Accessing the Oligomeric State of the SecYEG Protein Secretion Channel Complex

by

Zeliang Zheng

Faculty Advisor: Dr. Donald B. Oliver

A Dissertation submitted to the Faculty of Wesleyan University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Molecular Biology and Biochemistry

Middletown, Connecticut

Feb., 2018
Acknowledgements

Right now is 2-11-2018 at 10pm that I write this. The first person that I want to thank is my advisor Donald Oliver. When I first came into the lab, I knew little about science. I want to thank Don for taking me into the lab, and being very patient with me. I remember once he and I talked for 5 hours straight in order to troubleshoot something. You can see how patient he has been. Once he told me that training a PhD student is just like parenting; it really is. That’s why I really want to thank him because I admit that I can’t do better than him, even as a parent in the future. The value of Ph.D., for me, is not really about science any more. It’s about that patient, persistent and hard-working nature of scientists that can be applied to anything in life. Don did a great job in showing me that. Many times he told me, with patient, persistent and hard work, you can achieve anything. It really is true. Probably, I will forget about the Sec system right after I get my degree, but I will never forget the things that I have learned from Don.

I want to thank my parents for being supportive and understanding. Even though at times, we were not on the same page, but one thing we have always agreed upon is that I need to finish the degree. Their success as parents and as immigrants is truly a miracle. Over the years, as I have noticed how hard it is to get just the PhD degree, I started to realize how hard their life has been. They came here as illegal immigrants and worked their way up. What a miracle. I used to think that immigrants like my parents are non-educated and I am better because I have an education. But after
encountering numerous obstacles during my PhD career, I realized that I am really not any better than anyone like my parents, or any other sanitation workers, taxi drivers and anyone who works hard to make a living. So I am very grateful to have wonderful parents

I want to thank my beloved wife and my parents-in-law who have supported my living during my thesis-writing period, when I was so broke and couldn’t afford meals. Without their support, I couldn’t make it this far. I really enjoyed the time when I was just writing in their restaurant in Indiana; whenever I was hungry, I just came out to the kitchen to grab foods. I am so grateful to marry my wife, a woman who is so humorous and caring, someone whom I also love dearly. Without my wife, my life would have been very boring. My wife has a hard-working, determined, persistent and curious nature, things that I took 7 years to learn, and am still learning. With these characters, she has been successful in pretty much anything she does. Being with her helps me reinforce the importance of things that I have learned from Don. So I am very grateful for having her in my life

I also want to thank my committee members, Manju, Ishita, and Rich. They gave me a hard time during my oral qualifier, but of course in a good way. They have inspired me to think deeper about science. It’s amazing how their brain works; they always can challenge me with clever questions that I can never think of. So thanks to these wonderful fellows who have challenged me and guided me through various stages of my Ph.D. career.
I want to thank Tithi Banergee, my beloved lab mate whom I will missed dearly. I am very sure that I will never meet a better lab mate than her. She and I have a very similar personality, that’s why we get along very well. We help each other unconditionally. We gossiped together, laughed together. Of course, we also spent times complaining too, when we are stressed. So am very blessed to be with her. She will always get VIP treatment from me whenever she needs it in the future.
Table of Contents

Chapter I

Introduction

1. Cell Structure-Eukaryote vs Prokaryote 1
2. Gram-positive Gram-negative Bacteria 2
3. Protein Secretion-the challenges 4
4. Protein Secretion Across the Cytoplasmic Membrane 5
   4.1 Sec Secretion System 5
   4.2 Tat Secretion System 7
5. Protein Secretion Beyond the Periplasm and Sec-independent Systems 8
   5.1 Sec-independent
      A. Type I Secretion System (T1SS) 9
      B. Type III Secretion System (T3SS) 10
   5.2 Sec-dependent
      A. Type II Secretion System (T2SS) 10
      B. Type V Secretion System (T5SS) 11
6. Sec-dependent Protein Translocation-general requirements 12
   6.1 Signal Sequence 13
   6.2 Protein Folding 14
   6.3 Energetics of Protein Translocation 15
7. Sec-dependent Protein Translocation-a detail look

7.1 Co-translational Translocation 17

7.2 Integration of Membrane Protein 20

7.3 Post-translational Translocation 22

A. Sorting and Targeting-the action of trigger factor (TF) 23

B. Sorting and Targeting-the action of SecB 25

C. SecA-the ATPase motor protein 27

i. SecA Structure and Function 27

ii. SecA Oligomerization-structural studies 29

iii. SecA Oligomerization During Protein Translocation 31

iv. SecA and Preprotein Binding 34

v. SecA Membrane Insertion Property 35

D. SecYEG Translocon-structure and Function 36

i. The Lateral Gate 38

ii. The Pore-ring and Plug Domain 39

iii. The Cytoplasmic Loops 40

iv. SecE 42

v. SecG 43

E. SecDF-YajC 44

F. YidC 40

8. SecY Oligomer-the unknown 49

9. Models of SeA-driven Translocation 54
Chapter II

Determination of the Oligomeric State of SecYEG Protein Secretion Channel Complex using in vivo Photo- and Disulfide-crosslinking

1. Abstract 70
2. Introduction 71
3. Experimental Procedures 75
4. Results 82
5. Discussion 102

Chapter III

SecYEG monomer function: a mutational study

1. Abstract 108
2. Introduction 110
3. Experimental Procedures 116
Chapter IV
Discussion

References

List of Figures

Chapter I

Figure 1  Schematic representation of a eukaryote and prokaryote  2
Figure 2  Schematic representation of Gram-negative and Gram-positive cell walls  2
Figure 3  Schematic representation of co-translational translocation and post-translational translocation  7
Figure 4  Schematic of prokaryotic Sec-dependent and Sec-independent secretion pathways  8
Figure 5  Crystal structure of Methanococcus jannaschii SRP with or without signal peptide  18
Figure 6  Roles of YidC in membrane protein biogenesis  21
Figure 7  Structure of Trigger Factor from *E.coli*.  24
Figure 8  Crystal structure of SecB from *Haemophilus influenzae* and *E.coli*  26
Figure 9  Structures of SecA ATPase  29
Figure 10  SecA dimer crystal structures from different organisms  31
Figure 11  Crystal structure of SecYEG homologue Sec61αβγ complex from *M.jannaschii*  38
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 12</td>
<td>Crystal structure of SecDF complex</td>
<td>47</td>
</tr>
<tr>
<td>Figure 13</td>
<td>The crystal structure of YidC from <em>E.coli</em>.</td>
<td>49</td>
</tr>
<tr>
<td>Figure 14</td>
<td>SecYEG monomer and dimer structures</td>
<td>51</td>
</tr>
<tr>
<td>Figure 15</td>
<td>Models of SecA-mediated protein translocation</td>
<td>58</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Incorporation of an unnatural amino acid</td>
<td>65</td>
</tr>
<tr>
<td>Figure 17</td>
<td>Site-specific incorporation of pBpA</td>
<td>67</td>
</tr>
</tbody>
</table>

**Chapter II**

| Figure 2.1 | Placement of pBpA at the dimer interface for *in vivo* photo-crosslinking  | 83   |
| Figure 2.2 | Verification of *in vivo* photo-crosslinking                               | 85   |
| Figure 2.3 | Detection of distinct SecYEG dimer states using *in vivo* photo-crosslinking | 87   |
| Figure 2.4 | Accumulation of MBP in the cytoplasm as a result of translocon jamming     | 89   |
| Figure 2.5 | Analysis of *secY* K434 mutant during OmpA-GFP-induced translocon jamming by *in vivo* photo-crosslinking | 90   |
| Figure 2.6 | Analysis of *secY* S111 mutant during OmpA-GFP-induced translocon jamming by *in vivo* photo-crosslinking | 92   |
| Figure 2.7 | Analysis of the *secY* S111 mutant during MalE-LacZ-induced translocon jamming by *in vivo* photo-crosslinking | 94   |
| Figure 2.8 | Analysis of *secY* Q212C mutant during OmpA-GFP-induced translocon jamming by disulfide crosslinking | 95   |
| Figure 2.9 | Analysis of OmpA-GFP-jammed SecY complex by co-immunoprecipitation         | 96   |
| Figure 2.10 | Analysis of *secY* K20 mutant during kasugamycin treatment by *in vivo* photo-crosslinking | 98   |
| Figure 2.11 | SecA foot-printing study of SecY dimer status                              | 101  |
## Chapter III

| Figure 3.1 | Construction of *secY, secE, and secYE* mutants affecting SecYEG Dimerization | 126 |
| Figure 3.2 | Analysis of growth rates of mutant strains | 130 |
| Figure 3.3 | Translocation efficiency of OmpA | 133 |
| Figure 3.4 | Translocation efficiency for MBP and GBP | 135 |
| Figure 3.5 | Membrane insertion efficiency of MalF-PSBT-I | 138 |
| Figure 3.6 | SecA-foot-printing of wild-type and *secYE* mutant | 141 |
| Figure 3.7 | *In vivo* photo-crosslinking of wild-type (WT) and *secYE* mutant (YE) | 144 |
Chapter I

Introduction

1. Cell Structure-Eukaryote vs Prokaryote

All living organisms are made up of a fundamental unit called the cell. Two types of cell architecture that give rise to an ever-ending diversity in all living organisms are the eukaryotic and prokaryotic cell structures. In its simplest form, both can be thought of as cytoplasm enclosed within a membrane. However, eukaryotes, which constitute all multi-cellular organisms such as plants and animals, are bigger and more complex than prokaryotes. They are characterized by the presence of multiple membranous organelles in the cytoplasm, each of which carries out a specific and essential function (Figure 1). Among these organelles, arguably the most important one is the nucleus that houses the genetic material which is packaged in multiple thread-like structures called chromosomes. Prokaryotes on the other hand are much smaller, and can range from 0.5-2.0µm. They lack organelles such as the nucleus; thus their genetic material, which is condensed without forming nucleosomes, is floating in the cytoplasm (Figure 1). Prokaryotes, such as bacteria, are one of world’s oldest forms of life and are found virtually everywhere in the world. For example, bacteria are involved in nitrogen fixation when living in the soil, and they are involved in the human immune response when living in the gut flora. This thesis focuses on bacteria and their protein transport systems.
2. **Gram-positive vs Gram-negative Bacteria**

Bacteria are categorized into Gram-negative and Gram-positive based on their ability to retain the crystal violet-iodine dye in a staining procedure developed by Christian Gram in 1884. Gram-positive bacteria have a much thicker peptidoglycan layer surrounding the inner membrane; therefore, they can retain the dye better and appear purple under the microscope (Figure 2). On the other hand, Gram-negative bacteria have a much thinner peptidoglycan layer sandwiched in between their inner and outer membranes; thus they do not retain the crystal violet-iodine dye. Besides the thickness of the peptidoglycan layer, another main difference between them is that Gram-negative bacteria have an outer membrane layer that Gram-positive bacteria lack.

**A. Gram-negative bacteria**

**B. Gram-positive bacteria**

*Figure 1. Schematic representation of a eukaryote (left) and prokaryote (right)*
The inner membrane (IM), which is composed of phospholipid and numerous essential membrane proteins, encloses the cytoplasm. Because of the lack of organelles in prokaryotes, important membrane-associated proteins such as the ones that are involved in energy transduction, lipid biosynthesis and protein secretion all reside in the IM [2]. In between the two membranes in Gram-negative bacteria, there is a concentrated gel-like matrix called the periplasm which hosts proteins of various functions such as polymer degradation (DNases, RNAses and proteases), transport (sugar, amino acid, and vitamin-binding proteins) and folding (DsbA and SurA). The periplasm is also the place where a thin layer of peptidoglycan resides (Figure 2). Because of its rigidity, peptidoglycan is what gives structural strength and determines cellular shape. When it is damaged, for example by the treatment of lysozyme, bacteria assume a spherical shape, resulting in what’s call spheroplasts, which are fragile and can lyse due to osmotic pressure. The outer membrane (OM) is connected to the peptidoglycan layer by the Braun lipoprotein, which is the most abundant protein in E.coli. Unlike IM, the OM contains phospholipids only in the inner leaflet. The outer leaflet is composed of glycolipids, mostly lipopolysaccharide (LPS), which helps to keep the structural integrity of bacteria. LPS is infamous because it can induce a strong immune response in humans when its lipid portion, LipA, is released into blood circulation, which can cause a potential fatal endotoxic shock [2]. OM is where all the β-barrel porins reside. Some of these porins include OmpF and OmpC which function
to allow passive diffusion of small molecules such as disaccharides and amino acids across the membrane, and LamB, which allows diffusion of maltose and maltodextrins [2]. Gram-positive bacteria don’t have an outer membrane, but instead, they have a much thicker peptidoglycan layer (20-80nm) surrounding the IM. Within the outer layer are specialized polysaccharides called teichoic acid and lipoteichoic acid which help to maintain the structure of the cell wall (Figure 2). Despite a thicker peptidoglycan layer, Gram- positive bacteria are more susceptible to agents such as antibiotics compared to Gram-negative bacteria. This is because they lack an outer membrane which serves as a selective barrier against various harmful agents.

3. **Protein Secretion-the challenges**

More than one-third of bacterial protein that is synthesized in the cytoplasm needs to be exported in order to acquire function [3]. These final destinations in the case of Gram-negative bacteria include the IM, periplasm, OM and the extra-cellular environment. Whether a protein is translocated into or across the IM is not an easy task as it is a tightly sealed lipid bilayer. Furthermore, translocation of a hydrophilic and charged polypeptide is not thermodynamically favorable due to the hydrophobic nature of phospholipids. To ensure efficient translocation, bacteria have evolved multiple translocation pathways dedicated to the export of various kinds of substrates. Even with this mechanism, to get to their final destinations isn’t easy. It involves intricate networks of protein-protein interactions that need to be fine-tuned. Initially proteins need to be kept either unfolded or folded depending on the type of
translocation pathway they use. The former state requires the binding of cytoplasmic chaperones that keep them in an unfolded, translocation-competent state. Next proteins need to be sorted and targeted to the right translocation machineries with high fidelity. They also need to activate these machineries and be translocated into the bacterial IM, periplasm OM or beyond in an ATP-dependent manner [3]. For proteins in the periplasm that need to get across the OM, they have to be further sorted, targeted to the right OM secretin before passing through the OM. Getting from the periplasm to the OM and beyond has to be fueled by conformational energy since the periplasm lacks ATP or other high-energy carriers [4].

4. **Protein Secretion Across the Cytoplasmic Membrane**

The first barrier for any protein export is the bacterial IM. Bacteria have evolved two universal systems for this layer: the general secretion (Sec) pathway and twin arginine (Tat) pathway. These two pathways are the most conserved and are found in all three kingdoms of life (Bacteria, Archea and Eukarya [5]). Proteins translocated through these pathways can reside in the IM, the periplasm, or even to the OM or beyond with the help of other distal secretion systems.

4.1 **Sec Secretion System**

The Sec secretion system is by far the most commonly used system for proteins that are targeted to the bacterial IM and periplasm. Various components of this system are
essential for cell viability. Details of how it works are described in the following sections. In brief, proteins that are targeted to the Sec system require an N-terminal cleavable signal sequence or uncleaved signal-anchor sequence and need to be kept in translocation-competent unfolded state. This system is divided into 1) the co-translational translocation pathway, which is largely dedicated to the translocation and membrane integration of integral membrane proteins, and 2) the post-translational translocation pathway, which is responsible for secretion of proteins into the periplasm (Figure 3). In the co-translational route, a signal-anchor sequence emerges from the ribosome-nascent-chain (RNC) complex and is recognized by the signal recognition particle (SRP) based on its high level of hydrophobicity. It then targets the RNC complex to its SecY-bound membrane receptor called FtsY. The SRP-FtsY complex stimulates each other’s GTPase activity, which causes subsequent transfer of RNC to the SecYEG translocon complex. The energy from translation is thought to drive integration of the membrane protein laterally from the SecYEG translocon complex into the membrane [6, 7]. In the post-translational route, a fully translated protein is recognized by cytoplasmic chaperones such as SecB, which helps to keep the bound substrate in a translocation-competent, unfolded state before delivering it to the SecY-bound ATPase motor protein called SecA. SecA then uses the energy of ATP binding and hydrolysis to drive translocation of preproteins through the SecYEG translocon into the periplasm [7, 8]. After this point, proteins will fold and assume a three-dimensional structure and either stay in the periplasm, integrate into the OM, or be secreted outside of the cell through a Type II Secretion System (described below and
in Figure 4). Alternatively, they can also stay unfolded and secrete themselves through a Type V Secretion System (described below and in Figure 4) [5, 9]. These terminal branches of the Sec system are described below.

**Figure 3.** Schematic representation of A), co-translational translocation and B), post-translational translocation [7]

### 4.2 Tat Secretion System

Unlike the Sec system, the Tat secretion system transports folded proteins. Protein substrate for this system contains a pair of “twin” arginine residues in the motif Ser-Arg-Arg at the N-terminal end of a cleavable signal peptide. In *E.coli*, there are three different proteins involved in this system, TatA, TatB and TatC. TatB and TatC bind the arginine motif of the signal peptide, which then recruits TatA to the cytoplasmic membrane where it forms a channel [5, 10]. Folded protein translocates through the
Tat system, which either remains in the periplasm or is transported out of the cell through a Type II secretion system as described below.

5 Protein Secretion Beyond the Periplasm and Sec-independent Systems

Prokaryotic protein secretion beyond the periplasm can generally be divided into two categories, Sec-independent (Type I, Type III, Type IV) and Sec-dependent pathways (Type II and Type V). The former pathways are one-step processes that translocate protein straight from the cytoplasm across the OM and into the extra-cellular space, or even into the cytoplasm of a Eukaryotic host. The latter pathways, in general, are two-step processes that first translocate substrates into the periplasm through Sec or Tat pathways (Figure 4). Comparisons of selected pathways are discussed below.
Figure 4. Schematic of prokaryotic Sec-dependent and Sec-independent secretion pathways [5].

5.1 Sec-independent

A. Type I Secretion System (T1SS)

T1SS has been found in a large number of Gram-negative bacteria, including pathogens of plants and animals, where it transports substrates in a one-step process across both the IM and OM (Figure 4). The well-known bacterial pathogen that uses this pathway includes *V. cholerae*, which secretes HlyA hemolysin protein that ruptures the red blood cell among other host targets. T1SS closely resembles a large family of ATP-binding cassette (ABC) transporters. The core of this secretion system is composed of three parts, an ABC transporter protein in the IM, a membrane fusion protein (MFP) that bridges the ABC transporter to the third component, and an outer membrane factor (OMF). Together, they are involved in export of small molecules such as antibiotics or macromolecules such as toxins out of the cell. The ABC transporter consists of a cytoplasmic nucleotide-binding domain and a transmembrane domain made out of six α-helices that form the IM pore. It catalyzes ATP hydrolysis to fuel the transport process, interacts with MFP and participates in substrate recognition. The MFP interacts with the ABC transporter and forms the channel by which substrates can travel to the OMF. T1SS often uses the multi-purpose protein called TolC as its OMF. TolC only assembles and forms the OM pore once the ABC transporter and MFP have assembled with substrate. After translocation, TolC
disassembles and can serve as an OM porin for other secretory systems. Substrates for T1SS are exported in an unfolded state and typically contain a C-terminal signal sequence that remains uncleaved during the process [5, 9, 11].

