



# VCU

Virginia Commonwealth University  
VCU Scholars Compass

---

Theses and Dissertations

Graduate School


---

2023

## The Type IV Pilus Secretin BfpB: Structural Analysis and Binding Interactions

Janay I. Little  
*Virginia Commonwealth University*

Follow this and additional works at: <https://scholarscompass.vcu.edu/etd>

 Part of the [Bacteriology Commons](#), [Molecular Biology Commons](#), and the [Pathogenic Microbiology Commons](#)

© Janay I. Little

---

Downloaded from

<https://scholarscompass.vcu.edu/etd/7269>

This Dissertation is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact [libcompass@vcu.edu](mailto:libcompass@vcu.edu).

# **The Type IV Pilus Secretin BfpB: Structural Analysis and Binding Interactions**

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of  
Philosophy at Virginia Commonwealth University.

by

**Janay I. Little**

Bachelor of Science in Chemistry, Hampton University, 2017

Director: Michael S. Donnenberg, MD

Professor of Internal Medicine, Microbiology and Immunology, Biochemistry and Molecular  
Biology, Senior Associate Dean for Research and Training, and Director of Medical Scientist  
Training Program

Virginia Commonwealth University

Richmond, Virginia

May 2023

## Acknowledgements

I would first like to thank my advisor, Dr. Michael Donnenberg. He was the first professor I spoke to about lab rotations. Despite my initial interest in pharmacology, when I saw his research presentation, I knew his lab was the one I wanted to experience, and I am glad I made that decision. Michael has allowed me to develop into the scientist I am today because of how he provided me with space to learn and think on my own but also knew when I needed more help. He was very patient with me and provided a supportive and encouraging environment to learn and grow as a person and become a critical thinker. There were many ups and downs with this project, but he never gave up on me or my capabilities. I am a stronger person because of it, I leave competent in my abilities as a scientist and mentor, and excited to learn more. I am forever grateful for the experience and training I received from him.

I am also very grateful for the lessons and training I received from Dr. Montserrat Samsó. She has invested in my project and training to the point where I consider her my second mentor. I am thankful for the skills she has instilled in me and for making her laboratory my second lab home. A significant portion of my project would not have been possible without her.

I would also like to thank my committee members, Dr. William Barton, Dr. Kimberly Jefferson, and Dr. Brian Wattenberg, for their incredible guidance, input, and suggestions during my time as a graduate student.

During my time in the Donnenberg lab, I met amazing colleagues who had a significant impact on my training. Dr. Jinlei Zhao was the first person I worked with in the lab and directly trained me until I was independent enough to work on my own. I thank him for his patience and guidance. To Dr. Pradip Singh, he has continuously provided guidance, advice, and encouragement during my entire time here at VCU. Aside from being a great colleague, he became a great friend,

and has a friend in me for life. I would also like to thank Dr. Kavita Iyer, who also trained me during my time in Dr. Samsó's lab. I hope they all continue to be the great scientists and mentors that they are.

There are many family members and friends who provided me with immense support. I am very blessed to have my parents, Eugene and Hope. Their continuous support has allowed me to achieve my goals and dreams. Thank you to my sister, Meagan, for always being there and coming to visit me during this journey. I also thank my friends who were only a phone call or a visit away when I needed them. Lastly, but certainly not least, I thank my fiancé, Darryl. He has been the rock to my support system during this entire journey. Through him, I have gained additional support and love that I never imagined. I cannot wait to see what is next for us in our lives together.

## Table of Contents

<b>Acknowledgements</b> .....	ii
<b>List of Tables</b> .....	vi
<b>List of Figures</b> .....	vii
<b>List of Abbreviations</b> .....	viii
<b>Abstract</b> .....	1
<b>Chapter I: Introduction</b> .....	3
Background and Rationale .....	3
Enteropathogenic <i>Escherichia coli</i> .....	4
Type IV Pili.....	6
T4P Functions.....	6
Classification .....	8
Assembly of T4Ps.....	9
The Bundle Forming Pilus of EPEC .....	12
Secretins .....	14
Hypothesis and Specific Aims .....	17
<b>Chapter II: The Structure of the secretin BfpB</b> .....	19
Introduction .....	19
Materials and Methods .....	20
Results .....	25

Discussion .....	44
<b>Chapter III: Characterization of the periplasmic protein BfpU .....</b>	<b>46</b>
Introduction .....	46
Materials and Methods .....	47
Results .....	53
Discussion .....	65
<b>Chapter IV: Interaction between BfpU and BfpB periplasmic domains .....</b>	<b>67</b>
Introduction .....	67
Methods .....	69
Results .....	74
Discussion .....	81
<b>Chapter V: Discussion .....</b>	<b>84</b>
<b>References .....</b>	<b>89</b>

## List of Tables

Table 1: Strains and plasmids used in chapter II .....	20
Table 2: Primers used in chapter II.....	21
Table 3: Summary of conditions for size exclusion chromatography .....	28
Table 4: Summary of secretin and substrate diameters .....	36
Table 5: Summary of predicted secondary structures of BfpB, terminal residues of BfpB peptide constructs, and molecular weights .....	38
Table 6: Strains and plasmids used in chapter III.....	47
Table 7: Primers used in chapter III.....	48
Table 8: Summary of auto-aggregation assay.....	63
Table 9: Strains and plasmids used in chapter IV.....	69
Table 10: Summary of surface plasmon resonance results of BfpU and BfpB N0+N3 .....	78

## List of Figures

Figure 1: Schematic of the T4P system and component nomenclature .....	11
Figure 2: Overview of secretin features and domains of secretin-reliant systems .....	16
Figure 3: Affinity chromatography of BfpB under various conditions .....	26
Figure 4: Evaluation of BfpB under various concentrations of SB3-14.....	31
Figure 5: BfpB cryo-EM data and model .....	35
Figure 6: Results of MyHits motif scan and AlphaFold structure of BfpB (Q9S142) .....	41
Figure 7: Purification of BfpB N0 and N0+N3 .....	42
Figure 8: Thermal stability of BfpB N0 and N0+N3.....	43
Figure 9: Purification of BfpU.....	54
Figure 10: Thermal stability of BfpU .....	56
Figure 11: Circular dichroism of BfpU.....	57
Figure 12: AlphaFold predicted structure of BfpU (Q47069) .....	58
Figure 13: Auto-aggregation assay with bfpU mutants .....	62
Figure 14: Purification of BfpU N89Y and N125S .....	64
Figure 15: Schematic of surface plasmon resonance experimental setup with BfpB N0+N3 and BfpU.....	73
Figure 16: Surface plasmon resonance response vs time plots with BfpB N0+N3 and BfpU .....	77
Figure 17: Surface plasmon resonance response vs time plots with BfpU IgG ( $\alpha$ -BfpU) and BfpU.....	79
Figure 18: Surface plasmon resonance response vs time plots with BSA and BfpU .....	80



## List of Abbreviations

AA	auto-aggregation
A/E	attaching/effacing
AMIN	amidase
BFP	bundle forming pilus
BSA	bovine serum albumin
CD	circular dichroism
CM5	carboxymethyl dextran
Com	competence type IV pili
DAEC	diffusely adhering <i>Escherichia coli</i>
EAEC	enteroaggregative <i>Escherichia coli</i>
EHEC	enterohemorrhagic <i>Escherichia coli</i>
EIEC	enteroinvasive <i>Escherichia coli</i>
EM	electron microscopy
EPEC	enteropathogenic <i>Escherichia coli</i>
ETEC	enterotoxigenic <i>Escherichia coli</i>
IM	inner membrane
IPTG	isopropyl $\beta$ -D-1-thiogalactopyranoside
KH	nuclear ribonucleoprotein K homology
LB	Luria-Bertani
LEE	locus of enterocyte effacement
Ni-NTA	nickel nitrilotriacetic acid
NMR	nuclear magnetic resonance

OM	outer membrane
pEAF EPEC	adherence factor plasmid
RMSD	root-mean-square deviation
RU	resonance units
SB3-14	3-(N,N-Dimethyltetradecylammonio)propanesulfonate
SPR	surface plasmon resonance
STEC	Shiga toxin-producing <i>Escherichia coli</i>
Tad	tight-adherence pili
TCP	toxin-coregulated pilus
T2S	type II secretion
T2SS	type II secretion system
T3S	type III secretion
T4P	type IV pilus
T4Ps	type IV pili
T4aPs	type IVa pili
T4bPs	type IVb pili
T4bP	type IVb pilus
T4cPs	type IVc pili

## Abstract

### **The Type IV Pilus Secretin BfpB: Structural Analysis and Binding Interactions**

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

By Janay I. Little

Director: Michael S. Donnenberg, MD

Professor of Internal Medicine, Microbiology and Immunology, Biochemistry and Molecular Biology, Senior Associate Dean for Research and Training, and Director of Medical Scientist Training Program

Enteropathogenic *Escherichia coli* (EPEC) causes severe diarrhea in young children. The type IV pilus (T4P) of EPEC, known as the bundle-forming pilus (BFP), plays an important role in EPEC pathogenesis. T4Ps are a family of surface appendages that are important for adhesion, colonization, biofilm formation, virulence, twitching motility and many other functions. One essential component of the BFP system is the secretin, BfpB. Secretins are a large family of integral outer membrane proteins found in T4Ps as well as type II and type III secretion systems, and filamentous phages. Details of the secretin structure have been limited to the overall shape, with atomic resolution of only the soluble amino-terminus domains, which impede our understanding of T4P biogenesis. The goals of this project are: 1) determine the structure of BfpB via cryo-electron microscopy; 2) define the amino-terminus domains of BfpB and its interactions with BfpU, an essential periplasmic protein of the system. We present a 7.1 Å resolution cryo-EM structure of BfpB, the first of a type IVb pilus secretin. Internal features suggest that BfpB is composed of 17 subunits with C17 symmetry. Structural and bioinformatic analyses suggests that

monomeric BfpB possesses two amino-terminal domains, N0 and N3, which allowed us to successfully purify BfpB N0 and N0+N3 domains for interaction studies. Additionally, we have successfully purified BfpU for structure determination and interaction studies. Also, a random mutagenesis approach was used for further characterization of BfpU. Furthermore, surface plasmon resonance suggests the possibility that BfpB and BfpU interact with affinity in the micromolar range, but this result must be interpreted cautiously in light of similar interactions between BfpU and proteins chosen as negative controls. Results from these studies will not only further our understanding of BFP biogenesis, but also enhance research for understanding other T4Ps and secretion systems that are confirmed virulence factors.

## Chapter I: Introduction

### Background and Rationale

Infectious diarrhea is one of the leading causes of mortality and morbidity worldwide in children under the age of five. An estimated 15 million cases and 500,000 deaths occur each year globally, with most case occurrences from low to middle income earning countries (1, 2). Though deaths in the U.S. are uncommon, an estimated 2.1 million outpatient visits, 254,000 emergency department visits, and 68,600 hospitalizations occur annually (3). Bacterial diarrhea management involves providing standard fluid and electrolyte therapy, keeping a watchful eye for potential complications, and administering antibiotics to children who exhibit moderate to severe illness (1). Multidrug resistance is a continuously worsening global health threat. The prevalence of antibiotic-resistant pathogens continues to increase, especially enteric pathogens with a high incidence of severe disease (4). The number of novel antibiotics approved by the U.S. Food and Drug Administration has decreased over the past 30 years (5). Most antibiotics commonly used can not only eliminate the targeted pathogen but also a significant portion of the host's microbiome. Normally, these commensal bacteria play a crucial role in regulating opportunistic pathogens. However, the depletion of commensals due to antibiotics usage can potentially increase secondary infection caused by opportunistic pathogens (6).

New mechanisms and targets for treating infections must be developed. Virulence factors allow bacteria to colonize specific niches in the host, cross barriers in tissues, and avoid the immune system. Research involving virulence factors and anti-virulence compounds provide valuable leads towards the development of innovative therapies because they are designed to block disease without inhibiting growth or killing the pathogen (7). The bacteria would be cleared by the host immune response with little impact on the normal human microbiome (8). These anti-

virulence compounds may not exert strong selective pressure for emergence of resistance. Of the virulence factors that pathogenic bacterial possess, molecular machines, which include secretion and piliation systems, are of interest because they are often localized to the cell surface (7, 8).

Bacterial molecular machines are complex assemblies of multiple proteins that exhibit a highly organized quaternary and tertiary structure. The functioning of these machines is heavily reliant on various protein-protein interactions. Inhibitors that target the interface between interacting partners are particularly promising, as both proteins would have to evolve to resist the inhibitor's effect. Additionally, the probability of off-target effects is low since most components of bacterial molecular machines lack eukaryotic homologs and are highly conserved across different pathogenic bacteria (9). Therefore, a thorough understanding of the steps involved in machine assembly and the interactions that take place could aid in the development of specific inhibitors that can serve as broad-spectrum antibiotics in the future. For these reasons, further understanding the type IV pilus (T4P) is important, as it is a molecular machine that is a confirmed virulence factor of many pathogenic bacteria.

### Enteropathogenic *Escherichia coli*

Pathogenic *Escherichia coli* are classified into pathotypes based on their virulence genes (10). There are six pathotypes associated with diarrhea, which are collectively known as diarrheagenic *E. coli*. These pathotypes include enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), Shiga toxin-producing *E. coli* (STEC) (also known as enterohemorrhagic *E. coli* or EHEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC), and diffusely adherent *E. coli* (DAEC) (11, 12). Of these pathotypes, EPEC is associated with acute and persistent diarrhea and showed a higher attributable risk of infant mortality compared to other enteric pathogens (1, 12, 13).

EPEC are distinguished from other *E. coli* by their inability to produce Shiga toxins and their ability to inflict attaching and effacing (A/E) lesions upon their host cells, which lead to loss of the host's microvilli followed by the formation of pedestals on the epithelial cell (14). All the genes necessary for the A/E lesions are located in a pathogenicity island known as the locus of enterocyte effacement (LEE) (15). There are many virulence factors encoded in the LEE, including a type III secretion (T3S) system. The T3S system is another molecular machine that spans both bacterial membranes and is composed of several proteins that form a needle which protrudes from the outer membrane (OM). The T3S system allows EPEC to transport proteins from the bacterial cytoplasm through to the OM of EPEC and into the cytoplasm of the attached host cell (16). Elsewhere in the genome of typical EPEC strains is the EPEC adherence factor plasmid (pEAF), encoding additional virulence factors such as the bundle-forming pilus (BFP), a member of the T4P family of surface appendages (17). BFPs are required for EPEC adherence, form the characteristic localized-adherence pattern on human epithelial cells, and are responsible for auto-aggregation (18, 19). These auto-aggregate clusters form when in liquid culture at 37 °C and contain hundreds or even thousands of bacteria visible to the naked eye (20).

The ability of typical EPEC to cause diarrhea is well established. However, the virulence of atypical EPEC is less certain (21, 22). Typical EPEC express BFPs while atypical EPEC do not (23–25). Association between atypical EPEC and moderate-to-severe diarrhea is not often observed, whereas typical EPEC was significantly associated with moderate-to-severe diarrhea during the first 2 years of life. Additionally, analyses limited to cases showed that typical EPEC was significantly associated with death in infants aged 0–11 months (26, 27). Studies show that BFPs are a virulence factor of EPEC (17, 19), therefore further characterization of BFP is necessary.

## Type IV Pili

Type IV pili (T4Ps) are the most widespread class of fimbriae, found in Gram-negative, Gram-positive bacteria and archaea (28). Many Gram-negative pathogens possess T4Ps, including *E. coli*, *Pseudomonas aeruginosa*, *Neisseria gonorrhoeae* and *N. meningitidis*, and *Vibrio cholerae*, *Salmonella enterica* serovar Typhi, and *Yersinia pseudotuberculosis* (28–32). They are long, flexible surface fibers several micrometers in length and approximately 45-100 Å in width (33). T4Ps are composed of thousands of subunits known as major pilins, and fewer copies of minor pilins that are localized to the pilus tip or scattered throughout the pilus fiber. T4Ps have several functions including host colonization, adhesion, biofilm formation, twitching motility, auto-aggregation, and DNA uptake (Stone and Abu Kwaik 1998; Kirn et al. 2000; Mattick 2002; Luke et al. 2004). T4Ps also contribute to virulence and pathogenicity in many bacterial species (19, 34, 38–46). Assembly and retraction of T4Ps is directed by a complex multicomponent machine that, in Gram-negative bacteria, spans both the inner membrane (IM) and OM. The proteins that work together for T4P biogenesis include pilin proteins, a polytopic IM protein, at least one nucleotide binding protein for pilus extension, a group of four proteins that span the IM, and an OM secretin (47, 48). The type II secretion system (T2SS), which transports proteins across the OM, is closely related to T4Ps, as it has similar components (49, 50). A key difference between the two systems is that the T2SS has pseudopilins that assemble into a pseudopilus, which may act as a piston to force protein cargo into the extracellular milieu (51).

## T4P Functions

T4Ps serve a range of important functions, with many bacteria using them to move along surfaces through a mechanism known as twitching motility (30, 52). This movement involves the extension and retraction of T4Ps and generation of high mechanical forces, up to 140 pN (53, 54)



The retraction ATPase is one of the most powerful biological motors described. In particular, the retraction of T4Ps is facilitated by the depolymerization of pilins from the pilus into the membrane, driven by cytoplasmic ATPases PilT (55). The regulation of twitching motility can vary between species and involves the localization of both pili and ATPases. In addition to facilitating twitching motility, T4Ps can also attach to a variety of surfaces through major or minor pilin subunits (56, 57). This ability to adhere to different surfaces is a key aspect of T4Ps' function as virulence factors in pathogens.

The ability of bacteria to adhere to surfaces is essential for their colonization of host tissues and formation of microcolonies. Potential host cell receptors for major pilins have been identified for several bacterial species including *P. aeruginosa*, *N. gonorrhoeae*, and EPEC (20, 58, 59). T4Ps facilitate both host cell interactions and bacterial cell-to-cell interactions, including auto-aggregation and microcolony formation. The ability of T4Ps to facilitate bacterial adherence also leads to the formation of biofilms, which enhance bacterial survival, infection, and resistance to host defenses and antimicrobial agents (33, 60). Microcolony formation has been observed to occur through direct lateral contact between pili (35).

T4Ps often facilitates natural transformation, which is a form of horizontal gene transfer (61–64). Mutations of the major and minor pilins have been described that affect DNA binding without affecting other T4P functions (64). DNA was observed binding to the pilus tip during uptake (65). The role of T4Ps in twitching motility, adherence, biofilm formation, and transformation all contribute to various ways in which T4Ps are associated with disease and pathogenicity in many bacterial species.

## Classification

T4Ps can be classified based on genetic, structural, and biochemical features. Some of the different classes include type IVa pili (T4aPs) and type IVb pili (T4bPs). Other proposed classes include tight-adherence (Tad) pili, which have gained recognition as T4cPs, competence T4Ps (Com systems), mannose-sensitive hemagglutinin pilus, and archaeal T4Ps (66). Many Gram-negative bacteria produce T4aPs, while T4bPs are found in enteric pathogens like EPEC and ETEC, *V. cholerae* (toxin-coregulated pilus, TCP), and *S. enterica* serovar Typhi (33, 67, 68). Prepilins from T4aPs generally have short N-terminal leader sequences (5-6 amino acids), whereas those from T4bPs tend to have longer leader sequences (15-30 amino acids) (33). T4bPs are composed of major pilin subunits that are usually longer than those of T4aPs, resulting in larger pilin structures and thicker pili. T4bPs typically contain variable N-methylated N-terminal amino acids (methionine, leucine, or valine), while T4aP pilins contain an N-methylated N-terminal phenylalanine (67).

The homology among certain T4P proteins and those of related systems implies a common evolutionary origin. Evidence suggests that the Last Universal Common Ancestor, which existed around 4 billion years ago, possessed a T4P (66). The most conserved proteins found in all T4P and related systems are the ATPase, the polytopic IM protein, and the pilin or pilin-like protein (66). A phylogeny based on these proteins reveals significant evolutionary events in T4P history, such as loss, duplication, fission, and horizontal transfer of T4P genes. An ancestral archaeal organism with a T4P containing two genes encoding the IM platform evolved into an adhesive archaeal pilus-like system and was later horizontally transferred to a bacterium (possibly Gram-negative), leading to the present Tad pilus systems. Further analysis showed that T4bP diverged early in T4P evolution within bacteria, followed by the divergence of the Com pilus systems from

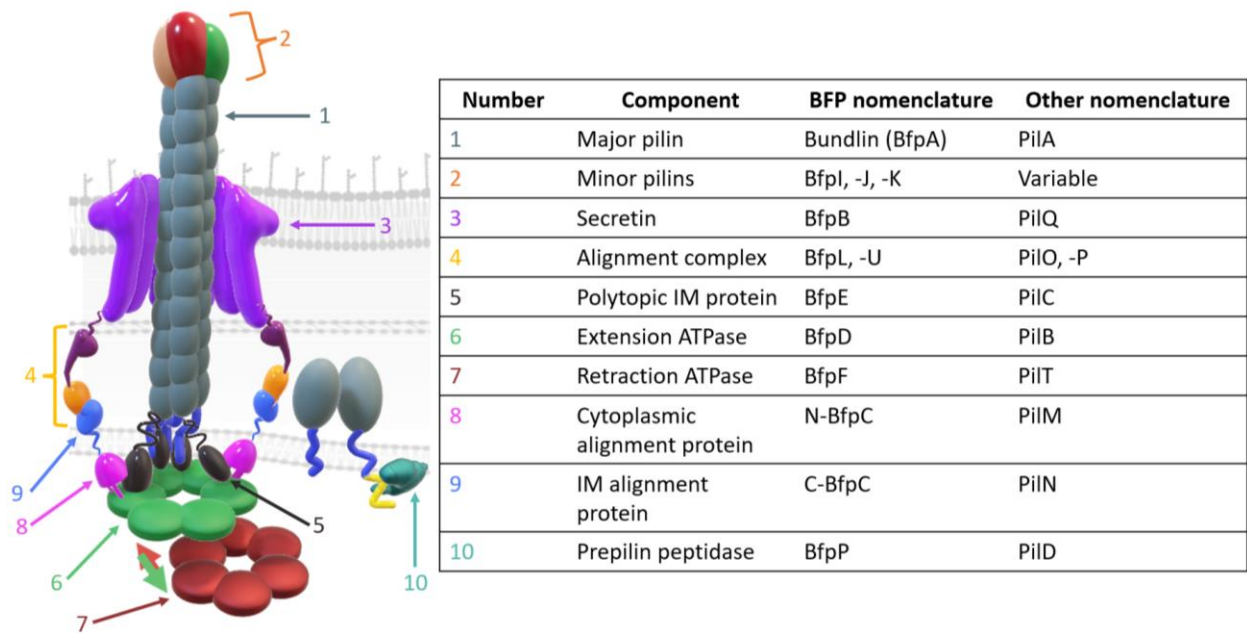
the T4aP, mannose-sensitive hemagglutinin pilus, and T2S systems. Intriguingly, this indicates that the T2SS is more closely related to T4aPs than T4bPs and that the Com system arose early in the evolution of bacterial T4Ps.

### Assembly of T4Ps

The process of assembling T4Ps in Gram-negative bacteria involves a complex biomolecular machine that spans both the IM and OM. Major pilin subunits (often called PilA), which are the building blocks of the pilus, are initially located in the IM as prepilin proteins (69). These prepilin proteins are then processed into mature pilins by the prepilin peptidase (PilD). An IM machinery, comprised of core proteins (PilC, PilM, PilN, PilO, and PilP) and associated ATPases, somehow extricates pilin from the membrane and assembles the mature pilin subunits into a helical structure. The N-terminal alpha helices of the pilins provide stability to the pilus superstructure through hydrophobic interactions (70). The associated cytoplasmic ATPases supply the energy for the pilus to extend (PilB) and retract (PilT) through polymerization and depolymerization of the major pilin subunits (71). When the pilus retracts, these subunits return to the membrane pool of pilins (69). Finally, the assembled pilus fiber passes through the OM secretin complex (PilQ) to reach the extracellular space (72). While these general findings reveal an overall interaction network, they do not provide specifics on how these proteins function in pilus assembly.

Defining distinct steps in T4P biogenesis and identifying the specific roles of individual proteins have been challenging. Most gene mutations result in a non-piliated phenotype without additional distinguishing characteristics, making it difficult to pinpoint the functions of each protein (57). Also, biochemical and structural biology studies have been hindered by the fact that many components are integral membrane proteins that are challenging to purify in their native

states. Currently, our understanding of pilus biogenesis primarily revolves around the initial steps of pilin monomer processing and translocation (73). However, details regarding subsequent steps and the proteins required for them are still unresolved.



*Figure 1: Schematic of the T4P system and component nomenclature. The numbers in the schematic correspond to the numbers listed in the table, which describe the various protein functions and names. “Other nomenclature” is the common name of the protein, though many T4P systems vary in names. Schematic adopted from (74).*

## The Bundle Forming Pilus of EPEC

The BFP is a paradigm system for T4bPs. While BFPs have many features in common with T4aPs, they differ in many ways. Located in pEAF, the *bfp* operon consists of 14 open reading frames (75, 76). When placed on a plasmid under the control of an inducible promoter, the *bfp* operon is sufficient for BFP biogenesis and auto-aggregation by laboratory *E. coli* strains (75, 77). All but two of these genes are required for expression of BFP (78). The exceptions are *bfpF*, which encodes the ATPase required for efficient pilus retraction (19, 79) and *bfpH*, which is classified as a pseudogene (78).

The major pilin subunit bundlin, encoded by the *bfpA* gene, is uniquely structured compared to T4a pilin subunits. Bundlin possesses an  $\alpha/\beta$  fold composed of four  $\alpha$ -helices that surround a seven-stranded, twisted, contiguous  $\beta$ -sandwich that is composed parallel and anti-parallel  $\beta$ -strands (80). Bundlin binds to *N*-acetylglucosamine glycan receptors on host cell surfaces (81). Also, bundlin has highly conserved residues as well as surface-exposed residues that differ among EPEC strains, likely due to evolutionary selective pressure (82), and this sequence variability may contribute to human reinfection (83).

The minor pilin subunits are encoded by the genes *bfpI*, *bfpJ*, and *bfpK*. These minor pilin genes are expressed at lower levels compared to *bfpA*. Like T4aP systems, mutations within the BFP minor pilins lead to a non-piliated phenotype (84, 85). Moreover, it was shown that the BFP minor pilins are dispersed throughout the pilus (84). The transmembrane prepilin peptidase BfpP cleaves pre-bundlin and the minor pilins at a specific N-terminal site to remove a leader sequence (84, 86). It is also suggested BfpP, like other prepilin peptidases, transfers a methyl group to the new amino terminus.

