Insight into the structural and functional characterization of T3SS's chaperone PcrG and stress response protein PspA


A THESIS

SUBMITTED FOR THE PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (SCIENCE)

OF

JADAVPUR UNIVERSITY

2022

BY

CHITTRAN ROY

STRUCTURAL BIOLOGY AND BIOINFORMATICS DIVISION

CSIR- INDIAN INSTITUTE OF CHEMICAL BIOLOGY

KOLKATA-700032

INDIA

सी.एस.आई.आर–भारतीय रासायनिक जीवविज्ञान संस्थान
वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद की एक इकाई
विज्ञान एवं प्रौद्योगिकी मंत्रालय के अधीन, एक स्वायत्त निकाय, भारत सरकार
4, राजा एस. सी. मल्लिक रोड, यादवपुर, कोलकाता - 700 032
CSIR - INDIAN INSTITUTE OF CHEMICAL BIOLOGY
A Unit of Council of Scientific & Industrial Research
An Autonomous Body, under Ministry of Science & Technology, Government of India
4, Raja S. C. Mullick Road, Jadavpur, Kolkata-700 032

CERTIFICATE FROM THE SUPERVISOR

This is to certify that the thesis entitled "Insight into the structural and functional characterization of T3SS's chaperone PcrG and stress response protein PspA" Submitted by Sri **Chittran Roy** who got his name registered on 21st September 2020 for the award of Ph. D. (Science) degree of Jadavpur University, is absolutely based upon his own work under the supervision of **Dr. Saumen Datta** and that neither this thesis nor any part of it has been submitted for either any degree/diploma or any other academic award anywhere before.

S Datta
17-6-2022

Dr. Saumen Datta
Senior Principal Scientist,
CSIR-IICB, Kolkata-700032

डॉ. सौमेन दत्ता / Dr. Saumen Datta
प्रधान वैज्ञानिक / Principal Scientist
सीएसआईआर-भारतीय रासायनिक जीवविज्ञान संस्थान
(वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद)
CSIR - Indian Institute of Chemical Biology
(Council of Scientific & Industrial Research)
4, राजा एस. सी. मल्लिक रोड / 4, Raja S. C. Mullick Road
कोलकाता-७०००३२ / Kolkata-700032

Phone : 2499-5700 (Pilot No. / EPABX No.) ● Website : www.iicb.res.in ● Telegram : LIVINGCELL
Fax : 91-33-2473-5197 (Dir) ● 91-33-2473-5368 (Dir) ● Fax : 91-33-2414-9475 (A.O.) ● 91-33-2429-8490 (A.O.)
91-33-2483-1983 (F&AO) ● 91-33-2483-1982 (SPO)

Preface

The study and findings presented in this thesis entitled ***"Insight into sequence structure and bioinformatics study of T3SS's chaperone and phage shock protein-A from Pseudomonas aeruginosa and Yersinia enterocolitica respectively"*** *were started in January 2016. This study is focused to further illuminating the pre-existing information regarding the structural and functional understanding of proteins inside the cell and their regulation.*

This thesis has been divided into five major chapters. ***Chapter I*** *contains the review of literature on the type three secretion system (T3SS) and phage shock protein (PSP) system.* ***Chapter II*** *describes detailed material and methods used to carry out this work.* ***Chapter III*** *contains the results and discussion of T3SS's chaperone protein, PcrG.* ***Chapter IV*** *contains the results and discussion of phage shock protein A (PspA). And finally,* ***Chapter V*** *contains the results of protein purification, crystallization, diffraction, and preliminary data solving for the domain of PspA (i.e., ΔPspA).* ***Chapter VI*** *contains the study of proteins' salt bridge energy, And,* ***Chapter VII*** *contains a summary.*

Due acknowledgment has been made for reporting the scientific observations and findings made by other scientists. I take all the responsibility for any kind of unintentional error or misinformation in this thesis.

Chittran Roy

(Chittran Roy)

Structural Biology & Bioinformatics Division

CSIR-IICB, Kolkata-700032

Dedicated to my beloved

“Mother”

Acknowledgment

First of all, I want to thank my mentor, ***Dr. Saumen Datta,*** *who gave me the opportunity to work under his guidance, without which it was an impossible task to pursue my doctoral studies. I am thankful for his continuous support and guidance which eased my PhD journey. His dedication and untiring efforts for science has been a source of inspiration for me. His vibrant presence always created a positive lab environment which boosted me and other students to work dedicatedly. I would like to take this opportunity to express my sincere gratitude to him for his all-time presence and his valuable advice to help me whenever I needed.*

Every gene requires the right set of factors for their expression, similarly in order to flourish as a researcher one cannot deny the importance of fellow lab colleagues. So, I would like to thank all the present and past members of the ***Datta's-lab*** *for making such a nice and ambient work environment.* ***Our lab may be considered an apt representation of India cause of*** *the rich diversity. I would like to take this opportunity to thank all the present and past members of Datta's-lab, starting from* ***Supratim Da, Rakesh Da, Abhishek Da, Pranab Da, Avishek, Rajeev, Basavraj, Arkoprova, Gourab, Atanu, Bidisha, Angira****,* ***and Maruf*** *for creating such a nice environment to work in and being such nice colleagues. I would like to thank* ***Arkaprava Das*** *for his support and his contribution during his short time stay in our lab.*

Personally, I want to thank my close friend, ***Sayanika*** *for her constant help and support. She would inspire me and through her words would give me the strength to face different challenges during my Ph. D. journey. I would always be indebted to her advice and scientific suggestions.*

I am really grateful to ***Anand and Rajeev*** *for being such good friends of mine. I am very much thankful for their valuable suggestions as and when required.*

I would like to thank my other close friends ***Arpan*** *and* ***Ruma****. Their presence and support have made my life more beautiful and enjoyable.*

I want to thank ***Koushik, Tridib, Dheeren, Dipayan, Kamran, Debojyoti****,* ***Anupam Da, Achinta****, and* ***Sunny*** *for all their direct and indirect help. I especially thank* ***Tridib*** for helping me many times since my graduation time.

I would like to thank my college friends Koushik, Sumanta, Shamim, Sabyasachi, Monojit, Subham, Avishek, Pallavi, Sukanya, Lopa, and Pritha as well as university friends

Sukanya, *Rifat, Arindam, Raihan, Yogesh, Debolina, Indrani**,** Antara for being always there for me and making my days memorable.*

I would like to acknowledge my school friends ***Sourav (Datta), Samiran, Sourav (Mukherjee), Bhabani, Rik, Krishanu, Sumit, Abhishek,.....*** *For making my school days wonderful.*

I want to convey my sincere regards to ***Dr. Arun Bandyopadhyay (Director CSIR-IICB)*** *and* ***all the past directors of CSIR-IICB*** *for giving me the opportunity to work at IICB.*

I express my sincere regards to ***all the scientists of CSIR-IICB*** *for their continuous help and good suggestion as and when required. Especially I would thank* ***Dr. Syamal Roy, and Dr. Jayati Sengupta*** *for their help and valuable suggestion.*

I am grateful to the ***technical staff members*** *of the central instrumentation facility (CIF) who always extended their helping hand during using the central instrumental facility.*

I am very much thankful to my college teacher ***Prof. Prithisadhan Bose, B B sir,*** *and* ***Mohit sir*** *who played a very crucial role at the time of my transition from B.Sc. to M.Sc.*

I am very much thankful to all teachers of my M.Sc. days, Particularly, I am thankful to ***Prof. Pranab Roy, Dr. Amal Kumar Bandyopadhyay Dr. Indrani Chandra, Dr. Sunil Kanti Mondal,*** *who played a very crucial supportive and guiding role during the time of transition from M.Sc. to PhD. I would heartily thank* ***Dr. Amal Kumar Bandyopadhyay,*** *who has always supported me and stood by me through thick and thin.*

I am also thankful to ***Prof. Tathagata Choudhuri,*** *for his encouraging words and support.*

I would like to thank ***Prof. Anupam Basu.*** *I have known him since childhood and seen him develop as a good scientific researcher. He has been a motivating force which guided me sail through the Ph. D. journey.*

Research is not complete without food. I would like to thank all the canteen staff for their generous behaviour during my stay at IICB. In this regard, ***Kashi Da, Avishek Da, and Mithu Da****'s names are worth a mention.*

Besides doing science I would like to acknowledge all my colleagues for finding time out of their busy lab schedule to play ***Cricket, Badminton, and carrom****. The times spent with them are very close to my heart and will always be remembered as the most memorable time of my life.*

I would like to offer my acknowledgments to ***Dushyant, Ritesh, Jeet, Sambit, Anirban, Purna, Chandrasobha, Ashmita, Sonali, Jyoti, Ayesha, Shirin, Krishna, Annek, Bhakta, Priya, Sumit, Bheem, Debleena, Debojyoti, Pinaki, Nirmal, Dipayan Sarkar, Bani Da, Nayan Da, Mithun Da, Som Da, Sarfaraz Da, Abdus, Kajal, Pulak, Chinmoy Da, Shiladitya,Peter, Subhadeep Da, Arijit, Samrat, Piyush Da, Arijeet Da, Snehasis, Ramkrishna, Parash,Sandip, Madan and Kartic Da*** *for their help whenever needed. This list would be incomplete if I miss acknowledging* ***3 no. Gate of Jadavpur University and Mamar chai er dokan****. These places worked as a stress-busting yard for me where I would spend a lot of my free time. In this regard, I would like to thank my accompanies-* ***Sayanika, Anand, Tridib, Koushik Sunny, Rajeev, Saroj, Ritesh, Kamran, Snehasish, Jeet, Saif, Debojyoti, Subrata, Kajal, Peter, Mithun Da, Som Da, Shameel Da, Tridib, Sayanika, Sarfaraj Da, Abdus and Dipayan.*** *The gathering and debate were never complete without a continuous supply of tea.*

I would like to express my heartfelt thanks to all my relatives for all their well wishes and love.

Baba, *your continuous support has helped me to complete my PhD. You have struggled a lot for giving me a better life to me. Thank you,* ***Baba****.*

Maa*, in words it's impossible to express my gratitude for you. Though, I am thanking you for giving me this life, and for your truly unconditional love, support, and care for me. It is your hard work and sacrifices that taught me how to chase my dream. You taught me how to overcome obstacles patiently- the best lesson of my life. Thank you,* ***Maa****.*

(Chittran Roy)

Tables of Contents

Page No.

ABBREVIATIONS

TTSS = Type three secretion system

PSP = Phage-shock-protein

PCR = Polymerase Chain Reaction

TAE = tris acetate-EDTA

PMSF = phenylmethylsulfonyl fluoride

SDS = Sodium Dodecyl Sulfate

PAGE = Polyacrylamide Gel Electrophoresis

NTA = Nitrilotriacetic acid

IM = Inner Membrane

OM = Outer Membrane

CD = Circular Dichrorism

FTIR = Fourier Transform Infrared Spectroscopy

BLI = Bio-Layer Interferometry

SB = Salt-bridge

GRAVY = Grand Average of Hydropathicity Index

MD = Molecular Dynamics

RMSD = Root Mean Square Deviation

RMSF = Root Mean Square Fluctuations

SASA = Solvent Accessible Surface Area

Rg = Radius of Gyration

IP = Isolated Pair

NU = Network Unit

SPC = Simple Point Charge

ASU = Asymmetric Unit

List of Tables

Page No.

List of Figures

Page No.

CHAPTER 1

1. Review of Literature

1. Review of Literature

Bacterial proteins are crucial because they frequently interact with the host cells (Nicod et al., 2017). They also assist bacteria in surviving in hostile settings such as extreme temperatures and pH fluctuations (Lund et al., 2014). Bacterial resistance to antibiotics and their survival results in the emergence of novel strains that pose a public health risk.

Understanding the pathogenic bacteria-host relationship is crucial nowadays (Casadevall & Pirofski, 2000). Apart from E. coli, *Pseudomonas aeruginosa* and *Yersinia enterocolitica* are useful models for studying host-pathogen interactions (Luo et al., 2017). As a result, drugs are used to inhibit bacterial cellular activity. *P aeruginosa* is a major antibiotic-resistant gram-negative organism that causes nosocomial infection (Rzhepishevska et al., 2014). Essentially, the bacteria take advantage of a person's weakened immune system to generate an infection, making them an opportunist. Pseudomonas uses the type three secretion system (T3SS) to inject anti-host effectors into the target cell (Halder et al., 2019). P aeruginosa has a Type 3 Secretion System (TTSS) that allows it to transfer its pathogenicity to the host membrane; it is a unique type of pathogenicity spreading mechanism that is usually seen in gram-negative bacteria. It is made up of 20 different types of different kinds of proteins (Yahr et al., 1997). The primary function of TTSS is the creation of injectosomes, which aid in the delivery of bacterial harmful effector proteins within host cells (Galle et al., 2012, Cornelis, 2006). The injectosome is a well-defined model that consists of a basal structure and a needle complex (Allmond et al., 2003). A translocon is generated at the tip of the needle by three translocator proteins: PcrV, PopB, and PopD. Two of these three translocators are hydrophobic (PopB and PopD), forming pores in the host cell membrane. On the other hand, the hydrophilic translocators (PcrV), also known as virulence factors or V-antigen, aid in the assembly of the hydrophobic translocators (Galle et al., 2012a). In the cytosol, the hydrophilic translocator PcrV (or V-antigen) interacts with PcrG, an effector chaperone of *P. aeruginosa's* Type 3 Secretion System (Cornelis et al., 2006).

Y. enterocolitica, on the other hand, causes gastroenteritis. The most common yersinia infection in humans, *Y enterocolitica*, causes gastroenteritis and lymphadenitis in the gastrointestinal tract and mesenteric lymph nodes (Luo et al., 2017). *Yersinia pestis*, a newly developed bacterium derived from Yersinia pseudotuberculosis, a slightly virulent enteric disease, is an excellent model genus for studying active pathogen evolution. In many cases, *Yersinia pseudotuberculosis* and *Y enterocolitica* serve as handy substitutes for research of Yersinia pathogenicity and infection cycles, as well as their interactions with immune and non-immune cells (Rzhepishevska et al., 2014). Phage-shock-protein (PSP) systems encode a stress response that is required for viability (Darwin et al., 2007). When the secretin component of *Yersinia enterocolitica* Ysc type III secretion system is formed, the production of this protein takes on an important physiological and physiological significance (Flores-Kim & Darwin, 2016b). In a mouse model of infection, *Y. enterocolitica* psp null mutants were found to be fully avirulent. Other known stress response systems in bacteria, such as RpoE and Cpx, are not activated when the PSP system is activated. This demonstrates the independent regulation of the PSP system as well as its unique regulation. The cytoplasmic membrane proteins PspB and PspC, which interact and probably work together to perform their regulatory role, are required for the stimulation of pspgene expression. However, the regulatory role of PspBC does not fully explain why they are required for survival during secretin-stress, implying that they have a second unrelated function (Srivastava et al., 2017a). For more information, the crystal structure of structural proteins gives new insight into the design of significant drugs against such kinds of a pathogen.

1.1 Type three secretion systems (T3SS)

P. aeruginosa chromosomes are involved in the control of the type III secretion mechanism (Zhu et al., 2016). At least six additional genes on the chromosome encode effector proteins and associated chaperones. The T3SS is made up of more than 20 proteins, many of which form oligomers and are membrane-bound (Diaz et al., 2011). During infection, bacteria detect an external cue from the host

environment and begin to construct the components of the secretion system, a process that must be coordinated in place and time. Assembly can be divided into several stages:

1.1.1 The needle complexes

The needle complex is the supramolecular structure that transports certain type III proteins from the bacterial cytosol to the external environment, passing through the bacterial cytoplasmic membrane, the peptidoglycan layer, and the outer membrane. *P. aeruginosa's* needle complex appears to be similar to that of *Yersinia spp*., Salmonella enterica, and *Shigella flexneri*, and is made up of two parts: a multi-ring base and a needle-like filament. The needle-like filament, which is made up of PscF subunits, is hypothesized to operate as a conduit for secreted factors as well as a sensor for host cell contact (Bergeron et al., 2016). These structures are 60–120 nm long and 6–10 nm broad, similar in size to *Shigella flexneri*, *Salmonella enterica*, and *Yersinia enterocolitica* needle-like filaments. Although some functional assignments may be deduced from the closely related Yersinia type III needle complex, little work has been done to describe the other components of the P. aeruginosa needle complex. PscN, for example, is assumed to be an ATPase that drives the P. aeruginosa secretion system and is controlled by PscL. With the aid of the lipoprotein PscW, PscC is a secretin-like protein that oligomerizes to form a channel across the bacterial outer membrane. According to research on its Yersinia homolog, PscP is most likely a molecular ruler that determines needle length. PscJ is thought to be a lipoprotein component of the needle complex's basal substructure. Despite modest improvements, more effort will be necessary to completely understand the intricate mechanism by which the needle complex plays its part in the secretion process.

1.1.2 The translocation apparatus

The translocation apparatus consists of a proteinaceous membrane hole that transfers effector proteins released by the needle complex across the host cell's plasma membrane. The translocation process is extremely efficient, with less than 0.1 percent of released effector proteins escaping into the

extracellular environment. P. aeruginosa T3SSs, like Yersinia, Salmonella, and Shigella spp. T3SSs, use three proteins for translocation: PopB, PopD, and PcrV (Banerjee et al., 2014).

1.1.3 Basal Structure: It goes through the inner membrane, periplasmic space, and outer membrane of the bacterium, affixing the needle to the bacterial membrane. The inner membrane ring of the basal body is formed by Pseudomonas lipoprotein PscD. PscC, a secretin-like protein, forms the outer membrane ring with PscI. PscJ forms a ring-like structure under the base. To energize the T3SS needle complex, the ATPase PscN is required. PscL is a negative regulator of PscN ATPase. The docking platform is created at the basal body's export gate by PscN/PscN-PscL/PscL.(Galle et al., 2012b).

1.1.4 Export Apparatus: In cooperation with the ATPase, an export apparatus promotes secretion via the needle's base by creating a docking platform for the chaperone substrate complex for export (Bergeron et al., 2016). Pseudomonas spp. have highly conserved export apparatus components that are integral membrane proteins (Aizawa et al., 2001). Pseudomonas has an export apparatus that includesPscD, PscS, PscU, PscT, and PcrD.

1.1.5 Chaperones of T3SS: Chaperones may make it easier for their corresponding protein partners to be stored in the bacterial cytoplasm and delivered to the secretion system. SpcS (previously known as Orf1) acts as a chaperone for both ExoS and ExoT and is essential for their maximum secretion. ExoU has a chaperone in the form of SpcU (Takaya et al., 2019). ExoY has yet to be assigned a chaperone.

1.1.6 Effectors of T3SS: PopB, PopD, PopK/A, PscF, YopM, PopJ/P, and PopT are some of the Pseudomonas outer proteins (Pops) found in *Pseudomonas aeruginosa*. They go for the cellular components that are responsible for the host cell's innate immunity. ExoY, ExsA, ExoU, and ExoS effectors are required for total pathogenicity of the bacterium (Khanppnavar & Datta, 2018).

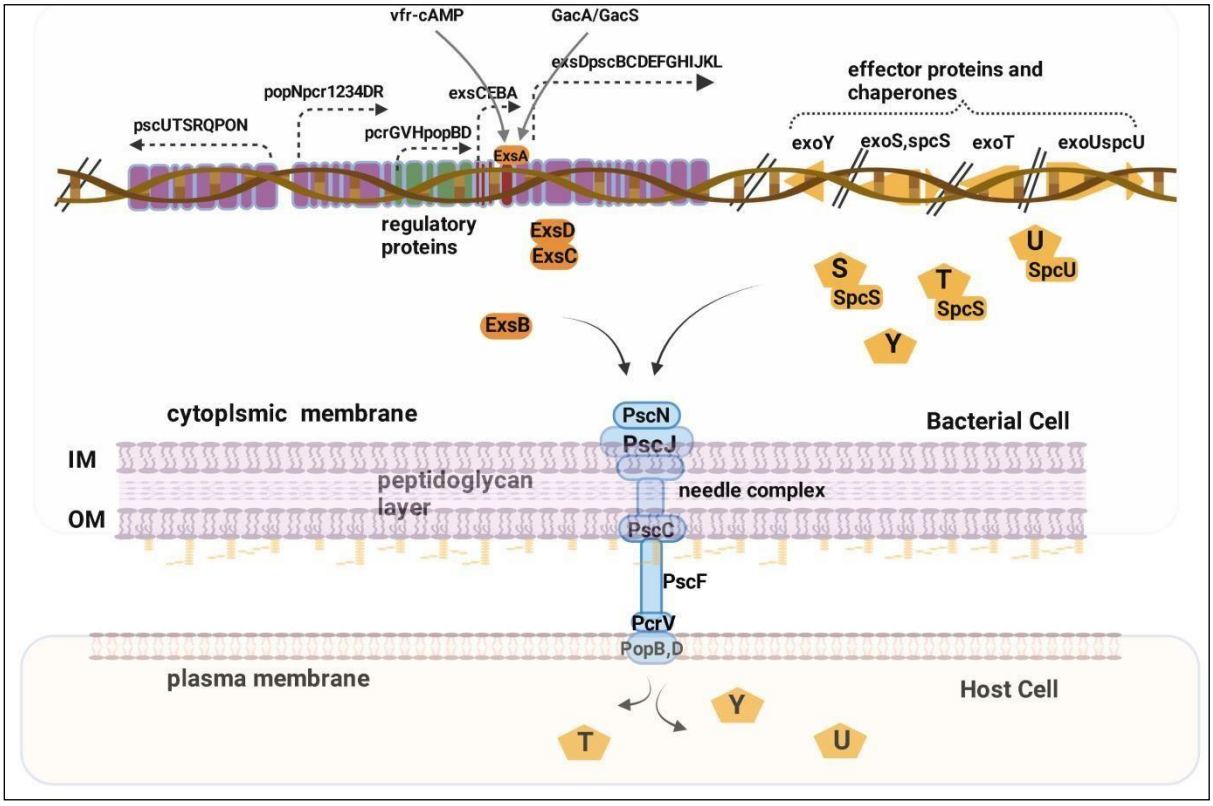


Fig 1.1 Type three secretion systems. T3SS is made up of five parts: the needle complex, the translocation machinery, regulatory proteins, chaperones, and effectors. Effectors and their chaperones are found throughout the bacterial genome, while structural and regulatory genes are found in five operons. The ultimate purpose of the T3SS structure is to inject effectors into host cells, destabilizing cell machinery and changing immune responses in order to increase bacterial survival rates (doi.org/10.1016/j.micres.2021.126719).

1.1.7 Pseudomonas chaperone protein PcrG

The needle tip protein chaperone PcrG regulates both the secretion of effectors from the bacterial cytoplasm into the host cytoplasm and the secretion activity. These two functions are tightly regulated by two distinct PcrG regions. As a result, PcrG is a multifunctional protein that, in addition to its export chaperone capabilities, also operates as a switch to break and make regulate effector secretion (Nanao et al., 2003). Experiments demonstrated that PcrG and PcrV interact 1:1 affinity complex (P. C. Lee et al., 2014), despite the fact that deletion of PcrG's 24 C terminal amino acids has no effect

on the formation of the affinity complex (Basu et al., 2014). Apart from the knowledge that it is a negative regulator of TTSS, nothing is known about the actual mechanism of PcrG regulation (Sundin et al., 2004). It has been shown that deletion of pcrG or pcrV causes partial dysregulation of effector secretion, whereas loss of both genes causes a high degree of effector secretion (P. C. Lee et al., 2010a) Scientists are currently trying to figure out how PcrG, as a tip chaperone, interacts with other TTSS regulatory proteins. 40aa of PcrG's N-terminal is enough to export its homologous substrate PcrV, which we've already looked at. The N terminal 40aa, on the other hand, plays no part in regulation. In experiments, PcrG fusion to the C terminal region of MBP has been shown to aid in the stability of the PcrG fragment. Other experimental evidence suggests that when PcrG binds to its corresponding protein (PcrV), its stability improves. Otherwise, it would be degraded.

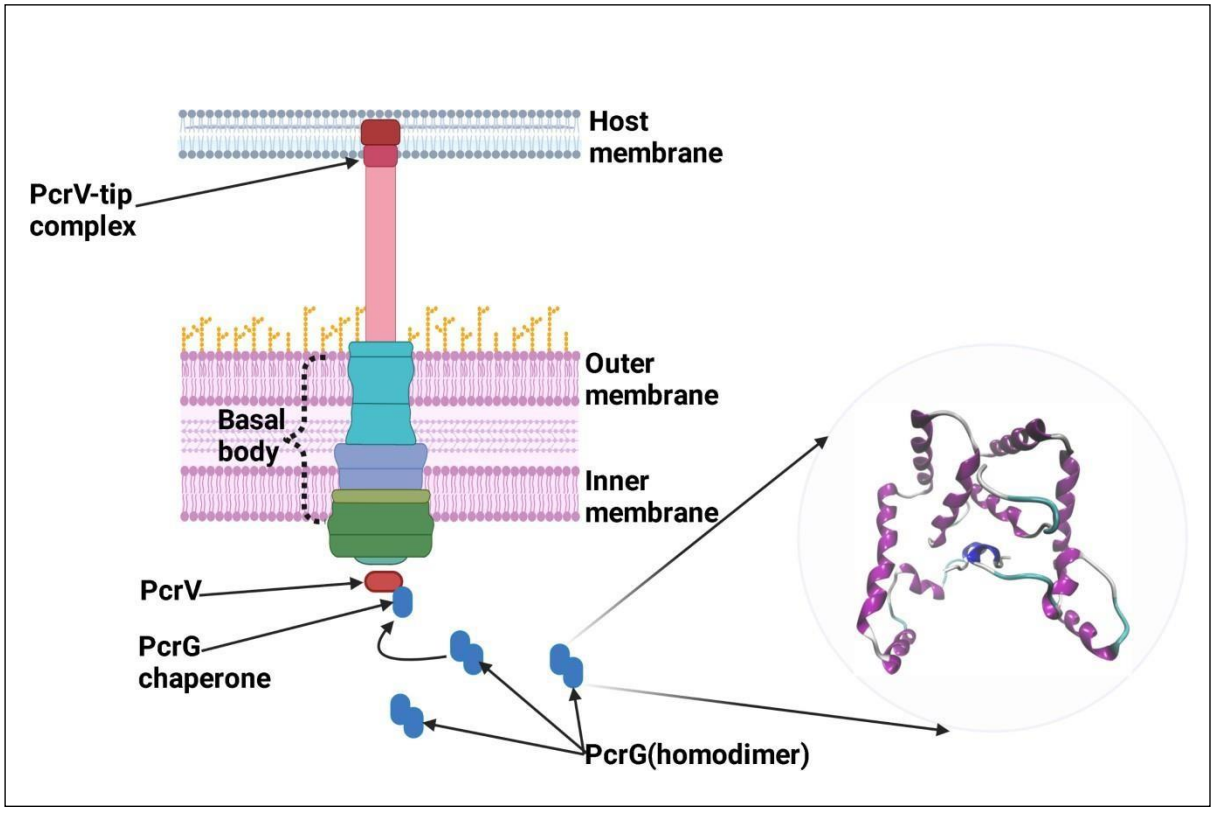


Fig 1.2 In a T3SS cartoon, the needle equipment is depicted. The needle apparatus consists of a basal body, a needle, a tip complex, and a translocon. An OM ring spans the bacterial inner membrane in the basal body, while an IM ring spans the bacterial outer membrane. The periplasmic region is reached by the neck, which is an extension of the OM ring.

