parts-per-million and ±0.6 Da accuracy for MS and MS/MS measurements respectively; cysteine carbamidomethylation as a fixed modification; N-terminal protein acetylation, methionine oxidation, duplex (3H2-Lys, 13C6-Arg) SILAC modifications as appropriate; peptide false discovery rate (FDR) was set at 1%. Quantitation was done using a mass precision of 2 ppm (three times the mass precision is used to create extracted ion chromatograms). After extracting each ion chromatogram, Proteome Discoverer runs several filters to check for, among other things, interfering peaks and the expected isotope pattern, and peptides that do not meet all the criteria are not used in calculating the final ratio for each protein. Analytical variability of SILAC data in the types of experiments performed here is typically <30% on average and biological variability was addressed in these experiments by performing at least three independent replicates of each experiment.

**Far Western** Purified bait proteins were run on an SDS-page gel and transferred to nitrocellulose. Full-length ZirS-HA and ZirU-HA were cloned as previously (13) and expressed in *S. typhimurium* strain SL1344. Either solubilized pellet or media containing ZirS-HA or ZirU-HA was used to blot the nitrocellulose matrix containing bait proteins. For controls, no HA-labeled protein was added. After binding the matrix was washed at least three times then incubated with anti-HA antibody (Cedar lane) followed by three more washes to remove unbound anti-HA antibody. In all steps the buffer was PBS. Experiments with controls were visualized by exposure on the same film.

**RESULTS**

*ZirS Structure* The NMR spectroscopically derived solution state structure of ZirS consists of eight β-strands and one 3_10 helix that fold into the canonical β-sandwich of the immunoglobulin superfamily (IgSF) (Figure 1A) (38). The ZirS structural ensemble is displayed in Figure 1B, the statistics of the calculation summarized in Table S1, and the assigned 15N-HSQC spectrum in Figure 2A. The ZirS mean structure was submitted to the Dali server search engine (33), and showed highest similarity to the Invasin subunit D3 of *Yersinia pseudotuberculosis* (39) with a Z-score of 9.9 and a rmsd of 2.5Å over 92 common backbone atoms. ZirS is also structurally related to Gelation Factor (Z-score 8.4), Intimin (Z-score 7.8), and Filamin (Z-score 7.5), as well as to a number of other IgSF members such as Complement C3 and C5 (Z-score 6.5) and Myosin binding proteins (Z-score 5.8).

The β-strands of the ZirS Ig domain fold to form two distinct faces on opposite sides of the molecule, with strands 1, 2, 4, 5, and 6 forming one β-sheet face and strands 3, 7, and 8 creating an opposing β-sheet face (Figure 1A). When analyzed for surface and sequence conservation with the Consurf server (40) and PSI-Blast (41),
one face of ZirS is highly conserved, whereas the other is variable (Figure 1C). Given that the Ig domain of ZirS behaves as a monomer in solution (Figure 2D) the former surface is not conserved for oligomerization of the IgSF domain itself, and thus is likely the binding site for a biological interaction partner. Furthermore, a computational analysis with APBS (Adaptive Poisson-Boltzmann Solver) (42) revealed that the conserved surface of ZirS is an electrostatic groove (comprised mainly of lysines and the polar side chains of Ser or Thr) flanked by two protruding areas, which trend toward a more negative potential (Figure 1D and Figure S1B) suggesting a largely electrostatic binding mode.

The observed ZirS IgSF domain spans residues 175-276 of the purified construct (Figure S1). Residues 136-174 of the NMR ensemble were omitted from Figure 1 for clarity as they do not to fold into an ordered structure. Although readily detectable, these 38 residues are disordered and mobile on a nsec-psec time-scale as indicated by their random coils ¹HN chemical shifts (Figure 2A) and by ¹H-¹³N heteronuclear NOE measurements (Figure 1E). DISOPRED (43) does not flag these residues as intrinsically disordered, however Jpred (32) fails to confidently predict any regular secondary structure for residues 150-174. These residues (150-174) may function as a flexible linker between the C-terminal IgSF domain of ZirS and the N-terminal domain of the protein (Figure 2C), a common feature in linked-domain proteins (44).