The *bfp* operon encodes two ATPases, BfpD for extension and BfpF for retraction. In contrast to the crystal structures of extension ATPases from other Gram-negative bacteria that contain T4aPs, high-resolution cryo-electron microscopy (cryo-EM) structures of BfpD revealed a six-fold symmetry instead of a two-fold symmetry (87). This difference may indicate a different catalytic mechanism. The specific interactions of BfpD with the polytopic IM protein and IM complex are not fully described. Thus, the mechanism by which the chemical energy of ATP hydrolysis is converted into mechanical energy to extract the pilin from the membrane is still not fully understood.

BfpE is the essential polytopic IM protein and PilC homolog. Extensive topology mapping suggested that it has four transmembrane domains (88). The protein BfpC spans the inner membrane once with large cytoplasmic and periplasmic domains. Despite its lack of sequence homology, the crystal structure of the N-terminal cytoplasmic domain of BfpC showed that it is a structural homolog of PilM (89). It is likely that the BfpC periplasmic C-terminus may be homologous to PilN which binds to PilM. Additionally, the cytoplasmic N-termini of both BfpC and BfpE were shown to co-elute with BfpD through nickel affinity chromatography (89). The protein BfpL is predominantly periplasmic with its N-terminus embedded in the IM (90) and has been shown to interact with the periplasmic C-terminus of BfpC. There are no sequence homologs of PilO in the BFP system however, BfpL may represent this protein.

The lipoprotein BfpB is the secretin of the BFP system and PilQ homolog. The N-terminus of BfpB resides in the periplasm and Cys18 serves as the outer membrane (OM) lipid anchor (91). A low-resolution structure of BfpB via negative stained transmission electron microscopy (TEM) suggested that BfpB is composed of 12 subunits with a gated pore (92). Also, localization studies showed that BfpB is distributed throughout the cell OM, rather than only at cell poles compared

to some other secretins (91). The proteins BfpG and BfpU are periplasmic however, both are partially localized to the OM in the presence of BfpB (93). Additional evidence of their interaction with BfpB was demonstrated by cross-linking and purification studies, which showed that all three proteins can be recovered after fusing any of them to an affinity tag. Although BfpG and BfpU do not share any sequence homology with other T4P proteins, it is possible that one of them may be an unrecognized PilP homolog.

### Secretins

Secretins are large multimeric OM channels essential to T4P, type II secretion (T2S) and type III secretion (T3S) systems as well as some filamentous bacteriophages (94). Proteins of the secretin family are typically composed of 12-15 monomers with cyclic symmetry that oligomerize into a ~1 MDa OM pore (95–97). These proteins are highly stable and resistant to heat, detergents and denaturing agents (98–100). They have plugged central channels, which open to permit extension of pili or export of other substrates. They have variable N-terminal domains in the periplasm with a conserved secretin domain in the C-terminus (101). The studied periplasmic domains of secretins have been solved by x-ray crystallography, but there are only a few resolved structures of the secretin domain. Most of the information known of secretins are from T2S and T3S systems.

The secretin domain is rich in  $\beta$ -sheets and forms a  $\beta$ -barrel (97, 102, 103). To date, studies have shown that the secretin monomer possesses four  $\beta$ -strands that contribute to the OM  $\beta$ -barrel and gated pore. The C-terminus also possesses an S domain which seems to be essential for OM localization and oligomerization however, this domain has only been observed in T2S and T3S system secretins (101). The secretin N-terminal domains are variable in length and sequence. It is predicted that these domains are tuned to the function of their system and the distinct IM and



periplasmic components they interact with. Studies of secretins have shown that these N-terminal domains include a TonB-dependent transduction domain (N0), multiple heterogeneous nuclear ribonucleoprotein K homology-like (KH-like) domains (N1-N5), and amidase N-terminal (AMIN) domains which have only been observed in T4P secretins (101, 104–107). The N0 domain adopts a  $\beta\alpha\beta\beta\alpha\beta\beta$  fold (96, 97, 108). Very few structures of secretins by cryo-EM were able to resolve the N0 domain, as it is thought to be flexible. The N1-N5 domains typically adopt a  $\beta\alpha\beta\beta\alpha$  fold. Of these various N-terminal domains, the N0 and N3 domains have been observed in every secretin thus far (109). Domain organizations of various secretins are summarized in figure 2. Additional structural studies are needed to further understand the mechanism of secretins.

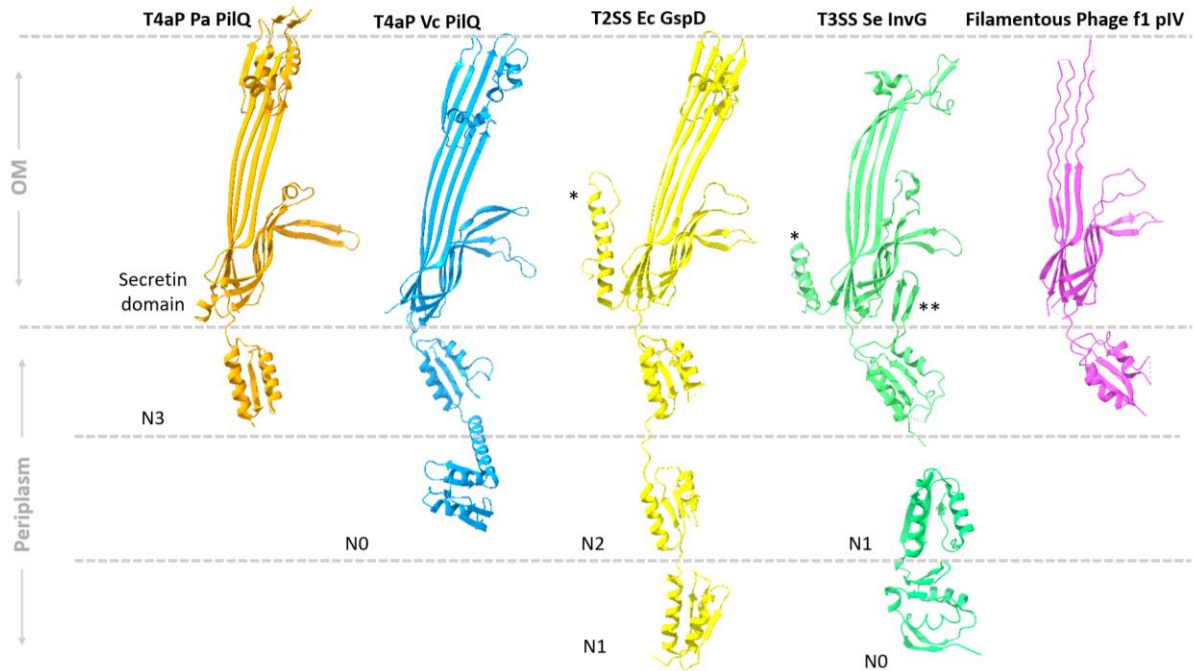


Figure 2: Overview of secretin features and domains of secretin-reliant systems. Monomers of secretin proteins from T4P, T2S, T3S, and filamentous phage systems. The secretin domain is highly conserved and consists of 4  $\beta$ -strands per monomer. N3 domain is next to the secretin domain and in the periplasm. Domain organization is variable thereafter, but usually ends with the N0 domain. N0 has not been solved by cryo-EM in most structures, except for T4aP PilQ of *V. cholerae*. Structures in figure: *P. aeruginosa* PilQ (6VE3, orange), *V. cholerae* (6W6M, blue), *E. coli* GspD (5WQ7, yellow), *S. enterica* InvG (6PEE for N3-secretin domains, 6PEM for N0-N1 domains, green), and f1 phage pIV (7OFH, magenta). \* Indicates S domain. \*\* Indicates N3  $\beta$ -hairpin.

## Hypothesis and Specific Aims

To further refine our understanding of BFP biogenesis, we sought to further characterize the secretin BfpB, BfpU, and their interactions with one another. Based on previous findings, I hypothesized that the structure of BfpB is a dodecamer with a carboxy-terminal beta barrel, has gate forming amino-terminal domain(s) for bundling and substrates to pass through, and the amino-terminal domain(s) interact with BfpU. The following specific aims were established to test this hypothesis: (1) to determine the structure of BfpB and (2) to determine whether BfpB interacts directly with BfpU. The structure of BfpB was characterized by cryo-electron microscopy and bioinformatics, where we present the first model of a T4bP secretin and established the putative amino-terminal domains. These studies have generated a new hypothesis that BfpB has a greater number of subunits to accommodate BFP, which is wider than other pili or substrates of other secretin reliant systems. Using biochemical, biophysical, and molecular biology methods, we were able to further characterize BfpU for interaction studies, and based on our findings we propose that BfpU is a structural homolog of PilP. PilP is known to interact with the T4aP secretin PilQ. For this and other reasons to be elaborated below, we hypothesize that BfpU interacts with BfpB. Lastly, using surface plasmon resonance, the interaction between BfpB and BfpU was studied. This is the first report of a quantitative interaction study between a secretin and its neighboring component.

Although the sequences of BfpB and BfpU differ from homologs in other species, it is important to investigate their functions to fully understand all T4P systems and their similarities in mechanisms. Understanding the secretin and other neighboring proteins in the periplasm will not only help further characterize the biogenesis of T4P, but also enhance research for understanding other secretion systems that require secretins, such as T2S and T3S which are

confirmed virulence factors. The secretin protein is a potential target for the development of new antimicrobial agents, as inhibiting its function could prevent T4P, T2S, and T3S assembly and disrupt the ability of bacteria to cause infections. Furthermore, using BFPs as a model for T4P biogenesis is advantageous due to its contribution to virulence in typical EPEC infections.

## Chapter II: The Structure of the secretin BfpB

### Introduction

Most of the knowledge available about secretins is from the T2S and T3S systems, but as of late, more information has emerged from T4aP secretins. The T4bP diverged early in the T4P evolution within bacteria and, despite the similarities between T4bP and T4aP, there may be critical and fundamental differences among related systems. The sequence homology among the secretin outer membrane family of proteins is generally low (20-70%) however, these proteins share a common architecture, consisting of a  $\beta$ -barrel structure that spans the outer membrane and a series of N-terminal periplasmic domains that interact with other proteins or ligands (101). It is important to study secretins of all system types and sub-classes to gain a clear understanding of their function and mechanisms. Currently, there are no known structures of T4bP secretins.

Using negative staining transmission electron microscopy, previous studies from the Donnenberg laboratory yielded a low-resolution ultrastructure of BfpB. Additionally, the results were interpreted to suggest that BfpB is a dodecamer, ~700 kDa with rotational symmetry, and a central pore (91). Previous work in the Donnenberg laboratory has also provided evidence that the first 171 amino acids of BfpB are periplasmic (93). To further understand and characterize BfpB, a higher resolution structure is required. In this chapter, a combination of molecular biology, bioinformatics, and cryo-electron microscopy were used to determine the structure of the T4bP secretin BfpB and to identify its periplasmic domains.

## Materials and Methods

### Strains, plasmids, and growth conditions.

All strains and plasmids used in this chapter are presented in Table 1. Bacterial strains were cultured in LB broth at 37°C. Antibiotics (ampicillin, 200 µg/mL; kanamycin 50 µg/mL) were added to select for or maintain plasmids. All bacterial cultures were grown at 37°C with agitation at 225 rpm.

Table 1: Strains and plasmids used in chapter II

Strain or Plasmid	Genotype or Description	Reference or Source
XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44</i> <i>relA1 lac</i> [F' <i>proAB lacIqZΔM15 Tn10</i> (TetR)]	Stratagene
BL21 (DE3)	<i>F<sup>-</sup>, dmc, ompT, hsdS(rB-mB-)</i>	Novagen
DH5α	<i>supE44ΔlacU169(φ80dlacZΔM15) hsdR17 recA1 endA1</i> <i>gyrA96 thi-1 relA1</i>	Gibco-BRL
pWS15	<i>bfpB-Strep</i> gene cloned into pASK-IBA3	(93)
pJIL003	<i>bfpB<sub>19-306</sub></i> with C-terminal His tag cloned into pET28a NcoI and XhoI sites	This chapter
pJIL004	<i>bfpB<sub>19-202</sub></i> with C-terminal His tag cloned into pET28a NcoI and XhoI sites	This chapter

Table 2: Primers used in chapter II

Primer Name	Sequence (5'-3')	Reference
JIL-002	GGCACCATGGGATCGGGTAATGGATTTTATAAAGATA ATCTTGGCG	This chapter
JIL-003	CCGTCTCGAGAGTTTCCTCGTTTGAAAAAGCAATCC	This chapter
JIL-005	GGCCTCGAGTCTTTCAAGCTGTGCATTCAGTGTATTAA TATATTCG	This chapter

### Generation of putative periplasmic domains of BfpB

FastCloning (110) was used to truncate BfpB from amino acid positions 19-306 (N0+N3) and 19-202 (N0) in pWS15 to generate pJIL003 and pJIL004, respectively. Primer pairs JIL-002 and JIL-003 generated pJIL004 and pairs JIL-002 and JIL-005 generated pJIL003. The polymerase chain reaction (PCR) products were digested with DpnI and purified. The purified PCR products and pET28a vector were restriction digested with NcoI and XhoI nucleases and then ligated. Ligated products were chemically transformed into *E. coli* DH5 $\alpha$  competent cells. The truncated *bfpB* mutations were confirmed by sequencing and expressed in *E. coli* BL21(DE3) as described below.

### BfpB Protein purification

For purification of BfpB, *E. coli* strain XL1-Blue (pWS15) was grown at 37°C in Luria-Bertani medium to an optical density at 600 nm (OD600) of 0.6 and induced with 0.2  $\mu$ g/mL of anhydrotetracycline at 37°C for 3 hours. Cells were harvested by centrifugation and chemically

lysed in buffer A (50 mM tris, 300 mM NaCl, pH 8.0) supplemented with protease inhibitors (Roche), 5% (w/v) 3-(N,N-Dimethyltetradecylammonio)propanesulfonate (SB3-14), 10 µg/mL DNase I and 100 µg/mL lysozyme. BfpB was purified by affinity chromatography on Strep-Tactin resin (Qiagen) column with buffer A + 0.04% (w/v) SB3-14. Fractions eluted with 10 mM desthiobiotin were analyzed by SDS page, combined, and concentrated with 100K (Amicon) molecular weight cut-off filter. The concentrated sample was further purified with Superose6 10/300 column with buffer A + 0.02% (w/v) SB3-14. Quality of purification and multimers were assessed via negative stain TEM.

For purification of BfpB N0+N3 and BfpB N0, *E. coli* strains BL21 (DE3) (pJIL003) and (pJIL004), respectively, were grown at 37°C in Luria-Bertani medium to an optical density at 600 nm (OD600) of 0.6 and induced with 1 mM of IPTG at 37°C for 2 hours. Cells were harvested by centrifugation and lysed by sonication in lysis buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 10 mM imidazole, pH 8.4) supplemented with protease inhibitors (Roche). BfpB N0+N3 and BfpB N0 were purified by affinity chromatography on nickel nitrilotriacetic acid (Ni-NTA) resin (Qiagen) column. Fractions eluted with 250 mM imidazole were analyzed by SDS-PAGE, combined, and concentrated with 10K (Amicon) molecular weight cut-off filter. Concentrated samples were further purified on a Sephacryl S100 column with PBS pH 7.4 or 20 mM HEPES, 150 mM NaCl, pH 7.4.

### **Cryo-electron microscopy**

Grids (300 mesh UltraAufoil -1.2/1.3 holey gold with 2 nm thin carbon (Quantifoil, Germany)) were glow discharged for 40s at 25 mA. Purified BfpB, 0.03 mg/ml, 4 µl was applied to the grids. Grids were blotted for 2 s with ash-free Whatman® Grade 540 filter paper in a Vitrobot Mark IV (ThermoFisher Scientific) and plunged into liquid ethane. Sample quality and distribution



was assessed on a Glacios electron microscope. Data acquisition was carried out in a Titan Krios transmission electron microscope (ThermoFisher Scientific) operated at 300 kV and counting mode, with a Gatan K3 detector and a 10 eV slit width Gatan Quantum Energy Filter (GIF) at 54,000 magnification. Datasets were collected in automated mode with the program Latitude (Gatan) with cumulative electron dose of  $50 \text{ e}^-/\text{\AA}^2$  applied over 40 frames.

### **Single-particle image processing**

Movies (9,248) collected for the BfpB datasets were processed in cryosparc2.15. Gain-normalization, movie-frame alignment, dose-weighting, and full and local motion correction were carried out with the patch motion correction. Global and local contrast transfer function values were estimated from non-dose weighted motion-corrected images using patch CTF module. Subsequent image processing operations were carried out using dose-weighted, motion corrected micrographs. One thousand manually picked particles were used to auto-pick about 255,000 particles. Extensive 2D classifications of 6 rounds yielded 33,096 pure particles which led to a 7.1 Å map. The reported resolution of the cryo-EM map is based on Fourier Shell Correlation (FSC) of 0.143.

### **Thermal stability assay**

Purified BfpB N0+N3 and BfpB N0 thermal stability were determined via GloMelt Thermal Shift Protein Stability kit (Biotium USA). Samples were prepared according to manufacturer's manual in phosphate buffered saline (PBS).

### *Reaction setup*

A working stock solution of 10X GloMelt dye was made with the 200X dye stock solution (supplied by kit) and PBS used for size exclusion chromatography. Peak size exclusion

chromatography eluate fractions of BfpB N0 and N0+N3 were used as protein stock solutions. Reactions were made in 20  $\mu$ L triplicates with protein final concentrations of 0.5, 1.0 and 1.5 mg/mL and GloMelt dye concentration of 1X. Goat  $\beta$ -actin IgG was used as a positive control with a final concentration of 0.1 mg/mL and 1X dye. The no protein control contained GloMelt dye (1X) in PBS; the blank was PBS. Reagents and solutions were kept in the dark on ice. Reactions were set up in qPCR plates with optical seals.

#### *Real-time PCR thermocycler setup*

Agilent Technologies (USA) Stratagene Mx3005P qPCR system was used. Acquisition was set in the FAM (excitation/emission: 470/510 nm) setting. Initial temperature was set at 25  $^{\circ}$ C for 30 s, with incremental increases of 0.5  $^{\circ}$ C every 30 s up to 95  $^{\circ}$ C.

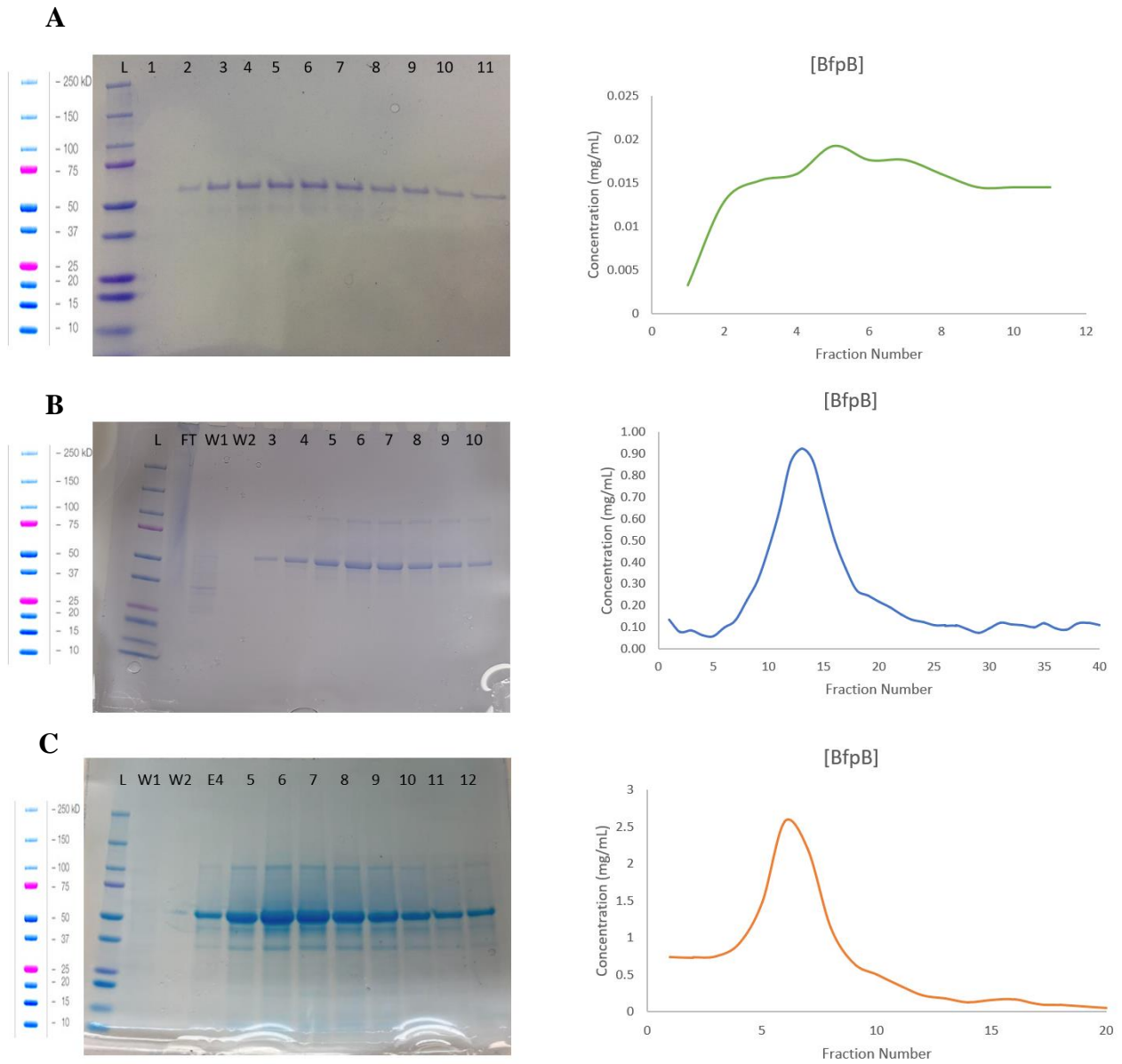
#### *Data analysis*

Data was analyzed via fluorescence vs temperature output. Blank and no protein control were subtracted from protein fluorescence readings to generate the final graph in Microsoft Excel. Melting temperature was calculated from the midpoint of sigmoidal curve of the graph.

## Results

### **Chemical lysis yields higher concentration of BfpB after affinity chromatography purification compared to physical lysis via sonication.**

The previously developed purification protocol for BfpB-Strep (93) involved physical lysing of the cells with a sonicator and solubilization of the membrane pellet with SDS, a chaotropic detergent. Under these conditions, the peak eluate fraction from Strep-Tactin affinity chromatography contained 0.2 mg/mL of BfpB from 2L culture of cells (figure 3A). In attempt to increase the protein yield from a 2L culture, homogenization with a French cell press was used. The yield of BfpB increased to ~0.92 mg/mL in the peak eluate fraction (figure 3B). There were also optimizations made to the column step, which included an increase in column size from 1 mL to 3 mL, and the concentration of desthiobiotin was increased from 5 mM to 10 mM in the elution buffer. Chemical lysis with a zwitterionic detergent, SB3-14 was then attempted to assess the impact on protein yield from a 2L culture. Under these conditions, the peak eluate fraction contained ~2.5 mg/mL of BfpB (figure 3C). This is a 11.5 -fold increase compared to the initial purification protocol; therefore, these conditions were used for subsequent purification by size exclusion chromatography.



*Figure 3: Affinity chromatography of BfpB under various conditions. (A) Purification results from the initial protocol (93) with SDS and sonicator, (B) with French press, and (C) chemical lysing with SB3-14. Top, SDS-PAGE of affinity chromatography. Bottom, elution profiles of affinity chromatography. BfpB is ~60 kDa, possible dimer observed between the 100K and 150K markers.*

### **Concentration of detergent impacts multimer to monomer ratio and protomer aggregation.**

In the previous purification protocol of BfpB-Strep (93), size exclusion chromatography was not implemented. A high concentration of 5% (w/v) SB3-14 was used for chemical lysing while the concentration of SB3-14 used during the affinity chromatography was 0.04%, which is roughly three times its critical micelle concentration. A minimum concentration of detergent is recommended for the samples moving forward for cryo-EM, as detergents often reduce the contrast within the micrographs and increase ice thickness (111). Therefore various concentrations SB3-14 were examined to find the best BfpB multimer to monomer ratio. The following concentrations of SB3-14 were tested: 0, 0.007, 0.014, 0.020, and 0.04%. Elution profiles and negative stain micrographs of each peak were analyzed to determine which condition yielded the most BfpB multimers (high concentration of ring shapes) with the least amount of aggregation (Figure 4 and Table 3). Based on the negative stain micrographs, 0.020% of SB3-14 had the least amount of aggregation and background, therefore cryo-EM was performed with samples from these conditions.

*Table 3: Summary of conditions for size exclusion chromatography.* The effect of detergent concentration used in the mobile phase was assessed for aggregates, background, and multimer concentration. The symbols +, ++, +++ indicate low, moderate, and high, respectively, while the symbol - denotes negligible.

<b>Detergent present</b>	<b>Aggregation</b>	<b>Background</b>	<b>Concentration</b>
0%	+++	-	+++
0.007%	+++	-	++
0.014%	++	+	++
0.02%	++	+	++
0.04%	+	+++	+

Figure 4.

