NMR STRUCTURAL STUDIES OF TYPE III SECRETION SYSTEM TIP PROTEINS

By

©2011
Bryce Nordhues

B.S., University of Kansas, 2008

Submitted to the graduate degree program in the Department of Molecular Biosciences and the Graduate Faculty of the of the University of Kansas in partial fulfillment of the requirements for the degree of
Master of Arts

Committee members:

Chairperson – Roberto N. De Guzman

Truman C. Gamblin

Yoshiaki Azuma

Date Defended: 06/08/2011

The Thesis Committee for Bryce Nordhues certifies
that this is the approved version of the following thesis:

NMR STRUCTURAL STUDIES OF TYPE III SECRETION SYSTEM TIP PROTEINS

Committee members:

Chairperson – Roberto N. De Guzman

Truman C. Gamblin

Yoshiaki Azuma

Date Approved: 06/14/2011

Acknowledgements

I would first like to thank Dr. Roberto De Guzman for his guidance as my research mentor for the last three years. He introduced me to the world of NMR and taught me how to be a good scientist. I would also like to thank him for his continuous support and encouragement in times of difficulty.

I would also like to thank Dr. Yoshiaki Azuma getting me hooked on research as an undergraduate student and for convincing me to apply to graduate school. Dr. Azuma was also a constant source of advice during my graduate career in regards to both my science career and my personal life and I cannot thank him enough.

I would like to thank all of my labmates in the De Guzman lab, including Dr. Yu Wang, Dr. Thenmalar Rathinavelan, Dr. David F. Estrada, Dr. Dalian Zhong, Dr. Nara Kongari, Srirupa Chatterjee, Sukanya Chaudhury, Sunhwan Jo, and Andrew McShan. This work could not have been completed without their help.

I would like to acknowledge members of the KU Protein Production Group for Biacore instrumental support including Dr. Philip Gao and Dr. Na Zhang. I would also like to acknowledge Dr. Asokan Anbanandam for his assistance with various NMR instruments and NMR data collection.

I would like to thank Dr. Truman C. Gamblin, Dr. Yoshiaki Azuma, and Dr. Roberto De Guzman for serving as my Masters thesis committee and for their feedback and advice in completing this document.

Finally, I want to thank my parents, my brother, my sister, and the rest of my family for always cheering me on. Their love and support is what has kept me going.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

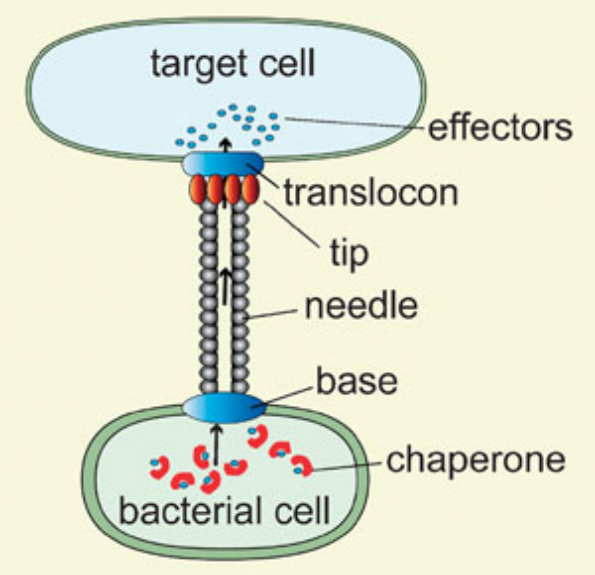
T3SS	Type III Secretion System
T3SA	Type III Secretion Apparatus
FRET	*Förster* Resonance Energy Transfer
GB1	*Streptococcus* protein G B1 domain
TEV	Tobacco Etch Virus
LB	Luria-Bertani
IPTG	Isopropyl-β-D-thiogalactopyrandoside
NMR	Nuclear Magnetic Resonance
ITC	Isothermal Titration Calorimetry
DMEM	Dulbecco's Modified Eagle's Medium
HSQC	Heteronuclear Single Quantum Coherence
CD	Circular Dichroism
TROSY	Transverse Optimized Relaxation Spectroscopy
CMC	Critical Micelle Concentration
BLAST	Basic Local Alignment Search Tool
SPR	Surface Plasmon Resonance
SUMO	Small Ubiquitin-Related Modifier
PIAS	Protein Inhibitor of Activated STAT
PARP	Poly(ADP-ribose) Polymerase

Chapter 1: Introduction to the Type III Secretion System

The type III secretion system (T3SS) has been identified as a critical factor for pathogenicity in more than twenty-five species of Gram-negative bacteria[1], which cause disease in both plants and animals (including humans).[2; 3] The resulting diseases are serious and potentially deadly, as exemplified by the recent European outbreak of the Shiga toxin-producing *Escherichia coli* (STEC). This strain of *E. coli*, which causes systemic infection and essentially leads to kidney failure, has killed 21 individuals and infected more than 2400, making it the largest *E. coli* outbreak in recorded history (www.cdc.gov). Outbreaks from other T3SS-harbouring bacteria have also been of recurrent concern for public health and safety, including those of *Salmonella* borne from contaminated vegetables in the United States.[4] Others among the species that cause illness in humans are *Pseudomonas aeruginosa, Yersinia* spp*., Burkholderia* spp*., and Shigella* spp. While several studies have indicated T3SS tip proteins as potential vaccination antigens, there are currently no approved vaccines for the pathogens listed above.[5; 6] In addition, a number of these species are considered potential agents for bioterrorism, such as *Burkholderia pseudomallei* and *Yersinia pestis*.[7] Despite the variance in infective strategies used by each of these species, they each use the T3SS to deliver bacterial effector proteins into their hosts to manipulate normal cellular function for bacterial advantage. A better understanding of this system is necessary to eliminate the threat these species pose to public health.

The structural component of the TTSS, the type III secretion apparatus (T3SA), is the conduit by which secreted proteins of the T3SS are transported from the bacterial cytosol into a eukaryotic host (Figure 1-1). This apparatus spans both the inner and outer

membranes of the bacteria as well as the host cell membrane. Multiple copies of more than 25 proteins strictly coordinate the assembly of the apparatus and subsequent effector protein secretion, including those of the needle apparatus, translocon, chaperones, and many regulatory proteins.[1; 2; 8] Each of these components are coded for in pathogenicity islands (PAIs), and while the organization within these pathogenicity islands is very similar between species, the resulting proteins exhibit differing levels of evolutionary divergence.[9]



Pathogen	Tip Protein	Chaperone
S. typhimurium	SipD	-
S. flexneri	IpaD	-
P. aeruginosa	PcrV	PcrG
Y. pestis	LcrV	LcrG

Figure 1-1. Generic illustration of the type III secretion system (T3SS) including the needle apparatus, translocon, effector proteins, and chaperones. (Figure is from reference 5). The tip proteins and their respective chaperones involved in this study are tabulated.

Following assembly of the needle apparatus and prior to effector secretion, intermediate activation steps involving translocon secretion and assembly have been proposed to involve environmental signaling.[10; 11; 12] However, bacterial secretion becomes fully activated upon cell contact and effector proteins are then delivered to the cytosolic surface of the target cell.[13; 14; 15] The roles of the secreted effectors vary widely

among different species, but commonly bear a resemblance to eukaryotic factors in order to manipulate target signal transduction.[2] Depending on the invading bacterial species, host responses to effector proteins can include actin cytoskeletal rearrangement and membrane ruffling (leading to bacterial intracellularization), altered vesicle trafficking, or evasion of the immune response by killing polymorphonuclear neutrophils and macrophages.[13; 16; 17] Despite the divergence in effector types and in strategies for pathogenesis, many of the T3SS components are conserved both structurally and functionally between species. Increasing knowledge of T3SS assembly and regulation has permit developments of live vectors for exogenous antigen delivery using species such as *S. typhimurium* and *P. aeruginosa*.[18; 19]

The tip protein, named for its localization to the needle tip (Figure 1-1), is a critical structural and functional component of the T3SS.[20; 21] Relevant tip proteins discussed herein are SipD in *Salmonella typhimurium*, IpaD in *Shigella flexneri*, BipD in *Burkholderia pseudomallei*, PcrV in *Pseudomonas aeruginosa*, and LcrV in *Yersinia Pestis*. In addition to their function in translocon assembly, tip proteins have been demonstrated to regulate secretion control in all species studied to date.[22] This regulation is coordinated with chaperones in the bacterial cytosol and with external environmental cues.[10; 11; 12] Structural studies of tip proteins and their binding partners are important in understanding T3SS-mediated infection and for potential drug design targeted at disrupting the assembly of the apparatus.

The contents of this thesis are focused on characterization of T3SS tip proteins and their binding partners. The interaction of bile salts with SipD, the *Salmonella* tip protein, is described in chapter 2 (adapted from reference 13) along with a model for the

differential bile salt response between *Salmonella* and *Shigella*. Chapter 3 is dedicated to structural characterization of PcrG in *Pseudomonas aeruginosa* and its binding interaction with the tip protein, PcrV. Chapter 4 includes a discussion of the T3SS studies described in chapters 2 and 3 as well as the contribution of this work to the field. Finally, using the same experimental template applied in the biophysical characterization of the T3SS proteins, an addendum is included describing the interaction between proteins involved in SUMOylation. These proteins are the *Xenopus laevis* E3 ligase, PIASy, and a SUMOylation target, PARP1.

All work described in this thesis, including experimentation and interpretation of results, is my own unless otherwise noted.

References

1. Cornelis, G. R. (2006) The type III secretion injectisome, *Nat. Rev. Microbiol. 4*, 811-825.
2. Hueck, C. J. (1998) Type III protein secretion systems in bacterial pathogens of animals and plants, *Microbiol. & Mol. Biol Rev*. 62, 379-433.
3. Wang, Y., Zhang, L., Picking, W. L., Picking, W. D., and De Guzman R. N. (2008) Structural dissection of the extracellular moieties of the type III secretion apparatus, *Mol. Biosyst*. 4, 1176.
4. De Guzman, RN. Bacterial Nanoinjectors. *Chem. Biol. News & Res.* 3, B87, 2008.
5. Titball, R. W., and Williamson, E. D. (2004) *Yersinia pestis* (plague) vaccines, *Expert. Opin. Biol. Ther.* 4, 965-73.
6. Tacket, C. O., Sztein, M. B., Losonsky, G. A., Wasserman, S. S., Nataro, J. P., Edelman, R., Pickard, D., Dougan, G., Chatfield, S. N., and Levine, M. M. (1997) Safety of live oral *Salmonella typhi* vaccine strains with deletions in htrA and aroC aroD and immune response in humans, *Infect. Immun.* 65, 452-456.
7. Ruppitsch, W., Stöger, A., Indra, A., Grif, K., Schabereiter-Gurtner, C., Hirschl, A., and Allerberger, F. (2006) Suitability of partial 16S ribosomal RNA gene sequence analysis for the identification of dangerous bacterial pathogens, *J. App. Microbiol.* 102, 852-859.
8. Blocker, A. J., Deane, J. E., Veenendaal, A. K. J., Roversi, P., Hodgkinson, J. L., Johnson, S., and Lea, S. M. (2008) What's the point of the type III secretion system needle?, *PNAS* 105, 6507-6513.

9. Hacker, J. and Kaper, J. B. (2000) Pathogenicity islands and the evolution of microbes, *Annu. Rev. Microbiol.* 54, 641-79.
10. Stensrud, K. F., Adam, P. R., La Mar, C. D., Olive, A. J., Lushington, G. H., Sudharsan, R., Shelton, N. L., Givens, R. S., Picking, W. L., and Picking, W. D. (2008) Deoxycholate interacts with IpaD of *Shigella flexneri* in inducing the recruitment of IpaB to the type III secretion apparatus needle tip, *J. Biol. Chem. 283*, 18646-18654.
11. Olive, A. J., Kenjale, R., Espina, M., Moore, D. S., Picking, W. L., and Picking, W. D. (2007) Bile salts stimulate recruitment of IpaB to the *Shigella flexneri* surface, where it colocalizes with IpaD at the tip of the type III secretion needle, *Infect. Immun. 75*, 2626-2629.
12. Wang, Y., Nordhues, B. A., Zhong, D., and De Guzman, R. N. (2010) NMR characterization of the interaction of the *Salmonella* type III secretion system protein SipD and bile salts, *Biochemistry* 49, 4220-4226.
13. Dacheux, D., Epaulard, O., de Groot, A., Guery, B., Leberre, R., Attree, I., Polack, B., and Toussaint, B. (2002) Activation of the *Pseudomonas aeruginosa* type III secretion system requires an intact pyruvate dehydrogenase aceAB operon, *Infect. Immun.* 70, 3973-3977.
14. Bahrani, F. K., Sansonetti, P. J. & Parsot, C. (1997). Secretion of Ipa proteins by *Shigella flexneri*: inducer molecules and kinetics of activation. *Infect Immun* **65**, 4005-10.
15. Rosqvist, R., Magnusson, K. E. & Wolf-Watz, H. (1994). Target cell contact triggers expression and polarized transfer of *Yersinia* YopE cytotoxin into mammalian cells. *EMBO J* **13**, 964-72.
16. Kaniga, K., Trollinger, D., and Galan, J. E. (1995) Identification of two targets of the type III protein secretion system encoded by the inv and spa loci of Salmonella typhimurium that have homology to the Shigella IpaD and IpaA proteins, *J.Bacteriol. 177*, 7078-7085.
17. Galan, J. E. (2001) *Salmonella* interactions with host cells: type III secretion at work, *Annu. Rev. Cell. Dev. Biol.* 17, 53-86.
18. Epaulard, E., Derouazi, M., Margerit, C., Marlu, R., Filopon, D., Polack, B., and Toussaint, B. (2008) Optimization of a type III secretion system-based *Pseudomonas aeruginosa* live vector for antigen delivery, *Clin. & Vac. Immun.* 15, 308-313.
19. Nishikawa, H. et. al (2006) In vivo antigen delivery by a *Salmonella typhimurium* type III secretion system for therapeutic cancer vaccines, *J. Clin. Invest.* 116, 1946-1954.
20. Mueller, C. A., Broz, P., Muller, S. A., Ringler, P., Erne-Brand, F., Sorg, I., Kuhn, M., Engel, A., and Cornelis, G. R. (2005) The V-antigen of Yersinia forms a distinct structure at the tip of injectisome needles, *Science 310*, 674-676.
21. Espina, M., Olive, A. J., Kenjale, R., Moore, D. S., Ausar, S. F., Kaminski, R. W., Oaks, E. V., Middaugh, C. R., Picking, W. D., and Picking, W. L. (2006) IpaD localizes to the tip of the type III secretion system needle of *Shigella flexneri*, *Infect. Immun. 74*, 4391-4400.

22. Lee, P., Stopford, C. M., Svenson, A. G., and Rietsch, A. (2010) Control of effector export by the *Pseudomonas aeruginosa* type III secretion proteins PcrG and PcrV, *Mol. Microbiol*. 75, 924-941.

Chapter 2: Characterization of the Interaction of the *Salmonella* T3SS Protein SipD and Bile Salts

INTRODUCTION

The type III secretion system (T3SS) is a complex and highly regulated scheme used by more than 25 Gram-negative bacteria for pathogenesis.[1] Examples are the *Salmonella* spp. and *Shigella* spp., which utilize the T3SS to deliver effector proteins into the cytosol of human epithelial cells leading to actin cytoskeletal rearrangement, membrane ruffling, and bacterial internalization.[2] Subsequent infection by these species results in dysentery (*Shigella)* or salmonellosis (*Salmonella*), each of which remain an ongoing problem in both developed and underdeveloped countries.

A critical structural and functional component of the T3SS is the tip protein; named for its localization to the tip of the T3SS needle apparatus.[3; 4] These tip proteins are known as SipD in *Salmonella typhimurium*, IpaD in *Shigella flexneri*, and BipD in *Burkholderia pseudomallei* and share a high degree of structural similarity in these species. These species each require their respective tip proteins for translocon secretion regulation, assembly, and for subsequent pathogenesis.[2; 5; 6] IpaD has recently been shown to interact with the small bile salts deoxycholate, taurodeoxycholate, and chenodeoxycholate[7] ; each of these salts further stimulated T3SS activation by recruitment of translocon protein, IpaB, to the *Shigella* cell-surface.[8] While SipD has also been shown to bind deoxycholate[7], *Salmonella* exhibits the opposite response by repressing invasion via transcriptional downregulation.[9; 10]

Bile salts are a major component of bile and normally serve as a biological surfactant for absorption of fats and vitamins.[11] Since bile salts are enriched in the

intestines[11], they are thought to serve as a logical environmental cue for enteric pathogens. However, it remains unclear why *Salmonella* and *Shigella* would have opposing responses to bile when IpaD[12] and SipD[13] crystal structures show such high similarity (C^{α} rmsd = 1.4Å). Stensrud et. al[7] previously used *Förster* Resonance Energy Transfer (FRET) analysis and computational docking simulations to predict a deoxycholate binding pocket in IpaD. This binding pocket was proposed to be located between the IpaD central coiled coil and N-terminal two-helix bundle.[7] In the experiments presented here, NMR methods are used to identify residues of SipD interacting with bile salts deoxycholate, taurodeoxycholate, and chenodeoxycholate. The results have identified a binding pocket distinct from that predicted for IpaD, which poses a potential rationale the opposing response of *Salmonella* and *Shigella* to bile salts. Point mutations were also generated and tested for the ability of *Salmonella* to invade human epithelial cells.

MATERIALS AND METHODS

Protein Expression and Purification

Wild-type *S. typhimurium* SipD (residues 39-343) was subcloned into pET-21a (Dr. Yu Wang, University of Kansas) to create a fusion with His_6 and the *Streptococcus* GB1 domain at the N-terminus (cleavage of the appended tag with the Tobacco Etch Virus (TEV) protease resulted in an N-terminal "GHM" cloning artifact). The construct was expressed in *E. coli* BL21(DE3) DNAY cells with 100 μg/mL carbenicillin and 30 μg/mL kanamycin in all growth media. Starter cultures were grown in 20 mL Luria-Bertani (LB) media overnight at 37 °C followed by centrifugation at 4000 rpm.

Unlabeled SipD$^{39\text{-}343}$ was grown by resuspending the starter culture in 1.0 L LB. Isotopically ^{15}N-labeled SipD$^{39\text{-}343}$ was acquired by resuspending the starter culture in 1.0 L of M9 minimal media supplemented with 1.0 g/L $^{15}NH_4Cl$. Perdeuterated ^{13}C,^{15}N-labeled SipD$^{39\text{-}343}$ was prepared by resuspending the starter culture in the above media dissolved in 99% (v/v) D_20 and supplemented with 2.0 g/L [^{13}C]glucose instead of unlabeled glucose (Dr. Yu Wang, University of Kansas). Cells were grown at 37 °C and induced with 1.0 mM isopropyl-β-D-thiogalactopyrandoside (IPTG) upon reaching A_{600} ~ 0.6-0.8. Growth was continued overnight at 37 °C in a shaking incubator and cells were harvested by centrifugation. Cells were then lysed by sonication, centrifuged (18200g, 10 min., 4 °C), and the supernatant was loaded onto a Ni^{2+}-affinity column (Sigma). Elutes from the Ni^{2+} affinity column containing the recombinant SipD were pooled and incubated with Tobacco Etch Virus (TEV)[14] protease overnight. The digested sample was loaded onto a Ni^{2+}-affinity column, which retained the cleaved His_6-GB1 tag and His_6-tagged TEV protease while passing the cleaved GHM-SipD$^{39\text{-}343}$. Proteins samples were concentrated using Amicon Ultra 3K (Millipore) and protein concentration was determined using absorbance at 280 nm.

^{15}N Amino Acid Specific Labeling

Eight different ^{15}N-amino acid specifically labeled SipD$^{39\text{-}343}$ constructs were expressed and purified to facilitate NMR assignments (Dr. Yu Wang, University of Kansas). These included ^{15}N-labeled Leu, Val, Ile, Ala, Phe, Tyr, Met, and Lys. Each sample was prepared as previously described.[15] Proteins were purified as described above.

Mutagenesis

Assignment of the four tryptophan side chain resonances of SipD (W135, W177, W234, and W290) was approached by introducing tyrosine point mutations using the QuikChange method (Stratagene). The recombinant ^{15}N-labelled SipD mutants were expressed and purified as described above.

Following NMR assignments, additional point mutations in SipD were designed based on the chemical shift perturbations resulting from bile salt titration. Full length *sipD* was first subcloned into plasmid pRK2[16] and residues (S96, F109, S110, L116, F117, Q124, D131, E133, Y149, Q165, L179, V187, K188, V256, Q306, T307, and L322) were mutated to alanines using the QuikChange method (Stratagene). Alanine substitutions were chosen to minimize changes in protein folding. All mutations were confirmed by DNA sequencing.

NMR Spectroscopy

NMR data were acquired at 30 °C with a Bruker Avance 800 MHz spectrometer equipped with a cryoprobe, processed with NMRPipe[17], and analyzed with NMRView[18]. Before NMR data collection, purified protein samples were dialyzed into NMR buffer (10 mM sodium phosphate, pH 6.5, 10 mM NaCl, and 10% (v/v) D_2O). Backbone assignments were achieved using 0.5 mM perdeuterated ^{15}N,^{13}C-SipD$^{39\text{-}343}$ to collect two-dimensional ^{1}H-^{15}N TROSY-HSQC[19] as well as three-dimensional TROSY-HNCA[20], TROSY-HNCACB[21], TROSY-HNCO[21], and TROSY-HN(CA)CO[21] (Dr. Yu Wang, University of Kansas). Two-dimensional ^{1}H-^{15}N TROSY-HSQC[19] were also collected for each ^{15}N-amino acid-specifically labeled sample. To assign tryptophan side chains, 2D ^{1}H-^{15}N HSQC or ^{1}H-^{15}N TROSY-HSQC[19] were collected for ^{15}N-labeled SipD$^{39\text{-}343}$

W135Y, W177Y, W234Y, and W290Y. For NMR chemical shift mapping, 2D ^{1}H-^{15}N TROSY-HSQC[19] were collected using ^{15}N-SipD$^{39-343}$ titrated with increasing concentrations of deoxycholate (Amresco), chenodeoxycholate (Sigma), taurodeoxycholate (Sigma), or cholate hydrate (Sigma). Samples were prepared as follows: 0.6 mM ^{15}N-SipD$^{39-243}$ was dialyzed overnight into 1.0 L NMR buffer containing increasing concentrations of deoxycholate (0.0, 0.7, 1.4, or 2.1 mM) or chenodeoxycholate (0.0, 0.3, 0.6, 1.3, or 2.6 mM) and 0.7 mM ^{15}N-SipD$^{39-343}$ in 1.0 L NMR buffer with taurodeoxycholate (0.0, 0.3, 0.7, 1.3, 2.6 mM) or cholate hydrate (0.0, 0.3, 0.7, 1.3, or 2.6 mM).

Isothermal Titration Calorimetry (ITC)

Microcalorimetric titrations were performed under isothermal conditions at 25 °C or 37 °C using a VP-ITC instrument (MicroCal Inc., Northampton, Ma, USA). Titrations were carried out under a range of buffering conditions, which tested varied salt concentration (10-100 mM NaCl), buffering reagents (Tris/HCl and phosphate buffer), as well as pH (6.5-8.0). The calorimetric cell (volume 1.4 mL) was loaded with 380 μM unlabeled SipD$^{39-343}$ and was titrated with small volumes (10 μL) of 2.5 mM deoxycholate for a total of 24-29 injections. Injections of 2.5 mM deoxycholate into buffer alone served as a control. Each injection was added at 5.0 μL/s with 240 s spacing to return the differential power to baseline. The cell was stirred throughout the procedure by constant rotation of the injection syringe at 310 rpm. Automated injections from a computer controlled syringe into the adiabatic Hastelloy® shield allowed highly sensitive measurements (0.1 ucal) of heat exchange. Raw data was analyzed using Origin 7.0 software (MicroCal Inc.).