Despite our efforts, attempts to over-express the ZirS N-terminal domain for structural studies were unsuccessful (Figure S1A). To gain insight into this domain, a structural model was calculated ab initio, using Rosetta (30,31). Only the lowest energy model was a fully folded globular domain, which predicted an α/β structure (Figure 2B and Figure 2C). Submission to the Dali server showed structural homology to members of the PAS-domain family (top Z-score 3.5, rmsd of 3.0 over 72 residues; Figure 2C), proteins that function as sensory components in signaling pathways (45,46), especially for oxygen (47). The Rosetta-generated structural model is thus attractive in the context of ZirS extracellular localization and oxidative stress (OxyR) regulated expression (13).

ZirS Ig-Domain Binding Partners To search for potential glycan and protein interaction partners, ZirS (136-276) was screened using carbohydrate arrays within the Consortium for Functional Glycomics (CFG) (http://www.functionalglycomics.org/), and via SILAC (stable isotope labeling with amino acids in cell culture) experiments (48) with both mammalian and Salmonella protein extracts. The results of the glycan array with common mammalian sugars did not show significant binding and can be viewed in the public repository at the CFG (Printed Array Version 4.1, H:1960). Likewise, SILAC experiments using HeLa extracts with GST-tagged ZirS (136-276) yielded weak hits, many of which were Myosin (Table 1). Myosin is a common contaminant with SILAC (37), and additional experiments showed little binding (Figure S2).

In contrast, SILAC experiments of GST-ZirS (136-276) with Salmonella extracts yielded significant heavy to light isotope ratio hits that were observed in all four replicates of the experiment (Table 1). These hits included both ZirS and ZirT, and the protein STM1670, encoded by an adjacent open reading frame (ORF) to ZirT (13). In line with the current nomenclature, STM1670 will be herein renamed ZirU.

ZirU is Secreted by ZirT Similar to both ZirS and ZirT, ZirU is predicted to be expressed under an OxyR promoter region by Virtual Footprint (49), and to contain an N-terminal cleavable sequence by the server SignalP (50) (Figure S3A). To probe secretion of ZirU by ZirT, we cloned ZirU into the vector pWSK29 with a C-terminal 2HA tag, added ZirT in trans on the vector pBR322 in Salmonella, and monitored the media for the presence of ZirU. In the presence of recombinant ZirT, ZirU was readily exported into the media, but was not exported by Salmonella containing only endogenous ZirT (Figure 3A). This is similar to what was observed previously for ZirS, and is thought to be due to the requirement of a particular stoichiometry (13) and/or a specific set of unmet environmental conditions. For further verification, we also compared the ZirT dependent secretion of ZirU to that of ZirS. Both ZirU and ZirS were secreted into the extracellular media when complemented with ZirT in a ΔzirT background. (Figure 3B). As
our results indicate that ZirU is a substrate for ZirT and a binding partner for ZirS, we will refer to the zirTS antivirulence pathway as the zir antivirulence pathway for simplicity.

**ZirU is Homologous to Immunoglobulin Super Family Members** The structure-based Robetta server (30) predicts ZirU to contain Ig-like domains (38-152 and/or 153-272) with a greater than 85% confidence (HHsearch score 2.2; (30,51)). This is supported by a sequence alignment between ZirS and ZirU showing that many of the residues in the ZirS IgSF hydrophobic core are conserved in ZirU (Figure S3B and S3D). In contrast, sequence analysis of ZirU using PSI-Blast (41) shows homology to the putative zinc finger protein 236 from *Homo sapiens* (e-value 4e-58, NP_0313713), although ZirU lacks a clear metal binding site (Figure S3C, Metaldetector (52)). To resolve this, we performed ICP-MS (inductively coupled plasma mass-spectrometry) to probe for potential metal binding with purified ZirU (38-281) and ZirS (136-276) (Figure 3C). As shown by the analysis, neither ZirU nor the ZirS IgSF domain co-purify with common biological metals, including zinc.

