NMR STUDIES OF MOLECULAR INTERACTIONS INVOLVED IN THE TYPE III SECRETION SYSTEM, SUMOYLATION, AND THE RNA BINDING PROTEIN HuR

By

Copyright 2016

Kawaljit Kaur

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

_____
Chairperson: Roberto N. De Guzman, Ph.D.

_____
Yoshiaki Azuma, Ph.D.

_____
Liang Xu, M.D., Ph.D.

_____
Mark Richter, Ph.D.

_____
Krzysztof Kuczera, Ph.D.

_____
Christopher Fischer, Ph.D.

Date Defended: July 8, 2016

The Dissertation Committee for Kawaljit Kaur

certifies that this is the approved version of the following dissertation:

NMR STUDIES OF MOLECULAR INTERACTIONS INVOLVED IN

THE TYPE III SECRETION SYSTEM, SUMOYLATION,

AND THE RNA BINDING PROTEIN HuR

Chairperson: Roberto N. De Guzman, Ph.D.

Date Approved: July 14, 2016

Abstract

Proteins are one of the most intriguing, versatile, and complex macromolecules in living systems. Proteins rarely function independently and perform their activities through a multitude of interactions with other proteins or molecules. Such molecular interactions are fundamental to almost all biological processes and their disruption is often associated with cellular irregularities and disease states. It is therefore of immense importance and interest to identify and characterize the binding interfaces of these biologically relevant molecular interactions.

NMR spectroscopy is unparalleled in its ability to monitor molecular interactions in solution at atomic level over a wide range of affinities. In this body of work, various NMR methods were successfully used to study and characterize the protein-ligand interactions of three discrete systems: the bacterial type III secretion system (T3SS), the post-translational SUMO modification system, and post-transcriptional regulation by the RNA binding protein HuR.

The T3SS is a macromolecular structure assembled by many Gram-negative bacterial pathogens, such as, *Shigella flexneri*, *Yersinia pestis*, and multi-drug resistant *Pseudomonas aeruginosa*, to cause infectious diseases. The structural component of the T3SS, the needle apparatus, consists of a base, a needle, a tip complex, and a translocon. Because the needle apparatus is exposed on the bacterial surface, present only among the pathogens, and essential for the virulence, disrupting the needle assembly is an attractive strategy for the development of novel anti-virulence drugs. However, this approach demands a detailed understanding of the protein-protein interactions involved in the needle assembly. Here, NMR methods were used to characterize the protein-protein interactions that are important in the assembly of the tip-translocon complex in the *Shigella* T3SS. Additionally, fragment-based screening was performed to identify small molecule binders of the tip proteins from *Yersinia* and *Pseudomonas* T3SS. The hits

were subsequently validated and characterized using NMR spectroscopy. Our results provide novel insight into the assembly of the needle apparatus and reveal the first small molecules that directly bind to the tip proteins of *Yersinia* and *Pseudomonas* T3SS.

Small ubiquitin-like modifier (SUMO) conjugation is a reversible post-translational modification process that can modulate biochemical and cell biological functions of the target protein substrate. SUMO E3 ligases are the enzymes that carry out the final step in SUMO conjugation pathway and facilitate the transfer of SUMO to the target protein. Prior to this work, SUMO binding by the PIAS family of E3 ligases was poorly understood. Here, using NMR spectroscopy, the protein-protein interactions involved in the SUMO-PIASy binding were characterized. The NMR binding studies surprisingly uncovered a novel SUMO-interacting motif in the E3 ligase PIASy, which was found to be essential for the ligase activity of PIASy.

The RNA binding protein HuR binds to adenine- and uridine-rich elements (AREs) located in the untranslated region of target mRNAs, regulating their stability and translation. HuR-ARE interaction contributes to carcinogenesis by stabilization of oncogenic mRNAs. HuR is overexpressed in a broad range of human cancers and associated with poor clinical outcome. *In vitro* and *in vivo* studies have demonstrated that HuR is an attractive therapeutic target. Drugs that disrupt HuR-ARE interaction could potentially inhibit cancer growth and persistence. Here, a fungal natural product azaphilone (AZA-9) was identified as a novel disruptor of HuR-ARE interaction using fluorescence polarization based screening. AZA-9 binding to HuR was validated and characterized by NMR methods. Results of NMR studies suggest that AZA-9 binds in the ARE-binding cleft of HuR, and thus competitively inhibits the HuR-ARE interaction.

The work presented in this dissertation illustrates the strength of NMR spectroscopy and its wide applicability as a tool to characterize and understand diverse interactions of proteins.

Dedication

I lovingly dedicate this work to my mother, Varsha, and father, Amarjit Singh.

Acknowledgements

I would first like to convey my heartfelt gratitude to my advisor, Dr. Roberto N. De Guzman. His unwavering support, constant encouragement, and valuable guidance in all matters professional and personal has led me to learn and accomplish more than I ever envisioned. He introduced me to the world of structural biology and NMR spectroscopy for which I am eternally grateful. He constantly propelled me to think beyond the convention while giving me space to develop as an independent researcher. He instilled in me the confidence and ability to communicate scientific observations in an eloquent and comprehensible manner. His constructive critiques have helped me immensely in refining my scientific writing skills. I greatly appreciate his patience while I wrote manuscripts at my own pace. Apart from molding my scientific career, he inspired me to pay more attention to my overall health and wellness. I will forever be indebted to Roberto for his tremendous contribution to my life.

I thank all the current and former lab members of the De Guzman lab for making the lab an enjoyable and collaborative place to work. I am thankful to Fernando Estrada for training me on lab techniques, and guiding me in "learning the ropes" of being a graduate student. I am so grateful for the support, friendship, and encouragement from fellow labmates, Andrew McShan, Supratim Dey, Sikta Patnaik, and Pallavi Guha Biswas. I acknowledge Srirupa Chatterjee, Sukanya Chaudhury, and Thenmalar Rathinavelan for helping me with research and answering my queries. I also thank all the outstanding undergraduate researchers, especially Mason Wilkinson and James Fields, who helped me tremendously in my experiments. I recognize Xiaoqing (Sarah) Wu from Xu Lab for her assistance with the HuR project.

Furthermore, I graciously thank my dissertation committee members for their guidance and the time to enhance my research with their expertise. I thank Dr. Yoshiaki Azuma for allowing me to take

lead on the SUMO project, Dr. Liang Xu for giving me the opportunity to work on the HuR project, Dr. Mark Richter for being a generous lab neighbor and letting me use his lab equipment and reagents and also for assistance with fluorescence spectroscopy, Dr. Krzysztof Kuczera for introducing me to protein modeling, and Dr. Chris Fischer for his help in analyzing results from the fluorescence polarization experiments. I also thank Dr. Peter Gegenheimer for his enormous help during my comprehensive oral examination and for investing in my growth as a scientist. I thank Dr. Asokan Anbanandam for technical assistance with NMR spectroscopy and Dr. Mario Rivera for teaching me the theoretical principles of NMR. I acknowledge the faculty of Department of Molecular Biosciences (MB) for presenting me with the opportunity to pursue my passion in Biochemistry and Biophysics, and also for giving me the chance to teach biology and biochemistry undergraduate labs, which led me to realize my love for teaching. I also recognize the members of the MB staff, especially Judith Harris, John Connolly and Linda Wiley for their assistance with the routine tasks. I would also like to thank the faculty of Department of Biophysics at Panjab University, my undergraduate alma mater, for inculcating in me a natural curiosity for scientific research that has allowed me to pursue a Ph.D.

I gratefully acknowledge the financial support from all the funding sources. I also thank the 'Kenneth Armitage', 'Borgendale', and 'Newmark' award committees for bestowing me with such great honors in teaching and research.

Lastly, I would like to thank my family for their endless love and support. I am indebted to my father, Amarjit Singh, and mother, Varsha, for believing in me and motivating me to dream big. I am fortunate to be a daughter of such loving and encouraging parents. I also thank my wonderful sisters, Puneet and Ravneet, for their enduring support and advice in helping me make pivotal decisions. I will always treasure the amusing anecdotes and the long talks we have had over the years. I also thank my brothers-in-law, Sarpreet and Upkar, for the unforgettable adventures during my trips to Australia. I am thankful to my three little nephews, Japgur, Angad, and Virajveer, for making me laugh during the tough

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ARE – Adenine- and uridine-rich element
AZA – Azaphilone
CD – Circular Dichroism
CSD – Chemical Shift Deviation
DMSO – Dimethyl Sulfoxide
DOC – Deoxycholate
DTT – Dithiothreitol
ELAV – Embryonic Lethal Abnormal Vision-like
FM – Fluorescein-Maleimide
FP – Fluorescence Polarization
HMQC – Heteronuclear Multiple Quantum Coherence
HSQC – Heteronuclear Single Quantum Coherence
HuR – Human antigen R
ILV – Isoleucine, Leucine and Valine
IPTG – Isopropyl-β-D-thiogalactopyranoside
MTSL – S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate
NOE – Nuclear Overhauser Effect
NOESY – Nuclear Overhauser Effect Spectroscopy
NMR – Nuclear Magnetic Resonance
PAGE – Polyacrylamide Gel Electrophoresis
PARP1 – Poly [ADP-Ribose] Polymerase 1
PBS – Phosphate-Buffered Saline
PCR – Polymerase Chain Reaction
PEI – Polyethyleneimine
PIAS – Protein Inhibitor of Activated STAT
PMSF – Phenylmethanesulfonylfluoride
PRE – Paramagnetic Relaxation Enhancement
RRM – RNA Recognition Motifs
SAR – Structure-Activity Relationship
SDS – Sodium Dodecyl Sulfate
SEC – Size Exclusion Chromatography
SIM – SUMO-Interacting Motif
SPR – Surface Plasmon Resonance
STD – Saturation Transfer Difference
SUMO – Small Ubiquitin-like Modifier
T3SS – Type III Secretion System
TEV – Tobacco Etch Virus
Topo IIα – Topoisomerase IIα
TROSY – Transverse Relaxation Optimized Spectroscopy
UTR – Untranslated Region

Chapter 1: Introduction

1.1. NMR studies of protein interactions

Structural and functional characterization of molecular interactions is fundamental to the understanding of many essential biological processes (1-3). Often times, such interactions are weak and transient, and therefore, difficult to investigate by biophysical methods such as pull-down, and crystallography. Nuclear Magnetic Resonance (NMR) spectroscopy is an exquisitely powerful tool for analyzing these weak biomolecular interactions in atomic details at physiological conditions (1-6).

1.1.1. NMR chemical shift mapping

One of the most widely used NMR methods for studying molecular interactions is called NMR chemical shift mapping (or differential chemical shift perturbation method) (7,8). It is a robust, reliable, and relatively straightforward technique that can provide information about both location and strength of the binding event from the same set of measurements (7-9). In protein interaction studies, NMR chemical shift mapping monitors changes in the chemical shifts of an isotopically labeled protein induced upon the addition of an unlabeled ligand/binding partner (*e.g.* a protein, nucleic acid, or a small molecule). NMR chemical shifts are extremely sensitive to the chemical environment of the nucleus, and therefore, NMR-active nuclei located at the binding interface typically experience largest changes in the chemical shifts upon complex formation, thus enabling determination of molecular interaction interfaces (6-9).

1.1.2. *NMR exchange regimes in protein-ligand interactions*

The appearance of the NMR spectra in protein interaction studies varies depending on binding affinity of the ligand ([10]). Three broad NMR chemical exchange conditions, namely, slow, intermediate, or fast, can be observed **(Figure 1-1)**. Slow exchange on NMR chemical shift time scale indicates a slow dissociation complex (tight-binding interactions) that rarely dissociates during the detection period of the NMR experiment ([5,8,10]). Two sets of resonances/peaks corresponding to the free and ligand-bound protein are observed. As the concentration of the ligand is increased, new peaks from the bound state grow in intensity while simultaneously peaks from the free protein gradually decrease in intensity until they disappear **(Figure 1-1a)**. Significant interconversion between the free and the ligand-bound state result in the intermediate exchange on NMR time scale. The signals become weaker and poorly defined ([5,8,10]). Peaks affected by the interaction show substantial reduction in the intensity and may even disappear during the titration. The peak broadening may also be accompanied by shifts in the peak positions **(Figure 1-1b)**. Lastly, fast dissociation complexes (or weak binding interactions) result in fast exchange on NMR time scale. The extremely fast interconversion between the free and ligand-bound forms during the detection period of the NMR experiment results in the observed chemical shift to be the weighted average of the chemical shifts for the free and bound form ([5,8,10]). The peaks move continuously from free to the bound position with increasing ligand concentrations **(Figure 1-1c)**. The three NMR chemical exchange regimes, slow, intermediate, and fast are typically associated with submicromolar, micromolar and high micromolar to millimolar dissociation constants, respectively ([5,6,8,10]). Irrespective of the kind of NMR chemical exchange regime, the bound state of the complex can be captured in the presence of excess unlabeled ligand ([5,10]).
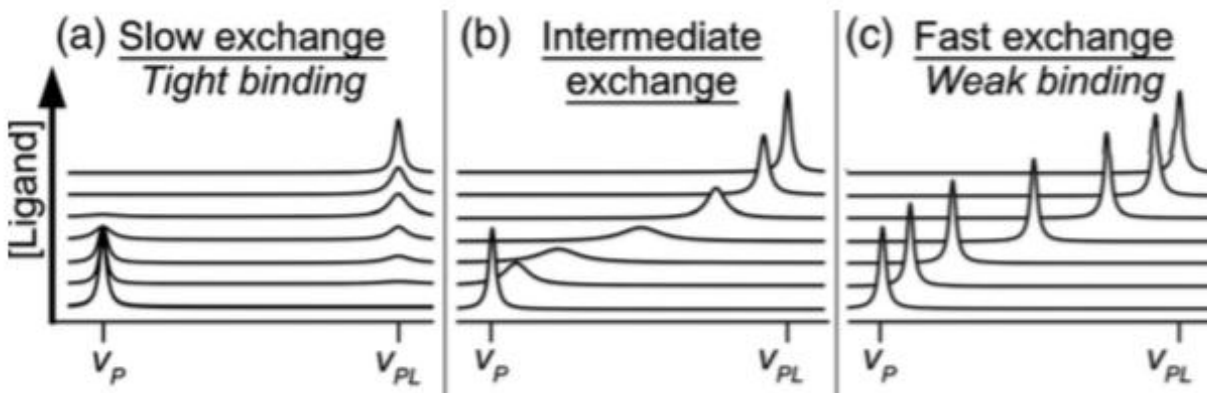
Figure 1-1. Exchange regimes on the NMR chemical shift time scale for protein-ligand complex formation. Distinct signals can be observed for the free and ligand-bound state of protein, at υ_P and υ_{PL}, respectively. The exchange regime **(a)** slow, **(b)** intermediate, or **(c)** fast, is strongly influenced by the ligand binding affinity. Figure adapted from ([10](#)).

1.1.3. Isotopic labeling for protein interaction studies

Characterization of the protein interactions using NMR spectroscopy requires labeling the protein of interest with stable NMR isotopes. Although the predominant hydrogen isotope ^{1}H isotope is NMR active, the most abundant nitrogen (^{14}N) and carbon (^{12}C) isotopes are NMR inactive and need to be substituted with NMR active nuclei (^{15}N and ^{13}C) through isotopic enrichment. Labeling with stable isotopes is routinely achieved by expressing the protein of interest in bacteria grown on appropriate minimal media with desired isotopically labeled nitrogen or/and carbon sources ([4-6](#)).

Based on the type of the experiment to be collected, protein size and quality of the NMR data, different labeling schemes can be utilized in the NMR studies of protein interactions. The labeling schemes can be broadly classified into two categories: uniform and selective labeling. Uniform labeling produces a protein with uniformly incorporated NMR active isotope(s) at all the sites, whereas with selective labeling, the protein is enriched with an isotope only at specific sites of interest. Selective labeling is generally used with large proteins (MW ~25 kDa or more) in order to reduce the spectral

overlap and complexity of the NMR datasets for unambiguous analysis (5,9). Different isotopic labeling strategies include:

Uniform ^{15}N labeling. This is the simplest and the cheapest form of isotopic labeling achieved by growing bacteria in media supplemented with ^{15}N-ammonium chloride as the nitrogen source. ^{15}N-labeled proteins can be used to acquire standard backbone amide correlation experiments such as 2D ^{1}H-^{15}N HSQC (or TROSY) to examine whether the protein is folded and to test the effect of different experimental conditions such as, temperature, pH, and buffers on spectral quality of the protein for optimization (1,5,11). Based on the data quality, uniform ^{15}N-labeled proteins can be used as a probe to investigate ligand interactions.

Uniform ^{15}N, ^{13}C and ^{2}H labeling. Along with ^{15}N-ammonium chloride as the nitrogen source, the bacterial growth media can be supplemented with ^{13}C-glucose as the carbon source to produce ^{15}N and ^{13}C double-labeled proteins. These are generally used for acquisition of 3D NMR experiments to perform resonance assignments, which is crucial information for identification of the residues involved in the interactions. For large proteins, instead of H_2O, the proteins are expressed in D_2O-based minimal media (^{2}H labeling) to improve the relaxation properties of protein and thereby, improving the overall NMR spectral quality. Additionally, when partial deuteration is not sufficient (especially for very large proteins, > 30-35 kDa), ^{13}C-glucose is often replaced with deuterated d_7-^{13}C-glucose to further enhance the deuteration levels in protein, and thus, gaining on NMR quality (4,5,8,9,11).

^{15}N amino acid specific labeling. NMR interaction studies of large proteins is often complicated by significant spectral overlap or poor resolution. In such cases, it is beneficial to label the backbone amides of few specific amino acids in the protein to decrease crowding in the spectra and monitor amino acids of interest without interference from other signals. Besides serving as a probe for protein interaction studies, ^{15}N amino acid specific labeling is also useful in the NMR assignment process. ^{15}N

amino acid specific labeling is achieved by supplying media with a specific ^{15}N-labeled amino acid of interest, while all the remaining amino acids are provided in the unlabeled form. However, due to metabolic conversion between amino acids in the biosynthetic pathway of the cell, the method can mis-incorporate ^{15}N label in undesired amino acids, also known as isotope scrambling. For amino acids, Ala, Ile, Leu, Val, Lys, Met, Phe, and Tyr, scrambling is minimal and specific labeling is routinely performed, but amino acids that are higher up or intermediates in the metabolic pathway, such as, Asp, Asn, Glu, and Gln, specific labeling can be particularly challenging due to severe interconversion (5,9,10).

Side chain methyl ILV (Ile, Leu, and Val) labeling. This method involves use of synthetic ^{13}C-keto acids as precursors to selectively label the side chain methyl groups of isoleucine (^{13}Cδ1), leucine (^{13}Cδ), and valine (^{13}Cγ) residues in the protein of interest. An extension (or variant) of the ILV labeling involving side chain methyl group labeling of alanine (^{13}Cβ), methionine, and/or threonine residues is also available. The ^{1}H-^{13}C NMR signals from side chain methyl groups of protein with inherently high sensitivity, sharp NMR resonances and favorable relaxation properties offer a powerful approach for the spectroscopic studies of large proteins. Since most protein interactions involve hydrophobic interfaces and methyl groups frequently occur at the hydrophobic cores of protein, sensitive ^{13}C methyl groups provide remarkable probes to qualitatively characterize protein-ligand interactions (5,9,12,13). Side chain methyl labeling can also be carried out in combination with the previously described ^{15}N backbone labeling to simultaneously investigate the effect of ligand binding on side chain and backbone groups using the same NMR sample.

1.1.4. Other NMR methods for characterizing protein-ligand interactions

Paramagnetic relaxation enhancement (PRE). In addition to NMR chemical shift mapping, a sensitive NMR method based on PRE can be employed to study weak protein-protein interactions. PRE studies require paramagnetic spin labels that are attached to one of the two interacting proteins through a surface-exposed cysteine residue. The PRE effect can be detected over a long range (up to 20-25 Å) using nitroxide spin labels, such as, MTSL. The spin-labeled protein is titrated with an isotopically labeled binding partner and residues at the interaction interface due to the close proximity to the paramagnetic label experience 'enhanced relaxation' and display significant reduction in the peak intensities, thus facilitating determination of the interaction surfaces (5,9,14,15).

Saturation transfer difference (STD) spectroscopy. In contrast to the previously described protein-detected NMR methods, STD NMR is a ligand-observed 1D-NMR technique used in the characterization of weak protein-ligand interactions. The method is based on the nuclear Overhauser effect (NOE) and involves observing the effect of selective saturation of the protein signals on the ligand signals (4,6,16). STD-NMR experiment is frequently used as a screening tool in the drug discovery process and is useful for identifying ligand moieties responsible for binding (16,17).

In this dissertation, I took advantage of various NMR methods, such as, chemical shift mapping, PRE, and STD, and different isotope labeling strategies, such as, ^{15}N, ^{13}C ILV, ^{15}N amino acid specific, and deuteration, to characterize molecular interactions involved in three distinct systems: the bacterial type III secretion system, the SUMO conjugation pathway, and the RNA binding protein HuR.

1.2. The type III secretion system (T3SS)

Many pathogenic Gram-negative bacteria assemble a protein nanoinjector called the type III secretion system (T3SS) to inject virulence effector proteins directly into the eukaryotic host cells to initiate infection ([18-20](#)). Bacteria harboring the T3SS include the causative agents of bacillary dysentery (*Shigella flexneri*), typhoid fever (*Salmonella typhimurium*), bubonic plague (*Yersinia pestis*), secondary hospital infections (*Pseudomonas aeruginosa*), and sexually transmitted diseases (*Chlamydia trachomatis*) ([21](#)). Together, they are responsible for millions of deaths worldwide each year. In the United States alone, these pathogens result in an estimated 2 million infections and 23,000 deaths per year ([22](#)). The rise of antibiotic resistance among these pathogens poses a major global health concern ([22](#)). Since the T3SS is indispensable in the pathogenesis of these clinically relevant Gram-negative pathogens, disrupting the assembly of the T3SS represents an attractive target for the development of novel anti-infectives ([23-25](#)).

1.2.1. Components of the T3SS

The T3SS consists of a needle apparatus, effector proteins, and chaperones **(Figure 1-2)**. The structural component of the T3SS, the needle apparatus, works together with an export apparatus and an ATPase complex to secrete the effector proteins directly into eukaryotic host cells. Chaperones serve as regulators of secretion by forming complexes with effectors and other T3SS structural proteins while inside the bacterial cytoplasm to maintain them in a pre-active state ([20](#),[26](#)). The needle apparatus is assembled from over 20 different proteins and is composed of a base that spans the bacterial membranes, an external needle with a 25 Å wide channel, a tip complex, and a translocon **(Figure 1-2B)** ([20](#),[27](#),[28](#)). The tip complex plugs the needle until contact with host cell is established ([29](#)). Upon host cell contact, a translocon is assembled between the needle tip complex and the host cell. The translocon creates a pore

8

in the host cell membrane to allow the passage of effector proteins into the host cell cytoplasm. Once inside the host cytoplasm, effectors modulate various host cell pathways, such as, cytoskeletal dynamics, and cell signaling, to enable the pathogen to invade, survive and multiply within the host environment (28,30). High-resolution atomic models of the base and the needle are currently available (30-33). However, the assembly of the tip complex and the translocon remain poorly understood.
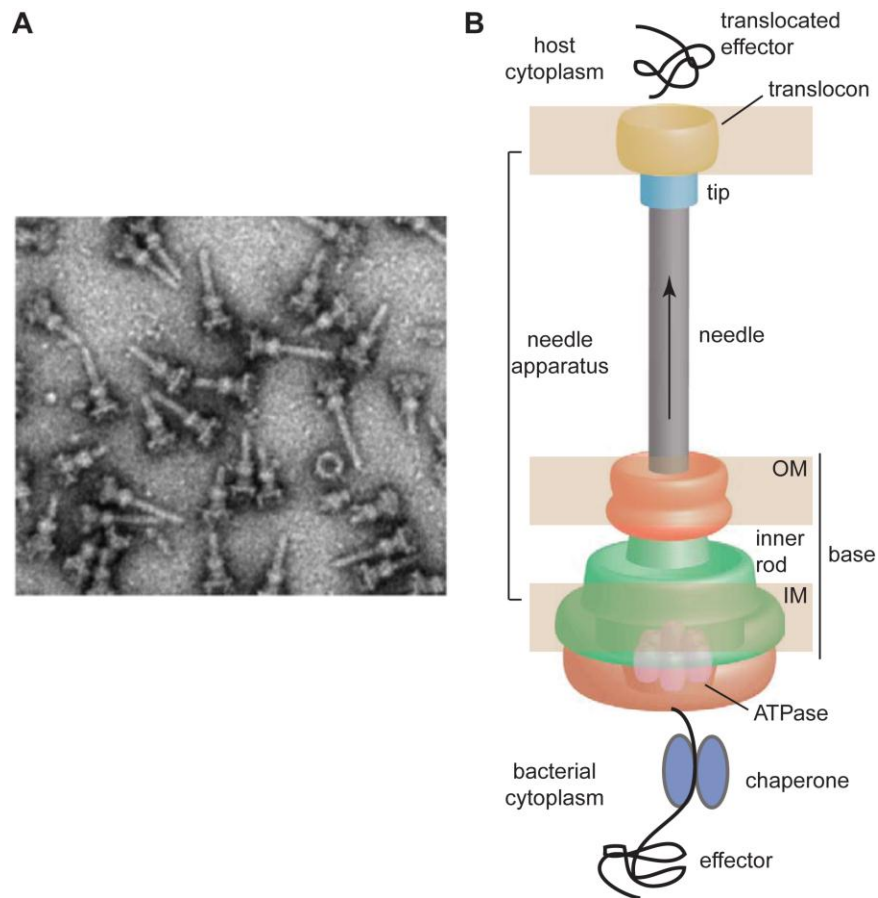
Figure 1-2. The bacterial type III secretion system. (A) Electron micrograph of purified *Salmonella* T3SS needle complexes (34). **(B)** Cartoon of the T3SS. The needle apparatus consists of a base, an extracellular needle, a tip complex, and a translocon. Gram-negative pathogens utilize the needle apparatus to deliver effector proteins directly into the host cytoplasm to initiate infectious diseases. The figure is adapted from (27).

1.2.2. The tip complex and the translocon

On top of the needle sits a tip complex formed by an estimated four to six copies of the tip protein **(Table 1-1)**, IpaD (in *Shigella*), SipD (in *Salmonella*), BipD (in *Burkholderia*), LcrV (in *Yersinia*), or PcrV (in *Pseudomonas*) ([20](#),[26](#),[28](#)). The multi-functional tip proteins play crucial roles in host cell sensing, translocon assembly, and effector secretion regulation ([28](#),[35](#)). The atomic structures of tip proteins are known ([36-39](#)). IpaD, SipD, and BipD (belonging to the *Shigella* family) show an overall oblong shape and contain an N-terminal α-helical hairpin, a long central coiled-coil, and a distal region with mixed α-β structural elements ([37](#),[38](#)). Dumbbell-shaped LcrV (from the *Yersinia* family) although contains the conserved central coiled-coil domain and a less-structured distal α-β domain, it lacks the N-terminal α-helical hairpin and instead, possess a globular N-terminal domain ([36](#),[39](#)). Specific functions have been proposed for these distinct domains. The highly conserved central coiled-coil region of the tip proteins has been shown to be important for interaction with the needle proteins ([15](#),[40](#)). The N-terminal α-helical hairpin found in the *Shigella* family functions as a self-chaperone to prevent premature oligomerization of the tip proteins within the bacterial cytoplasm ([37](#)). In the *Yersinia* family that lacks the self-chaperoning domain, small cytoplasmic proteins, LcrG and PcrG, have been shown to function as cognate chaperones for the tip proteins LcrV and PcrV, respectively ([41](#)). Other than the chaperone role, LcrG and PcrG also facilitate secretion of their respective tip proteins and play important roles in effector secretion regulation in the *Yersinia* and *Pseudomanas* T3SS ([26](#),[28](#)).

The tip complex serves as a platform for the formation of the translocon. The translocon is assembled from two integral membrane proteins, the major and the minor translocon protein, termed based on the difference in their molecular weights **(Table 1-1)** ([28](#)). IpaB and IpaC (in *Shigella*), SipB and SipC (in *Salmonella*), YopB and YopD (in *Yersinia*), and PopB and PopD (in *Pseudomonas*) are the major and the minor translocon proteins, respectively ([26](#),[28](#)). The translocon proteins are responsible for

10

the formation of the translocation pore in the target host cell membrane to deliver effector proteins into the host cell cytoplasm. The atomic structure of any full-length translocon protein is currently unknown, however partial crystal structures of a protease resistant soluble N-terminal ectodomain from major translocon proteins IpaB and SipB displayed the presence of α-helical anti-parallel coiled-coil motifs (42). The assembly mechanism and membrane insertion of the translocon proteins is not understood.

Bacteria	Disease	Base	Needle	Tip	Translocon
Shigella	Shigellosis	MxiD, MxiI, MxiJ, MxiG	MxiH	IpaD	IpaB, IpaC
Salmonella	Salmonellosis	InvG, PrgJ, PrgK, PrgH	PrgI	SipD	SipB, SipC
Yersinia	Bubonic Plague	YscC, YscI, YscJ, YscD	YscF	LcrV	YopB, YopD
Pseudomonas	Infection in immunocompromised	PscC, PscI, PscJ, PscD	PscF	PcrV	PopB, PopD

Table 1-1. List of homologous proteins of the T3SS needle apparatus from various bacteria.

Although in the absence of the tip protein, the translocon proteins can be efficiently secreted, they cannot insert into the host cell membrane (26), indicating the importance of the tip-translocon protein-protein interactions for the assembly of a functional translocon. The way in which the tip protein interacts with the translocon protein is poorly characterized. In this dissertation, I used NMR chemical shift mapping and PRE methods to characterize the interaction between the *Shigella* tip protein IpaD and the translocon protein IpaB. The results of IpaD-IpaB protein-protein interaction studies are described in Chapter 2.

The rise of antibiotic resistance in pathogens requires new targets for developing novel anti-bacterials (22). Because of the essential role of T3SS in bacterial virulence, it is an attractive target for

the development of novel anti-infectives ([23-25](#)). An initial step in the development of T3SS-based anti-infectives is the identification of small molecules that can bind to T3SS proteins. In this direction, SPR screening was performed to identify small molecule binders of the virulent tip proteins LcrV and PcrV. The identified hits were validated and characterized using ligand and protein-detected NMR methods. The results of these protein-small molecule interaction studies and the first methyl ILV assignments from the *Yersinia* family of tip proteins are presented in Chapter 3.

The key internal findings and recommendations for future T3SS molecular interaction studies are summarized in Chapter 4.

1.3. The small ubiquitin-like modifier (SUMO) conjugation

SUMOylation (or <u>s</u>mall <u>u</u>biquitin-like <u>mo</u>difier conjugation) is a reversible post-translational modification process that regulates diverse cellular processes, such as, transcription, replication, DNA repair, chromosome segregation, and signal transduction, and is essential for the survival of most organisms ([43](#),[44](#)). The three-dimensional structure of SUMO proteins closely resembles that of ubiquitin. However, they share less than 20% sequence identity and have different overall surface electrostatics ([45](#),[46](#)). SUMO is expressed by all eukaryotes, but is absent from bacteria and archaea. Organisms like yeast, *C. elegans*, and *D. melanogaster*, have a single SUMO gene, whereas plants and vertebrates express several SUMO proteins ([44](#),[45](#)). Humans express at least four SUMO isoforms, SUMO-1 through SUMO-4, each encoded by a different gene. SUMO-1, 2, and 3 are ubiquitously expressed and participate in SUMOylation, however, it remains unclear whether SUMO-4, which is mainly expressed in kidney and spleen, can be processed or conjugated to cellular proteins ([43-45](#)). SUMO-2 and SUMO-3 share ~97% sequence identity and are often referred to as SUMO-2/3. In contrast, SUMO-1 shares less

than 50% sequence identity to SUMO-2/3 (43-47). SUMO-1 and SUMO-2/3 are functionally distinct. While SUMO-2/3 can engage in SUMO-chain formation owing to the presence of SUMO-consensus motifs (ψKXE) in their N-terminal extensions, SUMO-1 lacks the SUMO-consensus site and is unable to form SUMO chains. Furthermore, SUMO-1 and SUMO-2/3 respond differently to stress conditions and display preferential substrate binding (43,44,46).

1.3.1. The SUMOylation pathway

Similar to the ubiquitination pathway, SUMO proteins are covalently conjugated to the target substrate by an enzymatic cascade involving E1, E2, and E3 enzymes **(Figure 1-3A)** (43,45). The process initiates with the proteolytic processing of the newly synthesized SUMO by SENPs (SUMO specific proteases) to reveal the C-terminal Gly-Gly (GG) motif in the 'mature' SUMO. The mature SUMO undergoes an ATP-dependent activation by SUMO E1 activating enzyme, Aos1/Uba2. Following activation, SUMO is transferred to the catalytic cysteine residue of the SUMO E2 conjugating enzyme, Ubc9, forming an E2~SUMO thioester complex. Ubc9 can then catalyze conjugation to the target substrate in an E3 ligase independent manner via recognition of the SUMO consensus motif (ψKXE, where ψ is a large hydrophobic residue, K is the acceptor lysine residue, X represents any residue, and E is an acidic residue) and form an isopeptide bond between the C-terminal glycine residue of SUMO and the acceptor lysine residue in the substrate. However, this step is usually facilitated by SUMO E3 ligases and is particularly important when SUMO conjugation has to occur at non-consensus lysine residues in the substrate. SUMO modification provides a platform for new protein-protein interactions. SUMO-modified substrates can interact with SUMO-binding proteins through their SUMO-interacting motifs (SIMs). SUMOylation is a reversible process, SENP proteases can de-conjugate SUMO from the substrate and free SUMO may be utilized for another round of modification **(Figure 1-3A)** (43-46).

1.3.2. The SUMO E3 ligase, PIASy

All SUMO modifications use the same E1 (Aos1/Uba2) and E2 (Ubc9) enzymes. However, several different types of SUMO E3 ligase enzymes have been identified, including members of the Siz/PIAS (Protein Inhibitor of Activated STAT) family, nuclear pore protein Nup358 or RanBP2, polycomb group protein Pc2, and recently categorized ZNF451 (48-54). SUMO E3 ligases act as catalysts that enhance the transfer of SUMO from charged E2 (E2~SUMO) to the target substrate. The largest group of SUMO E3 ligases, the Siz/PIAS family, is characterized by the presence of an SP-RING domain, which is predicted to resemble the RING domain in ubiquitin RING E3 ligases and function in an analogous manner (43,46,55). Members of the Siz/PIAS family include the yeast Siz proteins and human PIAS proteins (PIAS1, two variants of PIAS2, PIAS3, and PIAS4/y) that share important conserved functional domains **(Figure 1-3B)**. The N-terminal SAP (scaffold attachment factor-A/B, acinus and PIAS) domain of Siz/PIAS ligases is involved in AT-rich DNA binding, the PINIT domain is important for substrate recognition, the SP-RING domain related to the ubiquitin RING E3 ligases binds E2 Ubc9, and SIM together with the acidic domain at the C-terminus mediates non-covalent SUMO binding **(Figure 1-3B)** (43,52,55,56).

Several SUMO E3 ligases display specificity toward a particular SUMO isoform and even, the target substrate. PIASy (a member of Siz/PIAS E3 ligase family) preferentially conjugates SUMO-2/3 and not SUMO-1 to mitotic SUMOylation substrates (48,57,58). However, the molecular basis for this SUMO paralog specificity has not been determined. To gain an understanding of this phenomenon, I performed NMR interaction studies of PIASy and SUMO proteins, which lead to an unexpected discovery of a novel SUMO-interacting motif (SIM) in PIASy. The identification, characterization and implications of the new SIM of PIASy are presented in Chapter 5.

A

Maturation

Nature Reviews | Molecular Cell Biology

B

Siz/PIAS SUMO E3 ligases

SAP	PINIT	SP-RING	SIM S/DE
↓	↓	↓	↓
AT-rich DNA binding (nuclear localization)	substrate recognition	E2 (Ubc9) binding	SUMO interaction
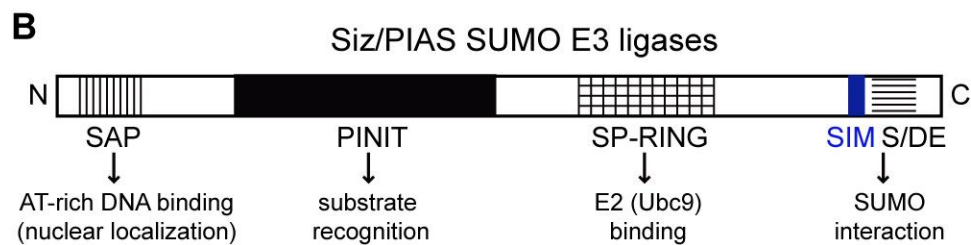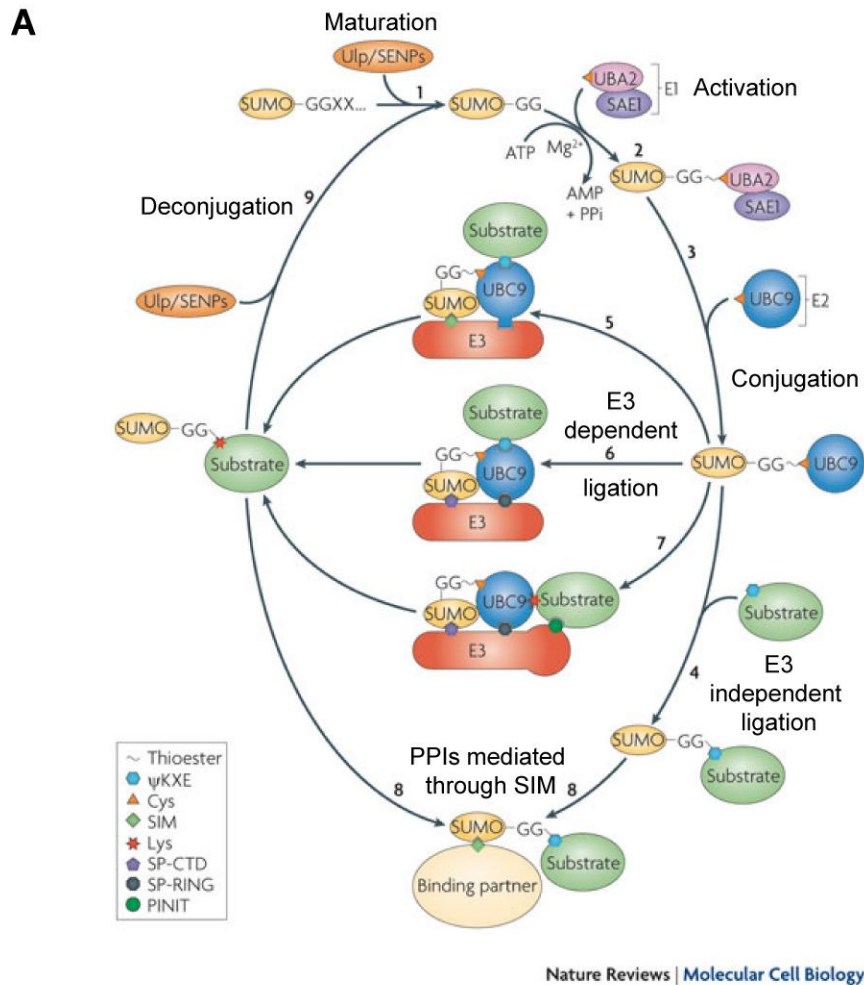
Figure 1-3. **The SUMO conjugation pathway and E3 ligases.** **(A)** SUMO conjugation requires sequential action of SUMO E1 activating, SUMO E2 conjugating, and SUMO E3 ligating enzymes. Although SUMO E2 can catalyze conjugation to a substrate in an E3-independent manner through recognition of SUMO consensus motifs (ψKXE), E3 ligases often enhance SUMO transfer and can impart substrate specificity for directing conjugation to non-consensus lysine residues. Figure adapted from (43). **(B)** Schematic showing conserved domains of the Siz/PIAS SUMO E3 ligase family: SAP (vertical lines), PINIT (solid black), SP-RING (grid), SUMO-interacting motif (SIM, blue), and S/DE (horizontal lines) and their roles in SUMOylation.

1.4. The RNA-binding protein HuR

Posttranscriptional regulation of gene expression is crucial for proper growth and development in all organisms. mRNA decay is one major factor that influences posttranscriptional gene expression (59,60). In mammalian cells, mRNA decay is tightly regulated by specific *cis* elements present in the RNA and *trans* acting factors. The most widespread and well-studied instability *cis* acting elements of mRNA decay are represented by the adenine- and uridine-rich elements (AREs) located in the 3'- or 5'-untranslated regions (UTRs). The *trans* acting RNA-binding proteins associate with these *cis* AREs and regulate the fate of target mRNAs (59-62).

1.4.1. HuR-ARE binding

The human antigen R (HuR, also called HuA) is among the most widely studied RNA binding proteins that bind to AREs of target mRNAs and modulate their stability and translation (59,62). HuR is a member of the Hu/ELAV (embryonic lethal abnormal vision-like) family of RNA-binding proteins. In contrast to the other three family members, namely, HuB, HuC, and HuD, which are exclusively found in neuronal tissues, HuR is ubiquitously expressed (59,63). HuR contains three classic RNA-recognition motifs, RRM1, RRM2, and RRM3 that are highly conserved among the family members. The two tandem N-terminal RRM domains, RRM1 and RRM2 (together referred as RRM1/2), are responsible for binding and stabilizing AREs contained in mRNAs (59,63,64). The third RRM, RRM3, is thought to bind the poly-A tail, maintain the stability of RNA-protein complex, and is responsible for cooperative assembly of HuR oligomers on RNA (65). The basic hinge region between RRM2 and RRM3 contains a nuclear-cytoplasmic shuttling sequence (HNS) that enables HuR to translocate from nucleus to cytoplasm (66). Posttranslational modifications, mainly phosphorylation, by kinases such as, cyclin-dependent kinase 1, protein kinase C, and p38 mitogen-activated protein kinase, can influence nucleocytoplasmic

transport and mRNA interactions of HuR ([67-69](#)). Crystal structures of free and ARE-bound HuR RRM1/2 are available, which indicate distinct open and close conformations of HuR RRM1/2, respectively ([64](#)).

1.4.2. HuR and Cancer

A number of cancer-related transcripts, including mRNAs of proto-oncogenes, growth factors, and cytokines, contain AREs, which confer a short mRNA half-life ([59,60](#)). Many of these mRNAs have been characterized as HuR targets ([59,67,69](#)). Binding of HuR to the AREs of these cancer-associated mRNAs generally lead to mRNA stabilization and increased translation. HuR has been shown to promote tumorigenesis by enabling multiple cancer phenotypes **(Figure 1-4)** ([67,69](#)). HuR regulated oncogenic mRNAs encode proteins that are implicated in different tumor processes, such as, cell proliferation, cell survival, reduced immune recognition, angiogenesis, invasion, and metastasis **(Figure 1-4)**. HuR is overexpressed in a wide variety of cancer. High levels of cytoplasmic HuR have been found in colorectal, gastric, lung, breast, ovarian, pancreas, and skin carcinoma ([67-72](#)). Elevated cytoplasmic accumulation of HuR correlates with high-grade malignancy and serve as a prognostic factor of poor clinical outcome in those cancers ([67,69-72](#)). In a mouse xenograft model, injection of carcinoma cells with elevated HuR produced significantly larger tumors than control population, and reduction in HuR expression through RNA interference resulted in a significant decrease in tumor size ([70](#)). Taken together, HuR presents an attractive drug target for novel cancer therapies ([67-69,73,74](#)).
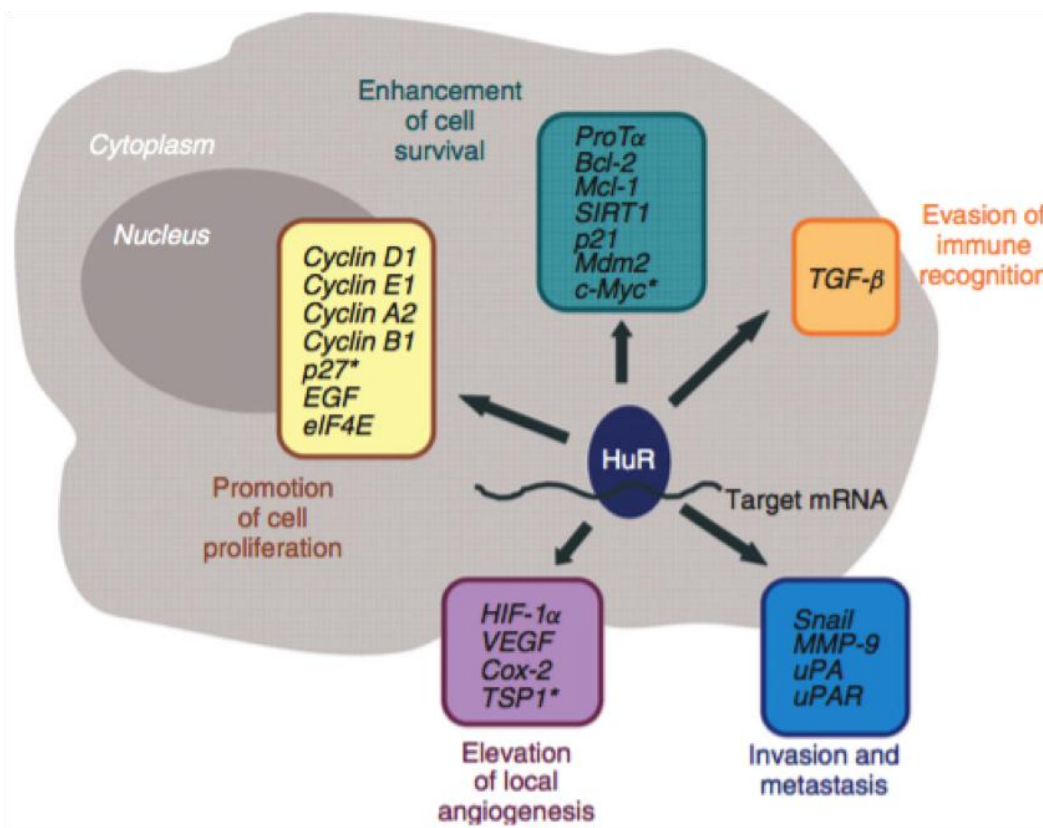
Figure 1-4. HuR regulated mRNAs implicated in establishing cancer traits. The target mRNAs of HuR are involved in all major cancer-acquired phenotypes, including cell proliferation, cell survival, reduced immune recognition, angiogenesis, invasion, and metastasis. Figure from (67).

I hypothesized that drugs that interfere with HuR-ARE interaction or disrupt the stabilizing effect of HuR upon mRNA targets could have dramatic effects on inhibiting cancer cell growth and progression (74). In this dissertation, using fluorescence polarization based screening, a fungal natural product was identified as an inhibitor of HuR-ARE interaction. I used NMR methods to validate and characterize small molecule and ARE binding to HuR. The inhibition and molecular interactions studies of HuR are presented in Chapter 6.

