Obtaining Structural Insights on Bacterial Protein Complexes Using Time-Resolved Hydrogen-Deuterium Exchange Mass Spectrometry

CRISTINA LENTO

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN CHEMISTRY
YORK UNIVERSITY
TORONTO, ONTARIO

AUGUST 2015

ABSTRACT

Persistent infections by *Pseudomonas aeruginosa* are initiated by interaction of a type IV pilus (T4P) with receptors on the mucosal cells of susceptible hosts. Here, we examine the structural changes occurring between the monomeric and dimeric states of ΔK122 using time-resolved electrospray ionization hydrogen-deuterium exchange mass spectrometry (TRESI-HDX-MS). Based on levels of deuterium uptake, the N-terminal α-helix and the loop connecting the second and third strands of the anti-parallel β-sheet contribute significantly to pilin dimerization. Conversely, the antiparallel β-sheet and αβ loop region exhibit increased flexibility, while the receptor binding domain retains a rigid conformation in the equilibrium state. Additionally, *Escherichia coli* are able to adapt to changing environmental conditions and develop antibiotic resistance through a process called F-plasmid conjugation, carried out through a type IV secretion system (T4SS). The F-T4SS protein TraF is of particular interest due to its involvement in pilus assembly to mediate the transfer of DNA. Dynamic analysis of a GST-TraF construct through TRESI-HDX-MS was performed to gain further insights on its structure. These studies have revealed that the C-terminal region predicted to contain the thioredoxin-like domain is quite structured compared to the more solvent accessible N-terminal region predicted to form a protein-protein interaction with companion T4SS protein TraH. Structural analysis of a GST-TraF construct is on-going to further characterize the regions responsible for protein-protein interaction and the elucidation of its three-dimensional structure.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Dr. Gerald Audette and Dr. Derek Wilson for giving me the opportunity to work in their labs and for their continual guidance and support throughout this project. The experience gained working with their research groups has been instrumental for the development of both my scientific knowledge and career. A special thanks also goes to my supervisory committee members, which includes Dr. Dasantila Golemi-Kotra and Dr. Andrew Eckford.

I would also like to thank the various lab members who have been there to help me both within and outside of the lab: Shaolong Zhu, Peter Liuni, Agnesa Shala, Ayat Yaseen, Fatima Alam-Portillo, Preet Gill, and Kerene Brown. You have all made spending time in the lab an unforgettable experience, and were never short of advice and support through the highs and lows of research. The experiences and friendships made throughout these years are those I will look back upon with fond memories. Above all, a special thanks goes to my mom and dad for their undying love, care and support, without which none of this would have been possible.

LIST OF ABBREVIATIONS

Amp – ampicillin
CIEX – cation exchange chromatography
CID – collision-induced dissociation
CNT – carbon nanotube
CRM – charged residue model
CV – column volume
EDTA – ethylenediaminetetraacetic acid
ESI – electrospray ionization
F-plasmid – fertility factor or sex factor plasmid
GST – glutathione S-transferase
HDX – hydrogen-deuterium exchange
i.d. – inner diameter
IMS – ion-mobility spectrometry
IPTG – isopropyl-B-D-thiogalactopyranoside
kDa - kilodaltons
LB – Luria Bertani
MBP – maltose binding protein
Mdr – multi-drug resistant
MPD – 2-methyl-2,4-pentanediol
Mpf – mating pair formation
MS – mass spectrometry
MWCO – molecular weight cut off
MWM – molecular weight marker
m/z – mass to charge ratio
OD_{600} – optical density at $\lambda = 600$
o.d. – outer diameter
pI – isoelectric point
PMMA – polymethyl methacrylate
PMSF – phenylmethylsulphonyl fluoride
PNT – protein nanotube
Q – quadrupole mass filter
RBD – receptor binding domain
SDS-PAGE – sodium dodecyl sulphate polyacrylamide gel electrophoresis
T4P – type IV pilus
T4SS – type IV secretion system
TOF – time-of-flight
tra – transfer
TRESI – time-resolved electrospray ionization

TABLE OF CONTENTS

CHAPTERS

LIST OF TABLES

CHAPTER 2

CHAPTER 4

CHAPTER 5

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

Chapter 1: Introduction

1.1. Multi-Drug Resistant Bacteria and Bacterial Infections

Bacteria are resilient organisms that have existed long before the rise of humans with their evolution being a well-documented and remarkable phenomenon. While advantageous to the organism, the adaptability of bacteria poses significant challenges to human health and allows for the rapid spread of infectious diseases, which make up 25% of annual deaths worldwide[1] – and this number is only predicted to get worse with the emergence of new pathogens and the increase of antibiotic resistance. It is therefore of no great surprise that research in understanding the intricate mechanisms behind bacterial pathogenesis is a critical field of research in the twenty-first century.

1.1.1. Bacterial Gene Transfer

The exchange of genetic material from a donor to a recipient cell is responsible for adaptation to changes in the environment and bacterial evolution. Most importantly to human health, it allows for the rapid spread of antibiotic resistant genes and the development of multi-drug resistant (Mdr) bacteria.[2] The physical process for the transfer of genetic material between cells occurs through one of three pathways: 1) uptake of DNA from the environment by transformation, 2) direct contact between cells by conjugation and 3) bacteriophage transduction (Figure 1.1).[2]

Transformation was the first type of horizontal gene transfer discovered,[2] whereby the intercellular transfer of non-specific DNA occurs (Figure 1.1a). It involves the direct uptake of naked DNA from the surroundings by competent bacterial cells through a permeable membrane. It is important to note that the proteins mediating this

process are chromosomally encoded and found in naturally transformable bacteria. In contrast, conjugation (Figure 1.1b) requires direct cell-to-cell connection whereby independent replicating elements in the form of plasmids encode for proteins that facilitate their own transfer. Finally, transduction is defined as DNA transfer mediated by independently replicating bacterial viruses known as bacteriophages (Figure 1.1c).[2] A portion of the research detailed herein (detailed in Chapter 4) has focussed on one of the bacterial secretion systems responsible for bacterial conjugation.

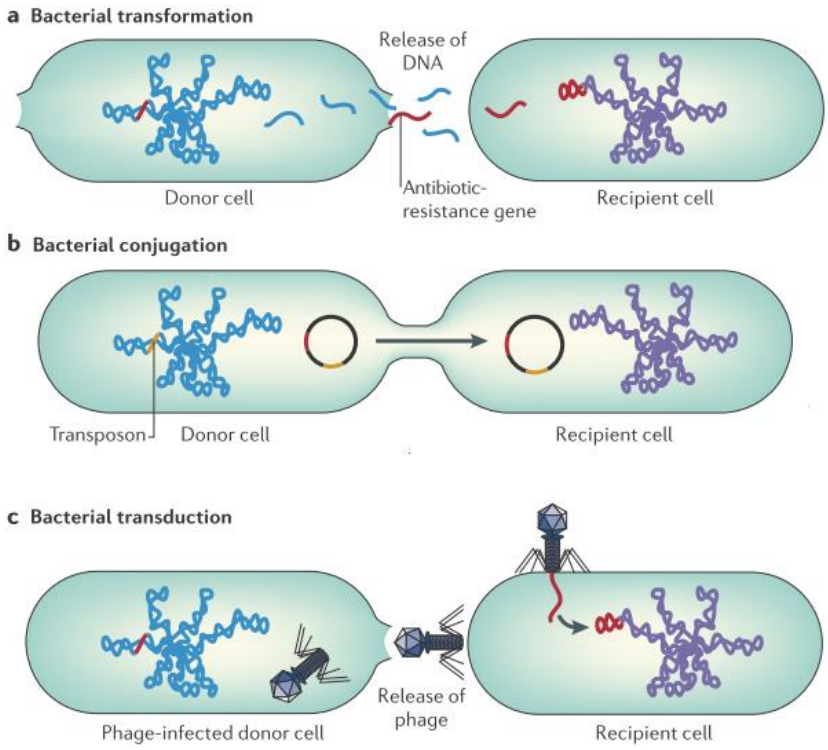


Figure 1.1: Horizontal gene transfer between bacteria. a) In transformation, the donor cell releases antibiotic-resistance genes into the extracellular matrix, which are taken up by the permeable membrane of the recipient. b) Conjugation occurs through direct cell contact between the donor and recipient cell. The plasmid that is transferred contains independent replicating elements and encodes for proteins mediating its own transfer. c) In transduction, DNA transfer is mediated by bacteriophages allowing for integration of antibiotic-resistance genes into the chromosome of the recipient cell (lysogeny). Adapted from Furuya and Lowy, 2006.[3]

1.2. Bacterial Secretion Systems

Bacteria have evolved to respond to their environmental conditions through the emergence of secretion systems that affect the transfer of macromolecules, DNA, or toxins across their cellular membranes. Different secretion systems are composed of specific genes and virulence factors and are numbered types I through VIII.[2,4] While type I and type V secretion systems are relatively simple and composed of 1 to 3 component proteins, types II, III and IV are much more complex. The type IV secretion system (T4SS) is the most complex and is composed of between 8 to 20 core proteins.[5] For these more complex secretion systems, individual studies on component proteins have started to piece together how proteins in the cytoplasm, inner membrane, periplasmic space, outer membrane and extracellular space come together in order to drive pilus assembly and DNA transfer.[5]

1.2.1. Type IV Secretion System (T4SS)

Bioinformatics has shown that T4SSs are evolutionarily related to bacterial conjugation systems, which are found in most gram-negative and gram-positive bacteria.[6-8] As mentioned, bacterial conjugation is the unidirectional transfer of single-stranded DNA between bacterial cells by a direct cell-to-cell connection, and is the main driving force for infection and the spread of antibiotic resistance genes. Another class of T4SSs are those involved in DNA release and uptake from the extracellular space without the need for a cell-to-cell connection or specific target cell. This relatively recently discovered subclass is composed of two DNA-uptake systems (found in *Campylobacter jejuni* and *Helicobactor pylori*) and one DNA-release system (found in *Neisseria*

gonorrhoeae).[9] T4SSs are also used for the transfer of DNA or protein substrates, known as effector or virulence factors, to target cells during the infection process.[10] Well-studied examples of such systems include the oncoprotein CagA in *H. pylori*, the pertussis toxin of *Bordetella pertussis*, and the oncogenic T_i plasmid of *Agrobacterium tumefaciens* (Figure 1.2).[9]

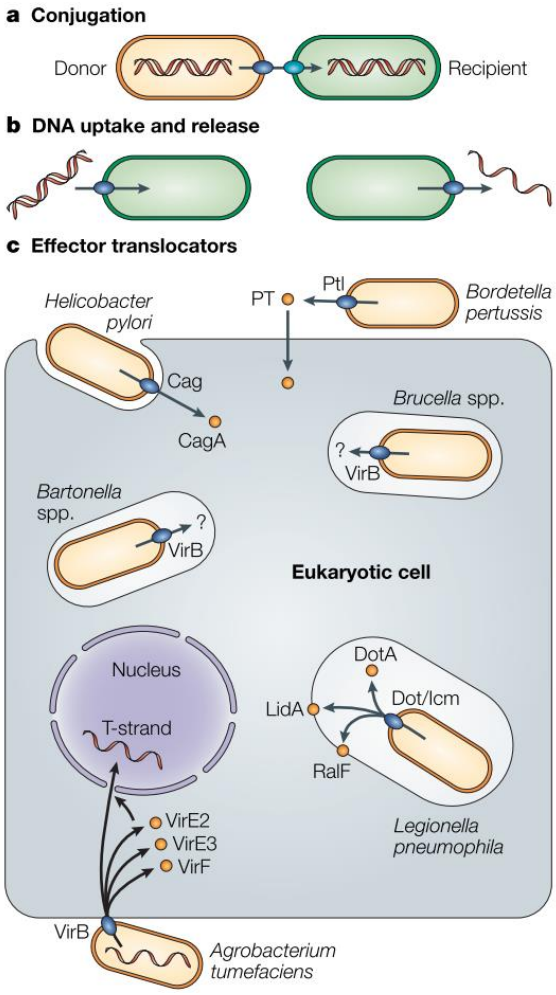


Figure 1.2: Schematic representation of the different T4SS mechanisms.
a) Conjugation systems deliver DNA to recipient cells via a direct cell-to-cell connection. b) DNA uptake and release systems exchange genetic material with the extracellular space. c) Effector translocators deliver DNA or protein substrates to eukaryotic cells during the infection process. Adapted from Cascales and Christie, 2003.[9]

Although they deliver different factors, many of the T4SSs are closely related to one another forming a similar translocation apparatus. The most studied T4SS is the tumour inducing (T_i) system found in *A. tumefaciens* consisting of 12 core proteins named VirB1-VirB11 and VirD4, and the *Escherichia coli* conjugative systems encoded by F, R1, and pKM101 plasmids with up to 20 conserved proteins.[6,8,11] Detailed structural studies of conjugative T4SSs are only recently becoming available despite the wealth of biochemical knowledge.[5,7-9,11,12] For example, in the T4SS encoded by the pKM101 plasmid, the core complex of TraF-N-O was studied by cryo-electron microscopy and X-ray crystallography. Together, the trimer was found to be present in 14 copies forming a 1.1 MDa complex enabling substrate passage.[13] It is believed that F-pilus assembly likely relies on a core complex that is structurally similar to that solved for the pKM101 system.

1.2.1.1. A Closer Look at the F-Plasmid Conjugative System

In 1946, Lederberg and Tatum first described the phenomenon of conjugation using *E. coli* as a model system.[14] *E. coli* is the most common pathogen affecting hospital patients causing various illnesses such as pneumonia, meningitis, and urinary tract infections often leading to death.[15] Bacterial conjugation is a plasmid-driven process[2,16,17] initiated by the interaction of the conjugative pilus of a donor cell with the cell surface of a recipient.[18] The extension and retraction of the pilus is a dynamic process[19] that actively identifies potential recipient cells. After attaching to a recipient, the pilus is retracted by the conjugative T4SS, resulting in the intimate association between donor and recipient cell, known as a mating pair.[20-22] Plasmid DNA is then transferred through a membrane-

associated supramolecular structure known as the mating pair formation (Mpf) complex, composed of T4SS proteins. Essentially, conjugative DNA transfer can be thought of as a protein secretion system that happens to also transport DNA.

While a class of conjugative plasmids are responsible for driving the process of bacterial gene transfer in *E. coli*, the first of these plasmids discovered and consequently the most well studied is the fertility factor or sex factor (F). *E. coli* cells containing the F-plasmid (F^+) act as donors while those lacking the F-plasmid (F^-) are recipients. Following stabilization of the mating pair by the Mpf complex, ATP-dependent transfer of the ssDNA occurs and the recipient cell is released. A newly synthesized complementary strand is produced with the recipient cell now becoming a potential donor (Figure 1.3).[18]

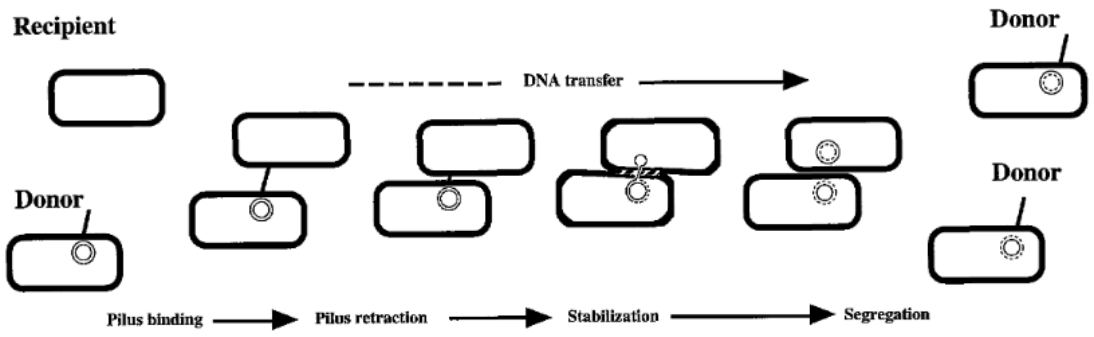


Figure 1.3: The stages of F mediated plasmid transfer. The pilus tip from the F^+ donor cells extends and attaches to the F^- recipient cell. The retraction (depolymerisation) of the F pilin subunit draws the donor and recipient in close proximity. The F-plasmid DNA replicates and the newly synthesized copy is transferred to a recipient cell with a copy always remaining in the donor cell. Adapted from Firth et al. 1996.[23]

The F-plasmid is approximately 100 kb in size, with a portion responsible for encoding proteins making up the T4SS. In fact, transfer (*tra* and *trb*) proteins making up the T4SS are encoded for by the transfer operon region spanning 33 kb of the F-plasmid (Figure 1.4).[24] Genetic studies have shown that the majority of F-*tra* proteins are

involved in pilus biogenesis, with other proteins involved in surface exclusion, mating-aggregate stabilization, regulation and DNA metabolism.[23] The F-pilus is composed primarily of the 121 amino acid *traA* gene, with processing of the precursor pilin subunit carried out by *traX* and *traQ*.[23] The pilus was observed to exhibit two subunit packing schemes using cryo-electron microscopy corresponding to a left-handed seven-start helix and a right-handed four-start helix.[25]

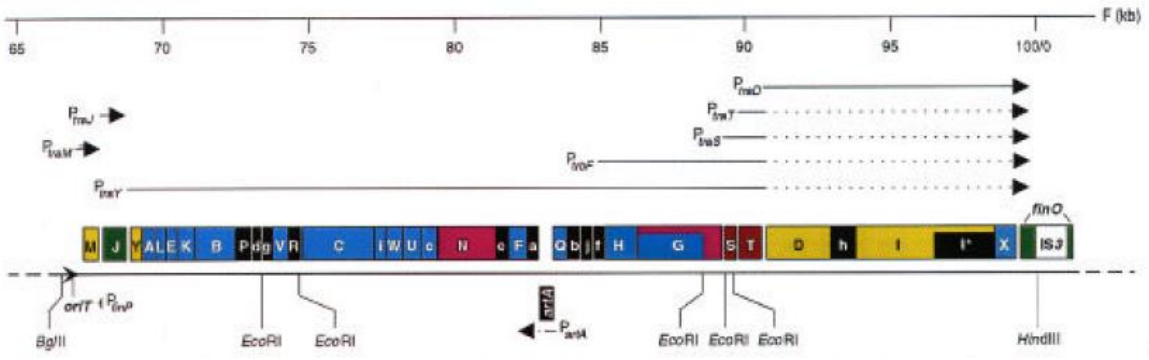


Figure 1.4: Physical and genetic map of the F-plasmid transfer region. Kilobase coordinates are indicated at the top, with transcription initiation of the promoter regions depicted with arrows. The functional class of each gene is depicted as coloured boxes; blue = pilus biogenesis, red = surface exclusion, magenta = mating-aggregate stabilization, green = regulation, yellow = DNA metabolism, black = unknown/nonessential. Capital and lowercase letters are used to label *tra* and *trb* genes, respectively. Adapted from Firth et al. 1996.[23]

T4SSs are membrane associated transport complexes spanning the inner and outer membrane of *E. coli* responsible for horizontal DNA transfer to target cells (Figure 1.5). There are a total of 9 hallmark proteins making up the F-plasmid T4SS and they are TraF, TraG (C-terminal domain), TraH, TraN, TraU, TraW, TrbB, TrbC, and TrbI.[8,26] Mutations in *traL*, -E, -K, -B, -V, -C, -W, -U, -F, -H, -G, *trbC* and *trbI* have been shown to create an accumulation of the pilin protein in the membrane, and therefore are required for pilus assembly.[27] Within the aforementioned proteins, an interaction group responsible for F-pilus extension is composed of TraF, TraH, TrbI, TraU, TraW, and

TrbB.[28] The mating pair is stabilized through interaction between TraN and TraG,[29] with the F-plasmid DNA being transferred from donor to recipient through a TraB-K-V core complex by the ATP-dependent coupling protein TraD.[2]

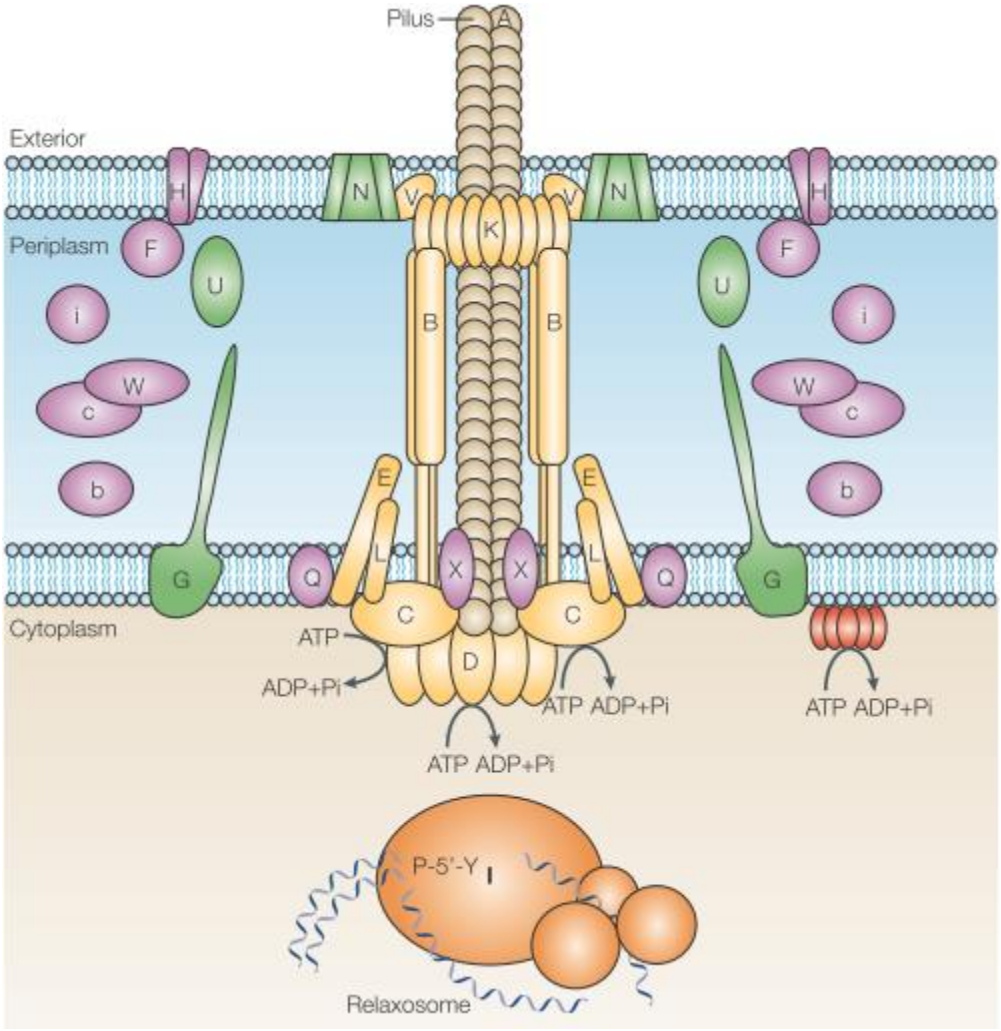


Figure 1.5: Signature T4SS proteins responsible for F-plasmid conjugation in gram negative bacteria. The double membrane spanning complex is composed of transfer (*tra* and *trb*) proteins responsible for mediating the passage of DNA between bacterial cells. Tra proteins are labelled in uppercase letters and Trb proteins are labelled in lowercase letters. The mating pair is stabilized through interaction between TraN and TraG, and the F-plasmid DNA is transferred from donor to recipient through a TraB-K-V core complex by the ATP-dependent coupling protein TraD. The pilus is composed of repeating TraA subunits processed by TraX and TraQ. An interaction group responsible for F-pilus extension is composed of TrbI, TraU, TraW, and TrbB. Proteins of interest TraF and TraH are located within this group with the processed periplasmic forms being 26 kDa and 48 kDa in size, respectively. Adapted from Frost et al. 2005.[2]

1.2.1.2. The TraF and TraH Interaction

TraF, one of six hallmark proteins of the F-like T4SS, is a 25.9 kDa protein linked with both pilus assembly and F-plasmid transfer. TraF contains a predicted thioredoxin domain but does not contain the characteristic CxxC active site of other thioredoxins;[30] there is currently no three-dimensional structure of TraF available, although its crystallization has been reported.[31] It is predicted that the thioredoxin domain is made up of at least three α-helices flanked by a four-stranded antiparallel β-sheet.[32] It is expected that TraF acts as a chaperone, assisting in other conjugative proteins such as TraH, -U and -N to achieve their correct conformation. A predicted structure of the C-terminal (a.a. 132 – 247) region of TraF is shown in Figure 1.6.

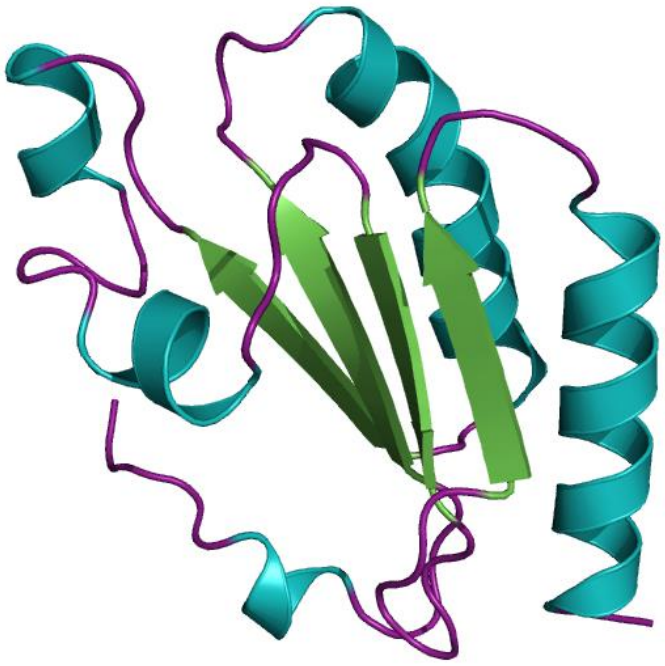


Figure 1.6: Predicted structure of the C-terminal TraF region. The structure of the C-terminal domain of TraF was predicted using the Phyre2[33] server, and is a homolog of TraF spanning amino acid residues 132-247. The image was rendered using Pymol.[34]

TraH is 47.8 kDa in size and one of three cysteine rich T4SS proteins also required for pilus assembly. TraH spans the periplasmic and outer membrane space and contains 6 cysteine residues suggesting potential disulfide bond isomerase activity.[35] In addition, it contains a C-terminal coiled-coil domain contributing to its oligomerization and interaction with other T4SS proteins.[8] There is currently no solved structure for the protein, as well as no solved homolog of high sequence similarity. Membrane hydrophobicity prediction software suggests that TraH is likely rich in amphipathic helices.