**ZirT Domain Delineation** ZirT residues 88-368 are homologous to the β-barrel region of Invasin, or β domain (13,19), and ZirT localizes to the OM where it facilitates the secretion of ZirS (13) and ZirU (Figure 3B). Furthermore, it is an interaction partner for the ZirS IgSF domain (Table 1), making ZirT a central element of the zir pathway. To explore these interactions we first characterized the domains of ZirT using biochemical and computational analysis (Figure 4 and Figure 4S). To determine the exact ZirT β domain, various constructs were purified as inclusion bodies and solubilized in SDS. Refolding of porins and other β-barrel membrane proteins can be induced by the addition of 2-methyl-2,4-pentanediol (MPD) to SDS and visualized on SDS-PAGE gel for the appearance of a faster migrating band (53) (Figure 4A and Figure S4A). A construct encompassing only the predicted minimal β domain region (80-368) or extension of the C-terminus (80-390) failed to properly refold (Figure S4A), whereas a construct spanning residues 1-390 showed a similar in-gel migration change as full length ZirT (13) (Figure 4B). Correct folding of ZirT (1-390) was monitored by circular dichroism spectroscopy (CD), which showed formation of an expected β-sheet structure and similar secondary structure partitions as predicted by Jpred (~42% β-strand and turns corresponding to ~12-16 β-strands) (Figure 4C and Figure 4D). As a final control, ZirT 1-390 maintained its fold when exchanged into other detergents such as C8E4 (Figure S4B).

The C-terminal region of ZirT (430-660) was analyzed using the homology modeling server Phyre (34) and Modeller (35), then verified by biochemical characterization of constructs spanning the predicted structural regions. Both tools yielded a model of ZirT spanning residues 431 to 647 with 99% confidence based on the segments D1 and D2 of Invasin from *Yersinia pseudotuberculosis* (39) (Figure 4E). Constructs spanning residues 430-660 and 430-647 were soluble, could be purified, and displayed a dispersed 15N-HSQC spectra characteristic of high β-sheet content (Figure S4A, Figure S4D). For completeness, ZirT (430-660) was also tested by ICP-MS but showed no evidence of zinc binding (Figure 3C).

**ZirT Topology** As ZirT residues 1-390 constitute an integral membrane β domain (Figure 4), we asked if the C-terminal domains of ZirT (430-660) are periplasmic or extend into the extracellular medium. To test this, ZirT 1-660 was cloned into pBAD with a thrombin cleavable C-terminal his-tag, and introduced into *Salmonella* (ΔinvG). ZirT expression was induced, followed by incubation of unlysed cells with thrombin and subsequent addition of PMSF to stop the proteolytic reaction. Membrane enriched fractions were solubilized, boiled, and ZirT visualized by anti-his-tag western blot (Figure 5). The untreated ZirT (0 thrombin) ran as two bands, one similar to purified inclusion bodies containing the sec-signal (verified by N-terminal sequencing), and a lower band that reflects N-terminal cleavage by the sec-system with the C-terminal his-tag still present (Figure 5A). ZirT that has yet to be acted upon by the sec-system (IB) retains its his-tag, whereas the his-tag of mature ZirT (Mem) is removed (Figure 5A and Figure 5B), suggesting that the C-terminus is extracellular. For further validation, ZirT induction was increased to 2hrs to allow the amount of mature ZirT to over-take the full-length material, which clearly showed a dose dependent
loss of the C-terminal his-tag (Figure 5C). Finally, as the IgSF domains of ZirT are extracellular, we probed for a mammalian interaction partner, but found no significant hits (Table S2).

**ZirT Self-Association requires the β Domain**

Outer membrane proteins such as Invasin and Intimin are believed to oligomerize on the cell surface to mediate multivalent adhesion (54,55). We tested if ZirT could also self-associate by examining the behavior of mature ZirT purified from the *Salmonella* OM by gel-filtration. ZirT injected at a concentration of 3 mg/mL, elutes as a peak close to the predicted size of a monomer as well as a second peak of large molecular size (Figure 6A). Immediately following gel filtration, the ZirT monomer fraction was observed by dynamic light scattering at two concentrations (Figure 6B). At a lower concentration (0.4 mg/mL) ZirT behaved as a single species with comparatively low polydispersity (26.7%) and had a predicted mass in agreement with the monomer seen in gel filtration. In contrast, a 2.5 fold more concentrated ZirT sample (1 mg/mL) displays higher polydispersity (44.1%) showing the presence of both the monomer and larger species. This suggests that ZirT can self-associate in a concentration dependent manner. Interestingly, the mean size of the larger species was ~300 kD, implying a possible trimer. For further verification, chemical cross-linking was also performed on ZirT within the membrane fraction (Figure S4C), which shows larger ZirT species that appear with increasing cross-linking agent.