1.5. References

1. Vaynberg, J., and Qin, J. (2006) Weak protein–protein interactions as probed by NMR spectroscopy. *Trends in Biotechnology* **24**, 22-27
2. Perkins, J. R., Diboun, I., Dessailly, B. H., Lees, J. G., and Orengo, C. (2010) Transient protein-protein interactions: structural, functional, and network properties. *Structure*
3. Ozbabacan, S. E. A., Engin, H. B., Gursoy, A., and Keskin, O. (2011) Transient protein-protein interactions. *Protein Eng. Des. Sel.* **24**, 635-648
4. Takeuchi, K., and Wagner, G. (2006) NMR studies of protein interactions. *Current Opinion in Structural Biology* **16**, 109-117
5. Marintchev, A., Frueh, D., and Wagner, G. (2007) NMR methods for studying protein-protein interactions involved in translation initiation. *Translation Initiation: Reconstituted Systems and Biophysical Methods* **430**
6. Bieri, M., Kwan, A. H., Mobli, M., King, G. F., Mackay, J. P., and Gooley, P. R. (2011) Macromolecular NMR spectroscopy for the non-spectroscopist: beyond macromolecular solution structure determination. *FEBS Journal* **278**, 704-715
7. Foster, M. P., Wuttke, D. S., Clemens, K. R., Jahnke, W., Radhakrishnan, I., Tennant, L., Reymond, M., Chung, J., and Wright, P. E. (1998) Chemical shift as a probe of molecular interfaces: NMR studies of DNA binding by the three amino-terminal zinc finger domains from transcription factor IIIA. *J Biomol NMR* **12**, 51-71
8. Williamson, M. P. (2013) Using chemical shift perturbation to characterise ligand binding. in *Progress in Nuclear Magnetic Resonance Spectroscopy*, Elsevier B.V.
9. Barrett, P. J., Chen, J., Cho, M.-K., Kim, J.-H., Lu, Z., Mathew, S., Peng, D., Song, Y., Van Horn, W. D., Zhuang, T., Sönnichsen, F. D., and Sanders, C. R. (2013) The Quiet Renaissance of Protein Nuclear Magnetic Resonance. in *Biochemistry*
10. Kleckner, I. R., and Foster, M. P. (2011) An introduction to NMR-based approaches for measuring protein dynamics. *Biochim. Biophys. Acta* **1814**, 942-968
11. Fernandez, C. (2003) TROSY in NMR studies of the structure and function of large biological macromolecules. in *Current Opinion in Structural Biology*
12. Hajduk, P. J., Augeri, D. J., Mack, J., and Mendoza, R. (2000) NMR-Based Screening of Proteins Containing 13C-Labeled Methyl Groups - Journal of the American Chemical Society (ACS Publications). in *Journal of American Chemical Society*
13. Tugarinov, V., and Kay, L. E. (2005) Methyl Groups as Probes of Structure and Dynamics in NMR Studies of High-Molecular-Weight Proteins. in *ChemBioChem*
14. Otting, G. (2010) Protein NMR Using Paramagnetic Ions. *Annu. Rev. Biophys.* **39**, 387-405
15. Rathinavelan, T., Tang, C., and De Guzman, R. N. (2011) Characterization of the interaction between the Salmonella type III secretion system tip protein SipD and the needle protein PrgI by paramagnetic relaxation enhancement. *J Biol Chem* **286**, 4922-4930
16. Viegas, A., Manso, J. o., Nobrega, F. L., and Cabrita, E. J. (2011) Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding. *J. Chem. Educ.* **88**, 990-994
17. Jahnke, W., and Widmer, H. (2004) Protein NMR in biomedical research. in *Cell. Mol. Life Sci.*
18. Hueck, C. J. (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiology and Molecular Biology Reviews* **62**, 379-433
19. Kubori, T., Matsushima, Y., Nakamura, D., Uralil, J., Lara-Tejero, M., Sukhan, A., Galan, J. E., and Aizawa, S. I. (1998) Supramolecular structure of the Salmonella typhimurium type III protein secretion system. in *Science*
20. Cornelis, G. R. (2006) The type III secretion injectisome. in *Nat Rev Micro*, Nature Publishing Group

21. Coburn, B., Sekirov, I., and Finlay, B. B. (2007) Type III Secretion Systems and Disease. in *Clin. Microbiol. Rev.*
22. CDC. (2016) Antimicrobial resistance.
23. Keyser, P., Elofsson, M., Rosell, S., and Wolf-Watz, H. (2008) Virulence blockers as alternatives to antibiotics: type III secretion inhibitors against Gram-negative bacteria. in *J Intern Med*
24. Duncan, M. C., Linington, R. G., and Auerbuch, V. (2012) Chemical inhibitors of the type three secretion system: disarming bacterial pathogens. in *Antimicrobial Agents and Chemotherapy*
25. May, A. E., and Khosla, C. (2013) Discovery and Mechanism of Type III Secretion System Inhibitors. in *Isr. J. Chem.*
26. Galán, J. E., Lara-Tejero, M., Marlovits, T. C., and Wagner, S. (2014) Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells. *Annu. Rev. Microbiol.* **68**, 415-438
27. Yip, C. K., and Strynadka, N. C. J. (2006) New structural insights into the bacterial type III secretion system. *Trends in Biochemical Sciences* **31**, 223-230
28. Chatterjee, S., Chaudhury, S., McShan, A. C., Kaur, K., and De Guzman, R. N. (2013) Structure and Biophysics of Type III Secretion in Bacteria. in *Biochemistry*
29. 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? *Proceedings of the National Academy of Sciences* **105**, 6507-6513
30. Kosarewicz, A., Konigsmaier, L., and Marlovits, T. C. (2012) The blueprint of the type-3 injectisome. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 1140-1154
31. Schraidt, O., and Marlovits, T. C. (2011) Three-dimensional model of Salmonella's needle complex at subnanometer resolution. *Science* **331**, 1192-1195
32. Loquet, A., Sgourakis, N. G., Gupta, R., Giller, K., Riedel, D., Goosmann, C., Griesinger, C., Kolbe, M., Baker, D., Becker, S., and Lange, A. (2012) Atomic model of the type III secretion system needle. *Nature* **486**, 276-279
33. Demers, J.-P., Habenstein, B., Loquet, A., Kumar Vasa, S., Giller, K., Becker, S., Baker, D., Lange, A., and Sgourakis, N. G. (2014) High-resolution structure of the Shigella type-III secretion needle by solid-state NMR and cryo-electron microscopy. *Nat Commun* **5**, 4976
34. Zhong, D., Lefebre, M., Kaur, K., McDowell, M. A., Gdowski, C., Jo, S., Wang, Y., Benedict, S. H., Lea, S. M., Galan, J. E., and De Guzman, R. N. (2012) The Salmonella Type III Secretion System Inner Rod Protein PrgJ Is Partially Folded. *J Biol Chem* **287**, 25303-25311
35. Hiromi Sato, D. W. F. (2011) Multi-Functional Characteristics of the Pseudomonas aeruginosa Type III Needle-Tip Protein, PcrV; Comparison to Orthologs in other Gram-negative Bacteria. in *Frontiers in Microbiology*, Frontiers Media SA
36. 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. in *Structure*
37. 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) Self-chaperoning of the Type III Secretion System Needle Tip Proteins IpaD and BipD. *J Biol Chem* **282**, 4035-4044
38. 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 Science* **20**, 75-86
39. Chaudhury, S., Battaile, K. P., Lovell, S., Plano, G. V., and De Guzman, R. N. (2013) Structure of the Yersinia pestis tip protein LcrV refined to 1.65 Å resolution. in *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, International Union of Crystallography
40. Lunelli, M., Hurwitz, R., Lambers, J., and Kolbe, M. (2011) Crystal Structure of PrgI-SipD: Insight into a Secretion Competent State of the Type Three Secretion System Needle Tip and its Interaction with Host Ligands. *PLoS Pathog* **7**, e1002163

41.   Lee, P.-C., Stopford, C. M., Svenson, A. G., and Rietsch, A. (2010) Control of effector export by the Pseudomonas aeruginosatype III secretion proteins PcrG and PcrV. *International Journal of Microbiology* **75**, 924-941

42.   Barta, M. L., Dickenson, N. E., Patil, M., Keightley, A., Wyckoff, G. J., Picking, W. D., Picking, W. L., and Geisbrecht, B. V. (2012) The structures of coiled-coil domains from type III secretion system translocators reveal homology to pore-forming toxins. *J Mol Biol* **417**, 395-405

43.   Gareau, J. R., and Lima, C. D. (2010) The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. in *Nat Rev Mol Cell Biol*, Nature Publishing Group

44.   Meulmeester, E., and Melchior, F. (2008) Cell biology: SUMO. in *Nature*

45.   Geiss-Friedlander, R., and Melchior, F. (2007) Concepts in sumoylation: a decade on. in *Nat Rev Mol Cell Biol*, Nature Publishing Group

46.   Johnson, E. S. (2004) Protein modification by SUMO.

47.   Hay, R. T. (2005) SUMO. *Mol. Cell* **18**, 1-12

48.   Azuma, Y., Arnaoutov, A., Anan, T., and Dasso, M. (2005) PIASy mediates SUMO-2 conjugation of Topoisomerase-II on mitotic chromosomes. in *EMBO J*

49.   Cappadocia, L., Pichler, A., and Lima, C. D. (2015) Structural basis for catalytic activation by the human ZNF451 SUMO E3 ligase. in *Nat Struct Mol Biol*

50.   Eisenhardt, N., Chaugule, V. K., Koidl, S., Droescher, M., Dogan, E., Rettich, J., Sutinen, P., Imanishi, S. Y., Hofmann, K., Palvimo, J. J., and Pichler, A. (2015) A new vertebrate SUMO enzyme family reveals insights into SUMO-chain assembly. in *Nat Struct Mol Biol*

51.   Merrill, J. C., Melhuish, T. A., Kagey, M. H., Yang, S.-H., Sharrocks, A. D., and Wotton, D. (2010) A role for non-covalent SUMO interaction motifs in Pc2/CBX4 E3 activity. in *PLoS One*

52.   Rytinki, M. M., Kaikkonen, S., Pehkonen, P., Jääskeläinen, T., and Palvimo, J. J. (2009) PIAS proteins: pleiotropic interactors associated with SUMO. in *Cell. Mol. Life Sci.*, SP Birkhäuser Verlag Basel

53.   Gareau, J. R. J., Reverter, D. D., and Lima, C. D. C. (2012) Determinants of Small Ubiquitin-like Modifier 1 (SUMO1) Protein Specificity, E3 Ligase, and SUMO-RanGAP1 Binding Activities of Nucleoporin RanBP2. *Journal of Biological Chemistry* **287**, 4740-4751

54.   Reverter, D., and Lima, C. D. (2005) Insights into E3 ligase activity revealed by a SUMO–RanGAP1–Ubc9–Nup358 complex. in *Nature*

55.   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. in *Mol. Cell*, Elsevier Ltd

56.   Ryu, H., and Azuma, Y. (2010) Rod/Zw10 Complex Is Required for PIASy-dependent Centromeric SUMOylation. *J Biol Chem* **285**, 32576-32585

57.   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. in *J Biol Chem*

58.   Ryu, H., Furuta, M., Kirkpatrick, D., Gygi, S. P., and Azuma, Y. (2010) PIASy-dependent SUMOylation regulates DNA topoisomerase II activity. in *J. Cell Biol.*

59.   Brennan, C. M., and Steitz, J. A. (2001) HuR and mRNA stability. in *Cell. Mol. Life Sci.*, Birkhäuser Verlag

60.   Ross, J. (1995) mRNA stability in mammalian cells. in *Microbiological reviews*

61.   Fan, X. C. (1998) Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the invivo stability of ARE-containing mRNAs. in *EMBO J*, EMBO Press

62.   Peng, S. S. Y. (1998) RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein. in *EMBO J*, EMBO Press

63.   Cheng, S. (1996) Cloning and Characterization of HuR, a Ubiquitously Expressed Elav-like Protein. in *Journal of Biological Chemistry*

64.   Wang, H., Zeng, F., Liu, Q., Liu, H., Liu, Z., Niu, L., Teng, M., and Li, X. (2013) The structure of the AREbinding domains of Hu antigen R (HuR) undergoes conformational changes during RNA binding. *Acta Cryst (2013). D69, 373-380* 1-8

65. Fialcowitz-White, E. J., Brewer, B. Y., Ballin, J. D., Willis, C. D., Toth, E. A., and Wilson, G. M. (2007) Specific protein domains mediate cooperative assembly of HuR oligomers on AU-rich mRNA-destabilizing sequences. in *Journal of Biological Chemistry*

66. Fan, X. C., and Steitz, J. A. (1998) HNS, a nuclear-cytoplasmic shuttling sequence in HuR. in *Proceedings of the National Academy of Sciences*

67. Abdelmohsen, K., and Gorospe, M. (2010) Posttranscriptional regulation of cancer traits by HuR. *Wiley Interdiscip Rev RNA* **1**, 214-229

68. Zucal, C., D'Agostino, V., Loffredo, R., Mantelli, B., NatthakanThongon, B. S. P., Lal, P., Latorre, E., and Provenzani, A. (2015) Targeting the Multifaceted HuR Protein, Benefits and Caveats. in *CDT*

69. Wang, J., Guo, Y., Chu, H., Guan, Y., Bi, J., and Wang, B. (2013) Multiple functions of the RNA-binding protein HuR in cancer progression, treatment responses and prognosis. *Int J Mol Sci* **14**, 10015-10041

70. de Silanes, I. L., Fan, J., Yang, X., Zonderman, A. B., Potapova, O., Pizer, E. S., and Gorospe, M. (2003) Role of the RNA-binding protein HuR in colon carcinogenesis. in *Oncogene*

71. Heinonen, M., Fagerholm, R., and Aaltonen, K. (2007) Prognostic Role of HuR in Hereditary Breast Cancer. in *Clinical Cancer Research*

72. Yi, X., Zhou, Y., Zheng, W., and Chambers, S. K. (2009) HuR expression in the nucleus correlates with high histological grade and poor disease-free survival in ovarian cancer. in *Australian and New Zealand Journal of Obstetrics and Gynaecology*

73. Wang, Z., Bhattacharya, A., and Ivanov, D. N. (2015) Identification of Small-Molecule Inhibitors of the HuR/RNA Interaction Using a Fluorescence Polarization Screening Assay Followed by NMR Validation. in *PLoS One*, Public Library of Science

74. Wu, X., Lan, L., Wilson, D. M., Marquez, R. T., Tsao, W.-c., Gao, P., Roy, A., Turner, B. A., McDonald, P., Tunge, J. A., Rogers, S. A., Dixon, D. A., Aubé, J., and Xu, L. (2015) Identification and Validation of Novel Small Molecule Disruptors of HuR-mRNA Interaction. in *ACS Chem. Biol.*

Chapter 2: Characterization of the *Shigella* Type III Secretion System Tip-Translocon Protein-Protein Interaction by NMR Spectroscopy

Kawaljit Kaur, Andrew C. McShan, Srirupa Chatterjee, and Roberto N. De Guzman

Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas 66045, USA
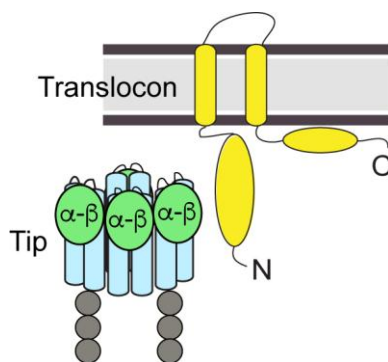
(The data presented in this chapter have been previously published and reformatted for this dissertation)

2.1. Abstract

Many Gram-negative pathogens, such as *Shigella*, *Salmonella*, *Yersinia*, and *Pseudomonas*, assemble the type III secretion system (T3SS) to inject virulence proteins directly into eukaryotic cells to initiate infectious diseases. The needle apparatus of the T3SS consists of a base, an extracellular needle, a tip protein complex, and a translocon. The atomic structure of the assembled tip complex and the translocon is currently unknown. Here, we show by NMR chemical shift mapping and paramagnetic relaxation enhancement (PRE) that the mixed α-β domain at the distal region of the *Shigella* tip protein IpaD interacts with the N-terminal ectodomain of the major translocon protein IpaB. Furthermore, incorporation of the extreme N-terminal residues of the translocon protein IpaB, previously thought to be important for the IpaD-IpaB interaction, had little change on the NMR-observed interaction. Our results reveal the binding surfaces involved in the tip–translocon protein–protein interaction and provide insight into the assembly of the needle apparatus of the T3SS.

2.2. Introduction

The T3SS plays a critical role in the pathogenesis of many Gram-negative bacteria, including the causative agent of bacillary dysentery, *Shigella flexneri* (1,2). To initiate infection, *Shigella* assembles the T3SS to inject virulence proteins directly into the target host cells (3). The needle apparatus of the T3SS consists of a base that spans the bacterial membranes (4-6), an external needle (7,8), and at the needle tip, a complex of an estimated 4-5 copies of the tip protein IpaD (in *Shigella* T3SS) (9). Upon host cell contact, the tip complex serves as a platform for the assembly of the translocon, which forms a pore in the host cell membrane that allows the passage of effector proteins directly into the host cell cytoplasm

24

(10). The atomic structure of the assembled tip complex and the translocon is currently unknown and how the tip protein interacts with the translocon protein is poorly understood.

The crystal structure of the tip protein IpaD shows an overall oblong shape with a long central coiled-coil juxtaposed to a N-terminal α-helical hairpin and a distal domain consisting of mixed α−helices and β−sheets (mixed α−β domain) **(Figure 2-1)** (11,12). The translocon is assembled from the major and minor translocon proteins (the qualifiers refer to their relative sizes; both are essential in the type III secretion). IpaB and IpaC are the major and minor translocon proteins of *Shigella,* respectively (4). The major translocon protein IpaB possesses structural features, such as, an N-terminal cytosolic ectodomain, a central hydrophobic region with two predicted transmembrane helices, and a C-terminal amphipathic region (13,14). The atomic structure for any full-length translocon protein is currently unknown, however, crystal structure of the N-terminal ectodomain of IpaB shows an α−helical trimeric antiparallel coiled-coil motif **(Figure 2-1)** (15).

Dickenson *et al*. (16) reported interaction between the IpaD and the N-terminal ectodomain of IpaB (residues 1-226) by FRET and showed that the interaction required presence of the bile salt deoxycholate (DOC) and that IpaB residues 11-27 were essential for the interaction. Here, we used NMR chemical shift mapping and paramagnetic relaxation enhancement (PRE) methods to identify the surfaces involved in the IpaD-IpaB protein-protein interaction. Our results indicate that the surface around the mixed α−β domain of the IpaD is involved in the interaction with the N-terminal domain of IpaB. Additionally, presence or absence of the extreme N-terminal residues of IpaB and DOC did not affect the IpaD-IpaB binding.
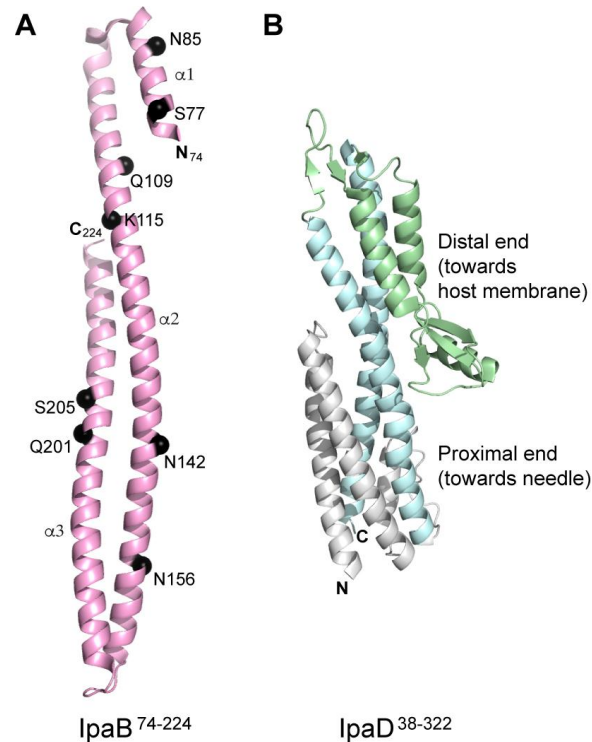
Figure 2-1. Crystal structure of IpaB$^{74-224}$ and IpaD$^{38-322}$. **(A)** The positions of the spin labels on IpaB used in the PRE studies are shown as black spheres. **(B)** Different domains of IpaD are colored as follows: N-terminal hairpin, gray; central coiled-coil, cyan; and distal mixed $\alpha-\beta$ region, green.

2.3. Methods

2.3.1. Protein expression and purification

IpaD (residues 38-332) wild type (referred as IpaD) and C322S mutant (referred as IpaD^{C322S}) were subcloned in the NdeI/SalI sites of a modified pET-21a expression vector. The IpaD construct contained an N-terminal His$_6$-tag followed by a tobacco etch virus (TEV) protease cleavage site. DNA corresponding to IpaB$^{76-308}$, IpaB$^{74-224}$, and IpaB$^{9-226}$ were PCR amplified and subcloned into the NdeI/XhoI sites of pET-22b. The IpaB constructs retained a non-cleavable C-terminal LEH$_6$ tag for

protein purification. Plasmids were freshly transformed in *E. coli* BL21 (DE3) and cells were grown in 1 L culture media containing 100 μg/ml carbenicillin. Unlabeled proteins were expressed in LB broth and ^{15}N-labeled proteins were obtained by cell growth in M9 minimal media supplemented with 1 g of ^{15}N-ammonium chloride. Bacteria were grown at 37 °C, induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at OD_{600} ~ 0.7-0.8, and cell growth was continued overnight at 15 °C. Cells were harvested by centrifugation and cell pellets were resuspended in 30 ml binding buffer (500 mM NaCl, 20 mM Tris-HCl pH 8.0, 5 mM imidazole). Cells were lysed by sonication in the presence of 0.1 mM phenylmethanesulfonylfluoride (PMSF). Cellular debris was removed by centrifugation at 13,000 rpm for 10 min, and 600 μl of 5% (v/v) polyethyleneimine was added to the supernatant to precipitate the nucleic acids. Following centrifugation (13,000 rpm, 10 min), the clear supernatant was loaded on a ~5 mL Ni^{2+}-affinity chromatography resin (Gold Biotechnology) and the Ni^{2+} column was washed with 100 ml binding buffer, followed by elution with 50 ml elution buffer (500 mM NaCl, 20 mM Tris-HCl pH 8.0, 250 mM imidazole). Fractions containing IpaD, were incubated overnight at room temperature with 250 μl of 0.04 mM recombinant TEV protease ([17]) in buffer (20 mM NaCl, 20 mM Tris-HCl pH 8.0, 0.5 mM EDTA, and 1 mM DTT) and the digest was passed through a Ni^{2+} column to separate the protein from the His_6-tag. IpaD retained a 3-residue 'GHM' cloning artifact. Purified proteins were dialyzed in NMR buffer (100 mM NaCl, 20 mM sodium phosphate pH 7.0) and concentrated using Amicon Ultra 10K centrifugal filter (Millipore) and protein concentrations were estimated by absorbance at A_{280}.

2.3.2. ^{15}N-amino acid-specific labeling of IpaD

^{15}N-leucine labeling was used in the PRE studies. The ^{15}N-leucine labeled IpaD was prepared following the published protocol with minor modification ([18]). Briefly, cells expressing IpaD from overnight 1 L LB culture were harvested and resuspended in 2X 500 ml M9 minimal media. The media was supplemented with 125 mg of ^{15}N-leucine and 300 mg of the remaining 19 unlabeled amino acids.

The starting OD_{600} was kept around 0.6-0.8 and cells were grown at 37 °C for 10 minutes followed by induction with 1 mM IPTG. Cell growth was continued for 4 hours at 37 °C or until OD_{600} was ~ 2.6 to 3.0. Recombinant proteins were purified as described above.

2.3.3. Engineering of cysteine mutations in IpaB[74-224] for spin labeling

Several non-conserved polar residues that are expected to be surface exposed in IpaB[74-224] were mutated into cysteine for the attachment of the spin label for PRE measurements. QuikChange site-directed mutagenesis was used to generate eight cysteine mutants in IpaB[74-224] residues Ser-77, Asn-85, Gln-109, Lys-115, Asn-142, Asn-156, Gln-201, and Ser-205 **(Figure 2-1)**. All mutations were confirmed by DNA sequencing. The cysteine mutant proteins were expressed and purified following the protocol described above. The purified cysteine-containing IpaB[74-224] protein was dialyzed in NMR buffer containing 5 mM DTT to prevent the formation of intermolecular disulfide bonds.

2.3.4. Spin labeling of IpaB[74-224]

A nitroxide spin label, MTSL (1-Oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate (Toronto Research Chemicals) was conjugated to the cysteine mutants of IpaB[74-224] following our published protocol ([19]). Briefly, 10 mg of MTSL was dissolved in 250 μl of acetone to make a 150 mM stock solution that was added at 7-fold molar excess to the protein for labeling. The extent of MTSL labeling was confirmed by ESI-MS.

2.3.5. Circular dichroism spectroscopy

CD spectra were acquired with a JASCO J-815 spectropolarimeter. Proteins were diluted to a final concentration of 1-2 μM in buffer (5 mM NaCl, 5 mM sodium phosphate pH 7.0). Spectra were acquired in triplicate at 20 °C with a scan rate of 50 nm/min. Thermal denaturation curves were acquired by monitoring the molar ellipticity at 222 nm over a temperature range of 20 °C to 80 °C with a temperature ramp rate of 2 °C/min.

2.3.6. NMR spectroscopy

NMR data were acquired on a Bruker Avance 800 MHz spectrometer equipped with a cryogenic triple-resonance probe, processed by NMRPipe ([20]), and analyzed using NMRView ([21]). Two-dimensional ^{1}H-^{15}N TROSY spectra were acquired using 0.2 mM ^{15}N-labeled IpaB$^{74-224}$ or ^{15}N-labeled IpaD (wild type or C322S mutant) in buffer (100 mM NaCl, 10 mM sodium phosphate pH 7.0, 10% D_2O) at 30 °C. PREs were measured by single time-point method ([22]). For PRE data acquisition, two 2D ^{1}H-^{15}N HSQC spectra were acquired on two NMR samples with identical NMR acquisition parameters. The first 2D ^{1}H-^{15}N TROSY spectrum was acquired using an ^{15}N-Leu labeled IpaD complexed with paramagnetic (or MTSL-spin labeled) IpaB$^{74-224}$. A second 2D ^{1}H-^{15}N TROSY spectrum was acquired using ^{15}N-Leu labeled IpaD complexed with diamagnetic IpaB$^{74-224}$. The protein concentrations, buffer conditions, acquisition and processing parameters were kept identical for the paramagnetic and diamagnetic NMR samples to ensure that the observed intensity changes were only due to the effect of the spin label on the tip protein. Typical PRE samples contained 0.4 mM ^{15}N-leucine IpaD mixed with 0.4 mM (paramagnetic or diamagnetic) IpaB$^{74-224}$ in PRE buffer (100 mM NaCl, 20 mM sodium phosphate pH 7.0, 10% D_2O) and the data were acquired at 32 °C. Published backbone amide assignments of IpaD

were used in the analysis ([23]). Residues perturbed during the binding studies were mapped onto the crystal structure of IpaD (PDB ID 2J0O) ([12]).

2.4. Results

2.4.1. *Expression and purification of protein constructs for NMR studies*

IpaD constructs were expressed and purified under native conditions. IpaD was well behaved in solution and has been previously characterized by NMR ([23]). The IpaB constructs, IpaB$^{76-308}$ and IpaB$^{74-224}$, were designed based on the proteolysis data ([24]), crystal structures ([15]), and the known chaperone binding sites ([25-27]). Purified IpaB$^{76-308}$ showed limited stability and solubility, whereas IpaB$^{74-224}$ showed better behavior in solution and was more amenable for NMR studies and thus used for the PRE studies. To test the significance of the extreme N-terminal residues of IpaB in the IpaD-IpaB protein-protein interaction ([16]), a longer IpaB construct, IpaB$^{9-226}$ was subcloned, expressed and purified under native conditions similar to the shorter IpaB constructs.

2.4.2. *NMR spectroscopy on the major translocon protein IpaB*

NMR spectroscopy was used to gain insights into the solution behavior of IpaB constructs. The 2D ^{1}H-^{15}N TROSY spectrum of IpaB$^{74-224}$ at 37 ^{o}C displayed overall good dispersion, but peaks were broad and poorly resolved **(Figure 2-2A)**. Increasing the acquisition temperature to 47 ^{o}C significantly improved the NMR spectrum of IpaB$^{74-224}$ **(Figure 2-2B)**. The peaks were distinct, sharp, and well resolved, however, only ~65-70% of the residues from the protein were represented in 2D ^{1}H-^{15}N TROSY

spectrum of IpaB[74-224], thus backbone assignment of IpaB[74-224] by 3D NMR would have been challenging and not carried out. Because the backbone resonances of the tip protein IpaD have already been assigned ([23]), and the translocon protein constructs gave less than ideal NMR spectra as shown above (**Figure 2-2**), we used ^{15}N-labeled IpaD with spin-labeled IpaB (or unlabeled IpaB) to characterize the tip-translocon protein-protein interaction by NMR.
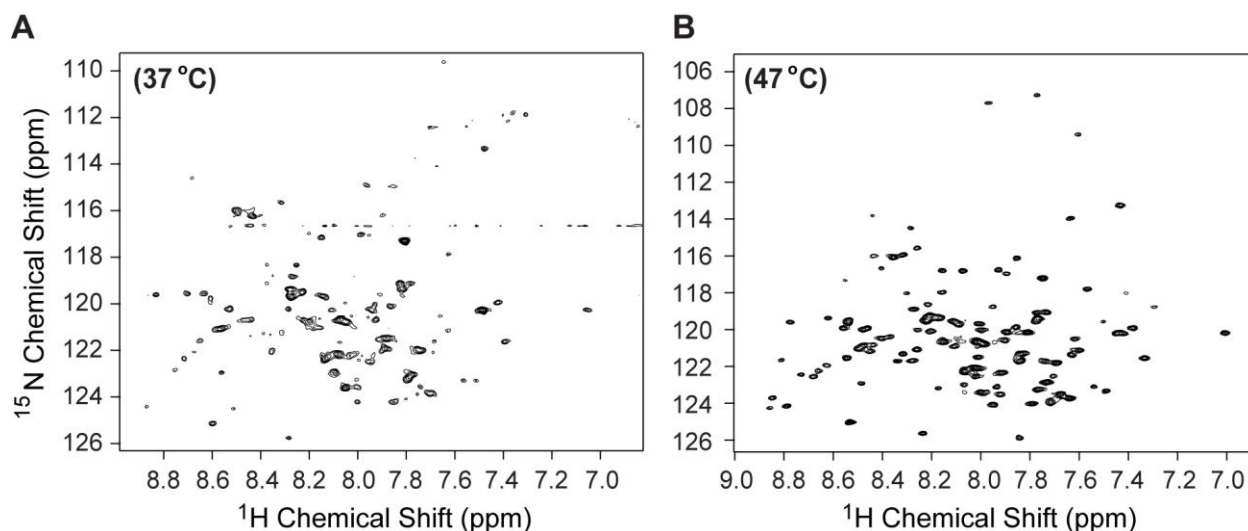
Figure 2-2. NMR spectroscopy on the N-terminal ectodomain of the major translocon protein IpaB.
2D [1]H-^{15}N TROSY spectra of **(A)** IpaB[74-224] at 37 °C, and **(B)** IpaB[74-224] at 47 °C.

2.4.3. Generation of IpaB[74-224] cysteine mutants for spin labeling

PRE measurements require attachment of paramagnetic spin labels at various positions. The site-directed spin labeling was achieved by engineering cysteine residues in specific locations on IpaB[74-224]. A total of eight cysteine point mutations in IpaB[74-224] were created for the attachment of the MTSL spin label **(Figure 2-1A)**. CD spectroscopy was used to assess the folding of the cysteine mutants and to ensure that the MTSL conjugation did not alter the overall secondary structure of the proteins. All the cysteine mutants before and after the attachment of MTSL spin label displayed CD spectra similar to the

wild type construct **(Figure 2-3)**. The extent of the spin labeling was confirmed by ElectroSpray Ionization Mass Spectrometry (ESI-MS) to be 100% with the expected mass increase of 185 Da upon MTSL conjugation **(Figure 2-4)**.
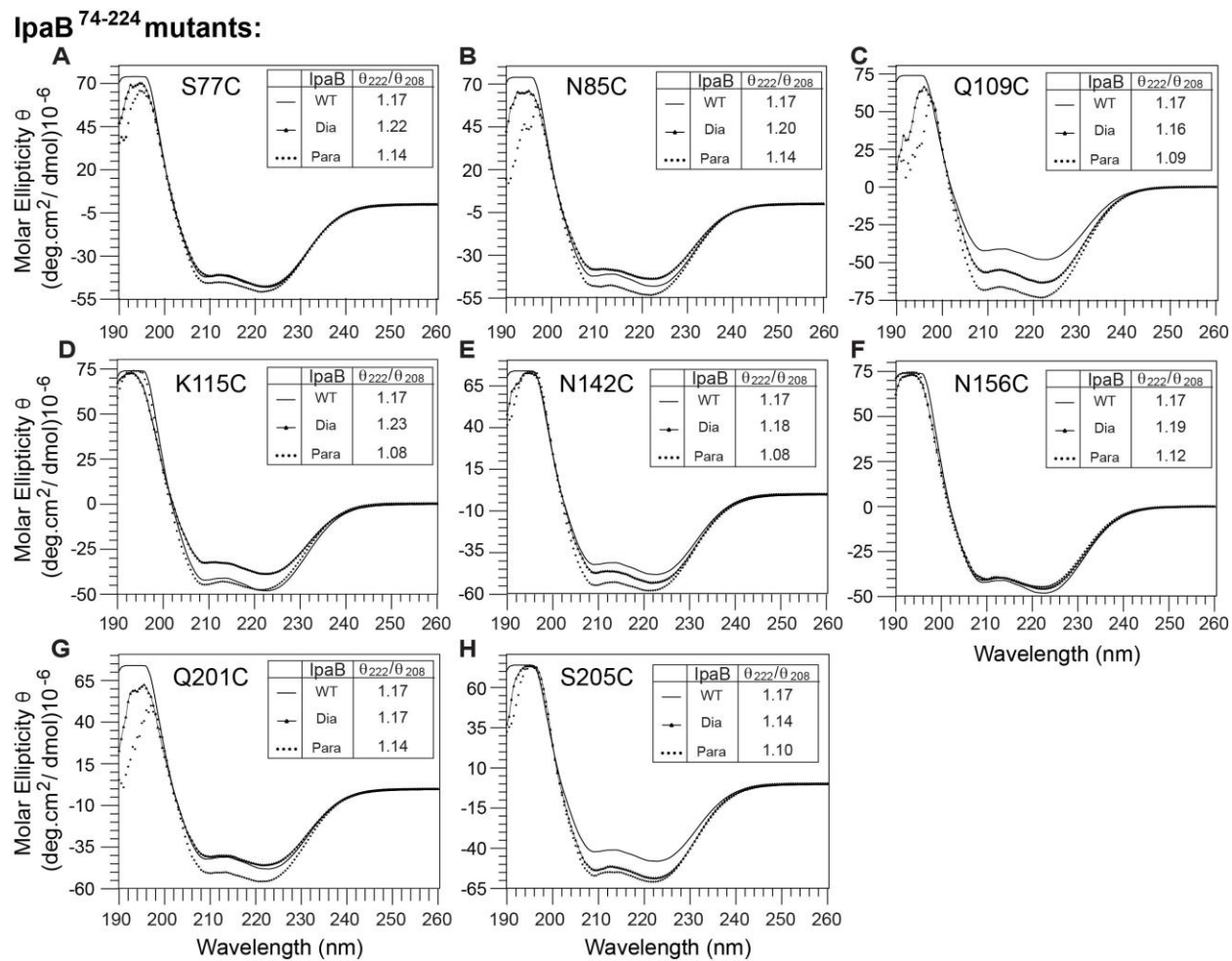
Figure 2-3. CD spectroscopy of IpaB$^{74-224}$ cysteine mutants used in the PRE studies. (A-H) The CD spectrum of wild type IpaB (WT) compared with cysteine point mutants before and after attachment of the MTSL spin label. (Legend: Dia, diamagnetic; Para, paramagnetic).
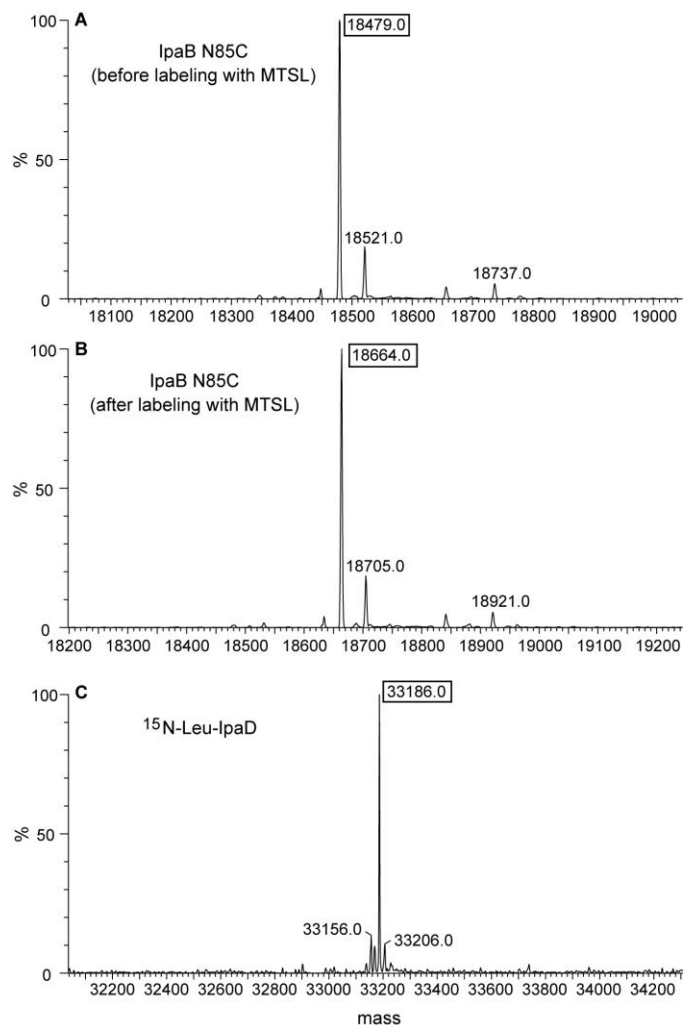
Figure 2-4. Representative electrospray ionization mass spectrometry (ESI-MS) on IpaB and IpaD proteins to confirm the extent of labeling. ESI-MS on IpaB N85C mutant **(A)** before MTSL-labeling (theoretical mass, 18479 Da), and **(B)** after MTSL labeling (theoretical mass, 18664). MTSL conjugation resulted in an expected increase in mass of 185 Da. **(C)** ESI-MS of ^{15}N-Leu labeled IpaD (theoretical mass, 33186).

2.4.4. PRE of ^{15}N-amino acid-specifically labeled IpaD with spin-labeled IpaB

IpaD (298 amino acids) is a large protein making the PRE analysis complicated. To reduce the complexity of the NMR datasets and enable the determination of subtle reductions in peak intensities upon protein-protein interaction, ^{15}N-leucine labeled IpaD was used in the PRE study. The leucine

residues of IpaD are distributed throughout the protein and provide coverage around all three domains: the N-terminal hairpin, the central coiled-coil, and the mixed α–β region. Corresponding to the 35 leucines present in the IpaD construct, selective ^{15}N-leucine labeling resulted in an increase of mass by 35 Da (33151 + 35 = 33186 Da, **Figure 2-4**). An optimal molar titration ratio of 1:1 for IpaD:IpaB was chosen for the PRE experiments on the basis of the extent of peak broadening observed at higher titration ratios (data not shown). PRE measurements were carried out by single time-point method ([22]). IpaD residues that lie close to the spin label should experience high PRE effect and show reduction in peak intensity in the paramagnetic datasets when compared to the control diamagnetic datasets ($I_{para}/I_{dia} < 1$). In contrast, IpaD residues residing away from the spin label should remain unaffected ($I_{para}/I_{dia} \sim 1$). Representative 2D NMR spectra used in the PRE determination for ^{15}N-Leu IpaD titrated with IpaB$^{74\text{-}224}$ are shown in **Figure 2-5**. The spin labels resulted in the reduction of the peak intensity for specific IpaD residues.
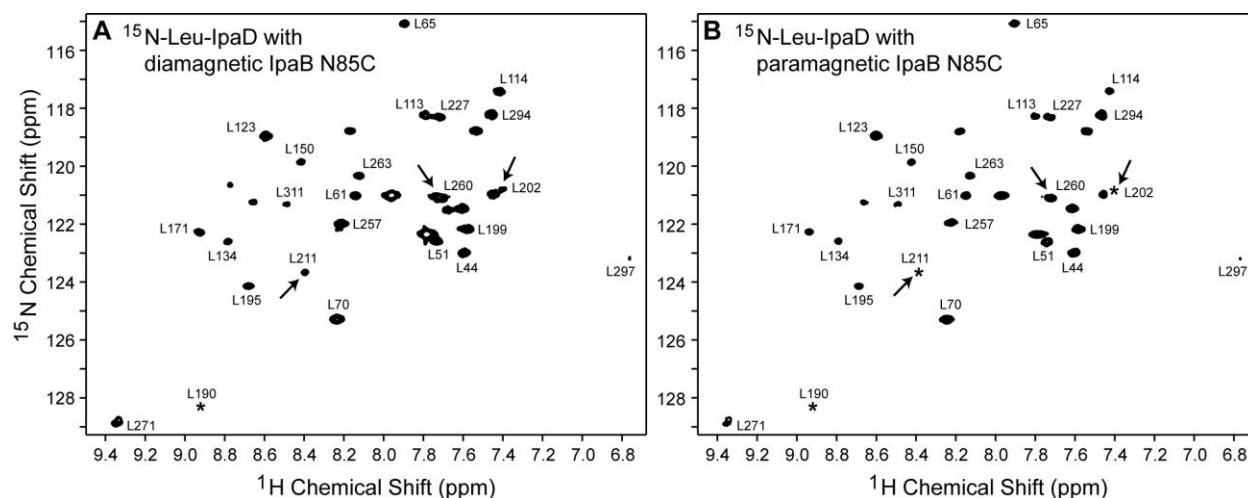
Figure 2-5. Representative 2D ^{1}H-^{15}N TROSY spectra of IpaD used in the PRE determination. ^{15}N-Leu IpaD complexed with (**A**) diamagnetic and (**B**) paramagnetic IpaB$^{74\text{-}224}$ N85C. Arrows point to the residues that showed reduction in peak intensity in the presence of spin label. Peaks marked with an asterisk were visible at a lower contour level.

2.4.5. PRE of Shigella IpaD-IpaB$^{74-224}$ interaction

Spin labels attached on IpaB$^{74-224}$ at positions Asn-156, Asn-142, Gln-201, Ser-205, and Ser-77 essentially resulted in weak to moderate PRE effect on IpaD **(Figure 2-6)**. With the exception of Leu-70 of IpaD, majority of the residues gave an I_{para}/I_{dia} ratio above 0.7 **(Figure 2-6A)**. Of these spin label positions, position Asn-156 displayed weakest PRE with a large number of peaks showing an I_{para}/I_{dia} ratio of ~1. Spin labels positioned at Asn-85, Gln-109, and Lys-115 (clustered towards the top N-terminal region of IpaB in the crystal structure, **Figure 2-6B**) produced strongest PRE effect on specific IpaD residues. Leu-70, Leu-202, Leu-211, and Leu-260 of IpaD experienced strongest PRE with I_{para}/I_{dia} values between 0.3-0.55 and Leu-199, Leu-227, Leu-257, and Leu-311 experienced moderate PRE with I_{para}/I_{dia} of ~0.6 **(Figure 2-6A)**. Spin labels close to the N-termini of IpaB$^{74-224}$ produced highest PRE effect and as the labels were moved down towards the end of the coiled-coil, PRE effect was minimized **(Figure 2-6)**. Irrespective of the spin label position, Leu-70 of IpaD gave highest PRE that could be due to its location on a highly flexible loop region. The affected residues mapped largely near the mixed $\alpha-\beta$ region of IpaD **(Figure 2-6C)**. A detailed analysis of PRE effects with individual spin-label sites has been shown in **Figure 2-7**. Our PRE results indicate that region around the mixed $\alpha-\beta$ domain of IpaD is the binding site for the N-terminal domain of IpaB.
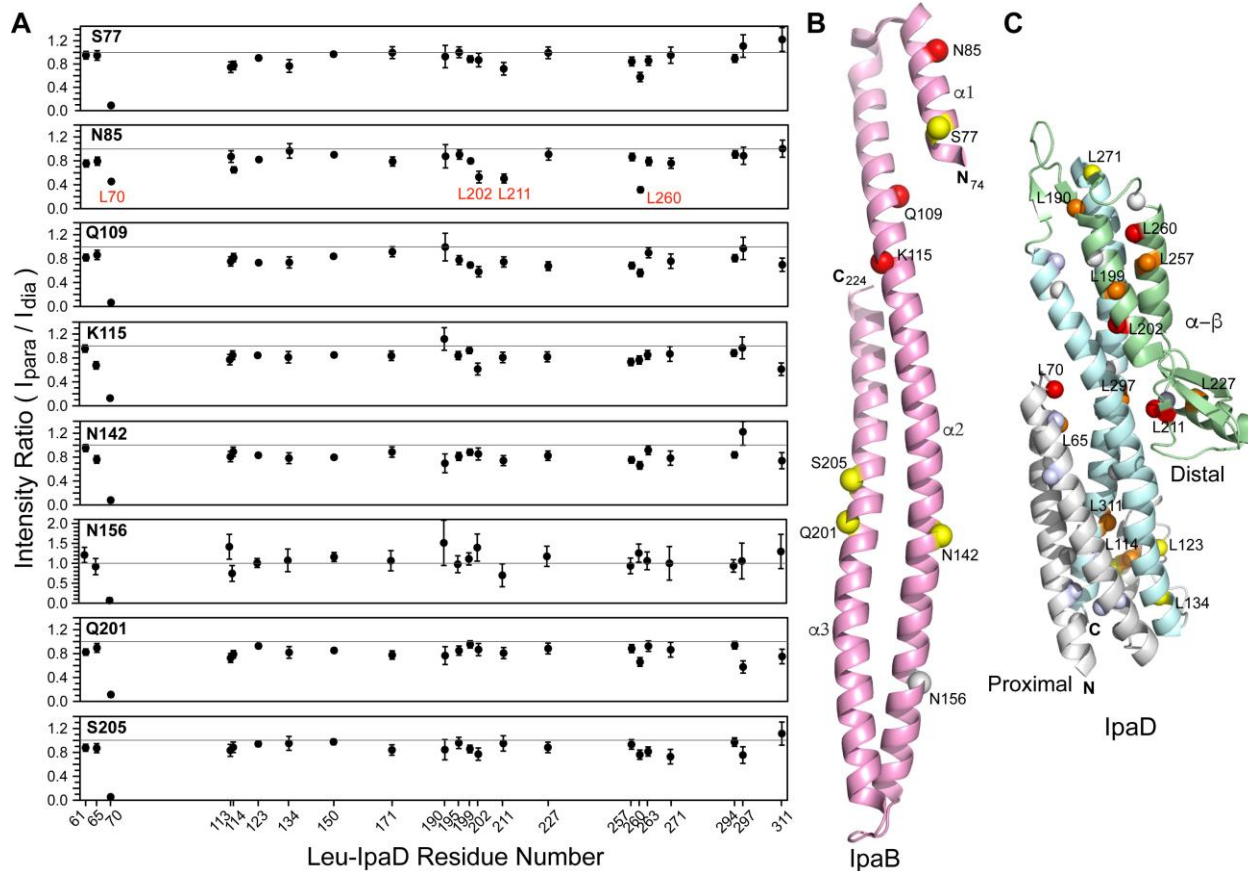
Figure 2-6. PRE results of ^{15}N-Leu IpaD and spin-labeled IpaB$^{74-224}$. **(A)** Each panel corresponds to one spin label position of IpaB$^{74-224}$. Circles in each panel represent I_{para}/I_{dia} intensity ratio for leucine residues of IpaD. The gray line at I_{para}/I_{dia} intensity ratio of 1.0 is taken as having no PRE effect. L70, L202, L211, and L260 of IpaD (labeled in red) experienced the strongest PRE effect. **(B)** The position of the spin labels (spheres) in IpaB$^{74-224}$ are colored based on the strength of the PRE as strong (red), moderate (yellow), and weak or none (gray). Spin label positions yielding strong PRE on IpaD (N85, Q109, K115) are clustered on the top N-terminal region of IpaB. **(C)** Results of PRE mapped onto the structure of IpaD. IpaD leucine residues that experienced the strongest PRE effect ($I_{para}/I_{dia} < 0.6$) are colored red, residues showing moderate PRE ($I_{para}/I_{dia} \sim 0.6\text{-}0.69$) are orange, weak ($I_{para}/I_{dia} \sim 0.7\text{-}0.79$) are yellow, and almost no effect ($I_{para}/I_{dia} > 0.8\text{-}1.0$) are white. Unassigned leucines are colored light blue. The affected leucine residues clustered near the distal region of IpaD.
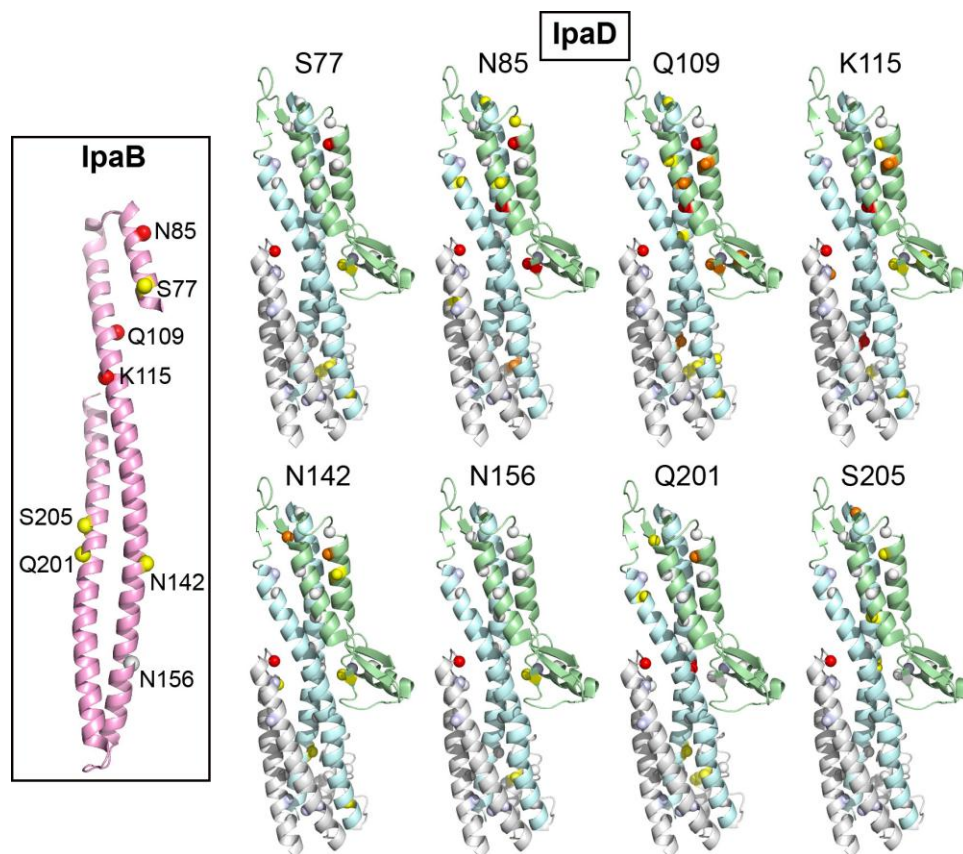
Figure 2-7. PRE effect produced by individual spin label positions in IpaD-IpaB interaction.
Results of the PRE experiments with spin-labeled IpaB[74-224] mapped onto the structure of IpaD. Color scheme is followed from Figure 2-6.