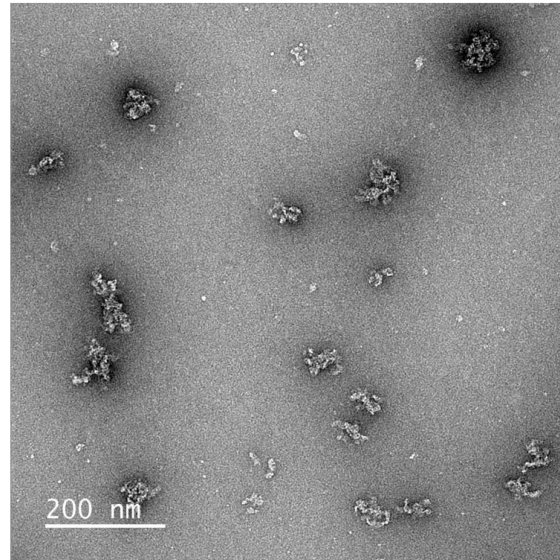
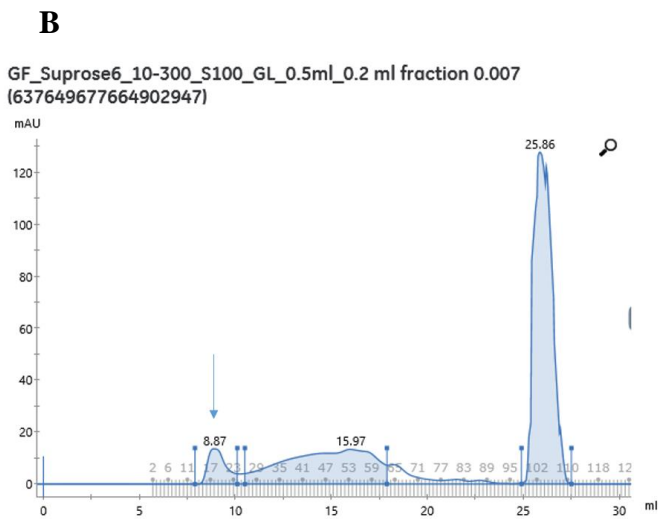
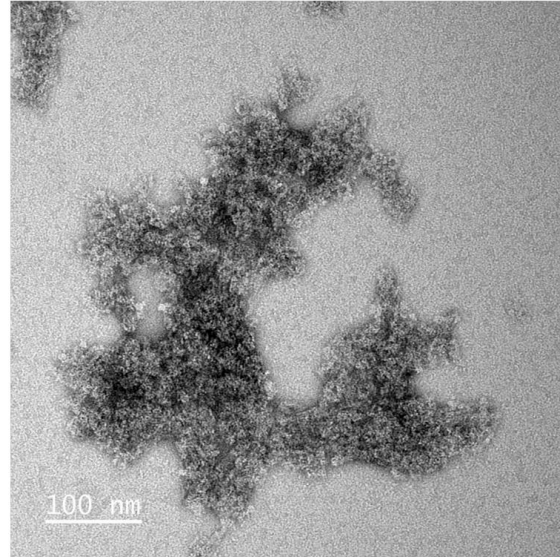
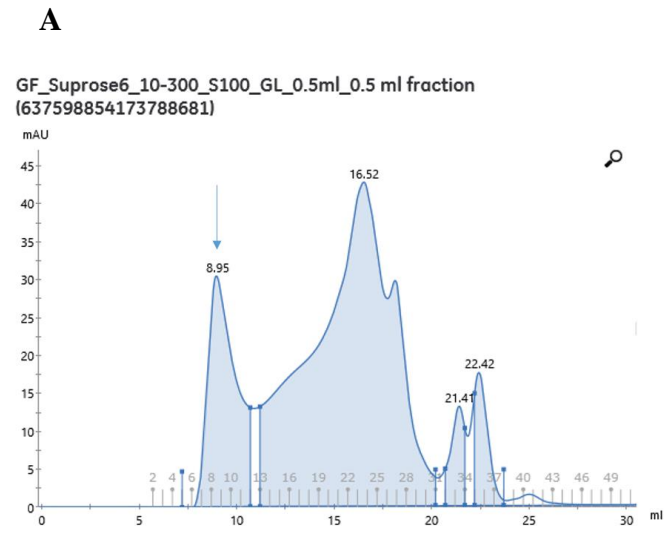
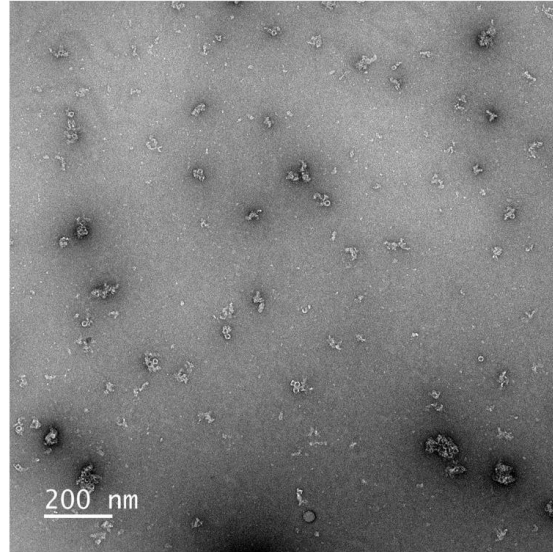
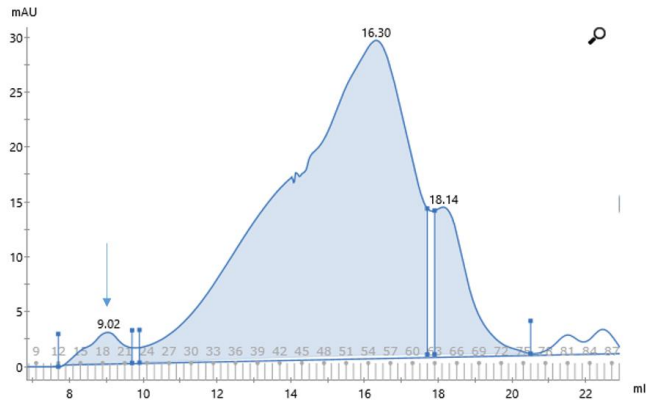


Figure 4, continued.

**C**

GF\_Suprose6\_10-300\_S100\_GL\_0.5ml\_0.2 ml fraction  
(637607522347223323)



**D**

GF\_Suprose6\_10-300\_S100\_GL\_0.5ml\_0.2 ml fraction  
(637626431113310313)

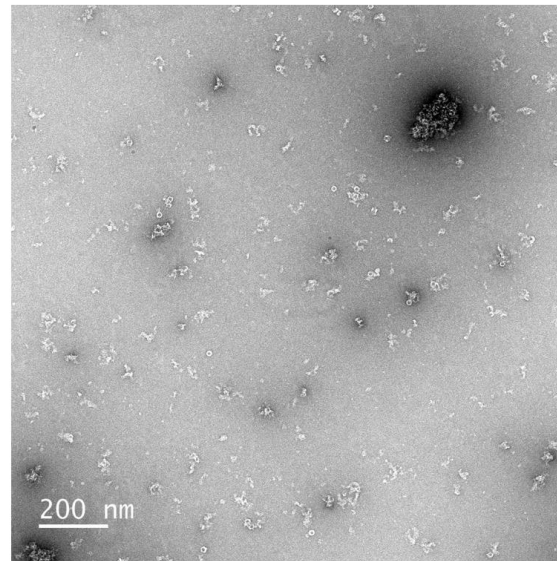
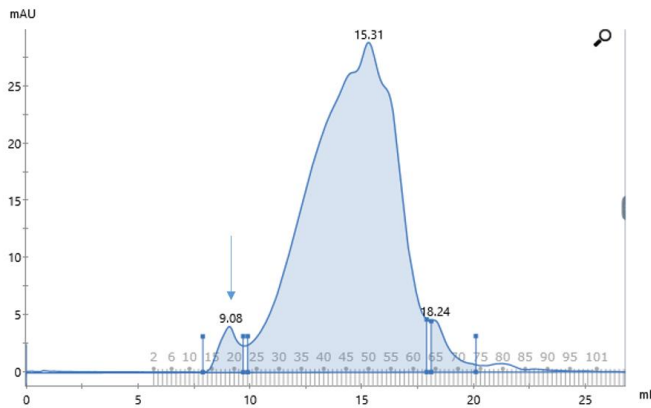




Figure 4, continued.

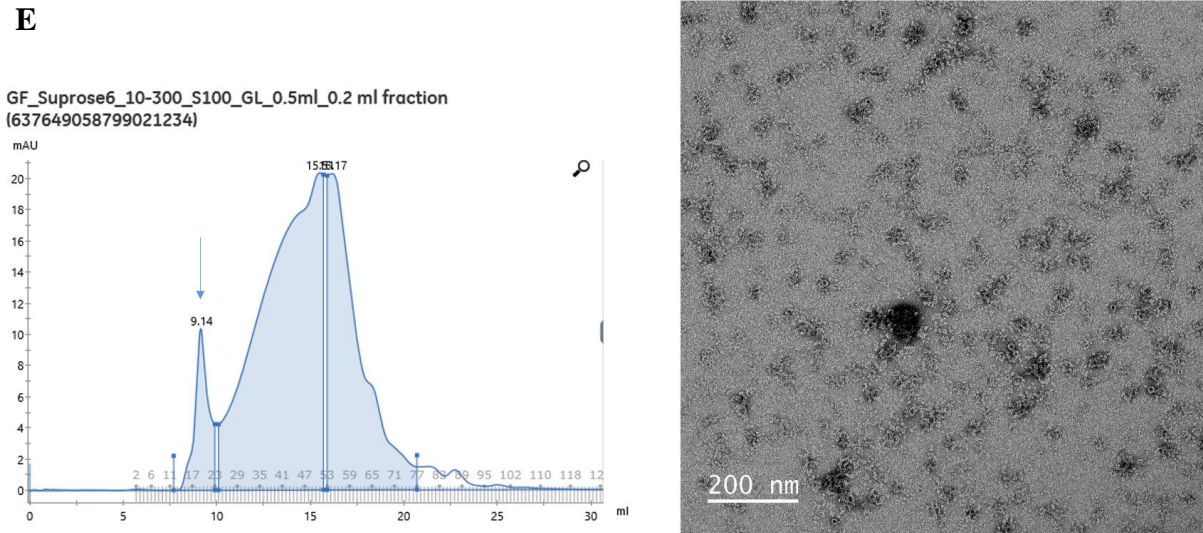


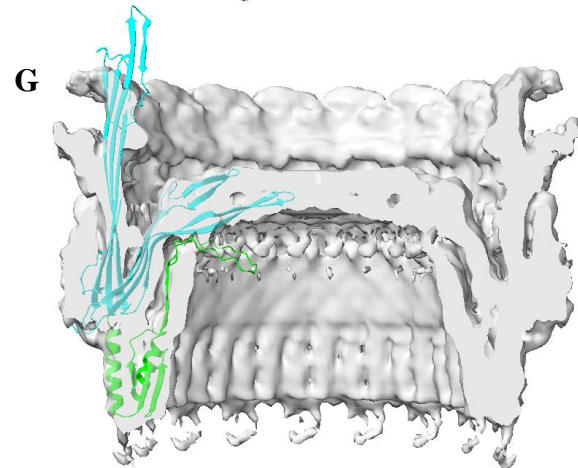
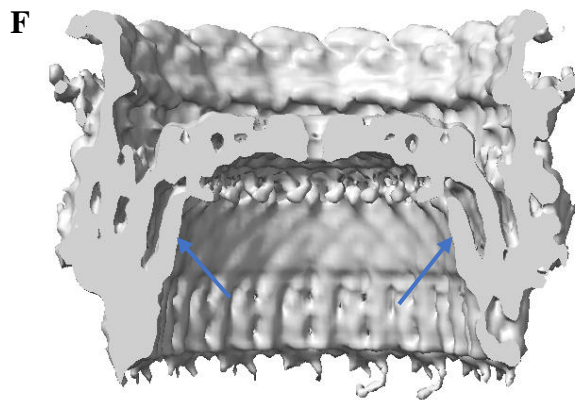
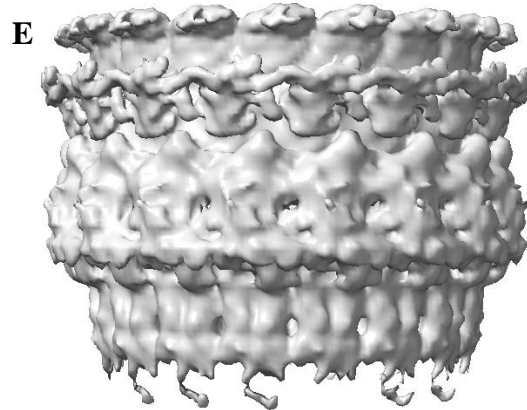
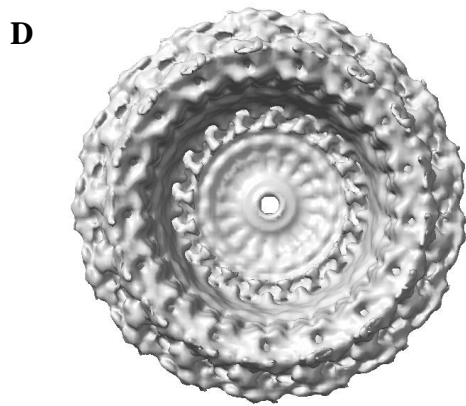
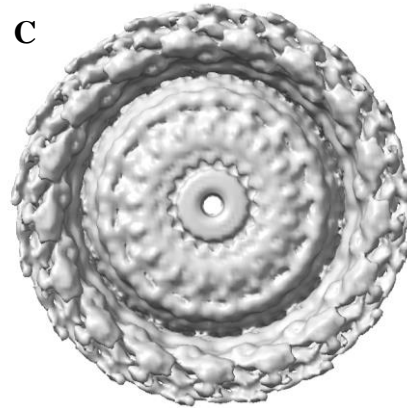
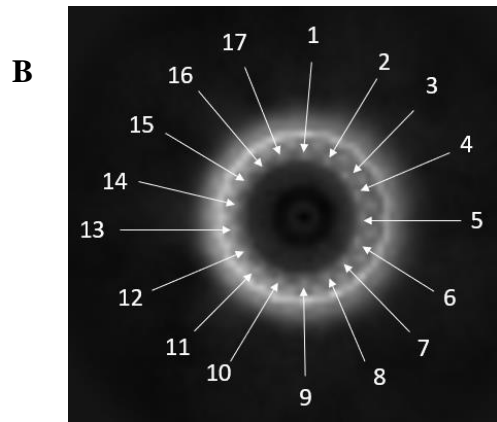
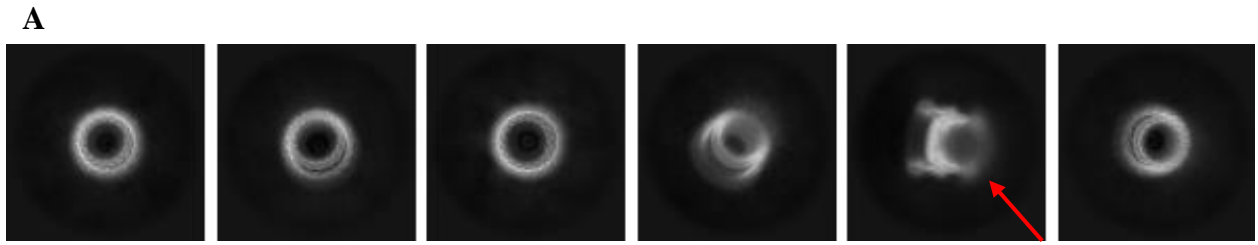
Figure 4: Evaluation of *BfpB* under various concentrations of *SB3-14*. Left column shows the size exclusion chromatography elution profile, right shows negative stained micrograph of peak with toroids (indicated by arrow on elution profile). (A) without *SB3-14*, (B) 0.007%, (C) 0.014%, (D) 0.020%, (E) 0.040%.

## **Cryo-electron microscopy of BfpB suggests it has 17-fold symmetry and N3 $\beta$ -hairpin.**

Figure 5 demonstrates the structure of the purified T4bP secretin BfpB by single-particle cryo-EM. The cryo-EM dataset consisted of 9,248 movies. Extensive rounds of 2D classifications were performed, resulting in the identification of 33,096 pure BfpB particles. The cryo-EM structure seems to have C17 symmetry and an overall resolution of 7.1 Å at FSC of 0.143. The features in the electron density map indicate that there are 3 domains present, which are likely the N0, N3, and secretin domains. The BfpB complex is ~125 Å in height (excluding N0 domain) and ~165 Å in width. One challenge we encountered was that BfpB has a preferential orientation that places the tops or bottoms of the protein on the grids with fewer side views. In addition, the total number of particles that could be used was relatively low. These limitations contributed to a lower resolution than might otherwise have been obtained.

This is the first time a secretin has been reported with symmetry (cyclic) or subunits greater than 15 by cryo-EM. To explain this result, features from cryo-electron density maps of other secretins were analyzed with respect to their function in other T4P, T2S, T3S and filamentous phage systems. The information is summarized in table 4. In the case of T4Ps, only T4aP secretins have been reported to date and these have C12-14 symmetry. Generally, T4aPs have a pilus diameter smaller than T4bPs, ranging from 45-70 Å. Of the T4bPs, the smallest diameter reported was 80 Å, suggesting that a larger secretin with a greater number of subunits is needed for transport through the OM. Due to the resolution, the model was not further refined with the predicted BfpB structure determined by AlphaFold. The experimental density map of BfpB shows a density right below the gate region (Figure 5D). Only one other secretin, InvG (PDB 6PEE) of the T3SS has a similar density (112). The experimental map of InvG and its solved structure shows that the N3

domain has a  $\beta$ -hairpin (extended loop) that fits that density. The predicted structure of BfpB also has an  $\beta$ -hairpin that would fit the experimental density once refined.



*Figure 5: BfpB cryo-EM data and model.* (A) Left, 2D class averages from cryoSPARC. A red arrow highlights a hazy area, likely due to protein flexibility. (B) Zoom of 2D class 3, highlighting the observed C17 symmetry (white arrows). (C-F) 3D reconstruction shown from left to right as top, bottom, side views and side view sliced through the center. Extra density observed right below the gate region (blue arrows). (G) BfpB predicted structure from AlphaFold (Q9S142) modelled into the experimental density map shown as a side view sliced through the center. Secretin domain colored in cyan, N3 domain colored in lime.

*Table 4: Summary of secretin and substrate diameters.* Diameters of secretins are based on the electron density maps from reported cryo-EM data. Outer diameter corresponds to the outer most points of the secretin  $\beta$ -barrel. Inner diameter corresponds to the inner diameter of the  $\beta$ -barrel. Substrate diameter corresponds to the diameter of the following: pilus (T4P), pseudopilus (T2SS), needle (T3SS), and virion (filamentous phage). Reference or PDB code in parentheses. Pa, *P. aeruginosa*; Tt, *T. thermophilus*, Vc, *V. cholerae*, Ec, *E. coli*, Se, *S. enterica*.

<b>System Type</b>	<b>Secretin Name</b>	<b># Subunits</b>	<b>Outer Diameter (Å)</b>	<b>Inner Diameter (Å)</b>	<b>Substrate Diameter (Å)</b>
<i>T4bP</i>	BfpB	17	136	122	85 (1ZWT)
	Vc TcpC	<i>unknown</i>			80 (3HRV)
	Se PilN	<i>unknown</i>			100 (1Q5F)
<i>T4aP</i>	Pa PilQ (6VE3)	14	112	103	60
	Tt PilQ (97)	13	115	100	45-70
	Vc PilQ (6W6M)	14	122	96	60
<i>T2SS</i>	Ec GspD (5WQ7)	15	120	100	70
	Vc GspD (5WQ8)	15	119	103	65 (cholera toxin)

Table 4, continued.

<i>T3SS</i>	Se InvG (6PEE)	15	125	105	70 (6ZNH)
<i>Filamentous Phage</i>	pIV (7OHF)	15	121	107	60 (2C0W)

**BfpB has two putative periplasmic domains.**

Initially, secondary structure prediction tools, such as JPred and Sable were used to identify the possible domains within BfpB. The first 17 residues of BfpB represent a class II signal peptide to be translocated to the OM (91). Residue C18 was excluded from the clones to prevent the lipidation/acylation. Table 5 summarizes the results and the ranges of BfpB constructs that were cloned by colleague Dr. Jinlei Zhao. Clones were confirmed by sequencing and expression analyses showed that the peptides were expressed. Purification under non-denaturing and denaturing conditions were attempted. However, protein was found in the flow through and not eluate fractions. This suggested that these peptides were not well folded or unstable.

Limited proteolysis was attempted with trypsin. BfpB has 42 trypsin cleavage sites. Not all sites should be accessible due to tertiary structures. Purified BfpB was exposed to trypsin (20 µg/mL) for 120 minutes, with samples collected at various time points. Over time, two fragments of BfpB remained resistant to trypsin degradation. MALDI-TOF mass spectrometry was performed to identify these peptides. The results were inconclusive.

*Table 5: Summary of predicted secondary structures of BfpB, terminal residues of BfpB peptide constructs, and molecular weights. Predicted secondary structures denoted by  $\alpha$ -helix (H) or  $\beta$ -sheet (S) and their respective residue ranges in parentheses. All peptides begin with residue S19 and their corresponding molecular weights. This table displays the predicted secondary structure of BfpB, ranging from residues 19-171. The constructs began with residue 19 to exclude the lipid anchor and leader sequence to prevent BfpB from being translocated to the outer membrane. Past residue G171, the predicted secondary structure consisted of majority of beta sheets, indicative of a beta barrel. Terminal residues were selected in regions that did not include a predicted helix or sheet (predicted linker regions).*

<i>Prediction Tool</i>	<i>Secondary Structures</i>				
Jpred	H(98-107)	S(112-115)	H(135-141)	-	S(164-168)
Sable	H(98-107)	S(113-117)	H(131-141)	H(149-158)	S(165-168)
PSIPRED	H(98-109)	S(113-115)	H(132-140)	H(151-157)	S(167-168)
<b>Terminal residue</b>	<b>G110</b>	<b>S123</b>	<b>K146</b>	<b>L164</b>	<b>G171</b>
<b>MW (kDa)</b>	<b>13.25</b>	<b>14.7</b>	<b>16.8</b>	<b>18.9</b>	<b>19.7</b>



Based on the solved structures of secretins, there are three domains that seem to be in common: the N0, N3 and secretin domains. The electron density map of BfpB suggests there are three domains. Additionally, bioinformatic analyses, including SIB Swiss Institute of Bioinformatics MyHits motif scan and AlphaFold suggest that BfpB has three domains, results summarized in figure 6. The range of amino acids that composes each predicted domain were relatively similar, therefore the confidence in this prediction was high. Although the MyHits motif scanner named the region 201-306 “N2”, we believe that this is the N3 domain as there is not a secretin without one and they have the same  $\beta\alpha\beta\alpha$  fold (96, 97, 108, 113, 114). Both prediction tools suggested that the periplasmic domain(s) end at Arg306 and the secretin domain ( $\beta$ -barrel) begins at Thr309. The signal peptide and lipid anchor were excluded from the N0 and N0+N3 constructs. Based on this information, the putative domains were cloned as follows: N0 includes residues S19-T202, N3 includes T202-R306, and N0+N3 ranges from S19-R306. N0, N3, and N0+N3 were successfully cloned and confirmed by sequencing.

The criteria followed to help identify the protein domains are that they form a compact three-dimensional structure and often can be independently stable and folded (115). Expression analyses showed that only N0 and N0+N3 were expressed, which were confirmed by immunoblot with a BfpB peptide antibody. Therefore, purification was attempted from a 2L culture for these constructs. Both N0 and N0+N3 were purified by Ni-NTA affinity chromatography in abundance, with a peak eluate fraction of 16.0 mg/mL and 45.3 mg/mL respectively (Figure 7A-B). The peak fraction from affinity chromatography was subjected to size exclusion chromatography, where the peak fraction of N0 was 1.6 mg/mL and N0+N3 was 16.0 mg/mL (Figure 7C-D). Thermostability of these purified proteins were assessed via GloMelt. The melting temperatures of N0 and N0+N3 were ~50 °C and ~55 °C, respectively (Figure 8). This data implies that BfpB N0 and N0+N3 are

both folded and stable. This information suggests that the putative domains of BfpB were identified as the N0 and N0+N3 constructs met the described criteria of a protein domain.

**A**

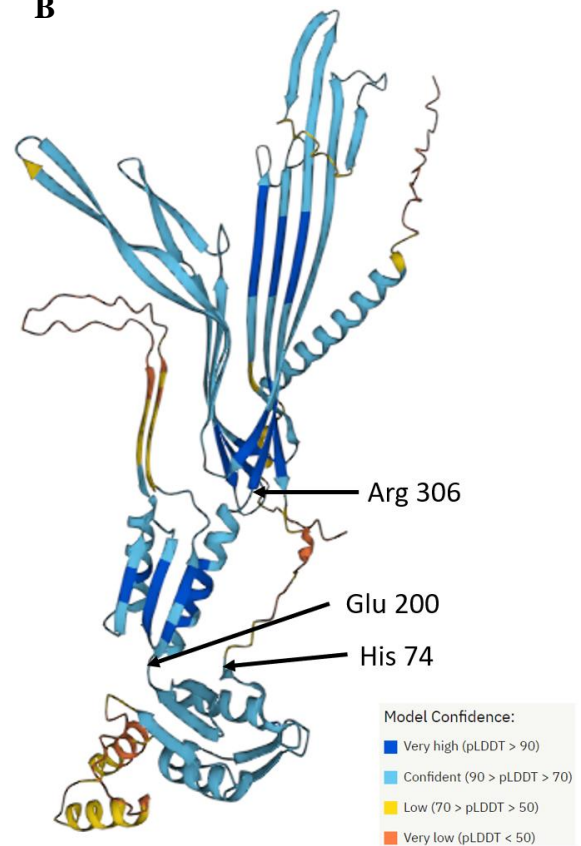
**SIB MyHits Motif Scan Top Results**

Residues 151-198: secretin and TonB N terminus short domain (N0 domain)

Residues 202-279: secretin N-terminal domain N2

Residues 309-546: secretin bacterial type II and III secretion system protein (secretin domain)

**B**



*Figure 6: Results of MyHits motif scan and AlphaFold structure of BfpB (Q9S142). (A) output of top matches from MyHits with high confidence. (B) AlphaFold predicted structure of BfpB. Colors correspond to confidence levels in the model. Arrows point to residues at the beginning of organized, folded regions in the model. AlphaFold produces a per-residue confidence score (pLDDT) between 0 and 100.*

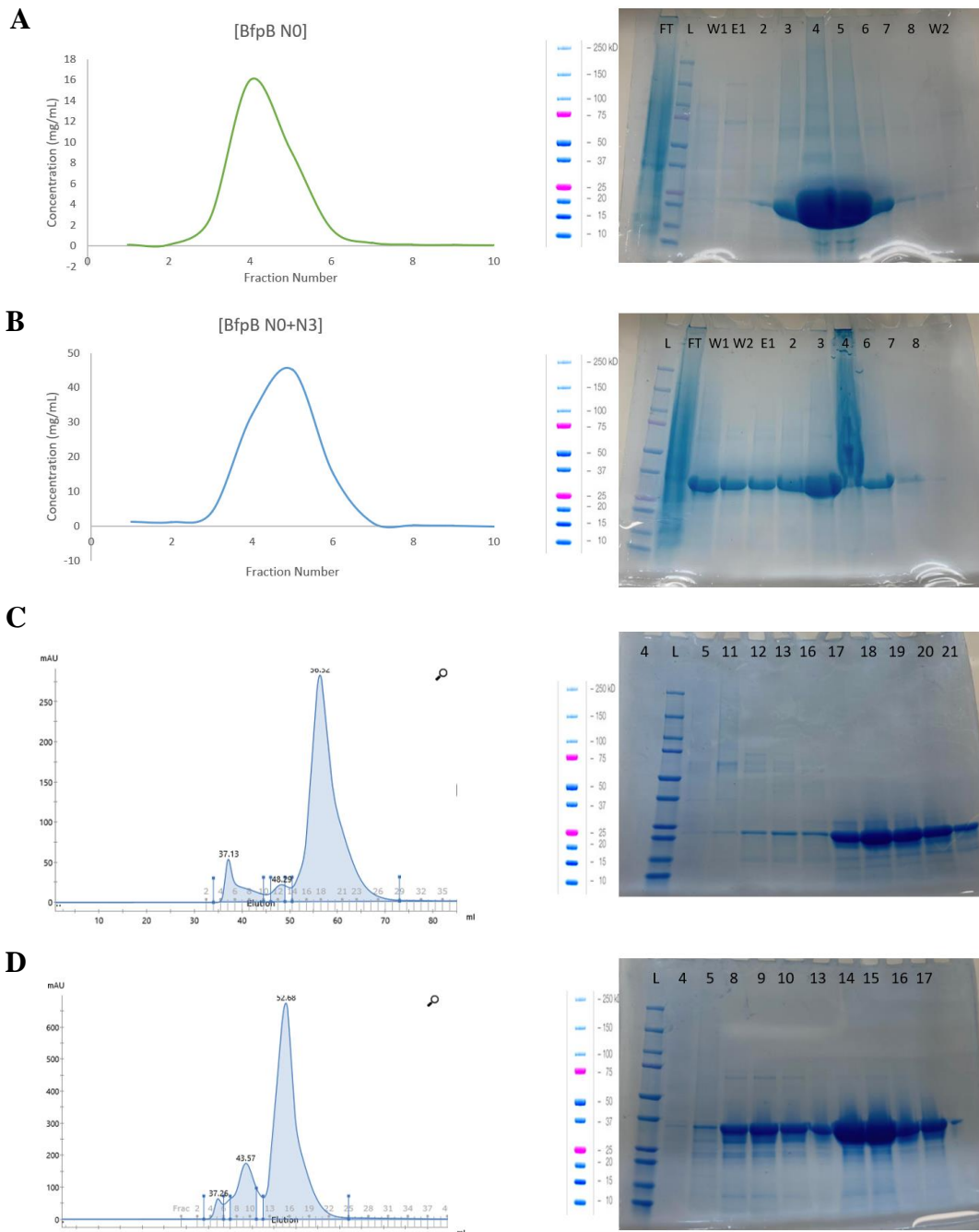


Figure 7: Purification of *BfpB* N0 and N0+N3. Ni-NTA affinity chromatography purification elution profile and SDS-PAGE of (A) *BfpB* N0 (~21 kDa) and (B) *BfpB* N0+N3 (~32 kDa). Size exclusion chromatography elution profile and SDS-PAGE of (C) *BfpB* N0 and (D) N0+N3. W= wash, E = elution.