We further explored the nature of ZirT self-association by examining if the IgSF-like domain (430-660) or the β domain (1-390) is the point of contact. ZirT (430-660) was examined by gel-filtration coupled static light scattering at 3 mg/mL but showed only a single well-defined monomeric species (Figure 6C). The refolded β domain was injected at 3 mg/mL and also eluted at a single peak, but had a predicted size of ~380 kD (Figure 6D) showing that it self-associates. As shown by the inset gel, the β domain remained folded (Figure 4) indicating that the result was not due to simple non-specific aggregation of unfolded protein.

**ZirS and ZirU Interact with the C-terminal Domain of ZirT**

The ZirS IgSF domain (Figure 1) is an interaction partner for both ZirU (Figure 3) and ZirT (Figure 4) as shown by SILAC (Table 1). To further explore these protein-protein interactions, purified GST-tagged constructs of the ZirS-IgSF domain (136-276), the Ig-like ZirU (38-281) and the extracellular ZirT IgSF domains (430-660) were used as bait in a far western experiment with the biological mature forms of ZirS and ZirU secreted from *Salmonella* (Figure 7). In agreement with the SILAC results (Table 1), wildtype ZirS interacts with both ZirU and the extracellular domain of ZirT (Figure 7A). Additionally, secreted ZirU also interacts with the Invasin domain of ZirT, but contrary to the SILAC results, does not interact with the ZirS IgSF domain (Figure 7B). Affinity tagged protein experiments are often directional, and unlike SILAC, a far-western first denatures the bait which can disrupt a biological interaction. Complementing these results, ZirU (38-281) and the ZirS IgSF domain were tested by both NMR titration and ITC showing a weak but identifiable interaction (Figure S5). Both techniques showed similar results, suggesting a dissociation constant in the mid-micromolar range (50-120 µM). An accurate number could not be obtained due to a concentration limit of the ZirS IgSF of ~1 mM.

The ZirS IgSF domain targets ZirS to ZirU and ZirT To test if the conserved surface of the ZirS IgSF (Figure 1C) is the contact point between ZirS and its binding partners, both conserved and unconserved surface residues were selected in full-length ZirS (Figure 7E), substituted with alanine, and probed for binding (Figure 7C and Figure 7D). Every mutant assayed was readily expressed at levels similar to wildtype (Figure S6C and S6D) but only the mutations of conserved residues seemed to significantly block the interaction with both ZirU and ZirT (Figure 7C). This data, in combination with the SILAC data (Table 1), show that ZirS and ZirU not only interact with each other after secretion but also bind to the extracellular domains of ZirT, with these interactions mediated in part by the conserved ZirS IgSF surface.

**DISCUSSION**

In this work, we structurally and biochemically characterize the known components of the *zir* antivirulence pathway, the secreted ZirS
and transporter ZirT, and identify a new protein component that interacts with the system, ZirU. We have solved the NMR solution structure of ZirS (136-276) (Figure 1) and found that the ZirS IgSF domain interacts with both ZirT and ZirU, by using SILAC mass-spectrometry methods supported by additional biochemical analysis (Table 1, Figure 3, Figure 4, and Figure 7). Similar to ZirS, ZirU is secreted by ZirT and is likely also an immunoglobulin super family fold. As ZirT appears central to this pathway, biochemical and biophysical characterization of this integral membrane protein were performed, allowing us to delineate the span and topology of the β domain and extracellular tandem Ig-like adhesin domains of ZirT (Figure 4 and Figure 5), and to show that ZirT can self-associate in a concentration dependent manner through contacts in its integral membrane β domain (Figure 6). Additionally, we observe that ZirT acts as a binding platform for ZirS and ZirU (Figure 7).