The aforementioned F-plasmid proteins TraF and TraH have been identified as forming an interaction group allowing for extension of the F pilus. This interaction group also includes proteins TrbI, TraU, TraW, and TrbB which are all predicted to be periplasmic with the exception of TrbI. Although located within the inner membrane, TrbI has a large periplasmic domain allowing for its involvement in the interaction group. It was also found that TraH is the most connected node with direct attachment to TraF, TraU and TrbI (Figure 1.7a).[28] Furthermore, mutations in TraF and TraH have been shown to abolish the ability of F^+ cells to form extended pili structures observed through electron microscopy. Yeast-two hybrid analysis has shown that the C-terminus of TraH contains a binding site responsible for its interaction with TraF.[26] It has been proposed that the region of the TraF protein responsible for interaction with TraH spans the segment 64-88 while TraH contains a TraF binding site between amino acid residues 315-458 on the C-terminus (Figure 1.7b).[28]

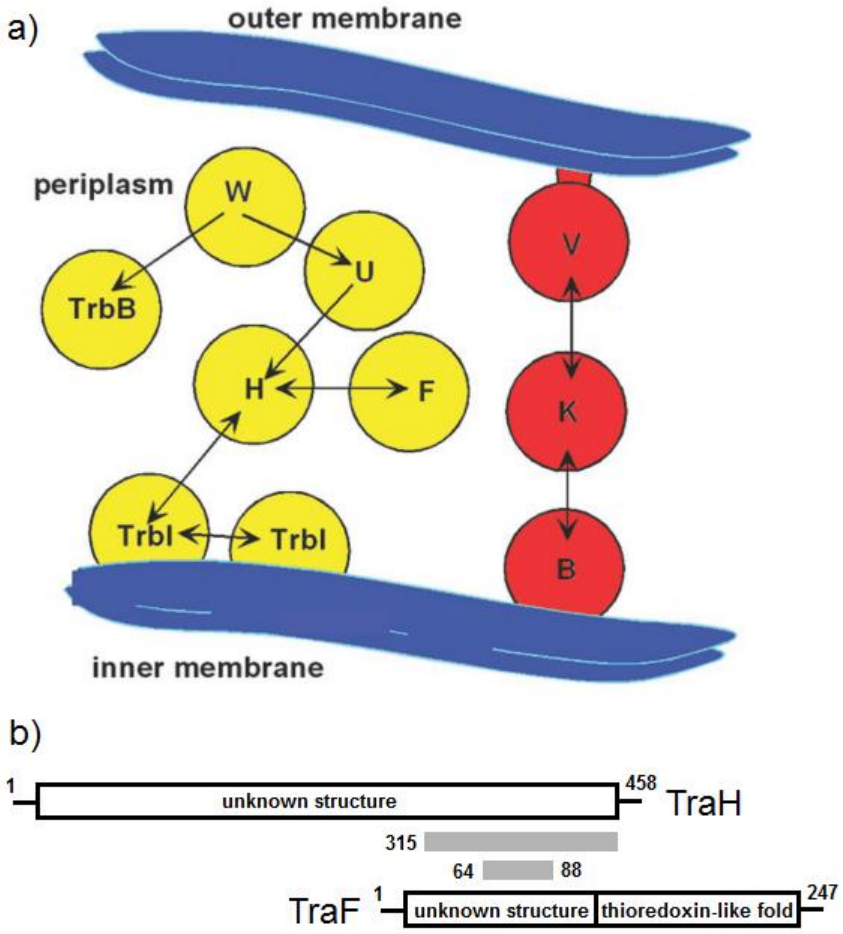


Figure 1.7: Protein interaction group of Tra proteins required for the formation of F pili. a) Arrows connect proteins shown to interact with one another through yeast-two hybrid analysis, with arrowheads pointing to the prey. Adapted from Harris and Silverman, 2004.[28] b) A prediction[28] of the segments of TraH and TraF that are proposed to interact with one another.

1.2.2. Type II Secretion System (T2SS)

First discovered in the 1980s in the genus *Klebsiella oxytoca*,[36] the type II secretion system (T2SS) is responsible for the specific transfer of folded periplasmic proteins in gram-negative bacteria. These secreted proteins released into the extracellular space or displayed on the cell surface allow the bacterium to survive and adapt to various environments ranging from deep-sea waters to animal and plant tissues.[37] The main

function of these exoproteins is for nutrient acquisition, as the majority are classed as hydrolytic enzymes that degrade biopolymers. In addition to these, T2SSs are responsible for promoting the secretion of toxins, such as the cholera toxin – a multimeric protein complex secreted by the human pathogen *Vibrio cholerae* causing serious and life-threating infections.[38] Up to 15 genes are responsible for encoding T2SS machinery in gram-negative bacteria, the components of which are currently the subject of many structural and functional studies.[37] A current understanding of the proteins involved in the T2SS of gram negative bacteria is shown in Figure 1.8.

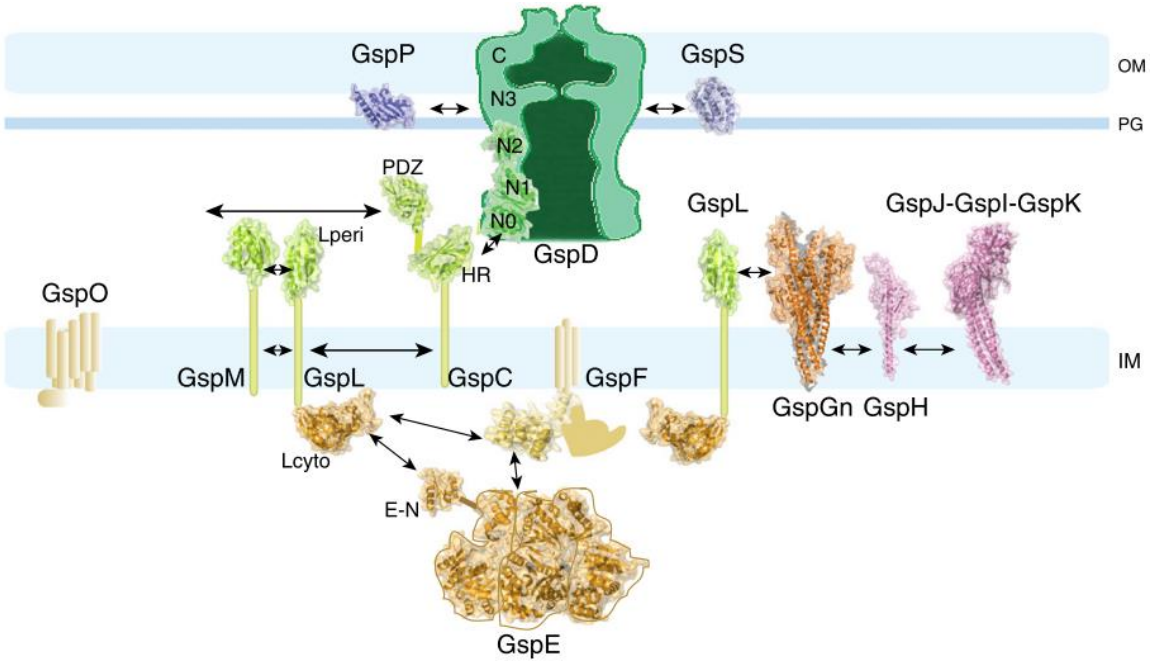


Figure 1.8: Localization and interactions of the T2SS in gram negative bacteria. Components of known structure are shown in cartoon representation, with proteins and parts of unknown structure represented as cylinders. Established interactions are depicted with double-ended arrows. Adapted from Mangayarkarasi and Francetic, 2014.[37]

T2SS are highly similar in composition and structure to systems that produce archaeal pili, flagella, as well as the bacterial type 4 pilus (T4P), suggesting a common

ancestral origin. The T4P is a surface-exposed nanofiber composed of multiple protein subunits that perform a range of bacterial functions.

1.2.2.1. The Type IV Pilus (T4P)

Opportunistic infections by *Pseudomonas aeruginosa* are prevalent in patients with compromised immune systems including those recovering from burn wounds[39-41] and organ transplants,[42,43] as well as in individuals suffering from cystic fibrosis,[44,45] acute leukemia[46] and HIV.[47] These persistent infections are initiated by interaction of a type IV pilus (T4P) with receptors on the mucosal cells of susceptible hosts.[48-53] For example, in *P. aeruginosa*, fibrous T4P extend from the poles of the bacterium and attach to glycosphingolipids asialo-GM_1 and asialo-GM_2 present on epithelial cells (Figure 1.9). In addition, *P. aeruginosa* accounts for 10% of all hospital infections including lung, urinary tract and blood infections.[54] T4P are also responsible for a variety of bacterial processes including surface motility,[52,55-58] microcolony and biofilm formation,[52,59-62] cell-host adhesion,[53] cell signalling[63] and DNA uptake (transformation).[64,65] Indeed, T4P are important structures found across a wide range of gram-negative and gram-positive bacteria, and disruption of the pilus leads to decreased bacterial virulence in many gram-negative pathogens.[40,51,52,66,67]

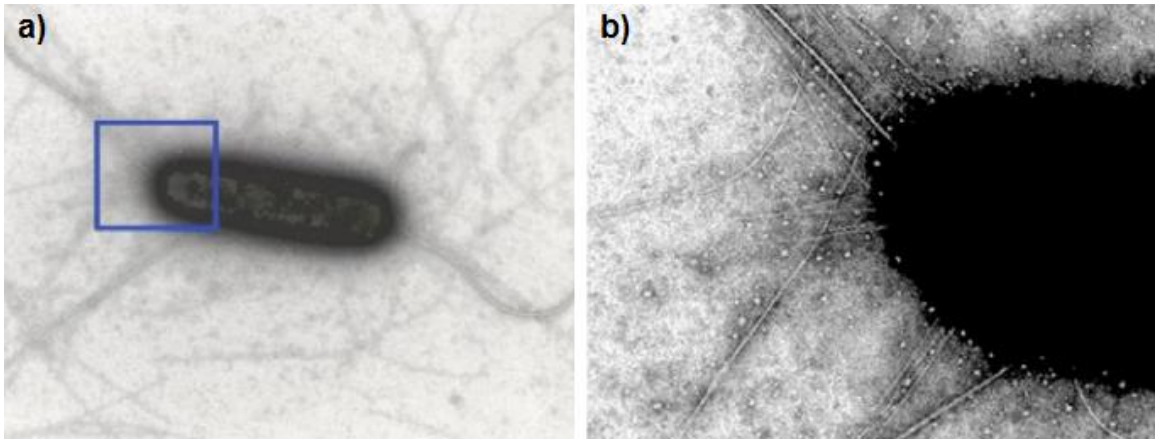


Figure 1.9: Negatively stained transmission electron micrograph of T4P from *P. aeruginosa*. a) The thin fibers that extend from the poles of the bacteria are T4P, with the thick fibers representing flagella. b) Magnified view of the bacterial pole, with a better view of the T4P fibers. Adapted from Audette and Hazes, 2007.[68]

1.2.2.2. The Assembly of T4P

As briefly mentioned earlier, T4P are assembled by systems that are evolutionarily related to T2SSs and archaeal systems. The assembly of T4P has several key players: the pilin subunit itself, an inner membrane prepilin peptidase, an assembly ATPase, an integral membrane protein, an outer membrane secretin, and a retraction ATPase.[69] The mechanism of T4P formation in *P. aeruginosa* is very similar to that of *N. gonorrhoeae* depicted in Figure 1.10.[70]

In *P. aeruginosa*, individual pilin subunits (PilA) are synthesized in the cytosol and are referred to as prepilins, that is, precursors with a leader peptide signalling transport to the inner membrane. The pilins remain anchored to the inner membrane by their hydrophobic α-helix prior to processing. The prepilin peptidase (PilD), which is associated with the cytoplasmic side of the inner membrane, is an aspartic protease responsible for removal of the leader peptide and addition of a methyl group to the N-terminal amine.[69] This process produces mature pilin subunits and facilitates their

transport through the integral membrane protein PilC. Assembly is driven through ATP hydrolysis by the ATPase PilB, which is recruited to the inner membrane by PilC. Secretins are multimeric structures composed of 12-14 subunits and often require the presence of a lipoprotein known as the pilot protein in order to ensure correct assembly and insertion into the outer membrane. A second ATPase is responsible for depolymerisation or retraction of the pilus, which allows for the twitching motility involved in bacterial translocation and rapid colonization of large surface areas.[70]

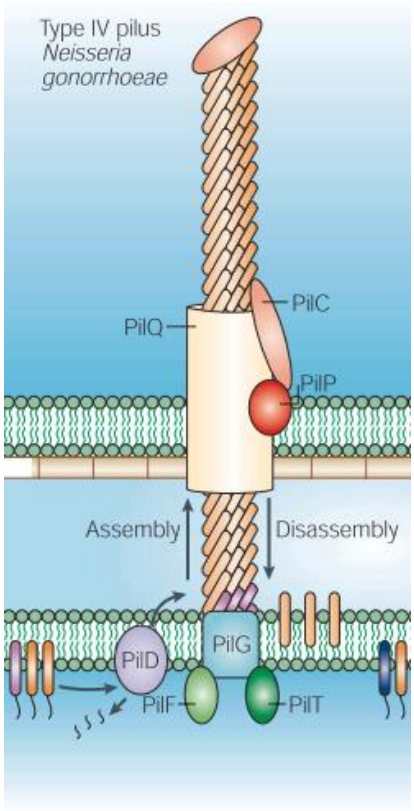


Figure 1.10: Schematic model of T4P formation in *N. gonorrhoeae*. Pilin subunits (PilE) are processed by the prepilin peptidase (PilD) prior to assembly of the pilus fibre. The ATPase (PilF) is responsible for pilus biogenesis with assistance from the polytopic membrane protein (PilG). The pilus crosses the outer membrane through the channel formed by multimeric secretin (PilQ) with the assistance of pilot protein (PilP). Adhesin protein (PilC) is associated with the outer membrane and is believed to stabilize the pilus filament. A second ATPase (PilT) is required for depolymerisation/retraction of the pilus as well as for twitching motility.[70]

1.2.2.3. The Structure of T4P

The T4P is a filamentous protein polymer of a single monomeric unit, the type 4 pilin. These subunits are arranged in a helical conformation with 5 subunits per turn, a 4.1 nm pitch and a diameter of approximately 6.0 nm.[53,71] The majority of type IV pilin subunits have a length ranging from 145-160 amino acids corresponding to a molecular weight of approximately 12-15 kDa.[55] Atomic models for pilins from several bacteria have been characterized, including: *P. aeruginosa* strains PAK, PAO, and K122-4, *N. gonorrhoeae* (GC) strain MS11, *V. cholerae* toxin coregulated pilin (TcpA), and the pilin from *Salmonella typhi*. Based on sequence similarities/differences in the pilin protein, there are two main pili sub-types, the IVa (T4aP) and IVb (T4bP) pili.[52] The T4b pilins are a heterogeneous group common in enteric species such as *V. cholerae* and *S. typhi* whereas T4a pilins, including those observed in *P. aeruginosa*, are more broadly distributed.[52] In addition, T4a pilins have a shorter leader sequence (5-6 amino acids) compared to T4b pilins (15-30 amino acids) and a shorter mature sequence (average of ~150 amino acids compared to ~190 amino acids for T4bP).[53] Structurally, the type IV pilin (PilA) monomer is comprised of an α-helix connected by a variable αβ-loop to a four-stranded antiparallel β-sheet; surface and cellular adherence is mediated through a conserved C-terminal loop known as the receptor binding domain (RBD) or D-region (Figure 1.11).[72-80]

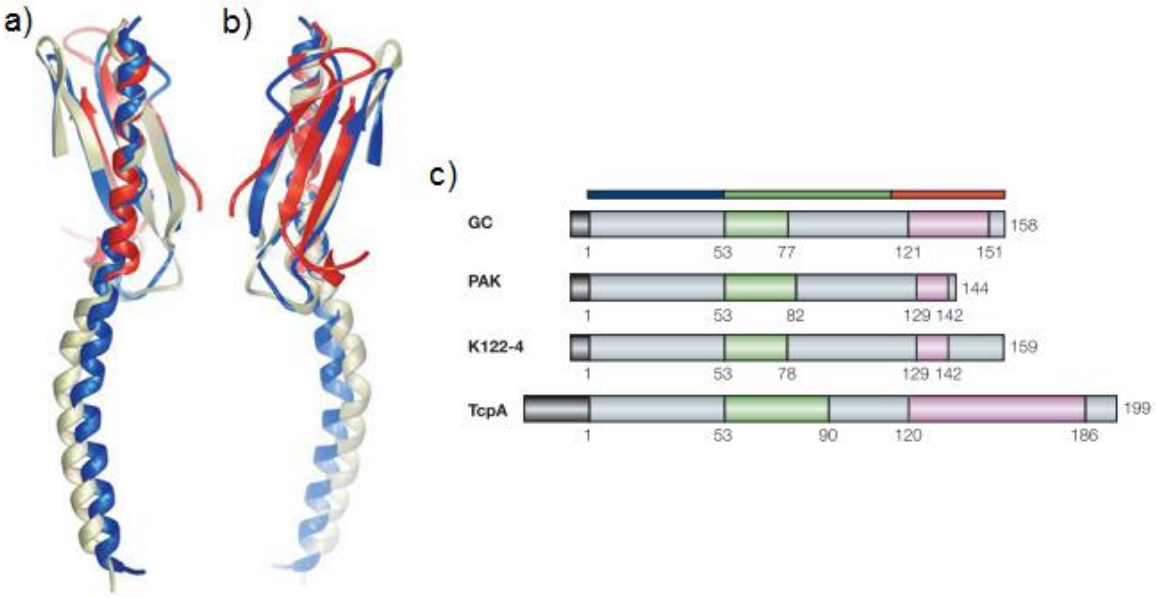


Figure 1.11: Conserved structural domains of T4P. The a) front and b) back view of superimposed PAK (blue), GC (white) and TcpA (red) pilins showing the conserved structural core and the variable αβ-loop and D-region. c) Structural alignment of T4Ps showing relative leader sequence lengths (black) and the positions of the αβ-loop (green) and D-region (magenta). The coloured bar above the GC pilin indicates areas of conserved (blue), variable (green) and hypervariable (red) sequence. Adapted from Craig et al. 2004.[53]

It is believed that there is a common pilus assembly mechanism across the different T4P, with the RBD identifying the same glycosphingolipid receptors. However, there is very little sequence homology when the primary sequences of various strains of T4P are compared. The main area of sequence conservation occurs at the N-terminal end of the α-helix with little to no sequence homology in the RBD. In fact, the only conserved amino acids in this domain are Cys 129 and Cys 142 forming a disulphide bond, as well as Pro 139 providing further structural rigidity.[81] In addition, the side chains of these cysteine and proline residues are buried within the core of the interior suggesting a structural rather than functional role.[31]

In the *P. aeruginosa* pilins, the N-terminal region of the α-helix (α1-N) is both highly conserved and hydrophobic whereas the C-terminal region of the helix (α1-C) is amphipathic and packs onto the β-sheet forming the globular domain.[73-76] The globular domain folds in the periplasm and disulfide bond formation is mediated by an oxidoreductase enzyme.[82,83] The pilins are assembled/disassembled by a membrane-spanning complex into a pilus several microns in length with an approximate outer diameter of 6-8 nm.[52,53,55,69,80,84-86] Cryo-electron microscopy[71,79,80] and fibre diffraction[87] studies of *P. aeruginosa* T4P have shown that assembly is through a three-start helical assembly of pilin monomers to the base of the pilus.[53,69,71]

1.2.2.4. Applications of T4P in Bionanotechnology

From a bionanotechnology perspective, T4P form robust nanofibers with the ability to bind biotic and abiotic surfaces via their tips, interactions which have been mapped to the D-region of the pilin.[50] It has been estimated that the attractive force between the native T4P tip and steel is in excess of 100 pN/molecular interaction[88] and for *in vitro* derived D-region peptides and protein nanotubes is in the range of 26-55 and 78-165 pN/molecular interaction, respectively.[89,90] Due to the self-assembling nature of T4P, functional nanostructures have been generated from native bacterial pili and explored for their potential use as biological nanowires. The development of protein-based nanofibres and nanotubes have several advantages when compared to their inorganic counterparts such as carbon nanotubes (CNTs), which are significantly more cytotoxic and pose compatibility issues. While the strength and stability of protein-based nanotubes are not expected to match those of CNTs, modifications can be made such as

the addition of disulphide bonds or the insertion of loop peptides. Studies have shown that cultures of *G. sulfurreducens* produce biofilms with one of the highest known current densities in microbial-based fuel cells,[91,92] and are capable of long-range metallic-like conductivity[93] and supercapacitor behaviour[94] through a T4P network. These properties make microbial biofilms and T4P exciting prospects for use as a low cost and environmentally sustainable form of energy storage. Additionally, the β-sheet and connecting loops that form the surface of T4P show extremely high sequence variability, which allows for the use of protein engineering strategies to design fibers with altered surface characteristics. Also, if we consider binding of T4P to biotic surfaces such as epithelial cells, this opens an exciting area for further research in therapeutics. As is the case with binding to abiotic surfaces, the D-region of the pilin is responsible for forming specific interactions with cellular glycolipids.[89] This receptor-specific interaction could allow for targeted drug delivery of the therapeutic-loaded T4P-based nanocarriers.

1.2.2.5. The ΔK122 Pilin from *P. aeruginosa* as a Model System

While there are several structural studies of pilins and T4P, the end points of assembly, structural studies of the initial to intermediate stages of pilus assembly have been more challenging. Current T4P models[53,69,71,80] place the α-helix within the interior of the fibril; T4P formation is driven by the hydrophobic α1-N region of the helix, and the β-sheet of the globular head domain defines the outer face of the pilus. In this way, the N-terminal α-helix is protected from identification by the immune system and acts as a conserved oligomerization domain.[68] The αβ-loop of the pilin is predicted to interact with neighbouring subunits due to its size and position with respect to the globular

domain,[53,69,71] while the D-region is predicted to be occluded and only exposed at the tip in order to bind receptors on epithelial cells.[50]

Recently, a truncated form of the pilin from *P. aeruginosa* strain K122-4 (ΔK122) lacking the conserved α1-N region of the helix has been shown to form T4P-like structures, so called protein nanotubes (PNTs), both in solution and at surfaces.[90,95-97] As this truncated form of the pilin lacks the predicted main driving force for pilus oligomerization, namely, the conserved α1-N region, other protein-protein interactions are required to stabilize the structure during the oligomerization process. It has also been observed that the ΔK122 pilin forms a monomer-dimer equilibrium in solution prior to fibril/PNT oligomerization upon incubation with hydrophobic initiators (Figure 1.12).[97]

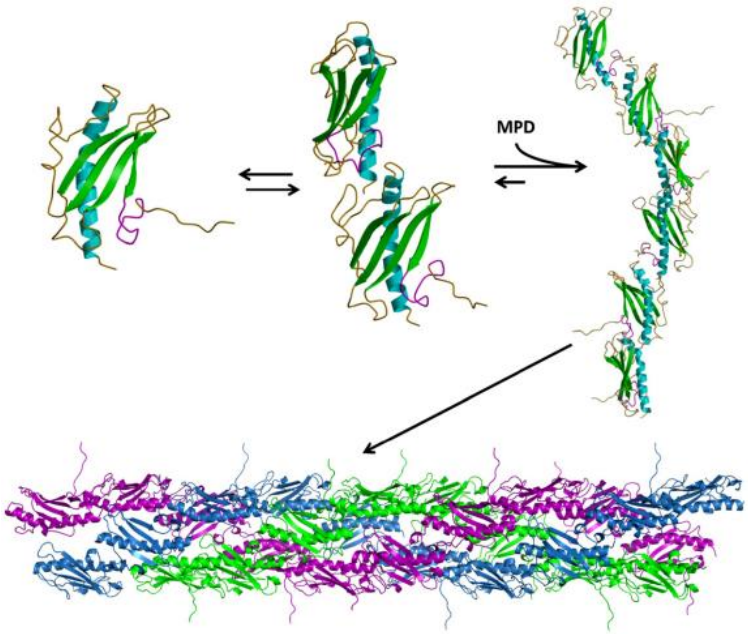


Figure 1.12: Schematic representation of the truncated pilin monomer and its proposed assembly mechanism. The common structural features of type IV pilins are highlighted on the ΔK122 monomer (PDB ID: 1QVE)[75]: the four-stranded antiparallel β-sheet (green), the truncated N-terminal α-helix (cyan), connecting loop regions (yellow), and the RBD (purple). The ΔK122 pilin exists as a monomer-dimer equilibrium and in the presence of hydrophobic imitators such as MPD forms fibrils that coalesce into PNTs by a 3-start assembly model. Adapted from Petrov et al. 2013.[97]

1.3. Time-Resolved Electrospray Mass Spectrometry as a Tool for Probing Protein Structure and Protein-Protein Interaction

Mass spectrometry (MS) has become one of the most sensitive bioanalytical tools for the structural characterization of (bio)molecules, driven by continual instrumental advancements since its origin stemming from the work of J.J. Thomson.[98,99] The development of soft ionization techniques, such as ElectroSpray Ionization (ESI) was a critical advance in mass spectrometry, paving the way for a multitude of new applications, particularly in biomolecular analysis.[100,101] A particular advantage of ESI was that it could be directly coupled to solution phase processes, making it suitable for integration with liquid separation techniques (principally HPLC)[102] and, of particular importance to time-resolved studies, rapid mixing.

In ESI-MS, ions are generated by passing solution through a capillary held at a high electric potential (2 – 6 kV). The high electric potential created at the tip of the capillary causes the dispersion of the sample creating a Taylor cone. In positive ion mode charging of the protein occurs via protonation, while in negative ion mode charging occurs via deprotonation.[103] It is widely accepted that globular species such as natively folded proteins follow the charged residue model (CRM) of gas phase ion formation. This model involves two major steps: the production of charged droplets followed by solvent evaporation.[104] As the radius of the droplet shrinks, the Coulomb force between the positive ions increases due to their proximity, with the electrostatic repulsion causing droplet fission.[105] This phenomenon is described by the Rayleigh equation; Q_{Ry}: charge on the droplet, ε_0: electrical permittivity, γ: surface tension, R: radius of the droplet.

$$Q_{Ry} = 8\pi(\varepsilon_0 \gamma R^3)^{1/2}$$

When the charge exceeds the Rayleigh limit, the droplets explosively dissociate producing charged ions containing single molecules of analyte (Figure 1.13).

