TIP AND TIP CHAPERONE INTERACTION IN *YERSINIA* AND *PSEUDOMONAS*

By

Pallavi Guha Biswas

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.

_____

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

_____

Joanna Slusky, Ph.D.

_____

Erik Holmstrom, Ph.D.

_____

Krzysztof Kuczera, Ph.D.

_____

David Weis, Ph.D.

Date Defended: 15 December 2021

The dissertation committee for Pallavi Guha Biswas certifies that this is
the approved version of the following dissertation:

TIP AND TIP CHAPERONE INTERACTION IN *YERSINIA* AND *PSEUDOMONAS*

_____

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

Date Approved: 12/28/21

Abstract

Spectroscopy is one of the most wildly used approaches to determine various properties of proteins. A wide range of spectroscopic methods are available to answer biological questions such as protein structure dynamics and protein-protein interactions at different sample states and timescales. In this dissertation, I have used electron paramagnetic resonance (EPR) spectroscopy to determine the structural dynamics of the tip chaperone proteins of the type III secretion system (T3SS). I have also determined protein-protein and protein-small molecule interactions of T3SS proteins using spectroscopic techniques such as Förster Resonance Energy Transfer (FRET), and nuclear magnetic resonance (NMR).

The T3SS macromolecular assemblage found in pathogenic Gram-negative bacteria such as *Salmonella typhimurium, Shigella flexneri, Yersinia pestis,* and *Pseudomonas aeruginosa* cause infections by transporting virulence effector proteins that modulate the host cell for the survival and propagation of the pathogens. The structural component of the T3SS is the needle complex which consists of the base, needle, tip complex, and translocon. The tip complex is assembled by multiple copies of the tip protein. In *Yersinia* and *Pseudomonas*, prior to the assembly of the needle complex, the tip proteins interact with small cytoplasmic tip protein chaperones LcrG and PcrG to prevent the premature secretion of the tip proteins LcrV and PcrV, respectively. There are no high-resolution structures of tip chaperone proteins. NMR spectroscopy shows LcrG and PcrG as partially folded alpha helices in contrast to the evidence in the literature which predicts the tip protein chaperones to be a coiled-coil. In this dissertation I have reported the different conformations of LcrG and PcrG in the free from and when bound to their cognate tip proteins LcrV and PcrV using EPR spectroscopy.

The interactions between tip proteins and tip chaperones are poorly characterized. The current knowledge of the interaction is based on predictions by other research groups that have shown that the tip protein chaperones have an alpha helical hairpin when bound to the tip proteins and interact with the coiled-coil domain present in the structures of the tip proteins. I have reported the model of LcrG-LcrV interaction using EPR spectroscopy and PcrG-PcrV interaction using FRET spectroscopy. This knowledge will be crucial in elucidating the structural domains of the tip proteins and tip chaperones involved in the interaction.

The *Yersinia* tip protein LcrV is exposed to the extracellular region between the bacteria and the host cell and it is essential for virulence as it interacts with the needle, translocon proteins, and tip chaperones. This biological relevance of the tip protein makes it an important target for the development of new antimicrobials. Currently no small molecules have been identified which bind to LcrV. In this dissertation I have reported the results of surface plasmon resonance (SPR) screening of two compound libraries which identified small molecules with benzene and pyridine moieties that interact with LcrV. The ligand binding surfaces of LcrV were identified by NMR spectroscopy.

The work presented in my dissertation illustrates the use of different spectroscopic techniques to determine protein structural dynamics, protein-protein, and protein-ligand interactions in *Yersinia* and *Pseudomonas* T3SS.

DEDICATION

I dedicate my work to all my Teachers; their guidance throughout my academic career has been instrumental in developing my scientific understanding.

Acknowledgements

First and foremost, I would like to thank my advisor, Dr. Roberto De Guzman, for giving me the opportunity to work in his lab and introducing me to the intriguing world of structural biology. His inputs were invaluable in expanding my limits of understanding the protein-protein interactions in the type III secretion system. He motivated me to think out of the box and supported me in all my scientific endeavors. His constructive advice helped me to improve my scientific communication. Under his guidance, I also got the multiple opportunities to train undergraduate students thereby allowing me to realize my potential as a mentor. I am forever grateful to Roberto.

I would also like to thank my current and former committee members for their guidance to help enhance my research with their expertise. I thank Dr. Erik Holmstrom for his time and collaboration in the FRET spectroscopy project. His continued guidance and advice helped me to understand the application of FRET spectroscopy in protein-protein interactions. I am grateful to him for introducing me to Mathematica, which helped me analyze my FRET spectroscopy data. I also thank Dr. Joanna Slusky for serving as my reader during my comprehensive exams, and her guidance helped me to navigate the process of grant writing. I would also like to thank Dr. Krzysztof Kuczera and Dr. David Weis for their valued inputs and suggestions in my doctoral research. I would especially like to extend my gratitude to Dr. Mark Richter for his valuable advice and collaboration that helped me start my FRET spectroscopy project and for letting me use his lab equipment at my convenience. I will forever miss his inspirational inputs which went a long way in making me a better biochemist. I would also like to thank my former committee member Dr. Eric Deeds, for serving as my chair during my comprehensive exams. His guidance during my grant writing process helped me improve my scientific writing. I extend my sincere

thanks to Dr. Likai Song from Florida State University for his collaboration and guidance in my EPR spectroscopy project. His support and cooperation in the data acquisition and analysis were invaluable for my research. I would also like to thank the members of the NMR core facility at KU - Dr. Justin Douglas and Dr. Minli Xing (former), for their assistance in setting up NMR experiments.

I thank all the former and current lab members of the De Guzman lab for providing me with a collaborative and fun environment to pursue my research. I want to thank Andrew McShan and Dr. Supratim Dey for training me during my lab rotations. Their mentorship and intellectual guidance were deciding factors for me to join this lab. I would also like to thank Kawaljit Kaur, Sikta Patnaik, and Amritangshu Chakravarty for their friendship and valued inputs that helped my research. I also want to thank our exceptional undergraduates — Mason Wilkinson, Shoichi Tachiyama, and Muhammed Ciftici for their invaluable contributions to my research. A massive shoutout to our new prep student Megan Hinrichsen for making my last few months in the lab truly enjoyable and helping me proofread my dissertation.

I thank the KU Molecular Biosciences department for allowing me to work as a graduate student. The Biochemistry and Biophysics program helped in strengthening my concepts which were invaluable for my doctoral research. I would also like to thank John Connolly, graduate advisor, and Cynthia Cowan, administrative associate at the department, for taking care of all my official requirements on time, which helped me concentrate on my graduate research.

My journey as a graduate student would not have been possible without the support of my remarkable graduate cohorts — Pushpa Itagi, Anupama Kante, and Nootan Pandey. Their friendship and encouragement have helped me through all my hardships during my doctoral research. I would also thank the KU meditation and mindfulness club for taking care of my

mental well-being and enhancing my leadership and interpersonal skills. I would also like to thank Aftan Jameson from the KU student wellness program for helping me stay fit and active, which helped with my research. A special mention to Hollie Hall from the KU international student association for helping me navigate my life as an international student and making Lawrence my home away from home.

Lastly, I would like to thank my parents Prasun Guha Biswas and Devi Priya Guha Biswas for encouraging me to pursue science and supporting me through all my endeavors. I thank them for providing me with insights from their life that helped me navigate my professional and personal life.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AF647 – Alexa Flour 647
BSBA – 2-(benzylsulfinyl) benzoic acid
CD – Circular Dichroism
CSD – Chemical Shift Deviation
CPM – 7-Diethylamino-3-(4'-Maleimidylphenyl)-4-Methylcoumarin
DEER – Double electron-electron resonance
DMSO – Dimethyl Sulfoxide
DOL – Degree of Labeling
DTT – Dithiothreitol
EDTBO – (3aR,7aS)-6-(6-ethoxypyridin-3-yl)-2,2-dimethyl-3a,4,5,7a-tetrahydro-2H-1,3-benzodioxol-5-one
EMAL – Eosin-5-Maleimide
EPR – Electron Paramagnetic Resonance
ESI – Electrospray Ionization
FM – Fluorescein C5 Maleimide
FRET – Förster Resonance Energy Transfer
GB1 – β1 Domain of Streptococcal Protein G
HSQC – Heteronuclear Single Quantum Coherence
ILV – Isoleucine, Leucine and Valine
IPTG – Isopropyl-β -D-thiogalactopyranoside
MS – Mass Spectrometry
MTSL – S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl) methyl methanesulfonothioate
NOE – Nuclear Overhauser Effect
NMR – Nuclear Magnetic Resonance
PBS – Phosphate-Buffered Saline
PE – 2-(pyridin-2-yl) ethan-1-ol
PEI – Polyethyleneimine
PELDOR – Pulsed electron double resonance
PM – N-(1-Pyrene) Maleimide
PMSF – Phenylmethanesulfonylfluoride
SDSL – Site Directed Spin Labeling
SPR – Surface Plasmon Resonance
T3SS – Type III Secretion System
TB – Terrific Broth
TCEP – Tris(2-carboxyethyl) phosphine hydrochloride
TEV – Tobacco Etch Virus
TROSY – Transverse Relaxation Optimized Spectroscopy

Chapter 1: Introduction

1.1 Type III Secretion System

Many pathogenic Gram-negative bacteria inject virulence effector proteins into the host cells by forming a nanoinjector known as the type III secretion system (T3SS)[1],[2],[3]. These bacteria are causative agents of infectious diseases such as bacillary dysentery (*Shigella flexneri)*, typhoid (*Salmonella typhimurium)*, bubonic plague (*Yersinia pestis*), and secondary nosocomial infections in immunocompromised patients (*Pseudomonas aeruginosa*)[4]. These bacteria cause millions of deaths worldwide[5]. The rise of antibacterial resistance among these pathogenic bacteria is a major public health concern[6]. The proteins involved in the assembly of the T3SS nanoinjectors are attractive targets for developing new antimicrobials[7],[8],[9].

1.1.1 *Components of the T3SS*

The T3SS consists of the needle complex, chaperones and effector proteins. The needle complex is the structural component of the T3SS along with the export apparatus and an ATPase complex to transport effector proteins into the host cell[10]. Chaperones regulate the secretion of effectors and other structural T3SS proteins by forming complexes in their pre-active states[3],[11]. The needle complex is a macromolecular assemblage of over 20 different proteins composed of the base, which spans the bacterial inner and outer membranes (Fig 1-1), a needle which forms a 25 Å wide channel in the intermembranous region between the bacteria and the host cell, a tip complex, and a translocon. The T3SS is shown as a cartoon in Fig 1-1[3],[12],[13]. The tip complex caps the needle until contact with the host cell is established[10],[11],[14],[15]. The tip complex is an environmental sensor for the deployment of the translocon upon contact with the host cell[10],[11]. The translocon forms a pore in the host cell membrane for the passage of the effector proteins. Inside the host cytoplasm, the effector

proteins modulate the host cell pathways such as cytoskeletal dynamics and cell signaling to enable the invasion of the pathogen for survival and propagation in the host cell[13],[16].

1.1.2 *Tip proteins in T3SS*

The tip complex has multiple functions, it plays a crucial role as host cell sensor, regulator of effector protein secretion, and as a platform for translocon assembly[13],[17]. The tip complex is positioned on top of the needle and is formed by an estimate of four to six copies of the tip proteins in *Burkholderia* BipD, *Salmonella* SipD, *Shigella* IpaD, *Yersinia* LcrV, and *Pseudomonas* PcrV[3],[11],[14],[16],[18]. The atomic structures of the tip proteins BipD, SipD, IpaD, and LcrV are known [18],[19],[20],[21]. The structures of IpaD, SipD, and BipD consists of an N-terminal alpha helical hairpin, a central coiled coil domain, and an α/β domain[18],[22] as shown in Fig 1-2C-E. The atomic structure of the *Yersinia* LcrV consists of the conserved coiled-coil domain and the α/β domain, but lacks the N-terminal alpha helical hairpin and possesses a globular N-terminal domain(Fig 1-2A)[21]. The highly conserved coiled coil domain is involved in the interaction with the needle proteins[23],[24]. The N-terminal alpha helical hairpin motif in BipD, IpaD and SipD function as a chaperone[18]. The tip proteins in *Pseudomonas* and *Yersinia* lack the N-terminal alpha helical hairpin domain but instead bind to small cytoplasmic proteins, known as the tip chaperone proteins[17].

The rise in antibacterial resistance among pathogens requires the development of alternate targets for developing novel anti-infectives[7],[8],[9],[14]. The role of the tip proteins in the assembly of the tip complex makes it an attractive target for development of new antimicrobials.

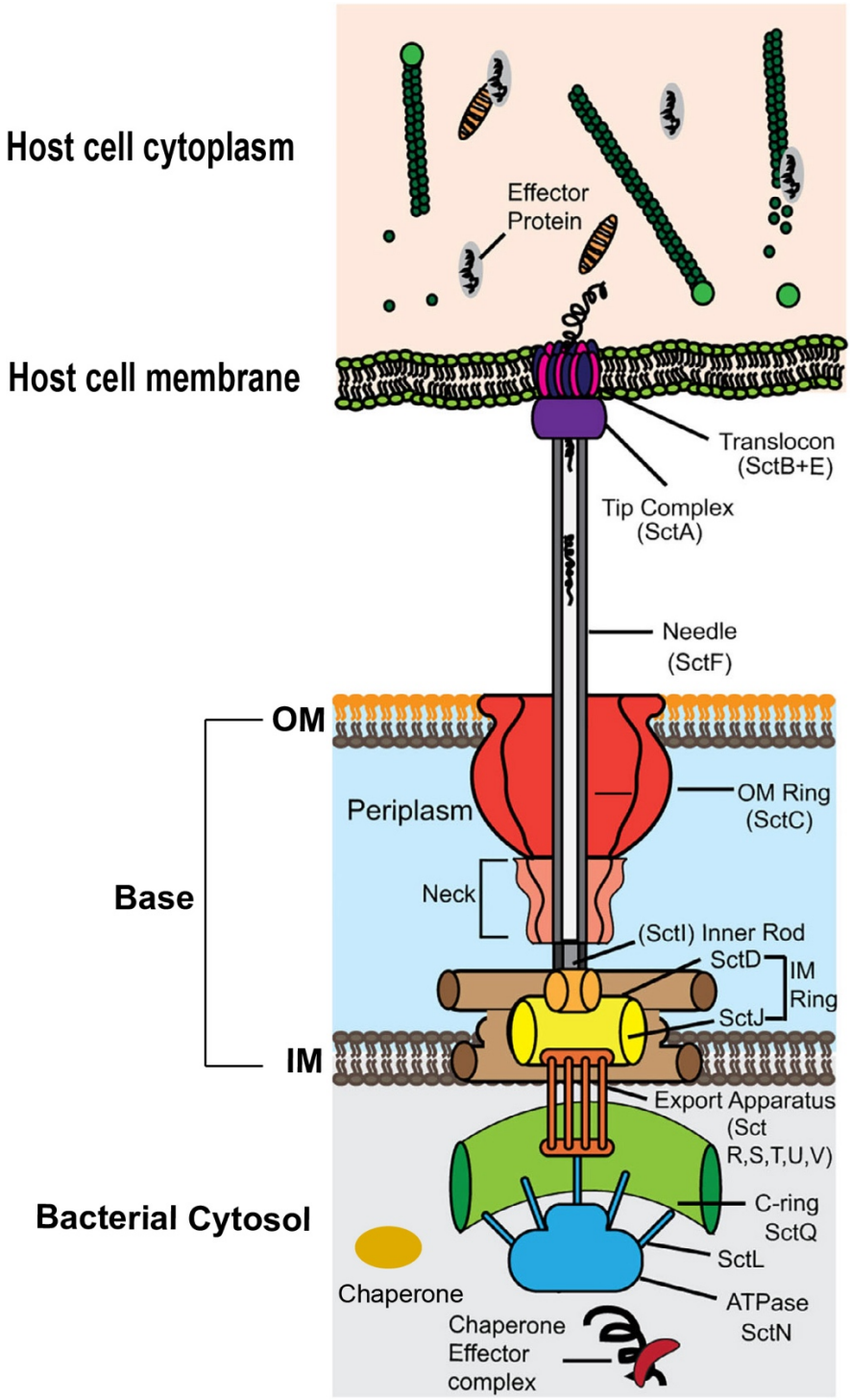

Figure 1-1 Molecular model of type III secretion system (T3SS). The needle complex is composed of the base, needle (solid grey channel), tip complex (purple) and, translocon (blue and pink). The chaperone (yellow oblong circle) prevents premature secretion of the T3SS components. The effector proteins are transported from bacterial cytosol through the export apparatus attached to the base to the host cell cytoplasm through the translocon. Figure adapted from **[10].**

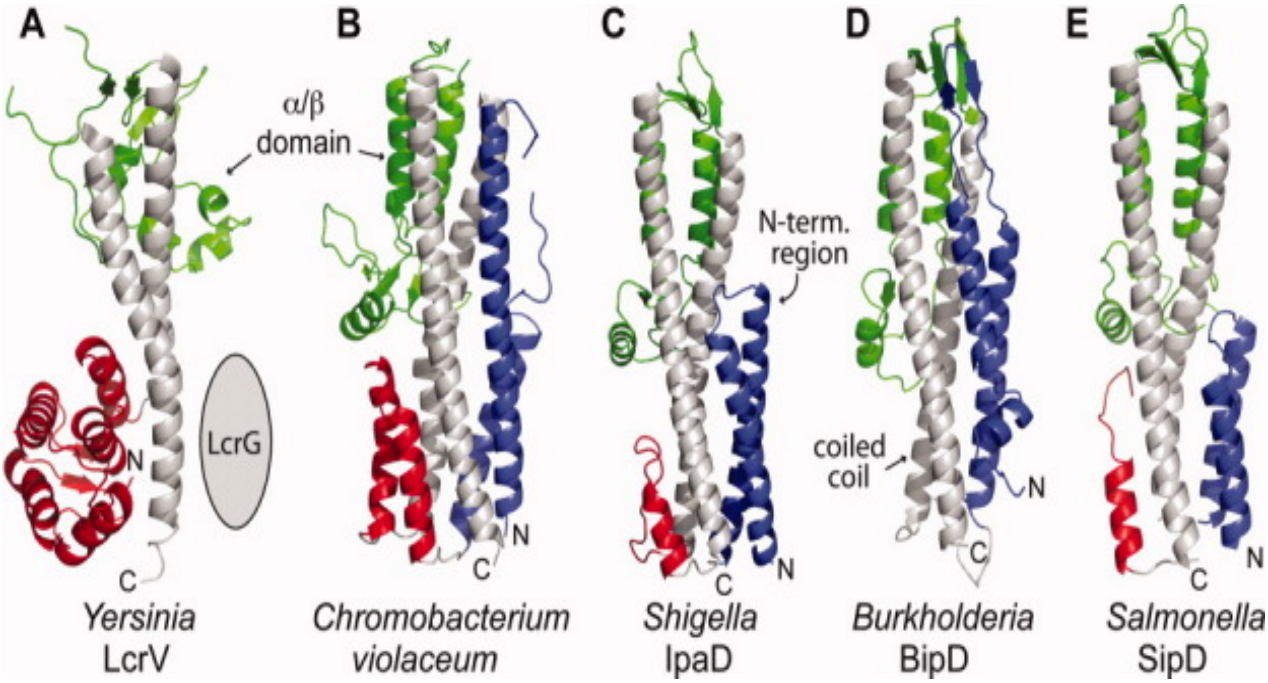

Figure 1-2 Crystal structures of T3SS tip proteins. (A) *Yersinia* LcrV **(B)** *Chromobacterium violaceum,* **(C)** *Shigella* IpaD **(D)** *Burkholderia* BipD and **(E)** *Salmonella* SipD. The C-terminal coiled coil(gray) domain and mixed α/β domain (green) are common structural features of T3SS tip proteins. The N-terminal region in **(A)** forms a globular domain of α-helices and β-strands (red), whereas in (B–E), the N-terminal region forms α-helical hairpins (blue). Figure from **[20].**

1.1.3 *Tip protein chaperones in T3SS*

Prior to the assembly of the tip complex, the tip proteins of *Yersinia* (LcrV) and *Pseudomonas* (PcrV) interact with small cytosolic proteins *Yersinia* LcrG and *Pseudomonas* PcrG, known as tip chaperone proteins. The tip chaperone proteins form a complex with their cognate tip proteins *Yersinia* LcrV and *Pseudomonas* PcrV to maintain their structural integrity in the bacterial cytosol during the assembly of the needle complex[17],[25]. The tip chaperone proteins also function as a regulator of the secretion of the effector proteins by interacting with the components of the base to prevent premature effector protein secretion[26],[27].There are currently no high resolution atomic structures for the tip chaperones LcrG and PcrG. There are several predictions in the literature which hypothesize that LcrG exists as a coiled-coil[28] and PcrG exists as a four helix bundle[29]. The structures determined by NMR shows that both LcrG and PcrG are partially folded alpha helical proteins connected by unstructured linkers[30],[31].

Based on the findings in the literature and the NMR structure data, I attempted to identify a population of LcrG and PcrG where the helices interact with each other using electron paramagnetic resonance (EPR) spectroscopy, in collaboration with Dr. Likai Song from Florida State University. My findings are presented in Chapter 2 and Chapter 4.

The interaction between the tip and tip chaperone proteins are poorly characterized. There are no high-resolution structures of the binding interaction of the tip and the tip chaperone protein. The current knowledge of this interaction has been derived from the comparison of the crystal structure of *Yersinia* tip proteins LcrV with the crystal structures of the tip proteins BipD, SipD and IpaD as shown in Fig 1-2B-E. The common structural features among all the tip proteins are the central coiled-coil domain and the mixed α/β domain, but the crystal structure of LcrV (Fig

1-2A) lacks the N-terminal alpha helical hairpin which is present in the crystal structure of BipD, IpaD and SipD. The N-terminal alpha helical hairpin structures of BipD, IpaD and SipD function as the chaperone[18]. Based on the functional similarity of the N-terminal alpha helical domain and the tip chaperones, it was hypothesized that both LcrG and PcrG have an alpha hairpin structure and interact with the central coiled domain of their cognate tip proteins LcrV and PcrV. However, the NMR data of the LcrG-LcrV or PcrG-PcrV interaction indicates the binding of the tip proteins LcrV or PcrV has an effect on the overall structure of the tip chaperone proteins LcrG or PcrG[30],[31].

Based on the NMR data of the LcrG-LcrV interaction and the current knowledge of the tip chaperone protein and the tip proteins, I determined the interaction of LcrG-LcrV complex using EPR spectroscopy in collaboration with Dr. Likai Song from Florida State University.
The interaction of PcrG-PcrV complex was determined using Förster resonance energy transfer (FRET) spectroscopy in collaboration with Dr. Mark Richter and Dr. Erik Holmstrom from the University of Kansas. My findings are presented in Chapter 3 and Chapter 5.

The structural models of LcrG-LcrV and PcrG-PcrV obtained from these studies and their further refinement has been discussed in Chapter 7.

1.2 EPR spectroscopy

Electron paramagnetic resonance (EPR) spectroscopy is a well-established method to study structural changes within proteins or interacting regions in protein–protein interactions [32].The paramagnetic label is attached to the protein by site directed spin labeling (SDSL). Since many biological macromolecules are diamagnetic, the EPR spectrum generated by spin label is the only signal in the sample.[33].

1.2.1 Basic principle of electron paramagnetic resonance (EPR) spectroscopy

The basis of EPR spectroscopy lies in the spin of an electron and its associated magnetic moment. When an electron is placed within an applied magnetic field, B_o, the two possible spin states of the electron have different energies[34]. The lower energy state occurs when the magnetic moment of the electron, μ_B, is aligned with the magnetic field and a higher energy state occurs where μ_B is aligned against the magnetic field as shown in Fig 1-3[34]. The two states are labeled by the projection of the electron spin, m_S, on the direction of the magnetic field, where $m_S = -1/2$ is the parallel state, and $m_S = +1/2$ is the antiparallel state as shown in Fig 1-3[34]. So, for a molecule with one unpaired electron in a magnetic field, the difference in energy states of the electron can be defined as:

$$\Delta E = h\nu = g\mu_B B_o$$

where g is the proportionality factor (or g-factor), μ_B is the Bohr magneton, Bo is the magnetic field[34]. An EPR spectrum is obtained by holding the frequency of radiation constant and varying the magnetic field. Absorption occurs when the magnetic field tunes the two spin states so that their energy difference is equal to the radiation. This is known as the field for resonance. The result of the absorption signal is presented as the first derivative of the absorbance spectrum as shown in Fig 1-4[34].

1.2.2 *Site directed spin labeling for EPR studies*

Nitroxides are stable free radicals with an unpaired electron delocalized in the N-O bond are required for EPR detection[33]. The most common spin labeling strategy for proteins uses cysteine point mutations followed by modification of the sulfhydryl group with spin labels like S-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate (MTSL) [33]. The modified side chain resulting from the reaction is designated as R1[33] as shown in Fig 1-5. The side chain R1 has become the spin label of choice in SDSL studies[35],[36] as it attaches at the vast majority of sites in many different proteins. Nitroxide spin labels are small and have been shown to have minimal effects on protein structures[33],[37],[38]. However, for SDSL, as for all labeling techniques, control experiments comparing wild type protein and labeled mutants are essential to exclude distortion of protein conformation and function due to the label[33].

1.2.3 *Protein-protein interaction determined by EPR spectroscopy.*

The EPR technique used by my collaborator, Dr. Likai Song from Florida State University, is continuous wave (CW) EPR spectroscopy to determine spin-spin distances between two spin labels attached to the proteins. CW-EPR is highly sensitive and can be used for samples with a volume of ~ 50 μL at concentrations at about 50 μM and detect spin-spin distances in the range of 8-25 Å[33] and the EPR spectra are collected at 150 K. The general scheme of the experiment is described in Fig 1-6. The two proteins are coupled with the MTSL spin label side chain (R1) and each of these proteins would generate a characteristic spectrum for R1 (Fig 1-6B). The two spectra added together would generate the control spectrum (Fig 1-6C).

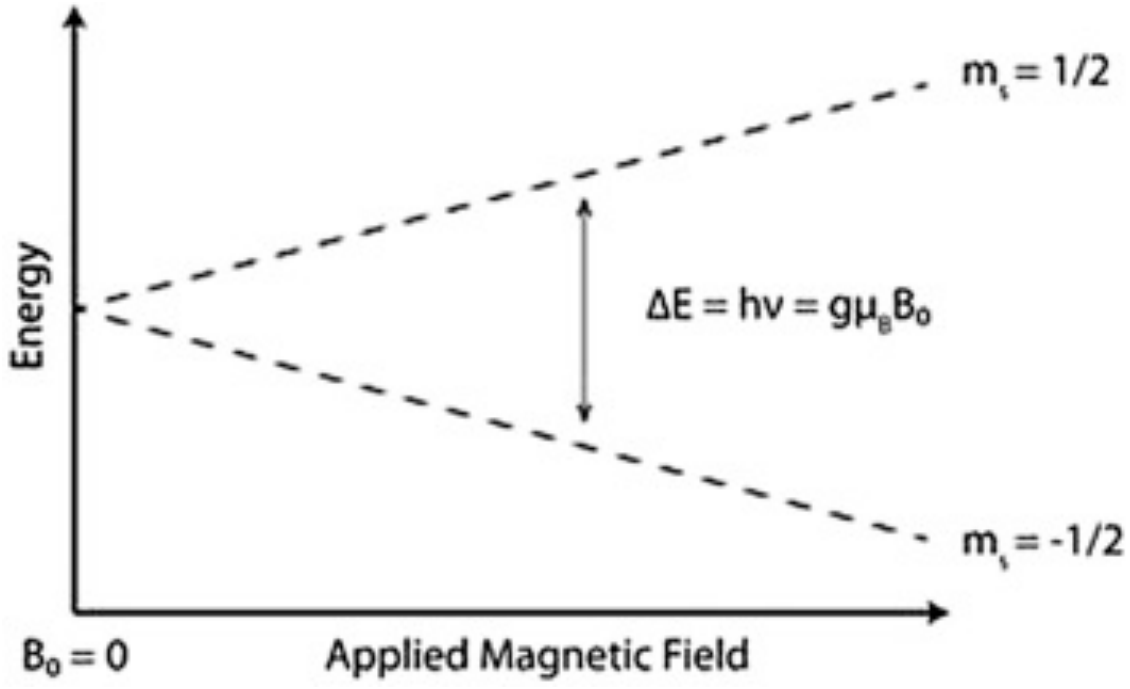

Figure 1-3. Principle of EPR. An external magnetic field Bo is applied the lower energy state occurs when the magnetic moment of the electron, μ_B, is aligned with the magnetic field and a higher energy state occurs where μ_B is aligned against the magnetic field. The two states are labeled by the projection of the electron spin, m_S, on the direction of the magnetic field, where $m_S = -1/2$ is the parallel state, and $m_S = +1/2$ is the antiparallel state. Figure from**[34].**

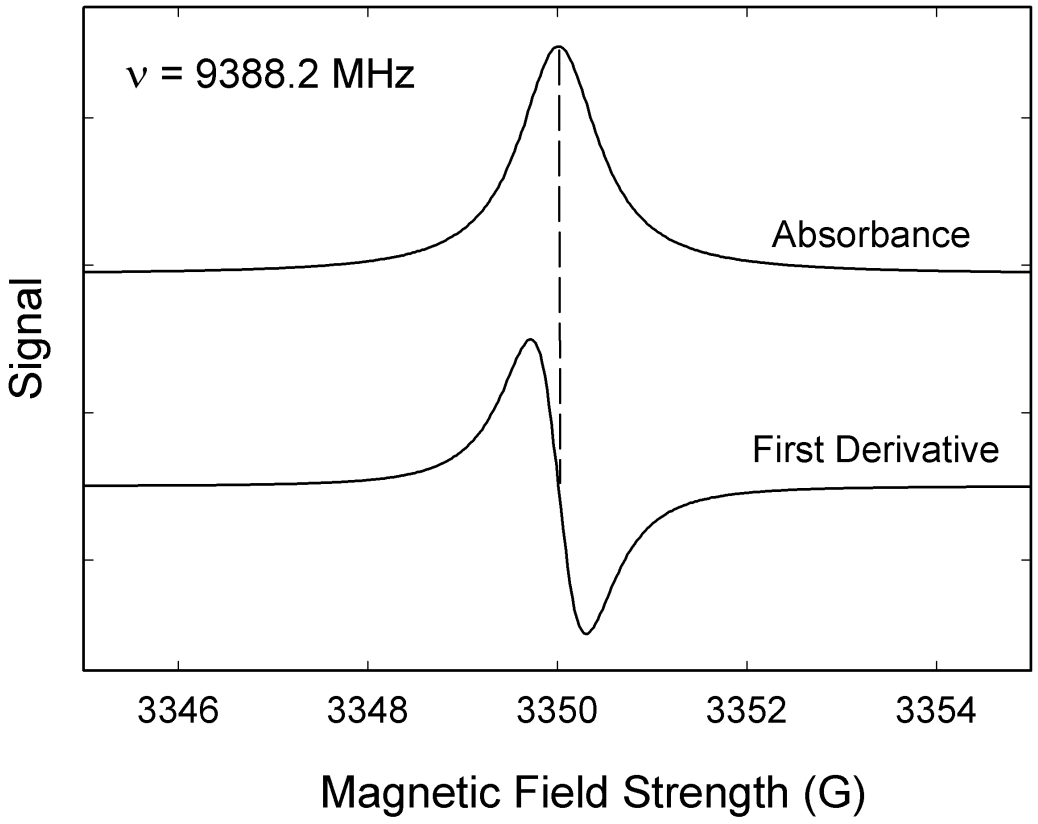

Figure 1-4. Example of an EPR signal acquired from absorption. The EPR signal is the first derivative of the absorption monitored at a radiation frequency of 9388.2 MHz. The EPR data is a plot of the signal(absorption) and magnetic field. Figure from**[34].**

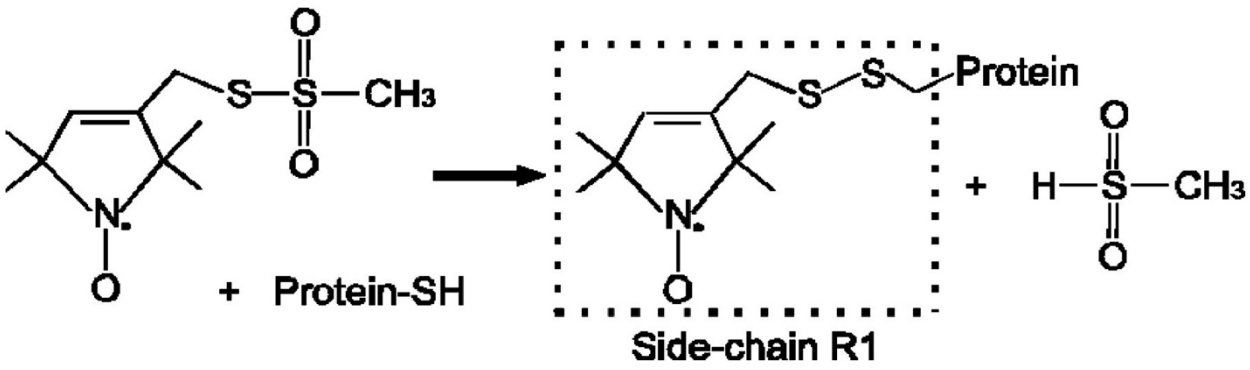

Figure 1-5 Site directed spin labeling reaction. MTSL reacts with the cysteine sulfhydryl side chain to produce the nitroxide side chain designated as R1. Figure adapted from[39].

The spectrum generated by the sample, which consists of two spin labels, would show spectral broadening compared to the control spectrum as result of dipole-dipole interaction due to the proximity of the spin labels, indicating interaction (Fig 1-6E)[40],[41]. If the two spin labels have a distance more than 25 Å the sample spectrum would overlay on the control spectrum as shown in Fig 1-5D. The distance information from the spectrum can be extracted from spectral deconvolution by Monte Carlo Simulations[42],[43],[44],[45],[46].

In this dissertation, I referred to CW-EPR spectroscopy as simply EPR spectroscopy and my collaborator has used this technique to determine the interaction between the two helices of LcrG and PcrG (Chapter 2 and Chapter 4) and determine the interaction of LcrG with the tip protein LcrV (Chapter 3).

1.3 Förster resonance energy transfer spectroscopy

Förster resonance energy transfer (FRET) is a technique used to determine the interaction between two molecules labeled with different fluorophores (the donor and the acceptor) by the transfer of energy from the excited donor to the acceptor. In biological applications, this technique is used to qualitatively map protein-protein interactions and determine distances between a single donor and acceptor molecule[47],[48].

1.3.1 Criteria for selection of donor and acceptor fluorophores

For FRET to occur, the donor and acceptor must be in close proximity for the FRET process to be efficient. FRET efficiency (E) is defined by the equation — $E = Ro^6 / (Ro^6 + r^6)$[49], where Ro is the Förster radius and r is the actual distance between the two fluorophores.

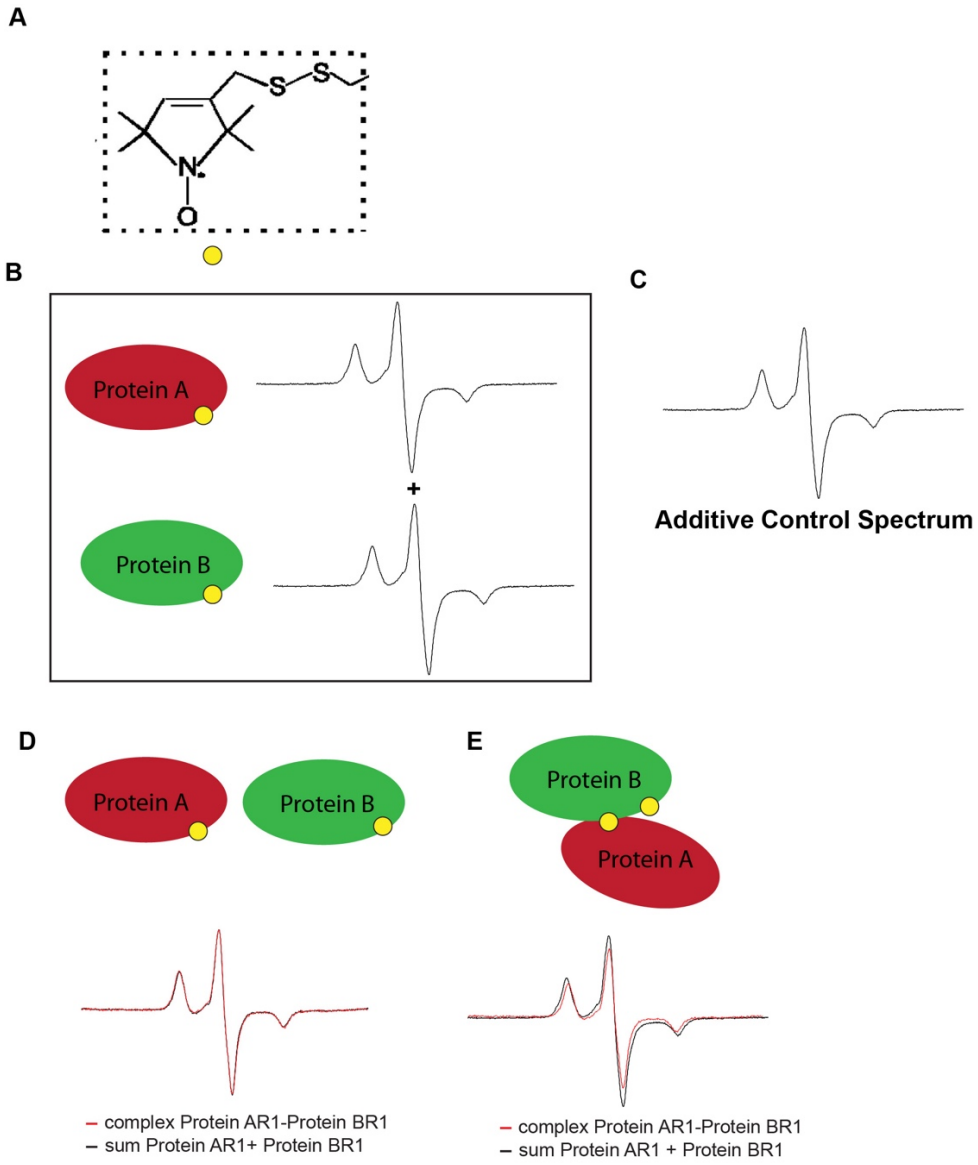

Figure 1-6 General scheme of EPR spectroscopy used to determine protein-protein interaction. (A) The spin label side chain (R1) is shown as yellow circle. **(B)** The characteristic MTSL side chain spectrum(black) generated by spin labeled Protein A and Protein B. **(C)** The additive control spectrum generated (black) by Protein A and Protein B. **(D)** EPR spectrum generated by the sample — Complex formed by Protein A and Protein B (red) overlays on the control spectrum (black) showing no interaction between the spin labels in protein A and protein B. **(E)** EPR spectrum generated by the sample — Complex formed by Protein A and Protein B (red) shows spectral broadening wherein the spectral height of the complex shows decrease compared to the control (black) spectrum indicating interaction between the spin labels in protein A and protein B.

The Förster radius is the distance at which 50% of the excitation energy is transferred from the donor to the acceptor, and the Ro value usually lies between 10-100Å[49].

Another criterion for FRET is that the absorption spectrum of the acceptor must overlap the fluorescence emission spectrum of the donor as shown in Fig 1-8A [50],[51]. If the spectral overlap between the fluorescence of the donor and the absorption of the acceptor is zero, there will be no energy transfer, even if the donor and acceptor are extremely close[52]. The dipole orientation of the donor and acceptor should be parallel to one another for FRET to occur[50].

In this dissertation, for the FRET studies, the donor is N-(1-Pyrene) Maleimide (PM) (Fig 1-7A) and the acceptor is 7-Diethylamino-3-(4'-Maleimidylphenyl)-4-Methylcoumarin (CPM) (Fig 1-7B) which satisfies these above mentioned criteria and the Förster radius (Ro) for this FRET pair is 30 Å[53],which is within the distance of interaction hypothesized for PcrG and PcrV.

1.3.2 *FRET spectroscopy of protein-protein interaction*

FRET spectroscopy based binding assays are performed to determine the protein-protein interaction and protein-ligand interaction. The FRET based assays are distinct from a number of other techniques as it has a higher sensitivity and throughput than standard isothermal calorimetry and does not require the sample immobilization and significant method development that is needed for SPR[54]. For the FRET experiments, the sum of the emission spectra of the donor and the emission spectra of the acceptor acquired at the donor excitation at 330 nm serves as the control.

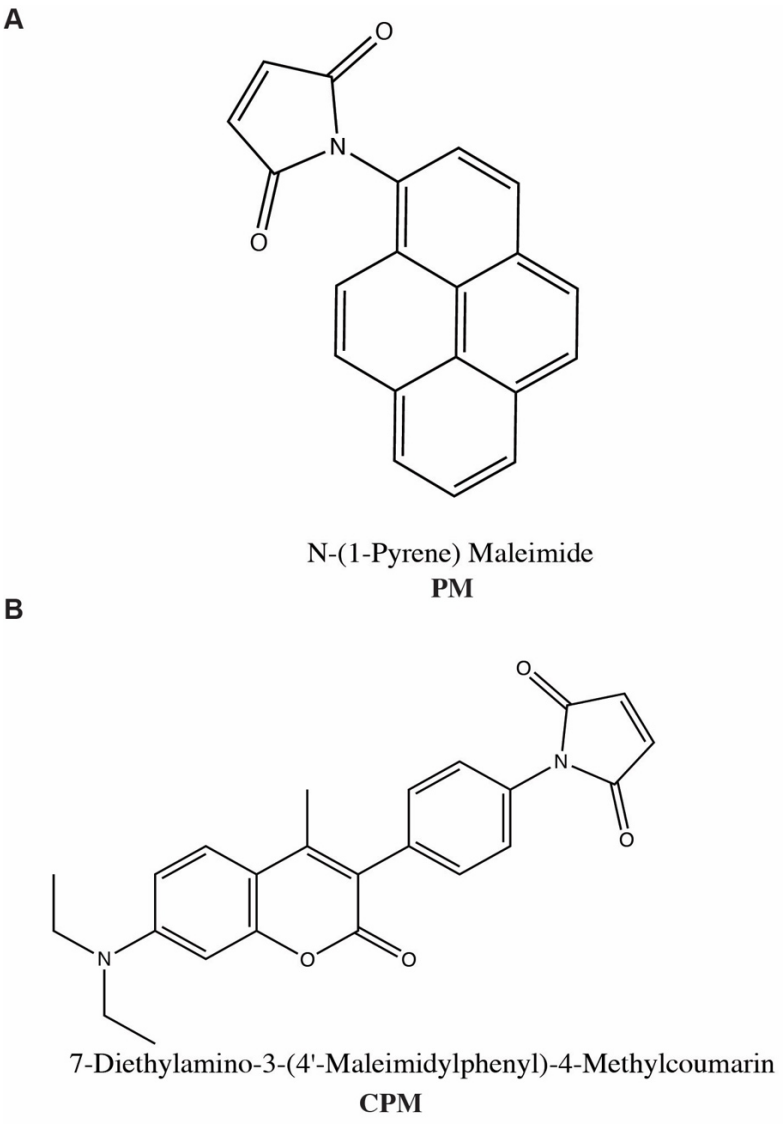

Figure 1-7 Fluorophores used for FRET spectroscopy. **(A)** N-(1-Pyrene) Maleimide (PM) used as donor. **(B)** 7-Diethylamino-3-(4'-Maleimidylphenyl)-4-Methylcoumarin (**CPM**) used as acceptor.

The emission spectrum generated by the sample of the protein-protein complex, compared to the control spectrum, shows the decrease in the donor maxima and increase in acceptor maxima as shown in Fig 1-8C, indicating interaction. This phenomenon occurs due to the transfer of energy from the donor to the nearby acceptor, which results in the deexcitation of the donor and excitation of the acceptor[55]. The FRET spectrum of protein A and protein B does not show any energy transfer compared to the control emission spectrum as shown in Fig 1-8B. This indicates that the donor and the acceptor fluorophores are not close when protein A and protein B do not interact. The analysis of the FRET spectrum obtained from donor and acceptor labeled proteins at different concentrations can be used to calculate the dissociation constant (k_d) of the complex[56],[57].

In this dissertation I have referred the FRET spectroscopy-based assays as FRET and used this technique to determine the interaction of donor labeled (PM) PcrG with acceptor labeled (CPM) labeled PcrV and also determine the dissociation constant (k_d).

1.4 NMR spectroscopy for protein-ligand binding interaction

The structural and functional characterization of molecular interactions are essential for understanding biological processes[58],[59],[60]. NMR spectroscopy is a tool to observe weak biomolecular interactions at physiological conditions at the residue level[58],[59],[60],[61],[62],[63]. NMR chemical shift mapping or chemical shift perturbation is a robust technique which can provide information about the location and strength of the binding event[64],[65],[66]. For protein-ligand binding, an isotopically labeled protein is titrated with the ligand in increasing molar ratios. The NMR active nuclei at the binding surfaces would undergo changes, thus enabling the determination of molecular interaction surfaces[63],[64],[65],[66].