B. Type III Secretion Systems (T3SS)

T3SS have also been found in a large number of Gram-negative bacterial pathogens. They are involved in export of a variety of protein substrates across both the IM and OM of bacteria and the host membrane, straight into the cytoplasm of the host cell in one step (Figure 4). Once inside the host, these proteins may modulate a variety of host cell functions, including immune and defense responses. T3SS have 9 core proteins that are highly conserved among different subclasses of T3SS, along with 10 to 20 proteins that play important effector functions. The overall structure of T3SS looks like a needle and syringe, composed of three main structural components: a base complex or basal body that contains a cytoplasmic ATPase domain and forms several ring-like structures that span the IM and OM; a needle that protrudes out from the OM into the extracellular space, and a translocon at the tip of the needle. Once the needle is in contact with the host membrane, the translocon will form a pore to allow secretion of proteins straight into the cytoplasm of the host. The secretion signal is embedded within the N-termini of substrates, and just like the signal sequence of T1SS substrates, is not cleavable. Once targeted to the basal body, they are secreted in an unfolded state in an ATP-dependent manner. In addition to fulfilling the export function, T3SS can
also act as an assembly and export system in the production of the bacterial flagellum [5, 9, 12].

5.2 Sec-dependent

A. Type II Secretion Systems (T2SS)

T2SS are responsible for secretion of mostly enzymes such as proteases, lipases, amylases and phosphatases into the extracellular environment, in a folded state. Since the T2SS channel only resides in the OM, its substrates need to have an N-terminal cleavable signal sequence and translocate through the IM first, either through the Sec or Tat pathway. For this reason, T2SS is also recognized as the main terminal branch of the Sec pathway. Those that are transported by the Sec pathway into the periplasm will need to be folded before targeting to the T2SS OM channel. T2SS are composed of at least 15 different proteins organized into three distinct functional regions: an OM complex or secretin where substrates traverse through the OM, a periplasmic pseudopilus that forms a pilus-like structure that connects the IM platform to the OM porin, and a cytoplasmic protein which interacts with ATP, presumably acting as an ATPase to fuel the transport in some manner. The OM porin belongs to the secretin superfamily, which is also required for type 4 pilus biogenesis and type III secretion. It forms a highly stable, possibly gated secretin of 12-14 subunits with a diameter of 5-10 nm: large enough to accommodate folded substrates. One model for secretion through T2SS posits that the growing pilus may push secreted molecules through the OM complex [5, 9, 13].
B. Type V Secretion Systems (T5SS)

Another terminal branch of the Sec pathway are T5SS, one of whose sub-types is called an autotransporter secretion system. Substrate for this pathway is interesting and unique in that it doesn’t require any dedicated secretion apparatus or membrane channel because it can translocate itself. T5SS can transport proteins with diverse functions, including proteases, toxins and adhesions. A typical autotransporter is composed of three domains: an amino-terminal signal sequence for secretion across the IM through the Sec machinery, an integral passenger or functional domain which is connected through a linker to a carboxy-terminal β-domain. After signal sequence cleavage and translocation into the periplasm, the β-domain forms a 12-stranded OM β-barrel porin to translocate its passenger domain, which is either retained on the bacterial surface or released into the extracellular environment by proteolysis [5, 9].

6 Sec-dependent Protein Translocation-general requirements

As aforementioned, protein secretion by the Sec pathway can be divided into two major routes: the co-translational pathway in which the translocation of integral membrane proteins through the SecYEG translocon is coupled to translation, and the post-translational pathway in which fully translated proteins are translocated into the periplasm and beyond, mediated by SecA. The former pathway is evolutionary conserved and is the major pathway for protein translocation into the ER membrane or the ER lumen in eukaryotes. It is not clear why bacteria possess the SecA pathway. It is thought that protein translation may be more rapid than translocation (particularly
at high ribosomal content); therefore, it is beneficial for fast-growing bacteria to have acquired this additional pathway to keep up with their metabolism [14]. Despite the differences in translocation mechanism, both pathways share two main pre-requisites for their substrates: they need to i) contain signal sequence and ii) remain in an unfolded state prior to translocation. The rest of this thesis will be dedicated to discussing mostly about the bacterial Sec system.

6.1 Signal Sequence

The signal sequence for the Sec pathway is approximately 18-30 amino acids in length and is generally located at the N-terminus of the newly synthesized secretory protein. There is very little sequence consensus among signal sequences, but they do share similar biochemical properties characterized by (i) 2 to 5 positively-charged amino acids on the N-terminus (ii) 7 to 15 consecutive hydrophobic amino acids in the middle and (iii) a polar C-terminus containing a signal peptidase processing site of 3 to 7 amino acids. Such a cleavage site contains a conversed Ala-X-Ala consensus, where X is any amino acid [3] and is cleaved at a later stage of translocation by signal peptidase on the trans-side of plasma membrane. The main function of the signal sequence is to target proteins into the Sec pathway, and the determining factor lies in its hydrophobicity. If it is highly hydrophobic (in the case of an integral membrane protein it is called a signal-anchor sequence), it will be recognized by the signal recognition particle (SRP) which targets proteins into the co-translational route; on the other hand, if it is less hydrophobic in the case of secretory proteins, it will be
recognized by SecB for targeting into the post-translational route. It was also shown that the signal sequence can slow down the folding kinetics of the mature region of preproteins and thus helps to maintain them in a translocation-competent state [15, 16]. It is worth mentioning also that unlike typical signal sequences, signal-anchor sequences for integral membrane proteins lack a signal peptide cleavage site and contain a longer hydrophobic core region.

Insertion of a signal sequence or signal-anchor sequence into the SecYEG translocon is the first step in membrane translocation. However, there are known suppressor mutations of SecY (prlA), SecE (prlG), SecG (prlH) and SecA (prlD) that can accommodate defective signal sequences [17-20]. The strongest of these are prlA mutations which are clustered in the lateral gate and the plug domains of SecY. Interestingly, such mutations lead to a more promiscuous translocon that can translocate preproteins lacking signal sequences [17]. It appears that these mutations stabilize the open conformation or destabilize the closed conformation of the translocon [18].

6.2 Protein Folding

Preproteins need to be kept in an unfolded, translocation-competent state before they can be translocated through the SecYEG translocon. In the co-translational pathway, this is achieved by coupling translocation to translation. The process starts when an emerging signal-anchor domain from the ribosome exit tunnel is recognized and bound by SRP, shielding it from water molecules. The signal-anchor domain was shown to
bind deeply into the methionine-rich hydrophobic pocket of the M domain of SRP [21-23]. In this state, protein synthesis is paused to prevent cytosolic exposure of mature sequences that could cause pre-mature folding [24]. Translation of the mature domain continues only when the ribosome nascent chain (RNC) complex is targeted to the translocon. For the post-translational pathway, fully synthesized presecretory protein is bound by the cytoplasmic chaperones such as SecB, which helps to keep preprotein in an unfolded state. When preprotein is handed over to SecA, the energy of ATP binding and hydrolysis drives its translocation across the SecYEG translocon [7]. In both pathways, hydrophobic regions of preproteins are kept away from excess exposure to water molecules prior to translocation. For hydrophobic membrane proteins, such a mechanism is especially important since contact with water molecules will lead to hydrophobic collapse and protein aggregation. Keeping substrates in an unfolded state is also favorable for SecYEG translocon function since the maximum expandable size of the channel can only accommodate unfolded protein [25].

6.3 Energetics of Protein Translocation

There are two major known sources of energy for Sec-dependent protein translocation. The most obvious one comes from SecA ATPase motor protein that harnesses the energy of ATP to drive translocation. The complete catalytic cycle of protein translocation permits step-wise translocation of ~5kDa of polypeptide. This is thought to be achieved in two steps: first, SecA binding to SecYEG translocon drives translocation of a 2.5 kDa segment; second, the binding of ATP drives translocation
of another 2.5kDa [26, 27], possibly through some kind of membrane insertion mechanism of SecA that presumably “pushes” the preprotein through the translocon. This is followed by ATP hydrolysis that leads to SecA membrane deinsertion and possibly dissociation from the translocon [28]. Such a cycle repeats itself until the entire preprotein is translocated. The exact mechanism of how ATP-induced conformational changes in SecA are coupled to protein translocation is still a controversy.

Besides ATP, proton motive force (PMF) also plays a stimulatory role. PMF is generated as a result of an electrochemical gradient across the bacterial inner membrane or the inner membrane of mitochondria and chloroplasts in eukaryotes. The electrochemical gradient itself is created by the electron transport chain that actively pumps protons out across the membrane, creating both a proton and electrical potential gradient. It has been suggested that PMF is the main driving force for translocation at the step after SecA-bound ATP is hydrolyzed, so that preprotein is prevented from backsliding [29]. In one study, a plug deletion mutant of SecY was characterized. This mutant behaved very similarly to the prl mutants mentioned above in that it can translocate preproteins with defective or missing signal peptides. It also showed enhanced SecA membrane insertion, a state that mimics the ATP-bound conformation [28, 30]. However, such a mutant is less dependent on PMF for translocation, suggesting that PMF and ATP act independently to facilitate translocation. PMF was also found to promote SecA membrane deinsertion from the membrane, a state that can also be achieved by ATP hydrolysis [28, 31]. Perhaps there is a certain degree of
crosstalk between the two energy networks, for example, PMF may regulates the ATP hydrolysis cycle.

7 Sec-dependent Protein Translocation-a detailed look

7.1 Co-translational Translocation

In bacteria, membrane proteins are targeted to the co-translational pathway for translocation into the inner membrane. The targeting process can be divided into i) recognition of the signal-anchor sequence by SRP; ii) targeting of RNC to its membrane receptor FtsY and iii) transfer of RNC to the SecYEG translocon.

Phase i begins when a newly emerging signal-anchor sequence from the ribosome exit tunnel is recognized and bound by SRP. Molecular crowding near the ribosome exit tunnel ensures efficient binding of SRP [32]. The hydrophobicity of the signal peptide seems to determine whether a protein is directed to the co-translational or post-translational route. The proof for this came from a study where when MBP or OmpA signal peptide was replaced with the first segment of a membrane protein called AcrB, they were both redirected from the SecB pathway to the SRP pathway [33]. Bacterial SRP is comprised of protein Ffh (a homologue of SRP54) bound to 4.5S SRP RNA [34]. Ffh has two functional domains connected by a flexible linker: a C-terminal M-domain that contains the signal peptide binding site, and an NG-domain that contains the binding site for ribosome, FtsY (the membrane-bound SRP receptor) and GTP (Figure 5A). SRP recognizes the signal peptide through binding with its methionine-rich hydrophobic site in the M domain (Figure 5B-C) [21-23]. Such binding causes a
$180^\circ$ rotation in the NG domain (compare A to B). In eukaryotes, once bound, protein synthesis is paused until the RNC is translocated to the SecYEG complex where translocation is coupled to translation [24]. This sequence of events is to ensure that there is no exposure of other hydrophobic TM helices to unfavorable interactions with surrounding water molecules. The N domain of SRP binds to ribosomal protein L23 near the peptide exit tunnel. This binding site overlaps with the SecY binding site, suggesting a mechanism by which SRP has to dissociate prior to translocation of RNC to the SecYEG complex [35-38].

Figure 5. Crystal structure of *Methanococcus jannaschii* SRP with or without signal peptide A) the structure of free SRP. Various domains are labeled. Structural rearrangement of NG domain in response to signal peptide binding is indicated. B) structural complex of SRP bound to a signal peptide. C) the detail of the signal peptide-binding site within the M domain [22]

In phase ii, SRP brings the RNC to its membrane receptor called FtsY. Without RNC, the weak electrostatic association between SRP and FtsY is mediated through the N sub-domain on each protein. Such an interaction can occur independent of GTP as the
G domain doesn’t interact extensively in this early intermediate [39, 40]. However, binding of RNC to SRP greatly enhances SRP’s affinity for FtsY, and creates a stable SRP-RNC-FtsY complex. This ensures a rapid delivery of RNC to the membrane [41]. The presence of the membrane further fortifies the interaction between the two NG domains. In vitro binding studies showed that FtsY preferentially binds to anionic phospholipids such as phosphatidylglycerol (PG) and cardiolipin [42, 43]. Interestingly, SecY was also shown to tightly associate with cardiolipin and acidic phospholipids, both of which are also required for SecA lipid and translocation ATPase activity [44, 45]. This suggests that sites in the membrane that are enriched for these phospholipids are favorable sites for protein translocation. Given its favorable interaction with acidic phospholipids, it is not surprising to find FtsY in contact with SecY [46, 47]. This provides a possible route for transferring of RNC to SecYEG for translation-coupled translocation.

In phase iii, SRP and FtsY trigger each other’s GTPase activity, which enables transfer of the RNC complex to the FtsY bound-SecYEG translocon. In order for this transition to be successful, the initial tight interaction between SRP and RNC needs to be weakened. This is achieved by the post-GTP hydrolysis conformational change in the NG domain that helps SRP to switch from an RNC-binding mode to a releasing mode, thus passing RNC to the SecYEG translocon [41]. GTP hydrolysis also leads to disassociation of SRP from FtsY. Interestingly, in the mammalian system, GTP hydrolysis is inhibited in the absence of a functional translocon, preventing disassociation of SRP from its receptor, suggesting that the Sec61 complex regulates
the GTPase cycle of the SRP-SR complex. This unique check-point ensures that no futile transferring of RNC occurs [48]. Translation resumes once SRP is docked onto the SecYEG translocon. Subsequent translation-coupled translocation ensures efficient one-way integration of membrane protein into the lipid bilayer, as compared to post-translational translocation where backsliding can occur. Insertion of the first TM helix, also called the signal anchor domain occurs at the hydrophobic groove just outside of the lateral gate of SecY/Sec61α. Integration of membrane protein into the lipid bilayer is facilitated by lipid partitioning and integral membrane chaperones [49]

7.2 Integration of Membrane Protein

The exact mechanisms of membrane protein insertion into the plasma membrane and subsequent folding into its final functional state are largely unknown, perhaps due to the immense diversity in size, structure and specific topological arrangement. It is well known that lipid partitioning plays a critical role in the lateral movement of membrane protein from the aqueous interior of the translocon channel into the membrane bilayer. However, getting into the membrane is just the first step towards final assembly. Before this, TM helices that exit the translocon must be in the right topology and subsequently make very specific contacts with other TM helices in order to obtain tertiary structure. One can also image that folding may need to be precisely coordinated with the translation process or aided subsequently by membrane chaperones. This is because two adjacent TM helices exiting the translocon may not make structural contacts in the final folded state; thus one helix may have to “wait”
and “skip” a few helices before its “soulmate” exits the translocon when folding can finally take place.

In both eukaryotes and prokaryotes, accessory proteins such as translocating-chain associating membrane protein (TRAM) and YidC have been implicated in membrane protein biogenesis. The support for this has largely come from their ability to crosslink to translocating substrates and the Sec translocon. They are also required for membrane insertion of a subset of integral membrane proteins. In particular, YidC can function independently from the Sec translocon to insert specific small membrane proteins with 1-2 TM segments [50, 51] (Figure 6).

![Figure 6. Roles of YidC in membrane protein biogenesis. A) SecY alone mediates the integration of membrane protein by lipid partitioning; B) some proteins need help from YidC for their membrane insertion. In this case, TM helices that are inserted and exit the SecYEG translocon are handed over to YidC first. YidC which helps with membrane folding can release TM into the membrane either one by one or as an assembled structure; C) YidC can act as an insertase on its own for a subset of membrane proteins; D) YidC can work sequentially, first by inserting helices on its own, then work with SecYEG to insert the rest of the protein. [52]

It is well established that membrane protein integration requires a high level of hydrophobicity. However, hydrophobicity must not be the only determining factor
since TM helices of certain proteins (e.g. bovine rhodopsin and voltage-gated potassium channels) are largely polar [50, 53]. This suggests that positional context within a large protein complex also plays an important role in membrane integration.

The topology of TM helices was initially thought to be governed by the positive-inside rule in which positive charges flanking a TM domain are statistically enriched in the cytoplasm [54, 55]. However, later it was found that the presence of PMF also influenced the topological determination [56]. This possibly is the reason why the positive-inside rule is less effective in eukaryote which lack PMF across the ER membrane. It is also unclear which TM helix determines the overall topology in multi-spanning membrane proteins. Early work suggested that two adjacent TM helices serve as initiation and stopping signals for translocation, respectively, and they have alternate topologies. Therefore, insertion of the first TM determines the topology of the rest of the TM helices. However, for some proteins such as human P-glycoprotein, the cooperative interaction between the first two TM helices is required for the correct topology [57]. Furthermore, the topology of a membrane protein of four TM helices can be inverted by introducing a single positive charge at the C terminus [58]. This suggests that the charged residue not only has topological effects on the adjacent TM helices, but also on the topology of more distant helices. This finding also suggests that membrane proteins don’t adopt a finely folded structure until all TM helices are integrated into the membrane.

7.3 Post-translational Translocation
The vast majority of secretory proteins of *E.coli* reach the SecYEG translocon complex post-translationally [59]. In this route, preproteins are targeted to the translocon either by binding of ribosome-associated chaperones such as trigger factor, a ubiquitous protein in bacteria, or SecB, which is only found in some proteobacteria such as *E.coli*. Targeting can also be achieved by binding of ribosome-associated SecA with or without the help of TF or SecB [3]. Post-translational protein secretion can be divided into three distinct steps: i) sorting and targeting by which preproteins are targeted to the SecYEG translocon through concerted actions of binding and releasing by SecA and/or associated chaperones; ii) translocation priming by which the signal sequence of a preprotein is first bound to the SecA two-helix finger sub-domain as a hairpin and then is subsequently inserted into the lateral gate, priming the SecYEG translocon for translocation; iii) translocation of the mature region of the preprotein by SecA through multiple rounds of ATP binding and hydrolysis.