***Salmonella* Invasion Assay**

The effect of the *sipD* point mutations on the ability of *S. typhimurium* to invade human epithelial cells was assessed using a modified gentamicin protection invasion assay.[22] The chosen line of cultured human epithelial cells was Henle 407 (American Type Culture Collection # CCL-6). Wild-type *S. typhimurium* (SL1344) was provided by Dr. Bradley Jones (University of Iowa) and a *sipD* knockout strain (*sipD*$^-$) was prepared as previously described.[13] The pRK2-*sipD* plasmids with either wild-type or mutant *sipD* were electroporated into the *S. typhimurium sipD*$^-$ strain and transformants were selected for trimethoprim resistance. Single colonies were inoculated into 10 mL LB supplemented with 25 mg/L trimethoprim, 50 mg/L ampicillin, and 50 mg/L chloramphenicol and grown at 37 °C overnight without aeration. Overnight cultures were diluted 1:10 in LB containing 1.0 mM IPTG +/- 2.0 mM deoxycholate and incubated at 37 °C without aeration for 3-3.5 hours (until O.D. ~0.6). Approximately 15 μL of the bacterial suspensions were added to the Henle 407 cells which had previously been grown to confluence in Dulbecco's Modified Eagle's Medium (DMEM) and 10% fetal bovine serum at 37 °C with 5% CO_2 in 24-well plates. Invasion progressed for 60 min at 37 °C in a 5% CO_2 chamber before the suspension was aspirated and washed 3× with DMEM containing 100 mg/mL gentamicin. The Henle 407 cells were then incubated in the DMEM-gentamicin solution for 90 min, aspirated, washed 1× with DMEM, and lysed with 1% Triton X-100 to release the internalized *Salmonella*. The number of resulting colonies, which correlate to invasiveness, was estimated by a serial dilution of the lysate followed by plating. All invasion assays were performed in triplicate.

RESULTS

Protein Expression and Purification

NMR experiments were initially performed using full length SipD$^{1-343}$ and the resulting HSQC spectra were of poor quality. However, it has been previously shown that homologous tip proteins are disordered at the amino terminus. Crystal structures of IpaD and BipD lacked electron density below residues 39 and 35, respectively.[12] Therefore, recombinant SipD$^{39-343}$ was expressed with a His_6-tagged GB1 domain fused to the N-terminus, purified using Ni^{2+} affinity chromatography, and digested with TEV protease to remove all but a three-residue N-terminal cloning artifact, GHM. Circular dichroism and thermal denaturation data showed that the recombinant SipD remained highly α-helical and maintained the core structure of SipD. Subsequently, HSQC data (acquired by Dr. Yu Wang, University of Kansas) for the truncated SipD revealed well-dispersed spectra, contrary to that of the full length protein (data not shown).

NMR Assignment

The 2D ^{1}H-^{15}N HSQC spectrum of recombinant SipD$^{39-343}$ was assigned nearly to completion using perdeuteration, TROSY NMR, and ^{15}N-amino acid-specific labeling (Dr. Yu Wang, University of Kansas). These experiments permitted identification of 274 of 295 non-proline backbone resonances; the remaining could not be assigned due to peak overlapping or broadening. In addition to backbone assignments, the side chains of all four tryptophan residues (W135, W177, W234, and W290) were assigned by individually mutating to tyrosines. Resulting changes in the overall TROSY and HSQC

spectra were minimal, which suggested limited folding changes and permitted side chain assignments (Figure 2-1).

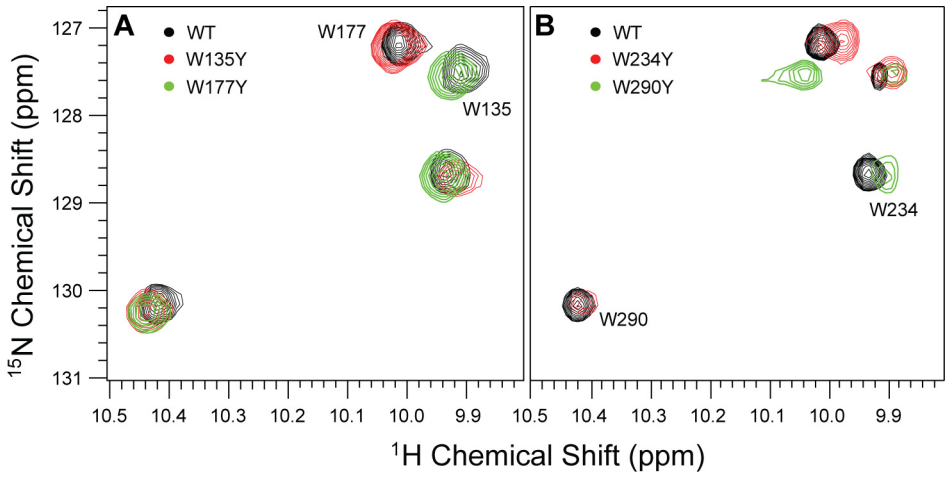


Figure 2-1. Sampling of the proton-nitrogen correlation spectra of wild-type SipD$^{39\text{-}343}$ and the four SipD$^{39\text{-}343}$ mutants W135Y, W177Y, W234Y, and W290Y. Both (**A**) 2D ^{1}H-^{15}N TROSY and (**B**) 2D ^{1}H-^{15}N HSQC spectra were used to assign the tryptophan side chains.

Despite the challenge in assigning all residues, 93% of the backbone C^{α}, C^{β}, and C' resonances were assigned and thereby permit the calculation of secondary C^{α}, C^{β}, and C' chemical shifts (Figure 2-2). The resulting secondary chemical shift data indicated a primarily α-helical secondary structure with several short β strands. In fact, the secondary structure of the recently obtained crystal[13] of SipD$^{39\text{-}343}$ corresponds very well with that predicted from the NMR secondary chemical shifts. This similarity suggests that the truncated protein exhibits little difference in 3D structure when in solution. Additionally, secondary structure alignment of the SipD[13] and IpaD[12] crystal structures reveal only minor variation in α-helix and β strand lengths and spacing between the two tip proteins (Figure 2-3).

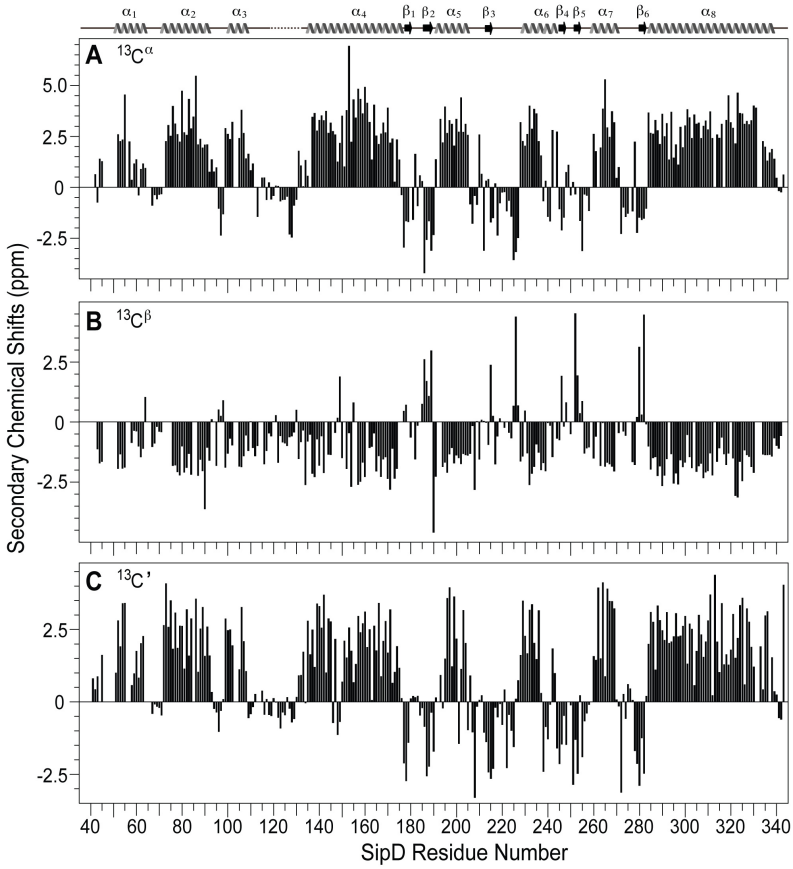


Figure 2-2. NMR secondary chemical shifts of C^{α} (**A**), C^{β} (**B**), and C' (**C**) of $SipD^{39\text{-}343}$ shown along with the secondary structures from the $SipD^{39\text{-}343}$ crystal (top). Secondary structure is denoted with solid lines (loops), wavy lines (helices), a dotted line (disordered loop), and arrows (β strand).

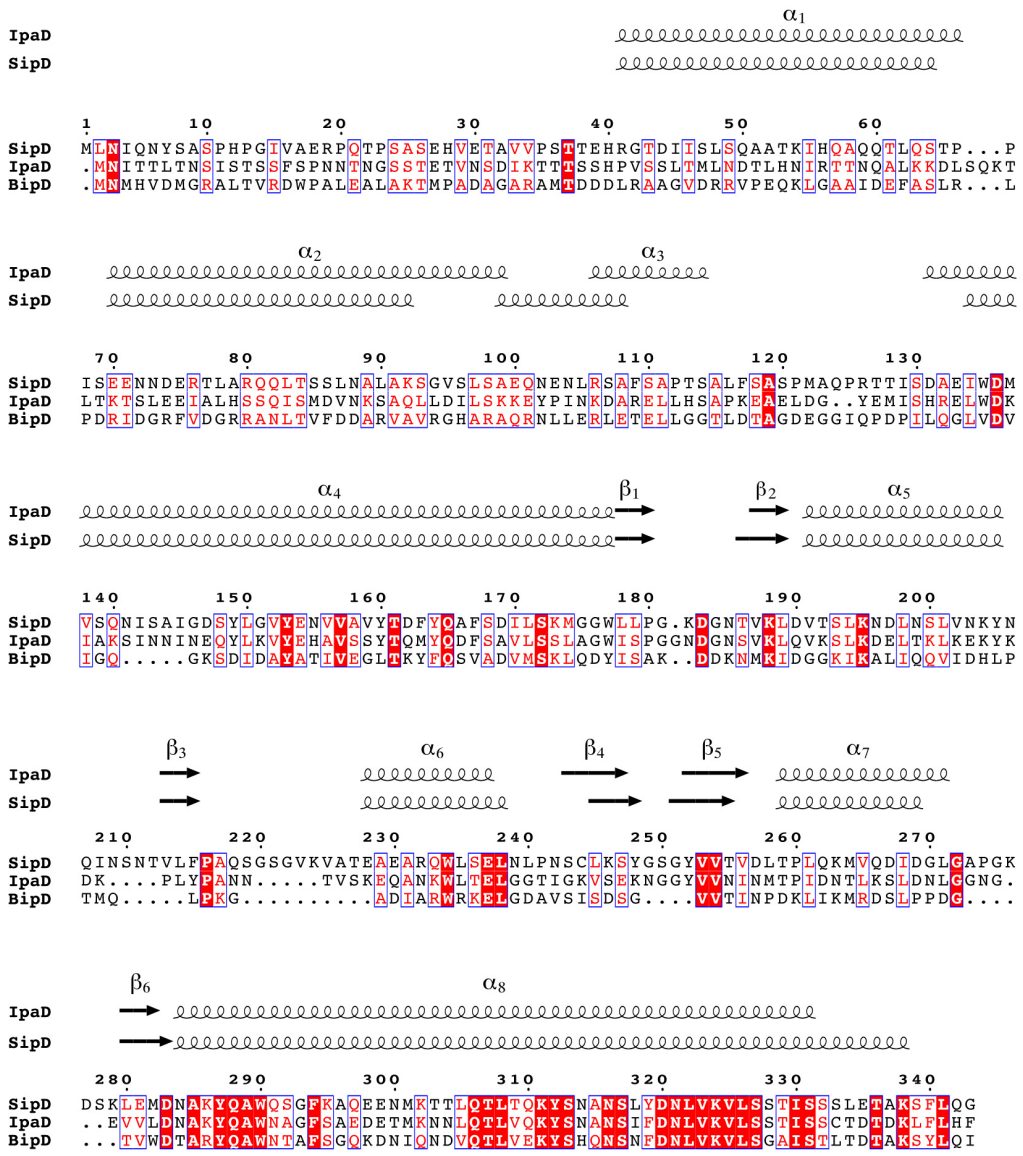


Figure 2-3. Sequences alignments of tip proteins SipD, IpaD, and BipD along with the secondary structures of IpaD[12] (PDB 2J0O) and SipD[13] (PDB 3NZZ) crystal structures. Secondary structures are representative of a composite of the asymmetric units from each protein. The sequences were aligned using ClustalW[24] and the figure was made using ESPript[25] by Dr. Yu Wang, University of Kansas.

NMR Titrations

NMR was used to identify the binding site on SipD$^{39-343}$ for bile salts (deoxycholate, taurodoxycholate, chenodeoxycholate, and cholate hydrate). Purified ^{15}N-SipD$^{39-343}$ was dialyzed into increasing concentrations of each bile salt followed by acquisition of 2D ^{1}H-^{15}N TROSY spectra at each bile concentration. The maximum ratios between bile salt and protein were limited by the critical micelle concentration (CMC) of each bile salt and by the high protein concentration requirement for quality NMR data. The bile salt CMCs were 5 mM, 2.5 mM, 7 mM, and 18 mM for deoxycholate[25], taurodeoxycholate[26], chenodeoxycholate[25], and cholate hydrate[25], respectively.

After superimposing the 2D ^{1}H-^{15}N TROSY spectra at each bile salt concentration, all but the cholate hydrate (Figure 2-7) proved to specifically interact with SipD$^{39-343}$. Upon titration with increasing quantities of deoxycholate (Figure 2-4), taurodeoxycholate (Figure 2-5), and chenodeoxycholate (Figure 2-6), many of the SipD$^{39-343}$ residues experienced gradual chemical shift changes suggesting an interaction in the fast-exchange NMR time scale. For each titration, all residues experiencing a change in chemical shift were categorized as strongly perturbed ($\Delta\delta_{HN} > 0.05$ppm) or significantly perturbed ($0.03\text{ppm} < \Delta\delta_{HN} < 0.05$ppm) based on the extent of peak migration.

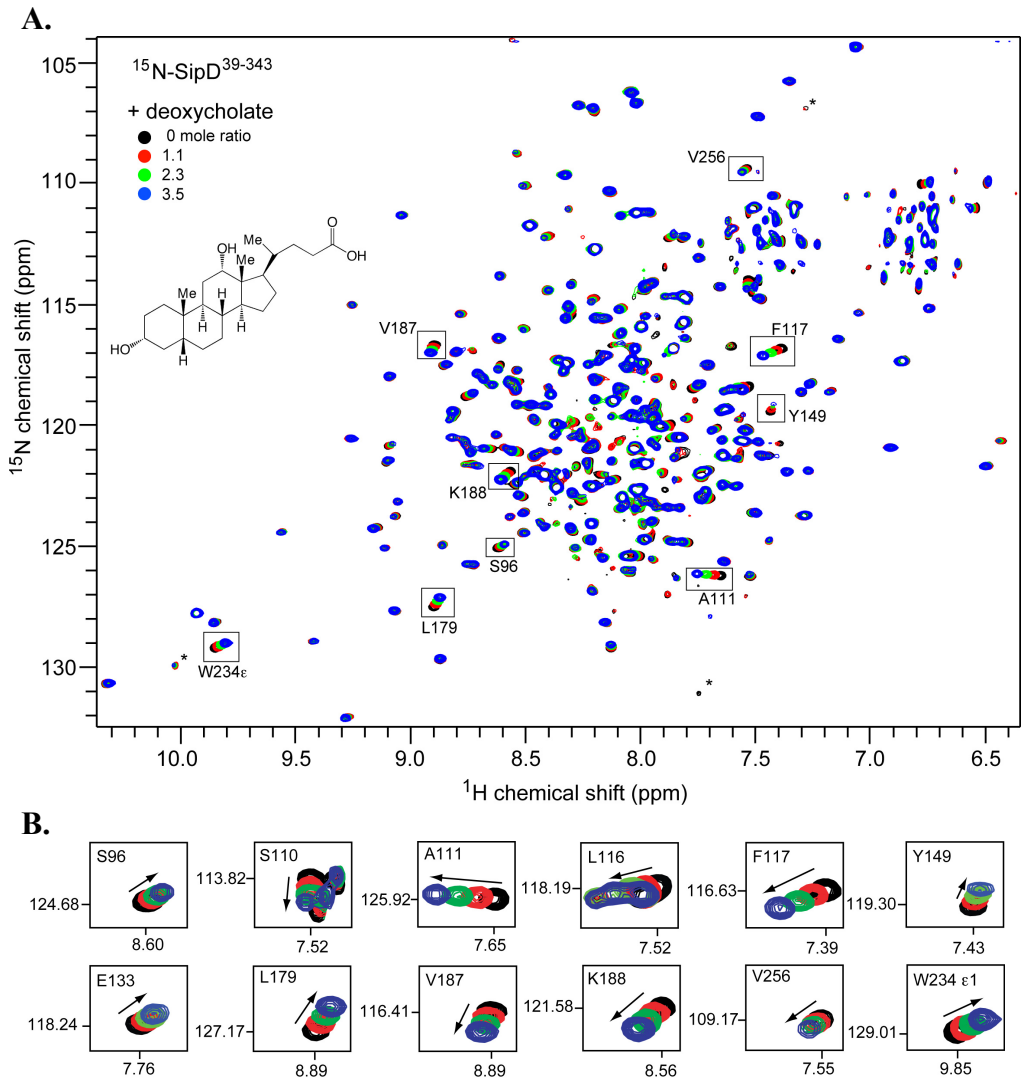


Figure 2-4. (**A**) The four overlaid 2D ^{1}H-^{15}N TROSY spectra of ^{15}N-SipD$^{39-343}$ titrated with increasing deoxycholate-protein ratios. An insert of the deoxycholate molecular structure is included, several of the affected residues are boxed with their assignment, and noise peaks are acknowledged (*). (**B**) An expanded view of perturbed residues including arrows to denote the direction of peak migration.

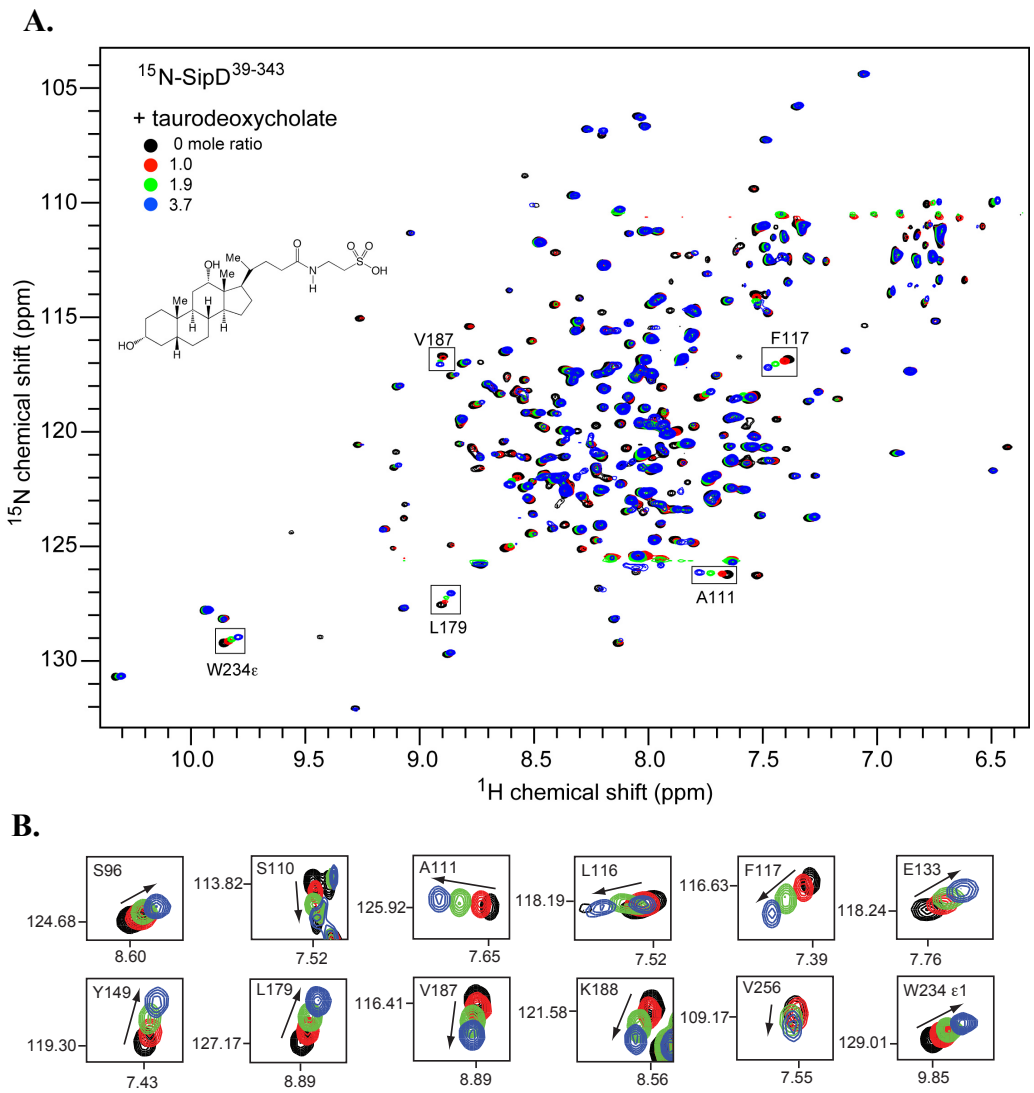


Figure 2-5. (**A**) The four overlaid 2D ^{1}H-^{15}N TROSY spectra of ^{15}N-SipD$^{39-343}$ titrated with increasing taurodeoxycholate-protein ratios. An insert of the taurodeoxycholate molecular structure is included and several of the affected residues are boxed with their assignment. (**B**) An expanded view of perturbed residues including arrows to denote the direction of peak migration.