2.4.6. Role of extreme N-terminal residues of IpaB and bile salt DOC in IpaD-IpaB interactions

We tested a longer IpaB fragment encompassing extreme N-terminal residues (IpaB 9-226) that were previously shown to be important for the IpaD-IpaB binding ([16]). ^{15}N and ILV labeled IpaD[C322S] was titrated with unlabeled IpaB[9-226] and the titrations were monitored by acquiring 2D ^{1}H-^{15}N TROSY and 2D ^{1}H-^{13}C HSQC experiments. Stepwise addition of unlabeled IpaB[9-226] to IpaD[C322S] resulted in a concentration dependent decrease in the peak intensities of specific IpaD[C322S] backbone amide (**Figure 2-8A**) and side chain ILV methyl resonances (**Figure 2-9A**). Importantly, there were residues that did not undergo signal broadening; even at the highest IpaD[C322S]:IpaB[9-226] titration ratio indicating the complex

formation did not result in protein aggregation. The intensity ratio ($I_{1:1}/I_{1:0}$ in amide titrations; and 1-$I_{1:1}/I_{1:0}$ in ILV titrations) for each non-overlapped IpaD[C322S] residue at an IpaD[C322S]:IpaB[9-226] ratio of 1:1 was plotted to identify the tip residues that were perturbed the most upon addition of IpaB[9-226] (**Figure 2-8B and 2-9B**) and mapped onto the crystal structure of IpaD (**Figure 2-8C and 2-9C**). In both backbone amide and side chain ILV methyl titrations, the affected residues clustered in the distal region of IpaD, similar to the PRE results obtained using a shorter IpaB[74-224] fragment.
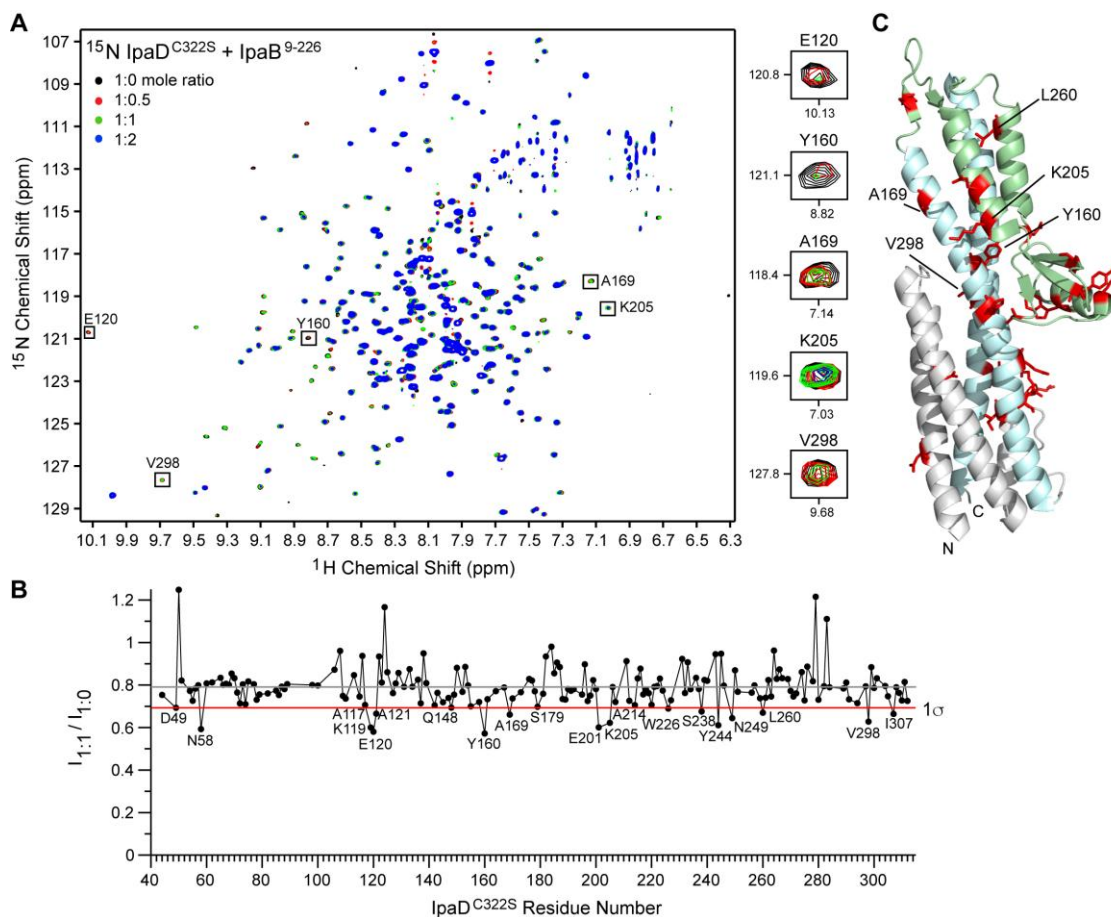
Figure 2-8. Presence of extreme N-terminal residues of IpaB did not alter the surfaces of IpaD involved in the binding of IpaB as detected by amide titrations. (**A**) Overlay of four 2D ^{1}H-^{15}N TROSY spectra of ^{15}N-labeled IpaD[C322S] with increasing concentrations of unlabeled IpaB[9-226]. (**B**) Plot of peak intensity ratio of complex ($I_{1:1}$) *vs* free form ($I_{1:0}$) for IpaD[C322S]:IpaB[9-226] complex with the average peak intensity ratio shown as gray line and one standard deviation from the average, 1σ, shown as red line. (**C**) Residues with peak intensity ratio ($I_{1:1}/I_{1:0}$) lower than 1σ mapped onto the crystal structure of IpaD and shown in red.
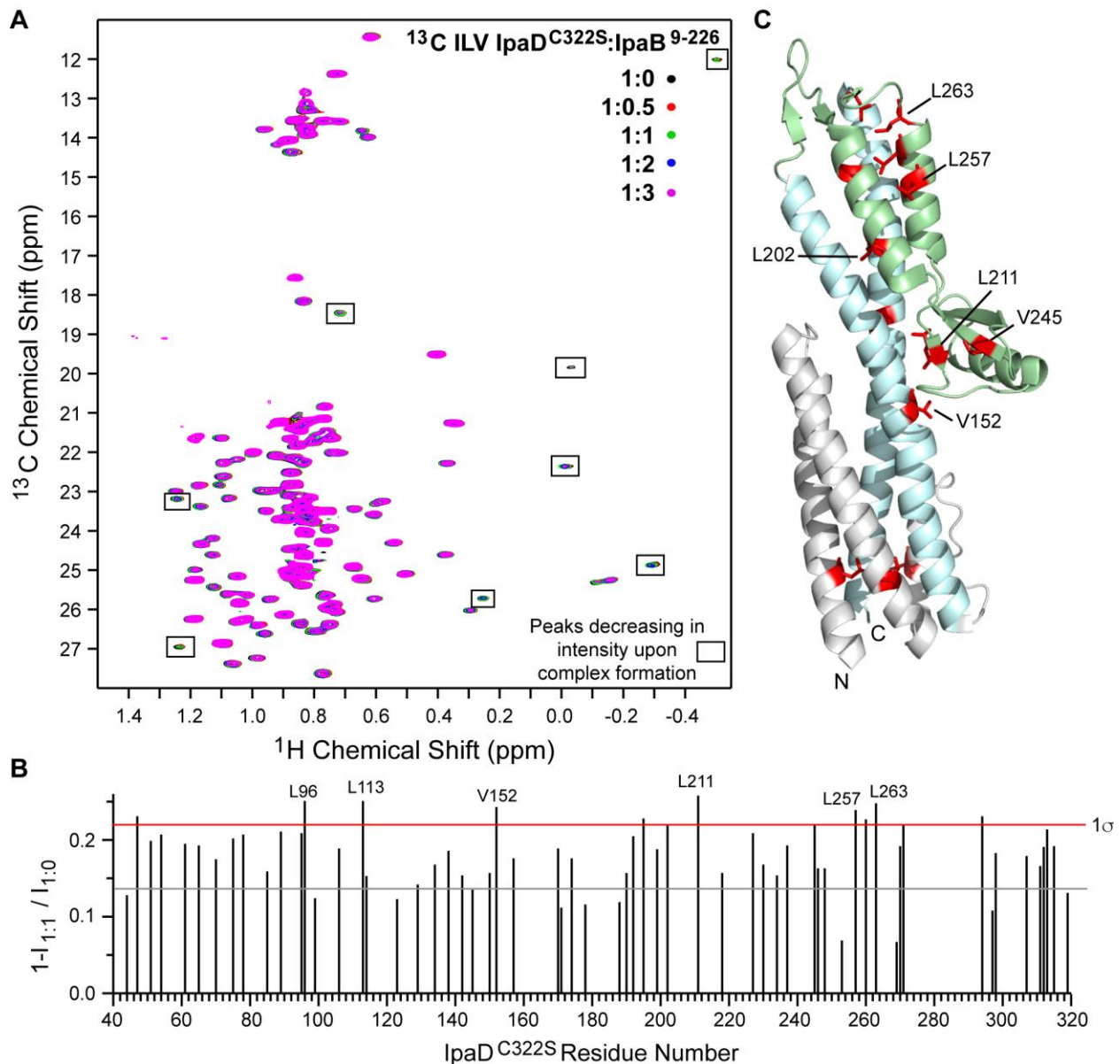
Figure 2-9. Extreme N-terminal residues of IpaB did not alter IpaD-IpaB interaction surfaces as detected by ILV methyl titrations. (**A**) Overlay of five 2D ^{1}H-^{13}C HSQC spectra of ^{13}C-ILV-labeled IpaD^{C322S} with increasing concentrations of unlabeled IpaB$^{9-226}$. (**B**) Peak intensity ratio of complex ($I_{1:1}$) *vs* free form ($I_{1:0}$) for IpaD^{C322S}:IpaB$^{9-226}$ complex plotted as a difference from peak intensity ratio of 1. Average peak intensity ratio is shown with a gray line and one standard deviation from the average, 1σ, shown as red line. (**C**) Residues with peak intensity ratio ($1-I_{1:1}/I_{1:0}$) greater than 1σ mapped onto the crystal structure of IpaD and shown in red.

Furthermore, bile salt DOC was reported to be essential for the interaction between IpaD and N-terminal ectodomain IpaB ([16]). To test this, we titrated ^{15}N and ILV labeled IpaD^{C322S} with unlabeled IpaB$^{9-226}$ in the presence of 0.4 mM DOC. Similar to the titration in the absence of DOC (**Figure 2-8 and**

2-9), we observed a concentration dependent decrease in the peak intensity of IpaD resonances in the presence of DOC **(Figure 2-10A and 2-11A)**. Analysis of the peak intensity ratio in the presence of DOC at a 1:1:2 molar ratio **(Figure 2-10B and 2-11B)** showed that the affected IpaD residues were primarily clustered at the distal region **(Figure 2-10C and 2-11C)**. Together, this suggests that DOC did not largely alter the affected IpaD and IpaB interaction surface or modulate the protein affinity as observed by NMR.

2.5. Discussion

An important step in type III secretion is the assembly of the translocon on the tip complex. The protein-protein interactions between the tip and the translocon proteins are crucial for the attachment of the bacterium to the host cell membrane and the delivery of effector proteins through the translocon pore ([28-30]). However, the atomic details of the protein-protein interaction governing the assembly of the tip-translocon structure remain poorly understood. The current hypothesis in the literature is that IpaD and IpaB form a heteropentameric complex consisting of 4 copies of IpaD and 1 copy of IpaB at the needle tip ([31-33]), and this IpaD:IpaB interaction is essential for host cell sensing, assembly of the translocon pore, and regulation of effector secretion into the host cell cytoplasm ([12],[30],[34]).
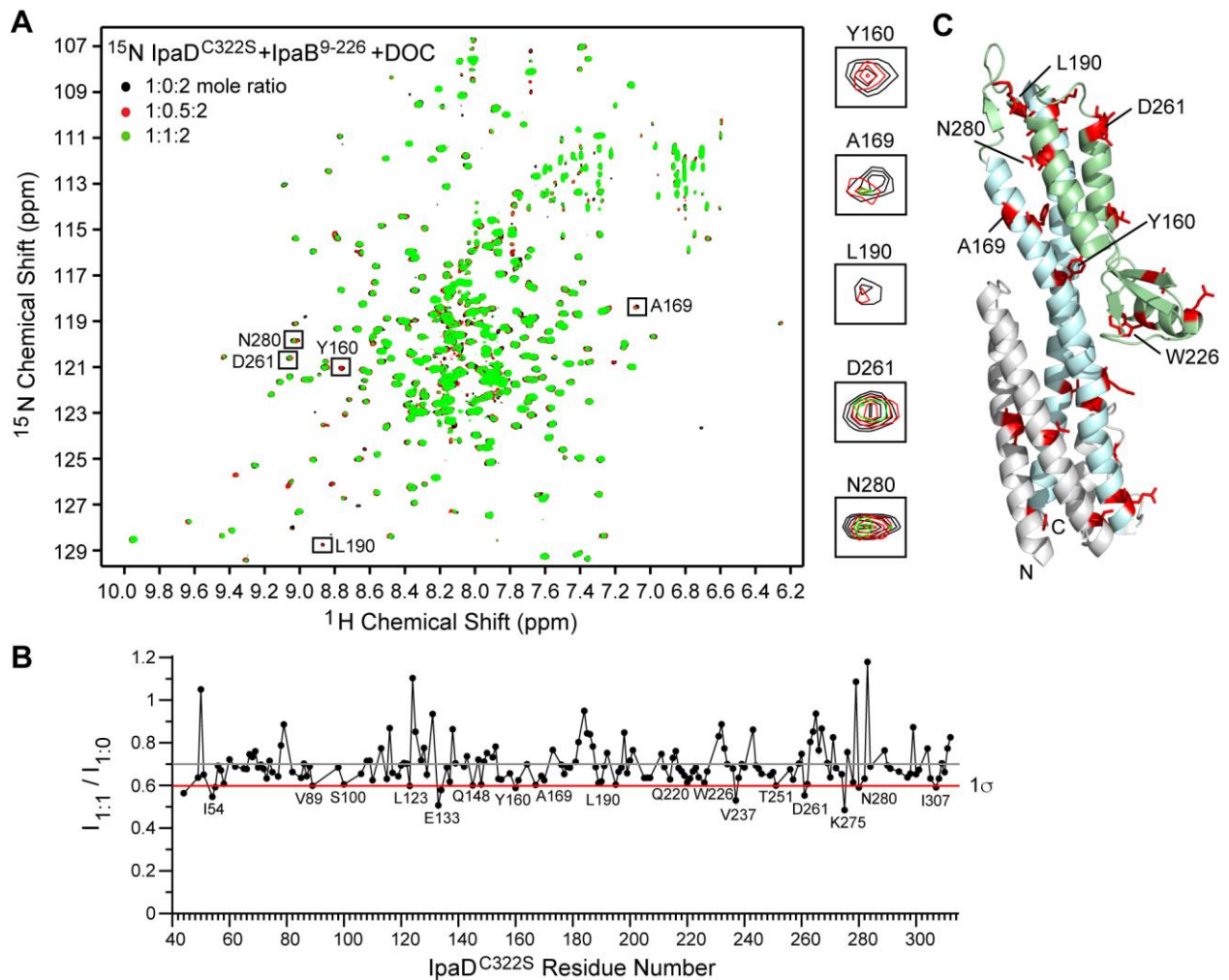
Figure 2-10. Significance of the bile salt DOC in IpaD-IpaB interaction. (**A**) Overlay of three 2D ^{1}H-^{15}N TROSY spectra of ^{15}N-labeled IpaD^{C322S} with increasing concentrations of unlabeled IpaB$^{9-226}$ in the presence of 0.4 mM DOC. (**B**) Plot of peak intensity ratio of complex ($I_{1:1}$) *vs* free form ($I_{1:0}$) for IpaD^{C322S}:IpaB$^{9-226}$ complex in presence of 0.4 mM DOC. The average peak intensity ratio is shown as gray line and one standard deviation from the average, 1σ, as red line. (**C**) Residues with peak intensity ratio ($I_{1:1}/I_{1:0}$) lower than 1σ mapped onto the crystal structure of IpaD (colored red).
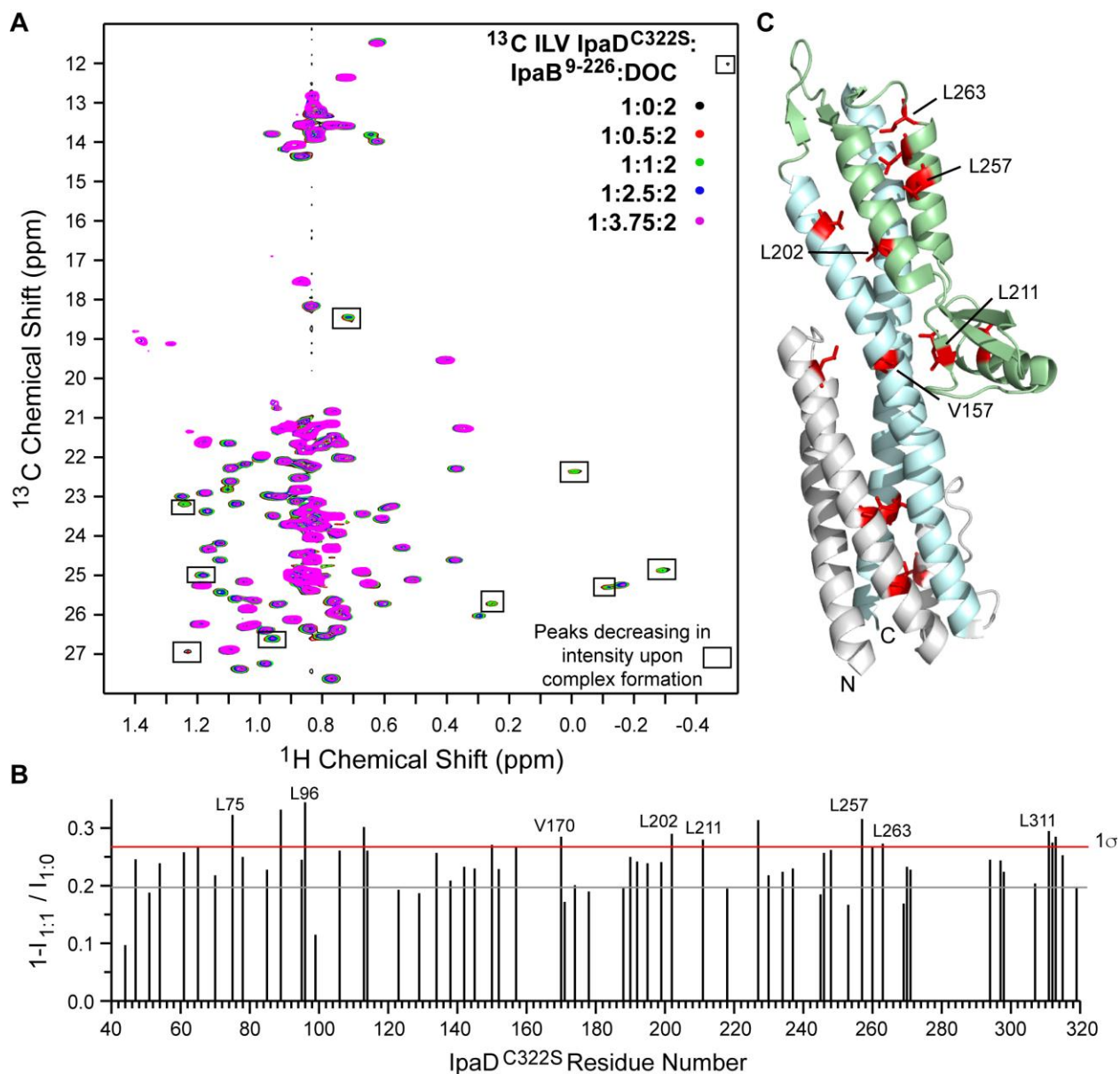
Figure 2-11. Bile salt DOC did not alter the IpaD-IpaB interaction surface. (**A**) Overlay of five 2D ^{1}H-^{13}C HSQC spectra of ^{13}C-ILV-labeled IpaD^{C322S} with increasing concentrations of unlabeled IpaB$^{9-226}$ in the presence of 0.4 mM DOC. (**B**) Peak intensity ratio of complex ($I_{1:1}$) *vs* free form ($I_{1:0}$) for IpaD^{C322S}:IpaB$^{9-226}$ complex in presence of 0.4 mM DOC plotted as a difference from peak intensity ratio of 1. The average peak intensity ratio is shown as gray line and one standard deviation from the average, 1σ, as red line. (**C**) Residues with peak intensity ratio ($1-I_{1:1}/I_{1:0}$) greater than 1σ mapped onto the crystal structure of IpaD (colored red).

There are conflicting results regarding the role of the bile salt DOC in triggering the presentation of IpaB at the tip complex. While several studies indicate exposure of IpaD to the bile salt deoxycholate

is needed for the recruitment of IpaB at the tip (16,34,35), others have shown the association of IpaD and IpaB at the needle tip without the need for DOC (32,33,36). Dickenson *et al*. (16) used fluorescence polarization and FRET assays to show that IpaD interacts with the N-terminal fragment of IpaB, and that the IpaD-IpaB interaction was detected only in the presence of high concentrations of DOC. Results of our NMR titrations and PRE studies presented here indicate that IpaB ectodomain interact with the cognate tip protein IpaD in the absence of deoxycholate and inclusion of extreme N-terminal residues of IpaB or DOC does not alter the IpaD-IpaB binding.

Results of fluorescence spectroscopy and NMR chemical shift mapping (**Figure 2-12**) show weak binding affinity at μM range between IpaD and IpaB (16). Such weak interactions probably provide these proteins sufficient conformational flexibility to perform their multi-functional roles in host environment sensing, translocon assembly, or for switching the system to an 'actively secreting' state. To gain insights into this weak interaction, we complemented the NMR chemical shift mapping with sensitive PRE methods, which requires paramagnetic spin labels that are conjugated to the protein through cysteine residues. Nuclear dipoles within ~15-20 Å of the spin label undergo increased relaxation, resulting in the reduction of the peak intensities of the residues lying close to the spin label. We engineered eight IpaB[74-224] cysteine mutants to attach the spin label and confirmed by CD spectroscopy that the cysteine mutation and the MTSL spin label did not alter the global fold of the proteins. Together our NMR and PRE results showed that the distal region of IpaD is the primary binding site for N-terminal ectodomain of IpaB.
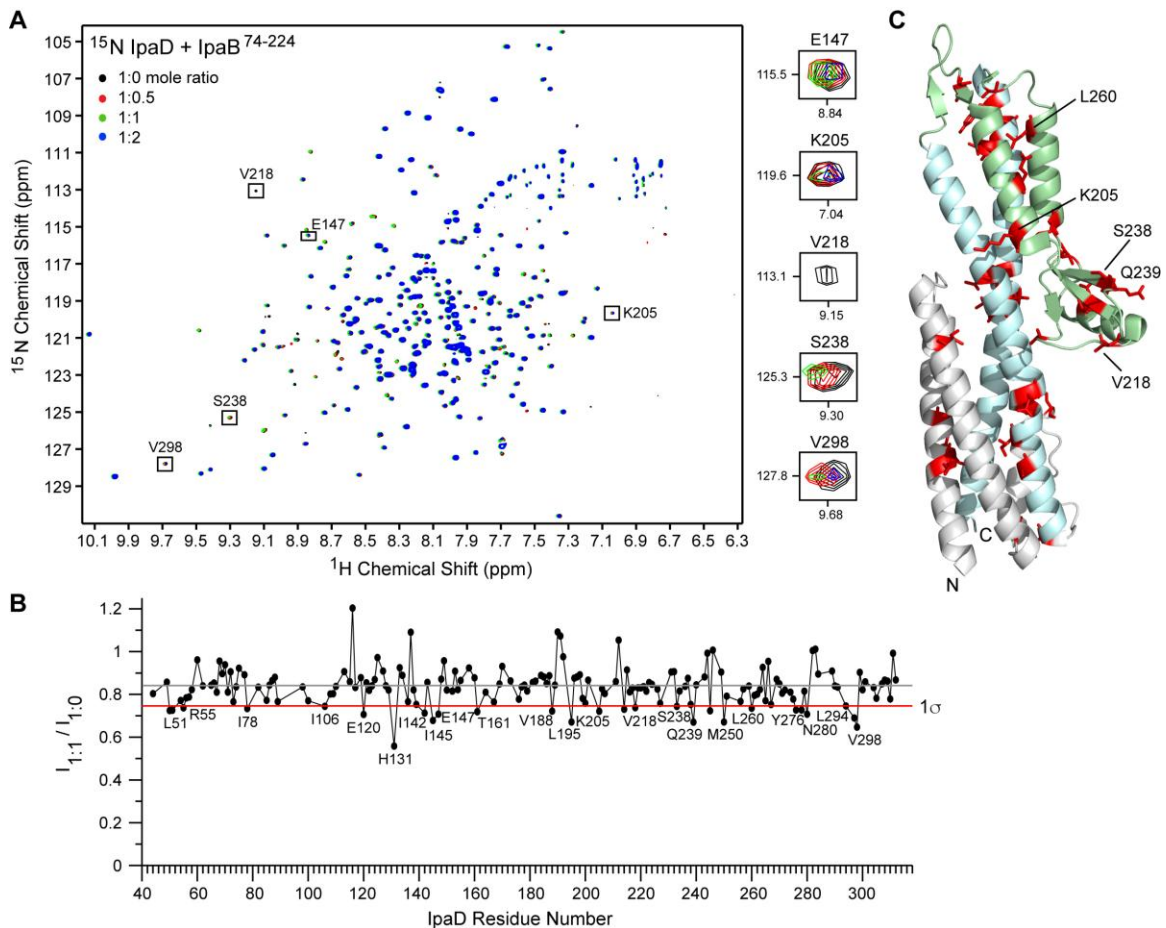
Figure 2-12. Characterization of IpaD-IpaB[74-224] interaction by NMR chemical shift mapping. (A) Overlay of four 2D ^{1}H-^{15}N TROSY spectra of ^{15}N-labeled IpaD with increasing concentrations of unlabeled IpaB[74-224]. **(B)** Plot of peak intensity ratio of complex ($I_{1:1}$) *vs* free form ($I_{1:0}$) for IpaD:IpaB[74-224] complex with the average peak intensity ratio shown as gray line and one standard deviation from the average, 1σ, shown as red line. **(C)** Residues with peak intensity ratio ($I_{1:1}/I_{1:0}$) lower than 1σ mapped onto the crystal structure of IpaD and shown in red.

Our results on the tip-translocon interaction are consistent with the previous results of Johnson *et al.* (12) showing that deletion of the mixed α−β domain of IpaD (residues 192 to 267) abolishes the localization of IpaB on the needle tip. The recent EM model of the *Shigella* tip-translocon complex, albeit at low resolution, shows that the mixed α−β domain of IpaD faces outside and lie close to the observed density for IpaB on the assembled tip (32). Additionally, the recent crystal structure of AopB (the homolog of IpaB in *Aeromonas*) indicates that the N- and C-terminal domains of AopB remain on the

extracellular portion upon membrane insertion, making these domains potential candidates for protein-protein interaction with the tip protein (37). Incorporating these with our previous results that the needle protein binds at the proximal end (or lower portion of the coiled-coil domain as depicted in Figure 2-1) of the tip protein (19,38), we propose a model where the N-terminal ectodomain of IpaB acts as an anchor to attach to the distal end of the tip protein IpaD at the mixed α–β domain, while the proximal end of the tip protein interacts with the needle **(Figure 2-13)**.
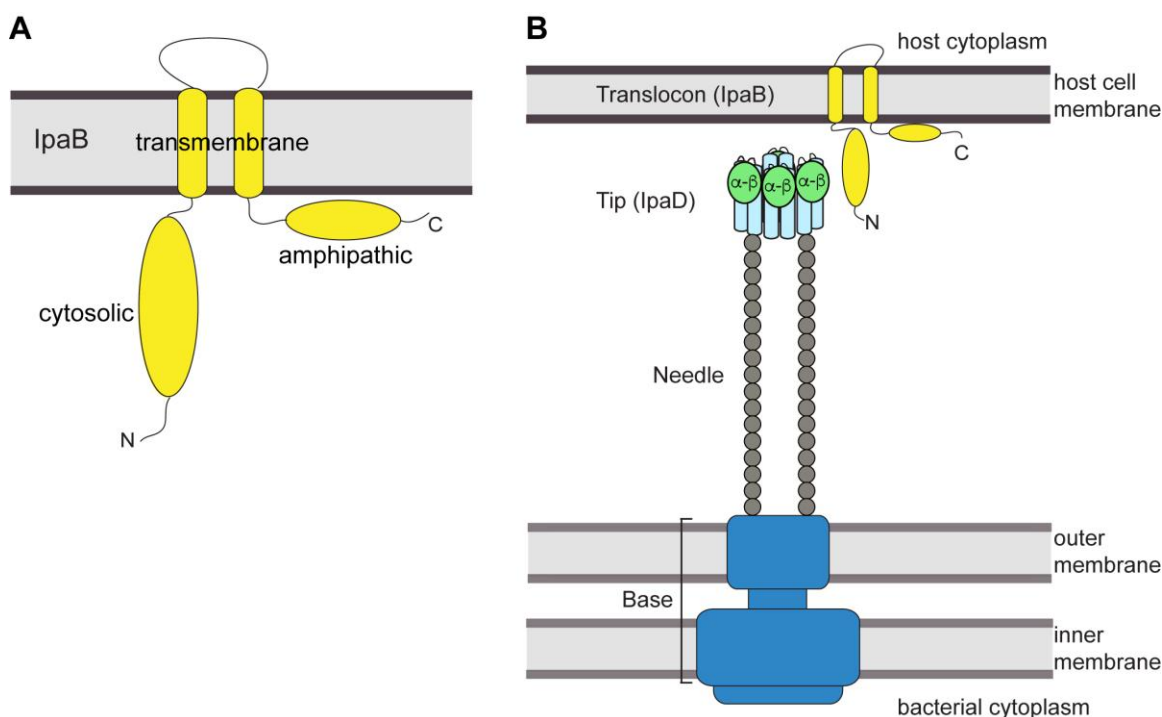
Figure 2-13. Model of the tip-translocon interaction. **(A)** Predicted membrane topology for the major translocon protein IpaB. **(B)** Proposed model for the interaction between the distal region of IpaD and the N-terminal ectodomain of IpaB in the assembly of the T3SS needle apparatus. The tip complex is represented based on the EM model of *Shigella* tip with the central coiled-coil domain shown in blue and the mixed α–β region in green.

In summary, our NMR binding studies confirmed a direct protein-protein interaction between the tip protein IpaD and the major translocon protein IpaB of *Shigella* in the absence of deoxycholate. Further, we have localized the interaction interface to the N-terminal cytosolic domain of IpaB and the

distal end of the tip protein IpaD. Similar studies performed on SipD-SipB, the IpaD-IpaB homologs in *Salmonella* also revealed that the distal region of SipD is involved in binding the N-terminal ectodomain of SipB. This suggests that the tip-translocon protein-protein interaction surfaces are conserved between *Shigella* and *Salmonella*. Our findings provide new insight into the interactions involved in the assembly of the T3SS needle apparatus.

2.6. References

1.	Coburn, B., Sekirov, I., and Finlay, B. B. (2007) Type III Secretion Systems and Disease. in *Clin. Microbiol. Rev.*
2.	Hueck, C. J. (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. in *Microbiology and Molecular Biology Reviews*
3.	Carayol, N., and Tran Van Nhieu, G. (2013) The Inside Story of Shigella Invasion of Intestinal Epithelial Cells. *Cold Spring Harb Perspect Med* **3**
4.	Chatterjee, S., Chaudhury, S., McShan, A. C., Kaur, K., and De Guzman, R. N. (2013) Structure and Biophysics of Type III Secretion in Bacteria. in *Biochemistry*
5.	Cornelis, G. R. (2006) The type III secretion injectisome. in *Nat Rev Micro*, Nature Publishing Group
6.	Galan, J. E. (2001) Salmonella interactions with host cells: type III secretion at work. in *Annu. Rev. Cell Dev. Biol.*
7.	Demers, J.-P., Habenstein, B., Loquet, A., Kumar Vasa, S., Giller, K., Becker, S., Baker, D., Lange, A., and Sgourakis, N. G. (2014) High-resolution structure of the Shigella type-III secretion needle by solid-state NMR and cryo-electron microscopy. in *Nat Commun*
8.	Loquet, A., Sgourakis, N. G., Gupta, R., Giller, K., Riedel, D., Goosmann, C., Griesinger, C., Kolbe, M., Baker, D., Becker, S., and Lange, A. (2012) Atomic model of the type III secretion system needle. in *Nature*
9.	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. *Infection and Immunity* **74**, 4391-4400
10.	Mueller, C. A., Broz, P., and Cornelis, G. R. (2008) The type III secretion system tip complex and translocon. in *Mol. Microbiol.*
11.	Barta, M. L., Guragain, M., Adam, P., Dickenson, N. E., Patil, M., Geisbrecht, B. V., Picking, W. L., and Picking, W. D. (2012) Identification of the bile salt binding site on ipad from Shigella flexneri and the influence of ligand binding on IpaD structure. *Proteins* **80**, 935-945
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) Self-chaperoning of the Type III Secretion System Needle Tip Proteins IpaD and BipD. in *J Biol Chem*
13.	Hume, P. J., McGhie, E. J., Hayward, R. D., and Koronakis, V. (2003) The purified Shigella IpaB and Salmonella SipB translocators share biochemical properties and membrane topology. in *Mol. Microbiol.*

14. McGhie, E. J. E., Hume, P. J. P., Hayward, R. D. R., Torres, J. J., and Koronakis, V. V. (2002) Topology of the Salmonella invasion protein SipB in a model bilayer. in *Mol. Microbiol.*

15. Barta, M. L., Dickenson, N. E., Patil, M., Keightley, A., Wyckoff, G. J., Picking, W. D., Picking, W. L., and Geisbrecht, B. V. (2012) The structures of coiled-coil domains from type III secretion system translocators reveal homology to pore-forming toxins. in *J Mol Biol*

16. Dickenson, N. E., Arizmendi, O., Patil, M. K., Toth, I., Ronald T, Middaugh, C. R., Picking, W. D., and Picking, W. L. (2013) N-Terminus of IpaB Provides a Potential Anchor to the ShigellaType III Secretion System Tip Complex Protein IpaD. in *Biochemistry*

17. 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

18. 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

19. Rathinavelan, T., Tang, C., and De Guzman, R. N. (2011) Characterization of the interaction between the Salmonella type III secretion system tip protein SipD and the needle protein PrgI by paramagnetic relaxation enhancement. in *J Biol Chem*

20. 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. in *J Biomol NMR*

21. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. in *Methods Mol. Biol.*

22. Gillespie, J. R., and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J Mol Biol* **268**, 158-169

23. Dickenson, N. E., Zhang, L., Epler, C. R., Adam, P. R., Picking, W. L., and Picking, W. D. (2011) Conformational Changes in IpaD from Shigella flexneri upon Binding Bile Salts Provide Insight into the Second Step of Type III Secretion. *Biochemistry* **50**, 172-180

24. Hayward, R. D., McGhie, E. J., and Koronakis, V. (2000) Membrane fusion activity of purified SipB, a Salmonella surface protein essential for mammalian cell invasion. *Mol. Microbiol.* **37**, 727-739

25. Adam, P. R., Patil, M. K., Dickenson, N. E., Choudhari, S., Barta, M., Geisbrecht, B. V., Picking, W. L., and Picking, W. D. (2012) Binding affects the tertiary and quaternary structures of the Shigella translocator protein IpaB and its chaperone IpgC. *Biochemistry* **51**, 4062-4071

26. Kim, B. H., Kim, H. G., Kim, J. S., Jang, J. I., and Park, Y. K. (2007) Analysis of functional domains present in the N-terminus of the SipB protein. *Microbiology (Reading, Engl.)* **153**, 2998-3008

27. Lokareddy, R. K., Lunelli, M., Eilers, B., Wolter, V., and Kolbe, M. (2010) Combination of Two Separate Binding Domains Defines Stoichiometry between Type III Secretion System Chaperone IpgC and Translocator Protein IpaB. in *Journal of Biological Chemistry*

28. Collazo, C. M., and Galan, J. E. (1997) The invasion-associated type III system of Salmonella typhimurium directs the translocation of Sip proteins into the host cell. *Mol. Microbiol.* **24**, 747-756

29. Lara-Tejero, M., and Galán, J. E. (2009) Salmonella enterica serovar typhimurium pathogenicity island 1-encoded type III secretion system translocases mediate intimate attachment to nonphagocytic cells. *Infection and Immunity* **77**, 2635-2642

30. Ménard, R., Sansonetti, P. J., and Parsot, C. (1993) Nonpolar mutagenesis of the ipa genes defines IpaB, IpaC, and IpaD as effectors of Shigella flexneri entry into epithelial cells. *J. Bacteriol.*

31. 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? in *Proceedings of the National Academy of Sciences*

32. Cheung, M., Shen, D.-K., Makino, F., Kato, T., Roehrich, A. D., Martinez-Argudo, I., Walker, M. L., Murillo, I., Liu, X., Pain, M., Brown, J., Frazer, G., Mantell, J., Mina, P., Todd, T., Sessions, R. B., Namba, K., and Blocker, A. J. (2015) Three-dimensional electron microscopy reconstruction and cysteine-mediated crosslinking provide a model of the type III secretion system needle tip complex. in *Mol. Microbiol.*

33. Veenendaal, A. K. J., Hodgkinson, J. L., Schwarzer, L., Stabat, D., Zenk, S. F., and Blocker, A. J. (2007) The type III secretion system needle tip complex mediates host cell sensing and translocon insertion. *Mol. Microbiol.* **63**, 1719-1730

34. 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. *Infection and Immunity* **75**, 2626-2629

35. 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

36. Shen, D. K., Saurya, S., Wagner, C., Nishioka, H., and Blocker, A. J. (2010) Domains of the Shigella flexneri Type III Secretion System IpaB Protein Involved in Secretion Regulation. *Infection and Immunity* **78**, 4999-5010

37. Nguyen, V. S., Jobichen, C., Tan, K. W., Tan, Y. W., Chan, S. L., Ramesh, K., Yuan, Y., Hong, Y., Seetharaman, J., Leung, K. Y., Sivaraman, J., and Mok, Y. K. (2015) Structure of AcrH–AopB Chaperone-Translocator Complex Reveals a Role for Membrane Hairpins in Type III Secretion System Translocon Assembly. *Structure*

38. Rathinavelan, T., Lara-Tejero, M., and Lefebre, M. (2014) NMR Model of PrgI–SipD Interaction and Its Implications in the Needle-Tip Assembly of the Salmonella Type III Secretion System. in *Journal of molecular biology*

Chapter 3: Identification and Characterization of the Binding of Small Molecules to LcrV and PcrV, the Protective Antigens of the Type III Secretion System

Kawaljit Kaur[1], Asokan Anbanandam[2], Mason Wilkinson[1], and Roberto N. De Guzman[1]

[1]Department of Molecular Biosciences and [2]Bio-NMR Core Facility, University of Kansas,

Lawrence, KS 66045, USA

3.1. Abstract

The rapid and alarming rise of antibiotic resistance in pathogens requires new targets for developing novel anti-infectives. The type III secretion system (T3SS) is essential in the pathogenesis of many medically important pathogens and represents an attractive target for the development of new antibiotics. *Yersinia* and *Pseudomonas*, the causative agents of plague and nosocomial infections, along with many other Gram-negative bacteria rely on the T3SS to inject virulence proteins directly into eukaryotic cells to initiate infection. An important step in rational design of anti-virulence drugs targeting T3SS is the identification of small molecules that can bind to T3SS proteins. LcrV and PcrV, the tip proteins from the *Yersinia* and *Pseudomonas* T3SS are critical virulence proteins and known protective antigens. In this study, through a surface plasmon resonance screen of 288 small and diverse fragments, we identified the binding of 5-hydroxyindole and 4-(4-methylpiperazino)aniline to LcrV, 3-amino-7-bromoquinoline to PcrV, and 3-(1H-pyrrol-1-yl)benzo[b]thiophene-2-carboxylic acid to both LcrV and PcrV. Further, NMR methods were utilized to characterize and identify the surface on tip proteins involved in the small molecule interaction. Our results demonstrate that LcrV and PcrV are promising druggable targets and the identified molecular fragments can potentially be used for the development of novel T3SS inhibitors.

3.2. Introduction

The type III secretion system (T3SS) is indispensable for virulence of a wide spectrum of clinically relevant bacterial pathogens, including *Yersinia pestis*, the etiologic agent of plague, *Pseudomonas aeruginosa*, the causative agent of secondary hospital infections, and other Gram-negative bacteria, such as, *Salmonella*, *Shigella*, and *Chlamydia* spp (1). With ever-increasing incidence of

antibiotic resistance in these pathogens, there is a pressing need for the development of new anti-bacterials (2,3). Targeting the virulence factors of pathogenic bacteria, rather than their survival, provides a promising alternative to the use of traditional bactericidal antibiotics. Since the T3SS is a major virulence determinant of many Gram-negative bacteria and exposed on the bacterial surface, disrupting the T3SS is an attractive target for developing novel anti-virulence drugs (2-5).

Bacterial pathogens utilize the T3SS to inject virulence effector proteins into their target host cells to initiate infectious diseases (6). The T3SS is composed of structural and nonstructural proteins (7). The structural component of the T3SS, the needle apparatus, is assembled from more than a dozen different proteins and consists of a base that spans the bacterial membranes, an external needle, and a tip complex. Upon host cell contact, a translocon is assembled between the needle tip complex and the host cell, which creates a pore in the host cell membrane and provides a gateway for translocation of effector proteins (8). The multi-functional tip proteins forming the needle tip complex are crucial for the bacterial virulence and responsible for host cell sensing, assembly of the translocon pore, and regulation of the secretion of effector proteins into the host cells (7,9). LcrV and PcrV (together referred as V-proteins) are the tip proteins belonging to the 'Ysc' family of the *Yersinia* and *Pseudomonas* T3SS, respectively (9,10). The crystal structure of LcrV is available and displays an overall dumbbell shape with an N-terminal globular domain, an elongated central coiled-coil, and a less-structured distal region (11,12). Although the structure of PcrV is currently not known, the I-TASSER model (13) show structural similarity to the homolog LcrV.

Both LcrV and PcrV are potent protective antigens and active or passive immunization against these V-proteins provides a high level of protection against T3SS-mediated lethal *Y. pestis* and *P. aeruginosa* infections (3,9,14). Indeed, vaccines based on these proteins are currently under various stages of development (3,9). Furthermore, for the past few years, tremendous efforts are being made to discover small molecule compounds that inhibit the virulence of *Yersinia* and *Pseudomonas* (15-20).

Although these compounds inhibit the T3SS, the mechanism of action and specific molecular targets within the T3SS for a vast majority remain unknown (4,5). Here, we show results of a surface plasmon resonance (SPR)-based fragment screening of 288 compounds that identified novel small molecule fragments that directly bind to the virulent V-proteins **(Figure 3-1 and Figure 3-2)**. Binding of the hit fragments to their cognate protein was further validated and characterized using both ligand and protein-observed NMR methods. This is the first report of small molecule fragments that are shown to bind to LcrV and PcrV and could be used as potential scaffolds for future drug design.

3.3. Methods

3.3.1. *Expression and purification of LcrV and PcrV*

The cloning, expression and purification of LcrV (residues 28-322, C273S) and PcrV (residues 25-294) have been previously described (11,21). Both constructs contained an N-terminal His$_6$-GB1 solubility tag followed by a tobacco etch virus (TEV) protease cleavage site. The expression plasmids were freshly transformed into *E. coli* BL21 (DE3) cells and grown in culture media containing 100 μg/mL carbenicillin. Unlabeled proteins for SPR screening were obtained by cell growth in 1 L of lysogeny broth (LB). For NMR studies, in addition to the traditional ^{15}N-labeling, proteins were simultaneously ^{13}C-labeled at the side chain ILV methyl groups (^{13}C labeling at Cδ1 methyl group of Isoleucine, geminal Cδ of Leucine, and Cγ of Valine). ILV and ^{15}N-labeled LcrV (or PcrV) were obtained by cell growth in 1 L of M9 minimal media supplemented with 1 g of ^{15}N-ammonium chloride and 3 g of glucose. Cells were grown at 37°C until OD$_{600}$ ~0.4 at which the growth medium was supplied with 60 mg of 2-ketobutyric acid-4-^{13}C (Sigma #571342) to ^{13}C label the Cδ1 isoleucine and 100 mg of 2-keto-3-(methyl-^{13}C)-butyric acid-4-^{13}C (Sigma #571334) to label Cδ leucine and Cγ valine methyl groups. For backbone

amide and side chain ILV methyl assignments, perdeuterated LcrV was prepared by cell growth in M9 minimal media in 1 L of D_2O. For the backbone amide assignments, uniform $^{15}N/^{13}C$-labeled perdeuterated LcrV was obtained by supplementing media with 1 g of ^{15}N-ammonium chloride and 2 g of either non-deuterated ^{13}C-D-glucose (U-$^{13}C_6$) or fully deuterated ^{13}C-D-glucose (U-$^{13}C_6$; 1,2,3,4,5,6,6-d_7). For the assignment of ILV methyl resonances, ^{15}N/ILV-labeled perdeuterated LcrV was prepared by supplying 1 g of ^{15}N-ammonium chloride, 2 g of deuterated D-glucose (1,2,3,4,5,6,6-d_7, Cambridge Isotope Laboratories, CIL, #DLM-2062), and appropriate ^{13}C ILV keto acids (as described above). At an OD_{600} of ~0.7, the culture was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Cell growth was continued overnight in a 15°C shaker to a final OD_{600} of ~2.8 (~2.0 for D_2O-based preparations).

Cells were harvested by centrifugation (4000 rpm, 10 min), resuspended in binding buffer (30 ml; 500 mM NaCl, 20 mM Tris-HCl pH 8.0, 5 mM imidazole), and lysed by sonication in the presence of 0.1 mM phenylmethanesulfonylfluoride (PMSF). Cellular debris was removed by centrifugation (13,000 rpm, 10 min), and 600 μl of 5% (v/v) polyethyleneimine (PEI) was added to the supernatant to precipitate the nucleic acids. After centrifugation (13,000 rpm, 10 min), the supernatant was loaded onto a Ni^{2+}-affinity column. The column was washed with binding buffer and the proteins were eluted with elution buffer (500 mM NaCl, 20 mM Tris-HCl pH 8.0, 250 mM imidazole). To remove the tag, fractions containing His_6-GB1-LcrV (or His_6-GB1-PcrV) were pooled and incubated overnight with 250 μL of 0.04 mM recombinant TEV protease at room temperature in buffer (20 mM NaCl, 20 mM Tris-HCl pH 8.0, 1 mM DTT, 0.5 mM EDTA). The digested protein was dialyzed in binding buffer and passed through a second round of Ni^{2+}-affinity purification. LcrV (or PcrV) eluted in the flow through and wash fractions, while the His_6-GB1 tag was retained in the column. Both LcrV and PcrV retained a 3-residue N-terminal cloning artifact 'GHM'. The purified proteins were either dialyzed in 1.05X phosphate-buffered saline (PBS, pH 7.4 for SPR screening) or NMR buffer (20 mM NaCl, 10 mM $NaPO_4$ pH 6.8). The proteins

were concentrated using Amicon Ultra 10K centrifugal filters (Millipore) and concentrations were measured by A_{280}.

3.3.2. Surface plasmon resonance (SPR) screening

A library of 288 compounds (Zenobia fragment library 2, San Diego, CA, http://www.zenobiafragments.com) was obtained as dry films in three 96-well microplates. Compounds were dissolved in 100% DMSO to produce a stock concentration of 100 mM. The working solution was made by diluting compounds to a final concentration of 1 mM in 5% DMSO and 1.05X PBS (running buffer).

SPR data were acquired at 25°C using a BIACORE 3000 instrument (GE Healthcare). The screening was carried out as previously with minor modifications (22). Briefly, the target proteins LcrV and PcrV were immobilized onto a CM5 sensor chip using standard amine-coupling chemistry. Activation, coupling and blocking steps were performed for 7 minutes each resulting in final 7,906 response units (RU) for LcrV and 6,673 RU for PcrV. BIACORE 3000 allowed the automated injection of compounds from 96-well plates. Compounds were injected into the flow cells and allowed to associate for 30 sec and dissociate for 60 sec. Running buffer was injected between each compound run to monitor carryover effects.

3.3.3. NMR spectroscopy

Saturation transfer difference (STD) NMR data were acquired at 30°C using a Bruker Avance III 600 MHz spectrometer equipped with a TXI-RT probe and processed using Topspin. Samples were prepared by mixing protein and compound at a molar ratio of 1:100 (40 μM protein + 4000 μM

54

compound). One-dimensional ^{1}H STD spectra were acquired by saturating the protein using 50 ms shaped saturation pulse over a period of 2 sec with center of the on-resonance pulse varied from -0.9 to -1.1 ppm for optimal STD signal, while the off-resonance pulse was maintained at 40 ppm. 1D ^{1}H NMR of the structures was predicted using ChemDraw.

Two-dimensional ^{1}H-^{15}N TROSY, ^{1}H-^{13}C ILV methyl HSQC, and three-dimensional HNCA, HNCACB, CBCA(CO)NH, and HMQC-NOESY-HMQC data on LcrV were acquired using a Bruker Avance 800 MHz spectrometer equipped with a TCI cryoprobe. Two-dimensional ^{1}H-^{15}N TROSY and ^{1}H-^{13}C ILV methyl HSQC spectra on PcrV were collected using Bruker Avance III 600 MHz spectrometer. The data were acquired at 30°C, processed using NMRPipe ([23]) and analyzed using NMRView ([24]). For backbone assignments, sample containing ~0.7 mM perdeuterated ^{15}N/^{13}C-labeled LcrV in 500 μL of NMR buffer with 10% D_2O was prepared. The ILV methyl resonances of LcrV were assigned following the published protocol ([25],[26]). A 500 μL sample containing ~0.35 mM perdeuterated ^{15}N/ILV-labeled LcrV in NMR buffer was lyophilized and re-suspended in 100% D_2O. Three-dimensional data were collected using a mixing time of 300 ms and a relaxation delay of 2 sec. To assist in the ILV assignment, twenty-four isoleucine to leucine point mutants in LcrV, and four isoleucine to leucine and leucine to alanine mutants in PcrV were made using site-directed mutagenesis (Stratagene). The mutants were expressed and purified as above and 2D ^{1}H-^{13}C HSQC spectra were acquired and compared with the wild type to assign the missing resonance. For NMR titrations, the compounds were dissolved in 100% d_6-DMSO (CIL). ILV and ^{15}N chemical shift mapping were performed by acquiring 2D ^{1}H-^{13}C HSQC and ^{1}H-^{15}N TROSY spectra on 0.29 mM ^{15}N/ILV-labeled LcrV (or 0.31 mM ^{15}N/ILV-labeled PcrV) titrated with increasing concentrations of the compound. All the titration samples contained 2% (v/v) d_6-DMSO and 10% D_2O. The weighted chemical shift deviations (CSD) were calculated using: $0.5[(\Delta H)^2 + (\Delta C/2)^2]^{1/2}$ for ILV titrations and $0.5[(\Delta H)^2 + (\Delta N/5)^2]^{1/2}$ for ^{15}N titrations

55

([27](#)). The binding dissociation constants were estimated by non-linear curve fitting of CSDs *vs.* compound concentration plots for representative residues using GraphPad Prism 5.0.

3.4. Results

3.4.1. SPR-based fragment screening

The expression and purification of LcrV (residues 28-322, C273S) and PcrV (residues 25-294) generated millimolar amounts of purified proteins. High stability and solubility of these proteins facilitated SPR screening followed by hit validation and characterization by NMR. SPR screening of 288 compounds from a structurally diverse fragment library (Zenobia fragment library 2) identified novel small molecules binding to LcrV and PcrV. Interestingly, 3-(1H-pyrrol-1-yl)benzo[b]thiophene-2-carboxylic acid (PBTCA) showed binding to both LcrV and the homolog, PcrV (**Figure 3-1A, 3-1D**). Additionally, screening results also displayed the binding of 5-hydroxyindole and 4-(4-methylpiperazino)aniline to LcrV (**Figure 3-1B, 3-1C**); and 3-amino-7-bromoquinoline to PcrV (**Figure 3-1E**).