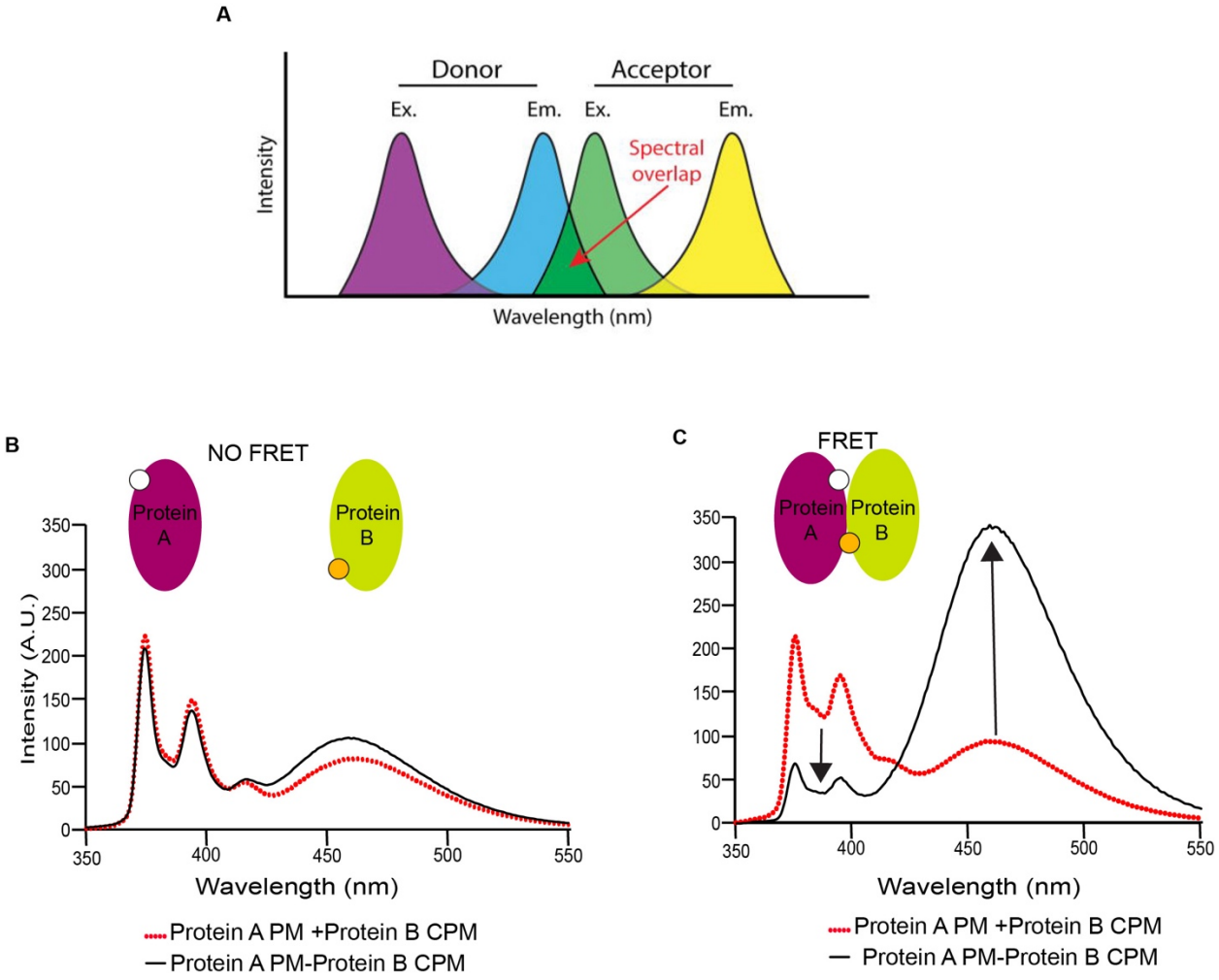

Figure1-8 General scheme of FRET spectroscopy used to determine protein-protein interaction. (A) The spectral overlap between donor emission and acceptor excitation spectra required for FRET experiments. Figure adapted from[67]. The donor and acceptor are depicted as white and yellow circles respectively. **(B)** The FRET spectrum of Protein A and Protein B (black solid line) shows little variance when overlayed on the control spectrum (dashed red line) showing that Protein A and Protein B do not interact. **(C)** FRET spectrum (black solid line) generated by the interaction of Protein A and Protein B shows decrease in donor maxima and increase in acceptor maxima as a result of energy transfer compared to the control spectrum (dashed red line), showing interaction between Protein A and Protein B.

1.4.1 NMR exchange regimes

The NMR spectra of proteins bound to the ligand vary in appearance depending on binding affinity of the protein ligand interaction[14],[68]. The broad three different chemical exchange states that can be observed are fast, intermediate, and slow. The slow exchange on NMR chemical shift timescale indicates the slow dissociation of the ligand and protein as a result of a tight binding interaction with binding constants in the nanomolar to sub micromolar range[14],[62],[63],[65],[68]. The complex rarely dissociates during the NMR experiment and two sets of resonances/peak are observed corresponding to the free and ligand state[14],[62],[65],[68]. The intensity of the free protein peaks gradually decreases and eventually disappears with increase in ligand concentration while new peaks for the bound state appear[14],[62],[65],[68] as shown in Fig 1-9(a). The intermediate exchange on NMR timescale exhibits interconversion between the free and ligand-bound state with binding constants in the micromolar range[14],[62],[63],[65],[68]. The peaks affected by the interaction show reduction in intensity and may even disappear during titration[14],[62],[65],[68] as shown in Fig 1-9(b). The shifts in peak position may be accompanied with peak broadening. Lastly, the fast chemical exchange in NMR timescale indicates fast complex dissociation or weak binding with binding constants in the high micromolar to millimolar range[14],[62],[63],[65],[68]. The fast interconversion between the free and ligand bound state during the NMR experiment results in the observed chemical shift as a weighted average of the free and bound form[14],[62],[65],[68]. The peaks move continuously from free to ligand bound positions with increasing ligand concentrations as shown in Fig 1-9(c). The bound complex can be captured in the presence of excess ligands irrespective of the type of NMR exchange regime[62],[68].

1.4.2 *Isotopic labeling of proteins for NMR data acquisition*

The abundant nuclei present in proteins are ^{1}H ,^{12}C and ^{14}N. The ^{12}C and ^{14}N nuclei are NMR inactive whereas the ^{1}H is NMR active. For the characterization of proteins by NMR spectroscopy, the NMR inactive nuclei are substituted by NMR active ^{13}C and ^{15}N through isotope enrichment. The protein of interest is expressed in *E. coli* using M9 minimal media supplemented with the desired isotope[61],[62],[63]. In this dissertation I have used two isotopic labeling schemes to study protein ligand binding which include:

Uniform ^{15}N labeling: During protein expression in *E. coli* the M9 media is supplemented with ^{15}N labeled ammonium chloride ($^{15}NH_4Cl$) as a nitrogen source[58],[62]. ^{15}N labeled proteins are used to acquire amide backbone correlation spectroscopy such as 2D ^{1}H-^{15}N HSQC (or TROSY) to study the changes in the protein overall[58],[62],[69]. ^{15}N-labeled proteins can be used to study as a probe to investigate ligand binding interactions[14].

Side chain methyl (^{13}C) labeling (ILV) labeling: Synthetic keto acid precursors are supplement to the *E. coli* used for protein expression to selectively label isoleucine ($^{13}C\delta1$), leucine ($^{13}C\delta$) and valine ($^{13}C\gamma1$)[70],[71],[72]. ILV labeled proteins can be used to study the interaction in the hydrophobic core of the protein and it serves as an excellent probe to study protein-ligand interaction due to reduced signal complexity[62],[66],[70],[71].

To study the interaction between the *Yersinia* LcrV and the small molecule from the Aube libraries I used NMR to find the ligand binding surfaces. The results of that study are presented in Chapter 6.

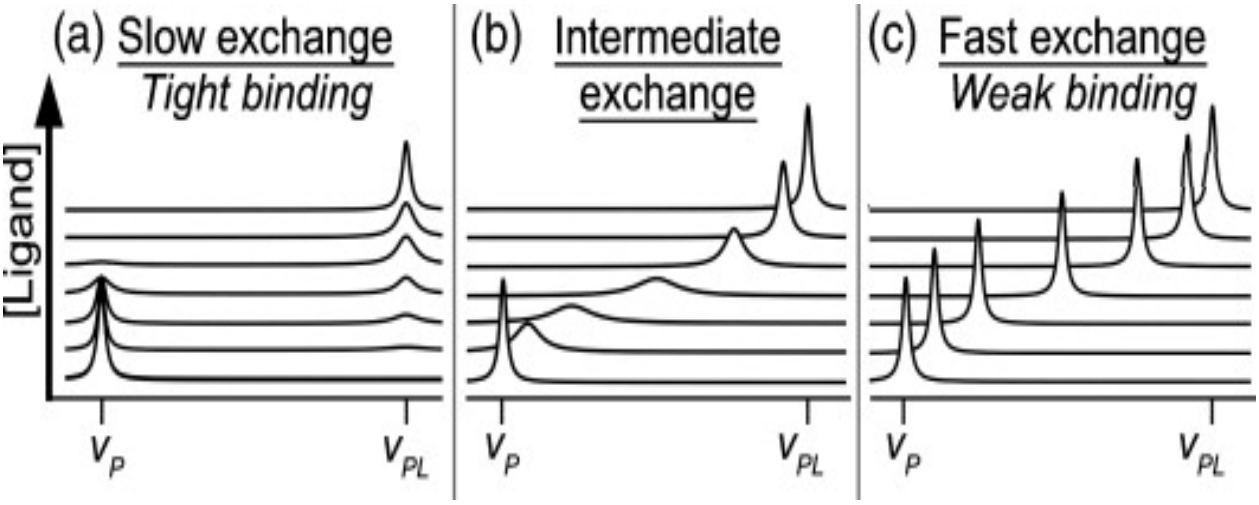

Figure 1-9 Exchange regimes on the NMR chemical shift time scale for protein-ligand complex. Signals of free and ligand-bound state of protein can be observed at V_P (frequency of resonance of the protein) and V_{PL} (frequency of resonance of the protein bound to ligand), respectively. The exchange regimes are (a) slow, (b) intermediate, or (c) fast. Figure from [14],[68].

1.5 Circular dichroism (CD) spectroscopy

Circular dichroism (CD) is an optical spectroscopic method that exploits the differential absorption of left and right-handed circularly plane polarized light by optically active molecules to determine their structural configuration[73]. For proteins, the electromagnetic spectrum detected in the far-UV wavelength range (190-260 nm) can be used to characterize the secondary structural content as α-helical, β-strand and random coil as shown in Fig 1-10[73].

Electromagnetic waves contain electric and magnetic field components that oscillate perpendicularly in the direction of a light beam's propagation.[74],[75] The directionality of these components defines the waves' polarization. In unpolarized light or white light, the electric and magnetic fields oscillate in many different directions[74],[75]. In linearly polarized light, the electromagnetic wave oscillates along a single plane while in circularly polarized light (CPL) two electromagnetic wave planes are at a 90° phase difference to one another and this plane rotates as the light beam propagates[74],[75]. Optically active molecules can be described by their chirality, or the asymmetry in a molecule's structure[74],[75]. The left-handed CPL will propagate through a chiral sample at a different speed than its right-handed counterpart[74],[75]. When a chiral molecule absorbs left- and right- handed CPL to different degrees the resulting electric field vector traces out an ellipse[74],[75]. The circularly polarized light travels through the sample, the phase relationship between the CPL waves changes and the linearly polarized wave is rotated, and the elliptically polarized light is now tilted as shown in Fig 1-11[75]. This ellipticity (θ) as a function of wavelength defines a circular dichroism spectrum[74],[75].

CD spectra are reported as ellipticity, θ and measured in units of mdeg[73]. Molar absorptivity (Δε) normalizes the circular dichroism spectra based on concentration and pathlength,

$$\Delta\varepsilon=\varepsilon l-\varepsilon r=\Delta A/(c*l)=\theta$$

where εl,-εr is the difference in the extent the left-(εl) and right-(εr) circularly polarized light components are absorbed at a given wavelength, ΔA is the difference in absorbance, c is the concentration (mol/L), and l is the pathlength(cm)[73].

In this dissertation I have used the CD spectroscopy to compare the overall change in the secondary structures of $LcrG^{7-73}$ (Chapter 2 and Chapter 3), LcrV (Chapter 3) and $PcrG^{9-76}$ (Chapter 4) after spin labeling with the MTSL side chain (R1) (Fig 1-5).

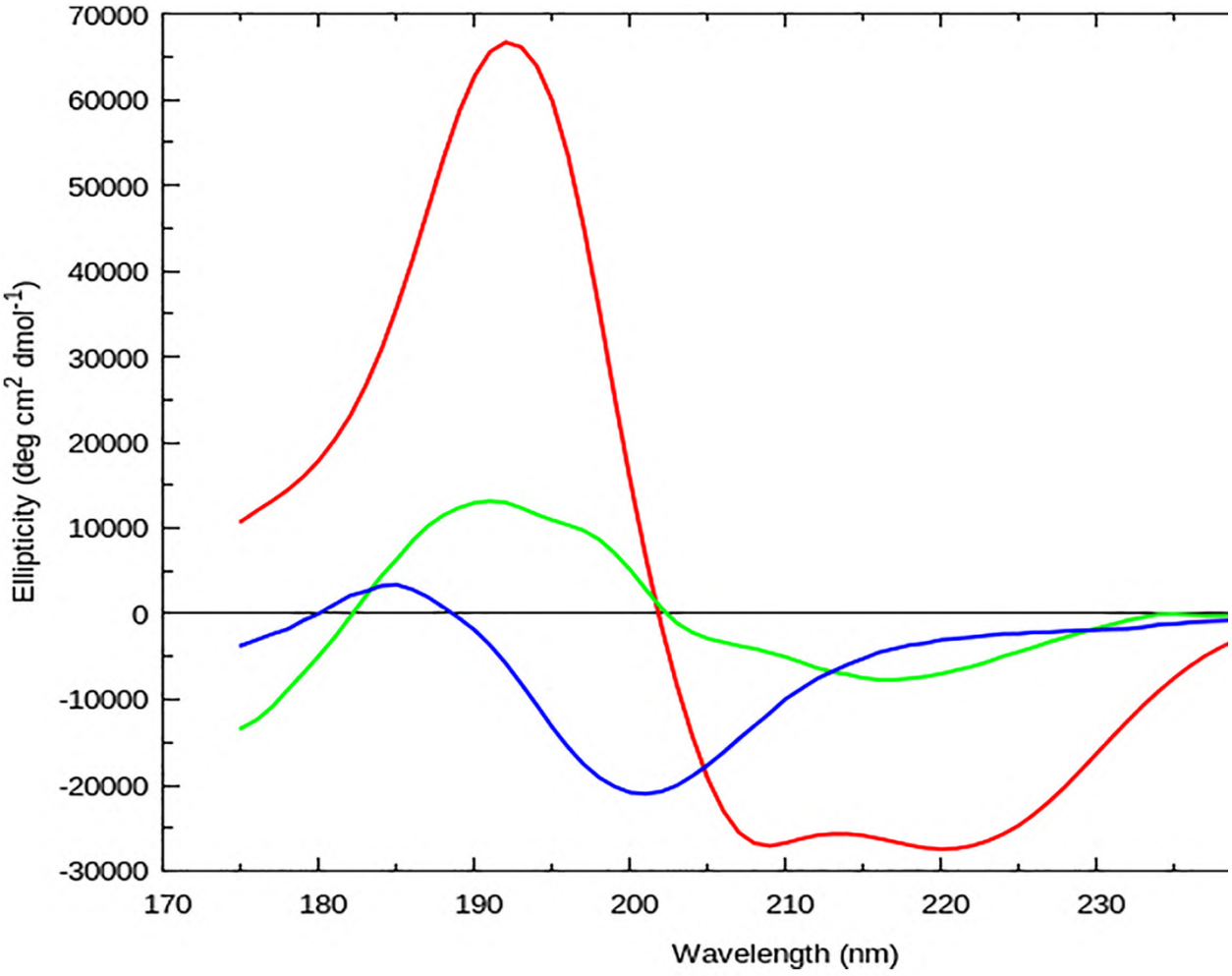

Figure 1-10. CD spectra of secondary structures in proteins. The solid red line shows the CD spectrum of a predominantly α-helical protein. The solid green line shows the CD spectrum of a predominantly β-sheet protein. The solid blue line shows the CD spectrum of a predominantly random coil protein. Figure from[76].

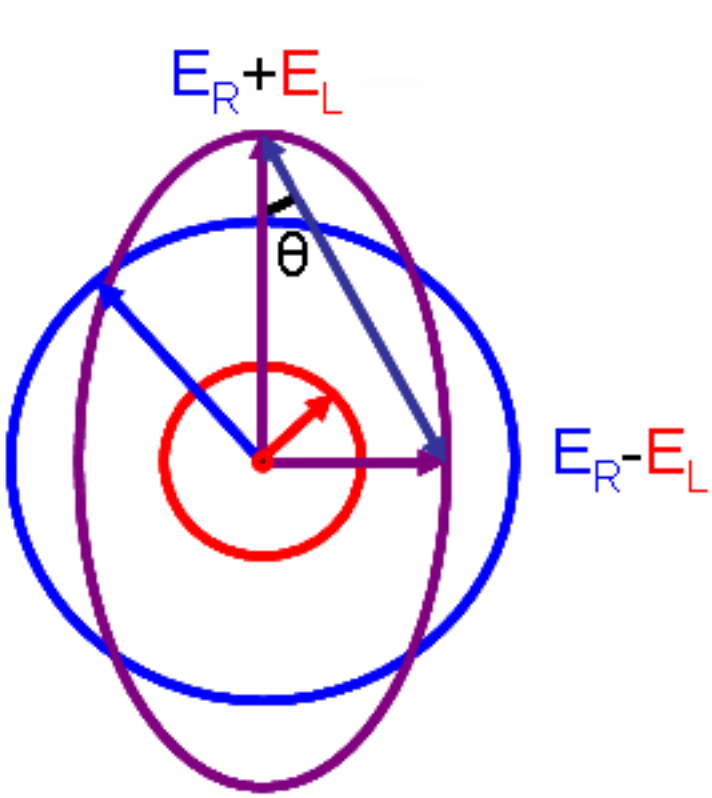

Figure 1-11. Elliptically polarized light. The elliptically polarized light shown in violet is composed of unequal contributions of the left (red) and right (blue) circular polarized light. The ER and EL are the magnitude of the electric field vectors of the of the left-circularly and right-circularly polarized light respectively. Figure from[75] .

1.6 References

1. Hueck, C.J., *Type III protein secretion systems in bacterial pathogens of animals and plants.* Microbiol. Mol. Biol. Rev., 1998. **62**(2): p. 379-433.
2. Kubori, T., et al., *Supramolecular structure of the Salmonella typhimurium type III protein secretion system.* Science, 1998. **280**(5363): p. 602-605.
3. Cornelis, G.R., *The type III secretion injectisome.* Nat. Rev. Microbiol., 2006. **4**(11): p. 811-825.
4. Coburn, B., I. Sekirov, and B.B. Finlay, *Type III secretion systems and disease.* Clin Microbiol Rev, 2007. **20**(4): p. 535-49.
5. CDC, *Summary of Notifiable Diseases, United States, 2007* MMWR, 2009. **56**(53): p. 1-94.
6. CDC, *Antibiotic Resistance Threats in the United States, 2013.* 2013.
7. Duncan, M.C., Linington, R. G., and Auerbuch, V., *Chemical inhibitors of the type three secretion system: disarming bacterial pathogens.* Antimicrobial Agents and Chemotherapy, 2012.
8. Keyser, P., Elofsson, M., Rosell, S., and Wolf-Watz, H., *Virulence blockers as alternatives to antibiotics: type III secretion inhibitors against Gram-negative bacteria.* J. Intern. Med., 2008.
9. May, A.E., and Khosla, C. , *Discovery and Mechanism of Type III Secretion System Inhibitors. in Isr.* J. Chem., 2013.
10. Dey, S., et al., *The type III secretion system needle, tip, and translocon.* Protein Sci, 2019. **28**(9): p. 1582-1593.
11. Galan, J.E., et al., *Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells.* Annu. Rev. Microbiol., 2014. **68**: p. 415-38.
12. Yip, C.K. and N.C. Strynadka, *New structural insights into the bacterial type III secretion system.* Trends Biochem.Sci., 2006. **31**(4): p. 223-230.
13. Chatterjee, S., et al., *Structure and biophysics of type III secretion in bacteria.* Biochemistry, 2013. **52**(15): p. 2508-17.
14. Kaur, K., *NMR Studies of Molecular Interaction Involved in Type III Secretion System,Sumoylation,and The RNA Binding Protein HuR. PhD. Diss. University of Kansas.* 2016.
15. Blocker, A.J., et al., *What's the point of the type III secretion system needle?* Proc. Natl. Acad. Sci. U.S.A., 2008. **105**(18): p. 6507-13.
16. Kosarewicz, A., Konigsmaier, L., and Marlovits, T. C., *The blueprint of the type-3 injectisome.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2012. **367**: p. 1140-1154.
17. Lee, P.C., et al., *Control of effector export by the Pseudomonas aeruginosa type III secretion proteins PcrG and PcrV.* Mol. Microbiol., 2010. **75**(4): p. 924-41.
18. Johnson, S., et al., *Self-chaperoning of the type III secretion system needle tip proteins IpaD and BipD.* J. Biol. Chem., 2007. **282**(6): p. 4035-4044.
19. Knight, M.J., et al., *Crystallization and preliminary X-ray diffraction analysis of BipD, a virulence factor from Burkholderia pseudomallei.* Acta Crystallograph.Sect.F.Struct.Biol.Cryst.Commun., 2006. **62**(Pt 8): p. 761-764.

20. Chatterjee, S., et al., *The crystal structures of the Salmonella type III secretion system tip protein SipD in complex with deoxycholate and chenodeoxycholate.* Protein Sci, 2011. **20**(1): p. 75-86.
21. Chaudhury, S., et al., *Structure of the Yersinia pestis tip protein LcrV refined to 1.65 Å resolution.* Acta Crystallogr. Sect. F, 2013. **69**(Pt 5): p. 477-81.
22. Chatterjee, S., et al., *The crystal structure of the Salmonella type III secretion system tip protein SipD in complex with deoxycholate and chenodeoxycholate.* Protein Sci., 2011. **20**: p. 75-86.
23. Rathinavelan, T., C. Tang, and R.N. De Guzman, *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., 2011. **286**(6): p. 4922-30.
24. Lunelli, M., et al., *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., 2011. **7**(8): p. e1002163.
25. Goure, J., Broz, P., Attree, O., Cornelis, G. R., and Attree, I., *Protective anti-V antibodies inhibit Pseudomonas and Yersinia translocon assembly within host membranes.* J. Infect. Dis. , 2005. **192**: p. 218-225.
26. Lee, P.C., et al., *Control of type III secretion activity and substrate specificity by the cytoplasmic regulator PcrG.* Proc. Natl. Acad. Sci. U.S.A., 2014. **111**(19): p. E2027-36.
27. Matson, J.S. and M.L. Nilles, *LcrG-LcrV interaction is required for control of Yops secretion in Yersinia pestis.* J. Bacteriol., 2001. **183**(17): p. 5082-91.
28. Lawton, D.G., et al., *Interactions of the type III secretion pathway proteins LcrV and LcrG from Yersinia pestis are mediated by coiled-coil domains.* J. Biol. Chem., 2002. **277**(41): p. 38714-22.
29. Basu, A., et al., *PcrG protects the two long helical oligomerization domains of PcrV, by an interaction mediated by the intramolecular coiled-coil region of PcrG.* BMC Struct Biol, 2014. **14**: p. 5.
30. Chaudhury, S., et al., *The LcrG tip chaperone protein of the Yersinia pestis type III secretion system is partially folded.* J. Mol. Biol., 2015. **427**: p. 3096-3109.
31. Chaudhury, S., et al., *Nuclear Magnetic Resonance Characterization of the Type III Secretion System Tip Chaperone Protein PcrG of Pseudomonas aeruginosa.* Biochemistry, 2015. **54**(43): p. 6576-85.
32. Feix, C.S.K.a.J.B., *Methods and Applications of Site-DirectedSpin Labeling EPR Spectroscopy.* Biological Magnetic Resonance, 2006: p. 251-281.
33. Drescher, M., *EPR in Protein Science Intrinsically Disordered Proteins.* Springer, 2012: p. 91-119.
34. Texas, U.o., *What is EPR?* https://sites.cns.utexas.edu/epr_facility/what-epr, 2021.
35. Langen R, O.K., Cascio 0, Hubbell WL, *Crystal structures of spin labeled T4 lysozyme mutants: implications for the interpretation of EPR spectra in terms of structure.* Biochemistry, 2000. **39**: p. 8396-8405.
36. Hubbell, W.L., D.S. Cafiso, and C. Altenbach, *Identifying conformational changes with site-directed spin labeling.* Nature Structural Biology, 2000. **7**(9): p. 735-739.
37. McHaourab HS, L.M., Hideg K, Hubbell WL, *Motion of spin-labeled side chains in T4 lysozyme, correlation with protein structure and dynamics.* Biochemistry, 1996. **35**: p. 7692-7704.

38. Alexander RS, N.S., Christianson OW, *Engineering the hydrophobic pocket of carbonic anhydrase-II.* Biochemistry 1991. **30**: p. 11064-11072.
39. Ani Der-Sarkissian, C.C.J., Jeannie Chen, Ralf Langen *Structural Organization of α-Synuclein Fibrils Studied by Site-directed Spin Labeling.* Journal of Biological Chemistry 2003. **278**(39): p. 37530-37535.
40. Scarpelli F, D.M., Rutters-Meijneke T, Holt A., Rijkers DTS, Killian JA, Huber M, *Aggregation of transmembrane peptides studied by spin-label EPR.* J. Phys. Chem. B., 2009. **113**: p. 12257- 12264.
41. Altenbach C, O.K., Trabanino RJ , Hideg K, Hubbell WL *Estimation of inter-residuedistances in spin labeled proteins at physiological temperatures : experimental strategies and practical limitations.* Biochemistry **40**: p. 15471- 15482.
42. Rabenstein MD, S.Y., *Determination of the distance between 2 spin labels attached to a macromolecule.* Proc. Natl. Acad. Sci. USA, 1995. **92**: p. 8239- 8243.
43. Steinhoff HJ, R.N., Thevis W, Lenz Y, Brandenburg O , Antson A, Dodson G, Wollmer, *Determination of interspin di stances between spin labels attached to insulin: comparison of electron paramagnetic resonance data with the X-ray structure.* Biophys. J., 1997. **73**: p. 3287-3298.
44. Bordignon, E., *Site-directed spin labeling of membrane proteins.* Top.Curr.Chern., 2011.
45. Sale K, S.C., Sharp KA, Hideg K, Fajer PG (2002), *Structural determination of spin label immobilization and orientation: a Monte Carlo minimization approach.* J .Magn. Reson., 2002. **156**: p. 104-112.
46. Fajer P, L.S., Liu YS, Peroza E, Budil D, Sale K, *Molecular modeling tools for dipolar EPR.* Biophys. J., 2004. **86**.
47. Watson BS, H.T., Eccleston JF, Davis C, Jameson DM, Johnson AE., *Macromolecular arrangement in the aminoacyl-tRNA.elongation factor Tu.GTP ternary complex. A fluorescence energy transfer study.* Biochemistry, 1995. **34**: p. 7904-7912.
48. Berney, C. and G. Danuser, *FRET or no FRET: a quantitative comparison.* Biophysical journal, 2003. **84**(6): p. 3992-4010.
49. Förster, T., *Intermolecular Energy Migration and Fluorescence.* 1948. **Ann. Phys.**(2).
50. Saltman LH, L.Y., Zaharias EM, Isberg RR, *A region of the Yersinia pseudotuberculosis invasin protein that contributes to high affinity binding to integrin receptors.* J. Biol. Chem. , 1996. **271**: p. 23438-44.
51. Bajar, B.T., Wang, E. S., Zhang, S., Lin, M. Z., & Chu, J. , *A Guide to Fluorescent Protein FRET Pairs. .* Sensors (Basel, Switzerland), 2016. **16**(1488).
52. Lakowicz, J.R., *Principles of Fluorescence Spectroscopy.* 1999(2).
53. Mark L. Richter, B.S., Richard E.McCarty , Gordon G. Hammes *Binding Stiochiometry and Structural Mapping of the ε Polypepetide of Choloroplast Coupling Factor 1.* Biochemistry, 1985. **25**: p. 5755-5763.
54. Schaap, M., et al., *Development of a steady-state FRET-based assay to identify inhibitors of the Keap1-Nrf2 protein-protein interaction.* Protein science : a publication of the Protein Society, 2013. **22**(12): p. 1812-1819.
55. Holmstrom, E.D.H., Andrea,Zheng, Wenwei,Nettels, Daniel,Best, Robert B.,Schuler, Benjamin, *Chapter Ten - Accurate Transfer Efficiencies, Distance Distributions, and Ensembles of Unfolded and Intrinsically Disordered Proteins From Single-Molecule FRET*, in *Methods in Enzymology*, E. Rhoades, Editor. 2018, Academic Press. p. 287-325.

56. Reed, *L. receptor-ligand dissociation* people.reed.edu. (n.d.), 2015.
57. Liao, J.M., V.; Dang, R.; Jiang, L., *Quantitative FRET(qFRET) Technology for the Determination of Protein–Protein Interaction Affinity in Solution.* Molecules, 2021. **26**.
58. Vaynberg, J.a.J.Q., *Weak protein-protein interactions as probed by NMR spectroscopy.* Trends Biotechnol, 2006. **24**: p. 22-27.
59. Perkins, J.R., Diboun, I., Dessailly, B. H., Lees, J. G., and Orengo, C., *Transient protein-protein interactions: structural, functional, and network properties.* Structure, 2010.
60. Ozbabacan, S.E.A., Engin, H. B., Gursoy, A., and Keskin, O., *Transient protein-protein interactions.* Protein Eng. Des. Sel., 2011. **24**: p. 634-648.
61. Takeuchi, K., and Wagner, G., *NMR studies of protein interactions.* Current Opinion in Structural Biology, 2006. **16**: p. 109-117.
62. Marintchev, A., Frueh, D., and Wagner, G., *NMR methods for studying protein-protein interactions involved in translation initiation.* Translation Initiation: Reconstituted Systems and Biophysical Methods, 2007. **430**.
63. Bieri, M., Kwan, A. H., Mobli, M., King, G. F., Mackay, J. P., and Gooley, P. R., *Macromolecular NMR spectroscopy for the non-spectroscopist: beyond macromolecular solution structure determination.* FEBS Journal, 2011. **278**: p. 704-715.
64. Foster, M.P., Wuttke, D. S., Clemens, K. R., Jahnke, W., Radhakrishnan, I., Tennant, L., and M. Reymond, Chung, J., and Wright, P. E., *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, 1998. **12**: p. 51-71.
65. Williamson, M.P., *Using chemical shift perturbation to characterise ligand binding.* Prog. Nucl. Magn. Reson. Spectrosc., 2013. **73**: p. 1-16.
66. Barrett, P.J., Chen, J., Cho, M.-K., Kim, J.-H., Lu, Z., Mathew, S., Peng, D., Song, Y., Van and W.D. Horn, Zhuang, T., Sönnichsen, F. D., and Sanders, C. R., *The Quiet Renaissance of Protein Nuclear Magnetic Resonance.* Biochemistry, 2013.
67. Broussard JA, G.K., *Research Techniques Made Simple: Methodology and Applications of Förster Resonance Energy Transfer (FRET) Microscopy.* J. Invest. Dermatol. , 2017. **137**: p. e185-e191.
68. Kleckner, I.R. and M.P. Foster, *An introduction to NMR-based approaches for measuring protein dynamics.* Biochim Biophys Acta, 2011. **1814**(8): p. 942-68.
69. Fernandez, C., *TROSY in NMR studies of the structure and function of large biological macromolecules.* Current Opinion in Structural Biology, 2003.
70. Hajduk, P., D. Augeri, and J. Mack, *NMR-based screening of proteins containing ^{13}C-labeled methyl groups.* J. Am. Chem. Soc., 2000. **122**: p. 7898-7904.
71. Tugarinov, V. and L.E. Kay, *Methyl Groups as Probes of Structure and Dynamics in NMR Studies of High-Molecular-Weight Proteins.* ChemBioChem, 2005. **6**(9): p. 1567-1577.
72. Tugarinov, V. and L.E. Kay, *Quantitative 13C and 2H NMR relaxation studies of the 723-residue enzyme malate synthase G reveal a dynamic binding interface.* Biochemistry, 2005. **44**(49): p. 15970-15977.
73. Miles, A.J. and B.A. Wallace, *Chapter 6 - Circular Dichroism Spectroscopy for Protein Characterization: Biopharmaceutical Applications*, in *Biophysical Characterization of Proteins in Developing Biopharmaceuticals*, D.J. Houde and S.A. Berkowitz, Editors. 2015, Elsevier: Amsterdam. p. 109-137.

74. Polavarapu, P.L.Z., C. X., *Vibrational circular dichroism: a new spectroscopic tool for biomolecular structural determination.* Fresenius Journal Anaytical Chemistry 2000. **366**: p. 727–734.
75. Hurlburt, N., *Circular Dichroism.* Chem. lib,tex. lib., 2020.
76. Rogers, D.M., et al., *Electronic Circular Dichroism Spectroscopy of Proteins.* Chem, 2019. **5**(11): p. 2751-2774.

Chapter 2: Structural Dynamics of *Yersinia* tip Chaperone Protein LcrG

2.1 Abstract

Yersinia pestis is the causative agent of bubonic plague, and it assembles the Type III Secretion System (T3SS) to inject virulence effector proteins into its target host cells. The tip protein LcrV and tip chaperone protein LcrG are essential proteins in *Yersinia* T3SS. Previous results by NMR spectroscopy showed that LcrG is a partially folded alpha helical protein with three alpha helices (helix α1, α2 and α3) that do not interact with each other. Other results in literature suggested that the alpha helices of LcrG may interact with each other. Here, we used Electron Paramagnetic Resonance (EPR) spectroscopy to determine the different conformations of LcrG. For EPR studies, we engineered cysteine mutations for the attachment of spin labels in LcrG. Our EPR results revealed that when LcrG is in the free form, there is a population of LcrG that has transient tertiary structures where the two helices α1 and α3 are in close contact with each other. Further, EPR revealed that, the two helices do not interact suggesting that LcrG exists in an open conformation, when bound to LcrV. EPR provided additional insight into the different conformations of LcrG in the free and when bound to LcrV.

2.2 Introduction

Yersinia pestis, the causative agent of Bubonic plague uses the T3SS to inject virulent effector proteins into target host cells to initiate infection [1],[2] The T3SS consists of the needle complex, chaperones, and effector proteins. The needle complex is a bacterial virulence protein-injector which is assembled from over 20 different proteins. It consists of a base, a needle, a tip and a translocon [3],[4]. The base is anchored at the inner and outer bacterial membranes[3]. The needle protrudes from the base to the intermembrane space between the bacteria and the host cell[3]. The needle is capped with the tip which consists of multiple copies of the tip protein[5].

The tip serves as a platform for the assembly of the translocon which forms a pore on the host cells membranes to transport virulence effector proteins called Yop effectors in *Yersinia*[5].

The tip in *Yersinia* is formed by 5 copies of the tip protein LcrV,[6] and functions as an environmental sensor for the deployment of the translocon. The crystal structure of LcrV[7],[8] shows structural similarity to the tip proteins in *Salmonella* SipD, *Shigella* IpaD and *Burkhloderia* BipD[9],[10],[11],[12]. At 37^{o}C, the T3SS genes[13] are activated by *Yersinia* to initiate the assembly of the base, needle and the tip[5]. This partially formed needle complex comes in contact with the eukaryotic host cells which triggers the final stages of assembly wherein the translocon proteins are released[14],[15],[16].

Prior to the assembly of the tip complex, the tip protein is tightly bound to a small protein LcrG (95 residues,8 kDa)[17],[18],[19],[20]. In this context, LcrG functions as a chaperone to LcrV[21],[22]. and the LcrG-LcrV binding is tight with a dissociation constant (k_d) of 140 nM[23] LcrG also acts as a regulator of Yop secretion[19],[20],[24],[25].

There are currently no crystal structures of LcrG. The NMR data indicated that LcrG is a partially folded protein which consists of three alpha helices (helices α1, α2 and α3) as shown in Fig 1-1A. The NOE based NMR data could not be detect any tertiary contacts between the helices. NMR also shows that LcrG has a dynamic nature with fast nanoscale motions[26]. Computational modelling predicted that LcrG existed as a coiled coil[27]. Other groups also proposed that LcrG forms a coiled-coil or an alpha hairpin upon binding with LcrV[28].

The dynamic nature of LcrG and the predicted coiled-coil structure upon binding with LcrV[26],[28] leads to the hypothesis that there could be a population of LcrG, where the helices are in close contact to each other. This closed conformation could not be determined by NOE based NMR experiments[26]. The NOE-based NMR experiments can detect distances up to 6Å,

so we use EPR spectroscopy where distances more than 6Å can be detected to determine if there is a population of LcrG where the two helices are close to each other that may not have been detected by NMR. We introduce spin labels on helix α1 and helix α3 and acquired the EPR data from free LcrG and the LcrG-LcrV complex. The EPR data revealed that there is a population of LcrG where the helices α1 and α3 are in close contact with other with a distance of 9Å. We designated this as the closed conformation of LcrG as shown in Fig 1-1C. The current models of LcrG-LcrV interaction suggested that LcrG would have a coiled or alpha hairpin [28] structure as depicted in Fig1-1B, but our EPR data suggested that when LcrV is bound to LcrG the spin-spin distances between helices α1 and α3 is greater than 25Å (Fig1-1D). This structural conformation of LcrG is designated as the open conformation. EPR provides us with additional insight into the different structural conformations of LcrG when in complex with LcrV.

2.2 Methods

2.3.1 Protein expression and purification

$LcrG^{FL}$ and $LcrG^{7-73}$ constructs were made as fusion proteins in pDZ1. pDZ1 were plasmids containing pET-21a along with His_6 –tagged GB1 protein and tobacco etch virus (TEV) protease cleavage site at the N-terminus[9]. GB1 is a β1 immunoglobin-binding domain of streptococcus protein G that enhances the solubility of the proteins. $LcrG^{7-73}$ has cysteine residue in the 34^{th} position so $LcrG^{7-73}$ was used as control. Site-directed mutagenesis by QuickChange (Stratagene) was used to introduce point mutation in $LcrG^{FL}$, $LcrG^{7-73}$ and $LcrG^{7-73}$ C34S to make $LcrG^{7-73}$ D65C and $LcrG^{7-73}$ C34S/D65C for attaching the spin labels.

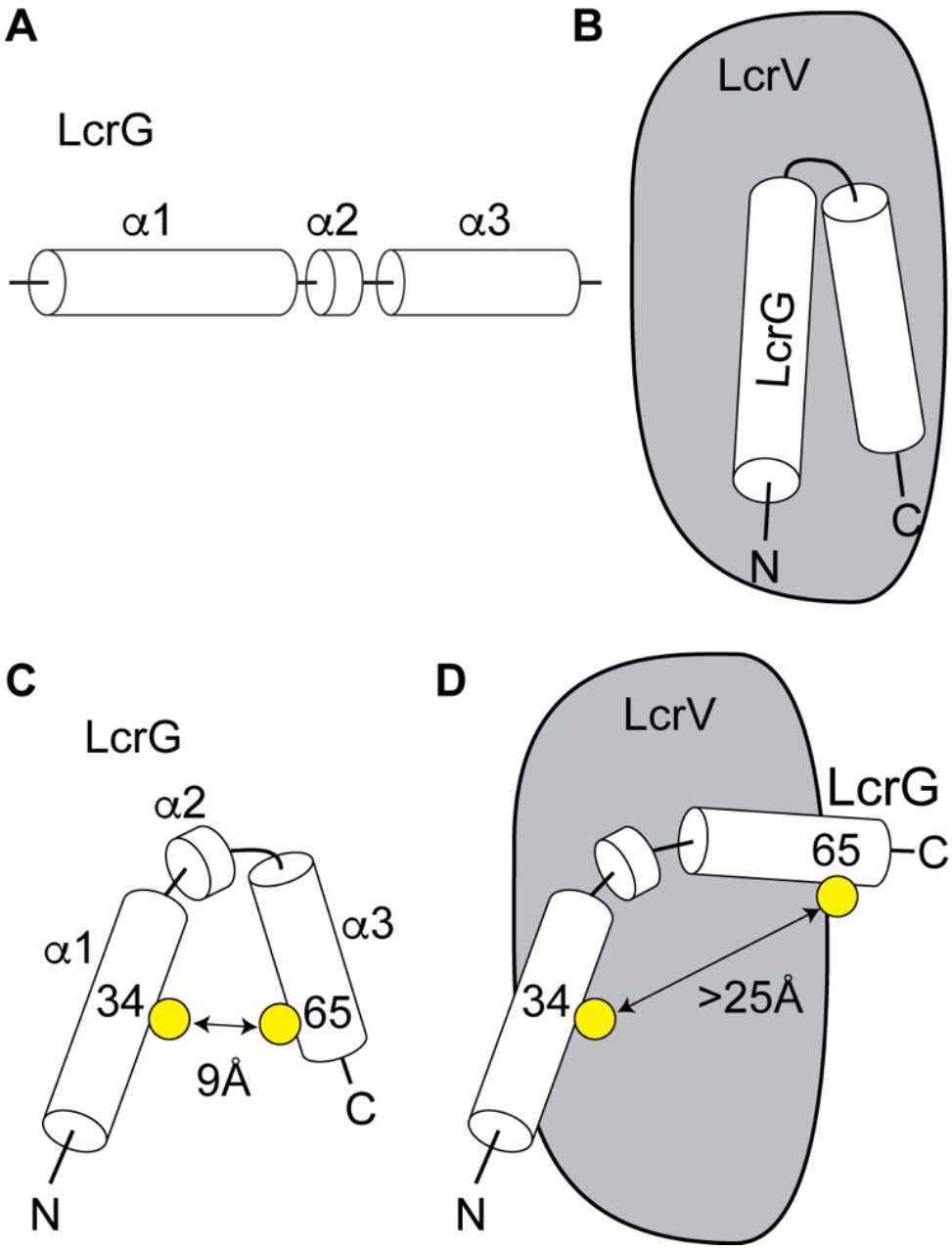

Figure 2-1. Models of the open and closed conformations of LcrG. (A) A model of LcrG based on NMR shows 3 helices in LcrG (helix α1, helix α2 and helix α3) that do not interact with each other. **(B)** A model of LcrG-LcrV interaction shows LcrG forms an alpha-helical hairpin upon binding to LcrV. **(C)** EPR spin-spin interaction revealed a closed conformation of LcrG where helix α1 and helix α3 are close to each other by 9 Å. This is designated herein as the closed conformation of LcrG. **(D)** EPR of LcrG-LcrV interaction revealed that helix α1 and helix α3 of LcrG move farther apart upon binding to LcrV, forming what is designated herein as the open conformation of LcrG

Plasmids for expressing LcrG constructs were transformed in *E. coli* BL21 (DE3) DNAY. A 10 mL Terrific Broth (TB) overnight starter culture was supplemented with antibiotics carbenicillin and kanamycin. This starter culture was used to inoculate 1 L of TB media, and cells are grown at 37 ºC in a shaker incubator at 200rpm. At A_{600} 0.7-0.8, cells were induced with 1 mM IPTG, and cell growth was continued at 15 ºC overnight. Cells were harvested by centrifugation (4000 rpm, 2392×*g*, 12 min, 4 ºC). The cell pellet was resuspended in 40 mL binding buffer (20mM Tris pH 8.0, 500 mM NaCl, 5mM imidazole) with 600 µL 10 mM PMSF.

The resuspended cells were sonicated for a total of 5 min 20 sec using a sonication cycle of 2 s pulse on and 30 s pulse off. The cell lysate is centrifuged at 13,000 rpm (13865×*g* in a Beckman JA-25.50 rotor) for 15 min at 4ºC. The supernatant was added to 700 µl of 5% PEI to precipitate nucleic acids, followed by centrifugation at 13,000 rpm (13865×*g*) for 15 min at 4ºC, the supernatant is loaded into a 5 mL nickel column, followed by three washes using 50 mL binding buffer. The protein eluted in 50 mL elution buffer (20 mM Tris pH 8.0, 500 mM NaCl, 250 mM imidazole). TEV protease (250 µL of 0.07 mM stock) was added, and the samples were dialyzed at room temperature in 1 L TEV buffer (50 mM Tris pH 8.0, 0.5 mM EDTA, 1 mM DTT, 100 mM NaCl) overnight, followed by further dialysis in 1L binding buffer for 4 hrs and in 2 L binding buffer overnight. The digest was loaded into a 5 mL nickel column, the flow through is collected, followed by 3 washes with 50 mL binding buffer. The fractions containing purified $LcrG^{7-73}$ constructs appears in the flow through and the first wash, whereas the His_6-GB1 tag was retained in the column which is eluted by the elution buffer. After digestion of the fusion proteins with TEV protease, purified LcrV and $LcrG^{7-73}$ constructs used herein contain 3residues (GHM) at their N-termini as cloning artifact.

Purified proteins were dialyzed twice in 1 L EPR buffer (10 mM $NaPO_4$,10 mM NaCl, pH 7.0), concentrated using Amicon 3k, and the protein concentration is estimated by absorbance at A_{280}.