The NMR solution structure of ZirS (136-276) revealed that this secreted protein is a member of the Immunoglobulin super family (IgSF), with significant structural homology to a large number of known protein binding modules. Closer inspection showed that one face of this domain exhibited high sequence conservation, suggestive of a potential interaction surface. Follow up experimentation with SILAC and site specific mutagenesis methods verified that this was indeed the case, showing that the ZirS-IgSF domain facilitates an interaction with ZirT and ZirU. To date, the bacterial immunoglobulin-like domains (BID or Big) (56) that have been studied are expressed in response to a host environment and are virulence factors. These BIDs enhance virulence by mediating cell host adhesion through a direct contact with host proteins or other targets on the host cell (39,55,57). This includes the Ig-like domains that form the pilus secreted by the chaperone-usher system, which self-associate by donor strand-complementation on the surface of enteric bacteria (58). Thus it was interesting to find that Salmonella also secrete BIDs, ZirS and ZirU, in response to the host environment as part of an antivirulence mechanism. It was known that regions of the ZirT transporter are homologous to Invasin/Intimin (13) and therefore may function in adhesion, but it was surprising to find that ZirT serves as the central binding platform for its own secreted IgSF substrates. This instead evokes a potential similarity to the chaperone-usher system in that the Zir proteins may self-associate into a large multiple Ig-domain adhesive structure, but unlike the pilus the zir pathway does not seem to use donor-strand complementation (59,60). Instead, based on our observations, the zir pathway appears to use multiple weak interactions between several multi-domain Ig-like proteins, possibly acting in a synergistic effect to construct a strong scaffold. Our data thus reveal a new function for the IgSF and BIDs, that of self-adhesion for antivirulence.

The zir pathway had been shown to be homologous but distinctive to type V secretion (13). Specifically, it was thought to mirror two-partner secretion (TPS) but have a separate mechanism for substrate recognition by the transporter. Here our SILAC results with the ZirS IgSF domain definitively separate any connection between this antivirulence secretion system and type V secretion. As we have found that an additional Salmonella protein ZirU not only interacts with ZirS but is also secreted by ZirT, this secretion system must either be classified as its own sub-family of type V or given its own secretion classification. We prefer the latter, as current bioinformatic results show a vast number of BIDs and associated operons of unknown function (56) that may secrete substrates in a manner similar to the zir pathway. As these systems are studied, a new secretion class that includes zir will likely be established. This likelihood is also exemplified by observations that ZirS and ZirU share no readily identifiable secretion signal for ZirT (Figure S3), suggesting a mode of transporter recognition unique from currently known mechanisms.

The newly identified ZirU is predicted to also contain at least one IgSF domain, which is supported by both computational modeling with Robetta (30) and sequence comparison to ZirS. Similar to ZirS, these domains are likely protein binding modules for interaction with ZirS and ZirT. However, more experimentation will be needed to search for any additional ZirU extracellular binding partners, and to understand the role of ZirU in antivirulence. For example, a multimeric assembly of the Zir proteins into an adhesive surface may be required to recognize
other partners. Likewise, although the ab initio model of ZirS (50-146) predicts a potential environmental sensor, the role of this domain also remains to be explored.

The discovery of ZirU as a component of the zir antivirulence pathway prompts the question if other genes encoded near ZirS and ZirT also participate in this system. The answer is likely yes, as STM1667, which we name here ZirR, is a Thioredoxin (13). This is especially convincing since ZirR is predicted to be a glutaredoxin (34), a known negative regulator of OxyR, which therefore may act together in regulation of the zir pathway. Another protein, STM1671 (encoded adjacent to ZirU), is a Rob-like transcription factor (as predicted by Phyre with 100% confidence) related to the activator SoxS involved in defense against oxidative damage in Enterobacteria (61). STM1666 (encoded adjacent to ZirR) is a predicted Ferredoxin-like fold (Phyre 100% confidence), similar to heme-degrading proteins such as LsdG (62). STM1665 is a putative phenylalanyl-tRNA synthetase (Phyre 100% confidence), proteins whose activity has been shown to be influenced by zinc (63). In concert, these surrounding genes may provide intracellular signaling cues coordinated to the self-adhesive functions and/or stimulation provided extracellularly by ZirUTS.