Following solvent evaporation, the ions then travel through the mass analyzer where they are separated based on their mass-to-charge (*m/z*) ratio. A detector converts the signal into a mass spectrum. Due to the multiple charging that occurs for the majority of proteins, the *m/z* values tend to fall within the range of most mass analyzers. ESI-MS can be used to obtain information on various protein properties including their three-dimensional structure,[106,107] posttranslational modifications,[108] amino acid sequence,[109] and noncovalent interactions that have been shown to persist in the gas phase.[110,111]

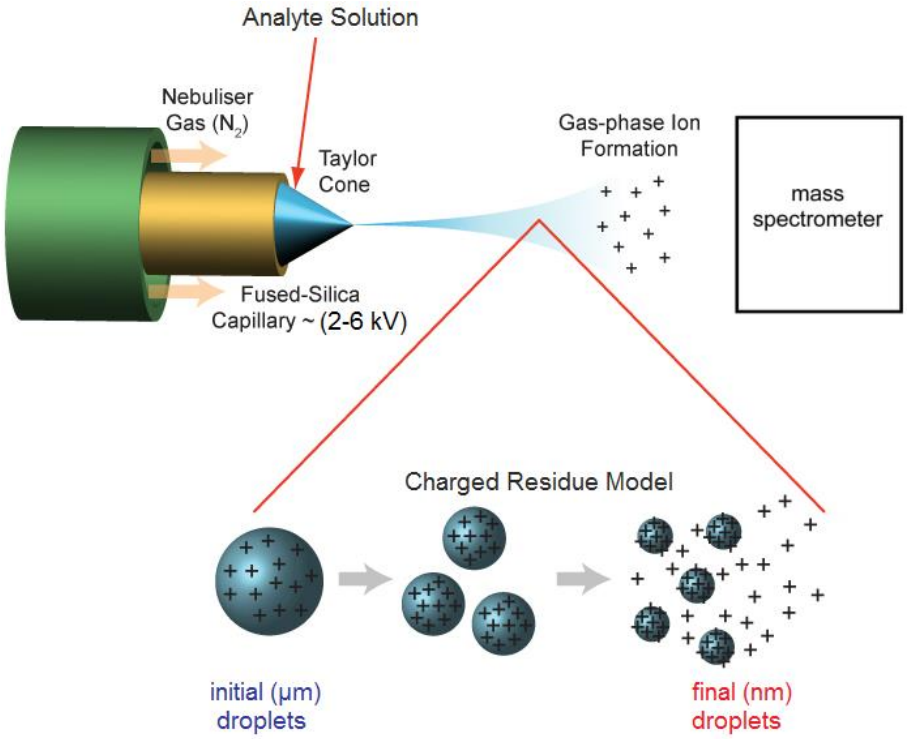


Figure 1.13: **Schematic depiction of an ESI source operated in positive mode.** Analyte solution passes through a metal capillary whose tip is held at a high voltage (2-6 kV). The electric potential created at the capillary tip causes the dispersion of the sample solution creating a Taylor cone. A sheath gas (N_2) that flows around the capillary results in more efficient nebulization of the analyte. The charged droplets diminish in size by solvent evaporation and pass through the sampling cone into the mass analyzer. Adapted from Lamond, 2010.[112]

Characterization of rapid biochemical processes and intermediates such as structural transitions during protein folding, ligand binding and enzymatic "pre-steady" states have proven challenging over the years, because of the need to monitor the reaction under pre-equilibrium conditions, which often persist for less than one second. Analytical techniques allowing for the detection of transient intermediates under pre-equilibrium conditions are referred to as time-resolved, implying that the analysis begins within milliseconds (or microseconds) of reaction initiation. Typically a rapid mixing apparatus is employed to efficiently mix the reactants of interest followed by transfer into an analytical cell. These set-ups usually consist of either stopped-flow or continuous-flow rapid mixers. Stopped-flow devices use a two-step approach, where reactants are first flushed at high flow rates through the mixer and observation cell clearing the contents from previous experiments. Once enough new volume of solution has entered the cell, the flow is stopped and observation begins.[113] Analysis of stopped-flow experiments are classically carried out using spectroscopic methods such as fluorescence and UV-visible absorbance.

Optical detection by fluorescence is suitable for the millisecond time scale mostly thanks to its high sensitivity. However, the main disadvantage of the method is low selectivity leading to an inability to study more than a few species simultaneously. Most optical studies involve the detection of one or two species assuming the availability of a chromophore, making them unsuitable for unravelling complex, multi-step reactions with a large number of reactive species. The use of artificial chromophoric substrate analogues is problematic because there is no guarantee that they exhibit the same kinetics as the natural substrate.[114] Alternatively, circular dichroism (CD) spectroscopy can be used to

study protein unfolding, or melting, as well as protein-protein interactions by measuring the changes that occur in the secondary and tertiary structure of proteins.[115] However, this method has low sensitivity and probes only global structure changes of the protein, making it difficult to pinpoint specific locations of change.

Time-resolved electrospray ionization mass spectrometry (TRESI-MS) allows for kinetic experiments to be carried out with a dead-time ranging from a few milliseconds to seconds.[116,117] The sensitivity and selectivity associated with this technique makes it a promising and often superior approach for studying reactions at sub-second time scales. In addition, TRESI-MS allows for the simultaneous detection of virtually all mass-distinguishable species. Incorporation of a kinetic mixer onto a microfluidic chip in combination with hydrogen-deuterium exchange (HDX) studies has allowed for the local dynamic analysis of proteins undergoing biologically-relevant conformational transitions.

1.3.1. Early Work

One of the earliest continuous-flow capillary-based devices consisted of a static mixing tee followed by a reaction capillary whose fixed internal volume regulated the reaction time (Figure 1.14a).[118] Different time points were obtained by 'switching out' the reaction capillary for one of different length or by varying the flow rates. One of the disadvantages of this system was that different time points could only be obtained in separate experiments. Stopped-flow rapid mixing was implemented for ESI-MS (where a reaction capillary replaces the optical cell) as an alternative that would allow continuous reaction monitoring, however, the open system caused considerable pressure when switching from the initial high-flow rate phase to the low flow-rate injection into the ESI

source, which ultimately had a substantially negative impact on time-resolution.[119] This shifted the focus to the advancement of TRESI based on continuous-flow rapid mixers, which ultimately allowed for substantially decreased dead-times.

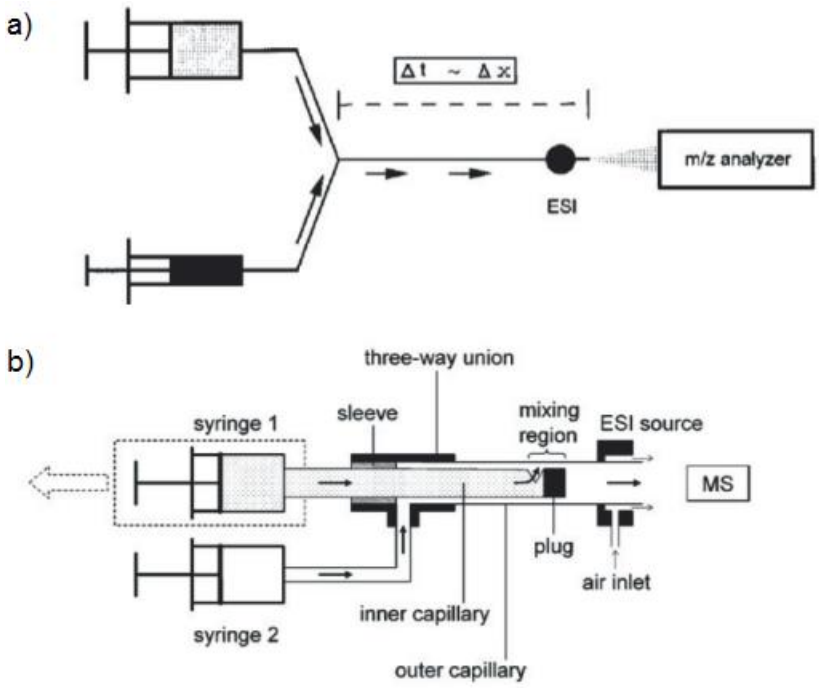


Figure 1.14: Schematic depiction of early continuous-flow capillary-based devices. a) Syringes carrying the reactants of interest meet and react within a static, or fixed, mixing tee. b) Adjustable reaction chamber within two concentric capillaries where multiple time points can be obtained by manual or automated pullback. Adapted from Konermann et al. 1997 and Wilson and Konermann, 2003.[118,120]

The 'fixed mixer' continuous-flow set-up was subsequently improved to include an adjustable reaction chamber in order to measure various reaction times in a single experiment (Figure 1.14b).[120] In this set-up, the reactants were flowed through concentric capillaries where the outer capillary served as the ESI source and solution was allowed to exit the inner capillary through a notch produced 2 mm from the inner capillary tip. The end of the inner capillary was initially lined up with the outer capillary, allowing for a

minimum observable reaction time of around 8 ms. Later, reaction times were monitored by withdrawing the inner capillary within the outer capillary, increasing the dead-volume between the mixing point and the onset of ESI. The capillary could be withdrawn continuously (kinetic mode), or in discrete steps (spectral mode) to allow for longer acquisitions at specific time points.

1.3.2. Microfluidics Coupled to TRESI-MS/HDX

Microfluidics refers to the etching of microscale features onto polymer, glass, or silicone substrates for the handling of small-volume liquid samples. The coupling of microfluidics to ESI-MS provides two main benefits. First, it allows for the possibility of integrating several electrospray emitters on the same chip and increasing the experimental throughput. Secondly, the chip can be used as a platform to carry out a variety of manipulations on the sample before analysis, including separation and proteolysis.[121] Early work in the fabrication of a microfluidic chip to study time-resolved reactions was undertaken by Rob and Wilson in 2009.[122] Here, a simple two-channel design was incorporated into a polymethyl methacrylate (PMMA) chip and used to study the unfolding intermediates of cytochrome *c*. The design has been modified over the years to accommodate a proteolysis chamber in order to carry out local HDX experiments.[123,124]

Briefly, HDX relies on the ability of backbone amide hydrogens to undergo exchange with deuterium atoms in solution.[125,126] At physiological pH, the exchange is base-catalyzed although acid-catalysis becomes dominant below pH 2.6. The rate of exchange is dependent on four main factors: pH, temperature, solvent accessibility and

intramolecular hydrogen bonding.[127] Tightly folded regions with extensive hydrogen bonding networks characteristic of α-helices and β-sheets are generally less dynamic than loops or disordered regions, and thus typically exhibit drastically reduced HDX rates (Figure 1.15).[128] Global HDX-MS experiments give information about the overall flexibility of proteins which can be monitored under various solution conditions and is often predictive of protein stability.

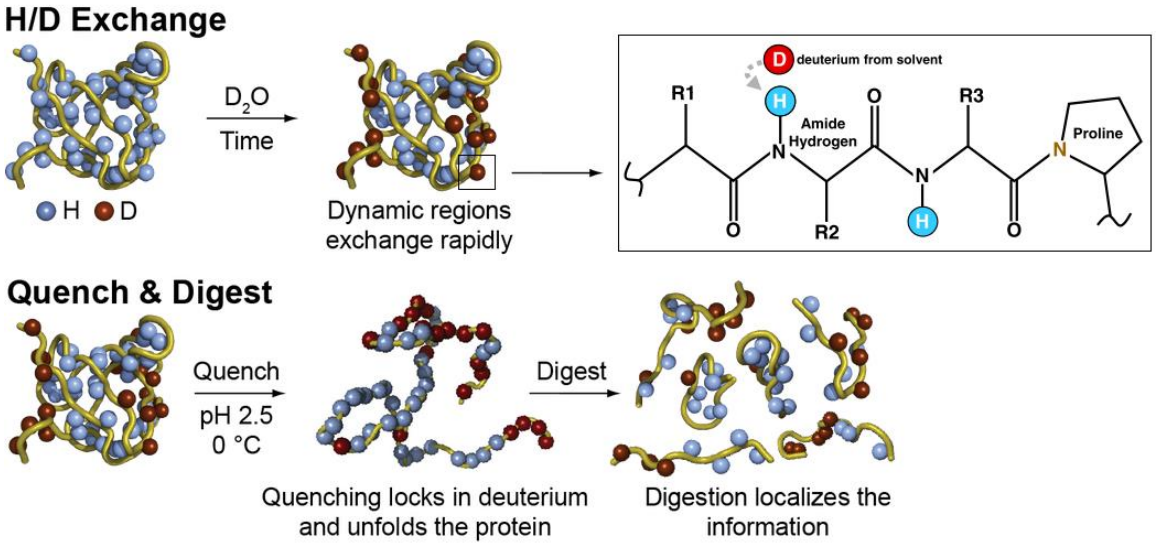


Figure 1.15: Schematic depiction of a protein undergoing hydrogen-deuterium exchange (HDX) prior to mass spectrometry analysis. The protein of interest is diluted into a solution of deuterium oxide (D_2O) where labile backbone amide hydrogens will exchange with deuterium. In highly dynamic unstructured regions, the exchange can be monitored on the millisecond to second time-scale while amides that are part of hydrogen bonding networks will exchange more slowly. HDX is followed by quenching of the reaction with acidification at approximately pH 2.5 and proteolysis with pepsin. This is referred to "local" HDX as it localizes the exchange to specific peptide fragments. Adapted from Weis, 2013.[129]

Alternatively, local HDX-MS experiments allow for the identification of specific stretches of amino acids undergoing dynamic change.[130] To achieve local TRESI-HDX measurements, the protein is incubated with deuterium at sub-second time scales and then

enters a proteolysis microreactor which is held at low pH in order to quench the exchange reaction. The proteolysis microreactor is functionalized with an acid protease, most commonly pepsin, which digests the labelled protein prior to direct transfer into the ESI source.[124] In both global and local experiments, the mass spectrometer monitors deuterium uptake as a function of time, either in the intact protein (global) or individual peptide fragments (local).[130] Typically, proteins under physiological conditions follow EX2 kinetics depicted in equation (1), where F and U are the folded and unfolded states subscripted with hydrogen and deuterium. Under EX2 kinetics the rate of refolding is much greater than the rate of unfolding ($k_{-1} >> k_2$) and thus the rate of deuteration is a function of solvent accessibility over time.[125]

$$F_H \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} U_H \xrightarrow[D_2O]{k_2} U_D \underset{k_1}{\overset{k_{-1}}{\rightleftharpoons}} F_D \quad (1)$$

TRESI-HDX, while still a recently emerging technique, has led to a wealth of knowledge on the dynamic nature of proteins which can be coupled to structural information acquired through NMR and X-ray crystallography. Monitoring backbone exchange at millisecond to second time scales allows for the characterization of conformational changes that occur in loop regions, molten globules, and intrinsically disordered proteins which are difficult to analyze using conventional HDX methods.

The TRESI-HDX device developed by Rob *et al.*[124] consists of a capillary-based rapid mixer integrated into a PMMA microfluidic chip etched with additional acid channels and a reaction chamber filled with pepsin agarose beads (Figure 1.16). Several systems have been explored with this apparatus including studies on weakly structured

and intrinsically disordered proteins. For example, a single time-point (100 ms) was used to obtain the HDX profile for the secondary structure of ubiquitin where there was a significant difference in the uptake levels between secondary structures and loop regions.[124]

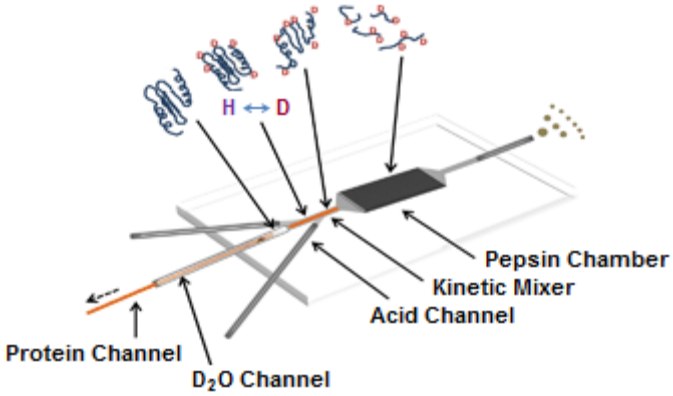


Figure 1.16: Experimental setup for time-resolved HDX-MS. A static kinetic mixer is integrated onto a microfluidic chip along with acid quenching channels and a downstream proteolytic chamber containing pepsin agarose beads. The distal capillary is used as an ESI source. Adapted from Rob et al. 2013.[131]

Similarly, conformational dynamics were studied in the weakly structured loop regions of cytochrome *c*. The analysis was based on 'segment-averaged' protection factors, which compare the observed rate of deuterium uptake to the rate that would be expected in the absence of structure (the latter 'intrinsic' rate has a known primary sequence dependence and can be calculated manually or using web tools).[132] The protection factors were in agreement with previous predicted properties of the loop regions, and were largely dependent on the extent of interaction with the heme prosthetic group.[124]

1.3.3. Applications

The microfluidic TRESI-HDX device can also be used to characterize proteins that currently do not have a solved atomic structure through the methods of X-ray crystallography and NMR. Many of these include intrinsically disordered proteins, such as the neuronal tau protein, which is involved in the formation of amyloid fibrils associated with Alzheimer's disease.[133] Currently, the tau protein is being studied under native (non-amyloidogenic) and hyperphosphorylated (amyloidogenic) conditions in order to localize which regions undergo dynamic change when transitioning into the pathogenic state.[134] One of the great advantages of TRESI-MS/HDX is that it is a very general approach that can be applied to virtually any protein and is not limited by size. For example, the same microfluidic device was used to study the large homotetrameric enzyme 3-deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) synthase, which is 154 kDa in size.[124] In addition, proteins that are difficult to study by NMR due to solubility issues can be analyzed using this technique.

1.4. Research Objectives

The main objective of this thesis is to gain a deeper understanding of how large protein complexes come together to drive important biological events using time-resolved electrospray ionization hydrogen-deuterium exchange mass spectrometry (TRESI-HDX-MS). This method has the ability to provide insight on protein structure – including conformational dynamics, solvent exposure, and areas of protein-protein interaction – together leading to a better understanding of overall function. As mentioned, HDX-MS is a sensitive technique that can be used to study protein-protein interactions of large

macromolecular protein complexes.[135,136] These interactions regulate virtually all aspects of cellular function and the aim is to understand the structural changes that occur upon interaction.

The use of local TRESI-HDX-MS will provide a more detailed structural characterization of TraF, a critical protein involved in the assembly of the F-plasmid T4SS apparatus, as well as a truncated type IV pilin protein (ΔK122) from *P. aeruginosa*. In conjunction, crystallography trials were performed on TraF as a solved crystal structure for the protein is currently not available. Ultimately, this information will provide a clearer picture of the individual components behind these complex protein assembly systems. This will lead to the targeted development of novel inhibitory agents to block bacterial conjugation as well as pili expression, assembly and attachment.

Chapter 2: Experimental Methods

2.1. Chemicals and Supplies

All chemicals were purchased from Sigma-Aldrich, Thermo Fischer Scientific, or BioBasic unless otherwise indicated, and were of ACS grade or higher. Ultrapure water was generated in-house on a Millipore Milli-Q Advantage A10 system or Barnstead Diamond Nanopure system from Thermo Scientific. The *E. coli* ER2507 and BL21 (DE3) cells as well as pT7.7-GST-TraF plasmids were obtained from laboratory stocks. The pMAL p2X and c5X plasmids along with amylose resin were purchased from New England Biolabs. Vivaspin concentrators and glutathione Sepharose 4B resin were purchased from GE Healthcare. All liquid chromatography was carried out using an ÄKTA Purifier 10 under the control of the Unicorn 5 software package. Purification columns were purchased from GE Healthcare and Bio-Rad. A Nikon SMZ1500-Fiber Lite MI-150 microscope was used to observe crystallization plates.

2.2. Expression and Protein Purification

2.2.1. *E. coli* ER2507 containing MBP-ΔK122 in pMAL-p2X

2.2.1.1. Cell Growth and Expression

Truncated monomeric type IV pilin from *P. aeruginosa* strain K122-4 (pilA, Δ1-28) was expressed as an MBP fusion protein under the control of a lac operon. It is of note that the ΔK122 pilin is truncated 28 amino acids to remove the hydrophobic first half of the N-terminal α-helix and improve solubility. Glycerol stocks transformed with vector pMAL-p2X encoding for MBP-ΔK122 in *E. coli* strain ER2507 were used to streak LB + Amp (100 μg/mL) agar plates and incubated overnight at 37°C. Overnight

cultures were made from single colonies where the cells were grown in Luria Bertani (LB) broth (1% Tryptone, 0.5% Yeast Extract, 1% NaCl) along with 1M glucose (1:100 v/v ratio with LB) and 50 µg/mL of ampicillin (1:1000 v/v ratio with LB) with shaking at 200 rpm at 37°C for 18 hours. The overnight cultures were transferred to 1L expression volumes at a 1:100 v/v ratio to LB and grown at 37°C to an optical density (OD_{600}) of 0.5 – 0.7 at 600 nm corresponding to mid-log phase growth. At this point, 1M of IPTG at a 1:1000 v/v ratio was added with continued induction for 3 hours at 30°C. Cell pellets were then collected through centrifugation of the cultures for 20 minutes at 6,000 x g at 4°C and stored at 4°C overnight.

2.2.1.2. Periplasmic Protein Release Using Osmotic Shock and Affinity Purification

The release of proteins from the periplasm was conducted using the osmotic shock method.[137] The process involves two steps, one in which the bacterial cell pellets are resuspended in a 50 mL solution of 10 mM tris(hydroxymethyl)aminomethane (Tris; pH = 7.4) and 20% sucrose. This allows for the movement of water from the periplasmic space outward into the extracellular environment of the cells due to the high concentration of sucrose. The resuspended cells were incubated on ice for 25 minutes with gentle shaking. The cells were then centrifuged at 7,000 x g for 20 minutes at 4°C. The second step involves removal of the supernatant and resuspension of the remaining pellets in 50 mL of 5 mM magnesium sulphate ($MgSO_4$). Transfer of the cells to a low osmotic strength solution and the sudden reduction of osmotic pressure causes the outer membrane to rupture, releasing the contents of the periplasmic space where MBP-ΔK122 has been exported to. The resuspended cells were incubated on ice for 25 minutes with

gentle shaking. The cells were then centrifuged at 35,000 x g for 20 minutes at 4°C. The supernatant was syringe filtered using a 0.45 µm membrane and purified by affinity chromatography. The periplasmic solution containing MBP-ΔK122 was loaded onto amylose beads equilibrated with 20 mM Tris-HCl pH 7.4, 200 mM NaCl and 1 mM EDTA. Following washing of the column with several columns of loading buffer, MBP-ΔK122 was eluted from the column with an elution buffer containing 20 mM Tris-HCl pH 7.4, 10 mM maltose and 1 mM EDTA. The flow through was collected in 45 mL fractions and the elution samples were collected in 5 mL fractions.

2.2.1.3. Cation Exchange Chromatography (CIEX)

Ion exchange chromatography separates proteins based on their net surface charge. While MPB is a large protein with an isoelectric point (pI) of 5.08, ΔK122 is a much smaller protein with a pI of 8.31. Due to the significant different in these values, ΔK122 will be positively charged in a Tris-HCl pH 7.4 buffer while MBP will be negatively charged. The ΔK122 protein was released from the MBP affinity tag through the process of trypsin digestion. The fusion protein was incubated on ice for 10 minutes at a 1:500 v/v ratio of trypsin to protein ratio. Proteolysis was quenched by the addition of phenylmethylsulphonyl fluoride (PMSF) protease inhibitor at a 10:1 v/v ratio of PMSF to trypsin. ΔK122 was purified from the digestion mixture by cation exchange chromatography (CIEX). The column was equilibrated with CIEX loading buffer followed by loading of the protein sample which was eluted in a linear 0 – 100% gradient with CIEX elution buffer (see Table 2.1). The flow through was collected in 45 mL fractions and the elution samples were collected in 5 mL fractions.

2.2.2. *E. coli* BL21 containing GST-TraF in pT7.7

2.2.2.1. Cell Growth and Expression

Glycerol stocks transformed with vector pT7.7 encoding for GST-TraF in *E. coli* strain BL21(DE3) were used to streak LB + Amp (100 µg/mL) agar plates and incubated overnight at 37°C. Cell growth and expression was undertaken as per Section 2.2.1.1.

2.2.2.2. Cell Re-suspension and Sonication

Cell pellets were resuspended in a 1:10 ratio of lysis buffer (see Table 2.1) until a homogenous solution was obtained. The cell suspension was then disrupted using the Fisher Model 500 Sonic Dismembrator with brief pulses on ice for 5 minutes (at 30% amplitude for 15 seconds on, 30 seconds off). The cytoplasmic fraction was collected through centrifugation at 35,000 x g for 30 minutes at 4°C. The supernatant was syringe filtered using a 0.45 µm membrane and purified by affinity chromatography.

2.2.2.3. Affinity Purification

The released cytoplasmic sample was loaded onto a glutathione Sepharose 4B column pre-equilibrated with 50 mM Tris-HCl pH 7.5, 150 mM NaCl and 1 mM EDTA. The column was then washed with three CV of loading buffer to remove the unbound sample. Finally, GST-TraF was eluted with two CV of elution buffer containing 20 mM Tris-HCl pH 8.0 and 10 mM glutathione. The flow through was collected in 45 mL fractions and the elution samples were collected in 5 mL fractions.

2.2.3. *E. coli* ER2507 containing MBP-TraH in pMAL-c5X

2.2.3.1. Cell Growth and Expression

Glycerol stocks transformed with vector pMAL-c5X encoding for MBP-TraH in *E. coli* strain ER2507 were used to streak LB + Amp (100 µg/mL) agar plates and incubated overnight at 37°C. Cell growth and expression was undertaken as per Section 2.2.1.1.

2.2.3.2. Cell Re-suspension and Sonication

Cell disruption was undertaken as per Section 2.2.2.2.

2.2.3.3. Affinity Purification

The released cytoplasmic sample was loaded onto amylose beads equilibrated with 10 mM Tris-HCl pH 7.5, 50 mM NaCl and 1 mM EDTA. Following washing of the column with 3 CV of loading buffer, MBP-TraH was eluted from the column with an elution buffer containing 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 10 mM maltose and 1 mM EDTA. The flow through was collected in 45 mL fractions and the elution samples were collected in 5 mL fractions. The purification buffer recipes used for the various protein constructs in this study can be found in Table 2.1.

2.2.3.4. MBP-TraH-TraF Pull Down Assay

The glutathione Sepharose 4B column with bound GST-TraF was pre-equilibrated with PreScission Protease Buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA,

1 mM DTT). PreScission Protease (GE Healthcare) was added in the amount of 2-units enzyme/100 µg of bound GST-TraF and incubated overnight at 4°C. The cleaved TraF protein was then collected and concentrated to 1 mg/mL before incubation with amylose beads bound to MBP-TraH. This was followed by a wash and elution step with maltose in order to determine if TraF would successfully bind to MBP-TraH.

Table 2.1: Purification and exchange buffers for proteins used in this study.

Protein Construct	Lysis/Loading Buffer	Elution Buffer	Buffer For MS/Crystallization Trials
MBP-K122 (Affinity)	20 mM Tris (pH = 7.4) 200 mM NaCl 1 mM EDTA	20 mM Tris (pH = 7.4) 10 mM maltose 1 mM EDTA	10 mM ammonium acetate (MS)
K122 (CIEX)	10 mM Tris (pH = 7.4)	10 mM Tris (pH = 7.4) 1 M NaCl	50 mM ammonium acetate (MS)
GST-TraF	50 mM Tris (pH = 7.5) 150 mM NaCl 1 mM EDTA	20 mM Tris (pH = 8.0) 10 mM glutathione	50 mM sodium phosphate (crystallization) 10 mM ammonium acetate (MS)
MBP-TraH	10 mM Tris (pH = 7.5) 50 mM NaCl 1 mM EDTA	10 mM Tris (pH = 7.5) 50 mM NaCl 10 mM maltose 1 mM EDTA	150 mM ammonium acetate (MS)

2.3. Protein Visualization

Fractions containing protein were further analyzed by loading onto a 12.5% polyacrylamide gel and subjected to electrophoresis (PAGE). Native-PAGE analysis of ΔK122 was performed as follows: purified protein (1 mg/mL) was mixed with 2X loading dye (62.5 mM Tris-HCl pH 6.8, 1% bromophenol blue, 25% glycerol). The native gel prepared contains a 12.5% acrylamide:bis-acrylamide (29:1) resolving gel, and a 4% stacking gel. After loading the protein samples, they were electrophoresed at 200V for 45 minutes using a native running buffer (25 mM Tris-HCl pH 8.3, and 192 mM glycine), and stained with Coomassie blue (40% methanol, 10% glacial acetic acid, 500 mg Coomassie blue G250) for visualization. For denaturing PAGE conditions, sodium dodecyl sulphate (SDS) was added to the loading dye and resolving/stacking gels. Thermo Scientific unstained Molecular Weight Marker (MWM) #26630 was used for size estimation.