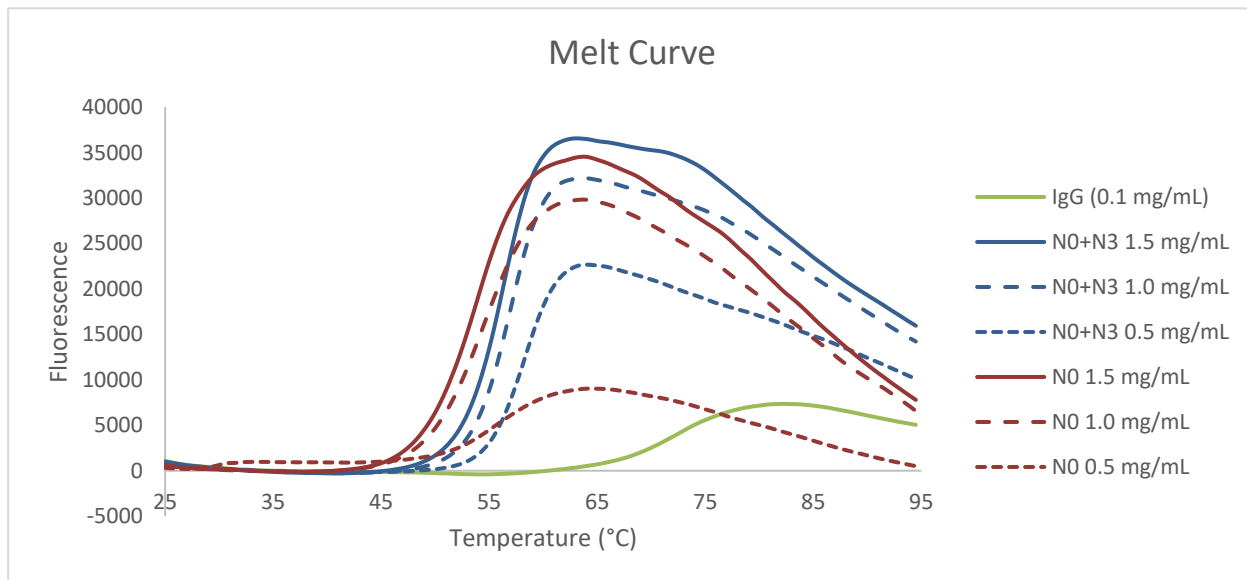


Figure 8: Thermal stability of BfpB N0 and N0+N3. Control IgG has a  $T_M$  of 75 °C while BfpB N0 and N0+N3 have a  $T_M$  of 50 and 55 °C, respectively. Fluorescence is directly proportional to protein unfolding or denaturation.

## Discussion

BfpB is a member of the secretin family of proteins and an essential protein of BFP. In these studies, the previously published purification protocol for BfpB (93) was optimized to avoid using SDS, a detergent known to denature proteins. These modifications yielded more protein compared to our previous protocol and uses a gentler, zwitterionic detergent. Zwitterionic detergents like SB3-14 can disrupt lipid-protein interactions while not disturbing protein-protein interactions (116). The protocol was also improved by including a second purification step, size exclusion chromatography, which helped isolate the BfpB toroids from monomers.

These studies presented the first 3D model of a T4bP secretin by cryo-EM. From this model, BfpB has many similarities with other secretins, but also possesses very distinctive features. Previous work in the Donnenberg laboratory determined the ultrastructure of BfpB at approximately 20 Å resolution, which displayed 12-fold symmetry from Markham rotational analysis (91). From the 2D class average of the current higher resolution images shown in figure 5, 17-fold symmetry was observed, suggesting BfpB has 17 subunits. This was an interesting observation as all secretins described to date range between 12-15 subunits. However, this information is based on the solved secretin structures from T2S, T3S, T4aP, and filamentous phage systems. The diameter of BfpB is wider compared to the other secretins. Generally, T4bPs are wider in diameter compared to T4aPs (table 4) (33). Our model of BfpB, the first T4bP secretin examined to this level of detail by cryo-EM, suggests that this subclass of secretins may have more subunits to allow a larger substrate to pass through. More studies involving the structures of T4bP secretins are needed to confirm this theory.

The electron density map of BfpB suggests there are three domains. The structure resolution was not high enough to confirm the motifs of the domains. However, bioinformatic

analyses also suggest the presence of three domains. Like other secretins, the C-terminus is where the conserved secretin domain resides, forming a gated  $\beta$ -barrel. Next to the secretin domain is the N3 domain, which has been observed in every secretin thus far. The predicted structure of BfpB from AlphaFold suggests that the N3 domain has an extended loop, also known as a  $\beta$ -hairpin. This feature has only been observed in the secretin InvG from the T3SS. The electron density map of BfpB has an interesting density right below the gate region that would fit this  $\beta$ -hairpin structure. Proximal to the suggested N3 domain is the N0 domain, which has also been observed in every secretin. It is often seen as a hazy area in the 2D class averages from cryo-EM of secretins. The N0 domain is often difficult to resolve as it is a flexible part of secretins, hence why it appears hazy (103, 114). While this low-resolution structure is an important first step in understanding the structure of BfpB, further studies are required to determine the exact boundaries of the domains, their folds, and, ideally, amino-acid assignments, so that potential mechanisms underlying its function can be revealed.

## Chapter III: Characterization of the periplasmic protein BfpU

### Introduction

Four of the 14 genes of the *bfp* operon have no known homologs in other T4P systems. One of these four genes is *bfpU*, which encodes a 17 kDa protein that has a class I signal peptide (75). Previous studies from the Donnenberg laboratory showed that BfpU is a soluble protein that is partially localized to the periplasm and essential for BFP biogenesis (117). Additionally, it was shown that BfpU co-localizes and cross-links with the secretin, BfpB and with BfpG, another essential BFP protein localized to the periplasm, and yeast two-hybrid analysis suggests that is a part of the OM assembly of BFP (118). Aside from this information, BfpU is not well characterized, and its function is still unknown.

While BfpU does not share any sequence homology with any other T4P biogenesis protein in other systems, it has similar properties to the protein PilP. PilP is a lipoprotein anchored in the IM (though primarily periplasmic) and has a similar structure to GspC, a component of the T2SS that has been shown to bind to its corresponding secretin, GspD. Previous studies in *P. aeruginosa* revealed that PilP interacts with the OM secretin PilQ (119, 120). In *M. xanthus*, an outside-in assembly pathway was observed with PilQ in the OM recruiting a subcomplex consisting of PilNOP, through the interaction between PilP and PilQ (121). These studies concluded that PilQ is required for the stability of PilMNOP complex and implies that the assembly of the IM subassembly complex requires the presence of PilQ and PilP. Furthermore, the lipidation site of PilP is not essential for its localization or function (122), suggesting that PilP might function in the periplasm in a similar manner as BfpU. In this chapter, a combination of molecular biology, biophysics, and bioinformatics were used to further characterize BfpU.



## Materials and Methods

### *E. coli* strains, plasmids and growth conditions.

All strains and plasmids used in this chapter are presented in Table 6. Bacterial strains were cultured in LB broth, Dulbecco's Modified Eagle Medium (DMEM), or M9 minimal media at 37°C. Ampicillin (200 µg/mL) was added to select for or maintain plasmids. All bacterial cultures were grown at 37°C with agitation at 225 rpm.

Table 6: Strains and plasmids used in chapter III

Strain or Plasmid	Genotype or Description	Reference or Source
BL21 AI <i>slyD</i> cm	<i>F<sup>-</sup> ompT hsdS<sub>B</sub> (r<sub>B</sub><sup>-</sup>m<sub>B</sub><sup>-</sup>) gal dcm araB::T7RNAP-tetA ΔslyD::cat</i>	(118)
SHuffle T7	<i>F<sup>+</sup> lac, pro, lacI<sup>q</sup> / Δ(ara-leu)7697 araD139 fhuA2 lacZ::T7 gene1 Δ(phoA)PvuII phoR ahpC* galE (or U) galK λatt::pNEB3-r1-cDsbC (Spec<sup>R</sup>, lacI<sup>q</sup>) ΔtrxB rpsL150(Str<sup>R</sup>) Δgor Δ(malF)3</i>	New England Biolabs
SHuffle T7 Express	<i>fhuA2 lacZ::T7 gene1 [lon] ompT ahpC gal λatt::pNEB3-r1-cDsbC (Spec<sup>R</sup>, lacI<sup>q</sup>) ΔtrxB sulA11 R(mcr-73::miniTn10--Tet<sup>S</sup>)2 [dcm] R(zgb-210::Tn10 --Tet<sup>S</sup>) endA1 Δgor Δ(mcrC-mrr)114::IS10</i>	New England Biolabs
DH5α	<i>supE44ΔlacU169(φ80dlacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i>	Gibco-BRL

EPEC E2348/69	Serotype O127 : H6 EPEC strain isolated from an outbreak in the UK	(123)
NH4 (EPEC)	E2348/69 $\Delta$ <i>hsrD</i>	(124)
UMD922 (EPEC)	E2348/69 <i>bfpU1::aphA-3</i> EPEC $\Delta$ <i>bfpU</i>	(118)
pAD07	Encodes C-terminal histidine tagged BfpU. Cloned into NcoI/SalI sites of pBAD24	(118)
pVAD111	Encodes BfpU without tag for native expression. Cloned into BamHI/HindIII sites of pTrc99A	This study
pTrc99A	Low copy <i>trc</i> promoter expression vector carrying the <i>lacI<sup>q</sup></i> gene	(125)
pJIL006	Encodes BfpU N89Y with C-terminal His tag generated from FastCloning with pAD07	This study
pJIL007	Encodes BfpU N125S with C-terminal His tag generated from FastCloning with pAD07	This study

Table 7: Primers used in chapter III

Primer Name	Sequence (5'-3')	Reference
BfpUN89Y.f	CATATGATACATATGAAAACACGATATACTTAATG	This study
BfpUN89Y.r	CGTGTTTTTCATATGTATCATATGAGATAAATGATCC	This study
BfpUN125S.f	CGATATCATTAGCGAACATTCTGTGGTGAGTGAC	This study
BfpUN125S.r	CAGAATGTTTCGCTAATGATATCGCCCTTACAG	This study

## **Random Mutagenesis via error prone PCR**

The plasmid used as template is pVAD111, which is *bfpU* cloned into pTrc99A. PCR was performed with New England Biolabs Taq polymerase in the presence of 0.15 mM MnCl<sub>2</sub> to decrease polymerase fidelity. Protocol was followed according to manual. Mutant *bfpU* PCR products were cloned into pTrc99A. Ligated products were chemically transformed into *E. coli* DH5 $\alpha$  strain, then electroporated into EPEC NH4 strain, then lastly electroporated into EPEC UMD922  $\Delta$ *bfpU*. The resulting strain and plasmids were named as follows, EPEC  $\Delta$ *bfpU* (pVAD111.x (x=101-300)) to express mutant BfpU.

## **Generation of BfpU N89Y and N125S**

FastCloning (110) was used to induce substitutions for asparagine codons at amino acid positions 89 and 125 of BfpU in pAD07. Primer pairs BfpUN89Y.f/r generated BfpU N89Y in pJIL006 and pairs BfpUN125S.f/r generated BfpU N125S in pJIL007. The PCR products were digested with DpnI and purified. The purified PCR products were chemically transformed into *E. coli* DH5 $\alpha$  competent cells. The BfpU mutants were confirmed by sequencing, electroporated into *E. coli* SHuffle T7, and expressed as described below.

## **Auto-aggregation Assay**

EPEC  $\Delta$ *bfpU* (pVAD111.x) and controls cultures were grown in DMEM supplemented with 1mM IPTG for 4 hours then analyzed via auto-aggregation (AA) assay for aggregates for 3 hours using Cytation 5 Cell Imaging plate reader. Cultures were triplicated in flat-bottom 48-well plates. Plasmids were isolated from confirmed AA negative strains plasmids and sequenced to confirm mutations.

## **Protein purification**

For purification of BfpU (pAD07), BfpU N89Y (pJIL006), and BfpU N125S (pJIL007), *E. coli* SHuffle T7 strains were grown at 37°C in Luria-Bertani medium to an optical density at 600 nm (OD<sub>600</sub>) of 0.6 and induced with arabinose (final concentration of 0.2%) at 37°C for 2 hours. Cells were harvested by centrifugation and lysed by sonication in lysis buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 10 mM imidazole, pH 8.4) supplemented with protease inhibitor (Roche). BfpU was purified by affinity chromatography on Ni-NTA resin (Qiagen) column. Fractions eluted with 250 mM imidazole were analyzed by SDS-PAGE, combined, and concentrated with 10K (Amicon) molecular weight cut-off filter. The concentrated sample was further purified on Sephacryl S100 column with PBS pH 7.4 or 20 mM HEPES, 150 mM NaCl, pH 7.4.

## **Thermal stability assay**

Purified BfpU thermal stability was determined via GloMelt Thermal Shift Protein Stability kit (Biotium USA). Samples were prepared according to manufacturer's manual in PBS.

### *Reaction setup*

A working stock solution of 10X GloMelt dye was made with the 200X dye stock solution (supplied by kit) and PBS used for size exclusion chromatography. Peak size exclusion chromatography eluate fraction of BfpU was used as the protein stock solution. Reactions were made in 20 µL triplicates with protein final concentrations of 0.5, 1.0 and 1.5 mg/mL and GloMelt dye concentration of 1X. Goat β-actin IgG was used as a positive control with a final concentration of 1.0 mg/mL and 1X dye. The no protein control was GloMelt dye (1X) in PBS and the blank

was PBS. Reagents and solutions were kept in the dark on ice. Reactions were set up in qPCR plates with optical seals.

#### *Real-time PCR thermocycler setup*

Agilent Technologies (USA) Stratagene Mx3005P qPCR system was used. Acquisition was set in the FAM (excitation/emission: 470/510 nm) setting. Initial temperature was set at 25 °C for 30 s, with incremental increases of 0.5 °C every 30 s up to 95 °C.

#### *Data analysis*

Data was analyzed via fluorescence vs temperature output. Blank and no protein control were subtracted from protein fluorescence readings to generate the final graph in Microsoft Excel. Melting temperature was calculated from the midpoint of sigmoidal curve of the graph.

### **Circular Dichroism**

The secondary structure of purified BfpU-His (0.5 mg/mL) in PBS was assessed using an Olis DSM 20 circular dichroism spectrophotometer and a 2 mm pathlength cylindrical cell (Hellma, #121.000-QS) was used at 20 °C. Sample was analyzed using a wavelength range of 200-250 nm. The blank, PBS, was subtracted from BfpU readings. Data were analyzed via BeStSel (Beta Structure Selection) online software.

### **Immunoblotting**

SDS-PAGE gels were run according to the manufacturer's instructions (Bio-Rad, USA) and transferred at 25 V for 3 min using a Trans-Blot® Turbo™ Transfer System (Bio-Rad, USA) to polyvinylidene fluoride membrane (Bio-Rad, USA) and blocked for 1 hour at room temperature in 5% milk/TBS-Tween. The blots were then incubated overnight at 4 °C with monoclonal mouse-

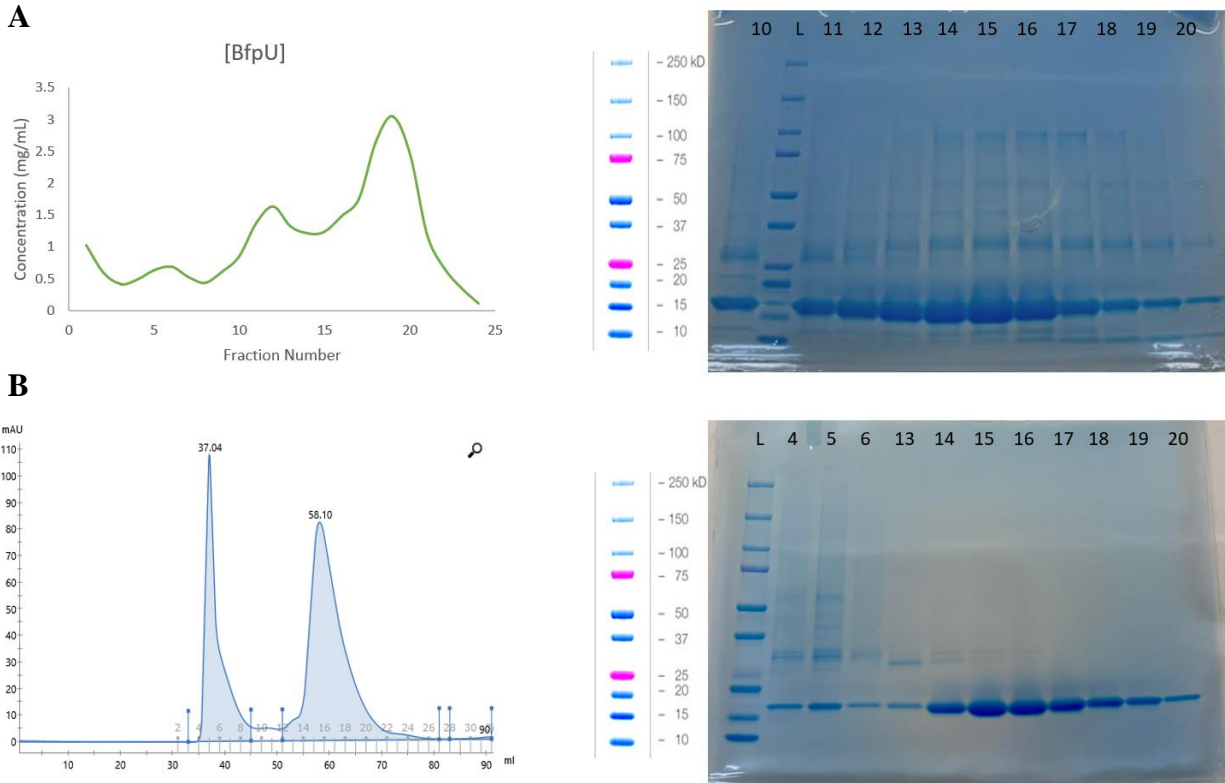
anti-BfpU (1:15,000) in 5% milk/TBS-Tween, washed 3 times for 5 min in TBS-Tween and incubated for 1 hour at room temperature with IRDye (680 nm nm)-conjugated anti-mouse secondary antibodies (Li-Cor Biosciences, Lincoln, NE). The blots were washed again 3 times for 5 min in TBS-Tween at room temperature and scanned with an Odyssey Western system (Li-Cor Biosciences).

## Results

### **Purification of BfpU reveals that a disulfide bond may be present.**

The previously developed expression and purification protocol for BfpU-His (118), involved overexpression in *E. coli* BL21 AI *slyD* strain. This strain uses a phage T7 polymerase that promotes tight regulation and strong overexpression of genes, and has *slyD* deleted, which encodes a histidine-rich protein that is a common contaminant of Ni-NTA affinity chromatography (126). Under these conditions, BfpU from 2L culture of cells would yield an abundance of protein that would precipitate after a few hours after affinity chromatography stored in 4 °C. The buffers used for the lysis and chromatography steps had salt and pH conditions that were ideally favorable for BfpU based on its isoelectric point of 6.6.

To further understand the reason why BfpU precipitated, we analyzed its primary structure. Upon closer look at the amino acid sequence, BfpU possibly has a disulfide bond between residues Cys 107 and Cys 119. To ensure proper protein folding and disulfide bond formation, *E. coli* SHuffle T7 strain, which facilitates accurate disulfide bond formation in the cytoplasm due to cytoplasmic expression of the oxidoreductase DsbC (127, 128), was used to over express BfpU. The assistance of DsbC is likely to reduce protein precipitation because proper disulfide bonds are essential for protein folding. Affinity chromatography was followed by size exclusion chromatography to remove contaminating proteins and buffer exchange. Purity was assessed by SDS-PAGE Coomassie staining (figure 9).



*Figure 9: Purification of BfpU. (A) Ni-NTA affinity chromatography purification elution profile and SDS-PAGE (B) Size exclusion chromatography elution profile and SDS-PAGE. BfpU MW ~17 kDa. FT = flow through/cell lysate, W = wash, E = elution fractions.*



### **Purified BfpU is stable and possesses secondary structure.**

A thermostability assay was performed with the purified BfpU to ensure that it can be used for downstream applications, such as the interaction studies with BfpB. The control,  $\beta$ -actin IgG, is a well folded, stable protein with a denaturation temperature of  $\sim 75$  °C. The various concentrations of BfpU suggested a denaturation temperature of  $\sim 80$  °C (figure 10). This is very stable, as many *E. coli* proteins begin to destabilize around 41 °C (129). To ensure that BfpU is in fact a well folded protein, circular dichroism (CD) was used to assess the secondary structures. The CD spectrum suggests that BfpU is primarily anti-parallel  $\beta$ -sheet (38%) and random coil (45%) (figure 11). Other features, such as turn and  $\alpha$ -helix, were in lower percentage. Though mostly random coil, BfpU is believed to be well folded because it has a high melting temperature and possibly has similar features to the crystal structure of PilP from *P. aeruginosa* (PDB 2LC4), which is mostly anti-parallel  $\beta$ -sheet and random coil. BfpU is possibly a structural homolog of PilP based on this data. The recent AlphaFold predicted structure of BfpU (figure 12) has a root-mean-square deviations of 0.786 Å (32 pruned pairs) and 5.382 Å (all pairs) when superimposed with PilP. The pruned atom pairs refer to the subset of aligned atoms that are used in the calculation of the RMSD, and these are typically selected to exclude regions of the structures that may be structurally variable or disordered, while all pairs include the pruned pairs as well as all other aligned atoms. Although the predicted structure shows a long  $\alpha$ -helix spanning residues L44-K79, the CD data indicates that BfpU is only approximately 3%  $\alpha$ -helical. Therefore, it is possible that this helix is shorter than predicted by AlphaFold and may consist of residues that occupy a different secondary structure in BfpU.

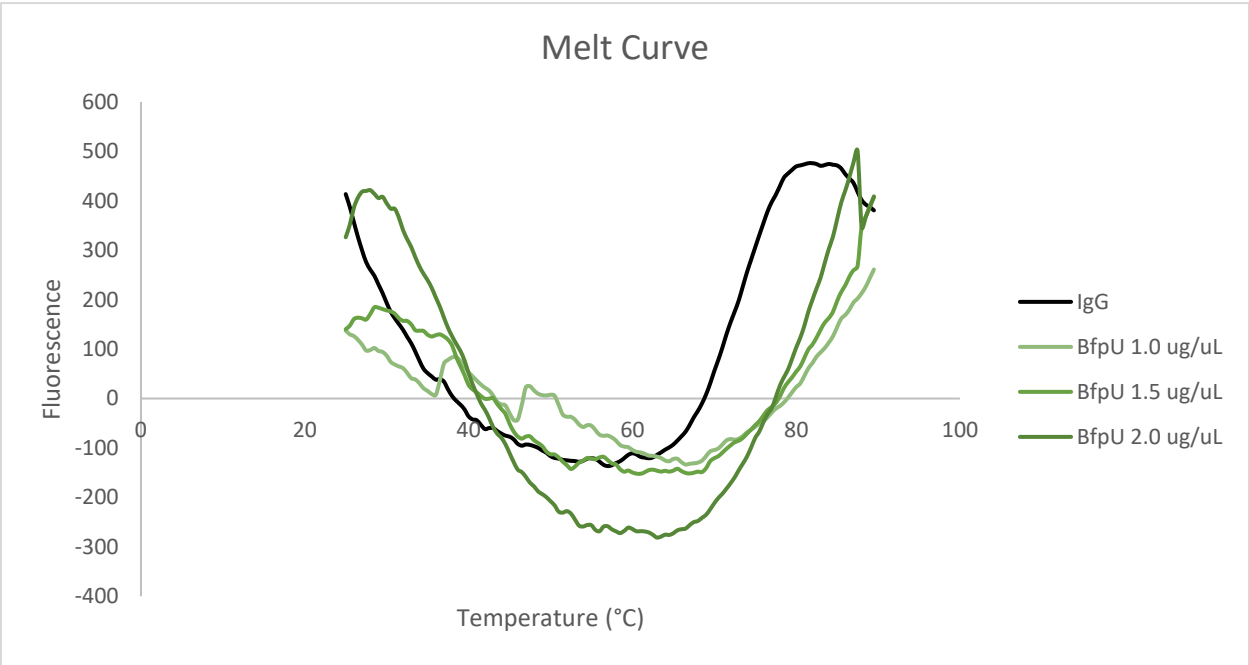


Figure 10: Thermal stability of BfpU. Control IgG has a  $T_M$  of ~75 °C while BfpU has a  $T_M$  of 80 °C. Fluorescence is directly proportional to protein unfolding or denaturation.