As described previously[7], LcrV was expressed as LcrV C273 (residues 28-322) was expressed as a fusion protein(His_6-GB1-TEV protease cleavage site-LcrV) in pDZ1.LcrV was purified by the same protocol as described above, except for the following modifications: after TEV digestion, the digest was loaded on a 10 ml nickel column, the flow through was collected and followed by wash 1 in 50 ml binding buffer, wash 2 with 50 ml (20mM Tris pH 8.0,500 mM NaCl, 25mM Imidazole) and wash 3 (20mM Tris pH 8.0,500 mM NaCl, 50mM Imidazole). The fractions flow through, wash 1 and wash 2, whereas the His_6-GB1 tag appeared in the elution fraction.

2.3.2 *Site directed spin labeling*

The nitroxide spin label MTSL (1-Oxyl-2,2,5,5-tetramethyl-3-pyrroline-methyl) methanethiosulfonate (Toronto Research Chemicals) was attached to the protein samples by Site Directed Spin Labeling (SDSL). 10 mg of MTSL was dissolved in 250ul of acetone to make a stock solution of 150mM. A 30-fold excess DTT was added to remove disulfide bonds from the protein samples and incubated for 30 minutes at room temperature. The excess DTT from the reduced samples was removed by a buffer exchange in the EPR buffer using NAP-5 columns. A 10-fold excess of MTSL was added to $LcrG^{FL}$, $LcrG^{7-73}$ and $LcrG^{7-73}$ C34S/D65C. 20-fold excess of MTSL was added to $LcrG^{7-73}$ D65C, $LcrG^{FL}$D65C and $LcrG^{FL}$ Q74C as it had 2 cysteines for double labeling. All the samples were incubated overnight and the nitroxide spin label (R1) was attached as a side chain to the cysteine residue as shown in Fig 2-5. Unbound MTSL was

removed by exchange in EPR buffer using a NAP-5 column. The attachment of the R1 spin label to the LcrG constructs was verified by ESI-Mass Spectroscopy.

2.3.3 *Circular dichroism spectroscopy*

CD spectra was acquired with a JASCO J-815 spectropolarimeter. Protein samples were diluted to a final concentration of ~ 5μM (0.05 mg/ml) in EPR buffer. Spectra were acquired from 190 nm to 260 nm in triplicate at 20^{o}C with a scan rate of 50 nm/min for unlabeled and labeled protein samples.

2.3.4 *EPR spectroscopy*

Spectra of single labeled LcrG $^{7-73}$C34R1, LcrG $^{7-73}$ C34S/D65R1, and double labeled LcrG$^{7-73}$ C34R1/D65R1 were acquired. The samples were in buffer containing 10 mM Phosphate buffer, pH 7.0, 10 mM NaCl. The final concentrations of all the LcrG constructs were 1.3 –0.5 mM in a volume of~ 500 μL.

The EPR spectra were recorded at 150 K on a Bruker EMX spectrometer at a microwave power of 0.02mW sweeping the magnetic field from 3,260 to 3,460 G at a frequency of 9.45 GHz. Spectra were recorded a minimum of two times; little to no variation was observed between trials. The spin–spin distances between the helices of LcrG was assessed by comparing the EPR spectrum of LcrG$^{7-73}$ C34R1/D65CR1 to the spectrum formed by adding that of LcrG$^{7-73}$C34R1 and that of LcrG$^{7-73}$ C34S/D65R1. The spin–spin distances of bound LcrG to LcrV was estimated by comparing the EPR spectrum of LcrV bound to LcrG$^{7-73}$ C34R1/D65CR1 to the spectrum formed by adding that of LcrV bound LcrG$^{7-73}$ C34R1 and that of LcrV bound to LcrG$^{7-73}$ C34S/D65CR1. The spin–spin distances between the helices LcrG^{FL} was estimated by

comparing the EPR spectrum $LcrG^{FL}$ C34R1/D65CR1 and $LcrG^{FL}$ C34R1/Q74CR1 to the spectrum of $LcrG^{FL}$ C34R1. The spin–spin distances was calculated using a Monte Carlo/simplex Gaussian convolution method[29].

2.4 Results

2.4.1 Protein expression and purification for EPR studies

LcrG constructs were expressed and purified under native conditions. $LcrG^{FL}$ and $LcrG^{7-73}$ was expressed in the soluble fraction and used previously [26],[30] to determine its structure by NMR. Site directed mutagenesis introduced cysteine mutations in the $LcrG^{7-73}$ and $LcrG^{7-73}$ C34S for the EPR studies to obtain $LcrG^{7-73}$ C34S D65C and $LcrG^{7-73}$ D65C. The purified $LcrG^{7-73}$ C34S D65C and $LcrG^{7-73}$ D65C was soluble and stable like $LcrG^{7-73}$. Site directed mutagenesis introduced cysteine mutations in the $LcrG^{FL}$ to obtain $LcrG^{FL}$ Q74C and $LcrG^{FL}$ D65C. The purified $LcrG^{7-73}$ C34S D65C, $LcrG^{7-73}$ D65C, $LcrG^{FL}$ Q74C and $LcrG^{FL}$ D65C were soluble and stable like $LcrG^{FL}$ and $LcrG^{7-73}$. The final concentrations of these proteins at A_{280} were ~0.9-1.3mM.

LcrV was also expressed and purified under native conditions[7] and the final concentration was ~0.5mM .

2.4.2 Electrospray ionization mass spectrometry (ESI-MS) shows site directed spin labeling of $LcrG^{7-73}$, $LcrG^{7-73}$ C34S/D65C and $LcrG^{7-73}$D65C

Fig 2-2 shows that LcrG constructs, $LcrG^{7-73}$ and $LcrG^{7-73}$ C34S/D65C formed dimers in the absence of a reducing agent. The theoretical mass of $LcrG^{7-73}$ was estimated to be 8009 Da., in Fig 1-3A majority of the protein molecules had a mass of 16018 Da., which is twice the

predicted mass of LcrG$^{7-73}$. When DTT was added to the protein solution followed by R1 spin label attachment, the mass of the sample was 8194Da, which was the theoretical mass of LcrG $^{7-73}$ (8009 Da), along with the R1 side chain with a mass of 185 Da, as shown in Fig 1-3B. In Fig1-3C, I observed the dimerization of the LcrG$^{7-73}$ C34S/D65C sample like LcrG$^{7-73}$. After addition of MTSL (Fig1-3D) majority of the protein molecules had a mass of 8166 Da, denoting that R1 is attached to LcrG$^{7-73}$ C34S/D65C with a mass of 185 Da. LcrG$^{7-73}$ D65C has two cysteines for MTSL side chain attachment. Fig 1-4A shows the mass of LcrG$^{7-73}$ D65C was estimated to be 7996 Da, after R1 spin label attachment, the mass was 8366Da (Fig 1-4B). The difference in mass before and after labeling is 370 Da, showing that two R1 spin labels were attached to LcrG$^{7-73}$ D65C.

2.4.3 *Circular dichroism spectroscopy*

The CD spectra for the LcrG constructs shows that the mutations did not change the overall secondary structure of the protein. The Molar Ellipticity ratio at 222nm and 208nm ($\theta_{222}/\ \theta_{208}$) for all LcrG constructs was 0.8 (Fig 2-4A,2-4B and 2-4C), suggesting that there were no changes in the secondary structures due to the introduction of the cysteine mutation.

After the addition of the R1 spin label, the labeled samples LcrG $^{7-73}$ C34R1, LcrG $^{7-73}$ C34S/D65R1 and LcrG $^{7-73}$ C34R1/D65R1 had a $\theta_{222}/\ \theta_{208}$ ratio of 0.78, 0.85 and 0.93 respectively, showing that the addition of the spin label did not change the overall secondary structure of the LcrG constructs.

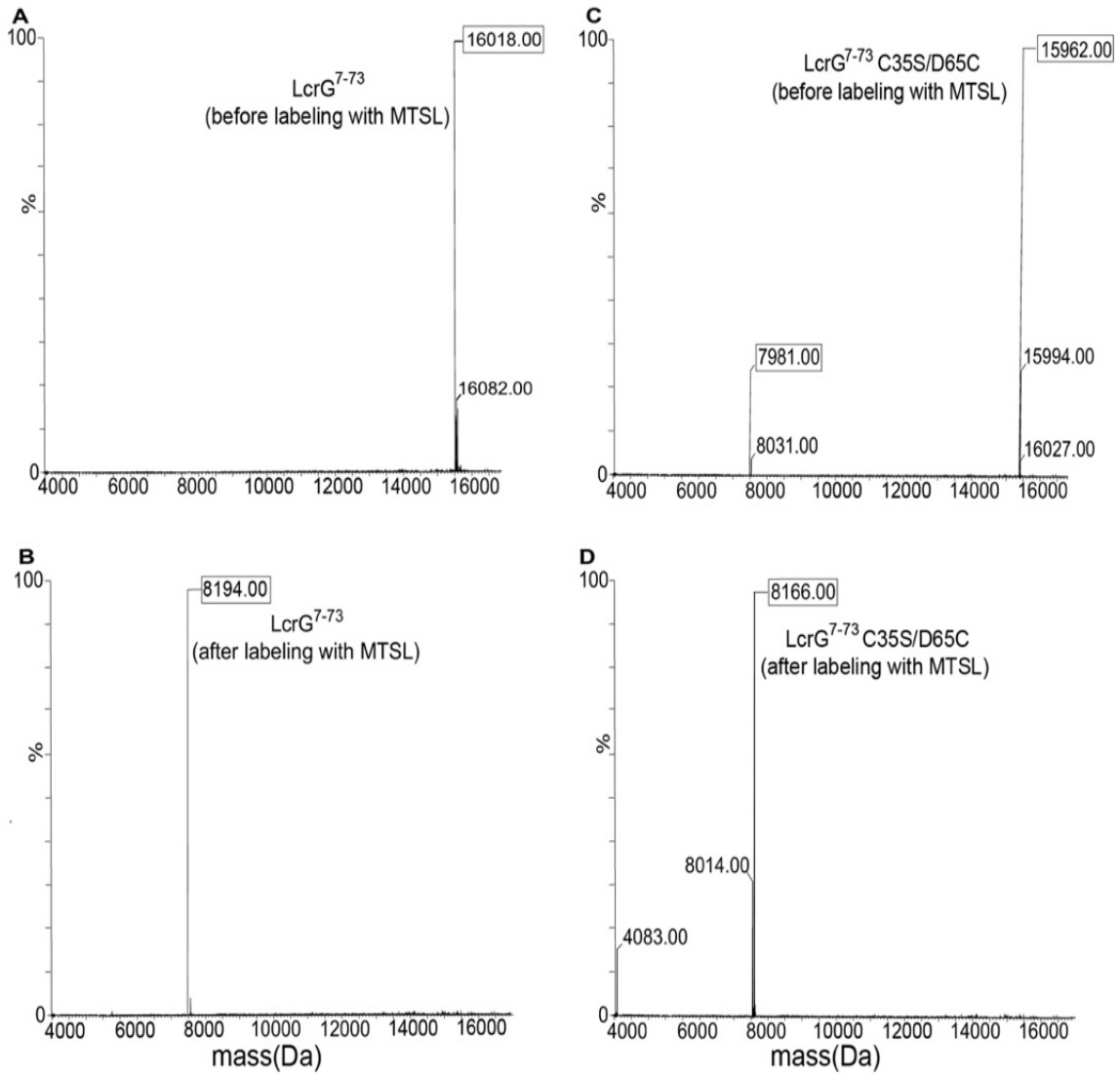

Figure 2-2. Electrospray ionization mass spectrometry (ESI-MS) of LcrG single mutant proteins to confirm the extent of labeling. ESI-MS on LcrG$^{7-73}$ and LcrG$^{7-73}$ C34S/D65C **(A)** before MTSL-labeling (theoretical MW, 8009 Da), and **(B)** after MTSL labeling (theoretical MW,8166 Da). (**C)** before MTSL-labeling (theoretical MW, 7981 Da), and **(D)** after MTSL labeling (theoretical MW,8166 Da). MTSL conjugation at one cysteine resulted in an expected increase in mass of 185Da.

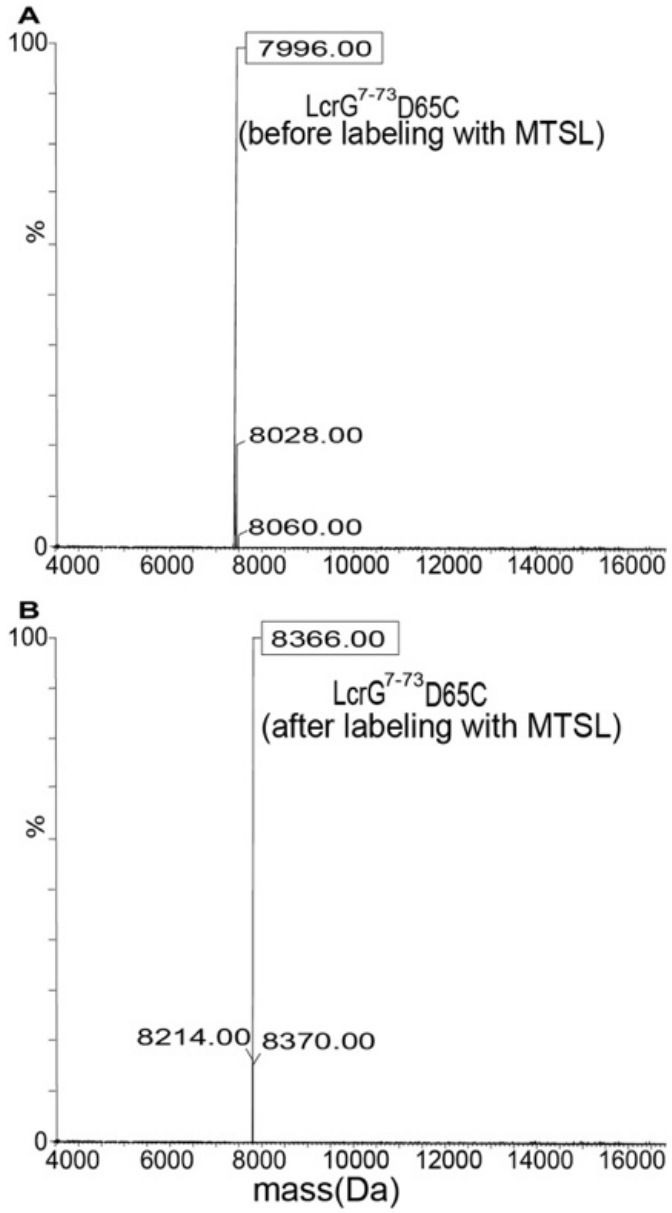

Figure 2-3. Electrospray ionization mass spectrometry (ESI-MS) of LcrG $^{7-73}$ D65C protein to confirm the extent of labeling. ESI-MS on LcrG $^{7-73}$ D65C mutant **(A)** before MTSL-labeling (theoretical MW, 7996Da), and **(B)** after MTSL labeling (theoretical MW,8366). MTSL conjugation at two cysteines resulted in an expected increase in mass of 370 Da.

2.4.4 LcrG has different structural conformations

Based on the capacity to detect long range contacts on lowly populated states, EPR spectroscopy was used to test our hypothesis that the helix α1 and helix α3 of LcrG interacted with each other. For the EPR studies, nitroxide spin label R1(Fig 2-5A) was covalently attached as a side to the cysteine residues by coupling of MTSL.

To determine the interaction between helix α1 and helix α3, three LcrG constructs were spin labeled and used for EPR studies — (1) $LcrG^{7-73}$ with a single cysteine residue C34 in helix α1 for single spin labeling. (2) $LcrG^{7-73}$ engineered with two point mutations (C34 and D65C) for single spin label attachment at position D65 on helix α3. (3) $LcrG^{7-73}$ engineered with a point mutation on D65C for attachment of spin labels at residue C34 on helix α1 and residue D65 on helix α3, as shown in Fig 2-5B. The LcrG constructs with single spin labels acted as the control for EPR spectroscopy whereas the construct containing the double spin labels was used as the sample to determine spin-spin distances.

EPR spectra acquired for single spin labeled LcrG on C34R1 and D65R1were added, which resulted in the black spectrum (Fig 1-6C). Results of EPR for the double spin labeled C34R1 and D65R1 (red spectrum, Fig 1-6C) showed spectrum broadening indicating spin-spin distance between C34R1(helix α1) and D65R1(helix α3). The distance between the two spin labels was calculated to be 9Å. Our EPR results suggested that there is a population of LcrG where helix α1 helix α3 were close to each other and we termed this spin-spin interaction as the closed conformation. This EPR data is the only experimental evidence which shows direct interaction between helix α1and helix α3 of LcrG.

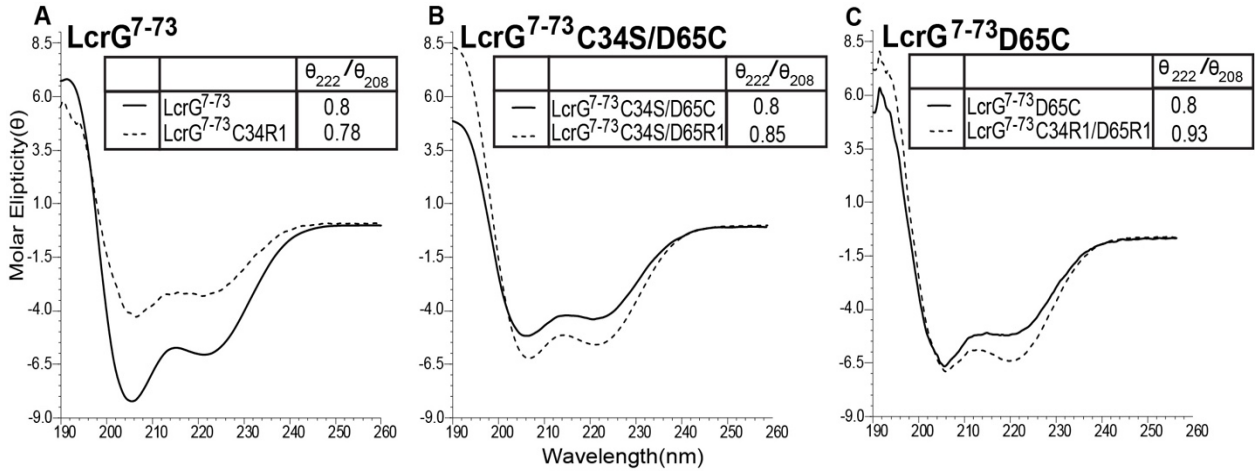

Figure 2-4. CD spectroscopy of LcrG constructs used for EPR studies. (A) CD spectra of $LcrG^{7-73}$ before and after ($LcrG^{7-73}$C34R1) MTSL side chain attachment. **(B)** CD spectra of $LcrG^{7-73}$ C34S/D65C before and after ($LcrG^{773}$C34S/D65R1) MTSL side chain attachment. **(C)** CD spectra of $LcrG^{7-73}$D65C before and after MTSL($LcrG^{7-73}$D65R1) side chain attachment.

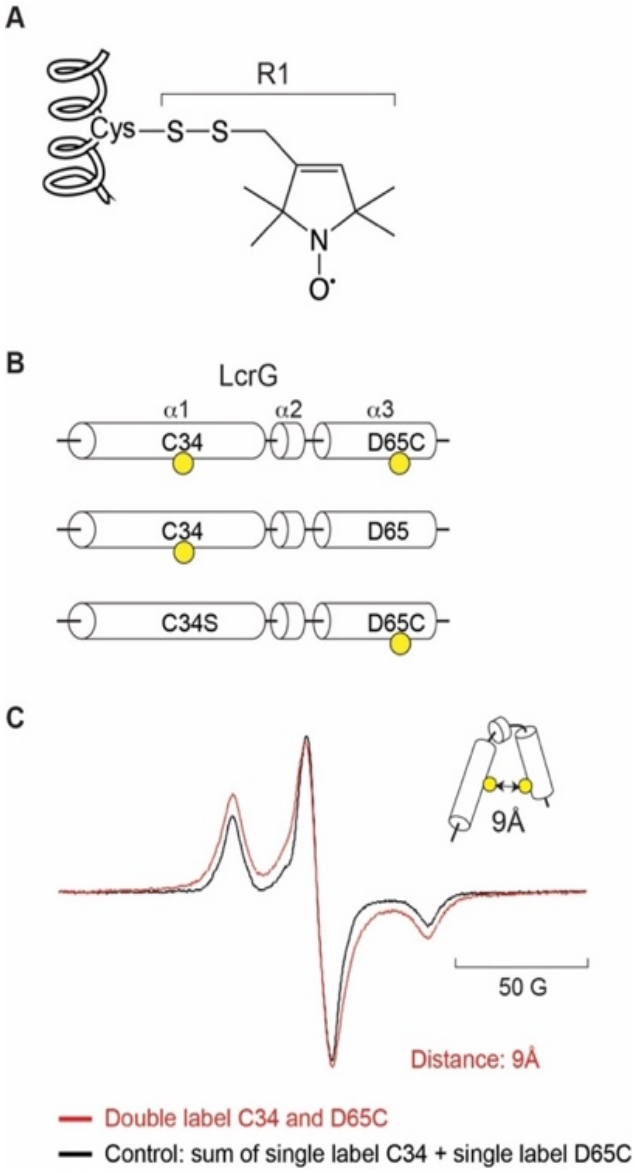

Figure 2-5. The closed conformation of LcrG as revealed by EPR. **(A)** MTSL spin label as an amino acid side chain (R1). **(B)** The LcrG spin-labeled constructs used to determine the spin-spin interaction between helix α1 and helix α3. **(C)** EPR spectra of the spin-spin interaction of $LcrG^{7-73}$ C34 and D65C.

EPR spectroscopy is also used to determine the spin-spin interaction between helix α1 and helix α3 when LcrG is bound to LcrV. The LcrG constructs used in the previous studies were complexed with unlabeled LcrV at a 1:1.2 molar ratio and the spin-spin interaction of LcrG C34R1 and D65R1 is determined.

The additive spectrum of the two single spin labeled proteins — LcrG C34R1 bound to LcrV and LcrG C34S/D65R1 bound to LcrV served as the control (Fig 2-6A). The EPR results indicated that when LcrG binds to LcrV, spin-spin distances between C34R1 and D65R1 disappear. The EPR spectrum of double spin labeled LcrG C34R1/D65R1bound to LcrV looked identical to the control spectrum (black spectrum Fig 2-6B), as the two spectra were superimposed.

Our EPR results indicate that when LcrG binds to LcrV the distance between the C34R1 on helix α1 and D65R1 on helix α3 is larger than 25Å.We designated this increase of spin-spin distance over 25Å between helix α1 and helix α3 as the open conformation of LcrG.

2.4.5 *EPR Spectroscopy of LcrG^{FL}*

The preliminary EPR spectra acquired for full length LcrG (LcrG^{FL}) showed no distance information as shown in Fig 2-7. The data acquired was inconclusive as I used only one control LcrG^{FL} C34 for the EPR studies. I planned to generate LcrG^{FL} C34S/D65C and LcrG^{FL} C34S/Q74C using LcrG^{FL} C34S as the template but the site directed mutagenesis experiments failed.

2.5 Discussion

EPR spectroscopy provides new insight into the structural conformations of tip LcrG. LcrG plays an important role in the virulence *Yersinia pestis* by functioning as the chaperone to the tip protein LcrV and as a regulator of Yop effect proteins. A homolog of LcrG is found in the

Pseudomonas T3SS known as PcrG. LcrG and PcrG form a family of tip chaperones and regulator of effector proteins which are essential for the virulence of *Yersinia pestis* and *Pseudomonas aeruginosa*. There are no crystal structures for the tip chaperone family of proteins [31]. Computational studies predicted the structures of LcrG and PcrG as a coiled coil and four helical bundle[27],[31]. Prior to our EPR studies, the current knowledge of the LcrG and PcrG comes from NMR spectroscopy which shows that LcrG and PcrG have secondary alpha helices but lack tertiary contacts[26, 31]. The NOE-based NMR experiments did not detect any long-range interactions between helix α1 and helix α3[26]. The NMR structure characterizes LcrG with three alpha helices which are connected to each other by disordered linker regions, and is highly dynamic showing fast nanoscale backbone motion[26]. The full-length LcrG ($LcrG^{FL}$) showed a poor quality two-dimensional (2D) ^{1}H-^{15}N HSQC NMR spectrum where the amide peaks were barely resolved and the numbers of peaks were much less than expected from a 95-residue protein[26]. Based on the secondary structure predicted by PSIPRED[32], the disordered tails of LcrG were truncated, while maintaining the hydrophilic terminal residues to generate the truncated construct spanning residues Asp-7 to Arg-73 ($LcrG^{7-73}$)[30]. $LcrG^{7-73}$ somewhat improved the 2D ^{1}H-^{15}N HSQC spectrum as compared to $LcrG^{FL}$, but $LcrG^{7-73}$ did not yield an ideal NMR spectrum either[30]. Only 58% of the residues in the protein were represented in the 2D ^{1}H-^{15}N HSQC spectrum of $LcrG^{7-73}$ and the peaks were broadened and poorly resolved[30]. It was reasoned that this non-ideal NMR spectrum of $LcrG^{7-73}$ could be due to the presence of a native cysteine at position 34 (C34), and this C34 might be responsible for forming disulphide linkage and contribute to the aggregation of the protein[30]. Hence, the native cysteine (C34) was mutated into serine and thereby creating the constructs $LcrG^{7-73}$

C34S[30]. LcrG$^{7-73}$ C34S yielded a better quality ^{1}H-^{15}N HSQC spectrum with sharp and well-resolved peaks[30].

The circular dichroism (CD) spectra of LcrG^{FL}, LcrG$^{7-73}$ and LcrG$^{7-73}$ C34S showed that the LcrG constructs are primarily α-helical proteins as demonstrated by the characteristic minima at 222 and 208 nm[26],[30]. The θ_{222}/ θ_{208} ratio of these constructs were ~0.8[26],[30]. The Yop secretion assay showed that LcrG$^{7-73}$ C34S successfully complemented the Δ*lcr*G mutation by restoring the secretion of effector proteins in YopM in *Yersinia pestis* like the wild type LcrG[30]. Based on the structural similarities and function of LcrG^{FL}, LcrG$^{7-73}$ and LcrG$^{7-73}$ C34S, we used these constructs to generate cysteine mutations to identify the interaction of helix α1and helix α3 in LcrG using EPR.

Our EPR data acquired from the LcrG$^{7-73}$, LcrG$^{7-73}$ C34S/D65C and LcrG$^{7-73}$ D65C constructs show that there is a population in LcrG where helix α1 and helix α3 are in close contact. The spin-spin distance between the two helices is calculated to be 9Å, thus forming a close conformation.

The function of LcrG is to maintain the structural integrity of LcrV in the bacterial cytosol during the assembly of the needle complex. The LcrG-LcrV complex shows tight binding with nanomolar affinities. PcrV, the homolog of LcrV in *Pseudomonas aeruginosa* forms a tight binding complex with PcrG[[33],[34],[35].There are currently no crystal structures of the LcrG-LcrV and the PcrG-PcrV complexes. The structural model of LcrG and LcrV complex is derived from known crystal structures of the tip proteins which include *Shigella* IpaD, *Salmonella* SipD, *Burkholderia* BipD *Yersinia* LcrV, [9],[10],[11],[12].The common structural features of the tip proteins include a mixed α/β region and a long central coiled coil domain at the C-termini of the proteins[7]. A major difference among the tip proteins is the N-terminal

alpha helical hairpin which is common among the IpaD, SipD and BipD but is missing in LcrV (Fig 1-2(A), Chapter 1)[9]. LcrG is an alpha helical protein which was predicted to interact with LcrV in a manner similar to the N-terminal alpha hairpin of the tip protein SipD[28](Fig1-2(E), Chapter 1)[9]. This prediction led to our hypothesis that when LcrG would interact with LcrV the helices α1 and α3 would be in close contact with each other (Fig 2-1B). Our EPR results show that when LcrG is bound to LcrV, instead of helix α1 and helix α3 forming an alpha hairpin, the distance between the spin labels attached to C34 to helix α1 and helix α3 D65C have a distance greater than 25Å.This is called the open conformation of LcrG.

In summary, EPR provides additional insight into the different structural conformation of LcrG. LcrG in its free form has a population where helix α1 and helix α3 are in a closed conformation with a distance of 9Å. When LcrG is bound to LcrV, helix α1 and helix α3 are in an open conformation with a distance more than 25Å.

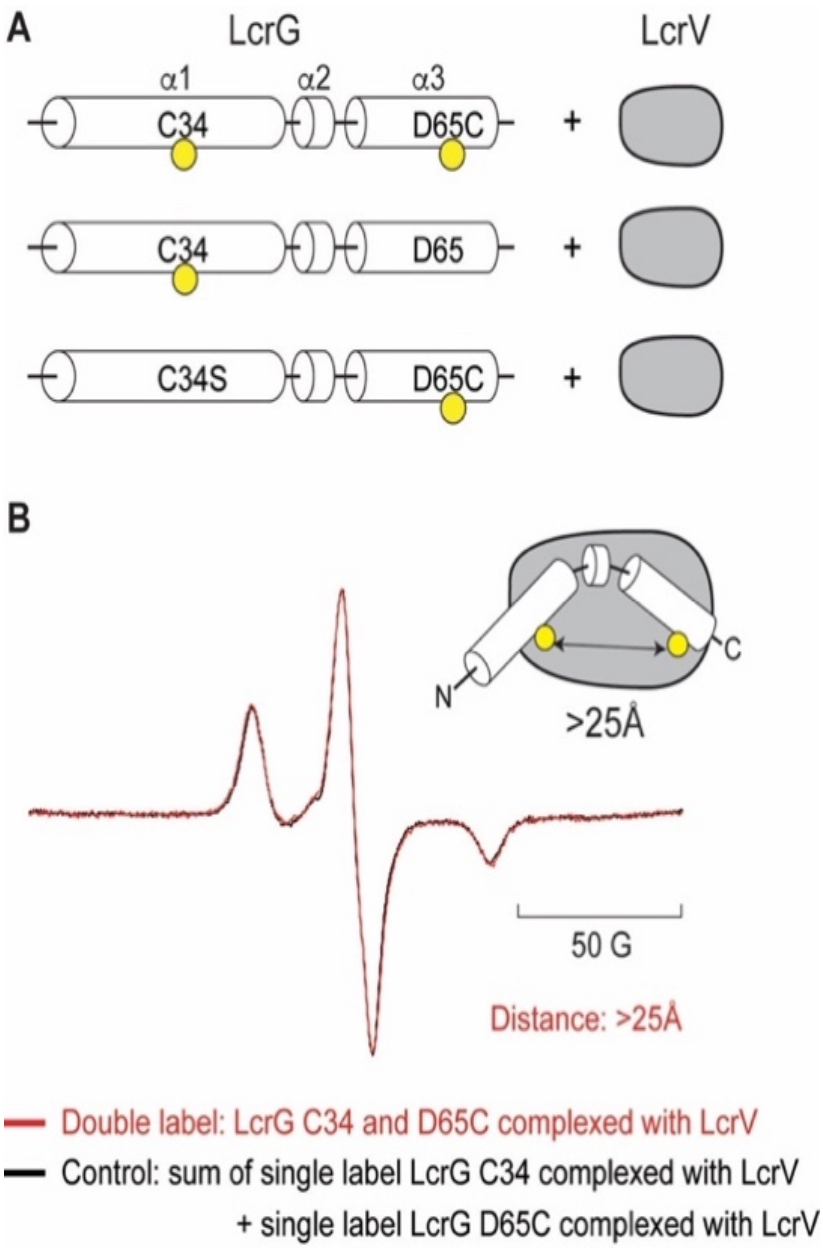

Figure 2-6. The open conformation of LcrG when bound to LcrV. **(A)** The LcrG spin-labeled constructs used in complex with LcrV. **(B)** EPR spectra of the spin-spin interaction of $LcrG^{7-73}$ C34 and D65C in complex with LcrV.

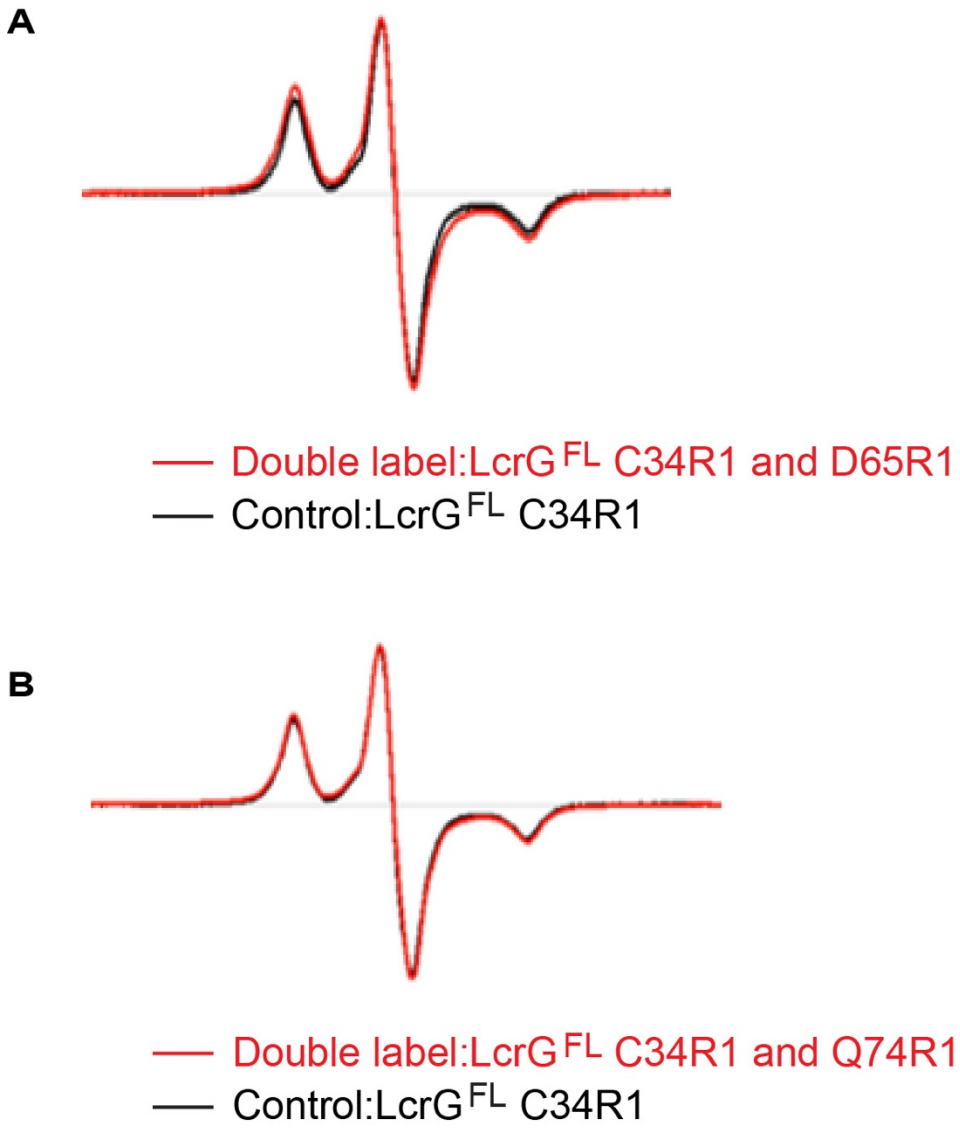

Figure 2-7. EPR did not detect any spin-spin interaction for LcrG^{FL} C34R1 and D65R1 and LcrG^{FL} C34R1 and Q74R1. R1 is the MTSL side with the paramagnetic center. The black control spectrum is generated by LcrG^{FL} spin labeled with MTSL side chain R1. The red spectrum is the sample generated by the two spin labels (MTSL) side chains attached to two helices of LcrG^{FL}. **(A)** EPR spectra of the spin-spin interaction of LcrG^{FL} C34R1 and D65R1. **(B)** EPR spectra of the spin-spin interaction of LcrG^{FL} C34R1 and Q74R1.

2.6 References

1. Cornelis, G.R., *The Yersinia Ysc-Yop 'type III' weaponry.* Nat Rev Mol Cell Biol, 2002. **3**(10): p. 742-52.
2. Cornelis, G.R., *Yersinia type III secretion: send in the effectors.* J. Cell Biol., 2002. **158**(3): p. 401-408.
3. Dey, S., et al., *The type III secretion system needle, tip, and translocon.* Protein Sci, 2019. **28**(9): p. 1582-1593.
4. Galan, J.E., et al., *Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells.* Annu. Rev. Microbiol., 2014. **68**: p. 415-38.
5. Mueller, C.A., et al., *The V-antigen of Yersinia forms a distinct structure at the tip of injectisome needles.* Science, 2005. **310**(5748): p. 674-6.
6. Broz, P., et al., *Function and molecular architecture of the Yersinia injectisome tip complex.* Mol. Microbiol., 2007. **65**(5): p. 1311-20.
7. Chaudhury, S., et al., *Structure of the Yersinia pestis tip protein LcrV refined to 1.65 Å resolution.* Acta Crystallogr. Sect. F, 2013. **69**(Pt 5): p. 477-81.
8. Derewenda, U., et al., *The structure of Yersinia pestis V-antigen, an essential virulence factor and mediator of immunity against plague.* Structure, 2004. **12**(2): p. 301-306.
9. Chatterjee, S., et al., *The crystal structure of the Salmonella type III secretion system tip protein SipD in complex with deoxycholate and chenodeoxycholate.* Protein Sci., 2011. **20**: p. 75-86.
10. Lunelli, M., et al., *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., 2011. **7**(8): p. e1002163.
11. Johnson, S., et al., *Self-chaperoning of the type III secretion system needle tip proteins IpaD and BipD.* J. Biol. Chem., 2007. **282**(6): p. 4035-4044.
12. Erskine, P.T., et al., *High Resolution Structure of BipD: An Invasion Protein Associated with the Type III Secretion System of Burkholderia pseudomallei.* J. Mol. Biol., 2006. **363**(1): p. 125-136.
13. Cornelis, G., J.C. Vanootegem, and C. Sluiters, *Transcription of the yop regulon from Y. enterocolitica requires trans acting pYV and chromosomal genes.* Microb Pathog, 1987. **2**(5): p. 367-79.
14. Espina, M., et al., *IpaD localizes to the tip of the type III secretion system needle of Shigella flexneri.* Infect. Immun., 2006. **74**(8): p. 4391-4400.
15. Olive, A.J., et al., *Bile salts stimulate recruitment of IpaB to the Shigella flexneri surface, where it colocalizes with IpaD at the tip of the type III secretion needle.* Infect. Immun., 2007. **75**(5): p. 2626-2629.
16. Lara-Tejero, M. and J.E. Galan, *Salmonella enterica serovar typhimurium pathogenicity island 1-encoded type III secretion system translocases mediate intimate attachment to nonphagocytic cells.* Infect. Immun., 2009. **77**(7): p. 2635-42.
17. Nilles, M.L., et al., *Yersinia pestis LcrV forms a stable complex with LcrG and may have a secretion-related regulatory role in the low-Ca^{2+} response.* J. Bacteriol., 1997. **179**(4): p. 1307-16.

18. Sarker, M.R., et al., *The Yersinia Yop virulon: LcrV Is required for extrusion of the translocators YopB and YopD.* J. Bacteriol., 1998. **180**(5): p. 1207-1214.
19. Sarker, M.R., et al., *LcrG is required for efficient translocation of Yersinia Yop effector proteins into eukaryotic cells.* Infect. Immun., 1998. **66**(6): p. 2976-9.
20. Nilles, M.L., K.A. Fields, and S.C. Straley, *The V antigen of Yersinia pestis regulates Yop vectorial targeting as well as Yop secretion through effects on YopB and LcrG.* J. Bacteriol., 1998. **180**(13): p. 3410-20.
21. Lee, V.T., C. Tam, and O. Schneewind, *LcrV, a substrate for Yersinia enterocolitica type III secretion, is required for toxin targeting into the cytosol of HeLa cells.* J. Biol. Chem., 2000. **275**(47): p. 36869-75.
22. DeBord, K.L., V.T. Lee, and O. Schneewind, *Roles of LcrG and LcrV during type III targeting of effector Yops by Yersinia enterocolitica.* J. Bacteriol., 2001. **183**(15): p. 4588-98.
23. Lawton, D.G., et al., *Interactions of the type III secretion pathway proteins LcrV and LcrG from Yersinia pestis are mediated by coiled-coil domains.* J. Biol. Chem., 2002. **277**(41): p. 38714-22.
24. Skrzypek, E. and S.C. Straley, *LcrG, a secreted protein involved in negative regulation of the low-calcium response in Yersinia pestis.* J. Bact., 1993. **175**(11): p. 3520-28.
25. Matson, J.S. and M.L. Nilles, *LcrG-LcrV interaction is required for control of Yops secretion in Yersinia pestis.* J. Bacteriol., 2001. **183**(17): p. 5082-91.
26. Chaudhury, S., et al., *The LcrG tip chaperone protein of the Yersinia pestis type III secretion system is partially folded.* J. Mol. Biol., 2015. **427**: p. 3096-3109.
27. Pallen, M.J., G. Dougan, and G. Frankel, *Coiled-coil domains in proteins secreted by type III secretion systems.* Mol. Microbiol., 1997. **25**(2): p. 423-5.
28. Blocker, A.J., et al., *What's the point of the type III secretion system needle?* Proc. Natl. Acad. Sci. U.S.A., 2008. **105**(18): p. 6507-13.
29. Fajer P., B., L.,Song,L., *Practical Pulsed Dipolar ESR (DEER). In ESR Spectroscopy in Membrane Biophysics*. Biological Magenetic Resonance. Vol. 27. 2007, Boston,MA: Springer.
30. Chaudhury, S., *Structural studies of chaperones and chaperone-tip interactions from the type III secretion systems of Yersinia and Pseudomonas*, in *Molecular Biosciences*. 2013.
31. Chaudhury, S., et al., *Nuclear Magnetic Resonance Characterization of the Type III Secretion System Tip Chaperone Protein PcrG of Pseudomonas aeruginosa.* Biochemistry, 2015. **54**(43): p. 6576-85.
32. McGuffin, L.J., K. Bryson, and D.T. Jones, *The PSIPRED protein structure prediction server.* Bioinformatics, 2000. **16**(4): p. 404-5.
33. Lee, P.C., et al., *Control of effector export by the Pseudomonas aeruginosa type III secretion proteins PcrG and PcrV.* Mol. Microbiol., 2010. **75**(4): p. 924-41.
34. Nanao, M., et al., *Type III secretion proteins PcrV and PcrG from Pseudomonas aeruginosa form a 1:1 complex through high affinity interactions.* BMC Microbiol., 2003. **3**: p. 21.
35. Allmond, L.R., et al., *Protein binding between PcrG-PcrV and PcrH-PopB/PopD encoded by the pcrGVH-popBD operon of the Pseudomonas aeruginosa type III secretion system.* Infect. Immun., 2003. **71**(4): p. 2230-3.

Chapter 3: Model of LcrG-LcrV Interaction

3.1 Abstract

Yersinia pestis is a Gram-negative bacterium which causes bubonic plague. *Yersinia* uses the Type III Secretion System (T3SS) to target its host cell with the virulence effector proteins. The T3SS consists of the needle complex, chaperone and effector proteins. The needle complex has the base, needle, tip complex and translocon. The *Yersinia* tip is made up of the tip proteins LcrV. During assembly of the needle complex LcrV interacts with its chaperone LcrG. The binding mode of LcrG and LcrV is unknown. The model of interaction of LcrG-LcrV is based on the crystal structures from of tip proteins of *Salmonella* SipD, *Shigella* IpaD and *Burkholderia* BipD, suggest that LcrG would have an alpha helical hairpin when it is bound to LcrV. Our previous EPR data revealed that LcrG has an open conformation when it is bound to LcrV. To identify the binding interaction of LcrV and LcrG, we made cysteine mutations on LcrG and LcrV for the attachment of the spin-label MTSL. Our EPR data showed that LcrG interacts with the N-terminal domain of LcrV. Our EPR results provide insight on the binding interaction of LcrG and LcrV.

3.2 Introduction

Bubonic plague is caused by the Gram-negative bacterium, *Yersinia pestis*. *Yersinia pestis* deploys the T3SS to form a nanoinjector to transport effector proteins from the bacteria to the host cells. The T3SS consists of the needle complex, chaperone, tip complex and translocon [1],[2]. The needle complex is a macromolecular assemblage made up of over a dozen of different proteins. The components of the needle complex are the base, a needle, a tip and a translocon[3],[4]. The base is attached to the inner and outer membranes of the bacterial cell membrane[3]. The needle protrudes from the base into the inter membranous region between the

bacteria and the host cell[3]. The needle is capped by the tip complex which is formed by multiple copies of the tip protein and provides a platform for the assembly of the translocon[5]. The translocon forms the pore in the host cell membrane through which the *Yersinia* Yop effector proteins are transported into the host cells[6].

The tip acts an environmental sensor for the assembly of the translocon and is made up of 5 copies of the tip protein LcrV[5],[7]. The crystal structure of LcrV shows similarity with the tip proteins following-*Salmonella* SipD, *Shigella* IpaD, and *Burkholderia* BipD[8],[9],[10],[11],[12],[13]. The T3SS genes[14] are activated to initiate the assembly of the base, needle and tip at 37^{o}C[5]. When the bacteria come in contact with the host cell and this contact is somehow sensed and detected by the tip complex and the bacteria then activates the assembly of the translocon proteins[15],[16].