The IM and OM rings are connected via the inner rod. An export apparatus and an ATPase complex are found in the basal body. The needle, which traverses the extracellular area, is capped by a tip complex and a translocon. The translocon produces a translocation pore within the host cell membrane. The bacterial effectors are delivered directly into the host's cytoplasm using the needle device. In the cytoplasm, the chaperone protein (PcrG) is dimeric, but in the presence of PcrV, it interacts in a 1:1 ratio (doi.org/10.1186/1472-6807-14-5).

1.2 Phage Shock Protein System in *Yersinia*

Yersinia enterocolitica possesses a selectively permeable cell membrane that controls substance entrance and exit by creating a proton-motive force (Nikaido, 2003,Wilharm et al., 2004). The Psp Response is thought to be centered on the dissipation of this proton-motive force. The Psp stress response in *Y. enterocolitica* is primarily produced by secretin mislocalization (Srivastava et al., 2017b), among other causes (Flores-Kim & Darwin, 2016, Srivastava et al., 2017b). The Type 3 Secretion System (T3SS) in *Y. enterocolitica* is made up of secretins, a class of outer membrane proteins that form 12-15 homomeric rings (Korotkov et al., 2011). The Psp operon system is encoded by PspF-PspABCDycjXF and PspGin *Y.enterocolitica* (Srivastava et al., 2017b). PspA is one of the most essential regulatory proteins involved in the induction of the PSP response (Joly et al., 2009). PspA stays linked to PspF, a transcription activator, in uninduced states (Joly et al., 2010). As a result, PspA prevents transcription of the enhancer-binding protein PspF (Heidrich & Brüser, 2018). PspA can attach to the changed PspBC complex after induction owing to various stresses such as mislocalization of secretins or dissipation of proton motive force. This frees PspF, which subsequently triggers the transcription of different ESR genes via activation of the PspA operon and the recently discovered PspG via a cascade of downstream signalling (Lloyd et al., 2004). PspB and PspC are integral cytoplasmic membrane proteins that operate as stress-relieving regulatory responses that are changed by induction and sequester PspA from PspF (Flores-Kim & Darwin, 2015). Many tests,such as the two-hybrid studies in E. coli in vivo, have demonstrated the connection between PspA

and PspF. PspF lacks a negatively regulating N terminal domain, so negative regulation occurs only when PspA is bound. This demonstrates that PspF is constitutively active in vivo, which has been demonstrated in PspA null mutants. PspA binds to PspF and suppresses its ATPase activity, which is required for the creation of an open complex at its target promoter (Joly et al., 2009). Coimmunoprecipitation studies (Yamaguchi S, Reid DA, Rothenberg E, Darwin AJ., 2013) show that as a result, PspB and PspC are required for most inducers to activate the Psp response (Flores-Kim& Darwin, 2016a). Because both the N terminal and the C terminal are entrenched in the cytoplasm, it has been hypothesized that PspB contains a single transmembrane helix whereas PspC has two (Jones et al., 2003). According to mutation studies, PspA binds at the C terminal of PspC (Gueguen et al., 2009). The ability of PspA to bind to PspC is one of the most important elements affecting the Psp response. In addition to connecting to PspA, PspC also binds to PspB's cytoplasmic domain.However, other findings have shown that PspA can separate from PspF without the participation of PspB and PspC, such as in the heat shock response in E. coli. According to genome analyses, PspA and PspF homologs have been discovered in other organisms (Huvet et al., 2011), but not PspB or PspC.According to José Flores-Kim and Andrew J. Darwin (2016), PspA may directly detect certain IM signals, prompting it to attach to the membrane and release PspF. As a result, another theory for the PSP system's operation is that PspB and PspC may not be sensors at all. They might be present in E. coli and Y. enterocolitica to allow their psp systems to respond to new signals. PspA clearly plays a function in maintaining the PMF of the Inner Membrane (IM) based on the evidence so far, but how it does so remain a mystery. A few theories have been derived from the tests thus far, although these are not definitive explanations. PspA generates oligomers on induction, and these oligomers cover the inner surface of the IM, providing a scaffold and reducing the IM's permeability (Kobayashi et al., 2007). Another hypothesis is that PspA oligomer mobility is essential to maintaining the PMF and is dependent on the cytoskeletal proteins MreB and RodZ (Engl C. et al., 2009). In PspBC null mutants, a mislocalized secretin decreases the PMF of E. coli and *Y. enterocolitica* and kills the

bacterium. When PspA is functioning alone, reduction of PMF cannot kill the cell, but mislocalization of secretin does. Other than diminished PMF in secretin mislocalization, catastrophic IM permeability and cytoplasmic shrinkage induce cell death. As a result, PspB and PspC can prevent cell death caused by secretin mislocalization on their own. As a result, PspB and C, like PspA, have a dual role and aid in the bacterial system's stress relief. However, the method through which PspB, PspC, and PspA deliver this stress-relieving activity remains unclear. Thus, obtaining the definitive answer to the role of these critical regulatory protein crystal structures would provide us with the whole picture and reveal the mysteries hidden inside these cryptic bacterial systems.

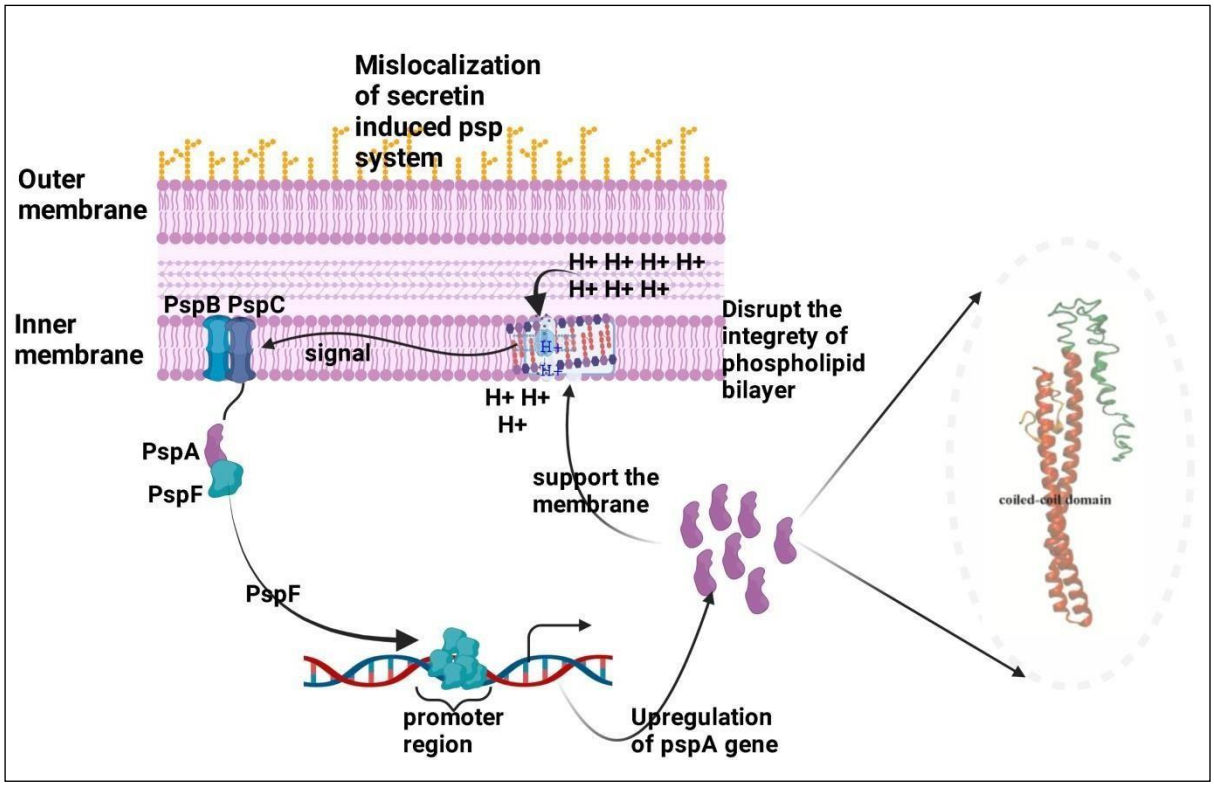


Fig 1.3 Phage Shock Proteins. The transcription factor PspF is required for increased psp gene expression. However, under non-inducing conditions, PspF is blocked by contact with another cytoplasmic protein, PspA. PspA relocates to the cytoplasmic membrane in response to a Psp-inducing stimulation, allowing PspF to stimulate psp gene expression. The integral membrane proteins PspB and PspC, which may sense an inducing trigger and sequester PspA through direct contact, are required for PspA translocation. The following rise of psp gene expression raises PspA concentration, allowing it to directly contact the membrane, possibly for physiological purposes (DOI:10.1007/s12275-012-1578-7).

1.2.1 Phage shock protein A

PspA is a single polypeptide subunit protein with a size of 26 kDa and 221 amino acids. It has one tryptophan (W71) and five phenylalanine (F4, F7, F168, F171, F197) side chains, with no disulfide bond (one cysteine, i.e., C64). Furthermore, about a quarter of the amino acid residues are acidic or basic. Heat shock, osmotic shock, ethanol, a deficiency in protein transport across the cytoplasmic membrane, and overproduction of an integral membrane protein, such as secretin, all stimulate the expression of pspA (Flores-Kim & Darwin, 2016b, P. C. Lee et al., 2010a, Joly et al., 2010). It's also been hypothesized that the psp system prevents the proton motive force from dissipating (Darwin & Miller, 2001, Eriksson et al., 2003). The activation of the *Y enterocolitica* PSP system appears to include an alteration in PspA position, as per experimental studies (Beloin et al., 2004). PspA forms an inhibitory complex with PspF in the cytoplasm, but during stress, PspA interacts with membrane-bound PspBC proteins. PspA, on the other hand, is soluble in non-Psp-inducing circumstances, but after induction of the Psp system, the bulk of it is now linked with the membrane (Joly et al., 2010). PspA protein has been linked to preventing a variety of physiological defects, including a defect in prolonging the stationary phase at highly alkaline pH (Yamaguchi & Darwin, 2012) or reducing proton motive force due to proton leakage across the cytoplasmic membrane, sec-dependent protein translocation (Heidrich & Brüser, 2018, Lloyd et al., 2004, Flores-Kim & Darwin, 2015), and membrane integrity mislocalization. However, the specific mechanism through which PspA performs this physiological role is uncertain.

CHAPTER 2

2. General Methodologies

2.1 Agarose gel electrophoresis of DNA:

Stock solutions:

Ethidium bromide solution: Ethidium bromide was typically prepared as a 10 mg/ml stock solution in water and stored in dark bottles at room temperature. After electrophoresis, gels were soaked in ethidium bromide solutions (0. 5 g/ml) for 30 minutes before visualization. b) Gel-loading buffer composition: 50 percent glycerol, 0.1 M EDTA, and 0.025 percent bromophenol blue were mixed together.1 mL Bromophenol Blue (11.5% w/v)87.5% glycerol2.9 mL EDTA (0.5 M)10 liters of glass distilled water Gel loading buffer (1.1 mL) was kept at 4°C for up to one month. 34. c) 50X TAE running buffer: TAE buffer containing 2M Tris and 50mM EDTA (pH 8.0), with glacial acetic acid added to achieve pH 7.6; Tris-base 24.3 g EDTA (0.5M stock) Glacial acetic acid, 10 mL. To make a final capacity of 100 ml, 5.6 ml of Milli Q water was added (pH 7.6).

2.2 The casting of the gel:

To make a mould, the open ends of the plastic tray that came with the electrophoresis device were sealed. On the table, the mold was set horizontally. In an Erlenmeyer flask, the needed amount of powdered agarose was mixed with a determined volume of 1x tris acetate-EDTA (TAE) electrophoresis buffer (typically to achieve a 1% solution of agarose). The slurry was heated until the agarose was completely dissolved, then cooled to 60^0C. The edges of the mold were sealed with a tiny amount of the agarose solution using a Pasteur pipette. The seal was left to dry. When the agarose was added, the comb was positioned 0.5 to 1 mm above the plate to ensure that a full well was produced. The remaining agarose was poured into the mould, making sure there were no air bubbles beneath or between the comb's teeth. The thickness of the gel ranged between 3 and 5 mm. The comb and sealing bridges were carefully removed after the gel had been set (30 to 45 minutes at room temperature), and the gel was installed in the electrophoresis tank.

Electrophoresis of DNA in agarose:

The gel was covered with enough electrophoresis buffer to cover it to a depth of roughly 1 mm. Gel-loading buffer was combined with DNA samples. Using a micropipette, the mixture was carefully injected into the slots of the submerged gel. At each end of the gel, marker DNAs of known size (e.g., Hind III digested DNA) were added. To allow the DNA to migrate toward the anode. The voltage was administered at a rate of 1 to 5 V/cm. The gel was run until the bromophenol blue and xylene cyanol moved to the proper distances. The gel was withdrawn from the tank, soaked in an ethidium bromide solution, and then photographed under ultraviolet light.

2.2.1 Purification of DNA using PCR:

The PCR purification Kit (Qiagen Cat. No. 28104) was used for this purpose.

Kit contains PCR clean-up columns with a collection tube. Nucleic acid-binding buffer, Washing buffer, and elution buffer.

Procedure:

1. 5 volumes of nucleic acid binding buffer were added to the PCR process and well mixed by vortexing.
2. In a tabletop centrifuge, the mixture was transferred to a PCR clean-up column placed in a collecting tube, and spun for 1 minute at 13,000 rpm. The flow-through was thrown out.
3. 500 µl of wash buffer was added to the column and centrifuged at the maximum speed for 1 min.
4. The flow-through was discarded, and the column was centrifuged again at maximum speed for 1 minute to eliminate any remaining wash buffer in the column.
5. After discarding the collecting tube, the column was put on a sterile microfuge tube and allowed to dry for 5 minutes. 6. For DNA elution, 50 l of water (pH 8–8.5) was employed. The water was boiled to 37°C before being poured into the column's base. After a 2-minute

incubation period at room temperature, the column was spun at maximum speed for 2 minutes in a microfuge to extract the pure DNA. The DNA-containing supernatant was kept at -20 degrees Celsius.

2.2.2 Purification of DNA fragments by agarose gel using Gel Extraction Kit.

Qiagen Cat. No. 28704 was used for this purpose.

2.2.3 Kit Contents

Elute Binding Column, Gel Solubilization Solution, Column Preparation Solution, Washing buffer, and elution buffer.

Procedure:

1. To decrease the gel volume, the DNA fragment of interest was removed from the agarose gel using a clean, sharp scalpel or razor blade.
2. The gel slice was weighed after being placed in a clean, sterile microcentrifuge tube.
3. 3 gel volumes of the gel solubilization solution were applied to the gel slice and incubated at 50- 60^0C for 10 minutes, or until the gel slice was fully dissolved, with periodic vortexing.
4. GenElute Binding Column G was placed in one collecting tube to create the binding column. Each binding column received 500µl of the Column Preparation Solution and was centrifuged for 1 minute. The flow-through was discarded.
5. 1 gel volume of 100 percent isopropanol was added after the gel slice was completely dissolved and blended until homogeneous.
6. After loading the binding column, which was constructed in a 2 ml collection tube, the solubilized gel solution combination was put into the column and centrifuged for 1 minute. The flowthrough was thrown out.

7. The binding column was filled with 700µl of wash solution (diluted from Wash Solution Concentrate G according to the manufacturer's instructions), centrifuged for 1 minute, and the flowthrough was discarded. To remove any remaining wash solution, the binding column was inserted back into the collecting tube and centrifuged for 1 minute. 8. The binding column was placed in a sterile microcentrifuge tube for the next step. In the middle of the membrane, 50 l of sterile water (pH 8.0) was introduced and incubated for 1 minute. 9. To elute DNA from the column, it was centrifuged for 1 minute.

2.2.4 Digestion and ligation of vector as well as PCR product

The vector and PCR product were digested with the appropriate restriction enzyme (RE) for cloning. The digestion was completed after 3–4 hours at 370°C, and the products were purified using the above-mentioned purification kit. After that, ligation was performed by mixing digested PCR and vector in the proper ratio and leaving it at 22^0C for 2-3 hours. A DH5 competent cell was created from the ligation product (clone). The desired gene's primer and restriction enzyme are listed in table 2.1.

Table 2.1 primers, restriction enzyme, and vector utilized for cloning of *pspA, ΔpspA,* and *ΔyspK gene.*

Gene name	primers	Restriction enzyme	vector
pspA *(Roy et al.)*	F5’ GATCAGATCTCATGGGTATTTTTTCTCGTTTTG R5’ TCGACTCGAGTTATTGCGCTGAATTCATTTTTG	Bgl2 and Xho1	pET Duet1
ΔpspA	F5’GGAATTCCATATGGGTATTTTTTCTCGTTTTGCC R5’CCGCTCGAGCTGATGACGCAG	Nde I & Xho I	pET 22b
ΔyspK	F5’ ATATACATATGGGTAAGCTAATAGGGGAAGG R5’ TGTGCTCGAGTAAATATTTCCTGTTTAATAATG	Nde I & Xho I	pET 22b

2.2.5 E. coli competent cells preparation

E. coli DH5 (or any other strain) was cultured in Luria broth overnight at 37 degrees Celsius. 100 mL Luria broth was inoculated with 1% of this culture. At 37 °C, cells were grown to an absorbance of 0.5-0.6 at 600 nm. The cells were then placed in an aseptic centrifuge tube and harvested at 4 °C for 5 minutes at 4,000 rpm (Sorvall SS-34 rotor). Cells were gently resuspended in 20ml of100mM MgCl2 and kept on ice for 10min. centrifuge again at 3500rpm for 15min. again resuspend the pellet in 20ml of 100mM CaCl2 and kept the ice for 15min at 3000rpm. Furthermore, resuspend the pellet in 2ml of 100mM CaCl2 with 15%v/v glycerol. Aliquot the above 2ml of competent cell and kept at -80^0 C for a long time.

2.2.6 Transformation of competent E. coli cells:

All procedures were carried out under stringent aseptic conditions. A total of 100 ml of a cooled suspension of competent cells were placed in a pre-chilled sterile Falcon tube for each transformation (17 mm x 100 mm). By progressively pouring DNA (60-70 ng) into the cell suspension at a volume of 5 l, DNA (60-70 ng) was added to each tube. For 30 minutes, the tubes were kept on ice. The tubes were heated for 1 minute in a non-circulating water bath at 42 °C. The tubes were placed on ice for 2 minutes after being heat shocked. The tubes were then put on a 37 °C shaker for 45min to allow the cells to recover and express the antibiotic resistance marker encoded by the plasmid. After that, 40ul of the culture was plated on Luria agar with the necessary antibiotics. After that, the plates were inverted and incubated for 14 to 16 hours at 37^0C.

2.2.7 Plasmid DNA isolation:

2.2.7.1 Plasmid DNA preparation on a small scale with the Qiagen Spin Plasmid Isolation kit:

1. In a mini centrifuge, 1-3 ml of overnight cell culture were centrifuged at 10,000 x g. To remove excess media, the supernatant was drained out and the tube was wiped upside-down on a paper towel.

2. The cell pellet was resuspended in 25 µl of cell resuspension solution with care.
3. 250 l of cell lysis solution was added, and the tube was inverted four times to mix it. The cell suspension was quickly cleaned.
4. By inverting the tube, 350µl of the neutralizing solution was added and stirred.
5. To obtain a tight pellet, the lysate was centrifuged at 10,000 x g in a microcentrifuge for 15 minutes or longer.
6. The column was used for each miniprep. The column was put on a collecting tube, and the cleaned lysate was added to it and centrifuged for 1 minute at maximum speed.
7. After discarding the flow-through, 500 l of wash buffer was added and centrifuged at 10,000 x g for 1 minute.
8. To remove any remaining wash buffer, the column was put on the collecting tube and centrifuged at 10,000 x g for 1 minute.
9. The column was put in a sterile microcentrifuge tube and allowed to dry for 5 minutes.
10. To elute DNA, 50 l of sterile water (pH 8.5) was immediately added to the column and centrifuged at 10,000 x g for 1 min.
11. The DNA was eluted and dried in a Savant Speed Vac centrifugal evaporator before being reconstituted in 20 l of T buffer and kept at-20 °C.

2.3 Estimation of protein

The Lowry and Bradford test technique was used to evaluate the protein content.

2.3.1 Lowry's method: Stock solutions:

Solution A: 4.5 gm sodium carbonate, 1 gm sodium dodecyl sulfate in 100 ml of distilled water.

Solution B: 2 gm anhydrous copper sulfate in 100 ml distilled water.

Solution C: 4.5 g sodium potassium tartarate in 100 mL distilled water (solution C). Within two months, all of the remedies had been implemented.

Reagent A was prepared by combining 9.8 ml of solution A, 0.1 ml of solution B, and 0.1 ml of solution C to make 10 ml of this reagent.

Reagent B: Dilute 1 volume of Folin-Ciocalteau phenol reagent with 1 volume of distilled water to make Reagent B. Reagents A and B were freshly produced immediately before use. The volume of the sample proteins and the reference protein, bovine serum albumin (BSA), comprising between 4 and 20 g of protein, were adjusted to 100 l in each tube. Each tube received 1 mL of reagent A. Vortexing was used to fully mix the solutions, which were then incubated at 37°C for 20 minutes. Following that, 100 l of reagent B was added to each tube, the contents were well mixed, and the tubes were incubated at 37^0C for another 20 minutes. In an Ultrospec II spectrophotometer, the absorbance of each protein solution was measured at 740 nm (Amersham Biosciences). The absorbance values of the standard protein solutions were used to create a standard curve, and the concentration of the protein was calculated using the standard curve.

2.3.2 Bradford method:

Stock solutions

Standard protein solution: 0.5 mg of BSA per mL of distilled water is the standard protein solution.

Bradford reagent: Commercially available Coomassie Brilliant Blue solution (1 litre contains 100 ml of Coomassie Brilliant Blue G250 dissolved in 50 ml of 95 percent ethanol and 100 ml of 85 percent phosphoric acid).

Procedure

The volume of the experimental samples and the reference protein, bovine serum albumin (BSA), comprising 1–10 g of protein, were adjusted to 100 l with 0–15 M NaCl in separate tubes. As a

control, a tube with no protein was used. Each tube received 1 mL of Bradford reagent. Vortexing was used to completely mix the solutions, which were then allowed to stand for 2 minutes at 25°C. At 595 nm, the absorbance of each protein solution was measured. Using the absorbance values of the standard protein solutions, a standard curve was created, and the concentration of the protein in the unknown samples was calculated using the standard curve.

2.3.3 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE):

SDS-PAGE was used to separate protein samples based on their molecular mass using the Laemmli technique.

2.3.3.1 Stock solutions:

a) Acrylamide solution or solution A: To make a 30% acrylamide solution, 29.2 gm of acrylamide and 0.8 gm of N, N/-methylene-bis-acrylamide were dissolved in 30 ml of distilled water. Distilled water was used to bring the volume up to 100 ml. The solution was filtered and placed in an amber-colored container for safekeeping.

b) Separating gel buffer or solution B: 1.5 M Tris-HCl, pH 8.8 was prepared by dissolving 18.13 gm of Tris base in 90 ml of distilled water, adding 0.4 gm of SDS, adjusting the pH to 8.8 with 3 M HCl, and bringing the final volume of the solution up to 100 ml with glass distilled water.

c)Stacking gel buffer or solution C: 0.5 M Tris-HCl, pH 6.8 was prepared by dissolving 6.0 gm tris base in 80 ml distilled water, adding 0.4 gm SDS, adjusting the pH to 6.8 with 3 M HCl, and bringing the final volume of the solution to 100 ml with glass-distilled water.

d) Ammonium persulfate (APS): 20 mg of APS was dissolved in 400 l of glass distilled water to make a 5% solution. Glycerol is a 50% glycerol solution. To make a 50 percent solution, combine 50 mL of glycerol with 50 mL of glass distilled water.

e) TEMED – N, N, N/, N/, -tetramethylethylenediamine: APS was made fresh for maximum effects, while the other stock solutions were kept in the refrigerator and utilized within a month.

2.3.3.2 Working solutions:

f) Separating gel solution: The stock solutions were combined in the following quantities for a 10% gel: Solution A 3000µl Solution B 2250µl distilled water 1250µl 50% glycerol 2500µl TEMED 5µl 5% APS 70µl.

g) Stacking gel solution: The following proportions were used to combine the stock solutions: Solution A is 900µl, whereas Solution C is 1500µl. 3600µl of water, distilled just before polymerization, TEMED (6 l 5% APS, 35 l APS and TEMED were added to each solution.

h) Electrophoresis buffer: The electrophoresis buffer was prepared by dissolving the following ingredients in 25 mM Tris-HCl. TRIS-HCl, 190 mM glycine, and 0.1 percent SDS are the following:3.0 gm tris-baseGlycine 14.3 gm SDS1 gm to make a final capacity of 1000 ml, glass distilled water was added.

i) Laemmli SDS sample denaturing buffer: 100 mM Tris-HCl, 25 mM dithiothreitol (DTT), 50% glycerol, 10% SDS, and 0.025 percent bromophenol blue dissolveIn 1.25 mL Bromophenol blue, combine 48 mg Glycerol, 0.5 mL DTT, and 240 mg SDS.0.75 mL (12.5 percent) water5 litres The sample denaturing buffer was kept at -20°C for up to one month.

2.3.3.3 Procedure:

Polymerization: For gel electrophoresis, Miniprotean gel equipment was used (BioRad). The glass plates were put together according to the instructions provided by the manufacturer. The separating gel solution, comprising APS and TEMED, was poured up to three-fourths of the total height between the two glass plates. The stacking gel solution was carefully poured on top of the separating gel to

avoid disturbing the separating gel's level. In the stacking gel solution, a 0.75 mm thick 10-well teflon comb was introduced. The gel solutions were allowed to polymerize for 20–30 minutes at room temperature before the comb was carefully removed. The wells were inundated with water, extensively cleansed to eliminate unpolymerized acrylamide, and then replaced with water until ready to be loaded. The bottom and top chambers of the equipment were filled with electrophoresis buffer. The gels were inserted into the chamber without causing any bubbles to form beneath the glass plates.

Sample preparation: For full denaturation, samples were mixed with a fifth of their volume of SDS sample denaturing buffer and cooked for 3 minutes in a water bath.

Sample loading and electrophoresis: Individual wells in the stacking gel were filled with different samples. To determine the molecular weights of sample proteins, standard marker proteins with known molecular masses were loaded. The electrophoresis buffer was then carefully poured into each well. The electrophoresis chamber was connected to a power supply, and electrophoresis was performed for 45 minutes at a constant voltage of 200 volts, or until the bromophenol blue dye front was 0.5 cm above the gel's base.

2.3.3.4 Protein staining and destaining: 0.3 gm of brilliant blue R250 was dissolved in 100 ml of the destaining solution to make the staining solution (0.3 percent). The solution was filtered and kept in an amber-colored container using Whatman No. 1 filter paper.