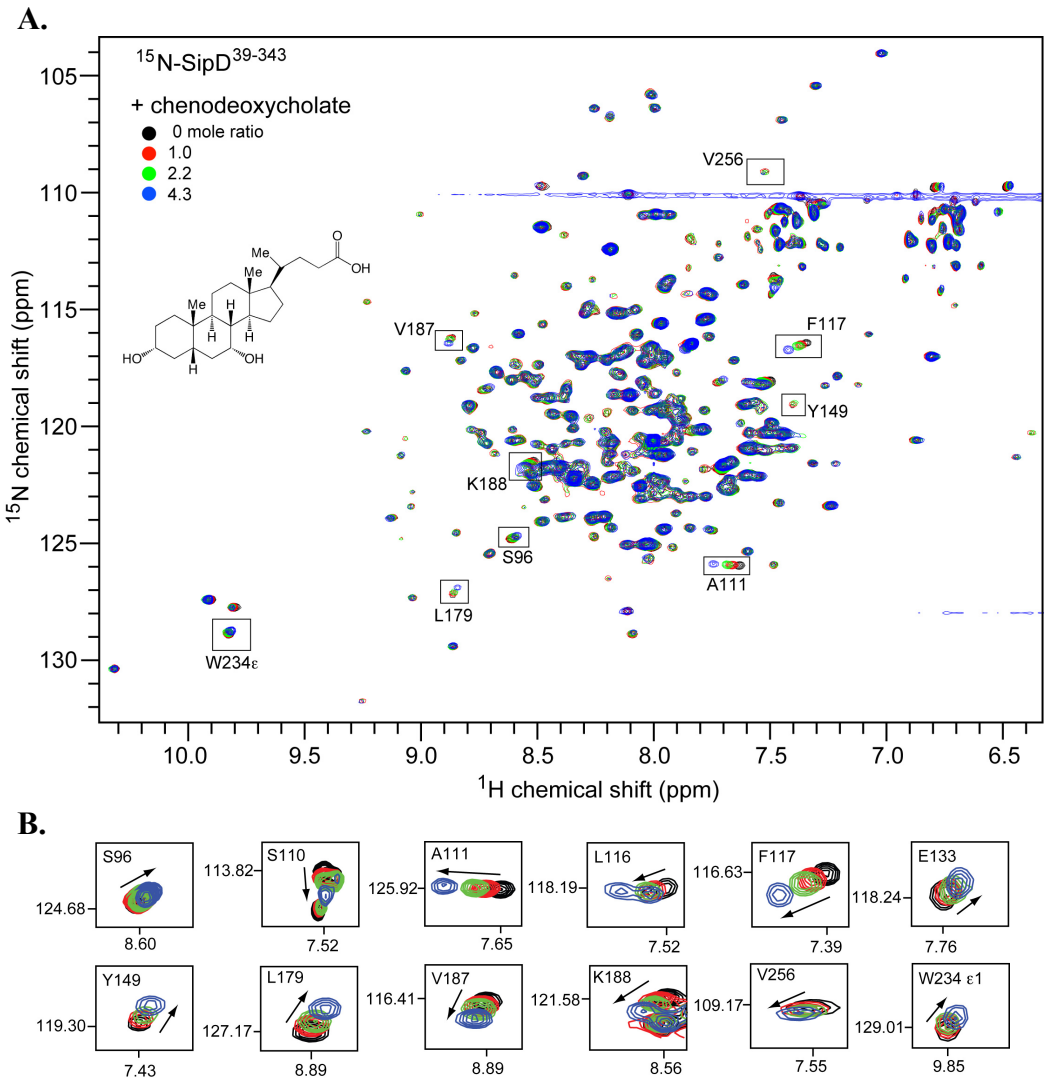


Figure 2-6. (**A**) The four overlaid 2D ^{1}H-^{15}N TROSY spectra of ^{15}N-SipD$^{39-343}$ titrated with increasing chenodeoxycholate-protein ratios. An insert of the chenodeoxycholate molecular structure is included and several of the affected residues are boxed with their assignment. (**B**) An expanded view of perturbed residues including arrows to denote the direction of peak migration.

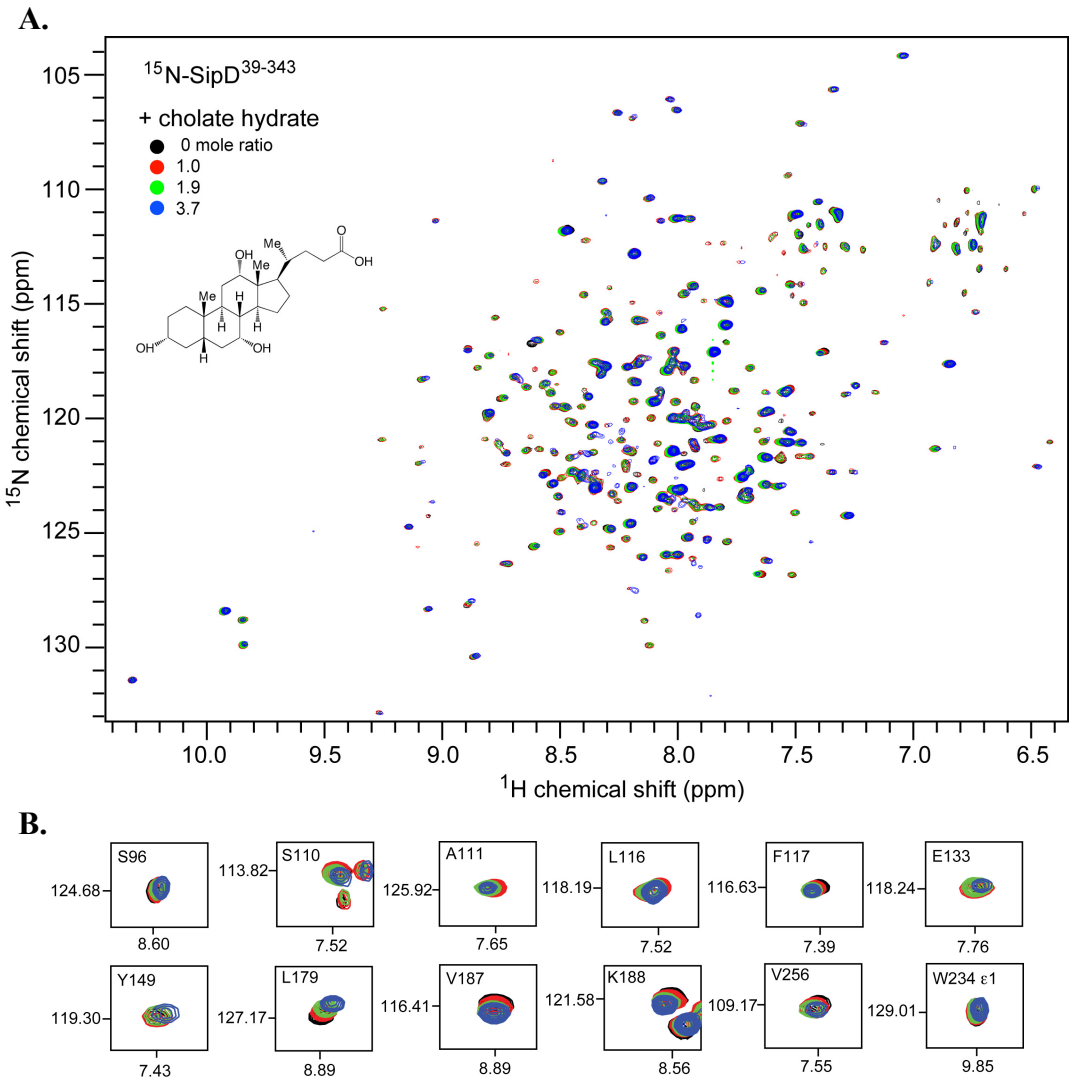


Figure 2-7. (**A**) The four overlaid 2D ^{1}H-^{15}N TROSY spectra of ^{15}N-SipD$^{39-343}$ titrated with increasing cholate hydrate-protein ratios. An insert of the cholate hydrate molecular structure is included and several of the affected residues are boxed with their assignment. (**B**) An expanded view for comparison with perturbed residues in **Figures 2-4**, **2-5**, and **2-6**.

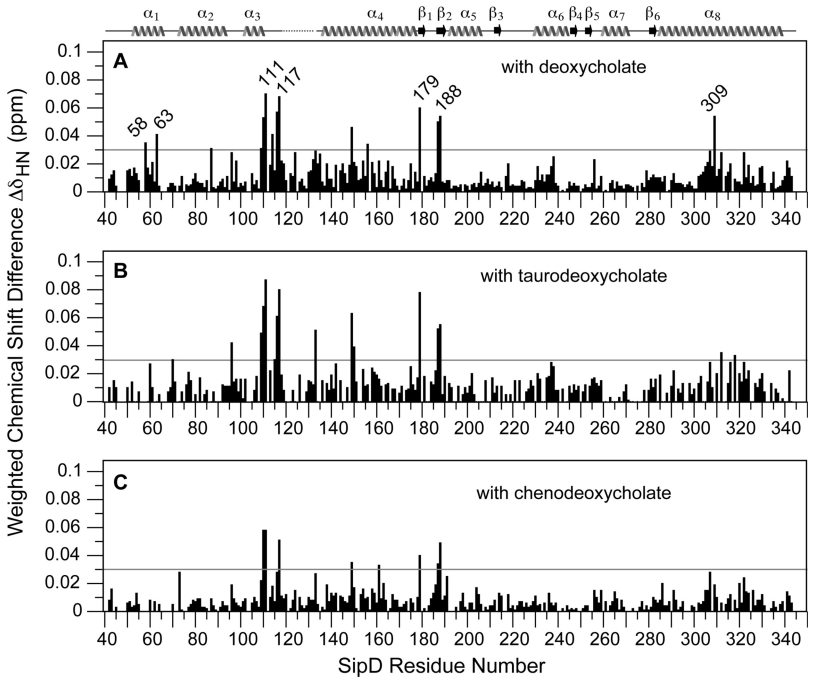


Figure 2-8. Weighted chemical shift differences ($\Delta\delta_{HN}$) of $SipD^{39-343}$ upon final titrations with deoxycholate (**A**), taurodeoxycholate (**B**), and chenodeoxycholate (**C**) ($\Delta\delta_{HN} = [{}^{1}/_{2}(\delta_H{}^2 + {}^{1}/_{25}\delta_N{}^2)]^{1/2}$ (27)). The average $\Delta\delta_{HN}$ of the deoxycholate titration was 0.011 ppm with a standard deviation (σ) of 0.011 ppm. Any $\Delta\delta_{HN}$ value more than 2σ greater than the average (or 0.033 ppm) regarded as significant. Horizontal lines (drawn at $\Delta\delta_{HN} = 0.03$ ppm) illustrate this distinction.

After quantifying the $\Delta\delta_{HN}$ for each residue in the SipD$^{39-343}$ sequence, the interaction profiles for the bile salts appeared very similar (excluding that of cholate hydrate). Deoxycholate produced the greatest chemical shift changes and the residues most strongly perturbed ($\Delta\delta_{HN}$ > 0.05ppm) were S110, A111, L116, F117, Y149, L179, V187, K188, and T309 (Figure 2-8A). Other significantly perturbed residues (0.05ppm > $\Delta\delta_{HN}$ > 0.03ppm) included A58, Q63, L87, F109, and S114 (Figure 2-8A). Residues strongly perturbed by taurodeoxycholate ($\Delta\delta_{HN}$ > 0.05ppm) were similar to those from deoxycholate with the addition of E133 and the exception of T309 (Figure 2-8B). Taurodeoxycholate also significantly perturbed (0.05ppm > $\Delta\delta_{HN}$ > 0.03ppm) residues E70, S96, F109, A115, L150, Y312, and L318 (Figure 2-8B). Finally, while chenodeoxycholate produced an interaction profile similar to that of deoxycholate and taurodeoxycholate, its effects on chemical shift changes were much less intense. Chenodeoxycholate produced strong perturbations ($\Delta\delta_{HN}$ > 0.05ppm) on residues S110, A111, and F117 and significantly perturbed (0.05ppm > $\Delta\delta_{HN}$ > 0.03ppm) residues Y149, T161, L179, V187, and K188 (Figure 2-8C).

Using the published crystal structure of SipD$^{39-343}$, residues that were strongly perturbed cluster into two regions of the protein (Figure 2-9). The larger of these clusters sits primarily in the loop between residues 110 and 134 and also includes the proximal Y149 on helix α_4 and the W234 side chain (deoxycholate and chenodeoxycholate only) on helix α_6. The smaller strongly perturbed cluster sits distally in β strands β_1 and β_2. Finally, a set of minor perturbations occurs in the N-terminal 2-helix bundle including A58, Q63, and L87 (deoxycholate only) as well as E70 and S96 (chenodeoxycholate only).

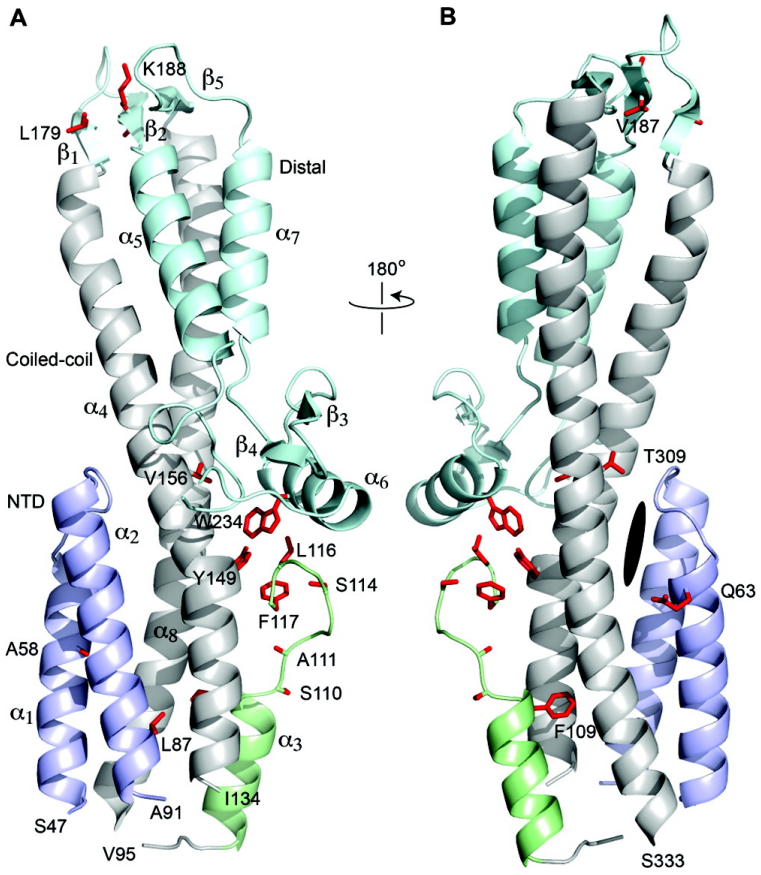


Figure 2-9. Ribbon illustration of the SipD crystal structure.[13] Residues with significant chemical shift changes ($\Delta\delta_{HN} > 0.03$ppm) from deoxycholate titrations are depicted in red and the corresponding region of binding in IpaD is represented as a black oval (as predicted by computer docking). Ribbon coloration illustrates the N-terminal to C-terminal transition.

Isothermal Titration Microcalorimetry

The protein and bile salt concentrations used in the microcalorimetric experiments were based on the CMC limitations of deoxycholate as well as the "experimental K window" (Figure 2-11) that would permit accurate K_D calculations. To avoid significant heat exchange from micelle disruption as the concentrated salt was injected and diluted in the adiabatic cell, the injection syringe was loaded with 2.5 mM deoxycholate (ca. half of the reported CMC[26]). Figures 2-10a and 2-10b illustrate a control titration performed by injecting deoxycholate into buffer alone (100 mM NaCl, 20 mM $NaHPO_4$, pH 7.5). While the starting bile salt concentration was below the reported CMC, there was still an appreciable quantity of heat absorbed (~7 kcal/mol) by initial injections, suggesting that even 2.5 mM deoxycholate retains a small number of micelles that dissociate upon injection. Since the K_D of the $SipD^{39-343}$-deoxycholate interaction has been reported (using fluorescence polarization) to be in the micromolar range[7], the critical parameter c of the isotherm was already in the low end of the experimental K window (Figure 2-11**)**. Therefore, the background heat exchanged from micelle dissociation could not be further reduced by using lower concentrations of deoxycholate. Figure 2-12 illustrates the resulting isotherm from a titration of 380 μM $SipD^{39-343}$ with 2.5 mM deoxycholate. The heat exchanged from interaction of the tip protein and bile salt clearly does not outweigh that of the micelle dissociation. Combined, these two events convolute the calorimetric titration curve, which does not resemble a typical isotherm. Subtracting the heat absorbed by micelle dissociation from each injection, Figure 2-13 shows the contribution of $SipD^{39-343}$-deoxycholate binding to the overall isotherm. The curve could not be fit to a one- or two-site binding model (chi^2>4500). Additional trials testing other salt

concentrations, pH, buffers, and DTT (10 mM NaCl, 10 mM NaHPO4, pH 7.0; 100 mM NaCl, 20 mM Tris/HCl, pH 8.0; 100 mM NaCl, 1.0 mM DTT, 20 mM NaHPO4, pH 7.5) resulted in very small heat exchange that could not be fit to one-site or two-site models (data not shown). Thus, ITC is not a useful method to determine the SipD-bile interaction.

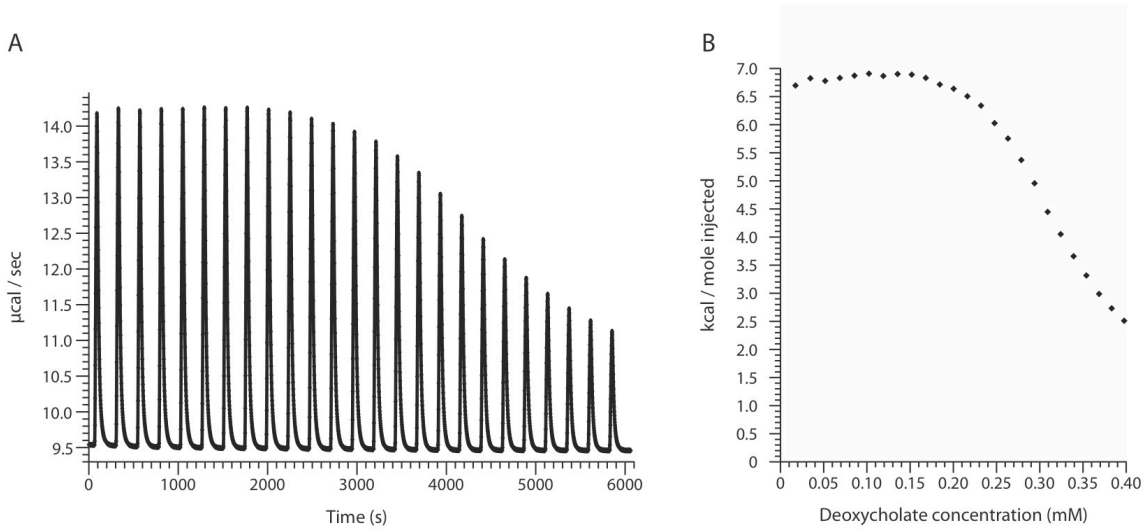
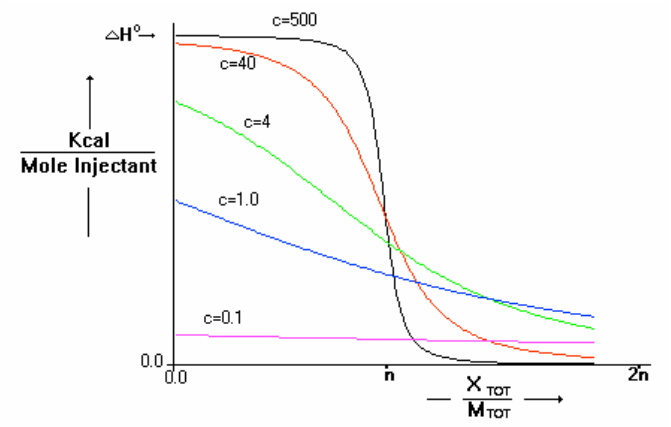


Figure 2-10. (A) Raw data from a series of 25 10µL injections of 2.5mM deoxycholate into buffer at 25°C. (B) Titration curve based on integration of raw injection heats with respect to time.



$$\mathbf{c = KM_{tot}n}$$

K - binding association constant
$\mathbf{M_{tot}}$ – initial macromolecule concentration
n - stoichiometric ratio

Figure 2-11. The experimental K window **(**VP-ITC User's Manual, MicroCal**)** illustrates the range of the unitless constant c permitting ideal measurement of the binding association constant K_A. Ideal isotherms for K_D approximation have a value of c between 5 and 500 **(**VP-ITC User's Manual, MicroCal**)**.

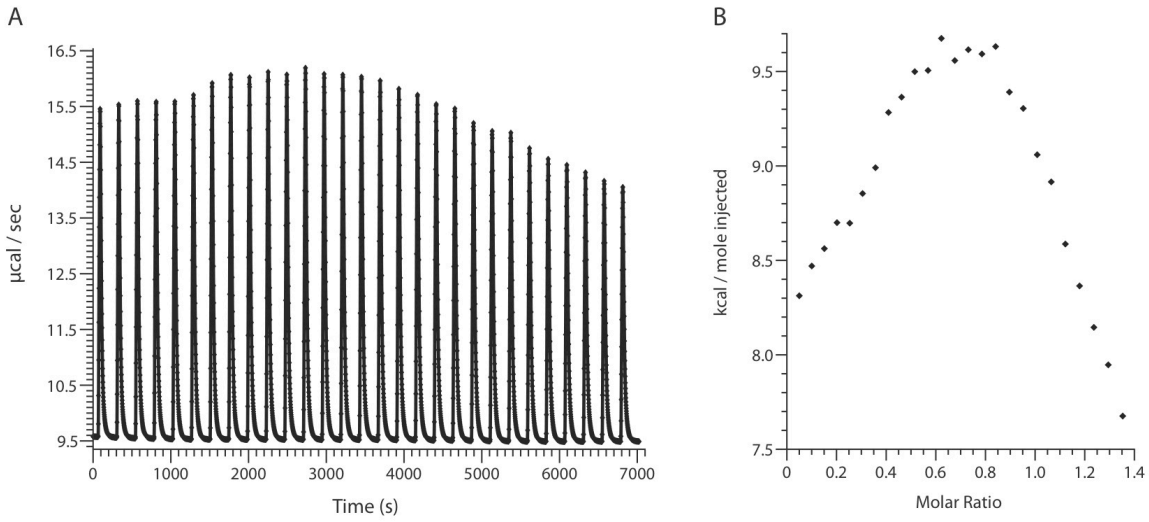


Figure 2-12. (A) Raw data from 10µL injections of 2.5mM deoxycholate into 380µM SipD$^{39-343}$ at 25°C. (B) Titration curve based on integration of raw injection heats with respect to time without subtracting background heat of micelle dissociation.

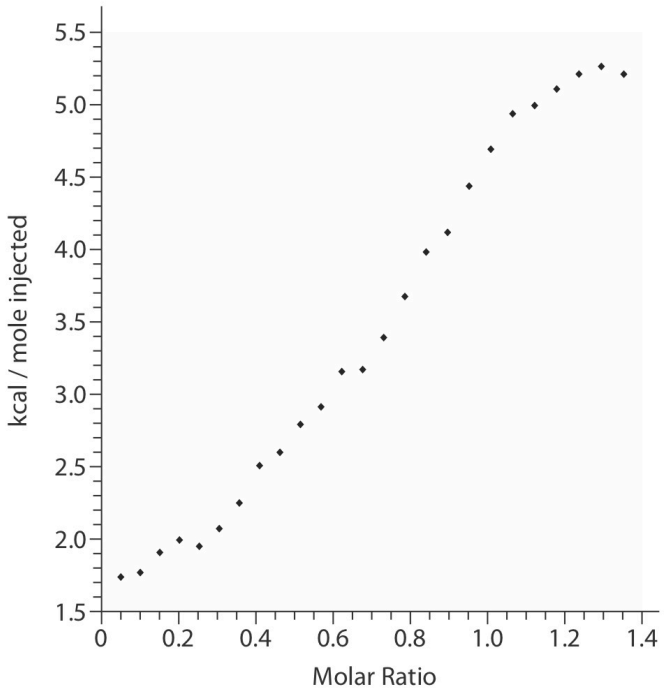


Figure 2-13. Integrated curve of SipD$^{39-343}$ titrated with 2.5 mM deoxycholate after subtraction of background heat from micelle dissociation.