Analysis of structure-activity relationship (SAR) of the hit compounds (**1, 4, 8, and 11, Figure 3-2**) with non-binder analogs within the library enabled the determination of chemical groups important for the binding. The pyrrole moiety of PBTCA (**1, Figure 3-2**) was required for binding LcrV and PcrV as its removal (**2, Figure 3-2**) abolished the interaction with both the proteins (**Figure 3-3A, 3-3B**). Likewise, the thiophene carboxylic acid alone was incapable in binding LcrV and PcrV (**3**). The hydroxyl group of 5-hydroxyindole (**4**) was important to bind LcrV since its removal (**5**) or modification with an aza or a cyano group (**6, 7**) interfered with the binding. Regarding the binding of the piperazinoaniline scaffold to LcrV (**8**), the 4-methyl group of piperazine was necessary as its replacement with a

morpholine ring (**9**) or removal (**10**) resulted in loss of binding. For the quinoline scaffold binding to PcrV (**11**), substituting bromine at position 7 with fluorine (**12**) abrogated the interaction **(Figure 3-3C)**. Similarly, removal of the amine and the bromide groups and introduction of a carboxylic acid (**13**) also disrupted the binding. Our initial SAR studies suggest that specific molecular interactions between the small molecule ligand and the cognate protein are required to establish binding.
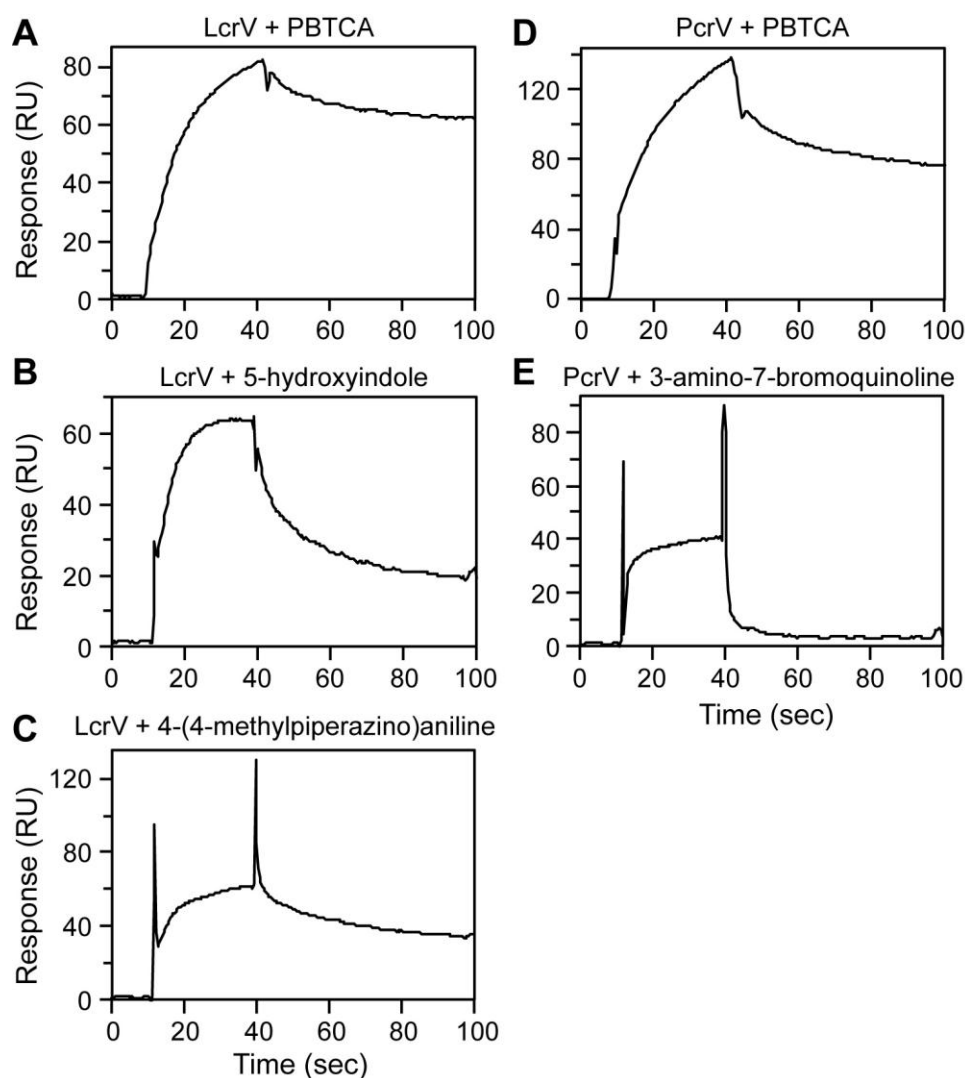
Figure 3-1. SPR screening of LcrV and PcrV. SPR sensorgrams displaying the binding of **(A)** 3-(1H-pyrrol-1-yl)benzo[b]thiophene-2-carboxylic acid (PBTCA), **(B)** 5-hydroxyindole, and **(C)** 4-(4-methylpiperazino)aniline to LcrV, **(D)** PBTCA, and **(E)** 3-amino-7-bromoquinoline to PcrV.
Data courtesy of Dr. Asokan Anbanandam.
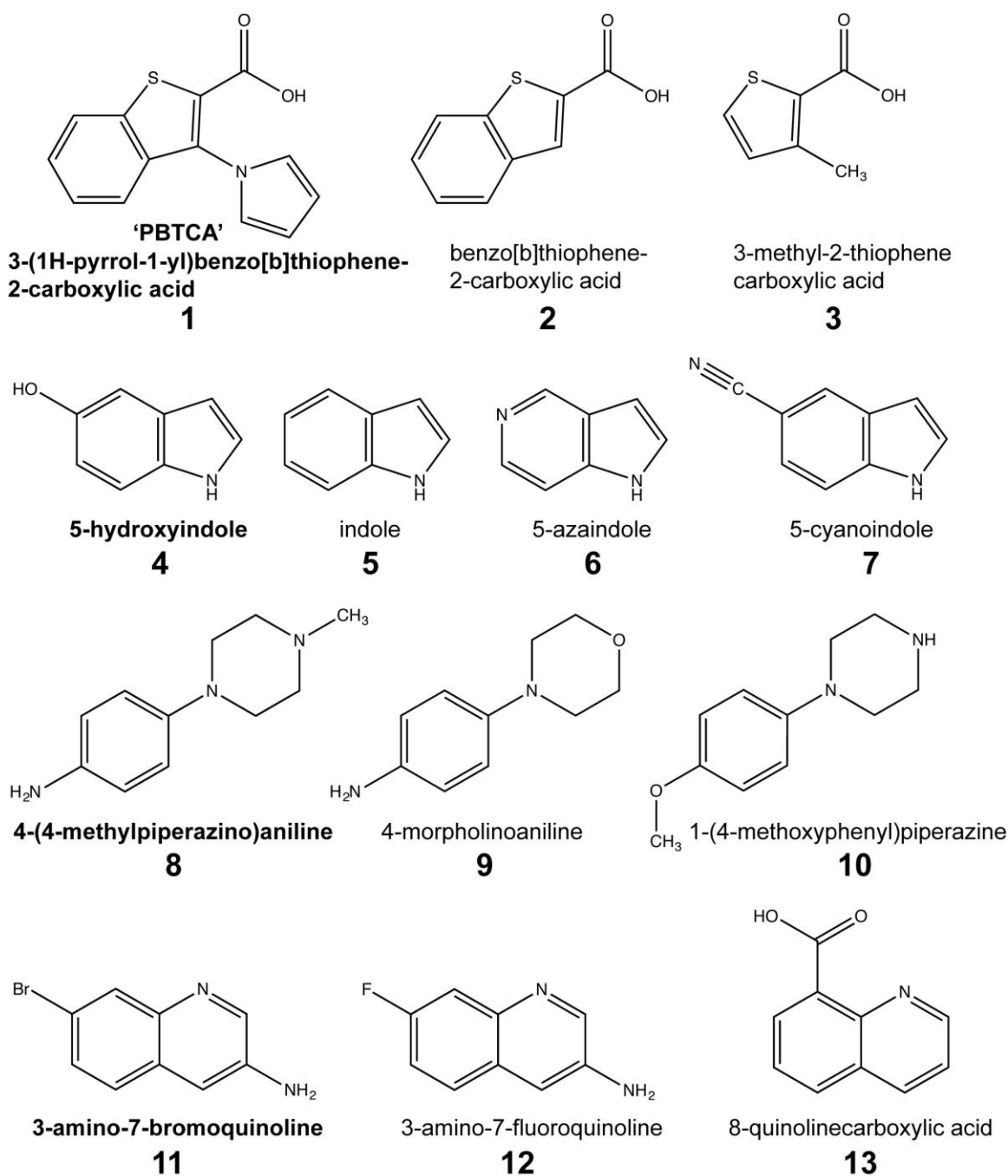
Figure 3-2. Structures of the fragment hits and their analogs. SPR screening showed the binding of PBTCA (**1**) to both LcrV and PcrV, 5-hydroxyindole (**4**) and 4-(4-methylpiperazino)aniline (**8**) to LcrV, and 3-amino-7-bromoquinoline (**11**) to PcrV. The non-binding analogs highlight the important chemical groups required for binding to LcrV or PcrV. The structures were made using ChemDraw.

3.4.2. Validation of fragment hits using STD NMR

Ligand-observed saturation transfer difference NMR (STD NMR) is a powerful tool to quickly validate the fragments identified as hits in the initial screening and to recognize ligand moieties important for the binding. Results of STD NMR confirmed the direct interaction of PBTCA, 5-hydroxyindole, and 4-(4-methylpiperazino)aniline to LcrV **(Figure 3-4A, 3-4B, 3-4C)**, and PBTCA and 3-amino-7-bromo-quinoline to PcrV **(Figure 3-4D, 3-4E)**. As expected, the off-resonance reference spectra showed NMR signals from both compound and protein **(Figure 3-4, *top panel*)**. Importantly, strong STD signals were readily observable for all the five hits **(Figure 3-4, *bottom panel*)**. Since most of the NMR signals corresponding to the compound remained above noise level in all the resultant STD spectra **(Figure 3-4)**, this suggests that the entire molecule is binding or is in close proximity of the target protein.

3.4.3. NMR assignments of LcrV

Chemical shift assignment is a critical first step in NMR-based protein interaction studies. LcrV (with 295 residues) is a large protein, making the NMR backbone amide assignment challenging. Side chain methyl ILV (Isoleucine, Leucine, and Valine) groups with high sensitivity and sharp NMR resonances offer useful probes for spectroscopic studies of high molecular weight proteins (28). The 25 isoleucine, 31 leucine, and 14 valine residues of LcrV are distributed among the 3 distinct parts of the protein: the N-terminal globular region, the central coiled-coil, and the distal region, and provide coverage of the entire protein. The 2D ^{1}H-^{13}C HSQC spectrum of ILV labeled LcrV showed 109 of the expected 115 peaks (1 peak for each Ile, and 2 peaks for each Leu and Val residue, **(Figure 3-5)**. Using a combination of site-directed mutagenesis (Ile to Leu mutations, **Figure 3-6**) and through space ^{1}H-^{1}H nuclear Overhauser effects (NOEs) obtained from 3D HMQC-NOESY (25) and analyzed using the crystal structure of LcrV (11) **(Figure 3-7)**, near complete methyl ILV assignments were obtained

(**Figure 3-5**, 67 assigned of the total 70 ILV residues; V63, L75, and L85 remain unassigned due to missing or overlapping peaks). Although challenging, 53 of the total 295 backbone amide peaks were also assigned using 3D HNCA, 3D HNCACB and 3D CBCA(CO)NH **(Figure 3-8)**. These are currently the only available methyl ILV and backbone amide assignments for the entire LcrV family of T3SS tip proteins and would facilitate molecular interaction studies of LcrV.
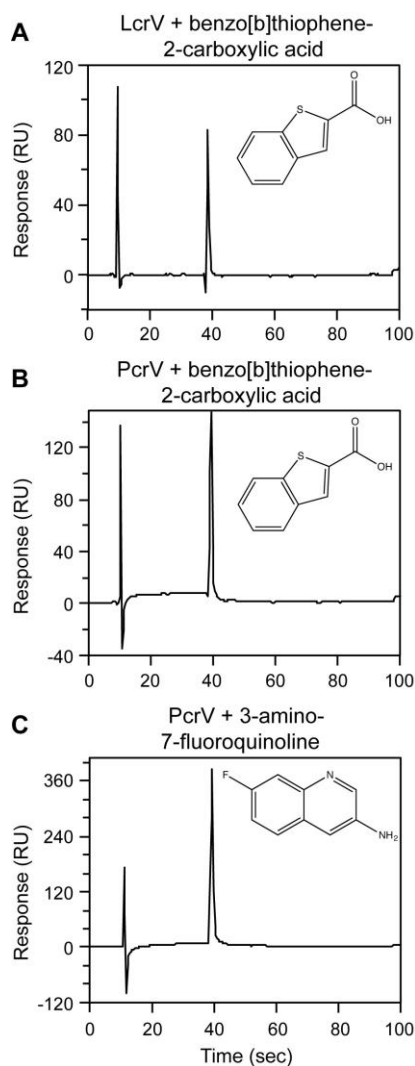
Figure 3-3. SPR sensorgrams of representative non-binding analogs of the hit fragments from LcrV and PcrV. Benzo[b]thiophene-2-carboxylic acid, an analog of PBTCA (**1**, Figure 3-2) shows no binding to **(A)** LcrV, or **(B)** PcrV. **(C)** 3-amino-7-fluoroquinoline, an analog of 3-amino-7-bromoquinoline (**11**, Figure 3-2) shows no binding to PcrV.
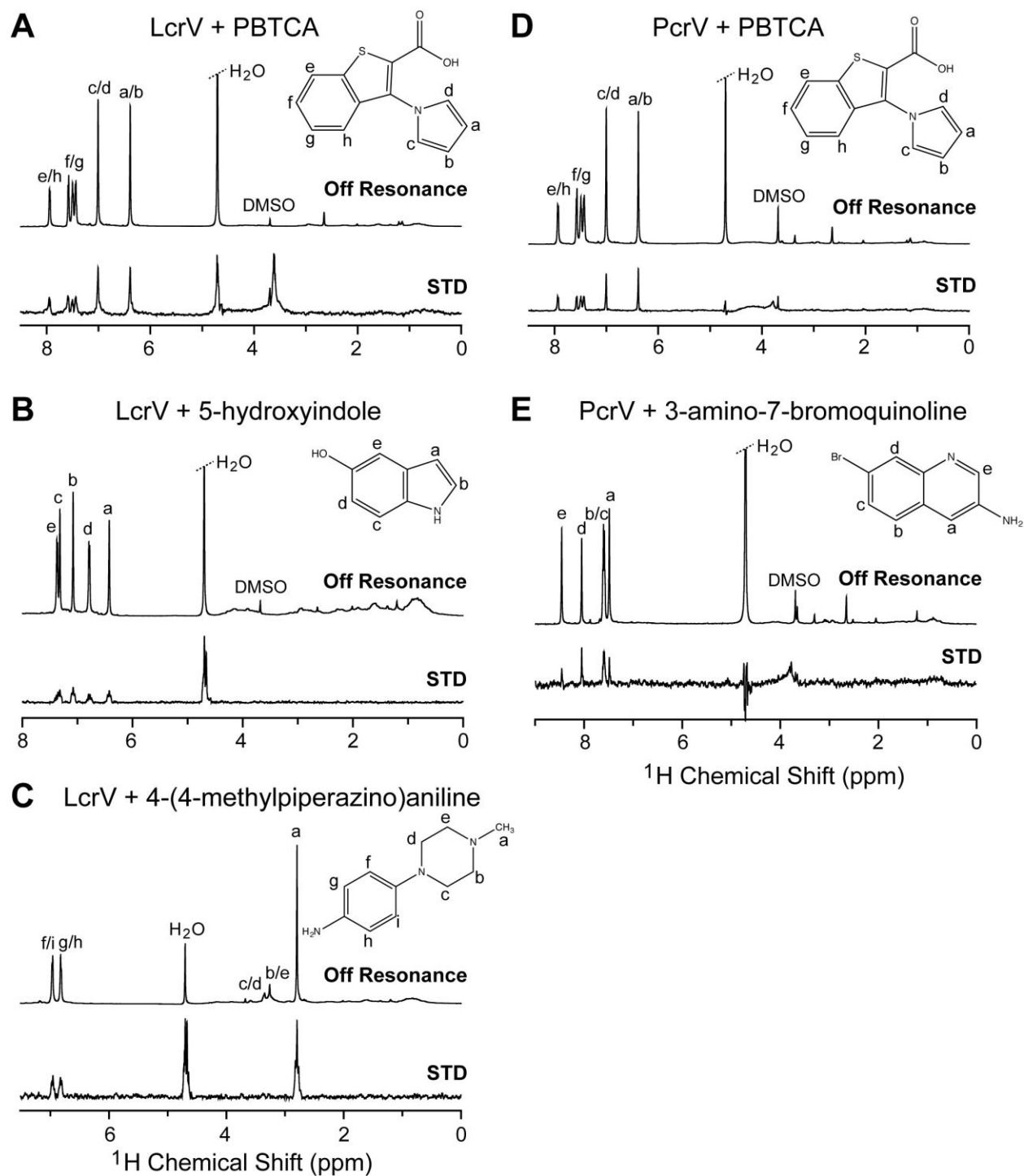
Figure 3-4. STD NMR of the fragment hits. **(A)** PBTCA, **(B)** 5-hydroxyindole, and **(C)** 4-(4-methylpiperazino)aniline with LcrV, **(D)** PBTCA, and **(E)** 3-amino-7-bromoquinoline with PcrV. (*Top panel*) Off resonance spectra, and (*Bottom panel*) STD NMR spectra.
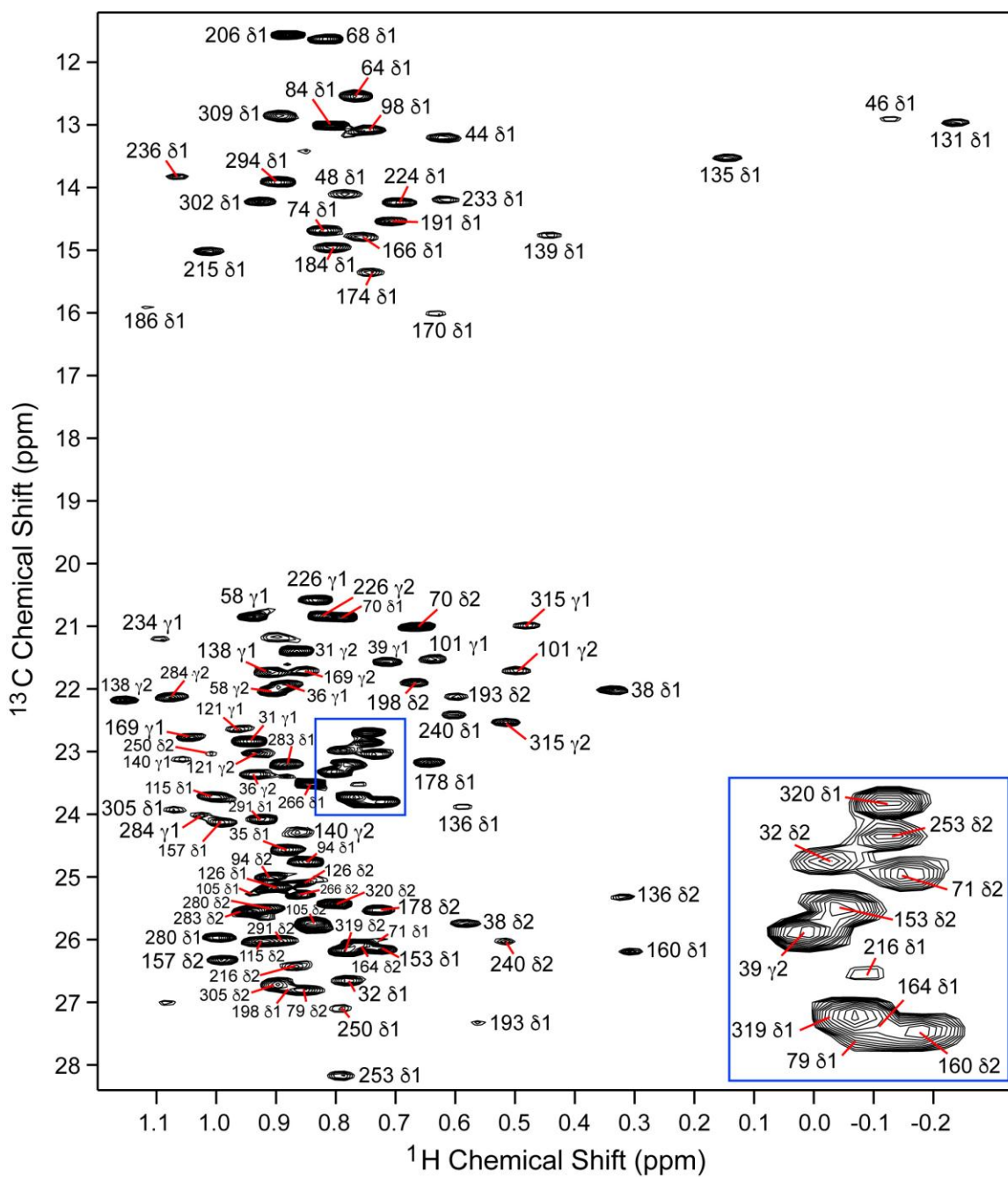
Figure 3-5. Assigned 2D ^{1}H-^{13}C ILV HSQC spectrum of LcrV.
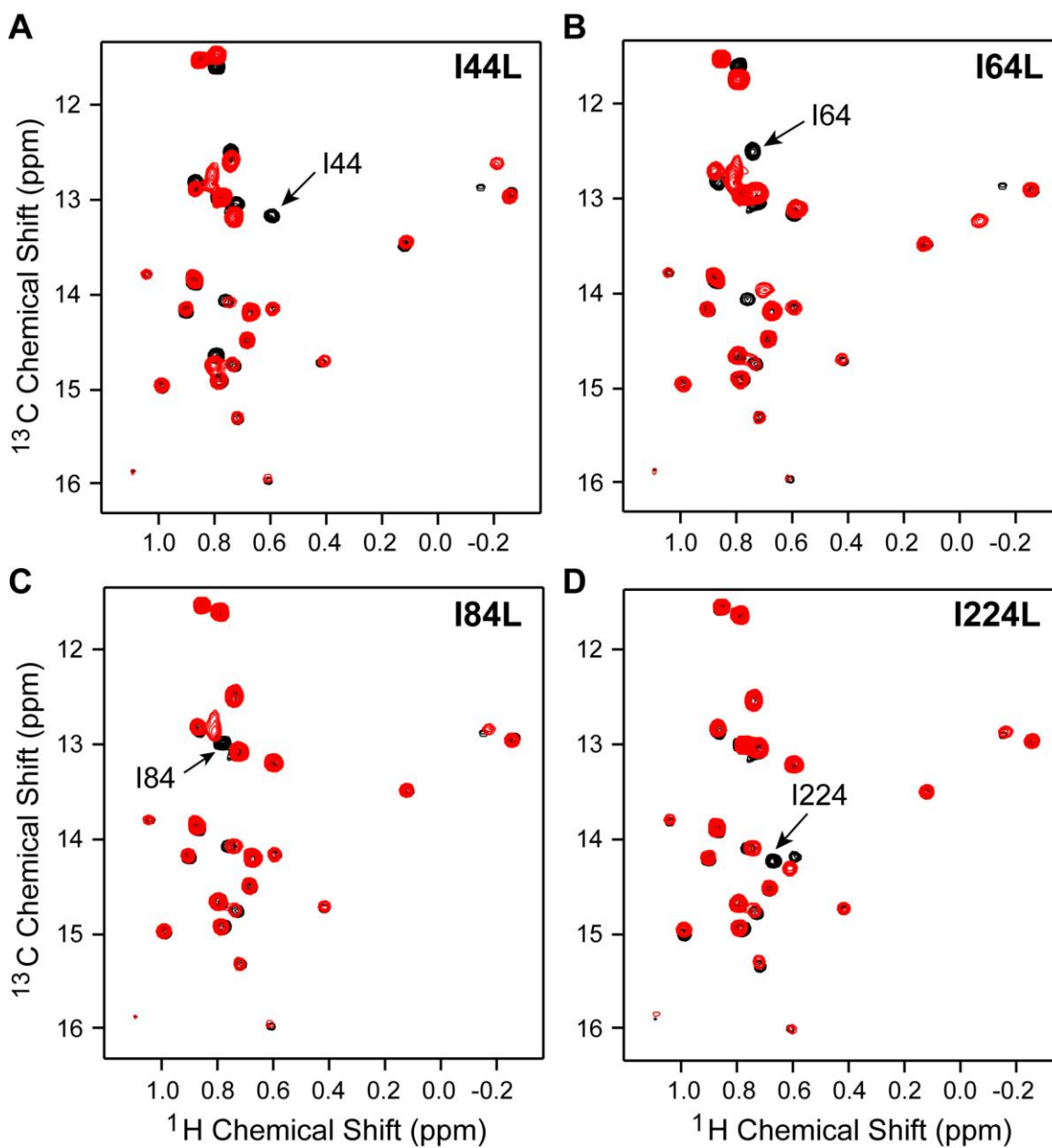
Figure 3-6. Site-directed mutagenesis approach used in assigning isoleucine $^{13}C\delta1$ methyl peaks of LcrV. (A-D) 2D ^{1}H-^{13}C HSQC spectra of selected isoleucine to leucine mutants of LcrV (red) overlaid with the wild type reference spectra (black). The single missing peak corresponding to the mutated residue is labeled.
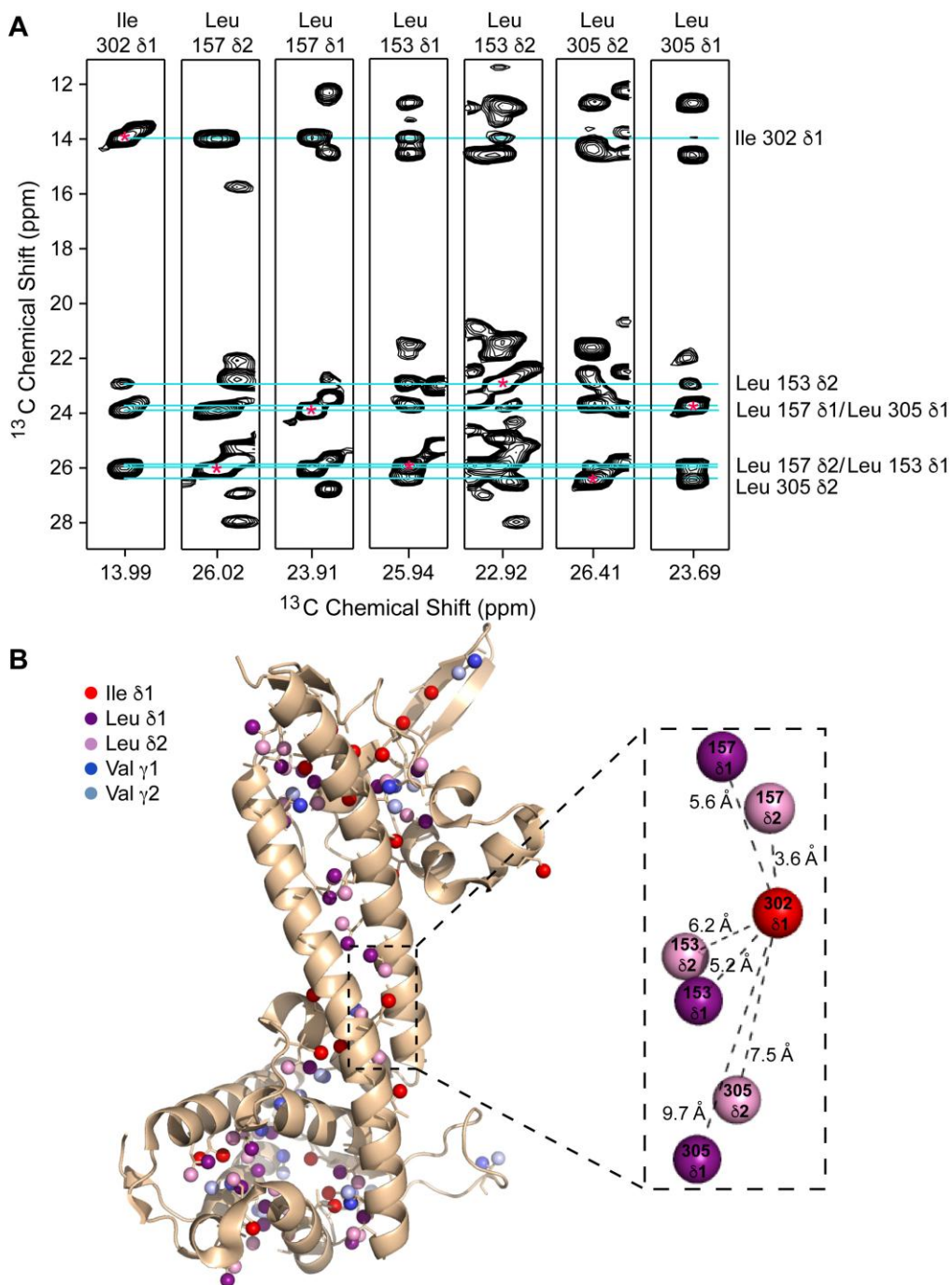
Figure 3-7. The nuclear Overhauser effect (NOE) used in the assignment of ILV methyl resonances of LcrV. **(A)** Representative strips from a 3D ^{1}H-^{13}C-^{13}C HMQC-NOESY-HMQC dataset of perdeuterated ILV labeled LcrV. Asterisks mark the diagonal peak and NOEs with other residues are shown with blue lines. **(B)** Crystal structure of LcrV (PDB ID: 4JBU) with the ILV methyl groups shown in spheres (Ile Cδ1, red; Leu Cδ, magenta & pink; Val Cγ, blue & light blue). An expansion of a region near I302 shows correlation of distance information with NOESY data to assign methyl peaks.
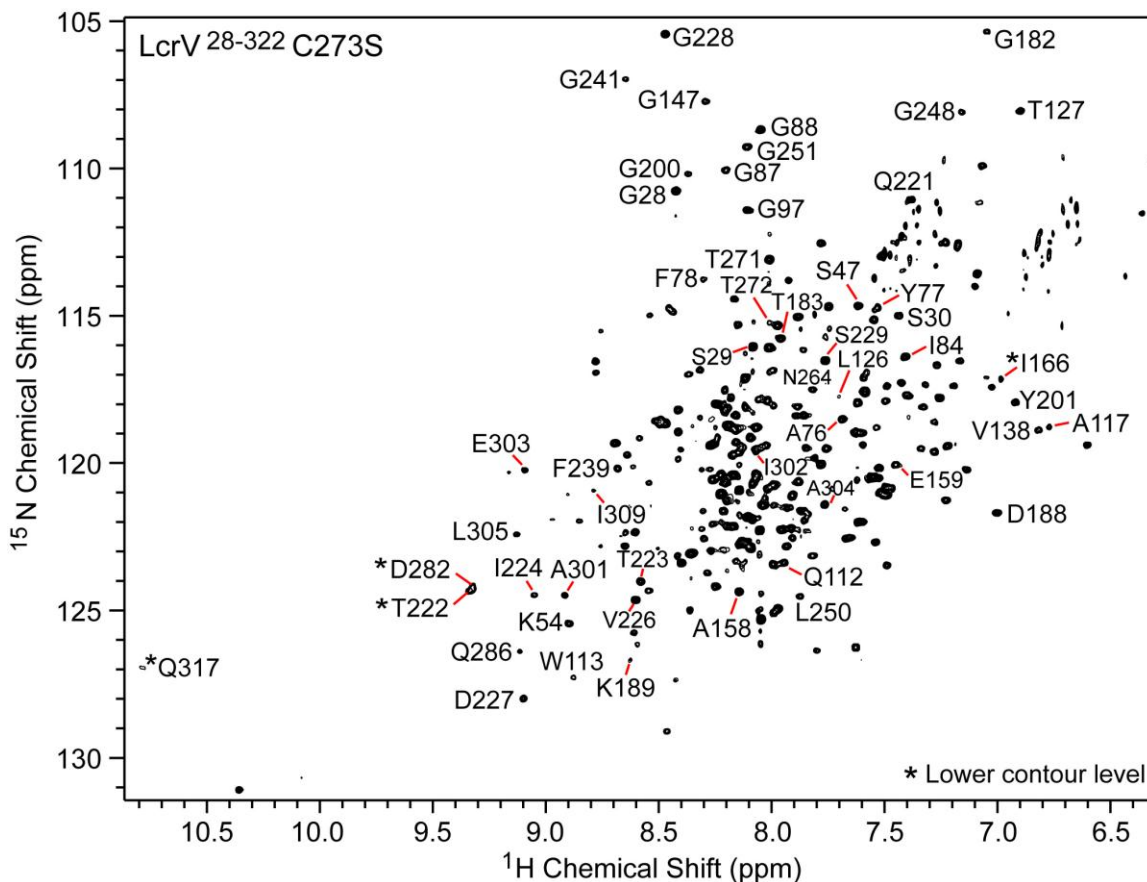
Figure 3-8. Assigned 2D ^{1}H-^{15}N TROSY spectrum of LcrV. Residues indicated with asterisk appeared at a lower contour level. Partially assigned by Dr. Narsimulu Kongari.

3.4.4. NMR titrations of ILV and 15N-labeled LcrV with PBTCA

To determine the binding site of PBTCA on LcrV, ILV and ^{15}N-labeled LcrV was titrated with increasing concentrations of PBTCA at molar ratios of 1:0, 1:4, 1:8, and 1:14. Simultaneous ILV and ^{15}N labeling of LcrV enabled titration to be monitored with two different probes. First, 2D ^{1}H-^{13}C HSQC spectra were acquired for various titration points to investigate the effect of PBTCA on the ILV methyl resonances of LcrV **(Figure 3-9A)**, and second, 2D ^{1}H-^{15}N TROSY spectra were acquired to query the effect of interaction on ^{15}N backbone amides of LcrV **(Figure 3-9B)**. The stepwise addition of PBTCA

resulted in significant chemical shift perturbations of specific LcrV resonances. Both the ILV and amide-based titrations showed concentration dependent changes in the peak positions for the affected LcrV residues (*e.g.* I64, I174, L240, L320, S47, G87, G88, and Q112), indicating interaction to be in fast exchange NMR time scale **(Figure 3-9C, 3-9D)**.

To map the LcrV/PBTCA interaction interface, we calculated the weighted chemical shift deviation (CSD) for each assigned non-overlapped ILV **(Figure 3-10A)** and amide **(Figure 3-10B)** peak of LcrV. Residues that displayed significant CSD (average plus a standard deviation) varied from hydrophobic (L35, I64, L79, L126, I174, I184, and L240) to polar uncharged (S47, Y77, T183, and T272) residues. Results of the ILV titrations complemented backbone amide titrations and together indicated that the affected residues map largely on two surfaces on LcrV: bottom of the N-terminal globular domain formed by residues L35, I64, Y77, L79, and L126; and top of the central coiled-coil involving residues I174, T183, I184, L240, and T272 **(Figure 3-10C)**. Although CSDs of residues at the top of the central coiled-coil did not reach saturation by our final titration point, based on the CSDs of residues located in the N-terminal globular region (S30, L126, and T127), the binding dissociation constant of 7 ± 1 mM could be estimated for the LcrV/PBTCA interaction **(Figure 3-10D)**. Accordingly, it can be speculated that PBTCA primarily binds near the bottom of the N-terminal globular domain, whereas the top of the coiled-coil probably presents a weaker secondary binding site for the LcrV/PBTCA interaction.

3.4.5. *NMR titrations of LcrV with 5-hydroxyindole and 4-(4-methylpiperazino)aniline*

Besides PBTCA, LcrV also displayed binding to two other scaffolds, 5-hydroxyindole and 4-(4-methylpiperazino)aniline, as shown by SPR **(Figure 3-1B, 3-1C)** and STD NMR **(Figure 3-4B, 3-4C)**.
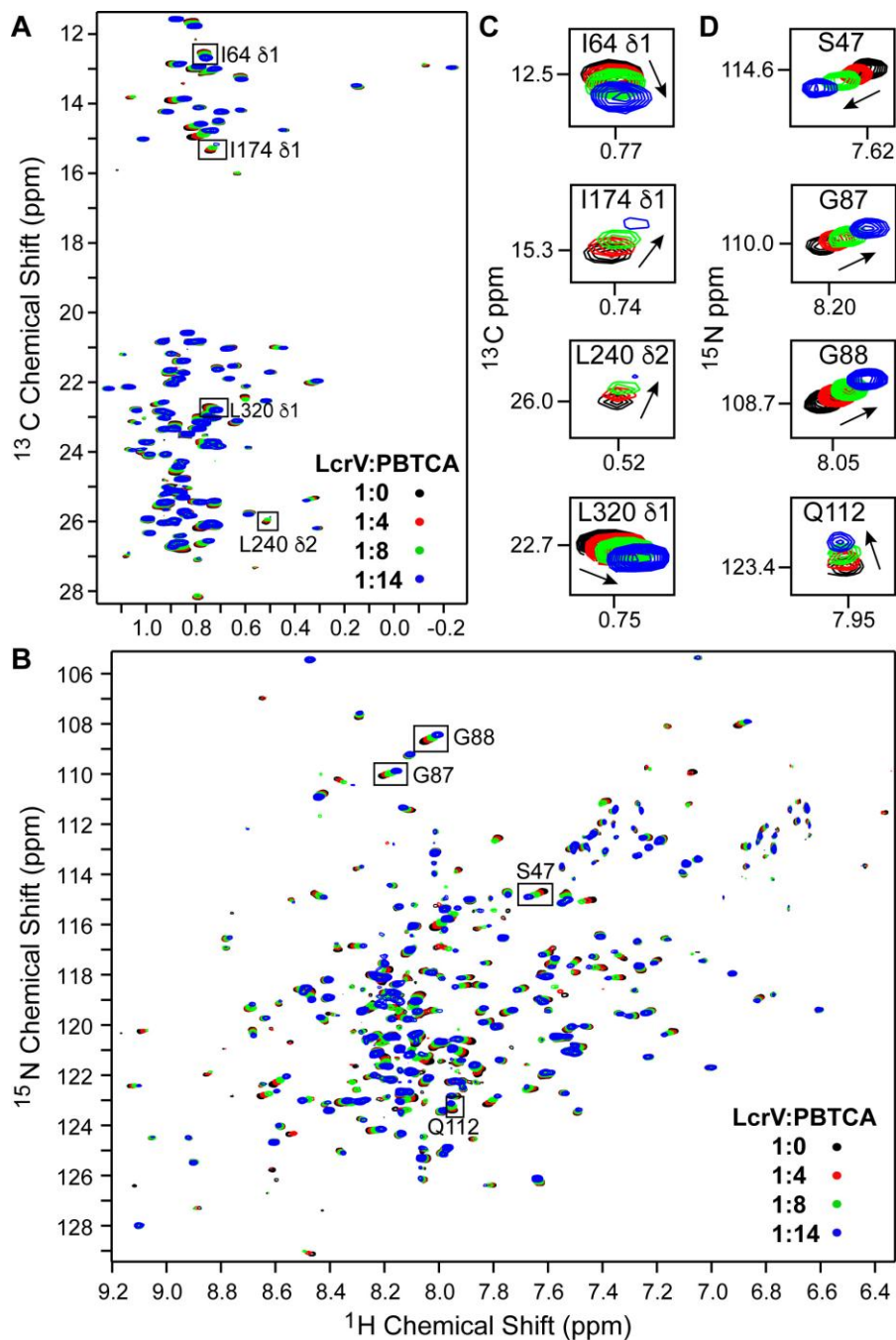
Figure 3-9. ILV and ^{15}N NMR titrations of LcrV with PBTCA. Overlay of four **(A)** 2D ^{1}H-^{13}C HSQC spectra, and **(B)** 2D ^{1}H-^{15}N TROSY spectra of ILV and ^{15}N-labeled LcrV titrated with increasing molar ratios of PBTCA. Expanded sections of selected LcrV residues affected by the interaction with PBTCA in **(C)** ILV and **(D)** ^{15}N NMR titrations.
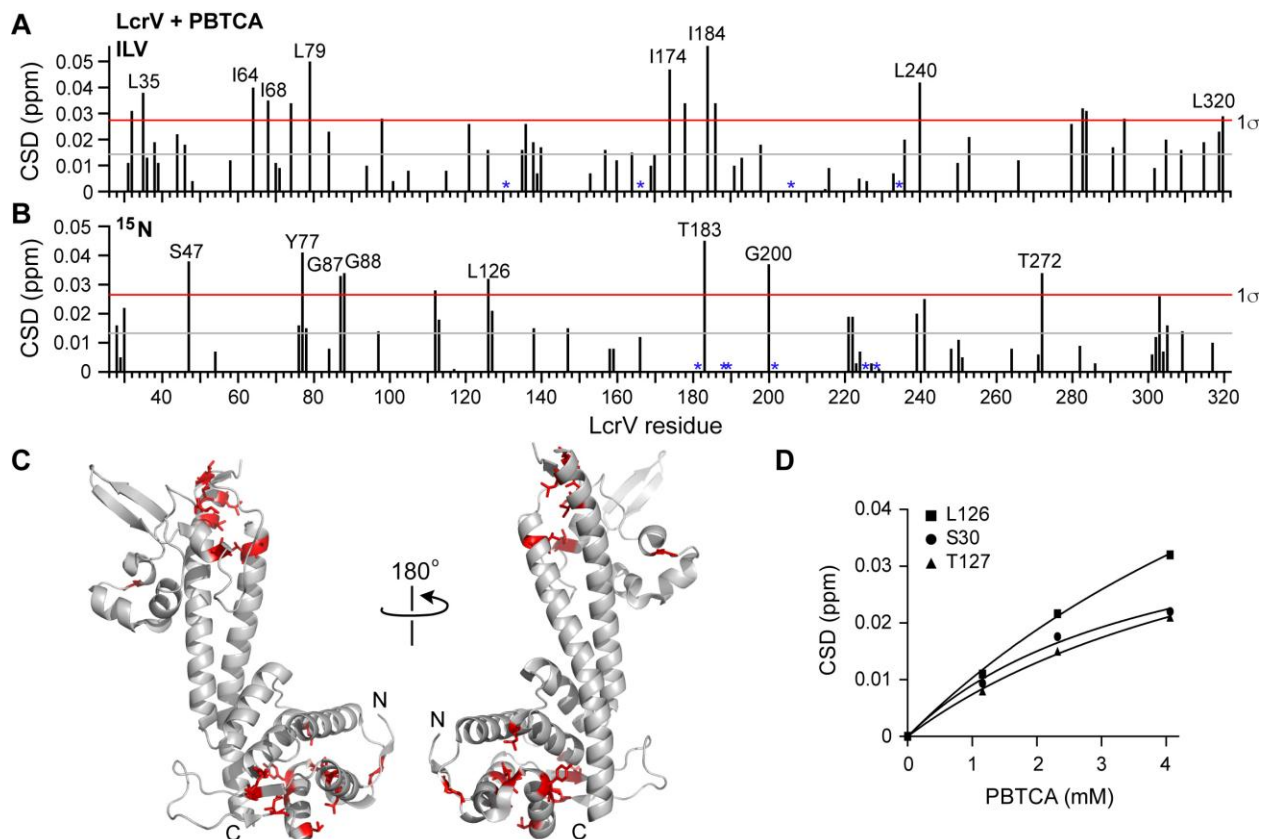
Figure 3-10. Analysis of the interaction of LcrV with PBTCA. Plots of weighted chemical shift deviation (CSD) from **(A)** ILV and **(B)** ^{15}N titrations of LcrV with PBTCA. Gray and red lines correspond to the mean and one standard deviation (1σ) from the mean, respectively. Residues with no deviation (zero CSD value) are shown with asterisk. **(C)** LcrV residues strongly affected by the binding of PBTCA shown in red. **(D)** Plot of CSD *vs.* concentration of PBTCA for selected residues.

NMR titrations of LcrV with 5-hydroxyindole at protein to compound ratios of 1:0, 1:25, 1:50, and 1:92 resulted in concentration dependent deviation in the peak positions of specific LcrV residues as shown by ^{13}C HSQC **(Figure 3-11A and Figure 3-12A)** and ^{15}N TROSY **(Figure 3-11B and Figure 3-12B)**, suggesting interaction in fast exchange NMR time scale. Analysis of the CSDs for ILV **(Figure 3-12C)** and amide **(Figure 3-12D)** titrations identified that majority of the affected residues reside in the N-terminal globular domain of LcrV and few lie near the middle of the coiled-coil **(Figure 3-12E)**. Except for some distant residues, such as, V140, L157, and I294, residues showing significant CSDs mainly clustered around the bottom of the N-terminal globular region, the region identified earlier important for

PBTCA binding **(Figure 3-10C)**. Based on nonlinear regression fits of CSDs *vs* 5-hydroxyindole concentration for some representative affected residues (L38, G87, I309, and Q317), the dissociation constant of 29 ± 5 mM was determined for the binding of 5-hydroxyindole to LcrV **(Figure 3-11C)**.

Similar to the ILV and amide titrations of LcrV with PBTCA **(Figure 3-9)** and 5-hydroxyindole **(Figure 3-11)**, addition of 4-(4-methylpiperazino)aniline at molar ratios of 1:0, 1:12, 1:25, and 1:50 induced distinct chemical shift deviations of specific LcrV resonances **(Figure 3-13A, 3-13B)**. Representative set of residues affected by 4-(4-methylpiperazino)aniline from the ILV (L38, I46, V101, and L178) and ^{15}N (S47, W113, G147, and G200) titrations exhibited peaks in fast exchange NMR time scale **(Figure 3-14A, 3-14B)**. Plots of the CSDs **(Figure 3-13C, 3-13D)** identified that the LcrV residues strongly perturbed by 4-(4-methylpiperazino)aniline are primarily confined towards the bottom of the N-terminal globular domain. The binding dissociation constant of 13 ± 1 mM was estimated based on the CSDs of some significantly affected residues (I46, V138, G147, and T272, **Figure 3-13E**). Together, based on the interaction studies of LcrV with PBTCA, 5-hydroxyindole and 4-(4-methylpiperazino)aniline, region near the bottom of the N-terminal globular domain of LcrV is expected to be the "hotspot" for small molecule fragment binding.

3.4.6. PcrV binds to PBTCA and 3-amino-7-bromoquinoline

In addition to LcrV, we also screened for fragments binding to PcrV, the LcrV homolog in *P. aeruginosa*. Interestingly, PBTCA, the fragment identified earlier as binding to LcrV, also showed binding to PcrV **(Figure 3-1D, 3-4D)**. Apart from PBTCA, SPR and STD NMR also displayed the binding of 3-amino-7-bromoquinoline to PcrV **(Figure 3-1E, 3-4E)**. Our efforts to assign the methyl ILV resonances of PcrV utilizing the previously described 'NOE' based assignment approach have so far been

unsuccessful due to lack of high-resolution structure of PcrV. However, using site-directed mutagenesis, some ILV assignments of PcrV have been obtained **(Figure 3-15)**.
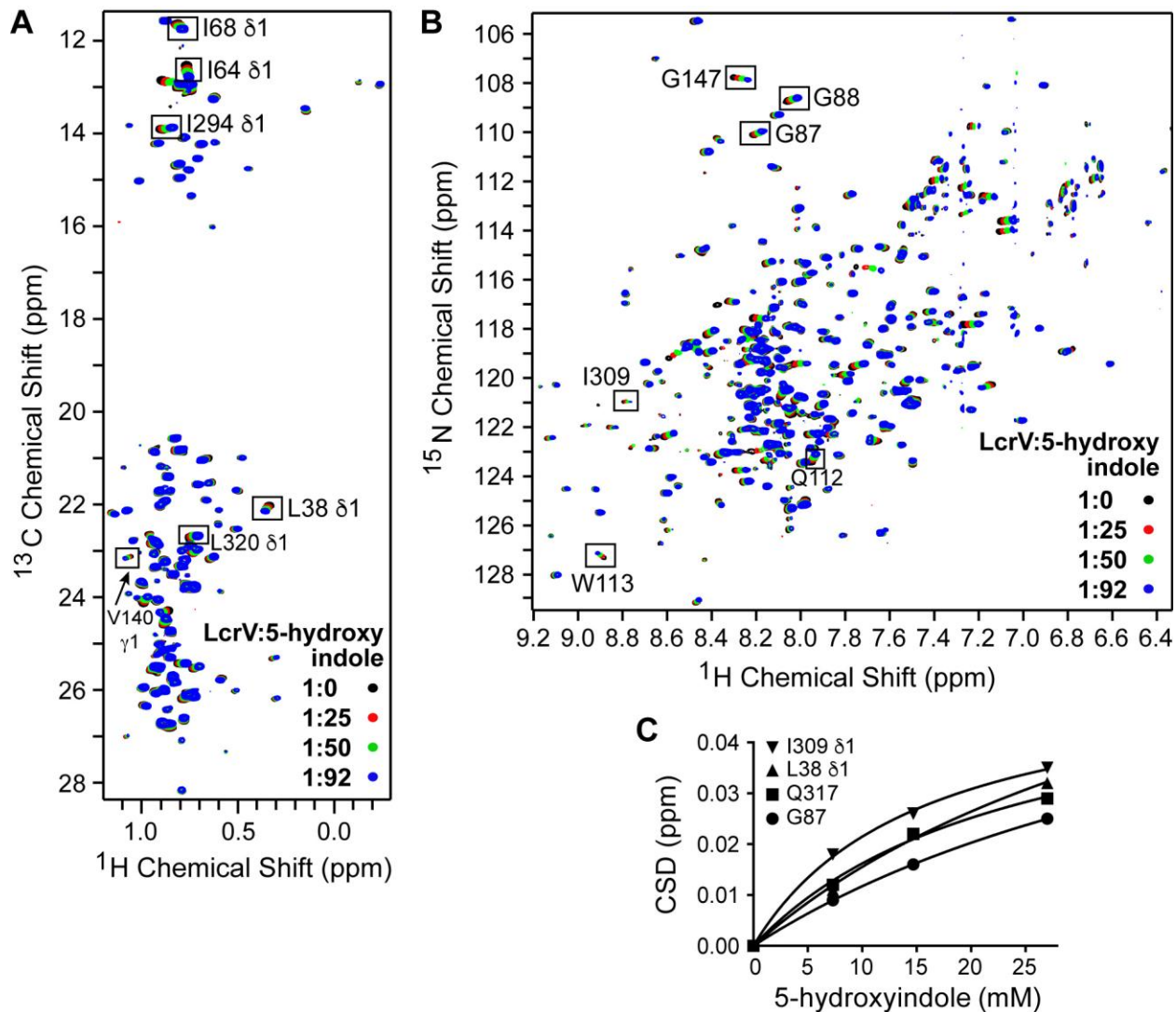
Figure 3-11. ILV and ¹⁵N NMR titrations of LcrV with 5-hydroxyindole. Overlay of four **(A)** 2D ¹H-¹³C HSQC spectra, and **(B)** 2D ¹H-¹⁵N TROSY spectra of ILV and ¹⁵N-labeled LcrV titrated with increasing molar ratios of 5-hydroxyindole. Expanded views of boxed residues are shown in Figure 3-12. **(C)** Plot of CSD *vs.* concentration of 5-hydroxyindole for selected affected residues of LcrV.
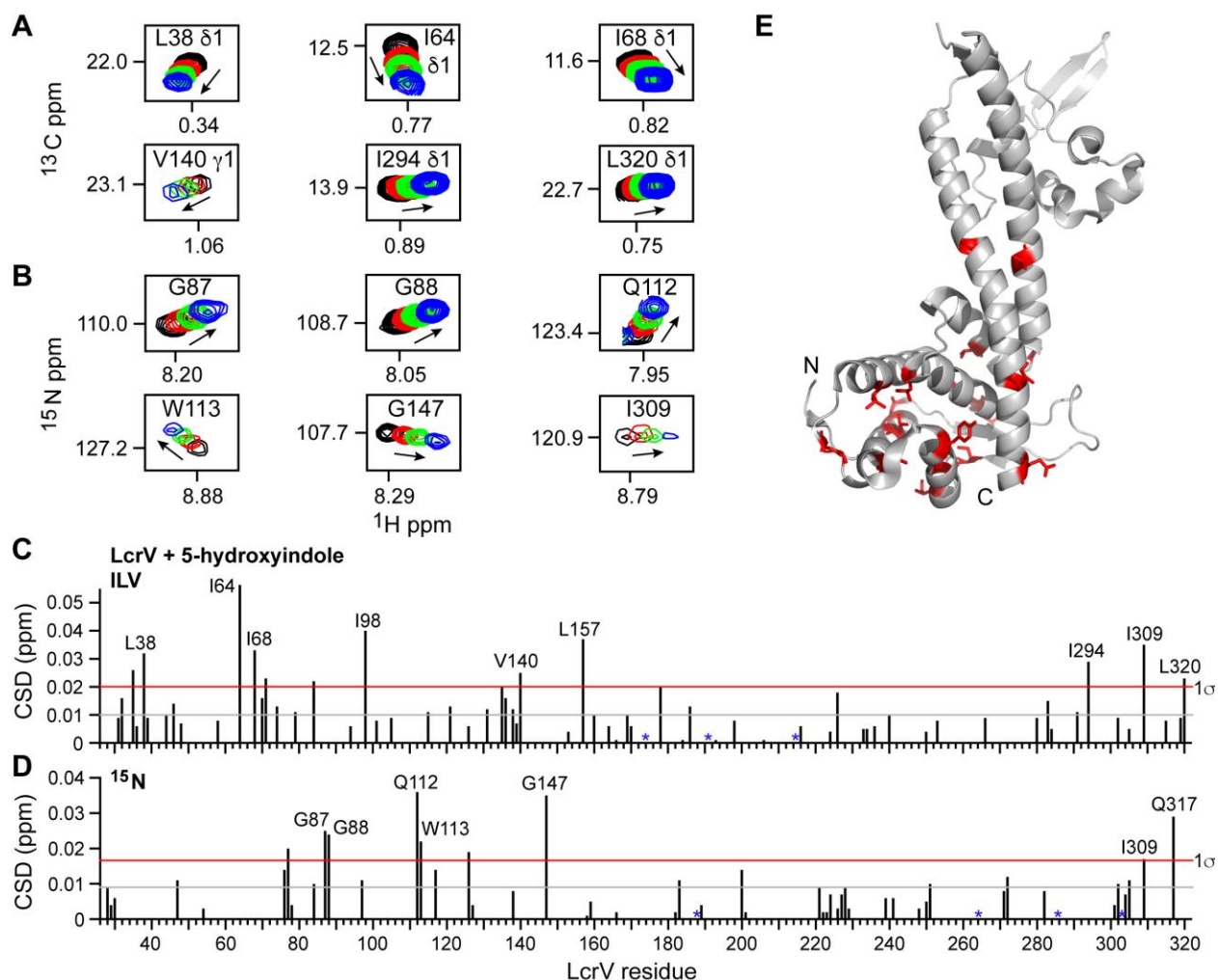
Figure 3-12. Analysis of the interaction of LcrV with 5-hydroxyindole. Representative LcrV residues affected by interaction with 5-hydroxyindole in (**A**) ILV and (**B**) ^{15}N NMR titrations. Plots of CSD from (**C**) ILV and (**D**) ^{15}N titrations of LcrV with 5-hydroxyindole. Gray and red lines correspond to the mean and one standard deviation (1σ) from the mean, respectively. Residues with no deviation (zero CSD value) are shown with asterisk. (**E**) LcrV residues strongly affected by the binding of 5-hydroxyindole highlighted in red.
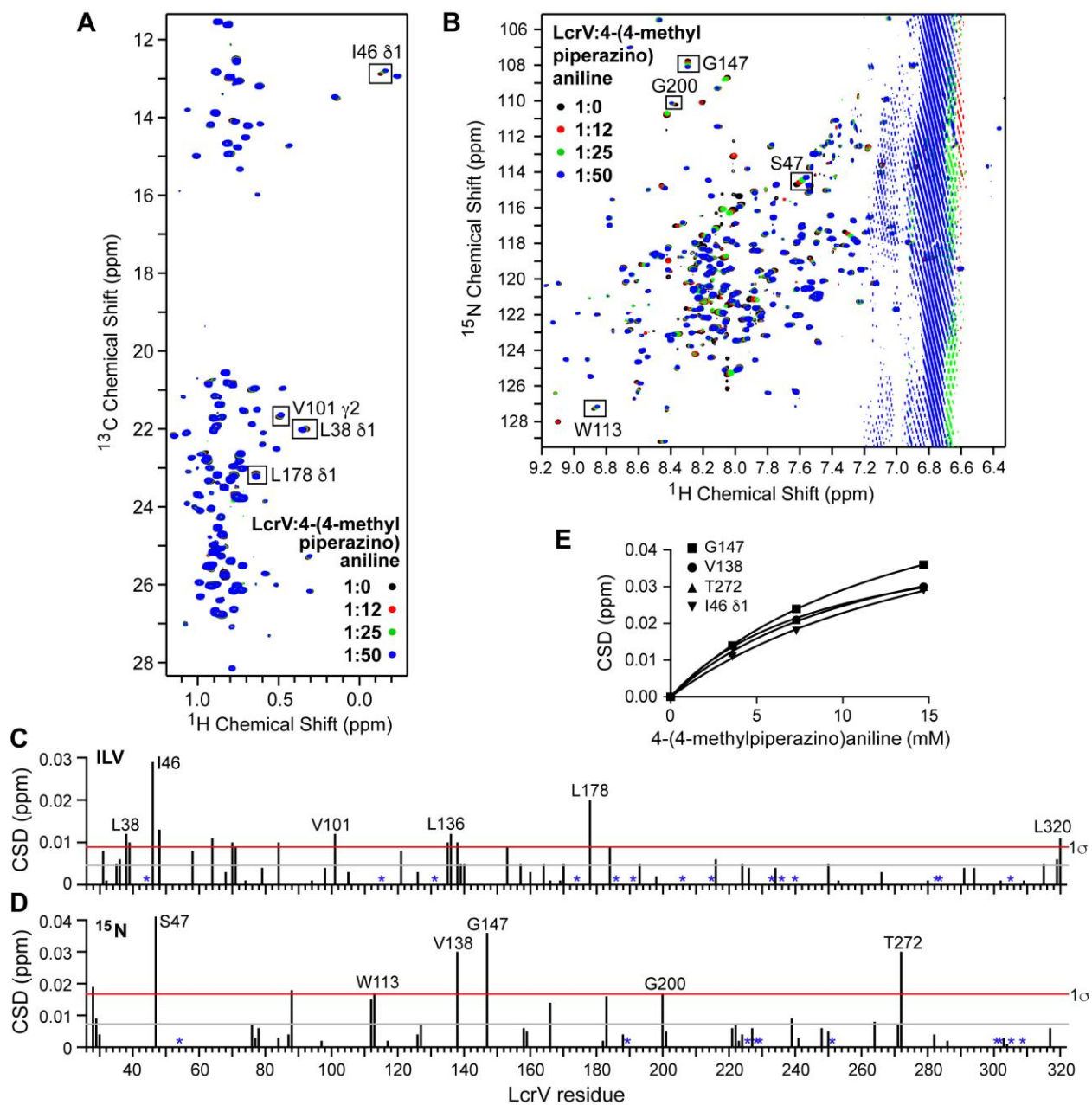
Figure 3-13. ILV and ^{15}N NMR titrations of LcrV with 4-(4-methylpiperazino)aniline. Overlay of four **(A)** 2D ^{1}H-^{13}C HSQC spectra and **(B)** 2D ^{1}H-^{15}N TROSY spectra of ILV and ^{15}N-labeled LcrV titrated with increasing mole ratio of 4-(4-methylpiperazino)aniline. Expanded views of boxed residues are shown in Figure 8 of the main text. The vertical signals at ~6.8 ppm in (B) are from compound. Plots of CSD from **(C)** ILV and **(D)** ^{15}N titrations of LcrV with 4-(4-methylpiperazino)aniline. Gray and red lines correspond to the mean and one standard deviation (1σ) from the mean, respectively. Residues with no deviation (zero CSD value) are shown with asterisk. **(E)** Plot of CSD *vs.* concentration of 4-(4-methylpiperazino)aniline for representative affected LcrV residues.
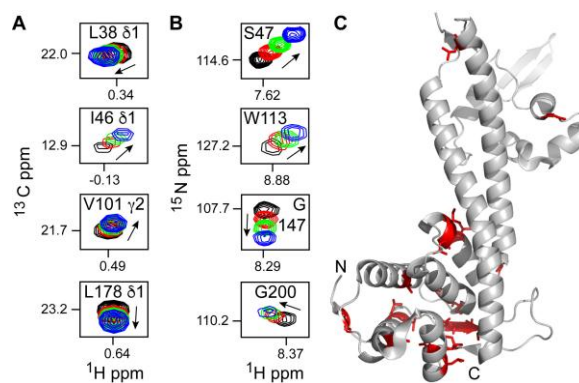
Figure 3-14. NMR titration results of LcrV with 4-(4-methylpiperazino)aniline. Expanded sections of representative LcrV residues affected by the interaction of 4-(4-methylpiperazino)aniline in **(A)** ILV and **(B)** ^{15}N NMR titrations. **(C)** LcrV residues affected by the binding of 4-(4-methylpiperazino)aniline are colored red.
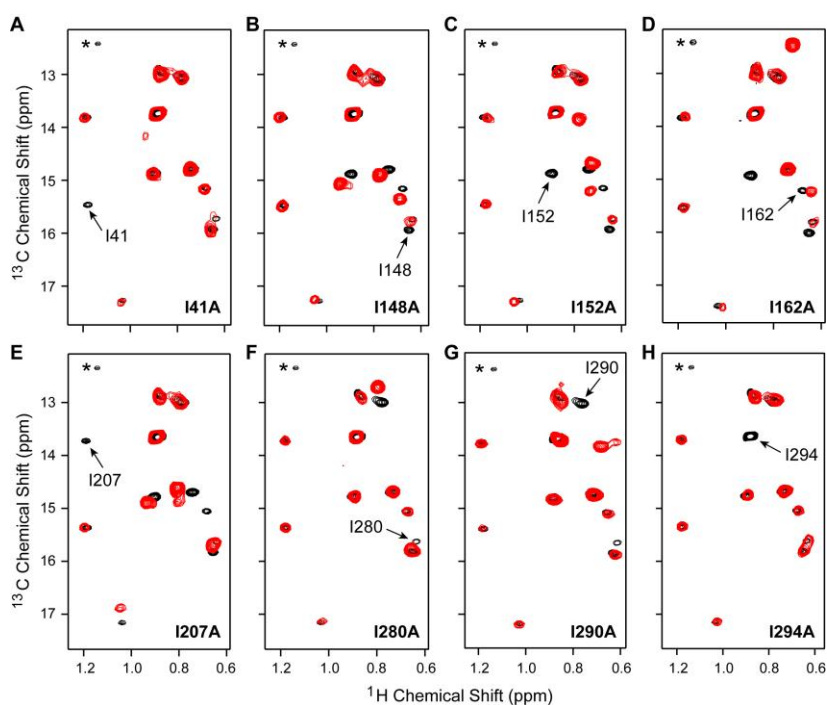
Figure 3-15. Site-directed mutagenesis approach used in assigning ILV methyl peaks of PcrV. (A-H) 2D ^{1}H-^{13}C HSQC spectra of selected isoleucine to alanine mutants of PcrV (red) overlaid with the wild type reference spectra (black). The missing peak corresponding to the mutated residue is labeled. Occasionally, ILV residues in close proximity of the mutated residue showed chemical shift perturbation. Asterisk indicates a folded leucine peak of PcrV.