A. Sorting and Targeting-the action of trigger factor (TF)

Trigger factor is a 48kDa protein ubiquitously expressed in bacteria, with a cytoplasmic concentration around 40-50µM, about several times more abundant than SecA (5.7-8.2 µM) and SecB (4-20 µM) [3, 8]. It has three distinct domains, the N-terminal domain that is involved in ribosome binding, the C-terminal domain that is the site of main chaperone activity, and a PPlase domain (peptidyl-prolyl-isomerase) that exhibits auxiliary chaperone activity (Figure 7A) [60]. In the cytoplasm, TF can bind to vacant ribosomes, or self-dimerizes. However, its affinity for RNC is the
highest, with a $K_d$ value about 0.04-0.7μM [61-64]. This ensures that TF can distinguish between the vacant ribosome and the actively translating ribosome. TF-ribosome binding is mediated by the binding of its N-terminal domain to the ribosomal protein L23 (Figure 7B-C), the same site for the binding of SRP, suggesting a mutually exclusive binding and recognition mechanism. However, two independent studies also reported that TF and SRP can be detected next to each other at the ribosome tunnel exit [65, 66]. TF recognizes both the signal peptide and mature domain of the nascent chain, when about 100 amino acids have been translated [67]. Binding occurs at several sites in TF that are located within its PPIase, N-terminal and carboxy-terminal domains, which are enriched with both hydrophilic and hydrophobic residues. This along with great structural flexibility enables TF to bind to diverse sets of substrates [68-70]. Despite this, TF has higher preference for binding to nascent chains that are enriched in basic and aromatic residues, with a positive net charge [71]. This is very different from SRP, which tends to bind highly hydrophobic signal anchor sequences [72]. Such differential binding ensures that proteins that are bound by TF are eliminated from binding of highly competitive SRP. Despite its important substrate binding activity, TF is not essential to cell viability [73]. Binding of TF helps to keep nascent chains in an unfolded translocation-competent state before delivering it to SecA. The exact mechanism of how that occurs still remains elusive.
Figure 7. Structure of Trigger Factor from *E. coli*. A) Structural features of TF. Various domains include: (i) the head domain (PPlase domain) that serves an auxiliary chaperone function, (ii) C-terminal (arm1 and arm 2) domain that houses the main chaperone activity, and (iii) the N-terminal (tail) domain that is involved in ribosome binding [60]. B) Modeled structure of RNC-bound TF; C) the N-terminal domain of TF binds to ribosomal proteins L23 and L29 near the peptide exit tunnel [68].

B. Sorting and Targeting-the action of SecB

Post-translational targeting may also be done through the action of SecB. *E. coli* uses SecB for export of only approximately 4% of its secretory proteins [3]. This cytoplasmic chaperone is a homotetramer of 69kDa that is composed of a dimer of dimers [74]. The basic monomer is composed of four antiparallel β-strands on top of two α-helices (Figure 8A). The dimer is stabilized through hydrogen bonding between β1 of two protomers (Figure 8B). A tetramer forms by packing two α-helices of each dimer in between two eight-stranded antiparallel β-sheets (Figure 8C-E). Two peptide-binding subsites are found in each monomer: subsite 1 that is deep and enriched in aromatic residues, and subsite 2 that is shallower and enriched in hydrophobic residues (Figure 8E). Together, these subsites allow binding of ~20 amino acid-long peptide [75, 76]. A peptide-scanning experiment suggested that a SecB-binding motif in a
polypeptide is ~9 residues long and is enriched in aromatic and basic residues, whereas acidic residues are disfavored [77]. The aromatic nature of site 1 would be a perfect binding site that stabilizes substrate through ring-ring stacking interaction [74].

Figure 8. Crystal structure of SecB from *Haemophilus influenzae* (A-D) and *E.coli* (E). SecB monomer (A), dimer (B) and homotetramer (C-E) are shown. Each SecB dimer contains two proposed peptide binding sites (E, indicated as S1 and S2). The final SecB structure is formed by packing two α-helices of each dimer against each other, sandwiched in between two eight-stranded β-sheets (C, top view, D-E side view). The final proposed deep and continuous peptide-binding site is represented by the hatched rectangle (D) [74, 75]

SecB binds to non-native proteins with low specificity but high affinity [3]. Early work suggested that the rate of folding of wild-type MBP was faster when SecB was absent and greatly retarded when excess SecB was present [78]. When the folding and unfolding pathway of MBP was studied using optical tweezers and MD simulations, it was also found that SecB only binds to the extended or molten globule-like structure [79]. Therefore, it is well understood that SecB binding keeps proteins in an unfolded, translocation-competent state.
After binding to preprotein, SecB must deliver it to either cytoplasmic or SecYEG-bound SecA ATPase protein. In the presence of preprotein, the affinity of SecB for membrane-bound SecA is very high (K_d of ~10nM) [80, 81]. A major basis of SecA-SecB binding is the electrostatic interaction between the positively-charged amino acids on the C-terminus of SecA and the negatively-charged amino acids on the solvent-exposed side of the β-sheet of SecB [75]. Binding of SecB to SecA is transient, as binding of ATP to SecA leads to disassociation of SecB, suggesting that SecB is released very early during translocation [80]. The basis of substrate transfer is less well known. It has been suggested that binding of the C-terminal α-helices of SecB to the dimer interface of SecA disrupts the SecA dimer, which causes the PPXD domain of SecA to be in a more open conformation to receive protein substrate [82]. Like TF, SecB is not essential to cell viability because deletion of it only resulted in a cold-sensitive phenotype in *E. coli* [83]

C. SecA-the ATPase Motor Protein

After receiving preprotein from SecB or TF, the rest of translocation is driven by SecA ATPase motor protein utilizing the energy of ATP. SecA is not only capable of translocating secretory protein, but it is also involved in translocating integral membrane proteins with large periplasmic domains [84]. It is found in both the cytosol, either free or ribosome-bound, and the membrane as SecYEG-bound, both of which are capable of binding to preprotein. In addition to serving as an ATPase, it has two additional functions: one, it acts as a chaperone that binds to preprotein in the cytosol
and keeps it unfolded [3], and two, as a targeting factor (similar to SRP) that binds to
the ribosomal protein L23 at the peptide exit tunnel and to the translating polypeptide
before targeting it to the SecYEG translocon [85].

i. SecA Structure and Function

SecA belongs to the superfamily 2 of DExH/D proteins, which includes various
helicases and nucleic acid modifying enzymes because its motor domain contains all
nine conserved motifs of DNA/RNA helicases included in this family [86, 87]. Like
all other proteins in this family, SecA has two RecA-folds, NBD1 (nucleotide-binding
domain 1) and IRA2 (intramolecular regulator of the ATPase, also known as NBD2),
both contain the universally conserved Walker A and B motifs (Figure 9) [88]. ATP is
sandwiched in between the two NBD domains during binding and hydrolysis. A highly
conserved salt-bridge called gate 1 is formed between the positively charged Arg566
and the negatively charged Asp212 residues at the base of the motor domain. It
controls the opening and closure of the nucleotide-binding cleft between NBD1 and
NBD2, and propagates the binding of preprotein in the preprotein-binding domain
(PBD). This ensures that substrate binding is coupled to the ATPase cycle [89]. What
sets SecA and other proteins in this family apart is the presence of additional domains
linked to the motor domains. In the case of SecA, two extra domains that confer its
functional specificity are PBD, which extends from NBD1 and the C-terminal region
[8, 86, 90]. PBD is the site that crosslinked to protein substrate, hence it is also called
the preprotein-crosslinking domain, or PPXD. This domain undergoes a ~80° rigid
body rotation towards NBD2 in the SecYEG-bound state (Figure 9C). This forms a clamp that captures protein substrate and positions it right above the SecY channel for translocation [91]. The C-terminal region can be further divided into 4 domains: i) the helical scaffold domain (HSD) which is a long $\alpha$ helix that interacts with all four domains of SecA and helps to transmit conformational change in the NBD1 and NBD2 domains to the C-terminal region in response to ATP binding and hydrolysis; ii) the helical wing domain, the function of which is unknown; iii) the IRA1, also known as the 2 helix-finger (2HF), which has been implicated as a “piston” in the proposed power-stroke mechanism of protein translocation (discussed in more detail in a later section) [91]. It is also implicated in the regulation of SecA ATPase activity, as SecA mutant N68 lacking IRA1 showed an elevated ATPase activity [92]; iv) the extreme C-terminal flexible tail that contains a zinc-finger motif that binds to zinc ions, lipid and SecB [93, 94]. It also functions as an inhibitor of signal peptide binding as it can self-associate with the signal peptide-binding site [95].
Figure 9. Structures of SecA ATPase. Crystal structures of SecA alone (A, PDB ID 3JUX), bound to signal sequence (B, PDB ID 3JV2) and bound to SecYEG (C, PDB ID 3DIN). Various domains of SecA are labeled A: NBD1 (orange), NBD2 (yellow), HSD (green), IRA1 (purple), PPXD (red) and HWD (blue). A conformational change of PPXD from wide open (A) to open (B) to closed (C) states in response to SecA ligand binding is indicated with a dash line (the clamp region).

ii. SecA Oligomerization-structural studies

One important property of SecA is its ability to form different dimers. Given its abundance inside the cell, which is about 5.7-8.2 μM [8] and its Kd value for self association of 0.5-1μM (in physiological salt and temperature [96]), it is expected to exist mostly as a dimer. There are eight high resolution SecA structures that have been solved by X-ray crystallography [91, 97-104] and one by nuclear magnetic resonance (NMR)[95]. All of these structures displayed SecA as a dimer (either within the unit cell or by crystal contacts), each with unique positioning of each protomer. Most dimers are arranged in an anti-parallel fashion (Figure 10A-D) [98, 100, 101, 105, 106], except one with protomers oriented in a parallel fashion (Figure 10E) [99]. It is hard to imagine all of these are physiological dimers. At least, some of these might be due to artifacts arising from crystallization conditions, for example, high salt, the condition known to monomerize SecA. Salt induces monomerization because the dimer forms by electrostatic and hydrophobic interactions between two protomers [96-98, 105, 107]. It is possible that during the crystallization process, SecA might have dissociated first due to the high salt condition and then reassembled into a dimer using crystal packing restraints [86, 108]. SecA is known to play multiple roles in the Sec pathway as mentioned above. Different dimeric forms may have functional importance.
specific to particular roles that SecA plays. Furthermore, SecA is also known to be a structurally flexible protein. Therefore, it is not a surprise that its various domains might pack slight differently in each crystal structure.

Figure 10. SecA dimer crystal structures from different organisms. SecA dimers from B. subtilis (A and D), E.coli (B), M. tuberculosis (C), and T. thermophilus (E) are shown, with one protomer colored cyan and the other protomer colored grey. The corresponding PDB ID associated with each structure is indicated. All dimers are formed using different dimer interfaces, with the two protomers arranged in an antiparallel fashion, except the one from T.thermophilus (E), which is arranged in a parallel fashion [109].
iii. SecA Oligomerization during Protein Translocation

The oligomeric form of SecA is affected by ligands and other factors that it encounters during protein translocation. For example, liposomes containing E.coli phospholipids monomerized SecA dimer as evident by FRET and crosslinking experiments. In addition, the synthetic signal peptide KRR-LamB, but not the mutant peptide KRR-LamBΔ78 monomerized SecA [110]. In contrast, the same signal peptide was shown to redimerize phospholipid-monomerized SecA [111] or had no effect [95]. A number of studies have tried to artificially populate the monomer or dimer in order to study their function. For example, studies utilizing covalently-linked SecA dimers to perform the in vitro protein translocation assay found that dimers were as active as WT SecA. These dimers were linked in different orientations, for example, through cysteine disulfide bonds using either native cysteine located at the C-terminal tail of SecA [112] or engineered cysteine at the SecA dimer interface [113], or by gene fusion in tandem [114]. Collectively, these experiments suggest that SecA dimer dissociation is not necessary for protein translocation. Likewise, experiments have also been done to investigate SecA monomer function. For example, a SecA mutant in which the 11 N-terminal amino acids were removed, which shifted the equilibrium towards the monomer in solution, showed a defect in vitro for SecA translocation ATPase activity and was non-functional in vivo at more normal expression levels [115-117]. The landmark study that helps to resolve these conflicting studies came from the co-crystal structure of the SecA and SecYEG complex, crystallized in the presence of ADP-BeF₅, the molecule that mimics the intermediate state of ATP hydrolysis. Therefore, this
complex presumably represents the intermediate state of SecA-driven protein translocation. Although substrate was not included, binding of a single SecA to SecYEG caused extensive structural changes in the complex, with marked opening of the lateral gate. Importantly, SecA PBD underwent an ~80Å rotation towards NBD2 compared to the B. subtilis SecA structure. This created a continuous groove formed by the SecA PDB and NBD2 domains along with SecYEG, which appears adequate for passage of substrate from SecA to SecYEG. Thus it was proposed that monomeric SecA is the minimal functional unit of SecA in the translocon-bound state [91]. However, it should be noted that high salt and detergent conditions were used during crystallization, so the possibility of SecA dimer binding cannot be entirely excluded [96, 110].

A number of SecA dimer structures as well as crosslinking data are not compatible with SecA-SecYEG crystal structures. For example the residues implicated in dimerization of M.tuberculosis SecA antiparallel dimer are close to or within the SecA-SecYEG binding surface [97]. Likewise residues 611 and 637 used to create the artificial SecA dimer that showed in vitro translocation ATPase activity also localized at the binding interface of SecA and SecYEG [113]. These residues are not capable of participating in dimer formation, at least when SecA is bound to SecYEG during protein translocation. Recently, an in vivo study was performed utilizing a combination of photo and cysteine-crosslinking techniques to investigate SecA dimer in the cytosol and in the SecYEG-bound translocation-active states. It was found that a discrete antiparallel 1M6N-like Hunt dimer (Figure 10D) populates the pool in both cases,
suggesting the exclusive function of a Hunt-like dimer [109]. This finding matches well with the crystal structure of the SecA-SecYEG complex because the residues lining the Hunt dimer interface are also exposed to the cytoplasm, making it possible to bind to another SecA protomer [91, 101, 118]. This evidence suggests that a Hunt-like dimer is likely to be the physiological dimer. However, considering SecA’s various functions inside the cell, other dimer states may still be functionally important but present at lower level. With regards to this possibility, it has been suggested that various combinations of hydrophobic and electrostatic interactions allow protomer sliding, thus generating multiple distinct SecA dimers that can be categorized into i) major electrostatic dimer, ii) a minor salt-resistant dimer, and iii) SecYEG and signal-peptide stimulated “triggered dimer”. All of these dimers maybe functionally relevant during protein translocation and difficult to detect by in vivo crosslinking [119].

iv. SecA and Preprotein Binding

SecA binds to both signal peptides and mature domains of preproteins. An earlier study showed that SecA interacts with preproteins by recognizing the positive charge at the amino terminus of the signal peptide [120]. A later study stressed the importance of hydrophobic core recognition. In recent years, several studies have focused on finding the signal sequence binding site in SecA, using approaches such as FRET[121-123] and photo-crosslinking [124]. A study by Gelis et al. determined the NMR structure of SecA bound to a signal peptide. This study revealed a signal peptide binding site that is enriched with mostly hydrophobic amino acids surrounded by some negatively
charged amino acids. This enables SecA to accommodate signal peptides of different sequences through extensive hydrophobic and minor electrostatic interactions [95]. Such a mode of recognition matches well with the typical biochemical properties of the signal peptide. A recent study employed FRET to map the binding site of the signal peptide plus the early mature region on the SecA-SecYEG complex. What’s unique about this study is that the preprotein was fused to the C-terminus of SecA, thus enabling rapid and potentially more stable binding. This study showed that the signal peptide and early mature region formed a hairpin parallel to the 2HF [125], in marked contrast to the NMR structure, which showed perpendicular binding of signal peptide to the 2HF [95]. In addition, this study also identified an early mature region-binding site in immediate vicinity to the signal peptide binding site. This is similar to another recent study that showed that SecA contains what’s called a patch A site that accommodates mature sequences. Such a site is distinct and located adjacent to the distal end of the signal peptide-binding site [126]. Collectively, these studies speak to a mechanism by which signal peptide and mature regions are recognized differently by SecA.

v. SecA Membrane Insertion Property

In vitro studies have demonstrated that SecA possesses lipid-binding properties. Such binding is enhanced in the presence of anionic phospholipid and high temperature [127]. Interestingly, regions of SecA, including parts of the N-terminal, central and C-terminal portions of SecA were found to be exposed to the periplasmic side of the
inner membrane. In particular, the regions around amino acid residue 300 and 350 of the PPXD were periplasmically exposed only when ATP was added, suggesting an ATP-dependent membrane insertion character of SecA [128]. Economou et al. have also shown that the 30kDa C-terminal region of SecA became protease resistant during the in vitro protein translocation assay. Such a fragment of SecA formed in an ATP, preprotein, acidic phospholipid and SecYEG-dependent manner, the same requirement as endogenous protein translocation. In addition, the insertion of the 30kDa region into the membrane was directly coupled to protein translocation [129]. These observations suggest that this fragment of SecA might undergo membrane insertion and deinsertion, possibly serving a power-stroke function to “push” preprotein across the channel. However, looking at all of the literature suggests that the integral membrane form of SecA may be a larger phenomenon than just the 30kDa region and membrane cycling during active protein translocation; other regions of SecA can integrate into the membrane in a translocation-independent fashion (e.g. a 48 KD domain of SecA was found persistently embedded in the membrane regardless of translocation status [130]). A broader study looking at different regions of SecA is needed to define the overall structural form of integral membrane SecA.

D. SecYEG Translocon-structure and Function

The core component of the Sec machinery is the heterotrimeric, universally conserved SecYEG or Sec61αβγ translocon complex that resides in the inner membrane of Bacteria or Archaea and the ER membrane of Eukarya. Earlier mapping of various
components of the Sec translocon was done through cysteine crosslinking [131-133].

The main structural insight came from a high resolution X-ray crystal structure of Sec61αβγ from Archaea *Methanococcus jannaschii*. Since the Sec translocon is universally conserved in its sequence (e.g. *M. jannaschii* shows 50% sequence similarity to Eukarya and Bacteria species) this structure is probably representative of all structures. This turned out to be true when compared to newer crystal structures from *P. furiosus* [134], *T. thermophilus* [135] and *T. maritima* [91]. One difference is that the bacterial SecG (β) subunit has an additional TM segment that precedes the one that is common to all SecY/Sec61 complexes [136].

The translocon resembles an hourglass with larger aqueous openings facing the cytoplasm and periplasm and narrowing toward the center. The cytoplasmic opening is about 20 to 25Å with a few C-terminal cytoplasmic loops (C5 and C6) protruding out from the plane of membrane (Figure 11A). The core component, α subunit (SecY), is made up of ten TM helices arranged in two clamshell-like structures, with TM1-5 and TM6-10 each forming one clamshell. Two portions are related to each other by a pseudo-symmetry through a two-fold rotational axis vertical to the plane of membrane (Figure 11B) and is held together by an amphipathic γ subunit (SecE), which has its TM crossing through the membrane diagonally, making multiple contacts with SecY TM helices (Figure 11A and C). The narrow middle region of the translocon is called the pore-ring, which is lined with hydrophobic amino acids with their side chains pointing toward the interior of the channel. This creates a narrow opening of roughly
5-8Å, suggesting that the crystal structure is in the closed conformation. Right below the pore ring, there is a small helix called the plug domain, which is a part of the broken TM2 (Figure 11D). Structural movement of this domain is coupled to channel activation and preprotein translocation. The β subunit (SecG) resides just a little outside of TM1 and TM4 (Figure 11B-C) [136]. The function of it is not well known, although it is dispensable in E.coli with its deletion resulting in a cold sensitive phenotype.