To make 50% methanol, 10% acetic acid, and 40% water, 500 mL of methanol, 100 mL of acetic acid, and 400 mL of glass distilled water were used.Following electrophoresis, the gels were gently mixed in a 0.3 percent Coomassie blue solution for 20 minutes at room temperature.After that, the gels were destained in a destaining solution. Destaining was done until the background was sufficiently clean and the protein bands could be seen clearly. Finally, the gels were submerged in Millipore water.

2.4 Purification of protein

After successful transformation, a single colony was selected from each plate and inoculated into 5ml of antibiotic-laced LB medium and cultivated overnight. Primary cultures were put into 500ml of LB medium that had been prepared separately. Protein synthesis was stimulated by 1mMIPTG and incubation was prolonged for 4 hours after 4 hours of vigorous shaking, when cell density reached an OD600 of 0.6. Bacterial cells were extracted and resuspended in sonication buffer after centrifugation at 6000rpm for 10 minutes. Cell debris was removed by centrifugation at 12000 rpm for 45 minutes after the resuspended bacterial cells were lysed using a sonicator (Q-Sonica 125). Protein was overexpressed and soluble, so the supernatant was treated with Ni-NTA agarose beads for 45 minutes, then pre-equilibrated with equilibration solution and passed through a Ni-NTA purification column. Proteins (His-tagged) were eluted using elution buffer after a thorough wash in the washing buffer. SDS-PAGE was used to verify the purity of all protein samples. A HiLoad 16/60 Superdex 200 pg Size Exclusion Chromatography (SEC) column was equilibrated. Throughout the chromatography, a flow rate of 0.8 ml/min was maintained and protein (2–5 mg/ml) was introduced into the column. Few no-His tagged protein was purified by Ion-exchanged chromatography followed by SEC. SDS-PAGE was used to collect and examine fractions corresponding to the SEC peaks (Figure 2II, 2I). The following gel filtration markers were employed for SEC column calibrations (molecular weight and related elution volumes are identical): ferritin (440 kDa51 ml), aldolase (158 kDa62 ml), albumin (66 kDa76 ml), carbonic anhydrase (29 kDa84 ml).

2.5 Crosslink reaction

Crosslinking is the best option to understand if a protein has a tendency to form dimers, trimers, or higher-order oligomers. The cross-linking procedure was carried out for 5 minutes at 27 °C using freshly manufactured 2% glutaraldehyde (Fischer Scientific). Prior to the crosslink experiment, the protein sample was dialyzed in phosphate buffer (20 mM Phosphate, pH 8, 120 mM NaCl). When

boiling the sample, the reaction was halted by adding Tris SDS-PAGE buffer, followed by SDS-PAGE analysis.

2.6 Native gel electrophoresis

Another good way to figure out if a protein tends to form dimers, trimers, or higher-order oligomers is to run it using native-PAGE. Proteins are separated in native PAGE based on the charge, shape, and size of their tertiary structure. SDS was not present in the 10% polyacrylamide gel used for native gel electrophoresis. The entire procedure took place in a cold room.

2.7 Biophysical analysis of protein

CD (circular dichroism), FTIR (Fourier transform infrared spectroscopy), and fluorescence spectroscopy were used to determine the protein's secondary structure. BLI (bio-layer interferometry) was used to look for protein-protein interactions, while AFM provides information on protein organization.

2.7.1 Far UV CD spectroscopy study

Far UV CD spectra were recorded by Jasco J-815 spectrophotometer. Spectra were recorded for 20µM of proteins from 240 to 190nm with a scan speed of 10mm/min. 1cm path length cuvette was used. Buffer Spectrum was subtracted from the protein CD spectra. Experimental data were analyzed by the K2D2 web server (Louis-Jeune et al., 2012).

2.7.2 FTIR spectra

For the Fourier-transform infrared spectroscopy (FTIR) experiment, protein concentrations of 20 mM were utilized. To enable reading, 10 samples were drop-cast onto an FTIR substrate. Before each spectrum, baseline subtractions were performed as a buffer. Curve fitting shape was used to deconvolve raw spectra in the amide I region (1600 cm^{-1} to 1700 cm^{-1}) using Prism software. The

criteria suggested by Chirgadze&Nevskaya (Nevskaya & Chirgadze, 1976) and Krimm&Bandekar (Krimm & Bandekar, 1986) were used to assign peck values.

2.7.3 Fluorescence spectroscopy study

Sample fluorescence spectra were measured at room temperature using a spectrofluorometer by excitation at 295 nm. A 10 mm path length quartz cuvette was used in all experiments. For fluorescence measurements, both excitation and emission slits were 5 nm and the scan speed was set at 200 nm/ min.

2.7.4 Biolayer interferometry

In biolayer interferometry (BLI) experiments, the OctetRED96e with Ni-NTA coated biosensor tips was used. A sample volume of 200 L was employed on 96 well plates during the experiment, which was maintained agitated. The Ni-NTA biosensor tips were hydrated in a buffer containing 60 mM Tris pH 8.0, 60 mM NaCl, and 1% glycerol throughout all of the tests. The biosensor tips were dipped in buffer solution without any analyte for baseline adjustments. After adjusting the baseline with a buffer solution comprising tris-buffer (pH 8), the biosensor tips were incubated for 300 seconds in the well containing the ligand (PspF) at a concentration of 40M for ligand binding to the biosensor tips. After removing any excess ligand by dipping the ligand-bound biosensor tips in the buffer solution for 60 seconds, the biosensor tips were incubated for 120 seconds in wells containing escalating concentrations of the untagged analyte (PspA) ranging from 2.5 to 70 M. The tips were submerged in the buffer solution for another 180 seconds to explore the dissociation process. His-tagged Protein, which bound 40 uM to Ni-NTA biosensor tips, was the analyte in the Biolayer interferometry assays. The tips were then immersed in various concentrations of cognate protein-containing solutions. Despite the fact that His-tag binding affinity to Ni-NTA is pH-dependent, there is no discernible dissociation pattern.

2.7.5 AFM experiment

A 5M pure protein sample was incubated for 1 minute before being diluted at least 100 times in 0.22m filtered milliQ, and 2l of this sample was placed in the center of a 20mm diameter MUSCOVITE MICA chip (Electron Microscopy Sciences, Hatfield, PA). The sample was left to air dry at an ambient temperature. On a Pico Plus 5500 AFM instrument, the acoustic alternative current (AAC) mode was evaluated. Microfabricated silicon cantilevers with a nominal spring force constant of 21–98 N/m were tested using nanosensors (PPP-NCL). The cantilever scanned the surface at a speed of 0.8-1.2 m/s with a resonance frequency of 168.4 kHz. Micrographs of the surface topology of the protein samples were produced after scanning each sample multiple times, and photos were processed using Pico View 1.10.1 software (Agilent Technologies, USA) and Pico Image Advanced software for 3D rendering and height profile analysis. All AFM experiments were repeated at least three times, with newly pure materials each time.

2.8 Bioinformatic study of protein

2.8.1 Sequence retrieval, generation of protein model

The protein was predicted using the SOPMA online program and the UniProt public protein database (UniProt ID: Q9I326 for PcrG and A0A0E1NET1 for PspA) respectively. The secondary structure was predicted using a FASTA file, and the amino acid sequences were obtained from an online database. Both protein sequences were fed into I-Tasser as input files in FASTA format (A. Roy et al., 2010). The top ten threading templateswere assigned by I-Tasser's programs. The best model was rendered in the VMD molecular graphic system (Humphrey et al., 1996) using the output files. The optimal frame with the lowest potential energy was chosen after the minimization. Various bioinformatics tools and software are used such as RAMPAGE and ERRAT were used to check the structural assessment and stereochemical analysisof both modelled proteins.

2.8.2 Docking simulation

ClusPro, an online docking service, was used to bind proteins together. By sampling billions of conformers, this server conducts rigid-body docking of two proteins. Interface refers to the area where two or more proteins interact with one another.

2.8.3 Molecular dynamics

GROMACS 4.6.2 was used to run the simulations (van der Spoel et al., 2005). The simulation experiments used the docked complex following ClusPro as an input structure. For the protein monomer or dimer, the CHARMM27 force field was employed. A TIP3P, where the box size is 1.2 nm, was used to place the protein-protein complex in a unit cell and fill it with water. Additional ions were introduced to neutralize the system, and energy was reduced by using the steepest descent approach. For electrostatic, van der Waals interaction, and energy computations, GROMACS uses the particle mesh Ewald (PME) approach. The complexes were equilibrated using NVT and NPT for 500 and 500 ps, respectively, after minimization. The backbone is still subjected to position limitations. A Nose-Hoover thermostat has been installed in lieu of the Berendsen thermostat, resulting in a more accurate ensemble of kinetic energy. The system was then put through a 40-ns MD manufacturing cycle. MGRACE was used to create the C-alpha backbone and RMSF graphs.

2.8.4 Umbrella sampling for binding free energy calculation

The dimer model structure was put in a rectangular box with dimensions adequate to fulfill the minimal image convention and enable room for pulling simulations to take place along the z-axis in order to build an equilibrated final simulated structure for the pulling simulations. The solvent was simple point charge (SPC) water31, and 100 mM NaCl was added, along with neutralizing counter ions. The model structure was equilibrated in two phases after the steepest descent minimization. Simulating for 100 ps under a constant volume (NVT) ensemble was the initial phase. The

temperature was maintained at 310 K using the Berendsen weak coupling method (32) and protein and nonprotein atoms were connected to separate temperature coupling baths. After NVT equilibration, 100 ps of constant pressure (NPT) equilibration was carried out, with weak coupling32 used to keep pressure isotropically at 1.0 bar.

Restraints were removed from chain A to chain B after equilibration. For the pulling simulation, Chain A was employed as an immovable reference. Using a spring constant of 800 kJ mol^{-1} nm^{-2} and a pull rate of 0.01 nm ps^{-1}, Chain B was drawn away from the core structure along thez-axis (Figure 1) over 500 ps. Snapshots of these trajectories were used to build the umbrella sampling's beginning configurations. The sample windows were distributed asymmetrically, with window spacing of 0.1 nm up to 2 nm COM separations and 0.2 nm beyond 2 nm COM separations.The use of 30 windows and a total of 60 ns of simulated time for umbrella sampling came from this spacing, which allowed for more information at a lower COM distance. The weighted histogram approach was used to analyze the data (WHAM).

2.8.5 Protein's salt-bridge and its energy estimation:

Non-covalent interactions such as hydrogen bonds, salt bridges, and aromatic-aromatic (AA) contacts play a critical role in sustaining biological systems by interacting between two or more proteins and within protein structures. The interface is a specific area where two or more proteins interact with one another. The inter-atomic SB and A-A distance in inter-chain proteins is determined using the equation below by ASBAAC, automatic Salt-Bridge and Aromatic-Aromatic Calculator (C. Roy & Datta, 2018), which is a completely automated, user-friendly program (fig 2.1A).

$$dist = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2 + (z_a - z_b)^2}$$

Intra-chain salt-bridge energy calculations are also important for better understanding protein folding across domains. Salt bridges are electrostatic interactions between groupings of opposing charges that have a particular electrostatic interaction. The net interaction energy (ΔΔGnet) of a salt bridge is divided into three energy terms: bridge (ΔΔGbrd), desolvation (ΔΔGdsolv), and protein (ΔΔGprot), the latter two of which can only be estimated using computational methods. Computation of salt-bridge energy terms utilizing the Poison-Boltzmann-Equation (PBE) solver technique is a time-consuming procedure that requires determining protein-specific salt bridges, input parameters, and hydrophobic isosteres-mutated charge-radius files before computing energy-terms. ADSBET (Nayek et al., 2015) uses the "APBS" technique to calculate component and net energy terms for any number of salt bridges in any number of structure files and then redirects the result to excel. The output is valuable not only for applying these energy terms after the run but also for calculating the net energetic contribution of networked salt bridges. It runs in a UNIX-like environment and supports CYGWIN. Overall, ADSBET gives precise energetics of salt bridges, from different crystal structures, which can also be used in computational structural biology. Using the PDB2PQR (Dolinsky et al., 2007) and APBS (Bruce et al., n.d.) programs, the isolated pair approach was employed to derive the isolated-pair salt-energy bridge's terms (Nayek et al., 2015, Roy et al., 2020). Because NU, unlike IP, has more than two salt-bridge partners, the energy terms for NU salt-bridge were derived using the NUM approach (AK et al., 2020). The total of the component energy terms is the net energy of a salt bridge. The NACCESS software was used to determine average accessibility (Hubbard: NACCESS: Program for Calculating Accessibilities - Google Scholar, n.d.).

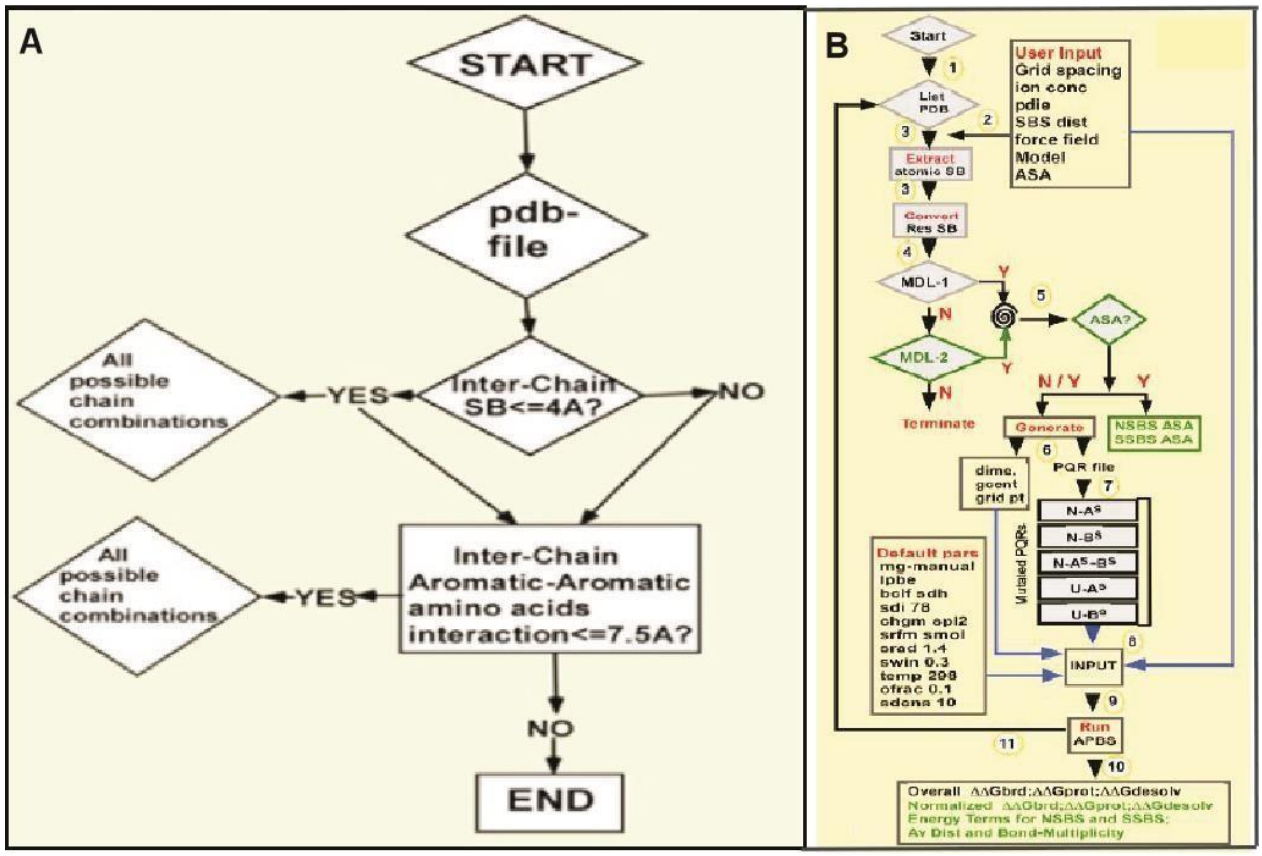


Fig 2.1. Flowchart of software's function. **A.** flowchart illustrating ASBAAC's function the scanning of the pdb file is the first function. If the distance between the atomic SBs is less than 4 (YES), ASBAAC calculates all possible chain combinations and proceeds to the next step. If the distance between the atomic SBs is greater than four (i.e., NO), it comes out and advances to the next level. The program then moves on to the next lap to calculate the pi-pi interaction. If the distance is less than 7.5 (YES), it calculates all possible chain combinations until the program ends (doi: 10.6026/97320630014164). **B.** N-AS and U-As are acidic side-chains in the folded and unfolded state of protein respectively. Additional features in ADSBET2 over its earlier version are shown in green color (doi:10.6026/97320630011413).

2.9 Protein crystallization

Protein crystallization was discovered by chance around 150 years ago and developed as a potent purification tool and a show of strength in chemical purity in the late nineteenth century (Giegé, 2013). The formation of a solution that is supersaturated in the macromolecule but does not significantly perturb its natural state is required for protein crystallization. Supersaturation is achieved

by adding mild precipitating agents such as neutral salts or polymers, as well as manipulating various parameters such as temperature, ionic strength, and pH (McPherson & Gavira, 2014). Factors that can affect the structural state of the macromolecule, such as metal ions, inhibitors, cofactors, or other conventional small molecules, are also important in the crystallization process (Weber, 1991).

2.9.1 Stable domain crystallization and data collection of PspA

Microcon centrifugation was used to concentrate the sample obtained directly from gel filtration to 10 mg/mL (Amicon, 10 kDa cutoff). By mixing protein and reservoir solution, the sitting-drop vapor diffusion techniques were used to screen for suitable crystallization conditions. Commercial crystal screens (Wizard, Jena Bioscience), were used. At 4°C (cold room), purified protein samples were incubated. Following the failure of extensive crystallization attempts to produce diffraction-quality crystals. All conditions are described in table 2.2. crystallization was further optimized. 500 nl of protein solution (10 mg/ml) was mixed with 500µl of reservoir solution containing 28 percent (w/v) PEG 400, 200 mM CaCl and 0.1 M HEPES, pH 7.5, at 277 K. After 3 days, rod-shaped crystals formed. The crystals were diffracted using *in house* Bruker machine, equipped with PHOTON III CCD detector. A total of 720 frames were acquired with a crystal-to-detector distance of 136 mm. Each image was exposed for 30 seconds, and the oscillation range was kept at 0.5^0C. The Proteum-4 software was used to index, integrate, and scale data. The crystals were from the P32 space group. P3221 was identified as the point group based on systematic absences. Using PspA (PDB-ID: 4whe) as the search model, the first attempts to determine phases via molecular replacement in PHASER (McCoy et al., 2007a) were unsuccessful. Rosetta model building coupled molecular replacement (MR-Rosetta) (Terwilliger et al., 2012a) may provide a partial solution later. After manual model creation in Coot (Emsley & Cowtan, 2004) and alternate macro-cycles of iterative refinement using PHENIX.refine (Terwilliger et.al., 2012b) , the entire model was achieved. MolProbity and other validation tools in

PHENIX were used to perform stereochemical and statistical validation of the model. PyMOL (org & 2002, n.d.) was used to construct structural figures.

Table 2.2 Δ PspA protein crystals occur under particular conditions.

No.	Precipitation Reagent	Buffer	Salt	
1.	10% w/v PEG6000	100mM HEPES pH 7.5		5% v/v MPD
2.	20% w/v PEG8000	100mM MES pH 6.0	200mM Calcium acetate	
3.	20% v/v PEG550 MME	100mM BICINE pH 9.5	100mM NaCl	
4.	20%w/v PEG8000	100mM CAPS pH 10.5	200mM NaCl	
5.	20% v/v PEG400	100mM HEPES pH 7.5	200mM CaCl2	

2.9.2 Crystallization trial of PcrG and PcrG-PcrV complex

Microcon centrifugation was used to concentrate purified proteins (or protein complexes) acquired through gel filtering. Sitting-drop vapor diffusion techniques were utilized to screen for optimal crystallization conditions by mixing protein and reservoir solution. Crystal screens from Wizard and Jena Bioscience as well as systematic grid-screening methods were used. On the crystal screen, some tiny-shaped (Fig. 2.2A1, A2), as well as brush-shaped (Fig. 2.2B1, B2) crystals, formed, but these were not reproduced following severe optimization.

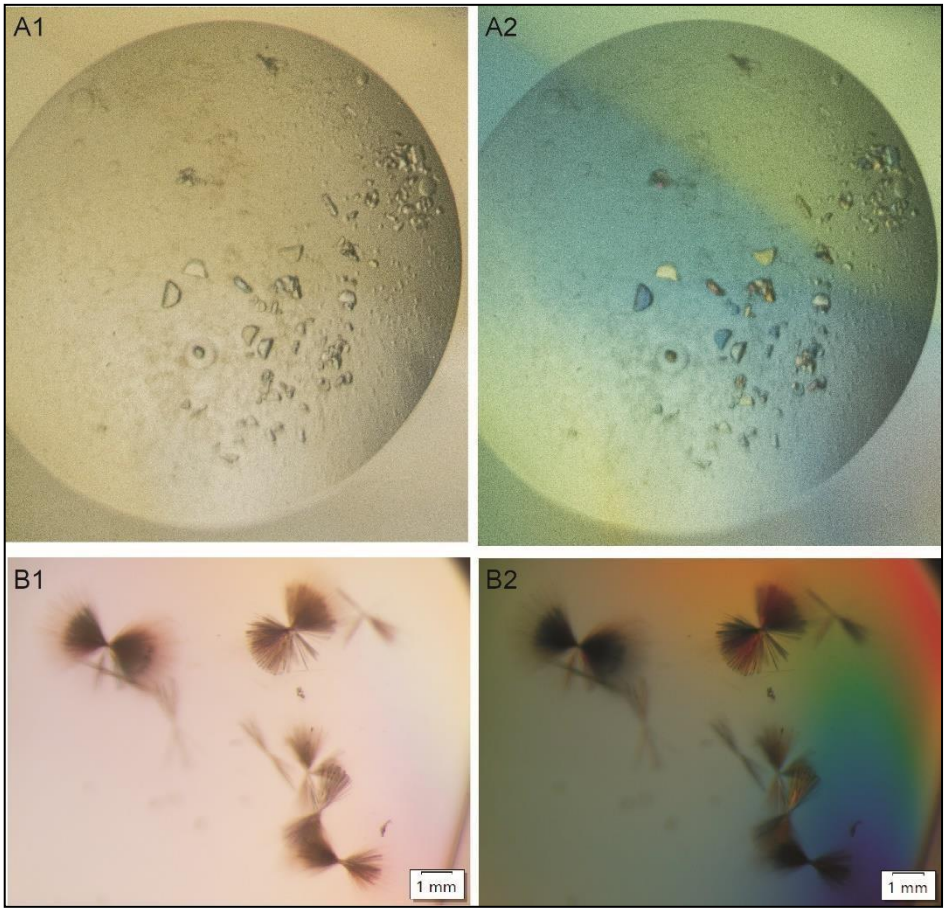


Fig 2.2 Crystals of PcrG and PcrG-PcrV complex. After one month in the phosphate buffer condition, tiny shaped crystals of PcrG formed. However, after four months in the presence of 100 mM Imidazole (pH 8), 33%v/v MPD and 200mM MgCl2, a few brush-shaped crystals of the PcrG-PcrV complexcrystal appeared.

CHAPTER 3

3. The binding energy of the dimeric structure of PcrG, a type three secretionsystem's tip chaperon of *Pseudomonas aeruginosa*, has been revealed via umbrella sampling.

3.1 Introduction:

The gram-negative bacteria *Pseudomonas aeruginosa* is an opportunistic and strong pathogen that causes acute infections in immunocompromised people (Sadikot et al., 2005). It is mostly responsible for nosocomial pneumonia and other infections caused by burns, wounds, urinary tract infections, and cystic fibrosis (Bassetti et al., 2018). The Type 3 Secretion System (TTSS) is used by *Pseudomonas aeruginosa* to transmit pathogenicity into the host cell. It's a one-of-a-kind pathogenicity-spreading mechanism found predominantly in gram-negative bacteria (Halder et al., 2019). The injectosome, which is made up of 20 highly conserved proteins (Yahr et al., 1997) and aids in the transport of bacterial toxic effector proteins into host cells, is an unavoidable and necessarycomponent of TTSS (Galle et al., 2012b, Cornelis, 2006). The injectosome is a well-defined modelthat includes a base structure and a needle complex (Lombardi et al., 2019). A translocon is generatedat the tip of the needle by three translocator proteins: PcrV, PopB, and PopD. Two of these translocators (PopB and PopD) are hydrophobic, forming holes in the host cell membrane, while thehydrophilic translocators (PcrV), also known as virulence factors or V-antigen, aid in the assembly of the hydrophobic translocators (Goure et al., 2004). PcrV (or V-antigen) is a hydrophilic translocator that interacts with PcrG in the cytosol, an effector chaperone of P. aeruginosa's Type 3 Secretion System (Allmond et al., 2003). The needle tip protein chaperone PcrG regulates both the secretion activity of apparatus and the secretion of effectors from the bacterial cytoplasm into the hostcytoplasm (Lee et al., 2010). These two roles are tightly regulated by two distinctareas. In addition to its export chaperone activities, PcrG also acts as a break and make switch for regulating effector secretion (Lee et al., 2014). Although deletion of the 24 C-terminal amino acids of PcrG does not affect the functional activity through complex formation, experimental studiesindicate that PcrG and PcrV interact at a 1:1 ratio (Nanao et al., 2003; Sundin et al., 2004). Apart from the knowledge that it is a negative regulator of TTSS, nothing is known about the actual mechanism of PcrG regulation (Sundin et al., 2004; P. C. Lee et al., 2010b). The deletion of pcrG or

pcrV led to a partial deregulation of effector secretion, but the removal of both genes resulted in a high degree of effector secretion (P. C. Lee et al., 2010b). Though PcrG's 40 N-terminal amino acids are enough to export its homologous substrate PcrV (11), the N terminal 40aa plays no function in regulation (P. C. Lee et al., 2014). In experiments, PcrG fused to the C terminal region of MBP has been shown to aid in the stability of PcrG fragments. PcrG is a 98-amino-acid polypeptide chain with a single component. It has one tryptophan (W37) side chain, one tyrosine (Y7) side chain, one phenylalanine (F51) side chain, and no disulfide bond. Furthermore, charged residues make up roughly a third of all amino acids. The NMR structure of PcrG also indicated that there is no tertiary structure, but that it does include a secondary structure dominated by two lengthy helices (Chaudhury et al., 2015).

With the aforementioned concerns in mind, we addressed the protein's structure, folding, and functions using a multidisciplinary approach. For the first time, the protein's dimeric state in solution, its specific and overall interface stability, and the issues of its most flexible region are depicted. We believe that this work will have an impact that can be applied to other similar systems.