Stepwise titration of ILV and ^{15}N labeled PcrV with PBTCA at 1:0, 1:4, and 1:8 molar ratios and monitored by ^{13}C HSQC **(Figure 3-16A)** and ^{15}N TROSY **(Figure 3-17A)** showed distinct chemical shift perturbations of specific PcrV resonances in a concentration dependent manner. In both ILV and amide titrations, along with progressive changes in the peak positions for some affected PcrV residues (*boxed* peaks depicting fast exchange phenomenon), other PcrV residues affected upon PBTCA binding showed decrease in their peak intensities (peaks shown by *arrows* indicative of intermediate exchange NMR time scale, **Figure 3-16A, Figure 3-17A**). Likewise, NMR titrations with 3-amino-7-bromoquinoline at protein to compound ratios of 1:0, 1:2, and 1:4 showed both the deviations in peak positions and reduction in the peak intensities for the affected residues (**Figure 3-16B and Figure 3-17B**). Even though most of the NMR assignments for PcrV are currently not available, it is evident that the resonances affected by PBTCA and 3-amino-7-bromoquinoline are largely same in both ILV (**Figure 3-16A** compared with **3-16B**) and ^{15}N amide titrations (**Figure 3-17A** compared with **3-17B**), suggesting that both fragments recognize a similar surface on PcrV. Using the assigned ILV resonances, L105, L121, I280, and I290 were identified as strongly affected residues in both PBTCA and 3-amino-7-bromoquinoline titrations **(Figure 3-16A, 3-16B)**. Mapping the affected residues onto the I-TASSER model of PcrV ([13](#)) highlight region near the N-terminal domain of PcrV **(Figure 3-16C)**. Our results suggest that similar to the fragment binding pocket of LcrV, region near the N-terminal globular domain of PcrV likely provides the interaction site for small molecule binding.

3.5. Discussion

The T3SS V-tip proteins, LcrV and PcrV are crucial virulence factors of *Y. pestis* and *P. aeruginosa*, respectively ([9,14](#)). Since these proteins are exposed on the bacterial surface and indispensable for the proper assembly and functioning of the T3SS, they represent an attractive target for

the development of anti-virulence drugs ([5](#),[29-32](#)). Here, we identified and characterized the binding of three novel small molecules, PBTCA, 5-hydroxyindole, and 4-(4-methylpiperazino)aniline to LcrV, and two molecules, PBTCA and 3-amino-7-bromoquinoline to PcrV. The fragments identified in this work represent the first small molecules shown to bind to LcrV and PcrV **(Figure 3-1 and Figure 3-2)** and could potentially be used as scaffolds toward rational design of inhibitors of the T3SS.
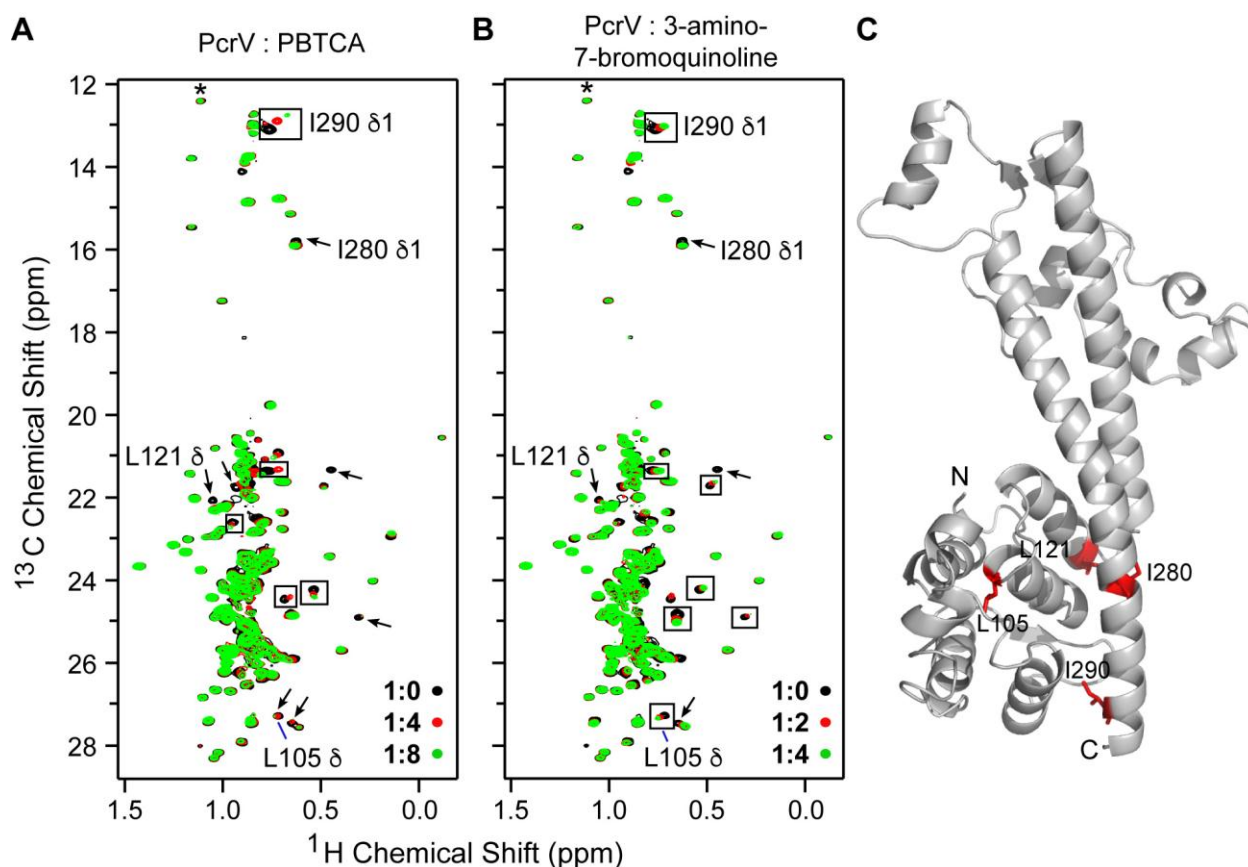
Figure 3-16. ILV titrations of PcrV with PBTCA and 3-amino-7-bromoquinoline. Overlay of three 2D ^{1}H-^{13}C HSQC spectra of ILV-labeled PcrV titrated with increasing mole ratios of **(A)** PBTCA and **(B)** 3-amino-7-bromoquinoline. **(C)** I-TASSER model of PcrV with residues affected in titrations with PBTCA and 3-amino-7-bromoquinoline shown in red.
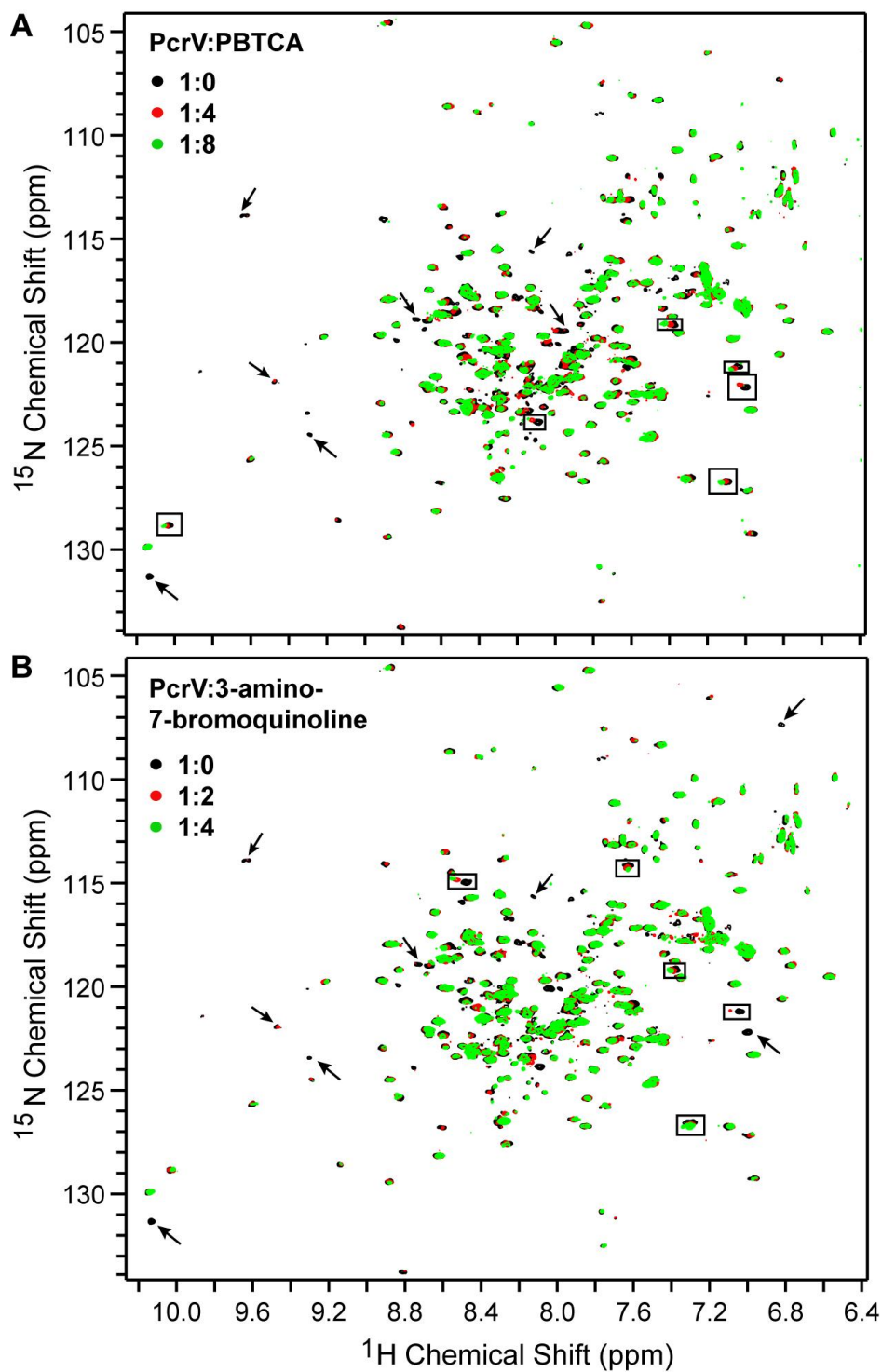
Figure 3-17. **¹⁵N NMR titrations of PcrV with PBTCA and 3-amino-7-bromoquinoline.** Overlay of three 2D ¹H-¹⁵N TROSY spectra of ¹⁵N-labeled PcrV titrated with increasing molar ratios of **(A)** PBTCA and **(B)** 3-amino-7-bromoquinoline. Residues displaying changes in peak positions are indicated with a box and residues with peak intensity reduction are marked with an arrow.

NMR spectroscopy is unparalleled in its ability to detect molecular interactions and characterize ligand binding at the atomic level over a wide range of affinities, and thus, valued highly in the drug discovery research (33-35). However, NMR characterization of molecular interactions of T3SS V-tip proteins has not yet been feasible largely due to lack of resonance assignments for this protein family. NMR assignment of backbone amides of the V-proteins is particularly challenging because of their large molecular weight (LcrV, 34 kDa and PcrV, 31 kDa), poor NMR data quality with significant signal overlap and peak broadening, and tendency to aggregate at high protein concentrations (PcrV at concentrations above 0.4 mM shows substantial peak broadening compared to a lower concentration sample). Perdeuteration of LcrV in combination with TROSY-based NMR experiments (36) resulted in sensitivity gains and resolution enhancement, allowing partial backbone amide assignment of LcrV **(Figure 3-8)**. However, increasing the total deuteration by replacement of non-deuterated D-glucose (U-$^{13}C_6$) as the carbon source to fully deuterated D-glucose (U-$^{13}C_6$; 1,2,3,4,5,6,6-d_7) did not further simplify spectral analysis to facilitate additional backbone amide assignments. Since methyl groups occur frequently in the hydrophobic cores of the protein and many protein-protein and protein-small molecules interactions often involve hydrophobic interfaces (28), we used sensitive ILV side chain methyl groups as alternative probes in the NMR studies of the V-proteins. Similar to the amide moieties, perturbations of ILV methyl groups can be used to determine ligand binding and map the interaction surfaces (28,37). Correlation of methyl ^1H-^1H NOEs with the distance information from our previously reported crystal structure of LcrV (11) enabled near complete methyl ILV assignments of LcrV **(Figure 3-5)**, which were successfully utilized here to characterize the interaction of LcrV with small molecules. The assignments provided herein are the only available NMR assignments for LcrV family of tip proteins and would assist in further molecular interaction studies of LcrV.

Analysis of the interaction of LcrV with PBTCA **(Figure 3-10)**, 5-hydroxyindole **(Figure 3-11)**, and 4-(4-methylpiperazino)aniline **(Figure 3-14)** reveal that all three hit fragments predominantly perturbed a similar surface on LcrV, bottom of the N-terminal globular domain. In addition to the N-

terminal domain, PBTCA also affected residues near top of the central coiled-coil of LcrV **(Figure 3-10)**. Since chemical shift perturbations can occur because of number of direct and indirect structural and dynamic effects, the other affected surface of LcrV could be a result of conformational changes induced upon the binding of PBTCA in the N-terminal globular domain or represent a secondary binding site for PBTCA. Even though with the limited number of assignments for PcrV, the observation that PBTCA and 3-amino-7-bromoquinoline affected residues near the N-terminal domain of PcrV highlight the importance of this region in small molecule binding **(Figure 3-16)**. Together, the data implies that surface near the bottom of the N-terminal domain of V-proteins represent a "hotspot" for binding small molecule fragments. Interestingly, the N-terminal globular domain of the V-proteins forms the basal structure of the needle tip complex and provides a crucial surface for the interaction with the YscF needle protein ([9],[31]). Furthermore, previous biophysical studies have shown the important role of the N-terminal domain in regulating the overall stability of the V-proteins ([38]). Additionally, the N-terminal domain has also been recognized to determine the efficiency of pore formation by acting as an assembly platform for the functional insertion of translocators into the host cell membrane ([39]). Given these significant roles of the N-terminal domain of V-proteins, our fragment binding site represents a biologically relevant hotspot that could be exploited to interfere with the assembly of the T3SS.

NMR titrations of small molecules with their cognate protein indicated interactions primarily in fast exchange regime with dissociation constants in the mM range **(Figure 3-9, Figure 3-11, Figure 3-14, and Figure 3-16)**. Such weak binding affinity by the compounds was expected given these are only small fragments (MW ~ 200 Da) that need to be optimized to generate a lead compound with higher affinity ([34],[35]). Fragment follow-up strategies such as, linking, merging, and/or chemical modification should help in the development of more potent compounds ([34]). Interestingly, the phenylene and the methyl piperazine moiety from our fragment hit 4-(4-methylpiperazino)aniline (**8**, **Figure 3-2**), upon combination with the quinoline moiety from our hit 3-amino-7-bromoquinoline (**11**, **Figure 3-2**) would provide a scaffold chemically related to INP1750 [5-nitro-7-((4-phenylpiperazine-1-yl-)methyl)quinolin-

8-ol], a compound that has been previously identified as a potent inhibitor of T3SS in *Y. pseudotuberculosis* and *C. trachomatis* (19). Although the specific molecular target of INP1750 in T3SS is currently unknown, presence of the phenyl + piperazine methyl + quinoline based scaffold could be associated with the inhibition of the tip protein of T3SS. It further demonstrates that the scaffolds identified in this work are promising starting structures that could be used in the development of specific inhibitors of the T3SS. Recently, 5-hydroxyindole (**4**, **Figure 3-2**) was shown to bind to SipD, the LcrV homolog in the *Salmonella* T3SS (22). Since most of the T3SS proteins are conserved across multiple bacterial pathogens, it is not surprising that the identified hits might interact with homologs of the target protein. In fact, T3SS inhibitors, such as acylhydrazides and thiazolidinones are effective against a wide range of Gram-negative bacterial pathogens, indicating a conserved T3SS protein as their likely target (15,40).

In summary, we have identified the first class of small molecules that bind to LcrV (PBTCA, 5-hydroxyindole, and 4-(4-methylpiperazino)aniline) and PcrV (PBTCA and 3-amino-7-bromoquinoline). These scaffolds provide a good starting point for fragment elaboration to enable the development of potent lead compounds targeting the T3SS.

3.6. References

1.   Coburn, B., Sekirov, I., and Finlay, B. B. (2007) Type III Secretion Systems and Disease. in *Clin. Microbiol. Rev.*
2.   Gu, L., Zhou, S., Zhu, L., Liang, C., and Chen, X. (2015) Small-Molecule Inhibitors of the Type III Secretion System. *Molecules* **20**, 17659-17674
3.   May, A. E., and Khosla, C. (2013) Discovery and Mechanism of Type III Secretion System Inhibitors. in *Isr. J. Chem.*
4.   McShan, A. C., and De Guzman, R. N. (2015) The bacterial type III secretion system as a target for developing new antibiotics. *Chem Biol Drug Des* **85**, 30-42
5.   Duncan, M. C., Linington, R. G., and Auerbuch, V. (2012) Chemical inhibitors of the type three secretion system: disarming bacterial pathogens. in *Antimicrobial Agents and Chemotherapy*

6.   Kubori, T., Matsushima, Y., Nakamura, D., Uralil, J., Lara-Tejero, M., Sukhan, A., Galan, J. E., and Aizawa, S. I. (1998) Supramolecular structure of the Salmonella typhimurium type III protein secretion system. in *Science*

7.   Chatterjee, S., Chaudhury, S., McShan, A. C., Kaur, K., and De Guzman, R. N. (2013) Structure and Biophysics of Type III Secretion in Bacteria. in *Biochemistry*

8.   Diepold, A., and Wagner, S. (2014) Assembly of the bacterial type III secretion machinery. *FEMS Microbiol. Rev.* **38**, 802-822

9.   Hiromi Sato, D. W. F. (2011) Multi-Functional Characteristics of the Pseudomonas aeruginosa Type III Needle-Tip Protein, PcrV; Comparison to Orthologs in other Gram-negative Bacteria. in *Frontiers in Microbiology*, Frontiers Media SA

10.   Cornelis, G. R. (2006) The type III secretion injectisome. in *Nat Rev Micro*, Nature Publishing Group

11.   Chaudhury, S., Battaile, K. P., Lovell, S., Plano, G. V., and De Guzman, R. N. (2013) Structure of the Yersinia pestistip protein LcrV refined to 1.65 Å resolution. in *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, International Union of Crystallography

12.   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. in *Structure*

13.   Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40

14.   Goure, J., Broz, P., Attree, O., Cornelis, G. R., and Attree, I. (2005) Protective anti-V antibodies inhibit Pseudomonas and Yersinia translocon assembly within host membranes. *J. Infect. Dis.* **192**, 218-225

15.   Kauppi, A. M., Nordfelth, R., Uvell, H., Wolf-Watz, H., and Elofsson, M. (2003) Targeting Bacterial Virulence. *Chemistry & Biology* **10**, 241-249

16.   Kim, O. K., Garrity-Ryan, L. K., Bartlett, V. J., Grier, M. C., Verma, A. K., Medjanis, G., Donatelli, J. E., Macone, A. B., Tanaka, S. K., Levy, S. B., and Alekshun, M. N. (2009) N-hydroxybenzimidazole inhibitors of the transcription factor LcrF in Yersinia: novel antivirulence agents. *J. Med. Chem.* **52**, 5626-5634

17.   Harmon, D. E., Davis, A. J., Castillo, C., and Mecsas, J. (2010) Identification and Characterization of Small-Molecule Inhibitors of Yop Translocation in Yersinia pseudotuberculosis. *Antimicrob. Agents Chemother.* **54**, 3241-3254

18.   Swietnicki, W., Carmany, D., Retford, M., Guelta, M., Dorsey, R., Bozue, J., Lee, M. S., and Olson, M. A. (2011) Identification of Small-Molecule Inhibitors of Yersinia pestis Type III Secretion System YscN ATPase. *PLoS One* **6**, e19716

19.   Enquist, P.-A., Gylfe, Å., Hägglund, U., Lindström, P., Norberg-Scherman, H., Sundin, C., and Elofsson, M. (2012) Derivatives of 8-hydroxyquinoline—antibacterial agents that target intra- and extracellular Gram-negative pathogens. *Bioorganic & Medicinal Chemistry Letters* **22**, 3550-3553

20.   Aiello, D., Williams, J. D., Majgier-Baranowska, H., Patel, I., Peet, N. P., Huang, J., Lory, S., Bowlin, T. L., and Moir, D. T. (2010) Discovery and characterization of inhibitors of Pseudomonas aeruginosa type III secretion. *Antimicrob. Agents Chemother.* **54**, 1988-1999

21.   Chaudhury, S., Nordhues, B. A., Kaur, K., Zhang, N., and De Guzman, R. N. (2015) Nuclear Magnetic Resonance Characterization of the Type III Secretion System Tip Chaperone Protein PcrG of Pseudomonas aeruginosa. *Biochemistry*, 151021074758005

22.   McShan, A. C., Anbanandam, A., Patnaik, S., and De Guzman, R. N. (2016) Characterization of the Binding of Hydroxyindole, Indoleacetic acid, and Morpholinoaniline to the Salmonella Type III Secretion System Proteins SipD and SipB. *ChemMedChem*

23.   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. in *J Biomol NMR*

24. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. in *Methods Mol. Biol.*

25. Xiao, Y., Lee, T., Latham, M. P., and Warner, L. R. (2014) Phosphorylation releases constraints to domain motion in ERK2.

26. McShan, A. C., Kaur, K., Chatterjee, S., Knight, K. M., and De Guzman, R. N. (2016) NMR Identification of the Binding Surfaces Involved in the Salmonella and Shigella Type III Secretion Tip-Translocon Protein-Protein Interactions. *Proteins*

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. in *Nat Struct Biol*

28. Tugarinov, V., and Kay, L. E. (2005) Methyl Groups as Probes of Structure and Dynamics in NMR Studies of High-Molecular-Weight Proteins. in *ChemBioChem*

29. Goure, J., Pastor, A., Faudry, E., Chabert, J., Dessen, 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. *Infection and Immunity*

30. Mueller, C. A., Broz, P., and Cornelis, G. R. (2008) The type III secretion system tip complex and translocon. *Mol. Microbiol.* **68**, 1085-1095

31. Mueller, C. A., Broz, P., Müller, 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

32. Keyser, P., Elofsson, M., Rosell, S., and Wolf-Watz, H. (2008) Virulence blockers as alternatives to antibiotics: type III secretion inhibitors against Gram-negative bacteria. in *J Intern Med*

33. Barrett, P. J., Chen, J., Cho, M.-K., Kim, J.-H., Lu, Z., Mathew, S., Peng, D., Song, Y., Van Horn, W. D., Zhuang, T., Sönnichsen, F. D., and Sanders, C. R. (2013) The Quiet Renaissance of Protein Nuclear Magnetic Resonance. in *Biochemistry*

34. Jahnke, W., and Widmer, H. (2004) Protein NMR in biomedical research. in *Cell. Mol. Life Sci.*

35. Pellecchia, M., Bertini, I., Cowburn, D., Dalvit, C., Giralt, E., Jahnke, W., James, T. L., Homans, S. W., Kessler, H., Luchinat, C., Meyer, B., Oschkinat, H., Peng, J., Schwalbe, H., and Siegal, G. (2008) Perspectives on NMR in drug discovery: a technique comes of age. *Nat Rev Drug Discov* **7**, 738-745

36. Fernandez, C. (2003) TROSY in NMR studies of the structure and function of large biological macromolecules. in *Current Opinion in Structural Biology*

37. Hajduk, P. J., Augeri, D. J., Mack, J., and Mendoza, R. (2000) NMR-Based Screening of Proteins Containing 13C-Labeled Methyl Groups - Journal of the American Chemical Society (ACS Publications). in *Journal of American Chemical Society*

38. Espina, M., Ausar, S. F., Middaugh, C. R., Baxter, M. A., Picking, W. D., and Picking, W. L. (2007) Conformational stability and differential structural analysis of LcrV, PcrV, BipD, and SipD from type III secretion systems. *Protein Sci.* **16**, 704-714

39. Broz, P., Mueller, C. A., Müller, S. A., Philippsen, A., Sorg, I., Engel, A., and Cornelis, G. R. (2007) Function and molecular architecture of the Yersinia injectisome tip complex. *Mol. Microbiol.* **65**, 1311-1320

40. Felise, H. B., Nguyen, H. V., Pfuetzner, R. A., Barry, K. C., Jackson, S. R., Blanc, M.-P., Bronstein, P. A., Kline, T., and Miller, S. I. (2008) An inhibitor of gram-negative bacterial virulence protein secretion. *Cell Host Microbe* **4**, 325-336

This page intentionally left blank.

Chapter 4: Concluding Remarks and Future Directions for T3SS Project

The T3SS is critical in the pathogenesis of many Gram-negative bacteria that infect humans (1-3). The structural component of the T3SS is a needle apparatus that function as a syringe to inject effector proteins from bacterial cytoplasm directly into the target eukaryotic cells. The T3SS needle apparatus is assembled from over 20 different proteins and is composed of a base that spans the bacterial membranes, an extracellular needle, a tip protein complex, and a translocon that punctures a hole in the host cell membrane (4). Since the visualization of this sophisticated bacterial nanomachine almost two decades ago, remarkable progress has been made towards determining high-resolution structures of various components of the needle apparatus (1,3-8). However, the assembly mechanism, specifically, the atomic details of protein-protein interactions involved in the assembly of the needle apparatus remain poorly understood. High-resolution models of the base and the needle are now available, but, much less is known regarding the tip-translocon protein-protein interactions (7,8). Therefore, the first aim of the T3SS research described in this dissertation was to fill this gap by gaining insights into the interactions of the *Shigella* tip protein IpaD and the major translocon protein IpaB. Furthermore, because the T3SS is crucial in virulence, surface-exposed, and highly conserved in structure and function among different pathogenic bacteria, disrupting the needle apparatus of T3SS is an exciting target for the development of novel anti-infectives (2,9,10). Thus, the second aim of the T3SS research presented in this dissertation was directed at identifying and characterizing small molecule binders of the T3SS proteins that could be used as potential starting structures in the drug design of inhibitors of T3SS.

4.1. IpaD-IpaB tip-translocon protein-protein interaction

4.1.1. Key findings and significance

The translocon is assembled from the major and the minor translocon proteins. IpaB, the major translocon protein of *Shigella*, is an integral membrane protein that contains an N-terminal cytosolic ectodomain, a central conserved hydrophobic region with two transmembrane helices, and a C-terminal amphipathic region ([11]). The transmembrane and the C-terminal amphipathic regions have been shown to be important for the membrane insertion of the major translocon proteins ([11],[12]). Prior to this work, the role of the cytosolic N-terminal ectodomain in the assembly of the translocon was unclear. Based on the topology of IpaB, we hypothesized that the soluble N-terminal ectodomain could be responsible for the attachment of the translocon to the tip complex. NMR methods were used to study this protein-protein interaction, as summarized in Chapter 2. NMR chemical shift mapping revealed a direct interaction between the tip protein IpaD and the N-terminal ectodomain of the major translocon protein IpaB. Changes in the NMR peaks brought about by IpaD-IpaB complex formation were primarily in the intermediate exchange regime on NMR timescale, indicating a weak 'µM' range binding affinity. Therefore, I used sensitive PRE methods utilizing spin labels to gain further insights into this weak interaction ([13],[14]). Results of NMR chemical shift perturbations complemented PRE studies and they together indicated that the distal region of the tip protein IpaD (the mixed α-β domain) is the primary binding site for the N-terminal ectodomain of IpaB (Chapter 2).

Such detailed characterization of the surfaces involved in the tip-translocon protein-protein interactions on a per-residue resolution level was unavailable for any type III secretion system. This work has expanded our current knowledge about the assembly of the needle apparatus of the T3SS. Our results have further provided structural basis for understanding previous mutational studies that showed deletion

of the distal region of IpaD abolishes the localization of IpaB on the needle tip ([15]). Additionally, our results are in agreement with the recent model of the *Shigella* tip complex derived by electron microscopy that displays the distal region of IpaD points outside in the assembled tip, and, thus available to mediate tip-translocon protein-protein interactions ([16]). However, in contrast to the previous fluorescence polarization studies indicating essential role of bile salt deoxycholate (DOC) in IpaD-IpaB binding ([17]), our IpaD-IpaB interaction studies in the absence or presence of DOC showed no major alteration in the protein binding affinity or the interaction surfaces.

4.1.2. *Future directions to understand the assembly of the needle apparatus*

The tip-translocon interaction studies presented in this dissertation combined with the previous data on the needle-tip interaction ([18],[19]) put forward a needle-tip-translocon interaction model centered on the tip protein IpaD, where the proximal end of the IpaD (bottom of the coiled-coil) binds with the needle protein, and the distal domain of IpaD interacts with the cytosolic ectodomain of IpaB (Chapter 2). However, to gain better appreciation of the intricate assembly mechanism of the needle apparatus and understand the role of the minor translocon protein in assembling a functional translocon, further testing is mandatory.

Firstly, although the binding surfaces on IpaD that are involved in the interaction with IpaB have been well characterized, the reverse is not explicitly known. PRE spin labeling on IpaB suggested region near the top of the trimeric coiled-coil of the N-terminal ectodomain of IpaB as important for the binding of IpaD. Future experiments following reverse labeling, *i.e.*, isotopically labeled IpaB titrated with unlabeled or spin-labeled IpaD would help in clear elucidation of the interaction surfaces of IpaD on the N-terminal domain of IpaB. Towards this direction, given the IpaD-IpaB interaction is weak, we used sensitive side chain methyl isoleucine labeling ([20],[21]) and tested various N-terminal constructs of IpaB to

find an ideal construct for NMR studies. The longer N-terminal constructs, IpaB residues 9-226 and IpaB residues 76-308, containing flexible domains showed poor NMR data quality **(Figure 4-1A, B)**. In contrast, IpaB residues 74-224, a protease-resistant fragment lacking disordered regions, gave an ideal NMR spectrum with all the 14 expected Ile resonances **(Figure 4-1C)**. These findings suggest IpaB residues 74-224 is an ideal construct that could be utilized for reverse NMR binding studies to characterize the full spectrum of IpaD and N-terminal ectodomain IpaB interactions.

Secondly, a major concern of using the N-terminal ectodomain of the major translocon protein IpaB in the tip interaction studies is that this fragment, although essential for IpaD binding, may not contain the entire surface of IpaB required for the IpaD binding. It is important to test how the full-length IpaB interacts with IpaD and vice-versa to understand the translocon assembly. Additionally, both major and minor translocon proteins are required to assemble a functional translocon (22,23). To gain insights into the interactions of an assembled translocon, it would be of immense importance to characterize the protein-protein interactions of the minor translocon protein (IpaC in *Shigella*) in addition to the major translocon protein IpaB. This is challenging since the translocon proteins, IpaB and IpaC, are integral membrane proteins (4). However, full-length translocon proteins can be expressed with their cognate chaperone proteins (IpgC in *Shigella*) to obtain a soluble complex (24). Studies have shown that a very small region (the extreme N-terminus) of the translocon protein is actually needed to mediate interactions with the chaperone (25,26), justifying the use of chaperone-bound soluble translocon complexes in the structural characterization and protein-protein interaction studies. Towards this goal, I subcloned, co-expressed, and co-purified full-length major translocon protein IpaB with the chaperone IpgC, and full-length minor translocon protein IpaC with IpgC. Because of the poor expression of the *Shigella* proteins, I also subcloned, expressed, and purified *Yersinia* full-length major and minor translocon proteins, YopB and YopD, respectively, with the chaperone LcrH. Unfortunately, the *Yersinia* proteins were not stable enough and degraded upon purification. Preliminary structural and biophysical characterization was carried out on *Shigella* translocon proteins using CD spectroscopy, size exclusion chromatography, and
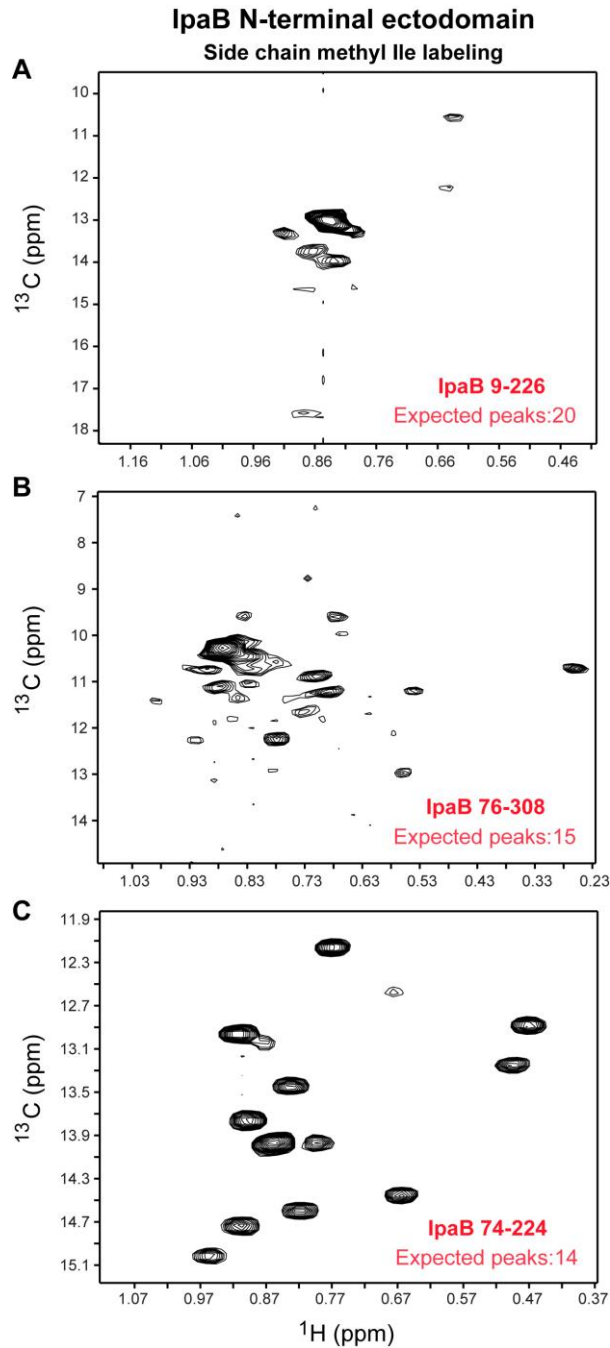
Figure 4-1. Two-dimensional ¹H-¹³C HSQC spectra of side chain methyl Ile-labeled IpaB N-terminal constructs. (**A**) IpaB residues 9-226, (**B**) IpaB residues 76-308, and (**C**) IpaB residues 74-224.

NMR spectroscopy. However, the full-length translocon-chaperone complexes, IpaB-IpgC and IpaC-IpgC, gave poor quality NMR data **(Figure 4-2)**. The translocon proteins can be separated from the

translocon-chaperone complexes using detergents, urea or pH treatment ([27],[28]). Upon pH change, most of the IpaC degraded, but IpaB could be stably dissociated from IpaB-IpgC complex and could be maintained in OPOE/LDAO detergents. Although ^{13}C-methyl isoleucine labeled full-length IpaB showed fewer peaks than expected, the data look promising with sharp and dispersed peaks (**Figure 4-2C**). Future experimental optimization, better NMR labeling schemes (side chain methionine labeling), deuteration, different detergent combinations, use of nanodiscs, and other techniques, such as, EPR and cryo-EM may prove helpful in the structural analysis and protein-protein interaction studies of the translocon proteins.
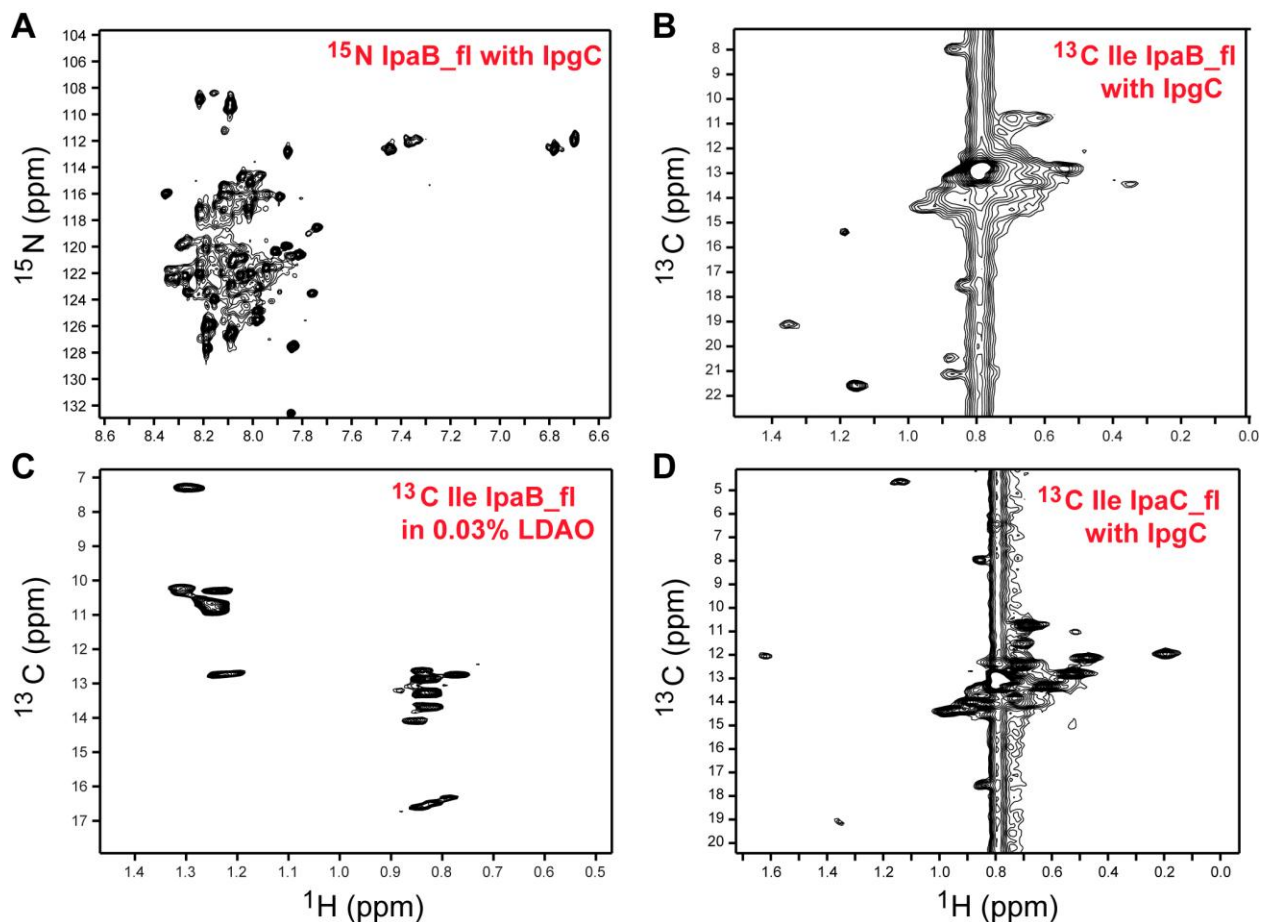
Figure 4-2. NMR spectroscopy on full-length translocon proteins from *Shigella flexneri* T3SS. (A) 2D ^{1}H-^{15}N TROSY spectrum of IpaB-IpgC complex, 2D ^{1}H-^{13}C HSQC spectra of Ile-labeled **(B)** IpaB-IpgC complex, **(C)** free IpaB in 0.03% LDAO, and **(D)** IpaC-IpgC complex.

4.2. Small-molecule binders of the T3SS tip proteins LcrV and PcrV

4.2.1. Key findings, significance, and future directions

Antibiotic resistance in Gram-negative bacteria is a looming public health threat that requires new targets for the development of novel anti-bacterials (29). Since the T3SS is essential for virulence and colonization within a host, exposed on the bacterial surface, and present only in the pathogens, disrupting the assembly of the T3SS is an attractive target for developing novel anti-virulence drugs (2,9,10). In this body of work, SPR-based fragment screening identified novel small molecules that bound to tip proteins LcrV (*Yersinia pestis*) and PcrV (*Pseudomonas aeruginosa*). Binding of the fragment hit to their cognate protein was validated and characterized using NMR spectroscopy, described in Chapter 3. Analysis of ^{15}N and ILV methyl-based NMR titrations of the tip proteins with the hit fragments revealed these small molecules, although chemically distinct, perturbed a similar surface – the N-terminal globular domain of the V-tip proteins (30), suggesting this surface to be a potential hotspot for drug interaction. Chemical inhibitors of the T3SS have been identified, but the specific molecular targets in the T3SS for most of the compounds are not known (9,31). The novel scaffolds that we have identified are the first small molecules shown to directly bind to the virulent tip proteins LcrV and PcrV and could potentially be developed into T3SS protein-protein interaction inhibitors.

Our compounds showed weak 'mM' range binding affinities to the tip proteins LcrV and PcrV. However, *in vivo* inhibition likely requires tighter binding compounds. Future fragment building studies would be needed to design high-affinity specific binders of the tip proteins (32). The potent compounds should subsequently be tested for their ability to disrupt protein-protein interactions and inhibit the T3SS in the *in vitro* binding assays and functional bacterial invasion assay.

4.2.2. Implications of the ILV methyl assignments of LcrV

NMR characterization of the binding of small molecules to the tip proteins LcrV and PcrV require resonance assignments (backbone amide or side chain methyl ILV) to identify the surface(s) of the tip proteins involved in the interaction. LcrV and PcrV are large proteins (~35 kDa) and tend to aggregate at high protein concentrations, making the NMR studies particularly challenging. Side chain ILV methyl groups with favorable NMR relaxation properties and high sensitivity offer phenomenal probes to study structure, dynamics, and protein-ligand interactions ([21]). Using mutagenesis, perdeuteration, ^{1}H-^{1}H NOEs from 3D-NOESY experiments, and a high-resolution crystal structure of LcrV, near complete ILV methyl assignments were obtained for LcrV (Chapter 3). The NMR assignments presented in this body of work are the only available assignments from the entire *Yersinia* family of the T3SS tip proteins. These assignments were successfully used in the identification and characterization of the LcrV-small molecule interaction surfaces and would facilitate further molecular interaction studies of the tip protein LcrV.