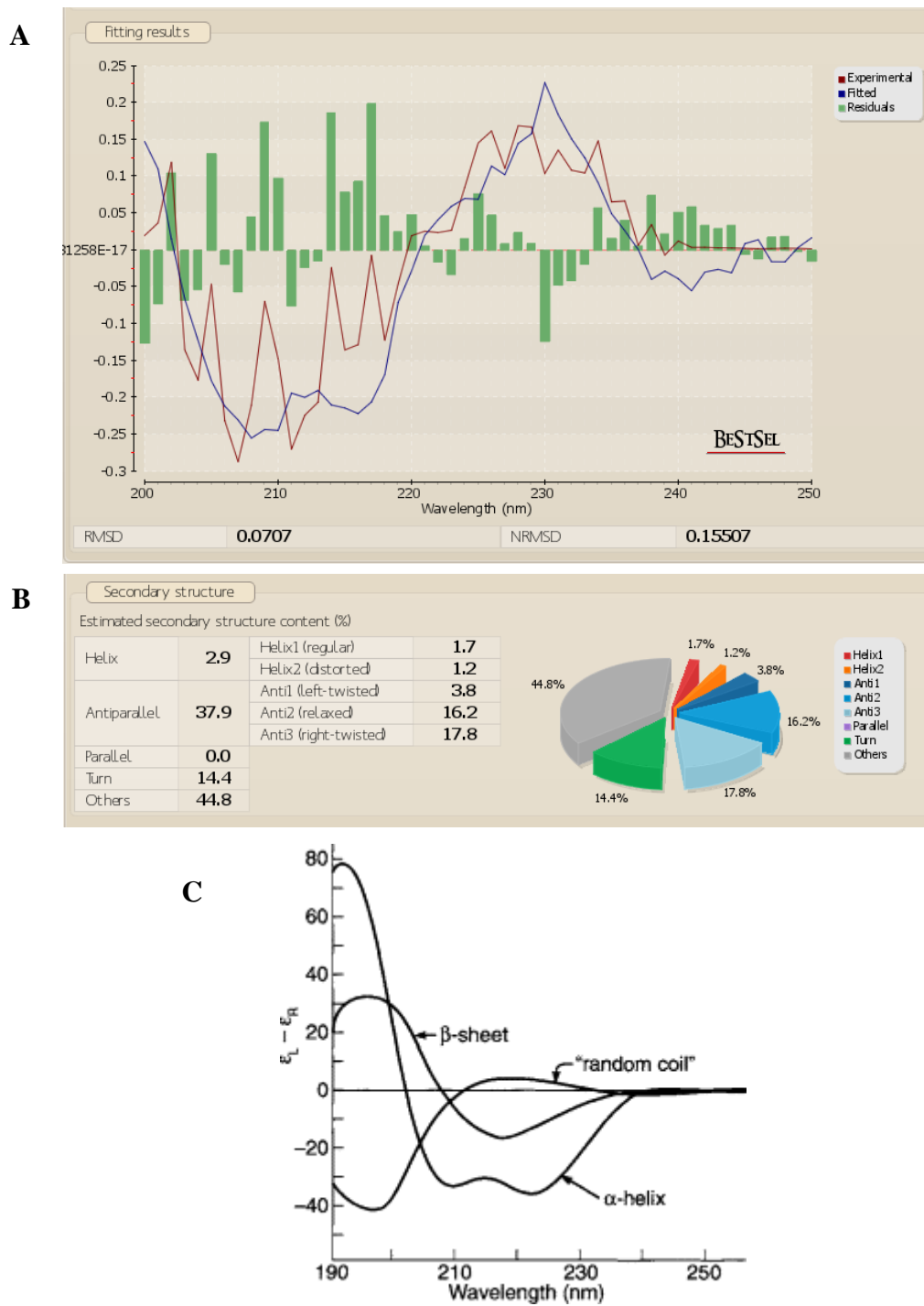
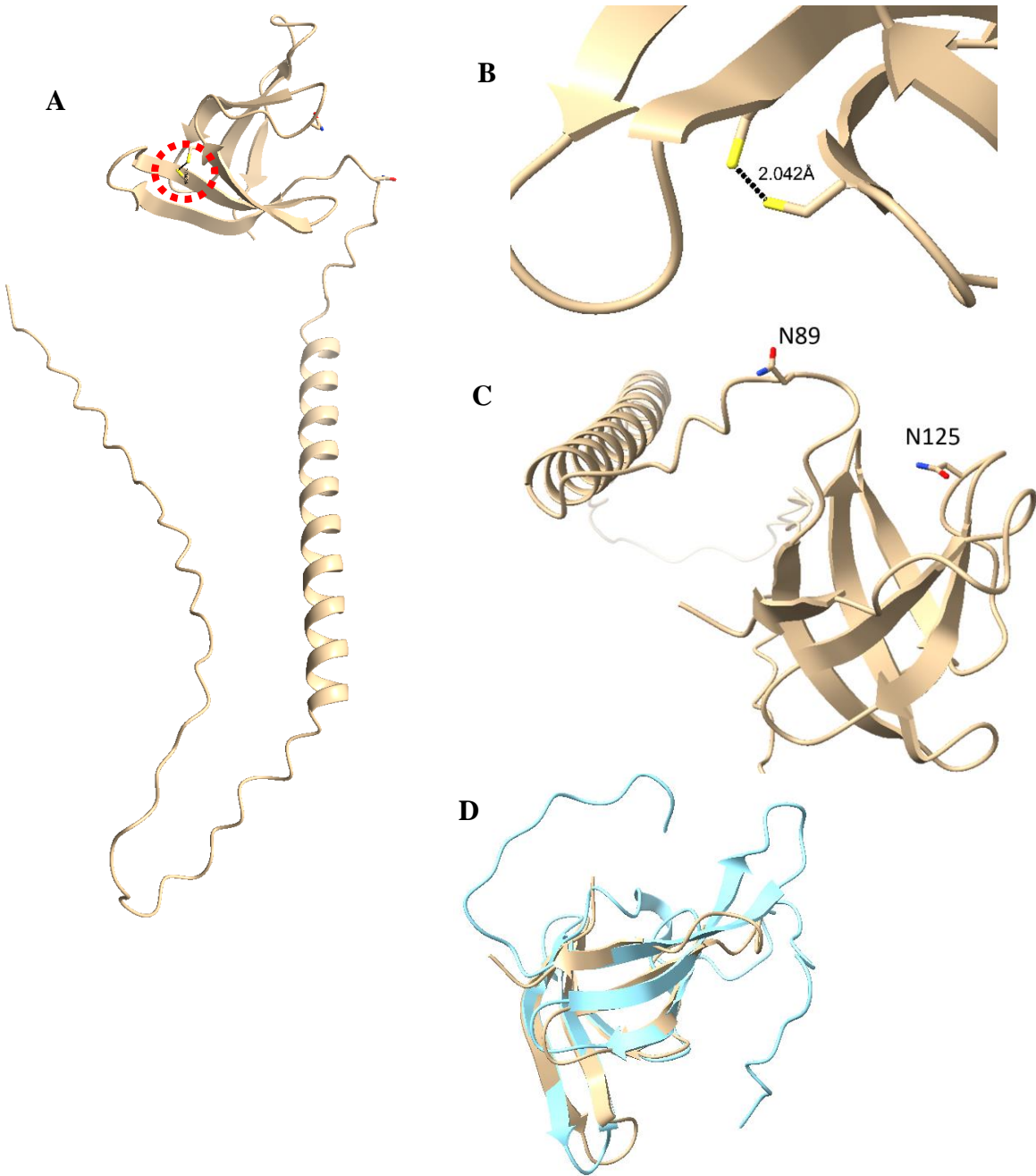


Figure 11: Circular dichroism of BfpU. (A) The spectrum of BfpU. (B) Analysis of secondary structure in percentage by Beta Structure Selection (BeStSel). (C) CD spectrum of “pure” secondary structures for comparison to (A) (130).



*Figure 12: AlphaFold predicted structure of BfpU (Q47069). (A) BfpU full length structure. Site of cysteine residues encircled in red. (B) Zoom of cysteine residues site and distance between C107 and C119. (C) Sites of AA negative mutations N89 and N125S. (D) PiIP (PDB 2LC4, cyan) superimposed with BfpU<sub>92-164</sub>. The residues involved with the β-sheets for both proteins were superimposed because that is what was resolved from the PiIP NMR structure.*

### **Preliminary screen of BfpU mutants highlight significant residues.**

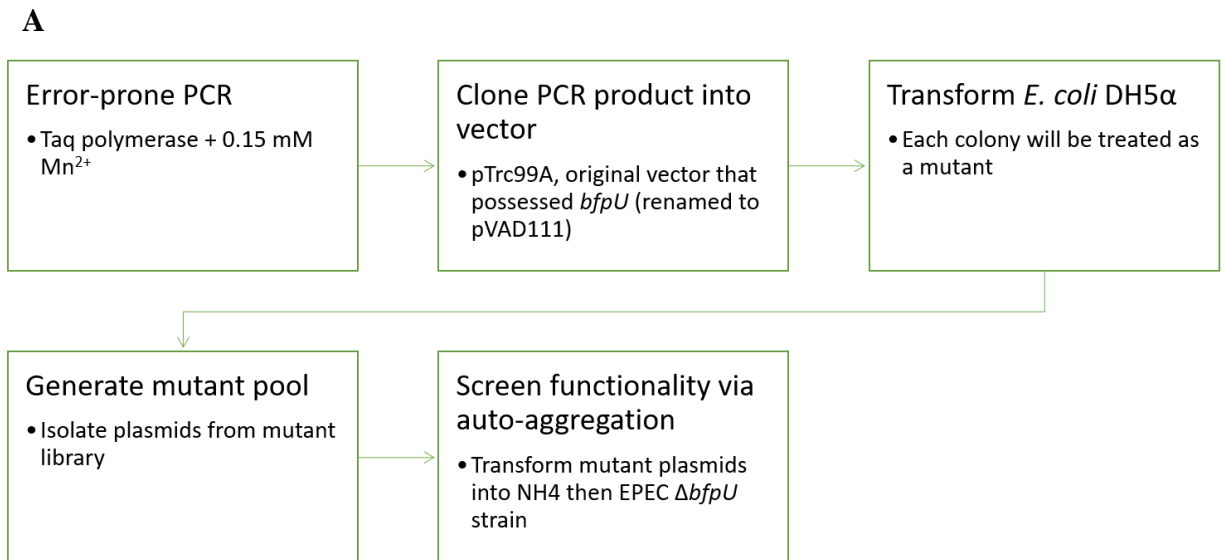
To further characterize BfpU, single random mutations were generated via error-prone PCR and assessed for functionality. EPEC has a specific phenotype based on functional BFPs, known as auto-aggregation (AA). Dense auto-aggregates containing hundreds or thousands of individual bacteria are readily visible by light microscopy. A high-throughput screen using the Cytation5 imager was developed based on the observation of this phenotype in EPEC  $\Delta bfpU$  complemented with mutant *bfpU*. The workflow for generating mutants is summarized in figure 13A. About 200 *bfpU* mutants were generated. Twelve mutants were randomly selected for a preliminary screen, sequenced on both DNA strands, screened for AA, and analyzed for *bfpU* expression. Wild-type EPEC and EPEC  $\Delta bfpU$  complemented with wild-type *bfpU* in pVAD111 were used for positive controls (AA positive). EPEC  $\Delta bfpU$  and EPEC  $\Delta bfpU$  complemented with pTrc99A (empty vector used to generate pVAD111) were used for negative controls (AA negative).

Of the twelve mutants, seven were confirmed to have single mutations (four did not prime well, suggesting the plasmid with *bfpU* was not present, and one had multiple mutations), which is summarized in figure 13B. The confirmed mutations are summarized in table 8. Of the eight mutants, three were AA negative 3/3 times, three were AA positive 3/3 times, and one was AA negative 2/3 times that they were screened in the assay. All of the mutants were then assessed via immunoblot to determine whether they expressed BfpU (figure 13D). The results of these analyses are summarized in table 8. There are a few interesting observations from these results. The first is that two lysine residues (K46 and K72) were mutated to arginine, a residue of similar properties, yet BfpU K72R was not expressed. Also, the mutations that caused an AA negative phenotype are

generally in the same region within the primary structure of BfpU. Though preliminary, these results suggest this region may be important for either the structure or binding sites of BfpU.

Mutants that were not expressed would generate an AA negative phenotype because BfpU, an essential protein, is not present to make functional BFP. Therefore, the mutations of particular interest are the ones that were expressed and AA negative, as these mutations would highlight residues of structural significance or related to interaction sites. Of the preliminary screen, mutations N89Y and N125S had these characteristics. The mutations from the screen were generated in pVAD111 which has the signal peptide and no purification tag. Thus, FastCloning (110) was used to generate a single mutation in pAD07 which expresses BfpU-His. The mutations were generated using the exact codons confirmed in the pVAD111 mutants. Both BfpU N89Y and N125S were expressed in *E. coli* SHuffle T7 and purified by Ni-NTA affinity chromatography. The peak eluate fraction of N125S was 2.35 mg/mL (figure 14B), which is comparable to the wild-type BfpU affinity chromatography peak of 3.0 mg/mL. BfpU N89Y purification resulted in a peak fraction of 0.90 mg/mL (figure 14A), which is significantly lower compared to wild-type and N125S. This decrease in expression may infer that that BfpU N89Y may have structural related significance.

Figure 13



**B**

MKKILFTFLFVVSSVNAEVNTSVDVSGKNDSRKEENKLPDVDLEKYKDELLSY  
QELQRKNMMLKLQSENQKLESEIKKSGSFISYDTNENTIYLM SIMTTPKKGLC  
ATIFNGAISRVCKGDII NEHSVVSDITGNYVEIRSINNISDPVQKIFLR

Figure 13, continued.

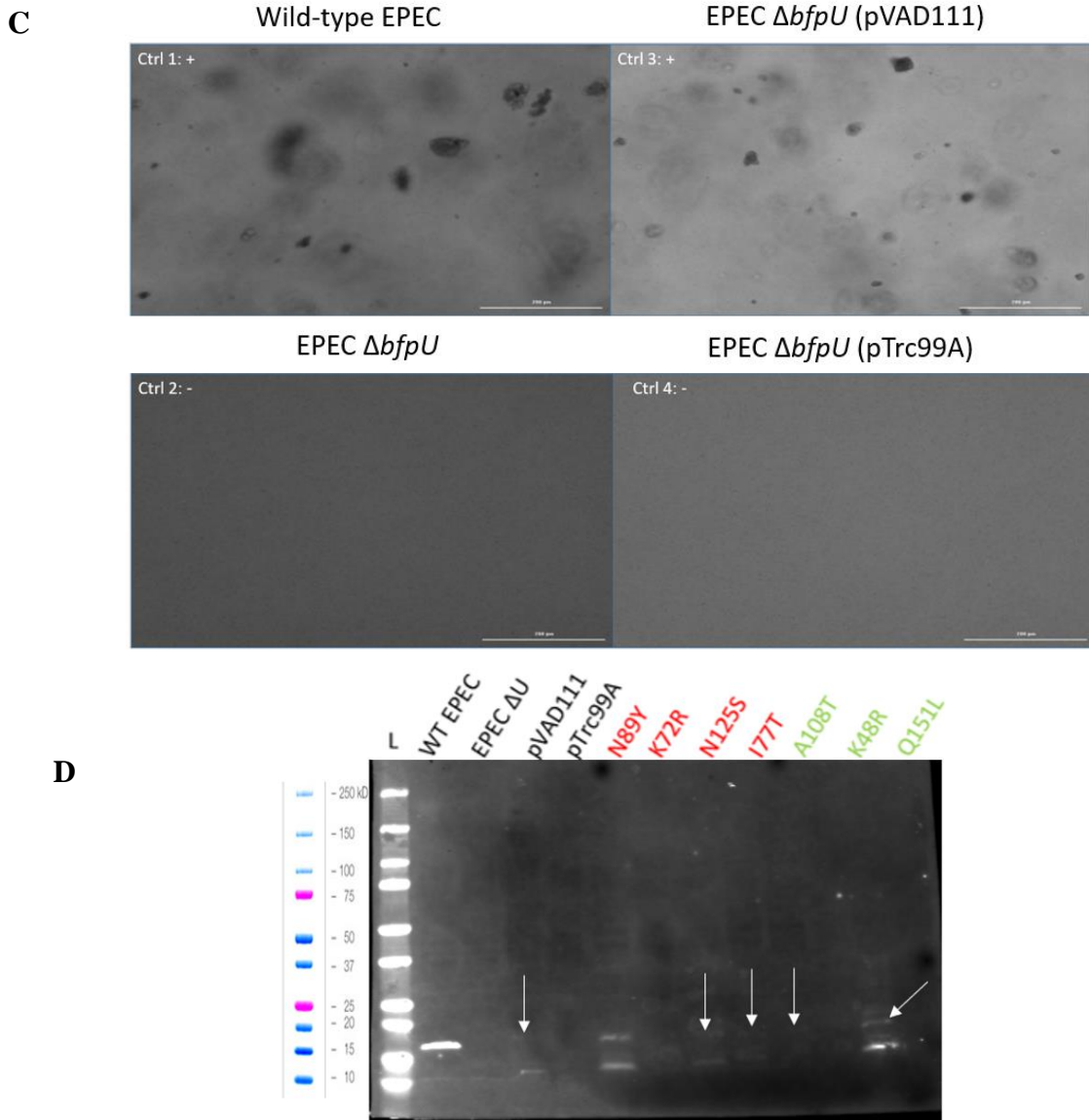
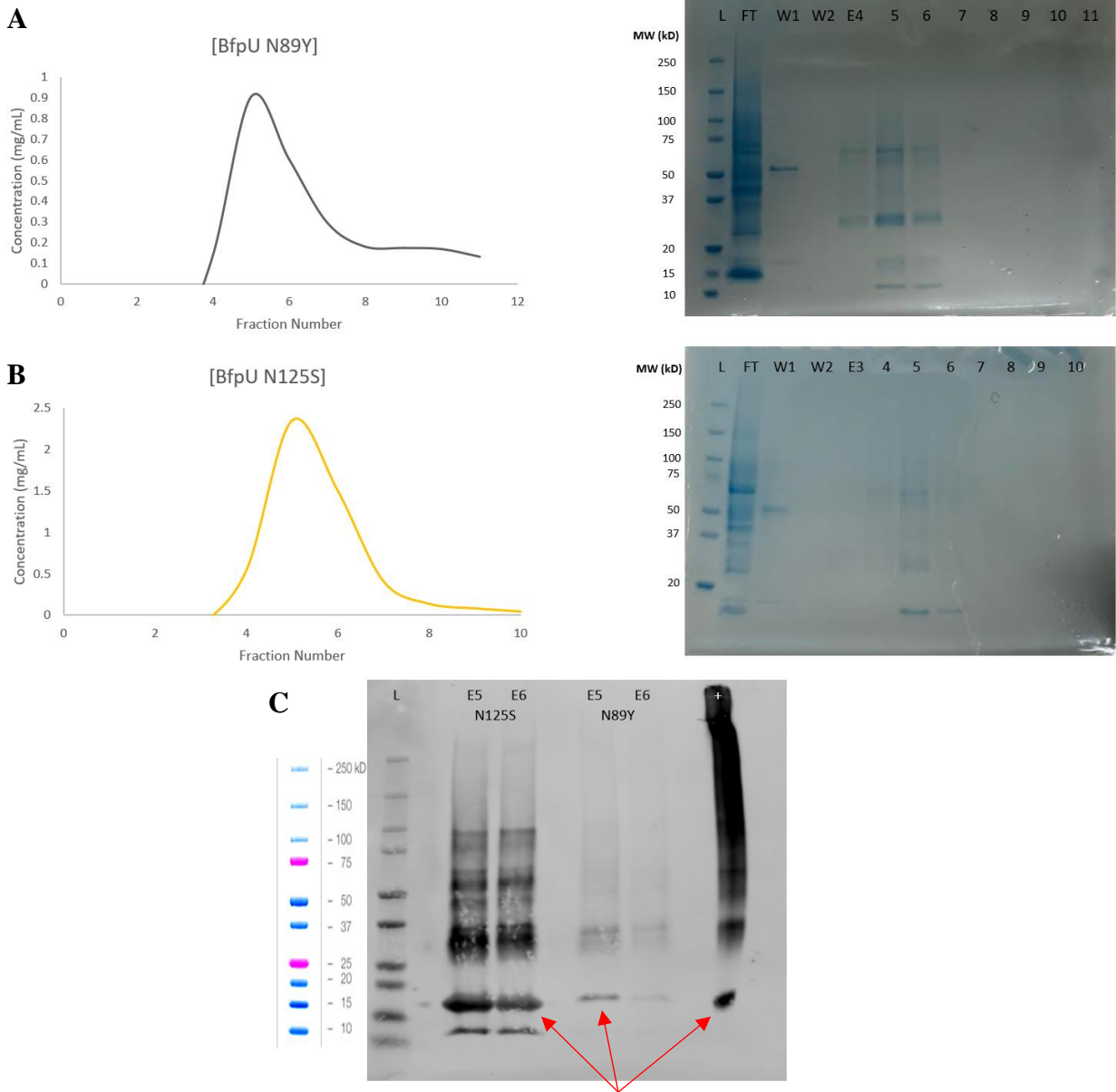


Figure 13: Auto-aggregation assay with *bfpU* mutants. (A) workflow of error-prone PCR, cloning and transformation. (B) Amino acid sequence of BfpU. Underlined region is the signal peptide. Residues in green were mutated and AA positive. Residues in red were mutated and AA negative. (C) examples of AA assay images capture with Cytation5. (D) immunoblot of *bfpU* expression in AA assay controls and mutants. BfpU MW ~17 kDa. White arrows highlights faint bands corresponding to BfpU.



Table 8: Summary of auto-aggregation assay.

<b>Mutant</b>	<b>AA by Cytation5 (x3)</b>	<b>Mutation</b>
pVAD111.101	-	N89Y
pVAD111.102	-	K72R
pVAD111.103	-	N125S
pVAD111.104	-(2/3)	I77T
pVAD111.105	+	A108T
pVAD111.106	+	K48R
pVAD111.107	+	Q151L



*Figure 14: Purification of BfpU N89Y and N125S. (A) Elution profile and SDS-PAGE of BfpU N89Y. (B) Elution profile and SDS-PAGE of N125S. (C) Immunoblot of eluate fractions 5 and 6 from affinity chromatography. Purified BfpU used as positive control (+). FT = flow through/cell lysate, W = wash, E = elution fractions. BfpU MW ~17 kDa (indicated by red arrows). Multiple bands detected by immunoblot possibly due to potential disulfide bond within BfpU.*

## Discussion

BfpU is an essential, periplasmic protein of the BFP system. In these studies, the previously published expression protocol for BfpU (118) was optimized to overexpress in the cytoplasm of SHuffle T7 to promote proper folding of BfpU. The periplasm of the cell is an oxidizing environment in which most proteins or cysteine residues would favorably form disulfide bonds, while the reducing environment of the cytoplasm of other *E. coli* strains would not (131). Additionally, overexpression in the periplasm is not ideal as it is a very small, concentrated environment. With the help of the SHuffle T7 strain, the oxidoreductase disulfide bond protein DsbC is present in the cytoplasm and can help with not only disulfide bond formation, but also favorable protein folding (127, 128, 132). The thermal stability assay showed that the recombinant expression and purification of BfpU from SHuffle T7 yielded a stable protein with a denaturing temperature of 80 °C. Additionally, the recent AlphaFold predicted structure of BfpU suggests that Cys 107 and Cys 119 are 2.04 Å apart from one another (figure 12), and disulfide bonds are usually 2.05 Å in length (132). However, further studies are needed to prove the presence of a disulfide bond within BfpU.

These studies presented evidence that BfpU is a structural homolog of PilP. Circular dichroism of BfpU (figure 11) suggested that it is primarily anti-parallel  $\beta$ -sheet and random coil just like PilP. The structure of PilP solved by NMR spectroscopy showed that it is a very flexible protein (due to the random coil) and has a  $\beta$ -sandwich fold, which consists of two up-and-down antiparallel  $\beta$ -sheets joined by hairpin loops (122, 133). The predicted AlphaFold structure of BfpU suggest the similar folds despite its lack of sequence homology to PilP or any other T4P proteins. Further studies are needed to determine the structure of BfpU and to confirm its structural homology to PilP.

BfpU is essential for BFP biogenesis (117). Random mutagenesis and auto-aggregation, a phenotype of EPEC that relies on functional BFP, helped identify possible residues that are important for BfpU function. A library of 200 *bfpU* mutants was developed and a preliminary screen via auto-aggregation assay highlighted seven mutations within BfpU. Three mutation sites were AA positive, which include Lys48Arg, Ala108Thr, and Gln151Leu. These residues may not influence BfpU structurally or may not be essential for interaction sites. Lys48 was mutated to Arg, which has similar properties so other mutations may need to be induced to further characterize that site. The three AA negative mutations were Lys72Arg, Asn89Tyr, and Asn125Ser. Interestingly, Lys72 was also mutated to Arg and could not restore BfpU function. These results infer that these AA negative residues are important for structural or interaction sites. Additional screening must be done to confirm all significant residues of BfpU.

From the AA assay mutant screen, two AA negative mutations expressed BfpU (figure 13D). These mutants, N89Y and N125S were cloned and expressed for Ni-NTA purification under the same conditions as wild-type BfpU-His (figure 14). BfpU N125S yielded a similar elution profile as wild-type BfpU-His. This mutation may not have influenced BfpU structurally, though additional studies are needed, such as thermal stability to compare to wild-type. Because of the N125S AA negative phenotype, N125 may be involved in an interaction site of BfpU. BfpU N89Y yielded significantly less protein compared to wild-type BfpU, suggesting this mutation may have a structural significance however, this does not rule out its possible role in BfpU interactions. Further studies are required to determine the roles of these residues for BfpU function.

## Chapter IV: Interaction between BfpU and BfpB periplasmic domains

### Introduction

In T4P systems, the secretin is known to interact with the IM alignment complex which spans the IM and periplasm. This type of interaction has been observed with the protein PilP. PilP is a lipoprotein anchored to in the IM (though primarily periplasmic) and has a similar structure to GspC, a component of the T2SS that has been shown to bind to its corresponding secretin GspD (134). Previous studies in *P. aeruginosa* revealed that PilP interacts with the OM secretin PilQ (119, 120). In *Myxococcus xanthus*, an “outside-in” assembly pathway was observed with PilQ in the OM recruiting a subcomplex consisting of PilNOP, through the interaction between PilP and PilQ (121). The concept of an “outside-in” pathway suggests that the machine assembly is initiated by formation of the secretin, followed by recruitment of the alignment subcomplex, which recruits the integral IM protein and cytoplasmic components. These studies concluded that PilQ is required for the stability of PilMNOP complex and implies that the assembly of the IM subassembly complex requires the presence of PilQ and PilP. In the BFP system, there is no protein that has recognizable sequence similarities with PilP. Information concerning the interactions between the secretin and other T4P components are limited to only a few species. Investigating interactions between the secretin and other BFP components in EPEC is also essential for understanding T4P biogenesis, especially of T4bP systems.