3.2 Results

3.2.1 In solution, helix-enriched PcrG exists as a dimer.

Pseudomonas aeruginosa PcrG (UniProt ID: Q9I326; Pfam ID: PF07216) is a type-III secretion apparatus component that negatively regulates the secretion of the host-infective protein. This protein's sequence and composition, which is capable of protein-protein interaction, are very distinctive (Fig. 3.1). This small protein (98 residues) is formed by the seventeenth type of residue (Lys, His, and Cys are absent). The protein has a helix tendency (Glu, Arg, Gln, Ala, Leu) in 2/3 of its residues, followed by a coiled structure (Fig. 1A, B), according to SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) analysis. The ProtParam analysis of the protein sequence ("The Proteomics Protocols Handbook," 2005) reveals that it has many interesting properties. Because two basic residues are missing from the PcrG sequence, the

acidic residue takes precedence (Fig. 3.1A), and the protein has a pI of 4.8. Despite the fact that the amounts of hydrophilic and hydrophobic residues are equal (50 percent each), the sequence'sGRAVY value is -0.58. It is also mentioned here that the amount of charged residue (30.6 percent) is significantly higher than that of polar residue (18 percent). Furthermore, Gly is the order structure breaker, which is why the 3D structure's flexibility is kept at a high level (9 percent). It is worth noting that the protein's sequence contains only one Trp and one Tyr. Furthermore, the protein's instability index (i.e., 52.12) is very low.

According to the final purification step of this protein using the gel filtration method, its molecular weight is close to 30 kD (Fig. 3.1C & inset). Running an SDS-PAGE in the absence and presence of glutaraldehyde, which acts as a cross-linker, reveals that the former corresponds to the monomeric form of the protein and the latter to the dimeric form (Fig. 3.1D, E). In this case, the gel filtration and cross-linked gel results are comparable. Furthermore, the protein's dimeric form is confirmed in the native gel (Fig. 3.1F).

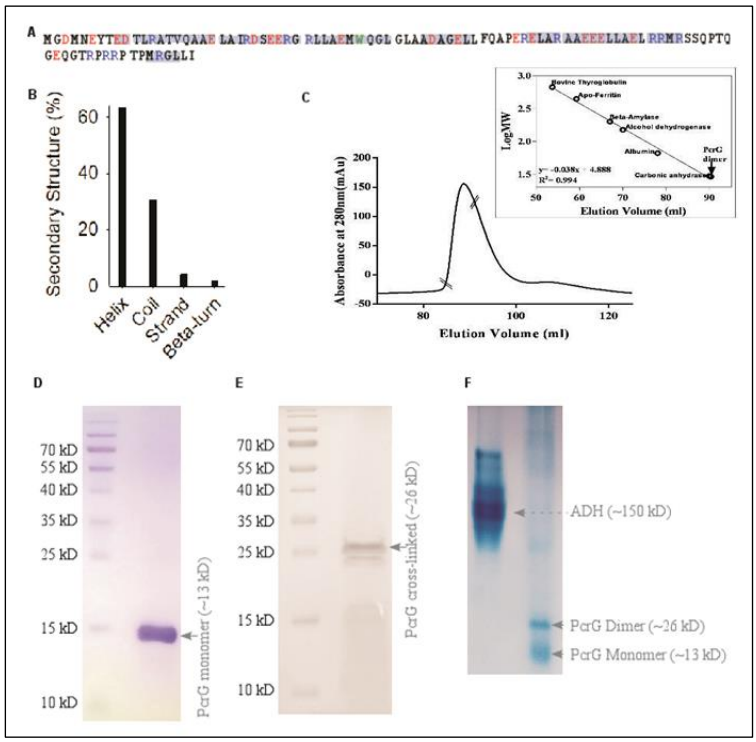


Fig 3.1 The gel and spectral properties indicate that PcrG exists in solution as a dimer. Sequence and SOPMA analysis of PcrG (Fig.1 A, B). The chromatogram of PcrG gel-filtration (elution maxima marked by dotted lines) and the log molecular weight vs. elution volume plot for standardproteins (inset) are shown (Fig.C). SDS-PAGE and cross-linked of the gel-filtered protein along with standard molecular weight markers (Fig.D&E) and Native gel along with alcohol dehydrogenase (ADH) as a marker (Fig.F).

Tryptophan fluorescence is an effective tool for studying the structure and dynamics of tryptophan's local environment in proteins (Vivian & Callis, 2001). In addition, the protein contains only one tryptophan (i.e., W37). Figure 3.2B depicts the protein's typical emission spectrum, which shows that the value of max is at 340 nm. In our secondary structure prediction, we discovered that the helix structure predominates in the protein (Fig. 3.2B). We performed a far-UV CD experiment on the purified protein to check the secondary structure in solution conditions. The characteristic double minima of ellipticity, one at 222 nm and the other at 210 nm, indicates that the protein has a helix structure (Fig. 3.2A). In addition to this helix structure, the protein's coil also predominance and is significant.

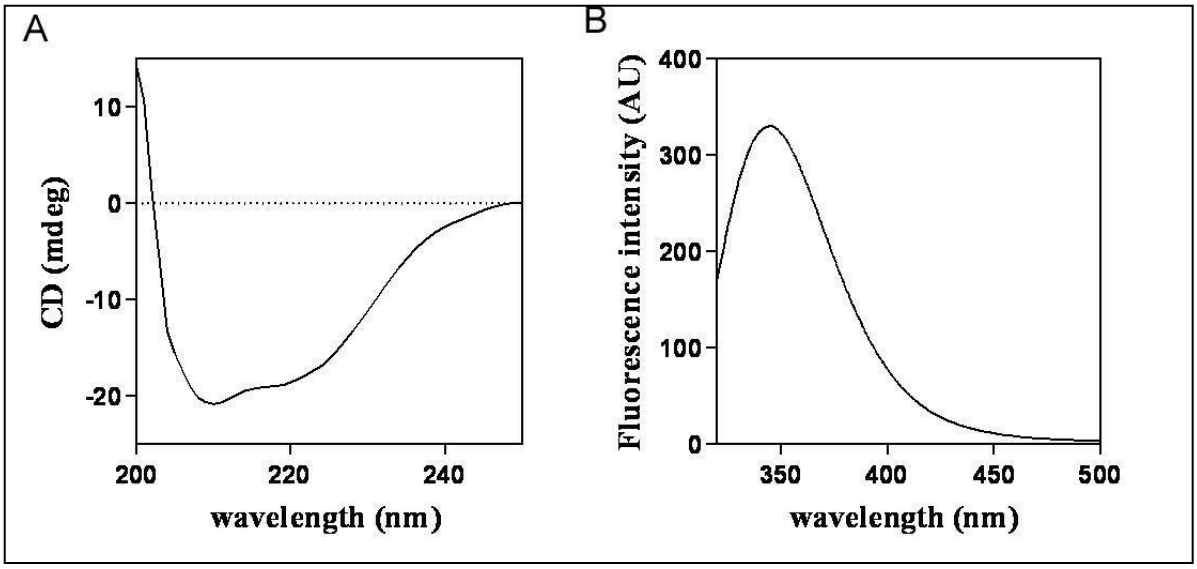


Fig 3.2 CD and fluorescence properties of PcrG. **A**, Typical fluorescence spectrum of the gel-filtered PcrG (10 μM) in 20 mM Tris-Cl buffer pH 8.0. **B**, Typical far-UV CD spectra of the gel-filtered PcrG (15 μM) in 20 mM Tris-Cl buffer, pH 8.0.

However, the protein's dimeric form in the solution is quite surprising. In this case, we want to know what the structure of this dimer is and how the interface of this dimer acquires stability in a dynamic state. To answer these questions, we used MDS and other methods to characterize the stability and other properties of the interface of the proteins in-silico dimer structure.

3.2.2 Certain residues and a salt bridge are required for the interface to be stable.

The interface is formed by the interaction of the inter-helix and the coil. Surprisingly, the majority of these interface residues are hydrophobic (Fig. 3.3B). In this section, we asked which of these residues has an interface property. To answer this question, we extracted the propensity of amino acid residues in an interface subunit (Fig. 3.3C). We can see that only a few residues have a proclivity to be at the interface, where polar (with Tyr having the highest proclivity) and hydrophobic residues predominate.

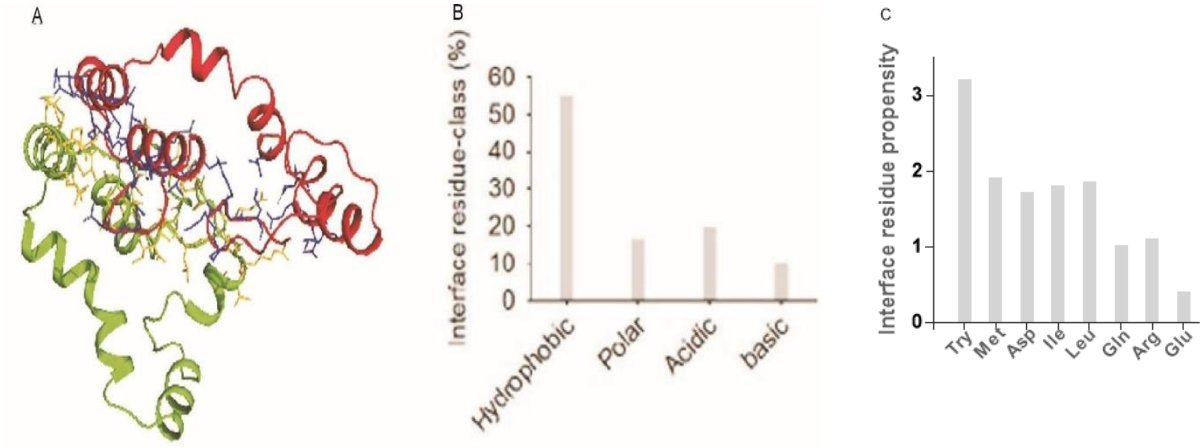


Fig 3.3 The homodimer interface of PcrG is very distinctive in terms of salt-bridge and residue type. A, the residues of the interface of the two subunits (red, A-chain and green, B-chain) are shown by different colors (blue, A-chain residue, and yellow, B-chain residue). **B**, The normalized frequency of the residue class of the interface (only for the A-chain of the homodimer). **C,** Interface residue propensity of dimeric PcrG.

However, in order to better understand the significance of salt bridges in interface stability, we conducted a salt bridge analysis (Table 3.1). Several points are worth mentioning here. To begin with, we can observe that the interface is made up of isolated (IP) and network (NU) salt bridges, with IP being the more common. Second, while these salt bridges have lower de-solvation energy, the bridge energy term is more favorable (Table 3.1). Finally, we can see that background

energy (**ΔΔ**Gbac) is lower in terminal salt bridges but higher in salt bridges created in the mid-region of the interface (Table 3.1). Finally, for salt bridges with higher background energy, the net stability, which is the sum of the above-mentioned component energy components, is larger.

Table 3.1 The interface forms a stable but limited number of IP and NU-type salt bridges. Salt-bridge (SB) component ($\Delta\Delta G_{dslv}$, $\Delta\Delta G_{bac,}$ and $\Delta\Delta G_{brd}$) and net ($\Delta\Delta G_{net}$) energy terms are presented along with bond-multiplicity (mul), average distance of interaction (av-dist) and average accessibility (avASA).

	Unit	**Component terms Kcal/mol**			**Kcal/mol**	**count**	**Å**	**Å²**
	SB	**$\Delta\Delta G_{dslv}$**	**$\Delta\Delta G_{bac}$**	**$\Delta\Delta G_{brd}$**	**$\Delta\Delta G_{net}$**	**mul**	**av_dist**	**avASA**
isolated	D3_R29	5.57	-1.31	-6.83	-2.57	4	3.42	35.4
	R24_D3	1.11	-0.53	-4.95	-4.37	4	4.04	66.2
	R31_E35	5.06	-8.82	-6.31	-10.07	5	4.07	33.5
	E35_R31	4.25	-9.13	-5.18	-10.06	3	3.92	35.4
network	R94-E63-R86-E64	14.35	-1.72	-22.32	-9.69	14	3.92	35.08

3.2.3 The stability of the interface of the dynamic state of the PcrG dimer was analyzed.

We were able to see the PcrG dimer from the actual solution state using a homology model and docking (Fig. 1A). The challenge here is whether this PcrG dimer structure can preserve its integrityin a dynamic state. To better understand this, we ran a dynamic for forty nanoseconds and performed some basic analysis (see Figure 3.4). The stability of the protein is determined by the RMSD pattern ofthe protein states occupied in the trajectory. After 5 ns, the trajectory pattern of the backbone RMSD of the PcrG dimer is nearly unchanged. Furthermore, the RMSD's overall average value (2.7) is extremely low.

Similarly, the RMSF value indicates the hardness and flexibility of each residue position's main chain in the protein states recorded throughout the route. Although the value of this RMSF after five nanoseconds is close to that of the normal native state (Sen Gupta et al., 2020), the hardness and flexibility of distinct regions of the PcrG dimer should be noted. The residues 15 to 40 (Fig 3.3A), which forms a Helix-bend-Helix-like structure and is involved in the compaction of the dimer interface (Fig. 3.3A), is stiffer than the rest of the yellow color. On the other hand, the residue area sixty to eighty, which forms a helix structure in connection with some coil structures and is furthest from the interface, is the most flexible. The borders of this stiffness and flexibility exist within the boundaries of the normal native state, based on the total average value (i.e., 0.30 nm).

As shown in Figure 3.4C, the SASA value is critical for comprehending the average assemblage of protein residues in connection to dynamic states. The SASA value of a protein is determined by the content of its residues as well as the nature of the residue (hydrophilic or hydrophobic). As previously stated, half of the PcrG residues are hydrophilic,while the other half are hydrophobic, with the former's side-chain interacting with the bulk solvent. The fact that the average SASA value is smaller (112 nm^2) than that of a typical monomeric protein(Gupta et al., 2020) could be due to PcrG's dimer state, which prevents bulk solvent from reaching the interface.

Rg values are used to understand the packing, compaction, and dimension of the dynamic states of a protein. The Rg value of PcrG is 1.46, and the variance throughout the trajectory is almost non-existent, suggesting that our protein is partially compact. This property means that the changes in these dynamic parameters (i.e., RMSD, SASA, and Rg) from the start to the end of the trajectory are essentially non-existent.

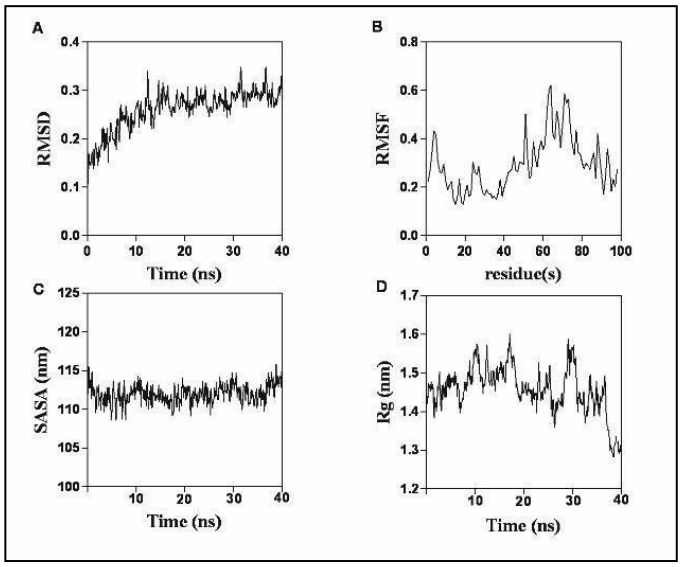


Fig 3.4 RMSD (A), RMSF (B), SASA (C) and Rg (D) of dimer PcrG. Here, the overall average value of the trajectory part of more than five nanoseconds in each case (A-D) is shown with the plot. The trajectory pattern of the PcrG dimer's backbone RMSD is nearly unchanged after 5 ns. The RMSF value indicates the rigidity and flexibility of each residue position in the protein states acquired along the trajectory. The SASA value indicates that half of the PcrG residues are hydrophilic and the other half are hydrophobic, where the former's side-chain interacts with the bulk solvent. Overall, the Rg value of PcrG is 1.46, with almost no variation along the trajectory, indicating that our protein is compact.

We used the umbrella sampling method to determine the overall strength of the interface (Fig. 3.5), and the strength of the PcrG interface is 54.07 Kcal/mol, which is fairly high. Because the interface containsa large number of hydrophobic residues, hydrophobic interactions, in addition to salt bridges, appearto play a role in the overall energetics of the PcrG contact.

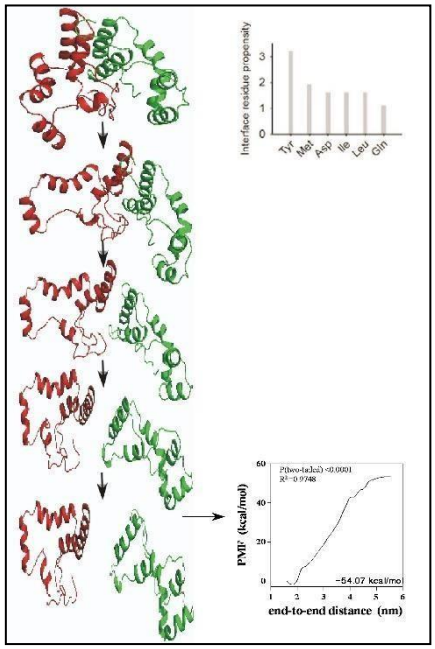


Fig 3.5 umbrella sampling and interface residues. In between are shown the initial and final state structures for umbrella sampling, as well as three intermediate structures. The potential mean force (PMF) is plotted against distance to calculate the overall interaction energy of the interface. The propensity of the interface residues in an initial state of dimeric PcrG also represents, that Tyr contribution is much higher than others.

3.3 Discussion

The study of the structure, stability, and interaction of the microorganisms' proteins that mediate the host's infection is extremely beneficial. PcrG from *P. aruginosa* is such a complex protein that many of these details remain unknown. We chose the current study with this in mind. Because the protein has a low residue count, it should be simple to characterize. As evidenced by the gel run, the protein is present in the solution as a dimer. Many chaperones are also present in the solution in dimeric form (Birtalan & Ghosh, 2001), according to other research (Birtalan & Ghosh, 2001; Ye et al., 2018). The fluorescence spectrum probes the dynamic state of the protein tertiary structure. The spectral characteristics in the native state indicate that each subunit's tryptophan exists in an asymmetric environment. The relative accessibility of the side-chain of this tryptophan is 162, according to the NACCESS study of the homology model of the protein dimer. As a result, we predict a mild intensity from this partially exposed form of tryptophan, which, if strong enough, could indicate the dimer's condition (i.e., two tryptophan probing). We confirmed that the protein is helix enriched using circular dichroism, which was followed by coil structure, as expected (Siddiqui et al., 2017). Several in-silico methods were used to get a deeper insight into the properties of the true solution state of PcrG. The protein sequence's physicochemical properties reveal a lot about its tertiary structure. The reason for the low GRAVY value is that bulky hydrophobic residue like Ile, Val, Phe, and others, which have large positive GRAVY values, are toward the bottom of the sequence. The instability index shows that the protein sequence's overall expected stability is lower than that of a normal monomeric protein (Rumfeldt et al., 2008). This shows that the parts of the interacting residue at the interface may be the source of the instability. This type of instability, which is used to achieve contact specificity (in this case, dimer formation), is not rare in other proteins. In general, it has been shown that the average accessibility of salt-bridge partners influences the results of component energy terms such as desolvation and bridge energy terms. The background energy term is mostly indifferent in this case (AK et al., 2020; S. Banerjee et al., 2021). In our example, all salt bridges are present in the

interface's surface region. As a result, the desolvation energy term should have a smaller value. The bridge energy, on the other hand, was predicted to be low (Kumar & Nussinov, 1999), which was not the case. The bridge energy, on the other hand, was predicted to be low, which was not the case. The recruitment of favored pairs of salt bridges out of all available pairs of salt bridges results in higher-than-expected bridge energy. It is indicated here that the RD and RE pairs, which are the pairs in our situation, have a bridge energy advantage (Blasie & Berg, 2004; Meuzelaar et al., 2016). Furthermore, because the partners of these selected pairs have a high helix propensity, they may enrich the helical structure at the interface. The presence of interchain helices at the interface is thought to aid in the compaction of the interface. Although salt bridges improved the interface's stability (C. Roy & Datta, 2018), there are only five salt bridges on the interface. Furthermore, hydrophobic residues, which are unrelated to salt-bridge mediated stability, have a greater interface. In other words, aside from the salt bridge, hydrophobic force is expected to play a considerable role in the stability of the interface due to the prevalence of the hydrophobic class of residues. We remark that the energy of the salt bridge was recovered using the APBS method, which is a Poisson-Boltzmann-based multiparameter numerical solver (AK et al., 2020). It's an ad hoc procedure that can diverge from precision in high salt and pH situations. It is also noted that in this calculation, no estimation of hydrophobic force is conceivable. During this calculation, we employed mild or default APBS run-time options.

Molecular modeling and docking techniques were used to create the dimer PcrG. We analyzed several molecular dynamic parameters over a forty nanosecond MDS trajectory to rule out the staticsense of the dimer structure and check the variation of the dimer state in a dynamic situation, despitethe fact that our model structure has been well qualified using several validation procedures. We alsoused the umbrella sampling approach to determine the overall interface stability of the dimer interface. The RMSD data not only highlights the MD trajectory's equilibrium and convergence, butthey may also suggest the dimer PcrG's excellent stability. After 5 ns, the initial and end states of the trajectory, including the intermediates, are nearly identical. Surprisingly, the RMSF data variance

aids in the identification of the dimer interaction site, which is a helix-bend-helix-like structure. The lower RMSF value in this region could be attributed to less fluctuation of interacting side-chains at both chains' interfaces. The lack of such inter-helix interactions could explain why the fluctuation for the higher RMSF value region is so much larger. It's worth noting that, while there is some variance in the RMSF values, it looks to be within the bounds of a typical MD trajectory (Akhilesh et al., 2021). Similarly, SASA and Rg result back up the idea that the native dimer state is quite stable during the MD run, resulting in a very stable structure.

It is quite difficult to comprehend the protein-protein interface ensemble stability. Notably, we only acquire the precise electrostatic contribution by estimating the interface salt-bridge net stability. Hydrophobic residues are more frequent than other residue types in the PcrG dimer. As a result, understanding ensemble stability not only allows us to understand why the PcrG dimer is so stable, but also allows us to understand the contributions of other components (hydrophobic, electrostatic, and so on). To approach a true ergodic system, we used the umbrella sampling technique, which carefully evaluates the sampling of states with similar ranges of variation in potential energy barriers (Torrie & Valleau, 1977). We used the same method as Dom et al. (2015) and calculated the total strength of the contact to be-54.07 Kcal/mol. It's worth noting that particular electrostatic stability alone contributes 30 Kcal/mol to interface stability, implying that hydrophobic and other forms of interactions account for the remaining 24 Kcal/mol. Although hydrophobic residues contribute to the above-mentioned electrostatic energy, their contribution is minor. We expected hydrophobic interactions to outnumber other interactions since they occur more frequently than other kinds. There are a couple of things to keep in mind here. First, the number of hydrophobic residues in each chain may induce an entropic loss of contact energy in order for the interface to be compactly packed. In the case of the salt bridge, such congestion is ruled out because there are only a few people who may be in the hotspot on the entire interface. Second, the APBS Poisson-Boltzmann electrostatic energy is an approximation that is dependent on the computation model as well as many run-time parameters.

3.4 Conclusions

Despite the fact that these procedures were obtained independently, they give the impression that the contact is quite stable, with contributions from both electrostatic and hydrophobic mechanisms.

CHAPTER 4

4. The finding of pH-dependent conformational changes in PspA in *Yersiniaenterocolitica* is facilitated by biophysical and computational approaches

4.1 Introduction

PspA, a member of the PspA/IM30 family of adaptor proteins, regulates pspA-E and pspG transcription with the help of RNAP σ54 mediated transcription activator and ATPase partner protein of the cytoplasm, PspF (Osadnik et al., 2015). Not only that, but in a cellular stress situation, PapA interacts with membrane-bound PspBC proteins (Osadnik et al., 2015; Srivastava et al., 2017b). PapA was identified at diverse places of the cell, such as at the inner or peripheral membrane and inthe cytoskeletal system, depending on the cell's state (normal or stress) and interaction with partner proteins. PspA is made up of three to four coiled-coil structures, the N-terminal of which contains 144 residues and interacts with PspF. Although the psp system has been widely explored in E. coli, it is also implicated in the cell defense systems of *Y. enterocolitica* and Salmonella. There are two unlinked loci in the Y. enterocolitica genome, pspF-pspABCDycjXF and pspG. The conserved pspFABC gene controls the cell's physiological activity (Flores-Kim & Darwin, 2016a). PspA deficiencyhas been linked to a variety of physiological defects. However, because of the cell's physiological state, it's unclear how PspA interacts with distinct partner proteins from various subcellular sites. Also relevant is the question of how the change in cell pH caused by proton leakage affects the structure of PspA.

Y. enterocolitica is nearly identical to that of E. coli (78 percent similarity), however, it lacks structural information. Leu, Ala (hydrophobic), charged residues (excluding Asp), and Gln (polar) all have high helix-propensity. pspA is a suitable system for fluorescence-based characterization since there is no Tyr and just a Trp (Hennig et al., 2017). Because PspA has a lot of titratable amino acid side chains, it's important to understand how pH variations alter the protein's structure.

We think that the knowledge gained from our research into the transduction of pH-dependent signals will throw light on PspA's unique role. This paper includes a description of the differential structural

properties of the end-states, including particular electrostatic strength, derived from pH-dependent molecular dynamic simulation.

4.2 Results

4.2.1. Physicochemical analysis and structure prediction of PspA and PspF

The 3D structure and interaction codes are linked in the protein sequence. To further comprehend the differences between the sequences of PspA and PspF, we evaluated their physicochemical characteristics. It has a substantially lower average hydrophobicity and a much greater average hydrophilicity than the latter (Fig. 4.2A, B). Similarly, to PspF (Fig. 4.2C), the 90-110 region of the sequence has a significant hydrophobic character, despite the fact that hydrophobic residues (HB) are slightly hampered due to the prevalence of charged residues (CR). A significant area (100 residues) in the center of the PspF sequence, on the other hand, displays substantial hydrophobic characteristics. Because of its C-terminal end, PspA's sequence includes more (18.6 percent) strong disorder-creating residues (S, E, R, P) than PspF's (13 percent). The secondary structure predictions of these two proteins provide quite different findings. While helix makes up the majority of PspA's secondary structure (95 percent), PspF's secondary structure is much more constant (Fig. 4.2D).

We used the ab-initio approach to identify the structure of these two proteins because there is none in the database for *Y.enterocolitica*. The structures of PspA and PspF were derived from their complete sequences. PspA's structure is predominantly made up of helices, as evidenced by secondarystructure prediction, which shows that the N-terminal helices (from residue 24 to residue 144) generate two distinct coiled-coil domain structures. It's worth mentioning that in the case of E. coli,a similar structure may be seen. The sequence from residue 144 to the C-terminal region contains pieces of small secondary structures, however, this area is quite disorganized (Fig. 4.3). PspF, on theother hand, is an enhancer-binding protein (EBP) that has at least two domains: a C-terminal domainwith helix and turns, and an N-terminal domain that includes the central domain with helix and beta structure (Fig 4.3). The PspA and PspF of *Y.enterocolitica* contain highly conserved loops of the N-

terminal domain, loop-1 and loop-2, that interact with RNAP-σ54. The impact of pH on the structure and structural components is shown below. Taken together, we can say that the typical traits observed in other species are well displayed in our two model structures.