Prior to the secretion of LcrV to assemble the *Yersinia* T3SS tip complex, it remains bound to a small chaperone protein LcrG in the bacterial cytoplasm. The tip-chaperone interactions in the *Yersinia* family of T3SS are important for the regulation of effector secretion ([33],[34]). LcrV-LcrG protein-protein interactions have been studied previously using NMR ([33],[35]), however, due to the lack of resonance assignments, interaction surfaces of the chaperone protein LcrG on the tip protein LcrV could not be determined. The ILV methyl assignments of LcrV, presented herein, allowed identification of the LcrV-LcrG binding surfaces. Analysis of the ILV methyl titrations revealed that a large surface on LcrV was perturbed upon the binding of LcrG **(Figure 4-3)**. This is not very surprising since LcrG is a partially folded and an intrinsically flexible protein ([35]), and flexible proteins usually provide a large surface area for interaction compared to globular proteins ([36]). Furthermore, previous interaction studies have shown that entire length of the chaperone protein LcrG is involved in binding to the tip protein LcrV and even

upon the LcrV-LcrG complex formation, LcrG remains flexible and does not undergo a disorder-to-order transition ([33],[35]). Taken together, it seems likely that the binding of LcrG to LcrV may involve a large interaction surface on LcrV. These results highlight the significance of the ILV methyl assignments of LcrV to understand and characterize the important molecular interactions of the T3SS.
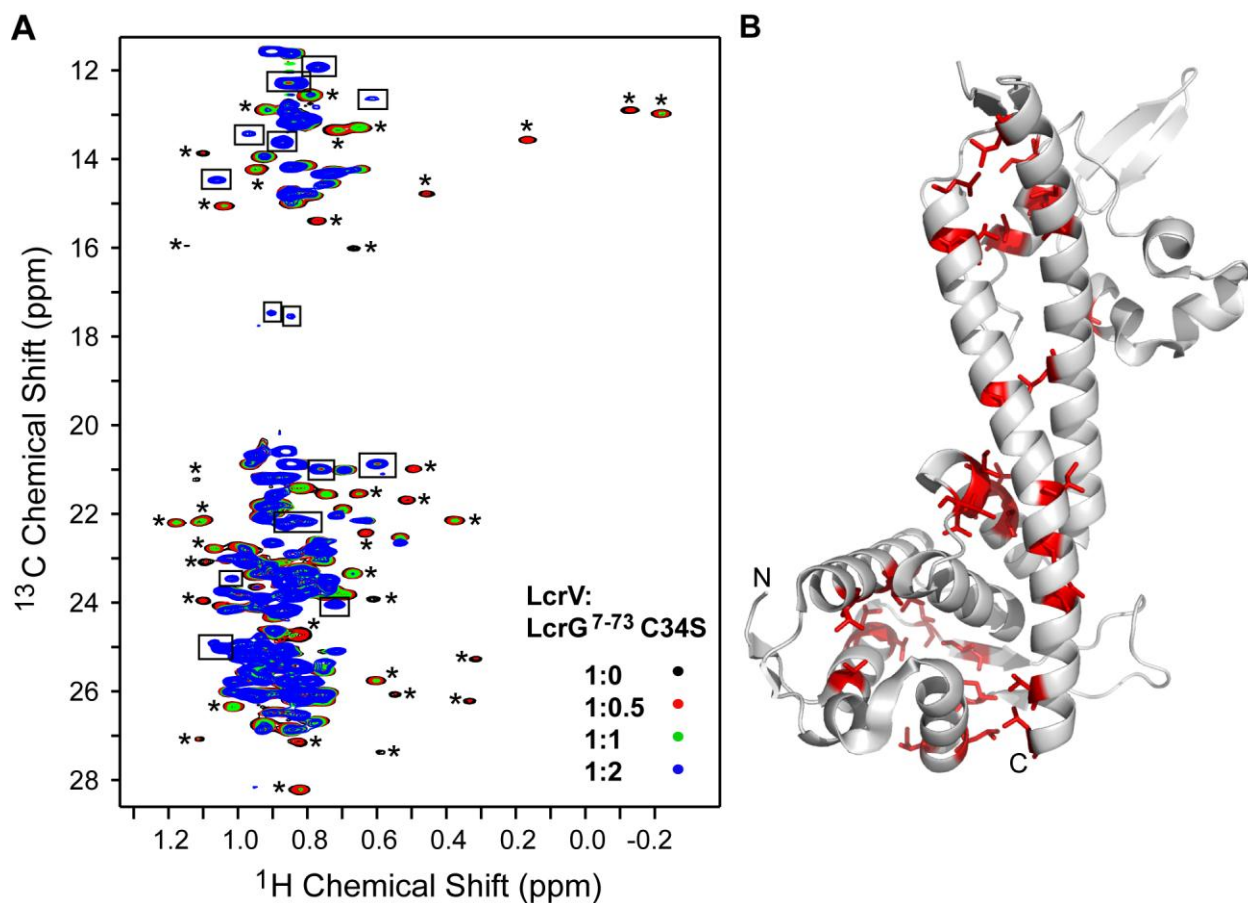
Figure 4-3. ILV titrations of tip protein LcrV with chaperone LcrG. (**A**) Overlay of four 2D ^{1}H-^{13}C HSQC spectra of ILV-labeled LcrV titrated with increasing mole ratios of LcrG. Affected residues displaying reduction in peak intensity are marked with an asterisk and new peaks appearing upon complex formation are shown with a box. (**B**) ILV resonances of LcrV affected upon the binding of LcrG mapped onto the crystal structure of LcrV, highlighted in red.

To better appreciate the structural basis of the tip-chaperone interactions of the T3SS, NMR titrations have also been performed on the *Pseudomonas aeruginosa* tip protein PcrV and the cognate chaperone PcrG, the homologs of *Yersinia pestis* LcrV and LcrG, respectively ([37]) **(Figure 4-4)**. Both side chain methyl ILV and backbone amide titrations show a large number of PcrV resonances being affected upon the binding of PcrG, similar to the LcrV-LcrG interactions. While some resonances display peak broadening, others show changes in the peak position. Additionally, new complex peaks upon the binding of PcrG are also evident **(Figure 4-4)**. Such complex NMR exchange behavior suggests that the chaperone binding is likely accompanied with conformational changes in the tip protein PcrV. Using the available ILV assignments of PcrV, the identified affected residues were mapped onto the bottom and top of the coiled-coil domain **(Figure 4-4B)**. A detailed characterization of this intricate interaction would require NMR (amide and ILV) assignments of the free and chaperone-bound PcrV (new peaks upon complex formation). In this direction, high-quality 3D-NOESY NMR data on perdeuterated methyl-ILV labeled PcrV has been collected **(Figure 4-5A)**. Given the large sizes of the tip proteins, perdeuteration was essential to observe the NOEs ([38],[39]). For comparison, a non-deuterated sample displayed broad resonances with minimal NOEs **(Figure 4-5B)**. However, in the absence of a high-resolution structure of PcrV, it is challenging to complete the methyl ILV assignments utilizing the NOESY data only. In future, the NOE restraints may be used in combination with computational modeling tools to build a de novo structure of PcrV.
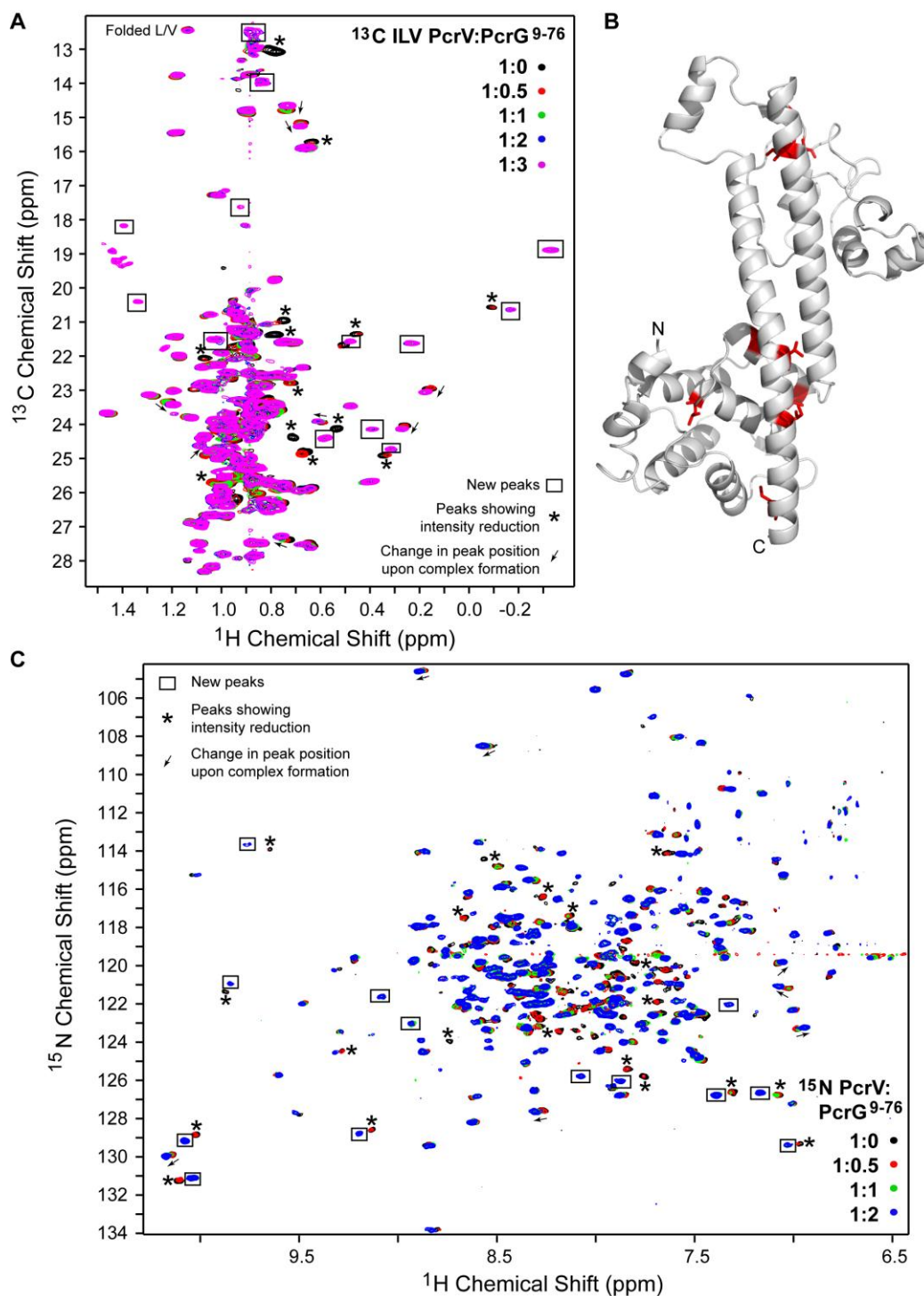
Figure 4-4. NMR determination of PcrV-PcrG tip-chaperone protein-protein interactions from the *Pseudomonas* **T3SS.** (**A**) ILV titrations of PcrV and PcrG. Affected residues are marked with an asterisk (peak intensity reduction), or arrow (change in peak position). New peaks are shown with a box. (**B**) ILV resonances affected upon PcrV-PcrG complex formation mapped onto the I-TASSER model of PcrV. (**C**) ^{15}N amide titrations of PcrV and PcrG. Affected resonances are marked as in **A**.
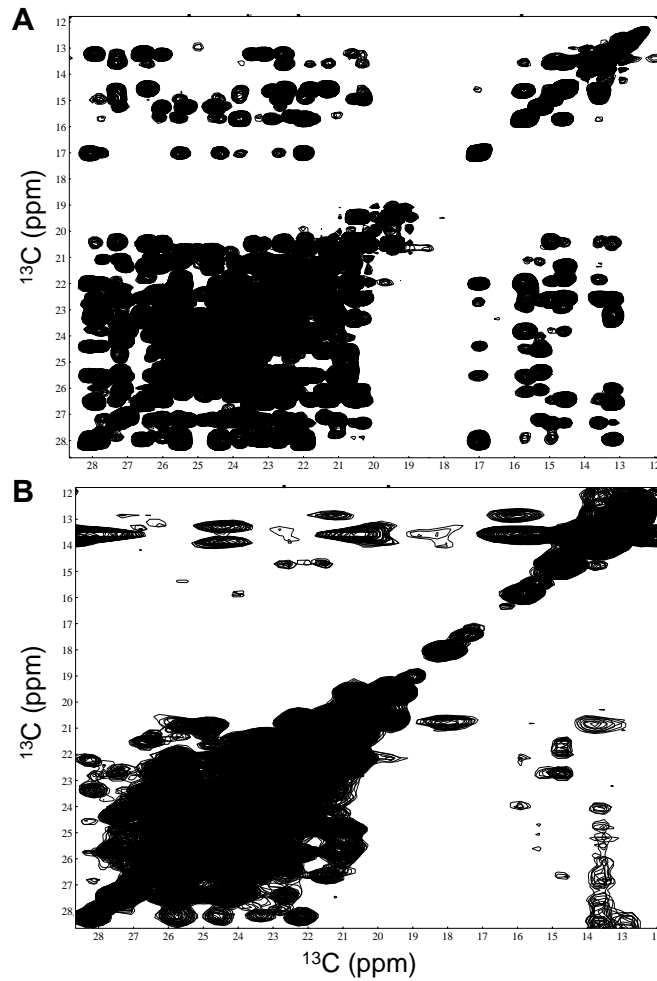
Figure 4-5. Perdeuteration is necessary with large proteins to observe NOEs. (A) A perdeuterated ^{13}C-ILV labeled sample of PcrV at 0.3 mM concentration, lyophilized and re-suspended in 100% D_2O, and **(B)** A non-deuterated ^{13}C-ILV labeled sample of PcrV at 0.3 mM concentration, lyophilized and re-suspended in 100% D_2O.

4.3. References

1.   Cornelis, G. R. (2006) The type III secretion injectisome. in *Nat Rev Micro*, Nature Publishing Group
2.   Coburn, B., Sekirov, I., and Finlay, B. B. (2007) Type III Secretion Systems and Disease. in *Clin. Microbiol. Rev.*

3.   Galán, J. E., Lara-Tejero, M., Marlovits, T. C., and Wagner, S. (2014) Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells. in *Annu. Rev. Microbiol.*

4.   Chatterjee, S., Chaudhury, S., McShan, A. C., Kaur, K., and De Guzman, R. N. (2013) Structure and Biophysics of Type III Secretion in Bacteria. in *Biochemistry*

5.   Kubori, T., Matsushima, Y., Nakamura, D., Uralil, J., Lara-Tejero, M., Sukhan, A., Galan, J. E., and Aizawa, S. I. (1998) Supramolecular structure of the Salmonella typhimurium type III protein secretion system. in *Science*

6.   Schraidt, O., and Marlovits, T. C. (2011) Three-dimensional model of Salmonella's needle complex at subnanometer resolution. in *Science*

7.   Kosarewicz, A., Konigsmaier, L., and Marlovits, T. C. (2012) The blueprint of the type-3 injectisome. in *Philosophical Transactions of the Royal Society B: Biological Sciences*

8.   Demers, J.-P., Habenstein, B., Loquet, A., Kumar Vasa, S., Giller, K., Becker, S., Baker, D., Lange, A., and Sgourakis, N. G. (2014) High-resolution structure of the Shigella type-III secretion needle by solid-state NMR and cryo-electron microscopy. in *Nat Commun*

9.   Keyser, P., Elofsson, M., Rosell, S., and Wolf-Watz, H. (2008) Virulence blockers as alternatives to antibiotics: type III secretion inhibitors against Gram-negative bacteria. in *J Intern Med*

10.   Duncan, M. C., Linington, R. G., and Auerbuch, V. (2012) Chemical inhibitors of the type three secretion system: disarming bacterial pathogens. in *Antimicrobial Agents and Chemotherapy*

11.   Hume, P. J., McGhie, E. J., Hayward, R. D., and Koronakis, V. (2003) The purified Shigella IpaB and Salmonella SipB translocators share biochemical properties and membrane topology. *Mol. Microbiol.* **49**, 425-439

12.   McGhie, E. J. E., Hume, P. J. P., Hayward, R. D. R., Torres, J. J., and Koronakis, V. V. (2002) Topology of the Salmonella invasion protein SipB in a model bilayer. *Mol. Microbiol.* **44**, 1309-1321

13.   Rodriguez-Castañeda, F., Haberz, P., Leonov, A., and Griesinger, C. (2006) Paramagnetic tagging of diamagnetic proteins for solution NMR. *Magn. Reson. Chem.* **44**, S10-S16

14.   Otting, G. (2010) Protein NMR Using Paramagnetic Ions. in *Annu. Rev. Biophys.*

15.   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) Self-chaperoning of the Type III Secretion System Needle Tip Proteins IpaD and BipD. in *J Biol Chem*

16.   Cheung, M., Shen, D.-K., Makino, F., Kato, T., Roehrich, A. D., Martinez-Argudo, I., Walker, M. L., Murillo, I., Liu, X., Pain, M., Brown, J., Frazer, G., Mantell, J., Mina, P., Todd, T., Sessions, R. B., Namba, K., and Blocker, A. J. (2015) Three-dimensional electron microscopy reconstruction and cysteine-mediated crosslinking provide a model of the type III secretion system needle tip complex. *Mol. Microbiol.* **95**, 31-50

17.   Dickenson, N. E., Arizmendi, O., Patil, M. K., Toth, I., Ronald T, Middaugh, C. R., Picking, W. D., and Picking, W. L. (2013) N-Terminus of IpaB Provides a Potential Anchor to the ShigellaType III Secretion System Tip Complex Protein IpaD. *Biochemistry* **52**, 8790-8799

18.   Rathinavelan, T., Lara-Tejero, M., and Lefebre, M. (2014) NMR Model of PrgI–SipD Interaction and Its Implications in the Needle-Tip Assembly of the Salmonella Type III Secretion System. *Journal of molecular biology*

19.   Rathinavelan, T., Tang, C., and De Guzman, R. N. (2011) Characterization of the interaction between the Salmonella type III secretion system tip protein SipD and the needle protein PrgI by paramagnetic relaxation enhancement. in *J Biol Chem*

20.   Kay, L. E. (2011) Solution NMR spectroscopy of supra-molecular systems, why bother? A methyl-TROSY view. *Journal of Magnetic Resonance* **210**, 159-170

21.   Tugarinov, V., and Kay, L. E. (2005) Methyl Groups as Probes of Structure and Dynamics in NMR Studies of High-Molecular-Weight Proteins. in *ChemBioChem*

22.   Faudry, E., Vernier, G., Neumann, E., Forge, V., and Attree, I. (2006) Synergistic pore formation by type III toxin translocators of Pseudomonas aeruginosa. *Biochemistry* **45**, 8117-8123

23. Myeni, S. K., Wang, L., and Zhou, D. (2013) SipB-SipC Complex Is Essential for Translocon Formation. *PLoS One* **8**, e60499

24. Birket, S. E., Harrington, A. T., Espina, M., Smith, N. D., Terry, C. M., Darboe, N., Markham, A. P., Middaugh, C. R., Picking, W. L., and Picking, W. D. (2007) Preparation and characterization of translocator/chaperone complexes and their component proteins from Shigella flexneri. *Biochemistry* **46**, 8128-8137

25. Lokareddy, R. K., Lunelli, M., Eilers, B., Wolter, V., and Kolbe, M. (2010) Combination of Two Separate Binding Domains Defines Stoichiometry between Type III Secretion System Chaperone IpgC and Translocator Protein IpaB. *Journal of Biological Chemistry* **285**, 39965-39975

26. Lunelli, M., Lokareddy, R. K., Zychlinsky, A., and Kolbe, M. (2009) IpaB-IpgC interaction defines binding motif for type III secretion translocator. *Proceedings of the National Academy of Sciences* **106**, 9661-9666

27. Chen, X., Choudhari, S. P., Kumar, P., Toth, R. T., Kim, J. H., Van Roosmalen, M. L., Leenhouts, K., Middaugh, C. R., Picking, W. L., and Picking, W. D. (2015) Biophysical Characterization of the Type III Secretion System Translocator Proteins and the Translocator Proteins Attached to Bacterium-Like Particles. *J Pharm Sci*

28. Romano, F. B., Rossi, K. C., Savva, C. G., and Holzenburg, A. (2011) Efficient Isolation of Pseudomonas aeruginosa Type III Secretion Translocators and Assembly of Heteromeric Transmembrane Pores in Model Membranes - Biochemistry (ACS Publications). *Biochemistry*

29. CDC. (2016) Antimicrobial resistance.

30. Chaudhury, S., Battaile, K. P., Lovell, S., Plano, G. V., and De Guzman, R. N. (2013) Structure of the Yersinia pestistip protein LcrV refined to 1.65 Å resolution. in *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, International Union of Crystallography

31. McShan, A. C., and De Guzman, R. N. (2015) The bacterial type III secretion system as a target for developing new antibiotics. in *Chem Biol Drug Des*

32. Scott, D. E., Ehebauer, M. T., Pukala, T., Marsh, M., Blundell, T. L., Venkitaraman, A. R., Abell, C., and Hyvönen, M. (2013) Using a fragment-based approach to target protein-protein interactions. *ChemBioChem* **14**, 332-342

33. De Guzman, R. N., Egan, S. M., Davido, D., Hefty, S., Richter, M., and Weis, D. (2013) Structural Studies of Chaperones and Chaperone-Tip interactions from the type III secretion systems of Yersinia and Pseudomonas.

34. DeBord, K. L., Lee, V. T., and Schneewind, O. (2001) Roles of LcrG and LcrV during Type III Targeting of Effector Yops by Yersinia enterocolitica. *J. Bacteriol.* **183**, 4588-4598

35. Chaudhury, S., de Azevedo Souza, C., Plano, G. V., and De Guzman, R. N. (2015) The LcrG Tip Chaperone Protein of the Yersinia pestis Type III Secretion System Is Partially Folded. *J Mol Biol* **427**, 3096-3109

36. Babu, M. M., van der Lee, R., de Groot, N. S., and Gsponer, J. (2011) Intrinsically disordered proteins: regulation and disease. *Current Opinion in Structural Biology* **21**, 432-440

37. Chaudhury, S., Nordhues, B. A., Kaur, K., Zhang, N., and De Guzman, R. N. (2015) Nuclear Magnetic Resonance Characterization of the Type III Secretion System Tip Chaperone Protein PcrG of Pseudomonas aeruginosa. in *Biochemistry*, American Chemical Society

38. Barrett, P. J., Chen, J., Cho, M.-K., Kim, J.-H., Lu, Z., Mathew, S., Peng, D., Song, Y., Van Horn, W. D., Zhuang, T., Sönnichsen, F. D., and Sanders, C. R. (2013) The Quiet Renaissance of Protein Nuclear Magnetic Resonance. in *Biochemistry*

39. Tzeng, S.-R., Pai, M.-T., and Kalodimos, C. G. (2012) NMR studies of large protein systems. *Methods Mol. Biol.* **831**, 133-140

This page intentionally left blank.

Chapter 5: NMR Identification of a New SUMO-Interacting Motif in SUMO E3 Ligase PIASy

Kawaljit Kaur, Yoshiaki Azuma, and Roberto N. De Guzman

Department of Molecular Biosciences, University of Kansas, Lawrence, KS 66045, USA

5.1. Abstract

SUMO conjugation is a reversible post-translational modification process implicated in the regulation of gene transcription, DNA repair, and cell cycle. SUMOylation depends on the sequential activities of E1 activating, E2 conjugating, and E3 ligating enzymes. SUMO E3 ligases enhance transfer of SUMO from the charged E2 enzyme to the substrate. PIASy, a member of the Siz/PIAS RING family of SUMO E3 ligases, is essential for mitotic SUMOylation of chromosomal proteins, yet the determinants required for SUMO conjugation by PIASy are not well defined. Here, using NMR chemical shift mapping and mutational analysis, we identified a new SUMO-interacting motif (SIM) in PIASy. *In vitro* SUMOylation assays illustrated significant role of the new SIM for a fully active SUMO E3 ligase. Our results provide novel insights into the mechanism of PIASy-mediated SUMO conjugation. With the identification of a new SIM, PIASy adds to the growing list of SUMO E3 ligases containing tandem SIMs important for the E3 ligase activity.

5.2. Introduction

Covalent attachment of small ubiquitin-like modifier (SUMO) to protein targets is an important post-translational modification that regulates multiple cellular processes, including transcription, nuclear transport, DNA repair, and chromosome segregation ([1],[2]). Vertebrates express at least three functional SUMO isoforms, namely, SUMO-1, SUMO-2, and SUMO-3 that can conjugate to other proteins. SUMO-1 shares less than 50% sequence identity to other isoforms, while SUMO-2 and SUMO-3 are virtually equivalent with 97% sequence identity and are commonly referred as SUMO-2/3 ([1-3]). Similar to ubiquitylation, SUMO proteins are typically conjugated to target lysine residue of the substrate by the sequential action of three enzymes: an E1 activating enzyme (Uba2/Aos1), an E2 conjugating enzyme

(Ubc9), and an E3 ligase. A SUMO E3 ligase stabilizes and optimally positions the thioester charged E2~SUMO complex to catalyze the transfer of SUMO from charged E2 to the substrate (2-5).

All SUMOylation reactions utilize the same E1 and E2 enzymes, however, several different classes of SUMO E3 ligases have been identified (2,3). Most known SUMO E3 ligases, including Siz and PIAS proteins, RanBP2, polycomb protein Pc2, SLX4, and recently characterized ZNF451 possess one or more SUMO-interacting motifs (SIMs) to mediate non-covalent interactions with SUMO (2,3,5-9). In most cases, SIMs are required for the E3 ligase activity and play an important role in positioning the donor SUMO from E2~SUMO complex for efficient SUMO discharge and catalysis. Typically, SIMs are short motifs that harbor a core sequence consisting of three or four hydrophobic residues (I, L, V) with an acidic or polar residue at position 2 or 3 (V/I-X-V/I-V/I or V/I-V/I-X-V/I/L) (10-13). The hydrophobic core of SIM is often flanked by a stretch of acidic amino acids that can promote electrostatic SUMO-SIM interactions (10,14). Structural studies have demonstrated that SIM can adopt a parallel or antiparallel β-strand conformation upon binding in the hydrophobic groove formed by β2-strand and α1-helix of SUMO proteins (5,10,14,15).

Members of the Siz/PIAS family of SUMO E3 ligases include human PIAS proteins (PIAS1, PIAS2α/xα and 2β/xβ, PIAS3, and PIAS4/y) and their yeast homologs (Siz1, Siz2, Mms21, and Zip3) (2,4,9,16). Siz/PIAS ligases contain several conserved regions **(Figure 5-1A)**, including an N-terminal SAP (scaffold attachment factor-A/B, acinus and PIAS) domain involved in AT-rich DNA binding, a PINIT domain important for substrate recognition and subcellular localization, a SP-RING domain related to ubiquitin RING E3 ligases to recruit E2 Ubc9, and a C-terminal SIM for SUMO binding (2,3,9).

Previous biochemical and structural studies in yeast SUMO E3 ligase Siz1 displayed a fragment comprising the PINIT, the SP-RING and a C-terminal region of SP-RING (named SP-CTD) as sufficient

to facilitate SUMO conjugation to the target substrates (4,17). The C-terminal SIM was dispensable for the Siz1 ligase activity, and instead, an acidic patch in the SP-CTD was suggested to be important for the donor SUMO positioning and activation of the E2~SUMO complex during Siz1-mediated SUMOylation (4). Intriguingly, other members of the Siz/PIAS ligase family do not contain a similar acidic domain in their SP-CTD region (4). We have previously shown that SUMO E3 ligase PIASy is essential for SUMO-2/3 conjugation of mitotic chromosomal proteins, DNA topoisomerase IIα (TopoIIα) and poly (ADP-ribose) polymerase 1 (PARP1) (16,18,19). To understand the SIM-dependence and mechanism of PIASy-mediated SUMOylation, we examined a series of PIASy truncations for their ability to enhance SUMOylation of TopoIIα and PARP1 using the established *in vitro* SUMOylation assay. The results assisted in the identification of a new SIM in PIASy. The new SIM was validated and characterized using NMR spectroscopy. Mutational analysis revealed significance of the new SIM in efficient PIASy-mediated SUMO conjugation.

5.3. Methods

5.3.1. Cloning, Expression, and Protein Purification

Xenopus laevis PIASy sharing 85% sequence similarity to the human homolog **(Figure 5-5)** was used in these studies. Design of expression plasmids for N-terminal His$_6$-tagged human SUMO-2-GG, SUMO-3-GG, and *Xenopus* full-length PIASy have been previously described (16,20). C-terminal truncations of PIASy (N480, N454, and N414, shown in Figure 1) and PIASy NMR constructs (287-454 and 287-501, shown in **Figure 5-2** and **5-3**) were amplified by PCR from full-length construct and subcloned into pET28a using EcoRI and XhoI restriction sites. PIASy fragments 409-454 and 409-501,

shown in **Figure 5-4** and **5-6**, were obtained by digesting PIASy 287-454 and 287-501 in pET28a with

BamHI/XhoI, respectively, and ligated into pET28a vector.
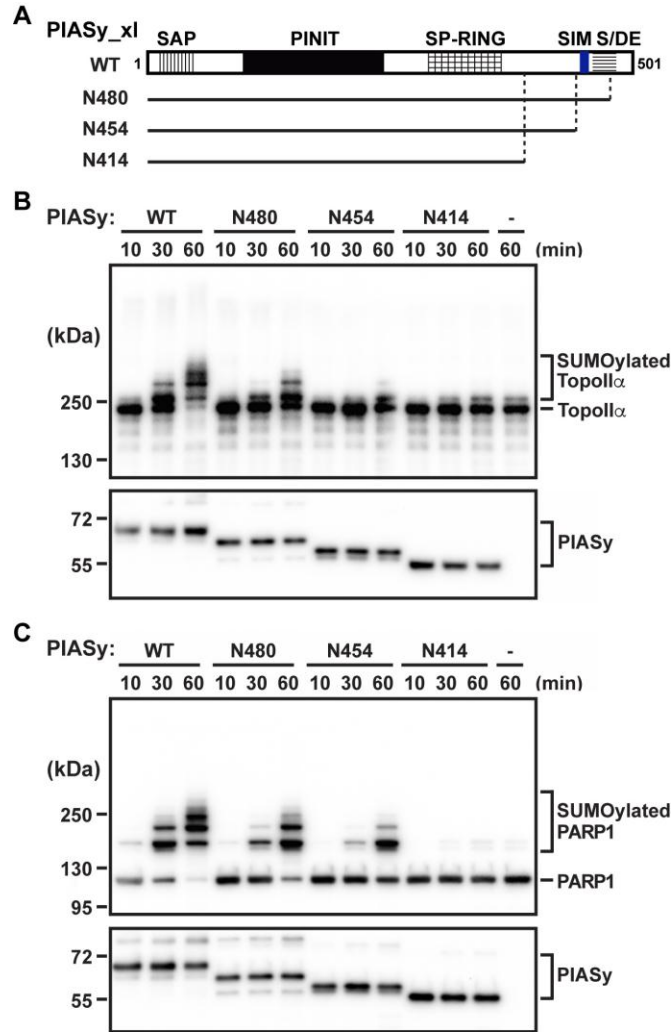
Figure 5-1. *In vitro* **SUMO E3 ligase activity of C-terminal deletions of PIASy.** (**A**) Schematic diagram with conserved domains of the Siz/PIAS E3 ligase family shown as: SAP (vertical lines), PINIT (solid black), SP-RING (grid), SIM (blue), and S/DE (horizontal lines). Solid bars below the schematic indicate PIASy truncation constructs tested in the SUMOylation assay. Time course for SUMO modification of (**B**) TopoIIα, and (**C**) PARP1 with the indicated PIASy truncation construct in the *in vitro* SUMOylation assay. SUMOylated and unmodified forms of TopoIIα and PARP1 are shown with bracket and bar, respectively. The bottom panel in **B** and **C** displays the input PIASy. Data courtesy of Dr. Yoshiaki Azuma.

Plasmids were freshly transformed in *E. coli* BL21 (DE3) cells and grown in 1 liter of culture media containing kanamycin at 30 μg/ml. ^{15}N-labeled SUMO-3 was expressed in M9 minimal media supplemented with 1 g of ^{15}N-ammonium chloride. Unlabeled SUMO-3 (for fluorescence polarization studies) and PIASy fragments (for NMR studies) were obtained by cell growth in LB media. PIASy constructs containing the RING domain were also supplied 0.1 mM $ZnSO_4$ before and after induction. Cells were grown at 37 °C until OD_{600} was ~ 0.7 to 0.8, induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and cell growth was continued overnight to a final OD_{600} ~ 2.5 (~2.1 for PIASy constructs) in a 15 °C shaker incubator. Cells were harvested by centrifugation (4000 rpm, 10 min), resuspended in ~ 30 ml binding buffer (500 mM NaCl, 20 mM Tris HCl pH 8.0, 5 mM imidazole), and lysed by sonication. Cellular debris was removed by centrifugation (13000 rpm, 10 min), and 600 μl of 5% (v/v) polyethyleneimine was added to the supernatant to precipitate the nucleic acids. After another centrifugation (13000 rpm, 10 min), the supernatant was loaded to a pre-charged Ni^{2+} column and His$_6$-tagged recombinant proteins were purified using standard nickel affinity chromatography procedure. Elution fractions containing purified protein were pooled and dialyzed in NMR buffer (100 mM NaCl, 10 mM $NaPO_4$ pH 6.8, and 5 mM DTT). Proteins were concentrated using Amicon Ultra 3K centrifugal filter (Millipore) and concentrations were determined by A_{280}. Recombinant proteins used in the SUMOylation assays, the E1 complex (Uba2/Aos1), E2 Ubc9, PIASy (full-length, truncation and mutant constructs), PARP1, TopoIIα, and SUMO-2 were expressed and purified same as previously ([16],[18-20]).

5.3.2. *Site-directed mutagenesis*

PIASy constructs with mutations in the SIM region (PIASy 409-454 mut, ΔO-SIM, ΔN-SIM, and ΔW-SIM, as shown in **Figure 5-4** and **5-7**) were constructed using QuikChange mutagenesis kit (Stratagene). All mutations were verified by DNA sequencing. Expression and purification of the mutants were carried out as described above.

5.3.3. *In vitro SUMOylation assay*

SUMOylation reactions were performed as previously ([18],[19]). Briefly, 5 μM SUMO-2, 2 mM ATP, 20 nM E1, 30 nM E2, 20 nM PIASy (different constructs as shown in figure), and 500 nM T7-tagged PARP1 or TopoIIα in reaction buffer (20 mM HEPES pH 7.8, 100 mM NaCl, 5 mM $MgCl_2$, 0.05% Tween-20, 5% glycerol, and 0.5 mM DTT) were incubated at 25 °C for 1 hour. At the indicated times, aliquots were removed, mixed with SDS-PAGE loading buffer and boiled. Samples were resolved by SDS-PAGE and analyzed by western blotting using HRP-conjugated anti-T7 monoclonal antibody.

5.3.4. *NMR Spectroscopy*

NMR data were acquired at 25 °C on a Bruker Avance III 600 MHz spectrometer equipped with TXI-RT probe. Data processing and analysis were carried out using NMRPipe ([21]) and NMRView ([22]), respectively. For NMR chemical shift mapping, 2D ^1H-^{15}N HSQC spectra were acquired using 0.2 mM ^{15}N-labeled SUMO-3 titrated with increasing molar ratios of unlabeled PIASy (indicated construct). All titration samples contained 10% (v/v) D_2O. The published backbone amide assignment of SUMO-2/3 was used in the analysis ([10],[23]). Weighted CSD was calculated using $0.5[(\Delta H)^2 + (\Delta N/5)^2]^{1/2}$ ([24]). Residues perturbed during the titrations were mapped onto the crystal structure of SUMO-3 (PDB ID: 1U4A).

5.3.5. *Fluorescence Polarization (FP)*

Purified SUMO-3 and PIASy 287-454 were dialyzed into 1X PBS buffer (phosphate buffered saline, pH 7.4) overnight. The single native cysteine of SUMO-3 (residue 47) was utilized to covalently attach the fluorophore, fluorescein-maleimide (FM), following manufacturer's protocol (Invitrogen).

Briefly, a 20-fold molar excess of FM prepared in dimethylsulfoxide was added to the TCEP-reduced

protein and allowed to react in dark at room temperature for ~ 3 hours. The unbound dye was removed

by extensive dialysis in PBS, followed by passage through a 3K centrifugal filter (Millipore). Increasing

concentrations of PIASy (0-360 μM) were titrated into 25 nM FM-labeled SUMO-3. Samples were

incubated at room temperature for 30 min and FP was measured with excitation and emission

wavelengths of 492 nm and 515 nm, respectively using a Varian Cary Eclipse Fluorescence

Spectrophotometer. G-factor corrected polarization values were obtained using formula: $p = (I_{VV} - G*I_{VH})/(I_{VV} + G*I_{VH})$. Background fluorescence was subtracted from the average of 5 data points and the

change in millipolarization (ΔmP) was plotted against increasing concentrations of PIASy. The

dissociation constant (K_d) was estimated by non-linear hyperbolic fitting of the data using GraphPad

Prism 5.0.

5.4. Results

5.4.1. *SUMO ligase activity of C-terminal truncation constructs of PIASy*

We have previously demonstrated the crucial role of PIASy in SUMO-2/3 modification of mitotic

substrates TopoIIα and PARP1 ([16],[18],[19]). To identify the determinants required for SUMO-2/3

modification by PIASy, we tested various C-terminal deletions of PIASy **(Figure 5-1A)** in the established

in vitro reconstituted SUMOylation assay. At physiological concentration of the SUMOylation enzymes,

PIASy N414 (a fragment lacking 87 C-terminal residues) was inefficient in SUMOylating both TopoIIα

and PARP1 **(Figure 5-1B, 5-1C)**. The extent of SUMO conjugation using N414 was almost similar to

the activity observed in E3-independent reaction (last two lanes, **Figure 5-1B, 5-1C**). Next, we examined

a longer PIASy construct, N454 that lacked the known C-terminal SIM but retained all the components

previously reported as sufficient for the ligase activity of Siz1, PIAS homolog in yeast ([4],[17]). Interestingly, though at a lower efficiency compared to full length PIASy (WT), SIM-less N454 was capable of SUMOylating both the mitotic substrates **(Figure 5-1B, 5-1C)**. Consistent with the previous observation in Siz1 ([4]), inclusion of the C-terminal SIM and the flanking acidic domain (PIASy N480) further enhanced the SUMO conjugation **(Figure 5-1B, 5-1C)**.

5.4.2. *PIASy lacking known C-terminal SIM interacts with SUMO*

Both PIASy N414 and N454 could recruit E2 Ubc9 through the SP-RING domain, however only N454 stimulated *in vitro* SUMO conjugation as an active ligase **(Figure 5-1)**. Previous studies have revealed that efficient E3-mediated SUMO conjugation requires optimal positioning of both E2 and SUMO from the E2~SUMO thioester complex ([4-7],[25]). To determine whether the SIM-less N454 orients the E2~SUMO complex for productive catalysis by additional stabilization of SUMO, we tested the direct binding between SUMO and the SIM-less PIASy using NMR spectroscopy. A shorter construct of PIASy, 287-454 **(Figure 5-2A)** due to better expression compared to N454, was more amenable to NMR and thus used in the interaction studies. ^{15}N-labeled SUMO-3 was titrated with increasing concentrations of unlabeled PIASy 287-454 at 1:0, 1:0.5, 1:1, and 1:2 molar ratios, and the titration was monitored by acquiring 2D ^{1}H-^{15}N HSQC spectra **(Figure 5-2B)**. The stepwise addition of SIM-less PIASy 287-454 resulted in peak broadening of specific resonances of SUMO-3 **(Figure 5-2B)** indicating interaction in the intermediate exchange NMR time scale. To identify the SUMO-3 residues that were perturbed the most by PIASy 287-454, we calculated the peak intensity ratio ($I_{1:1}/I_{1:0}$) for each non-overlapped SUMO-3 peak at a SUMO-3:PIASy 287-454 molar ratio of 1:1 **(Figure 5-2C)**. Residues displaying significant peak intensity reduction (average intensity minus a standard deviation) included S27, Q30, F31, K41, L42, A45, Y46, D62, E78, and T82. The affected residues were mapped onto the structure of SUMO-3 **(Figure 5-2D)**. Surprisingly, even though PIASy 287-454 lacks the known C-

terminal SIM, majority of the affected residues clustered together spatially in the β2-strand and α1-helix of SUMO-3, the well characterized binding surface for the SIM on SUMO ([10,12,14,15]). These results suggested existence of another SIM in PIASy within residues 287-454. To gain further insights into the SUMO-PIASy 287-454 interaction, we conducted a fluorescence polarization based binding assay with fluorescein-maleimide labeled SUMO-3. PIASy 287-454 interacted with SUMO-3 with an apparent K_d of ~ 81 μM **(Figure 5-2E)**. The value obtained is consistent with the observed intermediate exchange NMR behavior and previously reported dissociation constants for other SUMO-SIM interactions, which are in the 1-120 μM range ([6,10,12,15,26]).

In addition to PIASy 287-454, we also performed NMR titrations of ^{15}N SUMO-3 with unlabeled PIASy 287-501 **(Figure 5-3A)**. As expected, incorporation of the known C-terminal SIM and the acidic domain resulted in more pronounced chemical shift perturbations when compared to PIASy 287-454 **(Figure 5-3B)**. Stepwise addition of PIASy 287-501 caused concentration dependent decrease in intensities of SUMO-3 peaks. Some SUMO-3 residues along with peak intensity reduction also showed subtle shifts in peak positions (*e.g.* R35, K41, Y46, and N67) and several peaks broadened beyond detection during the titration **(Figure 5-3B)**. Although most SUMO-3 residues showed some degree of line broadening from the beginning of titration, which is not uncommon with the formation of a high molecular weight complex (SUMO-3, 13 kDa; PIASy 287-501, 28 kDa) slowing down the tumbling rate ([27]), specific residues were more affected than the others as illustrated by peak intensity ratio graph **(Figure 5-3C)**. The strongly affected SUMO-3 residues by PIASy 287-501 were identical to the residues affected by PIASy 287-454 and mapped to β2-strand and α1-helix, the SIM binding surface on the SUMO **(Figure 5-3D)**.
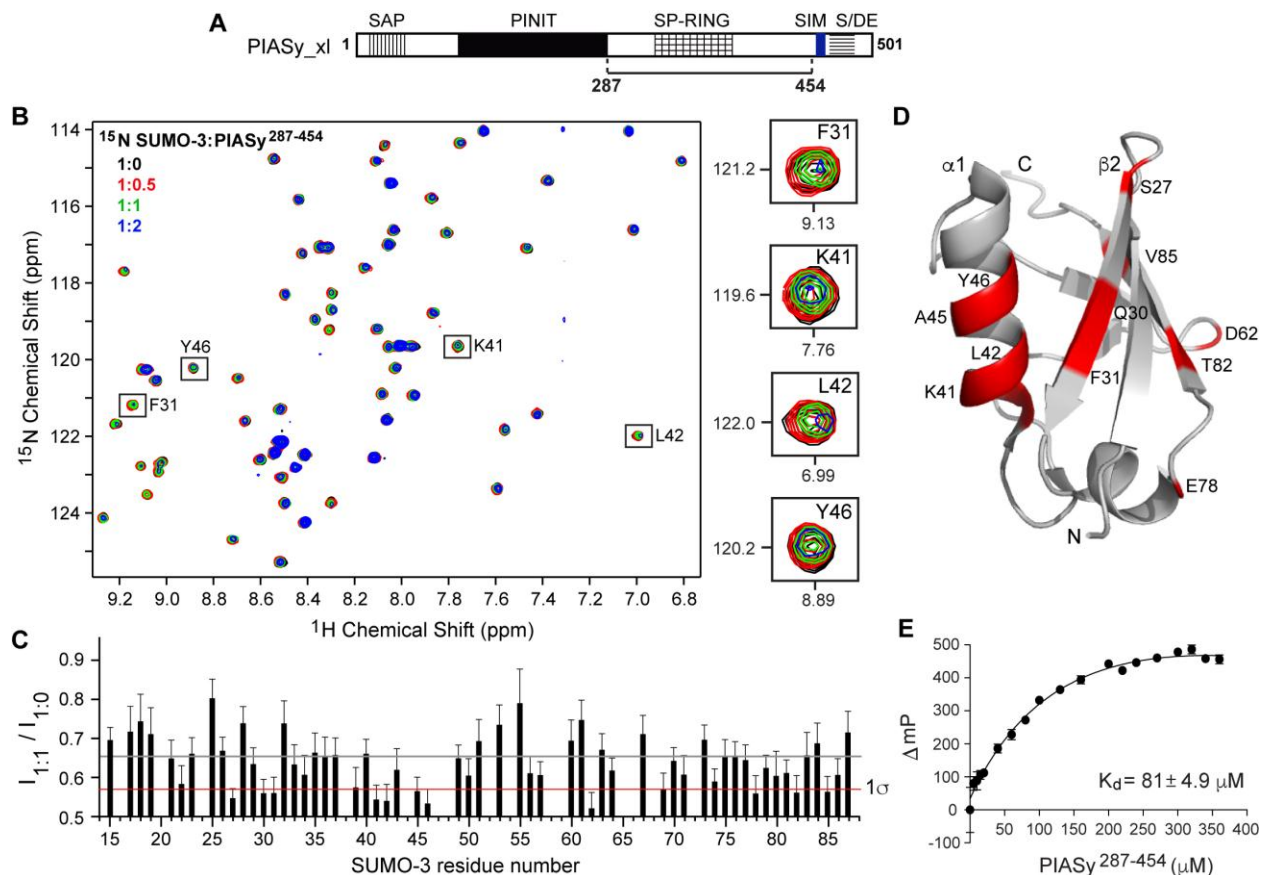
Figure 5-2. Analysis of the interaction of SUMO-3 with PIASy lacking the known SIM. **(A)** PIASy schematic as in Figure 1A, with solid bar indicating the fragment, 287-454, used in NMR and FP experiment. **(B)** 2D ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled SUMO-3 titrated with increasing molar ratios of PIASy 287-454. Expanded sections of representative SUMO-3 residues affected upon the binding of PIASy 287-454 are shown on right. **(C)** Plot of relative peak intensity for all assigned, non-overlapping SUMO-3 resonances in the ligand bound versus free form ($I_{1:1}/I_{1:0}$). Gray and red lines depict mean and one standard deviation from the mean (-1σ), respectively. **(D)** SUMO-3 residues displaying significant peak intensity reduction upon complex formation with PIASy 287-454 are highlighted red. **(E)** Fluorescence polarization binding assay of FM-labeled SUMO-3 titrated with increasing concentrations of PIASy 287-454.
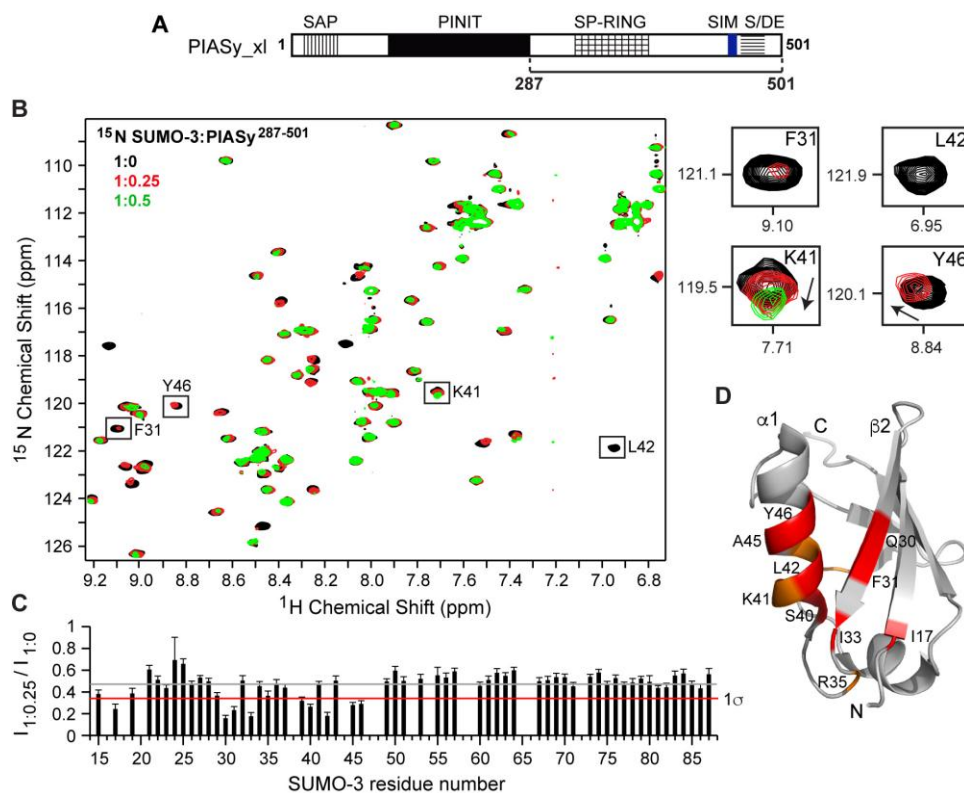
Figure 5-3. NMR binding study of SUMO-3 and PIASy 287-501. **(A)** Schematic of PIASy as in Figure 1A, with solid bar indicating the fragment used in the experiment. **(B)** 2D ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled SUMO-3 titrated with increasing amounts of PIASy 287-501. Expanded sections of selected SUMO-3 residues affected by PIASy 287-501 are shown on right. **(C)** Relative peak intensity graph for SUMO-3 resonances in the ligand bound versus free state ($I_{1:0.25}/I_{1:0}$), where, gray and red lines correspond to the mean and one standard deviation from the mean (-1σ), respectively. **(D)** Affected SUMO-3 residues upon the binding of PIASy 287-501 are highlighted (red, significant peak intensity reduction; orange, significant change in the peak position).

5.4.3. A new SIM in SUMO E3 ligase PIASy

NMR chemical shift mapping results for ^{15}N SUMO-3 titrations with PIASy 287-454 indicated an additional SIM in PIASy **(Figure 5-2)**. Using a SIM-prediction web server (28), we found PIASy 287-454 fragment to contain a high probability SIM (herein, termed as new SIM) within residues 417-435 **(Figure 5-4A)**. The new SIM showed complete sequence conservation across many different organisms from amphibian to the human PIASy **(Figure 5-5)**. Comparable to characteristic SIM sequences that bind

SUMO in antiparallel orientation ([6](#),[10](#),[11](#),[26](#)), the new PIASy SIM harbor a conserved hydrophobic core 'ILVL' that is preceded by few acidic and putative phosphorylatable serine residues **(Figure 5-4A)**. However, unlike a signature SIM, the new SIM lacks a long contiguous stretch of acidic residues. To confirm the role of new SIM in SUMO binding, we subcloned PIASy region containing the new SIM fragment (PIASy 409-454 'wt') and mutated three critical hydrophobic core SIM residues (PIASy 409-454 'mut') to alanines **(Figure 5-4A)**. In comparison to the earlier SP-RING containing fragments, PIASy 409-454 showed greater stability and solubility, and thus, was ideal for NMR studies. NMR binding experiments were performed using ^{15}N-labeled SUMO-3 with either 'wt' or 'mut' PIASy 409-454 new SIM fragment **(Figure 5-4)**. Besides gradual peak intensity reduction, addition of unlabeled 'wt' PIASy 409-454 resulted in concentration dependent changes in peak position for selective SUMO-3 resonances **(Figure 5-4B, 5-4C)**, indicating interaction at the edge of fast intermediate exchange NMR time scale. In contrast, NMR titrations with 'mut' PIASy 409-454 barely induced any discernible chemical shift perturbations on SUMO-3 **(Figure 5-4D, 5-4E)**, indicating the new SIM mutation interfered with the SUMO interaction. Analysis of weighted chemical shift deviations (CSD, **Figure 5-4F**) displayed that strongly affected residues upon binding of 'wt' PIASy 409-454 cluster clearly near the SIM binding groove of SUMO-3 **(Figure 5-4G)**. Together, the results highlight the importance of hydrophobic core of the new PIASy SIM in SUMO recognition.