2.4. Sample Concentration and Buffer Exchange

The proteins were concentrated to various target amounts using an appropriate MWCO (5K, 10K, or 30K) concentrator. The various protein samples in their respective elution buffers were centrifuged at 4,000 x g and at 4°C through the anisotropic semi-permeable membrane allowing solvents and low molecular weight solutes to escape while retaining high molecular weight proteins above the molecular weight cut-off. Protein concentration was determined using the absorbance at 280 nm and the following extinction coefficients ($\varepsilon_{280\ nm}$): MBP-ΔK122 = 1.47 mL/mg·cm, ΔK122 = 0.75 mL/mg·cm, GST-TraF = 0.93 mL/mg·cm and MBP-TraH = 0.78 mL/mg·cm. Unless

otherwise specified, all buffer exchanges were also performed using appropriate MWCO Vivaspin concentrators.

2.5. Electrospray Ionization Mass Spectrometry (ESI-MS) and Ion-Mobility Spectrometry (IMS)

Molecules of interest are introduced into the ionization source of the Waters SYNAPT G1 and analyzed in positive mode, where the ions acquire net positive charges. Charged droplets are produced through injection of the sample through a capillary and application of a positive potential to create an electric field. Solvent evaporation of the charged droplets leads to the formation of gas-phase ions. The ions enter the mass spectrometer through electrostatic attraction and a vacuum, then travel through the mass analyzer where they are separated based on their *m/z* ratio.[105,138]

Freshly purified ΔK122 was prepared for ESI-MS by buffer exchanging the protein into 50 mM ammonium acetate using Slide-Analyzer 3.5K Dialysis Cassettes (Thermo Scientific). The protein was dialyzed in a series of 2 hour, overnight, and 1 hour buffer exchanges. The sample was then collected for analysis on the Waters Synapt G1 High Definition Mass Spectrometer (Figure 2.1).

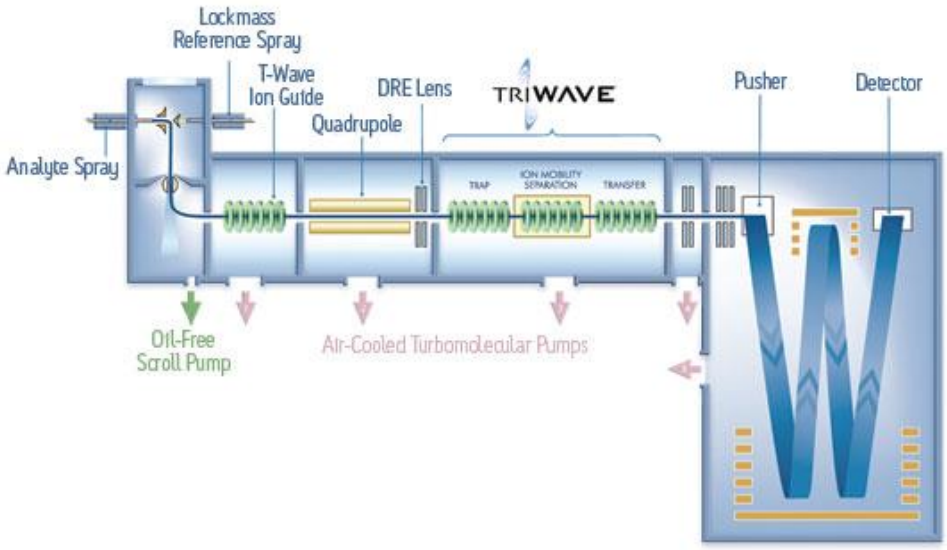


Figure 2.1: Schematic representation of the Synapt G1 High Definition Mass Spectrometer. The T-wave ion guide focuses incoming ions for mass selection by the quadrupole. Each travelling wave performs a specific function; with the first trapping and releasing parent ions, the second allowing for ion mobility separation, and the third responsible for transferring the ions to the TOF mass analyser. Ion mobility separation allows for the detection of ions not only by mass but also by size which is ideal for studying the topology of large protein complexes. Collision-induced-dissociation (CID) can be performed before or after ion mobility.

Following buffer exchange, ΔK122 concentrated to 100 µM was mixed in a 15:1 (v/v) ratio of protein to 2-methyl-2,4-pentanedial (MPD) and in a 3:1 (v/v) ratio of methanol to MPD. Readings were taken immediately after triggering and every two hours for 8 hours at a low mass (1000-3000 m/z) and high mass (3000-5000 m/z) range. The protein was injected at 5 µL/min at different settings optimized for each mass window. For the low mass region the following settings were used: ion source capillary voltage of 2.00 to 3.00 kV, Sampling Cone of 50, and Extraction Cone of 2.0. The high mass region was obtained under the same conditions except for an increased Sampling Cone of 200. Operating pressures (Torr) of the Trap: 7.87×10^{-3}, IMS cell: 3.18×10^{-4} and TOF: 4.40×10^{-7} remained the same between low and high mass readings. The backing pressure (Torr)

was increased for the high mass to allow for better desolvation: 1.87 (low mass), 4.20 (high mass).

2.6. Time-Resolved Electrospray Ionization Mass Spectrometry (TRESI-MS)

The Synapt G1 equipped with a custom-made time-resolved ESI source[120] was used to carry out global HDX on ΔK122 as well as monitoring of the oligomerization reaction at the millisecond time-scale. Briefly, a mobile polyamide-coated glass capillary with an outer diameter (o.d.) of 109.2 µm was inserted into a metal capillary with an inner diameter (i.d.) of 132.6 µm. A 2 mm notch was made from the end of the glass capillary in order to seal the end and allow for efficient mixing of the reagents. This kinetic mixer was integrated onto the Synapt G1 in order to carry out time-resolved hydrogen-deuterium exchange on the equilibrium state of ΔK122 (20 µM) as well as the triggered state (20 µM) after an incubation of 24 hours with 15% MPD. In addition, triggering of the ΔK122 protein on the millisecond time scale was attempted using the same device, using 2 µL of 20 µM ΔK122 with 2 µL of 30% MPD in methanol.

2.7. Microfluidic Device Fabrication

The microfluidic device was made on a blank poly(methyl methacrylate) or PMMA substrate purchased from Professional Plastics (Fullerton, CA) measuring 8.9 cm x 3.8 cm x 0.6 cm. The microfluidic channels outlining the protein channel, two acid channels, and the pepsin reaction chamber was etched onto the PMMA chip by laser ablation using the *Versa*Laser™ laser engraver (Universal Laser, Scottsdale, AZ). The microfluidic channel design was generated in CorelDraw X3 (Corel, Ottawa, ON) as

previously described[124] with a few modifications. A 28RW sized metal capillary was incorporated into the central channel of the device using a soldering iron. Two additional 27RW sized metal capillaries for acid delivery were incorporated in similar fashion on both sides of the central channel. A mobile polyamide-coated glass capillary with an outer diameter (o.d.) of 153 µm was inserted into the central 28RW metal capillary with an inner diameter (i.d.) of 178 µm. Again, a 2 mm notch was made from the end of the glass capillary. A deuterium oxide (D_2O) channel was connected to the central capillary via a T-junction (Figure 2.2). Polyimide coated glass capillaries were supplied by Polymicro Technologies (Phoenix, AZ) and metal capillaries (28RW, 27RW) were supplied by McMaster-Carr (Aurora, OH).

2.8. H/D Exchange of MBP-ΔK122, ΔK122 and GST-TraF

Once the chip is constructed, pepsin agarose beads are placed in the reaction well and activated using hydrochloric acid (pH 1.8) at a combined flow rate of 10 µl/min for one hour followed by acetic acid (pH 2.3) for one hour. Silicone rubber cut with an outline of the reaction well was placed between the fabricated PMMA chip and a blank chip. In this way, a liquid-tight seal can be obtained between the two blocks to prevent leaking. The sandwich is then placed in a metal clamp (LAC Machine & Tooling Limited, ON) to pressure-seal the microfluidic device. All acid, protein, and deuterium channels were connected to Harvard 11+ infusion syringe pumps (Holliston, MA) used to administer respective flow rates.

Freshly prepared protein is passed through the inner glass capillary, exits through the notch and mixes with the incoming deuterium. This creates an orthogonal flow from the inner glass capillary to meet the laminar flow coming through the outer channel to create efficient mixing through turbulent flow. The protein and deuterium were flowed at various ratios ranging from 0.5 uL protein and 3.5 uL D_2O to 2 uL protein and 2 uL D_2O. All deuterium uptake values were normalized to 100% for comparison. The total volume of this mixing region is approximately 13 nL (see Appendix). Previous work determined that the volume required for efficient mixing is approximately 10 nL, which is why placing the notch at 2 mm from the end of the capillary is required to allow for complete mixing before transfer to the reaction channel.[122] Moving the inner glass capillary within the central channel changes the reaction volume and allows one to track a reaction through various time points between mixing and ESI (Figure 2.2). In this study, labelling times ranging from 200 to 2000 ms were acquired (see Appendix).

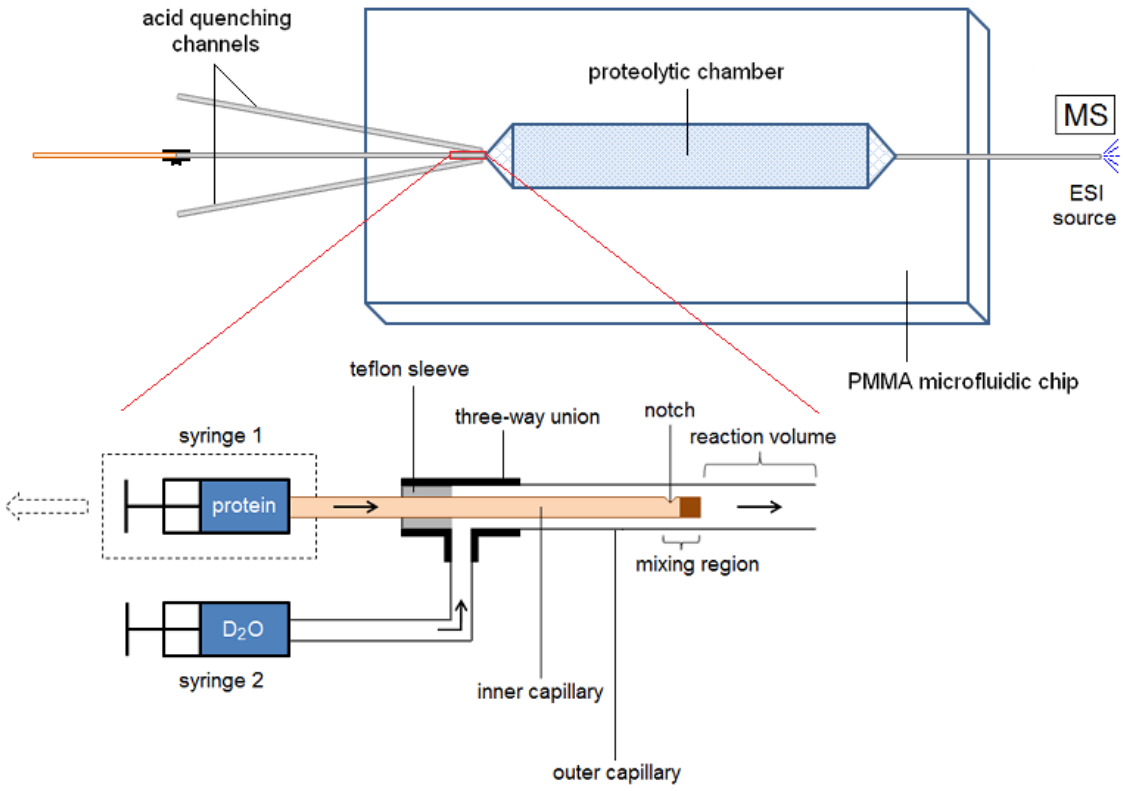


Figure 2.2: Experimental setup for time-resolved HDX-MS. The kinetic mixer is integrated into a microfluidic chip along with acid quenching channels and a downstream proteolytic chamber containing pepsin agarose beads. The distal capillary is used as an ESI source. A zoomed in view of the kinetic mixer with adjustable reaction volume is also shown. Two concentric capillaries are injected with reactants from syringe 1 (containing protein) and syringe 2 (containing deuterium). A notch is made 2 mm from the end of the inner capillary and plugged at the end allowing for protein to exit from the inner capillary and efficiently mix with the incoming deuterium. Changing the position of the inner capillary within the outer capillary changes the volume between mixing and ESI detection allowing for the continuous tracking of the reaction over various time points.

2.9. Data Acquisition

Time-resolved hydrogen-deuterium exchange mass spectrometry was carried out on a Q-Star Elite quadrupole time-of-flight (Q-TOF) instrument (MDS Analytical Technologies, Concord, ON). The instrument consists of three quadrupoles, Q0, Q1, and Q2, followed by a TOF analyzer (Figure 2.3). The instrument was operated in positive mode with a source voltage of 2600 V. Optimal positioning of the electrospray tip was

achieved using an adjustable stage. All data were acquired at a rate of $1s^{-1}$ and the samples were scanned over the 350 – 2000 *m/z* range.

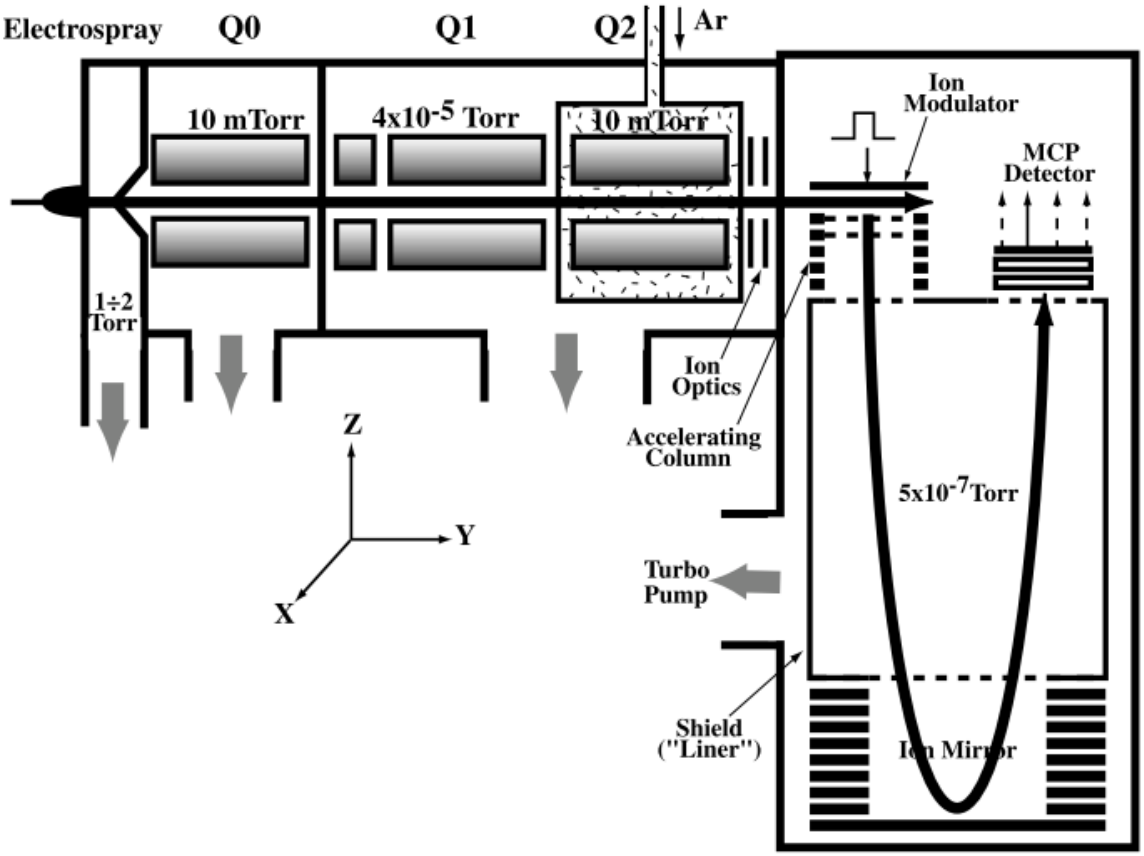


Figure 2.3: Schematic representation of a Q-TOF mass spectrometer. When Q0, Q1 and Q2 are set in RF-only mode, the ions pass through and the TOF analyzer records the spectrum. For MS/MS measurements, Q0 and Q2 are operated in the RF-only mode while Q1 is operated in the mass selective mode to transmit the interested parent ion into the collision cell Q2, where it undergoes CID. Adapted from Chernushevich et al. 2001.[139]

2.10. Data Analysis

The resulting protein digests of MBP-ΔK122, ΔK122, and GST-TraF were analyzed using the FindPept tool on the ExPASy Proteomics server (Swiss Institute of Bioinformatics, Basel). Search fields were set to pepsin (porcine A) at pH > 2 with a

mass tolerance of ±0.5 Da. The experimental deuterium uptake of each peptide obtained was calculated using the custom built software program (DJW, unpublished results). In parallel, the intrinsic rates of each peptide were calculated using SPHERE, an online server program for hydrogen-deuterium exchange rate estimation. The observed experimental and intrinsic uptake values were plotted as a function of reaction time and fit to a single exponential expression in order to obtain curves that best fit the k_{int} and k_{obs} rates of exchange. Error bars represent standard deviation values where at least 3 replicate runs were obtained. Due to the nature of the method and low sequence specificity of pepsin cleavage, low intensity peptides are sometimes lost during some intervals of the experiment and thus only 1 or 2 data points were obtained.

2.11. GST-TraF Crystallization Trials

The crystallization and preliminary diffraction analysis of TraF has been reported by Audette and colleagues.[31] However due to very small crystals at the time, suitably diffracting crystals of heavy atom derivatives were not obtained, and therefore additional information is required to solve for the TraF structure. To this end, crystallization trials of the fusion protein GST-TraF was attempted at room temperature (20°C). Initial crystallization experiments were set up using the Microlytics MCSG Core I-IV screening kit in 96-well sitting drop plates (Axygen Biosciences, Inc). In addition, a batch of purified GST-TraF (6.4 mg/mL) was sent to the Hauptman-Woodward Medical Research Institute (Buffalo, New York) for high-throughput crystallization screening. Crystallization solutions consist of varying concentrations of precipitating agents, salts, as well as differing physical conditions such as pH. Buffers that are commonly used

include Tris or HEPES to help maintain a particular pH, while a variety of different salts and precipitants such as polyethylene glycol (PEG) or ammonium sulphate serve to compete for association with water and excludes the protein from the solvent leading to phase separation.

Both sitting and hanging drop geometries were used for crystallization. For the 96-well screening plates, the sitting drop method was used. This technique involves the deposition of a 1 µL of protein with 1 µL of the crystallization reagent on an elevated platform in vapour equilibrium with a reservoir of the crystallization reagent. The concentration of the reagents in the sitting drop is initially half that of the reservoir concentration, and over time, vapour diffusion occurs. Vapour diffusion is the process whereby the reservoir will pull water from the droplet in a vapour phase until it reaches an equilibrium state. This process will also simultaneously increase the concentration of the protein sample increasing the supersaturation of the sample in the drop. The hanging drop technique is based on the same principles of vapour diffusion with the only difference being that the droplet is placed on a glass cover slide and inverted over the reservoir.

The fusion protein was concentrated to various concentrations ranging from 3.6 to 11.6 mg/mL. All sitting drops contained 1 µL of purified protein (either in elution buffer or exchanged into 50 mM sodium phosphate) with 1 µL of reservoir solution equilibrated over 100 µL of reservoir. When immediate precipitation was observed the addition of 1 µL of dH_2O took place. For the first two weeks, crystal plates were monitored daily using

a Nikon SMZ1500-Fiber Lite MI-150 light microscope. After the first two weeks, plates were monitored for crystal growth every 2-3 days. Optimization trials were conducted in 24 well plates using the hanging drop method and reservoir solutions were made according to crystallization conditions observed in the screening kits. Following identification of possible leads, protein concentration, precipitant concentration, pH, and temperature were varied within the optimization screens.

Results and Discussion

Chapter 3: Structural Changes During Dimerization of the Type IV Pilin from *Pseudomonas aeruginosa* Strain K122-4 Measured by Time-Resolved Hydrogen-Deuterium Exchange

3.1. Purification of MBP-ΔK122 and ΔK122

The truncated (Δ1-28) K122 pilin was previously cloned into the pMAL-p2X vector and transformed into *E. coli* strain ER2507.[140] Optimal expression of MBP-ΔK122 after induction with IPTG was observed in one-hour increments for a total of 3 hours at 30°C (data not shown). After lysis of the periplasmic components by osmotic shock, MBP-ΔK122 (~55 kDa) was purified using amylose affinity chromatography. As can be seen in Figure 3.1, the MBP-ΔK122 fusion eluted at a peak maximum absorption of approximately 1000 mAU and at 100% elution buffer containing 10 mM maltose. The blue line represents UV absorbance at 280 nm and the red line represents absorbance at 215 nm. The corresponding SDS-PAGE gel is shown in Figure 3.2. Since *E. coli* strain ER2507 is engineered not to express endogenous MBP, the secondary MBP band visualized on the gel (~42 kDa) is due to the cleavage of the fusion by proteases in the periplasm, and is often seen when expressing the MBP-pilins (GFA, personal communication). It was also found that performing the purification at 4°C decreased proteolysis of the MBP tag significantly.

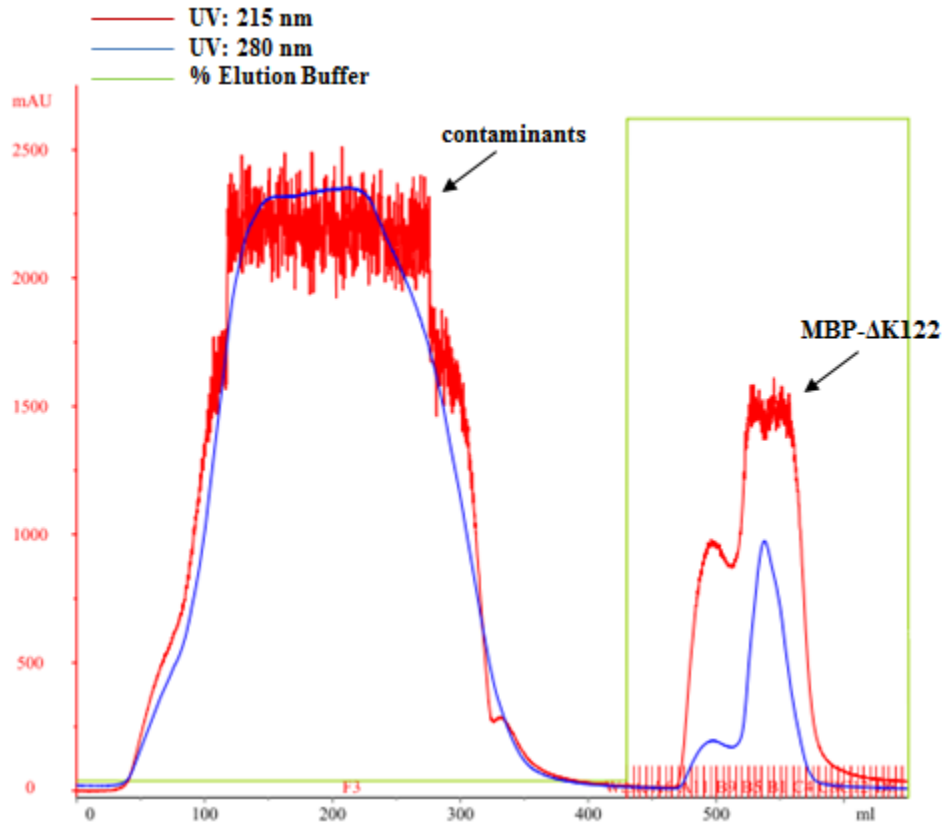


Figure 3.1: Affinity chromatogram of MBP-ΔK122. The MBP-ΔK122 protein eluted at a peak maximum absorption of ~ 1500 mAU and 100% elution buffer. The blue line is absorbance at 280 nm, the red line is absorbance at 215 nm and the green line is the elution buffer concentration.

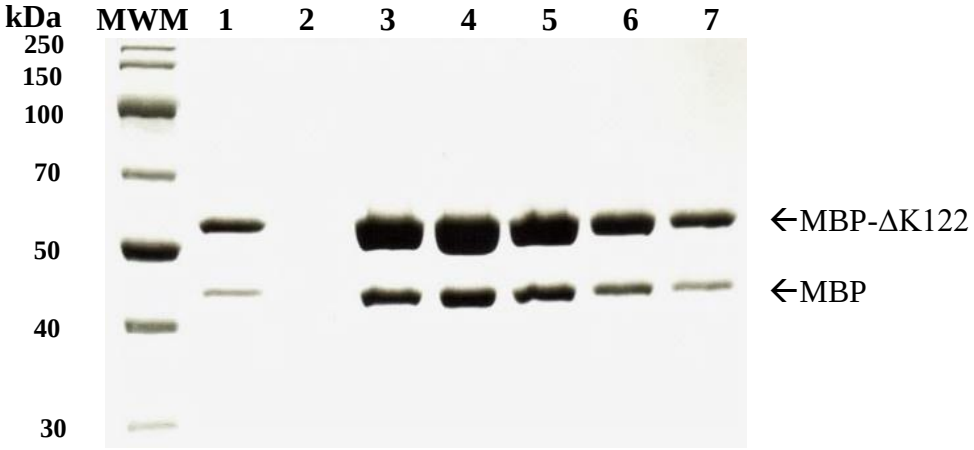


Figure 3.2: SDS-PAGE gel of MBP-ΔK122 affinity purification. Lane 1: periplasmic solution after osmotic shock. Lane 2: flow through. Lanes 3-7: elution samples containing the protein of interest MBP-ΔK122 at approximately 55 kDa and a secondary MBP band at approximately 42 kDa.

The pMAL-p2X plasmid contains a Factor Xa cut site to facilitate the cleavage of MBP from ΔK122, however, efficient cleavage occurs over several days of incubation. The Factor Xa cut site is a strict recognition sequence composed of Ile-Glu (or Asp)-Gly-Arg with cleavage occurring after the arginine as long as it is not followed by a proline.[141] Trypsin is another protease which has the ability to cleave on the carboxyl side of lysine or arginine residues. A previous study[140] on ΔK122 has shown that the protein itself is quite resistant to trypsin digestion when the protein-to-protease ratio is 500:1 and with a cleavage time of 10 minutes on ice before trypsin inactivation with PMSF. Therefore optimal cleavage at the Factor Xa cut site occurs releasing ΔK122 from MBP.

After digestion with trypsin, MBP was separated from ΔK122 using cation exchange chromatography (CIEX). CIEX employs a negatively charged resin to attract positively charged molecules. As mentioned, MBP is a large protein with a pI of 5.08 while ΔK122 is a much smaller protein in comparison with a pI of 8.31. Using a 10 mM Tris buffer maintained at pH 7.4, MBP (~42 kDa) becomes negatively charged and appears in the flow through, while the positively charge ΔK122 pilin binds to the negatively charged resin forming a strong ionic interaction. Elution occurs through displacement of the positively charged protein from the column using a linear 0–1 M NaCl gradient. As can be seen in Figure 3.3, ΔK122 (~13 kDa) eluted at a peak maximum absorption of approximately 2500 mAU and at approximately 22-25% elution buffer corresponding to 220–250 mM NaCl. The blue line represents absorbance at 280 nm and the red line represents absorbance at 215 nm. The corresponding SDS-PAGE gel is shown in Figure 3.4.