In *Yersinia, the* tip protein interacts with a small protein LcrG in the bacterial cytosol. The LcrV interacts with LcrG to maintain its structural integrity prior to the assembly of the tip. The LcrG-LcrV interaction shows tight binding with a K_d of 140nM[17]. LcrG also regulates the secretion of Yop proteins. The structure of LcrG determined by NMR shows that it consists of three alpha helices with no tertiary contacts and shows fast nanoscale motions[18].

Currently, there is no crystal structure of the LcrG-LcrV complex. The crystal structure of the other tip proteins: *Salmonella* SipD, *Shigella* IpaD and *Burkholderia* BipD have a N-terminal alpha helical hairpin which acts as self-chaperone[19]. This alpha helical hairpin structural motif is not present in LcrV. Comparing the crystal structures of the tip proteins with LcrV[20], it was hypothesized that LcrG would form an helical alpha hairpin similar to the N-terminal region of the tip proteins SipD, IpaD, and BipD. Previous NMR data acquired shows the dynamic nature of the LcrG-LcrV interactions where different regions of LcrG bound to LcrV with different

affinities[21]. Our EPR data also showed when LcrG is bound to LcrV, helix α1 and helix α3 are far apart with a distance larger than 25Å, forming an open conformation[22],as shown in Fig 2-6 of Chapter 2.

Based on the NMR and EPR data, I proposed some possible models of LcrV-LcrG interaction (Fig 3-1). To test these models of LcrV-LcrG interaction s by EPR, I engineered cysteine mutations in the N-terminal (helix α1) and C-terminal (helix α3) of LcrG and in the. N-terminal region of LcrV for the attachment of the spin label. EPR data acquired by my collaborator Dr. Likai Song for the Florida State University revealed that both the N-terminal and the C-terminal helices interacts with the N-terminal domain LcrV. Our EPR data provides insights the binding interaction of LcrG and LcrV.

3.3 Methods

3.3.1 Protein expression and purification

Site-directed mutagenesis by QuickChange (Stratagene) was used to introduce point mutations in LcrV to obtain LcrV Q37C and LcrV K137C. $LcrG^{7-73}$ has a native cysteine residue in the 34^{th} position so $LcrG^{7-73}$ was used as a construct for site directed spin labeling (SDSL). Point mutations were also introduced in the N-terminal helix of $LcrG^{7-73}$ C34S: $LcrG^{7-73}$ C34S/K28C and $LcrG^{7-73}$ C34S/A35C and C -terminal helix: $LcrG^{7-73}$ C34S/E60C and $LcrG^{7-73}$ C34S/D65C.

The proteins were expressed and purified as mentioned in the previous Chapter 2.3.1.

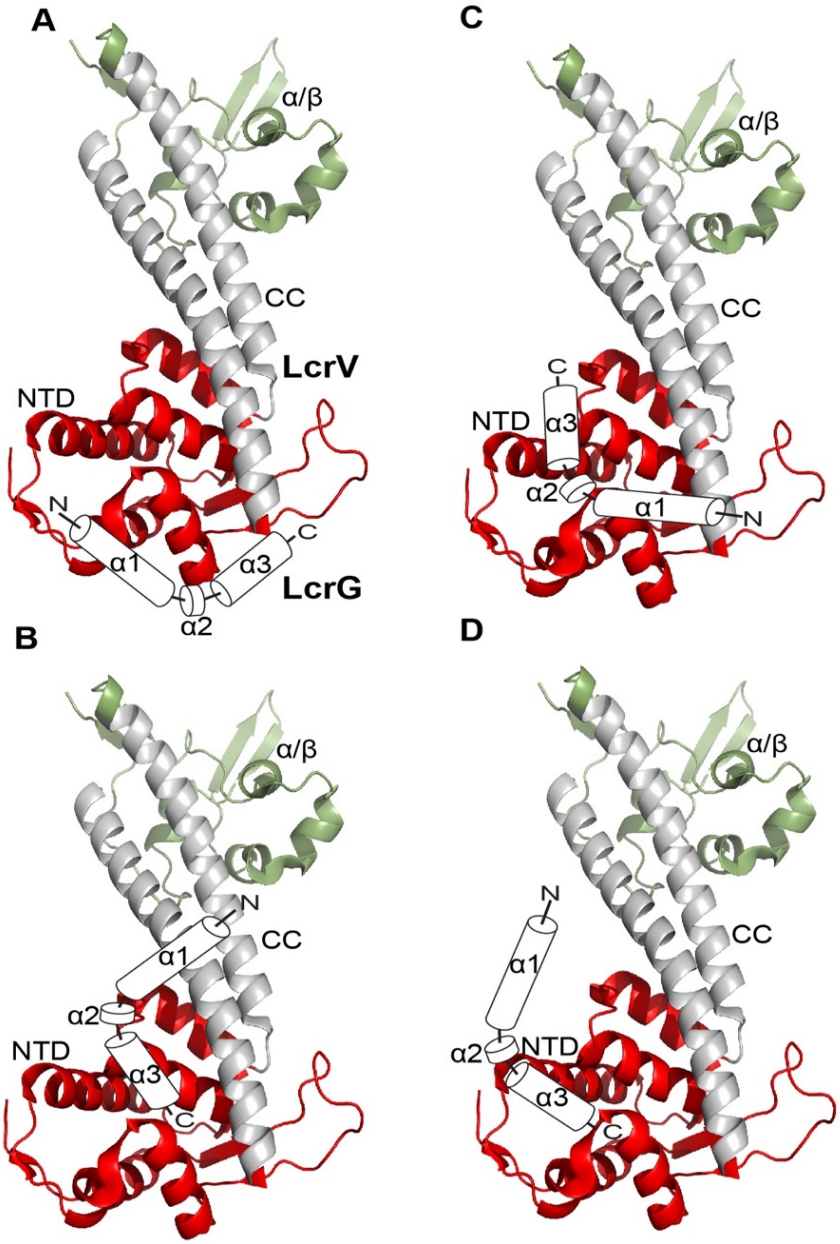

Figure 3-1. Proposed models of LcrV and LcrG interactions based on NMR and EPR data. **(A)** A proposed model showing interaction of N-terminal helix (helix α1) and C-terminal helix (helix α3) of LcrG with the N-terminal domain (NTD, shown in red) of LcrV. **(B)** A proposed model of LcrG-LcrV interaction where the C-terminal helix (helix α3) of LcrG interacts with the helical regions of NTD of LcrV and the N-terminal helix (helix1) of LcrG interacts with the β turn region of NTD and the coiled-coil (CC, shown in grey) of LcrV. **(C)** A proposed model of LcrG-LcrV interaction showing that the N-terminal helix of LcrG interacts with both the helices in the coiled-coil (CC) domain of LcrV; whereas LcrG helix α2 interacts with the unstructured regions of the LcrV NTD; and the C-terminal helix (helix α3) with the helical region of NTD of LcrV. **(D)** A proposed model of LcrV-LcrG interaction wherein the helix α2 and the C- terminal helix of LcrG interact with the helical region of the NTD of LcrV.

.

Table 3-1. Site-directed spin labeling of MTSL side chain R1 in LcrG and LcrV

Position of spin label	Number of spin label
$LcrG^{7-73}$	
C34S/K28R1	Single
C34R1	Single
C34S/A35R1	Single
C34S/E60R1	Single
C34S/D65R1	Single
LcrV	
Q37R1	Single
K137R1	Single

3.3.2 Site directed spin labeling of LcrG$^{7-73}$ and LcrV constructs

The LcrG and LcrV constructs was spin labeled with MTSL as described in the previous chapter. The spin labeled constructs used in this study are listed in Table 3-1.

3.3.3 Circular dichroism spectroscopy

CD data for the cysteine mutants of LcrG and LcrV along with spin labeled LcrG and LcrV were acquired in JASCO J-815 spectropolarimeter following similar protocols described in Chapter 2.3.3.

3.3.4 EPR spectroscopy

The EPR data for the spin labeled LcrG and LcrV constructs were acquired under identical conditions as mentioned in Chapter 2. The spin–spin distances was calculated using a Monte Carlo/simplex Gaussian convolution method[23].

3.4 Results

3.4.1 Protein expression and purification

Site directed mutagenesis generated four cysteine mutants in the LcrG $^{7-73}$ C34S construct. The constructs for the N-terminal helix were LcrG $^{7-73}$ C34S/K28C and LcrG $^{7-73}$ C34S/A35C. The C-terminal constructs were LcrG $^{7-73}$ C34S/E60C and LcrG $^{7-73}$ C34S/D65C. The cysteine mutant proteins showed similar expression and yield LcrG $^{7-73}$ C34S and LcrG $^{7-73}$ constructs as described previously[18]. Cysteine mutants of LcrV generated by site directed mutagenesis are LcrV Q37C and LcrV K137C had similar stability and yield as LcrV[8].

3.4.2 *ESI mass spectrometry confirms site directed spin labeling of LcrG$^{7-73}$ and LcrV constructs*

Electrospray ionization (ESI) mass spectrometry (Fig 3-2) confirmed the extent of spin labeling in the N-terminal constructs of LcrG, which dimerized in the absence of a reducing agent. The theoretical mass of LcrG$^{7-73}$ C34S/K28C was estimated to be 7967.6 Da. In Fig 3-2A, the mass of the sample was estimated to be 15935.20 Da, which is twice the predicted mass of LcrG$^{7-73}$ C34S/K28C. After addition of DTT and the MTSL spin label, the majority of the protein sample had a mass of 8152.7 Da as shown in Fig 3-2B, which confirmed the attachment of the R1 side chain with a mass of 185 Da to LcrG$^{7-73}$ C34S/K28C. Fig 3-2 C and Fig 3-2 E, showed that before MTSL labeling the samples LcrG$^{7-73}$ and LcrG$^{7-73}$ C34S/A35C formed dimers like LcrG$^{7-73}$ C34S/K28C and had an estimated mass of 16018 Da and 16049.40 Da respectively. After the addition of the MTSL spin label, the majority of the protein samples LcrG$^{7-73}$ and LcrG$^{7-73}$ C34S/A35C had a mass 8194 Da and 8209 Da respectively, denoting that the MTSL R1 side chain had been attached to the protein sample.

Fig 3-3 showed that, like the N-terminal constructs of LcrG, the C-terminal constructs also tend to dimerize (Fig 3-3A,3-3C). After MTSL spin-label attachment the mass of LcrG$^{7-73}$C34S/E60C is 8151.8 Da, which was the theoretical Mw of LcrG$^{7-73}$C34S/E60C (7966.7 Da) plus the mass of attached R1 side chain (185.1 Da) (Fig 3-3B). Fig 3-3D showed the mass of LcrG$^{7-73}$C34S/D65C as 8166 Da after the addition of MTSL showing, an increase of 185 Da (mass of R1 side chain) to its theoretical mass (7981 Da).

The extent of LcrV spin labeling as determined by ESI mass spectrometry is shown in Fig 3-4. The mass of LcrV Q37C is estimated to be 33885.6 Da, and the mass of LcrV K137C was 33885.6 Da (Fig 3-4A and Fig 3-4C). MTSL was added to these protein samples to attach the R1 side chain spin label. The mass of LcrV Q37C and LcrV K137C was estimated to be 34071.6

(Fig 3-4B) and 34071.0 (Fig 3-4D), respectively. The difference in the mass of LcrV Q37C before and after labeling is 186 Da, which denoted the addition of the MTSL spin label. Likewise, after labeling, the difference in mass of LcrV K137C, is also estimated to be 186 Da, showing the R1 attachment of the MTSL spin label.

3.4.3 *Circular dichroism spectroscopy*

CD spectroscopy was used to determine the effect of cysteine mutations in LcrG. CD spectra of the N-terminal and C-terminal constructs of LcrG showed no overall change in the secondary structures compared to $LcrG^{7-73}$. The Molar Ellipticity ratio at 222nm and 208nm ($\theta_{222}/\ \theta_{208}$) of all the constructs with cysteine mutations were estimated to be 0.8 (Fig 3-5A-E), which suggested that all the constructs had non-interacting helices similar to the wild type $LcrG^{7-73}$ construct, showing that the cysteine mutations did not affect the secondary structure. The addition of the MTSL by site directed spin labeling (SDSL) led to the attachment of the spin-label (R1).

CD spectroscopy was also used to determine the effect of R1 on secondary structure of LcrG. The labeled LcrG constructs had $\theta_{222}/\ \theta_{208}$ ratios from 0.78-0.88 (Fig 3-5A-E), showing the addition of the spin-label did not change the overall secondary structures of the LcrG constructs. There was no overall change in the secondary structure of LcrV Q37C and LcrV K137C upon the addition of the R1 spin label. The $\theta_{222}/\ \theta_{208}$ ratio was the same after the addition of the R1 side chain, as shown in Fig 3-6A-B.

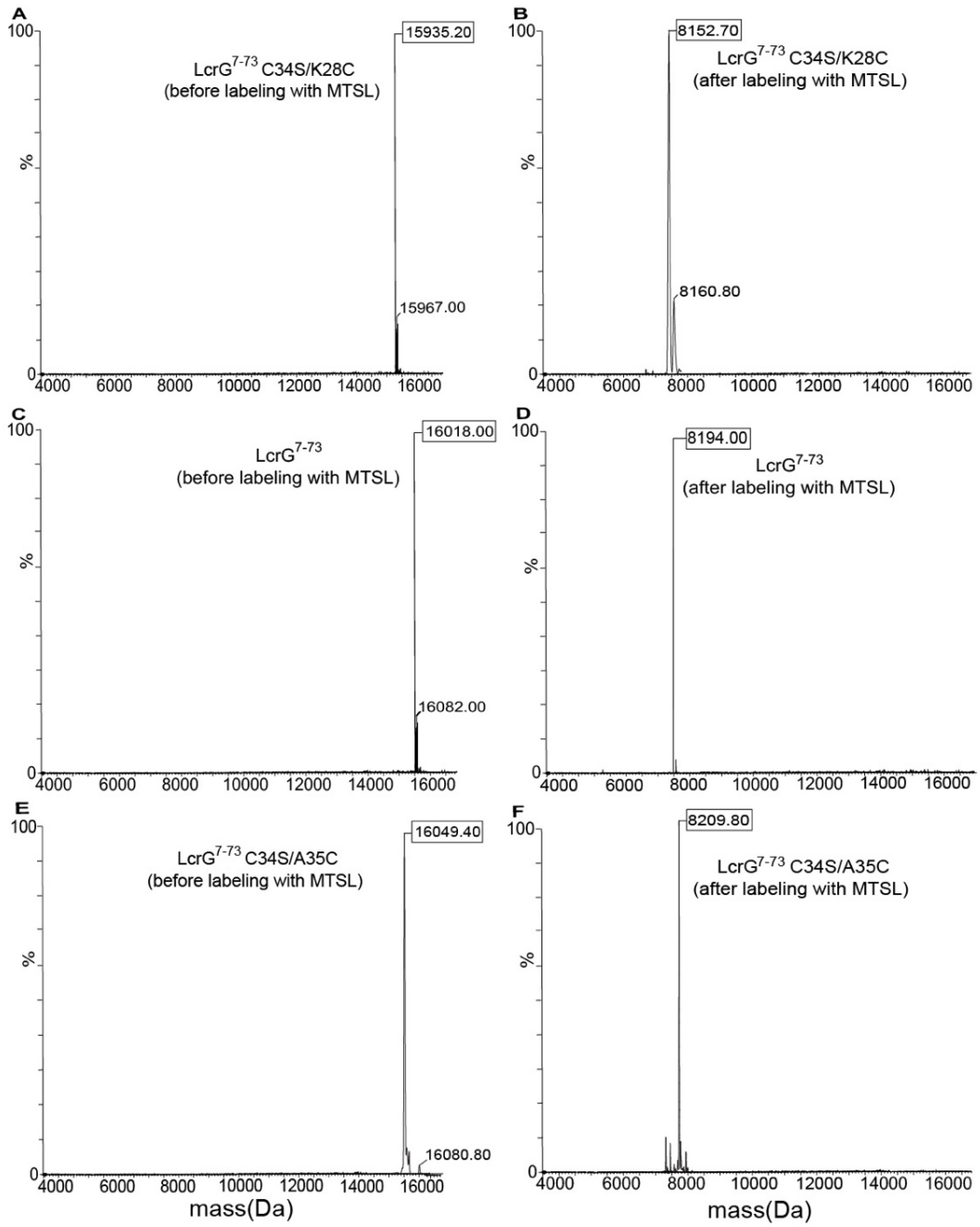

Figure 3-2. Extent of spin labeling of LcrG N-terminal helix by electrospray ionization mass spectrometry (ESI-MS). ESI-MS of LcrG$^{7-73}$C34S/DK28C, LcrG$^{7-73}$, LcrG$^{7-73}$C34S/A35C **(A)** before MTSL labeling (theoretical MW 7967.6 Da) (**B)** after MTSL labeling (theoretical MW 8152.7Da) **(C)** before MTSL labeling (theoretical MW 8009 Da) **(D)** after MTSL labeling (theoretical MW 8194 Da) (**E)** before MTSL labeling (theoretical MW 8024.6 Da) **(F)** after MTSL labeling (theoretical MW 8209.2 Da)

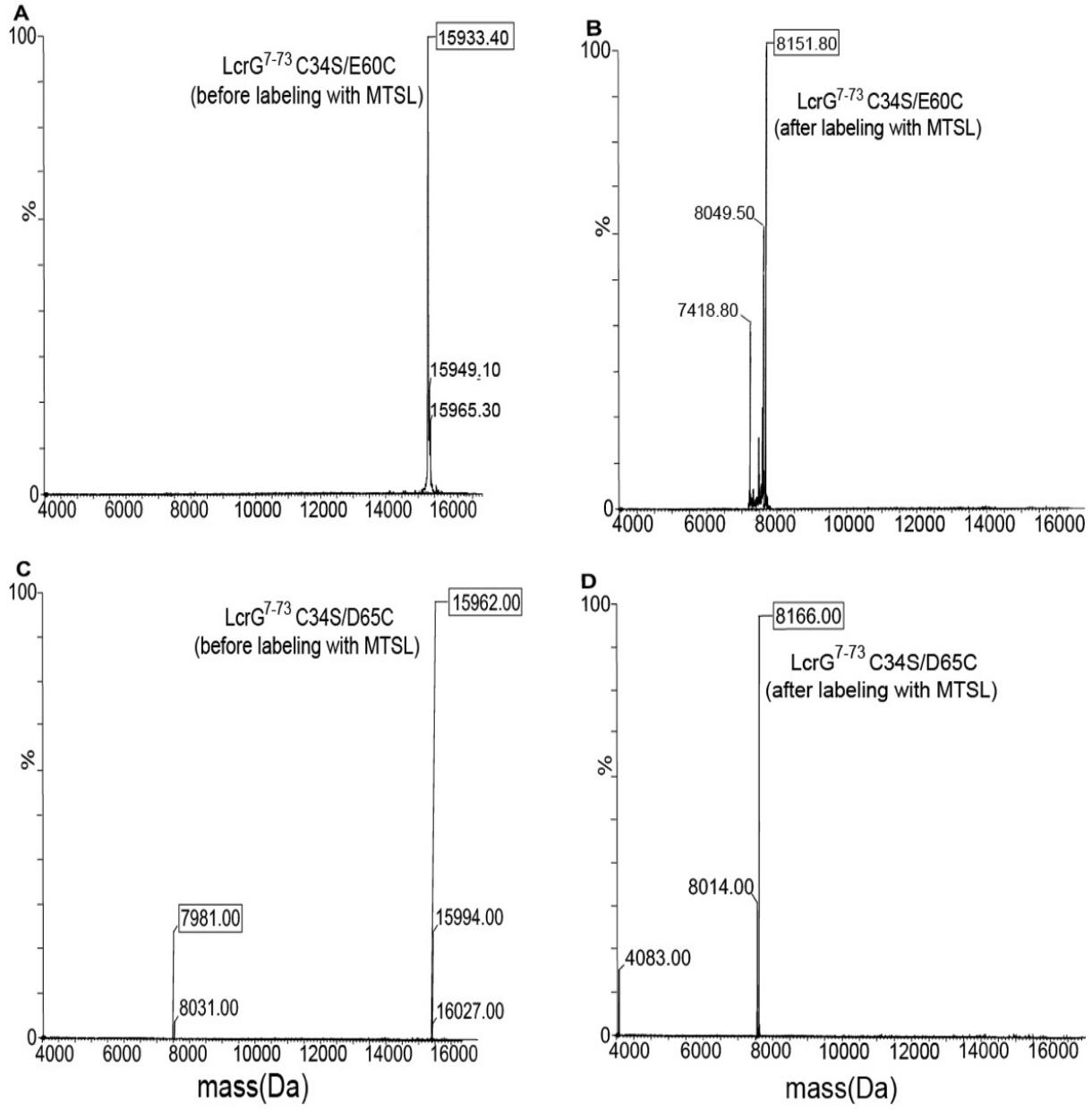

Figure 3-3. LcrG C-terminal helix spin labeling by electrospray ionization mass spectrometry (ESI-MS). **(A)** ESI-MS of LcrG$^{7-73}$C34S/E60C before MTSL labeling (theoretical MW 7966.7 Da) **(B)** ESI-MS of LcrG$^{7-73}$C34S/E60C after MTSL labeling (theoretical MW 8151.7 Da) **(C)** ESI-MS of LcrG$^{7-73}$C34S/D65C before MTSL labeling (theoretical MW 7981 Da) **(D)** ESI-MS of LcrG$^{7-73}$C34S/D65C after MTSL labeling (theoretical MW 8166 Da)

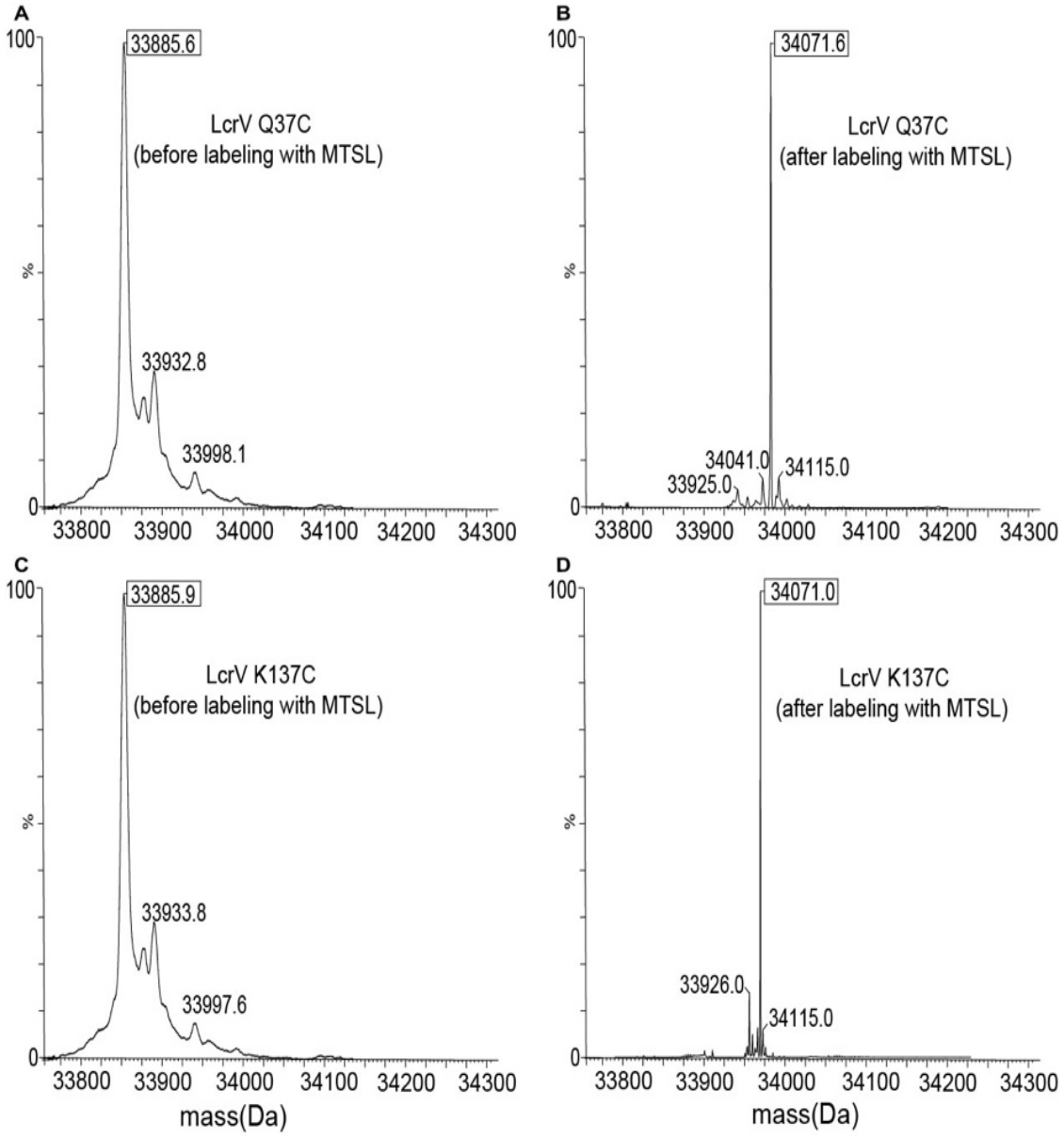

Figure 3-4. Extent of LcrV spin labeling as confirmed ESI-MS. ESI-MS of LcrV Q37C and LcrV K137C.**(A)** before MTSL labeling (theoretical MW 33885.6 Da) **(B)** after MTSL labeling (theoretical MW 34070.6 Da) **(C)** before MTSL labeling (theoretical MW 33885.9 Da) **(D)** ESI-MS of LcrV K137C after MTSL labeling (theoretical value 34077.9 Da)

3.4.4 EPR spectroscopy

EPR data showed that LcrG had an open conformation when bound to the tip protein LcrV, as shown in Fig 2-5 of Chapter 2. To determine how LcrG interacted with LcrV, we placed a spin-label on LcrG and another spin-label on LcrV.The LcrG residues K28 and A35 in theN-terminal helix (helix α1) and the residues E60 and D65 in the C-terminal helix (helix α3) (Fig 3-7A) were mutated to cysteine to generate four different constructs for spin labeling. The$LcrG^{7-73}$ construct was also used for spin labeling using the native cysteine in the 34^{th} position. The residues in the helix α3 (Q37) and helix α6 (K137) of the N-terminal domain of LcrV (Fig 3-8A) were mutated to cysteine to generate two different constructs for site directed spin labeling. The EPR spectra from the 7 spin-labeled LcrV and LcrG contracts in their uncomplexed form served as the control. For example, the spectra from LcrG $^{7-73}$ C34S/K28R1by itself and LcrV Q37R1 by itself served as control for the LcrG-LcrV complex formed by LcrG $^{7-73}$ C34S/K28R1 and LcrV Q37R1(Fig 3-8A).

The locations of the spin labels on LcrV were chosen based on previous results obtained by NMR, which showed that the many residues in the N-terminal domain of LcrV were perturbed upon titration of LcrV with LcrG[21] .A criterion for spin labeling LcrV residues was to select residues which were not directly involved in interaction with LcrG, so that attachment of the spin-label would not alter the protein-protein interaction .Another criterion was to choose residues that were surface exposed and non-conserved so the overall folding of LcrV was not affected. Based on the models proposed (Fig 3-1), the LcrV residue were separated by some distance to allow us to estimate the distance of how LcrG interacted with LcrV (the Cα – Cα is estimated to be 22 Å).

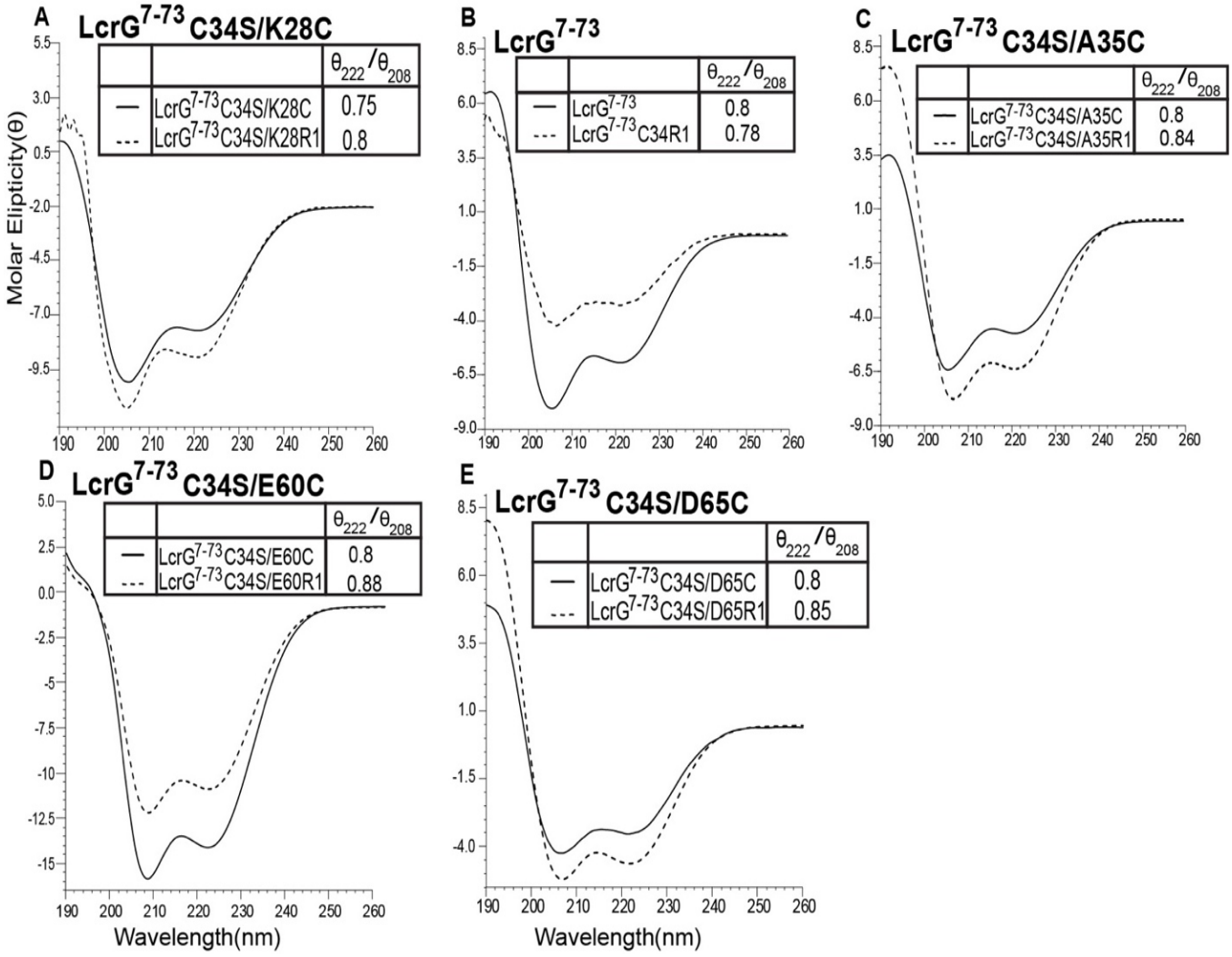

Figure 3-5. CD Spectroscopy of LcrG constructs used for LcrV-LcrG EPR studies. (A-E). CD spectra of $LcrG^{7-73}$ cysteine mutations before (solid line) and after the attachment of R1 side chain spin label by MTSL (dashed lines).

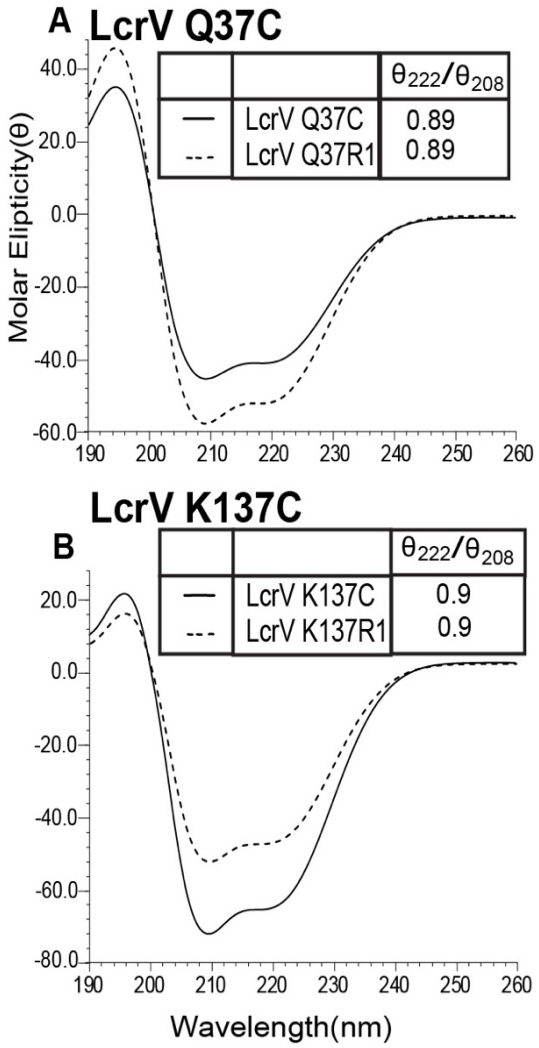

Figure 3-6. CD Spectroscopy of LcrV constructs used for LcrV-LcrG EPR studies. (A) CD spectra of LcrV Q37C before (solid lines) and after (LcrV Q37R1) MTSL side chain attachment (dashed lines)**. (B)** CD spectra of LcrVK137C before (solid lines) and after (LcrV K137R1) MTSL side chain attachment (dashed lines).

Ten different combinations of spin-spin interactions were tested for the LcrG-LcrV complex, of which only two combinations yielded a spin-spin distance below 25 Å-- LcrG $^{7-73}$ C34S/K28R1 to LcrV Q37R1 with a spin-spin distance of 12 Å (Fig 3-8A) and LcrG $^{7-73}$ C34S/D65R1 to LcrV K137R1 with a spin-spin distance of 18 Å (Fig 3-8 B).

In Fig 3-8A, the red spectrum generated by the LcrVQ37R1 -LcrG$^{7-73}$ C34S/K28R1 complex, showed spectral broadening and reduction in spectral height compared to the black control spectrum, which was generated by the addition of the LcrVQ37R1 spectrum and the LcrG$^{7-73}$ C34S/K28R1 spectrum. Similar spectral broadening was also observed for the LcrVK137R1 - LcrG$^{7-73}$ C34S/D65R1 complex (Fig 3-8B). The rest of the eight combinations tested did not show any significant spectral broadening, indicating the spin-spin distance to be more than 25 Å (Fig 3-8C).

Our EPR data suggested that the entire length of LcrG consisting of N-terminal helix (helix α1) and C-terminal helix (helix α3) is in close proximity to the LcrV N-terminal domain as shown in Fig 3-9C. Part of LcrG helix α1 was at a distance of 12 Å from helix α3 of the LcrV N-terminal domain (Fig 3-9B) and the LcrG helix α3 was at a distance of 18 Å from helix α6 of the LcrV N-terminal domain (Fig 3-9B).

3.5 Discussion

Our EPR results shown here provides new information regarding the binding model of *Yersinia* tip protein LcrV and tip chaperone protein LcrG. LcrV and LcrG are essential for the virulence of *Yersinia pestis.* LcrV forms the tip by capping the needle complex and providing a platform for the formation of the translocon[5],through which the effector proteins are transported into the host cell[5]. LcrG binds to LcrV during needle complex assembly at the

bacterial cytosol with nanomolar affinity[17]. The *Pseudomonas aeruginosa* tip protein PcrV, which is a homolog of LcrV also forms a tight binding complex with the *Pseudomonas* tip chaperone PcrG[24],[25],[26].

Currently, there are no high-resolution structures of the LcrV-LcrG complex or the PcrG-PcrV complex. The current structural model of the LcrV-LcrG interaction is based on the comparison of the crystal structures of the tip proteins *Salmonella* SipD, *Shigella* IpaD and *Burkholdaria* BipD to *Yersinia* LcrV[8],[11],[12],[15].Based on this hypothesis, LcrG is expected to interact with LcrV in a manner similar to the N-terminal alpha hairpin of IpaD[20].

Previous NMR data acquired by others [21] indicated that isoleucine, leucine and valine residues of LcrV showed chemical perturbation in the N-terminal domain, α/β domain and the coiled-coil domain (Fig 3-9A) suggesting that either these residues are involved in direct binding or they are undergoing conformational change upon binding to LcrG. The NMR data also show that LcrG binds to LcrV with different affinities based on the behavior of the NMR peaks, where the subset of the ILV peaks formed slow-exchange peaks and another subset formed the intermediate exchange peak[21]. These results suggested that LcrG-LcrV complex is highly dynamic. This NMR data serves as our guide to position EPR spin labels for our current studies of the MTSL labeled LcrV- MTSL labeled LcrG complex.

Previous EPR data I acquired showed that when LcrG is bound to LcrV, the N-terminal helix (helix α1) and C-terminal helix (helix α3) have a spin-spin distance more than 25Å, forming an open conformation, showing that LcrG does not form a coiled-coil (Fig 2-5, Chapter 2). The EPR data acquired for the MTSL labeled LcrG constructs and MTSL labeled LcrV complexes as shown in Fig 3-8A-B does not support the idea of the formation of a coiled-coil or alpha helical hairpin when LcrV is bound to LcrG. If LcrG forms a coiled-coil or an alpha helical

hairpin, we would expect spin labels on helix α1 and helix α3 to show short spin-spin distances to a spin-labeled LcrV, for example, on LcrV K137R1. However, my EPR data shows that the spin-label on LcrG helix α1 shows short spin-spin distances to LcrV Q37R1 but not to LcrV K137R1(Fig 3-7 B-C), and likewise a spin-label on LcrG helix α3 shows short spin-spin distances to LcrV K137R1 but not to LcrV Q37R1(Fig 3-8 B-C). The spin-spin distances estimated by EPR shows that the N-terminal helix (helix α1) and the C-terminal helix (helix α3) of LcrG interact with the LcrV N-terminal domain (NTD), as shown in a cartoon model in Fig 3-9C.

To summarize my findings, EPR provides new experimental evidence showing that LcrG does not interact with LcrV as coiled-coil or hairpin as hypothesized based on the crystal structure of tip proteins *Salmonella* SipD, *Shigella* IpaD and *Burkholderia* BipD[11],[12],[15].

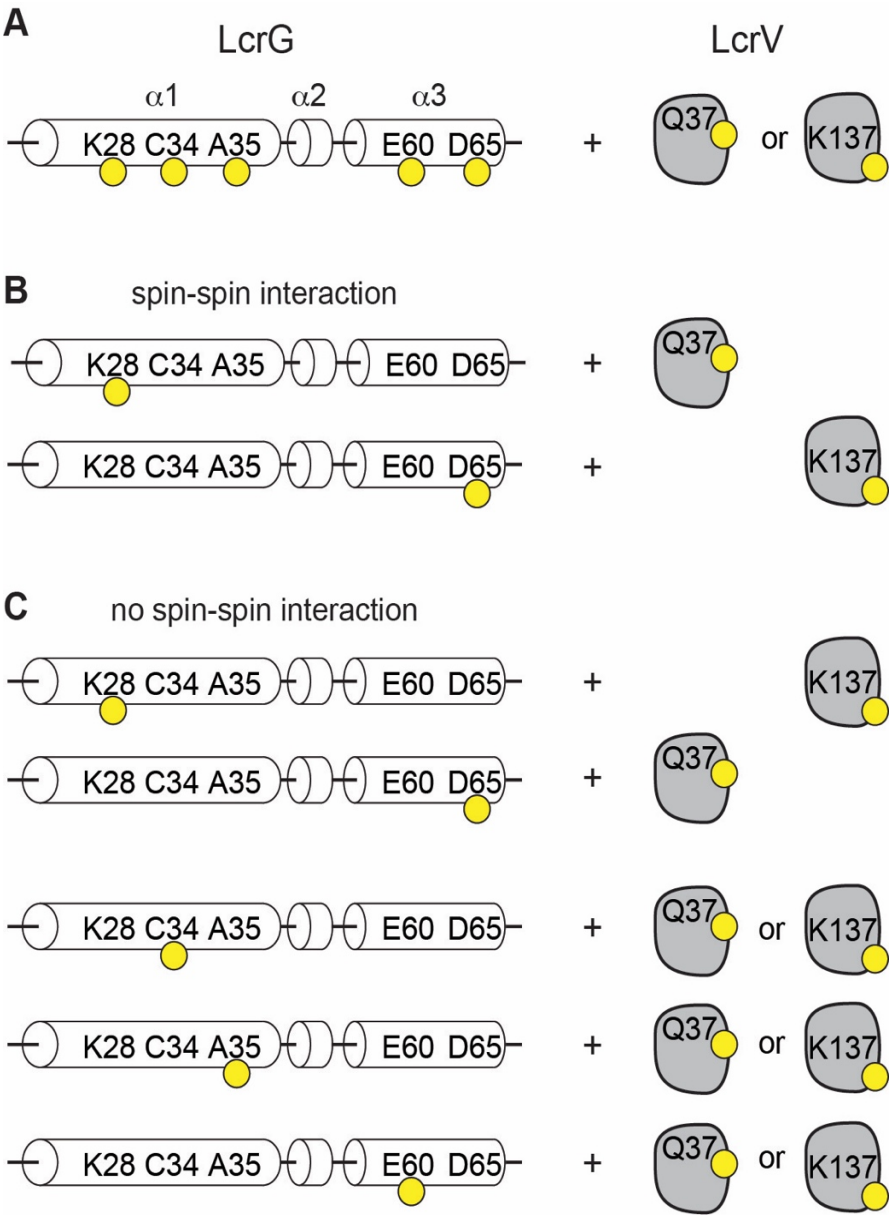

Figure 3-7. **Labeling scheme used in EPR determination of spin-spin interaction of the LcrG-LcrV complex.** (A) Summary of the residues in LcrG and LcrV that were spin-labeled. (B) Summary of results of EPR that detected spin-spin interaction between $LcrG^{7-73}$ C34S/K28R1 - LcrV Q37R1 complex, and $LcrG^{7-73}$ C34S/D65R1 – LcrV K137R1 complex.
(C) Summary of labeling scheme for the LcrG-LcrV complex were no spin-spin interactions were observed.

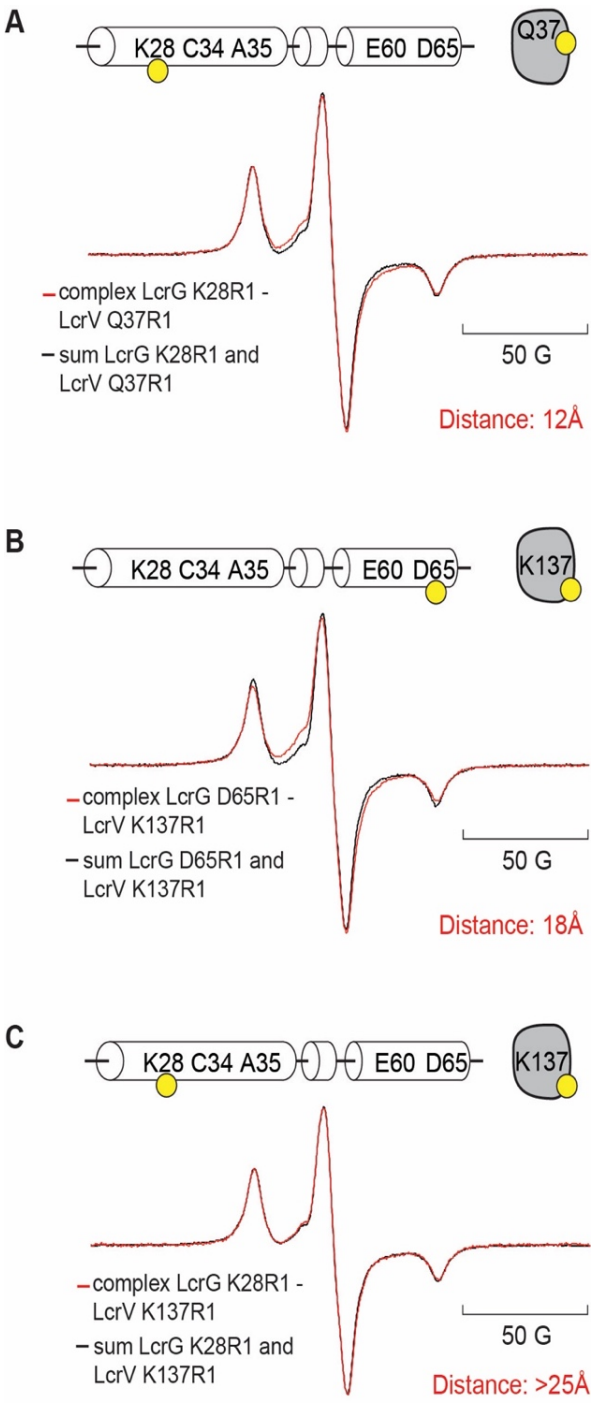

Figure 3-8. **EPR spectra of spin-labeled LcrG in complex with spin-labeled LcrV**. **(A)** and **(B)** EPR spectra of LcrG-LcrV complex that showed spin-spin interactions. **(C)** EPR spectra of LcrG-LcrV complex that did not show spin-spin interaction.