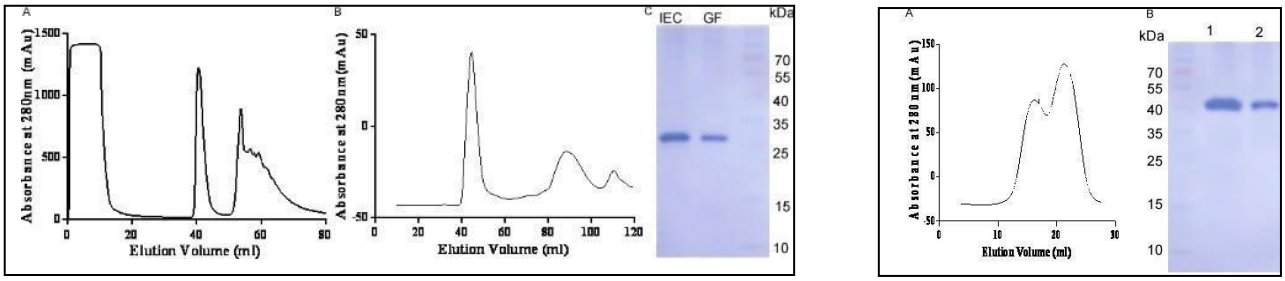


Fig 4.1. The oligomeric and monomeric states of PspA (left). A. Ion-exchange chromatography of PspA shows, that maximum elution of PspA came out in flow-through (arrow mark). B. Size exclusion chromatography separates oligomeric (peak 1) and monomeric (peak 2) states of PspA. C. Purity checked by SDS-PAGE.

The oligomeric and monomeric states of PspF (right). A. Size exclusion chromatography separates oligomeric (peak 1) and monomeric (peak 2) states of PspF. B. Purity checked by SDS-PAGE.

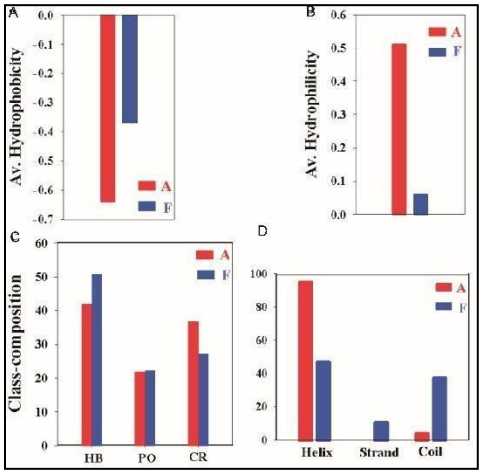


Fig 4.2 Physicochemical properties and model structure of PspA and PspF. Comparison of PspA and PspF's average hydrophobicity (A), average hydrophilicity (B), and amino acid class composition (C). Comparison between different types of the secondary structure of PspA and PspF (D).

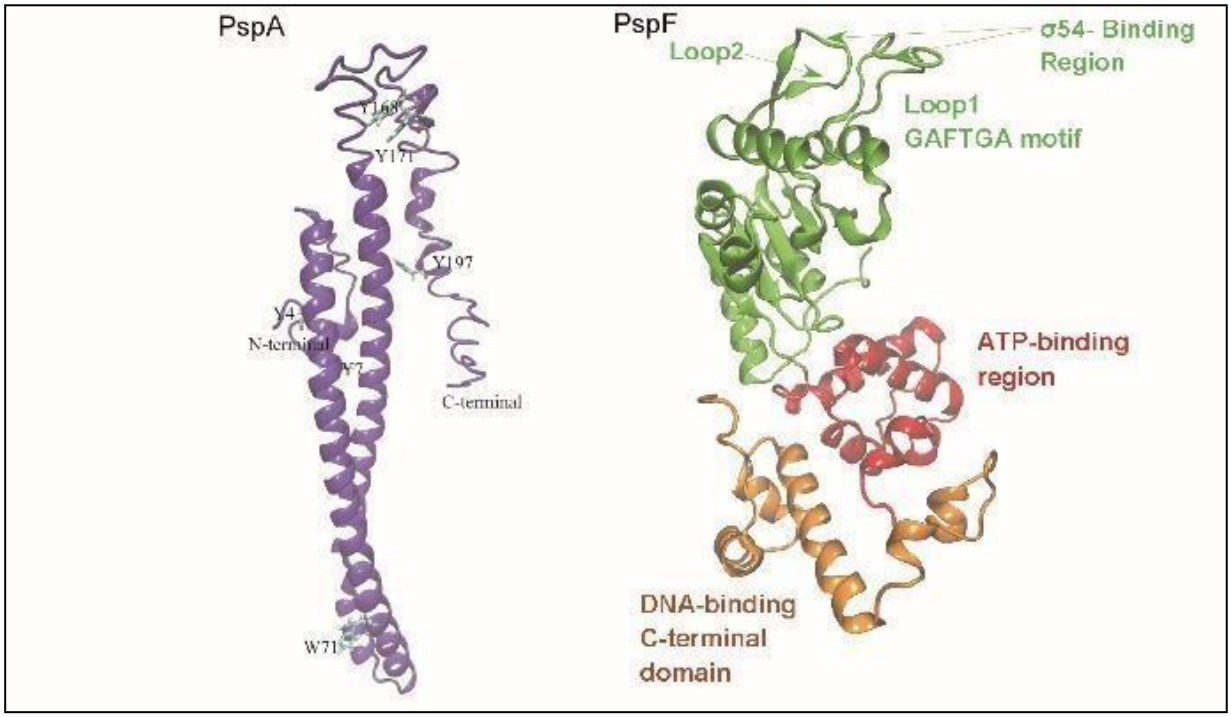


Fig 4.3. Predicted modelled structure of PspA and PspF.

4.2.2 The conformation of PspA secondary structure depends on the environmental pH

The secondary structure of a protein may be thought of as the determinant of its topology and conformation, and it can be precisely explored using circular dichroism and FTIR spectroscopy. We wonder how the shift in pH impacts secondary structures because it is followed by a significant change in PspA's conformation. The ellipticity ratio (CD222:CD208) graphic clearly shows the difference between this pH form and other pHs, where, θ_{222}: θ_{208} is one that indicates its 100% helix property (Banerjee et al., 2017). The helix-to-coil transition may be deduced from the fact that there are slightly negative minima at 210-220 nmat all pHs, which might imply a -like conformer. This type of conformation is most noticeablearound pH 7.3. At pH 6.3, a single minimum was recorded at 200 nm, indicating the disorderliness of the system.

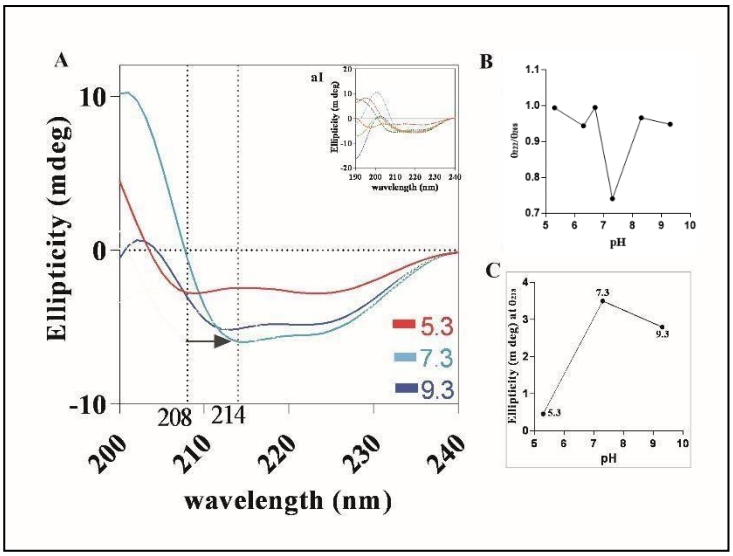


Fig 4.4 The secondary structure of PspA changes with pH. A. comparison of PspA far-ultraviolet CD spectra at pH 5.3 (red), 7.3 (cyan), and 9.3 (blue). B. The PspA 222 to 208 ratio varies with pH. C. Structure's relative value to pH (218 nm).

FTIR spectroscopy was used to investigate the secondary structure of PspA at various pH levels, as shown in Figure 4.5. The CD results are supported by the FTIR results. Figure 4 shows the FTIR spectra of PspA at various pH levels in the 1600-1700 cm^{-1} region, which probes the vibrational frequency of distinct amide bonds in the protein's main chain that is associated with the secondary structure. PspA's secondary structure is most likely a-helical at pH 5.3 and 9.3 (wavelengths of 1653 cm^{-1} and 1651 cm^{-1}, respectively), while intermediate helical conformation (1639 cm^{-1} - 1648 cm^{-1}) was predominant at pH 6.3, 6.7, and 8.3 and was also weakly detectable at pH 7.3. However, at pH 7.3, considerable alterations have been detected (Fig. 4.5).

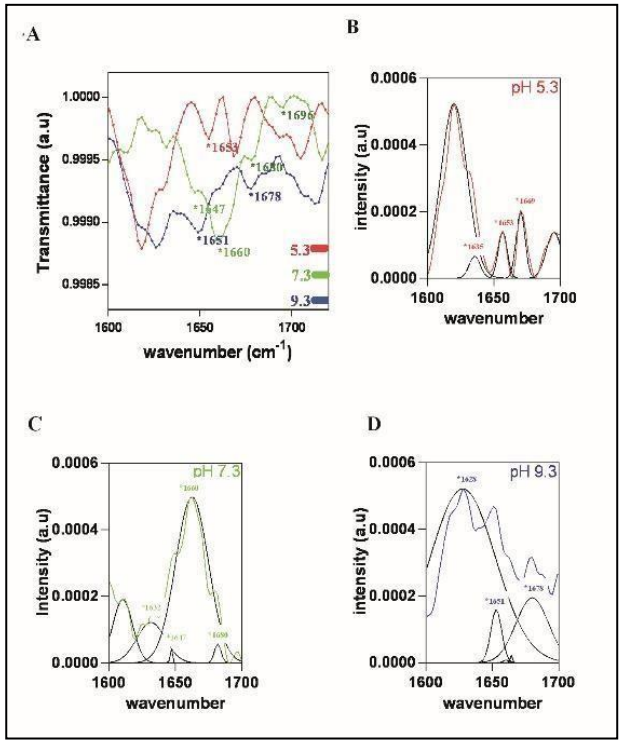


Fig. 4.5 Comparison of FTIR spectral characteristics of PspA at different pH. Spectra showing conformation of secondary structure where, α-helix, intermediate helix, β-turn, and β-sheet conformer denoted as an asterisk (*). Spectra of PspA at pH 7.3 (cyan) show β-turn and β-sheet-like conformation whereas at the pH of 5.3 (red) and 9.3 (blue) show α-helix and intermediate helix-like conformation.

Where the protein usually switches to a mix of extended ß-strand (1624 cm^{-1}, 1660 cm^{-1}, 1696 cm^{-1}) and ß-turn (1624 cm^{-1}, 1660 cm^{-1}, 1696 cm^{-1}) and ß-turn (1624 cm^{-1}, 1660 cm^{-1}, 1696cm^{-1}) (1680 cm^{-1}). We wanted to know how pH affects PspA's oligomeric form, which helpsit adapt to stress. In effect, we ran an AFM experiment that was pH-dependent.

4.2.3 Tendency of oligomerization of PspA at different pH conditions

The oligomeric state of PspA is determined by the location, contact, and function of the cell. In neutral pH, the lower oligo-form predominates over acidic and alkaline pHs. The findings of the pH-dependent AFM experiment have been used to demonstrate this. The shapes of acidic and basic pHs may also be compared to a bead in a chain when contrasted to neutral pH (Fig 4.6). Thatindicates PspA's neutral pH foam structure is most likely the best for interacting with PspF. PspA's pH-dependent secondary structure and oligomeric state, on the other hand, were detected.

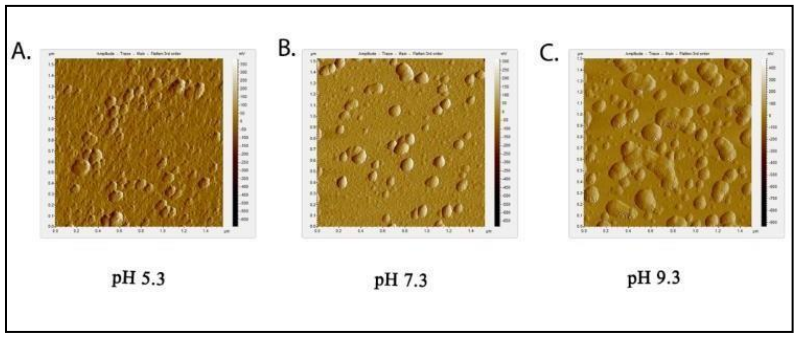


Fig. 4.6 Levels of oligomerization of PspA at different pHs. Compared to pH 7.3 (B), the level of oligomerization in pH 5.3 (A) and 9.3 (C) is visible. In pH 5.3 and 9.3 PspA looks like a higher-order chain like oligomers whereas less/scattered oligomeric nature is observed in pH 7.3.

4.2.4. Salt-bridge helps to maintain the overall tertiary structure of various structural conformers at different pH conditions

We used pH-dependent MDS to determine the influence of PspA confirmation on pH, and the results are shown in Fig. 4.7. There are a few aspects worth mentioning. The coiled-coil of the N-terminal interacts with PspF first. In pH 7.3, the bending of this coiled-coil structure of 24 to 144 residues is pH-dependent, extending from rough residue 58 to 110 (Fig 4.7AII). Despite the fact that this distinctive bending is present at all pH levels, the residue expansion of the bend zone and the bend curvature vary greatly. Based on these parameters, it can be concluded that the bending of pH 7.3 is more flexible than that of other pHs (Fig. 4.7AII). Although the helix structure is present in 61%, 65%, and 68% of the time in pH 5.3, pH 7.3, and pH 9.3, respectively. The Stride web server (Heinig & Frishman, 2004) was used for these investigations.Another feature that distinguishes Neutral pH from other pHs is the existence of an anti-parallel-like structure (Fig. 4.7BI).

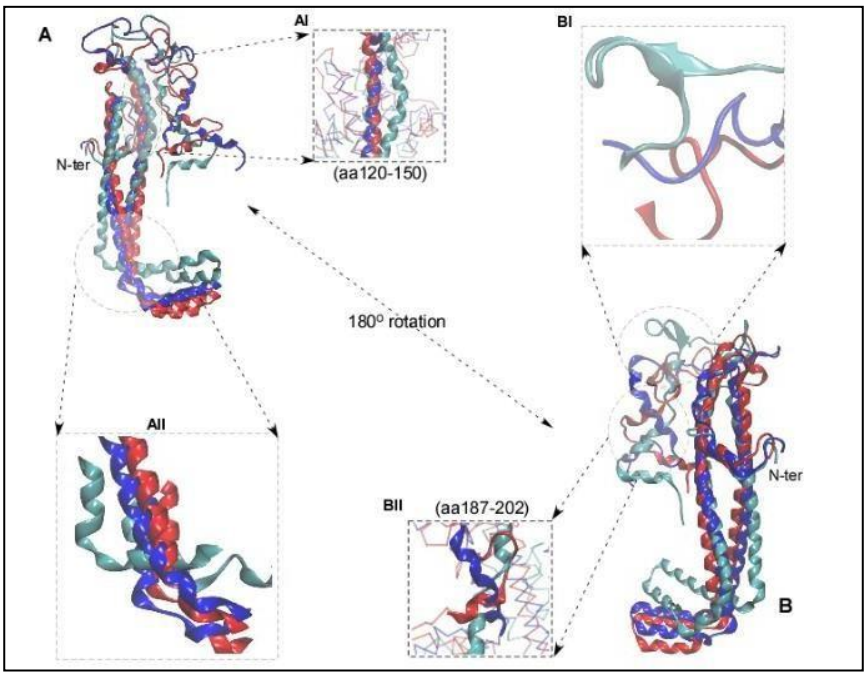


Fig. 4.7 Superimposition of different pH forms of PspA. A. Combined superimposition of pspA's pH 5.3 (red), pH 7.3 (cyan), and pH 9.3 (blue) models with an enlarged view of part of c-terminal (AI) and bend junction (AII) regions. B. Similar presentation of superposition of pH-models after 180° rotation with the highlight of specific C-terminal regions.

Second, we can see how the conformation changes as the pH changes. What effect does this conformational shift have on the salt bridge and its energetics? Salt bridges are divided into two categories: IP and NU. Table 4.1 demonstrates a large difference in the frequency of IP and NU salt-bridge types as pH changes. Because of the shift in the pH-dependent conformation, the specificity of the salt bridge is altered, and the overall frequency of the salt bridge in pHs 5.3, 7.3, and 9.3 is 19, 16, and 18, respectively. In other words, the variation of conformation with pH is supported by the change of salt bridge. The change in specificity is accompanied by a shift in the salt's position (core or surface).

Table 4.1 pH-dependent IP and NU salt-bridge frequency, overall and per-salt-bridge energetic (IP = isolated pair; NU = network unit).

protein	IP pair	NU pair	total pair	IP Energy	NU Energy	Total Energy	Energy/SB
pH5.3	6	13	19	-24.9	-57.91	-82.81	-4.35842
pH7.3	10	6	16	-53.79	-32.21	-86	-5.375
pH9.3	13	5	18	-68.42	-15.79	-84.21	-4.67833

Although the salt-bridge frequency at pH 7.3 is low, the per-salt-bridge stability point is the greatest (Table 4.1). The study of structures at pH 7.3 and pH 5.3 shows that the conformation of acidic and alkaline structures differs greatly from that of neutral pH. The challenge now is which of these structures binds PspF the most firmly. We've published the results of a binding experiment to help you understand (see below).

4.2.5 The binding stability of PspA with PspF is regulated by pH

The optical analytical approach of bio-layer interferometry was used to study the real-time interaction between PspA and PspF at varied pH settings (Fig. 4.8). PspA and PspF have an extremely high dissociation constant in pH 5.3, indicating quick dissociation. In this case, increasing the pH to neutral, 7.3 (also pH 6.7), results in the lowest average dissociation constant, implying the maximum contact affinity between PspA and PspF (1.55 M). Surprisingly, the interaction effects of these two proteins decrease at alkaline pH (pH=9.3). Taken together, the affinity of the interaction between PspA and PspF is strongest at near-neutral pH, whereas the largest and moderate deviations occur in acidic and alkaline circumstances, respectively. The best fit model of the interaction data between PspA and pspF indicates that their binding stoichiometry is 2: 1, meaning that one mole of PspF interacts with two moles of PspA.

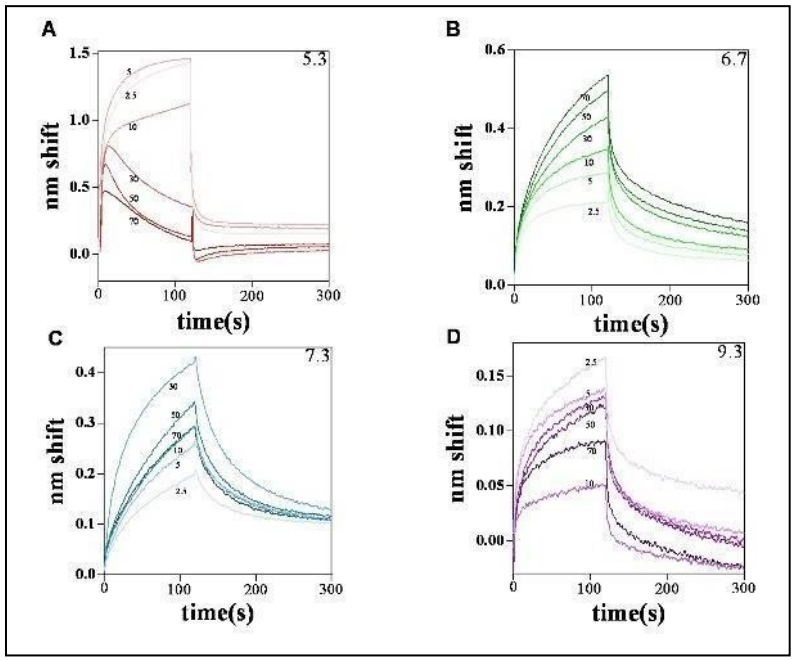


Fig. 4.8 An overview of the interaction between PspF and PspA at various pH levels. The binding affinity of PspF with PspA at different pH levels was determined using a binding kinetics assay using bio-layer interferometry. The interaction between PspA and PspF has the greatest affinity at pH 7.3 (cyan) and pH 6.7 (green), with the strongest and moderate deviation in acidic (pH 5.3) and alkaline (pH 9.3) conditions, respectively. The best fit model of pspA and pspF interaction data reveals that their binding stoichiometry is 2: 1, which means that one mole of pspF interacts with two moles of pspA.

4.3 Discussion

When a cell is exposed to stress, it may alter its normal metabolic state to defend itself. To transport protein and other biological components through the cell membrane in gram-negative bacteria, a multimeric secretin complex is required (Yamaguchi et al., 2010). Mis-localization of secretin occurs during extracellular stress or even under normal physiological conditions in the human pathogen *Y. enterocolitica*. *Y. enterocolitica* is a pathogen that causes many illnesses and deaths each year. Understanding the psp system appears to be relevant for both fundamental and applied considerations. For these reasons, we investigated the pH-dependent conformation change of PspA and the strength of its binding to PspF. Protrusion of cell protons occurs in response to extracellular stress, and this can be particularly sensitive to PaspA's structure and conformation.

We used the ab initio approach to create the structure since the structure codes are contained in the protein sequence and the structure database does not have the structure of the two proteins

(PspA and PspF). A coiled-coil motif appears to be beneficial in protein-protein interaction and oligomerization. Again, as with the C-terminal section of PspA, the disordered region can operate as a determinant of specificity in this interaction. pi-pi is abundant in the center of most proteins, but it lacks here demonstrate the relevance of the weak forces mentioned earlier. This is due to the fact that a disordered protein can interact with a partner protein to obtain thedesired shape and orientation (Widłak, 2013). Apart from specific-electrostatic (salt-bridge orion-pairs) interaction, determinants of specificity appear to be pi-cation, pi-anion, pi-alkyl, etc.,whose frequency is low or negligible. At pH 7.3, the bending and type of flexibility at the bend-junction of PspA's coiled-coil differ from other pHs. This type of bending is ideal for dockinga partner protein and indicates that this pH form's interaction with PspF is higher than other pHforms.

PspA and PspF interact to promote the lower oligomeric form of the former, while PspBC interacts to favor the higher form. This shift in conformation with pH, as seen in molecular dynamic data, circular dichroism, and FTIR spectroscopy, appears to be linked to the backbone's orientation and geometry. The specific explanation for this difference is unknown, although PspA's solution state suggests its interactivity has changed. The near-neutral pH has higher interaction than the others because of an anti-parallel ß-like structure in the C-terminal region, which is absent in the other pHs. The emergence of this structure at neutral pH appears to be the outcome of the coil to ß transition.

PspA is thought to be a cell stress-protective protein that may modify its shape in response to diverse forms of stress. If PspA attaches to PspF in the cytosol, the cell is in a normal metabolic condition; if it does not, it has a low oligomeric state, which changes when the cell membrane is subjected to extracellular stress. This can alter the conformation of RNAP-σ54 and act as a binary switch in terms of its transactivation. Acid and alkaline pHs may act as stress signals for PspA, which releases PspF from PspBC in the membrane during translocation. The titration

state of the side chain of these residues is affected by pH fluctuation because the prevalence of a charged residue is naturally greater in the sequence of PSpA. This difference affects salt bridges that are in stable conditions with a pH of near neutral (pH 7.3).

4.4 Conclusions

As a result, the overall stability of the protein fluctuates with changes in pH, potentially affecting the protein's global shape. We hypothesize that this PspA is a negative regulator of PspF at neutral pH and a helpful type of interaction with PspBC on the membrane at acidic or alkaline pH. Furthermore, it's possible that the pH shift in the in vitro condition can replicate the status of PspA in the cell's cytosol and membrane in the in vivo situation. We believe that by studying PspA's pH-dependent conformational shift, we will be able to learn more about how it transmits stress signals.

CHAPTER 5

5. Domain crystallization and diffraction of PspA, a phage shock protein A of

Yersinia enterocolitica

5.1 Introduction:

The phage shock protein (PSP) system is an adaptive response system in many bacteria to subvert various cellular stresses which may affect their cellular integrity. Phage shock protein A (PspA) is a critical player in the PSP system and was first reported in the case of filamentous phage-infected *Escherichia coli*. PspA belongs to the conserved PspA/IM30 protein family that includes proteins across bacteria (Jordan et al., 2006), archaea, and plant chloroplasts. PspA interacts with two sensor membrane proteins namely PspB and PspC, as well as with PspF (Dworkin et al., 2000), a transcriptional regulator. Similar to the pIV protein, which induces the *E. coli pspA* operon, activation of the PSP system in *Y. enterocolitica* results from the overproduction of YscC (a secretin family protein). Conclusively, the PSP response is an important envelope stress response (ESR) that has been associated with various significant phenotypes, including virulence in many pathogenic bacteria.

Although *Y. enterocolitica* is ~77 percent identical to E. coli, structural information is lacking. Helix-propensity is high in Leu, Ala (hydrophobic), charged residues (except Asp), and Gln (polar). As, PspA was discovered to connect with two tiny membrane proteins and putative sensors, PspB and PspC, as well as with PspF (Dworkin et al., 2000), the transcriptional activator of the system, resulting in a negative feedback loop as part of a stress-inducible system. So, crystallization and structure determination of PspA is very crucial for a better understanding of the mechanism of PSP system. Though the C-terminal part of PspA is very much flexible and disordered (through ProDOS server analysis, Fig 5.1C), so stable domain of PspA is also important for understanding the structural point of view. Coiled-coil domain is important for self assembles, which may help to support the membrane during stress conditions of the bacterial cell (mainly leakage of membrane).

One of the most common structural motifs in proteins is coiled coils, which are supercoiled bundles of a-helices (Hernandez Alvarez et al., 2008). According to conservative estimations, a-helical coiled coils include around 5% of all residues in proteins (Lupas et al., 1991). Protein symmetry, particularly homo-oligomerization, appears to facilitate protein crystallization. Crystallization trials have even utilized artificial symmetrization, which forces proteins into a dimeric state via disulfide bonds (Banatao et al., 2006). Though ΔPspA has no cysteine in protein sequence, however, in gel filtration data and from the final refinement model, we may suggest that ΔPspA exists in solution as a dimeric form.

5.2 Results:

5.2.1 In solution, ΔPspA exists as a dimer

The molecular weight of this protein is close to 18 kD, according to the final purification step by the gel filtration process (Fig. 5.1C & inset). Running an SDS-PAGE in the absence and presence of glutaraldehyde, a cross-linker reveals that the former belongs to the monomeric form of the protein and the latter relates to the dimeric form (Fig. 5.1C, D). The gel filtration andcross-linked gel outcomes are identical in this case.