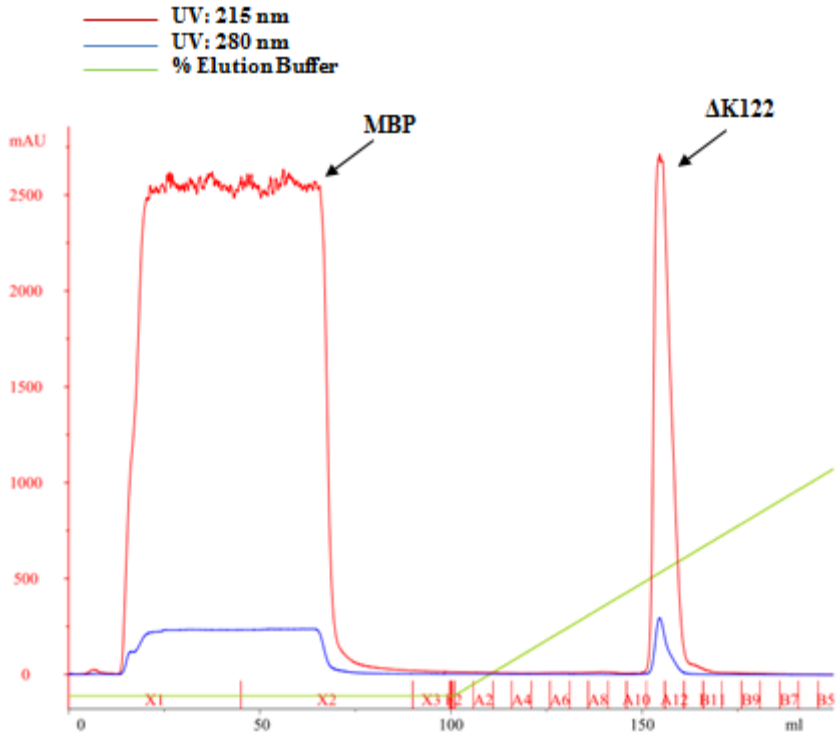


Figure 3.3: Cation exchange chromatogram of MBP-K122. The MBP protein eluted in flow-through samples X1-X2. The ΔK122 protein eluted at a peak maximum absorption of ~ 2500 mAU. The blue line is absorbance at 280 nm, the red line is absorbance at 215 nm and the green line is the elution buffer concentration.

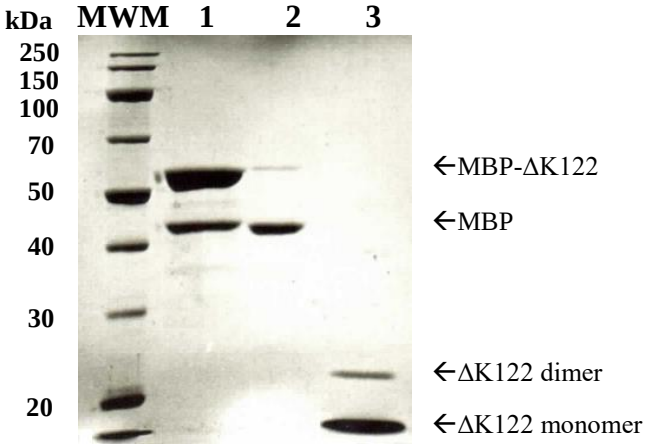


Figure 3.4: Native PAGE analysis of ΔK122 following cation exchange purification. Lane 1: pooled fraction from amylose resin purification. Lane 2: flow through of the cation exchange purification. Lane 3: purified ΔK122 (100 μM) showing the presence of a monomer (13 kDa) and dimer (26 kDa) equilibrium in solution.

3.2. Characterization of the Monomer-Dimer Equilibrium

A previous study by Petrov and colleagues showed that following removal of the MBP fusion partner, ΔK122 exists in a monomer-dimer equilibrium in solution and that upon incubation with 2-methyl-2,4-pentanediol (MPD), high molecular weight oligomers (fibrils and PNTs) are formed.[97] This monomer-dimer equilibrium can be observed in a native PAGE gel, which shows two distinct bands at approximately 13 and 26 kDa immediately after CIEX purification (Figure 3.4). To further confirm the presence of this equilibrium, native ESI-IMS-MS was employed. A sample of 10 µM ΔK122 shows two conformations present at the +4 monomeric charge state corresponding to an m/z charge of 3210, as well as low-abundance species at m/z 2869 and 3669, which correspond to ΔK122 dimers with odd charge states (Figure 3.5).

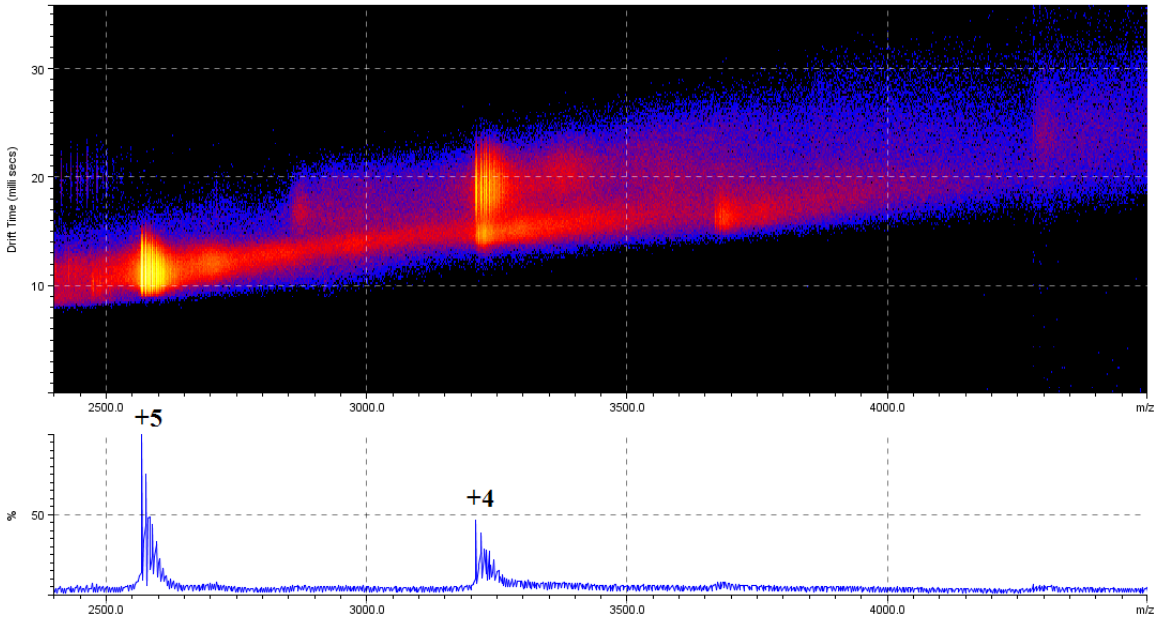


Figure 3.5: Observance of the monomer-dimer equilibrium by ESI-IMS-MS via a 3D Driftscope plot (*m/z* vs. drift time vs. intensity). Monomer peaks corresponding to the +5 (*m/z* = 2568) and +4 (*m/z* = 3210) charge states are predominant. In addition, there are faint dimer peaks present in between the monomer peaks in the spectrum as well as ion mobility plot.

3.3. Time-Resolved ElectroSpray Ionization Hydrogen-Deuterium eXchange (TRESI-HDX) Mass Spectrometry on the ΔK122 Pilin

The observation that the MBP tag facilitates the maintenance of the ΔK122 pilin as a monomer, and removal of the MBP results in ΔK122 entering into a monomer-dimer equilibrium is advantageous for the characterization of each state via local hydrogen-deuterium exchange. The MBP-ΔK122 fusion acts as the "time 0" monomer, while analysis of ΔK122 following removal of MBP provides details on the monomer-dimer equilibrium. Analysis of the digested MBP-ΔK122 fusion protein proved challenging due to spectral overlap of peptides originating from the large MBP tag. Despite these challenges, a 59% sequence coverage for the monomer was obtained. Removal of the tag and analysis of the monomer-dimer equilibrium state allowed for 73% sequence coverage (Figure 3.6).

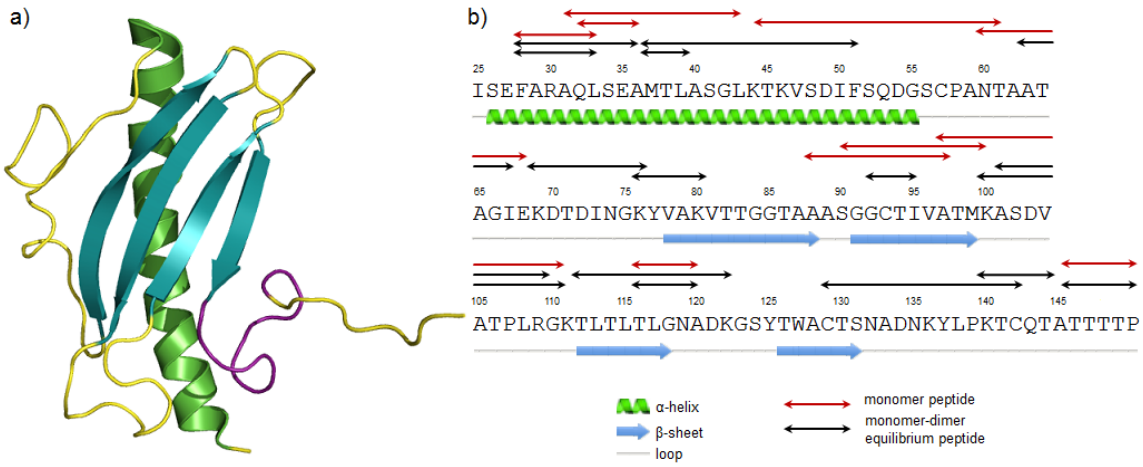


Figure 3.6: Structural elements of the ΔK122 pilin, and the peptides analyzed in this study. a) The common structural features of type IV pilins are highlighted on the ΔK122 monomer: the four-stranded antiparallel β-sheet (cyan), the truncated N-terminal α-helix (green), connecting loop regions (yellow), and the receptor binding D-region (purple). b) The secondary sequence is numbered taking into account the truncated portion of the α-helix. The peptic peptides obtained in this study are highlighted using red and black double ended arrows.

The mapped structure for the monomeric state of the pilin is colour-coded based on the deuterium uptake at the final time point obtained, corresponding to a labelling time of 2000 ms (Figure 3.7b). The N-terminal region of the amphipathic α-helix, in particular the peptide FARAQLSEA (a.a. 28-36), shows high dynamic flexibility; this peptide is close to Gly42 (Figure 3.6b), which induces a kink in the helix due to its increased stereochemical flexibility.[75] The observed dynamic flexibility of this peptide is consistent with NMR data for the ΔK122 pilin ($\Delta K122^{NMR}$),[74] which showed that the N-terminal end of the α1-C helix is deflected away from the β-sheet. In addition, terminal bulky substituents within α-helices are unable to fulfill the hydrogen bonding requirements due to steric hindrances. Accordingly, a comparison of $\Delta K122^{NMR}$ and crystallographically determined structures of the protein[75] reveals that an otherwise conserved hydrogen bond between the side chain O^{ε} of Gln 32 and the amide nitrogen of Ala 105 is missing. This could partially account for the observed high exchange rate as the helix is afforded more relative motion from the β-sheet and has less local i+4 hydrogen bond stability. These data suggest that the packing of the hydrophobic core in ΔK122, specifically the N-terminal end of the helix, is less tight thereby allowing for structural flexibility. In contrast, the C-terminal region of the helix exhibits relatively low deuterium uptake, suggesting that this end of the helix exhibits more stable hydrogen bonding and is more tightly packed alongside one or more strands of the β-sheet.

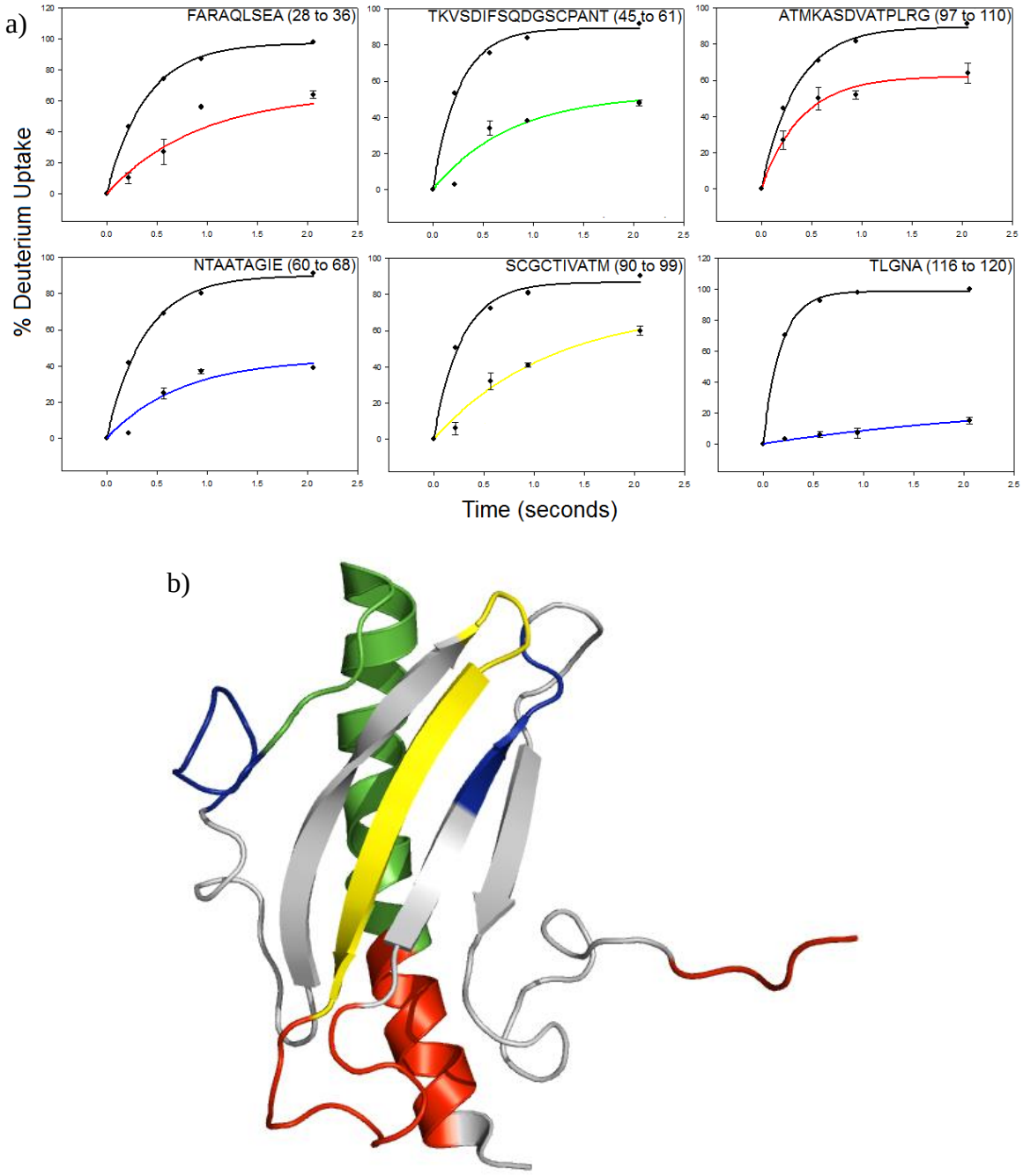


Figure 3.7: HDX analysis of the ΔK122 monomer. a) Representative kinetic plots of % deuterium uptake vs. time for 6 peptides from ΔK122. The intrinsic rate of the peptide is shown in black while the experimental uptake is shown in colour. Error bars represent standard deviation values where 3 or more replicates were obtained. b) Differing levels of deuterium uptake for the monomer mapped onto the structure of ΔK122. The measured profiles are coloured according to total deuterium uptake: red (61-100%), yellow (51-60%), green (41-50%) and blue (0-40%); regions for which no peptides were observed are coloured in grey.

The connecting loop between the second and third strands of the antiparallel β-sheet also exhibits some of the highest exchange rates in the monomer, which is not surprising due to the lack of a stabilizing H-bonding network between structural elements. The second strand of the sheet shows moderate-to-high deuterium uptake throughout. The moderate-to-high uptake may be a the result of the α-helix lying at a 45° angle relative to the surface of the β-sheet, as observed in $\Delta K122^{NMR}$,[74] allowing for more overall solvent exposure in parts of the network. The low level of exchange for the NTAATAGIE peptide (a.a. 60-68) located within the αβ-loop was unexpected, and suggests that it is tightly packed in the monomeric state forming a loop-protein packing interface. Finally, the C-terminal end of the protein shows complete deuterium uptake, which is consistent with NMR observations[74] suggesting that this region of the protein is highly mobile in solution. Representative kinetic plots for the monomeric peptides are shown in Figure 3.7a.

Following removal of the MBP tag, ΔK122 enters a monomer-dimer equilibrium; TRESI-HDX analysis of the protein results in a sequence coverage of 73% (Figure 3.8a). Examination of the relative deuterium uptake levels of the protein in this state can provide insight on the changes in dynamic flexibility, solvent exposure, and the domains likely involved in protein-protein interactions. Representative spectra of peptides obtained over the time course of the reaction exhibiting shifts in isotopic distribution upon exposure to deuterium are shown (Figure 3.8b). The amphipathic α-helix (the α1-C region of the protein) shows a considerable decrease in deuterium uptake compared to the monomer. Indeed, the FARAQLSEA (a.a. 28-36) peptide shows a decrease from 64% ± 2.29 to 49% ± 4.50, and the region spanning ASGLKTKVSDIF (a.a. 40-51) shows a

decrease from 48% ± 1.52 to 32% ± 2.12. These data suggest that stabilization of the truncated monomer is a result of interaction along the α-helix. Interestingly, the connecting loop region spanning KASDVATPLRG (a.a. 100-110) also shows a significant decrease in uptake, suggesting that it is either directly involved in protein-protein interaction within the dimer or serves to stabilize the interaction.

In the equilibrium state, the antiparallel β-sheet network shows variable exchange rates ranging from low to high deuterium uptake. In particular, the GCTI (a.a. 92-95) and TLGNA (a.a. 116-120) peptides exhibit increased uptake in the monomer-dimer equilibrium. This may be the result of destabilization of the hydrogen bonding network as the α-helix involved in protein-protein interactions moves away from the β-sheet. It is possible that the truncated helix shifts outward in order to form the interaction required for dimerization prior to fibril assembly. This is consistent with NMR data of the pilin,[74] which shows that the α-helix is less tightly packed on the β-sheet prior to oligomerization. In addition, the ATAGI (63-67) peptide within the αβ-loop shows high uptake relative to the monomer, indicating a state of increased dynamic flexibility. This could be the result of the loop having to move away in order to accommodate the incoming protein for dimerization. Representative kinetic plots of deuterium uptake over time for the peptides showing either a decrease or increase relative to the monomer are shown (Figure 3.8c).

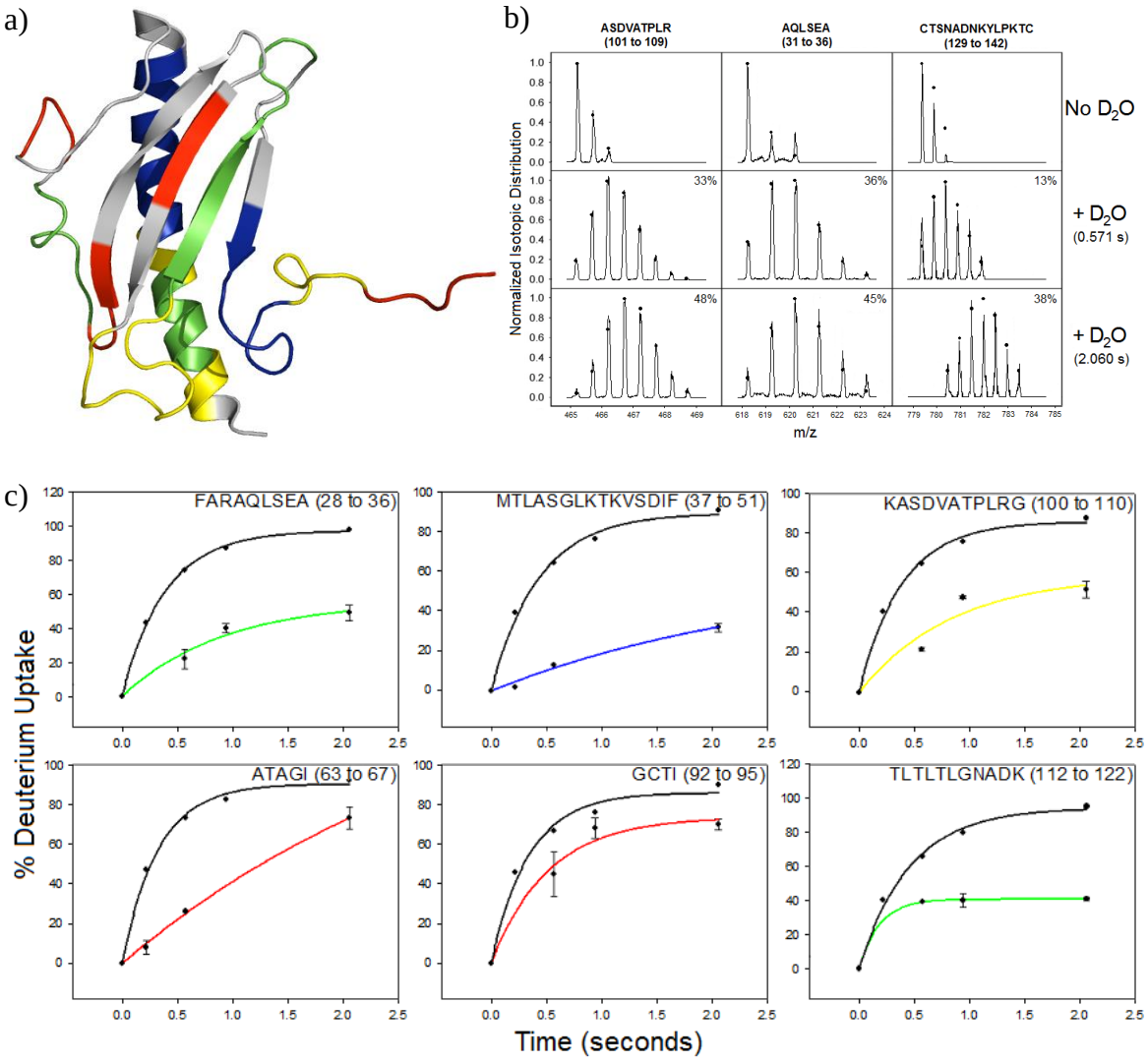


Figure 3.8: HDX analysis of the ΔK122 monomer-dimer equilibrium. a) Differing levels of deuterium uptake for the monomer-dimer equilibrium mapped onto the solution structure of ΔK122. The measured profiles are coloured according to total deuterium uptake: red (61-100%), yellow (51-60%), green (41-50%) and blue (0-40%); regions for which no peptides were observed are coloured in grey. b) Site-specific HDX analysis for the equilibrium state. Representative spectra of peptides for non-deuterated ΔK122 (top panel), and upon incubation with deuterium at 0.571 seconds and 2.060 seconds. Raw spectra exhibit shifts in isotopic distribution upon exposure to deuterium. Percent deuterium uptake is indicated on each spectrum. c) Representative kinetic plots of % deuterium uptake vs. time for 6 peptides from ΔK122. The intrinsic rate of the peptide is shown in black while the experimental uptake is shown in colour. The top three panels represent peptides that show a decrease in deuterium uptake in the monomer-dimer equilibrium while the bottom three panels represent peptides that show an increase in deuterium uptake. Error bars represent standard deviation values where 3 or more replicates were obtained.

One of the regions that was not analyzed in the monomeric protein but appears for the equilibrium state is the disulfide bonded receptor binding domain, or D-region. The D-region consists of a type I followed by a type II turn forming a V-shaped groove whose side chains protrude toward the protein interior, in particular toward the α1-C helix, stabilizing the double-turn conformation.[73,74] Accordingly, the peptide CTSNADNKYLPKTC (a.a. 129-142), which falls within the D-region, displays some of the lowest uptake levels of the protein. This indicates that the backbone amide hydrogens are hidden from the solvent and point up into the pocket, which correlates with crystallographic observations of ΔK122.[75] It is interesting to note that of the seven conserved residues in the globular domain of pilins originating from *Pseudomonas* three are located within the D-region, Cys 129 and Cys 142, which form a disulfide bridge that bounds the D-region, and Pro 139, which initiates the second β-turn of this region.[65,73,75] It is therefore not surprising that the disulfide bond and proline residue would increase the rigidity of this region as they restrict conformational flexibility. In previously solved pilin structures, the side chains of the cysteine and proline residues are buried towards the core of the protein, defining a packing interface.[75] This agrees with previous NMR data showing that the disulfide loop exhibits a rigid backbone conformation essential for attachment to host cell receptors.[142] In addition, interactions between the αβ-loop and D-region have been proposed for the K122-4 pilus based on the charge complementarity in these regions.[51,69,71,74,80] Following the D-region is the disordered C-terminal loop which continues to show high deuterium uptake at equilibrium, albeit slightly lower compared to it being fully saturated in the monomer.

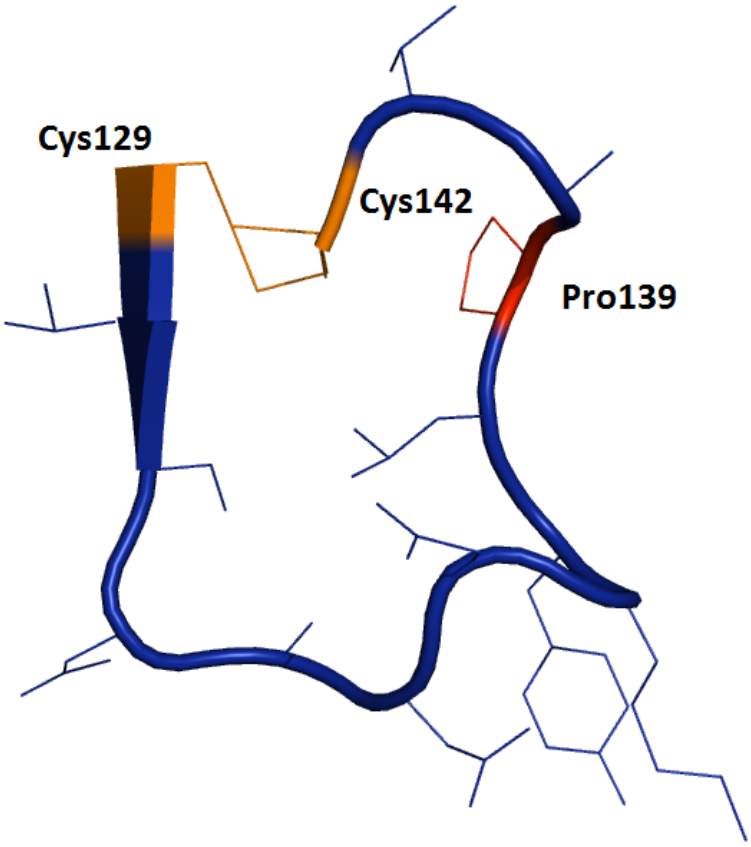


Figure 3.9: A close-up view of the CTSNADNKYLPKTC (a.a. 129-142) D-region peptide. Conserved peptides are highlighted, Cys129 and Cys142 forming a disulfide bridge and Pro139, all of which contribute to the structural rigidity of the D-region.