Multiple studies from the Donnenberg laboratory have suggested that BfpU interacts with BfpB (118). Using sucrose density gradient analysis, BfpU was found in the soluble, periplasmic fraction and the OM fraction in the presence of BfpB. In a  $\Delta bfpB$  mutant, BfpU was found mostly in periplasmic fraction. Additional evidence of an interaction was provided by a yeast two-hybrid system, which suggested that BfpU binds directly to the N-terminus of BfpB (residues 19-171).

Moreover, cross-linking assays revealed that complexes purified by affinity chromatography contain BfpB when the affinity tag is fused with BfpU and vice versa, suggesting that BfpB and BfpU are in close proximity to each other. Furthermore, the circular dichroism results (chapter III) suggest that BfpU has secondary structure characteristics similar to those of PilP. Finally, the recently released AlphaFold preliminary BfpU structure strongly suggests that BfpU is the PilP homologue in the BFP system. In this chapter, a combination of molecular biology and biophysics were used to further characterize the interaction between BfpU and BfpB.

## Methods

### Strains, plasmids, and growth conditions.

All strains and plasmids used in this chapter are presented in Table 9. Bacterial strains were cultured in LB broth at 37°C. Antibiotics (ampicillin, 200 µg/mL; kanamycin 50 µg/mL) were added to select for or maintain plasmids. All bacterial cultures were grown at 37°C with agitation at 225 rpm.

Table 9: Strains and plasmids used in chapter IV.

Strain or Plasmid	Genotype or Description	Reference or Source
SHuffle T7	<i>F' lac, pro, lacI<sup>q</sup> / Δ(ara-leu)7697 araD139 fhuA2 lacZ::T7 gene1 Δ(phoA)PvuII phoR ahpC* galE (or U) galK λatt::pNEB3-r1-cDsbC (Spec<sup>R</sup>, lacI<sup>q</sup>) ΔtrxB rpsL150(Str<sup>R</sup>) Δgor Δ(malF)3</i>	New England Biolabs
BL21 (DE3)	<i>F<sup>-</sup>, dmc, ompT, hsdS(rB-mB-)</i>	Novagen
pAD07	Encodes C-terminal histidine tagged BfpU. Cloned into NcoI/SalI sites of pBAD24	(118)
pJIL003	<i>bfpB<sub>19-306</sub></i> with C-terminal His tag cloned into pET28a NcoI and XhoI sites	Chapter II

## **Protein purification**

For purification of BfpB N0+N3, *E. coli* strain BL21 (DE3) (pJIL003) was grown at 37°C in Luria-Bertani medium to an optical density at 600 nm (OD600) of 0.6 and induced with 1 mM of IPTG at 37°C for 2 hours. For purification of BfpU, *E. coli* strain SHuffle T7 (pAD07) was grown at 37°C in Luria-Bertani medium to an optical density at 600 nm (OD600) of 0.6 and induced with arabinose (final concentration of 0.2%) at 37°C for 2 hours.

Cells were harvested by centrifugation and lysed by sonication in lysis buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 10 mM imidazole, pH 8.4) supplemented with protease inhibitor (Roche). BfpB N0+N3 and BfpB N0 were purified by affinity chromatography on nickel nitrilotriacetic acid (Ni-NTA) resin (Qiagen) column. Fractions eluted with 250 mM imidazole were analyzed by SDS-PAGE, combined, and concentrated with a 10K (Amicon) molecular weight cut-off filter. Concentrated samples were further purified on a Sephacryl S100 column with 20 mM HEPES, 150 mM NaCl, pH 7.4.

## **Surface plasmon resonance (SPR) spectroscopy**

All SPR experiments were carried out using a SR7500 Dual Channel instrument (Reichert Technologies, USA). Carboxymethyl dextran (CM5) sensor chips were obtained from Reichert Technologies. The amine coupling reagents and regeneration buffer (10 mM glycine, pH 2.0) were obtained from Cytiva Life Sciences, USA. The experimental setup is summarized in figure 15.

### *Immobilization of BfpU-His.*

Purified BfpU was immobilized onto a CM5 sensor chip surface via standard amine coupling chemistry using running buffer (20 mM HEPES, [pH 7.4], 150 mM NaCl, and 0.05% tween-20) at 25 °C, as per the manufacturer's recommendation. Two chips were immobilized with



purified BfpU from separate biological replicates. Briefly, carboxyl groups on the sensor chip were activated by injecting an equal mixture of 200 mM N-hydroxysuccinimide and 50 mM 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride at a flow rate of 10  $\mu\text{L}/\text{min}$ . BfpU (20  $\mu\text{g}/\text{mL}$ , chip #1; 40  $\mu\text{g}/\text{mL}$ , chip #2) was then diluted in 10 mM sodium acetate buffer pH 5.0 and injected over the surface at a flow rate of 10  $\mu\text{L}/\text{min}$ . This was followed by a 10 min injection of 1 M ethanolamine (pH 8) to block remaining activated sites on the chip and wash away any noncovalently bound BfpU.

### *BfpU binding analysis*

Analyte stock solutions were made to create 200  $\mu\text{L}$  samples of 1, 25, 50 (twice), 100 and 150  $\mu\text{M}$  for binding experiments with purified BfpB N0+N3-His and bovine serum albumin (BSA, negative control) (#A3294-100G, Sigma-Aldrich, USA). For these experiments, three trials were conducted with both chips using purified BfpB N0+N3 from separate biological replicates (a new purification per trial). Stock solution of BfpU monoclonal mouse IgG (positive control) (117) was made to create 200  $\mu\text{L}$  samples of 1, 25 (twice), 50  $\mu\text{M}$  for binding experiments. Stocks were diluted in the running buffer. Blank (running buffer) was injected at the beginning and end of analyte series. Injection time was 3 min, dissociation time 8.5 min, and regeneration 6 min. Flow rate was 25  $\mu\text{L}/\text{min}$ .

### **Data Processing and Curve Fitting.**

The data processing for kinetic and affinity for BfpU and the analytes binding interactions were performed using TraceDrawer 1.8 Software. All of the binding sensograms were collected at 25  $^{\circ}\text{C}$  prior to curve fitting. The final data obtained, which fit into a Langmuir 1:1 model, was double referenced using flow cell 1 and buffer blank subtractions. A 1:1 binding model is

commonly used for SPR analysis because it assumes a simple binding interaction between two molecules, where one molecule binds to a single binding site on the other molecule. In other words, the model assumes that the binding is stoichiometric, and that there is no cooperativity or allosteric effects involved. This model is also sufficient for analyzing the binding of small molecules to proteins or the binding of antibodies to antigens, which typically involve a single binding site. In these cases, the model provides a simple and accurate description of the binding kinetics and affinity. The dissociation constant,  $K_d$ , can be calculated by the following equation:

$$K_d = \frac{k_d}{k_a}$$

Where  $k_d$  is the dissociation rate ( $k_{off}$ ) in units of 1/s, and  $k_a$  is the association rate ( $k_{on}$ ) in units of 1/M·s.

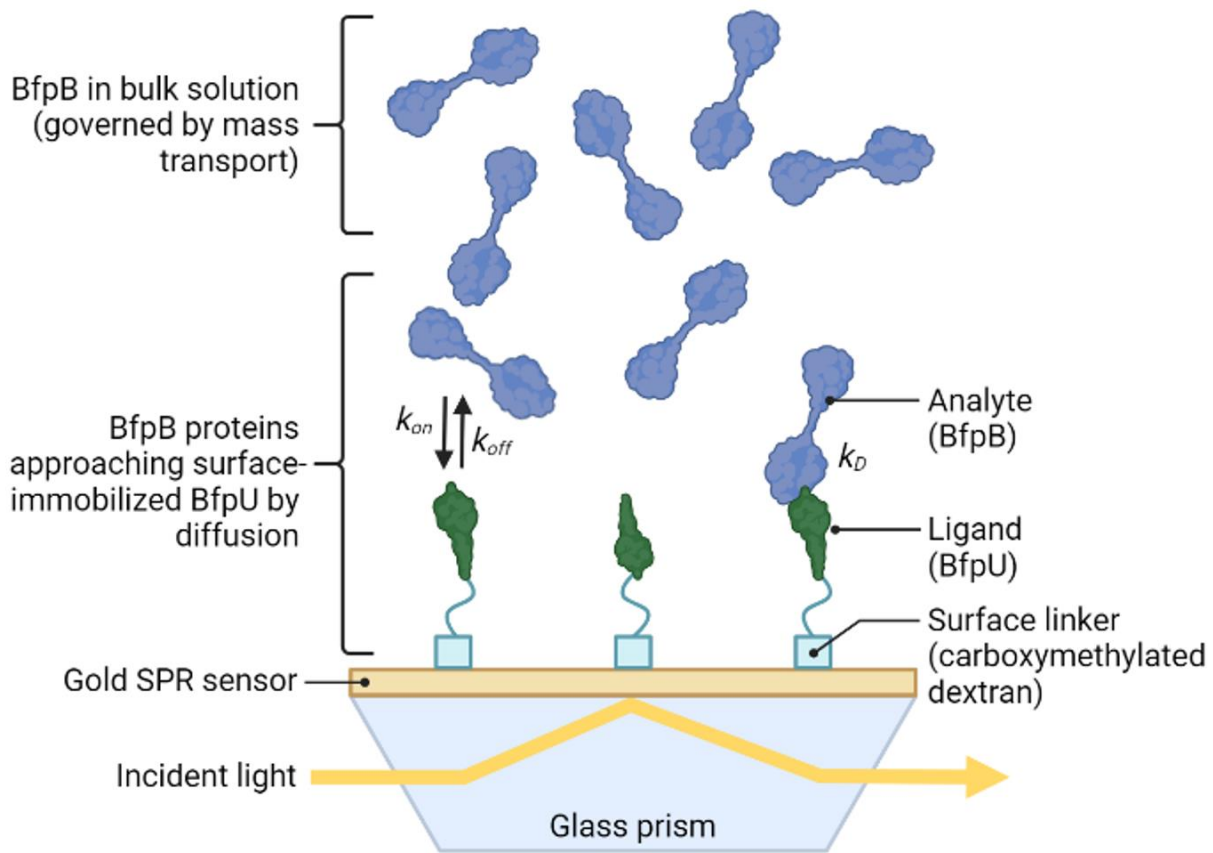


Figure 15: Schematic of surface plasmon resonance experimental setup with BfpB N0+N3 and BfpU. BfpU is immobilized to a CM5 chip via amine coupling in a random orientation, serving as the ligand. BfpB N0+N3 and the controls (BfpU IgG and BSA) are flowed over the chip, serving as the analytes in separate experiments. Analytes are free in solution and are able to interact with the ligand or chip surface. The mass of bound material is proportional to the change in the refractive index, and the SPR signal is dependent on the refractive index (135). As binding occurs, the change in SPR signal is measured in resonance units (RU). Schematic created using BioRender.

## Results

### **Real-time measurements and kinetic analyses of the interactions between BfpU, BfpB N0+N3 and controls via surface plasmon resonance.**

BfpB is a transmembrane protein with a soluble N-terminus. In chapter II, the residues 19-306 were predicted to represent the periplasmic domains of N0 and N3 of BfpB. This construct was purified in abundance and deemed stable. To avoid issues regarding solubility, aggregation, and buffer compatibility, the soluble BfpB N0+N3 construct was used in the interaction studies. This construct is also ideal because it is thought to reside in the periplasm along with BfpU. Figures 16-18 demonstrate the SPR binding experiments of immobilized BfpU with BfpB N0+N3, BfpU IgG (positive control), and BSA (negative control).

The trials of the BfpU-BfpB N0+N3 interaction showed that the concentration range of 1-150  $\mu\text{M}$  was ideal, as binding was detected and saturable. The binding experiments between BfpU and BfpB N0+N3 suggest that there is an interaction. Across all trials, the dissociation constant ranged from 3.9 to 8.8  $\mu\text{M}$ , with a mean ( $\pm$  SD)  $K_d$  of  $6.08 \pm 1.89 \mu\text{M}$ . These results are summarized in table 10. For the binding experiment between BfpU and BfpU IgG, a lower range of concentrations were used because the antibody precipitated at concentrations above 50  $\mu\text{M}$ . Within the range of 1-50  $\mu\text{M}$  saturation was achieved. The  $K_d$  of the antibody interaction was 133 and 134 nM for chip 2 and chip 1, respectively. Lastly, the same concentrations from the BfpB N0+N3 experiment were used for BSA. Unexpectedly, binding was detected between BfpU and BSA with  $K_d$  values of 16.2  $\mu\text{M}$  and 6.6  $\mu\text{M}$  for chips 1 and 2, respectively.

Figure 16

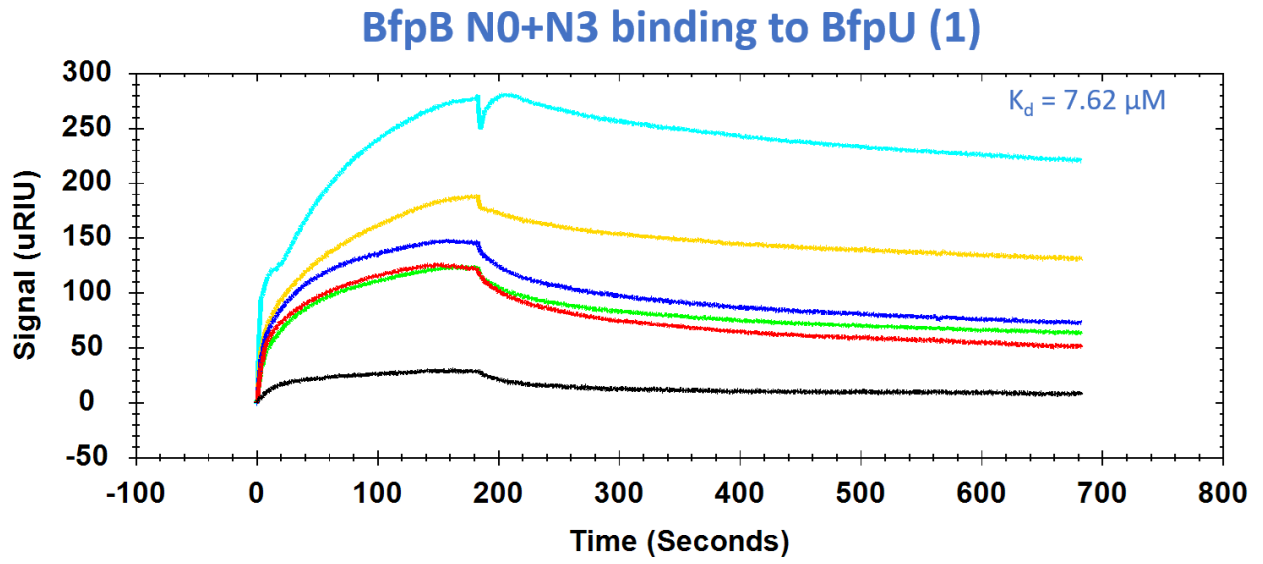
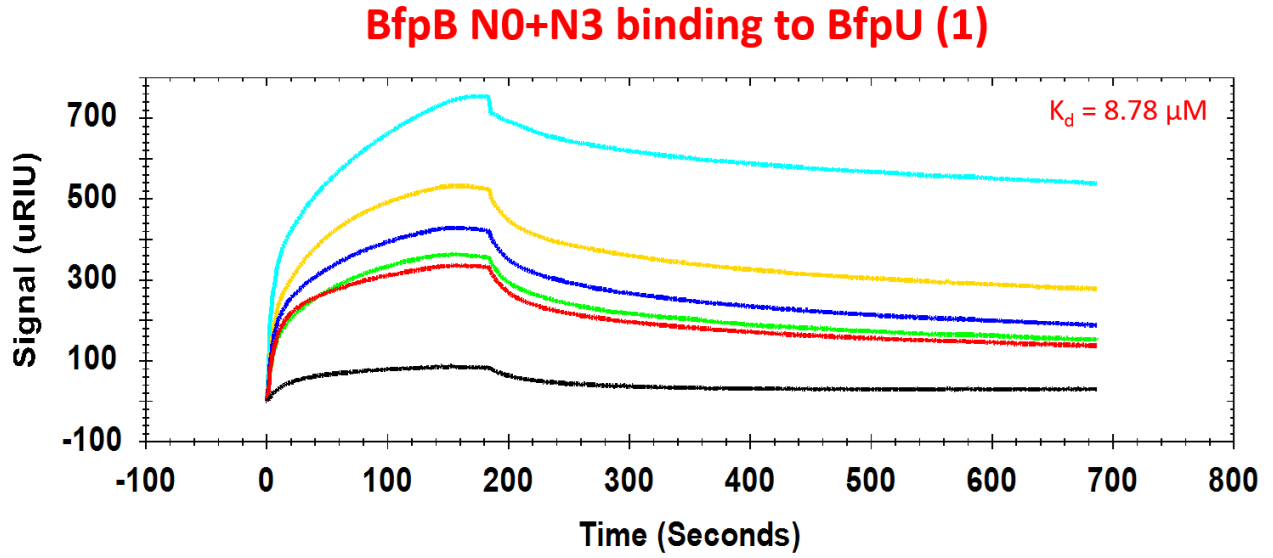
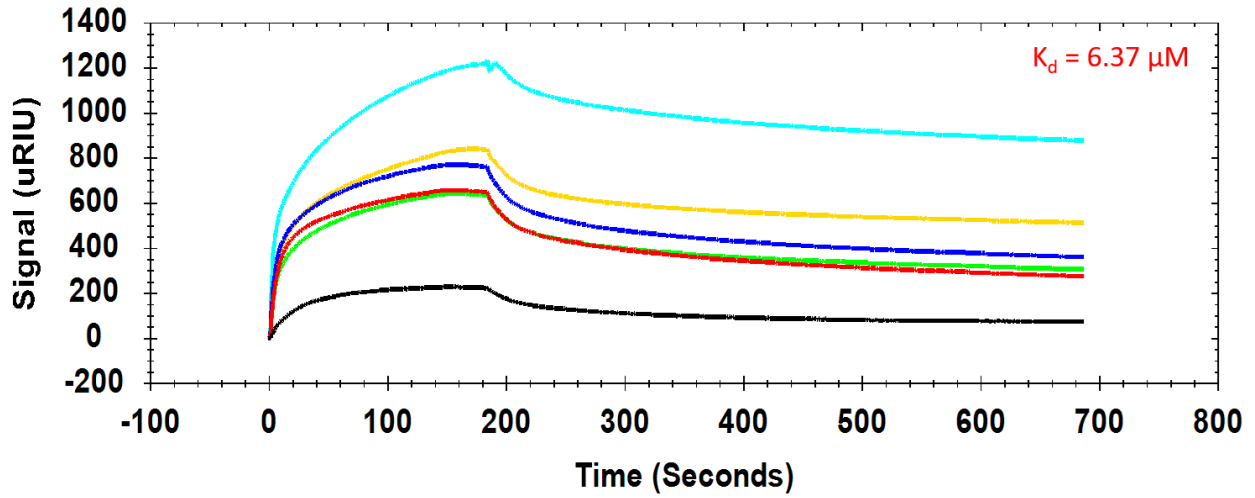


Figure 16, continued.

### BfpB N0+N3 binding to BfpU (2)



### BfpB N0+N3 binding to BfpU (2)

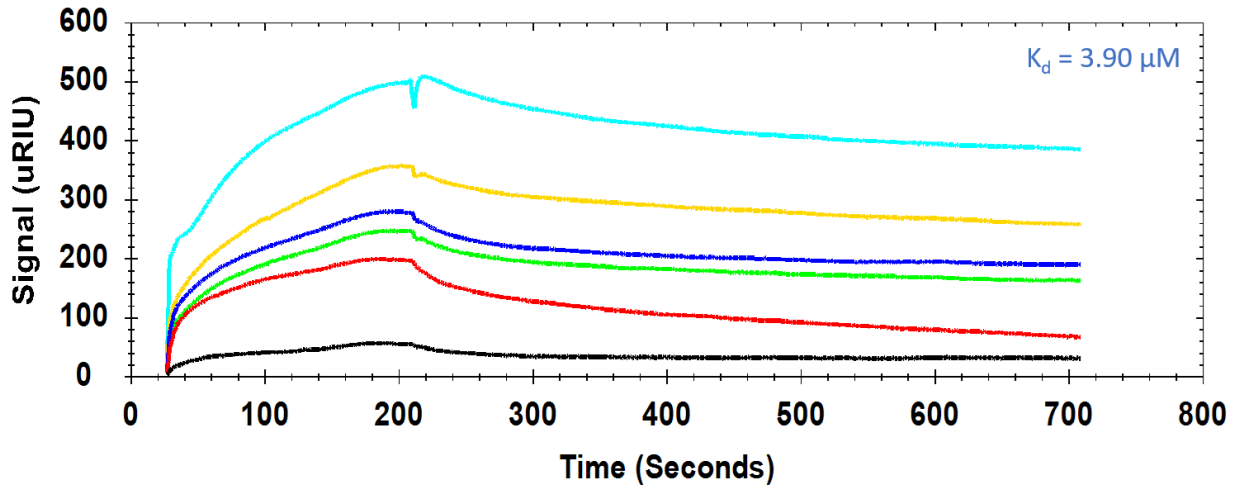


Figure 16, continued.

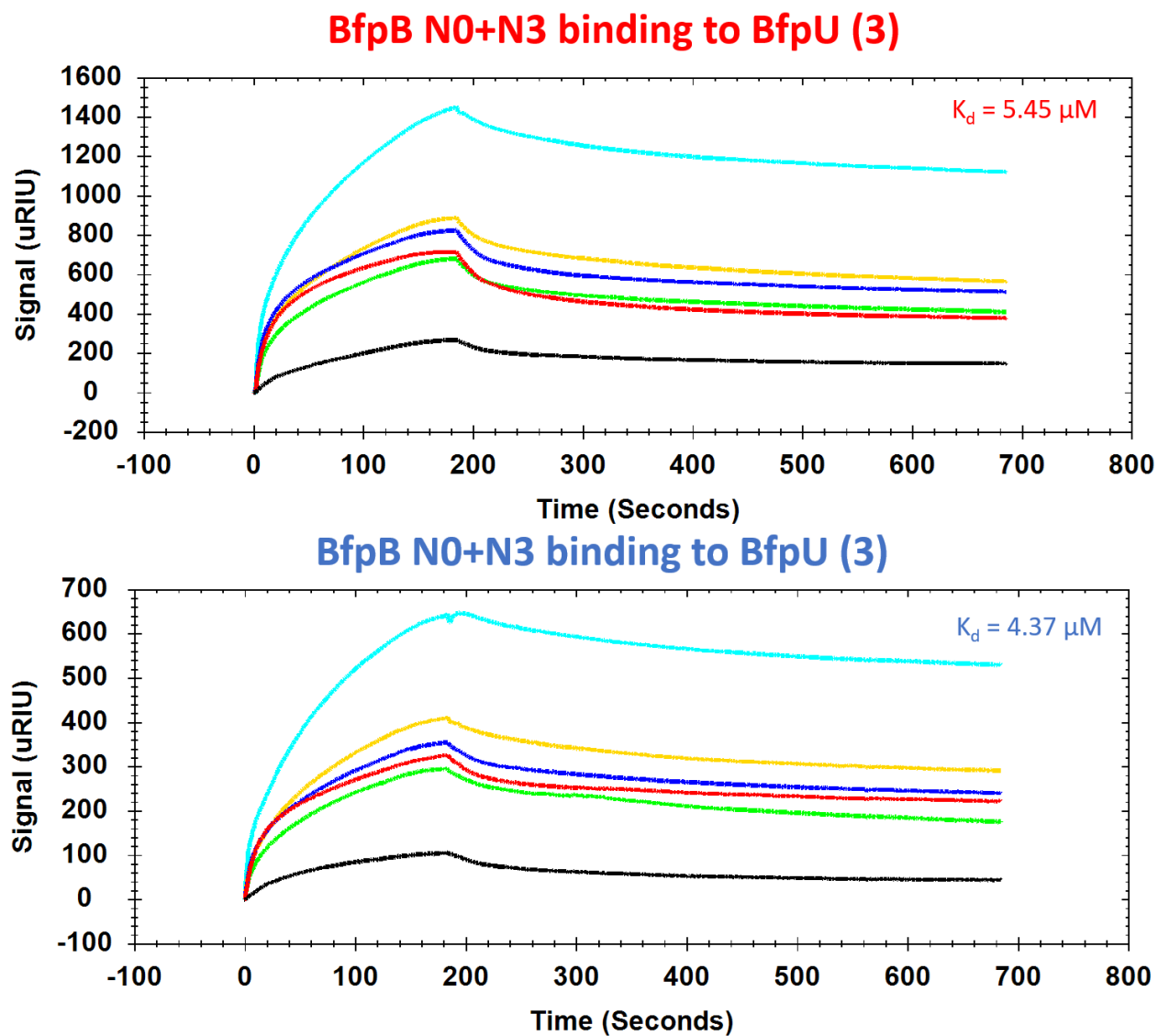


Figure 16: Surface plasmon resonance response vs time plots with BfpB N0+N3 and BfpU. Three trials were performed with separate, freshly purified BfpB N0+N3 as the analyte. BfpB N0+N3 concentrations of 1, 12.5, 25 (x2), 50, and 150  $\mu\text{M}$  were used. Chip 1 plots are labeled in red while chip 2 plots are labeled in blue. Both chips have BfpU immobilized from separate purifications. Trial numbers are in parentheses for both chips.  $K_d$  values were determined from one-to-one global fittings of the curves using TraceDrawer software.

Table 10: Summary of surface plasmon resonance results of BfpU and BfpB N0+N3. Chip 1 and 2 used purified BfpU from separate biological replicates. Values in red are from chip 1 and values in blue are chip 2. Each trial used a separate, freshly purified sample of BfpB N0+N3 for the experiment. The mean, standard deviation, and standard error were calculated using both chips and all trials.