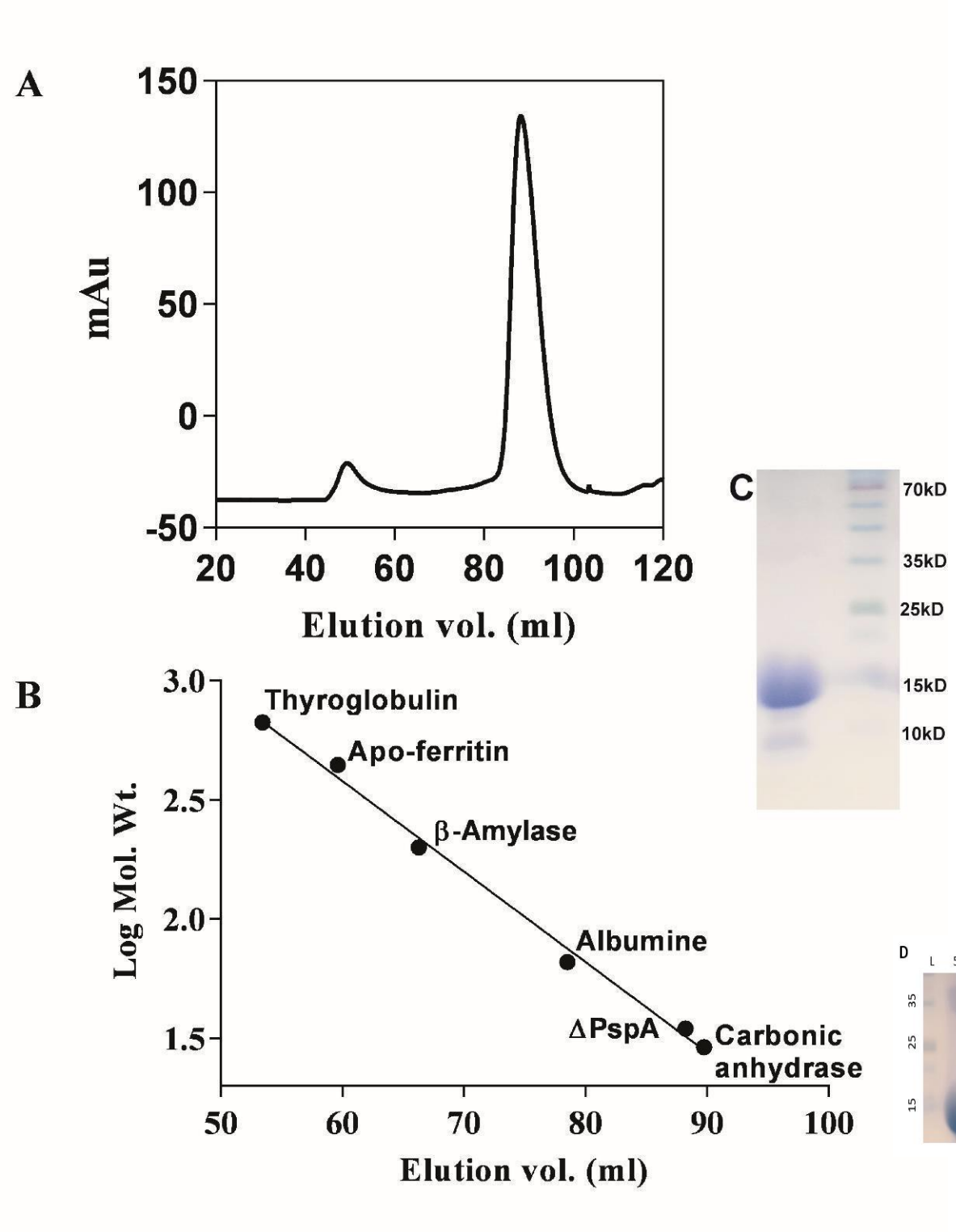


Fig 5.1 The SEC and standard plot. SEC indicates that ΔPspA exists in solution as a dimer. The chromatogram of ΔPspA gel-filtration (A) and the log molecular weight vs. elution volume plot for standard proteins are shown (B). purity checked by SDS-PAGE (C). Glutaraldehyde cross-linked of ΔPspA (D).

5.2 Crystal morphology and diffraction of ΔPspA crystal

Protein crystallization is a major bottleneck in the entire process of protein crystallography, and obtaining diffraction-quality crystals can be unpredictable and challenging at times, necessitating multiple rounds of high-throughput screening (M. J. Y. Lee et al., 2014). Researchers have been exploring the relationship between crystal structure and morphology for over a century. According to geometrical considerations and growth rates, the larger the interplanar distance dhkl, the lower the face's growth rate (hkl) (Plomp et al., 2002).

We arere aiming to make some diffract quality crystals here. A large number of crystals appeared on the crystal screen under various conditions. However, not all beautiful crystals (fig. 5.2) are diffracted; some are (fig. 5.3A, B, and C). Diffracted crystal has a rod-shaped form. In the sitting drop vapor diffusion method, the crystal emerged after 4 days. Cryo-solutions containing 10% glycerol were employed to increase diffraction quality (fig. 5.3D, E, F). However, diffraction quality did not improve.

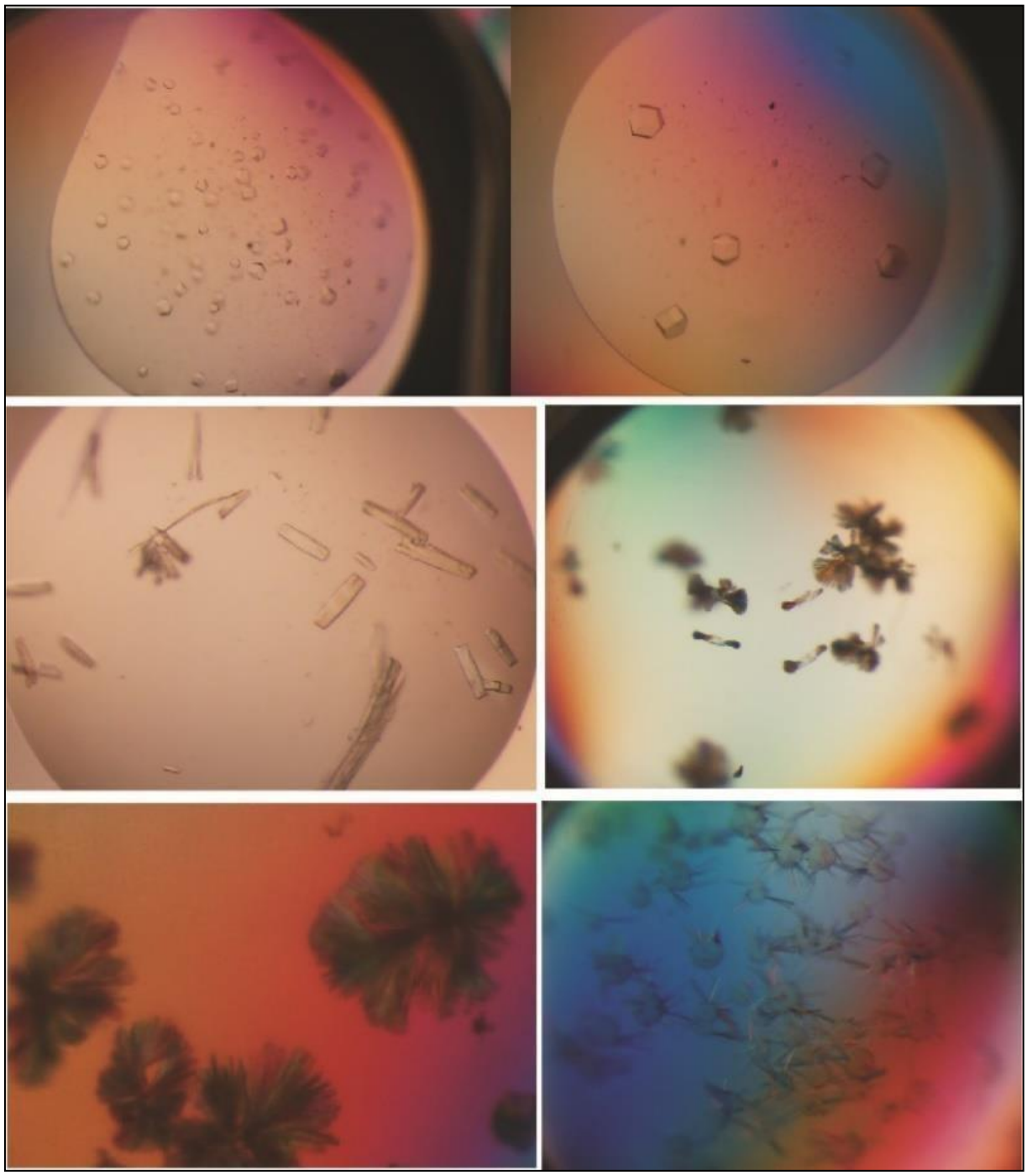


Fig 5.2 Protein crystals shape of ΔPspA. Crystals of various shapes (rod, hexagonal, spick, and brush-shaped) emerged during sitting drop vapor diffusion but were not diffracted. Conditions are described in the methodology section (Table 2.2).

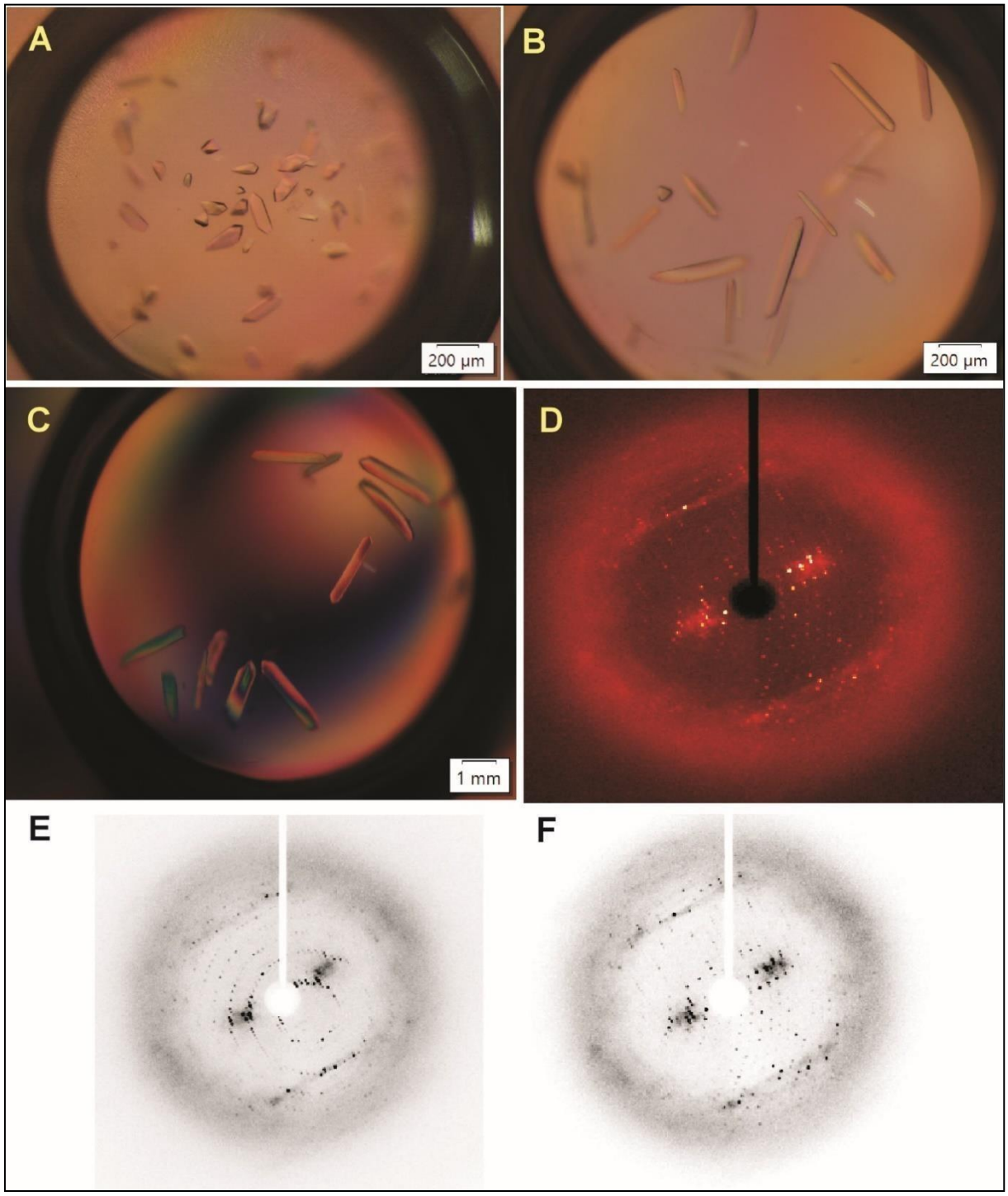


Fig 5.3 Crystals and diffractions pattern of ΔPspA. rod shape crystal appeared to appropriate buffer condition (PEG400, HEPES, and CaCl2, Table 2.2). D, E, F, showing the diffraction pattern of crystal.

5.3 Structure determination from diffraction data

Rosetta model building coupled molecular replacement (MR-Rosetta) (Terwilliger et al., 2012b) provides a partial solution of ΔPspA. After manual model creation in Coot (Emsley

& Cowtan, 2004) and alternate macro-cycles of iterative refinement using PHENIX.refine (McCoyet al., 2007b), the entire model was achieved. The structure was primarily solved at 4.38Å resolution in the P3221 space group (Table 5.1). MolProbity and other validation tools were used to perform stereochemical and statistical validation of the model (fig 5.4). PyMOL was used to construct structural figures.

Table 5.1: Data and refinement statistic from preliminary full wwPDB X-ray structure validation report

Space Group	**P 32 2 1**
Unit Cell	97.22 97.22 128.32 90.00^0 90.00^0 120.00^0
Resolution Range (Å)	23.47 – 4.38
Completeness (%)	97.5
Mean I/sigma(I)	10.47
Wilson B-factor ($Å^2$)	91.2
R-work	0.395
R-free	0.399
Macromolecule atoms	1949
Protein residues	243
Ramachandran outlier	0
Sidechain outliers	1.8%
Classcore	11
RSRZ outliers	0

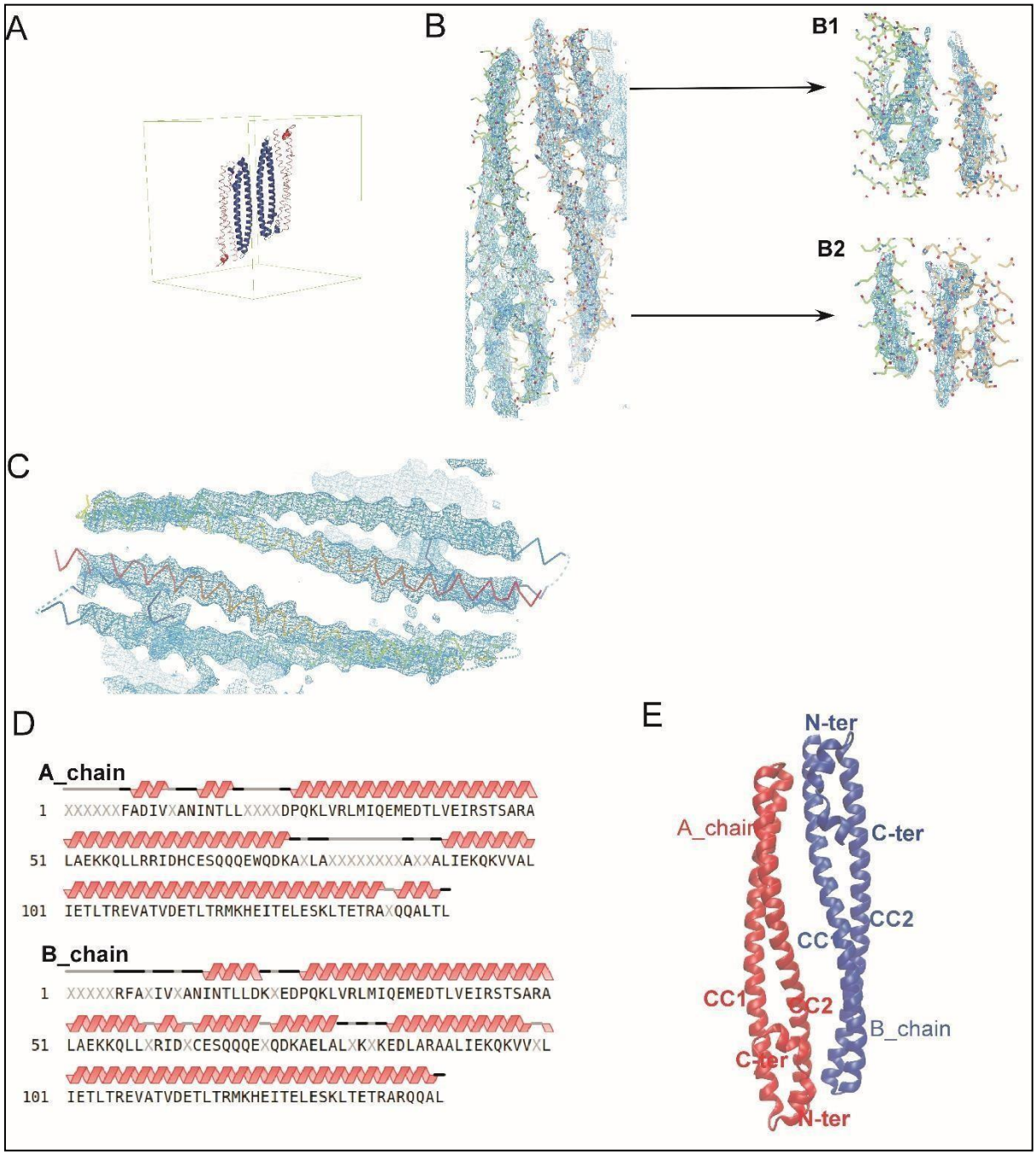


Fig 5.4 Structure of ΔPspA. There is 4 ASU content in the unit cell **(A)** 2F0-FC electron density map contoured at 2σ (B). the region displays probable interacting residues (B1, B2). The backbone is properly fitted in the electron density map (C). Loop and highly flexible portions are deleted (D). model of ΔPspA (C). the complete model was built by coot and pymol software (E).

Table 5.2 Interaction of inter-chain residues:

5.2.1 Hydrophobic Interactions within 5 Angstroms

Position	Residue	Chain	Position	Residue	Chain
79	**LEU**	**A**	**134**	**ALA**	**B**
134	**ALA**	**A**	**79**	**LEU**	**B**

5.2.2 Side Chain-Side Chain Hydrogen Bonds

DONOR				ACCEPTOR								
POS	**CHAIN**	**RES**	**ATOM**	**POS**	**CHAIN**	**RES**	**ATOM**	**MO**	**Dd-a**	**Dh-a**	**A(d-H-N)**	**A(a-O=C)**
16	A	ASN	ND2	84	B	GLU	OE2	2	3.23	3.26	79.43	999.99
72	A	GLN	OE1	131	B	GLU	OE2	2	3.46	4.24	37.00	999.99
84	A	GLU	OE2	16	B	ASN	OD1	2	3.24	3.57	63.43	999.99
94	A	LYS	NZ	131	B	GLU	OE2	-	3.09	9.99	999.99	999.99
130	A	THR	OG1	76	B	GLU	OE2	-	2.31	9.99	999.99	999.99
131	A	GLU	OE2	72	B	GLN	NE2	2	3.15	2.11	166.23	999.99
135	B	ARG	NH2	84	A	GLU	OE2	2	2.96	2.60	100.07	999.99
142	B	ARG	NH2	82	A	ASP	OD1	2	3.49	3.01	109.49	999.99

5.2.3 Ionic Interactions within 4 Angstroms

Position	Residue	Chain	Position	Residue	Chain
82	ASP	A	142	ARG	B
84	GLU	A	135	ARG	B

94	LYS	A	131	GLU	B
135	ARG	A	84	GLU	B

5.4 Discussions and conclusions:

ΔPspA crystallizes as a dimer, with each asymmetric unit cell containing four ΔPspA monomers (144aa in length). A loopy area exists in monomeric ΔPspA, which contains only the two predicted coiled-coil domains, CC1 and CC2, as well as the native N-terminal region. The backbone of the protein is properly fitted in the 2F0-FC electron density map, but the N-terminal and turn-region (75–85) are not. Therefore, with the help of Coot and Pymol, a complete model can be constructed. It has a pointed gourd-like form and is dimeric.

Two monomers are linked in opposite directions in Figure 5.4E. The hydrogen bond, salt bridge, and hydrophobic interaction all contribute to the structure's stability (Table 5.2). However, we may deduce that the benefit of such a coiled-coil structure is that it can be organized into a shape that supports the bacterial inner membrane during stressful conditions (e.g., mislocalization of secretin protein in *Y enterocolitica*).

CHAPTER 6

6. Proteins' salt bridge energy study among three domains of life

6.1 Introduction:

The majority of charged residues of the protein are present on the protein's surface in the tertiary structure, which helps to smooth out excess protein hydration, making it less hydrophobic and more flexible. As a reason, it promotes non-specific electrostatic interaction with salts in solution, which helps to offset the negative effects of salt (Tadeo et al., 2009). Asa result, the ion-pair or salt bridge is a substantial contributor to protein stability (Dill, 1990). When proteins adapt to harsh environments, such as high salt or high temperature, it becomes much more obvious. Other labs' experiments have shown that adding a single salt-bridge into a single mutation in the protein surface boosts protein stability in the tertiary structure (Hendsch& Tidor, 1994). As a result, a designed salt bridge on the protein's surface provides additional stability. The energy of a salt bridge may be divided into three main parts: columbic attraction of opposing charges, desolvation of opposite charges, and background interactions. Columbic attraction is always present in the tertiary structure of proteins, and the other two words may or may not play a role. The unfavorable desolvation of charges inside a salt bridge opposes the favorable charge-charge attraction, which is further controlled by charge-dipole interaction and the ionization behaviour of related charge groups. As a result, net ion-pair energy distributions may be stabilising (Horovitz & Fersht, 1992), destabilising (Dao-pin et al., 1991), or inconsequential. Using a dataset of 24 high resolution (1.5A) crystal structures from three domains of life, we examine the distribution of salt-bridge, its energy term, and their contribution to overall stability in the natural state of proteins (eukaryote, bacteria, and archaea). The frequency of five acidic and basic residues (such as Asp, Glu, Arg, His, Lys) creating salt bridges is estimated. We also use APBS techniques to determine the net contribution of the salt bridge, the tendency of arginine, histidine, and lysine to interact with acidic residues in proteins, and the average accessibility of the salt-bridge energy term.

6.2 Results:

6.2.1 Salt-bridge frequency versus intervening residues and potential pair

We constructed 10 intermediate residue group frequency distribution plots to investigate the position distribution of salt bridging partners in the protein sequence throughout the life of three domains. We've presented 8 servings of protein from each domain in this map, with the first 100 amino acids interval residue representing the Salt Bridge distribution. The frequencyof salt bridges reduces as the number of intervening residues increases, as demonstrated above.Despite the fact that the frequency of salt bridges has decreased as the intermediate residue hasincreased, archaea proteins have a tendency to build salt bridges in distant (40-50 & 50-60) intervening residues (C. Roy et al., 2020).

Figure 1 depicts the salts-bridge of all three domains. Arg-Glu, Lys-Glu, Arg-Asp, His-Glu, Lys-Asp, and His-Asp are the six salt-bridge potential pairs. In archaea, Arg-Glu and Lys-Glu have a greater potential to build salt bridges (fig 6.1). Arg-Asp and Lys-Asp generate greater frequency salt bridges in eubacterial proteins (fig 6.1), but His-Asp frequency contributes significantly to the eukaryote domain.

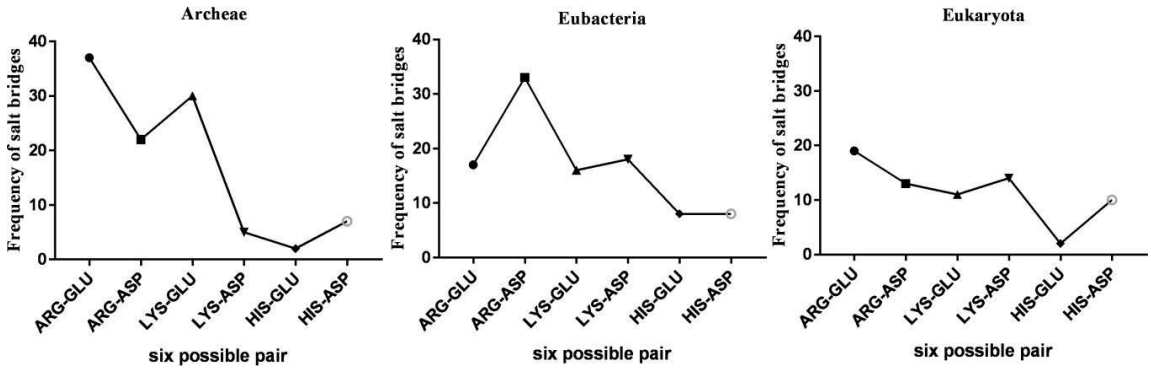


Fig 6.1: Across three domains of life, the frequency of salt bridges formed by each of six possible pairs

6.2.2 The average energy term in three domains of life

Salt-bridge interactions, like other non-covalent interactions, play an important role in protein stability. In this paper, we utilized the ADSBET2 (in section 2.8.9) to calculate the net salt bridge energy (ΔΔGnet/tot: sum of the above three-component terms) using all three related energy terms (ΔΔGdslv: desolvation; ΔΔGprt: backdrop; and ΔΔGbrd: bridge). ΔΔGdslv and ΔΔGprt are both indirect energy terms, with ΔΔGdslv being unfavorable due to charge desolvation during folding and Gprt being favourable or unfavourable. Because of the interaction of charges in the folded state of the protein, Gbrd is always advantageous. In three domains, we chose a homogenous data set. Archaea, as opposed to eubacteria and eukaryotes, appear to use more specialized electrostatic interactions. Overall total (ΔΔGnet) average network energy is higher than isolated in three domains, as shown in fig 3, while overall total (ΔΔGnet) salt-bridge energy is higher in archaea and eubacteria than in eukaryota core (buried). According to Roy et al., the overall stability of the salt-bridge in archaea (ΔΔGnet= -5.063.8) is substantially greater than in eubacteria (ΔΔGnet= -3.72.9) and eukaryotes (ΔΔGnet= -3.543.1). We have demonstrated that the ΔΔGtot term has an approximately symmetric distribution, encompassing both stable and unstable salt bridges, with the bulk of them lying in the stabilizing zone (ΔΔGtot). Quantitative analysis reveals that over 90% of salt bridges provide stability in archaea, compared to 78 percent in eubacteria and 71 percent in eukaryotes (Table 6.1). The distribution of bridging energy term and desolvation energy term occupies stabilizing and destabilizing zones, respectively, whereas background term has both stabilizing and destabilizing salt-bridge population.