\

Table 3-2. Results of LcrG-LcrV spin-spin distances determined by EPR spectroscopy

$LcrG^{7-73}$ R1	LcrV R1	Distance
C34S/K28	Q37	12 Å
C34S/D65	K137	18 Å
C34S/K28	K137	>25 Å
C34	Q37	>25 Å
C34	K137	>25 Å
C34S/A35	Q37	>25 Å
C34S/A35	K137	>25 Å
C34S/E60	Q37	>25 Å
C34S/E60	K137	>25 Å
C34S/D65	Q37	>25 Å

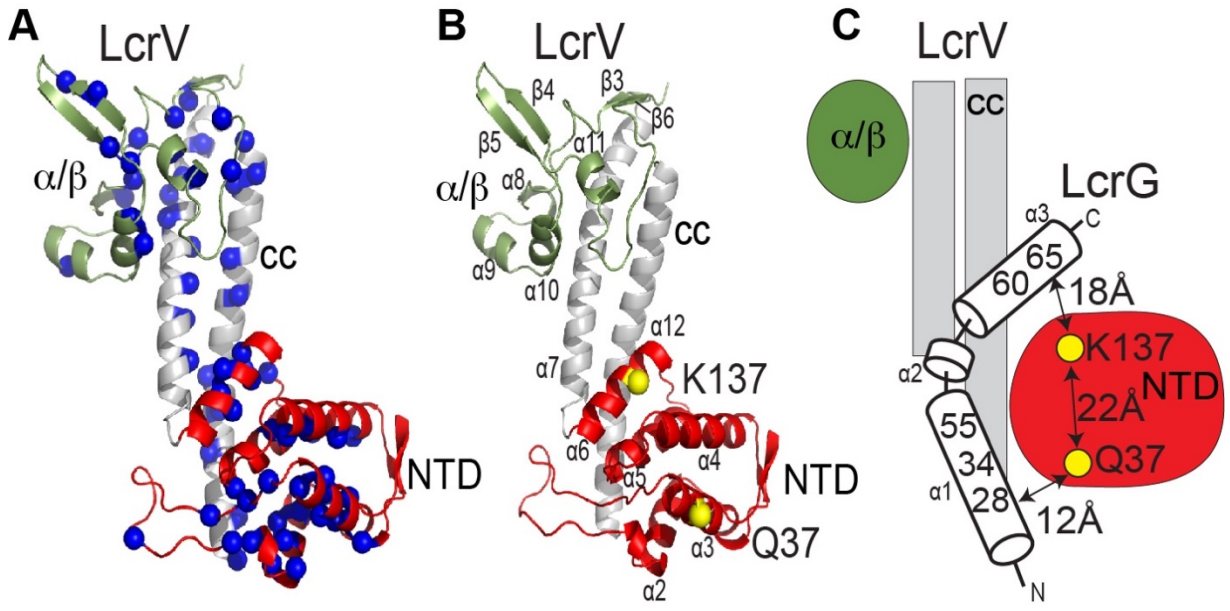

Fig. 3-9. **Model of LcrG-LcrV interaction from NMR and EPR**. **(A)**NMR data showing the affected ILV residues (blue spheres). **(B)** LcrV showing the MTSL labels on helix α3(Q37) and helix α6 (K137) of the NTD (red). **(C)** Cartoon of LcrG-LcrV interaction showing that LcrG helix α1 and helix α3 interact with NTD (red). The distance between the residues Q37 and K137 is 22 Å. LcrV has an N-terminal domain(NTD)(red),α/β domain(green) and a coiled-coil domain(CC) (grey).Figure adapted from [18],[21].

3.6 References

1. Cornelis, G.R., *The Yersinia Ysc-Yop 'type III' weaponry.* Nat. Rev. Mol. Cell Biol., 2002. **3**(10): p. 742-52.
2. Cornelis, G.R., *Yersinia type III secretion: send in the effectors.* J. Cell Biol., 2002. **158**(3): p. 401-408.
3. Dey, S., et al., *The type III secretion system needle, tip, and translocon.* Protein Sci, 2019. **28**(9): p. 1582-1593.
4. Galan, J.E., et al., *Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells.* Annu. Rev. Microbiol., 2014. **68**: p. 415-38.
5. Mueller, C.A., et al., *The V-antigen of Yersinia forms a distinct structure at the tip of injectisome needles.* Science, 2005. **310**(5748): p. 674-6.
6. Blocker, A., et al., *The tripartite type III secreton of Shigella flexneri inserts IpaB and IpaC into host membranes.* J.Cell Biol., 1999. **147**(3): p. 683-693.
7. Broz, P., et al., *Function and molecular architecture of the Yersinia injectisome tip complex.* Mol. Microbiol., 2007. **65**(5): p. 1311-20.
8. Chaudhury, S., et al., *Structure of the Yersinia pestis tip protein LcrV refined to 1.65 Å resolution.* Acta Crystallogr. Sect. F, 2013. **69**(Pt 5): p. 477-81.
9. Derewenda, U., et al., *The structure of Yersinia pestis V-antigen, an essential virulence factor and mediator of immunity against plague.* Structure, 2004. **12**(2): p. 301-306.
10. Chatterjee, S., et al., *The crystal structure of the Salmonella type III secretion system tip protein SipD in complex with deoxycholate and chenodeoxycholate.* Protein Sci., 2011. **20**: p. 75-86.
11. Lunelli, M., et al., *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., 2011. **7**(8): p. e1002163.
12. Johnson, S., et al., *Self-chaperoning of the type III secretion system needle tip proteins IpaD and BipD.* J. Biol. Chem., 2007. **282**(6): p. 4035-4044.
13. Erskine, P.T., et al., *High Resolution Structure of BipD: An Invasion Protein Associated with the Type III Secretion System of Burkholderia pseudomallei.* J. Mol. Biol., 2006. **363**(1): p. 125-136.
14. Cornelis, G., J.C. Vanootegem, and C. Sluiters, *Transcription of the yop regulon from Y. enterocolitica requires trans acting pYV and chromosomal genes.* Microb Pathog, 1987. **2**(5): p. 367-79.
15. Espina, M., et al., *IpaD localizes to the tip of the type III secretion system needle of Shigella flexneri.* Infect. Immun., 2006. **74**(8): p. 4391-4400.
16. Olive, A.J., et al., *Bile salts stimulate recruitment of IpaB to the Shigella flexneri surface, where it colocalizes with IpaD at the tip of the type III secretion needle.* Infect. Immun., 2007. **75**(5): p. 2626-2629.
17. Lawton, D.G., et al., *Interactions of the type III secretion pathway proteins LcrV and LcrG from Yersinia pestis are mediated by coiled-coil domains.* J. Biol. Chem., 2002. **277**(41): p. 38714-22.
18. Chaudhury, S., et al., *The LcrG tip chaperone protein of the Yersinia pestis type III secretion system is partially folded.* J. Mol. Biol., 2015. **427**: p. 3096-3109.

19. Chatterjee, S., et al., *The crystal structures of the Salmonella type III secretion system tip protein SipD in complex with deoxycholate and chenodeoxycholate.* Protein Sci, 2011. **20**(1): p. 75-86.
20. Blocker, A.J., et al., *What's the point of the type III secretion system needle?* Proc. Natl. Acad. Sci. U.S.A., 2008. **105**(18): p. 6507-13.
21. Kaur, K., *NMR Studies of Molecular Interaction Involved in Type III Secretion System,Sumoylation,and The RNA Binding Protein HuR,PhD. diss. University of Kansas.* 2016.
22. Pallavi Guha Biswas, Z.H., Pavanjeet Kaur, Kawaljit Kaur, Sukanya Chaudhury and M.W. Muhammed Ciftci, Justin Douglas, Likai Song, and Roberto N. De Guzman, *Model of Yersinia LcrG and LcrV Interaction Defined by EPR Spectroscopy.*
23. Fajer, P.G., L. Brown, and L. Song, *Practical Pulsed Dipolar ESR (DEER)*, in *ESR Spectroscopy in Membrane Biophysics. Biological Magnetic Resonance*. 2007, Springer, Boston, MA. p. 95-128.
24. Lee, P.C., et al., *Control of effector export by the Pseudomonas aeruginosa type III secretion proteins PcrG and PcrV.* Mol. Microbiol., 2010. **75**(4): p. 924-41.
25. Nanao, M., et al., *Type III secretion proteins PcrV and PcrG from Pseudomonas aeruginosa form a 1:1 complex through high affinity interactions.* BMC Microbiol., 2003. **3**: p. 21.
26. Allmond, L.R., et al., *Protein binding between PcrG-PcrV and PcrH-PopB/PopD encoded by the pcrGVH-popBD operon of the Pseudomonas aeruginosa type III secretion system.* Infect. Immun., 2003. **71**(4): p. 2230-3.

Chapter 4: EPR Shows that PcrG has Different Conformations

4.1 Abstract

Pseudomonas aeruginosa is the major cause of mortality among cystic fibrosis patients. *Pseudomonas* assemble a protein nanoinjector of the Type III Secretion System (T3SS) to inject virulence effector proteins into their target host cells. The T3SS nanoinjector consists of the needle complex, chaperones, and effector proteins. *Pseudomonas* PcrG is a protein that function as a chaperone to the tip protein, PcrV. Previous results by NMR and CD spectroscopy showed that PcrG lacks a tertiary structure and it is partially folded alpha helical proteins. Other results in the literature suggested that the two alpha helices of PcrG may interact with each other. Here, I and my collaborators used Electron Paramagnetic Resonance (EPR) spectroscopy to determine the conformations of PcrG. For EPR studies, we engineered cysteine mutations in PcrG for attachment of spin-labels. EPR results suggest that when PcrG is in free form, there is a population that showed transient tertiary structures where the two helices are in close contact with each other. When PcrG is bound to the tip protein, PcrV the two helices do not interact with each other anymore suggesting that PcrG exists in an open conformation. EPR thus provides additional insight into the different conformations of PcrG in the free form and when bound with PcrV.

4.2 Introduction

Pseudomonas aeruginosa is an opportunistic Gram-negative bacterium which is the lead cause of nosocomial infections among cystic fibrosis patients, AIDS and other immunocompromising diseases[1],[2]. The bacterium deploys the Type III Secretion System to initiate infection [3]. T3SS consists of the needle complex, chaperone and effector proteins. The needle complex is formed by over 20 different proteins, which consists of the base, needle , tip

complex and translocon.[4],[5]. The base is anchored to the bacterial outer and inner membrane[4]. The needle assembles in the intermembrane space between the bacteria and the host cell[4]. The tip caps the needle and provides the platform for the formation of the translocon[6]. The translocon forms a pore in the host cell membrane for the transport of virulence effector proteins[5].

The tip complex deploys the assembly of the translocon proteins upon contact with the host cells. The tip complex in *Pseudomonas aeruginosa* is assembled by multiple copies of the tip protein, PcrV (294 residues). PcrV shows 30 percent sequence identity and 60 precent sequence similarity with its *Yersinia* homolog LcrV[7],suggesting that PcrV and LcrV have similar structures[8].

Preceding the formation of the needle complex, in the bacterial cytosol, the tip protein PcrV, interacts with the tip chaperone protein, PcrG (98 residues)[9],[10],[11].Surface plasmon resonance(SPR) indicated that the PcrV- PcrG complex shows tight binding with a k_d of 26nM[3]. PcrG also interacts with PscO and PcrD to control the secretion of effector proteins[12]. NMR results indicated that that PcrG has two partially folded alpha helices (helix α1and helix α2) with no long-range distances between the helices (Figure 3-1) with fast nanoscale motions[3]. Structural prediction by computational modelling suggested that PcrG is a four helix bundle[13]. The *Yersinia* tip chaperone LcrG is predicted to form an alpha helical hairpin structure upon binding with LcrV, based on the structure of the N-terminal alpha hairpin domain of the *Shigella* tip protein IpaD[14]. PcrG is also hypothesized to have a similar coiled-coil topology upon binding with PcrV.

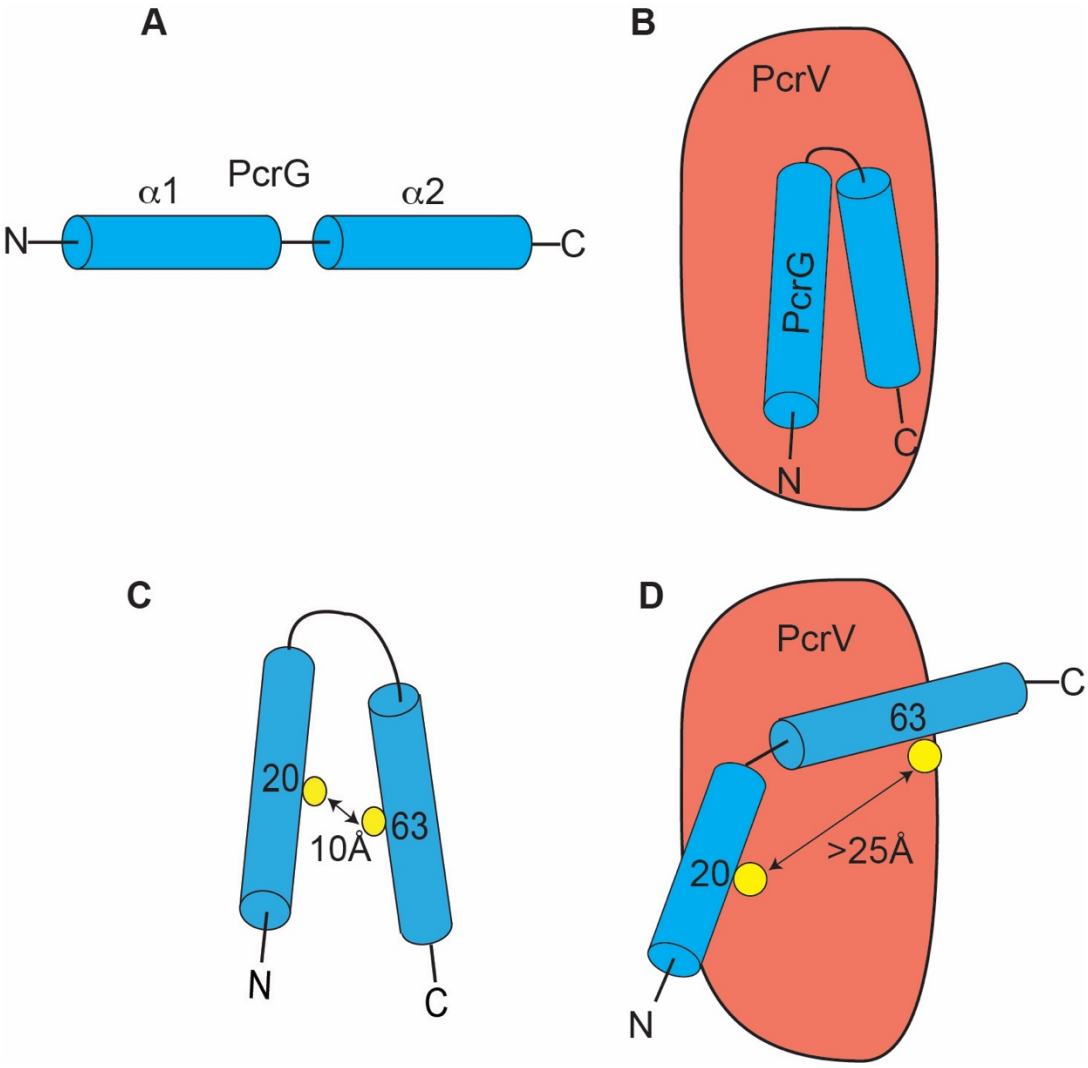

Figure 4-1 Models of the open and closed conformations of PcrG. (A) A model of PcrG based on NMR shows 2 helices in PcrG (helix α1 and helix α2) that do not interact. **(B)** A model of PcrG-PcrV interaction shows PcrG forms an alpha-helical hairpin upon binding to PcrV. **(C)** EPR spin-spin interaction revealed a closed conformation of PcrG where helix α1 and helix α2 are close to each other by 10 Å. This is designated as the closed conformation of PcrG. **(D)** EPR of PcrG-PcrV interaction revealed that helix α1 and helix α2 of PcrG move farther apart upon binding to PcrV. This is designated the open conformation of PcrG.

The structural dynamics of PcrG shown by NMR[3] and the predicted alpha helical hairpin structure upon binding with PcrV[3],[13], led to our hypothesis that there could be a population of PcrG wherein helix α1 interacts with helix α2. The long-range distances estimated by NOE based NMR can detect distances up to 6Å[15]. For detecting long range distances greater than 6Å, we used EPR spectroscopy. MTSL spin probes were attached to helix α1and helix α2, and the EPR data was acquired from PcrG in its free form and when bound to the tip protein PcrV. Our EPR data showed a population of PcrG where the helices α1 and α2 are in close contact with a spin-spin distance of 10Å forming the closed conformation of PcrG. The predicted models of PcrG-PcrV complex hypothesize that PcrG is alpha hairpin or a coiled-coil upon binding with PcrV[14]. Our EPR data shows that when PcrG is bound to PcrV, the distance between the helix α1 and helix α2 is higher than 25Å. Our EPR data provides additional insights into the different conformations of PcrG.

4.3 Methods

4.3.1 Protein expression and purification

$PcrG^{9-76}$ was used for biophysical characterization[3] instead of full length PcrG ,which aggregates in millimolar concentrations. $PcrG^{9-76}$ is a shorter construct comprising residues 9-76 and lacking the unstructured amino and carboxy termini. $PcrG^{9-76}$ is used for the current EPR studies. This shorter construct had all the structured alpha helical regions of PcrG as the structure characterized by NMR[3]. Site directed mutagenesis by Quickchange (Stratagene) was used to introduce cysteine mutations in helix α1 and helix α2 to generate $PcrG^{9-76}$ E20C and $PcrG^{9-76}$ E63C, respectively. The double cysteine mutant $PcrG^{9-76}$ E20C/E63C was also generated by site

directed mutagenesis to test the interaction between helix α1 and α2. These PcrG$^{9-76}$ constructs were expressed as fusion proteins (PcrG$^{9-76}$-TEVprotease cleavage site-His$_6$-GB1) in the plasmid pDZ3. After digestion with TEV protease, the used PcrG$^{9-76}$ constructs contain 6 residues (ENLYFQ) the C-terminus as an artifact[3].

Plasmids expressing PcrG$^{9-76}$ constructs were transformed in *E. coli* BL21(DE3) DNAY and grown in a culture media supplemented by carbenicillin and kanamycin. A 20 ml TB overnight culture starter was used to inoculate 1L TB, and cells were grown at 37 ºC in a shaker incubator. At A_{600} 07-0.8, cells were induced with 1mM IPTG and the cell growth was continued at 25 ºC for 4 hrs at a shaking speed of 120rpm. Cells were harvested at 4000 rpm by centrifugation (4000 rpm, 2392×*g*, 12 min, 4 ºC) in 25 ml binding buffer with 600 µl 10 mM PMSF. The cells were sonicated for a total of 5 mins 20 sec using the sonication cycle of 2s pulse and 30s pulse off. The cell lysate was centrifuged at 13,000 rpm(13865×*g*) for 15 min at 4 ºC. For the precipitation of nucleic acids in the supernatant a volume of 700 µl of 5 % PEI was added, followed by a centrifugation at 13,000 rpm for 15 mins and 4 ºC. The supernatant was loaded in a 5 ml nickel column, followed by 3 (50 ml) washes of the binding buffer. The PcrG$^{9-76}$ constructs eluted in the elution buffer. TEV protease (250 µl of 0.07 mM stock) was added, and the solution was dialyzed by at room temperature in 1 L.TEV buffer (100 mM Tris pH 8.0,0.5 mM EDTA,1 mM DTT, 100 mM NaCl) overnight. The proteins were further dialyzed in 1 L binding buffer for 4 hrs and in 3 L binding buffer overnight. The TEV digests are loaded in a 5ml nickel column, the flow through was collected and followed by 3 washes in 50ml binding buffer. The PcrG$^{9-76}$ constructs appear in the flow through and the first wash, whereas the His$_6$-GB1 tag appeared in the elution fraction. Purified proteins were dialyzed twice in 1 L EPR buffer (10 mM $NaPO_4$,

10mM NaCl, pH 7.0), concentrated using Amicon 3k, and the protein concentration was estimated by absorbance at A_{280}.

PcrV (residues 25-294) was expressed as a fusion protein (His_6-GB1-TEV protease cleavage site-PcrV) in pDZ1 was described previously[3].After TEV digestion, PcrV has a 3 residue (GHM) N-terminal cloning artifact. The purification of PcrV followed the same protocol as described above, except for the following modifications: after TEV digestion, the digest was loaded on a 5ml nickel column, the flow through was collected and followed by wash 1 in 50 ml binding buffer, wash 2 with 50 ml (20 mM Tris pH 8.0,500 mM NaCl, 25 mM Imidazole) and wash 3 with 50 ml (20 mM Tris pH 8.0,500 mM NaCl, 50 mM Imidazole). The fractions flow through, wash 1 and wash 2, the His_6-GB1 tag was found in the elution fraction.

4.3.2 *Site directed spin labeling*

The PcrG $^{9\text{-}76}$ constructs were spin labeled with MTSL as described in Chapter 2.3.2

4.3.3 *Circular dichroism spectroscopy*

The CD spectra was acquired for the $PcrG^{9\text{-}76}$ cysteine mutants along with spin labeled $PcrG^{9\text{-}76}$ as described in Chapter 2.3.3.

4.3.4 *EPR spectroscopy*

The EPR data was acquired spin labeled $PcrG^{9\text{-}76}$ constructs following similar protocols described in Chapter 2.3.4.

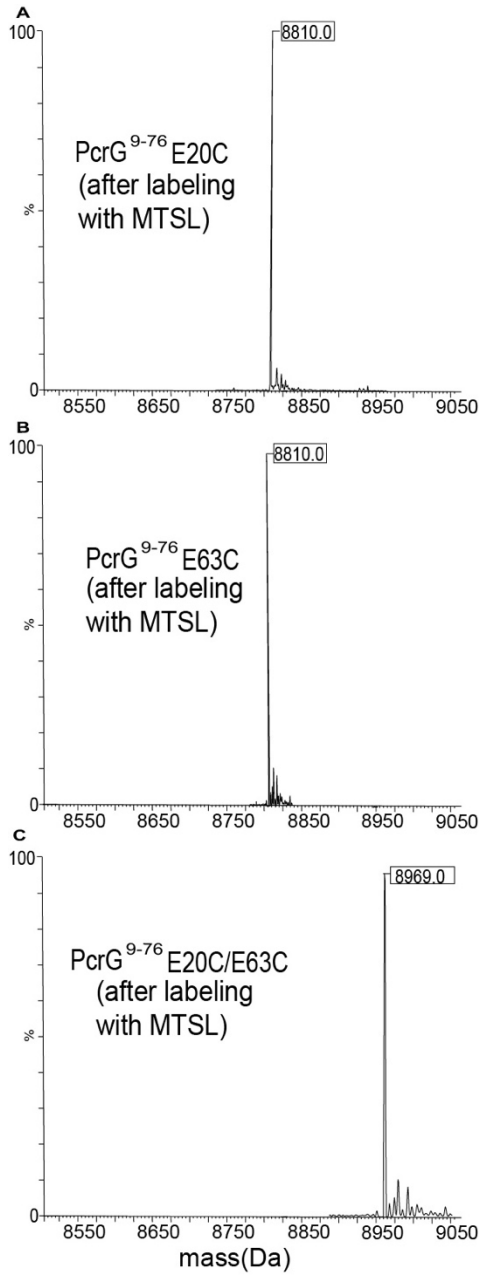

Figure 4-2. Electrospray ionization mass spectrometry (ESI-MS) to verify the attachment of MTSL label (A) PcrG $^{9-76}$ E20C (theoretical MW 8810.0 Da) with one MTSL spin-label **(B)** PcrG $^{9-76}$ E63C (theoretical MW 8810.0 Da) with one MTSL spin-label **(C)** PcrG $^{9-76}$ E20C/E63C with two MTSL spin-labels (theoretical MW 8969.0 Da).

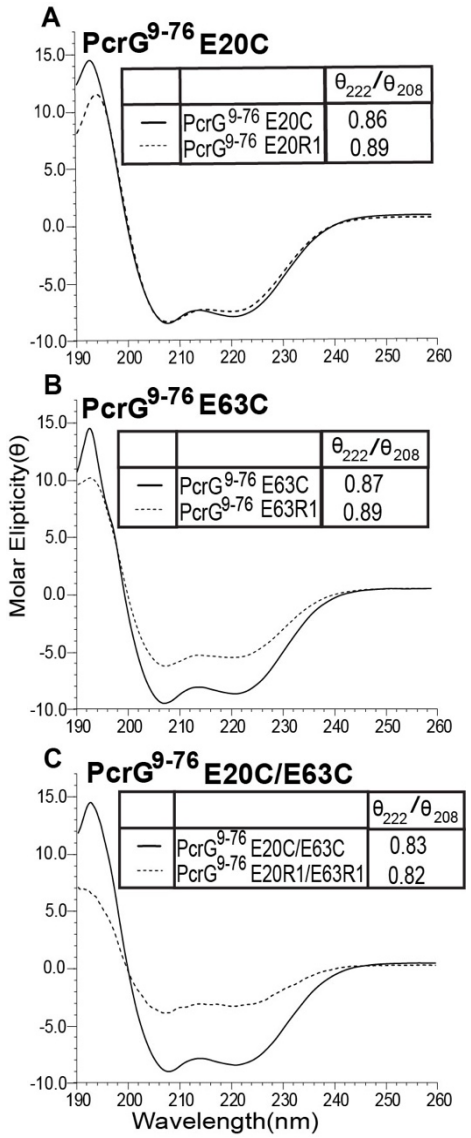

Figure 4-3. CD spectroscopy of PcrG constructs used for EPR studies. (A) CD spectra of PcrG$^{9-76}$ E20C before (solid line) and after (dashed line) (PcrG$^{9-76}$ E20R1) attachment of MTSL. **(B)** CD spectra of PcrG$^{9-76}$ E63C before (solid line) and after (PcrG$^{9-76}$ E63R1) (dashed line) attachment of MTSL side chain R1 **(C)** CD spectra before (solid line) PcrG$^{9-76}$ E20C/E63C and after attachment of MTSL side chain R1 (dashed line) (PcrG$^{9-76}$ E20R1/E63R1).

4.4 Results

4.4.1 Protein expression and purification

The site directed mutagenesis of PcrG$^{9-76}$ led to the formation of PcrG$^{9-76}$ E20C, PcrG$^{9-76}$ E63C, and PcrG$^{9-76}$ E20C/E63C. The final concentrations of these constructs are in the range 0.2-0.3mM with a similar expression and yield as PcrG$^{9-76}$[3]. PcrV are purified under native conditions as mentioned in[3] with a final yield of ~0.6 mM.

4.4.2 MTSL spin labeling of PcrG$^{9-76}$ constructs verified by electrospray ionization mass spectrometry (ESI-MS)

The theoretical mass of PcrG $^{9-76}$ E20C and PcrG $^{9-76}$ E63C is estimated to be 8626.0 Da using protparam server[18]. Site directed spin labeling (SDSL) added the MTSL side chain (R1) with an estimated mass in the range of 184-186 Da. Electrospray ionization (ESI) mass spectrometry showed that the addition of the R1 MTSL side chain to the cysteine residues of PcrG$^{9-76}$ E20C and PcrG$^{9-76}$ E63C increased the mass of the samples by 184 Da, resulting in the total mass of 8810 Da as shown in Fig 4-2 A-B.

The protein sample PcrG$^{9-76}$ E20C/E63C have two cysteine residues for spin labeling. The theoretical MW of PcrG$^{9-76}$ E20C/E63C is estimated to be 8599 Da. After the addition of two R1 spin-label side chains, the mass is estimated to be 8969 Da, increasing the mass of the protein sample by 370 Da as shown in Fig. 4-2C.

4.4.3 Circular dichroism spectroscopy shows no overall change in secondary structures of $PcrG^{9-76}$ constructs

The CD spectra acquired for the $PcrG^{9-76}$ constructs used for EPR studies shows that the cysteine residue mutations did not change the overall secondary structure of $PcrG^{9-76}$[3]. The constructs show a characteristic CD spectrum for alpha helical proteins[19], with minima at 222 nm and 208 nm. . The Molar Ellipticity ratio at 222nm and 208nm ($\theta_{222}/\theta_{208}$) for all the $PcrG^{9-76}$ constructs were in the range of 0.83-0.87, showing the alpha helices of the cysteine mutants are non-interacting.

CD data also showed that addition of the MTSL spin labels have did not affect the secondary structures of $PcrG^{9-76}$. The MTSL spin labeled samples $PcrG^{9-76}$ E20R1, $PcrG^{9-76}$ E20R1, and $PcrG^{9-76}$ E20R1/E63R1 had a $\theta_{222}/\theta_{208}$ ratio of 0.89, 0.89, and 0.82, respectively, denoting that the MTSL side chain did not affect the overall secondary structure of $PcrG^{9-76}$ constructs.

4.4.4 Electron paramagnetic resonance spectroscopy shows different conformations of $PcrG^{9-76}$ constructs

EPR spectroscopy can detect long range contacts in low populated states. We used EPR spectroscopy to test the hypothesis that the helix α1 and helix α2 of PcrG interact with each other. For the EPR studies, the MTSL spin-label (R1) was covalently attached to the cysteine residue, as shown in Fig 4-4A. Three cysteine mutants of PcrG were spin labeled and used to study the interaction between helix α1 and helix α2 — (1) A cysteine point mutation engineered at helix α1 at position E20 of $PcrG^{9-76}$ for single spin labeling. (2) A cysteine point mutation engineered at helix α2 (E63) of $PcrG^{9-76}$ for single spin labeling. (3) $PcrG^{9-76}$ engineered with a cysteine point mutation at residue E20 of helix α1 for spin labeling, followed by another cysteine point mutation at residue E63 of helix α2 for spin labeling. For EPR spectroscopy, the single spin

labeled constructs were used as control and the construct with two spin labels was used as the sample to detect the interaction between helix α1 and helix α2.

The EPR spectra generated by single spin labeled PcrG$^{9-76}$ E20R1 and PcrG$^{9-76}$ E63R1 was added, resulting in the black control spectrum shown in Fig 4-4C. The red double labeled spectrum from PcrG$^{9-76}$ E20R1/E63R1(Fig 4-4C) showed peak broadening and reduction in peak height, indicating short spin-spin distances. The distance between the two spin labels was estimated to be 10 Å. The EPR acquired by my collaborator showed that there is a population of PcrG where helix α1 and helix α2 are in close contact with each other. This spin-spin interaction was designated as the closed conformation of PcrG, as shown in Fig 4-4C. This EPR data was the first experimental data showing direct interaction between helix α1 and helix α2 of PcrG.

EPR spectroscopy was also used to determine the spin-spin interaction between helix α1 and helix α2 of PcrG when it is bound to the tip protein PcrV. The previous PcrG constructs used for EPR studies were complexed with unlabeled PcrV in a 1:1 molar ratio[3],[10] to determine the spin-spin distance between E20R1 and E63R1 of the PcrG$^{9-76}$ E20R1/E63R1 construct. The additive spectrum of the two single labeled constructs — PcrG$^{9-76}$ E20R1 complexed with PcrV and PcrG$^{9-76}$ E63R1 complexed with PcrV served as the control. The EPR results showed that when PcrG is bound to PcrV, the helix α1 and helix α2 of PcrG do not show any short spin-spin distances. The EPR spectrum of double labeled construct, PcrG$^{9-76}$ E20R1/E63R1bound to PcrV was identical to the control spectrum (black, Fig 4-5B), and the two spectra superimposed on each other. Our EPR results revealed that when PcrG is bound to PcrV, the spin-spin distance between helix α1 and helix α2 is greater than 25 Å. The increase in spin-spin distance between helix α1 and helix α2 of PcrG when bound to PcrV was designated as the open conformation of PcrG.

4.5 Discussion

EPR spectroscopy was used here to determine the different conformations of PcrG. PcrG plays an important role in the virulence of *Pseudomonas aeruginosa* as the chaperone to tip the protein PcrV and also as a regulator the transport of Pop effector proteins by interacting with PscO and PcrD[9]. The *Yersinia* T3SS has a protein homologous to PcrG known as LcrG. LcrG and PcrG form the family of tip chaperone proteins which are important for the transport of effector proteins in *Pseudomonas aeruginosa* and in *Yersinia pestis*[9],[20].There are currently no high resolution structures of LcrG and PcrG. Structural predictions by computational modelling suggests that PcrG and LcrG exists as a four helical bundle and coiled-coil[13],[21].The current data on the structures LcrG and PcrG comes from NMR spectroscopy, which shows that both LcrG and PcrG are partially folded alpha helical protein with fast nanoscale motions[3],[22].

The NMR data shows that PcrG has two alpha helices that are connected by an unstructured linker regions, and that NOE-NMR data did not detect any long range interaction between helix α1 and helix α2[3]. The NMR data using ^{15}N-labeled full-length PcrG ($PcrG^{FL}$,98 amino acids) yielded an NMR spectrum where the peaks were not well resolved and were overlapped. This could be due to the presence of disordered regions in PcrG[16]. To design a PcrG construct that will yield an ideal NMR spectrum, the program PSIPRED was used to predict the secondary structure of PcrG[16].The extreme N and C termini of PcrG were predicted by PSIPRED[17] to be random coils[16]. To remove the disordered terminal residues and include the structured regions of PcrG while maintaining the hydrophilic terminal residues for solubility, a PcrG construct was generated with residues Glu-9 to Ser-76 ($PcrG^{9-76}$)[16]. Circular dichroism (CD)

spectra of $PcrG^{FL}$ and $PcrG^{9-76}$ showed alpha helical content based on the minima at 222nm and a significant amount of disorder based on the predominant minima at. 208nm[3],[16]. The CD spectra confirmed that $PcrG^{FL}$ and $PcrG^{9-76}$ have nearly identical secondary structures because their CD spectra were nearly superimposable[3],[16]. Thus, the $PcrG^{9-76}$ truncation did not produce any appreciable changes to the secondary structure of PcrG[3],[16]. The SPR and NMR data of PcrG and PcrV interaction also showed that $PcrG^{FL}$ and $PcrG^{9-76}$ bind similarly to PcrV[3],[16]. In other words, the truncation did not alter the ability of PcrG to bind to PcrV and $PcrG^{FL}$ and $PcrG^{9-76}$ bound to identical residues of PcrV. The CD, SPR and NMR studies, shows that $PcrG^{9-76}$ is structurally and functionally similar to $PcrG^{FL}$ and can used for generating cysteine mutations for EPR studies.

Our EPR data shown here shows that there is a population of PcrG where the helix α1 and helix α2 are in close contact with each other. The spin-spin distance between the two helices is estimated to be 10 Å (Fig 4-1C), forming the closed conformation of PcrG. During the assembly of the needle complex, PcrG interacts with tip PcrV to maintain its structural integrity in the bacterial cytosol[9]. PcrG-PcrV form a tight binding complex[10],[11].The *Yersinia* homologs LcrV and LcrG also form a tight binding complex with nanomolar affinities[23]. Currently, there is no crystal structure of the PcrG-PcrV or the LcrG-LcrV complex. The model of interaction of the LcrG-LcrV complex is derived by comparison of the crystal structures of the following tip proteins *Salmonella* SipD, *Shigella* IpaD and *Burkholderia* BipD[24],[25],[26] to the crystal structure of *Yersinia* LcrV[27], with shared structural features including a mixed α/β domain and a long central coiled-coil region. The crystal structures of SipD, IpaD and, BipD have an N-terminal alpha helical hairpin domain which is missing in LcrV. Comparing the crystal structures of tip proteins, LcrG is hypothesized to form an alpha hairpin when bound to LcrV and

interact with the long coiled-coil domain[14]. Based on functional homology with the *Yersinia* T3SS proteins, this hypothesis is also extended to the PcrG-PcrV complex. Our EPR data shows that when PcrG is bound to PcrV, the spin-spin distance between helix α1 and helix α2 is greater than 25 Å, forming an open conformation (Fig 4-1D). Our EPR suggests PcrG does not form an alpha hairpin when bound to PcrV.

In conclusion, our EPR studies show the first direct experimental evidence that PcrG has different structural conformations in the free form and when bound to PcrV. PcrG in its free form has a closed conformation with a spin-spin distance of 10 Å. When PcrG is bound to PcrV, it forms an open conformation with a spin-spin distance greater than 25 Å.

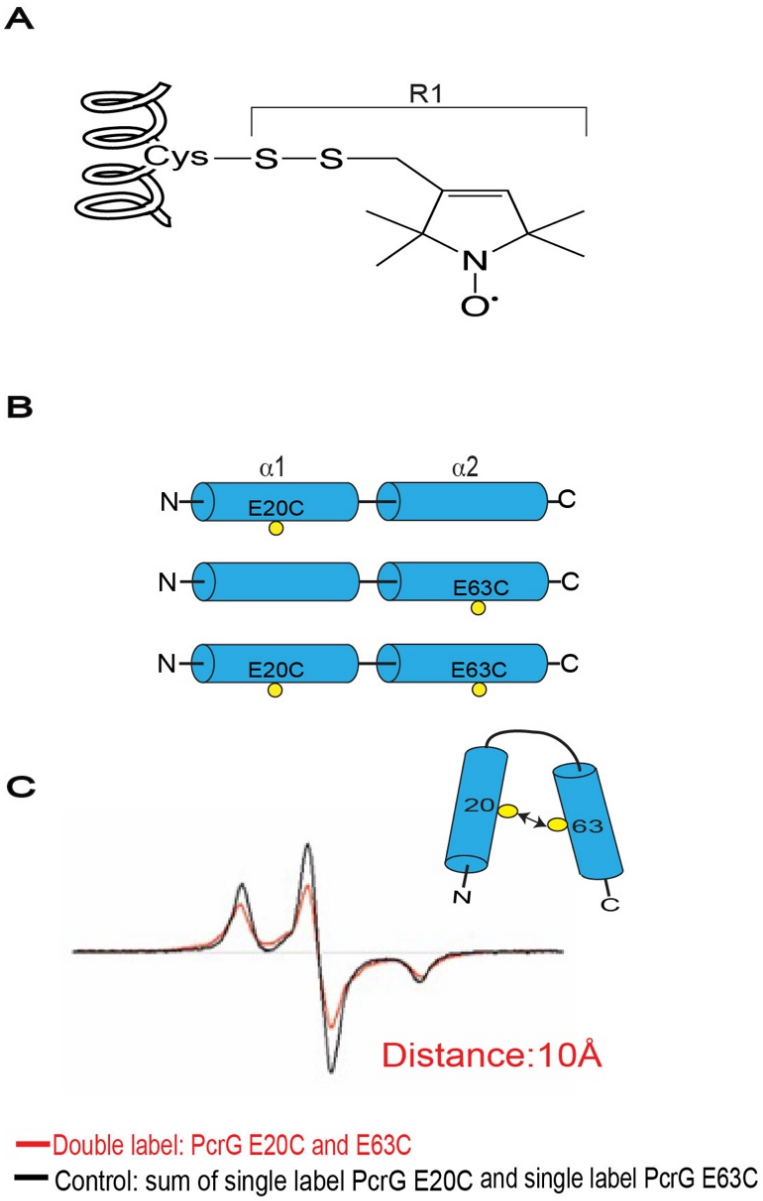

Figure 4-4. EPR reveals a closed conformation of PcrG. **(A)** MTSL spin-label as an amino acid side chain (R1). **(B)** The PcrG$^{9-76}$ spin-labeled constructs used to determine the spin-spin interaction between helix α1 and helix α2. **(C)** EPR spectra of the spin-spin interaction of PcrG$^{9-76}$ E20C and E63C.

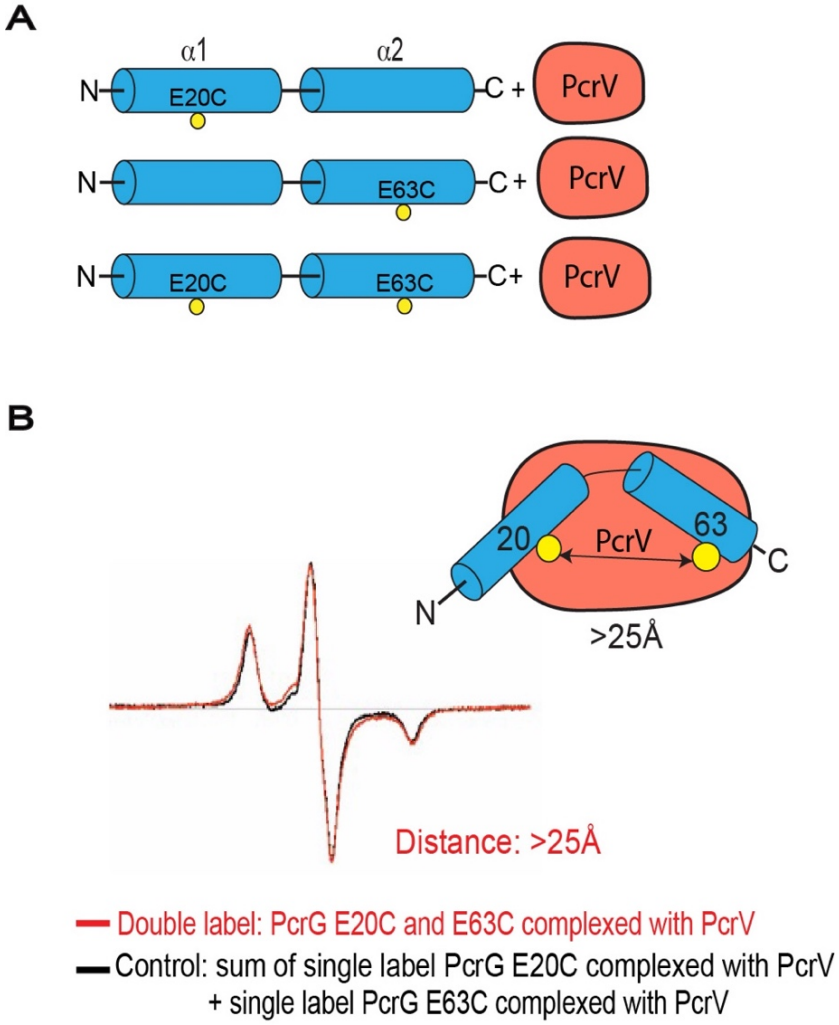

Figure 4-5. PcrG forms an open conformation upon binding with PcrV. (A) The PcrG$^{9-76}$ spin-labeled constructs used in complex with PcrV. **(B)** EPR spectra of the spin-spin interaction of PcrG$^{9-76}$ E20C and E63C in complex with PcrV.

4.6 References

1. Lyczak, J.B., C.L. Cannon, and G.B. Pier, *Establishment of Pseudomonas aeruginosa infection: lessons from a versatile opportunist.* Microbes Infect., 2000. **2**(9): p. 1051-60.
2. Engel, J. and P. Balachandran, *Role of Pseudomonas aeruginosa type III effectors in disease.* Curr. Opin. Microbiol., 2009. **12**(1): p. 61-6.
3. Chaudhury, S., et al., *Nuclear Magnetic Resonance Characterization of the Type III Secretion System Tip Chaperone Protein PcrG of Pseudomonas aeruginosa.* Biochemistry, 2015. **54**(43): p. 6576-85.
4. Galan, J.E., et al., *Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells.* Annu. Rev. Microbiol., 2014. **68**: p. 415-38.
5. Dey, S., et al., *The type III secretion system needle, tip, and translocon.* Protein Sci, 2019. **28**(9): p. 1582-1593.
6. Mueller, C.A., P. Broz, and G.R. Cornelis, *The type III secretion system tip complex and translocon.* Mol. Microbiol., 2008. **68**(5): p. 1085-95.
7. Stothard, P., *The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences.* BioTechniques, 2000. **28(6)**.
8. Chatterjee, S., et al., *Structure and biophysics of type III secretion in bacteria.* Biochemistry, 2013. **52**(15): p. 2508-17.
9. Lee, P.C., et al., *Control of effector export by the Pseudomonas aeruginosa type III secretion proteins PcrG and PcrV.* Mol. Microbiol., 2010. **75**(4): p. 924-41.
10. Nanao, M., et al., *Type III secretion proteins PcrV and PcrG from Pseudomonas aeruginosa form a 1:1 complex through high affinity interactions.* BMC Microbiol., 2003. **3**: p. 21.
11. Allmond, L.R., et al., *Protein binding between PcrG-PcrV and PcrH-PopB/PopD encoded by the pcrGVH-popBD operon of the Pseudomonas aeruginosa type III secretion system.* Infect. Immun., 2003. **71**(4): p. 2230-3.
12. Lee, P.C., et al., *Control of type III secretion activity and substrate specificity by the cytoplasmic regulator PcrG.* Proc. Natl. Acad. Sci. U.S.A., 2014. **111**(19): p. E2027-36.
13. Basu, A., et al., *PcrG protects the two long helical oligomerization domains of PcrV, by an interaction mediated by the intramolecular coiled-coil region of PcrG.* BMC Struct Biol, 2014. **14**: p. 5.
14. Blocker, A.J., et al., *What's the point of the type III secretion system needle?* Proc. Natl. Acad. Sci. U.S.A., 2008. **105**(18): p. 6507-13.
15. Hoffman, R.C., et al., *A simple method for the refinement of models derived from NMR data demonstrated on a zinc-finger domain from yeast ADR1.* J. Magn. Reson. Ser. B, 1993. **102**(1): p. 61-72.
16. Chaudhury, S., *Structural studies of chaperones and chaperone-tip interactions from the type III secretion systems of Yersinia and Pseudomonas*, in *Molecular Biosciences*. 2013.
17. McGuffin, L.J., K. Bryson, and D.T. Jones, *The PSIPRED protein structure prediction server.* Bioinformatics, 2000. **16**(4): p. 404-5.
18. Gasteiger E., H.C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A, *Protein Identification and Analysis Tools on the ExPASy Server;.* The Proteomics Protocols Handbook, Humana Press 2005(John M. Walker (ed)): p. 571-607

19. Greenfield, N., *Using circular dichroism spectra to estimate protein secondary structure.* Nat.Protoc.1, 2006: p. 2876–2890.
20. Cornelis, G.R., *Yersinia type III secretion: send in the effectors.* J. Cell Biol., 2002. **158**(3): p. 401-408.
21. Pallen, M.J., G. Dougan, and G. Frankel, *Coiled-coil domains in proteins secreted by type III secretion systems.* Mol. Microbiol., 1997. **25**(2): p. 423-5.
22. Chaudhury, S., et al., *The LcrG tip chaperone protein of the Yersinia pestis type III secretion system is partially folded.* J. Mol. Biol., 2015. **427**: p. 3096-3109.
23. Lawton, D.G., et al., *Interactions of the type III secretion pathway proteins LcrV and LcrG from Yersinia pestis are mediated by coiled-coil domains.* J. Biol. Chem., 2002. **277**(41): p. 38714-22.
24. Lunelli, M., et al., *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., 2011. **7**(8): p. e1002163.
25. Johnson, S., et al., *Self-chaperoning of the type III secretion system needle tip proteins IpaD and BipD.* J. Biol. Chem., 2007. **282**(6): p. 4035-4044.
26. Erskine, P.T., et al., *High Resolution Structure of BipD: An Invasion Protein Associated with the Type III Secretion System of Burkholderia pseudomallei.* J. Mol. Biol., 2006. **363**(1): p. 125-136.
27. Chaudhury, S., et al., *Structure of the Yersinia pestis tip protein LcrV refined to 1.65 Å resolution.* Acta Crystallogr. Sect. F, 2013. **69**(Pt 5): p. 477-81.