Salmonella Invasion Assay

To identify SipD residues that may be important for bacterial invasion or interaction with bile salts, a *Salmonella* invasion assay was used to assess the ability of SipD mutants to invade cultured human intestinal epithelial cells. A set of alanine point mutations of SipD were designed at residues strongly perturbed by NMR titrations with deoxycholate or significantly perturbed by more than one bile salt. These included perturbed residues on or proximal to the 110-134 loop, on the distal β1 and β2 strands, and others (shown in Figure 2-14). Relative percent invasion was determined with reference to the SipD-rescued strain in the absence of deoxycholate. Wild-type and SipD-rescued (*sipD*$^{+}$) strains of *Salmonella* both showed decreased invasiveness when incubated in the presence of deoxycholate. Of the 15 tested mutants, only 2 (L116A and Y149A) were found to maintain similar invasiveness in the presence and absence of deoxycholate (within 1 standard deviation). However, as in the case of all but 5 mutations, the L116A and Y149A exhibited reduced invasion levels in comparison to the referenced SipD-rescue. Therefore, while these 2 alanine substitutions may have abolished a deoxycholate binding site, residues L116 and Y149 may also play another functional role in *Salmonella* invasion. On the other hand, the 5 previously mentioned mutations (S96A, S110A, F117A, Q165A, and V256A) that had no effect on invasion efficiency in deoxycholate absence and failed to abolish deoxycholate mediated invasion reduction are likely of less functional significance.

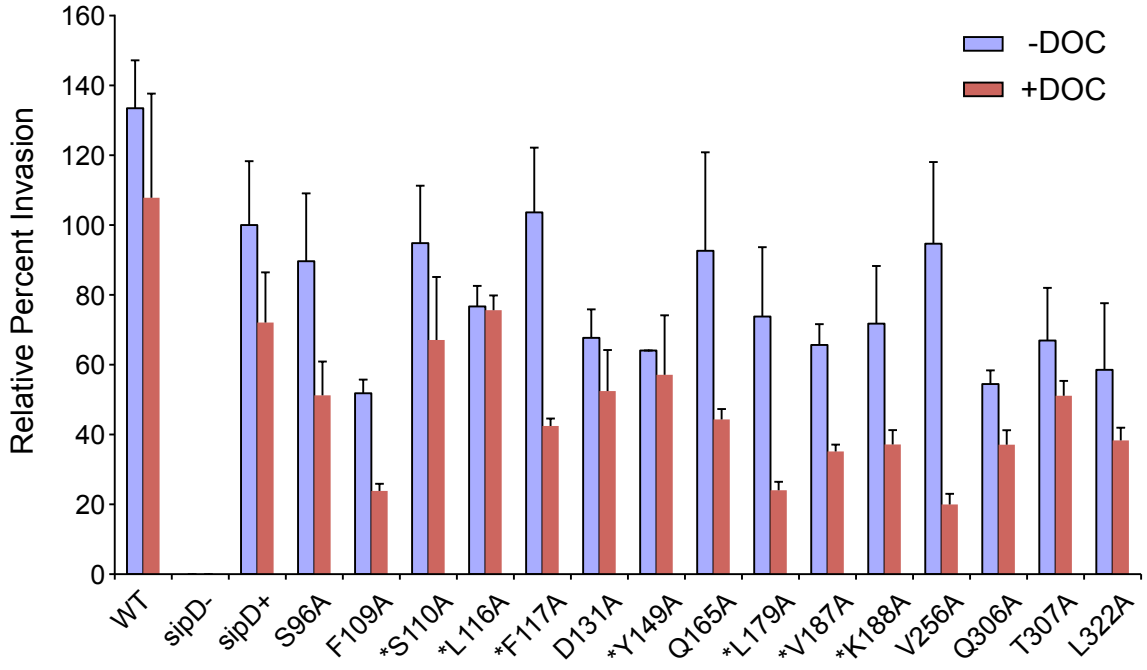


Figure 2-14. Results from the *Salmonella* invasion assay (legend: WT, wild-type strain SL1344, *sipD*$^{-}$, *sipD* null strain; *sipD*$^{+}$, wild-type *sipD* from plasmid pRK2-*sipD* introduced into null strain; S96A, F109A, S110A, L116A, F117A, D131A, Y149A, Q165A, L179A, Q165A, L179A, V187A, K188A, V256A, Q306A, T307A, and L322A are single point mutants of pRK2-*sipD* introduced into the null strain. Bar color indicates the presence or absence of deoxycholate (DOC) in the bacterial culture incubation step. Mutants of residues strongly perturbed in the NMR titration with DOC are indicated (*).

DISCUSSION

While previous studies have indicated that *Salmonella* and *Shigella* have opposing responses to bile salt exposure[7-10], a model detailing the differing mechanism of regulation remains unclear. Others have proposed that bile salts initiate a T3SS transcriptional repression cascade in *Salmonella* through the SirA-BarA two-component system[9], however Stensrud et. al[7] have shown that *Salmonella* and *Shigella* interact directly with deoxycholate with their respective tip-proteins, SipD and IpaD. While a deoxycholate binding site for IpaD has been proposed via computer docking[7], this study provides the first experimental identification of a bile salt binding site in SipD. This is also the first case of successful NMR characterization of a T3SS tip protein, which was previously impeded due to the large number of residues, poor solubility in NMR buffer, and disordered termini common to tip proteins. Based on the apparent disorder of homologous regions of BipD[12; 28] and IpaD[12] crystals, a 38-residue truncation from the SipD N-terminus was constructed and provided an excellent NMR spectrum for assignment. Nearly all of the SipD$^{39-343}$ backbone amide peaks were assigned using perdeuteration, TROSY, and amino acid-specific labeling. These included all of the backbone peaks showing changes in chemical shifts from the bile salt titrations (Figures 2-4, 2-5, and 2-6). The tryptophan side-chain peaks were also assigned, including W234 which was perturbed by deoxycholate and taurodeoxycholate, by individual tyrosine substitutions.

After mapping the residues perturbed by bile salts on the crystal structure of SipD$^{39-343}$ (Figure 2-9), a majority of the residues strongly perturbed by deoxycholate, taurodeoxycholate, and chenodeoxycholate were found in two clusters. The minor

cluster, albeit much smaller, is located in the β1 (L179) and β2 (V187 and K188) strands. The major cluster is found far from the minor cluster in the 110-134 loop (S110, A111, L116, F117) or facing the loop in the α4 (Y149) or α6 (W234-sc) helices. These NMR results suggest that the primary bile salt binding site in SipD sits between the 110-134 loop and the central coiled coil. Chemical shift changes in the distal β1 and β2 strands are likely due to small conformational changes or nonspecific interaction. In comparison to a previous study of the bile salt binding pocket in the homologous *Shigella* tip protein, IpaD[7], the proposed binding site in SipD is oriented on the opposite side of the central coiled-coil. Stensrud et. al[7] used two approaches to model deoxycholate binding in IpaD. FRET analysis suggested binding near the midpoint of the oblong protein and computer simulated binding (Autodock) identified a binding pocket between the N-terminal two-helix bundle and the central coiled coil. While the NMR proposed binding pocket for SipD also lies near the midpoint of the central coiled coil, there were no strong chemical shift changes in the region homologous to the proposed IpaD binding site. Only four residues in helices α1, α2, and α8 (A58, Q63, L87, and T309) were perturbed by deoxycholate titrations, however to a much less extent as that seen in the primary cluster.

To further examine the functional role of $SipD^{39-343}$ residues perturbed by NMR titrations, a *Salmonella* invasion assay was designed to identify mutants that could maintain wild-type invasion efficiency while abolishing deoxycholate binding. None of the mutants were able to fit both criterion, however, two mutants (L116A and Y149A) showed similar invasion levels in the presence and absence of deoxycholate. These residues both reside in the primary cluster of residues perturbed by NMR titrations with deoxycholate and in close proximity to one another. The remainder of the mutants tested,

including those in the region equivalent to the predicted deoxycholate binding site of IpaD, were unable to abolish deoxycholate mediated invasion reduction. Based on these observations (in combination with NMR titration results), it was concluded that SipD and IpaD interact with bile salts at different surfaces, despite strong similarity in their tertiary structures (C^{α} rmsd = 1.4 Å) (Figure 2-3). We propose that differences in bile salt binding between SipD and IpaD hold a causal relationship with the different responses of *Salmonella* and *Shigella* to bile salts. Both of these tip proteins have been shown to be present at the bacterial surface prior to host cell contact and hold a logical position for an environmental sensor.[4; 30]

Despite the successful identification of a bile salt binding site in $SipD^{39-343}$, the use of truncated SipD thwarted an analysis of bile salt interaction with the N-terminal 38 residues. However, the importance of this N-terminal region has previously been attributed to its necessity for secretion of SipD through the T3SS needle.[31; 32] Another limitation in the experiments herein was the inability to determine the K_D of the $SipD^{39-343}$ interaction with bile salts by means of ITC or NMR titrations due to the CMCs of bile salts. For ITC, the deoxycholate concentration in the injection syringe could not be reduced below the threshold for micelle formation without affecting the accuracy of K_D estimation from the isotherm. With two distinct processes occurring in the adiabatic cell (SipD-deoxycholate and micelle disruption), it is clear why the resulting isotherm appears convoluted.

REFERENCES

1. Cornelis, G. R. (2006) The type III secretion injectisome, *Nat. Rev. Microbiol. 4*, 811-825.
2. Kaniga, K., Trollinger, D., and Galan, J. E. (1995) Identification of two targets of the type III protein secretion system encoded by the inv and spa loci of Salmonella typhimurium that have homology to the Shigella IpaD and IpaA proteins, *J.Bacteriol. 177*, 7078-7085.
3. Mueller, C. A., Broz, P., Muller, S. A., Ringler, P., Erne-Brand, F., Sorg, I., Kuhn, M., Engel, A., and Cornelis, G. R. (2005) The V-antigen of Yersinia forms a distinct structure at the tip of injectisome needles, *Science 310*, 674-676.
4. Espina, M., Olive, A. J., Kenjale, R., Moore, D. S., Ausar, S. F., Kaminski, R. W., Oaks, E. V., Middaugh, C. R., Picking, W. D., and Picking, W. L. (2006) IpaD localizes to the tip of the type III secretion system needle of *Shigella flexneri*, *Infect. Immun. 74*, 4391-4400.
5. Stevens, M. P., Haque, A., Atkins, T., Hill, J., Wood, M. W., Easton, A., Nelson, M., Underwood-Fowler, C., Titball, R. W., Bancroft, G. J., and Galyov, E. E. (2004) Attenuated virulence and protective efficacy of a *Burkholderia pseudomallei bsa* type III secretion mutant in murine models of melioidosis, *Microbiology 150*, 2669-2676.
6. Picking, W. L., Nishioka, H., Hearn, P. D., Baxter, M. A., Harrington, A. T., Blocker, A., and Picking, W. D. (2005) IpaD of Shigella flexneri is independently required for regulation of Ipa protein secretion and efficient insertion of IpaB and IpaC into host membranes, *Infect. Immun. 73*, 1432-1440.
7. Stensrud, K. F., Adam, P. R., La Mar, C. D., Olive, A. J., Lushington, G. H., Sudharsan, R., Shelton, N. L., Givens, R. S., Picking, W. L., and Picking, W. D. (2008) Deoxycholate interacts with IpaD of *Shigella flexneri* in inducing the recruitment of IpaB to the type III secretion apparatus needle tip, *J. Biol. Chem. 283*, 18646-18654.
8. Olive, A. J., Kenjale, R., Espina, M., Moore, D. S., Picking, W. L., and Picking, W. D. (2007) Bile salts stimulate recruitment of IpaB to the *Shigella flexneri* surface, where it colocalizes with IpaD at the tip of the type III secretion needle, *Infect. Immun. 75*, 2626-2629
9. Prouty, A. M., and Gunn, J. S. (2000) *Salmonella enterica* serovar typhimurium invasion is repressed in the presence of bile, *Infect. Immun. 68*, 6763-6769.
10. Prouty, A. M., Brodsky, I. E., Manos, J., Belas, R., Falkow, S., and Gunn, J. S. (2004) Transcriptional regulation of Salmonella enterica serovar Typhimurium genes by bile, *FEMS Immunol. Med. Microbiol. 41*, 177-185.
11. Cabral, D. J., and D. M. Small (1989) Physical chemistry of bile, p. 621–662. In S. G. Schultz, J. G. Forte, and B. R. Bauner (ed.), Handbook of physiology, vol. 3. American Physiology Society, Bethesda, MD.
12. Johnson, S., Roversi, P., Espina, M., Olive, A., Deane, J. E., Birket, S., Field, T., Picking, W. D., Blocker, A. J., Galyov, E. E., Picking, W. L., and Lea, S. M. (2007) Selfchaperoning of the type III secretion system needle tip proteins IpaD and BipD, *J. Biol. Chem. 282*, 4035-4044.

13. Chatterjee, S., Zhong, D., Nordhues, B. A., Battaile, K. P., Lovell, S., and De Guzman, R. N. (2011) The crystal structures of the *Salmonella* type III secretion system tip protein SipD in complex with deoxycholate and chenodeoxycholate, *Protein Sci.* 1, 75-86.
14. Geisbrecht, B. V., Bouyain, S., and Pop, M. (2006) An optimized system for expression and purification of secreted bacterial proteins, *Protein Expr. Purif. 46*, 23-32.
15. Wang, Y., Nordhues, B. A., Zhong, D., and De Guzman, R. N. (2010) NMR characterization of the interaction of the *Salmonella* type III secretion system protein SipD and bile salts, *Biochemistry* 49, 4220-4226.
16. Kenjale, R., Wilson, J., Zenk, S. F., Saurya, S., Picking, W. L., Picking, W. D., and Blocker, A. (2005) The needle component of the type III secretion of *Shigella* regulates the activity of the secretion apparatus, J. Biol. Chem. 280, 42929-37.
17. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277-293.
18. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules, *Methods Mol. Biol. 278*, 313-352.
19. Czisch, M., and Boelens, R. (1998) Sensitivity enhancement in the TROSY experiment, *J. Magn. Reson. 134*, 158-160.
20. Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wuthrich, K. (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins, *Proc. Natl. Acad. Sci. USA 95*, 13585-13590.
21. Salzmann, M., Wider, G., Pervushin, K., Senn, H., and Wuthrich, K. (1999) TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins, *J. Am. Chem. Soc. 121*, 844-848.
22. Osiecki, J. C., Barker, J., Picking, W. L., Serfis, A. B., Berring, E., Shah, S., Harrington, A., and Picking, W. D. (2001) IpaC from *Shigella* and SipC from *Salmonella* possess similar biochemical properties but are functionally distinct, *Mol. Microbiol.* 42, 469-481.
23. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequences alignment through sequence weightin, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res.* 22, 4673-80.
24. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript, *Bioinformatics* 15, 305-308.
25. Simonovic, B. R., and Momirovic, M. (1997) Determination of critical micelle concentration of bile acid salts by micro-calorimetric titration, *Mikrochimica Acta 127*, 101-104.
26. DeLong, L. J., and Nichols, J. W. (1996) Time-resolved fluorescence anisotropy of fluorescent-labeled lysophospholipid and taurodeoxycholate aggregates, *Biophys J. 70*, 1466-1471.
27. Grzesiek, S., Bax, A., Clore, G. M., Gronenborn, A. M., Hu, J. S., Kaufman, J., Palmer, I., Stahl, S. J., and Wingfield, P. T. (1996) The solution structure of HIV-1 Nef reveals an unexpected fold and permits delineation of the binding surface for the SH3 domain of Hck tyrosine protein kinase, *Nat. Struct. Biol. 3*, 340-345.

28. Erskine, P. T., Knight, M. J., Ruaux, A., Mikolajek, H., Sang, N. W., Withers, J., Gill, R., Wood, S. P., Wood, M., Fox, G. C., and Cooper, J. B. (2006) High Resolution Structure of BipD: An Invasion Protein Associated with the Type III Secretion System of *Burkholderia pseudomallei*, *J. Mol. Biol. 363*, 125-136.
29. van Velkinburgh, J. C., and Gunn, J. S. (1999) PhoP-PhoQ-regulated loci are required for enhanced bile resistance in *Salmonella* spp, *Infect. Immun. 67*, 1614-1622.
30. Lara-Tejero, M., and Galan, J. E. (2009) *Salmonella enterica* serovar typhimurium pathogenicity island 1-encoded type III secretion system translocases mediate intimate attachment to nonphagocytic cells, *Infect. Immun. 77*, 2635-2642.
31. Karavolos, M. H., Roe, A. J., Wilson, M., Henderson, J., Lee, J. J., Gally, D. L., and Khan, C. M. (2005) Type III secretion of the Salmonella effector protein SopE is mediated via an N-terminal amino acid signal and not an mRNA sequence. J. Bacteriol. 187, 1559–1567.
32. Arnold, R., Brandmaier, S., Kleine, F., Tischler, P., Heinz, E., Behrens, S., Niinikoski, A., Mewes, H. W., Horn, M., and Rattei, T. (2009) Sequence-based prediction of type III secreted proteins. *PLoS Pathog.* 5, e1000376.

Chapter 3: PcrG is a Partially Structured Protein

INTRODUCTION

Pseudomonas aeruginosa is an opportunistic Gram-negative pathogen that is most frequently associated with nosocomial infections in patients with cystic fibrosis, AIDS, cancer, or other ailments that compromise the immune system.[1] To initiate infection, *P. aeruginosa* utilizes the type III secretion system (T3SS) to inject effector proteins into host cells.[2] Four of these effector proteins have been identified (ExoT/U/S/Y)[3] and have been shown to effectively evade primary immune response by destroying macrophages and polymorphonuclear neutrophils (PMNs)[4]. In fact, the simple assembly of the T3SS translocon in these cells has been demonstrated to perforate cell membranes sufficiently for oncotic cell death.[5] Two proteins that have been discovered to play critical roles in assembly and regulation of the T3SS apparatus in *P. aeruginosa* are the tip protein, PcrV, and its proposed chaperone, PcrG.[6; 7]

Effector proteins in *P. aeruginosa* are not secreted until induced by either cell contact or calcium depletion.[4] Using a β-galactosidase assay, Lee et. al[8] demonstrated that deletion of either PcrG or PcrV results in constitutive effector secretion, even in the absence of calcium depletion or cell contact, and that the effect of these deletions is additive. In the same study, PcrV export and assembly at the needle tip was shown to be critical for effector secretion regulation and the presence of PcrG is required for efficient PcrV secretion. Results from several labs have confirmed that PcrG and PcrV directly interact[9; 10], however, a point mutant disrupting this interaction revealed that secretion control is still maintained[8]. Therefore, a model has been proposed where PcrG plays a role in facilitating PcrV secretion, but also holds a role in effector secretion regulation

that is independent of PcrV-binding. This model is notably different from the current model for the highly sequentially related homologs in the *Yersinia* T3SS (LcrG and LcrV, respectively).[11]

Two unique LcrV binding domains on LcrG have been separately proposed by yeast two-hybrid studies[12] and NMR chemical shift mapping[13]. One aim of this study was to identify whether the PcrV binding domain on PcrG is conserved with previously proposed LcrV binding domains on LcrG. Instead, NMR chemical shift mapping suggested that PcrV induces a global change in PcrG structure that likely results as a combination of PcrV-binding interactions and binding-induced conformational changes. In addition, the first experimentally derived information on sequence specific secondary structure identified the regions of PcrG that adopt either α–helical or random coil conformation.

MATERIALS AND METHODS

Protein Expression and Purification

Two constructs of wild-type *P. aeruginosa* PcrG (full length and residues 9-76) were PCR amplified from cDNA and subcloned into a pET-21a expression vector with a C-terminal fusion to a TEV protease cleavage site and His_6-tagged GB1 domain. Resulting plasmids were expressed in BL21(DE3)-DNAY competent *E. coli* cells as follows: Recently transformed cells were grown in 10 mL starter cultures of LB media overnight at 37 °C followed by centrifugation at 4000 rpm. For unlabeled protein, cell pellets were resuspended in 1.0 L LB; whereas isotopically labeled protein (^{15}N,^{13}C and ^{15}N) was attained by resuspension into 1.0 L M9 minimal media supplemented with 2.0 g/L ^{13}C D-glucose and/or 1.0 g/L $^{15}NH_4Cl$. All media contained 30 μg/mL kanamycin and 100 μg/mL carbenicillin. Cells were then incubated at 37 °C, induced with 1.0 mM isopropyl-β-D-thiogalactopyrandoside (IPTG) at an O.D. of ~0.8, and cell growth was continued overnight at 15°C with aeration to a final O.D. of ~2.5. Cells were harvested by centrifugation, resuspended in binding buffer (500 mM NaCl, 20 mM Tris-HCl, 5 mM imidazole, pH 8.0), and lysed by sonication in the presence of 10 mg phenylmethyl sulfonyl fluoride (PMSF). The lysate was centrifuged again to remove cell debris and the supernatant was loaded onto a Ni^{2+}-affinity column. The Ni^{2+}-affinity column was washed with binding buffer and eluted with 250 mM imidazole. For NMR and CD spectroscopy application, pooled elutes were digested with recombinant Tobacco Etch Virus (TEV) protease[14] and a second Ni^{2+}-affinity purification as described[15]. Cleavage with TEV protease resulted in a C-terminal cloning artifact (-GSENLYFQ). Protein was

concentrated using Amicon Ultra 3K (Millipore) and concentration was determined by the Bradford method.

Wild-type *P. aeruginosa* PcrV (residues 25-294) was subcloned into a pET-21a expression vector with an N-terminal His_6-tagged GB1 domain fusion. The N-terminal 24 residues of PcrV were truncated based on the apparent disorder of the homologous region in the LcrV crystal structure[22]. Unlabeled protein was expressed and purified as described above including TEV proteolysis. TEV cleavage of the $PcrV^{25-294}$ construct resulted in an N-terminal cloning artifact (GHM-).

NMR Spectroscopy

NMR data were collected at 25°C using a Bruker Avance 800 MHz (at the University of Kansas) and a Varian 900 MHz (at the Rocky Mountain Regional 900 MHz NMR Facility, University of Colorado) spectrophotometer equipped with a cryoprobe, processed with NMRPipe[16], and analyzed using NMRView[17]. All protein samples were dialyzed overnight in NMR buffer (10 mM $NaHPO_4$, 10 mM NaCl, 10% (v/v) D_2O, pH 7.0). ^{15}N-$PcrG^{FL}$ (0.8 mM) and truncated ^{15}N-$PcrG^{9-76}$ (0.7 mM) were used to acquire initial 2D ^{1}H-^{15}N HSQC spectra[18]. For assignment of the ^{15}N-$PcrG^{9-76}$ carbon backbone, 0.6 mM ^{15}N,^{13}C-$PcrG^{9-76}$ was used to collect 2D ^{1}H-^{15}N HSQC, 3D HNCA, 3D HNCACB, and 3D CBCA(CO)NH.[18; 19; 20] Secondary structure was determined using the secondary chemical shifts of C^{α}, C^{β}, and C'.[21] For NMR chemical shift mapping, 2D ^{1}H-^{15}N HSQC spectra[18] were acquired for ^{15}N-$PcrG^{9-76}$ and ^{15}N-$PcrG^{9-76}$-GB1-6His titrated with increasing amounts of $PcrV^{25-294}$. Six samples were prepared in NMR buffer with 0.13 mM ^{15}N-$PcrG^{9-76}$ and varying $PcrV^{25-294}$ (0.00, 0.04, 0.08, 0.16, 0.32, and 0.65 mM).