In our refolding experiments with the ZirT β domain, we found it interesting that the sequence predicted minimal domain (88-368 or 88-390) was insufficient for proper refolding. Inclusion of the N-terminal elements (27-87) was required for folding and by inference structural integrity of the domain. From secondary structure prediction this N-terminal region is completely α-helical. A common feature of structurally related outer membrane domains is an α-helical plug in the central pore which functions to facilitate secretion of a substrate (64). As our topology results indicate that the ZirT Ig-like domains are extracellular, the N-terminal region of ZirT is likely the periplasmic point of interaction with ZirS and ZirU for secretion. Given that helices of ~20 amino acids are required to span the length of the membrane, one can reason that a third of the N-terminal helical region may cover the central lumen of the ZirT porin, with the remaining 40 amino acids extending into the periplasm to recognize secretion partners.

Given the previously observed inhibitory effects of zinc on the zir pathway (13), binding of the cation was an obvious first possibility to test. However, our ICP-MS data show that none of the available extracellular zir pathway components tightly associate with zinc. Although, the N-terminal domain was not included in this analysis due to poor solubility, one must also allow for the possibility that the zir pathway does not interact directly with zinc, but is rather regulated by the ion at the level of protein expression. During bacterial infection zinc and other metals are rapidly sequestered, a response known as nutritional immunity (65). Changing concentrations of zinc in the environment may thus allow Salmonella to regulate expression of the zir pathway and to cue antivirulence in response to encountering a host environment.

In light of the known biological regulation of the zir pathway combined with our structural and biochemical results, we can construct a first molecular model of this novel antivirulence mechanism (Figure 8). In our model, the acute phase response (sequestering of zinc) and anaerobic respiration by Salmonella in the small intestine (17) sustain the observed high zir pathway levels (13). ZirT secretes both ZirS and ZirU into the extracellular space, or possibly the Salmonella containing vacuole, where they interact with themselves and with the extracellular Ig-like domains of ZirT that has self-associated through the β domain due to high concentrations from sustained expression. This patch of adhesive ZirT would form a binding platform for the building concentration of Zir proteins secreted from an increasing Salmonella population. We predict, that a specific concentration of Zir complexes may provide a regulatory signal, the nature of which is still unknown, but could involve adhesion to other Salmonella in a manner similar to a biofilm or interaction with environmental solutes or host proteins through the ZirS N-terminal domain or perhaps ZirU. The build up of ZirS, and ZirU, would thus serve as a sensor of population or a change in the host environment to limit bacterial population growth and migration. Upon migration to systemic sites and exposure to reactive oxygen species from macrophages, OxyR
activity is engaged and the zir pathway silenced to allow for increased survival and replication of Salmonella. In this model, the observed population differences between wildtype and zir deletions at the systemic sites (13) would stem from the zir antivirulence regulation in the intestine. We suspect that ZirR may also function in the intestine to prevent shut down of the system by low levels of OxyR, with the genes surrounding the zir pathway relaying intracellular signaling cues timed to the events occurring at the cell surface.

Current research shows that the interaction between a pathogen and its host is an increasingly complex dynamic process that has been fine-tuned by co-evolution. The zir pathway studied here is no exception, and based on our research, exerts its regulatory effects through an intricate mechanism involving the interactions between several proteins. In the context of the structural and biochemical data presented here, the zir antivirulence pathway is a multicomponent immunoglobulin adhesion system key to the environmental conditions of the host.