The use of pilin-derived PNTs for biomedical applications is attractive due to their well-defined structures, assembly under physiologically relevant conditions, and easy manipulation through protein engineering approaches. However, in order to fully exploit these structures, an understanding of their assembly from monomer (pilin) to polymer (pilus) is required. Currently, there are structural models for the monomer (the pilin) and the polymer (the pilus), and the generation of the T4P from the monomeric pilin by the bacterium is reasonably well understood. On the other hand, the truncated ΔK122 pilin does not have (a) the bacterial machinery to guide assembly nor (b) the hydrophobic N-terminal region of the α-helix to hydrophobically drive pilin

oligomerization into PNTs. An experimentally derived structure of a full length pilin-derived PNT remains elusive, and the mechanism of oligomerization is less well understood, although it has been shown that it follows a fibril-mediated process.[97] Our observation that the pilin enters a monomer-dimer equilibrium upon purification away from its MBP fusion partner (Figure 3.4, 3.5) allowed us to explore a very early time point in the oligomerization process. It was not fully unexpected to observe the rigidity of the receptor binding D-region (Figure 3.9), especially noting that this region is structurally maintained for receptor binding despite sequence diversity among the *P. aeruginosa* pilins.[75] The observation that peptides from the N-terminal region of the α-helix, in particular the FARAQLSEA (a.a. 28-36) peptide, show a decrease in deuterium uptake and therefore reduced flexibility, indicates that while the full helix is not present, this portion of the helix plays an important role in the early points of pilin fibril formation. In addition, the increased stability of the loop connecting the second and third strands of the anti-parallel β-sheet suggests that it also plays a role in dimerization of the pilin, perhaps through increased interaction with the N-terminal region of the α-helix. Also, the increased flexibility observed in the αβ-loop region may be a result of a structural requirement induced by the altered packing of the α-helix onto the β-sheet in the dimer, resulting in the occlusion of the D-region in the PNT structure. Together, these observations point to regions of the pilin that may be useful to exploit in terms of varying oligomerization kinetics and or fibril/PNT stability; research is on-going to further characterize the structural and mechanistic requirements of pilin-derived PNT oligomerization, and to develop these structures for biomedical applications.

3.4. Discussion

In the current study, we analyzed the dynamics the ΔK122 pilin protein, both in the monomeric state (as a MBP-ΔK122 fusion protein) and in the pre-fibrillar dimeric state at the millisecond time scale using TRESI-HDX-MS. A summary of the regions showing significant increases and decreases between the two states is mapped onto the protein sequence (Figure 3.10). Results point toward the site of protein interaction mainly occurring along the amphipathic α-helix as well as possible involvement of a connecting loop. This is consistent with the α-helix having high sequence conservation among the pilins (for instance in *P. aeruginosa* strains K122-4, PAK, PAO, and *N. gonorrhoeae* strain MS-11).[74] In addition, variable exchange rates ranging from moderate to high in the β-sheet network and αβ-loop indicate that these regions show increased dynamic flexibility in the dimer. Finally, the D-region responsible for receptor binding shows low deuterium uptake in the monomer-dimer equilibrium and suggests both structural rigidity and solvent protection.

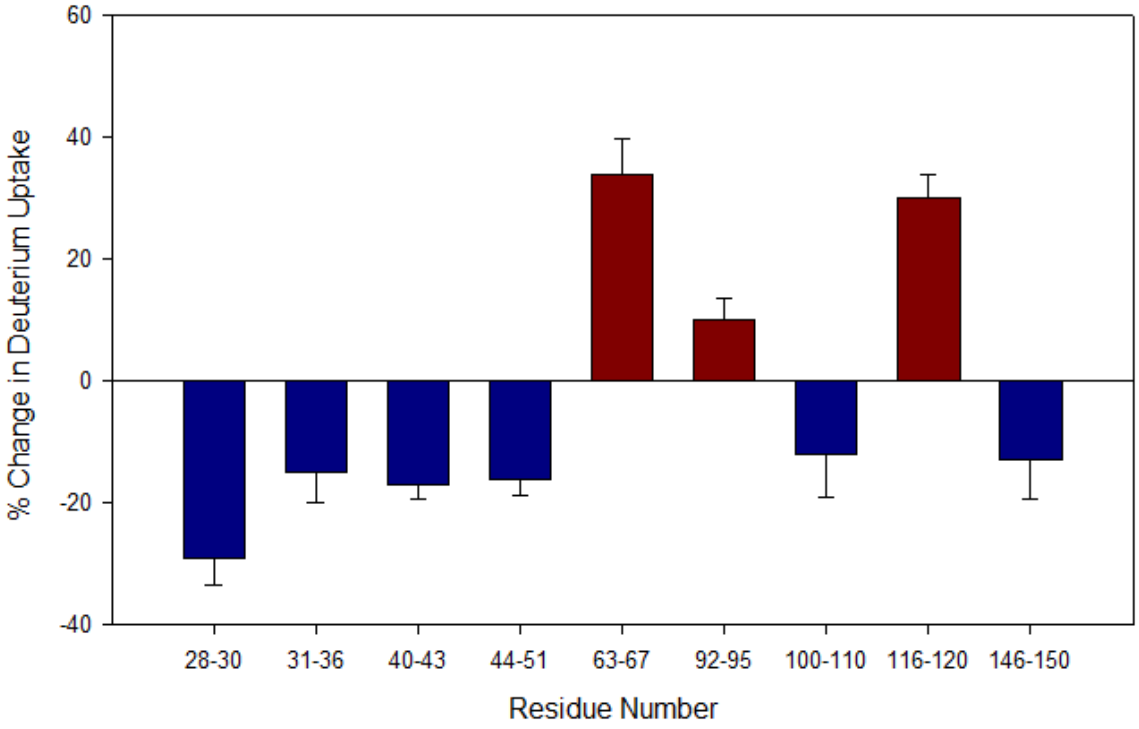


Figure 3.10: Localization of significant % changes in deuterium uptake occurring in the monomer-dimer equilibrium. Significant differences in the relative deuterium uptake in overlapping stretches of the protein are shown, with error bars representing the standard deviation of 3 or more replicates.

Overall, the results indicate that the truncated α-helix shifts in order to increase helix-helix interactions in the dimer, imparting stabilization of the dimeric precursor prior to coalescence into protein fibrils. The proposed interaction mechanism of dimerization (Figure 3.11) results in a stabilization of the α-helix through the protein-protein interaction interface. Movement of the helix away from the β-sheet network might account for the destabilization seen in peptides falling within this region.

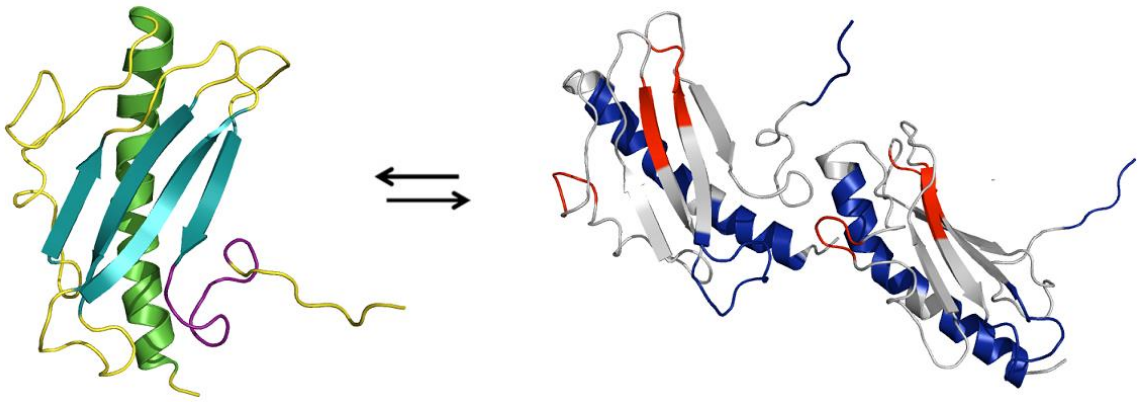


Figure 3.11: Proposed interaction mechanism of dimerization for the ΔK122 pilin. The structural elements of the monomer are coloured as described in Figure 1. In the dimer (right), peptides coloured in blue represent a significant decrease in uptake for the equilibrium state, and those in red represent a significant increase in uptake. Based on changes in hydrogen-deuterium exchange, interactions occur between the amphipathic α-helix in the dimer, as well as possible involvement of a connecting loop. The truncated helix shifts further outward in order to form the interaction required for dimerization, destabilizing the β-sheet network. The model of the dimer was generated using a 3-start helical assembly model for the T4P.[71]

Understanding the structural changes that occur when the protein enters its dimeric state is of great importance as it is the earliest intermediate leading up to PNT formation. Stabilization of the pilin dimer should result in more rapid fibril formation and PNT oligomerization. Characterization of the structural changes that occur during the formation of pilin-derived PNTs can also provide a greater understanding behind the mechanism of T4P formation as well as allowing for the development of these structures for applications in bionanotechnology. Efforts are on-going in understanding how the pilin oligomerizes from this equilibrium state to higher molecular weight species forming protein-derived nanofibrils and eventually PNTs.

Chapter 4: Initial Stage Oligomerization of the K122-4 Type IV Pilin

4.1. Analysis of Triggered ΔK122 Using ESI-IMS-MS

In the previous chapter, the ΔK122 monomer-dimer equilibrium that occurs following the removal of the MBP-fusion tag and prior to the addition of a hydrophobic trigger molecule to solution was explored. The next step in gaining an understanding of the mechanism of pilin-derived PNT formation was to explore the oligomerization mechanism that occurs upon pilin incubation with a hydrophobic triggering solution leading to the formation of extended fibrils. A mass spectrum of the untriggered ΔK122 sample is shown in Figure 4.1, with monomeric charge states ranging from +8 to +3. The presence of the dimer is again confirmed at m/z 3668.5356 with a charge state of +7.

A previous study confirmed that optimal oligomerization of truncated ΔK122 occurs using MPD.[97] These results were confirmed using transmission electron microscopy as well as size exclusion chromatography in conjunction with multi-angle light scattering.[97] The goal of this study was to use ESI-IMS-MS in order to study the intermediates between the monomer-dimer equilibrium and higher order molecular weight species leading to fibril formation.

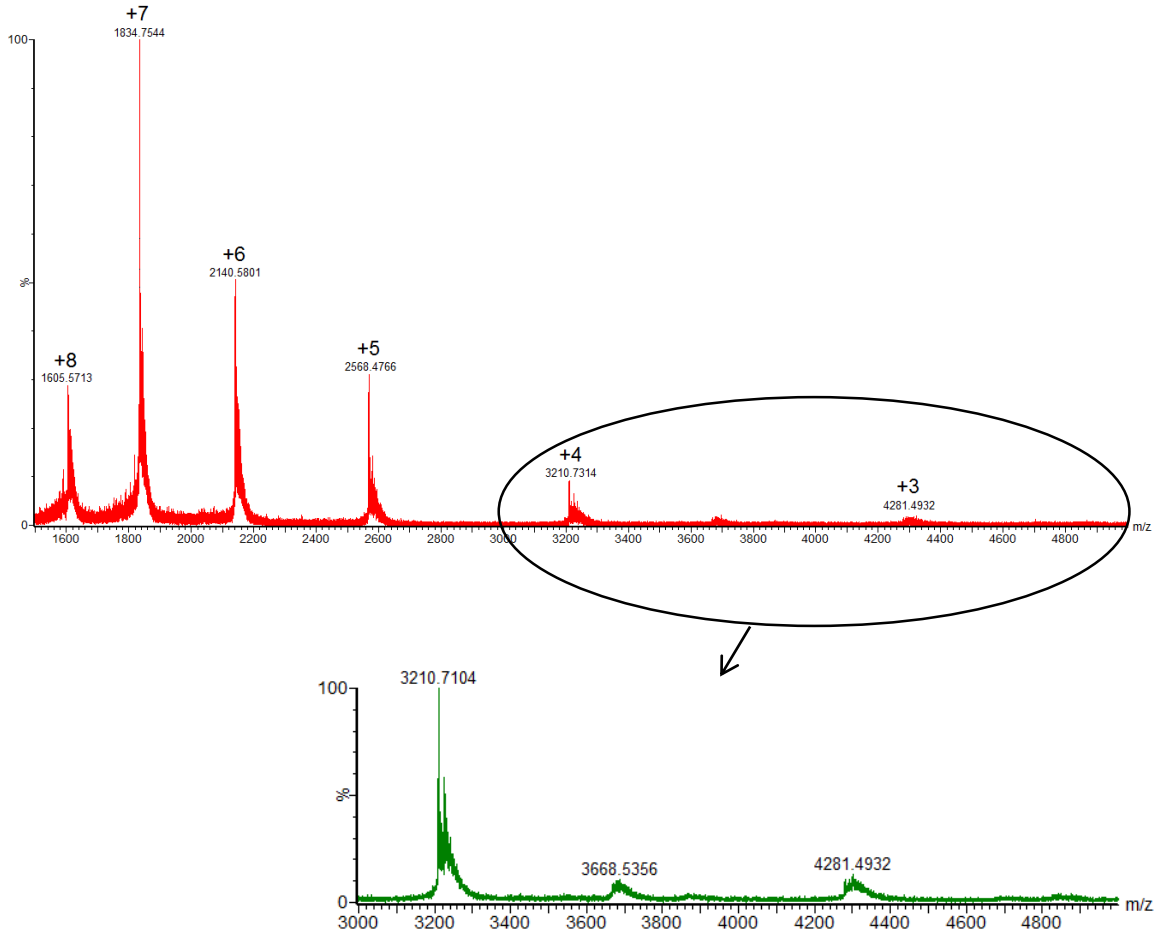


Figure 4.1: ESI mass spectrum of native ΔK122 at pH 6.9. Protonation states from +8 to +3 are observed for the monomer. Presence of the dimer is seen at m/z 3668.5356 with a charge state of +7.

ΔK122 concentrated to 20 µM was incubated with MPD at a 15:1 ratio to protein and monitored immediately after triggering (5-10 minutes), and hourly for up to 8 hours. Triggering with MPD alone caused various issues, including intercapillary clogging and extensive degradation of the protein. This degradation could be the result of increased shearing forces within the capillary as the protein oligomerizes.[143] Little to no significant changes were observed as the sample was analyzed at 2 hour intervals for up to 8 hours. In order to improve the spray and avoid aggregation, methanol was added along with

MPD, at a 3:1 ratio to MPD. High molecular weight species were in low abundance even after an 8 hour incubation period but some inferences can be made when analyzing the intermediate time points.

It should be noted that even without the addition of a triggering solution, ΔK122 oligomerizes to a small extent when stored in polypropylene tubes and to a lesser extent when stored in glass tubes. This can be seen in the spectrum acquired for untriggered protein which has a small population (ranging from 0.4% to 1.1%) of trimers to 14-mers. Previous crystallographic analysis of ΔK122 only shows the monomer, not any dimer or higher order fibril species.[75] Therefore, studying the process in solution allows us to look at the dynamic process that gets selected against in the crystallization process. At the 10-20 minute time point after triggering, the population of dimers in the 3000-5000 m/z range decrease to almost zero while the trimer population reaches 5.4%. At the 4 hour mark, there is a decrease in the trimer population with a steady rise in all other oligomeric species ranging from tetramers to 14-mers. A summary of the results for one of the 8 hour time trials is shown in Table 4.1. Comparison of these values to similar runs shows variation between the relative abundance of the oligomeric species, which suggests that the pathway to fibril formation does not follow a discreet step-by-step pathway *in vitro*. In addition, intercapillary clogging was a persistent issue that was never truly overcome even with the addition of methanol and acetonitrile within the triggering solution.

Noting the intercapillary clogging that was observed during the above experiments, it was thought that perhaps a better way to study the early intermediates leading to full fibril formation was to perform the experiment using a TRESI mixer, where the triggering solution is introduced at the millisecond to second time scale. The

results of this experiment are listed in Table 4.2. There were no observed increases in the relative abundance of higher molecular weight species across the time points measured (up to 4.82 seconds). This suggests that the formation of intermediates leading to fibrils likely occurs within the minute time scale and were therefore not detected using the TRESI mixer setup.

Table 4.1: Relative % of species before and after triggering with a 15:1 (v/v) ratio of MPD to protein and a 3:1 (v/v) ratio of methanol to MPD over an 8 hour time course.

	Monomer	Dimer	Trimer	Tetramer	Pentamer	Hexamer	Heptamer	Octamer	Nonamer	Decamer	14-mer
Size (kDa)	12,839.2	25,678.4	38,517.6	51356.8	64,196.0	77,035.2	89,874.4	102,713.6	115552.8	128,392	179,748.8
Relative % Abundance Before Triggering	100%	27%	0.4%	1.1%	0.9%	0.5%	0.25%	1.0%	1.1%	0.8%	0.9%
Relative % Abundance 10-20 min	100%	0.54%	5.4%	0.21%	0.19%	0.29%	0.29%	0.08%	.095%	.105%	0.23%
Relative % Abundance 4 hours	100%	10%	1.25%	0.56%	0.22%	0.60%	0.44%	0.25%	.37%	.52%	0.60%
Relative % Abundance 8 hours	100%	17%	9%	0.95%	0.33%	1.8%	1.15%	0.50%	1.5%	1.6%	0.46%

Table 4.2: Relative % of species before and after triggering with a 15:1 (v/v) ratio of MPD to protein over a milliseconds-seconds time course.

	Monomer	Dimer	Trimer	Tetramer	Pentamer	Hexamer	Heptamer	Octamer	Nonamer	Decamer	14-mer
Size (kDa)	12,839.2	25,678.4	38,517.6	51356.8	64,196.0	77,035.2	89,874.4	102,713.6	115552.8	128,392	179,748.8
Relative % Abundance Before Triggering	100%	45%	0.49%	0.009%	0.22%	0.36%	0.47%	1.25%	1.40%	0.80%	1.70%
Relative % Abundance 0.039 sec	100%	45%	1.15%	0.0095%	0.125%	0.53%	0.24%	0.80%	0.37%	0.39%	0.90%
Relative % Abundance 0.087 sec	100%	32%	0.80%	0.0098%	0.165%	0.49%	0.29%	0.80%	0.60%	0.53%	0.75%
Relative % Abundance 0.183 sec	100%	35%	0.37%	0.005%	0.04%	0.13%	0.21%	0.54%	0.32%	0.17%	0.22%
Relative % Abundance 0.518 sec	100%	43%	0.58%	0.005%	0.04%	0.20%	0.20%	0.51%	0.38%	0.14%	0.24%
Relative % Abundance 4.822 sec	100%	29%	0.28%	0.8%	0.06%	0.1%	0.29%	0.45%	0.31%	0.19%	0.48%

4.2. Global TRESI-HDX Mass Spectrometry

To try and obtain insight into the global changes in the structure of the monomeric population before and after the oligomerization process, hydrogen-deuterium exchange was performed on 20 µM protein in its monomer-dimer equilibrium state and 20 µM of the protein after incubation with MPD for 24 hours. Concentrations higher than this lead to intercapillary clogging.

The deuterium uptake for the monomeric charge state of +5 was compared between the two states (Figure 4.2). Upon incubation with 80% deuterium, the monomer in the monomer-dimer equilibrium state exhibited an average uptake of 32 deuterium atoms. On the other hand, the monomers in equilibrium with high-molecular weight fibril species in the oligomerized sample exhibited an average uptake of 16 deuterium atoms.

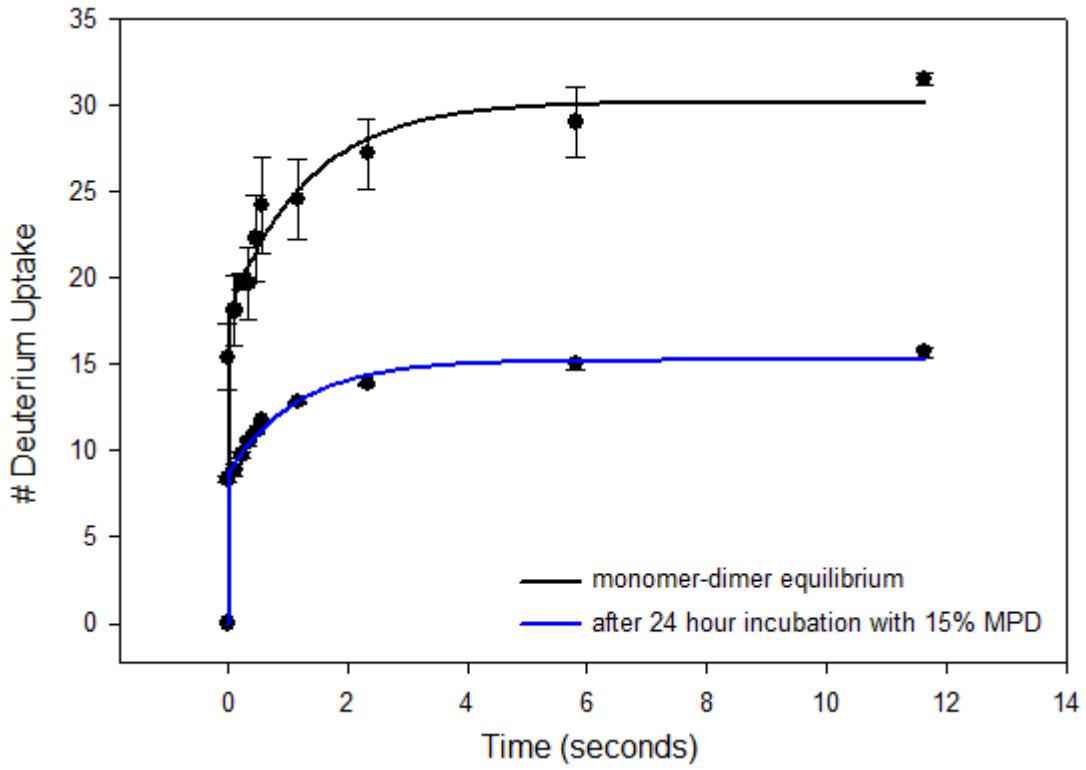


Figure 4.2: Deuterium uptake of ΔK122 in the monomer-dimer equilibrium and after 24 hours of incubation with 15% MPD.

4.3. Discussion

This study aimed to characterize the early stages of filament formation of ΔK122 before association into extended fibres by time course studies using ESI-IMS-MS both on the minutes-hours time scale as well as on the millisecond-second time scale. However, no conclusions can be made at this time about a discrete mechanism leading to fibril formation, as higher molecular weight species are not observed in great abundance and no reproducible trends were detected. The fact that the data could not be reproduced suggests that there might be multiple distinct pathways leading to fibril coalescence, and is much more complex without the presence of an *in vivo* T2SS.

One hypothesis for the low abundance of higher molecular weight species throughout the 8-hour time course suggests that perhaps the oligomers are broken apart during the ESI process. If the interactions that form the fibrils are not stable in the gas phase, it wouldn't be surprising that they are broken down to mostly monomeric species during ionization and thus not detected throughout the time trial.

Analysis of the global TRESI-HDX-MS data suggests that the oligomeric structures formed after 24 hours are more compact and have reduced areas available for deuterium exchange. Further work will need to be done in order to localize the areas affected in the oligomerized species compared to the monomer-dimer equilibrium. One problem that will need to be circumvented for local analysis is precipitation of the fibrils upon entering the acidic proteolysis chamber.

Chapter 5: Structural Studies of TraF, and the TraF-TraH Interaction

5.1. Expression and Purification GST-TraF

Optimal expression of GST-TraF after induction with IPTG was observed in one-hour increments for a total of 4 hours at 30°C (data not shown). GST-TraF (~51 kDa) was purified using a column packed with glutathione Sepharose 4B beads. As can be seen in Figure 5.1, the GST-TraF protein eluted at a peak maximum absorption of approximately 2250 mAU and 100% elution buffer. The blue line represents absorbance at 280 nm and the red line represents absorbance at 215 nm. The corresponding SDS-PAGE gel is shown in Figure 5.2.

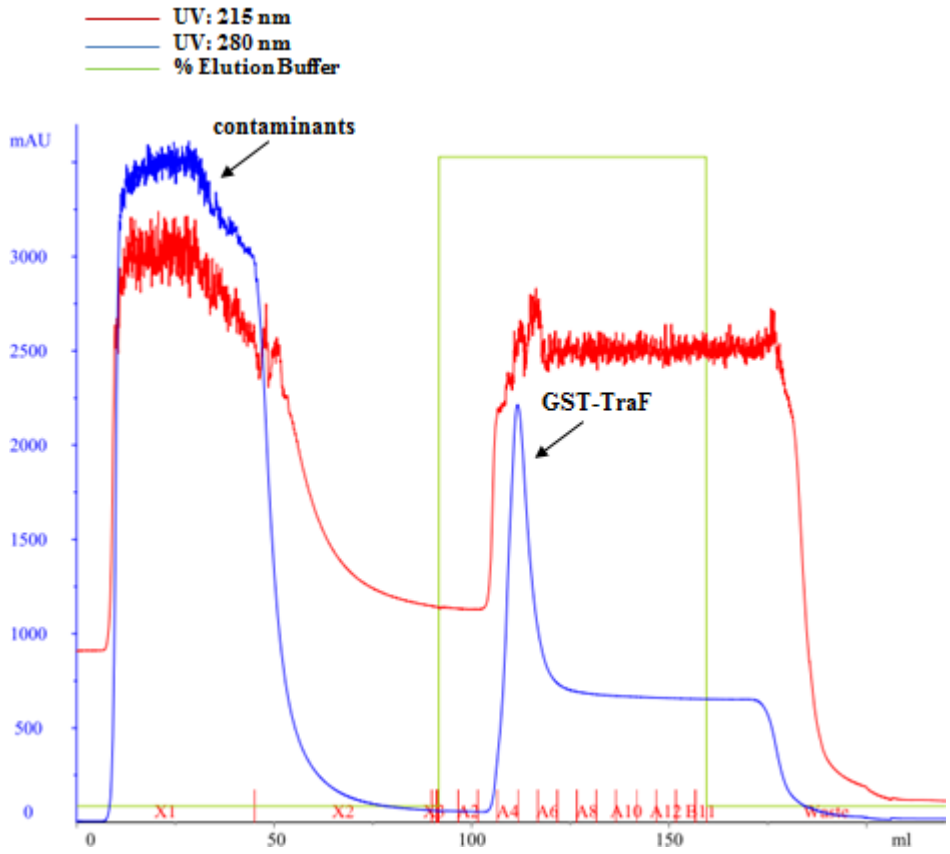


Figure 5.1: Affinity chromatogram of GST-TraF. The GST-TraF protein eluted at a peak maximum absorption ~ 2250 mAU. The blue line is absorbance at 280 nm, the red line is absorbance at 215 nm and the green line is the elution buffer concentration.