Four samples were prepared in NMR buffer with 0.39 mM ^{15}N-PcrG$^{9-76}$-GB1-His$_6$ and varying PcrV$^{25-294}$ (0.00, 0.20, 0.40, and 0.80 mM).

Circular Dichroism Spectroscopy

Full length PcrG and the truncated construct, PcrG$^{9-76}$, were cleaved by TEV protease, concentrated to 5.8 μM, and dialyzed in buffer (10 mM $NaHPO_4$, 10 mM NaCl, pH 7.0) for all CD spectroscopy and thermal denaturation experiments. Using a JASCO J-815 Spectropolarimeter, CD spectra were collected in triplicate from 195 to 260 nm at 50nm/min. Thermal denaturation curves were observed at 222 nm from 20-80°C at a ramp rate of 30°C/h. Control curves from buffer alone were subtracted in both experiments.

Surface Plasmon Resonance Spectroscopy

Apparent dissociation constants (K_D) between PcrG^{FL} or PcrG$^{9-76}$ and PcrV$^{25-294}$ were measured by means of surface plasmon resonance (SPR) using a Biacore 3000 instrument (Biacore, Uppsala, Sweden). A 1:1 (v/v) mixture of 400 mM 1-ethyl-3- (3-dimethylaminopropyl) carbodiimide (EDC) (Sigma) and 100 mM N-hydroxysuccinimide (NHS) (Sigma) was prepared and immediately injected into two flow cells of a CM5 chip to activate the carboxymethylated dextran surface. Lyophilized anti-polyhistidine antibodies (R&D Systems, Inc.) were resuspended to 25 μg/mL in immobilization buffer (10 mM sodium acetate, pH 5.0) and injected into the 2 activated flow cells for amine coupling. Remaining uncoupled surfaces were inactivated by injection of blocking buffer (1 M ethanolamine, pH 8.5) (Sigma). Activation, coupling, and blocking steps were

performed for 7 min at a flow rate of 10 μL/min resulting in anti-polyhistidine direct coupling of ~14,000 response units (RUs) in both flow cells.

Biacore assays were performed at 25°C using freshly prepared and degassed running buffer (10 mM HEPES, 150 mM NaCl, 0.005% (v/v) Surfactant P-20, pH 7.4). Unlabeled $PcrG^{FL}$ or $PcrG^{9-76}$ (still fused to C-terminal GB1-His_6 tag) was dialyzed overnight in running buffer and immobilized to the anti-polyhistidine coated surface of flow cell 2 by injection of 6-9 μL of 1 μg/mL protein at 10 μL/min. Varying concentrations (2-32 nM) of the $PcrV^{25-294}$ analyte (TEV cleaved and dialyzed overnight in running buffer) were then injected in both flow cells at 40 μL/min. Immobilization of the ligand and binding of analyte was repeated for each $PcrV^{25-294}$ concentration followed by surface regeneration using 20 μL pulses of regeneration buffer (10 mM glycine, pH 2.0). K_D values were calculated using BIAevaluation 4.1 software only after subtracting the flow cell 1 sensorgram (control) from flow cell 2 to account for non-specific binding of analyte to the surface.

As an additional control, the above steps were repeated using the TEV cleavage GB1-His_6 tag as the ligand immobilized to the anti-polyhistidine coated surface to verify that the fusion tag did not contribute to binding. Tests for mass transfer effect were also performed to validate the selected analyte flow rate (40 μL/min).

RESULTS

Defining a PcrG Domain for NMR Characterization

LcrG is the most closely related homolog of PcrG and shares with it a 41% sequence identity (Figure 3-1). Two unique LcrV binding domains on LcrG have been separately proposed by yeast two-hybrid studies[12] and NMR chemical shift mapping (unpublished, Sukanya Chaudhury, University of Kansas)[13]. Results of the yeast two-hybrid study suggested that the LcrG N-terminus residues, from residues D7 to T40 (corresponding to D10 to A42 in PcrG), are required for LcrV interaction, whereas results from NMR suggested the same for residues G50 to R73 (A53 to S76 in PcrG) (Figure 3-1). Each of these suggested binding domains include regions that are highly conserved between LcrG and PcrG. While *in vivo* substitutions of PcrG for LcrG in *Yersinia pestis* have been unable to transcomplement normal T3SS regulation, both proteins bind LcrV.[12] Thus, PcrG likely maintains the LcrG binding motif for its respective adaptor protein. Initial NMR spectra, aiming to identify the residues of PcrG interacting with PcrV, used ^{15}N-labeled PcrG^{FL} and resulted in very poor quality data (Figure 3-3a). Using several different secondary structure prediction servers, including ITASSER, DSC, MLRC, and PHD, it became clear that both termini of PcrG consist of random coils (Figure 3-2). In order to include the structured domains of PcrG while maintaining hydrophilic terminal residues for solubility, a PcrG truncation was subcloned with residues 9 -76 fused to a C-terminal His_6-tagged GB1 domain. This truncation also retains the regions of PcrG homologous to previously proposed adaptor protein binding domains in LcrG.

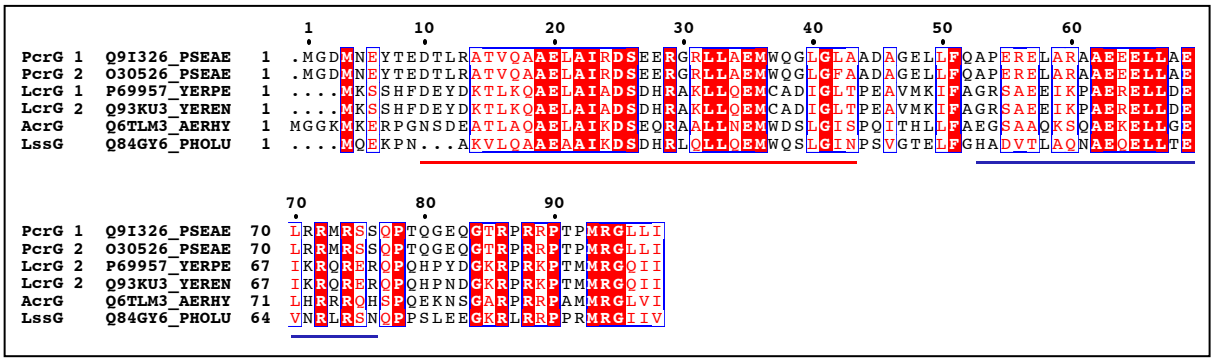


Figure 3-1. Sequence alignment of a protein BLAST with PcrG (Swiss-Prot ID: Q9I326_PSEAE) as the query sequence. Blast results were aligned with the query using ClustalW (Thompson, J. D., Higgins, D. G., and Gibson, T. J. *Nucleic Acids Res. 22*, 4673, 1994) and the figure was made using ESPript (Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. *Bioinformatics 15*, 305, 1999). The red and blue lines denote regions homologous to reported LcrV binding domains on LcrG using yeast two-hybrid assays[12] and NMR chemical shift mapping[13], respectively.

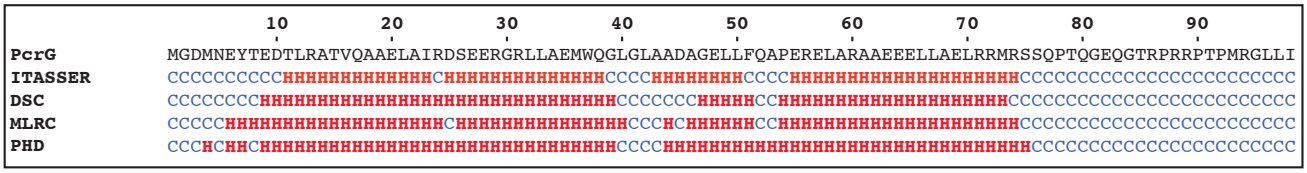


Figure 3-2. Secondary structure predictions for PcrG (Swiss-Prot ID: Q9I326) based on calculations from ITASSER (http://zhanglab.ccmb.med.umich.edu/I-TASSER) as well as DSC, MLRC, and PHD (http://npsa-pbil.ibcp.fr/cgibin/npsa_automat.pl?page=/NP-SA/npsa_seccons.html). Residues predicted to exist as part of a random coil or α-helix are denoted with letter "C" or "H", respectively.

Protein Expression and Purification

Full length PcrG and truncated PcrG$^{9-76}$ were both overexpressed in *E. coli* and purified by Ni^{2+}-affinity chromatography. Purified fractions were then cleaved by a His_6-tagged TEV protease and products were loaded onto a Ni^{2+}-affinity column to purify the cleaved PcrG. Purified PcrG^{FL} and PcrG$^{9-76}$ were then concentrated to 0.8 mM and 0.7 mM, respectively, and 2D ^{1}H-^{15}N HSQC spectra were collected (Figure 3-3). While the full length protein provided a poor spectrum with broad and overlapped peaks, the PcrG$^{9-76}$ truncation afforded a spectrum suitable for NMR characterization based on the presence of sharp and well resolved peaks. A comparison between the circular dichroism (CD) spectra of the two proteins verifies that the PcrG$^{9-76}$ truncation did not produce any appreciable changes to the secondary structure (Figure 3-4a). In addition, thermal denaturation curves of PcrG^{FL} and PcrG$^{9-76}$ have no clear inflection points, indicating that neither construct contains a folded core.

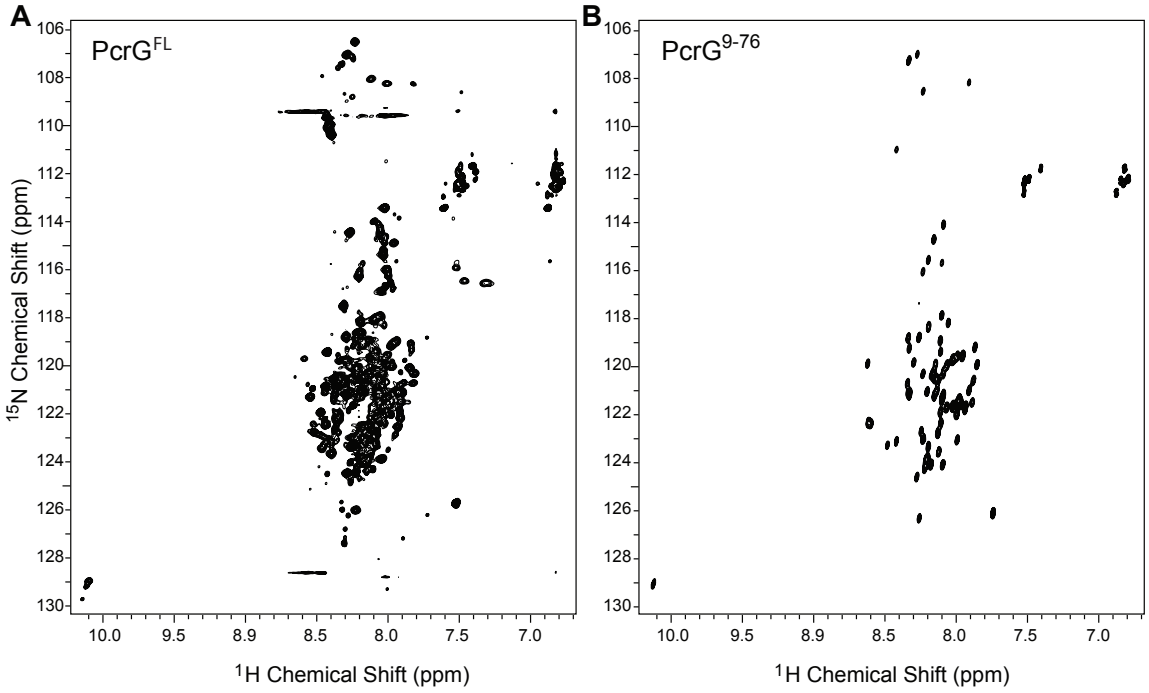


Figure 3-3. 2D ^{1}H-^{15}N HSQC spectrum of (A) PcrG^{FL} and (B) truncated PcrG$^{9-76}$.

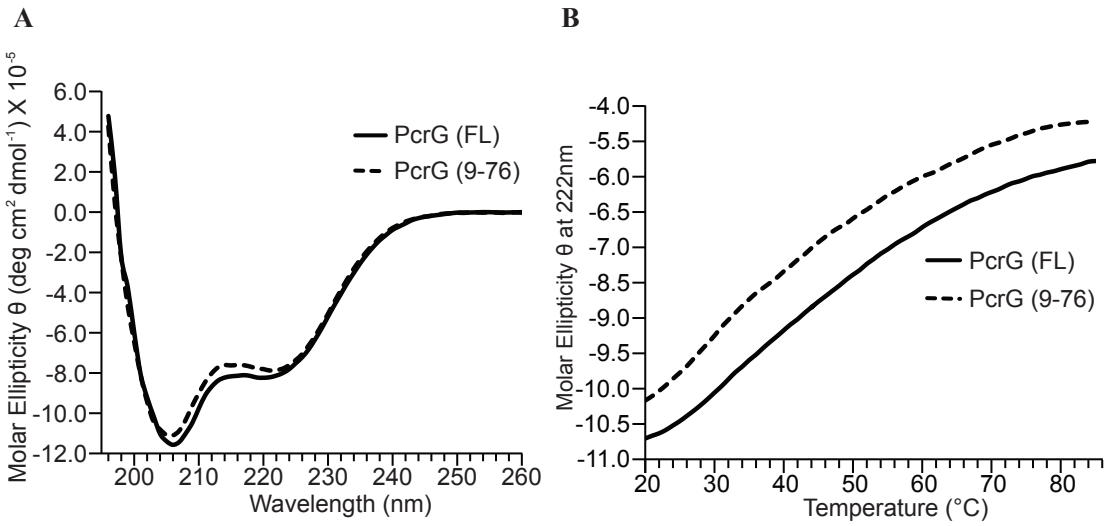


Figure 3-4. (A) CD spectra acquired at 25°C of PcrG^{FL} and PcrG$^{9-76}$ showing that the proteins are composed partially of α-helices and partially of random coils. (B) CD thermal denaturation plots of PcrG^{FL} and PcrG$^{9-76}$ lack inflection points, suggesting the proteins do not have a core structured domain.

NMR Assignment

The 2D ^{1}H-^{15}N HSQC spectrum of PcrG$^{9-76}$ was assigned nearly to completion using 3D HNCACB (Figure 3-5a) and 3D HNCA (Figure 3-5b) experiments. Examples of the carbon backbone connectivity are illustrated in Figure 3-5. The backbone amides of 54 residues, or approximately 80.1% of the 67 non-proline residues of PcrG$^{9-76}$ were successfully assigned as well as all 8 residues of the C-terminal cloning artifact GSENLYFQ (Figure 3-6). The side chain peak of the single tryptophan, W37, was also assigned. The remaining 13 non-proline residues that could not be assigned due to peak overlaps or peak broadening were Glu (9), Glu (28), Arg (29), Leu (32), Leu (33), Met (36), Leu (49), Leu (50), Phe (51), Arg (56), Glu (57), Glu (65), and Glu (66). Secondary chemical shift data was calculated for the assigned residues of PcrG$^{9-76}$ (Figure 3-7). The data indicates that the protein contains three major α–helical regions with short interspersed residues that lack secondary structure. This observation correlates well with

the CD data collected for PcrG$^{9-76}$ (Figure 3-4). The CD spectrum contains two minima with a 222/208-nm amplitude ratio of approximately 0.7 (indicative of a partially unfolded and partially α–helical protein).

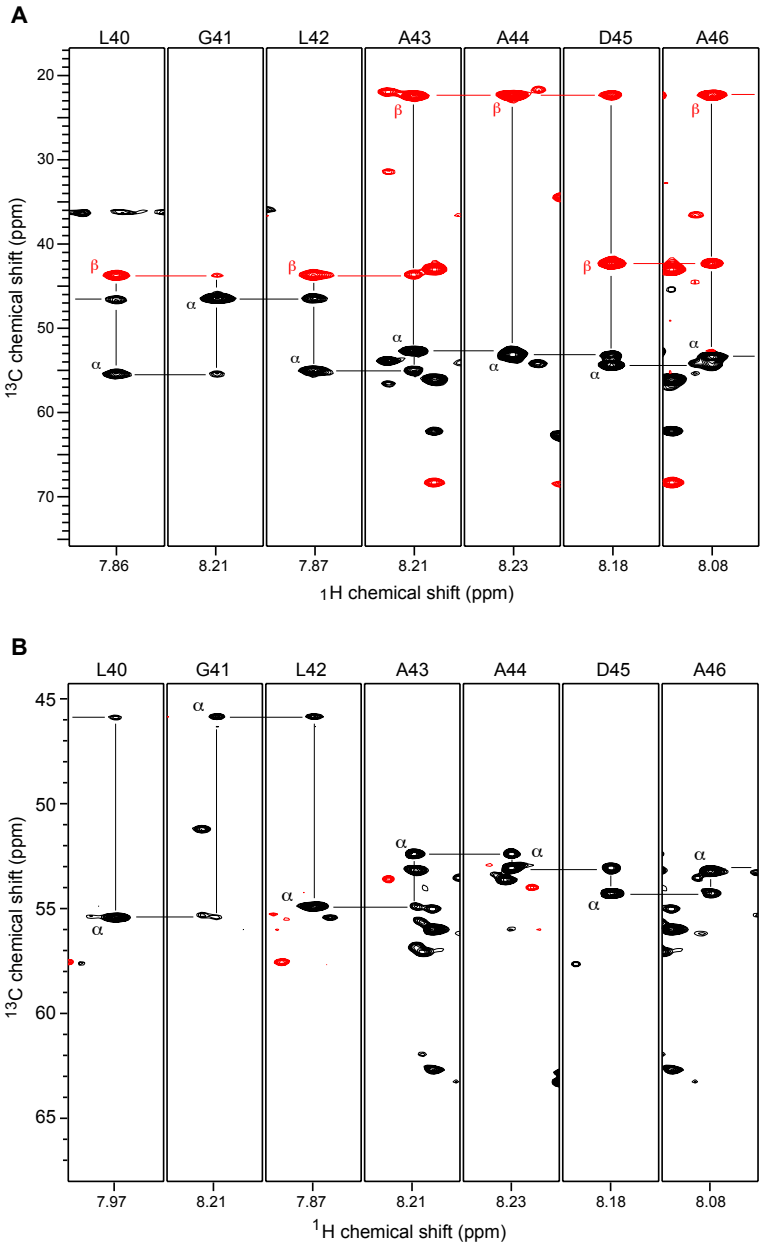


Figure 3-5. Representative portions of the 3D (A) HNCACB and (B) HNCA spectra of ^{15}N, ^{13}C-PcrG$^{9-76}$ spanning residues L40 to A46 to illustrate the backbone connectivity identified for PcrG$^{9-76}$ assignment. The peaks of lower intensity in each strip belong to residue (i-1).

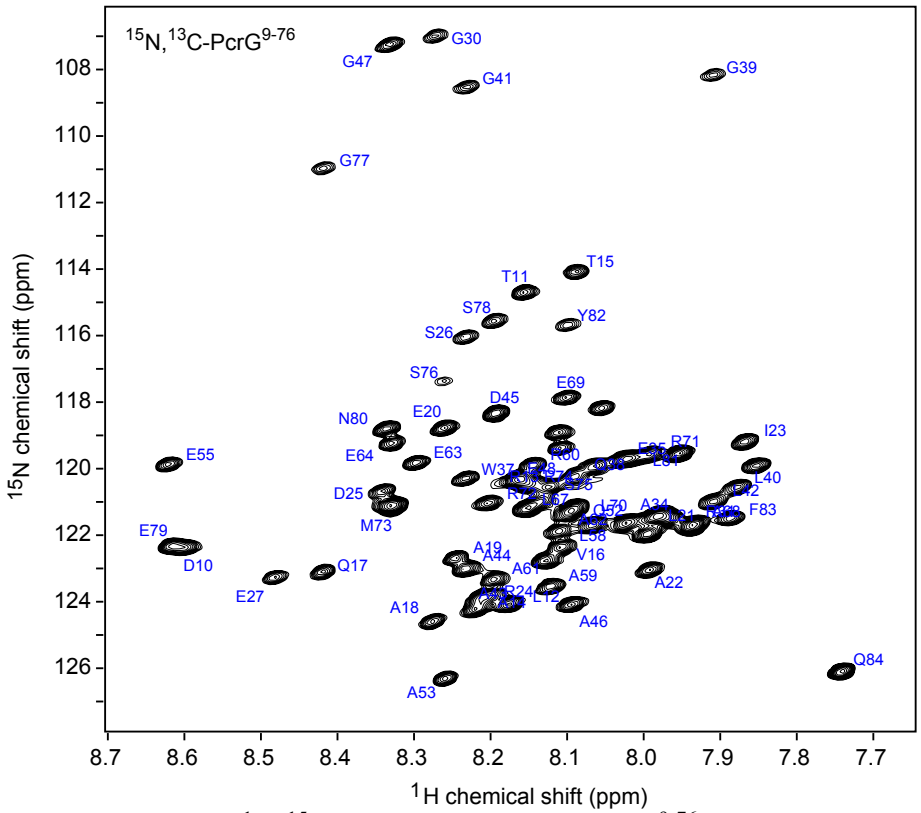


Figure 3-6. Assigned 2D ^{1}H-^{15}N HSQC spectrum of PcrG$^{9-76}$. Residue names are indicated in blue next to their corresponding peaks.

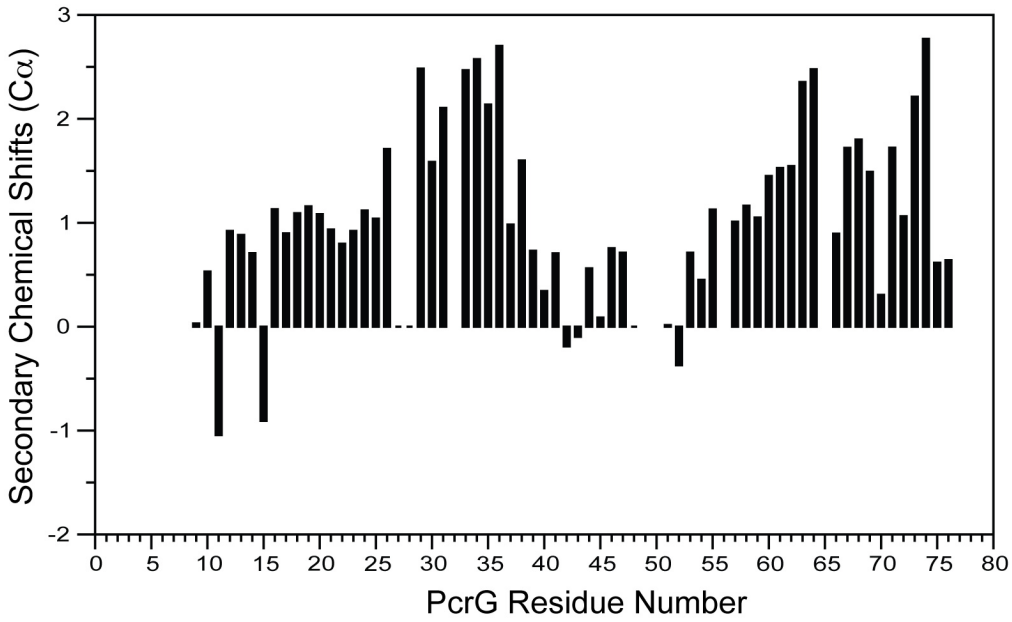


Figure 3-7. The C$^{\alpha}$ secondary chemical shifts of PcrG$^{9-76}$. Contiguous regions with positive secondary chemical shifts are typical of α–helical regions.