REFERENCES
5. Marlovits, T. C., and Stebbins, C. E. (2010) Type III secretion systems shape up as they ship out Curr Opin Microbiol 13, 47-52


38. Harpaz, Y., and Chothia, C. (1994) Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains J Mol Biol 238, 528-539


FOOTNOTES

Acknowledgements The authors thank Dr. Harold Jerome Coyne III for NMR tutelage, Dr. Emilie Lameignere for help with light scattering, Matthew Solomonson for help with ICP-MS, Dr Fred Rosell of the LMB spectroscopy and kinetics hub for assistance with the circular dichroism measurements, Wanyin Deng for assistance with far westerns, and Dustin King for help with Cross-linking. This work was funded by the Human Frontier Science Program (HFSP), the Michael Smith Foundation for Health Research (MSFHR), the Canadian Institute of Health Research (CIHR) and Howard Hughes Medical Institute. Instrument support was provided by the CIHR, the Canadian Foundation for Innovation (CFI), the British Columbia Knowledge Development Fund (BCKDF), the UBC Blusson Fund, and the Michael Smith Foundation for Health Research (MSFHR).


Conflict of Interest The authors declare that they have no conflict of interest.

The abbreviations used are: IgSF, Immunoglobulin Super Family; SILAC, Stable Isotope Labeling of Amino acids in Cell culture; ICP-MS, inductively coupled plasma mass-spectrometry; BID and Big, bacterial immunoglobulin-like domains.

FIGURE LEGENDS

Figure 1. Structure of the ZirS C-terminal IgSF Domain. (A) A ribbon diagram of a low energy representative of the structural ensemble of ZirS residues 175-276 showing two rotated views. β-sheets are colored in blue and the single 3₁₀ helix in red. (B) Overlay of the 10 member ensemble of the lowest energy ZirS structures. (C) The surface representation of the structures displayed in panel A, colored for sequence conservation by the Consurf server. Red is highly conserved, pink shows areas of greater than average conservation, and grey areas of no conservation (Consurf score of 5 or less). (D) The first surface in Panel C colored by electrostatic potential generated by APBS (Adaptive Poisson-Boltzmann Solver) and the PARSE forcefield. The potential is contoured from -2 kT/e (red) to 2 kT/e (blue). (E) ¹H-¹⁵N heteronuclear NOE relaxation data for ZirS recorded with a 600 MHz spectrometer at 20°C. Decreasing NOE values result from increasing mobility of the ¹⁵N-¹H bond vector on the sub-nsec timescale. Missing data points correspond to prolines or residues with overlapping signals. Graphics were generated using Pymol (66).
Figure 2. ZirS Domain Structure and N-terminal Ab initio Model. (A) Assigned $^{15}$N-HSQC of ZirS C-terminal IgSF domain residues 136-276. (B) Domain map of ZirS based on the ZirS IgSF structure, expression trials, and ab initio model. (C) An ab initio model of ZirS (50-146) generated by Rosetta. $\alpha$-helices are colored in red and $\beta$-strands in blue with an overlay of the model with an HTR-like protein (pdb code 3fc7) in yellow. (D) Static gel-filtration coupled light scattering of ZirS (136-276) in the NMR buffer. The inset is the SDS-PAGE gel of purified ZirS.

Figure 3. ZirU is part of the $zir$ Pathway. (A) Secretion of ZirU by ZirT in cultured Salmonella as visualized by western blot. P represents the cell fraction and S the filtered media. Lanes with ZirT + indicate that ZirT has been added in trans. DnaK is shown as a negative secretion control. (B) As in (A) with the secretion of both ZirS and ZirU by ZirT in a Salmonella $zirT$ deletion strain, SL1344ΔzirT. (C) Metal binding analysis of the $zir$ pathway by ICP-MS. The data show the average metal observed for each protein from three measurements. GST is the negative control and NDM-1 the positive control for zinc binding. The data is graphed by shaded box.

Figure 4. Domain Structure of ZirT Shows a $\beta$ Domain Linked to a Soluble IgSF Fold. (A) Cartoon representation of the ZirT domains. (B) Refolding of the ZirT $\beta$ domain by SDS and MPD as visualized by coomassie stained SDS-PAGE gel. Bands corresponding to folded and unfolded species are indicated. (C) MPD-induced ZirT $\beta$ domain refolding monitored by circular dichroism spectroscopy. (D) Deconvolution of the recorded CD data with Dichroweb. The secondary structure fraction for ZirT (1-390) as predicted by Jpred was calculated by dividing the number of residues predicted to have secondary structure elements by the total residues. Unordered denotes no secondary structure prediction. Error bars show the standard deviation from three independent experiments (E) Homology model of ZirT residues 430-660 as calculated by Phyre and Modeller. A530 is in the middle of the linker region between the two Ig folds. Residues C-terminal to K647 are predicted as disordered.