Figure 11. Crystal structure of SecYEG homologue Sec61αβγ complex from M.jannaschii. A) overall-structure of Sec61αβγ resembles an hour-glass with larger openings towards the cytoplasm and periplasm and narrowing towards the center. The cytoplasmic loops that are important for SecA and ribosome binding are labeled. B) Ten TM helices of the α-subunit forms two clamshell-like structures that open in the lateral gate region towards the membrane. C) View from the cytosol. The γ subunit forms a hinge that embraces two halves of the α subunit, making multiple contacts with different TM helices of the α subunit. D) Pore-ring residues are indicated by a yellow stick model. Displacement of the plug domain away from the channel center towards the γ subunit during translocation is indicated.
i. The Lateral Gate

Two clamshell-like structures of SecY can open into the lipid bilayer at a site called the lateral gate, so that integral membrane proteins can exit the channel into the membrane. Regardless of co-translational or post-translational translocation, the ability of the channel to open is the key to its function, as the channel with the lateral gate fixed in a closed state by a disulfide bond showed a dramatic decrease in translocation efficiency [137]. The lateral gate is formed mainly by TM2b and TM7 and partially by TM7, 8 and 3 [136]. The structures of SecYEG bound to SecA or the ribosome without substrate showed marked widening of the lateral gate, presumably priming the channel for subsequent signal peptide insertion [91, 138, 139]. An early photo-crosslinking experiment in a yeast in vitro system has shown that the signal sequence of prepro-α-factor (ppαF) formed an alpha helix of approximately two turns, contacted by both TM7 and TM2b. Importantly, every residue in the signal peptide could also be crosslinked to phospholipids, suggesting that the site of signal peptide insertion is in between the aqueous pore and the lipid bilayer [140]. This work was confirmed by the recent crystal structures of an active post-translational, SecA-substrate engaged, SecY complex [141] as well as the cryoEM structure of a mammalian, ribosome-bound, Sec61 complex containing a short secretory protein [142]. These structures showed that the signal or signal-anchor sequence is inserted at the lateral gate in a 45° angle relative to the plane of membrane. The hydrophobic segment of the signal sequence sits in a groove outside of the lateral gate contacting
the phospholipids. The overall similarity between post- and co-translational signal sequence insertion suggests a universally conserved mechanism of signal sequence recognition. The position of the signal sequence suggests that lipid partitioning is the mechanism for membrane integration [49, 143].

ii. The Pore-ring and Plug Domain

Translocation of integral membrane and secretory proteins are accompanied by major structural changes in the SecYEG translocon. However, it must not occur at the expense of its structural integrity. Both the pore-ring and plug domain serve gating function to restrict the channel from small molecules and ion leakiness during resting and translocating states. Indeed, it was found that four out of six hydrophobic residues of the pore-ring surrounded the translocating polypeptide forming a seal [141]. However, there is also evidence suggesting that some ions, especially anions such as Cl⁻, but not cations can permeate during in vitro SecA-mediated translocation [144]. In the resting state, the plug domain is located towards the center of the translocon, close to the pore-ring, and also participates in gating function. It has to relocate, possibly towards SecE, [141, 145] to allow passage of secretory proteins into the periplasm. However, the plug domain is not essential for cell viability. E.coli that expressed a plug-less mutant of SecYEG was viable, but it showed a more loosely-assembled channel, with prlA like phenotypes such as reduced requirements for PMF and WT signal peptide for efficient translocation [17, 30]. Therefore, the plug domain
seems to be involved in stabilizing various subunits of the translocon as well as the proof-reading function to make sure that preproteins have correct signal sequences. Remarkably, the crystal structure of a plug-deleted mutant from *M.jannaschii* showed new plug formation, although it lacked many interactions that normally stabilized the closed channel. This result suggests that the plug domain is important and explains why cells harboring the plug-less SecY mutant are still viable and able to translocate substrates with defective or deleted signal peptides [146].

iii. The Cytoplasmic Loops

The crystal structure of Secαβγ showed that the cytoplasmic loops towards the C-terminal end of Secα protrude out from the membrane into the cytoplasm. These include the C4 loop between TM 6 and 7, the C5 loop between TM 8 and 9 and the extreme C6 tail (Figure 11A) [136]. These sites are the major contact points between SecY and its binding partners SecA and the ribosome. The importance of these loops was first demonstrated in an early proteolysis experiment that showed that when the C5 loop of Secαβγ-containing microsomes was digested, the ribosome could no longer bind to the microsomal membrane [147]. Mutants of SecY deleted for the last six or more amino acids in the C6 tail showed significant defects in *in vitro* protein translocation [148]. Mori *et al.* utilized an *in vivo* photo-crosslinking approach to investigate the interaction between SecA and SecY cytoplasmic loops. They identified cytoplasmic loops C2, C4, C5 and C6 as the sites of SecA interaction. Interestingly, interaction with the C6 tail, but not other loops, was enhanced by treatment of the
E.coli cell with NaN₃, the agent known to arrest the ATPase cycle of SecA and stabilize the membrane-inserted state by prohibiting de-insertion. This led the authors to propose that there are two modes of SecA-SecY interaction: i) static interaction between SecA and SecY C2/C4/C5 loops, and ii) dynamic interaction between SecA and the C6 tail of SecY [149].

Recent advances in structural analysis have finally confirmed these earlier biochemical studies. The crystal structure of an anti-SecY Fab fragment bound to SecYE, which mimics SecA binding, showed major contacts with the SecY C5 loop, particularly with residue Pro352, which is universally conserved among different SecY species [135, 136]. A random mutagenesis study in the C5 loop identified the same residue to be absolutely essential for translocation function [150]. The importance of C5 was confirmed in the co-crystal structure of the SecA-SecYEG complex without substrate. The structure shows extensive contacts between the PPXD of SecA and the C4 and C5 loops of SecY, although some interaction with the N-terminal C2 loop was also observed. Likewise, the cryo-EM structure of ribosome-bound SecYEG also suggests extensive contacts between C4 and C5 of the α subunit of Sec61αβγ and ribosomal proteins L23 and L29 located close to the peptide exit tunnel [138, 139]. The binding of SecA and ribosome to the C-terminal loops causes widening of the lateral gate, and in the case of SecA, also displacement of the plug domain towards the front of the channel. Such binding presumably primed the translocon for insertion of a signal sequence. In summary, the cytoplasmic loops seem to be an allosteric trigger that
relays the binding of SecA and ribosome into motion of the lateral gate that primes the channel for subsequent translocation.

iv. SecE

SecE is a 14kDa protein in *E. coli* that embraces the two halves of the clamshell-like structures in SecY. It is composed of three TM helices with its amino-terminal region residing in the cytoplasm and its carboxy-terminal region residing in the periplasm [151]. In other organisms, for example *M. jannaschii*, the SecE homologue γ subunit of Secαβγ is composed of only two TM helices (Figure 11). This final TM is homologous to TM3 of *E.coli* SecE, which is the only essential TM segment. Sequence of *E.coli* and other organisms identified TM3 as the only shared segment with modest sequence conservation. The importance of TM3 was proven in an *in vivo* complementation experiment when the SecE homologue from *T.maritima* was shown to be able to complement the growth of a SecE null *E.coli* mutant, an otherwise inviable strain [152]. The crystal structure shows that SecE makes multiple contacts with SecY TM helices, holding the two halves of SecY together (Figure 11); therefore it seems that SecE is required for the correct assembly of the SecYEG translocon. Indeed, in the absence of SecE, SecY is rapidly degraded by the protease FtsH. In the FtsH mutant strain, overproduction of SecY is detrimental to cell viability as uncomplexed SecY is not being effectively degraded [153]. In this regard, it makes sense that SecE is over-expressed inside the cell compared to SecY, presumably as a way to make sure not only that each SecY gets one copy of SecE, but also that each
SecY is assembled into a functional translocon rapidly so that the limited pool of SecY is not being degraded.

v. SecG

E. coli SecG is 12kDa, the smallest subunit of the SecYEG translocon. Similar to SecE, the number of TM helices that constitute SecG varies among organisms, with little sequence conservation [7]. The co-crystal structure of the SecA-SecYEG complex shows limited contacts between SecA and SecG [91]. SecG is not essential to cell viability, but it has been shown to enhance protein translocation at low temperature (20°C or lower) as well as when PMF is absent [154, 155]. An earlier study also isolated prlH mutations in SecG, which allow secretion of preproteins with defective or missing signal sequences, much like other prl mutations [19]. Much of the proposed translocation function of SecG has come from a proteolysis study of in vitro protein translocation. It was shown that the cytoplasmic loop of SecG was protected from proteolytic digestion when SecA was arrested in its membrane-inserted state by treatment with AMP-PNP or azide, both of which are known to stabilize membrane insertion by inhibiting the SecA de-insertion step, but not when ATP was used. This led the authors to propose a model in which a topological inversion of SecG is coupled to the translocation-dependent membrane insertion and deinsertion cycle of SecA [156]. One caveat of this study is that the binding of SecA and preprotein might have induced a conformational change in the cytoplasmic loop of SecG that rendered it resistant to proteolytic digestion. The legitimacy of this finding has been disputed by
a later study that showed that when SecG was topologically fixed to SecY, *in vitro* protein translocation was not hampered [157]. However, the study didn’t address whether such a topologically-fixed SecG was able to enhance protein translocation under low temperature when SecG function is most important. More studies need to be performed to reveal the function of SecG.

E. SecDF-YajC

SecDF-YajC is a heterotrimeric protein complex that associates transiently with the SecYEG translocon. Although SecDF is not technically an essential protein, their inactivation led to a severe protein secretion defect and strong growth inhibition [7]. Earlier biochemical studies suggested that SecDF is involved in regulation of SecA membrane cycling, as depletion of SecDF in inverted membrane vesicle (IMV) led to a drastic decrease in the membrane-inserted state of SecA under translocation conditions [28, 158]. SecDF has also been implicated in PMF-dependent translocation [7]. The first high resolution crystal structure of SecDF came from *Thermus thermophilus*. The structural complex shows that the core transmembrane component is composed of 12 TM helices, with SecD and SecF each contributing six helices (Figure 12A-B). The two proteins are arranged in a pseudo-symmetrical manner with two major periplasmic domains called P1 and P4 rooted from TM2 and TM8 of SecD and SecF, respectively. P1 is further divided into the P1 head and P1 base domains which are connected by two flexible loop structures. The original structure was in the F conformation, but another conformation, I, which is related to F by a ~120° rigid
body rotation of the P1 head domain towards the periplasm, was also modeled [159, 160]. The P1 domain may be coupled to the movement of protein substrate as the P1 head could interact with the substrate [159].

The newest crystal structure of SecDF confirmed the existence of the I form (Figure 12C). MD simulation showed that deprotonation of the conserved Asp365 in the TM helix of SecD allowed the formation of a single file, hydrogen bonded, water chain (Figure 12D), which provides a pathway for conducting protons across the membrane. The same study also confirmed that the head group in the I form could crosslink to the protein substrate through in vivo photo-crosslinking. Putting all of these studies together, the authors proposed a model for SecDF-driven protein translocation (Figure 12E): In the initial F form, the P1 head domain is located more towards the membrane and binds to the periplasmic portion of the protein substrate that has already been translocated (Step 1). Transition from F to I form causes a “dragging” of the substrate towards the periplasm (Step 2). Deprotonation of Asp365 causes an influx of a proton from the periplasm into the cytoplasm, and in conjunction with this, the P domain releases the substrate and resets itself for the next round of substrate binding (Step 2) [161]. Further experiments are needed to validate and elucidate the detailed function of SecDF. Of particular interest would be to determine the crystal structure of substrate-bound SecDF to determine the nature of P1-substrate binding, and to address whether proton conducting ability is required for P1 domain movement and the “dragging” motion or not.
YajC is an 8kDa integral membrane protein with one TM helix and a large cytoplasmic domain. Little is known about its function in protein translocation. Studies have shown that YajC is not directly involved in protein translocation, and it is certainly not essential for cell growth. Its effects on protein translocation may be indirect because in the absence of YajC, the expression level of SecDF is greatly reduced [162], resulting in protein secretion defects. More studies are required to elucidate the function of this mysterious protein that appears to play a regulatory role.

**Figure 12. Crystal structure of SecDF complex.** A) Crystal structure of the full-length SecDF from *Thermus thermophilus* in the F conformation. Six TM helices of SecD and SecF are colored in red and blue, respectively. The P1 domain rooted from SecD is indicated. B) TM helices of SecD and SecF are arranged in a pseudo-symmetrical manner with the same colored scheme. C) Crystal structure of the full-length SecDF from *Deinococcus radiodurans* in the I conformation. P1 undergoes a ~120° rigid body rotation in comparison to A. D) MD simulation showing single file, hydrogen bonded, water chain (blue spheres) as a result of deprotonation of amino
acid Asp365. E) The proposed model of SecDF-enhanced protein translocation [159, 161]

F. YidC

*E.coli* YidC is a 61kDa protein of the inner membrane that interacts with the Sec machinery to assist in the insertion and membrane partitioning of integral membrane proteins [163]. It is a member of the Oxa family of membrane proteins, which consist of Oxa1p in mitochondria, and Alb3 in the thylakoid membrane of chloroplasts. YidC can also act as an insertase on its own, independent from SecYEG [164]. The protein was initially discovered as a part of Sec translocase when it was shown to crosslink to the signal anchor sequence of an actively translocating polytopic membrane protein called FtsQ [163]. YidC was co-purified with SecDF-YajC complex, suggesting that SecDFyajC might serve as bridge between SecYEG and YidC [165]. The site of interaction between the two complexes was later found to be between YidC and SecF. Such interaction is not essential as YidC lacking residues 24-264 of its periplasmic P1 domain, which was found not to interact with SecF, was fully functional for insertion of a YidC substrate, the protein subunit c of F1Fo ATPase [166]. This also suggests that the main function of YidC lies in its TM helices. Indeed, the main substrate-binding site was mapped to the vicinity of TM3 and TM5 [167].

The crystal structure of *E.coli* YidC shows that it has six TM helices and one periplasmic domain P1 rooted from TM1 (Figure 13A). The six TM helices are tightly packed together towards the periplasmic side, but more loosely packed towards the
cytoplasmic side, resulting in a hydrophilic groove opening towards the cytoplasm. Such a groove is in line with hydrophilic residues and a previous substrate-binding site that mapped in the vicinity of this groove (Figure B-C) [168]. The major targeting route of YidC-dependent substrates is the SRP pathway [7, 169]. After targeting and insertion of TM helices into the lateral gate, YidC can promote membrane partitioning. Indeed, in a recent study using in vivo photo-crosslinking, YidC was found to interact with four TM helices of the lateral gate. These interactions were not influenced by SecA binding, rather a new interaction interface was formed in response to RNC binding. This suggests that RNC, but not SecA, induced the conformational changes at the SecY-YidC interface which might promote lateral movement of substrate TM helices into the lipid bilayer [170]. This makes perfect sense considering that the action of YidC is through biogenesis of membrane proteins, not secretory proteins. Ironically, the crystal structures showed that SecA binding [91] caused more lateral gate opening than ribosome binding [138, 139], which predicts that the former will perturb SecY-YidC interaction more. More studies are needed in order to elucidate the exact function of YidC in Sec-dependent translocation.
Figure 13. The crystal structure of YidC from *E.coli*. A) The crystal structure of *E.coli* YidC. The P1 domain and six TM helices are colored blue and yellow, respectively. B) The substrate-contacting residues in the TM domain are boxed. C) Close-up view of the boxed areas showing residues bound to Sec-independent substrates colored green (Pf3 coat protein), pink (MifM), brown (Foc) and dark green (both Pf3 protein and Foc) [168]

8 SecY Oligomer-the unknown

Similar to SecA, the functional oligomeric state of SecYEG has been an ongoing debate. In the simplest case, a monomer may suffice for ribosome, SecA and/or substrate binding, translocon activation and subsequent translocation. By contrast, if the dimer is the functional unit, then it is likely to require more regulatory signals for the timely and correct assembly of the relevant ternary complex as well as collaboration between each protomer to perform its specific function during the transport process. This is a challenging topic to address because various binding factors during protein translocation will likely modulate its oligomeric state. Indeed, the oligomeric state of detergent-solubilized SecYEG in the absence of any ligand
revealed that SecYEG is in a dynamic equilibrium between monomer, dimer and even higher oligomers. These different forms have been observed through techniques such as density centrifugation [171], analytical ultracentrifugation [172], native gel-electrophoresis [173]. What makes this kind of study even more challenging is that there are two types of dimer orientations observed through various crosslinking approaches, and they are termed front-to-front (FtF) and back-to-back (BtB) based on the location of the lateral gate relative to the dimer interface (Figure 13). In the FtF dimer, two protomers contact each other through the lateral gate, potentially forming a composite channel that could translocate larger, partially folded substrates (Figure 13B). The BtB dimer is characterized by two SecE proteins forming the protomers major contact point, with the lateral gate at the ends of each protomer facing into the lipid bilayer (Figure 13C). This conformation is more suitable for translocation of membrane proteins that need unobstructed lateral opening to exit into the membrane. The physiological relevance of these dimers still remains elusive because crystallographic or cryo-EM studies often resulted in monomer-biased structures. Scientists have used multiple alternative approaches to address the functional oligomeric state of SecYEG. Although such studies have resulted in more disagreements in the field, they are generally agreeing that the functional state of SecYEG falls into two camps: monomer or dimer.
Figure 14. SecYEG monomer and dimer structures. A) SecYEG homologue SecαβƳ from *M. jannaschii* as described in Figure 11A. B-C) the proposed SecYEG dimers based on previous cysteine crosslinking studies: B) Front-to-front SecYEG dimer with the lateral gate of each protomer facing each other. TM7 and TM2b of the lateral gate are colored yellow and green, respectively. C) Back-to-back dimer with SecE (colored in red) of each protomer facing each other. The lateral gates are facing outwards towards the membrane.

There are several lines of evidence that suggest that SecYEG functions as monomer. The most direct evidence came from multiple crystal structures or cryo-EM structures of SecYEG bound to SecA or ribosome, with or without substrate. While most structures revealed only a single copy of SecYEG, a cryo-EM structure from Mitra *et al.*’s study showed a second non-translocating copy, and Zimmer *et al.* mentioned the second protomer being lost during the purification of their SecA-SecYEG structural complex [91, 174-176]. One caveat of these studies is the use of high amounts of detergent, which can monomerize the channel complex. In support of monomer function, when two different fluorescently labeled SecYEG proteins were reconstituted into giant unilamellar vesicles at a physiological concentration and *in*
protein translocation was initiated and arrested with an OmpA-dihydrofolate reductase substrate (OmpA-DhfR), no FRET signal was observed. Since this substrate contains full-length OmpA, this result suggests that monomer is sufficient for translocation initiation and elongation [177, 178]. In another study, when the protein substrate OmpA-GFP was used to arrest the translocon in a translocation-intermediate state, the dimer was shown to disassociate based on the loss of cysteine crosslinking, suggesting that the monomer is the functional unit during protein translocation [179]. However, this result could be an artifact due to a conformational change induced by the binding of SecA and OmpA-GFP to SecYEG, which could cause two cysteines to move out of favorable geometry for crosslinking. The legitimacy of this finding was tested in our study.