NMR Titrations

NMR chemical shift mapping was used to characterize the PcrG-PcrV interaction. Purified ^{15}N-labeled PcrG$^{9-76}$ was combined with increasing concentrations of PcrV$^{25-294}$ in NMR buffer and 2D ^{1}H-^{15}N HSQC spectra were acquired for each PcrV:PcrG molar concentration ratio (0.0, 0.25, 0.5, 1.0, 2.0 and 4.0). The 2D ^{1}H-^{15}N HSQC spectra for each titration ratio were superimposed to determine the changes in the 2D ^{1}H-^{15}N HSQC of PcrG as it binds to PcrV (Figure 3-8a). With the exception of only 8 of the 75 backbone amide peaks, nearly all peaks showed significant reduction in intensity as the molar ratio of titration increased. The 8 peaks that maintained constant intensity throughout the titration corresponded to cloning artifact residues G77, S78, E79, N80, Q84, and 3 unassigned residues.

A second NMR titration was performed with purified ^{15}N PcrG$^{9-76}$-GB1-His$_6$ (His$_6$-tagged GB1 domain was left intact with the initial purpose of aiding solubility) combined with increasing concentration of PcrV$^{25-294}$. 2D ^{1}H-^{15}N HSQC spectra were acquired at each PcrV:PcrG molar concentration ratio (0.0, 0.5, 1.0, and 2.0) and then superimposed (Figure 3-8b). A PcrV:PcrG molar ratio of 2.0 was sufficient for binding saturation. While the PcrG$^{9-76}$-GB1-His$_6$ spectrum has not been assigned, it very closely resembles that of its TEV cleaved counterpart (Figure 3-8a) along with 65 additional peaks from the appended GB1-His$_6$ tag. As the concentration of PcrV$^{25-294}$ increases, it is clear that the new GB1-His$_6$ residues remain inert and that the same subset of PcrG$^{9-76}$ peaks reduce in intensity as seen in the first titration.

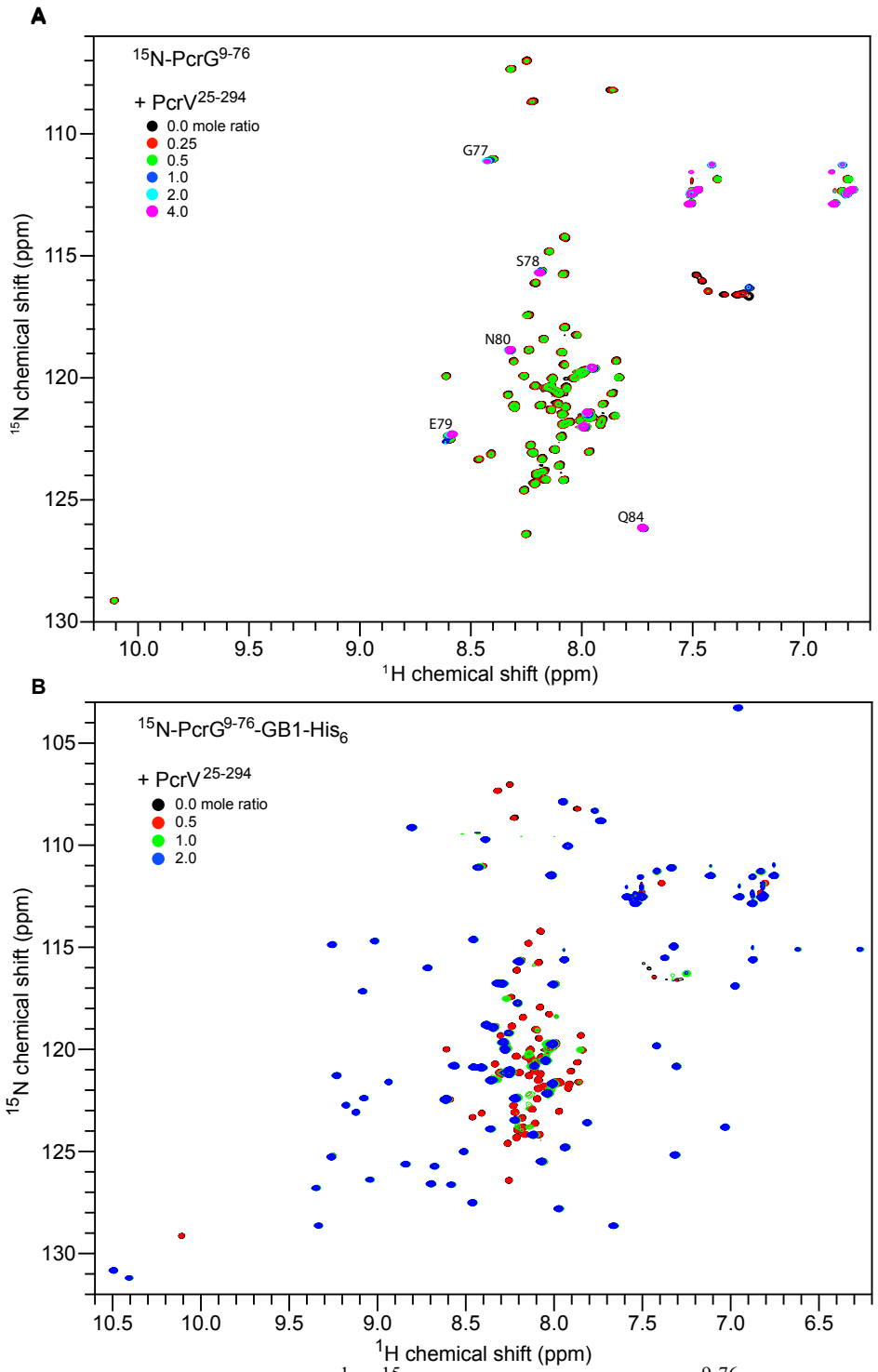


Figure 3-8. (A) Six overlaid 2D ^{1}H-^{15}N HSQC spectra of PcrG$^{9-76}$ with increasing ratios of PcrV$^{25-294}$. Assigned backbone amide peaks that remained unperturbed by the titration are labeled. (B) Four overlaid 2D ^{1}H-^{15}N HSQC spectra PcrG$^{9-76}$-GB1-His$_6$ with increasing ratios of PcrV$^{25-294}$.

Surface Plasmon Resonance Spectroscopy

Surface plasmon resonance (SPR) was used to evaluate the binding affinity and binding kinetics between $PcrV^{25-294}$ and PcrG (full length and 9-76). PcrG constructs were fused to C-terminal GB1-His_6 tags to facilitate immobilization to anti-His_6 antibody coated CM5 flow cells. Injecting of $PcrV^{25-294}$ into flow cells with immobilized PcrG produced a clear binding response that increased proportionally with increasing analyte concentration. Using BIAevaluation 4.1 software, data for the interaction between $PcrV^{25-294}$ and full length PcrG was best fit to a 1:1 Langmuir model ($\chi^2 = 0.314$) (Figure 3-9a). The K_D for this interaction was calculated to be 2.65×10^{-8} M (26.5 nM). This high affinity binding is very similar to that previously described by Nanao et. al[9] in which SPR was used to evaluate binding between full length proteins, PcrG and PcrV. Thus, the N-terminal 24-residue truncation of PcrV has little effect on the interaction. The binding also occurred with rapid kinetics, as previously described[9], with $k_a = 5.62 \times 10^5$ $M^{-1}s^{-1}$ and $k_d = 1.53 \times 10^{-2}$ s^{-1}.

The truncations made in $PcrG^{9-76}$ did not appear to have a significant effect in its PcrV-binding capacity. In fact, the sensorgram for $PcrV^{25-294}$ binding to immobilized $PcrG^{9-76}$-GB1-His_6 was fit to a 1:1 Langmuir model ($\chi^2 = 2.53$) and the calculated binding affinity was slightly stronger than that for $PcrG^{FL}$ ($K_D = 2.37 \times 10^{-8}$ M or 23.7 nM) (Figure 3-9b). Also, the binding kinetics were found to be more rapid than seen for $PcrG^{FL}$, though less than an order of magnitude, with $k_a = 1.02 \times 10^6$ $M^{-1}s^{-1}$ and $k_d = 2.42 \times 10^{-2}$ s^{-1}. These minimal changes in binding suggest that truncated $PcrG^{9-76}$ retains the PcrV-interaction domain.

To account for possible interference of the GB1-His_6 tag with the PcrG-PcrV interaction, a control sensorgram was collected with identical conditions from the previous experiment, but with GB1-His_6 as the ligand (Figure 3-10a). The binding response to $PcrV^{25\text{-}294}$ was insignificant and did not increase with respect to $PcrV^{25\text{-}294}$ concentration. Therefore, the appended GB1-His_6 tag remained inert in the interaction. As an additional control, a test for mass transfer effect was performed by varying the $PcrV^{25\text{-}294}$ analyte flow rate from 5-80 μl/min over immobilized $PcrG^{FL}$-GB1-His_6 (Figure 3-10b). No resulting differences were observed in the association phase of these sensorgrams, which confirms the lack of a mass transfer effect and validates the selected analyte flow rate of 40 μl/min.

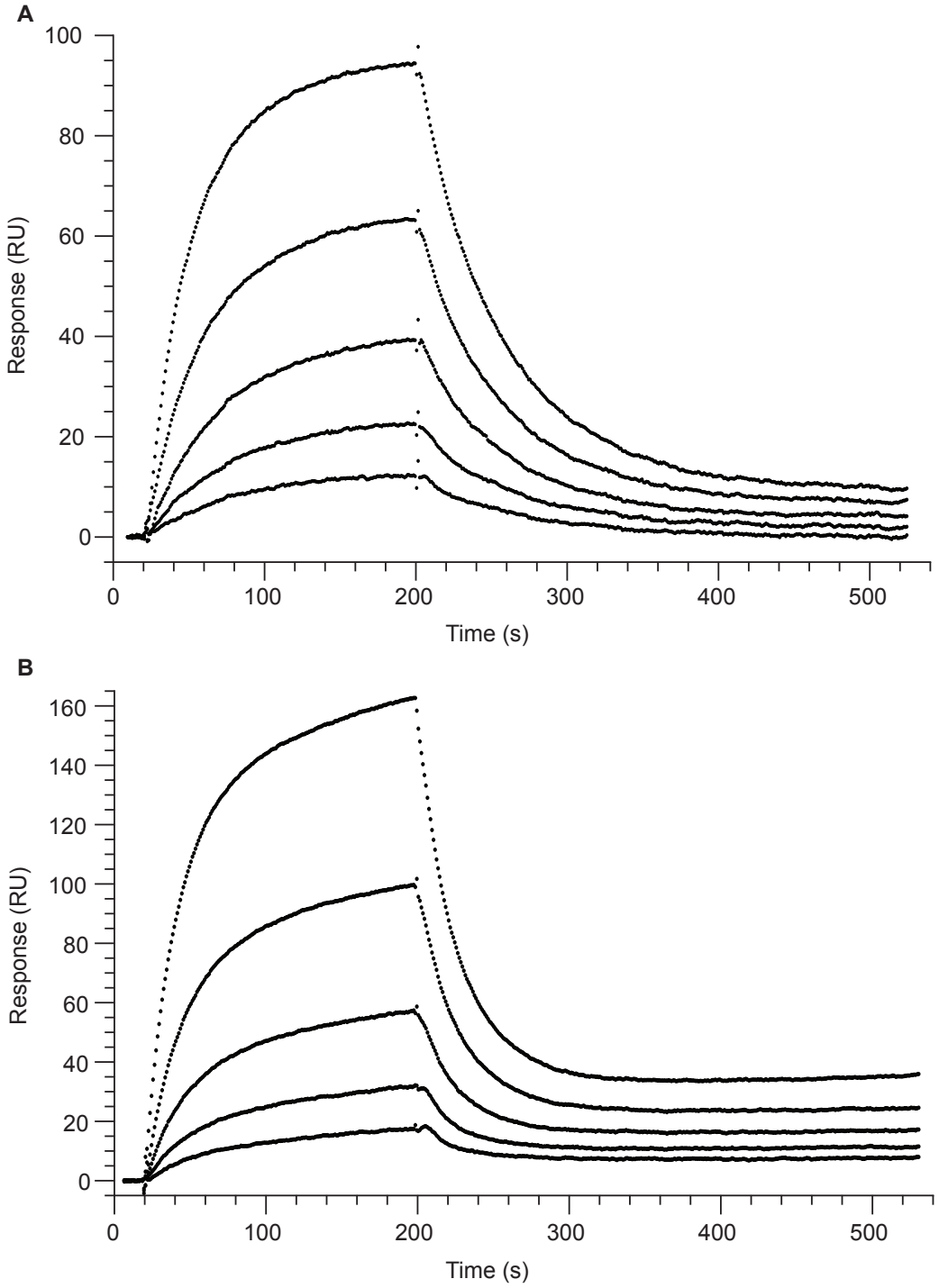


Figure 3-9. SPR sensorgrams for $PcrV^{25\text{-}294}$ binding to $PcrG^{FL}$ and $PcrG^{9\text{-}76}$. A serial set of $PcrV^{25\text{-}294}$ concentrations (2-32 nM) were injected at 40 μl/min for 3 min over immobilized (A) $PcrG^{FL}$-GB1-His_6 (200 RU); (B) $PcrG^{9\text{-}76}$-GB1-His_6 (104 RU). Non-specific binding of the analyte to the anti-His_6 antibody coated surface and the bulk effect from the buffer were determined from flow cell 1 and subtracted from the raw data.

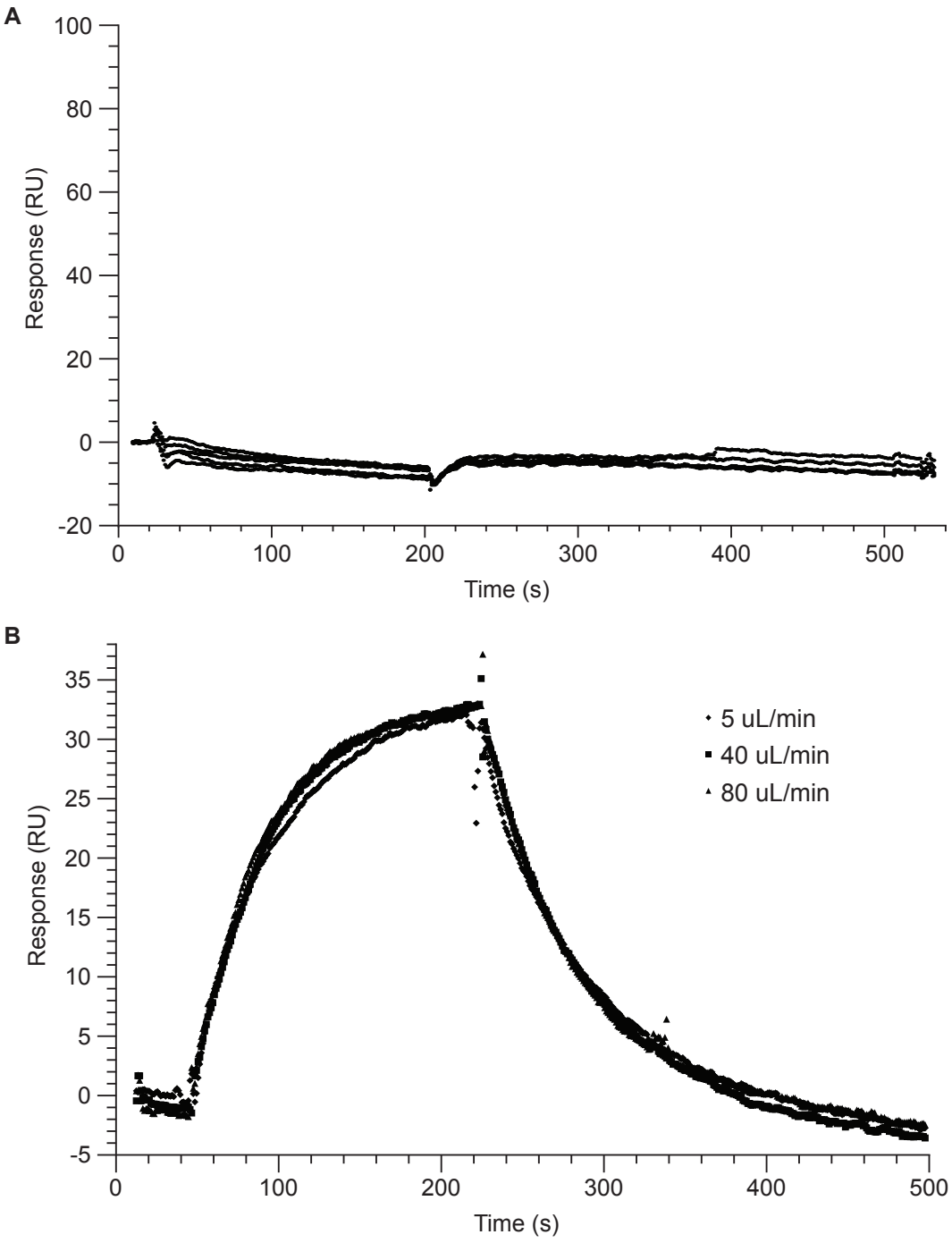


Figure 3-10. (A) SPR sensorgram control for $PcrV^{25\text{-}294}$ binding to a GB1-His_6 fusion tag. A serial set of $PcrV^{25\text{-}294}$ concentrations (2-32 nM) were injected at 40 μl/min for 3 min over immobilized GB1-His_6 (108 RU). (B) Injection of 8 nM $PcrV^{25\text{-}294}$ at varying flow rates over immobilized $PcrG^{FL}$-GB1-His6 to test for mass transfer effect.

DISCUSSION

Recent results have pointed to the essential role of PcrG in control of effector protein secretion in *Pseudomonas aeruginosa* as well as to its probable role as an export chaperone of the tip protein, PcrV.[7; 8] Direct binding between PcrG and PcrV has been demonstrated[9; 10], however, a clear PcrV binding region on PcrG has yet to be identified. The closest homolog of PcrG, LcrG of the *Yersinia* T3SS, has also been attributed with a role in secretion regulation and has been shown to interact with its respective tip protein, LcrV.[11] However, the mechanisms by which PcrG and LcrG mediate secretion regulation in their respective type III secretion systems differ substantially.[8; 12] The aim of this study was to identify whether the PcrV binding region on PcrG is conserved with the previously reported tip protein binding on its close homolog, LcrG. Instead, NMR experiments indicate that PcrV induces a global change in PcrG structure that likely results as a combination of PcrV-binding interactions and binding-induced conformational changes. In addition, the first experimentally derived information on sequence specific secondary structure identified the regions of PcrG that adopt either α–helical or random coil conformation.

Based on sequence conservation, predicted random coil domains, and previously proposed tip protein binding domains in LcrG, a suitable PcrG truncation was constructed for NMR interaction studies. Approximately 80.1% of the $PcrG^{9-76}$ 2D ^{1}H-^{15}N HSQC spectrum was successfully assigned by use of 3D NMR experiments. Assignment was followed by a titration of ^{15}N-$PcrG^{9-76}$ with $PcrV^{25-294}$, which resulted in peak intensity reduction of 67 of the 75 backbone amide peaks. Notably, only 3 of the 8 residues that remained inert in the interaction are part of the $PcrG^{9-76}$ sequence and the remaining 5

belong to the C-terminal cloning artifact. While a majority of the peaks disappeared upon binding saturation with $PcrV^{25-294}$, this was not accompanied by the reappearance of new peaks, which is common for interactions in the intermediate exchange NMR time scale. For this reason, no structural information can be extracted with regard to the bound-state of PcrG. However, the observation that ~89.3% of the $PcrG^{9-76}$ peaks are reduced suggests that PcrV binding induces radical changes in PcrG structure. This is not unprecedented for natively unstructured proteins, which in most observed cases undergo binding induced folding transition.[23] The notion of a global structural transition is also supported by previous computational predictions that suggest LcrV interacts with LcrG via an LcrG coiled-coil domain.[11; 12]

In addition to identifying LcrV-interacting residues, NMR assignment of PcrG also elucidated the protein's sequence specific secondary structure. Secondary chemical shift data indicates roughly three stretches of α-helical regions (residues 10-26, 29-41, and 53-76) separated by short unstructured regions. While this data and the NMR titration data were limited by the use of truncated PcrG, the purported PcrV binding domain[12; 13] was retained. The validity of this truncation study is also supported by CD and SPR data. Circular dichroism showed that the $PcrG^{FL}$ and $PcrG^{9-76}$ spectra overlaid very closely and therefore maintained similar ratios of secondary structure composition. Results from the SPR binding interaction verified that the PcrG truncation sustained relatively unchanged binding affinity for $PcrV^{25-294}$ as well as similar binding kinetics. In fact, the binding affinity was slightly higher than that of $PcrG^{FL}$, which might be attributed to the removal of non-participating residues that may otherwise reduce binding accessibility.

Despite the experimental advantages provided by truncating the 8 N-terminal residues and C-terminal 22 residues of PcrG, this prevented characterization of structure or binding within these regions. The study was also limited in that chemical shift mapping cannot decipher between PcrG residues directly binding PcrV and those involved in binding induced conformational change. Future studies might consider alternative methods to identify interacting residues of the full length protein, such as a yeast two-hybrid study (successfully applied in the *Yersinia* system[12]).