<b>Trial #</b>	<b>K<sub>d</sub> (μM)</b>	<b>Mean</b>	<b>St Dev</b>	<b>St Error</b>
1	8.78	6.08	1.89	0.77
	7.62			
2	6.37			
	3.9			
3	5.45			
	4.37			



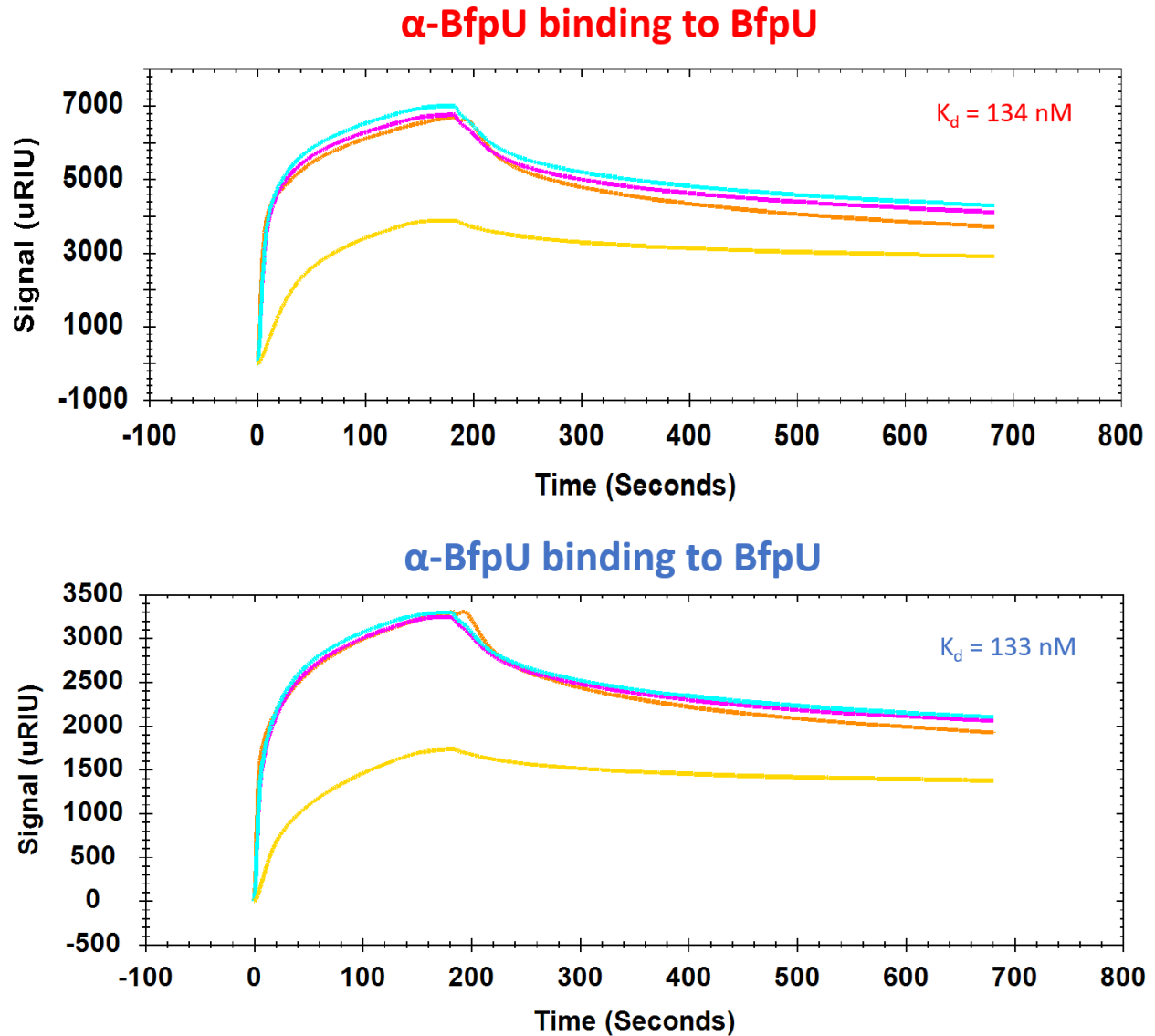


Figure 17: Surface plasmon resonance response vs time plots with BfpU IgG ( $\alpha$ -BfpU) and BfpU. BfpU IgG concentrations of 1, 25 (x2), and 50  $\mu\text{M}$  were used. Chip 1 and 2 are labeled in red and blue, respectively.  $K_d$  values were determined from one-to-one global fittings of the curves using TraceDrawer software.

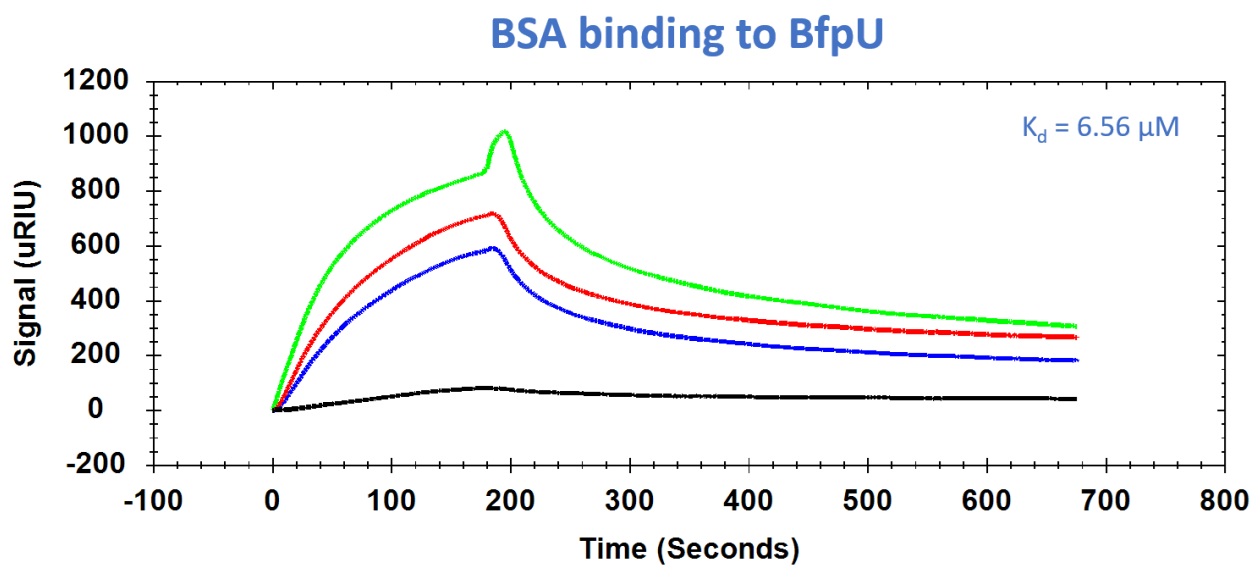
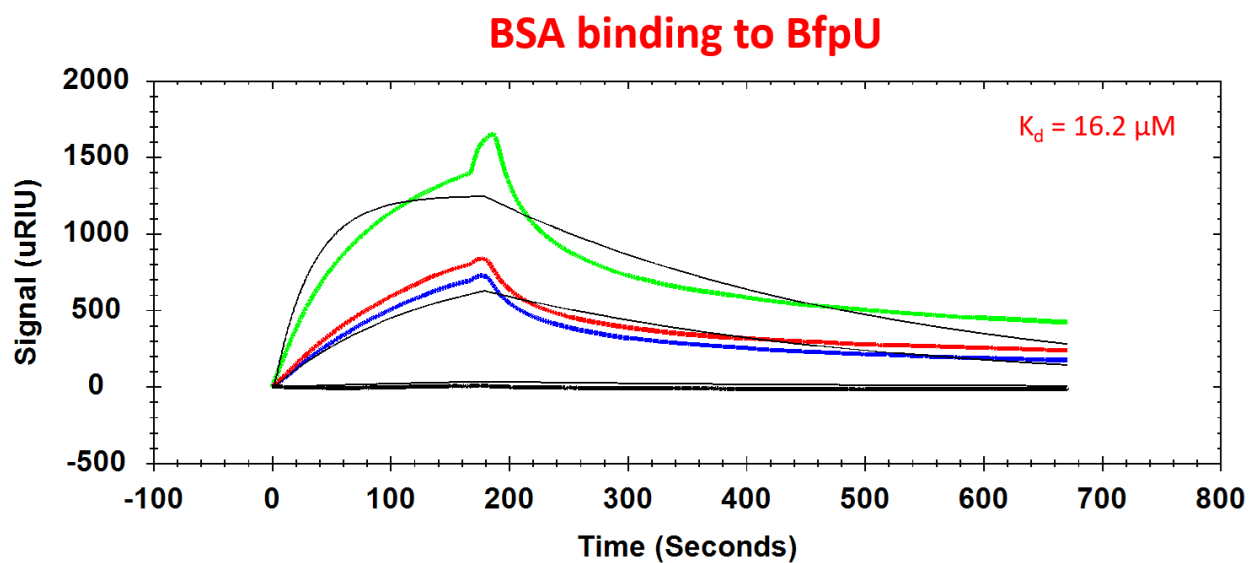


Figure 18: Surface plasmon resonance response vs time plots with BSA and BfpU. BSA concentrations of 1, 50 (x2), and 150  $\mu\text{M}$  were used. Chip 1 and 2 are labeled in red and blue, respectively.  $K_d$  values were determined from one-to-one global fittings of the curves using TraceDrawer software.

## Discussion

BfpB is the secretin of the BFP system, and it is likely to interact with neighboring periplasmic BFP proteins, as seen in other T4P systems. Our data implies that BfpU may be a PilP homolog, suggesting that BfpU should interact with the PilQ homolog BfpB. The carboxyl terminus of BfpB is imbedded in the OM, so its putative periplasmic domains of N0+N3 were used to assess the interaction with BfpU. Preliminary studies using isothermal titration calorimetry suggested that BfpU and BfpB N0+N3 interact. However, the heat released from the interaction did not overcome the “heat of dilution” from the concentrated BfpB N0+N3 that was titrated into the reaction cell, so the affinity could not be accurately calculated (data not shown). As a result, this interaction was characterized by surface SPR because it is a label free, non-destructive way to characterize ligand binding, uses less protein and provides real time data such as on and off rates of an interaction (135, 136).

Here we report the first quantitative interaction between a T4P system secretin and its neighboring counterpart using SPR. These experiments showed that BfpU interacts with BfpB N0+N3 with an average  $K_d$  of  $6.08 \pm 0.77 \mu\text{M}$ . It is difficult to gauge whether its  $K_d$  is reasonable because BfpU is soluble and can move in three dimensions while BfpB is anchored to the membrane and can move in two dimensions (137). Also, the periplasm is a concentrated, gelatinous environment. Considering these conditions, it is likely that BfpU and BfpB do interact, but this interaction might not require high affinity. The positive control experiment with monoclonal BfpU IgG worked well with a high  $K_d$  of 133.5 nM. However, the negative control BSA showed binding at similar concentrations to BfpB N0+N3, casting doubt on the specificity of the observed BfpU-BfpB interaction. Of note, we originally chose the N1 domain of the cytoplasmic extension ATPase BfpD as the negative control, reasoning that it should not interact

with BfpU, but recorded similar results (data not shown). An ideal negative control for SPR would be a deactivated version of the ligand (e.g., phosphorylated protein vs dephosphorylated) or an analog of the ligand, but neither of these is available for BfpU. If the interaction between BfpB and BfpU exists, it may have a low affinity, which could make it challenging to accurately measure by SPR.

The use of SPR for assessing protein interactions may lead to other potential problems. This technique requires the protein of interest to be fixed onto a chip which can mask key residues and possibly alter protein structure. BfpU was immobilized to the chip such that the bound particles are randomly oriented. The ratio of BfpU bound with the potential interaction site being exposed versus unexposed is not clear. Further optimization of the conditions needs to be done to improve the SPR experiments. The running buffer has a significant impact on real-time binding traces. The buffer used for the experiments was composed of 20 mM HEPES, 150 mM NaCl, and 0.05% tween-20 at pH 7.4. HEPES is typically used for SPR experiments as it is a reliable choice for many reasons (138, 139). It has a high buffering capacity in the pH range generally used in SPR experiments (pH 6-8). Also, HEPES is a zwitterionic buffer, so it can help reduce non-specific binding of biomolecules to the SPR sensor surface, which can interfere with the measurement of specific binding events (139, 140). Furthermore, HEPES has a low affinity for metal ions, which is important for SPR experiments that use metal-coated sensor surfaces (140). The NaCl concentration used in these experiments was common, however, other concentrations should be tried. Salt affects the ionic strength of a solution, which influences solubility and electrostatic interactions. The addition of detergents like Tween-20 often helps reduce non-specific binding from occurring, but higher concentrations may be tried to improve the current conditions. Furthermore, BSA was used as a negative control in these experiments because BSA is often added

to the running buffer to decrease non-specific binding and because it typically interacts with small, charged compounds or atoms, but not other proteins (141). Once another negative control is established, BSA should be added into the running buffer instead of being used as a control. Moreover, pH values should be tried as it can induce a charge effect on the experimental environment. The pI of BfpU is 6.6 and the pH of the buffer was 7.4. At this pH, BfpU would have a net charge between -0.84 to -0.41. Under these same conditions, BfpB N0+N3 would have a net charge between -7.46 to -6.34. At pH 7.4, BSA has a net charge of -9 (142). Both the ligand and analytes were negatively charged under these conditions which likely influenced the results. Aside from the net charges, pH may affect the folding or conformation of these proteins which can interfere with interactions. Various conditions should be tried to find the optimum buffer for these studies.

Given the results obtained, we cannot at this time confirm the hypothesis that BfpU binds specifically to BfpB.

## Chapter V: Discussion

The objective of this dissertation was to elucidate the structure of secretin BfpB, and to characterize the interaction between BfpB and BfpU. Investigating the functions of BfpB and BfpU is crucial to comprehensively understand T4P systems and their mechanisms, even though their sequences are different from homologs in other species. This will also enhance research for understanding other secretin reliant systems, like T2S and T3S, which are confirmed to be virulence factors. Inhibiting the function of the secretin protein could prevent the assembly of T4P, T2S, and T3S, making it a potential target for the development of new antimicrobial agents that could help limit bacterial infections. Additionally, utilizing BFPs as a model for T4P biogenesis is advantageous due to their role in virulence in typical EPEC infections.

The first part of this dissertation work was to determine the structure of BfpB. BfpB needed to be purified under optimal conditions for structure determination by cryo-EM. Several optimizations were made to an earlier purification protocol (93). It was shown that chemical lysing with a zwitterionic detergent prior to the column steps increases the yield of purified BfpB compared to sonication or homogenization via French Press. The affinity chromatography step was improved by increasing the column size and the concentration of desthiobiotin to elute bound BfpB-Strep molecules. Also, a second purification step via size exclusion chromatography with zwitterionic detergent present improved the sample quality for cryo-EM by isolating the toroids from monomers.

The cryo-EM data yielded interesting results about the structure of BfpB despite the low resolution of 7.1 Å. One the 2D class averages appears to have C17 symmetry, suggesting that BfpB is composed of 17 subunits. This is unusual, and the other known secretins have 12-15 subunits with a symmetry of C12-15 (101). In theory, T4bP secretins such as BfpB may be wider

to accommodate wider substrates compared to other secretin-reliant systems. An improved resolution would help confirm the number of subunits and cyclic symmetry of BfpB. Also, the 3D reconstruction of BfpB shows a density right below the gate of the  $\beta$ -barrel that is not typically seen in other secretins. Only one other secretin possesses this extra density, and it was resolved to be an extended  $\beta$ -hairpin of the N3 domain (112). Intriguingly, the predicted structure of BfpB from AlphaFold has a similar  $\beta$ -hairpin that would fit in to this density. AlphaFold and SIB MyHits motif scanner predicts BfpB to have three domains and the 3D reconstruction suggest the same: N0, which is poorly resolved in the structure, N3, and the secretin domain. The putative N0 and N0+N3 domains were successfully purified, are stable, and appear by thermal stability assays to fold correctly. It is likely that the secretin domain begins after Arg306 in BfpB. However, the exact boundaries and folds of each domain remain unknown. Additional studies, which include improving the resolution of the BfpB structure, are needed. We hope to improve the structure of BfpB with analysis of additional data that have very recently been collected.

The major pitfall of these structural studies was the paucity of BfpB particles from the cryo-EM data set, which yielded a low resolution. There are multiple reasons contributing to the low particle number and resolution. Thin carbon coated grids were used; particles can preferentially stick to the carbon layer and fail to partition into holes (143, 144). Also, particles can adopt preferred orientations within the vitrified ice layer, thus limiting the number of unique views. This was the case with BfpB, as the top projection dominated the 2D classification. One way to overcome this is by tilting the stage at an angle to possibly capture other orientations of the protein. Furthermore, blotting conditions could be further optimized to yield the best vitreous ice conditions for the particles.

Further optimizations could be made to the purification protocol to possibly increase the number of BfpB toroids. The detergent is the key factor of the monomer to protomer ratio as seen in chapter II. Various detergent concentrations were tried for size exclusion chromatography step but not at the lysing or affinity chromatography step. A lower concentration of detergent may be helpful for further improving the ratio of multimers prior to size exclusion. Also, co-purification of BfpB with an OM associated protein may help with multimer formation. There are two groups of proteins often associated with secretins which are pilotins and accessory proteins. Pilotins are involved in the localization and assembly of secretins and are found in T4aP, T2S and T3S systems (101). To date, they have not been described in T4P competence systems, filamentous phage, and T4bP systems. Accessory proteins usually are involved in the secretin stability (105, 145). Without accessory proteins, secretins were observed to be reduced in number, only in the monomeric form, or degraded. BfpG was claimed to be an accessory protein (146). Not much information is known about accessory protein mechanisms of BfpG. It is unclear whether BfpG is truly an accessory protein as we were able to purify and isolate BfpB complexes. If in fact an accessory protein, BfpG could possibly help improve the number of BfpB particles needed for cryo-EM.

The second part of this dissertation work was to determine whether BfpB interacts with BfpU directly. Considering the lack of information known about BfpU, further characterization was needed prior to the interaction studies. BfpU is a periplasmic protein that has no known homologs and is essential for BFP biogenesis (117, 147). Prior to expression in SHuffle, BfpU would precipitate after affinity chromatography. Over expression in *E. coli* SHuffle allowed for stable, folded BfpU. BfpU was shown to have a denaturation temperature of 80 °C. Circular dichroism data suggested that BfpU is primarily anti-parallel  $\beta$ -sheet and random coil, which is similar to the structure of PilP. PilP is an IM protein, yet primarily periplasmic, and interacts with



the secretin PilQ of other T4P systems. Furthermore, the predicted structure of BfpU from AlphaFold shows a similar structure to PilP (RMSD of 0.850 Å) despite its lack of sequence homology. The predicted structure shows that the cysteine residues of BfpU are ~2 Å apart, which is close enough to form a disulfide bond (131). Random mutagenesis has also helped further characterize BfpU, but additional screening is needed to confirm important residues. Moreover, additional structural studies are needed to help determine the function of BfpU and if it is an unrecognized PilP homolog. We hope to determine the structure of BfpU via NMR.

We hypothesized that BfpU is a PilP homolog, therefore it would interact with BfpB as PilP interacts with PilQ. Multiple studies from the Donnenberg laboratory implied that there is an interaction between BfpU and BfpB (91, 93). To quantitatively characterize this interaction, SPR was used. The results suggest that the putative periplasmic domains BfpB N0+N3 interact with BfpU with micromolar affinity. However, these results were undermined when the negative control BSA was shown to interact with BfpU with similar affinity. It is possible that buffer optimizations might improve the specificity of these binding experiments. These optimizations include salt concentration, pH, detergent concentration, and the addition of BSA to the running buffer. Also, higher concentrations of BfpU may help with increasing the amount of BfpU immobilized to the chips for more accurate readings. Moreover, legitimate negative controls are needed. Although non-specific binding is possible, it is probable that there is at least a relatively weak interaction between BfpU and BfpB, given that BfpB is located in the OM and partially extends into the periplasm, while BfpU has limited freedom of movement in the periplasmic space.

The work in this dissertation has improved the information known of the BFP system and T4bP systems. However, more work is necessary to answer the remaining questions. Is BfpB composed of 17 subunits? Are T4bP secretins wider and possess more subunits than other

subclasses of secretins? What functions do the periplasmic domains serve? With which domain of BfpB does BfpU interact? Is BfpU a structural homolog of PilP? If BfpU is a homolog of PilP, with what other BFP proteins does it interact? Continuation of structural studies including cryo-EM, NMR, and functional studies via mutagenesis will help answer these questions.

## References

1. Kotloff KL. 2022. Bacterial diarrhoea. *Curr Opin Pediatr* 34:147–155.
2. Getachew HB, Dahl RM, Lopman BA, Parashar UD. 2018. Rotavirus Vaccines and Health Care Utilization for Diarrhea in US Children, 2001 to 2015. *Pediatr Infect Dis J* 37:943–948.
3. Newall AT, Leong RN, Reyes JF, Curns AT, Rudd J, Tate J, Macartney K, Parashar U. 2021. Rotavirus Vaccination Likely to Be Cost Saving to Society in the United States. *Clin Infect Dis* 73:1424–1430.
4. Micoli F, Bagnoli F, Rappuoli R, Serruto D. 2021. The role of vaccines in combatting antimicrobial resistance. *Nat Rev Microbiol* 19:287–302.
5. Talbot GH, Jezek A, Murray BE, Jones RN, Ebright RH, Nau GJ, Rodvold KA, Newland JG, Boucher HW. 2019. The Infectious Diseases Society of America’s 10 × ’20 Initiative (10 New Systemic Antibacterial Agents US Food and Drug Administration Approved by 2020): Is 20 × ’20 a Possibility? *Clin Infect Dis* 69:1–11.
6. Hotinger JA, Gallagher AH, May AE. 2022. Phage-Related Ribosomal Proteases (Prps): Discovery, Bioinformatics, and Structural Analysis. *Antibiotics (Basel)* 11.
7. Duménil G. 2019. Type IV Pili as a Therapeutic Target. *Trends Microbiol* 27:658–661.
8. Clatworthy AE, Pierson E, Hung DT. 2007. Targeting virulence: a new paradigm for antimicrobial therapy. *Nat Chem Biol* 3:541–548.

9. Peabody CR, Chung YJ, Yen MR, Vidal-Ingigliardi D, Pugsley AP, Saier MH. 2003. Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology (Reading)* 149:3051–3072.
10. Kaper JB, Nataro JP, Mobley HLT. 2004. Pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2:123–140.
11. Winstead A, Hunter JC, Griffin PM. *Escherichia coli, Diarrheagenic - Chapter 4 - 2020 Yellow Book | Travelers' Health | CDC*.  
<https://wwwnc.cdc.gov/travel/yellowbook/2020/travel-related-infectious-diseases/escherichia-coli-diarrheagenic>. Retrieved 21 March 2023.
12. Spina A, Kerr KG, Cormican M, Barbut F, Eigentler A, Zerva L, Tassios P, Popescu GA, Rafila A, Eerola E, Batista J, Maass M, Aschbacher R, Olsen KEP, Allerberger F. 2015. Spectrum of enteropathogens detected by the FilmArray GI Panel in a multicentre study of community-acquired gastroenteritis. *Clinical Microbiology and Infection* 21:719–728.
13. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, Wu Y, Sow SO, Sur D, Breiman RF, Faruque AS, Zaidi AK, Saha D, Alonso PL, Tamboura B, Sanogo D, Onwuchekwa U, Manna B, Ramamurthy T, Kanungo S, Ochieng JB, Omore R, Oundo JO, Hossain A, Das SK, Ahmed S, Qureshi S, Quadri F, Adegbola RA, Antonio M, Hossain MJ, Akinsola A, Mandomando I, Nhampossa T, Acácio S, Biswas K, O'Reilly CE, Mintz ED, Berkeley LY, Muhsen K, Sommerfelt H, Robins-Browne RM, Levine MM. 2013. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *The Lancet* 382:209–222.

14. Donnenberg MS, Kaper JB, Finlay BB. 1997. Interactions between enteropathogenic *Escherichia coli* and host epithelial cells. *Trends Microbiol* 5:109–114.
15. McDaniel TK, Jarvis KG, Donnenberg MS, Kaper JB. 1995. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc Natl Acad Sci U S A* 92:1664.
16. Jarvis KG, Girón JA, Jerse AE, Mcdaniel TK, Donnenberg MS, Kaper JB. 1995. Enteropathogenic *Escherichia coli* contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation. *Proc Natl Acad Sci U S A* 92:7996–8000.
17. Khursigara C, Abul-Milh M, Lau B, Girón JA, Lingwood CA, Barnett Foster DE. 2001. Enteropathogenic *Escherichia coli* Virulence Factor Bundle-Forming Pilus Has a Binding Specificity for Phosphatidylethanolamine. *Infect Immun* 69:6573.
18. Girón JA, Ho ASY, Schoolnik GK. 1991. An inducible bundle-forming pilus of enteropathogenic *Escherichia coli*. *Science* 254:710–713.
19. Bieber D, Ramer SW, Wu CY, Murray WJ, Tobe T, Fernandez R, Schoolnik GK. 1998. Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic *Escherichia coli*. *Science* 280:2114–2118.
20. Hyland RM, Sun J, Griener TP, Mulvey GL, Klassen JS, Donnenberg MS, Armstrong GD. 2008. The bundlin pilin protein of enteropathogenic *Escherichia coli* is an N-acetyllactosamine-specific lectin. *Cell Microbiol* 10:177–187.

21. Levine MM, Nalin DR, Hornick RB, Bergquist EJ, Waterman DH, Young CR, Sotman S, Rowe B. 1978. *Escherichia coli* strains that cause diarrhoea but do not produce heat-labile or heat-stable enterotoxins and are non-invasive. *Lancet* 1:1119–1122.
22. Trabulsi LR, Keller R, Tardelli Gomes TA. 2002. Typical and atypical enteropathogenic *Escherichia coli*. *Emerg Infect Dis* 8:508–513.
23. Afset JE, Bevanger L, Romundstad PL, Re Bergh K. 2004. Association of atypical enteropathogenic *Escherichia coli* (EPEC) with prolonged diarrhoea. *J Med Microbiol* 53:1137–1144.
24. Cohen MB, Nataro JP, Bernstein DI, Hawkins J, Roberts N, Staat MA. 2005. Prevalence of diarrheagenic *Escherichia coli* in acute childhood enteritis: A prospective controlled study. *Journal of Pediatrics* 146:54–61.
25. Robins-Browne RM, Bordun AM, Tauschek M, Bennett-Wood VR, Russell J, Oppedisano F, Lister NA, Bettelheim KA, Fairley CK, Sinclair MI, Helland ME. 2004. *Escherichia coli* and community-acquired gastroenteritis, Melbourne, Australia. *Emerg Infect Dis* 10:1797–1805.
26. Levine MM, Edelman R. 1984. Enteropathogenic *Escherichia coli* of classic serotypes associated with infant diarrhea: epidemiology and pathogenesis. *Epidemiol Rev* 6:31–51.
27. Okeke IN. 2009. Diarrheagenic *Escherichia coli* in sub-Saharan Africa: status, uncertainties and necessities. *J Infect Dev Ctries* 3:817–842.
28. Pelicic V. 2008. Type IV pili: e pluribus unum? *Mol Microbiol* 68:827–837.