Chapter 5: Binding Affinity of the PcrG-PcrV Complex determined by FRET Spectroscopy

5.1 Abstract

Pseudomonas aeruginosa, the major cause of lung infections among cystic fibrosis patients, assembles a protein nanoinjector of the Type III Secretion System (T3SS) to inject virulence effector proteins into their target host cells. The T3SS nanoinjector consists of the needle complex, chaperones, and effector proteins. The needle complex comprises of a base, a needle, a tip complex and a translocon. PcrG functions as the chaperone to the tip protein, PcrV in *Pseudomonas*. Previous NMR and SPR data showed that PcrG-PcrV complex has tight binding interaction. How PcrG-PcrV interact with each other is currently unknown. Here, I used Forster Resonance Energy Transfer (FRET) to determine the mode of binding and the binding affinity of the PcrG-PcrV complex. To do FRET, I and my collaborators labeled $PcrG^{9-76}$ A66C with the donor N-(1-Pyrene) Maleimide (PM), and we labeled PcrV V120C with the acceptor, 7-Diethylamino-3-(4'-Maleimidylphenyl)-4-Methylcoumarin (CPM). We determined that the dissociation constant (k_d) for the PcrG-PcrV complex to be 15 ± 2 nM. Our results shows that $PcrG^{9-76}$ A66C binds to PcrV V120C.

5.2 Introduction

Pseudomonas aeruginosa is a Gram-negative bacterium frequently associated with the nosocomial infections in patients with cystic fibrosis, AIDS, pneumonia, cancer and other immunocompromised diseases with a mortality rate of 40% in case of acute infection[1],[2],[3]. The bacterium infects the host cell by deploying the Type III Secretion System (T3SS). T3SS consist of the needle complex, chaperones and effectors proteins[4],[5]. The needle complex is assembled by over 20 different proteins which form a base, a needle, a tip and a translocon[5]. The base is embedded in the inner and outer bacterial membrane and the needle protrudes from

the base forming the needle in the intermembranous space between the pathogen and the host cell[6]. The tip complex caps the needle and provides the platform the assembly of the translocon[7]. The pore formed by the translocon proteins in the host cell membrane transports virulent effector proteins for infection[4].

The tip complex is an environmental sensor for the deployment of the translocon proteins when the bacteria come in contact with the host cell[8]. The T3SS genes are activated for initiating the assembly of the base, needle, and tip complex[9]. When the bacteria come in contact with the host cell membrane, it leads to translocon assembly for effector protein transport[5]. The *Pseudomonas* tip complex consists of 5 copies of the tip protein PcrV[10]. There are no high resolution structures of PcrV, however, an I-TASSER model[11] shows that PcrV has structural similarity with the *Yersinia* homolog LcrV[12].

PcrV forms a tight binding complex with the tip chaperone protein, PcrG which functions as a chaperone prior to the assembly of the needle complex[13],[14]. PcrG also regulates Pop effector secretion by interaction with the needle complex proteins PscO and PcrD[15].The structure of PcrG determined by NMR spectroscopy shows that PcrG is formed two partially folded alpha helices connected by an unstructured linker[16].

Previous SPR data shows that PcrV binds to PcrG in a 1:1 molar ratio[16].There are no high resolution structures of the PcrG-PcrV interaction. The current knowledge of the PcrG-PcrV binding is derived from the hypothesis made for the *Yersinia* LcrG-LcrV complex, where LcrG is predicted to be a coiled-coil when bound to LcrV[17]. Previous NMR data also shows that PcrV induces a global change in the structure of PcrG, indicating that all the residues of PcrG are involved in binding interaction or show binding induced conformation changes[16]. My current EPR data has also revealed that the N-terminal and C-terminal helices of PcrG are at a spin-spin

distance greater than 25 Å, indicating that PcrG has an open conformation when bound to PcrV (Fig 4-5, Chapter 4).

Based on the NMR data[16], my previous EPR data for the LcrV-LcrG interaction (Fig 3-6, Chapter 3), and the conformational change of PcrG upon binding with PcrV (Fig 4-5, Chapter 4), I propose some possible models of PcrG-PcrV interaction (Fig 5-1) to be verified by FRET. For the donor and acceptor fluorophore attachment, I engineered cysteine mutation on the N-terminal helix (helix α1) and C-terminal helix (helix α2) of PcrG and the N-terminal domain of PcrV to determine the binding interaction of PcrG and PcrV. I collaborated with Dr. Mark Richter and Dr. Erik Holmstrom (University of Kansas) to acquire FRET data. Our FRET data shows that C-terminal PcrG construct $PcrG^{9-76}$ A66C interacts with N-terminal domain of PcrV construct PcrV V120C with a dissociation constant (k_d) of 15 ± 2 nM, showing tight binding interaction for PcrG-PcrV complex.

5. 3 Methods

5.3.1 Protein expression and purification

Site directed mutagenesis by Quick Change (Stratagene) was used to introduce the following cysteine point mutations in $PcrG^{9-76}$ — E20C, R24C, E63C, and A66C for the attachment of the donor fluorophores. Cysteine mutations were also introduced in PcrV to generate PcrV E76C and PcrV V120C for attaching the acceptor fluorophores. The PcrG and PcrV constructs were expressed and purified as described in Chapter 4.3.1. The purified proteins were dialyzed in 5L of Phosphate-Saline Buffer (PBS) pH 7.4, concentrated using Amicon 3k, and the protein concentration was estimated by absorbance at A_{280}.

5.3.2 *Labeling of donor and acceptor fluorophore*

The $PcrG^{9-76}$ and PcrV constructs were in Phosphate-Buffer Saline (PBS), pH 7.4 with a concentration of 15uM each. A 50-fold-molar excess of Tris(2-carboxyethyl) phosphine hydrochloride (TCEP) was added and incubated for 2 hours to reduce disulfide bonds in the protein samples. After reduction, excess TCEP was removed by exchange using PD-10 desalting column before the labeling reaction. A 4-fold molar excess of the donor and acceptor fluorophore dyes were added to the PcrG and PcrV constructs, respectively. The labeling reaction proceeded for 4 hours at room temperature with rocking or intermittent light vortexing.

The labeled proteins were separated from excess dyes by two consecutive PD-10 desalting column exchanges with PBS buffer. The spectrophotometer was used to identify dye-labeled proteins by the characteristic absorbance peak for the donor and the acceptor . The degree of labeling (DOL) was estimated by using the equation below[18]:

$$\mathbf{DOL= (A_{max} * \varepsilon_{prot})/(A_{280}-A_{max}.CF_{280}) *\varepsilon_{max}}$$

A_{max}, the absorbance of the fluorophore at wavelength maxima, $\varepsilon_{prot,}$ the molar extinction coefficient of the protein used for fluorophore labeling, A_{280}, the absorbance of the protein used for fluorophore labeling, CF_{280} correction factor of the fluorophore at 280 nm, and ε_{max} the molar extinction coefficient of the fluorophore.

5.3.3 *Förster resonance energy transfer (FRET) spectroscopy using FM and AF647*

The protein-protein interaction between Fluorescein C5 Maleimide (FM)-labeled $PcrG^{9-76}$ and $LcrG^{7-73}$ constructs and Alexa Flour(AF)-647 labeled LcrV and PcrV constructs was determined using FRET. Emission spectra were collected between 500nm to 800 nm with an excitation wavelength of 492nm, the excitation and emission slit width of 5nm, and an integration time of 0.1seconds, using a Varian Cary Eclipse Fluorescence Spectrophotometer at 25°C.

5.3.4 *Förster resonance energy transfer (FRET) spectroscopy using CPM and EMAL*

FRET was used to determine protein-protein interaction between CPM-labeled $PcrG^{9-76}$ constructs and Eosin-5-Maleimide (EMAL)-labeled PcrV constructs. Emission spectra were collected between 400 nm to 600 nm with an excitation wavelength of 384nm, the excitation and emission slit width of 5nm, and an integration time of 0.1seconds, using a Varian Cary Eclipse Fluorescence Spectrophotometer at 25°C.

5.3.5 *Förster resonance energy transfer (FRET) spectroscopy using PM and CPM*

PM-labeled $PcrG^{9-76}$ constructs and CPM-labeled PcrV constructs were used to determine protein-protein interaction using FRET. Emission spectra were collected between 350 nm to 550 nm with an excitation wavelength of 330 nm, the excitation and emission slit width of 5nm, and an integration time of 0.1seconds, using a Varian Cary Eclipse Fluorescence Spectrophotometer at 25°C.

5.3.6 Calculation of the dissociation constant (k_d)

The PcrG$^{9-76}$-PM and PcrV-CPM binding curves were plotted using Mathematica programming suite and fitted in the following equation to calculated the k_d[19],[20].

$$Y = initial + \frac{(final - initial)\ L + R + kd \pm \sqrt{(L + R + kd)^2 - 4LR}}{2R}$$

Where Y was the D/A ratio, L is the concertation of the PcrG$^{9-76}$-PM, R was the concentration of the PcrV-CPM, and kd was the dissociation constant, initial was the efficiency of the D/A ratio in the absence of any interaction and final was D/A ratio when the all molecules are in a complex.

5.4 Results

5.4.1 Protein expression and purification

The PcrG$^{9-76}$ and PcrV cysteine mutants for fluorophore labeling were engineered using site directed mutagenesis and were expressed and purified under native conditions. The constructs showed similar expression and yield as PcrG$^{9-76}$ and PcrV[16]. The final yield for the PcrG$^{9-76}$ constructs were ~ 0.1-0.25mM and the PcrV constructs had a final yield of ~ 0.6mM.

5.4.2 Donor and acceptor fluorophore labeling

The degree of labeling (DOL) of the donors-FM, CPM and PM labeled to PcrG$^{9-76}$ and LcrG$^{7-73}$ and the acceptors-AF-647, EMAL and CPM labeled to PcrV and LcrV proteins as been listed below in Table 5-1, 5-2 and 5-3.

5.4.3 FRET data acquired from Donor FM and Acceptor AF647

Fluorescein C5 Maleimide (FM) was used as the donor to label $PcrG^{9-76}$ E20C and $PcrG^{9-76}$ R24C. Alexa Fluor 647-C2-Maleimide (AF647) was used acceptor to label PcrV V120C[21]. The Ro value for the donor (FM) -acceptor (AF647) was estimated to be 56 Å. Based on the Ro value, this FRET pair could be used to determine the long-range distances between the donor labeled $PcrG^{9-76}$ constructs and the acceptor labeled PcrV V120C construct. The concentration $PcrG^{9-76}$ (labeled with the donor (FM)), and PcrV (labeled with the acceptor (AF-647)) was calculated to be 1μM.

The results of the FRET data acquired as shown in Fig 5-2 A-B as the emission spectra of $PcrG^{9-76}$ (labeled with the donor (FM)), the emission spectra of PcrV V120C (labeled with the acceptor (AF647)), and the emission spectra of the $PcrG^{9-76}$ (labeled with the donor (FM)) – PcrV (labeled with the acceptor (AF647)) complex. The three emission spectra were obtained using an excitation wavelength of 492nm. The peak at 520nm is the emission maxima of the donor, whereas the peak at 667nm is the emission maxima of the acceptor. The red plot, served as control, was the sum of two spectra — the donor's emission spectrum and the acceptor's emission spectrum. Upon complex formation, the emission spectra in Fig 5-2 A-B (black) showed a decrease in intensity of the donor maxima but no increase in the acceptor maxima, denoting that there is no energy transfer between the donor and the acceptor, as shown in Table 5-4.

To verify the energy transfer between the donor (FM) and acceptor (AF647), I labeled the $LcrG^{7-73}$ K28C with FM and LcrV Q37C with AF647. My EPR data for the LcrG-LcrV

interaction, as mentioned in Chapter 3.4.4, showed that MTSL labeled LcrG $^{7-73}$ K28C and LcrV Q37C were close upon complex formation. The results of the FRET data acquired for the LcrG$^{7-73}$ K28C FM and LcrV Q37C AF647 show a decrease in the donor maxima but no increase in the acceptor maxima, as shown in Fig 5-3, indicating that there is no energy transfer. This FRET data suggested that LcrG$^{7-73}$ K28C FM and LcrV Q37C AF647 do not interact, or these two residues are far apart when the complex is formed. Further analysis showed that the emission spectrum of the donor FM and excitation spectrum of the acceptor AF647 hardly show any spectral overlap, as shown in Fig 5-4, denoting low energy transfer between the donor and acceptor. So, this donor FM and acceptor AF647 pair could not be used to determine the interaction between PcrG and PcrV.

5.4.4 *Labeling with Donor CPM and Acceptor EMAL*

The dye 7-Diethylamino-3-(4'-Maleimidylphenyl)-4-Methylcoumarin (CPM**)** was used as the donor to label PcrG$^{9-76}$ E20C, PcrG$^{9-76}$ R24C, PcrG$^{9-76}$ E63C, and PcrG$^{9-76}$ A66C. Eosin-5-Maleimide (EMAL) was used acceptor to label PcrV V120C and PcrV E76C. The Ro value for the donor (CPM) -acceptor (EMAL) was estimated to be 45Å. The donor emission spectrum of CPM overlaps with the acceptor excitation spectrum of EMAL.

The results of the FRET data acquired as shown in Fig 5-5A-B, 5-6A-B are the emission spectra of PcrG$^{9-76}$ (labeled with the donor (CPM)), the emission spectra of PcrV (labeled with the acceptor (EMAL)) and the emission spectra of the PcrG$^{9-76}$ (labeled with the donor (CPM)) – PcrV (labeled with the acceptor (EMAL)) complex. The concentration PcrG$^{9-76}$ (labeled with the donor (CPM)), and PcrV (labeled with the acceptor (EMAL)) was calculated to be 1μM. The three emission spectra were obtained using an excitation wavelength of 384nm. The peak at

460nm is the emission maxima of the donor, whereas the peak at 547nm is the emission maxima of the acceptor. The red plot, served as control, was the sum of two spectra — the donor's emission spectrum and the acceptor's emission spectrum. Upon complex formation, the emission spectra in Fig 5-5 A-B, 5-6 A-B (black) showed a decrease in intensities of the donor maxima and the acceptor maxima, denoting that there is no energy transfer between the donor and the acceptor, as shown in Table 5-5. This lack of energy transfer between the donor labeled $PcrG^{9-76}$ constructs and acceptor labeled PcrV constructs shows no interaction or these residues are far apart when the PcrG-PcrV complex is formed. This FRET pair did not show any energy transfer for the PcrG-PcrV interaction.

The Ro value of this FRET pair is beyond the range of the PcrG-PcrV interaction that I was trying to detect. Further analysis of the degree of labeling (DOL) of the donor and acceptor shown in Table 5-2, indicted that a few more rounds of PD-10 column exchanges are required to remove the unbound fluorophores. The labeling protocol for this donor-acceptor pair needs to be optimized for FRET.

5.4.5 *FRET shows $PcrG^{9-76}$ A66C interacts with PcrV V120C*

Our previous EPR data acquired by my collaborator, Dr. Likai Song from Florida State University showed that in the case of the *Yersinia* LcrG-LcrV complex, LcrG did not form a coiled-coil when bound to LcrV[22] (Fig 3-6,Chapter 3). The EPR data acquired when PcrG is bound to PcrV showed the N-terminal (helix α1) and C-terminal (helix α2) had a spin-spin distance of more than 25 Å, which we designated as the open conformation of PcrG (Fig 4-5, Chapter 4). Based on our new EPR findings, I decided to determine the binding interaction of the PcrG-PcrV complex by FRET. For the FRET experiments, the $PcrG^{9-76}$ constructs — E20C,

R24C, E63C, and A66C were labeled with the donor fluorophore (PM), and the PcrV constructs — PcrV E76C, and PcrV V120C were labeled with the acceptor fluorophore (CPM). The emission spectra acquired had donor and acceptor concentrations of 1μM each (Fig 5-12).

The results of the FRET data acquired is shown in Fig 5-7 and Fig 5-8 as the emission spectra of $PcrG^{9-76}$ (labeled with the donor (PM)), the emission spectra of PcrV (labeled with the acceptor (CPM)), and the emission spectra of the $PcrG^{9-76}$ (labeled with the donor (PM)) – PcrV (labeled with the acceptor (CPM)) complex. The three emission spectra were obtained by using an excitation wavelength of 330 nm. The peaks at 375 nm and 395 nm are the emission maxima of the donor, whereas the peak at 460 nm is the emission maxima of the acceptor. The red plot, served as control, was the sum of two spectra — the emission spectrum of the donor and the emission spectrum of the acceptor. The black plot is the emission spectrum of the $PcrG^{9-76}$ (PM)-PcrV (CPM) complex. We observed a strong energy transfer for the complex formed by $PcrG^{9-76}$ A66 (PM) and PcrV V120 (CPM) as shown in Fig 5-7D. Upon complex formation of $PcrG^{9-76}$ A66C (PM) and PcrV V120C (CPM), the emission spectrum of the complex (black plot) showed decreased intensities of the donor peaks at 375 nm and 395 nm, followed by increased intensity of the acceptor peak at 460 nm. We did not observe energy transfer for the complexes formed by $PcrG^{9-76}$ E20C, R24C or E63C(PM), and PcrV V120C (CPM) (Fig 5-7A-C). Upon complex formation of $PcrG^{9-76}$ E20C, $PcrG^{9-76}$ R24C or $PcrG^{9-76}$ E63C (PM) and PcrV V120 (CPM), the emission spectrum of the complex (black plot) showed no decrease the intensities of the donor peaks at 375 nm and 395 nm or increase the acceptor peak at 460 nm. The results of Fig 5-7 suggested that among the four residues of $PcrG^{9-76}$ labeled with PM, $PcrG^{9-76}$ A66C is closer to PcrV V120C CPM upon complex formation, than PM labeled $PcrG^{9-76}$ E20C, R24C, or E63C.

Fig 5-8A-D shows no energy transfer between the complexes formed by PcrG$^{9-76}$ E20C, PcrG$^{9-76}$ R24C, PcrG$^{9-76}$ E63C, or PcrG$^{9-76}$ A66C (PM) with PcrV E76C (CPM). The emission spectra of PcrG$^{9-76}$ E20C, PcrG$^{9-76}$ R24C, PcrG$^{9-76}$ E63C or PcrG$^{9-76}$ A66C (PM), and PcrV E76C (CPM) complexes did not show a decrease in the donor peak intensities at 375 nm and 395 nm or an increase in the peak intensity of the acceptor at 460nm. The FRET data in Fig 5-8A-D shows that different constructs of PcrG$^{9-76}$ labeled with PM were far in distance upon complex formation with PcrV E76C labeled with CPM.

The relative intensities of the different PcrG$^{9-76}$ constructs E20C, R24C, E63C, or A66C labeled with the donor (PM) and PcrV construct V120C labeled with the acceptor (CPM) is shown in Fig 5-9. The relative intensity was determined by comparing the normalized spectra of the different PcrG$^{9-76}$ (PM)-PcrV (CPM) complexes. A normalized spectrum was obtained by dividing the intensities measured from the wavelength of 350 nm to 550 nm by the sum of all the intensities of the emission spectrum of the PcrG$^{9-76}$ (PM)-PcrG (CPM) complex.

The blue, red, green, and black plots depict the normalized curve from PcrG$^{9-76}$ E20C, PcrG$^{9-76}$ R24C, PcrG$^{9-76}$ E63C, and PcrG$^{9-76}$ A66C (PM) with two constructs of PcrV V120C (Fig 5-9). The normalized emission spectra PcrG$^{9-76}$ A66C (PM)-PcrV V120(CPM) had the lowest intensity for the donor maxima at 375 nm and 395 nm, and the highest intensity for the acceptor maxima at 460 nm, whereas the PcrG R24C (PM)-PcrV V120C (CPM) had the highest intensity for the donor maxima and the lowest intensity for the acceptor maxima, showing that the PcrG$^{9-76}$ A66C (PM) is the closest, and PcrG$^{9-76}$ R24C (PM) is the farthest from PcrV V120C (CPM). Fig 5-9B shows that PcrG$^{9-76}$ E20C, PcrG$^{9-76}$ R24C, and PcrG$^{9-76}$ A66C (PM) were relatively far apart from PcrV E76C (CPM) or the PcrG 9-76 (PM) constructs did not bind to PcrV E76C (CPM).

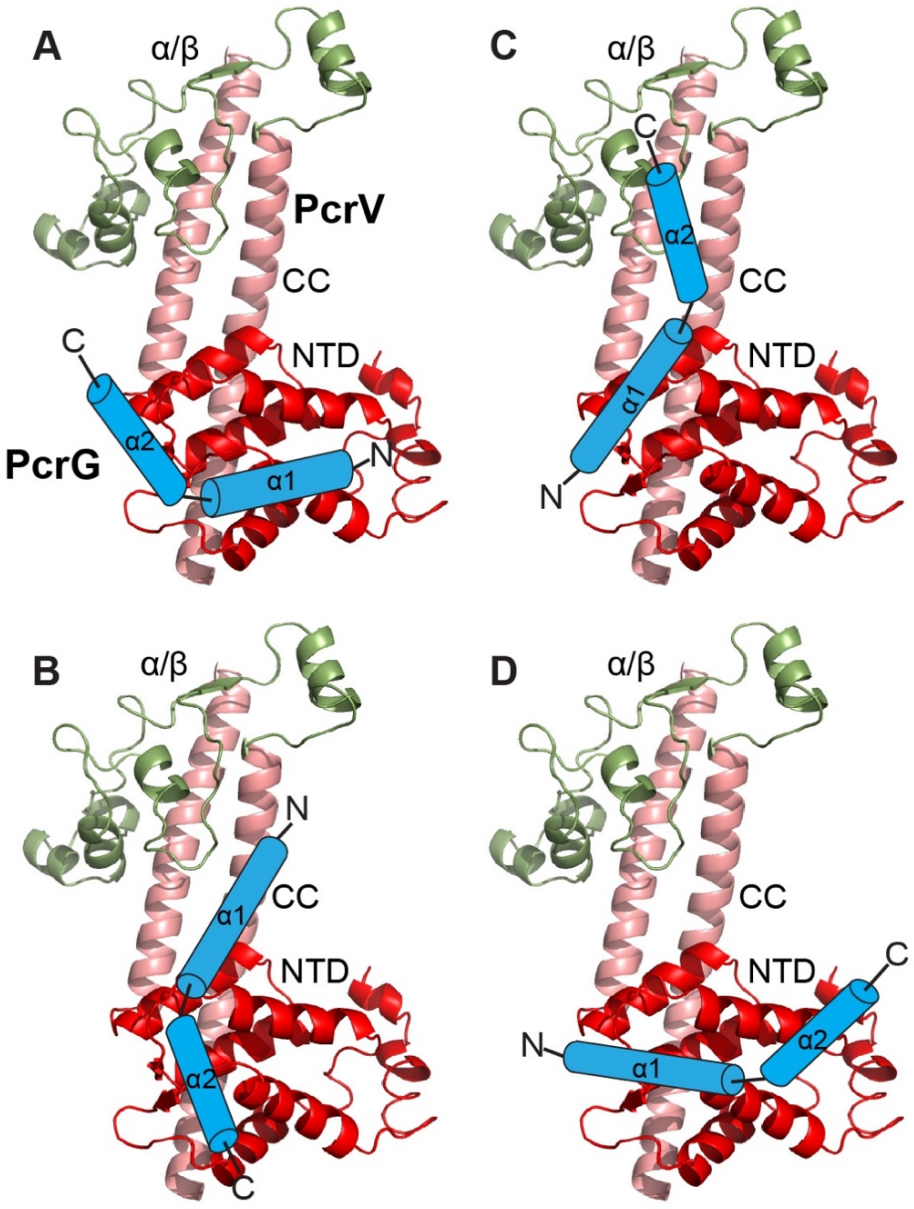

Figure 5-1. Proposed models of PcrG-PcrV interactions. (A) A proposed model showing interaction of N-terminal helix (helix α1) and C-terminal helix (helix α2) of PcrG with the N-terminal domain (NTD, shown in red) of PcrV. **(B)** A proposed interaction model of PcrG-PcrV showing that the N-terminal helix (helix α1) of PcrG interacts with the helices of the central coiled-coil (CC, shown in salmon) domain of PcrV whereas C-terminal (helix α2) interacts with helical region of N-terminal domain (NTD, red) of PcrV. **(C)** A Proposed model of PcrG-PcrV interaction where the C-terminal helix (helix α2) of PcrG interacts with central coiled domain (CC, salmon) and the N-terminal helix (helix α1) of PcrG interacts with the N-terminal domain (NTD, red) of PcrV. **(D)** A proposed model of PcrG-PcrV interaction where the N-terminal (helix α1) and C-terminal (helix α2) of PcrG interacts with the N-terminal domain (NTD, red) of PcrV. The orientation of helix α1 and helix α2 is reversed compared to **(A).**

The normalized emission spectra of $PcrG^{9-76}$ R24C (PM)-PcrV E76C (CPM) had the lowest

The normalized emission spectra of PcrG$^{9-76}$ R24C (PM)-PcrV E76C (CPM) had the lowest intensity at the donor maxima 375 nm and 395nm whereas PcrG$^{9-76}$ A66C (PM)-PcrV E76C (CPM) had the highest intensity at the donor maxima. The normalized spectra of PcrG$^{9-76}$ E20C (PM)-PcrV E76C (CPM), PcrG$^{9-76}$ R24C (PM)-PcrV E76C (CPM) and PcrG$^{9-76}$ E63C (PM)-PcrV E76C (CPM) overlay on the acceptor maxima. The PcrG$^{9-76}$ A66C (PM)-PcrV E76C (CPM) spectrum showed a decrease.

5.4.4 *PcrG$^{9-76}$A66C forms a complex with PcrV V120C with a k_d of 15±2 nM*

The FRET data from Fig. 5-2D shows that PcrG$^{9-76}$ A66C PM is close to PcrV V120C CPM. Fig 5-10 and 5-11 show the complex formation of PcrG$^{9-76}$ A66C PM -PcrV V120C CPM and the dissociation constant (k_d) for this interaction. Fig 5-10A shows emission spectra acquired from the PcrG$^{9-76}$ A66C PM -PcrV V120 CPM at decreasing concentrations of the donor labeled PcrG$^{9-76}$ A66C and acceptor labeled PcrV V120C ranging from 0.75 uM to 0.73 nM. The normalized spectra of different emission spectra obtained from Fig 5-5A is shown in Fig 5-10B, the intensity at donor maxima 375nm and 395nm was higher than the acceptor maxima at 460nm at concentrations from 23.43 nM to 0.73nM, showing less energy transfer between the donor and acceptor. This denoted most of the PcrG$^{9-76}$ A66C PM-PcrV V120C CPM complex had dissociated in solution. The complex dissociation is also verified by dividing the intensities of the PcrG$^{9-76}$ A66C PM —PcrV V120 CPM spectrum at different concentrations by their corresponding intensity at the donor maxima, 375 nm. Fig 5-11A shows that from the concentration of 23.43 nM to 0.73 nM, the intensity ratio at the acceptor maxima,460nm, is less than one. This decrease in the acceptor to donor intensity ratio showed that the PcrG$^{9-76}$ A66C PM and PcrV V120C CPM were monomeric and no longer forming a complex.

Table 5-1 Degree of labeling (DOL)

Position of fluorophore Label	Degree of Fluorophore Labeling (Moles of dye per molecule of protein)
$PcrG^{9-76}$	
E20C FM	1.03
R24C FM	1.2
PcrV	
V120C AF647	1.3
$LcrG^{7-73}$	
K28C FM	0.83
LcrV	
Q37C AF647	0.76

Table 5-2 Degree of labeling (DOL) of $PcrG^{9-76}$ and PcrV constructs

Position of fluorophore Label	Degree of Fluorophore Labeling (Moles of dye per molecule of protein)
$PcrG^{9-76}$	
E20C CPM	1.5
R24C CPM	1.2
E63C CPM	1.4
A66C CPM	1.5
PcrV	
E76C EMAL	2.3
V120C EMAL	3.4

Table 5-3 Degree of labeling (DOL) of $PcrG^{9-76}$ and PcrV constructs

Position of fluorophore Label	Degree of Fluorophore Labeling (Moles of dye per molecule of protein)
$PcrG^{9-76}$	
E20C PM	1.03
R24C PM	0.75
E63C PM	0.25
A66C PM	0.48
PcrV	
E76C CPM	0.93
V120C CPM	1.2

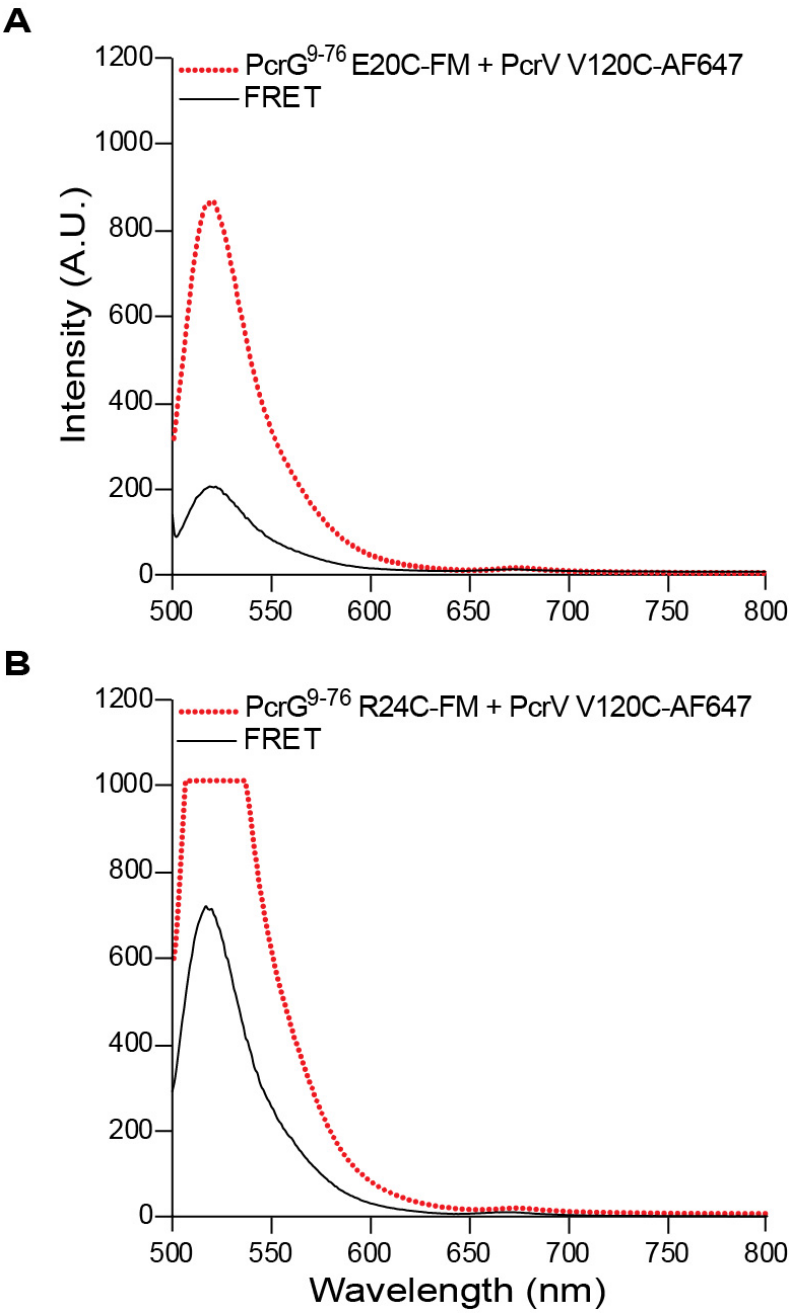

Figure 5-2 PcrG$^{9-76}$ FM constructs do not interact with PcrV V120C AF647. (A) The FRET data for PcrG$^{9-76}$ E20C FM and PcrV V120C AF647. The red dashed line is the control spectrum-sum of emission spectra of PcrG$^{9-76}$ E20C labeled with FM and PcrV V120C labeled with AF647. The black spectrum is the emission spectrum of the complex PcrG$^{9-76}$ E20C FM-PcrV V120C AF647. **(B)** The FRET data for PcrG$^{9-76}$ R24C FM and PcrV V120C AF647. The red dashed line is the control spectrum-sum of emission spectra of PcrG$^{9-76}$ R24C FM and PcrV V120C labeled with AF647. The black spectrum is the emission spectrum of the complex PcrG$^{9-76}$ R24C FM-PcrVV120C AF647.

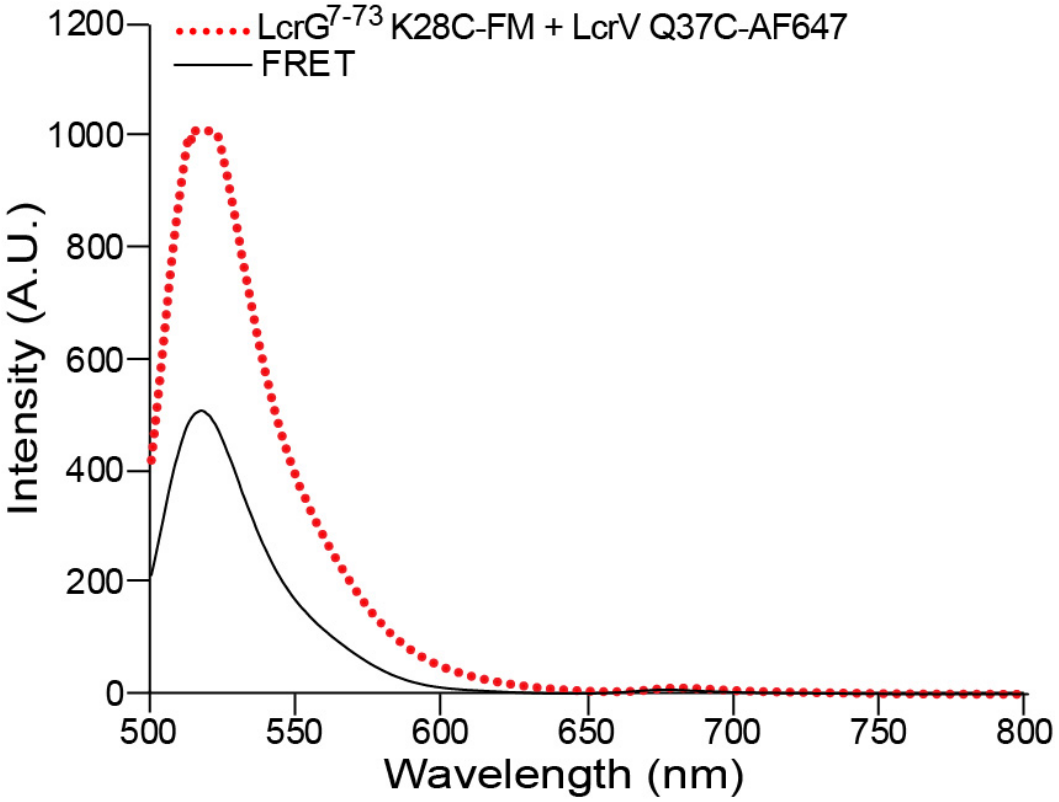

Figure 5-3 LcrG$^{7-73}$ K28C FM do not interact with LcrV Q37C AF647. The FRET data for LcrG$^{7-73}$ K28C FM and LcrV Q37C AF647. The red dashed line is the control spectrum-sum of emission spectra of LcrG$^{7-73}$ K28C FM and LcrV Q37C AF647. The black spectrum is the emission spectrum of the complex LcrG$^{7-73}$ K28C FM-LcrV Q37C AF647.

Table 5-4 FRET data for FM and AF647 FRET pair

Donor	**Acceptor**	**FRET**
$PcrG^{9-76}$ E20C FM	PcrV V120C AF647	Not observed
$PcrG^{9-76}$ R24C FM	PcrV V120C AF647	Not observed
$LcrG^{7-73}$ K28C FM	LcrV Q37C AF647	Not Observed

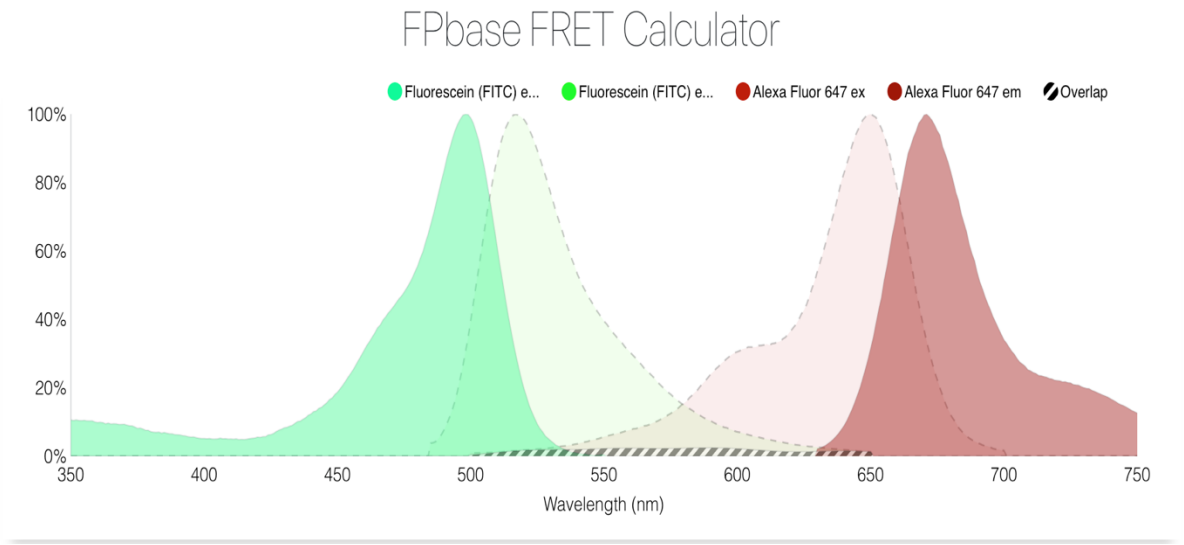

Figure 5-4 Fluorescein (FM) emission spectrum and Alexa Flour 647 (AF 647) excitation spectrum show little overlap. The solid green spectrum is excitation spectrum of Fluorescein (FM). The dashed green spectrum is emission spectrum of Fluorescein (FM). The dashed maroon spectrum is excitation spectrum of AF674. The solid maroon spectrum is emission spectrum of AF674. Spectra analyzed using FPbase FRET calculator[23].

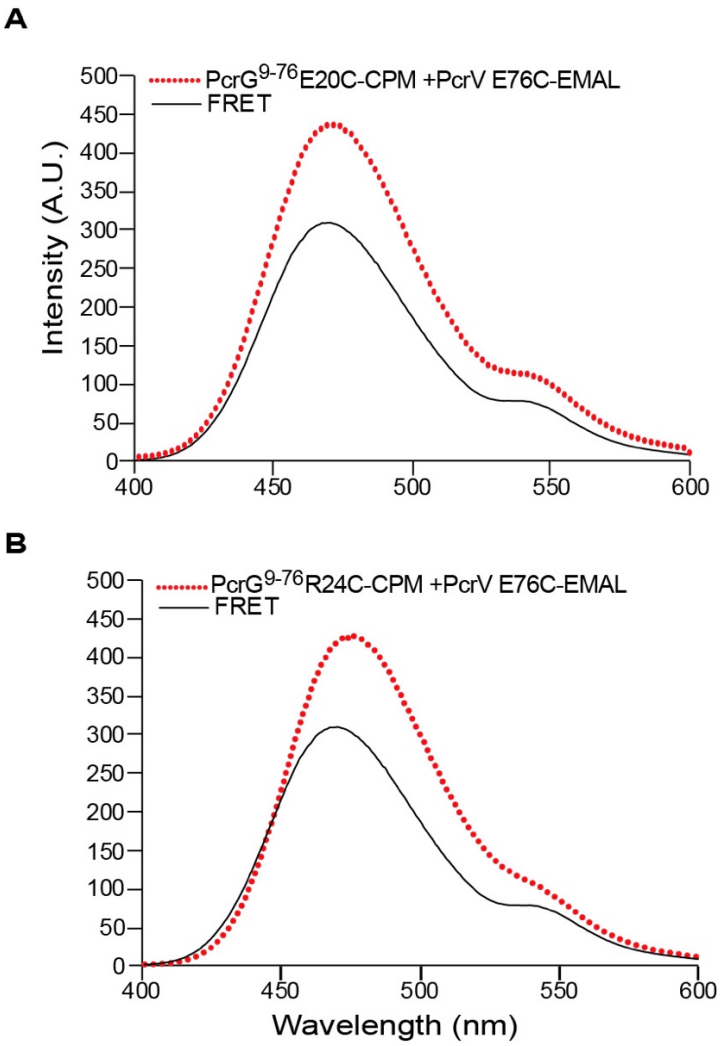

Figure 5-5 PcrG$^{9-76}$ CPM constructs do not interact with PcrV E76C EMAL. (A) The FRET data for PcrG$^{9-76}$ E20C CPM and PcrV E76C EMAL. The red dashed line is the control spectrum-sum of emission spectra of PcrG$^{9-76}$ E20C labeled with CPM and PcrV E76C labeled with EMAL. The black spectrum is the emission spectrum of the complex PcrG$^{9-76}$ E20C CPM-PcrV E76C EMAL. **(B)** The FRET data for PcrG$^{9-76}$ R24C FM and PcrV V120C AF647. The red dashed line is the control spectrum-sum of emission spectra of PcrG$^{9-76}$ R24C CPM and PcrV E76C labeled with EMAL. The black spectrum is the emission spectrum of the complex PcrG$^{9-76}$ R24C CPM-PcrV E76C AF647.

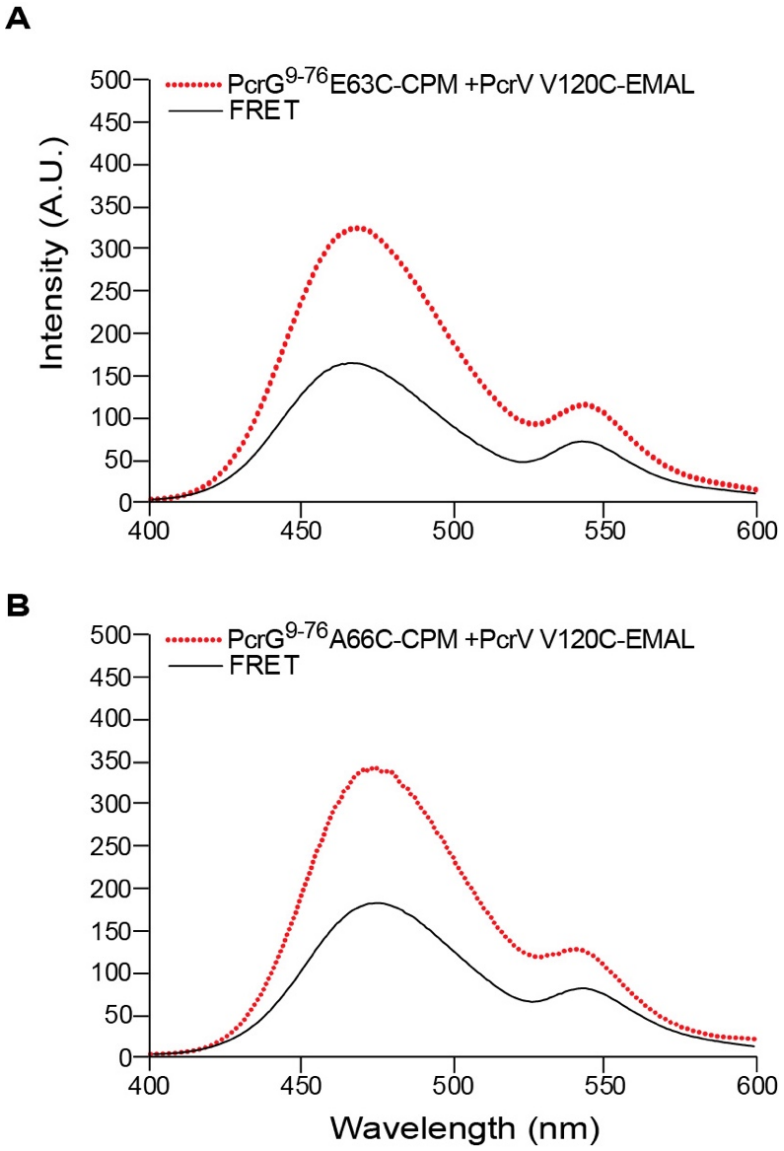

Figure 5-6 PcrG$^{9-76}$ CPM constructs do not interact with PcrV V120C EMAL. (A) The FRET data for PcrG$^{9-76}$ E63C FM and PcrV V120C EMAL. The red dashed line is the control spectrum-sum of emission spectra of PcrG$^{9-76}$ E20C labeled with CPM and PcrV V120C labeled with EMAL. The black spectrum is the emission spectrum of the complex PcrG$^{9-76}$ E20C CPM-PcrVV120C EMAL. **(B)** The FRET data for PcrG$^{9-76}$ A66C CPM and PcrV V120C EMAL. The red dashed line is the control spectrum-sum of emission spectra of PcrG$^{9-76}$ A66C CPM and PcrV V120C labeled with EMAL. The black spectrum is the emission spectrum of the complex PcrG$^{9-76}$ R24C CPM-PcrV V120C EMAL.