REFERENCES

1. Lyczak, J. B., Cannon, C. L., and Pier, G. B. (2000) Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist, *Microbes Infect.* 2, 1051-1060.
2. Hauser, A. R. (2009) The type III secretion system of *Pseudomonas aeruginosa*: infection by injection, *Nat. Rev. Microbiol.* 7, 654-665.
3. Yahr, T. L., Vallis, A. J., Hancock, M. K., Barbieri, J. T., and Frank D. W. (1998) ExoY, an adenylate cyclase secreted by the *Pseudomonas aeruginosa* type III system, *Proc. Natl. Acad. Sci* 95, 13899-13904.
4. Dacheux, D., Epaulard, O., de Groot, A., Guery, B., Leberre, R., Attree, I., Polack, B., and Toussaint, B. (2002) Activation of the *Pseudomonas aeruginosa* type III secretion system requires an intact pyruvate dehydrogenase aceAB operon, *Infect. Immun.* 70, 3973-3977.
5. Dacheux, D., Goure, J., Chabert, J., Usson, Y., and Attree, I. (2001) Pore-forming activity of type III system-secreted proteins leads to oncosis of Pseudomonas aeruginosa-infected macrophages, *Mol. Microbiol.* 40, 76-85.
6. Goure, J., Pastor, A., Faudry, E., Chabert, J., Sessen, A., and Attree, I. (2004) The V antigen of *Pseudomonas aeruginosa* is required for assembly of the functional PopB/PopD translocation pore in host cell membranes, *Infect. Immun.* 72, 4741-4750.
7. Sundin, C., Thelaus, J., Bröms, J. C., Forsberg, Å. (2004) Polarisation of type III translocation by *Pseudomonas aeruginosa* requires PcrG, PcrV, and PopN, *Microbial Pathogenesis* 37, 313-322.
8. Lee, P., Stopford, C. M., Svenson, A. G., and Rietsch, A. (2010) Control of effector export by the *Pseudomonas aeruginosa* type III secretion proteins PcrG and PcrV, *Mol. Microbiol.* 75, 924-941.
9. Nanao, M., Ricard-Blum, S., Di Guilmi, A. M., Lemaire, D., Lascoux, D., Chabert, J., Attree, I., and Dessen, A. (2003) Type III secretion proteins PcrV and

PcrG from *Pseudomonas aeruginosa* form a 1:1 complex through high affinity interactions, *BMC Microbiol.* 3:21.
10. Allmond, L. R., Karaca, T. J., Nguyen, V. N., Wiener-Kronish, J. P., and Sawa, T. (2003) Protein binding between PcrG-PcrV and PcrH-PopB/PopD encoded by the *pcrGVH-popBD* operon of the *Pseudomonas aeruginosa* type III secretion system, *Infect. Immun.* 71, 2230-2233.
11. Lawton, D. G., Longstaff, C., Wallace, B. A., Hill, J., Leary, S. E. C., Titball, R. W., and Brown K. A. (2002) Interactions of the type III secretion pathway proteins LcrV and LcrG from Yersinia pestis are mediated by coiled-coil domains, *J. Biol. Chem.* 277 , 38714-38722.
12. Matson, J., S., and Nilles, M., L. (2002) Interaction of the *Yersinia pestis* type III regulatory proteins LcrG and LcrV occurs at a hydrophobic interface, *BMC Microbiol.* 2:16.
13. Chaudhury, S. and De Guzman, R. N., University of Kansas (to be published).
14. Geisbrecht, B. V., Bouyain, S., and Pop, M. (2006) An optimized system for expression and purification of secreted bacterial proteins, *Protein Expr. Purif.* 46, 23-32.
15. Wang, Y., Nordhues, B. A., Zhong, D., and De Guzman, R. N. (2010) NMR characterization of the interaction of the *Salmonella* type III secretion system protein SipD and bile salts, *Biochemistry* 49, 4220-4226.
16. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* 6, 277-293.
17. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules, *Methods Mol. Biol.* 278, 313-352.
18. Grzesiek, S. and Bax, A. (1993) The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements, *J. Am. Chem. Soc.* 115, 12953-54.
19. Grzesiek, S, Dobeli, H., Gentz, R., Garotta, G., Labhardt, A. M., Bax, A. (1992) 1H, 13C, and 15N NMR backbone assignments and secondary structure of human interferon-gamma, *Biochemistry* 31, 8180-8190.
20. Wittekind, M., and Mueller, L. (1993) HNCACB, A high sensitivity 3D NMR experiment to correlate amide proton and nitrogen resonances with the α-carbon and β-carbon resonances in proteins, *J. Magn. Reson.* 101B, 201-205.
21. Wishart, D. S. and Nip, A. M. (1998) Protein chemical shift analysis: a practical guide. *Biochem Cell Biol* 76, 153-163.
22. Derewenda, U., Mateja, A., Devedjiev, Y., Routzahn, K., M., Evdokimov, A. G., Derewenda, Z. S., and Waugh, D. S. (2004) The structure of *Yersinia pestis* V-antigen, an essential virulence factor and mediator of immunity against plague, *Structure* 12, 301-306.
23. Fink, A. L (2005) Natively unfolded proteins, *Curr. Opin. Struct. Biol.* 15, 35-41.

Chapter 4: Discussion of Type III Secretion System Projects

The type III secretion system (T3SS) is a complex set of regulatory and structural protein machinery common to many Gram-negative bacteria for virulence.[1] Many of these bacterial species are human pathogens and cause a variety of infectious diseases.[2] These resulting diseases can be fatal and range from chronic infections of the lungs in cystic fibrosis patients from *Pseudomonas aeruginosa* infection to gastroenteritis from *Salmonella typhimurium*.[1; 3] Each of these species uses the T3SS to deliver bacterial effector proteins into the host cell cytosol in order to manipulate normal host cell functions.[1; 4] The purpose of these host cell alterations varies widely between bacterial species, including prevention of phagocytosis, evasion of host immune response, or even bacterial intracellularization.[4] This variation in consequences for the host can largely be attributed to the many unique effector proteins between species (more than 100 have been identified[4]), however, the proteins components of the type III secretion apparatus (T3SA) used to transfer these effectors are both structurally and functionally conserved.[5] As there are still gaps in the current knowledge of how some of these T3SS proteins interact to regulate T3SA assembly and effector secretion, both structural and functional studies of these proteins are essential. In the work presented in this thesis, NMR studies and biophysical methods were used to characterize the interactions of T3SS tip proteins of *S. typhimurium* and *P. aeruginosa* with previously identified binding partners implicated in secretion control.

The binding of bile salts to the *Shigella* tip protein, IpaD, was recently suggested to play a role in activating the *Shigella* T3SS by inducing the secretion of the translocon protein, IpaB.[6] This is significant in that bile salts are a major component of bile and are

present in the stomach and intestines where *Shigella* infects humans.[7] *Salmonella* are another enteric pathogen and the *Salmonella* tip protein, SipD, was also recently shown to bind bile salts.[8] However, *Salmonella* exhibit a response opposite to that of *Shigella,* whereby growth in the presence of bile reduces invasiveness.[9; 10] This difference in response to bile salts is poorly understood.

After finding that SipD$^{39\text{-}343}$ provides a high quality NMR spectrum, SipD$^{39\text{-}343}$ was titrated with three different prevalent bile salts. Each of the bile salts perturbed a nearly identical set of SipD residues that primarily cluster in a loop between the N-terminal two-helix bundle and the central coiled-coil. While several other residues were also perturbed, the major cluster of interacting residues lies on the opposite side of the central coiled-coil from where the structurally similar tip protein, IpaD, was predicted to bind bile salts by computer simulated docking.[8] Additional testing to verify the interacting residues of SipD was attempted using isothermal calorimetry (ITC) and a *Salmonella* invasion assay. However, ITC was limited by the low CMC of the bile salts and results of the invasion assay were inconclusive. In the invasion assay, only single point mutations of SipD were introduced and were potentially not severe enough to disrupt bile salt binding. Based on the NMR titration data, it is proposed that the alternative bile salt binding sites for the two tip proteins may account for the differing bile salt responses between the two species. The reason for the differing responses remains unclear, however it may correlate to different spatial preferences in the gut for invasion. The tip protein-bile salt interaction would serve as a logical system for spatial orientation within the gut since 1) the tip proteins of both species have been shown to localize to the bacterial surface before cell contact[11; 12], 2) the tip proteins are the only

identified direct binding partners of bile salts, and 3) both species are exposed to bile near their target tissue for invasion, and 4) reabsorption and recycling of bile salts in the intestines leads to a concentration gradient.[7]

The tip protein, PcrV, of *Pseudomonas aeruginosa* and its proposed cytoplasmic chaperone, PcrG, have been clearly demonstrated with roles in effector secretion control.[13; 14; 15] For this species, secretion is normally initiated by cell contact or depletion of Ca^{2+}.[16] However, deletions of either PcrG or PcrV lead to secretion in the absence of these events.[15] PcrG and PcrV directly interact[17; 18] and PcrG has been shown to facilitate PcrV secretion.[15]

For further investigation, NMR was used to characterize this interaction using a PcrG truncation (residues 9-76) that provided an excellent NMR spectrum. The truncation was demonstrated to have little effect on the PcrG secondary or tertiary structure using circular dichroism (CD) and thermal denaturation. Also, surface plasmon resonance (SPR) verified minimal changes to the PcrG-PcrV binding affinity and binding kinetics when PcrG is truncated. Approximately 80.1% of the $PcrG^{9-76}$ HSQC spectrum was successfully assigned and these assignments were subsequently used for secondary chemical shift calculations and analysis of PcrV titrations. While PcrG had been previously observed to have partial α–helical structure when free in solution[17], secondary chemical shift analysis provided the first experimentally derived sequence specific structural information and identified approximately 3 major α–helical regions with flanking random coils. NMR titrations were performed to elucidate whether the interaction with PcrV results in a PcrG binding-induced folding event, which is common for natively unfolded proteins[19], and to identify the interacting residues of PcrG.

Titration with PcrV resulted in peak reductions of nearly all PcrG residues. One interpretation of this change, though unlikely, is that nearly all of the $PcrG^{9-76}$ residues directly bind PcrV. Such a large surface area for binding would be supported by the high binding affinity identified by SPR. However, if PcrG retains the α–helical regions from its unbound structure (moreover, if it adopts tertiary structure upon binding), it is less plausible that all PcrG residues could constitute a single binding surface. Instead it is proposed that the global effect on PcrG residues from PcrV binding results as a combination of direct residue contacts with PcrV and binding-induced conformational changes in PcrG. Since PcrV binding leads to a global change in the PcrG HSQC spectra, the utility of the NMR titration is limited. An alternate approach might be a yeast two-hybrid study, which has been used to identify the tip protein binding region of the close homolog of PcrG in *Yersinia pestis*, LcrG. Identification of these interacting residues will be critical in understanding the mechanism facilitating PcrV secretion.

In summary, the T3SS studies presented in this thesis have contributed to the further understanding of secretion control mediated by tip proteins and their binding partners. The study of the PcrG-PcrV interaction elucidated a step critical to *Pseudomonas aeruginosa* tip protein secretion and the study of the SipD interaction with bile salts identified a binding site that may account for the differing responses of *Salmonella* and *Shigella* to bile. Characterization of these interactions has potential towards facilitating the intelligent design of novel antibacterial agents that might disrupt proper assembly and regulation of the type III secretion system.

References

1. Hueck, C. J. (1998) Type III protein secretion systems in bacterial pathogens of animals and plants, *Microbiol. & Mol. Biol Rev*. 62, 379-433.
2. Coburn, B., Sekirov, I. and Finlay, B. B. (2007) Type III secretion systems and disease. *Clin. Microbiol. Rev.* 20, 535-49.
3. Lyczak, J. B., Cannon, C. L., and Pier, G. B. (2000) Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist, *Microbes Infect.* 2, 1051-1060.
4. Cornelis, G. R. (2006) The type III secretion injectisome, *Nat. Rev. Microbiol. 4*, 811-825.
5. Blocker, A., Komoriya, K. and Aizawa, S. (2003) Type III secretion systems and bacterial flagella: insights into their function from structural similarities. *Proc. Natl. Acad. Sci.* 100, 3027-30.
6. Olive, A. J., Kenjale, R., Espina, M., Moore, D. S., Picking, W. L., and Picking, W. D. (2007) Bile salts stimulate recruitment of IpaB to the *Shigella flexneri* surface, where it colocalizes with IpaD at the tip of the type III secretion needle, *Infect. Immun. 75*, 2626-2629.
7. Cabral, D. J., and D. M. Small (1989) Physical chemistry of bile, p. 621–662. In S. G. Schultz, J. G. Forte, and B. R. Bauner (ed.), Handbook of physiology, vol. 3. American Physiology Society, Bethesda, MD.
8. Stensrud, K. F., Adam, P. R., La Mar, C. D., Olive, A. J., Lushington, G. H., Sudharsan, R., Shelton, N. L., Givens, R. S., Picking, W. L., and Picking, W. D. (2008) Deoxycholate interacts with IpaD of *Shigella flexneri* in inducing the recruitment of IpaB to the type III secretion apparatus needle tip, *J. Biol. Chem. 283*, 18646-18654.
9. Prouty, A. M., and Gunn, J. S. (2000) *Salmonella enterica* serovar typhimurium invasion is repressed in the presence of bile, *Infect. Immun. 68*, 6763-6769.
10. Prouty, A. M., Brodsky, I. E., Manos, J., Belas, R., Falkow, S., and Gunn, J. S. (2004) Transcriptional regulation of Salmonella enterica serovar Typhimurium genes by bile, *FEMS Immunol. Med. Microbiol. 41*, 177-185.
11. Espina, M., Olive, A. J., Kenjale, R., Moore, D. S., Ausar, S. F., Kaminski, R. W., Oaks, E. V., Middaugh, C. R., Picking, W. D., and Picking, W. L. (2006) IpaD localizes to the tip of the type III secretion system needle of *Shigella flexneri*, *Infect. Immun. 74*, 4391-4400.
12. Lara-Tejero, M., and Galan, J. E. (2009) *Salmonella enterica* serovar typhimurium pathogenicity island 1-encoded type III secretion system translocases mediate intimate attachment to nonphagocytic cells, *Infect. Immun. 77*, 2635-2642.
13. Dacheux, D., Goure, J., Chabert, J., Usson, Y., and Attree, I. (2001) Pore-forming activity of type III system-secreted proteins leads to oncosis of Pseudomonas aeruginosa-infected macrophages, *Mol. Microbiol.* 40, 76-85.
14. Goure, J., Pastor, A., Faudry, E., Chabert, J., Sessen, A., and Attree, I. (2004) The V antigen of *Pseudomonas aeruginosa* is required for assembly of the functional PopB/PopD translocation pore in host cell membranes, *Infect. Immun.* 72, 4741-4750.

15. Lee, P., Stopford, C. M., Svenson, A. G., and Rietsch, A. (2010) Control of effector export by the *Pseudomonas aeruginosa* type III secretion proteins PcrG and PcrV, *Mol. Microbiol*. 75, 924-941.
16. Dacheux, D., Epaulard, O., de Groot, A., Guery, B., Leberre, R., Attree, I., Polack, B., and Toussaint, B. (2002) Activation of the *Pseudomonas aeruginosa* type III secretion system requires an intact pyruvate dehydrogenase aceAB operon, *Infect. Immun.* 70, 3973-3977.
17. Nanao, M., Ricard-Blum, S., Di Guilmi, A. M., Lemaire, D., Lascoux, D., Chabert, J., Attree, I., and Dessen, A. (2003) Type III secretion proteins PcrV and PcrG from *Pseudomonas aeruginosa* form a 1:1 complex through high affinity interactions, *BMC Microbiol.* 3:21.
18. Allmond, L. R., Karaca, T. J., Nguyen, V. N., Wiener-Kronish, J. P., and Sawa, T. (2003) Protein binding between PcrG-PcrV and PcrH-PopB/PopD encoded by the *pcrGVH-popBD* operon of the *Pseudomonas aeruginosa* type III secretion system, *Infect. Immun.* 71, 2230-2233.
19. Fink, A. L (2005) Natively unfolded proteins, *Curr. Opin. Struct. Biol.* 15, 35-41.

Addendum: NMR Characterization of the *Xenopus laevis* PARP1-PIASy Interaction

INTRODUCTION

Post-translational modification by SUMO (small ubiquitin-related modifier) has been studied for nearly 14 years, over which time many SUMOylation targets and functions have been identified.[1] SUMO is found in all eukaryotic cells and SUMOylation has been discovered to have roles in altering the stability, activity, localization, and interactions of its target proteins.[2] The process of SUMOylation entails three important steps carried out by three intermediate enzymes (E1 activating enzymes, E2 conjugating enzymes, and E3 ligases).[1; 2] While the same E1 and E2 enzymes, respectively Aos1/Uba2 and Ubc9 in *Xenopus laevis*, participate in all SUMOylation, many unique E3 enzymes have been identified and are recognized for mediating target specific SUMO conjugation.[3] One such enzyme, from the family of protein inhibitors of activated STAT (PIAS), is PIASy and has been shown to mediate critical mitotic chromosomal protein SUMOylation for efficient chromatin separation during anaphase.[4] More recently, Ryu et. al (5) identified a novel mitotic SUMOylation substrate, PARP1.

Poly(ADP-ribose) polymerases (PARP) catalyze post-translational modification by covalently attaching chains of poly(ADP-ribose) to substrates (known as PARylation). PARylation is closely tied to proper mitotic function. PARP1 localizes to condensed chromatin with particular concentration near the centromere[5] and its inhibition has been demonstrated to prevent mitotic spindle formation[6]. Ryu et. al (5) found PARP1 SUMOylation significantly reduces PARylation of mitotic chromatin. The same study discovered that PARP1 SUMOylation in *Xenopus* egg extracts required PIASy and that

this E3 ligase both enhanced SUMOylation efficiency and promoted modification site specificity.[5] Together, these findings suggest that PIASy-dependent SUMOylation of PARP1 may have a critical role in mitotic regulation.

In order for PIASy to mediate SUMOylation, a 4-protein complex including Ubc9-SUMO-2/3, PIASy, and a target protein is formed to facilitate isopeptide bond formation between SUMO and a particular lysine residue on the target. While domain analysis of one E3 ligase, Siz1, has recently elucidated this conjugation step in yeast[7], little structural information is available for the vertebrate system. In this study, NMR titrations were used in an attempt to identify the interacting residues of *Xenopus laevis* PARP1 and PIASy. While PARP1 potentially contains more than one PIASy binding domain, $PARP1^{375-650}$ was identified as a minimum domain for E3-dependent SUMOylation (unpublished, Gada Al-Ani, University of Kansas). NMR titration of a fragment within this domain, $PARP1^{503-650}$, indicated that approximately 9 PARP1 residues have contacts with PIASy. Since the titration was performed in the absence of Ubc9-SUMO-2/3 or other SUMOylation machinery, the results suggest that PARP1 interacts with PIASy prior to formation of the 4-protein complex of the SUMO conjugation step. Further investigation will be required to determine if other portions of the minimal domain for PIASy-dependent sumoylation also interact and to identify these residues.

MATERIALS AND METHODS

Protein Expression and Purification

PARP1 fragments (residues 375-650, 375-525, 503-650, and 514-639) were PCR amplified from *Xenopus laevis* cDNA and subsequently subcloned into pET-29a or into pET-21a (as an N-terminal or C-terminal fusion to a TEV recognition sequence and His_6-tagged GB1 domain as described above). All constructs were expressed in BL21(DE3) DNAY cells using LB media or ^{15}N M9 minimal media, as described above. Cells were harvested by centrifugation, resuspended in binding buffer (500 mM NaCl, 20 mM Tris-HCl, 5mM imidazole, pH 8.0), and lysed by sonication in the presence of 10 mg phenylmethyl sulfonyl fluoride (PMSF). The lysate was centrifuged again to remove cell debris, supernatant was loaded onto a Ni^{2+}-affinity column (Sigma), and elutes were pooled. PARP1 samples expressed from the pET-21a plasmid were used in the NMR studies with and without TEV proteolysis (which removed the His_6-tagged GB1 domain). Samples to be digested with protease followed the first Ni^{2+} affinity purification with addition of TEV protease and overnight dialysis as described[8]. After TEV proteolysis, protein samples were loaded onto a Ni^{2+} affinity column and the cleaved PARP1 was collected in the flow-through fractions.

Unlabeled $PIASy^{FL}$ protein samples for NMR titration of PARP1 were provided by Dr. Azuma (University of Kansas). All proteins samples for PARP1 and PIASy were concentrated using Amicon Ultra 3K (Millipore) and concentration was determined using absorbance at 280 nm.

NMR Spectroscopy

NMR data were collected at 25°C using Bruker Avance 800 MHz equipped with a cryoprobe, processed with NMRPipe[9], and analyzed using NMRView[10]. All protein samples were dialyzed overnight in NMR buffer (10-150 mM $NaHPO_4$, 10 mM NaCl, 10% (v/v) D_2O, pH 7.0) prior to data collection. ^{15}N-PARP1$^{375-525}$ (~40 μM) and ^{15}N-PARP1$^{503-650}$ (~30 μM) were used to collect initial 2D ^{1}H-^{15}N HSQC[11] spectra. After subcloning PARP1$^{503-650}$ into pET-21a, two additional 2D ^{1}H-^{15}N HSQC[11] spectra were acquired using TEV cleaved samples (~90 μM each). These samples differed by the cloning artifact that remained after TEV proteolysis (an N-terminal "GHM" or C-terminal "GSENLYFQ"). The sample with the N-terminal "GHM" also retained a 26 residue C-terminal tail as a cloning artifact. The final 2D ^{1}H-^{15}N HSQC[11] spectrum collected used ^{15}N-PARP1$^{514-639}$ (~80 μM) with the N-terminal cloning artifact "GHM".

For NMR chemical shift mapping, 2D ^{1}H-^{15}N HSQC[11] were collected using ^{15}N-PARP1$^{503-650}$-GSENLYFQ or ^{15}N-GHM-PARP1$^{503-650}$ (+26 C-terminal residues) titrated with increasing amounts of PIASy^{FL} as follows: 3 samples of 75 μM ^{15}N-PARP1$^{503-650}$-GSENLYFQ were prepared in NMR buffer (150 mM NaCl, 10 mM NaCl, 10% (v/v) D_20, pH 7.0) with increasing concentrations of PARP1^{FL} (0, 37.5, and 75 μM); 3 samples of 45 mM ^{15}N-GHM-PARP1$^{503-650}$ (+26 C-terminal residues) were prepared in NMR buffer with increasing concentrations of PARP1^{FL} (0, 45, and 116 μM).

RESULTS

PARP1 Expression and Solubility

The PARP1 fragments used herein were based on previous *in vitro* SUMOylation studies that identified the minimum domain required for PIASy-mediated modification by SUMO-2/3 (Gada Al-Ani, University of Kansas). This minimum domain, $PARP1^{375\text{-}650}$, was then further divided to more narrowly identify PIASy interaction surfaces by NMR (Figure A-1).

NMR experiments were initially performed using $PARP1^{375\text{-}525}$ and $PARP1^{503\text{-}650}$. While each of the PARP1 fragments illustrated in Figure A-1 were successfully expressed and purified, overnight dialysis in a standard low salt NMR buffer (10 mM NaCl, 10 mM $NaHPO_4$, pH 7.0) led to significant precipitation. $PARP1^{375\text{-}650}$ and $PARP1^{503\text{-}650}$ completely fell out of solution and $PARP1^{375\text{-}525}$ was only stable at ~40 μM. Increasing the ionic strength of the buffer to 50 mM permit initial 2D ^{1}H ^{15}N-HSQC data acquisition for the two smaller fragments (data not shown). The resulting HSQC spectrum of $PARP1^{375\text{-}525}$ exhibited highly overlapped peaks in a narrow proton range. On the other hand, $PARP^{503\text{-}650}$ provided a well-dispersed spectrum, but was of too poor quality for resonance assignments.