Other biochemical studies have pointed to the dimer as the functional unit. Interactions with SecA, preprotein, ATP and acidic phospholipid, all of which are required for protein translocation, have been shown to promote SecYEG dimerization, or even tetramer formation [44, 180]. Furthermore, covalently-linked SecYEG was capable of complementing the SecY temperature sensitive strain [181], suggesting the dimer is functional in vivo. One study aimed to access the functional size of the SecYEG channel. In this study, the precursor of OmpA was conjugated to rigid spherical tetraarylmethane derivatives of different diameters at a unique cysteine residue and in vitro protein translocation was performed. It was observed that the SecYEG translocon could expand to at least 22-24Å to accommodate translocation of a bulky proOmpA derivative [137], which is difficult to reconcile with an estimate of the maximum
expandable channel size of 16Å based on molecular dynamic simulation [25]. This result could be explained if SecYEG functions as a dimer, particularly as a front-to-front dimer in which two protomers can potentially fuse to form a composite channel. In another study, SecA interaction with SecYEG was mapped using an in vivo photocrosslinking technique, and it was found that SecY interacts with both the amino and carboxy terminal region of SecA. The only way to account for all of the observed interactions is to form a front-to-front dimer [182]. This result also alluded to a hypothesis that translocation occurs through one copy of SecA interacting with two copies of SecY, which itself is a controversy in the field. In an in vivo complementation experiment, two different SecY mutants, one defective in SecA binding, and the other defective in translocation, were co-expressed. Remarkably, the two defective SecY copies could form a functional dimer [183]. This result suggests a functional division between two protomers, for example, one copy serving as a SecA binding site, while the second copy translocates the preprotein. If this is the case, then it would be interesting to know how such a functional division is established, especially under in vivo conditions when two SecYs are initially functionally equivalent.

Further complicating the matter is a resent report that monomer or dimer appears to be utilized depending on the substrate being translocated. For example, by using radiolabeled preGBP or proOmpA that contained a stabilized C-terminal loop as substrates to arrest SecYEG in a translocation-intermediate state, it was found that GBP required twice the number of active translocon units compared to MBP [184]. The simplest explanation would be that dimer and monomer are required to translocate
GBP and MBP, respectively, although higher order oligomers are also possible. In order for this to happen, the cell would require a mechanism to sense the identity of protein substrates so that correct assembly of SecY can be achieved. Alternatively, the assembly of the dimer can take place at an early step after one protomer binds to the substrate.

9 Models of SecA-driven Translocation

Over decades of research, many models have been proposed to answer how SecA harnesses the energy of ATP to drive protein translocation. These models are designed to fit previous structural and functional studies. In general, these models can be categorized as follows: those that invoke a power stroke within SecA, those that invoke biased diffusion, or that invoke quaternary interaction between multiple SecA molecules [185]. All of these models assume that SecA undergoes some kind of conformational change during the process, but such conformational changes often lack experimental proof. Thus, more work is required to elucidate the mechanism of SecA-driven translocation.

9.1 Power-stroke Model

In the power stroke model, SecA harnesses the energy of ATP binding and hydrolysis to induce a certain conformational change in itself that the region in contact with preprotein substrate undergoes successive cycles of membrane insertion and deinsertion to push the preprotein across the SecYEG translocon (Figure 15A) [7, 86, 186]. Early evidence that supported this mechanism was the discovery that a protease-
resistant 30kDa domain of SecA inserted and deinserted into the membrane during in vitro protein translocation [28]. While this finding is compelling, membrane insertion of SecA seems to be a part of a larger phenomenon of SecA-lipid binding properties, as it was later discovered that a 48 kDa domain of SecA was persistently embedded in the membrane regardless of the protein translocation status of SecA [130]. The finding that translocation occurs in a stepwise manner in which 20-25 amino acids are translocated per ATP binding and hydrolysis cycle is also consistent with this model [27, 187]. Additional evidence for this model came from the co-crystal structure of the SecA-SecYEG complex. In this structure, the 2HF of SecA is partially inserted at the mouth of the translocon channel, suggesting a possible role for pushing preprotein substrate across the channel [91]. Indeed, it was found that the loop at the tip of the 2HF interacts with the substrate polypeptide chain right at the SecY cytoplasmic entrance [188]. However, nothing is known about the depth of 2HF insertion during protein translocation. One can imagine that deep insertion of SecA into the SecYEG channel is not favorable as it could potentially damage the structural integrity of SecYEG. However, in a recent study by Banerjee et al., SecA was shown, through in vivo photo-crosslinking, to insert deeply into the SecYEG translocon during OmpA-GFP jamming, with multiple contacts being made at the pore-ring and the lateral gate regions [109]. Whether such insertion was due to the 2HF or not still remains unknown. One study suggests that insertion doesn’t need to be made by the 2HF because when the 2HF was immobilized onto a cytoplasmic loop of SecY, in vitro protein translocation was not hampered [189]. However, it would be better if the 2HF
was immobilized onto a TM helix, as a cytoplasmic loop could still allow a certain
degree of mobility. Therefore, it is only fair to say from this study that major
movement in the 2HF domain is not necessary for protein translocation.

Recently, the power stroke model was updated by Bauer et al. to allow passive sliding
of substrate. In this model, ATP binding causes the tip of the 2HF to come into contact
with a subset of amino acids in the substrate and “drag” it toward the entrance of the
SecYEG channel. The 2HF doesn’t interact with every amino acid strongly, so the
power stroke doesn’t always allow effective translocation. When non-interacting
amino acids are encountered, passive sliding will allow the substrate to slide in either
direction. When ATP is hydrolyzed, the 2HF retracts, allowing passive sliding to
dominate. The next round of power stroke action will provide forward movement. In
this model, the purpose of the 2HF is to drive unidirectional movement of substrate
into the periplasm. Since SecA spends most of its time in the ADP-bound state, passive
sliding is a major contributor to translocation [190]. In addition to power stroke and
diffusion motions, the aforementioned binding of the P1 domain of SecDF to the
substrate on the trans-side of the membrane may also ensure unidirectional and
efficient translocation (Figure 12E), along with final substrate protein foldings. In the
eukaryotic cell, binding of the ATP-dependent Bip protein to the translocating
polypeptide on the luminal side of the ER membrane prevents backward sliding [191].
Thus, SecDF may be functionally analogous to Bip, in that it provides directionality
to protein translocation.
9.2 Inchworm Model

A more refined version of the power stroke model is the inchworm model (Figure 13B). This model was initially proposed for the monomeric DEAD helicases that move along DNA by means of their two nucleotide-binding sites [86]. Since SecA also belongs to the super family of DEAD helicases, a similar model was also proposed. This model is best understood by looking at the DNA helicase PcrA, which contains two RecA-like domains just like SecA, and two additional DNA binding domains 1B and 2B that constantly switch between high affinity and low affinity DNA binding states in response to ATP binding and hydrolysis. This allows one domain, for example 1B, to be in tight association with DNA, while the 2B domain binds more loosely. Upon ATP binding and hydrolysis, the loosely associated 2B domain will disassociate and undergoes a power stroke motion to bind to the position ahead on the DNA. In this new position, 2B domain will bind tightly with DNA, while the previous tight binding 1B domain will weaken its binding and dissociate. Another round of ATP binding and hydrolysis allows the 1B domain to move ahead of
Figure 15. Models of SecA-mediated protein translocation. A) Power-stroke model; B) Inchworm model; C) Brownian-ratchet model; D) Peristalsis model; E) Subunit-recruitment model; F) Reciprocating-piston model [86, 192, 193]

domain 2B [194, 195]. Despite two DEAD motor domains on SecA with a high level of sequence homology with PcrA, SecA only contains one mature region, substrate-binding site. Thus, this model would best fit for SecA functioning as dimer. Although
there are a number of SecA dimer structures available now, no one structure contains substrate in it. In addition, there is great controversy in the field as to whether SecA works as monomer or dimer, and if a dimer, which of the many dimer states is the functional form. Thus, the validity of this model is still far from proven.

9.3 Brownian-ratchet Model

Although the power-stroke model explains the step-wise translocation of protein substrate, it is hard to imagine how the 2HF or any other domain of SecA can interact with virtually any polypeptide sequence to drive translocation. Such sequence-insensitive nature of protein translocation is what led scientists to propose the Brownian-Ratchet model (Figure 13C). In this model, ATP binding in SecA causes a conformational change in SecY such that it allows polypeptide substrates to diffuse through the translocon [86]. Various models have been proposed in regards to how net forward movement is achieved. The originally proposed Brownian-Ratchet model suggests that segments of protein that have already been translocated are subject to folding and various modifications such as glycosylation and disulfide bond formation [196]. These modifications prohibit polypeptide backsliding, thus resulted in net forward movement.

Recent support for this model came from a molecular dynamics simulation which shows that binding of SecA and ATP cause widening of the SecYEG channel at the lateral gate, and in return, opening and closure of the lateral gate regulates nucleotide exchange in SecA. In this model, the ADP-bound, translocon-narrowed state allows
polypeptide to diffuse in both directions. Binding of ATP by SecA is triggered by a channel blockage on the cytoplasmic side, possibility through a bulky amino acid side chain. This leads to further opening of the channel to facilitate diffusion of such amino acids toward the trans-side. Once the blockage is cleared, ATP is hydrolysed, and the channel returns to its close narrowed state to allow passive diffusion. Backward sliding is prevented because the energy transducing step, which resolves the channel blockage, occurs only in the cytoplasm [197]. Although the Brownian Ratchet model explains why protein translocation can accommodate a large variety of protein substrates, backward sliding potentially slows down translocation. Furthermore, translocation is initiated by active, targeted insertion of a signal peptide into the lateral gate. In this regard, passive diffusion, which lacks active motion, would fail to initiate translocation.

9.4 Peristalsis Model

The Peristalsis model is similar to the Brownian-Ratchet model in that the proposed role of SecA is to regulate the opening and closure of the SecYEG translocaton so that substrate can diffuse through it by directed Brownian motion. The difference is that in this model, SecA dimer binds to SecYEG dimer, particularly the front-to-front dimer (Figure 13D). There hasn’t been too much evidence showing SecA dimer binding to SecYEG dimer. Frank Duong did show that the SecA dimer is capable of binding to the covalently linked, in vivo functional SecYE dimer [198]. However, the implication of this finding for the in vivo and in vitro stoichiometry of SecA and SecYEG binding
during translocation needs to be confirmed since the study was based on in vitro incubation of various components of protein translocation without substrates. The peristalsis model proposes that binding of SecA dimer to the front-to-front SecYEG dimer creates a large vestibule between the two protein complexes. Binding of ATP causes conformational change in the SecA dimer such that a new dimeric interface is formed. This leads to closure of the central opening in the original SecA dimer, thereby trapping the bound substrate, with concomitant opening of the SecYEG channel. This allows substrates to diffuse through the translocon by Brownian motion. The net forward movement is achieved by the closure of the central opening in the starting SecA dimer form. Hydrolysis of ATP reverses the conformational changes in SecA and SecY dimer [192, 199, 200]. This model relies on the functional importance of the SecA central opening and the front-to-front SecYEG dimer, both of which are controversial in the current state of the field. A central opening within the SecA dimer has been observed in two antiparallel SecA dimer crystal structures [97, 99], but its importance, especially in regulation of the opening and closure of the front-to-front SecYEG dimer needs to be further investigated.

9.5 Subunit-recruitment Model

This model was formulated based on the observation that the SecA monomer and dimer equilibrium is subject to change in response to binding of various translocation ligands. For example it was found that phospholipids induced monomerization of SecA, while signal peptide redimerized it. This led to the proposal that translocation
occurs through the recruitment of a second SecA protomer (Figure 13E) [111]. In this model the newly recruited protomer binds to the upstream sequence of the translocating polypeptide. Dimerization of SecA is what drives the substrate through the translocon [86]. The nature of this model suggests that substrate induces SecA dimerization. However, the signal peptide also led to monomerization of SecA in two other studies [110, 122]. One caveat is that this observation was made in the absence of SecYEG. It was found that in the presence of SecYEG and protein substrate, SecA didn’t monomerize [118].

9.6 Reciprocating-piston Model

Driessen’s team has integrated previous conflicting studies and put forward what’s called the Reciprocating-Piston model (Figure 13F) [86]. In this model, during the initiation step, SecA binds to SecYEG as an ADP-bound dimer (Step 1). As mentioned before, the cytoplasmic concentration of SecA is around 5.7-8.2µM [8] while its self association-dissociation constant is about 0.5-1µM [96]; therefore, it makes sense that SecA binds as dimer. In this SecA2-SecY complex, only one copy of SecA is bound to SecY (Step 2). Next, SecB chaperone brings the unfolded, translocation-competent substrate to SecA dimer through interaction with the C-terminal region of both SecA protomers (Step 3). When ADP is exchanged with ATP, a conformational change in SecA allows insertion of the signal peptide into the lateral gate of SecY as a hairpin loop. This step is possibly coupled to the membrane insertion of the 30kDa region of SecA [28], which may encompasses the 2HF region. Concomitantly, SecB is
dissociated from SecA, which completes the initiation process (Step 4). In the translocation step, ATP hydrolysis causes one protomer of SecA to dissociate from the substrate, leaving the second copy bound to both substrate and SecYEG, preventing backsliding (Step 5). This is coupled to SecA membrane deinsertion, which is facilitated by PMF (Step 6). PMF was also shown to act as an energy source to keep substrate moving forward in the absence of ATP [26]. The dissociated SecA can rebind or exchange with cytosolic SecA (Step 7), which causes an ATP-independent translocation of 2-2.5kDa of protein substrate (Step 8). This step resembles the subunit recruitment model in that the binding of the second copy of SecA upstream in the polypeptide sequence, in conjunction with dimerization with the SecYEG-bound SecA, compresses the substrate between the two SecA protomers, thereby forcing it through the translocon. In this step, before the binding of the next ATP molecule, Brownian motion can drive the movement of substrate forward. The binding of a new ATP molecule initiates another round of power stroke, accompanied by the insertion of SecA and translocation of another 2-2.5kDa preprotein segment (Step 9). Step 5-9 are repeated until complete translocation of the preprotein occurs.

10 21st Amino Acid Technology

With rare exception, the genetic codes of all known organisms specify the same 20 amino acids that form the building blocks for proteins synthesis [201]. 21st amino acid technology enables researchers to site-specifically incorporate 21st unnatural amino acids (UAA) into protein in vivo using a pair of specially evolved tRNA/tRNA
synthetase (Figure 16). Unlike common amino acids, UAAs are customized to contain unique moieties that allow functional conjugation to a variety of reagents suitable for studies of protein structure and function. In the case of our study, we incorporate an UAA called para-benzoylphenylalanine (pBpA) into SecY which enabled us to study SecY-SecA and SecY-SecY interactions through in vivo photo-crosslinking.

10.1 Incorporation of UAA

Site-specific incorporation of UAA relies on a pair of orthogonal tRNA/tRNA synthetase with minimum, if any, cross reactivity with the host (in our case E.coli) tRNA/tRNA synthetases. This means that the introduced tRNA cannot be recognized by the endogenous tRNA synthetase of E.coli, but it still can function efficiently in translation (orthogonal tRNA). This tRNA must also be able to incorporate UAA in response to a unique translation codon (usually an amber stop codon). Furthermore, the introduced tRNA synthetase must only be able to aminoacylate the orthogonal tRNA, but not any other endogenous tRNAs, with added UAA, but not any of the other 20 canonical amino acids in E.coli (orthogonal aminoacyl-tRNA synthetase) (Figure 16). The first such tRNA/tRNA synthetase pair that fits these criteria was derived from a tyrosyl pair (tRNA\textsuperscript{Tyr}-TyrRS) of the Archaea Methanococcus jannaschi (M.j), through rounds of positive and negative genetic selections [201]. The M.j pair was used because the recognition element of M.j tRNA\textsuperscript{Tyr} is different from that of E.coli, giving it the basic orthogonal character. Also the M.j TyrRS has a minimalist anti-codon binding domain, which makes it possible to alter its tRNA\textsuperscript{Tyr} anti-codon to CUA
with minimum loss of affinity to TyrRS. An amber stop codon was selected for incorporation of UAA because it is the least commonly found stop codon in *E.coli*. In fact, only 7% of *E.coli* genes end with an amber stop codon, and it rarely terminates the essential genes. Many *E.coli* strains also contain naturally occurring amber suppressor tRNA that have no adverse effects on cell physiology [202]

**Figure 16. Incorporation of an unnatural amino acid.** A) Canonical translation that is terminated by the recruitment of release factor 1 into a stop codon. B) Amber stop codon is being used to encode an UAA by a pair of orthogonal tRNA/tRNA synthetase, which doesn’t cross react with the endogenous tRNA/tRNA synthetases. C) The resulting protein contains a UAA incorporated at a designated spot. UAA can be selectively conjugated to different chemical agents for studies of protein structure and function [202].

In our experiment, an amber stop codon was engineered into various locations in SecY (Figure 17A). UAA pBpA (Figure 17B), which can be readily transported by *E.coli*, is added to the growing media. Once inside the cytoplasm, it is recognized by the orthogonal TyrRS only, and is amminoacylated onto the orthogonal tRNA\textsuperscript{Tyr}. tRNA\textsuperscript{Tyr} with its CUA anti-codon loop brings pBpA to the amber stop codon and adds it to the growing polypeptide chain. The end result is the production of SecY with pBpA incorporated at a very specific spot.
10.2 Activation of pBpA

pBpA can be activated at its carboxyl group when treated with ultra-violate light at a wavelength of ~365nm and forms a reactive di-radical, which can react with near-by C-H groups within 3 Å to form a covalent bond (Figure 17C). This is advantageous for in vivo study of protein-protein interaction because UV light at this range does not damage other amino acids since proteins generally absorb at a shorter wavelength around 240nm. Although the pBpA reaction is promiscuous in nature, it prefers to react with the electron-rich tertiary centers such as C\(_\gamma\)-H of leucine and C\(_\beta\)-H of valine, at a 108.9° degree angle of attack [203]. We used pBpA chemistry to study SecY-SecY interaction by treating *E.coli* cells with 365nm UV light. The result of such treatment can be seen in the form of a UV dependent shift in protein electrophoretic mobility when analyzed in an SDS-PAGE protein gel (Figure 17D). This method was first used to study SecA-SecY interaction by Mori *et al.* This study revealed multiple SecA interacting sites on the cytoplasmic loops of SecY and serves as a proof-of-principle for the feasibility of this approach for our study [149].
Figure 17. Site-specific incorporation of \( \text{pBpA} \). A) An amber mutation was engineered into SecY for site-specific incorporation of pBpA using a pair of orthogonal tRNA/tRNA synthetase. B) Chemical structure of pBpA. C) \textit{In vivo} treatment of \textit{E. coli} cell with UV light around 365nm induces di-radical formation in the carboxyl group of pBpA, which can react with a near-by CH group to form a covalent bond. D) UV-crosslinked SecY dimer can be visualized in SDS-PAGE gel showing a UV-dependent shift in its electrophoretic mobility.