Table 6.1: Calculation of percentiles for various salt bridges over three domains

Salt bridge type	Archaea	Eubacteria	Eukaryota
Stable	89 (88%)	82 (78%)	50 (67%)
Unstable	12 (12%)	23 (22%)	24 (33%)
Network	57 (56%)	35(33%)	31 (41%)
Isolated	44 (44%)	70 (67%)	43 (59%)
Buried/Core	42 (42%)	31 (30%)	34 (46%)
Exposed	59 (58%)	74 (70%)	40 (54%)

6.3 Discussions and Conclusions

Salt-bridge interactions, like other non-covalent interactions, play an important function in the stability of folded proteins. Because secondary structure, which controls protein topology, has remained more conserved throughout evolution, our findings in this context suggest that salt-bridging acidic and basic residues that are closer in sequence tend to generate more salt-bridges. However, a distant intervening remnant salt bridge may be seen in Archaea. Now the question is whether it is just preserved in archaea. It is warranted due to its protein adaptability in harsh environments.

We can show that 88 percent of the salt bridges extract from three domains are extremely stable in the archaea Salt-Bridge. Where the salt bridge in eubacteria is 78 percent and the salt bridge from eubacteria is 67 percent. The rest of the salt bridge is of the destabilizing kind. Because of the lower contribution of background and bridge energy, this salt-bridge population is becoming unstable. The bridge energy is nearly the same among salt-bridge categories, while the background energy varies substantially. Among all component terms, it (background energy) is the most sensitive in the protein microenvironment. Excessive charges and dipoles create a local environment surrounding each salt-bridge partner, determining the amount of their contribution to net stability. However, not all buried salt bridges are stable, according to our observations. In comparison to the other three domains, archaea have a 90 percent buried salt bridge that is stable, but eubacteria and eukaryotes have 75 percent and 59 percent burial stability, respectively. The difficulty now is how to overcome the desolvation penalty and

achieve net stability with a buried salt bridge. The buried salt bridge suffers greatly as a result of the significant desolvation penalty, whereas the entropic cost is minimal in the local salt-bridge interaction, resulting in an enthalpically favourable situation. Out of 42 buried salt bridges in our data set, 28 have developed a network. In eubacteria and eukaryotes, respectively, 18 out of 29 and 17 out of 34 buried salt-bridges form a network. In archaea, rather than eubacteria and eukaryotes, excess network salt-bridge is produced in protein to overcome desolvation penalty and acquire average net stability with hidden salt-bridge. As a result, the excess network salt-bridge performs an expected function in maintaining protein structural stability in a hostile environment, such as that seen in archaea.

CHAPTER 7

7.1 Summary

Nowadays, it is critical to comprehend the pathogenic bacteria-host connection (Casadevall & Pirofski, 2000). In addition to E. coli, Pseudomonas aeruginosa and Yersinia enterocolitica are effective models for researching host-pathogen interactions (Luo et al., 2017). As a result, medicines are used to stop bacteria from dividing. Bacteria, in essence, take advantage of a person's weaker immune system to cause an illness, making them an opportunist. Pseudomonas and Yersinia use the type three secretion system (T3SS) to inject anti-host effectors into the target cell (Halder et al., 2019). Both bacteria have a Type Three Secretion System (TTSS), which allows them to transfer disease to the host membrane; this is a unique pathogenicity spreading mechanism seen only in gram-negative bacteria.

PcrG (in Pseudomonas) and LcrG (in Yersinia) are needle tip protein chaperones that control the secretion of effectors from the bacterial cytoplasm into the host cytoplasm as well as the secretion activity of the apparatus. Export of effectors via the T3SS is energized primarily by the pmf. As, *Y. enterocolitica* also possesses T3SS (Ysc-Yop) which is known to spread pathogenicity into the host cell. Ysc-Yop T3SS production induces the Psp (pspA-F & pspG) system, which then mitigates T3SS-induced envelope stress. The PspA (PA3731, in *Pseudomonas*) protein assists in the maintenance of the membrane when a secretin component (YscC) is mislocalized. However, it is becoming increasingly obvious that they are intertwined with the function of systems used during host infection in particular, as well as bacterial survival in this setting. So, understanding such proteins' structure, and how the salt bridge is important for maintaining their stability during pathogenesis or stress conditions, is crucial for better evaluation of bacterial behavior.

Reference

Aizawa, S.-I. (2001). Bacterial flagella and type III secretion systems. *FEMS Microbiology Letters*, *202*(2), 157–164. https://doi.org/10.1111/J.1574-6968.2001.TB10797.X

AK, B., RN, U. I., & N, H. (2020). Salt-bridges in the microenvironment of stable protein structures. *Bioinformation*, *16*(11), 900–909. https://doi.org/10.6026/97320630016900

Akhilesh, Baidya, A. T. K., Uniyal, A., Das, B., Kumar, R., & Tiwari, V. (2021). Structure-based virtual screening and molecular dynamics simulation for the identification of sphingosine kinase-2 inhibitors as potential analgesics. https://doi.org/10.1080/07391102.2021.1971559

Allmond, L. R., Karaca, T. J., Nguyen, V. N., Nguyen, T., Wiener-Kronish, J. P., & Sawa, T. (2003). Protein binding between PcrG-PcrV and PcrH-PopB/PopD encoded by the pcrGVH-popBD operon of the Pseudomonas aeruginosa type III secretion system. *Infection and Immunity*, *71*(4), 2230–2233. https://doi.org/10.1128/IAI.71.4.2230-2233.2003

Banatao, D. R., Cascio, D., Crowley, C. S., Fleissner, M. R., Tienson, H. L., & Yeates, T. O. (2006). An approach to crystallizing proteins by synthetic symmetrization. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(44), 16230–16235. https://doi.org/10.1073/PNAS.0607674103

Banerjee, A., Dey, S., Chakraborty, A., Datta, A., Basu, A., Chakrabarti, S., & Datta, S. (2014). Binding mode analysis of a major T3SS translocator protein PopB with its chaperone PcrH from Pseudomonas aeruginosa. *Proteins*, *82*(12), 3273–3285. doi.org/10.1002/PROT.24666

Banerjee R, Sheet T. Ratio of ellipticities between 192 and 208 nm (R1): An effective electronic circular dichroism parameter for characterization of the helical components of proteins and peptides. Proteins. 2017 Nov;85(11):1975-1982. doi: 10.1002/prot.25351. Epub 2017 Aug 7. PMID: 28707342.

Banerjee, S., Gupta, P. S. sen, Islam, R. N. U., & Bandyopadhyay, A. K. (2021). Intrinsic basis of thermostability of prolyl oligopeptidase from Pyrococcus furiosus. *Scientific Reports*, *11*(1). https://doi.org/10.1038/S41598-021-90723-4

Bassetti, M., Vena, A., Croxatto, A., Righi, E., & Guery, B. (2018). How to manage Pseudomonas aeruginosa infections. *Drugs in Context*, *7*. https://doi.org/10.7573/DIC.212527

Basu, A., Das, U., Dey, S., & Datta, S. (2014). PcrG protects the two long helical oligomerization domains of PcrV, by an interaction mediated by the intramolecular coiled-coil region of PcrG. *BMC Structural Biology*, *14*(1). https://doi.org/10.1186/1472-6807-14-5

Beloin, C., Valle, J., Latour-Lambert, P., Faure, P., Kzreminski, M., Balestrino, D., Haagensen, J. A. J., Molin, S., Prensier, G., Arbeille, B., & Ghigo, J. M. (2004). Global impact of mature biofilm lifestyle on Escherichia coli K-12 gene expression. *Molecular Microbiology*, *51*(3), 659–674. https://doi.org/10.1046/J.1365-2958.2003.03865.X

Bergeron, J. R. C., Fernández, L., Wasney, G. A., Vuckovic, M., Reffuveille, F., Hancock, R. E. W., & Strynadka, N. C. J. (2016). The Structure of a Type 3 Secretion System (T3SS) Ruler Protein Suggests a Molecular Mechanism for Needle Length Sensing. *The Journal of Biological Chemistry*, *291*(4), 1676–1691.

Birtalan, S., & Ghosh, P. (2001). Structure of the Yersinia type III secretory system chaperone SycE. *Nature Structural Biology*, *8*(11), 974–978. https://doi.org/10.1038/NSB1101-974

Blasie, C. A., & Berg, J. M. (2004). Entropy-enthalpy compensation in ionic interactions probed in a zinc finger peptide. *Biochemistry*, *43*(32), 10600–10604. https://doi.org/10.1021/BI0363230

Bruce, N., Ganotra, G., & Wade, R. (n.d.). *First Part of the Practical on Brownian Dynamics Simulation Protein Electrostatic Potentials: Calculation with APBS and Analysis*. Retrieved March 6, 2022, from http://mcm.h-

Casadevall, A., & Pirofski, L. A. (2000). Host-pathogen interactions: basic concepts of microbial commensalism, colonization, infection, and disease. *Infection and Immunity*, *68*(12), 6511–6518. https://doi.org/10.1128/IAI.68.12.6511-6518.2000

Chaudhury, S., Nordhues, B. A., Kaur, K., Zhang, N., & de Guzman, R. N. (2015). Nuclear Magnetic Resonance Characterization of the Type III Secretion System Tip Chaperone Protein PcrG of Pseudomonas aeruginosa. *Biochemistry*, *54*(43), 6576–6585. https://doi.org/10.1021/ACS.BIOCHEM.5B00664

Cornelis, G. R. (2006). The type III secretion injectisome. *Nature Reviews Microbiology 2006 4:11*, *4*(11), 811–825. https://doi.org/10.1038/nrmicro1526

Dao-pin, S., Anderson, D. E., Baase, W. A., Dahlquist, F. W., & Matthews, B. W. (1991). Structural and thermodynamic consequences of burying a charged residue within the hydrophobic core of T4 lysozyme. *Biochemistry*, *30*(49), 11521–11529. https://doi.org/10.1021/BI00113A006

Darwin, A. J. (2007). Regulation of the phage-shock-protein stress response in Yersinia enterocolitica. *Advances in Experimental Medicine and Biology*, *603*, 167–177. https://doi.org/10.1007/978-0-387-72124-8_14

Darwin, A. J., & Miller, V. L. (2001). The psp locus of Yersinia enterocolitica is required for virulence and for growth in vitro when the Ysc type III secretion system is produced. *Molecular Microbiology*, *39*(2), 429–445. https://doi.org/10.1046/J.1365-2958.2001.02235.X

Diaz, M. R., King, J. M., & Yahr, T. L. (2011). Intrinsic and Extrinsic Regulation of Type III Secretion Gene Expression in Pseudomonas Aeruginosa. *Frontiers in Microbiology*, *2*(APR). https://doi.org/10.3389/FMICB.2011.00089

Dill, K. A. (1990). Dominant Forces in Protein Folding. *Biochemistry*, *29*(31), 7133–7155. https://doi.org/10.1021/BI00483A001/ASSET/BI00483A001.FP.PNG_V03

Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., & Baker, N. A. (2007). PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Research*, *35*(suppl_2), W522– 525.

Dworkin, J., Jovanovic, G., & Model, P. (2000). The PspA protein of Escherichia coli is a negative regulator of σ54- dependent transcription. *Journal of Bacteriology*, *182*(2), 311–319. https://doi.org/10.1128/JB.182.2.311-319.2000/ASSET/928B1ABD-0676-49FC-B1E3-1C926C067C75/ASSETS/GRAPHIC/JB0200737006.JPEG

Emsley, P., & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallographica. Section D, Biological Crystallography*, *60*(Pt 12 Pt 1), 2126–2132. https://doi.org/10.1107/S0907444904019158

Eriksson, S., Lucchini, S., Thompson, A., Rhen, M., & Hinton, J. C. D. (2003). Unravelling the biology of macrophage infection by gene expression profiling of intracellular Salmonella enterica. *Molecular Microbiology*, *47*(1), 103–118. https://doi.org/10.1046/J.1365-2958.2003.03313.X

Flores-Kim, J., & Darwin, A. J. (2015). Activity of a bacterial cell envelope stress response is controlled by the interaction of a protein binding domain with different partners. *The Journal of Biological Chemistry*, *290*(18), 11417–11430. https://doi.org/10.1074/JBC.M114.614107

Flores-Kim, J., & Darwin, A. J. (2016a). Interactions between the Cytoplasmic Domains of PspB and PspC Silence the Yersinia enterocolitica Phage Shock Protein Response. *Journal of Bacteriology*, *198*(24), 3367–3378. https://doi.org/10.1128/JB.00655-16

Flores-Kim, J., & Darwin, A. J. (2016b). The Phage Shock Protein Response. *Annual Review of Microbiology*, *70*, 83–101. https://doi.org/10.1146/ANNUREV-MICRO-102215-095359

Galle, M., Carpentier, I., & Beyaert, R. (2012a). Structure and function of the Type III secretion system of Pseudomonas aeruginosa. *Current Protein & Peptide Science*, *13*(8), 831–842. https://doi.org/10.2174/138920312804871210

Galle, M., Carpentier, I., & Beyaert, R. (2012b). Structure and Function of the Type III Secretion System of Pseudomonas aeruginosa. *Current Protein & Peptide Science*, *13*(8), 831. https://doi.org/10.2174/138920312804871210

Giegé, R. (2013). A historical perspective on protein crystallization from 1840 to the present day. *The FEBS Journal*, *280*(24), 6456–6497. https://doi.org/10.1111/FEBS.12580

Goure, J., Pastor, A., Faudry, E., Chabert, J., Dessen, A., & Attree, I. (2004). The V antigen of Pseudomonas aeruginosa is required for assembly of the functional PopB/PopD translocation pore in host cell membranes. *Infection and Immunity*, *72*(8), 4741–4750. https://doi.org/10.1128/IAI.72.8.4741-4750.2004

Gueguen, E., Savitzky, D. C., & Darwin, A. J. (2009). Analysis of the Yersinia enterocolitica PspBC proteins defines functional domains, essential amino acids and new roles within the phage-shock-protein response. *Molecular Microbiology*, *74*(3), 619–633. https://doi.org/10.1111/J.1365-2958.2009.06885.X

Halder, P. K., Roy, C., & Datta, S. (2019). Structural and functional characterization of type three secretion system ATPase PscN and its regulator PscL from Pseudomonas aeruginosa. *Proteins*, *87*(4), 276–288. https://doi.org/10.1002/PROT.25648

Heidrich, E. S., & Brüser, T. (2018). Evidence for a second regulatory binding site on PspF that is occupied by the C-terminal domain of PspA. *PloS One*, *13*(6). https://doi.org/10.1371/JOURNAL.PONE.0198564

Heinig, M., & Frishman, D. (2004). STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Research*, *32*(Web Server issue). https://doi.org/10.1093/NAR/GKH429

Hendsch, Z. S., & Tidor, B. (1994). Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Science: A Publication of the Protein Society*, *3*(2), 211–226. https://doi.org/10.1002/PRO.5560030206

Hennig, R., West, A., Debus, M., Saur, M., Markl, J., Sachs, J. N., & Schneider, D. (2017). The IM30/Vipp1 C-terminus associates with the lipid bilayer and modulates membrane fusion. *Biochimica et Biophysica Acta. Bioenergetics*, *1858*(2), 126–136. https://doi.org/10.1016/J.BBABIO.2016.11.004

Hernandez Alvarez, B., Hartmann, M. D., Albrecht, R., Lupas, A. N., Zeth, K., & Linke, D. (2008). A new expression system for protein crystallization using trimeric coiled-coil adaptors. *Protein Engineering, Design & Selection: PEDS*, *21*(1), 11–18. https://doi.org/10.1093/PROTEIN/GZM071

Horovitz, A., & Fersht, A. R. (1992). Co-operative interactions during protein folding. *Journal of Molecular Biology*, *224*(3), 733–740. https://doi.org/10.1016/0022-2836(92)90557-Z

Hubbard*: NACCESS: program for calculating accessibilities - Google Scholar*. (n.d.). Retrieved May 21, 2022, from https://scholar.google.com/scholar?hl=en&as_sdt=0,5&cluster=10212996008891395850

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of Molecular Graphics*, *14*(1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5

Huvet, M., Toni, T., Sheng, X., Thorne, T., Jovanovic, G., Engl, C., Buck, M., Pinney, J. W., & Stumpf, M. P. H. (2011). The evolution of the phage shock protein response system: interplay between protein function, genomic organization, and system function. *Molecular Biology and Evolution*, *28*(3), 1141–1155. https://doi.org/10.1093/MOLBEV/MSQ301

Joly, N., Burrows, P. C., Engl, C., Jovanovic, G., & Buck, M. (2009). A Lower-Order Oligomer Form of Phage Shock Protein A (PspA) Stably Associates with the Hexameric AAA+ Transcription Activator Protein PspF for Negative Regulation. *Journal of Molecular Biology*, *394*(4), 764. https://doi.org/10.1016/J.JMB.2009.09.055

Joly, N., Engl, C., Jovanovic, G., Huvet, M., Toni, T., Sheng, X., Stumpf, M. P. H., & Buck, M. (2010). Managing membrane stress: the phage shock protein (Psp) response, from molecular mechanisms to physiology. *FEMS Microbiology Reviews*, *34*(5), 797–827. https://doi.org/10.1111/J.1574-6976.2010.00240.X

Jones, S. E., Lloyd, L. J., Tan, K. K., & Buck, M. (2003). Secretion defects that activate the phage shock response of Escherichia coli. *Journal of Bacteriology*, *185*(22), 6707–6711. https://doi.org/10.1128/JB.185.22.6707-6711.2003

Jordan, S., Junker, A., Helmann, J. D., & Mascher, T. (2006). Regulation of LiaRS-dependent gene expression in bacillus subtilis: identification of inhibitor proteins, regulator binding sites, and target genes of a conserved cell envelope stress-sensing two-component system. *Journal of Bacteriology*, *188*(14), 5153–5166. https://doi.org/10.1128/JB.00310-06

Khanppnavar, B., & Datta, S. (2018). Crystal structure and substrate specificity of ExoY, a unique T3SS mediated secreted nucleotidyl cyclase toxin from Pseudomonas aeruginosa. *Biochimica et Biophysica Acta. General Subjects*, *1862*(9), 2090–2103. https://doi.org/10.1016/J.BBAGEN.2018.05.021

Kobayashi, R., Suzuki, T., & Yoshida, M. (2007). Escherichia coli phage-shock protein A (PspA) binds to membrane phospholipids and repairs proton leakage of the damaged membranes. *Molecular Microbiology*, *66*(1), 100–109. https://doi.org/10.1111/J.1365-2958.2007.05893.X

Krimm, S., & Bandekar, J. (1986). Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins. *Advances in Protein Chemistry*, *38*(C), 181–364. https://doi.org/10.1016/S0065-3233(08)60528-8

Kumar, S., & Nussinov, R. (1999). Salt bridge stability in monomeric proteins. *Journal of Molecular Biology*, *293*(5), 1241–1255. https://doi.org/10.1006/JMBI.1999.3218

Lee, M. J. Y., Faucher, F., & Jia, Z. (2014). Growth of Diffraction-Quality Protein Crystals Using a Harvestable Microfluidic Device. *Crystal Growth & Design*, *14*(7), 3179–3181. https://doi.org/10.1021/CG500450B

Lee, P. C., Stopford, C. M., Svenson, A. G., & Rietsch, A. (2010a). Control of effector export by the Pseudomonas aeruginosa type III secretion proteins PcrG and PcrV. *Molecular Microbiology*, *75*(4), 924–941. https://doi.org/10.1111/J.1365-2958.2009.07027.X

Lee, P. C., Stopford, C. M., Svenson, A. G., & Rietsch, A. (2010b). Control of effector export by the Pseudomonas aeruginosa type III secretion proteins PcrG and PcrV. *Molecular Microbiology*, *75*(4), 924–941. https://doi.org/10.1111/J.1365-2958.2009.07027.X

Lee, P. C., Zmina, S. E., Stopford, C. M., Toska, J., & Rietsch, A. (2014). Control of type III secretion activity and substrate specificity by the cytoplasmic regulator PcrG. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(19). https://doi.org/10.1073/PNAS.1402658111/SUPPL_FILE/PNAS.1402658111.SAPP.PDF

Lloyd, L. J., Jones, S. E., Jovanovic, G., Gyaneshwar, P., Rolfe, M. D., Thompson, A., Hinton, J. C., & Buck, M. (2004). Identification of a new member of the phage shock protein response in Escherichia coli, the phage shock protein G (PspG). *The Journal of Biological Chemistry*, *279*(53), 55707–55714. https://doi.org/10.1074/JBC.M408994200

Lombardi, C., Tolchard, J., Bouillot, S., Signor, L., Gebus, C., Liebl, D., Fenel, D., Teulon, J. M., Brock, J., Habenstein, B., Pellequer, J. L., Faudry, E., Loquet, A., Attrée, I., Dessen, A., & Job, V. (2019). Structural and functional characterization of the type three

secretion system (T3SS) needle of pseudomonas aeruginosa. *Frontiers in Microbiology*, *10*(MAR), 573. https://doi.org/10.3389/FMICB.2019.00573/BIBTEX

Louis-Jeune, C., Andrade-Navarro, M. A., & Perez-Iratxeta, C. (2012). Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins*, *80*(2), 374–381. https://doi.org/10.1002/PROT.23188

Lund, P., Tramonti, A., & de Biase, D. (2014). Coping with low pH: molecular strategies in neutralophilic bacteria. *FEMS Microbiology Reviews*, *38*(6), 1091–1125. https://doi.org/10.1111/1574-6976.12076

Luo, J., Dong, B., Wang, K., Cai, S., Liu, T., Cheng, X., Lei, D., Chen, Y., Li, Y., Kong, J., & Chen, Y. (2017). Baicalin inhibits biofilm formation, attenuates the quorum sensing-controlled virulence and enhances Pseudomonas aeruginosa clearance in a mouse peritoneal implant infection model. *PloS One*, *12*(4). https://doi.org/10.1371/JOURNAL.PONE.0176883

Lupas, A., van Dyke, M., & Stock, J. (1991). Predicting coiled coils from protein sequences. *Science (New York, N.Y.)*, *252*(5009), 1162–1164. https://doi.org/10.1126/SCIENCE.252.5009.1162

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007a). Phaser crystallographic software. *Journal of Applied Crystallography*, *40*(Pt 4), 658–674. https://doi.org/10.1107/S0021889807021206

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007b). Phaser crystallographic software. *Journal of Applied Crystallography*, *40*(Pt 4), 658–674. https://doi.org/10.1107/S0021889807021206

McPherson, A., & Gavira, J. A. (2014). Introduction to protein crystallization. *Acta Crystallographica. Section F, Structural Biology Communications*, *70*(Pt 1), 2–20. https://doi.org/10.1107/S2053230X13033141

Meuzelaar, H., Vreede, J., & Woutersen, S. (2016). Influence of Glu/Arg, Asp/Arg, and Glu/Lys Salt Bridges on α-Helical Stability and Folding Kinetics. *Biophysical Journal*, *110*(11), 2328–2341. https://doi.org/10.1016/J.BPJ.2016.04.015

Nanao, M., Ricard-Blum, S., di Guilmi, A. M., Lemaire, D., Lascoux, D., Chabert, J., Attree, I., & Dessen, A. (2003). Type III secretion proteins PcrV and PcrG from Pseudomonas aeruginosa form a 1:1 complex through high affinity interactions. *BMC Microbiology*, *3*, 1–9. https://doi.org/10.1186/1471-2180-3-21

Nayek, A., Gupta, P. S., Banerjee, S., Sur, V. P., Seth, P., Das, S., Ul Islam, R. N., & Bandyopadhyay, A. K. (2015). ADSBET2: Automated Determination of Salt-Bridge Energy-Terms version 2. *Bioinformation*, *11*(8), 413–415. https://doi.org/10.6026/97320630011413

Nevskaya, N. A., & Chirgadze, Y. N. (1976). Infrared spectra and resonance interactions of amide-I and II vibration of alpha-helix. *Biopolymers*, *15*(4), 637–648. https://doi.org/10.1002/BIP.1976.360150404

Nicod, C., Banaei-Esfahani, A., & Collins, B. C. (2017). Elucidation of host-pathogen protein-protein interactions to uncover mechanisms of host cell rewiring. *Current Opinion in Microbiology*, *39*, 7–15. https://doi.org/10.1016/J.MIB.2017.07.005

Nikaido, H. (2003). Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and Molecular Biology Reviews : MMBR*, *67*(4), 593–656. https://doi.org/10.1128/MMBR.67.4.593-656.2003

org, W. D. pymol., & 2002, undefined. (n.d.). The PyMOL molecular graphics system. *Ci.Nii.Ac.Jp*. Retrieved May 21, 2022, from https://ci.nii.ac.jp/naid/10020095229/

Osadnik, H., Schöpfel, M., Heidrich, E., Mehner, D., Lilie, H., Parthier, C., Risselada, H. J., Grubmüller, H., Stubbs, M. T., & Brüser, T. (2015). PspF-binding domain PspA1-144 and the PspA·F complex: New insights into the coiled-coil-dependent regulation of AAA+ proteins. *Molecular Microbiology*, *98*(4), 743–759. https://doi.org/10.1111/MMI.13154

Plomp, M., McPherson, A., & Malkin, A. J. (2002). Crystal growth of macromolecular crystals: correlation between crystal symmetry and growth mechanisms. *Journal of Crystal Growth*, *237–239*(1-4 I), 306–311. https://doi.org/10.1016/S0022-0248(01)01926-1

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, *5*(4), 725–738. https://doi.org/10.1038/NPROT.2010.5

Roy, C., & Datta, S. (2018). ASBAAC: Automated Salt-Bridge and Aromatic-Aromatic Calculator. *Bioinformation*, *14*(4), 164–166. https://doi.org/10.6026/97320630014164

Roy, C., Kumar, R., & Datta, S. (2020). Comparative studies on ion-pair energetic, distribution among three domains of life: Archaea, eubacteria, and eukarya. *Proteins*, *88*(7), 865–873. https://doi.org/10.1002/PROT.25878

Rumfeldt, J. A. O., Galvagnion, C., Vassall, K. A., & Meiering, E. M. (2008). Conformational stability and folding mechanisms of dimeric proteins. *Progress in Biophysics and Molecular Biology*, *98*(1), 61–84. https://doi.org/10.1016/J.PBIOMOLBIO.2008.05.004

Rzhepishevska, O., Hakobyan, S., Ekstrand-Hammarström, B., Nygren, Y., Karlsson, T., Bucht, A., Elofsson, M., Boily, J. F., & Ramstedt, M. (2014). The gallium(III)-salicylidene acylhydrazide complex shows synergistic anti-biofilm effect and inhibits toxin production by Pseudomonas aeruginosa. *Journal of Inorganic Biochemistry*, *138*, 1–8. https://doi.org/10.1016/J.JINORGBIO.2014.04.009

Sadikot, R. T., Blackwell, T. S., Christman, J. W., & Prince, A. S. (2005). Pathogen-host interactions in Pseudomonas aeruginosa pneumonia. *American Journal of Respiratory and Critical Care Medicine*, *171*(11), 1209–1223. https://doi.org/10.1164/RCCM.200408-1044SO

Siddiqui, M. Q., Choudhary, R. K., Thapa, P., Kulkarni, N., Rajpurohit, Y. S., Misra, H. S., Gadewal, N., Kumar, S., Hasan, S. K., & Varma, A. K. (2017). Structural and biophysical properties of h-FANCI ARM repeat protein. *Journal of Biomolecular*

Structure & Dynamics, *35*(14), 3032–3042. https://doi.org/10.1080/07391102.2016.1235514