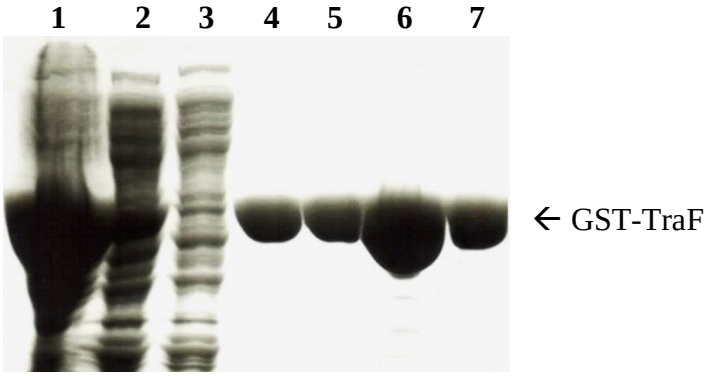


Figure 5.2: SDS-PAGE gel of GST-TraF affinity purification and concentration. Lane 1: pellet. Lane 2: Supernatant after sonication. Lane 3: Flow through. Lanes 4 and 5: elution samples containing the protein of interest GST-TraF at approximately 50 kDa. Lanes 6 and 7: GST-TraF concentrated to 11.6 mg/mL and 4.6 mg/mL.

Previous efforts to obtain the full structure solution using molecular replacement methods of TraF has been challenging due to the lack of high sequence identity with a related thioredoxin of known structure that would represent only 50% of the total scattering mass of the protein. To obtain structural insights on the full-length TraF, the protein was expressed as a GST-fusion, with the target of utilizing the GST tag as a handle for molecular replacement calculations. In addition, inclusion of the tag allows for a more soluble protein and concentration to levels required for mass spectrometric analysis. The purified fusion was subjected to both TRESI-HDX-MS analysis as well as crystallization trials in order to obtain structural insights.

5.2. Dynamic Analysis of GST-TraF

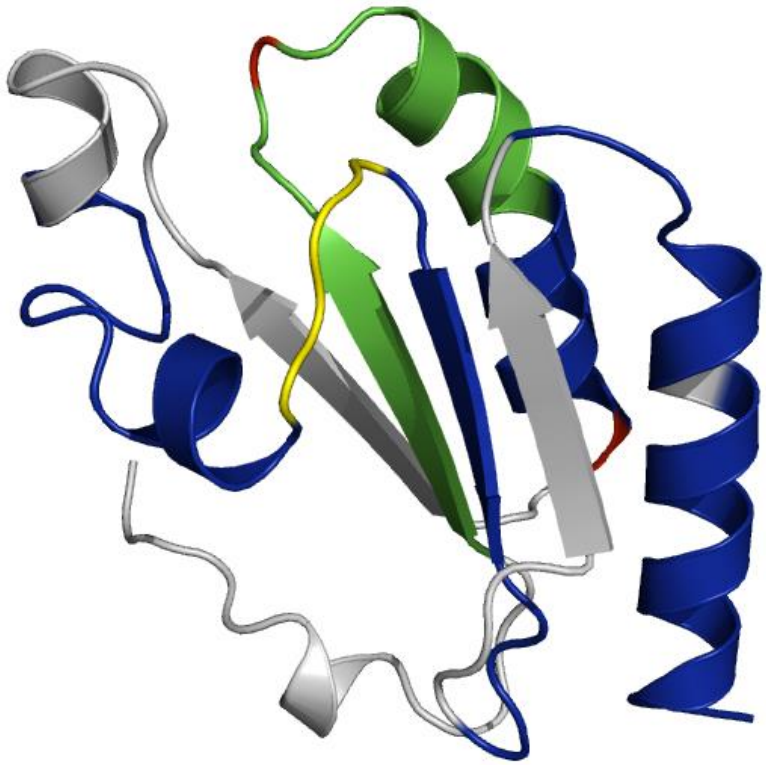
As mentioned previously (Section 1.2.1.2), TraF has a predicted thioredoxin-like fold placing it among the thioredoxin-like superfamily of enzymes. The thioredoxin fold consists of at least three α-helices, a four-stranded β-sheet and a CxxC active site. In TraF, the fold forms the C-terminal domain which is preceded by a domain of unknown

function.[31] A previous study on *E. coli* thioredoxins confirmed that residues incorporated in α-helices and β-sheets show low incorporation of deuterium while loops and turns had high deuterium incorporation.[144]

The isotopic distribution for each peptide was analyzed for each reaction time with combined runs yielding an overall sequence coverage of 75%. A homolog of closest sequence coverage for the C-terminal domain was acquired using PHYRE 2.0 (Protein Homology/analogY Recognition Engine V 2.0).[33] This homolog is the thioredoxin of C-type cytochrome biogenesis protein *DipZ* from *Mycobacterium tuberculosis*. Using the saturation levels of each peptide, a heat map of the structure for the C-terminal homolog is shown in Figure 5.3a. Similarly, a heat map of the predicted full-length secondary structure of TraF was generated and is shown in Figure 5.3b.

In general, the C-terminal region contains most of the peptides observed to have low deuterium exchange levels (0-55%). This is expected since it is the region of the protein predicted to contain a defined structure. The N-terminal region contains most of the peptides observed to take up moderate-high (66-75%) and high deuterium levels (76-100%). This indicates that the region is much more dynamic relative to the C-terminus, either due to lack of structure or by being more exposed to the solvent. In addition, peptides falling within the region hypothesized to interact with TraH (a.a. 47-70 and 68-73) show moderate-high and high deuterium uptakes respectively.

a)



b)

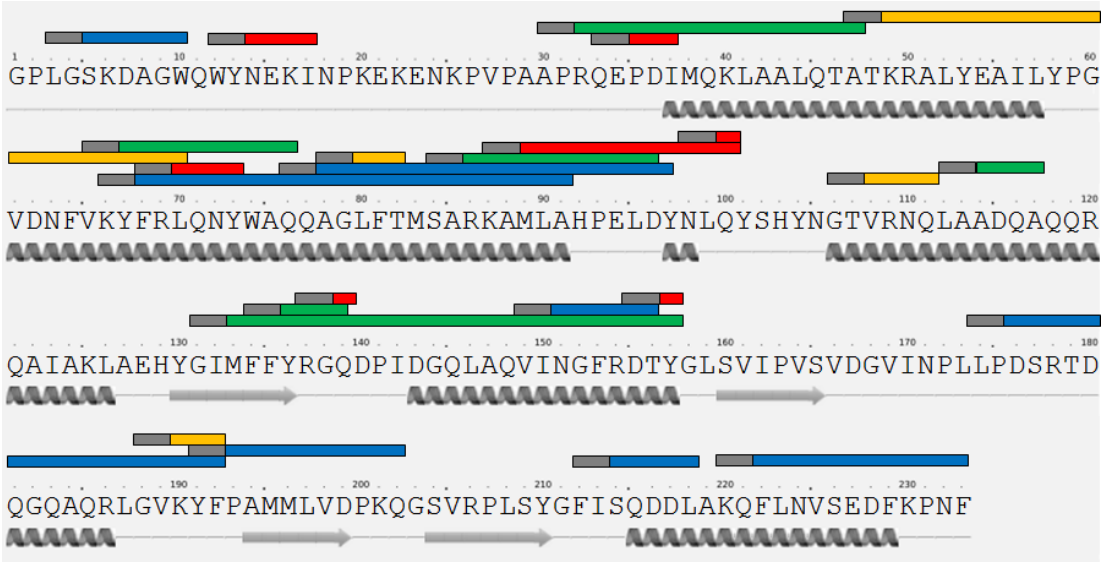
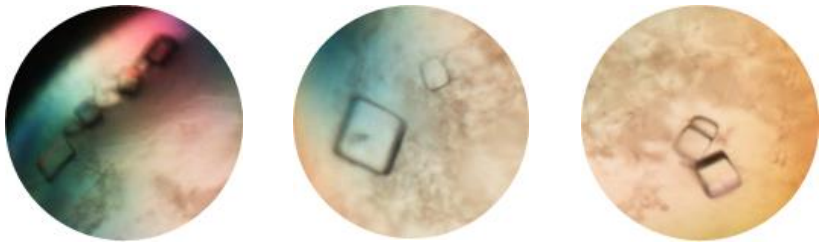


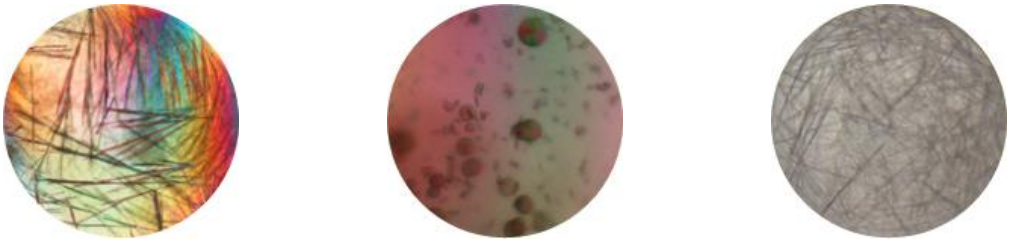
Figure 5.3: HDX analysis of GST-TraF. a) Deuterium exchange profile mapped onto the crystal structure of a C-terminal homolog for TraF. The structure is coloured based on the differences in observed amplitude of deuterium uptake: red (76-100%), yellow (66-75%), green (55-65%), and blue (0-55%). Regions that were not analyzed are coloured in grey. b) Deuterium exchange profile mapped onto the full-length secondary structure for TraF. Grey colouring indicates the rapid loss of deuterium label for the first two residues in each peptide due to back exchange.

5.3. Crystallization of GST-TraF

A complete list of preliminary GST-TraF crystal hits obtained during this research project can be found in Table 5.1. A few of the most promising hits are shown in Figure 5.4. Unfortunately, none of the initial crystal hits screened were suitable for collection of diffraction data. The hits obtained from the Hauptman-Woodward Medical Research Institute served as screens only and could not be screened for diffraction quality. Needle and spherulite crystals obtained in-house were not suitable for data collection and will need to be further optimized.



GST-TraF-02
MSCG Core III #86 (8.2 mg/mL)
0.1 M Ammonium Acetate pH 4.5
3.0 M Sodium Chloride



GST-TraF
MSCG Core II #39 (5.4 mg/mL)
0.2 M Lithium Sulfate
0.1 M CAPS:NaOH pH 10.5
2.0 M Ammonium Sulfate

GST-TraF
MSCG Core III #90 (8.2 mg/mL)
0.2 M Zinc Acetate
0.1 Imidazole:HCl pH 8.0
20% (v/v) 1,4-Butanediol

GST-TraF-Hauptman
Cocktail #15_C0938 (6.4 mg/mL)
0.1 M Manganese Chloride Tetrahydrate
0.1 M Bis-Tris Propane pH 7.0
80% (w/v) PEG 4000

Figure 5.4: GST-TraF protein crystals. The protein concentration and crystallization conditions are found within the figure.

Table 5.1: Description of conditions for preliminary crystal hits.

Sample Name (Date Plated)	X-tal Kit & Condition #	Concentration (mg/mL)	Salt	Buffer	Precipitant
GST-TraF-01 (Sep 18th, 2013)	MCSG Core I #64	11.6	0.2 M magnesium chloride	0.1 M sodium citrate:citric acid pH 5.5	40% (v/v) PEG 400
GST-TraF-01 (Sep 18th, 2013)	MCSG Core II #60	11.6	0.2 M lithium sulfate	0.1 M sodium cacodylate:HCl pH 6.5	30% (v/v) PEG 400
GST-TraF-02 (Sep 19th, 2013)	MCSG Core III #86	8.2	-	0.1 M sodium acetate pH 4.5	3.0 M sodium chloride
GST-TraF-02 (Sep 19th, 2013)	MCSG Core III #90	8.2	0.2 M zinc acetate	0.1 M imidazole:HCl pH 8.0	20% (v/v) 1,4-butanediol
GST-TraF-03 (Sep. 23rd, 2013)	MCSG Core I #54	5.4	0.2 M potassium sulfate	-	20% (w/v) PEG 3500
GST-TraF-03 (Sep. 23rd, 2013)	MCSG Core I #82	5.4	0.2 M sodium chloride	0.1 M CAPS:NaOH pH 10.5	20% (w/v) PEG 8000
GST-TraF-03 (Sep. 23rd, 2013)	MCSG Core II #5	5.4	0.2 M lithium sulfate	0.1 M CAPS:NaOH pH 10.5	1.2 M NaH_2PO_4/ 0.8 M K_2HPO_4
GST-TraF-03 (Sep. 23rd, 2013)	MCSG Core II #39	5.4	0.2 M lithium sulfate	0.1 M CAPS:NaOH pH 10.5	2.0 M ammonium sulfate
GST-TraF-03 (Sep. 23rd, 2013)	MCSG Core II #96	5.4	0.2 M zinc acetate	0.1 M sodium acetate:acetic acid pH 4.5	10% (w/v) PEG 3000
GST-TraF-04 (Oct. 8th, 2013)	MCSG Core I #15	3.6	0.2 M ammonium acetate	0.1 M Bis-Tris:HCl pH 5.5	25% (w/v) PEG 3350
GST-TraF-04 (Oct. 8th, 2013)	MCSG Core I #20	3.6	0.17 M sodium acetate	0.085 Tris:HCl pH 8.5	25.5% (w/v) PEG 4000 15% (w/v) glycerol
GST-TraF-04 (Oct. 28th, 2013)	MCSG Core I #17	3.6	0.2 M magnesium chloride	0.1 M Tris:HCl pH 8.5	25% (w/v) PEG 3350
GST-TraF-05 (Nov. 5th, 2014)	Cocktail Generation #15_C0938	6.4	0.1 M manganese chloride tetrahydrate	0.1 M Bis-Tris propane pH 7.0	80% (w/v) PEG 4000
GST-TraF-05 (Nov. 5th, 2014)	Cocktail Generation #15_C0785	6.4	0.1 M lithium chloride	0.1 M TAPS pH 9.0	40% (w/v) PEG 1000

5.4. Towards Structural and Protein Interaction Studies with MBP-TraH

Optimal cell expression of MBP-TraH after induction with IPTG was observed in one-hour increments for a total of 3 hours at 30°C (data not shown). MBP-TraH (~90 kDa) was purified using a gravity column packed with amylose resin at 4°C. Performing the purification at a lower temperature decreased both the degradation of the fusion as well as the cleavage of MBP. The corresponding SDS-PAGE gel is shown in Figure 5.5.

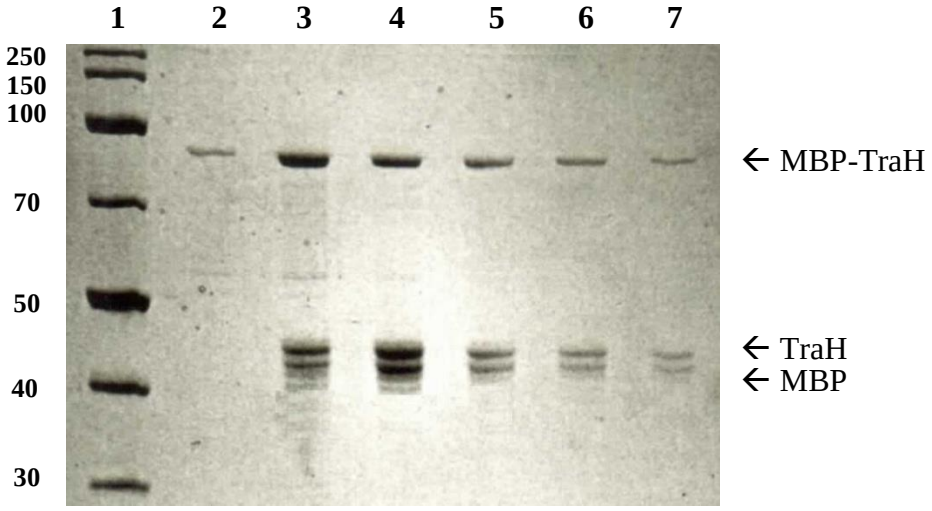


Figure 5.5: SDS-PAGE analysis of MBP-TraH affinity purification. Lanes 2-7: elution samples containing the protein of interest MBP-TraH at approximately 90 kDa.

Previous studies have failed to show a clear interaction between GST-TraF and MBP-TraH using various affinity pull-down methods as well as through size exclusion chromatography. It was hypothesized that the N-terminal tags interfere with the binding of these two proteins, where the C-terminal region of TraH interacts with the N-terminal region of TraF. In the current study, following on-column cleavage of the GST tag from TraF and incubation with amylose beads attached to MBP-TraH, no interaction was observed. Optimization of various conditions for the pull-down assay will need to be

undertaken including varying the pH, salt concentrations, and perhaps the addition of micelles or detergents to aid in resembling membrane conditions. Following elucidation of favourable interaction conditions, the goal is to localize the area of protein-protein interaction using TRESI-HDX-MS.

5.5. Discussion

The GST-TraF and MBP-TraH protein constructs were successfully expressed and purified using affinity chromatography. The main objective of the research was to gain structural insights on these two proteins that are predicted to interact with one another. Unfortunately, crystallization trials of GST-TraF have not yet yielded diffraction quality crystals and the structure of the protein could not be solved at this time. Dynamic analysis on the construct did reveal some broad insights on the structure of the protein, with the C-terminal domain predicted to be more structured compared to the N-terminal domain. Although a yeast-two hybrid assay has demonstrated the interaction between TraH and TraF,[28] conditions must be optimized in order to reproduce this interaction *in vitro* and localize the area of interaction.

Chapter 6: Conclusion and Future Work

6.1. Conclusion

The current study was successful in using TRESI-HDX-MS for the structural analysis of key proteins from two biological systems. Atomic structures for the ΔK122 dimer and TraF have not been able to be solved through X-ray crystallography, and thus were studied by mass spectrometry in order to gain insight on the dynamics of the protein in solution and gather some conclusions on areas of protein-protein interaction, as well as overall structure.

The main research objective of this Master's thesis was to obtain insights on the oligomerization mechanism behind the ΔK122 pilin, and to identify sites of protein-protein interaction in the earliest intermediate leading up to fibril formation. Localizing the areas that are essential in stabilizing the dimeric state allows for use of the pilin in the field of nanobiotechnology and vaccine development. For example, engineered protein nanotubes can be used as biological nanowires as well as for drug delivery systems, biosensors, and probes. In addition, the proposed interaction thought to occur between TraF and TraH was explored, and in order to study the structural characteristics of TraF, both TRESI-HDX-MS and crystallization trials were attempted. Insights on the N- and C-terminal domains of TraF confirm a more structured C-terminal domain which likely houses a thioredoxin-like fold. In addition, several promising conditions for the optimal crystallization of GST-TraF were identified.

6.2. Future Work

To complement the data obtained from HDX-MS, structure determination from a crystallized protein is ideal. Future studies should be centered on the crystallization of a truncated and soluble TraF construct preferably containing the less dynamic and more structured C-terminus. In order to gain a better understanding of TraH, it is clear that optimization of its storage buffer needs to occur in order to prevent it from degradation as well as aggregation following purification. Due to its predicted membrane associated domain consisting of hydrophobic α-helices and coiled-coil domains, precipitation occurs upon removal of the soluble MBP tag. In order to stabilize the protein, the use of nanodiscs may have to be explored, allowing for a more native environmental condition and friendlier mass spectrometry conditions compared to detergents and liposomes. Due to the extensive cleavage of the MBP tag, it is not optimal to carry out crystallization studies while there is a large amount of the fusion present. In order to carry out crystallographic screenings on the TraH protein, it might be necessary to replace the MBP tag with a His-SUMO tag. Additional work needs to be undertaken in order to identify the optimal conditions in which a robust interaction occurs between TraF and TraH perhaps using isothermal calorimetric experiments, and only when these conditions are found can the analytical methods discussed within this thesis (TRESI-HDX-MS and X-ray crystallography) be applied in order to obtain localized structural insights on the protein complex.

Chapter 7: References

1. Morens, D. M.; Folkers, G. K.; Fauci, A. S., The challenge of emerging and re-emerging infectious diseases. *Nature* **2004,** *430* (6996), 242-9.
2. Frost, L. S.; Leplae, R.; Summers, A. O.; Toussaint, A., Mobile genetic elements: the agents of open source evolution. *Nature Reviews. Microbiology* **2005,** *3* (9), 722-32.
3. Furuya, E. Y.; Lowy, F. D., Antimicrobial-resistant bacteria in the community setting. *Nature Reviews. Microbiology* **2006,** *4* (1), 36-45.
4. Chen, I.; Christie, P. J.; Dubnau, D., The ins and outs of DNA transfer in bacteria. *Science* **2005,** *310* (5753), 1456-60.
5. Hazes, B.; Frost, L., Towards a systems biology approach to study type II/IV secretion systems. *Biochimica et Biophysica Acta* **2008,** *1778* (9), 1839-50.
6. Christie, P. J., Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. *Molecular Microbiology* **2001,** *40* (2), 294-305.
7. Christie, P. J.; Atmakuri, K.; Krishnamoorthy, V.; Jakubowski, S.; Cascales, E., Biogenesis, architecture, and function of bacterial type IV secretion systems. *Annual Review of Microbiology* **2005,** *59*, 451-85.
8. Lawley, T. D.; Klimke, W. A.; Gubbins, M. J.; Frost, L. S., F factor conjugation is a true type IV secretion system. *FEMS Microbiology Letters* **2003,** *224* (1), 1-15.
9. Cascales, E.; Christie, P. J., The versatile bacterial type IV secretion systems. *Nature Reviews. Microbiology* **2003,** *1* (2), 137-49.
10. Backert, S.; Meyer, T. F., Type IV secretion systems and their effectors in bacterial pathogenesis. *Current Opinion in Microbiology* **2006,** *9* (2), 207-17.
11. Christie, P. J.; Whitaker, N.; Gonzalez-Rivera, C., Mechanism and structure of the bacterial type IV secretion systems. *Biochimica et Biophysica Acta* **2014,** *1843* (8), 1578-91.
12. Fronzes, R.; Christie, P. J.; Waksman, G., The structural biology of type IV secretion systems. *Nature Reviews. Microbiology* **2009,** *7* (10), 703-14.
13. Fronzes, R.; Schafer, E.; Wang, L.; Saibil, H. R.; Orlova, E. V.; Waksman, G., Structure of a type IV secretion system core complex. *Science* **2009,** *323* (5911), 266-8.
14. Lederberg, J.; Tatum, E. L., Gene recombination in *Escherichia coli*. *Nature* **1946,** *158* (4016), 558.
15. Yagi, T.; Kurokawa, H.; Senda, K.; Ichiyama, S.; Ito, H.; Ohsuka, S.; Shibayama, K.; Shimokata, K.; Kato, N.; Ohta, M.; Arakawa, Y., Nosocomial spread of cephem-resistant *Escherichia coli* strains carrying multiple Toho-1-like beta-lactamase genes. *Antimicrobial Agents and Chemotherapy* **1997,** *41* (12), 2606-11.
16. Juhas, M.; Crook, D. W.; Hood, D. W., Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. *Cellular Microbiology* **2008,** *10* (12), 2377-86.
17. Thomas, C. M.; Nielsen, K. M., Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews. Microbiology* **2005,** *3* (9), 711-21.

18. Anthony, K. G.; Sherburne, C.; Sherburne, R.; Frost, L. S., The role of the pilus in recipient cell recognition during bacterial conjugation mediated by F-like plasmids. *Molecular Microbiology* **1994,** *13* (6), 939-53.
19. Clarke, M.; Maddera, L.; Harris, R. L.; Silverman, P. M., F-pili dynamics by live-cell imaging. *PNAS* **2008,** *105* (46), 17978-81.
20. Achtman, M., Mating aggregates in *Escherichia coli* conjugation. *Journal of Bacteriology* **1975,** *123* (2), 505-15.
21. Durrenberger, M. B.; Villiger, W.; Bachi, T., Conjugational junctions: morphology of specific contacts in conjugating *Escherichia coli* bacteria. *Journal of Structural Biology* **1991,** *107* (2), 146-56.
22. Lederberg, J., Conjugal pairing in *Escherichia coli*. *Journal of Bacteriology* **1956,** *71* (4), 497-8.
23. Firth, N.; Ippen-Ihler, K.; Skurray, R. A., Structure and function of the F factor and mechanism of conjugation. 2nd ed.; ASM Press: Washington, DC, 1996; pp 2377-2401.
24. Frost, L. S.; Ippen-Ihler, K.; Skurray, R. A., Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiological Reviews* **1994,** *58* (2), 162-210.
25. Wang, Y. A.; Yu, X.; Silverman, P. M.; Harris, R. L.; Egelman, E. H., The structure of F-pili. *Journal of Molecular Biology* **2009,** *385* (1), 22-9.
26. Arutyunov, D.; Arenson, B.; Manchak, J.; Frost, L. S., F plasmid TraF and TraH are components of an outer membrane complex involved in conjugation. *Journal of Bacteriology* **2010,** *192* (6), 1730-1734.
27. Ippen-Ihler, K. A.; Minkley, E. G., Jr., The conjugation system of F, the fertility factor of *Escherichia coli*. *Annual Review of Genetics* **1986,** *20*, 593-624.
28. Harris, R. L.; Silverman, P. M., Tra proteins characteristic of F-like type IV secretion systems constitute an interaction group by yeast two-hybrid analysis. *Journal of Bacteriology* **2004,** *186* (16), 5480-5.
29. Audette, G. F.; Manchak, J.; Beatty, P.; Klimke, W. A.; Frost, L. S., Entry exclusion in F-like plasmids requires intact TraG in the donor that recognizes its cognate TraS in the recipient. *Microbiology* **2007,** *153* (Pt 2), 442-51.
30. Elton, T. C.; Holland, S. J.; Frost, L. S.; Hazes, B., F-like type IV secretion systems encode proteins with thioredoxin folds that are putative DsbC homologues. *Journal of Bacteriology* **2005,** *187* (24), 8267-77.
31. Audette, G. F.; Holland, S. J.; Elton, T. C.; Manchak, J.; Hayakawa, K.; Frost, L. S.; Hazes, B., Crystallization and preliminary diffraction studies of TraF, a component of the *Escherichia coli* type IV secretory system. *Acta Crystallographica. Section D, Biological Crystallography* **2004,** *60* (Pt 11), 2025-7.
32. Raina, S.; Missiakas, D., Making and breaking disulfide bonds. *Annual Review of Microbiology* **1997,** *51*, 179-202.
33. Kelley, L. A.; Sternberg, M. J., Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols* **2009,** *4* (3), 363-71.
34. Schrodinger, LLC, The PyMOL Molecular Graphics System, Version 1.3r1. 2010.
35. Hemmis, C. W.; Berkmen, M.; Eser, M.; Schildbach, J. F., TrbB from conjugative plasmid F is a structurally distinct disulfide isomerase that requires DsbD for redox state maintenance. *Journal of Bacteriology* **2011,** *193* (18), 4588-97.