Table 5-5 FRET data for CPM and EMAL FRET pair

Donor	Acceptor	FRET
$PcrG^{9-76}$ E20C CPM	PcrV E76C EMAL	Not observed
$PcrG^{9-76}$ R24C CPM	PcrV E76C EMAL	Not observed
$PcrG^{9-76}$ E63C CPM	PcrV V120C EMAL	Not Observed
$PcrG^{9-76}$ A66C CPM	PcrV V120C EMAL	Not Observed

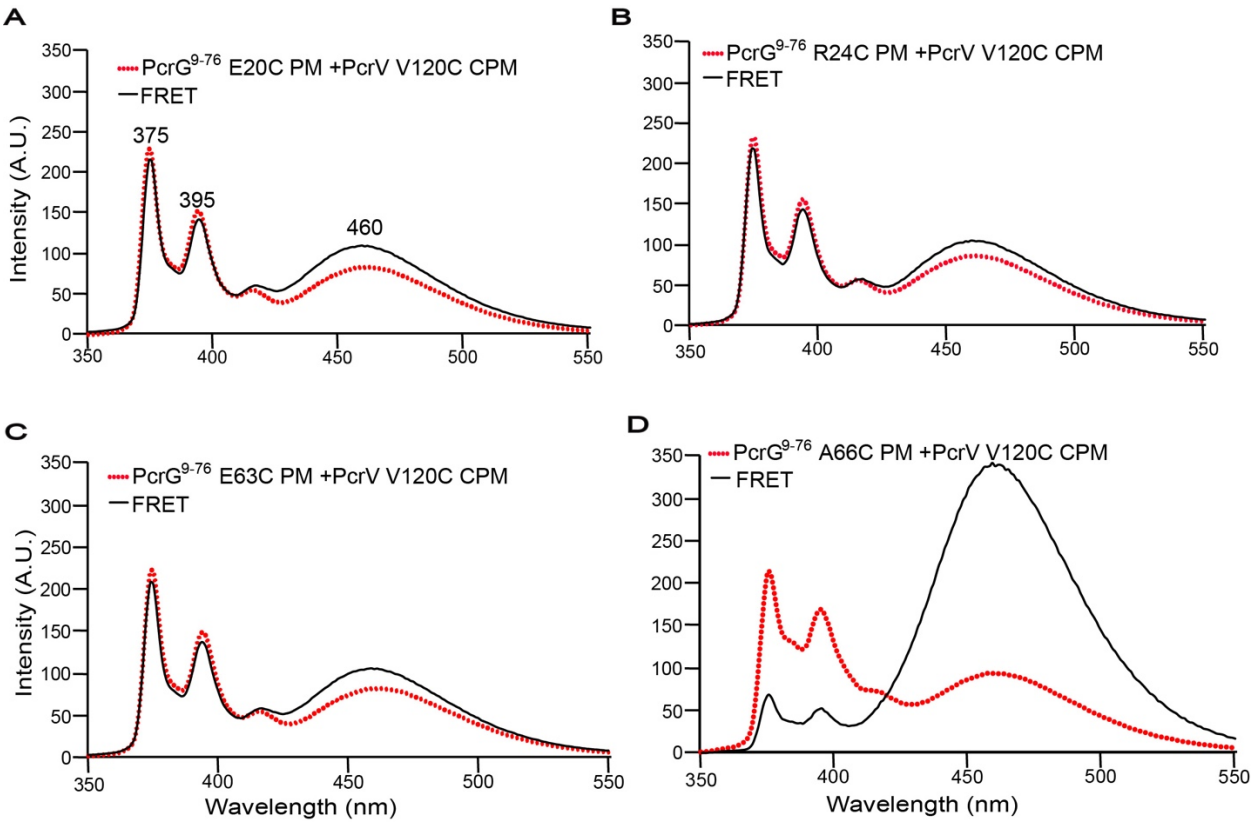

Figure 5-7. PcrG$^{9-76}$ A66C PM Interacts with PcrV V120C CPM. (A) The FRET data for PcrG$^{9-76}$ E20C PM and PcrV V120C CPM. **(B)** The FRET data for PcrG$^{9-76}$ R24C PM and PcrV V120C CPM. **(C)** The FRET data for PcrG$^{9-76}$ E63C PM and PcrV V120C CPM. **(D)** The FRET data for PcrG$^{9-76}$ A66C PM and PcrV V120C CPM.

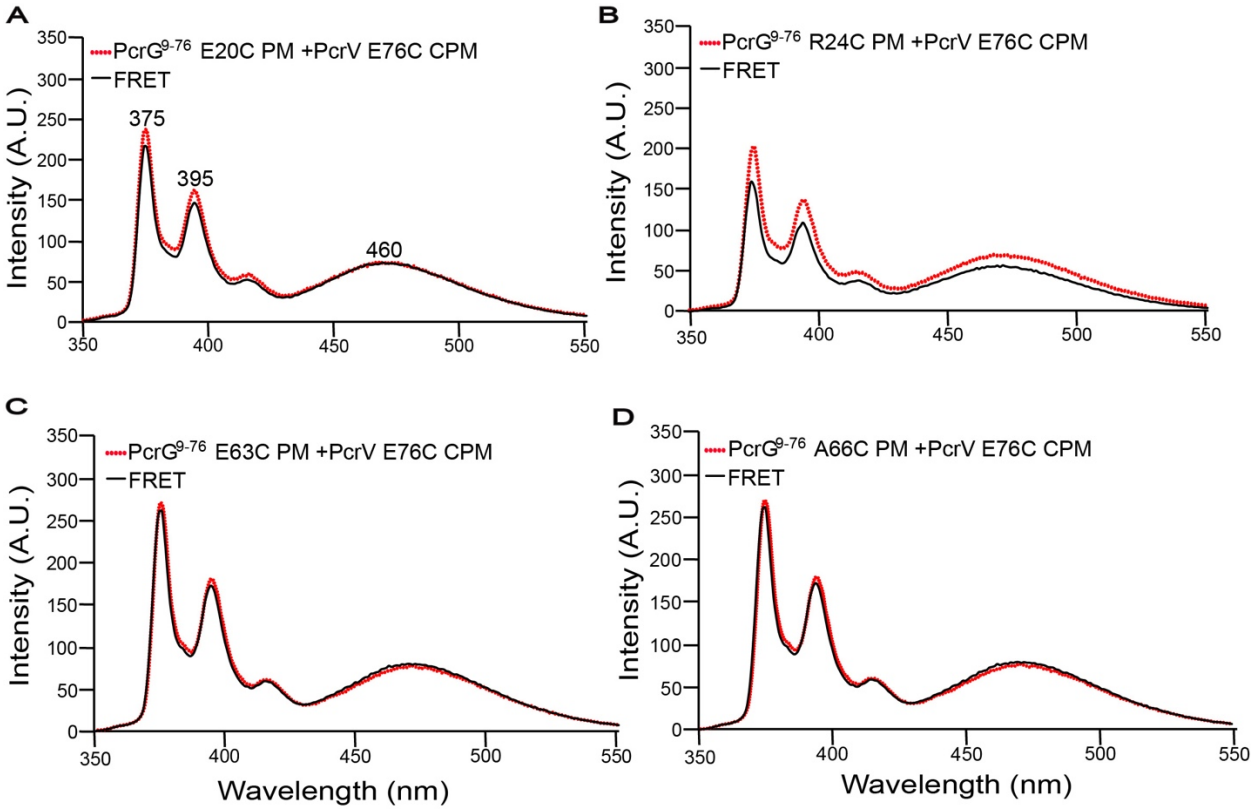

Figure 5-8. PcrG$^{9-76}$ PM constructs do not interact with PcrV E76C CPM (A) The FRET data for PcrG$^{9-76}$ E20C PM and PcrV E76C CPM. (B) The FRET data for PcrG$^{9-76}$ R24C PM and PcrV E76C CPM. (C) The FRET data for PcrG$^{9-76}$ E63C PM and PcrV E76C CPM. (D) The FRET data for PcrG$^{9-76}$ A66C PM and PcrV E76C CPM.

The acceptor to donor intensity ratio of the different $PcrG^{9-76}$ A66C PM-PcrV V120C CPM spectra was plotted against the concentration of PcrG or PcrV to determine the k_d of this interaction. The k_d of this interaction was determined to be 15±2 nM, as shown in Fig 5-11B by fitting the binding curve shown in the equation mentioned in the section 5.3.4.

5.5 Discussion

Our FRET data provides additional experimental evidence into the mode of binding interaction of the PcrG-PcrV complex. PcrG and PcrV are proteins of the T3SS required for the virulence of *Pseudomonas aeruginosa*[1]. PcrV, the tip protein forms a pentameric complex known the tip. The tip caps the intermembranous needle and provides a platform for the assembly of the translocon which forms a pore for the effector protein transport[24]. During the formation of the needle complex PcrV interacts with PcrG, the tip chaperone protein present in the bacterial cytosol[4]. PcrG also regulates effector protein transport by interacting with Pseudomonas T3SS proteins PscO and PcrD[15].

Currently there are no atomic structures of the tip protein and tip chaperone protein complex. Computational studies have predicted that PcrG has a coiled-coil conformation when bound to PcrV[25]. Studies by other groups[17] have predicted that the *Yersinia* tip chaperone, LcrG has an alpha hairpin structure. This hypothesis compares the crystal structures of tip proteins — *Salmonella* SipD, *Shigella* IpaD, and *Burkholderia* BipD with the crystal structure of *Yersinia* LcrV[17]The crystal structure of LcrV lacks an N-terminal alpha hairpin structure that functions as a self-chaperone for tip proteins — SipD, IpaD, and BipD[26],[27],[28]. The N-terminal alpha helical hairpin interacts with the central coiled-coil domain of the tip proteins. Based on this comparison, LcrG is hypothesized to form an alpha helical hairpin structure when bound to

LcrV. The structural and functional similarity of *Yersinia* T3SS proteins with *Pseudomonas* T3SS proteins extends this hypothesis to the PcrG-PcrV complex.

Results of EPR spectroscopy suggested that the N-terminal and C-terminal helices of LcrG are at a distance of more than 25 Å from each other, suggesting that LcrG has an open conformation when bound to LcrV (Fig 2-6 ,Chapter 2)[22]. The EPR data for the interaction of spin labeled LcrG and LcrV also shows a model where LcrG interacts uniquely to the different regions of the N-terminal domain of LcrV, further confirming the open conformation of LcrG when bound to LcrV[22](Fig 3-9, Chapter 3).

Our previous NMR data has shown that the overall structure of *Pseudomonas* PcrG undergoes a global change upon binding with PcrV[16]. Previous SPR and yeast-two hybrid data also show that PcrG interacts with PcrV with a molar ratio of 1:1[13],[16]. Recent findings from our EPR data (Fig.4-5, Chapter 4) show that as the N-terminal and C-terminal of PcrG have a spin-spin distance greater than 25 Å when bound to PcrV suggesting that PcrG has an open conformation when bound to PcrV.

I proposed several models of PcrG-PcrV interaction (Fig 5-1A-D) and these models were verified by FRET. The cysteine mutations designed for the FRET studies were guided by the EPR data of LcrG-LcrV interaction (Fig 3-9 Chapter 3), NMR data of PcrG[16], and EPR data showing the different conformations PcrG (Fig.4-4and Fig 4-5, Chapter 4). The PcrG $^{9\text{-}76}$ construct, which consists of the helical regions of PcrG is selected for designing cysteine mutations for donor (PM) fluorophore attachment in the N-terminal helix (helix α1) — PcrG$^{9\text{-}76}$ E20C and PcrG$^{9\text{-}76}$ R24C, and the C-terminal helix (helix α2) — PcrG$^{9\text{-}76}$ E63C and PcrG$^{9\text{-}76}$ A66C (Fig 5-12).

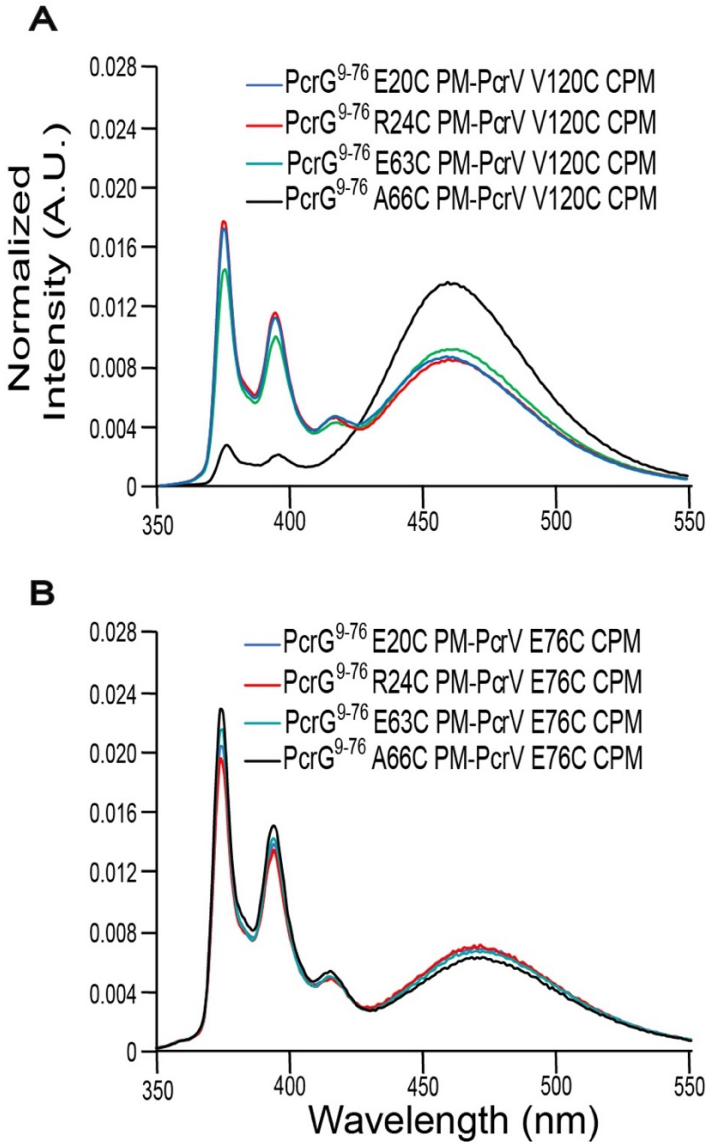

Figure 5-9. PcrG$^{9-76}$ A66C PM is closer to PcrV V120C CPM (A) Relative distance of PcrG$^{9-76}$ constructs from PcrV V120C. (B) Relative distance of PcrG$^{9-76}$ constructs from PcrV E76C.

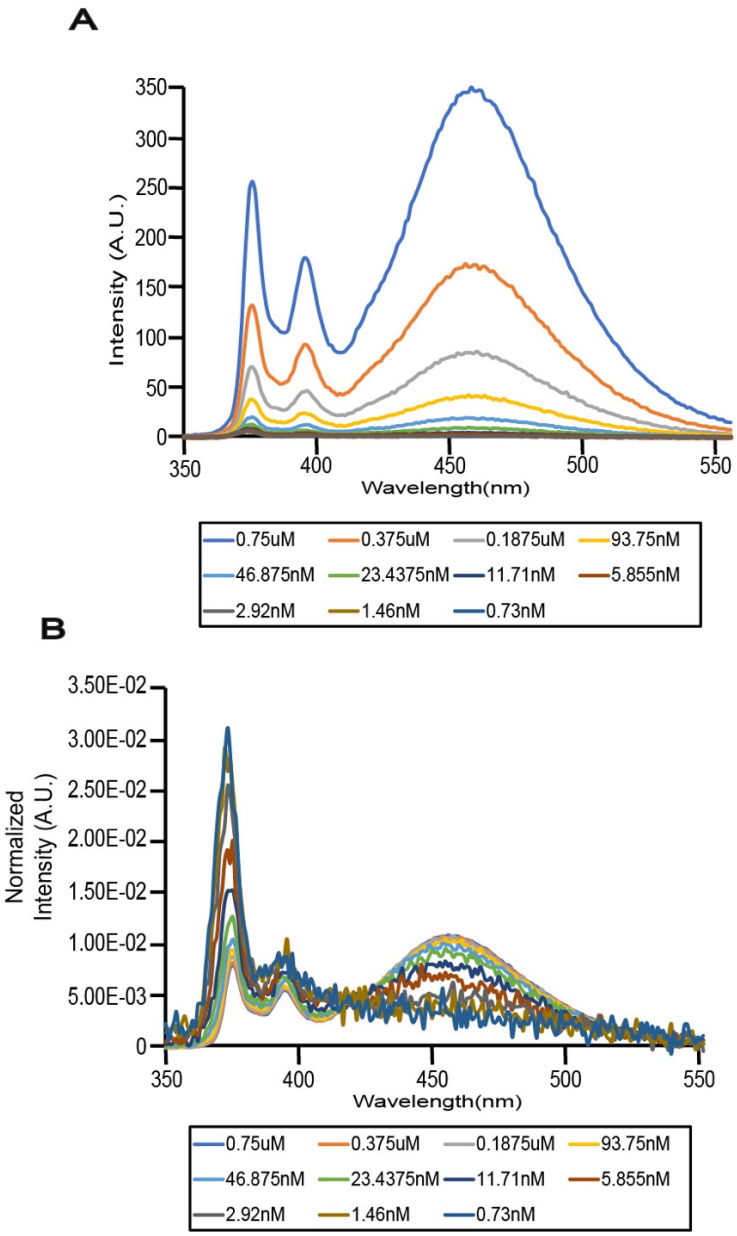

Figure 5-10. PcrG$^{9-76}$ A66C PM binds to PcrV V120C CPM. (A) FRET data of PcrG$^{9-76}$ A66C PM and PcrV V120C CPM acquired at decreasing concentrations donor and acceptor to determine the concentration at which the complex dissociates. **(B)** Normalized curve of PcrG$^{9-76}$ A66C PM and PcrV V120C CPM interaction.

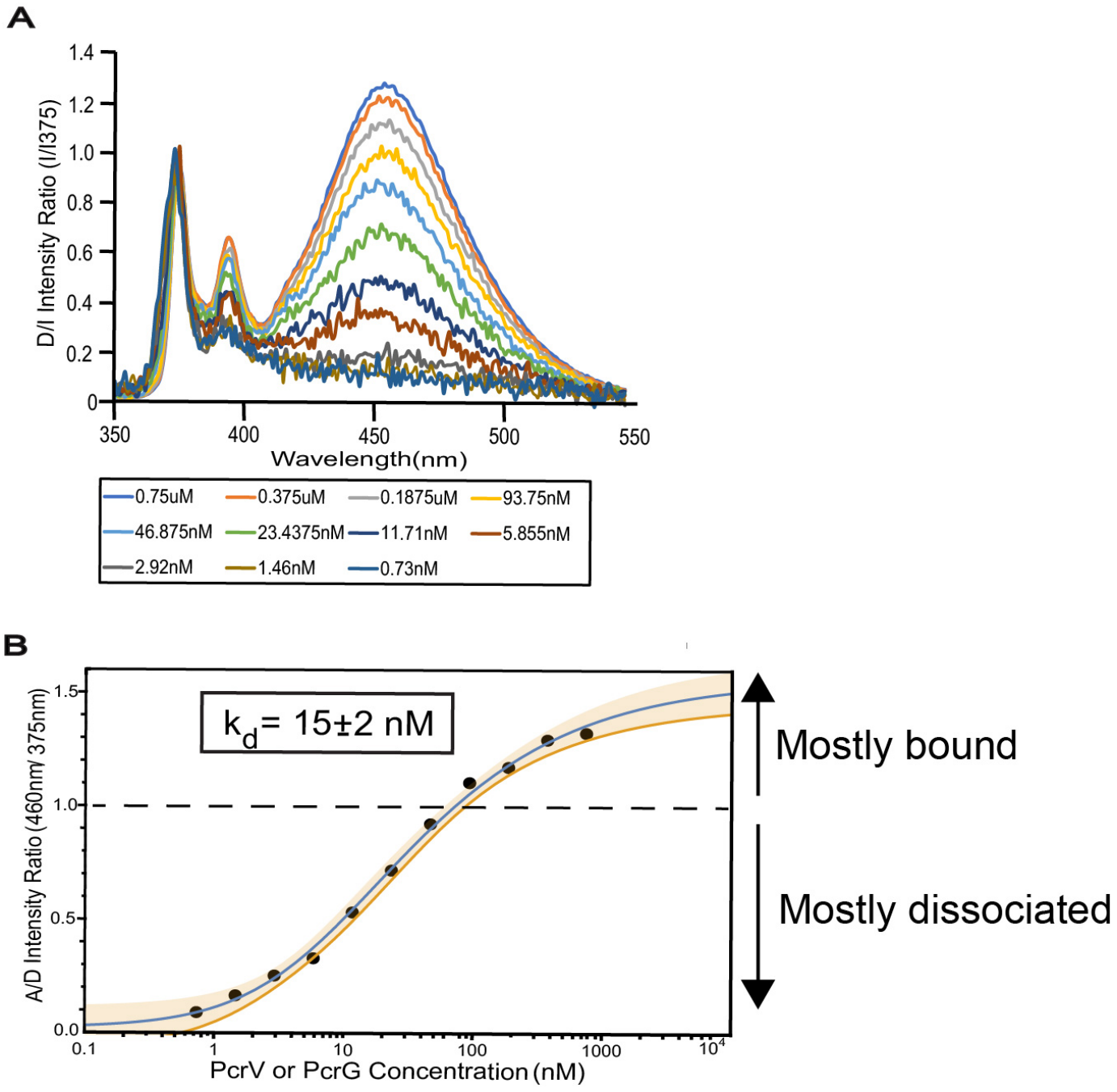

Figure 5-11. PcrG$^{9-76}$ A66C PM forms a complex with PcrV V120C. (A) Normalized curve of PcrG$^{9-76}$ A66C PM and PcrV V120C CPM with donor maxima =1 **(B)** Data fitting of PcrG$^{9-76}$ A66C and PcrV V120C shows binding with k_d= 15± 2 nM.

The two residues E76 and V120 selected for cysteine mutations in the N-terminal domain of PcrV with a Cα-Cα distance of 22 Å, which would separate the residues for studying the interaction between PcrG and PcrV as shown in Fig 5-13B.

The FRET spectra obtained from a donor (PM) labeled $PcrG^{9-76}$ constructs and acceptor (CPM) labeled PcrV constructs show that the complex formed by the donor (PM) labeled $PcrG^{9-76}$ A66C and acceptor (CPM) labeled PcrV V120C showed energy transfer. The donor maxima at 375 nm and 395 nm in the black FRET spectrum shows a decrease in intensity compared to the control spectrum (red, Fig 5-8D). In contrast, the intensity of the acceptor maxima increases in the FRET spectrum (black, Fig 5-8D) compared to the red control spectrum. This energy transfer confirms the proximity of the donor and acceptor in the PcrG-PcrV complex, as shown by the cartoon model in Fig 5-13B. This interaction has a dissociation constant (k_d) of 15± 2 nM (Fig 5-11B).

The FRET spectra acquired from the seven remaining complexes do not show any change in energy transfer compared to the control (Fig 5-7 A-C, Fig 5-8 A-D). This data suggests that when the PcrG-PcrV complex is formed, the $PcrG^{9-76}$ donor (PM) and PcrV acceptor (CPM) labeled residues are far apart, or the fluorophore attachment to PcrG and PcrV disrupted complex formation.

In summary, my FRET provides additional insight into the binding interaction of the PcrG-PcrV complex.

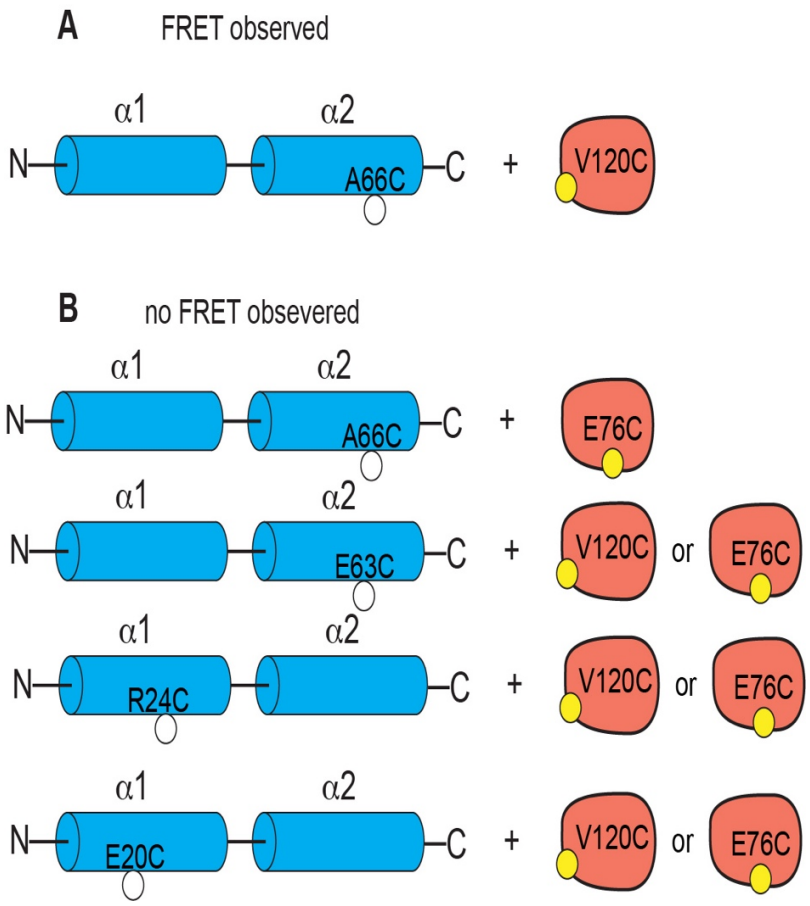

Figure 5-12. Summary of PcrG$^{9-76}$ and PcrV proteins for the FRET studies. (A) PcrG$^{9-76}$ A66C PM and PcrV V120C CPM show strong energy transfer. **(B)** PcrG$^{9-76}$ donor (PM) and PcrV acceptor (CPM) constructs showing no energy transfer. PcrG$^{9-76}$ construct depicted with two helices α1 and α2 as cylinders (blue) labeled with donor PM, shown as white spheres. PcrV is depicted as an oblong circle (salmon) labeled with acceptor CPM shown as yellow spheres.

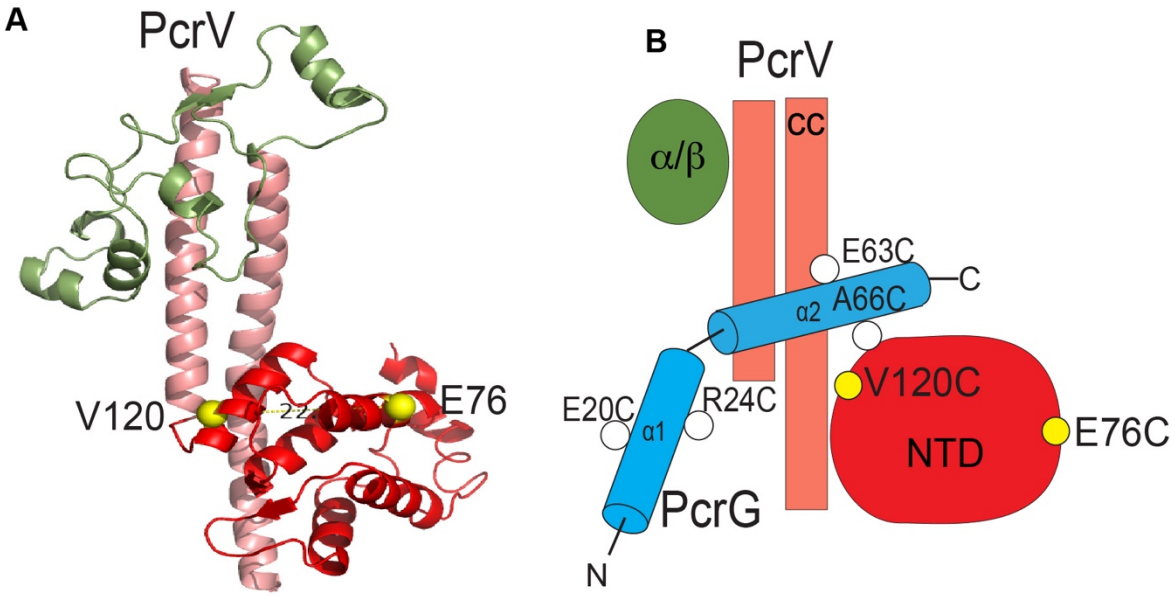

Figure 5-13. Final model of PcrG-PcrV interaction derived from EPR and FRET data. I-TASSER model of PcrV containing the N-terminal domain (NTD, red), α/β mixed region (green) and central coiled-coil domain(salmon). **(A)** Model of PcrV with a Cα-Cα distance between E76 and V120 calculated as 22 Å. **(B)** Cartoon of $PcrG^{9-76}$ A66C PM-PcrV V12C CPM interaction showing that helix α2 of $PcrG^{9-76}$ interacts with N-terminal domain (NTD) (red) domain of PcrV).

5.6 Reference

1. Lyczak, J.B., C.L. Cannon, and G.B. Pier, *Establishment of Pseudomonas aeruginosa infection: lessons from a versatile opportunist.* Microbes Infect., 2000. **2**(9): p. 1051-60.
2. Engel, J. and P. Balachandran, *Role of Pseudomonas aeruginosa type III effectors in disease.* Curr. Opin. Microbiol., 2009. **12**(1): p. 61-6.
3. Rosenfeld, M., B.W. Ramsey, and R.L. Gibson, *Pseudomonas acquisition in young patients with cystic fibrosis: pathophysiology, diagnosis, and management.* Curr.Opin.Pulm.Med., 2003. **9**(6): p. 492-497.
4. Galan, J.E., et al., *Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells.* Annu. Rev. Microbiol., 2014. **68**: p. 415-38.
5. Dey, S., et al., *The type III secretion system needle, tip, and translocon.* Protein Sci, 2019. **28**(9): p. 1582-1593.
6. Marlovits, T.C., et al., *Structural insights into the assembly of the type III secretion needle complex.* Science, 2004. **306**(5698): p. 1040-1042.
7. Veenendaal, A.K., et al., *The type III secretion system needle tip complex mediates host cell sensing and translocon insertion.* Mol. Microbiol., 2007. **63**(6): p. 1719-30.
8. Matson, J.S. and M.L. Nilles, *LcrG-LcrV interaction is required for control of Yops secretion in Yersinia pestis.* J. Bacteriol., 2001. **183**(17): p. 5082-91.
9. Mueller, C.A., et al., *The V-antigen of Yersinia forms a distinct structure at the tip of injectisome needles.* Science, 2005. **310**(5748): p. 674.
10. Cornelis, G.R. and F. Van Gijsegem, *Assembly and function of type III secretory systems.* Annu. Rev. Microbiol., 2000. **54**: p. 735-774.
11. Zhang, Y., A.K. Arakaki, and J. Skolnick, *TASSER: an automated method for the prediction of protein tertiary structures in CASP6.* Proteins, 2005. **61 Suppl 7**: p. 91-98.
12. Chaudhury, S., et al., *Structure of the Yersinia pestis tip protein LcrV refined to 1.65 Å resolution.* Acta Crystallogr. Sect. F, 2013. **69**(Pt 5): p. 477-81.
13. Nanao, M., et al., *Type III secretion proteins PcrV and PcrG from Pseudomonas aeruginosa form a 1:1 complex through high affinity interactions.* BMC Microbiol., 2003. **3**: p. 21.
14. Allmond, L.R., et al., *Protein binding between PcrG-PcrV and PcrH-PopB/PopD encoded by the pcrGVH-popBD operon of the Pseudomonas aeruginosa type III secretion system.* Infect. Immun., 2003. **71**(4): p. 2230-3.
15. Lee, P.C., et al., *Control of effector export by the Pseudomonas aeruginosa type III secretion proteins PcrG and PcrV.* Mol. Microbiol., 2010. **75**(4): p. 924-41.
16. Chaudhury, S., et al., *Nuclear Magnetic Resonance Characterization of the Type III Secretion System Tip Chaperone Protein PcrG of Pseudomonas aeruginosa.* Biochemistry, 2015. **54**(43): p. 6576-85.
17. Blocker, A.J., et al., *What's the point of the type III secretion system needle?* Proc. Natl. Acad. Sci. U.S.A., 2008. **105**(18): p. 6507-13.
18. H Hamouda, R.M., Marwa F, *Rapid Detecttion of Bovine Herpes Virus Type 1 (Bohv-1) in Egyptian Dromadary Camels.* Appl. Vet. Sci. , 2020.
19. Fersht, A., *Structure and Mechanism in Protein Science:A Guide to Enzyme Catalysis and Protein Folding.* 1999.

20. Reed, *L. receptor-ligand dissociation* people.reed.edu. (n.d.), 2015.
21. McShan, A.C., *Elucidation of the Needle-Tip and Tip -Translocon Interactons of the Salmonella SP-1 Type III Secretion System and Identification of Small Molecule Binders of the Tip and Translocon Protiens. PhD. Diss. University of Kansas.* 2016.
22. Pallavi Guha Biswas, Z.H., Pavanjeet Kaur, Kawaljit Kaur, Sukanya Chaudhury and M.W. Muhammed Ciftci, Justin Douglas, Likai Song, and Roberto N. De Guzman, *Model of Yersinia LcrG and LcrV Interaction Defined by EPR Spectroscopy.*
23. Lambert, T., *FPbase: a Community-Editable Fluorescent Protein Database.* Nature Methods, 2019. **16**: p. 277–278.
24. Mueller, C.A., P. Broz, and G.R. Cornelis, *The type III secretion system tip complex and translocon.* Mol. Microbiol., 2008. **68**(5): p. 1085-95.
25. Basu, A., et al., *PcrG protects the two long helical oligomerization domains of PcrV, by an interaction mediated by the intramolecular coiled-coil region of PcrG.* BMC Struct Biol, 2014. **14**: p. 5.
26. Espina, M., et al., *IpaD localizes to the tip of the type III secretion system needle of Shigella flexneri.* Infect. Immun., 2006. **74**(8): p. 4391-4400.
27. Johnson, S., et al., *Self-chaperoning of the type III secretion system needle tip proteins IpaD and BipD.* J. Biol. Chem., 2007. **282**(6): p. 4035-4044.
28. Chatterjee, S., et al., *The crystal structure of the Salmonella type III secretion system tip protein SipD in complex with deoxycholate and chenodeoxycholate.* Protein Sci., 2011. **20**: p. 75-86.

Chapter 6: Small Molecule interaction with *Yersinia* tip protein LcrV determined by NMR Spectroscopy

6.1 Abstract

The rise in antibiotic resistance among different bacterial species requires the discovery of novel antimicrobials. *Yersinia pestis* is a Gram-negative bacterium, which causes bubonic plague. *Yersinia* uses the T3SS to inject virulence effector proteins into its target host cells. The *Yersinia* tip protein LcrV is essential for virulence and can be a potential target for developing antimicrobials. For antimicrobial development, it is important to identify small molecules that can bind to LcrV. The screening of two libraries of 544 fragments using surface plasmon resonance (SPR) identified 3 compounds that could potentially bind to LcrV. The three compounds are the following — 2-(benzylsulfinyl) benzoic acid (BSBA), (3aR,7aS)-6-(6-ethoxypyridin-3-yl)-2,2-dimethyl-3a,4,5,7a-tetrahydro-2H-1,3-benzodioxol-5-one (EDTBO) and 2-(pyridin-2-yl) ethan-1-ol (PE). I used Nuclear magnetic resonance (NMR) spectroscopy to identify the binding regions of BSBA in LcrV. This study identifies a new small molecule which binds to the *Yersinia* tip protein LcrV.

6.2 Introduction

Yersinia pestis is a pathogenic bacterium that is the main causative agent of bubonic plague[1]. Gram-negative bacteria such as *Salmonella, Shigella, Burkholderia ,Yersinia,* and *Pseudomonas* use the T3SS to assemble a nanoinjector for the transport of virulence effector proteins that cause infection in the host cell[2],[3]. Antibacterial resistance among pathogens[4],[5] has necessitated the need for the developing new alternate microbial targets. The T3SS is an attractive target for developing new antimicrobials because of its essential role in virulence, its exposure to the bacterial surface, and its presence only among pathogens[6].

The T3SS is used to inject virulence effector proteins into the host cells to manipulate the host signaling pathways for the survival and propagation of the bacteria.[2],[6]. The structural component of the T3SS is an assemblage of over 20 different proteins known as the needle complex[6], which comprises the base, needle, tip complex, and translocon. The base is anchored to the bacterial inner and outer membrane[3].The needle is assembled in the intermembranous space between the bacteria and the host cell[2]. The needle is capped by the tip complex, which provides the platform for the assembly of the translocon proteins[2]. The translocon proteins assemble to form a pore in the host cell membrane for the transport of virulent effector to the host cells for infection[2]. In *Yersinia,* the tip complex is formed by the assemblage of multiple copies of the tip protein, LcrV[7]. When the bacteria come in contact with the host cells, the tip complex acts as an environmental sensor for the deployment of the translocon proteins to form a pore for the passage of virulent effector proteins to the host cell for infection[2],[6]. The crystal structure of LcrV consists of an N-terminal domain, a mixed α/β region, and a central coiled domain[8].

LcrV is a potential candidate for active and passive immunization against V-proteins for providing high level protection against the T3SS mediated pathogen *Yersinia pestis* infections[9],[10]. New small molecules have also been discovered which inhibit the virulence of *Yersinia pestis*. Although these compounds inhibit the T3SS of *Yersinia,* their specific targets and mechanism of binding remain unknown[11],[12],[13],[14]. In collaboration with the Center of Biomedical Research Excellence in Protein Structure and Function at the University of Kansas, my collaborator, Anne Cooper used surface plasmon resonance (SPR) based fragment screening using the Aube 440 and Aube 104 small molecule libraries to identify new small molecules which bind to LcrV. The interaction surface of the small molecule BSBA on LcrV

was determined by NMR spectroscopy. NMR shows the ligand binding surfaces of LcrV, which could be potential targets for new developing new antimicrobials.

6.3 Methods

6.3.1 Expression and purification of LcrV

The cloning, expression and purification of LcrV (residues 28-322) has been previously described[8]. The plasmid containing LcrV was transformed in *E. coli* BL21(DE3) DNAY and grown in a culture medium supplemented with carbenicillin and kanamycin. Unlabeled LcrV was grown in 1L TB for SPR screening. Protein for NMR spectroscopy were simultaneously labeled with ^{15}N and ILV (the methyl groups isoleucine, leucine and valine) by growing cells in M9 minimal media supplemented with 3g/L of glucose and 1g/L ^{15}N-ammonium chloride (Sigma)[6]. Cells were grown at 37 ºC and 200 rpm in a shaker incubator till O.D.$_{600}$ ~0.4 at which point the media was supplemented with 60 mg/L of 2-ketobutyricacid-4-^{13}C sodium salt hydrate (#571342, Sigma) to label the Cδ1 of isoleucine and 100 mg/L of 2-keto-3-(methyl-^{13}C)-butyric-4-^{13}C acid sodium salt (#571334, Sigma) to label Cδ1 of leucine and Cψ valine methyl groups. At O.D.$_{600}$ 0.8 the cells were induced with 1mM isopropyl-β-D-thiogalactopyranoside (IPTG) and cell growth was continued overnight at 15 ºC to a final O.D.$_{600}$ ~2.8-3.0.
Cells were harvested by centrifugation (4000 rpm, 2392×*g*, 12 min, 4 ºC). The cell pellet was resuspended in 40 mL binding buffer (20mM Tris pH 8.0, 500 mM NaCl, 5mM imidazole) with 600 µL 10 mM PMSF. The resuspended cells were sonicated for a total of 5 min 20 sec using a sonication cycle of 2 s pulse on and 30 s pulse off. The cell lysate was centrifuged at 13,000 rpm (13865×*g* in a Beckman JA-25.50 rotor) for 15 min at 4ºC. To the supernatant was added to 700

µl of 5% PEI to precipitate nucleic acids, followed by centrifugation at 13,000 rpm (13865×*g*) for 15 min at 4ºC. The supernatant was loaded into a 5 mL nickel column, followed by three washes using 50 mL binding buffer. The protein eluted in 50 mL elution buffer (20 mM Tris pH 8.0, 500 mM NaCl, 250 mM imidazole). TEV protease (250 uL of 0.07 mM stock) was added, and the sample was dialyzed at room temperature in 1 L TEV buffer (50 mM Tris pH 8.0, 0.5 mM EDTA, 1 mM DTT, 100 mM NaCl) overnight, followed by further dialysis in 1L binding buffer for 4 hrs and in 2 L binding buffer overnight. The digest was loaded on a 10 ml nickel column, the flow through was collected and followed by wash 1 in 50 ml binding buffer, wash 2 with 50 ml (20mM Tris pH 8.0, 500 mM NaCl, 25 mM Imidazole) and wash 3 (20mM Tris pH 8.0, 500 mM NaCl, 50 mM Imidazole). LcrV appeared in the fractions flow through, wash 1 and wash 2, whereas the His_6-GB1 tag appeared in the elution fraction.

The purified proteins were dialyzed in 1L 1X Phosphate-Saline Buffer (PBS) pH 7.4 for SPR screening or in 1L NMR Buffer (20mM NaCl,10mM $NaPO_4$) for NMR spectroscopy. The proteins were concentrated using Amicon Ultra 10X centrifugal filter (Millipore) and the concentration at was determined by A_{280}[8].

6.3.2 *SPR screening of Aube 440 and Aube 104 small molecule libraries*

The surface resonance plasmon (SPR) screening of small molecule libraries Aube 440 and Aube110 was done in collaboration with the Center of Biomedical Research Excellence in Protein Structure and Function at the University of Kansas as described previously[6]. LcrV was covalently immobilized to the CM5 sensor chip surface (GE Healthcare, #BR100399) by standard amine coupling chemistry using the amine coupling kit (GE Healthcare #BR-1106-33) in the running buffer 1X PBS. The small molecule fragments were injected for 60 sec at a flow

rate of 60 μL/min followed by a dissociation period of an additional 60 sec. The unbound fragments were washed with a 1:1 solution of DMSO-water[15].

6.3.3 *NMR spectroscopy*

NMR data was acquired at 30 ^{o}C using a Bruker Avance 800 MHz spectrophotometer at equipped with a cryogenic triple probe, processed using NMRPipe[16] and analyzed using NMRView[17].Typical 2D ^{1}H–^{15}N TROSY acquisition parameters were 16 scans at 30 ppm ^{15}N sweep width centered at 118 ppm, whereas typical 2D ^{1}H–^{13}C HSQC acquisition parameters for ILV-labeled samples were 8 scans, 18 ppm ^{13}C sweep width centered at 18 ppm and 10 ppm ^{1}H sweep width centered at 4.69 ppm[6].

For NMR titrations, the compound was dissolved in 100% d_6-DMSO from Cambridge Isotope Laboratories (CIL). The 2D ^{1}H-^{13}C HSQC and ^{1}H-^{15}N TROSY spectra were acquired on 0.3 mM of ^{15}N/ILV-labeled LcrV titrated with increasing molar concentrations of the compound for mapping the ILV and ^{15}N chemical shifts. All the titrated samples contain 2% (v/v) d_6-DMSO and 10% D_2O. The chemical shift deviations were calculated for ILV titrations using this formula:

$0.5[\Delta\delta_H^2 + (\Delta\delta_C^2/4]^{1/2}$[18].