29. Häse CC, Mekalanos JJ. 1998. TcpP protein is a positive regulator of virulence gene expression in *Vibrio cholerae*. *Proc Natl Acad Sci U S A* 95:730–734.
30. Bradley DE. 1980. A function of *Pseudomonas aeruginosa* PAO polar pili: twitching motility. *Can J Microbiol* 26:146–154.
31. Collyn F, Léty MA, Nair S, Escuyer V, Ben Younes A, Simonet M, Marceau M. 2002. *Yersinia pseudotuberculosis* harbors a type IV pilus gene cluster that contributes to pathogenicity. *Infect Immun* 70:6196–6205.
32. Meyer TF, Billyard E, Haas R, Storzbach S, So M. 1984. Pilus genes of *Neisseria gonorrhoeae*: chromosomal organization and DNA sequence. *Proc Natl Acad Sci U S A* 81:6110–6114.
33. Giltner CL, Nguyen Y, Burrows LL. 2012. Type IV Pilin Proteins: Versatile Molecular Modules. *Microbiol Mol Biol Rev* 76:740.
34. Stone BJ, Abu Kwaik Y. 1998. Expression of multiple pili by *Legionella pneumophila*: identification and characterization of a type IV pilin gene and its role in adherence to mammalian and protozoan cells. *Infect Immun* 66:1768–1775.
35. Kirn TJ, Lafferty MJ, Sandoe CMP, Taylor RK. 2000. Delineation of pilin domains required for bacterial association into microcolonies and intestinal colonization by *Vibrio cholerae*. *Mol Microbiol* 35:896–910.
36. Mattick JS. 2002. Type IV pili and twitching motility. *Annu Rev Microbiol* 56:289–314.

37. Luke NR, Howlett AJ, Shao J, Campagnari AA. 2004. Expression of type IV pili by *Moraxella catarrhalis* is essential for natural competence and is affected by iron limitation. *Infect Immun* 72:6262–6270.
38. Tacket CO, Taylor RK, Losonsky G, Lim YU, Nataro JP, Kaper JB, Levine MM. 1998. Investigation of the Roles of Toxin-Coregulated Pili and Mannose-Sensitive Hemagglutinin Pili in the Pathogenesis of *Vibrio cholerae* O139 Infection. *Infect Immun* 66:692.
39. Hobbs MM, Sparling PF, Cohen MS, Shafer WM, Deal CD, Jerse AE. 2011. Experimental Gonococcal Infection in Male Volunteers: Cumulative Experience with *Neisseria gonorrhoeae* Strains FA1090 and MS11mkC. *Front Microbiol* 2.
40. Xicohtencatl-Cortes J, Monteiro-Neto V, Ledesma MA, Jordan DM, Francetic O, Kaper JB, Puente JL, Girón JA. 2007. Intestinal adherence associated with type IV pili of enterohemorrhagic *Escherichia coli* O157:H7. *J Clin Invest* 117:3519–3529.
41. Mazariego-Espinosa K, Cruz A, Ledesma MA, Ochoa SA, Xicohtencatl-Cortes J. 2010. Longus, a type IV pilus of enterotoxigenic *Escherichia coli*, is involved in adherence to intestinal epithelial cells. *J Bacteriol* 192:2791–2800.
42. Farinha MA, Conway BD, Glasier LMG, Ellert NW, Irvin RT, Sherburne R, Paranchych W. 1994. Alteration of the pilin adhesin of *Pseudomonas aeruginosa* PAO results in normal pilus biogenesis but a loss of adherence to human pneumocyte cells and decreased virulence in mice. *Infect Immun* 62:4118–4123.



43. Comolli JC, Hauser AR, Waite L, Whitchurch CB, Mattick JS, Engel JN. 1999. *Pseudomonas aeruginosa* gene products PilT and PilU are required for cytotoxicity in vitro and virulence in a mouse model of acute pneumonia. *Infect Immun* 67:3625–3630.
44. Eugène E, Hoffmann I, Pujol C, Couraud PO, Bourdoulous S, Nassif X. 2002. Microvilli-like structures are associated with the internalization of virulent capsulated *Neisseria meningitidis* into vascular endothelial cells. *J Cell Sci* 115:1231–1241.
45. Anderson MT, Dewenter L, Maier B, Steven Seifert H. 2014. Seminal plasma initiates a *Neisseria gonorrhoeae* transmission state. *mBio* 5.
46. Zhang XL, Tsui ISM, Yip CMC, Fung AWY, Wong DKH, Dai X, Yang Y, Hackett J, Morris C. 2000. *Salmonella enterica* serovar typhi uses type IVB pili to enter human intestinal epithelial cells. *Infect Immun* 68:3067–3073.
47. Ramer SW, Schoolnik GK, Wu CY, Hwang J, Schmidt SA, Bieber D. 2002. The Type IV Pilus Assembly Complex: Biogenic Interactions among the Bundle-Forming Pilus Proteins of Enteropathogenic *Escherichia coli*. *J Bacteriol* 184:3457.
48. Craig L, Forest KT, Maier B. 2019. Type IV pili: dynamics, biophysics and functional consequences. *Nat Rev Microbiol* 17:429–440.
49. Cianciotto NP. 2005. Type II secretion: a protein secretion system for all seasons. *Trends Microbiol* 13:581–588.
50. Korotkov K V, Sandkvist M, Hol WGJ. 2012. The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat Rev Microbiol* 10:336–51.
51. Sandkvist M. 2001. Biology of type II secretion. *Mol Microbiol* 40:271–283.

52. Wall D, Kaiser D. 1999. Type IV pili and cell motility. *Mol Microbiol* 32:01–10.
53. Maier B, Potter L, So M, Seifert HS, Sheetz MP. 2002. Single pilus motor forces exceed 100 pN. *Proc Natl Acad Sci U S A* 99:16012–16017.
54. Biais N, Ladoux B, Higashi D, So M, Sheetz M. 2008. Cooperative retraction of bundled type IV pili enables nanonewton force generation. *PLoS Biol* 6:907–913.
55. Maier B, Wong GCL. 2015. How Bacteria Use Type IV Pili Machinery on Surfaces. *Trends Microbiol* 23:775–788.
56. Burrows LL. 2005. Weapons of mass retraction. *Mol Microbiol* 57:878–888.
57. Berry JL, Pelicic V. 2015. Exceptionally widespread nanomachines composed of type IV pilins: the prokaryotic Swiss Army knives. *FEMS Microbiol Rev* 39:134–154.
58. Hazes B, Sastry PA, Hayakawa K, Read RJ, Irvin RT. 2000. Crystal structure of *Pseudomonas aeruginosa* PAK pilin suggests a main-chain-dominated mode of receptor binding. *J Mol Biol* 299:1005–1017.
59. Källström H, Liszewski MK, Atkinson JP, Jonsson AB. 1997. Membrane cofactor protein (MCP or CD46) is a cellular pilus receptor for pathogenic *Neisseria*. *Mol Microbiol* 25:639–647.
60. Donlan RM, Costerton JW. 2002. Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev* 15:167–193.
61. Chen I, Dubnau D. 2004. DNA uptake during bacterial transformation. *Nat Rev Microbiol* 2:241–249.

62. Wolfgang M, Van Putten JPM, Hayes SF, Koomey M. 1999. The *comP* locus of *Neisseria gonorrhoeae* encodes a type IV prepilin that is dispensable for pilus biogenesis but essential for natural transformation. *Mol Microbiol* 31:1345–1357.
63. Friedrich A, Prust C, Hartsch T, Henne A, Averhoff B. 2002. Molecular analyses of the natural transformation machinery and identification of pilus structures in the extremely thermophilic bacterium *Thermus thermophilus* strain HB27. *Appl Environ Microbiol* 68:745–755.
64. Ellison CK, Dalia TN, Vidal Ceballos A, Wang JCY, Biais N, Brun Y V., Dalia AB. 2018. Retraction of DNA-bound type IV competence pili initiates DNA uptake during natural transformation in *Vibrio cholerae*. *Nature Microbiology* 2018 3:7 3:773–780.
65. Cehovin A, Simpson PJ, McDowell MA, Brown DR, Noschese R, Pallett M, Brady J, Baldwin GS, Lea SM, Matthews SJ, Pelicic V. 2013. Specific DNA recognition mediated by a type IV pilin. *Proc Natl Acad Sci U S A* 110:3065–3070.
66. Denise R, Abby SS, Rocha EPC. 2019. Diversification of the type IV filament superfamily into machines for adhesion, protein secretion, DNA uptake, and motility. *PLoS Biol* 17:e3000390.
67. Craig L, Pique ME, Tainer JA. 2004. Type IV pilus structure and bacterial pathogenicity. *Nat Rev Microbiol* 2:363–378.
68. Craig L, Li J. 2008. Type IV pili: paradoxes in form and function. *Curr Opin Struct Biol* 18:267–277.

69. Ellison CK, Kan J, Dillard RS, Kysela DT, Ducret A, Berne C, Hampton CM, Ke Z, Wright ER, Biais N, Dalia AB, Brun Y v. 2017. Obstruction of pilus retraction stimulates bacterial surface sensing. *Science* 358:535–538.
70. Melville S, Craig L. 2013. Type IV Pili in Gram-Positive Bacteria. *Microbiology and Molecular Biology Reviews* 77:323–341.
71. Bulyha I, Schmidt C, Lenz P, Jakovljevic V, Höne A, Maier B, Hoppert M, Søgaaard-Andersen L. 2009. Regulation of the type IV pili molecular machine by dynamic localization of two motor proteins. *Mol Microbiol* 74:691–706.
72. Gold VA, Salzer R, Averhoff B, Kühlbrandt W. 2015. Structure of a type IV pilus machinery in the open and closed state. *Elife* 4.
73. Strom MS, Lory S. 1992. Kinetics and sequence specificity of processing of prepilin by Pild, the type IV leader peptidase of *Pseudomonas aeruginosa*. *J Bacteriol* 174:7345–7351.
74. Singh PK, Little J, Donnenberg MS. 2022. Landmark Discoveries and Recent Advances in Type IV Pilus Research. *Microbiology and Molecular Biology Reviews* <https://doi.org/10.1128/MMBR.00076-22>.
75. Stone KD, Zhang HZ, Carlson LK, Donnenberg MS. 1996. A cluster of fourteen genes from enteropathogenic *Escherichia coli* is sufficient for the biogenesis of a type IV pilus. *Mol Microbiol* 20:325–337.

76. Sohel I, Puente JL, Ramer SW, Bieber D, Wu CY, Schoolnik GK. 1996. Enteropathogenic *Escherichia coli*: identification of a gene cluster coding for bundle-forming pilus morphogenesis. *J Bacteriol* 178:2613–2628.
77. Vogt SL, Nevesinjac AZ, Humphries RM, Donnenberg MS, Armstrong GD, Raivio TL. 2010. The Cpx envelope stress response both facilitates and inhibits elaboration of the enteropathogenic *Escherichia coli* bundle-forming pilus. *Mol Microbiol* 76:1095–1110.
78. Anantha RP, Stone KD, Donnenberg MS. 2000. Effects of bfp mutations on biogenesis of functional enteropathogenic *Escherichia coli* type IV pili. *J Bacteriol* 182:2498–2506.
79. Anantha RP, Stone KD, Donnenberg MS. 1998. Role of BfpF, a member of the PilT family of putative nucleotide-binding proteins, in type IV pilus biogenesis and in interactions between enteropathogenic *Escherichia coli* and host cells. *Infect Immun* 66:122–131.
80. Ramboarina S, Fernandes PJ, Daniell S, Islam S, Simpson P, Frankel G, Booy F, Donnenberg MS, Matthews S. 2005. Structure of the Bundle-forming Pilus from Enteropathogenic *Escherichia coli*. *Journal of Biological Chemistry* 280:40252–40260.
81. Humphries RM, Griener TP, Vogt SL, Mulvey GL, Raivio T, Donnenberg MS, Kitov PI, Surette M, Armstrong GD. 2010. N-acetyllactosamine-induced retraction of bundle-forming pili regulates virulence-associated gene expression in enteropathogenic *Escherichia coli*. *Mol Microbiol* 76:1111–1126.
82. Blank TE, Zhong H, Bell AL, Whittam TS, Donnenberg MS. 2000. Molecular variation among type IV pilin (bfpA) genes from diverse enteropathogenic *Escherichia coli* strains. *Infect Immun* 68:7028–7038.

83. Fernandes PJ, Guo Q, Donnenberg MS. 2007. Functional consequences of sequence variation in bundlin, the enteropathogenic *Escherichia coli* type IV pilin protein. *Infect Immun* 75:4687–4696.
84. Martinez de la Peña CF, De Masi L, Nisa S, Mulvey G, Tong J, Donnenberg MS, Armstrong GD. 2015. BfpI, BfpJ, and BfpK Minor Pilins Are Important for the Function and Biogenesis of Bundle-Forming Pili Expressed by Enteropathogenic *Escherichia coli*. *J Bacteriol* 198:846–856.
85. Giltner CL, Habash M, Burrows LL. 2010. *Pseudomonas aeruginosa* minor pilins are incorporated into type IV pili. *J Mol Biol* 398:444–461.
86. Zhang HZ, Lory S, Donnenberg MS. 1994. A plasmid-encoded prepilin peptidase gene from enteropathogenic *Escherichia coli*. *J Bacteriol* 176:6885.
87. Nayak AR, Singh PK, Zhao J, Samsó M, Donnenberg MS. 2022. Cryo-EM Structure of the Type IV Pilus Extension ATPase from Enteropathogenic *Escherichia coli*. *mBio* <https://doi.org/10.1128/MBIO.02270-22>.
88. Blank TE, Donnenberg MS. 2001. Novel Topology of BfpE, a Cytoplasmic Membrane Protein Required for Type IV Fimbrial Biogenesis in Enteropathogenic *Escherichia coli*. *J Bacteriol* 183:4435–4450.
89. Yamagata A, Milgotina E, Scanlon K, Craig L, Tainer JA, Donnenberg MS. 2012. Structure of an essential type IV pilus biogenesis protein provides insights into pilus and type II secretion systems. *J Mol Biol* 419:110–124.

90. de Masi L, Szmacinski H, Schreiber W, Donnenberg MS. 2012. BfpL is essential for type IV bundle-forming pilus biogenesis and interacts with the periplasmic face of BfpC. *Microbiology (Reading)* 158:2515–2526.
91. Lieberman JA, Frost NA, Hoppert M, Fernandes Paula J, Vogt SL, Raivio TL, Blanpied TA, Donnenberg MS. 2012. Outer membrane targeting, ultrastructure, and single molecule localization of the enteropathogenic *Escherichia coli* type IV pilus secretin BfpB. *J Bacteriol* 194:1646–1658.
92. Lieberman JA, Petro CD, Thomas S, Yang A, Donnenberg MS. 2015. Type IV Pilus Secretins Have Extracellular C Termini. *mBio* 6.
93. Daniel A, Singh A, Crowther LJ, Fernandes PJ, Schreiber W, Donnenberg MS. 2006. Interaction and localization studies of enteropathogenic *Escherichia coli* type IV bundle-forming pilus outer membrane components. *Microbiology (N Y)* 152:2405–2420.
94. Bayan N, Guilvout I, Pugsley AP. 2006. Secretins take shape. *Mol Microbiol* 60:1–4.
95. Tosi T, Estrozi LF, Job V, Guilvout I, Pugsley AP, Schoehn G, Dessen A. 2014. Structural similarity of secretins from type II and type III secretion systems. *Structure* 22:1348–1355.
96. Korotkov K V., Gonen T, Hol WGJ. 2011. Secretins: dynamic channels for protein transport across membranes. *Trends Biochem Sci* 36:433–443.
97. D’Imprima E, Salzer R, Bhaskara RM, Sánchez R, Rose I, Kirchner L, Hummer G, Kühlbrandt W, Vonck J, Averhoff B. 2017. Cryo-EM structure of the bifunctional secretin complex of *Thermus thermophilus*. *Elife* 6.

98. Hardie KR, Lory S, Pugsley AP. 1996. Insertion of an outer membrane protein in *Escherichia coli* requires a chaperone-like protein. *EMBO J* 15:978.
99. Linderoth NA, Model P, Russel M. 1996. Essential role of a sodium dodecyl sulfate-resistant protein IV multimer in assembly-export of filamentous phage. *J Bacteriol* 178:1962.
100. Nouwen N, Stahlberg H, Pugsley AP, Engel A. 2000. Domain structure of secretin PulD revealed by limited proteolysis and electron microscopy. *EMBO J* 19:2229.
101. Majewski DD, Worrall LJ, Strynadka NC. 2018. Secretins revealed: structural insights into the giant gated outer membrane portals of bacteria. *Curr Opin Struct Biol* 51:61–72.
102. Weaver SJ, Ortega DR, Sazinsky MH, Dalia TN, Dalia AB, Jensen GJ. 2020. CryoEM structure of the type IVa pilus secretin required for natural competence in *Vibrio cholerae*. *Nat Commun* 11.
103. Hay ID, Belousoff MJ, Dunstan RA, Bamert RS, Lithgow T. 2018. Structure and Membrane Topography of the *Vibrio*-Type Secretin Complex from the Type 2 Secretion System of Enteropathogenic *Escherichia coli* <https://doi.org/10.1128/JB.00702-17>.
104. Hospenthal MK, Costa TRD, Waksman G. 2017. A comprehensive guide to pilus biogenesis in Gram-negative bacteria. *Nat Rev Microbiol* 15:365–379.
105. Koo J, Burrows LL, Lynne Howell P. 2012. Decoding the roles of pilotins and accessory proteins in secretin escort services. *FEMS Microbiol Lett* 328:1–12.
106. Spreter T, Yip CK, Sanowar S, André I, Kimbrough TG, Vuckovic M, Pfuetzner RA, Deng W, Yu AC, Finlay BB, Baker D, Miller SI, Strynadka NCJ. 2009. A conserved



- structural motif mediates formation of the periplasmic rings in the type III secretion system. *Nat Struct Mol Biol* 16:468–476.
107. Noinaj N, Gumbart JC, Buchanan SK. 2017. The  $\beta$ -barrel assembly machinery in motion. *Nat Rev Microbiol* 15:197–204.
108. Korotkov K V., Pardon E, Steyaert J, Hol WGJ. 2009. Crystal Structure of the N-Terminal Domain of the Secretin GspD from ETEC Determined with the Assistance of a Nanobody. *Structure* 17:255–265.
109. Filloux A, Voulhoux R. 2018. Multiple Structures Disclose the Secretins' Secrets. *J Bacteriol* 200:e00702-17.
110. Li C, Wen A, Shen B, Lu J, Huang Y, Chang Y. 2011. FastCloning: a highly simplified, purification-free, sequence-and ligation-independent PCR cloning method <https://doi.org/10.1186/1472-6750-11-92>.
111. Kampjut D, Steiner J, Sazanov LA. 2021. Cryo-EM grid optimization for membrane proteins. *iScience* 24.
112. Hu J, Worrall LJ, Vuckovic M, Hong C, Deng W, Atkinson CE, Brett Finlay B, Yu Z, Strynadka NCJ. 2019. T3S injectisome needle complex structures in four distinct states reveal the basis of membrane coupling and assembly. *Nat Microbiol* 4:2010–2019.
113. Salzer R, D'Imprima E, Gold VAM, Rose I, Drechsler M, Vonck J, Averhoff B. 2016. Topology and Structure/Function Correlation of Ring- and Gate-forming Domains in the Dynamic Secretin Complex of *Thermus thermophilus*. *J Biol Chem* 291:14448–14456.

114. Conners R, McLaren M, Łapińska U, Sanders K, Stone MRL, Blaskovich MAT, Pagliara S, Daum B, Rakonjac J, Gold VAM. 2021. CryoEM structure of the outer membrane secretin channel pIV from the f1 filamentous bacteriophage. *Nature Communications* 2021 12:1 12:1–14.
115. Dawson N, Sillitoe I, Marsden RL, Orengo CA. 2017. The Classification of Protein Domains. *Methods Mol Biol* 1525:137–164.
116. Kundlacz T, Bender J, Schmidt C. 2021. Effects of non-ionic and zwitterionic detergents on soluble proteins during native mass spectrometry experiments. *Int J Mass Spectrom* 468:116652.
117. Schreiber W, Stone KD, Strong MA, DeTolla LJ, Hoppert M, Sonnenberg MS. 2002. BfpU, a soluble protein essential for type IV pilus biogenesis in enteropathogenic *Escherichia coli*. *Microbiology (Reading)* 148:2507–2518.
118. Daniel A, Singh A, Crowther LJ, Fernandes PJ, Schreiber W, Sonnenberg MS. 2006. Interaction and localization studies of enteropathogenic *Escherichia coli* type IV bundle-forming pilus outer membrane components. *Microbiology (N Y)* 152:2405–2420.
119. Ayers M, Sampaleanu LM, Tammam S, Koo J, Harvey H, Howell PL, Burrows LL. 2009. PilM/N/O/P proteins form an inner membrane complex that affects the stability of the *Pseudomonas aeruginosa* type IV pilus secretin. *J Mol Biol* 394:128–142.
120. Tammam S, Sampaleanu LM, Koo J, Manoharan K, Daubaras M, Burrows LL, Howell PL. 2013. PilMNOPQ from the *Pseudomonas aeruginosa* Type IV Pilus System Form a Transenvelope Protein Interaction Network That Interacts with Pila. *J Bacteriol* 195:2126.

121. Friedrich C, Bulyha I, Søggaard-Andersen L. 2014. Outside-in assembly pathway of the type IV pilus system in *myxococcus xanthus*. *J Bacteriol* 196:378–390.
122. Tammam S, Sampaleanu LM, Koo J, Sundaram P, Ayers M, Andrew Chong P, Forman-Kay JD, Burrows LL, Howell PL. 2011. Characterization of the PilN, PilO and PilP type IVa pilus subcomplex. *Mol Microbiol* 82:1496–1514.
123. Levine MM, Nalin DR, Hornick RB, Bergquist EJ, Waterman DH, Young CR, Sotman S, Rowe B. 1978. *Escherichia coli* strains that cause diarrhoea but do not produce heat-labile or heat-stable enterotoxins and are non-invasive. *Lancet* 1:1119–1122.
124. Hobson N, Price NL, Ward JD, Raivio TL. 2008. Generation of a restriction minus enteropathogenic *Escherichia coli* E2348/69 strain that is efficiently transformed with large, low copy plasmids. *BMC Microbiol* 8.
125. Amann E, Ochs B, Abel KJ. 1988. Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* 69:301–315.
126. Hu Y, Romão E, Vertommen D, Vincke C, Morales-Yáñez F, Gutiérrez C, Liu C, Muyldermans S. 2017. Generation of Nanobodies against SlyD and development of tools to eliminate this bacterial contaminant from recombinant proteins. *Protein Expr Purif* 137:64–76.
127. Bessette PH, Åslund F, Beckwith J, Georgiou G. 1999. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc Natl Acad Sci U S A* 96:13703–13708.

128. Levy R, Weiss R, Chen G, Iverson BL, Georgiou G. 2001. Production of correctly folded Fab antibody fragment in the cytoplasm of *Escherichia coli* *trxB* *gor* mutants via the coexpression of molecular chaperones. *Protein Expr Purif* 23:338–347.
129. Ghosh K, Dill K. 2010. Cellular Proteomes Have Broad Distributions of Protein Stability. *Biophys J* 99:3996.
130. Bhagavan NV. 2002. Three-Dimensional Structure of Proteins. *Med Biochem* 51–65.
131. Denoncin K, Collet JF. 2013. Disulfide Bond Formation in the Bacterial Periplasm: Major Achievements and Challenges Ahead. *Antioxid Redox Signal* 19:63.
132. Sun M an, Wang Y, Zhang Q, Xia Y, Ge W, Guo D. 2017. Prediction of reversible disulfide based on features from local structural signatures. *BMC Genomics* 18.
133. Golovanov AP, Balasingham S, Tzitzilonis C, Goult BT, Lian LY, Homberset H, Tønjum T, Derrick JP. 2006. The solution structure of a domain from the *Neisseria meningitidis* lipoprotein PilP reveals a new beta-sandwich fold. *J Mol Biol* 364:186–195.
134. Korotkov K V., Johnson TL, Jobling MG, Pruneda J, Pardon E, Héroux A, Turley S, Steyaert J, Holmes RK, Sandkvist M, Hol WGJ. 2011. Structural and functional studies on the interaction of GspC and GspD in the type II secretion system. *PLoS Pathog* 7.
135. Schuck P. 1997. Use of surface plasmon resonance to probe the equilibrium and dynamic aspects of interactions between biological macromolecules. *Annu Rev Biophys Biomol Struct* 26:541–566.

136. Tiwari PB, Bencheqroun C, Lemus M, Shaw T, Kouassi-Brou M, Alaoui A, Üren A. 2021. SPRD: a surface plasmon resonance database of common factors for better experimental planning. *BMC Mol Cell Biol* 22:1–8.
137. Hu J, Lipowsky R, Weikl TR. 2013. Binding constants of membrane-anchored receptors and ligands depend strongly on the nanoscale roughness of membranes. *Proc Natl Acad Sci U S A* 110:15283–15288.
138. Good NE, Winget GD, Winter W, Connolly TN, Izawa S, Singh RMM. 1966. Hydrogen ion buffers for biological research. *Biochemistry* 5:467–477.
139. de Mol NJ, Fischer MJE. 2010. *Surface Plasmon Resonance: Methods and Protocols*. *Life Sci* 255.
140. Vasconcelos MTSD, Azenha MAGO, Lage OM. 1996. Electrochemical evidence of surfactant activity of the hepes pH buffer which may have implications on trace metal availability to cultures in vitro. *Anal Biochem* 241:248–253.
141. Topală T, Bodoki A, Oprean L, Oprean R. 2014. Bovine Serum Albumin Interactions with Metal Complexes. *Clujul Medical* 87:215.
142. Baler K, Martin OA, Carignano MA, Ameer GA, Vila JA, Szleifer I. 2014. Electrostatic Unfolding and Interactions of Albumin Driven by pH Changes: A Molecular Dynamics Study. *J Phys Chem B* 118:921.
143. Chen J, Noble AJ, Kang JY, Darst SA. 2019. Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO. *J Struct Biol X* 1:100005.

144. Zi Tan Y, Baldwin PR, Davis JH, Williamson JR, Potter CS, Carragher B, Lyumkis D. 2017. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature Methods* 2017 14:8 14:793–796.
145. Silva YR de O, Contreras-Martel C, Macheboeuf P, Dessen A. 2020. Bacterial secretins: Mechanisms of assembly and membrane targeting. *Protein Sci* 29:893–904.
146. Schmidt SA, Bieber D, Ramer SW, Hwang J, Wu CY, Schoolnik G. 2001. Structure-function analysis of BfpB, a secretin-like protein encoded by the bundle-forming-pilus operon of enteropathogenic *Escherichia coli*. *J Bacteriol* 183:4848–4859.
147. Strong MA, Stone KD, Donnenberg MS, Schreiber W, DeTolla, Jr LJ, Hoppert M. 2015. BfpU, a soluble protein essential for type IV pilus biogenesis in enteropathogenic *Escherichia coli*. *Microbiology (N Y)* 148:2507–2518.