11 About this work

SecY has been shown to exist as a monomer, dimer and even higher oligomers in detergent. Which of these is the functional form inside the cell has been a great controversy. A number of structural studies support monomer function. Numerous other biochemical, biophysical and genetic approaches such as \textit{in vitro} protein translocation, crosslinking, FRET and \textit{in vivo} complementation assays have resulted in support of either monomer or dimer function or a mix of both as the functional unit. However, the validity of these results needs to be challenged and sorted out because
the SecYEG oligomeric state can be altered artificially by various experimental conditions such as high detergent and SecYEG expression level. Besides this, our study was also motivated by the study of Park et al., who showed that SecYEG dimer monomerized when it was arrested in a translocation-intermediate state by an OmpA-GFP substrate, as accessed by the loss of crosslinking of a dimer-specific cysteine reporter pair. However, this result could simply be due to artifacts of conformational change induced by the binding of SecA and OmpA-GFP that rendered two cysteines incapable of forming the disulfide bond. Thus the second chapter of my thesis describes the work that we undertook to more thoroughly investigate the oligomeric state of SecYEG under different conditions that mimic various steps of protein translocation using mainly in vivo photo-crosslinking-based approaches. Through this extensive analysis, it was found that SecY dimer participated in active protein translocation.

The above work provided strong evidence that SecY dimer is involved in a particular step of the translocation cycle. However, we didn’t know whether dimer remains throughout the entire translocation reaction or not, or if there is any possibility of monomer function alone or as a part of a monomer-dimer translocation cycle. In Chapter III of my thesis, we devised a series of experiments to get a better understanding of monomer function. We utilized E.coli strains that only expressed WT or putative monomeric SecYEG and investigated the secretion rate of three periplasmic proteins, maltose-binding protein (MBP), galactose-binding protein (GBP), and ribose-binding protein (RBP), and one outer membrane protein called
OmpA. Careful analysis of our results suggested no major difference in translocation function in the WT and putative monomer-producing secYE mutant strains. In order to more rigorously assess the monomeric state of the secYE mutant, we undertook additional studies, namely in vivo photo-crosslinking as well as SecA foot-printing. Our studies revealed that the translocons in the latter strain form dimers both during active translocation and in the resting state similar to the WT strain. Thus, more effort is needed to create a true SecYEG monomer-producing strain.
Chapter II

Determination of the Oligomeric State of SecYEG Protein Secretion Channel Complex using \textit{in vivo} Photo- and Disulfide-crosslinking

1. ABSTRACT

SecYEG protein of bacteria or Sec61αβϒ of eukaryotes is a universally conserved heterotrimeric protein channel complex that accommodates the partitioning of membrane proteins into the lipid bilayer as well as the secretion of proteins to the trans side of the plasma or endoplasmic reticular membrane, respectively. SecYEG function is facilitated by cytosolic partners, mainly a nascent chain-ribosome complex or the SecA ATPase motor protein. Extensive efforts utilizing both biochemical and biophysical approaches have been made to determine whether SecYEG functions as a monomer or a dimer, but such approaches have often generated conflicting results. Here we have employed site-specific \textit{in vivo} photo-crosslinking or cysteine crosslinking, along with co-immunoprecipitation or SecA foot-printing techniques to readdress this issue. Our findings show that the SecY dimer to monomer ratio is relatively constant regardless of whether translocons are actively engaged with protein substrate or not. Under the former conditions the SecY dimer can be captured associated with a translocon-jammed substrate, indicative of SecY dimer function. Furthermore, SecA ATPase can be crosslinked to two copies of SecY when the complex contains a translocation intermediate. Collectively, our results suggest that SecYEG dimers are functional units of the translocon.
2. INTRODUCTION

Up to one third of bacterial proteomes consist of proteins that must be inserted into or secreted across the plasma membrane. Such proteins are generally synthesized with one or more signal-anchor sequences that determine their ultimate fate and topology. Proteins that contain more hydrophobic signal-anchor sequences (typically membrane proteins) are directed into the co-translational translocation pathway, whereby their ribosome-nascent chain complexes are guided to the SecYEG channel complex by interaction of the signal recognition particle both with the emerging signal-anchor sequence and its membrane receptor located adjacent to the channel complex. After ribosome-translocon docking, it is thought that the elongation phase of translation provides energy for integration of proteins into the membrane along with the thermodynamics of protein folding in this environment [204]. On the other hand, proteins that contain moderately hydrophobic signal-anchor sequences (typically secretory preproteins) are directed to the post-translational translocation pathway [7]. In this case, incomplete or fully translated, but partially folded preproteins are bound by cytosolic chaperones (SecB in the case of Gram-negative bacteria like E. coli), keeping them competent for transport. SecB then targets the pre-protein to SecYEG-bound SecA protein, which recognizes both the signal peptide and SecB, and promotes the SecB-preprotein release step. In addition to facilitating preprotein targeting to the translocon, SecA is also a motor protein that utilizes its ATP-promoted conformational cycles to drive protein transport by a still poorly defined mechanism [81, 205]. A popular model based on the SecA-SecYEG x-ray structure suggests that the SecA two-
helix finger serves as a molecular ratchet to processively bind and insert substrate preproteins into the mouth of the SecYEG channel utilizing consecutive rounds of ATP hydrolysis [91, 190]. The nature and depth of SecA insertion into the translocon channel remains unclear since its initial discovery over two decades ago [205, 206].

The translocon channel complex is a universally conserved, heterotrimeric protein, Sec61αβϒ in eukaryotes or SecYEG in bacteria. It forms an hourglass-shaped structure composed of 10 transmembrane (TM) helices of Sec61α/SecY hinged by Sec61Y/SecE on one side, allowing for a clamp-like action that opens and closes the lateral gate region on the side opposite the hinge. Signal-anchor sequences insert into the lateral gate region of the channel [207], allowing for the subsequent release of membrane proteins into the lipid bilayer. The channel complex also contains a constriction pore at its center as well as a small helical plug domain that relocates to the channel exterior in order to open a path for the translocation of secretory proteins across the membrane [136, 208].

Early biochemical and structural work pointed to the existence of SecYEG dimers and their potential physiological importance in protein transport. For example, the purified SecYEG complex in detergent or reconstituted into proteoliposomes formed dimers and tetramers in a concentration dependent manner that was enhanced by the presence of translocation partners such as SecA, preprotein, or cardiolipin [44, 173, 180, 209, 210]. Disulfide crosslinking studies of inverted membrane vesicles demonstrated the presence of a SecYEG dimer complex by the formation of SecY-SecY or SecE-SecE homo-dimers [132, 211-213]. Moreover, when purified SecYEG
was reconstituted into nanodisks as either monomers or dimers, only the latter unit was capable of activating SecA translocation ATPase activity: a requirement for protein transport in vitro [183]. Furthermore, a genetically fused SecY dimer was found to display intragenic complementation both in vivo and in vitro: a result mostly simply understood by assuming a functional dimer state at some point in the protein translocation cycle [183, 214].

Indeed, considerable controversy exists as to whether the translocon channel complex functions as a monomer or a dimer. Recent crystal structures of the SecY complex with or without its associated SecA ATPase protein clearly reveal a protein-conducting channel within a single SecY subunit [91, 135, 136, 215]. While structures of the ribosome-nascent chain-SecYEG complex obtained by cryo-electron microscopy also support this view, they differ in the presence of a second, non-translocating SecYEG protomer within the complex [38, 175, 216, 217]. Likewise, single molecule studies have also lead to conflicting results on the active state of the translocon. For example, a fluorescence-based study using differentially labeled SecYEG proteins reconstituted into giant unilamellar vesicles failed to detect the presence of SecYEG dimers during SecA or protein substrate-bound conditions [177]. However another study, where SecYEG monomers and artificially stabilized dimers were compared, found that while monomers sufficed for SecA and preprotein binding, dimers were required for active transport [218]. An in vivo study was performed employing disulfide crosslinking at the SecY dimer interface to assess the oligomeric status of the translocon under translocation conditions that simulate ongoing protein
transport (by utilizing a substrate that jammed the translocon channel). The authors found that the SecY dimer level decreased precipitously when channels were jammed with appropriate substrate proteins, suggesting that protein transport occurs from a SecYEG monomer [179]. One limitation of this latter study, however, is that substrate jamming may have led to a conformational change of SecY that prevented disulfide crosslinking. These authors also functionally-disrupted SecY dimer assembly by mutagenesis at the dimer interface, further supporting the functional monomer view. Adding to the ongoing controversy, a recent study utilizing a highly efficient SecYEG reconstitution system found that preproteins apparently differed in their translocon stoichiometry requirement, where a monomer or dimer was required depending on the preprotein species under study [184].

Studies assessing the functional size of the translocon channel are difficult to reconcile with a functional monomer subunit that contains the observed channel dimensions. For example, a study employing pro-OmpA conjugated to rigid tetraarylmethane derivatives estimated a channel diameter of greater than 22-24 Å [137]. Similarly, a fluorescence-based study of the mammalian endoplasmic reticular channel dimensions utilizing quenchers of defined sizes placed the channel diameter within a range of 40-60 Å [219]. Both of these findings exceed the maximum expandable size of a single channel, and they would be more consistent with a dimer-based, composite channel.

Here we have utilized both in vivo photo-crosslinking and disulfide crosslinking approaches in order to investigate the functional oligomeric state of the
SecY channel complex, given the conflicting claims present in the literature. The former method was developed by Schultz and co-workers (for review, see Ref.[201]). It uses a specially engineered amber suppressor tRNA/tRNA synthetase pair that incorporates the photo-activatable phenylalanine derivative pBpA into the target protein of interest. Excitation by light at ~365 nm generates a triplet diradical that reacts with nearly C-H acceptors (within 5Å) to form a covalent crosslink. This technique offers a physiological method to study protein-protein interaction. The results presented below indicate that SecY dimers comprise at least part of the active protein translocational unit.

3. EXPERIMENTAL PROCEDURES

Chemicals, Media, Strains and Plasmids—LB (Miller) broth, IPTG and arabinose were purchased from Fisher Scientific. pBpA and maltose were purchased from Bachem and Difco, respectively. DDM was purchased from Anatrace, while anti-GFP beads were from ChromoTek. Quik-Change™ and Wizard Plus SV Miniprep DNA Purification System kits were obtained from Stratagene and Promega, respectively. Restriction enzymes were obtained from New England Biolabs, while the WesternBright™ Sirius enhanced chemi-luminescence kit was obtained from Advansta. Protease Inhibitor Cocktail was obtained from Sigma Aldrich. Most other common laboratory chemicals were obtained from the latter supplier or Fisher Scientific and were laboratory grade or better. Mouse anti-c-Myc monoclonal antibody and HRP-conjugated goat anti-mouse antibody were obtained from Genescript, while chicken anti-GFP antibody and HRP-conjugated goat anti-rabbit antibody were
obtained from Abcam. HRP-conjugated goat anti-chicken antibody was obtained from Jackson ImmunoResearch Laboratories. Peptide affinity-purified SecY antisera was prepared by hyper-immunizing rabbits to a peptide identical to the carboxyl terminus of SecY (CYESALKKANLKGYGR) conjugated to keyhole limpet hemocyanin by maleimide chemistry utilizing Tana Laboratories, LC (Houston, TX) for peptide synthesis, protein carrier conjugation, immunization, and peptide affinity purification of the antisera. *E. coli* BLR(λDE3) [F^{ompT} hsdS (rB^- mB^-) gal dcm Δ(srl-recA)306::Tn10 (Tet^R)] was obtained from Stratagene, while its isogenic BL26 (λDE3) ΔlacU169 recA^+ derivative was obtained from Bill Studier [220]. BL26.1(λDE3) is a recA::KAN version of this latter strain made by P1 transduction. MM18.7 is a recA derivative of MM18 containing the malE-lacZ72-47(Hyb) fusion that has been described previously [221]. MC4100.2(λDE3) is a recA1 srl::Tn10 derivative of MC4100 [222] that was lysogenized with λDE3, which has been previously described [220]. The pSup-pBpARS-6TRN plasmid encoding the *Methanococcus jannaschii* amber suppressor tRNA-tRNA synthetase pair that efficiently incorporates pBpA in place of an amber codon has been described previously [201]. A series of pCDFT7secYEG plasmids with secYEG under control of the T7 promoter with amino-terminal c-Myc tags on secY or secE or a carboxyl-terminal c-Myc tag on secG have been described [182]. The pBAD-OmpA-GFP plasmid carrying the arabinose-inducible, OmpA-GFP jamming chimera was obtained from Tom Rapoport [179]. pBAD-secA-OmpA-GFP was created by fusing secA onto the N-terminus of *ompA-GFP*. For this purpose, pT7-secA-his [115] was modified
using QuikChange™ mutagenesis by making synonymous codon changes eliminating two internal NcoI sites within secA, and a Ndel site at the secA start codon was changed into a NcoI site. This allowed the entire secA gene to be isolated as an NcoI-XhoI restriction fragment. pBAD-OmpA-GFP was modified using QuikChange™ mutagenesis to insert a XhoI site after the NcoI site that includes the ompA start codon. This allowed insertion of the secA fragment into the NcoI-XhoI cleaved pBAD-OmpA-GFP vector fragment to form the SecA-OmpA-GFP trimera, which was verified by DNA sequence analysis utilizing the University of Pennsylvania DNA Sequencing Facility. All codon substitutions within plasmid-borne secY, secE, secA, or ompA genes were made by QuikChange™ mutagenesis and also verified by DNA sequence analysis.

In Vivo Photo-crosslinking—A freshly struck out single colony of BLR(λDE3) containing pSup-BpARS-6TRN and pCDFT7secYmycEG or pCDFT7secYEmycG plasmids with the given secY or secE amber mutation, respectively, was inoculated into LB media supplemented with appropriate antibiotics (25 μg/ml chloramphenicol, 50 μg/ml streptomycin, and 100 μg/ml ampicillin as needed) and grown overnight at 37°C with shaking at 250 rpm. The overnight culture was diluted 1:50 into LB media supplemented with appropriate antibiotics and 1 mM pBpA and grown until OD_600 reached 0.3, when SecYEG was induced with IPTG at final concentration of 1 mM for 2 h. All subsequent steps were done at 4°C or on ice. To adjust for somewhat different cell densities, 4.8 OD_600 cell equivalent of each culture was harvested by sedimentation in a microfuge at 14,000 rpm for 5 min, washed with 5 ml PBS (10 mM sodium
phosphate, pH 7.5, 140 mM NaCl), and resuspended in 4 ml PBS buffer. 2 ml samples were UV irradiated on ice at 365 nm for 20 min using a Rayonet 2000 UV crosslinker (Southern New England Ultraviolet Company), while 2 ml non-irradiated samples served as negative controls. Each sample was sedimented, resuspended in 1 ml breakage buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.25 mM PMSF, 1 mM DTT, 100 µg/ml RNase, 100 µg/ml DNase, 60 µg/ml lysozyme, 1X Protease inhibitor cocktail), and cells were placed in a polycarbonate tube and disrupted with 30 sec bursts of a cup horn sonicator (Heat Systems) until near clarity. Unbroken cells were removed by sedimentation in a microfuge at 14,000 rpm for 5 min. Supernatant was isolated and sedimented in a Sorvall S120 AT2 rotor at 82,000 rpm for 30 min at 4°C to isolate the membrane fraction. Each membrane pellet was solubilized in 60 µl ABB buffer (5% SDS, 10 mM Tris-Cl, pH 8, 1 mM EDTA) with constant stirring for 1 h at 37°C, when 20 µl 4X sample buffer (8% SDS, 500 mM Tris-HCl, pH 6.8, 20% 2-mercaptoethanol, 60% glycerol, 0.02% bromophenol blue) was added, and stirring continued for an additional 10 min. 15 µl samples were loaded onto a 15% SDS-PAGE gel, which was run at 100 V at 4°C until the dye front reached the bottom. Western transfers were performed at 100 V for 1 h, and nitrocellulose membranes were blocked overnight with 10 ml TBS buffer (20 mM Tris-HCl, pH 7.5, 140 mM NaCl, 0.25% Tween 20) supplemented with 10% non-fat dry milk. The membrane was probed with mouse anti-c-Myc monoclonal antibody, followed by HRP-conjugated goat anti-mouse antibody, both at a 1:5,000 dilution. SecYY was probed separately with rabbit anti-c-term-SecY antibody and HRP-conjugated goat anti-rabbit antibody at a 1:5,000
dilution. Proteins were visualized using a WesternBright™ Sirius kit as described by the manufacturer.

**Translocon Jamming**—An overnight culture of MC4100.2(λDE3) containing pSup-pBpARS-6TRN, pCDFT7secYmycEG with a given secY amber mutation, and pBAD-OmpA-GFP plasmids was grown in LB containing appropriate antibiotics, 30 μM IPTG, and 0.2% maltose as described above for In Vivo Photo-crosslinking. It was then sub-cultured 1:100 into LB supplemented similarly except for addition of 1 mM pBpA and omission of chloramphenicol. The culture was grown until OD$_{600}$ reached 0.15, when arabinose was added to a final concentration of 0.2%. For each time point analyzed, 6 OD$_{600}$ cell equivalent of culture was harvested and prepared for UV irradiation and subsequent analysis as described above. For detection of the OmpA-GFP chimera, 5 μl of the final non-irradiated sample was analyzed by SDS-PAGE, and western blots were probed with chicken anti-GFP antibody and HRP-conjugated goat anti-chicken antibody each at a 1:5,000 dilution. The extent of translocon jamming was assessed by MBP fractionation. For each time point analyzed, 0.6 OD$_{600}$ cell equivalent was harvested by sedimentation in the microfuge at 14,000 rpm for 5 min, and resuspended in 200 μl 20% sucrose-0.03 M Tris-HCl, pH 8. Half the sample was saved as the total cell control, while the other half was spheroplasted by addition of 10 μl of 1 mg/ml lysozyme in 0.1M EDTA, pH 8 for 15 min. Treated cells were sedimented in the microfuge at 5,500 rpm for 5 min, and the supernatant (i.e. periplasmic fraction) was removed, while the cell pellet (cytoplasm/membrane fraction) was resuspended in 100 μl of buffer A (10 mM Tris-HCl, pH 7.5, 150 mM...
NaCl, 5 mM MgSO$_4$). All samples were boiled for 10 min in the presence of sample buffer, and 5 µl aliquots were analyzed by SDS-PAGE and western blotting with rabbit anti-MBP antibody and HRP-conjugated goat anti-rabbit antibody, both at 1:5,000 dilution. For quantification of SecY and MBP, ImageJ was used. Average pixel count of SecY monomer to dimer ratio (D/M) or cytoplasm/membrane MBP to total MBP ratio (C/T) was used and plotted with standard error measurement.

**In Vivo Cysteine Crosslinking**—An overnight culture of MC4100.2(λDE3) containing pCDFT7secYmycEG with a given secY Cys mutation and pBAD-OmpA-GFP plasmids was grown in LB containing appropriate antibiotics, 10 µM or 30 µM IPTG, and 0.2% maltose, sub-cultured in the absence of pBpA, and arabinose induced for 45 min as described above in Translocon Jamming. 3 OD$_{600}$ cell equivalent was harvested, resuspended in 2 ml protoplast buffer (100 mM sodium phosphate, pH 7.5, 5 mM EDTA, 10 mM phenanthroline), and treated with copper phenanthroline (180 mM phenanthroline, 60 mM CuSO$_4$, 50 mM NaH$_2$PO$_4$) to a final concentration of 300 µM [223] for 10 min at 30°C as indicated. The reaction was quenched by addition of n-ethylmaleimide and EDTA to a final concentrations of 1 mg/ml and 5 µM, respectively, for 10 min. For treatment with reducing agent, DTT was added to a final concentration of 60 mM for 10 min. Membrane was isolated and subsequently analyzed as described above for In Vivo Photo-crosslinking except that the final sample was solubilized in 40 µl ABB buffer and an equal volume of 2X non-reducing sample buffer (40 mM Tris–HCl, pH 7.8, 16 mM NaH$_2$PO$_4$, 2% (w/v) SDS, 50 µg/ml
bromophenol blue, 20 mM EDTA, 2 mg/ml N-Ethylmaleimide, and 0.10 g/ml sucrose).