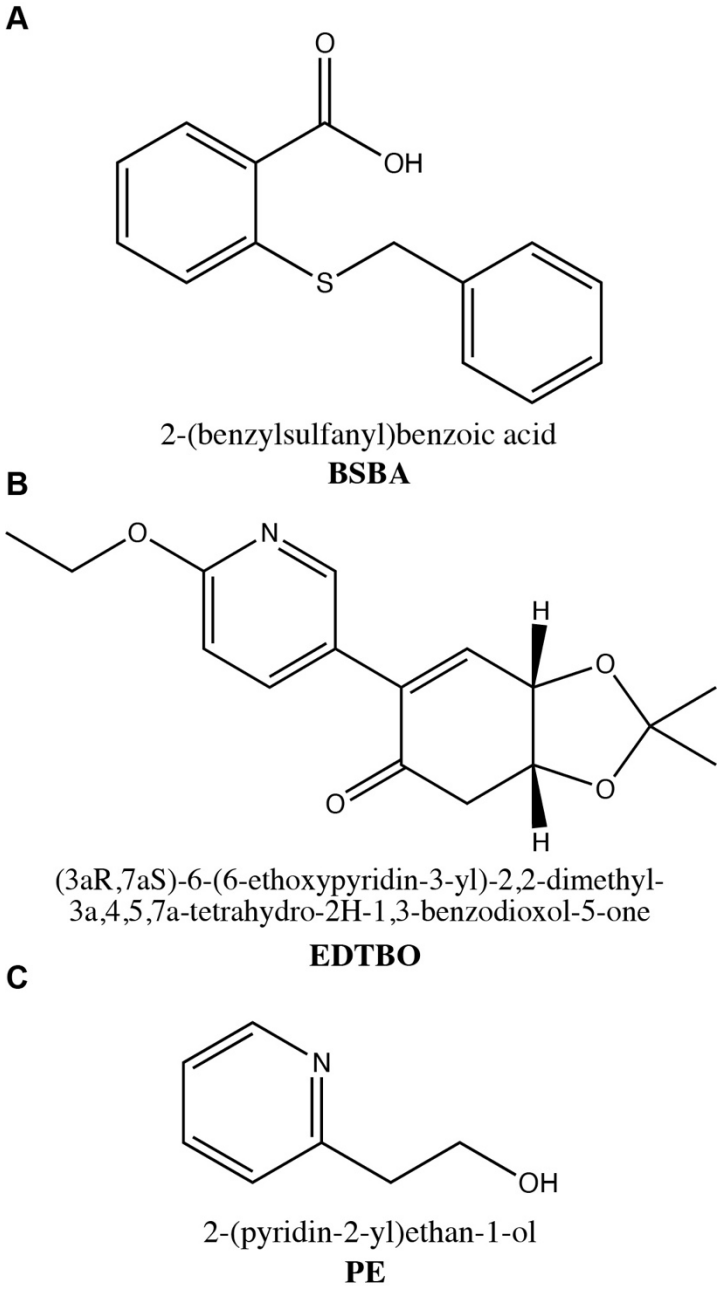

Figure 6-1. Structures of the fragment hits from SPR screening. SPR screening identified that **(A)** 2-(benzylsulfinyl) benzoic acid **(BSBA), (B)** (3aR,7aS)-6-(6-ethoxypyridin-3-yl)-2,2-dimethyl-3a,4,5,7a-tetrahydro-2H-1,3-benzodioxol-5-one **(EDTBO)**, and **(C)** 2-(pyridin-2-yl) ethan-1-ol **(PE)** bound to LcrV. Structures were draw using ChemDraw.

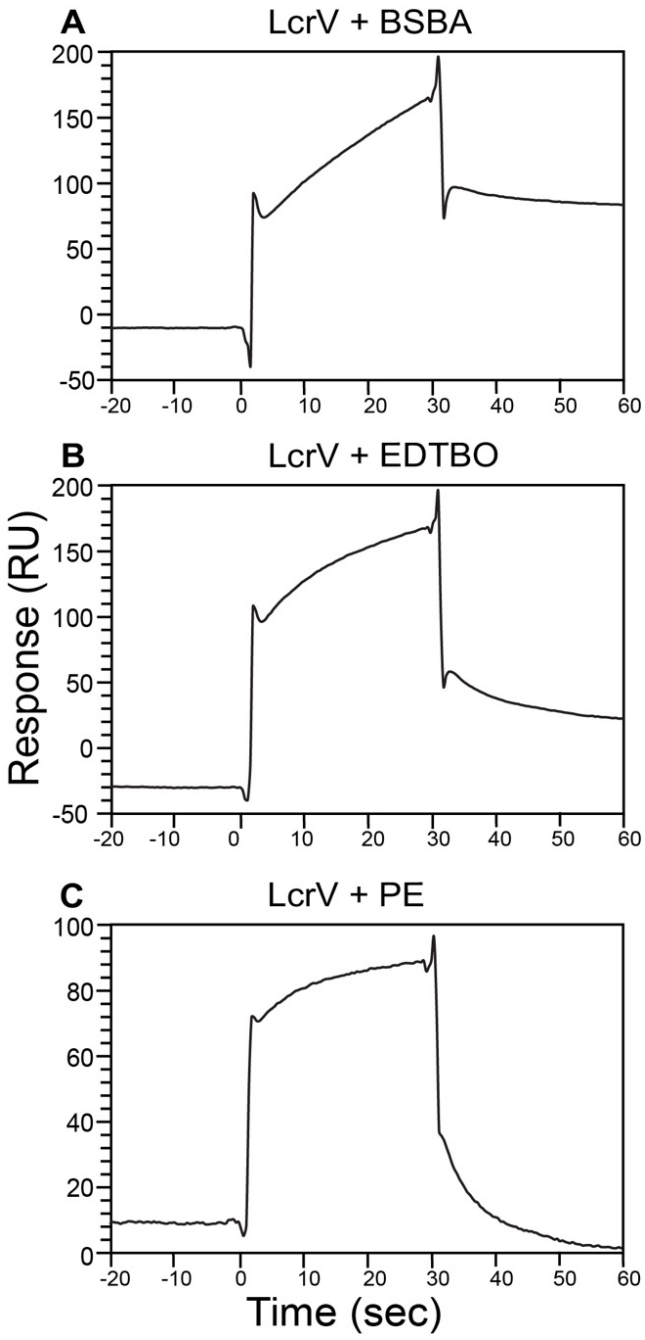

Figure 6-2. SPR screening results of LcrV. SPR sensorgrams displaying the binding of LcrV to**(A)** 2-(benzylsulfinyl) benzoic acid **(BSBA)**, **(B)** (3aR,7aS)-6-(6-ethoxypyridin-3-yl)-2,2-dimethyl-3a,4,5,7a-tetrahydro-2H-1,3-benzodioxol-5-one **(EDTBO)**, and **(C)** 2-(pyridin-2-yl) ethan-1-ol (**PE)**. SPR screening done by Anne Cooper.

6.4 Results

6.4.1 Expression and purification of LcrV

The expression and purification generated millimolar concentrations of unlabeled and isotopic ^{15}N/ILV labeled LcrV as described previously[8]. The final yield of the protein is ~0.5 mM. The high solubility of LcrV facilitates the SPR screening and the NMR characterization of binding surfaces.

6.4.2 Surface plasmon resonance (SPR) screening of Aube 104 and Aube 110 pibraries

SPR screening of the Aube 104 and Aube 440 identified three small-molecule fragments which bound to LcrV (Fig 6-2). The fragments had benzene (Fig6-1A) and pyridine scaffolds (Fig 6-1B-C).

6.4.3 NMR titration of LcrV and 2-(benzylsulfinyl) benzoic acid (BSBA) to determine binding surfaces

The binding site of BSBA on LcrV was determined by the ILV and ^{15}N-labeled LcrV titration of increasing concentrations of BSBA at molar ratios of 1:0, 1:4, 1:8, and 1:16. Simultaneous ILV and ^{15}N labeling of LcrV enabled titration to be monitored with two different probes[19]. 2D ^{1}H-^{13}C HSQC spectra were acquired for various titration points to observe the effect of BSBA on the ILV methyl resonances of LcrV (Fig. 6-3A), and 2D ^{1}H-^{15}N TROSY spectra were acquired to study the effect of BSBA on the ^{15}N backbone amides of LcrV (Fig. 6-3B). The ILV and amide-based titrations showed concentration dependent changes in the peak positions for the affected LcrV residues at G87, G88, T127, L137, L178, G228, and G240

(Fig. 6-3C-D), indicating the interaction to be in fast exchange of NMR timescale.

The LcrV/BSBA interaction interface was mapped by calculating the weighted chemical shift deviation (CSD) for the non-overlapped ILV (Fig 6-4 A) peaks of LcrV. Residues that displayed significant CSD (average plus a standard deviation) were L136 and L178. The ILV titrations and backbone amide titrations results indicated that the affected residues map on the bottom of the N-terminal globular domain, top of the central coiled-coil, and the turns of the α/β domain of LcrV (Fig 6-4B).

6.5 Discussion

LcrV is an essential virulence factor in *Yersinia pestis* that is indispensable for the proper assembly and functioning of the T3SS[1]. The exposure of LcrV to the bacterial surface makes it an attractive target for developing new antimicrobials. Here we identified new small molecules from the Aube 104 and Aube 440 libraries that showed binding to LcrV which could be used as scaffolds for developing new antimicrobials.

The Aube 104 library consists of 104 small molecules, which are found in substructures of known antibiotics and the Aube-440 consists of 440 small molecules that are derived as intermediates in synthesis of drug-like candidates. Results of SPR screening of these libraries showed that LcrV bound to only 3 compounds— 2-(benzylsulfinyl) benzoic acid (BSBA), (3aR,7aS)-6-(6-ethoxypyridin-3-yl)-2,2-dimethyl-3a,4,5,7a-tetrahydro-2H-1,3-benzodioxol-5-one (EDTBO) and 2-(pyridin-2-yl) ethan-1-ol (PE) in the Aube 440 library. These compounds are derivatives of benzene and pyridine scaffolds. Previous screening with the Zenobia library showed that LcrV binds to similar scaffolds[19], suggesting that these scaffolds are important for binding with LcrV.

We characterize the interaction of LcrV with BSBA by NMR spectroscopy. The chemical perturbations in the ILV and amide-based titrations show that BSBA binds to the bottom of the

N-terminal domain, top of the central coiled-coil, and the turns of the α/β mixed domain of LcrV (Fig 6-4B). Since chemical shift perturbations can occur because of direct and indirect structural and dynamic effects, the affected surfaces of LcrV could also be a result of conformational changes induced upon the binding of BSBA. Previous NMR based characterization of small molecules from the Zenobia library also shows LcrV binds to small molecules primarily at the bottom of the N-terminal domain[19]. The combination of the BSBA and Zenobia small molecule library screening data implies that the surface near the bottom of the N-terminal domain of LcrV is a hotspot for binding small molecule fragments. The N-terminal domain of LcrV forms the basal structure of the tip complex which interacts with needle, YscF[20] . 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[21]. Our recent EPR spectroscopy data also shows that the N-terminal domain of LcrV binds to the tip protein chaperone LcrG (Fig 3-9, Chapter 3). Given these significant roles of the N-terminal domain of LcrV, the small molecule BSBA binds biologically relevant site which could be exploited to disrupt the assembly of the *Yersinia* T3SS.

In summary, we identified a new small molecule that binds to LcrV. This small molecule can act as a starting point for creating new compounds which could be used to develop a new class of antimicrobials.

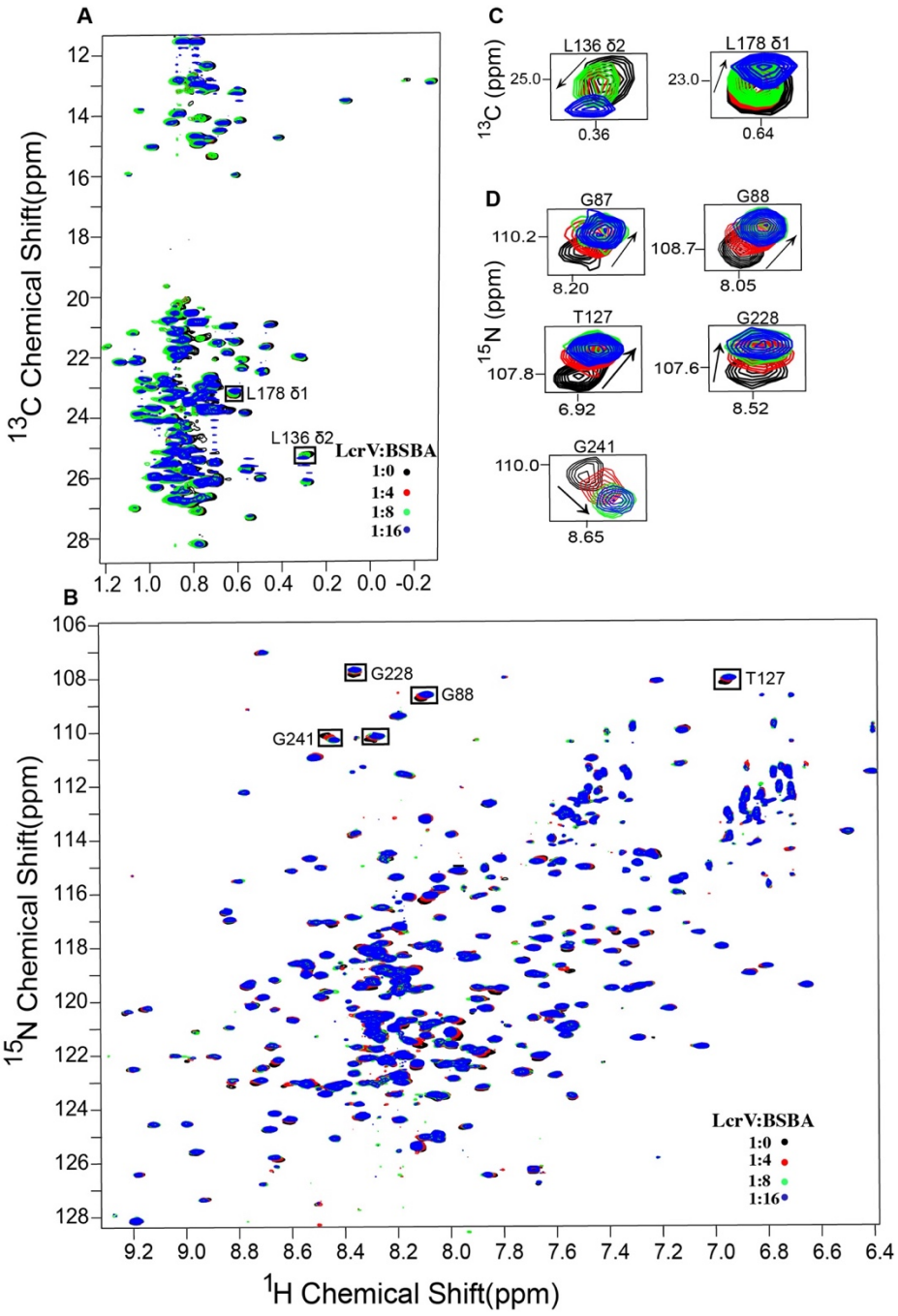

Figure 6-3. ILV and ^{15}N NMR titrations of LcrV with BSBA. **(A)** 2D ^{1}H-^{13}C HSQC spectra, and **(B)** 2D ^{1}H-^{15}N TROSY spectra of ILV and ^{15}N-labeled of LcrV titrated with increasing molar ratios of BSBA. Expanded sections of selected LcrV residues affected by the interaction with BSBA in **(C)** ILV and **(D)** ^{15}N NMR titrations.

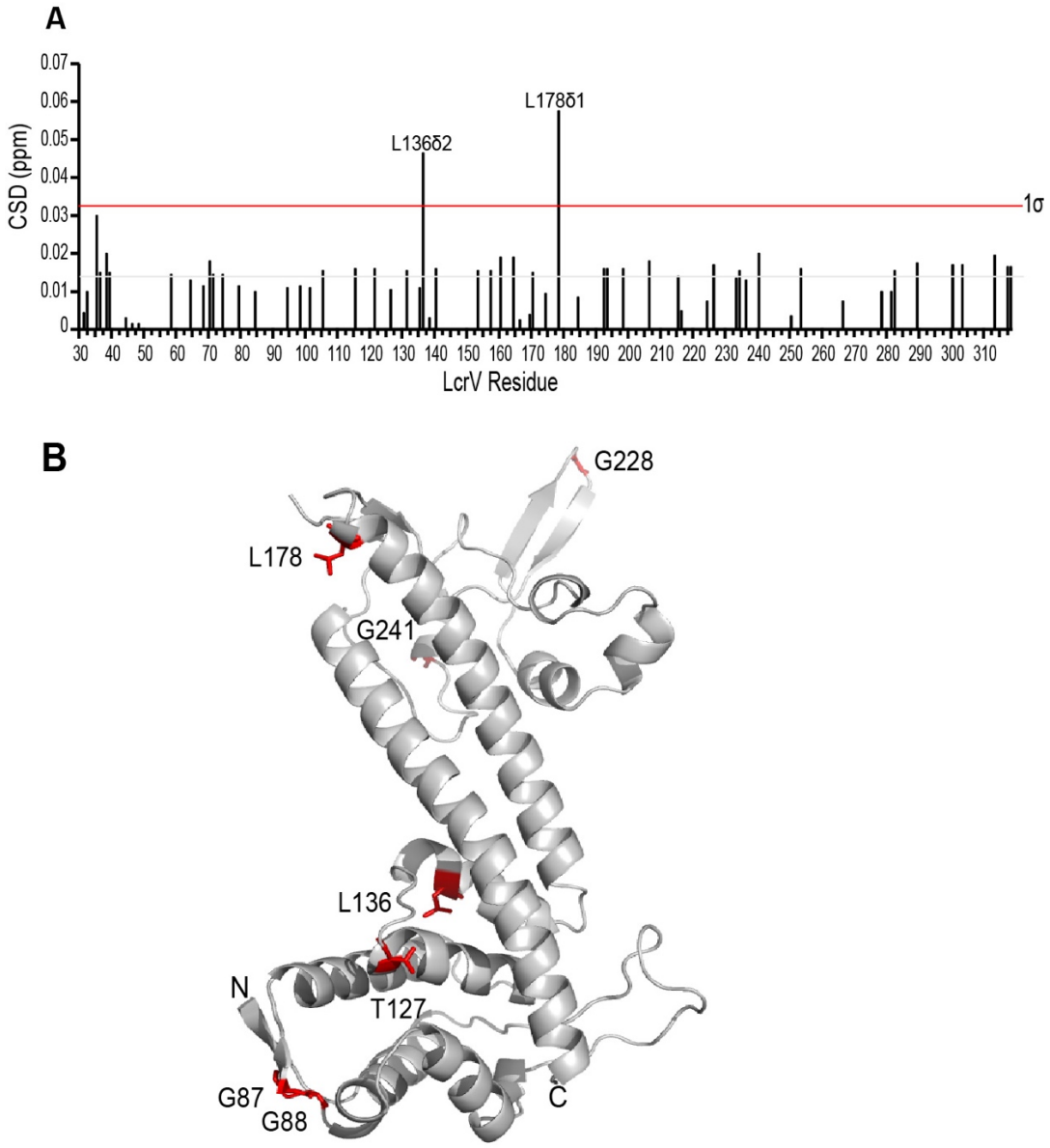

Figure 6-4. Analysis of the LcrV-BSBA interaction. **(A)** Plot of weighted chemical shift deviation from ILV. Grey and red lines correspond to the means and one standard deviation(1σ) from the mean, respectively. **(B)** LcrV residues affected by the binding of BSBA.

6.6 Reference

1. Cornelis, G.R. and F. Van Gijsegem, *Assembly and function of type III secretory systems.* Annu. Rev. Microbiol., 2000. **54**: p. 735-774.
2. Galan, J.E., et al., *Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells.* Annu. Rev. Microbiol., 2014. **68**: p. 415-38.
3. Dey, S., et al., *The type III secretion system needle, tip, and translocon.* Protein Sci, 2019. **28**(9): p. 1582-1593.
4. Wellington, E.M., et al., *The role of the natural environment in the emergence of antibiotic resistance in gram-negative bacteria.* Lancet Infect Dis, 2013. **13**(2): p. 155-65.
5. Davies, J. and D. Davies, *Origins and evolution of antibiotic resistance.* Microbiol Mol Biol Rev, 2010. **74**(3): p. 417-33.
6. McShan, A.C., et al., *Characterization of the Binding of Hydroxyindole, Indoleacetic acid, and Morpholinoaniline to the Salmonella Type III Secretion System Proteins SipD and SipB.* ChemMedChem, 2016. **11**(9): p. 963-971.
7. Mueller, C.A., et al., *The V-antigen of Yersinia forms a distinct structure at the tip of injectisome needles.* Science, 2005. **310**(5748): p. 674-6.
8. Chaudhury, S., et al., *Structure of the Yersinia pestis tip protein LcrV refined to 1.65 Å resolution.* Acta Crystallogr. Sect. F, 2013. **69**(Pt 5): p. 477-81.
9. Goure, J., et al., *Protective anti-V antibodies inhibit Pseudomonas and Yersinia translocon assembly within host membranes.* J.Infect.Dis., 2005. **192**(2): p. 218-225.
10. May, A.E., and Khosla, C. , *Discovery and Mechanism of Type III Secretion System Inhibitors. in Isr.* J. Chem., 2013.
11. Swietnicki, W., Carmany, D., Retford, M., Guelta, M., Dorsey, R., Bozue, J., Lee, M. S., and Olson, M. A. , *Identification of Small-Molecule Inhibitors of Yersinia pestis Type III Secretion System YscN ATPase.* PLoS One 2011. **6**.
12. Kauppi, A.M., et al., *Targeting bacterial virulence: inhibitors of type III secretion in Yersinia.* Chem. Biol., 2003. **10**(3): p. 241-249.
13. 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, *N-hydroxybenzimidazole inhibitors of the transcription factor LcrF in Yersinia: novel antivirulence agents.* J. Med. Chem, 2009. **52**: p. 5626-5634.
14. Harmon, D.E., Davis, A. J., Castillo, C., and Mecsas, *Identification and Characterization of Small-Molecule Inhibitors of Yop Translocation in Yersinia pseudotuberculosis.* Antimicrob. Agents Chemother., 2010 **54**: p. 3241-3254.
15. Chakravarty, A., *Elucidation of Protein-Protein Interaction and Protein-Small molecule Interactions in the Type III Secretion System. PhD. Diss.* 2021.
16. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes.* J. Biomol. NMR, 1995. **6**(3): p. 277-293.
17. Johnson, B.A., *Using NMRView to visualize and analyze the NMR spectra of macromolecules.* Methods Mol. Biol., 2004. **278**: p. 313-352.

18. Grzesiek, S., et al., *The solution structure of HIV-1 Nef reveals an unexpected fold and permits delineation of the binding surface for the SH3 domain of Hck tyrosine protein kinase.* Nat. Struct. Biol., 1996. **3**(4): p. 340-5.
19. Kaur, K., *NMR Studies of Molecular Interaction Involved in Type III Secretion System,Sumoylation,and The RNA Binding Protein HuR,PhD. diss. University of Kansas.* 2016.
20. Mueller, C.A., et al., *The V-antigen of Yersinia forms a distinct structure at the tip of injectisome needles.* Science, 2005. **310**(5748): p. 674.
21. Broz, P., et al., *Function and molecular architecture of the Yersinia injectisome tip complex.* Mol. Microbiol., 2007. **65**(5): p. 1311-20.

Chapter 7: Key Findings and Future Directions

7.1 Overview

The T3SS is essential for the pathogenesis of Gram-negative bacteria to cause infection in eukaryotic host cells[1],[2]. The T3SS consists of the needle complex, chaperone and effector proteins[3],[4],[5]. The needle complex is a macromolecular assemblage of over 20 different proteins consisting of the base, needle, tip complex, and translocon[3],[6]. The base is anchored to the inner and outer membrane of the bacteria. The needle protruding from the base forms a long filament in the intermembranous space between the bacteria and the host cell[3],[7]. The needle is capped by the tip complex which provides a platform for the assembly of the translocon[3],[7]. The tip complex is formed by the assembly of multiple copies of the tip protein[8]. The translocon assembles to form a pore through which virulence effector proteins are secreted into the host cell cytoplasm[3](Fig 1-1,Chapter 1). The chaperones present in the bacterial cytosol regulate the secretion of the effectors and other components of the T3SS[6],[8] (Fig 1-1, Chapter1-1).

The *Yersinia pestis* tip protein LcrV and the tip protein PcrV in *Pseudomonas aeruginosa* interacts with their cognate tip protein chaperones *Yersinia* LcrG and *Pseudomona*s PcrG in the bacterial cytosol during the assembly of the needle complex. There are no high resolution structures of the tip chaperone proteins and the structures characterized by NMR spectroscopy show that the tip chaperone proteins LcrG and PcrG are partially folded alpha helices[9],[10]. Computational modeling predicts that LcrG exists as a coiled coil[11] and PcrG as a four helix bundle[12]. The first aim of my dissertation was to detect the populations of LcrG and PcrG which may have tertiary contacts and also provide additional insight to conformational changes in the structures of LcrG and PcrG when they are bound to their cognate tip proteins LcrV and PcrV. To achieve this aim, I used EPR spectroscopy in collaboration with Dr. Likai Song from Florida State University.

The interaction between the tip and tip chaperone protein is poorly characterized. The current knowledge of this interaction predicts that the tip protein chaperones will form an alpha helical hairpin and interact with the central coiled-coil domain of the tip proteins (Fig 1-2,Chapter 1)[13],[14]. The previous NMR data shows that binding of the *Yersinia* tip chaperone induces chemical perturbations of the residues found in the N-terminal globular domain, the central coiled-coiled domain, and the mixed α/β domain suggesting that the overall structure is affected by the binding of LcrG[9],[15]. The NMR data of the PcrG-PcrV interaction shows that the binding of PcrV induces a global structural change in PcrG[10]. The second aim of my dissertation was investigating the binding interactions of *Yersinia* LcrV and LcrG using EPR spectroscopy and the binding interaction of *Pseudomonas* PcrG-PcrV using FRET spectroscopy in collaboration with Dr. Mark Richter and Dr. Erik Holmstrom from the University of Kansas.

The T3SS is crucial for virulence and it is highly conserved and surface exposed among different Gram-negative pathogenic bacteria. The disruption of the various protein-protein interactions in the needle complex is an attractive target for the development of new antimicrobials. In this dissertation I have identified new small molecules that bind to the *Yersinia* tip protein LcrV by SPR and NMR spectroscopy. These new small molecules can be used as lead compounds in the drug design of new inhibitors of T3SS.

7.2 Different Structural Conformations of *Yersinia* LcrG and *Pseudomonas* PcrG

7.2.1 Key findings

The EPR studies[16] of the different conformations of LcrG and PcrG are described in Chapter 2 and Chapter 4. The results of the EPR data for LcrG shows that when LcrG is in the free form, the N-terminal helix (helix α1) and the C-terminal helix (helix α3) have a spin-spin distance of 9 Å (Fig 7-1A) suggesting that the spin labeled helices form a closed conformation. The EPR data of LcrV bound to spin labeled LcrG shows the spin labeled helices have a distance greater than 25 Å suggesting the two helices form an open conformation as shown in Fig 7-1B.

The results of the EPR studies of PcrG also shows that in the free form PcrG exists in a closed conformation with a spin-spin distance of 10 Å (Fig 7-1C) and when bound to PcrV, the spin-spin interaction disappears and the distance between helices N-terminal helix (helix α1) and C-terminal helix (helix α2) is greater than 25 Å forming an open conformation as shown in Fig 7-1D.

In my dissertation I have reported the first experimental evidence suggesting that LcrG and PcrG have drastically different conformations in their free form and when bound to their binding partners LcrV and PcrV.

7.2.2 Future directions

The results of the EPR spectroscopy of LcrG and PcrG show that the distance information was obtained from one double spin labeled sample — LcrG C34R1/D65R1 and $PcrG^{9-76}$ E20R1/E63R1. Site directed mutagenesis has to be used to design new constructs for double spin

labeling in N-terminal and C-terminal helices of LcrG and PcrG for acquiring more spin-spin distances between the two helices using EPR.

The spin-spin distance between the N-terminal and C-terminal helices of LcrG and PcrG when bound to their cognate binding partners LcrV and PcrV could not be detected by the EPR spectroscopy technique used for the studies in Chapter 2 and Chapter 4. Alternatively, the EPR techniques like double electron-electron resonance (DEER) and pulsed electron double resonance (PELDOR) can be used instead as these techniques can detect longer spin-spin distances up to 4nm[17]. In the future, the acquisition of more spin-spin distances between N-terminal and C-terminal helices of LcrG and PcrG in the free form and when bound to LcrV and PcrV may be used to build structural models of LcrG and PcrG in combination with computational modeling tools[15].

7.3 The LcrG-LcrV and PcrG-PcrV Interaction

7.3.1 Key findings

The results from EPR spectroscopy of spin labeled LcrG and spin labeled LcrV are described in Chapter 3. In contrast to the prevailing hypothesis in the literature[13], results presented here indicate that LcrG does not form an alpha-helical hairpin upon binding to LcrV. The N-terminal helix (helix α1) of LcrG is at a spin-spin of 12 Å from the helix α3 of the N-terminal domain of LcrV, and the C-terminal helix (helix α3) of LcrG is at a spin-spin distance of 18 Å from helix

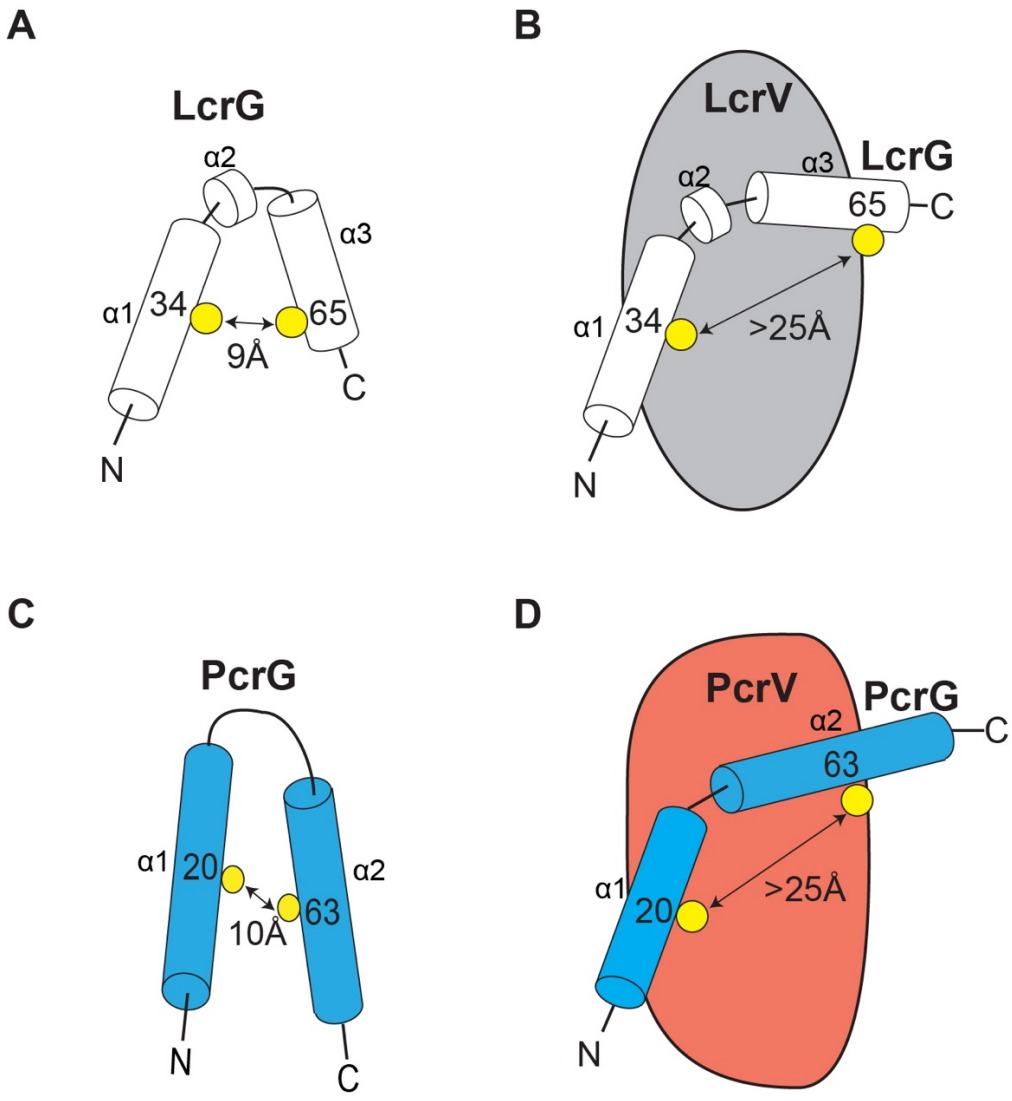

Figure 7-1 Models of the Open and Closed Conformations of LcrG and PcrG. (A) A model of LcrG based on EPR shows that the N-terminal helix (helix α1) and C-terminal helix (helix α3) have a spin-spin distance of 9 Å. **(B)** A model of LcrG-LcrV interaction based on EPR shows that the N-terminal helix (helix α1) and C-terminal helix (helix α3) have a spin-pain distance > 25 Å. **(C)** A model of PcrG based on EPR shows that the N-terminal helix (helix α1) and C-terminal helix (helix α2) have a spin-spin distance of 10 Å. **(D)** A model of PcrG-PcrV interaction based on EPR shows that the N-terminal helix (helix α1) and C-terminal helix (helix α2) have a spin-pain distance > 25 Å. **(A)** and **(C)** are designated as the closed conformation whereas **(B)** and **(D)** are designated as the open conformation.

α6 of the N-terminal domain of LcrV (Fig 7-2A-B), indicating the entire structure of LcrG interacts with N-terminal domain of LcrV.

The results of FRET spectroscopy of the PcrG-PcrV interaction as described in Chapter 5 shows that the donor labeled $PcrG^{9-76}$ A66C (PM) interacts with the acceptor labeled PcrV V120C (CPM) forming a tight binding complex with a k_d of 15±2 nM, indicating that the C-terminal helix (helix α2) of PcrG is close to the N-terminal domain of PcrV (Fig 7-2C).

7.3.2 Future directions of LcrG-LcrV interaction

The model of LcrG-LcrV interaction is derived from the spin-spin distances of spin labeled N-terminal helix — C34S/K28C and C-terminal helix — C34S/D65C of LcrG in complex with the spin labeled N-terminal domain constructs — Q37C and K137C of LcrV. I have used site directed mutagenesis to design new constructs of LcrV for the attachment of spin labels on the central coiled-coil domain — S151C, K176C, and S285C and the mixed α /β domain — N197C. The EPR data acquired from the spin labeled N-terminal and C-terminal helices of LcrG and the spin labeled coiled-coil and mixed α/β domain constructs of LcrV would add new spin-spin distances to the current model (Fig 7-2B). The new spin-spin distances would also provide further insight into the interaction of LcrG with the central coiled-coil and mixed α/β domains of LcrV.

7.3.3 Future directions of PcrG-PcrV interaction

The FRET spectroscopy data acquired for the PcrG-PcrV interaction only shows the complex formation of $PcrG^{9-76}$ A66C (PM) and PcrV V120C (CPM). The FRET data however, does not provide distance information between the two fluorophores. Alternate techniques such as single-molecule FRET spectroscopy[18] and EPR spectroscopy can be used to determine distances between the interacting residues of PcrG and PcrV. The distance information further improves the model validate the PcrG-PcrV interaction as shown in Fig 7-2C, and also determine the domains of PcrV which are involved in interaction with PcrG.

In this dissertation, I have used EPR and FRET spectroscopy to provide additional insight into the domains of LcrV and PcrV which interact with N-terminal and C-terminal helices of their cognate tip proteins LcrG and PcrG. *In-silico* methods and SPR can be used to screen novel small-molecule libraries to identify small molecules that bind to LcrV or PcrV and disrupt their interaction with tip proteins LcrG and PcrG respectively.

7.4 LcrV and Small Molecule Interaction

7.4.1 Key findings

The SPR based fragment screening of the Aube libraries shows that molecules with benzene and pyridine scaffolds bind to the *Yersinia* tip protein LcrV. The NMR characterization of 2 -(benzylsulfinyl) benzoic acid (BSBA) with LcrV shows that BSBA binds to the N-terminal globular domain of LcrV (Fig 6-4B, Chapter 6). The N-terminal globular domain of LcrV binds to the needle YscF[19] and the tip protein LcrG (Fig 7-2A), showing that BSBA binds to a biologically significant site of LcrV that can be used as a target to develop new antimicrobials.

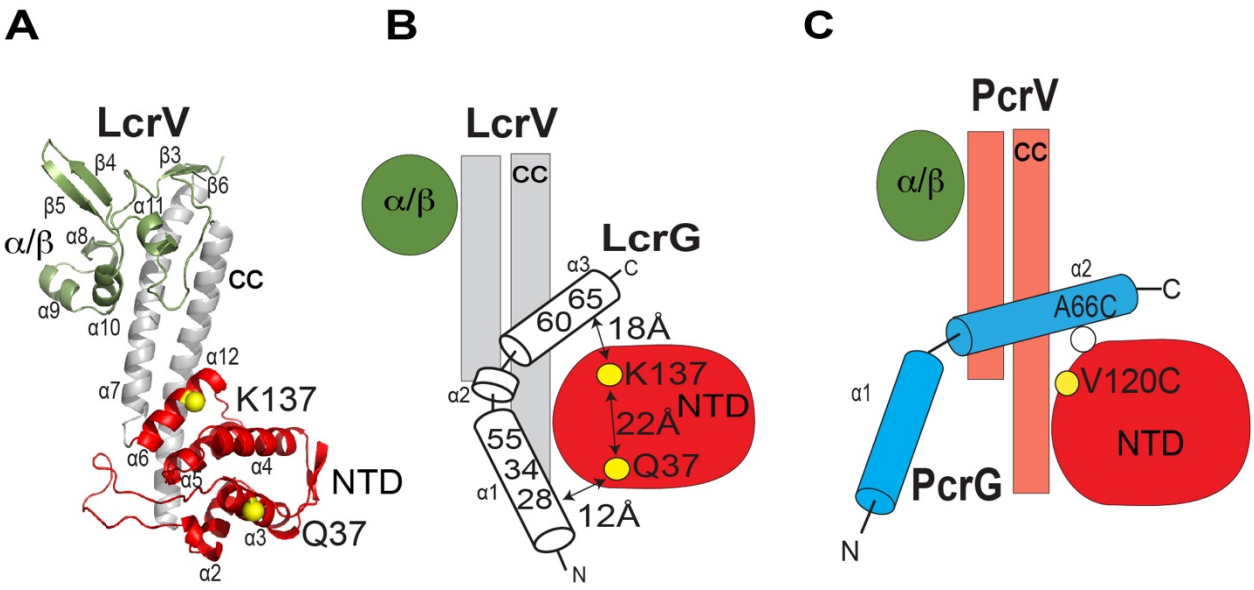

Fig. 7-2. **Models of LcrG-LcrV and PcrG-PcrV Interactions from EPR and FRET**. **(A)** LcrV showing the residues on helix α3(Q37) and helix α6 (K137) of the N-terminal domain (NTD) (red). **(B)** Cartoon of LcrG-LcrV interaction showing that LcrG N-terminal helix (helix α1) and C-terminal helix (helix α3) interact with NTD (red). LcrV has an N-terminal domain (NTD)(red), α/β domain(green), and a coiled-coil domain (CC) (grey). Model derived from EPR data. **(C)** Cartoon of PcrG A66C PM-PcrV V12C CPM interaction showing that the C-terminal helix (helix α2) of PcrG interacts with N-terminal domain (NTD) (red) domain of PcrV. PcrV has an N-terminal domain (NTD)(red), α/β domain(green), and a coiled-coil domain (CC) (salmon). Model derived from FRET data.

7.4.2 Future directions

The SPR screening of Aube library shows that the benzene and pyridine scaffolds bind to LcrV and similar scaffolds bind to LcrV from the Zenobia small molecule library[15]. Small molecules that bind to PcrV used by *Sundin et.al.* also consist of the same scaffolds[20]. PcrV should also be screened with the Aube libraries and LcrV also should be screened with the small molecules used by *Sundin et.al*.[20] to find the common scaffolds that bind to LcrV and PcrV. These chemical scaffolds identified in my SPR screens and the small molecules identified in the previous studies[15],[20] can be used as lead compounds to develop new antimicrobials which target the LcrG-LcrV and PcrG-PcrV interaction. Their ability to disrupt the complex formation should tested by FRET-based binding assays[21].

The interaction of BSBA with LcrV is indicative of weak binding. BSBA can be modified chemically in collaboration with medicinal chemists to improve its specificity and bioavailability[22],[23],[24]. BSBA can be used as a starting point to develop new anti-infectives which can target the *Yersinia* tip protein LcrV[22],[24].

7.5 References

1. Chakravortty, D., et al., *Formation of a novel surface structure encoded by Salmonella Pathogenicity Island 2.* EMBO J., 2005. **24**(11): p. 2043-2052.
2. Dortet, L., et al., *Pore-forming activity of the Pseudomonas aeruginosa type III secretion system translocon alters the host epigenome.* Nat Microbiol, 2018. **3**(3): p. 378-386.
3. Dey, S., et al., *The type III secretion system needle, tip, and translocon.* Protein Sci, 2019. **28**(9): p. 1582-1593.
4. Coburn, B., I. Sekirov, and B.B. Finlay, *Type III secretion systems and disease.* Clin Microbiol Rev, 2007. **20**(4): p. 535-49.
5. Cornelis, G.R. and F. Van Gijsegem, *Assembly and function of type III secretory systems.* Annu. Rev. Microbiol., 2000. **54**: p. 735-774.
6. Galan, J.E., et al., *Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells.* Annu. Rev. Microbiol., 2014. **68**: p. 415-38.
7. Burkinshaw, B.J. and N.C. Strynadka, *Assembly and structure of the T3SS.* Biochim Biophys Acta, 2014. **1843**(8): p. 1649-63.
8. Cornelis, G.R., *The type III secretion injectisome.* Nat. Rev. Microbiol., 2006. **4**(11): p. 811-825.
9. Chaudhury, S., et al., *The LcrG tip chaperone protein of the Yersinia pestis type III secretion system is partially folded.* J. Mol. Biol., 2015. **427**: p. 3096-3109.
10. Chaudhury, S., et al., *Nuclear Magnetic Resonance Characterization of the Type III Secretion System Tip Chaperone Protein PcrG of Pseudomonas aeruginosa.* Biochemistry, 2015. **54**(43): p. 6576-85.
11. Lawton, D.G., et al., *Interactions of the type III secretion pathway proteins LcrV and LcrG from Yersinia pestis are mediated by coiled-coil domains.* J. Biol. Chem., 2002. **277**(41): p. 38714-22.
12. Basu, A., et al., *PcrG protects the two long helical oligomerization domains of PcrV, by an interaction mediated by the intramolecular coiled-coil region of PcrG.* BMC Struct Biol, 2014. **14**: p. 5.
13. Blocker, A.J., et al., *What's the point of the type III secretion system needle?* Proc. Natl. Acad. Sci. U.S.A., 2008. **105**(18): p. 6507-13.
14. Chatterjee, S., et al., *The crystal structures of the Salmonella type III secretion system tip protein SipD in complex with deoxycholate and chenodeoxycholate.* Protein Sci, 2011. **20**(1): p. 75-86.
15. Kaur, K., *NMR Studies of Molecular Interaction Involved in Type III Secretion System,Sumoylation,and The RNA Binding Protein HuR. PhD. Diss. University of Kansas.* 2016.
16. Jennings-Antipov, L.D., L. Song, and R.J. Collier, *Interactions of anthrax lethal factor with protective antigen defined by site-directed spin labeling.* Proc Natl Acad Sci U S A, 2011. **108**(5): p. 1868-73.
17. Richard W, B.A., Sozudogru E, EI-Mkami H, Owen-Hughes T, Norman DG, *EPR distance measurements in deuterated proteins.* J.Magn. Reson., 2010. **207**: p. 164-167.
18. Holmstrom, E.D.H., Andrea,Zheng, Wenwei,Nettels, Daniel,Best, Robert B.,Schuler, Benjamin, *Chapter Ten - Accurate Transfer Efficiencies, Distance Distributions, and*

Ensembles of Unfolded and Intrinsically Disordered Proteins From Single-Molecule FRET, in *Methods in Enzymology*, E. Rhoades, Editor. 2018, Academic Press. p. 287-325.
19. Mueller, C.A., et al., *The V-antigen of Yersinia forms a distinct structure at the tip of injectisome needles.* Science, 2005. **310**(5748): p. 674.
20. Sundin C, S.M., Spjut S, Qin L, Elofsson M., *Identification of Small Molecules Blocking the Pseudomonas aeruginosa type III Secretion System Protein PcrV. .* Biomolecules, 2021. **11**(5).
21. Schaap, M., et al., *Development of a steady-state FRET-based assay to identify inhibitors of the Keap1-Nrf2 protein-protein interaction.* Protein science : a publication of the Protein Society, 2013. **22**(12): p. 1812-1819.
22. Duncan, M.C., Linington, R. G., and Auerbuch, V., *Chemical inhibitors of the type three secretion system: disarming bacterial pathogens.* Antimicrobial Agents and Chemotherapy, 2012.
23. Clatworthy, A.E., E. Pierson, and D.T. Hung, *Targeting virulence: a new paradigm for antimicrobial therapy.* Nat. Chem. Biol., 2007. **3**(9): p. 541-8.
24. Chakravarty, A., *Elucidation of protein-protein an protein-small molecule interactions in type III secretion system. PhD.,Diss.* 2021.