36. d'Enfert, C.; Ryter, A.; Pugsley, A. P., Cloning and expression in *Escherichia coli* of the Klebsiella pneumoniae genes for production, surface localization and secretion of the lipoprotein pullulanase. *The EMBO Journal* **1987,** *6* (11), 3531-8.
37. Nivaskumar, M.; Francetic, O., Type II secretion system: a magic beanstalk or a protein escalator. *Biochimica et Biophysica Acta* **2014,** *1843* (8), 1568-77.
38. Abendroth, J.; Murphy, P.; Sandkvist, M.; Bagdasarian, M.; Hol, W. G., The X-ray structure of the type II secretion system complex formed by the N-terminal domain of EpsE and the cytoplasmic domain of EpsL of *Vibrio cholerae*. *Journal of Molecular Biology* **2005,** *348* (4), 845-55.
39. Pruitt, B. A., Jr.; McManus, A. T.; Kim, S. H.; Goodwin, C. W., Burn wound infections: current status. *World Journal of Surgery* **1998,** *22* (2), 135-45.
40. D'Avignon, L. C.; Hogan, B. K.; Murray, C. K.; Loo, F. L.; Hospenthal, D. R.; Cancio, L. C.; Kim, S. H.; Renz, E. M.; Barillo, D.; Holcomb, J. B.; Wade, C. E.; Wolf, S. E., Contribution of bacterial and viral infections to attributable mortality in patients with severe burns: An autopsy series. *Burns* **2010,** *36* (6), 773-779.
41. Bang, R. L.; Sharma, P. N.; Sanyal, S. C.; Al Najjadah, I., Septicaemia after burn injury: a comparative study. *Burns* **2002,** *28* (8), 746-51.
42. Fishman, J. A.; Rubin, R. H., Infection in organ-transplant recipients. *The New England Journal of Medicine* **1998,** *338* (24), 1741-51.
43. Dummer, J. S.; Montero, C. G.; Griffith, B. P.; Hardesty, R. L.; Paradis, I. L.; Ho, M., Infections in heart-lung transplant recipients. *Transplantation* **1986,** *41* (6), 725-9.
44. Pier, G. B., Pulmonary disease associated with *Pseudomonas aeruginosa* in cystic fibrosis: current status of the host-bacterium interaction. *The Journal of Infectious Diseases* **1985,** *151* (4), 575-80.
45. Marchetti, F.; Giglio, L.; Candusso, M.; Faraguna, D.; Assael, B., Early antibiotic treatment of *pseudomonas aeruginosa* colonisation in cystic fibrosis: a critical review of the literature. *European Journal of Clinical Pharmacology* **2004,** *60* (2), 67-74.
46. Maschmeyer, G.; Braveny, I., Review of the incidence and prognosis of *Pseudomonas aeruginosa* infections in cancer patients in the 1990s. *European Journal of Clinical Microbiology & Infectious Diseases: Official Publication of the European Society of Clinical Microbiology* **2000,** *19* (12), 915-25.
47. Manfredi, R.; Nanetti, A.; Ferri, M.; Chiodo, F., *Pseudomonas* spp. complications in patients with HIV disease: an eight-year clinical and microbiological survey. *European Journal of Epidemiology* **2000,** *16* (2), 111-8.
48. Costerton, J. W.; Stewart, P. S.; Greenberg, E. P., Bacterial biofilms: a common cause of persistent infections. *Science* **1999,** *284* (5418), 1318-22.
49. Sheth, H. B.; Lee, K. K.; Wong, W. Y.; Srivastava, G.; Hindsgaul, O.; Hodges, R. S.; Paranchych, W.; Irvin, R. T., The pili of *Pseudomonas aeruginosa* strains PAK and PAO bind specifically to the carbohydrate sequence βGalNAc(1–4)βGal found in glycosphingolipids asialo-GM1 and asialo-GM2. *Molecular Microbiology* **1994,** *11* (4), 715-723.
50. Lee, K. K.; Sheth, H. B.; Wong, W. Y.; Sherburne, R.; Paranchych, W.; Hodges, R. S.; Lingwood, C. A.; Krivan, H.; Irvin, R. T., The binding of *Pseudomonas aeruginosa* pili to glycosphingolipids is a tip-associated event involving the C-

terminal region of the structural pilin subunit. *Molecular Microbiology* **1994,** *11* (4), 705-713.
51. Burrows, L. L., Weapons of mass retraction. *Molecular Microbiology* **2005,** *57* (4), 878-888.
52. Burrows, L. L., *Pseudomonas aeruginosa* twitching motility: type IV pili in action. *Annual Review of Microbiology* **2012,** *66* (1), 493-520.
53. Craig, L.; Pique, M. E.; Tainer, J. A., Type IV pilus structure and bacterial pathogenicity. *Nature Reviews. Microbiology* **2004,** *2* (5), 363-378.
54. Hancock, R. E. W.; Speert, D. P., Antibiotic resistance in *Pseudomonas aeruginosa*: mechanisms and impact on treatment. *Drug Resistance Updates* **2000,** *3* (4), 247-255.
55. Mattick, J. S., Type IV pili and twitching motility. *Annual Review of Microbiology* **2002,** *56*, 289-314.
56. Jin, F.; Conrad, J. C.; Gibiansky, M. L.; Wong, G. C. L., Bacteria use type-IV pili to slingshot on surfaces. *PNAS* **2011,** *108* (31), 12617-12622.
57. Conrad, Jacinta C.; Gibiansky, Maxsim L.; Jin, F.; Gordon, Vernita D.; Motto, Dominick A.; Mathewson, Margie A.; Stopka, Wiktor G.; Zelasko, Daria C.; Shrout, Joshua D.; Wong, Gerard C. L., Flagella and pili-mediated near-surface single-cell motility mechanisms in *P. aeruginosa*. *Biophysical Journal* **2011,** *100* (7), 1608-1616.
58. Skerker, J. M.; Berg, H. C., Direct observation of extension and retraction of type IV pili. *PNAS* **2001,** *98* (12), 6901-6904.
59. O'Toole, G.; Kaplan, H. B.; Kolter, R., Biofilm formation as microbial development. *Annual Review of Microbiology* **2000,** *54*, 49-79.
60. Barken, K. B.; Pamp, S. J.; Yang, L.; Gjermansen, M.; Bertrand, J. J.; Klausen, M.; Givskov, M.; Whitchurch, C. B.; Engel, J. N.; Tolker-Nielsen, T., Roles of type IV pili, flagellum-mediated motility and extracellular DNA in the formation of mature multicellular structures in *Pseudomonas aeruginosa* biofilms. *Environmental Microbiology* **2008,** *10* (9), 2331-2343.
61. Allesen-Holm, M.; Barken, K. B.; Yang, L.; Klausen, M.; Webb, J. S.; Kjelleberg, S.; Molin, S.; Givskov, M.; Tolker-Nielsen, T., A characterization of DNA release in *Pseudomonas aeruginosa* cultures and biofilms. *Molecular Microbiology* **2006,** *59* (4), 1114-1128.
62. Mikkelsen, H.; Sivaneson, M.; Filloux, A., Key two-component regulatory systems that control biofilm formation in *Pseudomonas aeruginosa*. *Environmental Microbiology* **2011,** *13* (7), 1666-1681.
63. Abraham, S. N.; Jonsson, A. B.; Normark, S., Fimbriae-mediated host-pathogen cross-talk. *Current Opinion in Microbiology* **1998,** *1* (1), 75-81.
64. Dubnau, D., DNA uptake in bacteria. *Annual Review of Microbiology* **1999,** *53*, 217-44.
65. van Schaik, E. J.; Giltner, C. L.; Audette, G. F.; Keizer, D. W.; Bautista, D. L.; Slupsky, C. M.; Sykes, B. D.; Irvin, R. T., DNA binding: a novel function of *Pseudomonas aeruginosa* type IV pili. *Journal of Bacteriology* **2005,** *187* (4), 1455-1464.

66. Bieber, D.; Ramer, S. W.; Wu, C. Y.; Murray, W. J.; Tobe, T.; Fernandez, R.; Schoolnik, G. K., Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic *Escherichia coli*. *Science* **1998,** *280* (5372), 2114-8.
67. Herrington, D. A.; Hall, R. H.; Losonsky, G.; Mekalanos, J. J.; Taylor, R. K.; Levine, M. M., Toxin, toxin-coregulated pili, and the toxR regulon are essential for Vibrio cholerae pathogenesis in humans. *The Journal of Experimental Medicine* **1988,** *168* (4), 1487-92.
68. Audette, G. F.; Hazes, B., Development of protein nanotubes from a multi-purpose biological structure. *Journal of Nanoscience and Nanotechnology* **2007,** *7* (7), 2222-9.
69. Craig, L.; Li, J., Type IV pili: paradoxes in form and function. *Current Opinion in Structural Biology* **2008,** *18* (2), 267-277.
70. Chen, I.; Dubnau, D., DNA uptake during bacterial transformation. *Nature Reviews. Microbiology* **2004,** *2* (3), 241-9.
71. Craig, L.; Volkmann, N.; Arvai, A. S.; Pique, M. E.; Yeager, M.; Egelman, Edward H.; Tainer, J. A., Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. *Molecular Cell* **2006,** *23* (5), 651-662.
72. Parge, H. E.; Forest, K. T.; Hickey, M. J.; Christensen, D. A.; Getzoff, E. D.; Tainer, J. A., Structure of the fibre-forming protein pilin at 2.6 A resolution. *Nature* **1995,** *378* (6552), 32-38.
73. Hazes, B.; Sastry, P. A.; Hayakawa, K.; Read, R. J.; Irvin, R. T., Crystal structure of *Pseudomonas aeruginosa* PAK pilin suggests a main-chain-dominated mode of receptor binding. *Journal of Molecular Biology* **2000,** *299* (4), 1005-1017.
74. Keizer, D. W.; Slupsky, C. M.; Kalisiak, M.; Campbell, A. P.; Crump, M. P.; Sastry, P. A.; Hazes, B.; Irvin, R. T.; Sykes, B. D., Structure of a pilin monomer from *Pseudomonas aeruginosa*: implications for the assembly of pili. *Journal of Biological Chemistry* **2001,** *276* (26), 24186-24193.
75. Audette, G. F.; Irvin, R. T.; Hazes, B., Crystallographic analysis of the *Pseudomonas aeruginosa* strain K122-4 monomeric pilin reveals a conserved receptor-binding architecture. *Biochemistry* **2004,** *43* (36), 11427-11435.
76. Kao, D. J.; Churchill, M. E. A.; Irvin, R. T.; Hodges, R. S., Animal protection and structural studies of a consensus sequence vaccine targeting the receptor binding domain of the type IV pilus of *Pseudomonas aeruginosa*. *Journal of Molecular Biology* **2007,** *374* (2), 426-442.
77. Nguyen, Y.; Jackson, S. G.; Aidoo, F.; Junop, M.; Burrows, L. L., Structural characterization of novel *Pseudomonas aeruginosa* type IV pilins. *Journal of Molecular Biology* **2010,** *395* (3), 491-503.
78. Hartung, S.; Arvai, A. S.; Wood, T.; Kolappan, S.; Shin, D. S.; Craig, L.; Tainer, J. A., Ultrahigh resolution and full-length pilin structures with insights for filament assembly, pathogenic functions, and vaccine potential. *Journal of Biological Chemistry* **2011,** *286* (51), 44254-44265.
79. Li, J.; Egelman, E. H.; Craig, L., Structure of the *Vibrio cholerae* type IVb pilus and stability comparison with the *Neisseria gonorrhoeae* type IVa pilus. *Journal of Molecular Biology* **2012,** *418* (1–2), 47-64.

80. Craig, L.; Taylor, R. K.; Pique, M. E.; Adair, B. D.; Arvai, A. S.; Singh, M.; Lloyd, S. J.; Shin, D. S.; Getzoff, E. D.; Yeager, M.; Forest, K. T.; Tainer, J. A., Type IV pilin structure and assembly: X-ray and EM analyses of *Vibrio cholerae* toxin-coregulated pilus and *Pseudomonas aeruginosa* PAK pilin. *Molecular Cell* **2003,** *11* (5), 1139-1150.
81. Hahn, H. P., The type-4 pilus is the major virulence-associated adhesin of *Pseudomonas aeruginosa*--a review. *Gene* **1997,** *192* (1), 99-108.
82. Peek, J. A.; Taylor, R. K., Characterization of a periplasmic thiol:disulfide interchange protein required for the functional maturation of secreted virulence factors of *Vibrio cholerae*. *PNAS* **1992,** *89* (13), 6210-4.
83. Zhang, H. Z.; Donnenberg, M. S., DsbA is required for stability of the type IV pilin of enteropathogenic *escherichia coli*. *Molecular Microbiology* **1996,** *21* (4), 787-97.
84. Forest, K. T.; Tainer, J. A., Type-4 pilus-structure: outside to inside and top to bottom – minireview. *Gene* **1997,** *192* (1), 165-169.
85. Hazes, B.; Frost, L., Towards a systems biology approach to study type II/IV secretion systems. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2008,** *1778* (9), 1839-1850.
86. Ayers, M.; Howell, P. L.; Burrows, L. L., Architecture of the type II secretion and type IV pilus machineries. *Future Microbiology* **2010,** *5* (8), 1203-1218.
87. Folkhard, W.; Marvin, D. A.; Watts, T. H.; Paranchych, W., Structure of polar pili from *Pseudomonas aeruginosa* strains K and O. *Journal of Molecular Biology* **1981,** *149* (1), 79-93.
88. Maier, B.; Potter, L.; So, M.; Long, C. D.; Seifert, H. S.; Sheetz, M. P., Single pilus motor forces exceed 100 pN. *PNAS* **2002,** *99* (25), 16012-7.
89. Giltner, C. L.; van Schaik, E. J.; Audette, G. F.; Kao, D.; Hodges, R. S.; Hassett, D. J.; Irvin, R. T., The *Pseudomonas aeruginosa* type IV pilin receptor binding domain functions as an adhesin for both biotic and abiotic surfaces. *Molecular Microbiology* **2006,** *59* (4), 1083-96.
90. Yu, B.; Giltner, C. L.; Van Schaik, E. J.; Bautista, D. L.; Hodges, R. S.; Audette, G. F.; Li, D. Y.; Irvin, R. T., A novel biometallic interface: high affinity tip-associated binding by pilin-derived protein nanotubes. *Journal of Bionanoscience* **2007,** *1* (2), 73-83.
91. Reguera, G.; McCarthy, K. D.; Mehta, T.; Nicoll, J. S.; Tuominen, M. T.; Lovley, D. R., Extracellular electron transfer via microbial nanowires. *Nature* **2005,** *435* (7045), 1098-101.
92. Yi, H.; Nevin, K. P.; Kim, B. C.; Franks, A. E.; Klimes, A.; Tender, L. M.; Lovley, D. R., Selection of a variant of *Geobacter sulfurreducens* with enhanced capacity for current production in microbial fuel cells. *Biosensors & Bioelectronics* **2009,** *24* (12), 3498-503.
93. Malvankar, N. S.; Vargas, M.; Nevin, K. P.; Franks, A. E.; Leang, C.; Kim, B. C.; Inoue, K.; Mester, T.; Covalla, S. F.; Johnson, J. P.; Rotello, V. M.; Tuominen, M. T.; Lovley, D. R., Tunable metallic-like conductivity in microbial nanowire networks. *Nature Nanotechnology* **2011,** *6* (9), 573-9.
94. Malvankar, N. S.; Mester, T.; Tuominen, M. T.; Lovley, D. R., Supercapacitors based on c-type cytochromes using conductive nanostructured networks of living

bacteria. *Chemphyschem: A European Journal of Chemical Physics and Physical Chemistry* **2012,** *13* (2), 463-8.
95. Audette, G. F.; vanSchaik, E. J.; Hazes, B.; Irvin, R. T., DNA-binding protein nanotubes: learning from nature's nanotech examples. *Nano Letters* **2004,** *4* (10), 1897-1902.
96. Lombardo, S.; Jasbi, S. Z.; Jeung, S.-K.; Morin, S.; Audette, G. F., Initial studies of protein nanotube oligomerization from a modified gold surface. *Journal of Bionanoscience* **2009,** *3* (1), 61-65.
97. Petrov, A.; Lombardo, S.; Audette, G., Fibril-mediated oligomerization of pilin-derived protein nanotubes. *Journal of Nanobiotechnology* **2013,** *11* (1), 24.
98. Griffiths, I. W., J. J. Thomson — the centenary of his discovery of the electron and of his invention of mass spectrometry. *Rapid Communications in Mass Spectrometry : RCM* **1997,** *11* (1), 2-16.
99. Budzikiewicz, H.; Grigsby, R. D., Mass spectrometry and isotopes: a century of research and discussion. *Mass Spectrometry Reviews* **2006,** *25* (1), 146-57.
100. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989,** *246* (4926), 64-71.
101. Yamashita, M.; Fenn, J. B., Electrospray ion source. Another variation on the free-jet theme. The *Journal of Physical Chemistry* **1984,** *88* (20), 4451-4459.
102. Whitehouse, C. M.; Dreyer, R. N.; Yamashita, M.; Fenn, J. B., Electrospray interface for liquid chromatographs and mass spectrometers. *Analytical Chemistry* **1985,** *57* (3), 675-9.
103. Banerjee, S.; Mazumdar, S., Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *International Journal of Analytical Chemistry* **2012,** *2012*, 282574.
104. Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S., Unraveling the mechanism of electrospray ionization. *Analytical Chemistry* **2013,** *85* (1), 2-9.
105. Kebarle, P.; Verkerk, U. H., Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrometry Reviews* **2009,** *28* (6), 898-917.
106. Sinz, A., Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrometry Reviews* **2006,** *25* (4), 663-82.
107. Jin Lee, Y., Mass spectrometric analysis of cross-linking sites for the structure of proteins and protein complexes. *Molecular BioSystems* **2008,** *4* (8), 816-23.
108. Larsen, M. R.; Trelle, M. B.; Thingholm, T. E.; Jensen, O. N., Analysis of posttranslational modifications of proteins by tandem mass spectrometry. *BioTechniques* **2006,** *40* (6), 790-8.
109. Ziady, A. G.; Kinter, M., Protein sequencing with tandem mass spectrometry. *Methods in Molecular Biology* **2009,** *544*, 325-41.
110. Przybylski, M.; Glocker, M. O., Electrospray mass spectrometry of biomacromolecular complexes with noncovalent interactions - New analytical perspectives for supramolecular chemistry and molecular recognition processes. *Angewandte Chemie International Edition* **1996,** *35* (8), 807-826.

111. Kaltashov, I. A.; Eyles, S. J., Studies of biomolecular conformations and conformational dynamics by mass spectrometry. *Mass Spectrometry Reviews* **2002,** *21* (1), 37-71.
112. Lamond, A. I. Electrospray Ionisation: How to Ionise Peptides in LC-MS. http://www.lamondlab.com/MSResource/LCMS/MassSpectrometry/electrosprayIonisation.php (accessed July 15, 2015).
113. Gomez-Hens, A.; Perez-Bendito, D., The stopped-flow technique in analytical chemistry. *Analytica Chimica Acta* **1991,** *242* (1), 147-177.
114. Northrop, D. B.; Simpson, F. B., Beyond enzyme kinetics: direct determination of mechanisms by stopped-flow mass spectrometry. *Bioorganic & Medicinal Chemistry* **1997,** *5* (4), 641-4.
115. Kelly, S. M.; Price, N. C., The use of circular dichroism in the investigation of protein structure and function. *Current Protein & Peptide Science* **2000,** *1* (4), 349-84.
116. Konermann, L.; Pan, J.; Wilson, D. J., Protein folding mechanisms studied by time-resolved electrospray mass spectrometry. *BioTechniques* **2006,** *40* (2), 135, 137, 139 passim.
117. Rob, T.; Wilson, D. J., Time-resolved mass spectrometry for monitoring millisecond time-scale solution-phase processes. *European Journal of Mass Spectrometry* **2012,** *18* (2), 205-14.
118. Konermann, L.; Collings, B. A.; Douglas, D. J., Cytochrome c folding kinetics studied by time-resolved electrospray ionization mass spectrometry. *Biochemistry* **1997,** *36* (18), 5554-9.
119. Kolakowski, B. M.; Simmons, D. A.; Konermann, L., Stopped-flow electrospray ionization mass spectrometry: a new method for studying chemical reaction kinetics in solution. *Rapid Communications in Mass Spectrometry : RCM* **2000,** *14* (9), 772-6.
120. Wilson, D. J.; Konermann, L., A capillary mixer with adjustable reaction chamber volume for millisecond time-resolved studies by electrospray mass spectrometry. *Analytical Chemistry* **2003,** *75* (23), 6408-14.
121. Lion, N.; Rohner, T. C.; Dayon, L.; Arnaud, I. L.; Damoc, E.; Youhnovski, N.; Wu, Z. Y.; Roussel, C.; Josserand, J.; Jensen, H.; Rossier, J. S.; Przybylski, M.; Girault, H. H., Microfluidic systems in proteomics. *Electrophoresis* **2003,** *24* (21), 3533-62.
122. Rob, T.; Wilson, D. J., A versatile microfluidic chip for millisecond time-scale kinetic studies by electrospray mass spectrometry. *Journal of the American Society for Mass Spectrometry* **2009,** *20* (1), 124-30.
123. Liuni, P.; Rob, T.; Wilson, D. J., A microfluidic reactor for rapid, low-pressure proteolysis with on-chip electrospray ionization. *Rapid Communications in Mass Spectrometry : RCM* **2010,** *24* (3), 315-20.
124. Rob, T.; Liuni, P.; Gill, P. K.; Zhu, S.; Balachandran, N.; Berti, P. J.; Wilson, D. J., Measuring dynamics in weakly structured regions of proteins using microfluidics-enabled subsecond H/D exchange mass spectrometry. *Analytical Chemistry* **2012,** *84* (8), 3771-9.
125. Wales, T. E.; Engen, J. R., Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrometry Reviews* **2006,** *25* (1), 158-70.

126. Konermann, L.; Pan, J.; Liu, Y. H., Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chemical Society Reviews* **2011,** *40* (3), 1224-34.
127. Morgan, C. R.; Engen, J. R., Investigating solution-phase protein structure and dynamics by hydrogen exchange mass spectrometry. *Current Protocols in Protein Science* **2009,** *Chapter 17*, Unit 17.6 1-17.
128. Katta, V.; Chait, B. T., Conformational changes in proteins probed by hygrogen-exchange electrospray ionization mass spectrometry *Rapid Communications in Mass Spectrometry : RCM* **1991,** *5* (4), 214-217.
129. Weis, D. D. Hydrogen-Deuterium Exchange Mass Spectrometry. https://mvsc.ku.edu/content/hydrogen-deuterium-exchange-mass-spectrometry (accessed July 15, 2015).
130. Resetca, D.; Wilson, D. J., Characterizing rapid, activity-linked conformational transitions in proteins via sub-second hydrogen deuterium exchange mass spectrometry. *The FEBS Journal* **2013,** *280* (22), 5616-25.
131. Rob, T.; Gill, P. K.; Golemi-Kotra, D.; Wilson, D. J., An electrospray ms-coupled microfluidic device for sub-second hydrogen/deuterium exchange pulse-labelling reveals allosteric effects in enzyme inhibition. *Lab on a Chip* **2013,** *13* (13), 2528-32.
132. Zhang, Y. Z. SPHERE - A Server Program for Hydrogen Exchange Rate Estimation. http://www.fccc.edu/research/labs/roder/sphere/ (accessed July 15, 2015)
133. Serrano-Pozo, A.; Frosch, M. P.; Masliah, E.; Hyman, B. T., Neuropathological alterations in Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine* **2011,** *1* (1), a006189.
134. Zhu, S.; Wilson, D. J., Characterization of residual structure in the instrinsically disordered protein tau. In *61st Annual Conference of the American Society for Mass Spectrometry*, Minneapolis, MN, USA, 2013.
135. Marcsisin, S. R.; Engen, J. R., Hydrogen exchange mass spectrometry: what is it and what can it tell us? *Analytical and Bioanalytical Chemistry* **2010,** *397* (3), 967-72.
136. Percy, A. J.; Rey, M.; Burns, K. M.; Schriemer, D. C., Probing protein interactions with hydrogen/deuterium exchange and mass spectrometry-a review. *Analytica Chimica Acta* **2012,** *721*, 7-21.
137. Neu, H. C.; Heppel, L. A., The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplasts. *The Journal of Biological Chemistry* **1965,** *240* (9), 3685-92.
138. Ho, C. S.; Lam, C. W.; Chan, M. H.; Cheung, R. C.; Law, L. K.; Lit, L. C.; Ng, K. F.; Suen, M. W.; Tai, H. L., Electrospray ionisation mass spectrometry: principles and clinical applications. *The Clinical Biochemist. Reviews / Australian Association of Clinical Biochemists* **2003,** *24* (1), 3-12.
139. Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A., An introduction to quadrupole-time-of-flight mass spectrometry. *Journal of Mass Spectrometry : JMS* **2001,** *36* (8), 849-65.
140. Audette, G. F.; Irvin, R. T.; Hazes, B., Purification, crystallization and preliminary diffraction studies of the *Pseudomonas aeruginosa* strain K122-4 monomeric pilin.

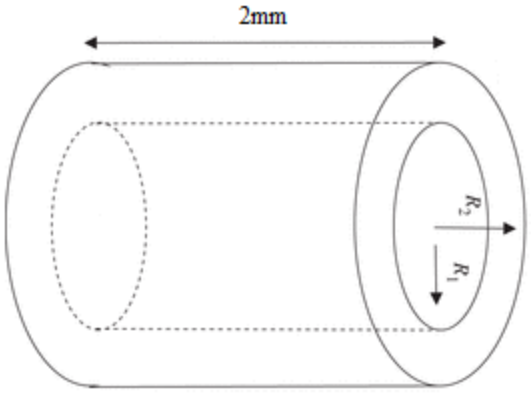
Acta Crystallographica. Section D, Biological Crystallography **2003,** *59* (Pt 9), 1665-7.
141. Walz, D. A.; Hewett-Emmett, D.; Seegers, W. H., Amino acid sequence of human prothrombin fragments 1 and 2. *PNAS* **1977,** *74* (5), 1969-72.
142. Suh, J. Y.; Spyracopoulos, L.; Keizer, D. W.; Irvin, R. T.; Sykes, B. D., Backbone dynamics of receptor binding and antigenic regions of a *Pseudomonas aeruginosa* pilin monomer. *Biochemistry* **2001,** *40* (13), 3985-95.
143. Charm, S. E.; Wong, B. L., Enzyme inactivation with shearing. *Biotechnology and Bioengineering* **1970,** *12* (6), 1103-9.
144. Kim, M. Y.; Maier, C. S.; Reed, D. J.; Deinzer, M. L., Site-specific amide hydrogen/deuterium exchange in *E. coli* thioredoxins measured by electrospray ionization mass spectrometry. *Journal of the American Chemical Society* **2001,** *123* (40), 9860-6.

Appendix

Sample Calculations:

1) Mixing Time

Inner Glass Capillary (o.d. values for radius):
h = 2 mm notch
R_1 = .153/2 = .0765 mm
Area = $\pi r^2 h$
= 0.0368 mm^3



Outer Metal Capillary (i.d. values for radius):
h = 2 mm notch
R_2 = .178/2 = 0.089 mm
Area = $\pi r^2 h$
= 0.0498 mm^3

Difference in Area = 0.013 mm^3

1 mm^3 = 10^3 nL

Mixing Volume = 13 nL

With a total flow rate (protein + D_2O) of 4 μL/min = 66.67 nL/sec

Mixing Time = 13 nL ÷ 66.67 nL/sec = 0.195 sec ~ 200 ms

2) Reaction Time

For 5 mm pullback,	Added Volume	=	π $(0.089)^2$ x 5 mm
		=	0.124 mm^3 = 124 nL
	Total Time	=	Mixing Time + Pullback
		=	0.195 sec + (124 nL ÷ 66.67 nL/sec)
		=	2.06 sec