*OmpA-GFP-SecY Co-immunoprecipitation*—MC4100.2(λDE3) containing the pCDFT7secYmycEG with the secY Q212C mutation and pBAD-OmpA-GFP plasmids was utilized. Cell growth, cysteine crosslinking, and membrane isolation were as described for *In Vivo Cysteine Crosslinking*. Each membrane pellet was solubilized in 60 μl TSGM buffer (20 mM Tris-HCl, pH 8, 130 mM NaCl, 10% glycerol and 2 mM MgCl₂) supplemented with 2% (wt/vol) DDM for 3 h at 4°C with constant stirring. Insoluble material was removed by sedimentation at 82,000 rpm for 30 min. Anti-GFP beads were pre-equilibrated by washing three times consecutively with 700 μl dilution buffer (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.5 mM EDTA) by successive sedimentation steps in a microfuge at 5,500 rpm for 2 min. Solubilized membrane was diluted 1:20 into TSGM buffer, when 30 μl anti-GFP beads were added, followed by incubation at 4°C for 2 h. Beads were sedimented and washed three times consecutively with 1 ml TSGM buffer supplemented with 0.6 mM DDM. For elution, 40 μl elution buffer (0.2 M glycine, pH 2.5) was added to the washed beads, which were agitated for 2 min and re-sedimented. The eluate was removed and quickly neutralized with 6 μl of 1 M Tris base. The purification was monitored by taking samples before, during, and after the purification, which were analyzed by SDS-PAGE and western blotting with c-Myc or GFP antibodies.

*Translocon Clearing*—An overnight culture of BL26.1(λDE3) containing pSup-pBpARS-6TRN, pCDFT7secYmycEG with the secY K20 amber mutation, and
pBAD-lacZ plasmids was grown as described above for In Vivo Photo-crosslinking. The culture was grown until OD$_{600}$ reached 0.20, when kasugamycin was added to a final concentration of 2 mg/ml. For each time point analyzed, 3 OD$_{600}$ cell equivalent of culture was harvested and prepared for UV irradiation and subsequent analysis of protein as described above for In Vivo Photo-crosslinking.

**SecA Foot-printing**—MC4100.2(λDE3) containing pSup-pBPARS-6TRN, pCDFT7secYmycEG with the secY Cys68 mutation, and pBAD-secA-OmpA-GFP with the secA 59 Amber and ompA Cys21 mutations was grown and arabinose induced as described above for Translocation Jamming. 20 ml of culture was harvested 45 min after arabinose addition and subjected to photo-crosslinking or cysteine crosslinking as described in In Vivo Photo-crosslinking or In Vivo Cysteine Crosslinking. For double crosslinking, cysteine crosslinking was performed first, cells were sedimented in the microfuge at 14,000 rpm for 5 min, decanted and resuspended in 2 ml PBS prior to photo-crosslinking. After photo-crosslinking, final sample treatments were performed as indicated for In Vivo Cysteine Crosslinking. Western blots were probed with either anti-GFP antibody or anti-c-Myc antibody as described above.

4. **RESULTS**

In Vivo Photo-crosslinking Methodology—SecYEG protein is known to exist as monomers and dimers. Two types of SecYEG dimers have been proposed to date based largely on previous disulfide crosslinking studies: a back-to-back dimer largely stabilized through interactions between SecE, and a front-to-front dimer stabilized by
interactions around the lateral gate region of SecY [132, 211, 212, 224]. Less evidence exists in support of the front-to-front dimer orientation [179]. We investigated the formation of both dimers using in vivo photo-crosslinking by engineering the incorporation of pBpA at various locations at the dimer interface (Figure 2.1). SecYEG dimer structures were obtained by manually docking two SecYEG monomers together based on the previous cysteine crosslinking data [132, 211]

FIGURE 2.1. Placement of pBpA at the dimer interface for in vivo photocrosslinking. A-C: Two protomers of Thermotoga maritima SecYEG were manually docked into a front-to-front conformation based on cysteine crosslinking studies [132, 211]. SecY, SecE and SecG are colored green, red, and cyan, respectively. TM2b and 7 of the lateral gate are colored orange. Sites with amber mutations for incorporation of pBpA are shown as magenta spheres. Views from the A, cytoplasm, B, along the plane of membrane, and C, periplasm are shown. D-F: Similar to A-C except protomers were docked into a back-to-back conformation. Magenta and blue spheres indicate the amber mutations on SecY and SecE subunits, respectively. Two negative controls, shown in purple spheres, are amber mutations engineered into regions distant to either dimer interface.

We performed in vivo photo-crosslinking on secY amber mutant K20 in order to demonstrate the ability of this technique to provide credible information on SecY dimer status. Several criteria were used to assess the success of this approach. First,
we noted that the production of full-length SecY was dependent on inclusion of pBpA in the media (Figure 2.2, compare lanes 2 and 7). The low residual level of SecY production in the absence of pBpA was probably due to a minor level of mis-incorporation of other amino acids by the engineered tRNA synthetase/amber suppressor tRNA. Second, we noted that SecY dimer formation was largely dependent on UV exposure (compare lanes 1 and 2). Third, the presence of the amber codon was necessary for efficient UV-dependent crosslinking, since the low residual level of SecY dimer observed in the wild-type strain was not particularly UV dependent (compare lanes 1 and 2 with lanes 10 and 11). Fourth, the identity of SecY monomers and dimers was confirmed by their heat-dependent aggregation (lane 3 and 12) [225], the failure of an isogenic strain lacking the c-Myc tag on SecY to react with c-Myc antibody in the western blot (lanes 15 and 16), co-migration of a genetically-fused SecY dimer with our crosslinked species (lanes 1 and 17), and finally the susceptibility of disulfide crosslinked mono-cysteine SecY dimer to DTT treatment to produce monomers (see Disulfide Crosslinking Studies below). We conclude that our in vivo photo-crosslinking methodology can detect bona fide SecY interprotomer interactions.

*SecY Forms Both Front-to-front and Back-to-back Dimers*—After confirming the feasibility of our approach, we proceeded to confirm the existence of front-to-front and back-to-back SecY dimers by in vivo photo-crosslinking of appropriate mutant strains. Both dimer forms were detectable by this approach (Figure 2.3). In addition, certain secY or secE mutants produced two or even three possible SecY or SecE dimers
(e.g. Figure 2.3A G296 and 2.3D G110, respectively) that ran as tightly spaced bands, consistent with the promiscuous crosslinking.

**FIGURE 2.2 Verification of in vivo photo-crosslinking.** Each strain was grown as indicated in the presence or absence of 1 mM pBpA until OD$_{600}$ reached 0.3, when SecYEG was induced or not with IPTG at a final concentration of 1 mM for 2 h. Cells were harvested, resuspended in PBS buffer, and treated with UV irradiation (350-365nm) for 20 min where indicated. Cell membranes were isolated and analyzed by western blotting using c-Myc antibody as described in Experimental Procedures. Samples in lanes 3 and 12 were heated at 100°C for 5 min prior to loading on the gel to induce SecY aggregation. YY, indicates a genetically-fused SecY dimer used as a marker [181]. Anti-SecY peptide antiserum was used to detect this latter species.

chemistry of pBpA and the separation by SDS-PAGE of dimers with differing degrees of asymmetry in the position of their crosslink. Alternatively, one or more of these species could represent SecY crosslinked to another membrane protein that associates peripherally with the translocon. Consistent with previously identified sites within the SecY-SecY interface, we found strong crosslinks that support the existence of the front-to-front dimer when amber codons were positioned at residues S111, N151 and F154 (Figure 2.3A) [179]. An additional eight novel sites supporting this dimer
orientation were also found. We detected an additional major crosslinked species (Figure 2.3A, indicated by a star) migrating between the monomer and dimer forms of SecY, which likely corresponds to SecY interacting with either SecE or SecG since both proteins are part of the heterotrimeric complex. To investigate this possibility, we created c-Myc-tagged versions of either SecE (SecYE_{myc}G) or SecG (SecYEG_{myc}) in a secY Y122 mutant. Photo-crosslinking of the relevant strains demonstrated that this novel species was in fact a SecY-SecG complex (Figure 2.3B). Indeed, when docking SecY in a front-to-front orientation, SecG is the nearest neighbor to some of the engineered amber sites (Figure 2.1A-C). For other secY mutations at this interface, we cannot exclude the possibility of SecY interacting with additional protein partners outside of the heterotrimeric complex, although a previously documented SecY-YidC interaction [226] is unlikely here given the large size predicated for this complex (~95 kDa).
FIGURE 2.3. Detection of distinct SecYEG dimer states using in vivo photo-crosslinking. Each strain carrying the indicated amber mutation was grown in the presence of 1 mM pBpA until OD$_{600}$ reached 0.3, when SecYEG was induced with IPTG at a final concentration of 1 mM for 2 h. Cells were harvested, resuspended in PBS buffer, and treated with UV irradiation for 20 min as indicated. Cell membranes were isolated and analyzed by western blotting using c-Myc antibody as described in Experimental Procedures. Analysis of (A and B) secY amber mutations in the front-to-front (FTF) dimer interface, (C and D) secY and secE mutations, respectively, in the back-to-back (BTB) dimer interface, and (E) secY mutations that lie at neither interface. Certain samples (boil) were heated at 100°C for 5 min to induce SecY aggregation. YY, indicates a genetically-fused SecY dimer used as a marker [181]. The star indicates a SecY-SecG crosslinked complex.

We found two novel sites within SecY that support the back-to-back interface, K20 and K434, which formed strong crosslinks (Figure 2.3C). Amber mutations at
previously identified secE sites also resulted in strong SecE-SecE crosslinks (Figure 2.3D), consistent with the back-to-back dimer [132, 211]. When amber mutations were engineered at two sites within secY that are remote from either dimer interface, no crosslinked SecY dimer was observed (Figure 2.3E), consistent with the specificity of our analysis. Together these results show the utility of our approach for detecting both types of SecYEG dimers in vivo.

*SecY Dimer Levels Are Refractory To Translocon Jamming*—We next set out to determine whether SecY dimers are involved in ongoing protein translocation or not. We first investigated how cellular SecY dimer levels change when they are engaged in active protein transport. Since protein translocation occurs at a very rapid speed (on the order of one preprotein translocated per sec per translocon based on E. coli secretory protein content, doubling time, and translocon number), we made use of an inducible OmpA-GFP chimera that contains a rapidly folding GFP variant fused to an OmpA signal sequence to jam translocon channels in order to mimic a translocation intermediate state [179]. If SecY functions exclusively as a monomer during protein transport, then we would expect to see a decrease in SecY dimer levels over time as more translocon channels become jammed with the OmpA-GFP substrate. Translocon jamming was assessed by the cytoplasmic accumulation of MBP precursor (Figure 2.4). Furthermore, in order to minimize the potential for artificial dimerization due to SecY overexpression, we titrated SecY levels using different concentrations of IPTG in order to find a low expression level that still allowed us to visualize our starting and potentially diminishing SecY dimer signal throughout the experiment. We found that
30 μM IPTG, which modestly over-expressed SecY (Figure 2.5A), meets this criteria, and it was used for all subsequent experiments.

**Figure 2.4. Accumulation of MBP in the cytoplasm as a result of translocon jamming.** Under normal translocation conditions (A) pre-MBP is targeted to the translocon and rapidly secreted into the periplasm as mature MBP. When the translocon is jammed by either OmpA-GFP (B) or MalE-LacZ (C), pre-MBP builds up in the cytoplasm.

We picked mutants that produced the strongest photo-crosslinking signals for the back-to-back (K434 and K20) or front-to-front (S111) dimers and jammed them over a 60 min time course. Samples were collected at different time points to access the change in SecY dimer levels, as well as the extent of translocon jamming. For assessment of the former quantity, we utilized the SecY dimer to monomer ratio (D/M) in order to control for minor variations in SecY amounts or detection. Similarly, for assessment of the latter quantity, we utilized the cytosolic/membrane associated MBP to total MBP ratio (C/T) for the same reason. To perform this latter measurement cells
were spheroplasted with lysozyme/EDTA and sedimented, and the supernatant (containing periplasmic proteins including secreted MBP) and pellet (containing cytosolic/membrane proteins including retained preMBP/MBP) were isolated and quantified by western blotting. We found

**FIGURE 2.5. Analysis of secY K434 mutant during OmpA-GFP-induced translocon jamming by in vivo photo-crosslinking.** (A) Titration of SecY expression with IPTG. The secY K434 mutant was grown in the presence of 1 mM pBpA until OD$_{600}$ reached 0.15, when IPTG was added to the indicated final concentration and growth continued for another 60 min. Cells were harvested and membranes were isolated and analyzed by western blotting using anti-SecY peptide antiserum. (B) View of *T. Maritima* SecYEG docked in the back-to-back conformation with the same color scheme as Figure 2.1. Residue K434 is shown as magenta spheres. (C-F) The secY K434 mutant was grown in the presence of 1 mM pBpA, 30 μM IPTG and 0.2% maltose until OD$_{600}$ reached 0.15, when the OmpA-GFP chimera was induced by adding arabinose to a final concentration of 0.2%. Cells were harvested at the indicated time points post jamming, and exposed to UV irradiation as indicated. A wild-type (WT) strain was used in parallel as a control. (C) Western blot of cell membranes probed with c-Myc antibody. (D) Western blot of cell membranes probed with GFP antibody. (E) Western blot of cells divided into cytoplasm/membrane (C) and periplasmic (P) fractions, compared to total (T) cell input, probed with MBP antibody.
(F) Quantification of SecY dimer to monomer ratio (D/M) or cytoplasm/membrane MBP to total MBP ratio (C/T) during jamming. The average results from three experiments are plotted with standard error measurements. The error bars for the D/M points are too small to be seen on this scale.

that the cellular SecY D/M ratio remained unchanged during the course of jamming for the secY K434 mutant (Figure 2.5C and F). Such jamming was very effective as evident by the rapid buildup of MBP precursor in the cytoplasm (Figure 2.5E), which saturated between 30 and 40 min (Figure 2.5F). This correlated well with the build-up of OmpA-GFP during this time frame (Figure 2.5D). A similar pattern was also observed for other secY mutants, including K20 (data not shown) and S111 (Figure 2.6) that report on the back-to-back or front-to-front dimer conformation, respectively. Interestingly, there was a subtle increase in the SecY D/M ratio for the latter mutant (Figure 2.6E), possibility due to more efficient photo-crosslinking at this position in the translocon-jammed conformation. To make sure that our results were not unique to the OmpA-GFP fusion, we also carried out similar experiments using a strain that carries a chromosomal MalE-LacZ fusion (Figure 2.4C). This chimera contains the first third of preMBP fused to β-galactosidase, and its ability to jam protein transport has been previously demonstrated [227]. We found that the SecY D/M ratio remained unchanged under conditions of MalE-LacZ jamming as well (Figure 2.7).
FIGURE 2.6. Analysis of secY S111 mutant during OmpA-GFP-induced translocon jamming by in vivo photo-crosslinking. (A) View of *T. Maritima* SecYEG docked in the front-to-front conformation with the same color scheme as Figure 2.1. Residue S111 is shown as magenta spheres. (B-E) The secY S111 mutant was grown in the presence of 1 mM pBpA, 30 μM IPTG and 0.2% maltose until OD<sub>600</sub> reached 0.15, when the OmpA-GFP chimera was induced by adding arabinose to a final concentration of 0.2%. Cells were harvested at the indicated time points post jamming, and exposed to UV irradiation as indicated. A wild-type (WT) strain was used in parallel as a control. (B) Western blot of cell membranes probed with c-Myc antibody. (C) Western blot of cell membranes probed with GFP antibody. (D) Western blot of cells divided into cytoplasm/membrane (C) and periplasmic (P) fractions, compared to total (T) cell input, probed with MBP antibody. (E) Quantification of SecY dimer to monomer ratio (D/M) or cytoplasm/membrane MBP to total MBP ratio (C/T) during jamming. The average results from three experiments are plotted with standard error measurements.

Disulfide Crosslinking Studies—In order to determine if the foregoing results were unique to photo-crosslinking, we performed similar experiments utilizing cysteine disulfide crosslinking. For this purpose we used a cysteine at secY Q212 (Figure 2.8A): a known hot spot for forming SecY back-to-back dimers by disulfide crosslinking [179]. SecY dimer levels were assessed in the secY Q212C mutant before
and after translocon jamming utilizing the OmpA-GFP chimera as before. We obtained a result similar to our \textit{in vivo} photo-crosslinking analysis, namely SecY dimer levels remained relatively constant independent of translocon jamming (Figure 2.8B). This result was not caused by artificial dimerization of SecY due to overexpression, since we obtained a similar result when SecY was expressed around chromosomal levels using 10\(\mu\)M IPTG (Figures 2.8B and 2.5A). It was also not caused by ineffective OmpA-GFP jamming since jamming led to an efficient buildup of MBP in the cytoplasm (Figure 2.8C). Collectively, these results suggest a common phenomenon, namely the retention of the SecY dimer during translocation-arrested conditions.