Srivastava, D., Moumene, A., Flores-Kim, J., & Darwin, A. J. (2017a). Psp Stress Response Proteins Form a Complex with Mislocalized Secretins in the Yersinia enterocolitica Cytoplasmic Membrane. *MBio*, *8*(5). https://doi.org/10.1128/MBIO.01088-17

Srivastava, D., Moumene, A., Flores-Kim, J., & Darwin, A. J. (2017b). Psp stress response proteins form a complex with mislocalized secretins in the yersinia enterocolitica cytoplasmic membrane. *MBio*, *8*(5). https://doi.org/10.1128/MBIO.01088-17/ASSET/F9B19982-0F6C-47A9-B4CB-4A77EC27036A/ASSETS/GRAPHIC/MBO0041734840007.JPEG

Sundin, C., Thelaus, J., Bröms, J. E., & Forsberg, Å. (2004). Polarisation of type III translocation by Pseudomonas aeruginosa requires PcrG, PcrV and PopN. *Microbial Pathogenesis*, *37*(6), 313–322. https://doi.org/10.1016/J.MICPATH.2004.10.005

Takaya, A., Takeda, H., Tashiro, S., Kawashima, H., & Yamamoto, T. (2019). Chaperone-mediated secretion switching from early to middle substrates in the type III secretion system encoded by Salmonella pathogenicity island 2. *The Journal of Biological Chemistry*, *294*(10), 3783–3793. https://doi.org/10.1074/JBC.RA118.005072

Terwilliger, T. C., di Maio, F., Read, R. J., Baker, D., Bunkóczi, G., Adams, P. D., Grosse-Kunstleve, R. W., Afonine, P. v., & Echols, N. (2012a). phenix.mr_rosetta: molecular replacement and model rebuilding with Phenix and Rosetta. *Journal of Structural and Functional Genomics*, *13*(2), 81–90. https://doi.org/10.1007/S10969-012-9129-3

Terwilliger, T. C., di Maio, F., Read, R. J., Baker, D., Bunkóczi, G., Adams, P. D., Grosse-Kunstleve, R. W., Afonine, P. v., & Echols, N. (2012b). phenix.mr_rosetta: molecular replacement and model rebuilding with Phenix and Rosetta. *Journal of Structural and Functional Genomics*, *13*(2), 81. https://doi.org/10.1007/S10969-012-9129-3

Torrie, G. M., & Valleau, J. P. (1977). Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, *23*(2), 187–199. https://doi.org/10.1016/0021-9991(77)90121-8

van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. C. (2005). GROMACS: fast, flexible, and free. *Journal of Computational Chemistry*, *26*(16), 1701–1718. https://doi.org/10.1002/JCC.20291

Weber, P. C. (1991). Physical principles of protein crystallization. *Advances in Protein Chemistry*, *41*(C), 1–36. https://doi.org/10.1016/S0065-3233(08)60196-5

Widłak, W. (2013). *Protein Structure and Function*. 15–29. https://doi.org/10.1007/978-3-642-45361-8_2

Wilharm, G., Lehmann, V., Krauss, K., Lehnert, B., Richter, S., Ruckdeschel, K., Heesemann, J., & Trülzsch, K. (2004). Yersinia enterocolitica type III secretion depends on the proton motive force but not on the flagellar motor components MotA and MotB. *Infection and Immunity*, *72*(7), 4004–4009. https://doi.org/10.1128/IAI.72.7.4004-4009.2004

Yahr, T. L., Mende-Mueller, L. M., Friese, M. B., & Frank, D. W. (1997). Identification of type III secreted products of the Pseudomonas aeruginosa exoenzyme S regulon. *Journal of Bacteriology*, *179*(22), 7165–7168. https://doi.org/10.1128/JB.179.22.7165-7168.1997

Yamaguchi, S., & Darwin, A. J. (2012). Recent findings about the Yersinia enterocolitica phage shock protein response. *Journal of Microbiology (Seoul, Korea)*, *50*(1), 1–7. https://doi.org/10.1007/S12275-012-1578-7

Yamaguchi, S., Gueguen, E., Horstman, N. K., & Darwin, A. J. (2010). Membrane association of PspA depends on activation of the phage-shock-protein response in Yersinia enterocolitica. *Molecular Microbiology*, *78*(2), 429–443. https://doi.org/10.1111/J.1365-2958.2010.07344.X

Ye, F., Yang, F., Yu, R., Lin, X., Qi, J., Chen, Z., Cao, Y., Wei, Y., Gao, G. F., & Lu, G. (2018). Molecular basis of binding between the global post-transcriptional regulator CsrA and the T3SS chaperone CesT. *Nature Communications*, *9*(1). https://doi.org/10.1038/S41467-018-03625-X

Zhu, M., Zhao, J., Kang, H., Kong, W., Zhao, Y., Wu, M., & Liang, H. (2016). Corrigendum: Modulation of Type III Secretion System in Pseudomonas aeruginosa: Involvement of the PA4857 Gene Product. *Frontiers in Microbiology*, *7*(JUN). https://doi.org/10.3389/FMICB.2016.00881

PUBLICATIONS

(Accepted)

The Protein Journal

Biophysical and computational approaches to unravel pH-dependent conformational change of PspA assist PspA-PspF complex formation in Yersinia enterocolitica

--Manuscript Draft--

Manuscript Number:	JOPC-D-22-00004R1
Full Title:	Biophysical and computational approaches to unravel pH-dependent conformational change of PspA assist PspA-PspF complex formation in Yersinia enterocolitica
Article Type:	Original Research
Keywords:	Phage shock protein A; pH; intermediate helix (3/10-helix); salt-bridge
Corresponding Author:	Saumen Datta, Ph.D. CSIR-IICB: Indian Institute of Chemical Biology CSIR Kolkata, West Bengal INDIA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	CSIR-IICB: Indian Institute of Chemical Biology CSIR
Corresponding Author's Secondary Institution:	
First Author:	Chittran Roy
First Author Secondary Information:	
Order of Authors:	Chittran Roy
	Rajeev Kumar
	Maruf Hossain
	Arkaprava Das
	Saumen Datta, Ph.D.
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	In enteropathogen Yersina enterocolitica , the genes encoding phage shock proteins are organized in an operon (pspA-E), which is activated at the various types of cellular stress (i.e. extracytoplasmic or envelop stress) whereas, PspA negatively regulates PspF, a transcriptional activator of pspA-E and pspG , and is also involved in other cellular machinery maintenance processes. The exact mechanism of association and dissociation of PspA and PspF during the stress response is not entirely clear. In this concern, we address conformational change of PspA in different pH conditions using various in-silico and biophysical methods. At the near-neutral pH, CD and FTIR measurements reveal a ß-like conformational change of PspA; however, AFM measurement indicates the lower oligomeric form at the above-mentioned pH. Additionally, the results of the MD simulation also support the conformational changes which indicate salt-bridge strength takes an intermediate position compared to other pHs. Futhermore, the bio-layer interferometry study confirms the stable complex formation that takes place between PspA and PspF at the near neutral pH. It, thus, appears that PspA's conformational change in adverse pH conditions abandon PspF from having a stable complex with it and thus, the latter can act as a trans-activator. Taken together, it seems that PspA alone can transduce the cell's adverse signal by changing its conformation.

Received: 23 October 2019 | Revised: 6 January 2020 | Accepted: 25 January 2020
DOI: 10.1002/prot.25878

RESEARCH ARTICLE

PROTEINS WILEY

Comparative studies on ion-pair energetic, distribution among three domains of life: Archaea, eubacteria, and eukarya

Chittran Roy | Rajeev Kumar | Saumen Datta

Structural Biology and Bioinformatics Division, Council of Scientific and Industrial Research—Indian Institute of Chemical Biology, Kolkata, West Bengal, India

Correspondence
Saumen Datta, Structural Biology and Bioinformatics Division, Council of Scientific and Industrial Research—Indian Institute of Chemical Biology, 4, Raja S. C. Mullick Road, Jadavpur, Kolkata 700032, West Bengal, India.
Email: saumen_datta@iicb.res.in

Funding information
ICMR; Council of Scientific and Industrial Research

Peer Review
The peer review history for this article is available at https://publons.com/publon/10.1002/prot.25878.
[Correction added on 29th Feb 2020, after first online publication: Peer review history statement has been added.]

Abstract

Salt-bridges play a unique role in the structural and functional stability of proteins, especially under harsh environments. How these salt-bridges contribute to the overall thermodynamic stability of protein structure and function across different domains of life is elusive still date. To address the issue, statistical analyses on the energies of salt-bridges, involved in proteins' structure and function, are performed across three domains of life, that is, archaea, eubacteria, and eukarya. Results show that although the majority of salt-bridges are stable and conserved, yet the stability of archaeal proteins ($\Delta\Delta G_{net}$ = −5.06 ± 3.8) is much more than that of eubacteria ($\Delta\Delta G_{net}$ = −3.7 ± 2.9) and eukarya ($\Delta\Delta G_{net}$ = −3.54 ± 3.1). Unlike earlier study with archaea, in eukarya and eubacteria, not all buried salt-bridge in our dataset are stable. Buried salt-bridges play surprising role in protein stability, whose variations are clearly observed among these domains. Greater desolvation penalty of buried salt-bridges is compensated by stable network of salt-bridges apart from equal contribution of bridge and background energy terms. On the basis proteins' secondary structure, topology, and evolution, our observation shows that salt-bridges when present closer to each other in sequence tend to form a greater number. Overall, our comparative study provides insight into the role of specific electrostatic interactions in proteins from different domains of life, which we hope, would be useful for protein engineering and bioinformatics study.

KEYWORDS
accessible surface area, ion-pair, PDB, PDB2PQR, salt-bridge

1 | INTRODUCTION

Proteins maintain an intricate balance between rigidity and flexibility for function. The majority of charged residues of protein that form salt-bridges are present on proteins' surface, reducing excess protein hydration and making it less hydrophobic and more flexible. It also helps to overcome the deleterious effect of salt by promoting non-specific electrostatic interaction with salts in solution.[1,2] So, ion-pair or salt-bridge is one of the major contributors to the stability of protein structure and function.[3-5] It is much more visible when the proteins are adapted in the extremes of physical and chemical environments such as high salt and high temperature.[6-9] It has been demonstrated that by inserting a single salt-bridge by mutation in the protein surface increases protein stabilization in the tertiary structure.[10] So, engineered salt-bridge on the protein surface may contribute to the excess of stability.

The energy of salt-bridge can be subdivided into three components, that is, columbic attraction between opposite charge, bridge interactions ($\Delta\Delta G_{brd}$), their desolvation ($\Delta\Delta G_{dslv}$), and background ($\Delta\Delta G_{prt}$) interactions. Columbic attraction is always contributing in tertiary structure of proteins and other two terms could be either contributing or costly. The favorable charge-charge attraction within a salt-bridge is often opposed by unfavorable desolvation of charges and is further modulated by charge-dipole interaction as well as by the ionization behavior of nearest charged group.[11] Thus, net ion-pair energy distribution could either be stabilizing[12-15] or destabilizing[11,16,17] or insignificant.[18]

Received: 15 June 2018 | Revised: 29 October 2018 | Accepted: 12 December 2018

DOI: 10.1002/prot.25648

WILEY PROTEINS

RESEARCH ARTICLE

Structural and functional characterization of type three secretion system ATPase PscN and its regulator PscL from *Pseudomonas aeruginosa*

Pranab Kumar Halder | Chittran Roy | Saumen Datta

Structural Biology and Bioinformatics Division, Council of Scientific and Industrial Research-Indian Institute of Chemical Biology, Kolkata, West Bengal, India

Correspondence
Saumen Datta, Structural Biology and Bioinformatics Division, Council of Scientific and Industrial Research-Indian Institute of Chemical Biology, 4 Raja SC Mullick Road, Jadavpur, Kolkata 700032, West Bengal, India.
Email: saumen_datta@iicb.res.in

Funding information
Department of Science and Technology, Ministry of Science and Technology, Grant/Award Number: SR/SO/BB-036/2015, dated 24/03/2015; University Grant Commission (UGC)

Abstract

Type Three Secretion Systems (T3SS) from many gram-negative bacteria utilize ATPases for the translocation of effector proteins into the eukaryotic host cells through injectisome. Cytosolic regulators effectively control the action of these ATPases. PscN from *Pseudomonas aeruginosa* was an ATPase which was regulated by an uncharacterized PscL. Here we have bioinformatically, biochemically, and biophysically characterized PscN as a T3SS ATPase and PscL as its regulator. In solution, PscN exists predominantly as oligomer and hydrolyzes ATP with V_{max} of 3.9 ± 0.2 μmol/min/mg and K_m 0.93 ± 0.06 mM. Hexameric structure of PscN was observed under AFM and TEM in the presence of ATP. PscL was dimeric in solution and interacted with PscN strongly in Ni-NTA pull-down assay and SPR analysis. PscL was shown to downregulate PscN ATPase activity up to 80% when mixed with PscN in 1:2 ratio (PscN:PscL). SEC data reconfirm the PscN–PscL interaction stoichiometry (ie, 1:2 ratio) which can also be visualized under AFM. In the present study, we have also found out the existence of an oligomeric form of the PscN–PscL heterotrimeric complex. PscL being the regulator of PscN and interacts to form this conformation, which may play an important role too in the regulation of T3SS utilized by *Pseudomonas aeruginosa*. For structural aspect, three dimensional in silico models of PscN, PscL, and PscN–PscL were generated. So, in short, present study tried to enlighten both the structural, functional and mechanistic insights into the action of PscN–PscL complex in T3SS mediated pathogenic pathway.

KEYWORDS

atomic force microscopy, ATPase–regulator complex, hexamer, molecular modeling, transmission electron microscopy

1 | INTRODUCTION

Many gram-negative bacteria possess several different kinds of machinery to secrete their effector proteins into eukaryotic host cells and modulate host cell function, thereby enhancing bacterial survivability and pathogenicity. Till date nine such translocation machineries have been identified, namely type I-IX secretion systems.[1–3] These systems are distinguished mainly by the translocation mechanism of their secreted protein. A wide range of pathogens, like *Yersinia*, *Pseudomonas*, *Shigella*, *Xanthomonas*, *Bordetella*, *Erwinia*, and *Escherichia coli* are equipped with Type Three Secretion System (T3SS) apparatus, a needle-like structure (injectisome) what protrudes out from the bacterial surface, pierces through the eukaryotic membrane, and finally reaches into the eukaryotic cell cytosol.[4,5] This macromolecular apparatus is composed of over 20 different proteins that altogether form three basic components, namely, a large cytoplasmic complex, a transmembrane basal body, and an extracellular needle.[6] Apart from the apparatus, a functional T3SS requires other proteins like effectors, chaperones, translocators, and regulatory proteins. Component proteins of T3SS injectisome are highly conserved among different bacterial clans, but their effectors vary significantly.[7] These effectors are kept in semi-unfolded and inactive form inside the bacterial cytosol by a set of proteins known as a chaperone.[8,9] Attachment of the injectisome with the host cell membrane triggers the delivery of unfolded effectors through a hollow conduit formed inside

BIOINFORMATION
Discovery at the interface of physical and biological sciences

Open access

BIOMEDICAL INFORMATICS

www.bioinformation.net

Volume 14(4)

Software

ASBAAC: Automated Salt-Bridge and Aromatic-Aromatic Calculator

Chittran Roy & Saumen Datta*

[1]Structural Biology and Bioinformatics Division, Council of Scientific and Industrial Research – Indian Institute of Chemical Biology, 4 Raja SC Mullick Road, Jadavpur, Kolkata-700032, West Bengal, India; Saumen Datta – E-mail: saumen_datta@iicb.res.in; Telephone Number: +913324995896 Fax: +913324735197, +913324723967; *Corresponding Author

Received March 19, 2018; Revised April 5, 2018, Accepted April 5, 2018, Published April 30, 2018

doi:10.6026/97320630014164

Abstract:
Biological systems are made of complex networks non-covalent interactions observed among protein-protein, protein-DNA, protein-lipid complexes using hydrogen-bonds, salt-bridges, aromatic-aromatic, van der Waals (vdW), hydrophobic-interactions and several others using distance criteria. Hence, large-scale data analysis is required to understand the principles of biological complex formation. Therefore, it is of interest to analyze non-covalent interaction namely, salt-bridge and aromatic-aromatic contacts in known and modeled protein complex structures. Here, we describe ASBAAC for automatic calculation of salt-bridges and aromatic-aromatic contacts in protein complexes. This software tool is fast, robust and user-friendly for large-scale analysis of inter-chain salt bridges and aromatic-aromatic contact in protein complexes.

Availability: ASBAAC is available for free at https://sourceforge.net/projects/asbaac

Abbreviation: S-B = Salt-Bridge, A-A = Aromatic-Aromatic

Keywords: salt-bridges, aromatic-aromatic contacts, protein, program, software

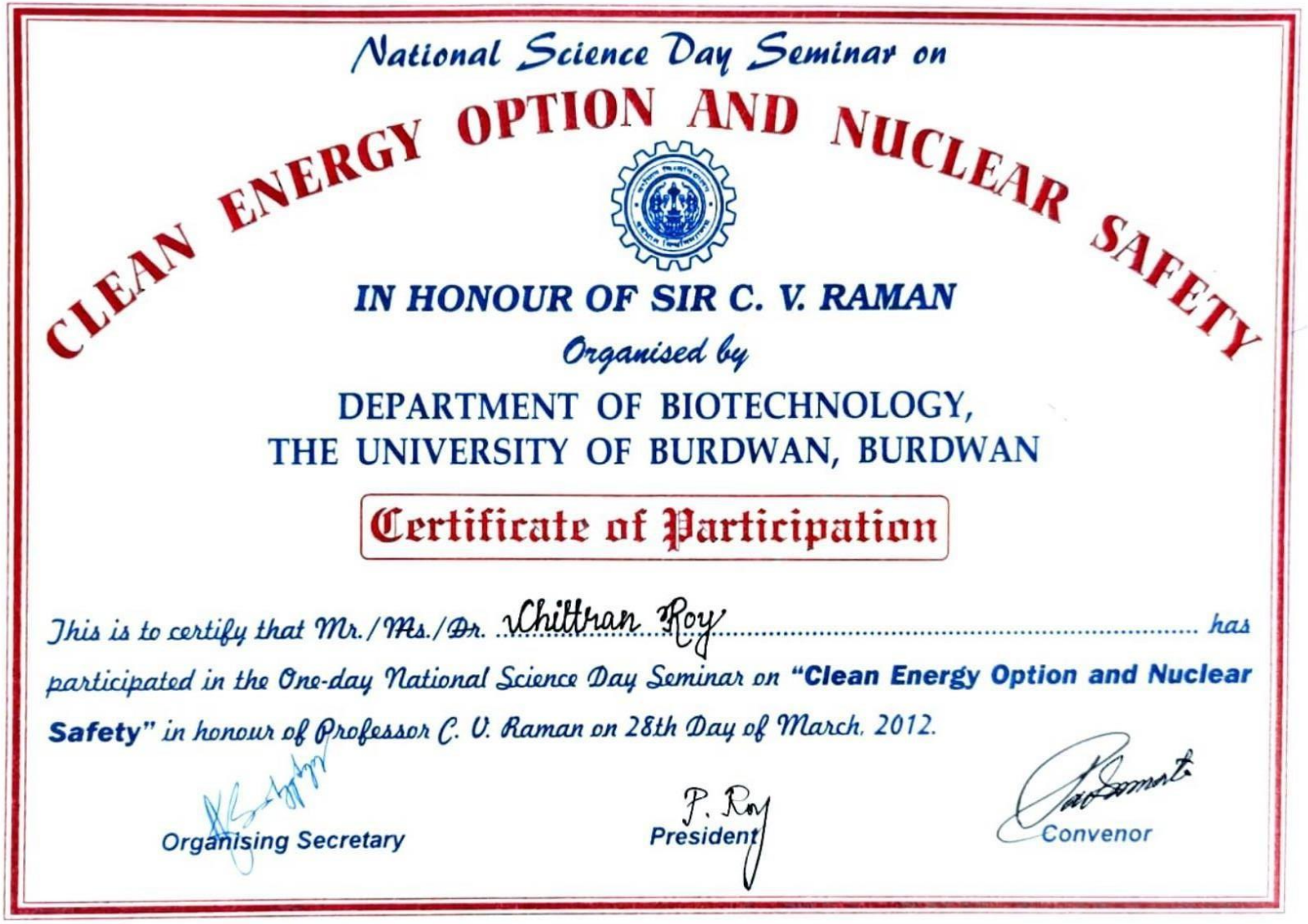
Background:
Interaction between two or more proteins play a crucial role in maintaining cellular systems through non-covalent interactions such as hydrogen bonds, salt bridges and aromatic-aromatic (A-A) contacts. A particular region where two or more proteins interact with each other is called the interface **[1].** Knowledge of amino acids, which are involved in the formation of salt-bridges and A-A contacts with other interacting interfaces between various proteins, is important for the better understanding of protein-protein binding. Salt-bridges most often arise from the anionic amino acids (aspartic acid or glutamic acid) and the cationic amino acids (lysine, arginine or histidine) **[2].** The interaction with these functional groups plays an important role in the formation of the structure and function of proteins. Formation of Salt Bridge between two residues occurs at a distance of 4A° **[3].** There are numerous salt-bridge interaction strategies such as simple or complex, isolated or networked **[4],** intra- helical, coiled, strand or inter-helical, coiled and strand **[5].** Mutagenesis studies and nuclear resonance technique reveal that the contribution of salt-bridge plays an important role in the overall stability of proteins.

Apart from salt-bridge, the aromatic residue interactions also play crucial role in protein stabilization, protein-protein recognition, ligand binding and protein folding **[6].** On average 60% of aromatic side chains of protein involves in the formation of aromatic pair and 80% of which involves in the formation of a network of three or more interacting aromatic side chains **[7].** Interaction occurs when pi rings range is between 4.5 to 7.5 A° assisting in the formation of pi-pi stacking **[8].**

There is about fifty thousand (until February 2018) high-resolution (≤ 2.5 A°) multi-chain protein structure available at the protein data bank (PDB) used for the analysis of salt-bridges and aromatic-aromatic contacts. There has been extensive research on hydrogen bonds. However, data on salt-bridges and aromatic-

ISSN 0973-2063 (online) 0973-8894 (print)

CONFERENCES/SEMINAR/ORAL PRESENTATION



National Science Day Seminar on

CLEAN ENERGY OPTION AND NUCLEAR SAFETY

IN HONOUR OF SIR C. V. RAMAN

Organised by

DEPARTMENT OF BIOTECHNOLOGY,
THE UNIVERSITY OF BURDWAN, BURDWAN

Certificate of Participation

This is to certify that Mr./Ms./Dr. Chittran Roy has participated in the One-day National Science Day Seminar on **"Clean Energy Option and Nuclear Safety"** in honour of Professor C. V. Raman on 28th Day of March, 2012.

Organising Secretary

P. Roy
President

Convenor

UGC Sponsored National Seminar

Journey of Chemistry through Life

This is to certify that Chittran Roy of the University of Burdwan has participated and presented a paper entitled x

in the national seminar held at the college premises on and from 23rd February to 25th February, 2012.

Secretary & Principal

Convener

Organized by
Department of Chemistry
Dr. B. N. Dutta Smriti Mahavidyalaya
Hatgobindapur, Burdwan

In Collaboration with
M. P. Birla Institute of Fundamental Research
Kolkata

Certificate

Presented to

CHITTRAN ROY

for participation in the

Training Programme On Laboratory Safety : Biosafety, Chemical Safety, Radiation Safety & Fire Safety.

Organized by

CSIR-Indian Institute of Chemical Biology, Kolkata

1st February 2016

D. Bhattacharyya
Dr. D. Bhattacharyya
Chairman

Dr. R. K. Bhadra
Chairman

Dr. A. Bandyopadhyay
Chairman

M. C. Debnath
Dr. M. C. Debnath
Organizing Secretary

Prof. Samit Chattopadhyay
Director

Enrolment No. : PhD CW / 2020-21/17

CSIR-Indian Institute of Chemical Biology

(An autonomous body, under the Ministry of Science & Technology, Government of India)

Certificate

(Courses offered as per UGC guidelines, July 2009)

This is to certify that Mr. / Ms. Chittran Roy

has successfully completed the Ph.D Course Work conducted by CSIR- IICB for the session 2020-21

S. Bhattacharyya

Chairperson
Academic Affairs Committee

Arun Bandyopadhyay

Director
CSIR-Indian Institute of Chemical Biology

University of North Bengal

ENLIGHTENMENT TO PERFECTION

Bioinformatics Facility

CERTIFICATE

This is to certify that

Chittran Roy

virtually attended ***"Prof. Asim Bothra Memorial Webinar"*** *held on 30th September 2020, organized by Bioinformatics Facility, University of North Bengal.*

30-09-2020

ARNAB SEN
Prof. Arnab Sen
Coordinator
Bioinformatics Facility
University of North Bengal

Certificate of Participation

This is to certify that

Dr/Mr/Ms Chittran Roy participated in the "International Symposium On Bioinformatics And Artificial Intelligence In Covid-19 Era And Beyond", organized by Centre For Computational Biology & Bioinformatics, Amity Institute Of Biotechnology, Amity University Uttar Pradesh, Noida Campus, held on November 24 - 25, 2020. He / She also gave Oral presentation.

Dr. Puniti Mathur
Professor & Head
Centre for Computational Biology and Bioinformatics

Prof. (Dr.) Chanderdeep Tandon
Director
Amity Institute of Biotechnology

https://www.amity.edu/aib/Academic.asp

This is a computer generated certificate and does not require signature.

INDIAN IMMUNOLOGY SOCIETY
1972

47th Annual Conference of The Indian Immunology Society

Virtual

Organized in collaboration with

CSIR- Indian Institute of Chemical Biology, Kolkata

18th -19th December, 2021

This is to certify that ....Chittran Roy.... *has participated in the 47th Annual conference of The Indian Immunology Society-Virtual.*

Prof. Amit Awasthi
Secretary-IIS

Prof. Sunil K. Arora
President-IIS

Certificate No.: B/0421/11670/W20

THE BIOMICS
Revolutionising Life Sciences

Certificate

This is to certify that

Chittran Roy

Student, Structural biology and bioinformatics
Csir- iicb

has *participated* in the

International Virtual Conference on Frontiers in Biological Research

from **Feb 15th to 21st, 2021**

organised by **Department of Botany, St. Joseph's College (Autonomous), Tiruchirappalli** and **The Biomics, Bengaluru.**

The Biomics

Dr. S. Anusuya
Organising Secretary

Dr. K. Rajan
Organising Secretary

Principal
SJC, Trichy, TN

https://www.thebiomics.com/node/341

IBS 2022

Certificate

This is to certify that

Chittran Roy

has presented an E-poster in the 44th Indian Biophysical Society Meeting entitled

Conceptual Advances in Biophysics and its Applications

held at ***Advanced Centre for Treatment, Research and Education in Cancer,*** *Navi Mumbai, India from* ***30th March to 1st April, 2022***

Dr. Prasanna Venkatraman	Dr. Sudeep Gupta	Prof. Madan Rao	Dr. Kakoli Bose
Convener, IBS	Director, ACTREC	President, IBS	Convener, IBS