Next, we analyzed the binding of SUMO-3 with PIASy 409-501, construct encompassing old SIM and the acidic domain **(Figure 5-6A)**. Similar to the earlier titrations, identical SIM-binding residues of SUMO-3 were most strongly perturbed upon the addition of PIASy 409-501 **(Figure 5-6B, 5-6C, 5-6D)**. Comparison of SUMO-3 titrations with PIASy 409-454 (new SIM alone) and PIASy 409-501 (both SIMs) indicated that presence of both SIMs together conferred a higher magnitude of change in the chemical shift perturbations of selective SUMO-3 resonances, notably, even when using a much lower ligand concentration. Important SUMO-3 SIM-binding residues, such as, Q30, F31, L42, A45, and Y46,

111

showed severe peak intensity reduction starting at 1:0.25 molar ratios and complete peak broadening by 1:0.5 titration point (**Figure 5-6B, 5-6C**), indicating binding in the intermediate exchange NMR regime.
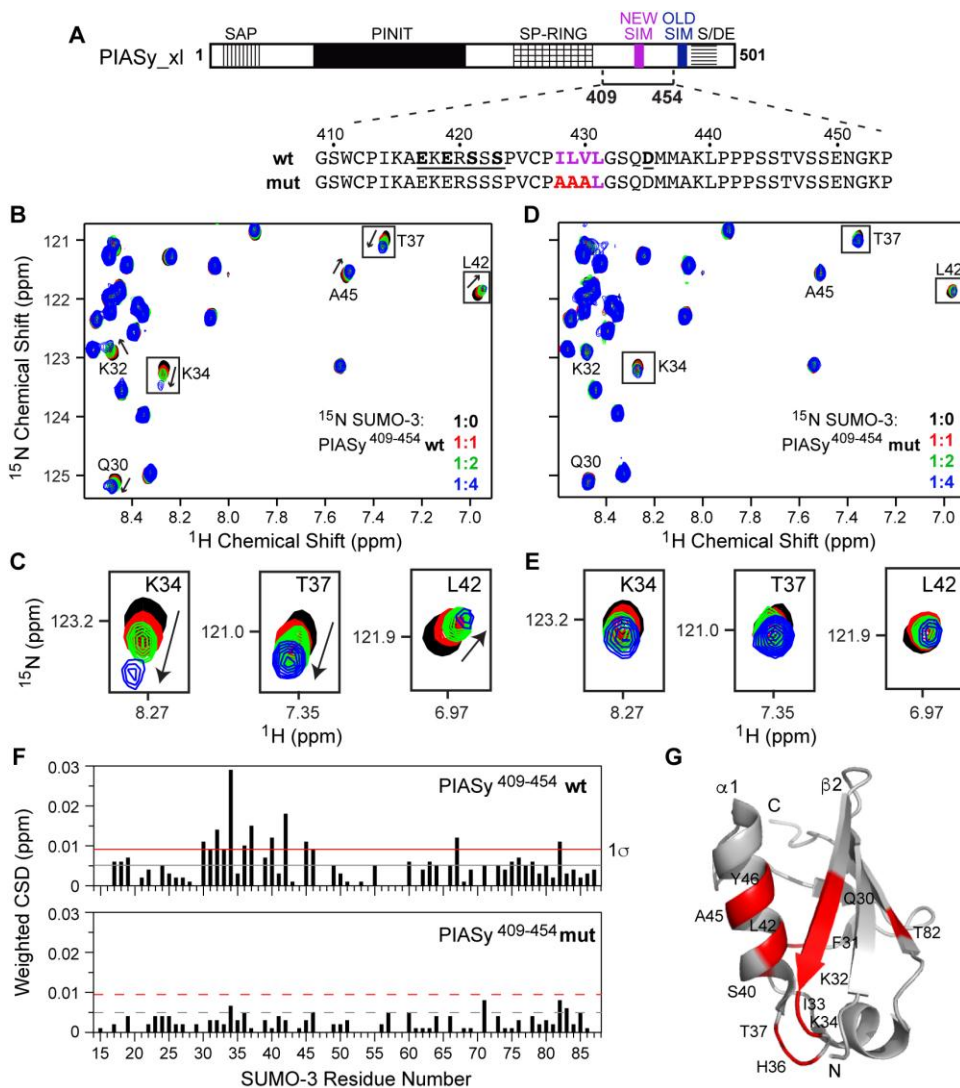
Figure 5-4. Characterization of the binding of SUMO-3 and the new SIM of PIASy. (**A**) PIASy schematic as in Figure 1A with new SIM in magenta. Solid bar corresponds to the PIASy fragment, 409-454, used in the experiment. Sequence of the wild type 'wt' and mutant 'mut' PIASy construct is shown. (**B**) A section of 2D ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled SUMO-3 titrated with increasing molar ratios of wt PIASy 409-454 demonstrates specific SUMO-SIM binding. (**C**) Expanded sections of selected SUMO-3 residues affected upon wt PIASy 409-454 binding. (**D**) and (**E**) Similar to (**B**) and (**C**) for SUMO-3 titrations with mut PIASy 409-454. (**F**) Plots of weighted CSD from titrations of SUMO-3 with wt PIASy 409-454 (*top*) and mut PIASy 409-454 (*bottom*). Gray and red lines correspond to the mean and one standard deviation ($+1\sigma$) from the mean, respectively. (**G**) Strongly affected SUMO-3 residues by the binding of wt PIASy 409-454 are colored red.
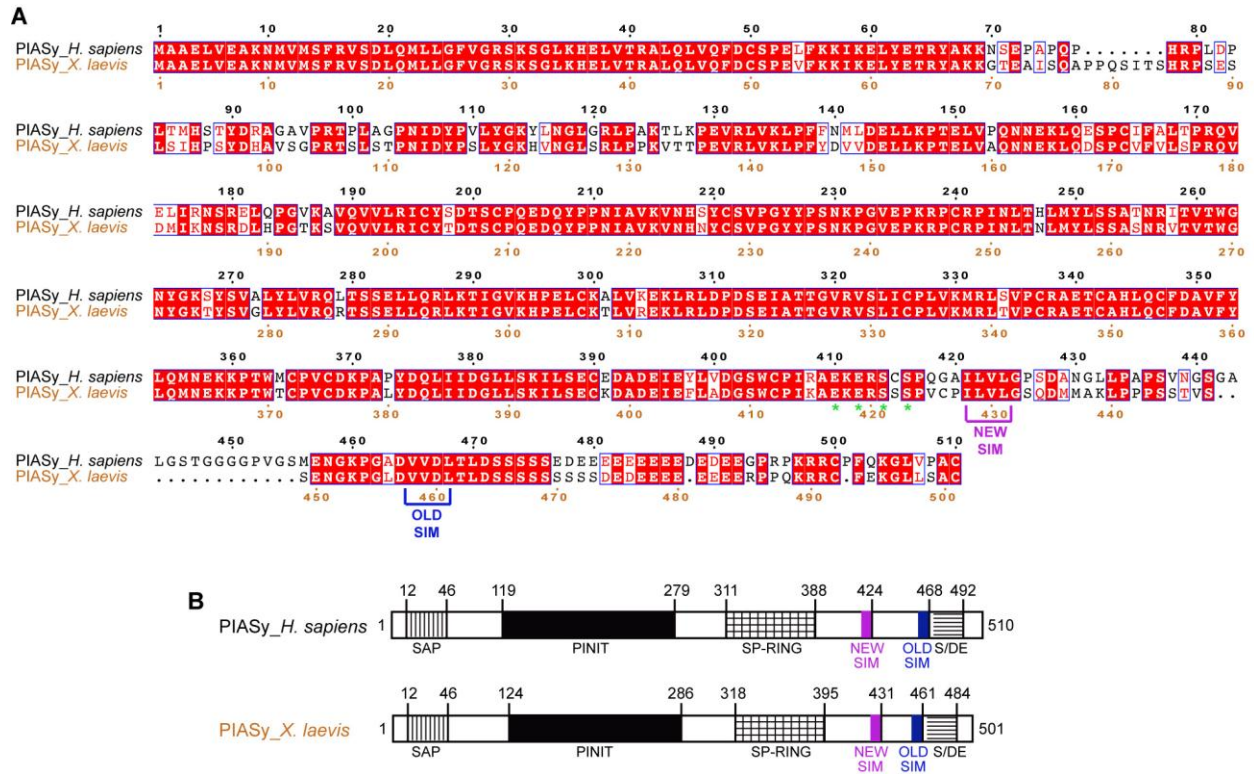
Figure 5-5. Comparison of human and *Xenopus* PIASy. (A) Sequence alignment of PIASy from *Homo sapiens* and *Xenopus laevis*. Numbers above and below the alignment correspond to human (black) and *Xenopus* (orange) PIASy, respectively. The old known SIM is shown in blue and the new SIM identified in this work is indicated in magenta. Conserved serine and acidic residues preceding the new SIM are marked with green asterisk. **(B)** Schematic of human and *Xenopus* PIASy displaying conserved domains of Siz/PIAS family of SUMO E3 ligases. Numbers indicate boundaries between the domains.

5.4.4. New SIM is crucial in PIASy-mediated SUMOylation

To evaluate the significance of the SIMs for the E3 ligase activity of PIASy, we mutated three essential core SIM residues to alanines to generate ΔO-SIM (mutation in the old known SIM), ΔN-SIM (mutation in the new SIM), and ΔW-SIM (a double SIM mutant) in full-length PIASy as shown in **Figure 5-7A**. The mutants were examined for their ability to facilitate SUMO conjugation of PARP1 in the *in vitro* SUMOylation assay. Interestingly, mutation of the old-SIM mutant (ΔO-SIM) did not interfere with the role of PIASy and exhibited high SUMOylation levels comparable to the wild type (WT) PIASy

(**Figure 5-7B, 5-7C**). However, the new SIM mutant (ΔN-SIM) was significantly less efficient in mediating SUMO conjugation (**Figure 5-7B, 5-7C**). Indeed, there was almost negligible SUMOylation using ΔN-SIM after 30 min of the reaction time. Furthermore, a combinatorial mutation of old and new SIM (ΔW-SIM) nearly abolished the ligase activity of PIASy (**Figure 5-7B, 5-7C**). Collectively, our findings suggest that the new SIM may have important implications in PIASy-mediated SUMO-2/3 conjugation.
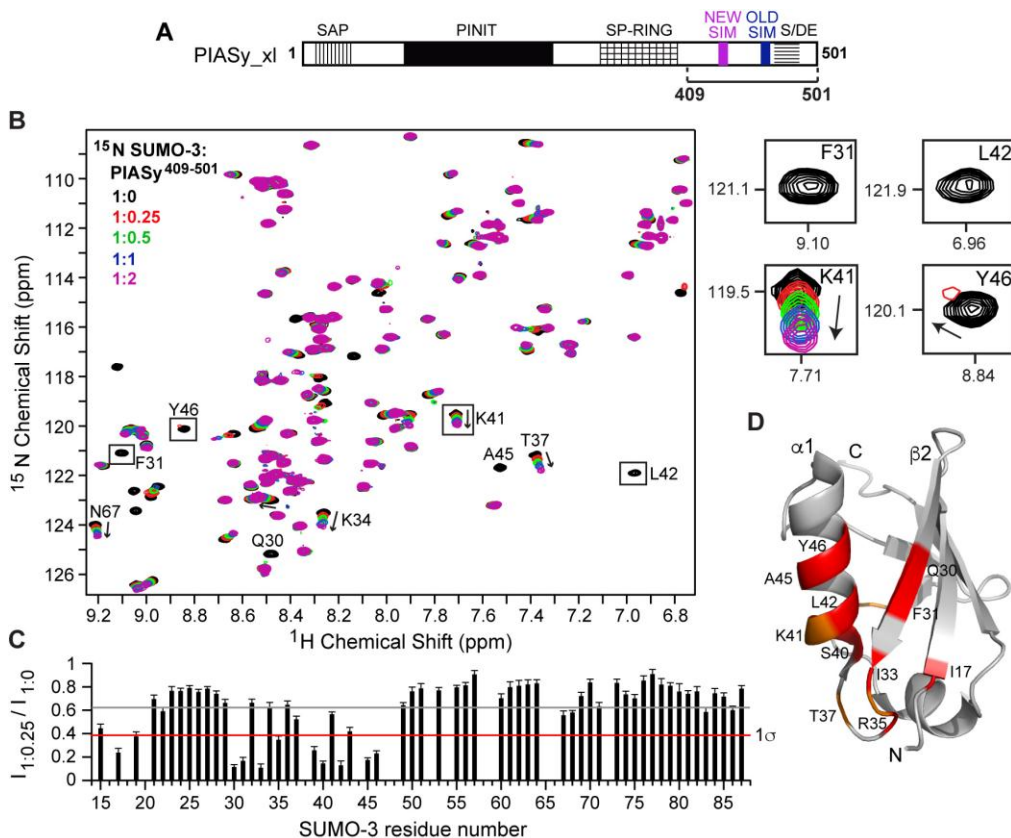
Figure 5-6. NMR titrations of SUMO-3 and PIASy 409-501. (**A**) Schematic of PIASy as in Figure 5-4, with solid bar indicating fragment used in the experiment. (**B**) 2D ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled SUMO-3 titrated with increasing amounts of PIASy 409-501. Expanded sections of representative SUMO-3 residues affected by the binding of PIASy 409-501 are on right. (**C**) Relative peak intensity plot for SUMO-3 resonances in the ligand bound versus free state ($I_{1:0.25}/I_{1:0}$). Gray and red lines depict mean and one standard deviation from the mean (-1σ), respectively. (**D**) Affected SUMO-3 residues upon complex formation with PIASy 409-501 are highlighted (red, residues displaying significant peak intensity reduction; orange, residues with significant change in the peak position).
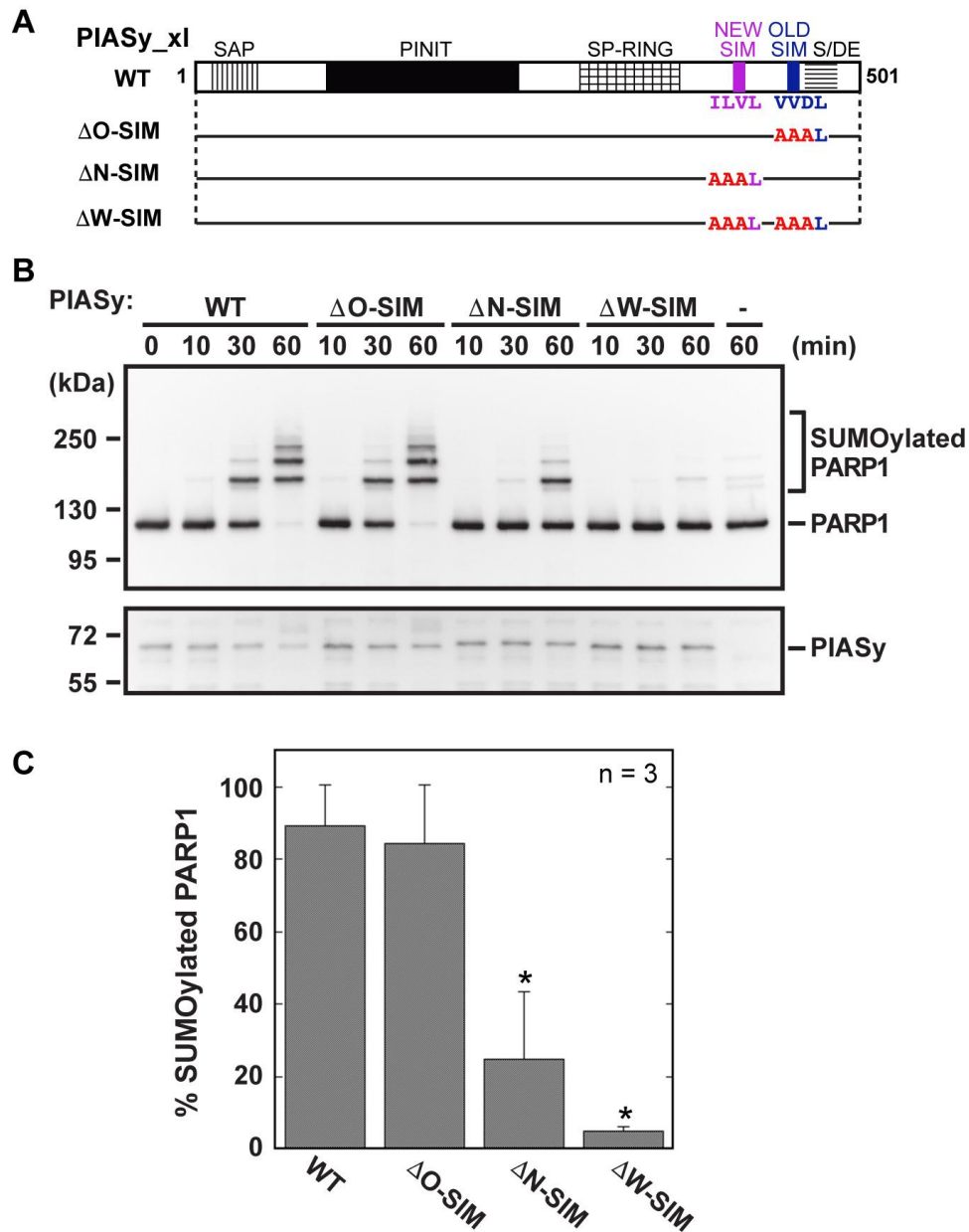
Figure 5-7. Significance of the new SIM of PIASy. **(A)** Schematic of PIASy as in Figure 5-4, with solid bars illustrating different SIM mutations examined in the SUMOylation assay. **(B)** Time course for SUMO modification of PARP1 with the indicated PIASy SIM mutant in the *in vitro* SUMO conjugation assay. SUMOylated and unmodified forms of PARP1 are marked with bracket and bar, respectively. The PIASy input is shown in the bottom panel. **(C)** Quantification of the SUMO conjugation assay shown in **B** at the final time point. Assay was performed in triplicate. Error bars represent one standard deviation and asterisk indicates statistically significant differences from the WT activity.
Data courtesy of Dr. Yoshiaki Azuma.

5.5. Discussion

In the current study, the combination of NMR binding studies and *in vitro* SUMOylation assays revealed a new SIM in PIASy that was found to be essential for efficient PIASy-dependent SUMO-2/3 conjugation. In contrast to the old SIM of PIASy (or the known SIM of other PIAS family members), the new SIM contains an uninterrupted core of four hydrophobic residues and lacks the conventional extended acidic residue stretch **(Figure 5-5)**. Previous SUMO-SIM interaction studies have described the hydrophobic core of SIM as a key mediator for SUMO binding ([10],[12],[13]). Consistent with these findings, our NMR studies demonstrate that hydrophobic core residues of the new SIM are sufficient to establish binding to SUMO-3 and mutation of these critical residues abolishes binding **(Figure 5-4)**. Additionally, previous studies have highlighted the role of acidic residue tract surrounding the hydrophobic core of SIM in SUMO paralog binding specificity ([10]). While presence of the negative charges in SIM has been shown to contribute more toward SUMO-1 binding, SIMs lacking the acidic stretch preferentially bind SUMO-2 ([8],[10]). Absence of the acidic residue stretch in the new SIM, together with the significant role of new SIM in PIASy-mediated SUMO conjugation **(Figure 5-7)**, provides a likely explanation for the high preference of PIASy in mediating SUMO-2/3 conjugation over SUMO-1 to substrates TopoIIα and PARP1 ([16],[18],[19]).

In addition to the covalent interaction between E2 Ubc9 and SUMO (E2~SUMO thioester complex), the backside of Ubc9 can also recruit a second molecule of SUMO (E2-SUMO) through non-covalent interactions ([29],[30]). Recent structural and biochemical studies have shown that SUMO E3 ligase ZNF451 exploit these interactions to preferentially interact with two molecules of SUMO (donor SUMO from E2~SUMO[D] and backside E2-SUMO) via its tandem SIMs ([6],[7]). The first SIM in conjunction with the inter-SIM region positions the E2~SUMO[D] thioester complex for productive catalysis while the second SIM acts as an anchor to bind to the backside SUMO to facilitate poly-SUMOylation ([6],[7]). Intriguingly, although PIAS proteins use the SP-RING domain to directly bind to E2

Ubc9, studies suggest that the SIM of PIAS proteins can also interact with the backside E2-SUMO to form a non-covalent ternary complex ([31]). Taken together with the finding of the new SIM in PIASy, it seems plausible that analogous to ZNF451, the tandem SIM of PIASy may also simultaneously interact with the donor SUMO (E2~SUMO^D) and the backside SUMO to mediate efficient SUMOylation. In line with this notion, results of our NMR interaction studies with PIASy containing both the SIMs showed improved SUMO binding affinity **(Figure 5-6)** compared to the shorter construct consisting of new SIM only **(Figure 5-4)**. Specific SUMO-3 SIM-interacting residues including Q30, F31, L42, A45, and Y46, showed intense peak broadening even with very small amount of PIASy 409-501 (construct containing both the SIMs), however, these residues were in fast intermediate exchange NMR regime with the new SIM alone, suggesting weaker binding compared to PIASy 409-501. Since both SIMs individually bind to the same surface on SUMO-3 through classical SUMO-SIM interactions ([31]), stronger perturbations observed with both SIM containing PIASy 409-501 is probably reflective of a synergistic effect occurring upon the binding of different SUMO-3 molecules to the tandem SIM. Additionally, our *in vitro* SUMOylation assays confirmed that PIASy stimulated SUMO conjugation in a SIM-dependent manner and both SIMs are important for effective ligase activity. Although, mutation of the new SIM (ΔN-SIM) rendered SUMOylation slow and inefficient, probably implying its role in the donor SUMO (E2~SUMO^D) positioning and chain initiation, mutation of both SIMs together (ΔW-SIM, **Figure 5-7**) or deletion of the tandem SIM region (N414, **Figure 5-1**) eliminated PIASy E3 ligase activity. In concordance, presence of both SIMs together resulted in the robust ligase activity of PIASy (N480, **Figure 5-1**). Altogether, it is tempting to speculate that the SP-RING domain and the new SIM of PIASy may configure the E2~SUMO^D thioester complex and the old SIM may engage the backside SUMO for efficient catalysis and SUMO conjugation by PIASy. The functional importance of the new SIM in configuring donor SUMO is also in agreement with the previous Siz1 studies as both the new SIM of PIASy and Siz1 SP-CTD acidic patch regarded important for optimal donor SUMO positioning are located in the similar regions of the proteins ([4]).

Lastly, the N-terminal of the new SIM includes few conserved acidic and serine residues (**Figure 5-4**) corresponding to the CK2 consensus phosphorylation site ([32]). Phosphorylation of these residues could create a phosphoSIM that may enhance SUMO-SIM interactions by allowing additional electrostatic interactions ([15]). This may also influence an antiparallel bound orientation for the new SIM by asymmetrically augmenting negative charges toward the N-terminal of the hydrophobic core of the new SIM ([10],[13],[26]).

In conclusion, we have identified a new SIM (second C-terminal SIM) in PIASy that is critical for the efficient ligase activity and distinct from the other PIAS family proteins that contain a single SIM on their C-terminus. Future *in vivo* analysis with mutant PIASy lacking individual SIM will provide further insight into the functional importance of multi-SIMs of E3 ligase PIASy.

5.6. References

1. Meulmeester, E., and Melchior, F. (2008) Cell biology: SUMO. in *Nature*
2. Gareau, J. R., and Lima, C. D. (2010) The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. in *Nat Rev Mol Cell Biol*, Nature Publishing Group
3. Geiss-Friedlander, R., and Melchior, F. (2007) Concepts in sumoylation: a decade on. in *Nat Rev Mol Cell Biol*, Nature Publishing Group
4. 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. in *Mol. Cell*, Elsevier Ltd
5. Reverter, D., and Lima, C. D. (2005) Insights into E3 ligase activity revealed by a SUMO–RanGAP1–Ubc9–Nup358 complex. in *Nature*
6. Cappadocia, L., Pichler, A., and Lima, C. D. (2015) Structural basis for catalytic activation by the human ZNF451 SUMO E3 ligase. in *Nat Struct Mol Biol*
7. Eisenhardt, N., Chaugule, V. K., Koidl, S., Droescher, M., Dogan, E., Rettich, J., Sutinen, P., Imanishi, S. Y., Hofmann, K., Palvimo, J. J., and Pichler, A. (2015) A new vertebrate SUMO enzyme family reveals insights into SUMO-chain assembly. in *Nat Struct Mol Biol*
8. Merrill, J. C., Melhuish, T. A., Kagey, M. H., Yang, S.-H., Sharrocks, A. D., and Wotton, D. (2010) A role for non-covalent SUMO interaction motifs in Pc2/CBX4 E3 activity. in *PLoS One*
9. Rytinki, M. M., Kaikkonen, S., Pehkonen, P., Jääskeläinen, T., and Palvimo, J. J. (2009) PIAS proteins: pleiotropic interactors associated with SUMO. in *Cell. Mol. Life Sci.*, SP Birkhäuser Verlag Basel

10.   Hecker, C. M. (2006) Specification of SUMO1- and SUMO2-interacting Motifs. *J Biol Chem* **281**, 16117-16127

11.   Jardin, C., Horn, A. H. C., and Sticht, H. (2015) Binding properties of SUMO-interacting motifs (SIMs) in yeast. *J Mol Model* **21**, 50

12.   Song, J., Durrin, L. K., Wilkinson, T. A., Krontiris, T. G., and Chen, Y. (2004) Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14373-14378

13.   Song, J., Zhang, Z., Hu, W., and Chen, Y. (2005) Small Ubiquitin-like Modifier (SUMO) Recognition of a SUMO Binding Motif. *Journal of Biological Chemistry*

14.   Sekiyama, N. N., Ikegami, T. T., Yamane, T. T., Ikeguchi, M. M., Uchimura, Y. Y., Baba, D. D., Ariyoshi, M. M., Tochio, H. H., Saitoh, H. H., and Shirakawa, M. M. (2008) Structure of the small ubiquitin-like modifier (SUMO)-interacting motif of MBD1-containing chromatin-associated factor 1 bound to SUMO-3. *J Biol Chem* **283**, 35966-35975

15.   Chang, C.-C., Naik, M. T., Huang, Y.-S., Jeng, J.-C., Liao, P.-H., Kuo, H.-Y., Ho, C.-C., Hsieh, Y.-L., Lin, C.-H., Huang, N.-J., Naik, N. M., Kung, C. C.-H., Lin, S.-Y., Chen, R.-H., Chang, K.-S., Huang, T.-H., and Shih, H.-M. (2011) Structural and functional roles of Daxx SIM phosphorylation in SUMO paralog-selective binding and apoptosis modulation. *Mol. Cell* **42**, 62-74

16.   Azuma, Y., Arnaoutov, A., Anan, T., and Dasso, M. (2005) PIASy mediates SUMO-2 conjugation of Topoisomerase-II on mitotic chromosomes. in *EMBO J*

17.   Takahashi, Y., and Kikuchi, Y. (2005) Yeast PIAS-type Ull1/Siz1 is composed of SUMO ligase and regulatory domains. *J Biol Chem* **280**, 35822-35828

18.   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. in *J Biol Chem*

19.   Ryu, H., Furuta, M., Kirkpatrick, D., Gygi, S. P., and Azuma, Y. (2010) PIASy-dependent SUMOylation regulates DNA topoisomerase II activity. in *J. Cell Biol.*

20.   Azuma, Y., Arnaoutov, A., and Dasso, M. (2003) SUMO-2/3 regulates topoisomerase II in mitosis. *J. Cell Biol.* **163**, 477-487

21.   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. in *J Biomol NMR*

22.   Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. in *Methods Mol. Biol.*

23.   Ding, H., Xu, Y., Chen, Q., Dai, H., Tang, Y., Wu, J., and Shi, Y. (2005) Solution Structure of Human SUMO-3 C47S and Its Binding Surface for Ubc9. *Biochemistry* **44**, 2790-2799

24.   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. in *Nat Struct Biol*

25.   Plechanovova, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H., and Hay, R. T. (2012) Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **489**, 115-120

26.   Namanja, A. T., Li, Y.-J., Su, Y., Wong, S., Lu, J., Colson, L. T., Wu, C., Li, S. S. C., and Chen, Y. (2012) Insights into High Affinity Small Ubiquitin-like Modifier (SUMO) Recognition by SUMO-interacting Motifs (SIMs) Revealed by a Combination of NMR and Peptide Array Analysis. *J Biol Chem*

27.   Williamson, M. P. (2013) Using chemical shift perturbation to characterise ligand binding. in *Progress in Nuclear Magnetic Resonance Spectroscopy*, Elsevier B.V.

28.   Zhao, Q., Xie, Y., Zheng, Y., Jiang, S., Liu, W., Mu, W., Liu, Z., Zhao, Y., Xue, Y., and Ren, J. (2014) GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. *Nucleic Acids Research* **42**, W325-W330

29.   Knipscheer, P., van Dijk, W. J., Olsen, J. V., Mann, M., and Sixma, T. K. (2007) Noncovalent interaction between Ubc9 and SUMO promotes SUMO chain formation. *EMBO J* **26**, 2797-2807
30.   Capili, A. D., and Lima, C. D. (2007) Structure and analysis of a complex between SUMO and Ubc9 illustrates features of a conserved E2-Ubl interaction. *J Mol Biol* **369**, 608-618
31.   Mascle, X. H., Lussier-Price, M., Cappadocia, L., Estephan, P., Raiola, L., Omichinski, J. G., and Aubry, M. (2013) Identification of a non-covalent ternary complex formed by PIAS1, SUMO1, and UBC9 proteins involved in transcriptional regulation. *J Biol Chem* **288**, 36312-36327
32.   Stehmeier, P., and Muller, S. (2009) Phospho-regulated SUMO interaction modules connect the SUMO system to CK2 signaling. *Mol. Cell* **33**, 400-409

Chapter 6: The Fungal Natural Product Azaphilone-9 Inhibits HuR-mRNA Interaction

Kawaljit Kaur[1], Xiaoqing Wu[1], James Fields[1], David K. Johnson[2], Lan Lan[1], Miranda Pratt[1],

Amber D. Somoza[3], Clay C. C. Wang[3], John Karanicolas[1,2], Berl R. Oakley[1], Liang Xu[1],

and Roberto N. De Guzman[1]

[1]Department of Molecular Biosciences and [2]Center for Computational Biology, University of Kansas,

Lawrence, KS 66045, USA.

[3]Department of Chemistry and Department of Pharmacology and Pharmaceutical Sciences, School of

Pharmacy, University of Southern California, Los Angeles, CA 90089

6.1. Abstract

The RNA-binding protein Hu antigen R (HuR) binds to adenine- and uridine-rich elements (ARE) in the 3'- or 5'-untranslated region (UTR) of target mRNAs. The HuR-ARE interactions stabilize many oncogenic mRNAs that play important roles in tumorigenesis. Thus, small molecules that interfere with the HuR-ARE interaction could potentially inhibit cancer cell growth and progression. We used fluorescence polarization (FP) based competition assay to identify that the compound azaphilone-9 (AZA-9) derived from the fungal natural product asperbenzaldehyde is a potent inhibitor of HuR-ARE interaction. Results from surface plasmon resonance (SPR) verified the direct binding of AZA-9 to HuR. NMR methods mapped the RNA-binding interface of HuR and identified the involvement of critical RNA-binding residues in binding AZA-9. Further, computational methods modeled the binding of AZA-9 in the RNA-binding cleft of HuR. Our results show that AZA-9 blocks key ARE-binding residues of HuR and disrupt HuR-ARE interactions.

6.2. Introduction

The limited lifetime and subsequent decay of messenger RNA (mRNA) is an important mechanism for posttranscriptional regulation of gene expression. In mammalian cells, mRNA decay is dependent on both *cis* elements located in the RNA and *trans* acting regulatory factors such as RNA-binding proteins. AU-rich elements (ARE) in 5'- or 3'-untranslated region (UTR) of mRNAs are the most common *cis* elements that promote rapid degradation of mRNAs ([1,2](#)). Specific RNA-binding proteins can bind to these AREs and either accelerate decay or protect mRNA from degradation ([1-4](#)).

The RNA-binding protein Hu antigen R (HuR), a ubiquitous member of the ELAV/Hu protein family, binds and stabilizes ARE-containing mRNAs that encode oncoproteins, cytokines, growth factors and transcription factors (3-7). HuR is a multi-domain protein containing three RNA-recognition motifs, RRM1, RRM2, and RRM3, with each RRM comprising of about 80 amino acids. High affinity binding of HuR to ARE of mRNA is accomplished via its two tandem N-terminal RRM, RRM1 and RRM2 that are separated by a 7-residue inter-domain linker (8). The third RRM of HuR, RRM3, along with the basic hinge region that connects RRM2 with RRM3 mediate cooperative assembly of HuR oligomers on RNA (9). Although HuR is predominantly nuclear, the protein rapidly translocates to the cytoplasm in response to stimuli mediated by a nucleo-cytoplasmic shuttling sequence located in the hinge region (10).

HuR is overexpressed in a wide variety of cancer, including colon, ovarian, brain, breast, cervical, and pancreas (7,11-13). HuR promotes tumorigenesis by binding to cancer-associated ARE-containing mRNAs, which encode proteins implicated in tumor cell proliferation, cell survival, angiogenesis, invasion, and metastasis (7,14-16). HuR binds to the AREs of the oncogene Musashi1 (Msi1) and anti-apoptotic proteins, Bcl-2 and XIAP, thereby up-regulating their expression and activating the Wnt/Notch signaling pathway and inhibiting apoptosis (14,17,18). Disrupting HuR-ARE interaction is thus an attractive strategy in developing new cancer therapeutics (15,19-21).

Here, using a fluorescence polarization-based screening (21), we identified azaphilones (a class of compounds derived from the fungal natural product asperbenzaldeyde) (22,23) as potent inhibitors of HuR-ARE interaction *in vitro*. We characterized the HuR binding of the most potent azaphilone derivative, azaphilone-9 (AZA-9), by surface plasmon resonance (SPR), nuclear magnetic resonance (NMR), and computational modeling. AZA-9 disrupts HuR-ARE interaction by competitive binding in the RNA-binding cleft of HuR.

6.3. Methods

6.3.1. Protein expression and purification

The protein expression and purification of full length HuR (326 residues) and HuR RRM1/2 (residues 18 - 186) have been described ([21]). For NMR studies, in addition to ^{15}N-labeling, we also performed ILV-labeling, where the Isoleucine $C\delta 1$ and the geminal Leucine $C\delta$ and Valine $C\gamma$ methyl groups are ^{13}C-labeled by growing *E. coli* in M9 minimal media supplied with the appropriate ^{13}C alpha keto acids. His$_6$-tagged HuR RRM1/2 labeled with ^{15}N and ILV together was prepared by expression in *E. coli* BL21 (DE3) grown in 1 liter of M9 minimal media supplemented with 1 g of ^{15}N-ammonium chloride and 3 g of glucose at 37 °C. At OD$_{600}$ ~0.4, the growth medium was supplied with 60 mg of 2-ketobutyric acid-4-^{13}C (Sigma #571342) to label the $^{13}C\delta 1$ methyl group of isoleucine and 100 mg of 2-keto-3-(methyl-^{13}C)-butyric acid-4-^{13}C (Sigma #571334) to label the two leucine $^{13}C\delta$ and the two valine $^{13}C\gamma$ methyl groups ([24]). Approximately 1 hour later (at an OD$_{600}$ of ~0.8), the culture was induced with 0.7 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and cell growth was continued overnight in a 15 °C shaker incubator to a final OD$_{600}$ ~2.5. Cells were harvested by centrifugation (4000 rpm, 10 min), resuspended in binding buffer (30 ml; 500 mM NaCl, 20 mM Tris-HCl pH 8.0, 5 mM imidazole), and lysed by sonication in the presence of 0.1 mM phenylmethanesulfonylfluoride (PMSF). Cellular debris was removed by centrifugation (13,000 rpm, 10 min), and 600 μl of 5% (v/v) polyethyleneimine was added to the supernatant to precipitate the nucleic acids. Following centrifugation (13,000 rpm for 10 min), the supernatant was purified by nickel affinity chromatography. The supernatant was loaded to the Ni^{2+} column, washed with binding buffer and he His$_6$-tagged HuR RRM1/2 was eluted with elution buffer (500 mM NaCl, 20 mM Tris-HCl pH 8.0, 250 mM imidazole). The HuR RRM1/2 construct used herein retained an N-terminal His$_6$-tag. Purified protein was dialyzed in buffer (100 mM NaCl, 10 mM NaPO$_4$ pH 6.8) and concentrated using Amicon Ultra 3K centrifugal filter (Millipore). Protein concentration was measured by A$_{280}$.

6.3.2. ILV assignment

Mutagenesis was used to assign the ILV resonances by introducing conservative mutations (I→L, L→I, and V→I) in HuR RRM1/2. The QuikChange kit (Stratagene) was used to introduce site-directed mutants and mutations were verified by DNA sequencing. Proteins obtained from cell growth in 200 mL of M9 minimal media supplied with the appropriate ^{13}C-alpha keto acid was enough to obtain 2D ^{1}H-^{13}C HSQC to assign the ILV peaks of the 12 isoleucine and L39, L61, V66, and L138 residues of HuR RRM1/2.

6.3.3. Chemicals and reagents

Synthetic RNA oligos were from Dharmacon. For fluorescence polarization (FP) or AlphaLISA (ALPHA) assay, fluorescein-tagged or 3'-biotin modified ARE^{Msi1} oligo derived from the 3'-UTR of Musashi RNA-binding protein 1 (Msi1) with the sequence 5'-GCUUUUAUUUAUUUUG-3' was used. For NMR studies, the 11-mer ARE^{c-fos} RNA oligo (5'-AUUUUUAUUUU-3') identical to the c-fos RNA sequence used in the crystal structure of HuR-RNA complex ([8]) was used. Prior to the addition of RNA to the protein, the RNA was heated at 95 °C for 5 min followed by immediate cooling on ice for 5 min. The azaphilone derivatives used herein were obtained by semisynthetic diversification of asperbenzaldehyde, which was purified from a strain of *Aspergillus nidulans* that was engineered to overproduce this compound as described elsewhere ([22],[23]). Compounds were dissolved in dimethyl sulfoxide (DMSO) to form 10 mM stock solutions; for NMR studies, deuterated dimethyl sulfoxide (d$_6$-DMSO) was used.

6.3.4. Biochemical assays

FP and ALPHA competition assays for screening HuR inhibitors and validation of hits were carried out as reported ([21]). Briefly, compounds with increasing doses were added to plate wells prior to the addition of pre-formed protein-ARE^{Msi1} complex. This was followed by the addition of streptavidin-

coated donor beads and nickel chelate (Ni-NTA) acceptor beads in the ALPHA assay. Measurements were taken using a BioTek Synergy H4 hybrid plate reader (Biotek, Winooski, VT) after incubating for 2 hours at room temperature. IC_{50}, the drug concentration causing 50% inhibition, was calculated by sigmoid fitting of the dose response curve using GraphPad Prism 5.0. K_i values were calculated following method of Nikolovska-Coleska *et al.* ([25]). Percent inhibition was calculated by comparing to the DMSO (0% inhibition) and labeled free RNA only (100% inhibition) controls.

Surface plasmon resonance (SPR) datasets were acquired using a BIACORE 3000 (GE Healthcare) at 20°C as described ([21]). Briefly, HuR was immobilized into a CM5 chip by amine-coupling chemistry and AZA-9 dissolved in buffer (20 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.05% p20 (v/v), 5% DMSO (v/v)) was injected into the flow cell at a flow rate of 60 μL/min. The mixtures complexes were allowed to associate for 400 sec and dissociate for 160 sec. SPR sensorgrams were generated using Scrubber2 (BioLogic Software, Australia).

6.3.5. NMR spectroscopy

Proteins for NMR were dissolved in NMR buffer (100 mM NaCl, 10 mM $NaPO_4$ pH 6.8, 10% D_2O). ^{15}N NMR data were acquired using Bruker Avance 800 MHz spectrometer with a TCI cryoprobe. ILV ^{13}C NMR data were acquired using on a Bruker Avance III 600 MHz spectrometer. NMR data were acquired at 25 °C, processed using NMRPipe ([26]) and analyzed using NMRView ([27]). For ^{15}N and ILV chemical shift mapping, 80 μM ^{15}N/ILV-labeled HuR RRM1/2 was titrated with unlabeled ARE[c-fos] RNA at increasing molar ratios of 1:0, 1:0.5, 1:1, and 1:1.7. AZA-9 was titrated at 1:1 and 1:2 molar ratios into 50 μM ^{15}N/ ILV-labeled HuR RRM1/2 in NMR buffer with 10% d_6-DMSO. Higher protein concentrations or higher molar ratios of AZA-9 resulted in sample precipitation. The ^{15}N titrations were monitored by acquiring 2D ^1H-^{15}N TROSY spectra, and the ILV titrations were monitored by acquiring

2D ^{1}H-^{13}C HSQC spectra. The published backbone amide assignments of HuR RRM1/2 (BMRB entry 26628) ([19]) and the ILV assignments made herein were used in the NMR analysis.

6.3.6. Computational modeling

A docked model of how AZA-9 may bind to HuR RRM1/2 was built using the crystal structure of HuR RRM1/2 in complex with RNA (PDB 4ED5) ([8]). FRED (version 3.0.1) from OpenEye Software was used for molecular docking ([28]). For each of the biological units present in PDB 4ED5, the RNA was removed and a receptor was built using APOPDB2RECEPTOR (OpenEye Scientific Software, Santa Fe, NM) with residue Y63 specified as the active site residue. Conformers of AZA-9 were generated with OMEGA (version 2.5.1.4) ([29]) and docked into each of the prepared receptors using FRED. Full-atom minimization of the top-scoring model was then carried out using ROSETTA ([30]).

6.4. Results

6.4.1. *Protein expression and purification of HuR and HuR RRM1/2*

Full length HuR and a shorter HuR construct consisting only of the two tandem RRM1 and RRM2 (RRM1/2; residues 18-186) were overexpressed in *E. coli* BL21 (DE3) and purified under native conditions by Ni^{2+}-affinity chromatography. Typically, ~28 mgs of pure unlabeled or ^{15}N/ILV-labeled recombinant HuR RRM1/2 protein was obtained per liter of LB or M9 minimal medium. The HuR RRM1/2 was soluble in 100 mM NaCl, 10 mM sodium phosphate pH 6.8 buffer at a concentration of ~6 mg/ml (~0.27 mM protein concentration), beyond which it precipitated.

6.4.2. *Identification of AZA-9 as an inhibitor of HuR-RNA interaction*

A high throughput screen based on fluorescence polarization (FP) of HuR and a 16-mer ARE^{Msi1} RNA oligomer from the 3'-UTR of Musashi1 (Msi1) mRNA identified novel small molecule inhibitors of HuR-ARE interactions ([21]). This FP-based assay ([21]) was used here to screen ~2000 compounds from NCI (Diversity Set II, natural product set and approved oncology drugs) plus in-house compounds to identify inhibitors of HuR-ARE^{Msi1} interaction (screening data not shown). This screen identified the in-house fungal natural product azaphilones as potent inhibitors of HuR-ARE^{Msi1} interaction (**Figure 6-1**). Azaphilones are derived from a fungal natural product, asperbenzaldehyde, which exhibit diverse biological activities, including cytotoxic, anti-inflammatory, anti-proliferative, and anti-tumorigenic ([22],[31],[32]). Eight out of nine azaphilone derivatives (AZA-7 through AZA-15) screened in this assay showed significant dose-dependent inhibition of HuR-ARE^{Msi1} interaction (**Figure 6-1**). Similar inhibitory effects were also observed using other HuR target mRNAs, ARE^{Bcl2} and ARE^{XIAP} (data not shown). AZA-9 (**Figure 6-2A, B**) with a K_i value of 0.15 µM (*n*=3) was the most potent compound

128

among the nine azaphilone derivatives and thus was chosen for further characterization by surface plasmon resonance (SPR) and NMR methods. SPR confirmed the direct binding of AZA-9 to HuR (**Figure 6-2C**). Upon injections of increasing concentrations of AZA-9 on immobilized HuR, SPR sensorgrams showed increased optical response in a dose-dependent manner (**Figure 6-2C**).
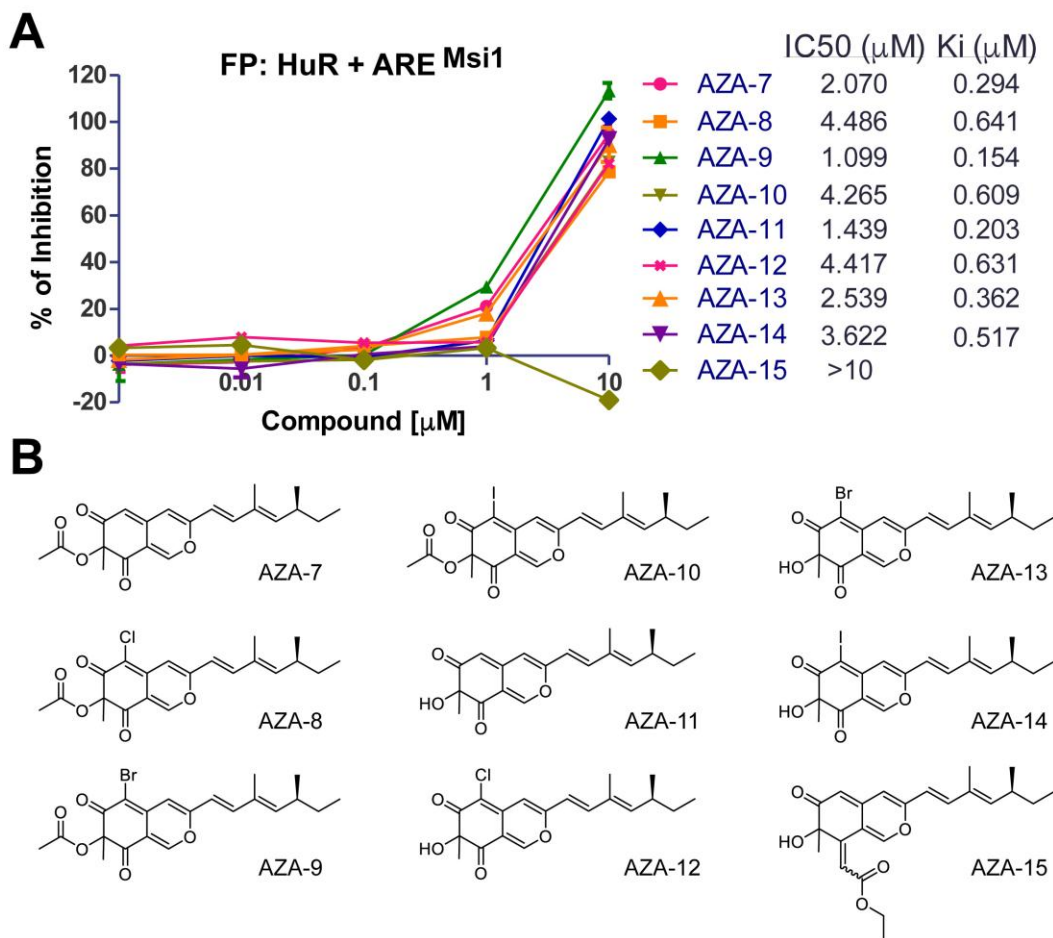
Figure 6-1. Azaphilones inhibit HuR-RNA interaction. (**A**) Dose-response curves of azaphilone compounds disrupting HuR-ARE^{Msi1} binding in FP assay using 10 nM HuR protein and 2 nM fluorescein-labeled Msi1 RNA. Data are representative of three independent experiments; IC_{50} and K_i values are mean from three independent experiments. (**B**) Structures of azaphilone compounds tested in FP assay. Data courtesy of Dr. Xiaoqing Wu.
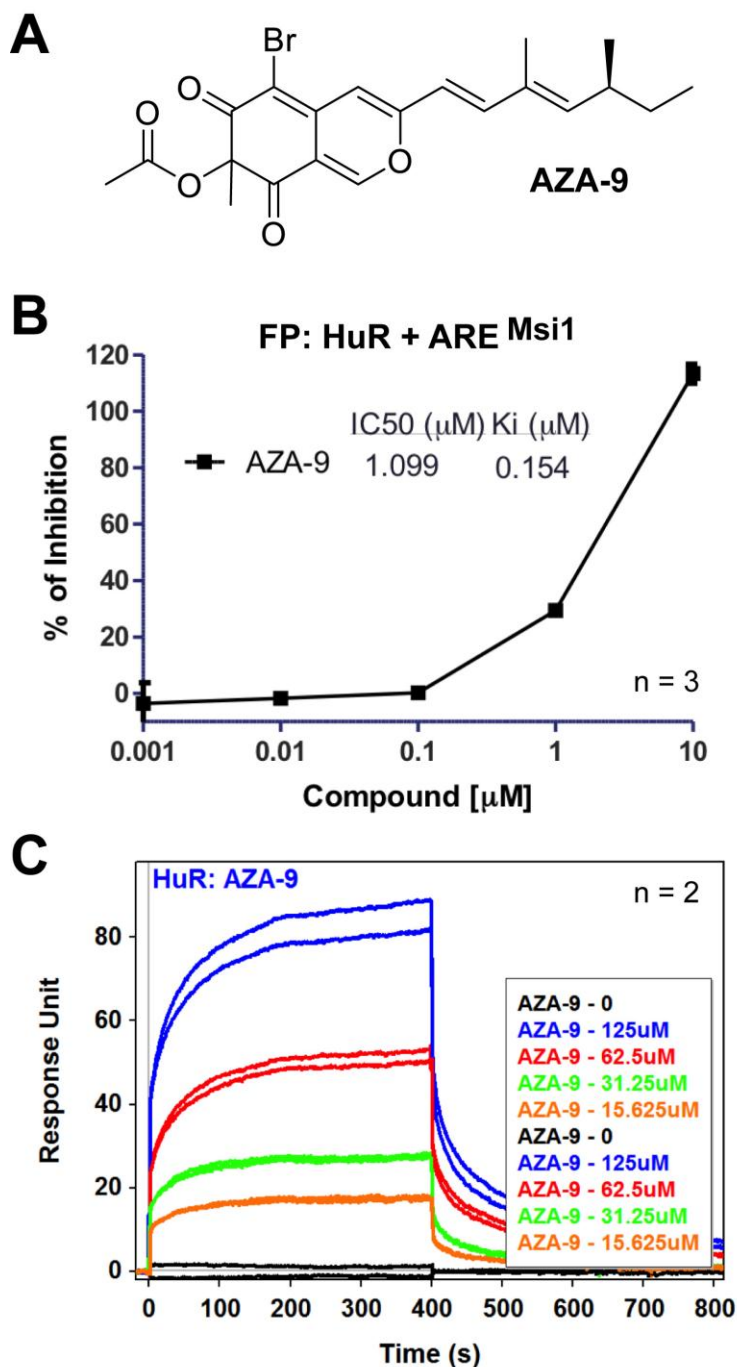
Figure 6-2. AZA-9 inhibits HuR-ARE^{Msi} RNA interaction and binds directly to HuR. (**A**) Structure of AZA-9. (**B**) Dose-response curve of AZA-9 disrupting HuR-ARE^{Msi1} binding in a fluorescence polarization-based assay using 10 nM HuR and 2 nM fluorescein-labeled ARE^{Msi1} (n = 3; number of independent experiments). (**C**) SPR sensorgrams of AZA-9 injected at increasing concentrations of 0-125 uM into a flow cell containing immobilized HuR (n=2). Data courtesy of Dr. Xiaoqing Wu (FP inhibition assay) and Dr. Lan Lan (SPR).

6.4.3. Use of HuR RRM1/2 in this study

Full length HuR gave non-ideal NMR spectra that made further NMR characterization challenging ([19]), we therefore used a shorter version of HuR comprising the tandem RRM1/2 domains (HuR RRM1/2). The smaller HuR RRM1/2 construct recapitulated the RNA binding and the inhibitory effect of AZA-9 as seen in the full length HuR. Results of FP assay ([21]) showed that AZA-9 significantly inhibited HuR RRM1/2-ARE[Msi1] binding with a sub-micromolar K_i value (**Figure 6-3A**). Likewise, results of AlphaLISA assay (ALPHA) ([21]) demonstrated dose-dependent HuR RRM1/2-ARE[Msi1] inhibition by AZA-9 (**Figure 6-3B**).
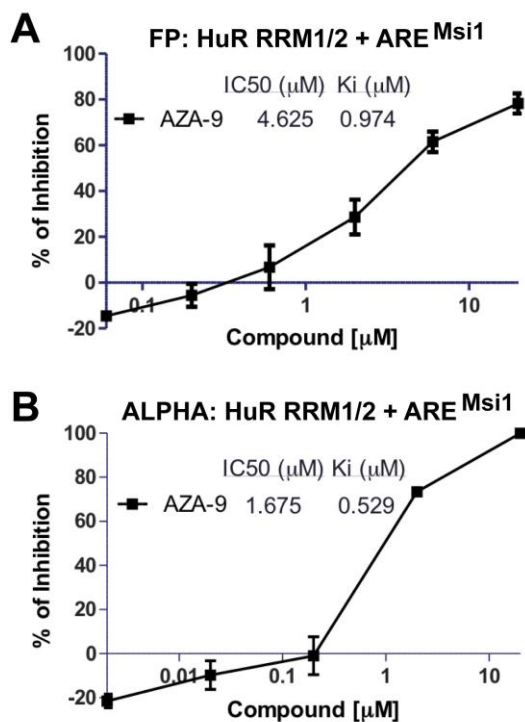
Figure 6-3. Compound AZA-9 displays dose-dependent inhibition with HuR RRM1/2 in FP and AlphaLISA assays. Dose-response curves of compound AZA-9 disrupting HuR RRM1/2-ARE[Msi1] binding in **(A)** FP assay using 25 nM HuR RRM1/2 protein and 2 nM fluorescein-labeled Msi1 RNA; and **(B)** ALPHA assay using 100 nM HuR RRM1/2 protein and 25 nM biotin-labeled Msi1 RNA. **A** and **B** are representative of three independent experiments. IC_{50} and K_i values are mean from three independent experiments. Data courtesy of Dr. Xiaoqing Wu.