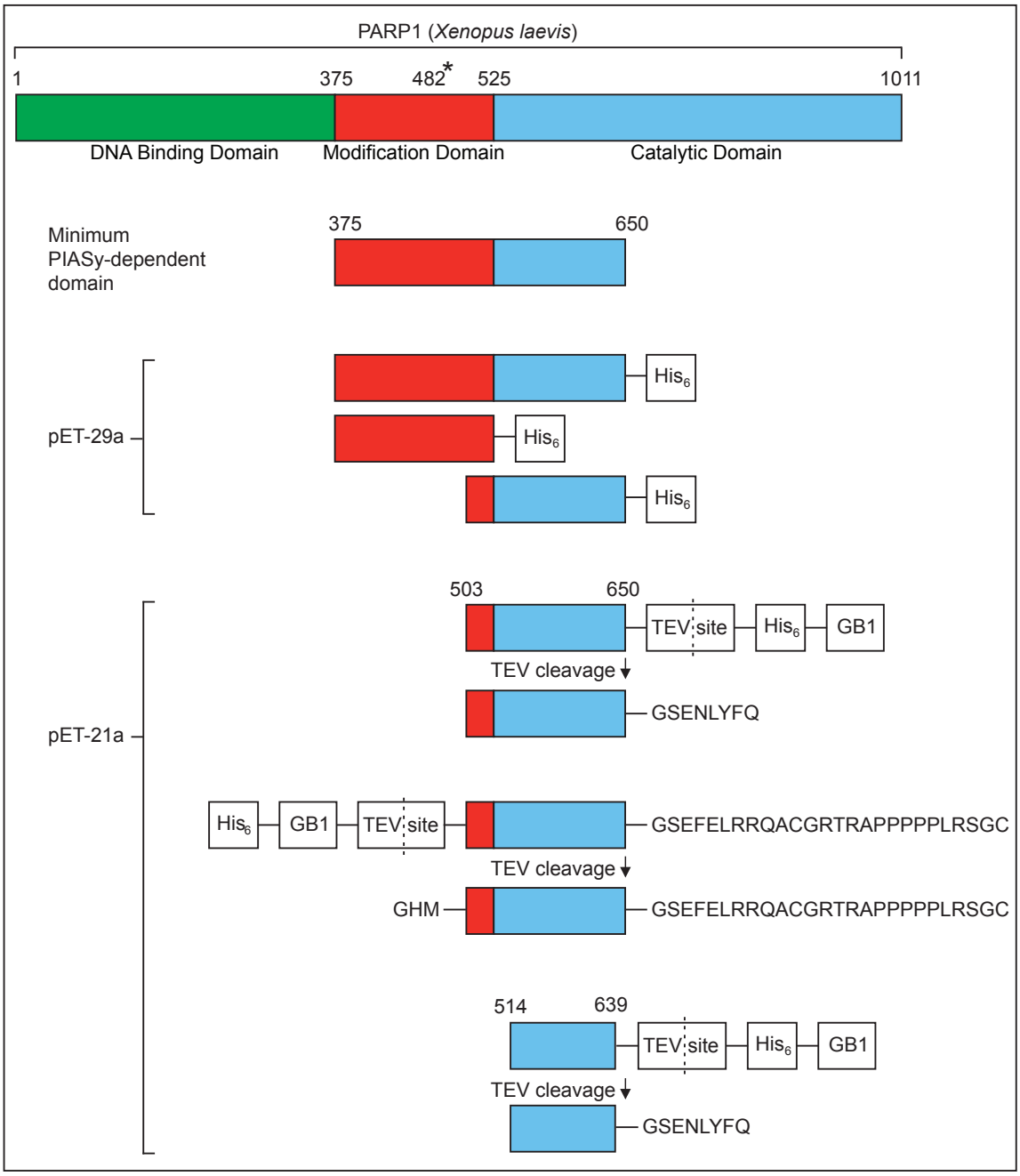


Figure A-1. PARP1 fragments derived from *Xenopus laevis* PARP1^{FL} for NMR analysis. The lysine that forms an isopeptide bond with SUMO2/3 is indicated with an asterisk.

To facilitate solubilization, PARP1$^{503-650}$ was subcloned into pET-21a vectors resulting in an N-terminal or C-terminal fusion to GB1, the highly soluble G1 domain of *Streptococcus* protein G. These constructs remained much more stable in solution, but the GB1 fusion also led to expression as a dimer. Subsequent removal of the GB1 domain by TEV proteolysis eliminated dimerization. While TEV cleavage negated the purpose of the new constructs, it provided the proteins with smaller cloning artifacts and maintained higher solubility than the original constructs. Several NMR buffer conditions were tested to maximize saturation concentration (Table A-1).

Table A-1. Solubility testing for PARP1$^{503-650}$-GSENLYFQ including variable pH and buffer choice. Saturation concentration refers to concentration measurements taken after samples ceased to precipitate. An asterisk indicates the chosen buffer.

Buffer	pH	Saturation Concentration (mM)
10 mM $NaHPO_4$, 50 mM NaCl, 5% (v/v) D2O	5.5	0.0845
10 mM $NaHPO_4$, 50 mM NaCl, 5% (v/v) D20	6.0	0.0909
*10 mM $NaHPO_4$, 50 mM NaCl, 5% (v/v) D20	6.5	0.0891
10 mM $NaHPO_4$, 50 mM NaCl, 5% (v/v) D20	7.0	0.0925
10 mM $NaHPO_4$, 50 mM NaCl, 5% (v/v) D20	7.5	0.0951
5 mM Tris/HCl, 50 mM NaCl, 5% (v/v) D20	7.5	0.0630
5 mM HEPES, 50 mM NaCl, 5% (v/v) D20	6.8	0.0610
5 mM HEPES, 50 mM NaCl, 5% (v/v) D20	7.5	0.0800

NMR Titrations

NMR chemical shift mapping was used to identify residues of PARP1 that interact with PIASy during PIASy-mediated PARP1 SUMOylation. ^{15}N-PARP1$^{503-650}$-GSENLFQ was prepared in the chosen NMR buffer (Table A-1) and provided an excellent initial spectrum (Figure A-2a). NMR titrations were then performed by preparing ^{15}N-PARP1$^{503-650}$-GSENLYFQ with increasing concentrations of PIASy^{FL} and 2D ^{1}H-^{15}N HSQC were acquired for each PIASy:PARP1 concentration ratio (0.0, 0.5, and 1.0) (Figure A-2b). Approximately 9 different PARP1$^{503-650}$ peaks show clear reduction in peak intensity as the titration ratio increased (indicated by arrows in Figure A-2b). This is characteristic of interactions in the slow-exchange NMR time scale. Despite the observation that 9 residues in this fragment interact with PIASy, residue assignments were not possible due to the high concentration requirement for 3D NMR experiments.

In spite of the 26-residue C-terminal extension of the GHM-PARP1$^{503-650}$ construct, NMR titrations of this construct with PIASy^{FL} still proved useful in supporting titration results shown in Figure A-2b. In addition to the appearance of new peaks, many chemical shift changes in the 2D ^{1}H-^{15}N HSQC indicate that the lengthened C-terminus made small changes in protein folding (Figure A-3a). Despite these changes, 6 of the peaks affected by PIASy correspond to those of the previous titration (Figure A-3b). These residues have not yet been assigned, but these consistent results support the specificity of the interaction.

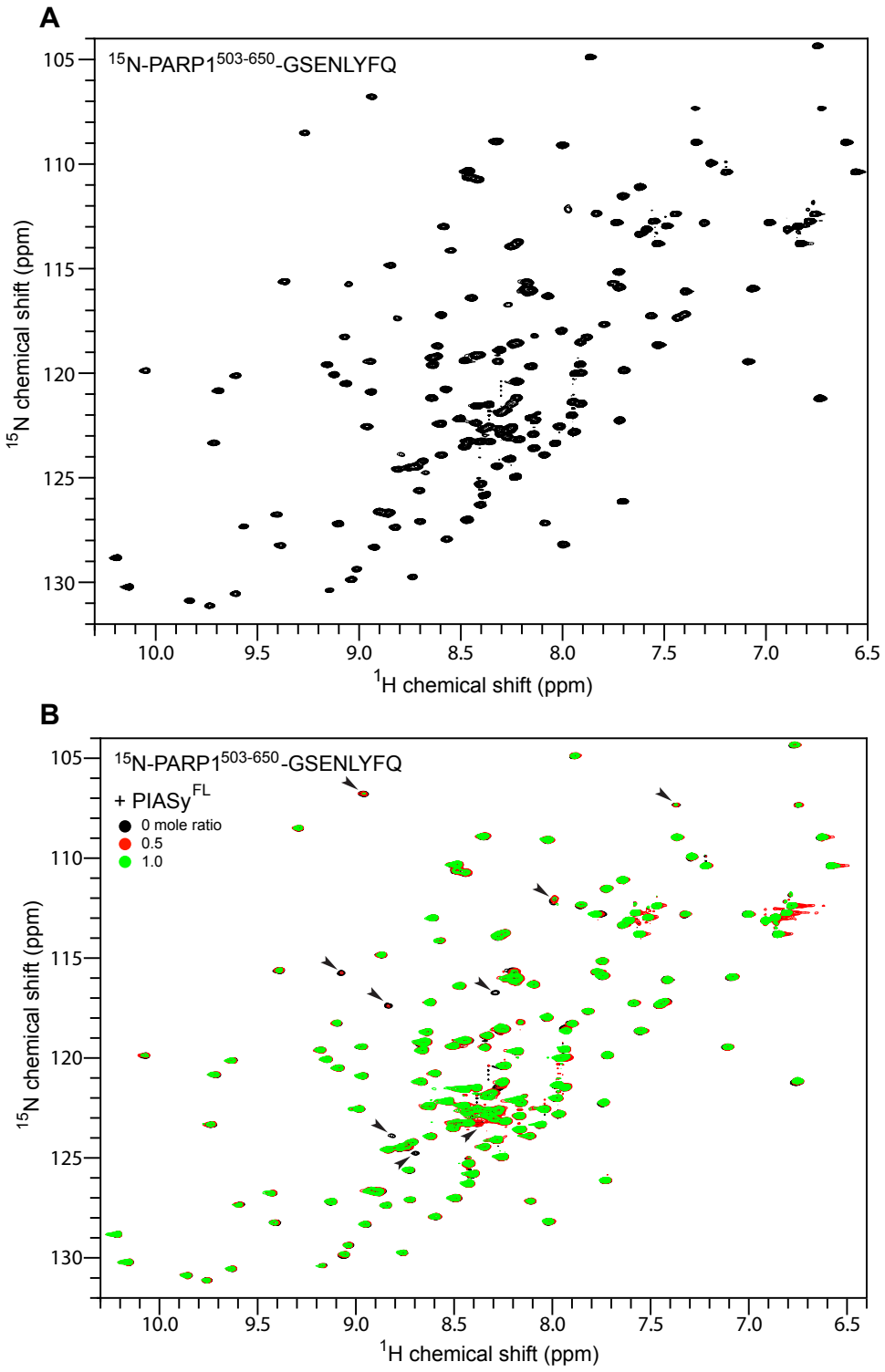


Figure A-2. (**A**) 2D ^{1}H-^{15}N HSQC spectrum of free ^{15}N-PARP1$^{503-650}$-GSENLYFQ. (**B**) Three overlaid 2D ^{1}H-^{15}N HSQC spectra of ^{15}N-PARP1$^{503-650}$-GSENLYFQ titrated with PIASy^{FL}. Spectra are color coded according to the protein molar ratio. Arrows denote peaks exhibiting intensity reduction.

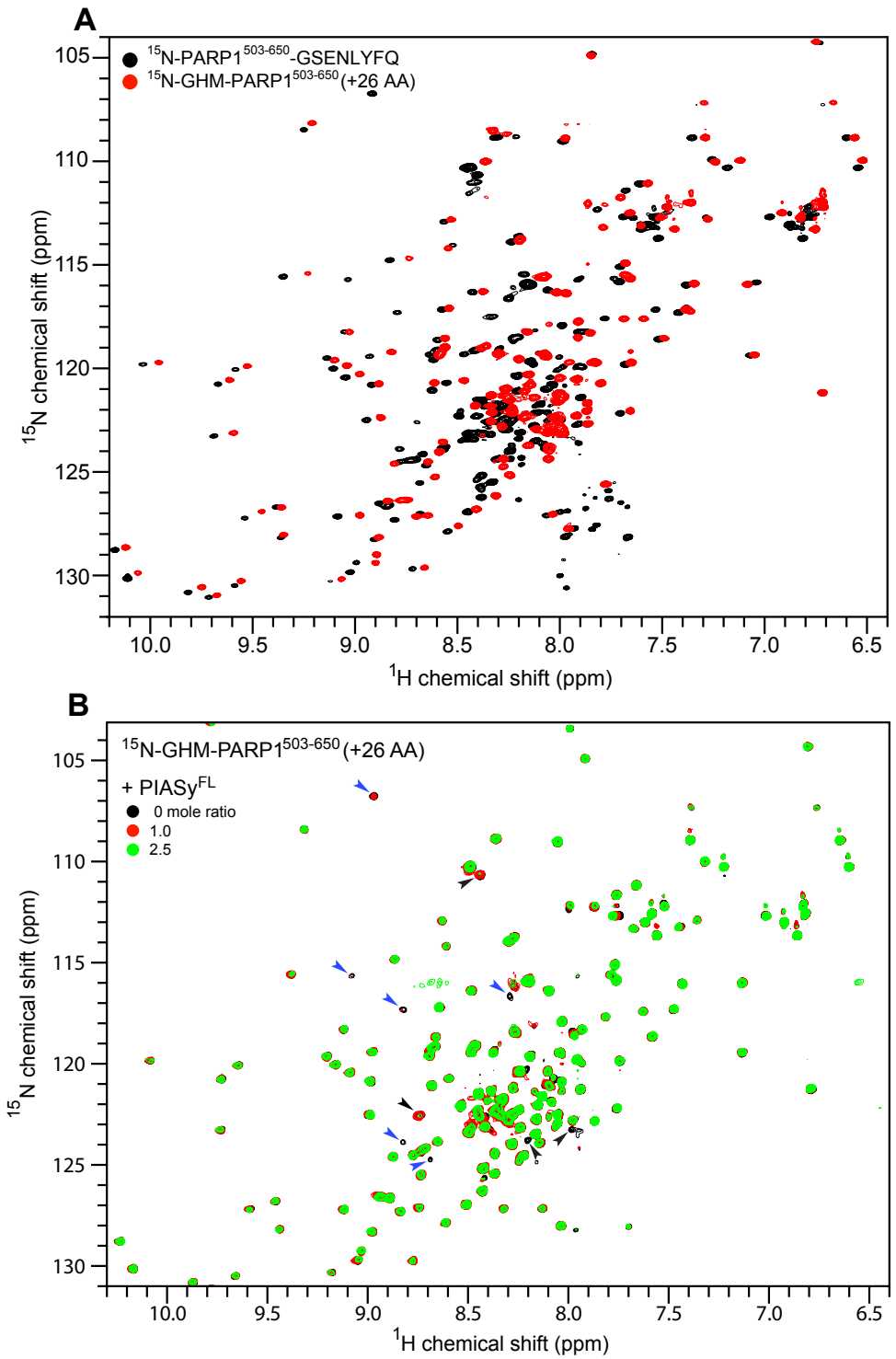


Figure A-3. (**A**) 2D ^{1}H-^{15}N HSQC spectra of free ^{15}N-GHM-PARP1$^{503-650}$ (with additional 26 residue C-terminal tail) in comparison with ^{15}N-PARP$^{503-650}$-GSENLYFQ. (**B**) Three overlaid 2D ^{1}H-^{15}N HSQC spectra of ^{15}N-GHM-PARP1$^{503-650}$ (with additional 26-residue C-terminal tail) titrated with PIASy^{FL}. Arrows denote peaks exhibiting intensity reduction and blue arrows denote peaks corresponding to those reduced in Figure A-2b.

An additional attempt was made to attain a PARP1 fragment stable enough in NMR solution for 3D experimentation. The structure of human PARP1 (residues 518-643) was previously solved by NMR techniques and has been deposited at the Protein Data Bank with accession code 2CR9.[12] This fragment closely corresponds to the *Xenopus laevis* PARP1$^{503\text{-}650}$ constructs tested in this work and a sequence alignment confirmed an 81% sequence identity with *Xenopus laevis* residues 514-639 (Figure A-4). This fragment was expressed with an N-terminal GB1 tag, cleaved with TEV protease, and dialyzed overnight in the NMR buffer chosen by Nagashima et. al[12] (20 mM Tris/HCl, 100 mM NaCl, 1mM DTT, 0.02% (v/v) NaN_3, 10% (v/v) D_20, pH 7.0). The sample initially maintained a high concentration (~0.8 mM) and 2D ^{1}H-^{15}N HSQC data were acquired (Figure A-5). The resulting spectrum was well dispersed with little overlap, however, the sample slowly precipitated to ~80 μM over the course of 5 days. 3D data were not collected.

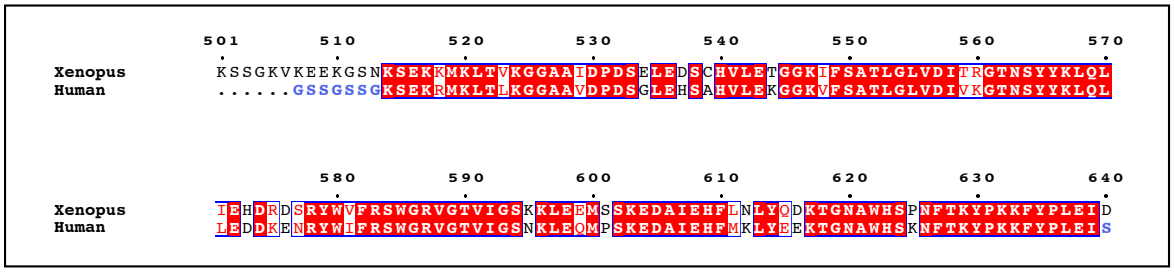


Figure A-4. Alignment of the *Xenopus laevis* (Swiss Prot ID: Q2NLA7) and human (Swiss Prot ID: P09874) PARP1 DNA sequences corresponding to the published structure of PARP1 fragment 2CR9. Blue letters indicate cloning artifact from 2CR9 (not part of wild-type human parp1).

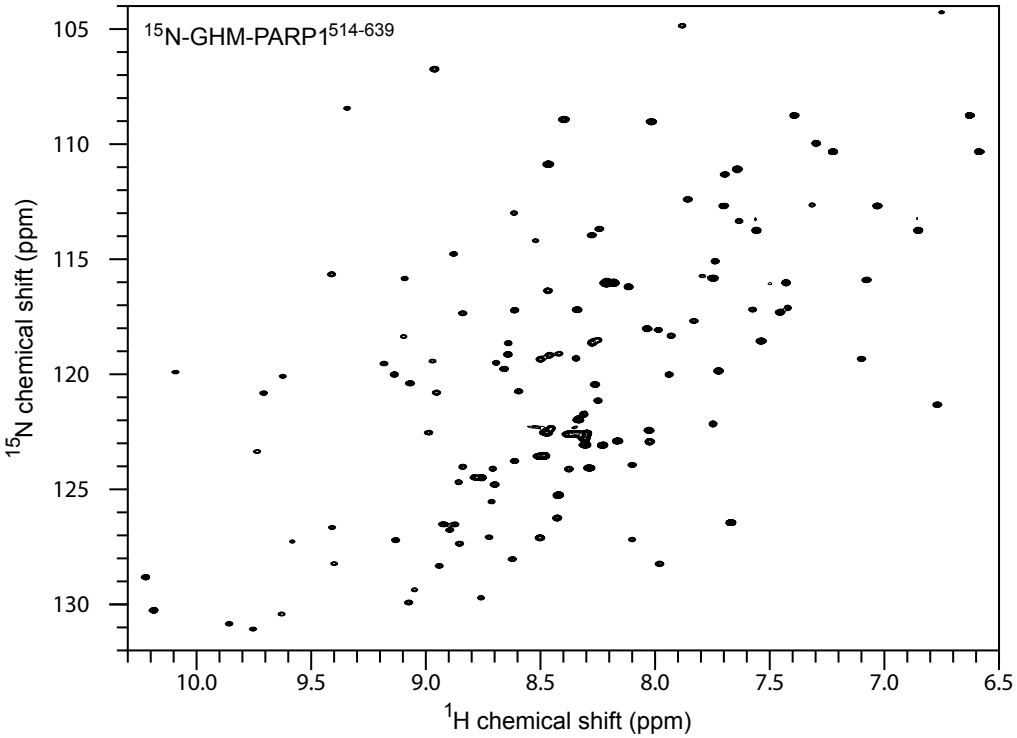


Figure A-5. 2D ^{1}H-^{15}N HSQC spectrum of ^{15}N-GHM-PARP1$^{514-639}$.

DISCUSSION

The major finding of the NMR experiments herein is that the E3 ligase, PIASy, interacts with PARP1 prior to formation of the 4-protein quaternary complex (SUMO, Ubc9, PIASy, and PARP1) involved in the SUMOylation conjugation step. The identified interaction domain is located within residues 503-650 of *X. laevis* PARP1, which may or may not contain the entire interaction surface. This conclusion is based on the PARP1-PIASy NMR titration where approximately 9 unidentified peaks within the 15N-PARP1$^{503\text{-}650}$-GSENLYFQ spectrum exhibited significant intensity reduction. The interaction was observed in the absence of the Ubc9-SUMO-2/3 complex or other SUMO conjugation machinery. Unfortunately, the low solubility of the PARP1 constructs used in the NMR titrations prevented the acquisition of the 3D NMR experiments needed for spectrum assignment.

Recent work identified a larger PARP1 domain (residues 375-650) required for PIASy-dependent SUMOylation (Gada Al-Ani, University of Kansas) that includes the predominantly modified residue, lysine 482.[5] Additional fragmentation of this domain (unpublished, Gada Al-Ani, University of Kansas) or conservative mutation[5] of lysine 482 has been shown to remove PIASy-dependent SUMOylation specificity for this site. Taking this into account, it is possible that the observed 9-residue interaction of PIASy with PARP1$^{503\text{-}650}$ in this work would be altered with the addition of the PARP1 minimum domain remainder. However, analysis of this addition was not possible by means of NMR due to solubility limitations and the unstructured nature of PARP1 residues 375-525. In the case that the absence of PARP1 residues 375-502 altered the PIASy interaction domain specific for lysine 482 modification, the observed 9-residue

interaction may correlate to a secondary PIASy-interaction domain to mediate an alternate SUMOylation site. PIASy has been shown to increase SUMOylation capacity of multiple PARP1 lysine residues, in addition to promoting residue specificity, which supports the likelihood of multiple PIASy-interaction domains on PARP1.[5] The same theme has also been observed in yeast where Siz1 (PIASy paralogue) has been shown to mediate SUMOylation of multiple residues of a single target protein, PCNA.[7]

In summary, the work presented here suggests that the PIASy E3 ligase independently associates with PARP1 prior to complex formation with Ubc9-SUMO-2/3. Future work may address folding and solubility restrictions by use of limited proteolysis and mass spectrometry to identify independently folded structural domains. Solubility has proven to be a critical barrier that must be addressed both for quality NMR titrations of the minimum PARP1 domain and for assignment of those residues interacting with PIASy. Identification of these residues will be essential in understanding the mechanism of PIASy-dependant SUMO conjugation.

REFERENCES

1. Geiss-Friedlander, R., and Melchior, F. (2007) Concepts in sumoylation: a decade on, *Nat. Rev. Mol. Cell Biol.* 8, 947-956.
2. Meulmeester, E., and Melchior, F. (2008) Cell biology: SUMO, *Nature* 452, 709-711.
3. Hochstrasser, M. (2001) SP-RING for SUMO: New functions bloom for a ubiquitin-like protein, *Cell* 107, 5-8.
4. Azuma, Y., Arnaoutov, A., Anan, and T., Dasso, M. (2005) PIASy mediates SUMO-2 conjugation of Topoisomerase-II on mitotic chromosomes, *EMBO* 24, 2172-2182.
5. Ryu, H., Al-Ani, G., Deckert, K., Kirkpatrick, D., Gygi, S. P., Dasso, M., and Azuma, Y. (2010) PIASy mediates SUMO-2/3 conjugation of poly(ADP-ribose) polymerase 1 (PARP1) on mitotic chromosomes, *J. Biol. Chem.* 285, 14415-23.
6. Chang, P., Jacobson, M., K., and Mitchison, T. J. (2004) Poly(ADP-ribose) is required for spindle assembly and structure, *Nature* 432, 645-649.
7. Yunus, A. A., and Lima, C. D. (2009) Structure of the Siz/PIAS SUMO E3 ligase Siz1 and determinants required for SUMO modification of PCNA, *Mol Cell* 35, 669-682.
8. Geisbrecht, B. V., Bouyain, S., and Pop, M. (2006) An optimized system for expression and purification of secreted bacterial proteins, *Protein Expr. Purif.* 46, 23-32.
9. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* 6, 277-293.
10. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules, *Methods Mol. Biol.* 278, 313-352.
11. Grzesiek, S. and Bax, A. (1993) The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements, *J. Am. Chem. Soc.* 115, 12953-54.
12. Nagashima, T., Hayashi, F., and Yokoyama, S. RIKEN structural biology (to be published).