Figure 5. The ZirT IgSF Domains are Extracellular. (A) Western blot of the membrane fractions of cells with ZirT expression induced for 1 hr after whole cells were treated with thrombin. IB represents purified inclusion bodies and the amount of Thrombin used in each experiment indicated above the panel. Each lane consists of an upper band corresponding to unprocessed protein and a lower band corresponding to processed protein inserted into Salmonella membranes. (B) Quantification of Panel A performed in triplicate. IB represents the upperband and Mem the lower band. (C) Repeat experiment of Panel A but at 2hrs post induction. Only the processed lower band is observed.

Figure 6. ZirT Self-Associates (A) Chromatogram representing the gel-filtration profile of purified ZirT from Salmonella membranes. The elution volumes are the average with the standard deviation of three separate purifications. The inset shows a Coomassie stained SDS-PAGE gel of pure ZirT. The bottom panel lists the standard elution volumes of known proteins and the calculated size of ZirT. (B) Dynamic light scattering of ZirT at two different concentrations, 0.4 mg/mL in grey and 1 mg/mL in black. The calculated diameter, polydispersity, predicted molecular weight (MW) in kilodaltons, and percentage of mass is listed in the box next to the curve. (C) Gel filtration profile of the ZirT Invasin domain (430-660) with the calculated static light scattering size indicated on the figure. The inset shows purified ZirT Invasin domain. (D) Gel filtration profile of the ZirT $\beta$ domain (1-390). The elution volume is the average with the standard deviation of three separate purifications. The inset shows a Coomassie stained SDS-PAGE gel of pure ZirT 1-390. The bottom panel lists the calculated size of ZirT (1-390).

Figure 7. ZirS and ZirU Bind the C-terminal Domains of ZirT. (A) Far western experiment with ZirS-HA. GST was used as a control and GST-ZirU or GST-ZirT(430-660) used as bait. A western blot without treatment with ZirS-HA is also shown next to ZirS-HA isolated from Salmonella as a control. (B) The equivalent experiment as described in Panel A but with ZirU-HA and GST-ZirS and GST-
ZirT(430-660) used as bait. P indicates cell pellet and S the media. Bait concentrations were equivalent. (C) Far-western binding experiments with mutants of the conserved surface of the ZirS IgSF domain. (D) Far-western binding experiments with mutants of unconserved residues within the ZirS IgSF domain. (E) Structure of the ZirS IgSF domain with the positions of the mutants from panels C and D shown. M264 is on the opposite side of the displayed surface.

**Figure 8.** Mechanistic Model for the zir Antivirulence Pathway. Cartoon representation of the structural and biochemical data presented in Figures 1-7. Proteins are represented by colored ellipse and labeled by structural domain. The zir operon and surrounding genes are represented by a block arrow at the bottom of the figure. ZirRTSU are expressed during shedding and infection of the host gastrointestinal tract to regulate *Salmonella* growth. ZirT (orange) secretes both ZirS (blue) and ZirU (green), which are released into the extracellular space and can self-associate. ZirT maintains an equilibrium in the OM, forming an Ig-rich adhesion surface for ZirS and ZirU. Upon the build up of enough ZirT, ZirS, and ZirU from infecting *Salmonella* a regulatory signal is engaged. This event may involve unknown protein targets (light brown) from *Salmonella* or the host that interact with the ZirS N-terminal domain or ZirU. Expression of the zir genes is negatively regulated by OxyR and zinc and may involve the thioredoxin protein ZirR (red). The adjacent genes STM1666, STM1665, and STM1671 encode predicted redox pathway proteins (see text) making them candidates for the cytoplasmic signaling events of the zir pathway.
Figure 1
Figure 2
Figure 3
Figure 4
Figure 5
Figure 6
Figure 7
Figure 8