6.4.4. NMR titrations of ^{15}N-labeled HuR RRM1/2 with $ARE^{c\text{-}fos}$

The 2D ^{1}H-^{15}N TROSY spectrum of HuR RRM1/2 **(Figure 6-4A)** closely resembled the reported 2D ^{1}H-^{15}N HSQC spectrum ([19]), thereby allowing use of the reported backbone amide assignments of HuR RRM1/2 ([19]) in our analysis. Overall, the 2D ^{1}H-^{15}N spectrum of HuR RRM1/2 is well dispersed with ~175 sharp, well-resolved peaks as expected for the construct **(Figure 6-4A)**. To gain insights into the interaction of HuR RRM1/2 and RNA in solution, we used the same 11-mer *c-fos* RNA oligo (ARE$^{c\text{-}fos}$) used in the co-crystallization of HuR RRM1/2 ([8]) for the NMR characterization of protein-RNA interaction. ^{15}N-labeled HuR RRM1/2 was titrated with increasing concentrations of unlabeled ARE$^{c\text{-}fos}$ at 1:0.5, 1:1, and 1:1.7 molar ratios, and the titration was monitored by acquiring 2D ^{1}H-^{15}N TROSY spectra. The stepwise addition of ARE$^{c\text{-}fos}$ induced mainly peak broadening of specific HuR RRM1/2 resonances **(Figure 6-4A)** indicating that the interaction occurred in intermediate exchange NMR time scale for the backbone amides. Protein-RNA contacts have been identified in the co-crystal structure of HuR RRM1/2-ARE$^{c\text{-}fos}$ ([8]). Some essential HuR residues identified contributing to the specific recognition of RNA substrate through side-chain, main-chain and/or stacking interactions include Y26, R97, I103, Y109, and R153 ([8]). These residues showed significant peak intensity reduction during the NMR titration **(Figure 6-4B)**. To identify the amino acids involved in ARE$^{c\text{-}fos}$ binding, we calculated the peak intensity ratio ($I_{1:1}/I_{1:0}$) for each non-overlapped peak at an HuR RRM1/2:RNA molar ratio of 1:1 **(Figure 6-4C)**. Residues with peak intensity ratio lower than average intensity minus one standard deviation were mapped onto the co-crystal structure of the protein-RNA complex **(Figure 6-4E)**. HuR RRM1/2 residues with significant peak intensity reduction cluster together in the β-strands, surrounding loops and linker region of RRM1/2 and form the RNA-binding surface of RRM1/2 **(Figure 6-4E)**. The RNA interaction surface determined by NMR is consistent with the crystal structure of HuR-RNA complex.

Although the HuR RRM1/2 residues directly in contact or in close proximity with the RNA showed reduction in peak intensities as described above, several residues distant from the RNA-binding site showed changes in peak positions in the presence of RNA (**Figure 6-4A**). Inspection on these peaks at lower contour level showed appearance of a new peak at a slightly different frequency with increasing concentrations of RNA (**Figure 6-4D**). The original peak from the RNA-free form of HuR RRM1/2 gradually broadens, while the new peak emerging from the RNA-bound form progressively gains intensity with increasing amounts of RNA. HuR RRM1/2 residues displaying such shifts were mapped onto the structure of RNA-bound HuR (**Figure 6-4E**, residues colored *yellow*). Because these residues are situated ~9-13 Å apart from the RNA-binding site, such perturbations can be most directly interpreted a result of conformational change upon RNA binding. The appearance of two peaks corresponding to HuR RRM1/2 residues distant from the RNA-binding site in the presence of RNA is suggestive of slow conformational switching of HuR upon RNA binding. The allosteric effects observed in solution are in agreement with crystal structures and SAXS analysis of RNA-free *versus* bound forms of HuR that demonstrate conformational changes play a major role in formation of a stable compact HuR RRM1/2-RNA complex ([8],[33]).

6.4.5. *NMR assignment of ILV-labeled HuR RRM1/2*

ILV labeling offer additional probes in NMR studies because of their high sensitivity. Similar to the amide signals shown above, perturbations of ^{13}C methyl peaks can be used to report on protein ligand interactions, conformational changes, structure and dynamics ([24],[34],[35]). Further, side-chain interactions are crucial for the RNA recognition of HuR and methyl-containing residues such as isoleucine, leucine, and valine (ILV) occur near the RNA cleft. We therefore used ILV-labeled HuR RRM1/2 here to probe the side-chain protein-RNA interaction. His$_6$-tagged HuR RRM1/2 contains 12 isoleucine, 15 leucine and 13 valine residues (the His-tag contributes 1 valine and 2 leucines). The 2D ^1H-^{13}C methyl HSQC

spectrum of HuR RRM1/2 **(Figure 6-5A)** showed 12 Ile peaks (corresponding to each of the δ1 methyl group of 12 isoleucine) within the spectral window ~8-16 ^{13}C ppm and 28 pairs of Leu and Val peaks (corresponding to the two δ1 and δ2 methyl groups of 15 leucines; and the two γ1 and γ2 methyl groups of 13 valines) within 19-28 ^{13}C ppm range. To assign the 16 ILV residues used herein, fifteen point mutations were introduced in HuR RRM1/2 (with isoleucine mutated to leucine; and leucine or valine mutated into isoleucine), and the resultant protein expressed and purified under native conditions. Comparable to the wild type construct, the mutant proteins were soluble at ~6 mg/ml in 100 mM NaCl, 10 mM sodium phosphate pH 6.8 buffer. To illustrate the assignment method to assign I23 for example, comparison of the 2D ^{1}H-^{13}C HSQC spectrum of the I23L point mutant that was selectively ^{13}C-labeled at the isoleucine δ1 methyl with the spectrum of the wild type protein **(Figure 6-6A)**, enabled the unambiguous assignment of the ^{13}C δ1 methyl resonance of I23. All the 12 isoleucine residues were assigned in addition to 3 leucines (L39, L61, and L138) and a valine residue, V66 **(Figure 6-5A)**. These 16 ILV probes are strategically located and provide overall coverage of the HuR RRM1/2 structure.

6.4.6. *Titrations of ILV-labeled HuR RRM1/2 with ARE$^{c\text{-}fos}$*

To monitor the effect of RNA binding on ILV methyl resonances of HuR RRM1/2, 2D ^{1}H-^{13}C HSQC spectra were acquired on various titration samples. Of the assigned side chain methyl resonances, peaks corresponding to I23, I52, and L61 of RRM1, I103 of inter-domain linker region, and I133, L138, and I152 of RRM2 disappeared from their free position and reappeared at a different frequency in the spectrum as new peaks for the protein-RNA complex **(Figure 6-5A)**. This indicates that the side-chains of these residues either directly mediates tight RNA-binding or are present in close proximity of the RNA and thus, experience strong perturbation in their local chemical environment upon RNA-binding. The affected ILV residues (I23, I52, L61, I103, I133, L138, and I152) delineate the RNA-binding cavity of HuR RRM1/2 and lie within 3-4 Å of the RNA substrate **(Figure 6-5C)**.
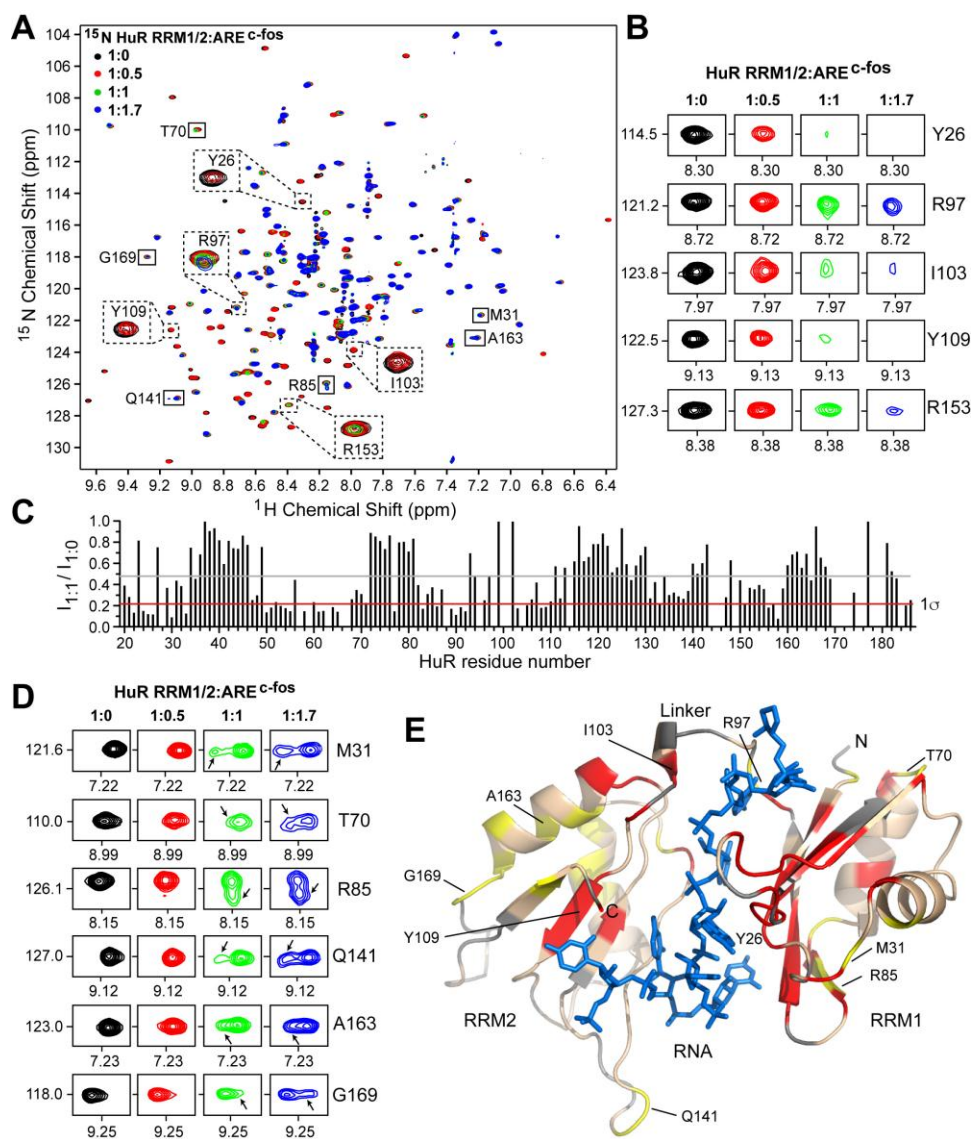
Figure 6-4. Amide titrations of HuR RRM1/2 with RNA. (**A**) Overlay of four 2D ^{1}H-^{15}N TROSY spectra of ^{15}N HuR RRM1/2 titrated with increasing molar ratios of ARE$^{c\text{-}fos}$ RNA. Representative residues showing peak broadening (dashed box) and residues displaying changes in peak positions (solid box) are shown. (**B**) Representative residues displaying peak broadening upon RNA binding are shown at similar contour levels. (**C**) Plot of relative peak intensity for all non-overlapping HuR RRM1/2 resonances in the ligand bound versus free state ($I_{1:1}/I_{1:0}$). Gray and red lines correspond to the mean and one standard deviation from the mean (1σ), respectively. (**D**) Representative residues that displayed changes in peak positions are shown at similar contour level at individual titration points. While the original peak (RNA-free form) gradually decreases in intensity, a new peak (shown by arrow) appears and progressively gains intensity with increasing concentrations of RNA. (**E**) Results of NMR titrations mapped onto the co-crystal structure of HuR-ARE$^{c\text{-}fos}$ complex (PDB 4ED5) and colored as follows: RRM1/2 residues with peak intensity ratio ($I_{1:1}/I_{1:0}$) lower than 1σ (red), residues with new peaks shown in **D** (yellow), unassigned residues (gray), and RNA (blue).
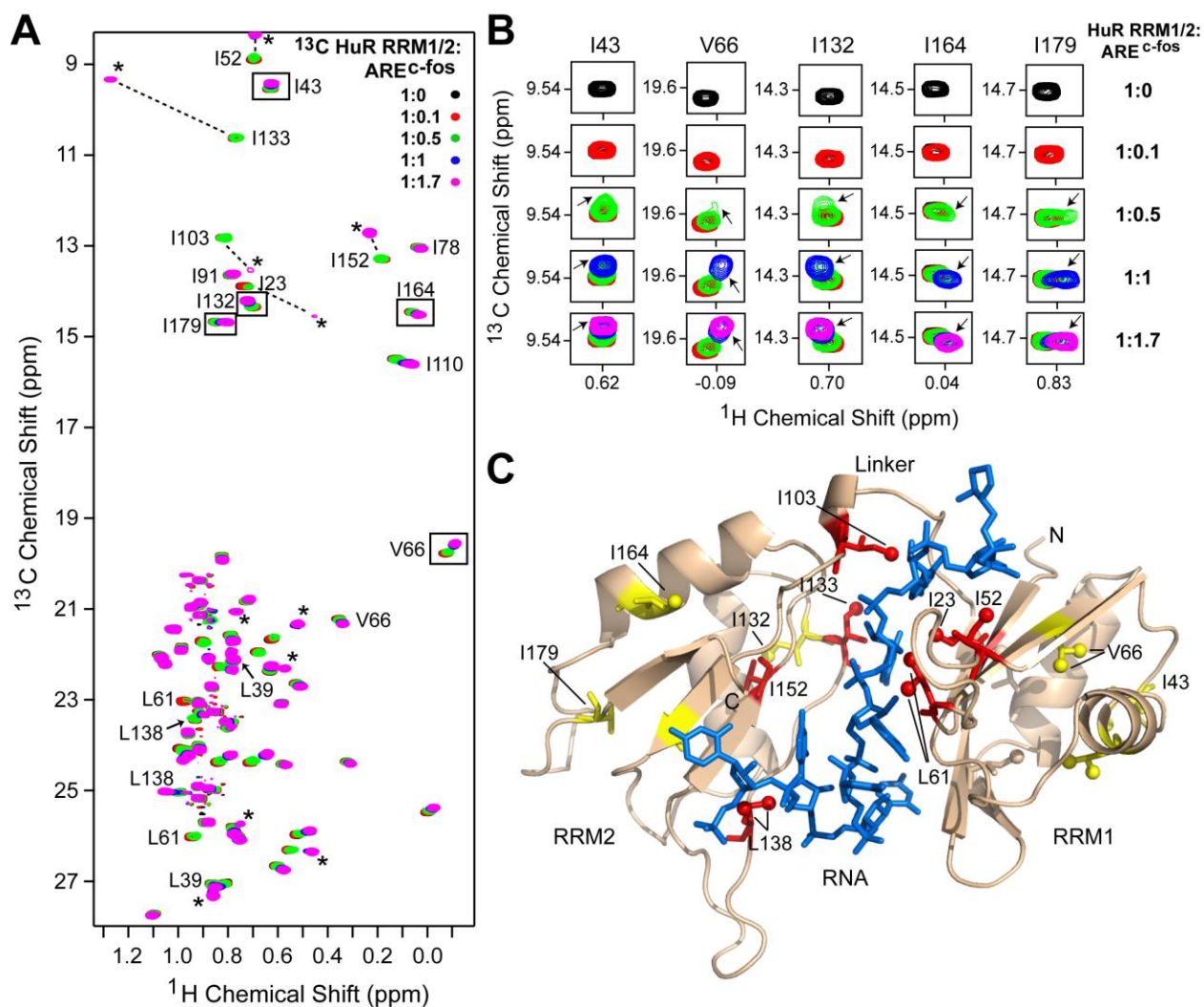
135

Figure 6-5. ILV titrations of HuR RRM1/2 with RNA. (A) Overlay of five 2D ^{1}H-^{13}C HSQC spectra of ILV-labeled HuR RRM1/2 titrated with increasing concentrations of unlabeled ARE[c-fos]. Assigned ILV peaks, the new peaks appearing upon complex formation (*), and Some representative ILV residues away from the RNA-binding site that showed chemical shift perturbations (boxed) are indicated. **(B)** The boxed peaks in **A** are shown at similar contour levels at individual titration points. The new peak emerging with increasing concentrations of RNA is shown by an *arrow*. **(C)** Results of NMR titration mapped onto the crystal structure of HuR-ARE**c-fos** complex (PDB 4ED5). Protein and RNA are represented as in Figure 6-4. Assigned ILV residues (red) are shown as sticks (Ile) and with spheres (Leu and Val) and colored as follows: RRM1/2 residues showing new complex peaks (red) and residues with chemical shift perturbations as indicated in **B** (yellow).

136
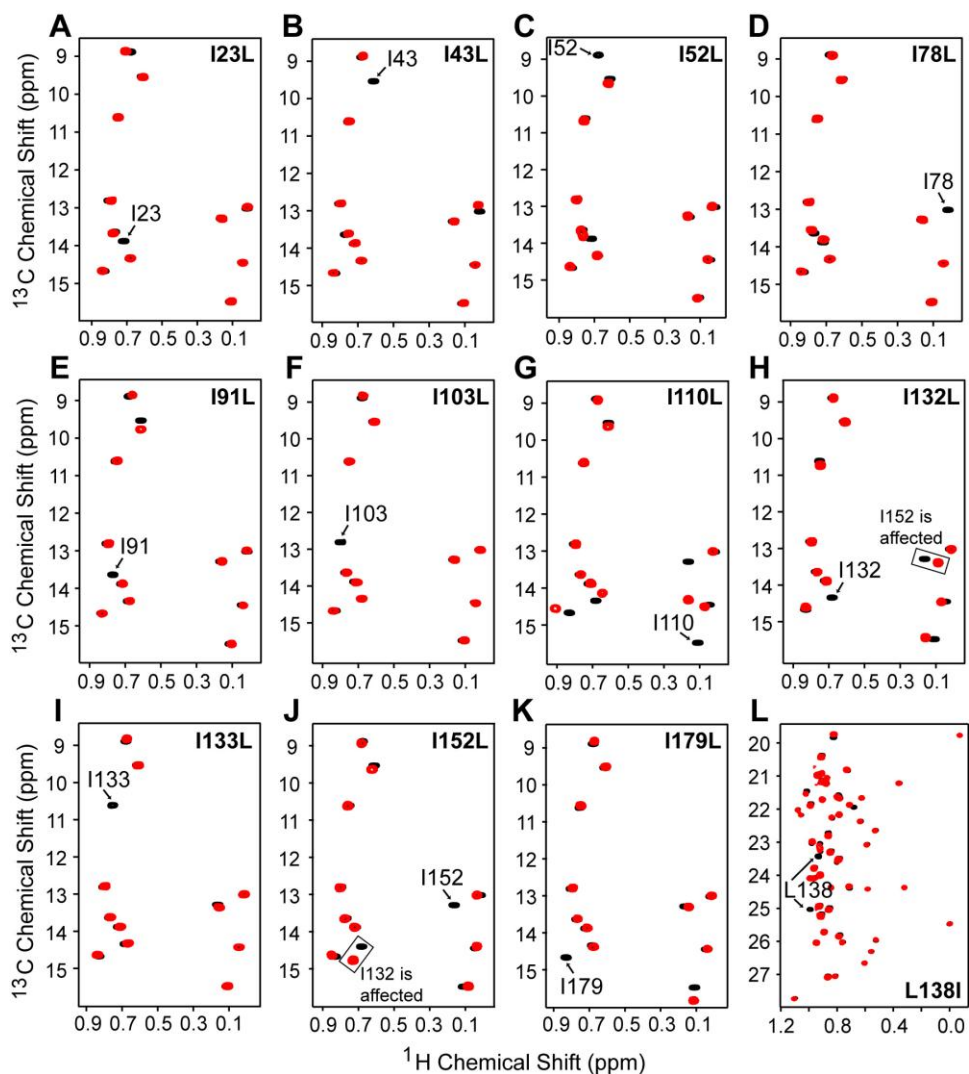
Figure 6-6. Single amino acid site-directed mutagenesis approach used in assigning selective methyl peaks of HuR RRM1/2. 2D ^{1}H-^{13}C methyl HSQC spectra for HuR RRM1/2 mutants (red) overlaid with the wild type reference spectra (black). **(A-K)** A single missing ^{13}Cδ1 methyl peak for Ile; and **(L)** Two ^{13}Cδ1 and ^{13}Cδ2 methyl peaks for Leu corresponding to the mutated residue are labeled.

Analogous to the allosteric effects observed in the backbone amide titrations **(Figure 6-4D)**, several residues, such as L39, I43, V66 of RRM1, and I110, I132, I164, I179 of RRM2, with the side chain methyl group that are positioned ~12-18 Å from the RNA binding site also showed changes in their chemical shift positions **(Figure 6-5B)**. The original free peak disappeared and a new peak representing the RNA-bound form of HuR appeared at a slightly different position **(Figure 6-5B)**. After 1:1 complex

is reached, it was observed that the addition of RNA merely adds to the intensity of the peak of the bound form. Residues with such perturbations are highlighted as *yellow* sticks in **Figure 6-5C**. The appearance of new peaks for such distant residues in the RNA-bound form confirms conformational changes upon RNA-binding. The results of side-chain ILV methyl titrations complement backbone amide titrations and together, they support that the β-sheet region of HuR RRM1/2 is the RNA binding surface and RNA binding is accompanied by conformational changes in HuR.

6.4.7. NMR titrations of HuR RRM1/2 with AZA-9

The interaction of AZA-9 with HuR RRM1/2 was characterized by NMR methods. Titrations of ^{15}N/ILV-labeled HuR RRM1/2 with AZA-9 resulted in concentration dependent reduction in the ^{15}N and ILV peak intensities of HuR RRM1/2 **(Figure 6-7)**, indicating complex formation on an intermediate exchange time scale. In addition to the decrease in peak intensities, the side chain methyl groups of some HuR RRM1/2 residues, such as I103, L138 **(Figure 6-7D)**, also showed chemical shift deviations upon binding of AZA-9. Further, key RNA-binding residues, Y26, R97, I103, Y109, and R153 **(Figure 6-4A)**, exhibited significant peak broadening with increasing doses of AZA-9 in the ^{15}N-titrations **(Figure 6-7B)**. Specifically, residue R97 in the inter-domain linker region (whose intensity reduced by ~50% at equimolar concentration of AZA-9, **Figure 6-7C**) has been previously reported as important for RNA binding, RNA recognition and high affinity HuR-RNA complex formation ([8]). Comparable to the results of the amide titrations, results of the ILV titrations of other RNA-binding residues, such as I52, L61, I103, and L138, identified earlier **(Figure 6-5A)** also displayed significant reduction in peak intensities upon addition of AZA-9 **(Figure 6-7D)**. A plot of the peak intensity ratio ($I_{1:1}/I_{1:0}$) in the bound and free form at 1:1 molar ratio **(Figure 6-7C, 6-7E)** revealed that the NMR resonances of the major RNA-binding residues of HuR, including K55, G62, R97, I103, L138, and R153, were significantly perturbed by AZA-9. Residues that were significantly perturbed during titrations were mapped on the structure of RNA-

bound HuR **(Figure 6-7F)**. Complex formation with AZA-9 primarily affected a cluster of RNA-binding residues located near the inter-domain linker region of HuR **(Figure 6-7F)**. NMR results indicate that AZA-9 interacts at the same binding pocket that HuR RRM1/2 uses to bind its target RNA.

6.4.8. *In silico docking of AZA-9 in the RNA cleft of HuR RRM1/2*

Molecular docking studies were performed to gain further insight into the binding mode of compound AZA-9 to HuR. **Figure 6-8A** represents the top-scoring computational model of compound AZA-9 bound HuR RRM1/2 generated using FRED ([28](#)) following full-atom minimization with ROSETTA ([30](#)). Consistent with the NMR-derived binding site of AZA-9 **(Figure 6-7E)**, molecular docking confirmed a feasible binding mode for compound AZA-9 in the RNA binding cleft of HuR near the inter-domain linker region **(Figure 6-8A)**. Several hydrophobic and positively charged HuR residues, such as Y26, K55, R97, and R153 line the pocket for AZA-9 and potentially stabilize the protein-ligand complex through electrostatic, hydrogen bond, hydrophobic, and pi-stacking interactions **(Figure 6-8A)**. Together, results of NMR titrations and molecular docking indicate that compound AZA-9 disrupts HuR-RNA interaction by competitively binding in the RNA cleft of HuR **(Figure 6-8)**.

6.5. Discussion

HuR-ARE interaction ([1](#),[4-6](#)) contributes to carcinogenesis by stabilizing the mRNAs of oncogenes ([7](#),[11](#),[14](#),[17](#),[18](#)), thus, finding inhibitors of HuR-ARE interaction could contribute in the development of new cancer therapy ([19](#),[21](#)). So far, there has been limited success in the discovery of HuR inhibitors that competitively bind to HuR that directly disrupt the HuR-ARE interactions ([19-21](#)). Currently, the most potent HuR inhibitor known, MS-444, is a bacterial natural product isolated from

Actinomyces sp. microbial broths, and MS-444 inhibits HuR-RNA interaction by interfering with HuR homodimerization ([20]). Here, we identified a new class of compounds, azaphilones **(Figure 6-1)** ([31],[32]) and in particular, AZA-9 **(Figure 6-2)**, as novel inhibitors of HuR-ARE interaction.

Azaphilones are derived from the fungal secondary metabolite asperbenzaldehyde ([22],[23]). Fungi-derived natural products are excellent sources of pharmaceuticals and many fungal secondary metabolites show anti-cancer properties that inhibit cell proliferation, angiogenesis, and tumorigenesis ([31],[36],[37]). The two rings of azaphilones form the isochromene scaffold, and this scaffold is present in the previously identified methyl-benzoisochromene scaffold of the chrysanthones secondary fungal metabolites isolated from another fungi, *Ascochyta chrysanthemi* ([37]). Chrysanthones were reported to have anti-proliferative, anti-tumorigenic and anti-angiogenic properties, however, their specific molecular target was not determined ([37]). The presence of the isochromene scaffold plus the observed anti-cancer properties of chrysanthones could be associated to HuR inhibition.

The tandem RRM1/2 of HuR is the minimal domain needed for binding AREs. Our results of the backbone amide titrations showed significant peak broadening for the ARE-binding residues similar to what was reported by Wang et al. ([19]) **(Figure 6-4)**, however, we also observed additional new slow exchange peaks in the ILV titrations **(Figure 6-5)**. This differing NMR exchange behavior could be due to the direct interactions of the protein side chains with the RNA substrate and their dominant role in the formation of tight HuR RRM1/2-ARE complex ([8]). Additionally, we observed allosteric effects in HuR RRM1/2 occurring upon RNA binding **(Figure 6-4D and Figure 6-5B)**. Conformational changes in HuR RRM1/2 on binding the RNA substrate have been previously reported ([8],[33]). These conformational changes contribute in the formation of a stable, high-affinity HuR RRM1/2-RNA complex. Consistent with the crystal structures and SAXS analysis ([8],[33]), our NMR results identified the specific residues (such as M31, T70, R85, Q141, A163, and G169) that are involved in the slow conformational switching in HuR RRM1/2 upon RNA binding **(Figure 6-4 and Figure 6-5)**.
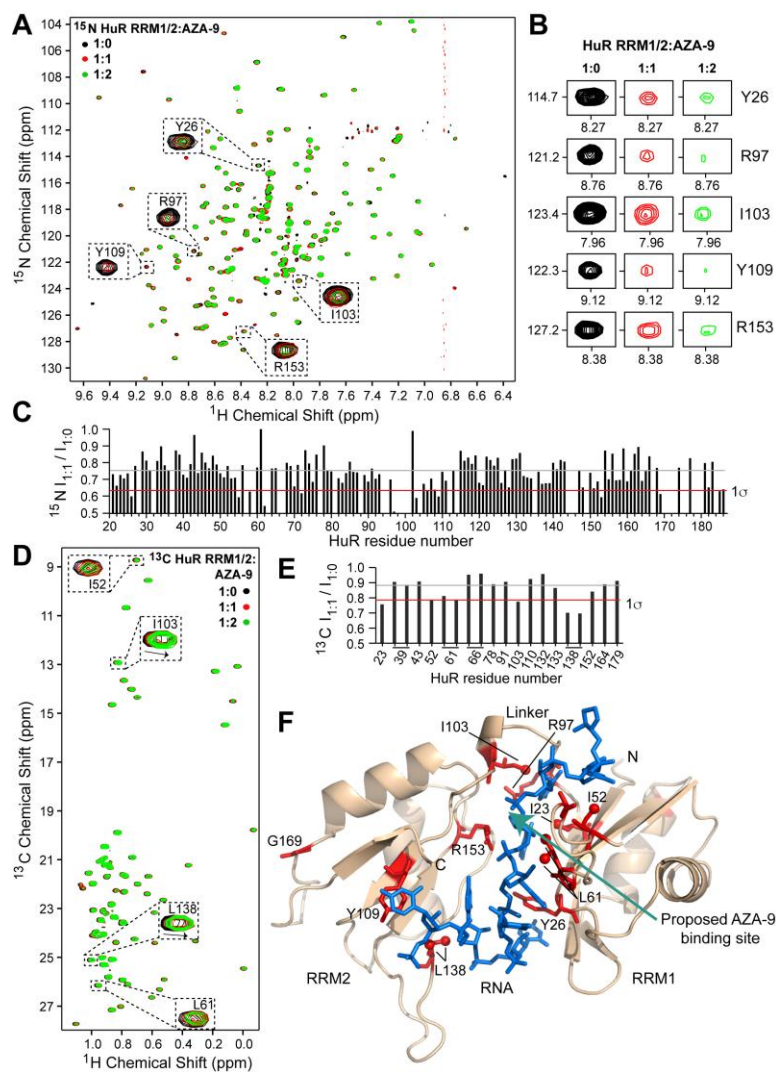
Figure 6-7. NMR titration of HuR RRM1/2 with AZA-9. (**A**) Overlay of three 2D ^{1}H-^{15}N TROSY spectra of ^{15}N HuR RRM1/2 titrated with increasing molar ratios of AZA-9. The peaks of some critical RNA-binding residues that undergo significant line broadening upon addition of AZA-9 are shown (expanded dashed box). (**B**) Representative residues showing peak broadening upon titration of AZA-9 are shown at similar contour levels at individual titration points. (**C**) Relative peak intensity plot for non-overlapping amide resonances of HuR RRM1/2 in the ligand bound versus free state ($I_{1:1}$ /$I_{1:0}$). (**D**) Overlay of three 2D ^{1}H-^{13}C HSQC spectra of ILV-labeled HuR RRM1/2 titrated with increasing molar ratios of AZA-9. Analogous to ^{15}N-titrations, ILV methyl groups of RNA-binding residues showed peak broadening with a few residues such as I103 and L138 also displaying chemical shift deviations (dashed box). (**E**) Relative peak intensity plot for assigned ILV methyl HuR RRM1/2 resonances in the ligand bound versus free state ($I_{1:1}$ /$I_{1:0}$). (**F**) Results of titrations mapped onto the co-crystal structure of HuR-ARE$^{\text{c-fos}}$ complex (PDB 4ED5). Protein and RNA are shown as in Figure 6-4 and 6-5, and depicted as follows: RRM1/2 residues with peak intensity ratio ($I_{1:1}$ /$I_{1:0}$) lower than 1σ (red) and the proposed AZA-9 binding site (arrow). Most of the AZA-9 affected residues are key RNA-binding residues. (**C,E**) Gray and red lines correspond to the mean and one standard deviation from the mean (1σ), respectively.
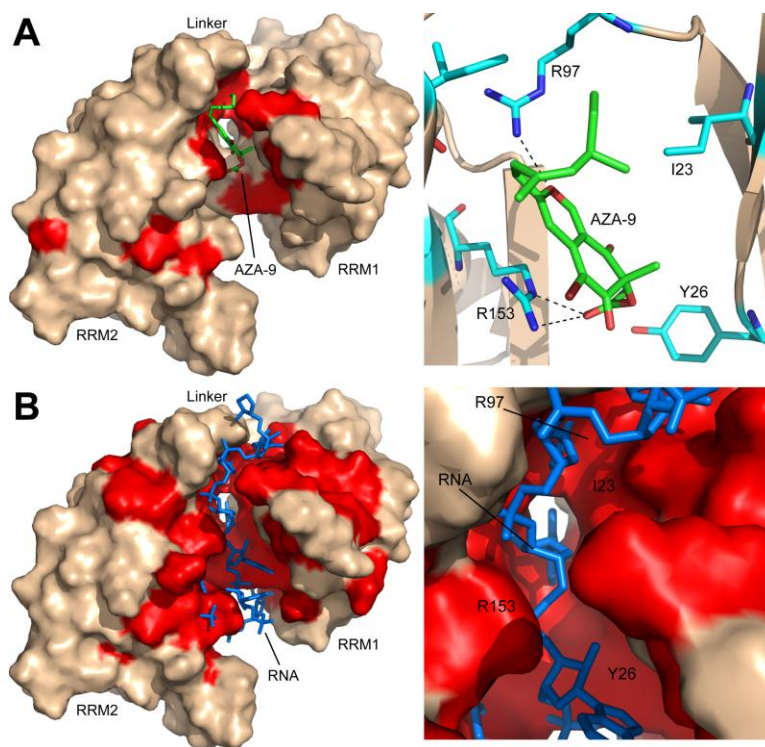
Figure 6-8. Molecular docking identified a possible binding pocket for AZA-9 in the RNA-cleft of HuR. **(A)** Computational model of AZA-9 bound to HuR RRM1/2, with protein shown in surface representation and AZA-9 in sticks. HuR RRM1/2 residues affected in AZA-9 NMR titrations are colored *red*. An expanded view of the drug-binding pocket is shown in the right. Residues involved in AZA-9 binding are displayed as sticks, with hydrogen bonds shown as dashed lines. **(B)** Surface representation of HuR RRM1/2 bound to RNA (*blue* sticks) for comparison with the AZA-9 model. Residues affected upon RNA titrations are highlighted in *red*. An expanded view of the RNA-pocket is shown in the right. Residues (Y26, K55, L61, R97, R153) perturbed by AZA-9 in the NMR titrations are critical RNA-binding residues. Computational modeling by Dr. David Johnson.

The crystal structures of the free and RNA-bound HuR RRM1/2 ([8]) suggest major structural rearrangements in the relative orientation of RRM1 with respect to RRM2 upon RNA binding (**Figure 6-9**). For example, RRM2 have to swing about 41 Å to reposition itself vis-à-vis RMM1 upon RNA binding (**Figure 6-9**). This major conformational change in HuR RRM1/2 upon RNA binding is not reflected in the results of our NMR titrations as well as the results of the amide titrations of Wang et al. ([19]). For such major conformational rearrangements of the two RRM domains, one would expect major changes in the peak positions in the ^{15}N TROSY spectra of the free and RNA-bound HuR RRM1/2

(**Figure 6-4**).   Instead, we observed essentially similar peak positions in the ^{15}N TROSY of free and RNA-bound HuR RRM1/2 (**Figure 6-4**), with the peak intensities of the RNA-bound form progressively weakening upon addition of more RNA.   In the ILV-titrations (**Figure 6-5**), there were indeed new slow-exchange peaks for RRM2 isoleucine residues (I103, I133, I152) but the rest of the isoleucines in RRM2 (I110, I164, and I179) were essentially in similar (fast exchange) peak positions as the free form suggesting the changes in the peak positions and intensities observed by NMR are due to RNA-binding rather than the major conformational rearrangements of the two domains upon RNA-binding.   Our NMR results suggest that RRM1 and RRM2 are somewhat 'pre-formed' for RNA-binding with the two domains already close together and poised to accept the RNA.   Upon RNA-binding, the side chains and loops of HuR RRM1/2 'wiggle' to accommodate and interact with the RNA.   SAXS results suggest two populations of free HuR RRM1/2 where one population has an extended structure (with a size of 74 Å) and another population that has a more compact structure (with a size of 56 Å) ([33]).   Upon RNA binding, the HuR RRM1/2 becomes even more compact (with a size of 51 Å).   This suggest that SAXS is able to trap two populations of free HuR RRM1/2 whereas our NMR results suggest an average conformation that is somewhat poised for RNA binding.

Our efforts to co-crystallize AZA-9 with HuR RRM1/2 have been so far unsuccessful, thus, we used NMR methods to characterize how AZA-9 interacts with HuR RRM1/2.   Results of NMR titrations showed that AZA-9 essentially perturbs the same HuR RRM1/2 residues involved in binding RNA (**Figure 6-7**).   NMR titrations with AZA-9 affected specific RNA-binding residues of HuR RRM1/2 (**Figure 6-7**).   The previously identified key ARE-binding residues, including Y26, L61, R97, I103, Y109, and R153 that make side chain, main chain and/or stacking interactions with the ARE substrate ([8]), showed significant perturbations upon the addition of AZA-9 (**Figure 6-7A, 6-7B**).   Particularly, R97 in the inter-domain linker region showing strongest peak intensity reduction has been identified through mutagenesis as a critical residue required for high affinity ARE binding ([8]).   Together, results of both backbone amide and ILV titrations indicate that AZA-9 predominantly affects a surface near the inter-

domain linker region in the RNA cleft of HuR **(Figure 6-7F)**. Although few ARE-binding residues (Y109, L138) in the lower region of the RNA-cleft also showed perturbations, this could be due to conformational changes induced by AZA-9 upon binding in the upper portion of the cleft or a second cryptic AZA-9 binding site, others cluster primarily at a single surface on HuR RRM1/2.
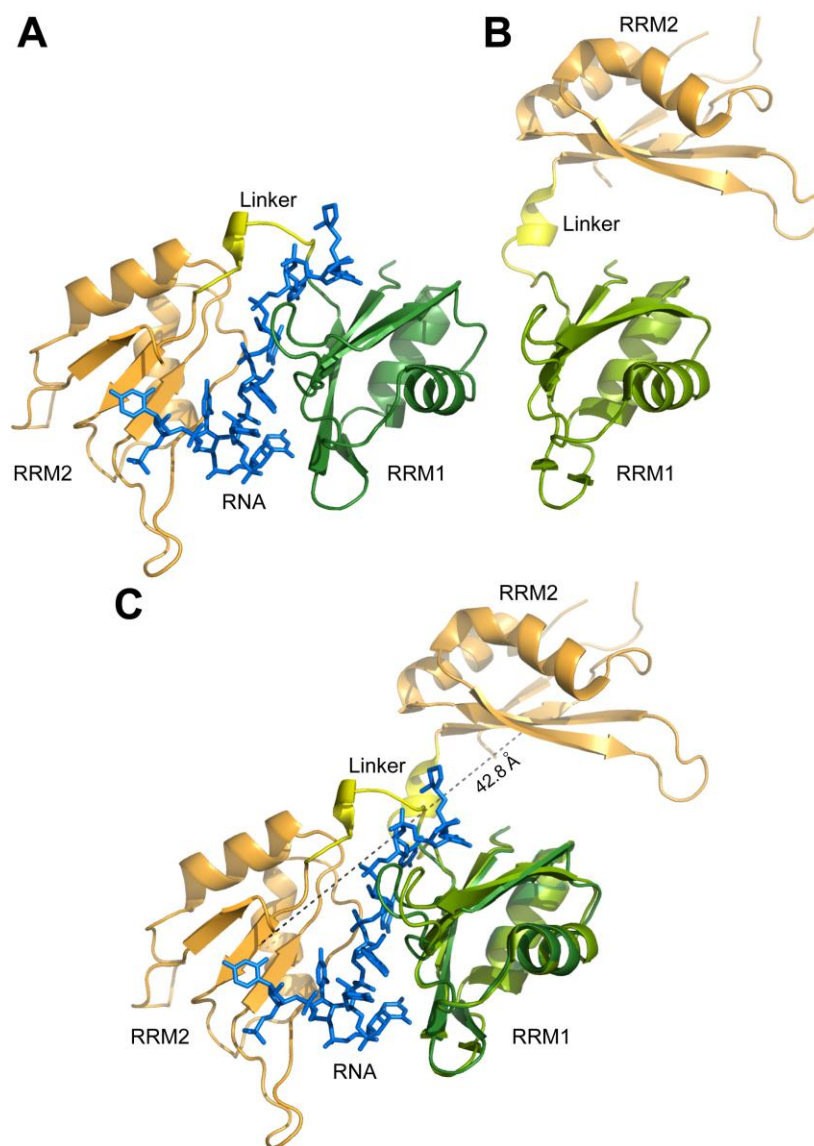

Figure 6-9. Comparison of the crystal structure of (**A**) RNA-bound, PDB ID: 4ED5; and (**B**) RNA-free HuR RRM1/2, PDB ID: 4EGL. (**C**) Superposition of RRM1 from the RNA-bound and RNA-free forms of HuR RRM1/2. RRM1 is colored *green*, RRM2 is *orange*, and RNA is *blue*.

In agreement with the NMR-derived binding site of AZA-9, results of molecular docking positioned AZA-9 in the ARE-binding cleft near the inter-domain linker region of HuR RRM1/2 (**Figure 6-8A**). Computational modeling suggested that AZA-9 is well situated in the binding pocket surrounded by several positively charged and hydrophobic residues to enable hydrophobic, hydrogen bond, and/or electrostatic interactions. In the computational model, the long hydrophobic tail of AZA-9 lies adjacent to the methyl side chain of I23, and runs roughly parallel to the aliphatic chain of R97 to promote hydrophobic interactions; the oxygen atom of the pyran ring in AZA-9 positioned such that it can form a hydrogen bond with the guanidinium group of R97; and the ester carbonyl of AZA-9 lies very close to the guanidinium side chain of R153 to form a salt bridge. The ability of AZA-9 to interact with R97 is particularly significant as R97 is a key residue for stable HuR-ARE complex formation as identified by crystallography. Overall, our results show that AZA-9 competes with target mRNAs for binding in the RNA cleft of HuR RRM1/2

To summarize, we discovered azaphilones, which are derived from fungal natural products, as novel disruptors of HuR-RNA interaction. Results of SPR, NMR and molecular docking confirmed that the most potent azaphilone derivative, AZA-9, inhibited HuR-mRNA interaction by competing directly for the RNA-binding site in HuR RRM1/2.

6.6. References

1. Brennan, C. M., and Steitz, J. A. (2001) HuR and mRNA stability. *Cell. Mol. Life Sci.* **58**, 266-277
2. Ross, J. (1995) mRNA stability in mammalian cells. *Microbiol. Mol. Biol. Rev.* **59**, 423-450
3. Fan, X. C. (1998) Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the invivo stability of ARE-containing mRNAs. *EMBO J* **17**, 3448-3460
4. Peng, S. S. Y. (1998) RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein. *EMBO J* **17**, 3461-3470

5.   de Silanes, I. L., Zhan, M., Lal, A., Yang, X., and Gorospe, M. (2004) Identification of a target RNA motif for RNA-binding protein HuR. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 2987-2992

6.   Cheng, S. (1996) Cloning and Characterization of HuR, a Ubiquitously Expressed Elav-like Protein. *J. Biol. Chem.* **271**, 8144-8151

7.   Dixon, D. A., Tolley, N. D., King, P. H., Nabors, L. B., McIntyre, T. M., Zimmerman, G. A., and Prescott, S. M. (2001) Altered expression of the mRNA stability factor HuR promotes cyclooxygenase-2 expression in colon cancer cells. *J. Clin. Invest.* **108**, 1657-1665

8.   Wang, H., Zeng, F., Liu, Q., Liu, H., Liu, Z., Niu, L., Teng, M., and Li, X. (2013) The structure of the AREbinding domains of Hu antigen R (HuR) undergoes conformational changes during RNA binding. *Acta Cryst.* **D69**, 373-380

9.   Fialcowitz-White, E. J., Brewer, B. Y., Ballin, J. D., Willis, C. D., Toth, E. A., and Wilson, G. M. (2007) Specific protein domains mediate cooperative assembly of HuR oligomers on AU-rich mRNA-destabilizing sequences. *J. Biol. Chem.* **282**, 20948-20959

10.   Fan, X. C., and Steitz, J. A. (1998) HNS, a nuclear-cytoplasmic shuttling sequence in HuR. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15293-15298

11.   de Silanes, I. L., Fan, J., Yang, X., Zonderman, A. B., Potapova, O., Pizer, E. S., and Gorospe, M. (2003) Role of the RNA-binding protein HuR in colon carcinogenesis. in *Oncogene*

12.   Yi, X., Zhou, Y., Zheng, W., and Chambers, S. K. (2009) HuR expression in the nucleus correlates with high histological grade and poor disease-free survival in ovarian cancer. *Aus. NZ J. Ob. Gyn.* **49**, 93-98

13.   Heinonen, M., Fagerholm, R., and Aaltonen, K. (2007) Prognostic Role of HuR in Hereditary Breast Cancer. *Clin. Cancer Res.* **13**, 6959-6963

14.   Vo, D. T., Abdelmohsen, K., Martindale, J. L., Qiao, M., Tominaga, K., Burton, T. L., Gelfond, J. A. L., Brenner, A. J., Patel, V., Trageser, D., Scheffler, B., Gorospe, M., and Penalva, L. O. F. (2012) The Oncogenic RNA-Binding Protein Musashi1 Is Regulated by HuR via mRNA Translation and Stability in Glioblastoma Cells. *Mol. Cancer Res.* **10**, 143-155

15.   Zucal, C., D'Agostino, V., Loffredo, R., Mantelli, B., NatthakanThongon, B. S. P., Lal, P., Latorre, E., and Provenzani, A. (2015) Targeting the Multifaceted HuR Protein, Benefits and Caveats. *Curr. Drug Targets* **16**, 499-515

16.   Denkert, C., Koch, I., von Keyserlingk, N., Noske, A., Niesporek, S., Dietel, M., and Weichert, W. (2006) Expression of the ELAV-like protein HuR in human colon cancer: association with tumor stage and cyclooxygenase-2. *Mod. Pathol.* **19**, 1261-1269

17.   Zhang, X., Zou, T., Rao, J. N., Liu, L., Xiao, L., Wang, P.-Y., Cui, Y.-H., Gorospe, M., and Wang, J.-Y. (2009) Stabilization of XIAP mRNA through the RNA binding protein HuR regulated by cellular polyamines. *Nuc. Acids Res.* **37**, 7623-7637

18.   Ishimaru, D., Ramalingam, S., Sengupta, T. K., Bandyopadhyay, S., Dellis, S., Tholanikunnel, B. G., Fernandes, D. J., and Spicer, E. K. (2009) Regulation of Bcl-2 Expression by HuR in HL60 Leukemia Cells and A431 Carcinoma Cells. *Mol. Cancer Res.* **7**, 1354-1366

19.   Wang, Z., Bhattacharya, A., and Ivanov, D. N. (2015) Identification of Small-Molecule Inhibitors of the HuR/RNA Interaction Using a Fluorescence Polarization Screening Assay Followed by NMR Validation. *PLoS One* **10**, e0138780

20.   Meisner, N.-C., Hintersteiner, M., Mueller, K., Bauer, R., Seifert, J.-M., Naegeli, H.-U., Ottl, J., Oberer, L., Guenat, C., Moss, S., Harrer, N., Woisetschlaeger, M., Buehler, C., Uhl, V., and Auer, M. (2007) Identification and mechanistic characterization of low-molecular-weight inhibitors for HuR. *Nat. Chem. Biol.* **3**, 508-515

21.   Wu, X., Lan, L., Wilson, D. M., Marquez, R. T., Tsao, W.-c., Gao, P., Roy, A., Turner, B. A., McDonald, P., Tunge, J. A., Rogers, S. A., Dixon, D. A., Aubé, J., and Xu, L. (2015) Identification and Validation of Novel Small Molecule Disruptors of HuR-mRNA Interaction. *ACS Chem. Biol.* **10**, 1476-1484

22. Paranjape, S. R., Riley, A. P., Somoza, A. D., Oakley, C. E., Wang, C. C. C., Prisinzano, T. E., Oakley, B. R., and Gamblin, T. C. (2015) Azaphilones Inhibit Tau Aggregation and Dissolve Tau Aggregates in Vitro. *ACS Chem. Neurosci.* **6**, 751-760

23. Somoza, A. D., Lee, K.-H., Chiang, Y.-M., Oakley, B. R., and Wang, C. C. C. (2012) Reengineering an azaphilone biosynthesis pathway in Aspergillus nidulans to create lipoxygenase inhibitors. *Org. Lett.* **14**, 972-975

24. Tugarinov, V., and Kay, L. E. (2005) Methyl Groups as Probes of Structure and Dynamics in NMR Studies of High-Molecular-Weight Proteins. *Chembiochem : a European journal of chemical biology* **6**, 1567-1577

25. Nikolovska-Coleska, Z., Wang, R., Fang, X., Pan, H., Tomita, Y., Li, P., Roller, P. P., Krajewski, K., Saito, N. G., Stuckey, J. A., and Wang, S. (2004) Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarization. *Anal. Biochem.* **332**, 261-273

26. 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

27. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.* **278**, 313-352

28. McGann, M. (2011) FRED Pose Prediction and Virtual Screening Accuracy. *J. Chem. Inf. Model.* **51**, 578-596

29. Hawkins, P. C. D., Skillman, A. G., Warren, G. L., Ellingson, B. A., and Stahl, M. T. (2010) Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J. Chem. Inf. Model.* **50**, 572-584

30. Leaver-Fay, A., Tyka, M., Lewis, S. M., Lange, O. F., Thompson, J., Jacak, R., Kaufman, K., Renfrew, P. D., Smith, C. A., Sheffler, W., Davis, I. W., Cooper, S., Treuille, A., Mandell, D. J., Richter, F., Ban, Y.-E. A., Fleishman, S. J., Corn, J. E., Kim, D. E., Lyskov, S., Berrondo, M., Mentzer, S., Popović, Z., Havranek, J. J., Karanicolas, J., Das, R., Meiler, J., Kortemme, T., Gray, J. J., Kuhlman, B., Baker, D., and Bradley, P. (2011) ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Meth. Enzymol.* **487**, 545-574

31. Osmanova, N., Schultze, W., and Ayoub, N. (2010) Azaphilones: a class of fungal metabolites with diverse biological activities. *Phytochemistry Rev.* **9**, 315-342

32. Gao, J.-M., Yang, S.-X., and Qin, J.-C. (2013) Azaphilones: chemistry and biology. *Chem. Rev.* **113**, 4755-4811

33. Kim, H. S., Wilce, M. C. J., Yoga, Y. M. K., Pendini, N. R., Gunzburg, M. J., Cowieson, N. P., Wilson, G. M., Williams, B. R. G., Gorospe, M., and Wilce, J. A. (2011) Different modes of interaction by TIAR and HuR with target RNA and DNA. *Nucleic Acids Res.* **39**, 1117-1130

34. Kreishman-Deitrick, M., Egile, C., Hoyt, D. W., Ford, J. J., Li, R., and Rosen, M. K. (2003) NMR Analysis of Methyl Groups at 100−500 kDa: Model Systems and Arp2/3 Complex †. *Biochemistry* **42**, 8579-8586

35. Hajduk, P., Augeri, D., and Mack, J. (2000) NMR-based screening of proteins containing ^{13}C-labeled methyl groups. *J. Am. Chem. Soc.* **122**, 7898-7904

36. Evidente, A., Kornienko, A., Cimmino, A., Andolfi, A., Lefranc, F., Mathieu, V., and Kiss, R. (2014) Fungal metabolites with anticancer activity. *Nat. Prod. Rep.* **31**, 617-627

37. Giannini, G., Penco, S., Pisano, C., Riccioni, T., Nasini, G., and Candiani, G. (2003) Chrysanthones, a new source of fungal metabolites with potential antitumor and antiangiogenesis properties. *Fitoterapia* **74**, 323-327