Our data thus far suggests that SecY dimers are retained and likely participate in active protein translocation. However, since we examined total cellular SecY protein in the foregoing analysis, one could argue that these experiments simply captured a residual pool of inactive SecY dimer. Given the rapidity and efficiency of our translocon-jamming methodology, we don’t feel that this explanation is likely. However, to directly address this issue, we performed a
FIGURE 2.7. Analysis of the secY S111 mutant during MalE-LacZ-induced translocon jamming by in vivo photo-crosslinking. (A-C) MM18.7 containing the pSup-pBpARS-6TRN and pCDFT7secYmycEG plasmids with the secY S111 amber mutation was grown in the presence of 1 mM pBpA and 30 μM IPTG until OD$_{600}$ reached 0.15, when the MalE-LacZ chimera was induced by adding maltose to a final concentration of 0.2%. Cells were harvested at the indicated time points post jamming, and exposed to UV irradiation as indicated. A wild-type (WT) strain was used in parallel as a control. (A) Western blot of cell membranes probed with c-Myc antibody. (B) Western blot of cells divided into cytoplasm/membrane (C) and periplasmic (P) fractions, compared to total (T) cell input, probed with MBP antibody. (C) Quantification of SecY dimer to monomer ratio (D/M) or cytoplasmic/membrane MBP to total MBP ratio (C/T) during jamming. The average results from three experiments are plotted with standard error measurements. The error bars for the D/M points are too small to be seen on this scale.

co-immunoprecipitation experiment to directly examine the nature of the OmpA-GFP-jammed SecY pool. Accordingly, a jamming experiment was performed with the secY
Q212C mutant, followed by *in vivo* disulfide crosslinking. Next, cells were broken and total membrane proteins were isolated after solubilization with DDM, which has been shown to preserve at least some

![Diagram](image1)

FIGURE 2.8. **Analysis of secY Q212C mutant during OmpA-GFP-induced translocon jamming by disulfide crosslinking.** (A) View of *T. Maritima* SecYEG docked in the back-to-back conformation with the same color scheme as Figure 2.1. Magenta spheres indicate the Q212 residue mutated to cysteine. (B-C) The *secY* Q212C mutant was grown in the presence of 10 μM or 30 μM IPTG and 0.2% maltose until OD<sub>600</sub> reached 0.15, when the OmpA-GFP chimera was induced by adding arabinose to a final concentration of 0.2%. Cells were harvested 45 min post jamming and treated with CuPhe3 and DTT as indicated. (B) Western blot of cell membranes probed with c-Myc antibody. (C) Western blot of cells divided into cytoplasm/membrane (C) and periplasmic (P) fractions, compared to total (T) cell input, probed with MBP antibody.

translocon associations [173]. Anti-GFP beads were utilized to specifically bind OmpA-GFP-associated membrane proteins. Controls indicated the specificity of the system, since SecY protein was not observed in the Anti-GFP bead eluate when the
strain remained uninduced (Figure 2.9, lanes 15-21). By contrast, during arabinose induction, both SecY dimers and monomers were observed in the eluate fraction along with OmpA-GFP (Figure 2.9, lanes 1-7). The identity of the disulfide-linked SecY dimer was confirmed by its sensitivity to reducing agent either within the total solubilized cell membranes or within the eluate itself (Figure 2.9, compare lanes 1 and 8, or 6 and 7, respectively). These results indicate that a significant fraction of total cellular SecY dimer pool is associated with OmpA-GFP. However, the experiment was limited by SecY recovery, since OmpA-GFP dissociated from the majority of both forms of SecY, presumably due to its instability in detergent (Figure 2.9, lanes 2 and 9). In several repetitions of this experiment we found that SecY dimers were more readily pulled down than monomers, suggesting that OmpA-GFP forms a more stable complex with the former species (data not shown). The possibility that SecY dimers formed in vitro under our conditions appeared unlikely based on a parallel study [173]. Together, these results strongly suggest a direct involvement of SecY dimers in active translocation.

FIGURE 2.9. Analysis of OmpA-GFP-jammed SecY complex by co-immunoprecipitation. The secY Q212C mutant was grown in the presence of 30 μM IPTG to an OD_{600} of 0.15, when arabinose was added as indicated to a final concentration of 0.2%, and growth was continued for an additional 45 min. Cells were
harvested and treated with CuPhe3 to crosslink SecY prior to purification. Membranes were isolated and solubilized at 4°C for 3 h in TSGM buffer containing 2% (wt/v) DDM without (w/o βME) or with (w/ βME) 5% β-mercaptoethanol as indicated, and insoluble material was removed by high-speed sedimentation. Solubilized membrane protein was mixed with anti-GFP beads at 4°C for 2 h, when beads were sedimented, washed, and the sample eluted as described by the supplier. Western blot of the total input (I) of solubilized membrane protein, the flow-thru (FT) fraction, three consecutive washes (W1, W2, and W3), and specifically eluted protein without (E) or with (EβME) treatment with 5% β-mercaptoethanol are shown. Western blots were probed with c-Myc or GFP antibody as indicated. Given the rapidly-folding GFP domain, OmpA-GFP can run as a doublet [179].

**SecY Dimer Levels Are Refractory to Translocon Clearing**—The foregoing results contradict a recent study indicating that SecY monomerizes under translocon-jamming conditions [179]. This latter finding would predict that SecYEG dimers would tend to accumulate when the translocon channel is idle. In order to test this prediction, we utilized the translation initiation inhibitor kasugamycin in order to clear channels of any substrate proteins. This antibiotic inhibits translation initiation by perturbing the mRNA codon-tRNA anticodon interaction, thus preventing the binding of fmet-tRNA to the P-site of the 30S ribosomal subunit [228]. We first titrated kasugamycin under our standard growth conditions in an isogenic strain that contained an arabinose-inducible lacZ gene in order to determine an effective concentration for use. We found that treatment with 2 mg/ml of kasugamycin for 2 min was sufficient to inhibit 99% of subsequently induced β-galactosidase activity (Figure 2.10A). Given known translation elongation rates under these growth conditions (12-17 amino acids per sec [229]), even the longest nascent polypeptide chains should be released from ribosomes within the first couple of minutes after this treatment. Furthermore, given
the rapid *in vivo* protein translocation rates observed in *E. coli* with a wild-type secretion machinery [162], translocon channels should be cleared of all substrates after 5 min of kasugamycin treatment.

We compared SecY dimer levels before and after kasugamycin treatment by *in vivo* photo-crosslinking of the secY K20 mutant. The SecY dimer level remained relatively constant in cells whether they were treated with kasugamycin for 5 or 10 min or not at all (Figure 2.10B). This result, along with the results presented above, indicate that SecY dimer levels remain relatively constant in cells and appear to be mostly refractory to changes in protein secretion physiology.

**FIGURE 2.10. Analysis of secY K20 mutant during kasugamycin treatment by in vivo photo-crosslinking.** (A) Kinetics of kasugamycin inhibition. BL26.1(λDE3) containing pSup-pBpARS-6TRN, pCDFT7secYmycEG, and pBAD-lacZ was grown to an OD₆₀₀ of 0.2, when kasugamycin was added to a final concentration of 2 mg/ml at time 0; arabinose was added to a final concentration of 0.2% at the indicated times, and 40 min later β-galactosidase activity was measured. The average Miller units [230] of β-galactosidase activity of duplicate or triplicate samples are shown. The basal level of β-galactosidase activity in the untreated culture lacking both kasugamycin and arabinose was subtracted as background from all assays. (B) The secY K20 mutant was grown in the presence of 1 mM pBpA and 30 μM IPTG until OD₆₀₀ reached 0.2, when kasugamycin was added to a final concentration of 2 mg/ml. Cells were harvested at
the indicated times and exposed to UV irradiation as indicated. Cell membranes were isolated and analyzed by western blotting using c-Myc antibody.

SecA Interacts with Two Copies of SecY During Translocon Jamming—Given the general controversy surrounding SecY oligomeric function, we sought an independent method to address this issue. In particular, we have previously utilized a form of in vivo photo-crosslinking termed SecA foot-printing, where distinct halves of the SecA molecule (NBD-1-NBD-2 or PPXD-HSD-THF-HWD-CTL) were found to associate with individual protomers of the SecY dimer [182]. We reasoned that this method could be modified to address the present issue if SecA could be captured in a translocationally-active complex with SecY. In order to accomplish this task, we fused SecA at its non-essential carboxyl-terminus to the amino-terminus of OmpA-GFP, to form a SecA-OmpA-GFP trimera. Since this trimera should be capable of efficient SecA-dependent jamming, it was placed under the tightly-regulated araBAD promoter system. We next engineered crosslinking sites for both halves of SecA in order test for interactions with individual protomers of the SecY dimer. For this purpose, an amber mutation was made at residue 59 within NBD-1 of SecA, and a cysteine was introduced at residue 21 (C21) of the OmpA signal sequence, adjacent to the distal half of SecA within the trimera (Figure 2.11A). The former position has been shown to be a hot spot for SecA-SecY photo-crosslinking, while the latter position has been shown to efficiently crosslink to a cysteine at residue 68 (C68) of the plug domain of SecY [179, 182]. This latter interaction would also ensure that only SecA-OmpA-GFP translocation intermediates that had inserted the OmpA signal sequence into the lateral gate of a translocationally-active SecY protomer would be detected by our system.
We performed a jamming experiment with the strain containing the SecA-OmpA-GFP trimera, and we found effective translocon jamming based on cytoplasmic MBP accumulation (Figure 2.11B). We then performed a SecA foot-printing experiment, where SecA-SecY interaction was probed during jamming conditions by utilizing photo-crosslinking and/or cysteine disulfide crosslinking. When single crosslinking was utilized, we found that the SecA-OmpA-GFP trimera was crosslinked to a single SecY protomer (Figure 2.11, panels C and D, lane 1-4). The identity of the crosslinked species was verified by its UV or copper phenanthroline (CuPhe3) dependency, its sensitivity to reduction in the latter case, and its reactivity to either GFP or c-Myc antibody. The intensity of the crosslinked species was significantly greater with disulfide crosslinking, consistent with the generally higher efficiency of this method compared to pBpA-induced photo-crosslinking. We also verified that the Cys21-Cys68 interaction of the OmpA signal sequence with the SecY plug domain was specific, because when Cys21 was mutated to serine, no crosslinked complex was generated (data not shown). Importantly, when double crosslinking was performed, we detected the SecA-OmpA-GFP-SecYY triple complex migrating above the double
FIGURE 2.11. SecA foot-printing study of SecY dimer status. (A) Cartoon of SecA foot-printing study, where interaction of the NBD-1-NBD-2 (N domain) half of SecA with SecY1 or PPXD-HSD-THF-HWD-CTL (C domain) half of SecA with SecY2 are depicted along with the OmpA-GFP chimera attached to SecA. Sites for photo-crosslinking the SecA N domain (59) with SecY1 or for disulfide crosslinking the OmpA signal sequence (C21) to the plug domain (C68) of SecY2 are shown. (B-D) The secY Cys68 mutant carrying the pBAD-secA-OmpA-GFP plasmid with the ompA Cys21 and secA 653 Amber (panel B) or secA 59 Amber (panel C-D) allele was grown in the presence of 1 mM pBpA and 30 μM IPTG until OD\textsubscript{600} reached 0.15, when arabinose was added to a final concentration of 0.2%. 45 min later, cells were harvested and either fractionated (panel B) or subjected to crosslinking (panel C-D). (B) Western blot of cells divided into cytoplasm/membrane (C) and periplasmic (P) fractions, compared to total (T) cell input, probed with MBP antibody. (C-D) Western blot of isolated membranes from cells treated with UV irradiation, CuPhe3 and/or DTT
as indicated probed with GFP or c-Myc antibody as indicated. A large degradation fragment is visible immediately below the SecA-OmpA-GFP trimera.

SecA-OmpA-GFP-SecY complex (Figure 2.11, panels C and D, lane 5). The former complex completely disappeared, and the latter complex was substantially diminished as well, when the doubly crosslinked sample was treated with DTT, consistent with the majority of crosslinking coming from the disulfide method (Figure 2.11, panels C and D, compare lanes 5 and 6). Because we used a disulfide pair that captures a translocation intermediate at the trans side of the channel, we were confident that the two crosslinked complexes represent ones engaged in active protein translocation. In short, the results of the SecA foot-printing study indicate that a SecA protomer can interact with two copies of SecY when it is fused to a translocation intermediate: indicative of an active role for SecY dimers in ongoing protein transport. However, given the limitations inherent to our foot-printing method, this approach does not address whether some protein translocation also occurs from SecA-bound, SecY monomers or whether the passive SecY copy stays associated with the active one throughout the entire translocation cycle.

5. DISCUSSION

While the earliest studies suggested that the protein-conducting channel might reside at the subunit interface of SecYEG dimers or tetramers [209, 210], later studies visualized a protein-conducting channel within a single SecYEG protomer that has been suggested to be fully functional [91, 135, 136, 175, 215-217, 231]. However a
variety of other structural and functional studies during the past decade have also pointed to the existence of a second SecYEG subunit that appears to play a critical role in facilitating overall translocon function [38, 132, 183, 211, 212, 214, 218, 224].

Thus the primary focus of the present study was to investigate the oligomeric state of SecYEG protein during ongoing protein translocation utilizing several different approaches that might yield a consistent picture in this matter. Towards this end we utilized both in vivo photo-crosslinking and disulfide crosslinking methods along with a previously developed SecA foot-printing approach to address this question. We found that the SecY dimer pool remained relatively constant whether translocons were cleared of substrate by kasugamycin pre-treatment, or whether they were engaged in active protein translocation utilizing the translocon jamming technique. This latter result was not simply due to more limited sampling of a subset of inactive translocons, since SecY dimer could be isolated associated with the OmpA-GFP translocation intermediate. Likewise, the existence of a substrate-engaged SecY dimer was also demonstrated in our SecA foot-printing study, where crosslinking footprints were seen with both the translocating and non-translocating SecY subunits. Collectively, these results argue strongly for a role of the SecYEG dimer in ongoing protein transport.

Our results contradict a parallel study that also utilized disulfide crosslinking and substrate jamming to study the functional state of the SecYEG subunit within cells [179]. Indeed, one reason that we compared our photo-crosslinking method with disulfide crosslinking, including utilizing some identical mutations with the same
jamming OmpA-GFP chimera, was because of the different results noted. We also compared jamming promoted by the shorter OmpA-GFP chimera with that promoted by the longer MalE-LacZ chimera, and we found similar results in both cases. In addition, we performed our study at moderate or even near chromosomal SecY levels in order to minimize the chance of artificial dimerization due to SecYEG over-production. Indeed, one of the strengths of our in vivo photo-crosslinking approach is its relatively low crosslinking efficiency (~10%) that minimally distorts the natural monomer-dimer equilibrium present within the cell. While it remains possible that different degrees of jamming were obtained in the two studies due to the differing stains or growth physiologies employed, our co-immunoprecipitation experiment indicated that at least a subset of the SecY dimer pool detected was derived from jammed translocons associated with substrate. The exact cause of discrepancy between the two studies is not clear at present, although we note different strains and somewhat different protocols were utilized for this portion of our study. Beyond the good agreement obtained between our photo-crosslinking and disulfide crosslinking studies, our SecA footprinting study provided yet another independent method to evaluate this issue. Here again we obtained a clear indication of the existence of a translocationally-active SecY dimer, although we were unable to address the more complicated question as to whether the dimer was stable throughout the entire translocation process given the limitations inherent in our approach.

A number of explanations have been given for the formation of SecYEG dimers. It has been suggested that the non-translocating SecYEG subunit could
provide an additional binding platform for SecA or the ribosome [218]. In the case of SecA, for example, association of the NBD-1-NBD-2 half of SecA with the inactive copy of SecYEG would provide a stator during cycles of SecA-dependent insertion of the preprotein into the translocationally-active SecYEG subunit promoted by the PPXD-HSD-THF-HWD-CTL half of SecA. Indeed the NBD-1-NBD-2 half of SecA is the portion that displays high affinity binding specificity for SecYEG [232]. This hypothesis would also be consistent with in vivo or in vitro complementation results obtained with strains that contain two different secY alleles [183, 214], as well as the recently observed processive action of SecA ATPase during protein transport [190]. It would also provide a mechanism to switch between ribosome and SecA-promoted transport for integral membrane proteins with periplasmic domains that require both pathways for their biogenesis. Perhaps, dimerization allows SecYEG to transit from a resting state to a translocation-ready state, characterized by greater SecA binding affinity [183] and more flexibility of to the plug domain to accommodate preprotein [208]. Interestingly, a recent study hints at a more subtle type of dimer function where the nature of the preprotein substrate may determine whether transport occurs via a SecYEG monomer or dimer [184]. This phenomenon points to a possible structural signal within particular substrate proteins that facilitates dimer formation. Finally, it could be that the dimer could be a storage form of SecYEG protein, in keeping with its accumulation under conditions of over-production. However, our data does not favor this hypothesis, since we saw little change in the dimer pool whether translocons were empty of protein substrate during kasugamycin treatment or whether they were
largely occupied during translocon jamming conditions. This hypothesis is also inconsistent with reports that SecA and substrate promote SecYEG dimer assembly in vitro, suggesting a functional role of dimers rather than merely a storage form [173, 180].

The extensive set of crosslinking residues utilized in our photo-crosslinking study provides additional confirmation for the existence of the two SecYEG dimer forms in vivo, particularly for the front-to-front dimer that is supported by relatively few crosslinks [179]. Of course the physiological relevance of either dimer form remains unclear. The front-to-front dimer with the lateral gates facing one another could allow protomers to fuse in order to form a larger composite channel, which would permit expansion of the channel to dimensions more consistent with two previous estimates [137, 219]. Such a fused dimer would only be useful for the transit of partially folded, secretory proteins, since it would lack any lateral gate opening for the exit of membrane proteins into the lipid bilayer. On the other hand, the back-to-back dimer would be suitable for translocation of either integral membrane or secretory proteins, but it would be more limited in its potential for channel expansion. Its two outward facing lateral gates would be in an ideal position for interaction with YidC or other chaperones involved in membrane protein folding and assembly [226].

Positing different functions for the two observed SecYEG dimer forms seems appealing given the large diversity of substate proteins in bacterial proteomes, which differ greatly in their disposition of topogenic sequences as well as rate of delivery to the translocon.
Given the quantity and quality of seemingly opposing findings in this area, it seems reasonable to conclude that both views have merit. Thus the SecYEG monomer may indeed be the minimal functional unit of the translocon for many protein substrates, while the dimer may be a more optimal functional unit under a variety of conditions including those examined in the present study. Thus studies reporting exclusive dimer function may simply be reflective of the higher specific binding and transport activity of the dimer form of the translocon compared to its monomer counterpart, or they may relate to the type of substrate protein under investigation. A more complicated view posits that both forms of the translocon are required at different stages of the translocation cycle such that the monomer and dimer switch off with one another as needed. Clearly the initiation or elongation phases of protein transport place different requirements on the translocon as does the translocation of proteins with differing topologies or destinations. Thus additional approaches with a higher degree of sophistication and resolution will be required to address these two opposing viewpoints. State of the art techniques such as real time FRET and fluorescence microscopy may enable us to watch how SecYEG monomers and dimers work during active protein translocation in model systems or perhaps even within living cells.
Chapter III

SecYEG monomer function: a mutational study

1. ABSTRACT

Transport of membrane or secretory proteins into or across the bacterial plasma membrane employs the universally conserved SecYEG channel complex. Whether this complex functions as a monomer or dimer during transport remains controversial. Here we reexamined a secYE mutant whose translocons were reported to exist solely as monomers in vivo. We utilized arginine substitutions at the dimer interface between adjacent SecY or SecE protomers to disrupt the front-to-front or back-to-back dimers, respectively, which was confirmed by disulfide crosslinking. The resulting secY, secE, or secYE mutants were viable, although they had reduced growth rates compared to wild-type as secYEG expression was reduced. The translocation efficiency of the secYE mutant was modestly reduced for OmpA, maltose-binding protein, and galactose-binding protein compared to wild-type, but these differences largely disappeared when cultures with comparable SecY levels were compared. Finally, two independent methods were utilized to reassess the monomeric state of the secYE mutant translocon. A SecA foot-printing study employing a translocon-jamming SecA-OmpA-GFP trimera displayed crosslinking to two SecY copies, consistent with a dimer at arrested translocation sites in the mutant. Furthermore, site-specific photo-crosslinking revealed similar levels of front-to-front and back-to-back dimers in both the wild-type and secYE mutant. Thus while bulky arginine substitutions perturbed the
cysteines utilized to detect SecYEG dimers, they did not actually prevent dimerization in the mutant. Our study indicates that additional work will be needed to address the functional state of the SecYEG monomer and discusses potential in vivo approaches.

**IMPORTANCE** The universally conserved SecYEG channel complex is the major protein transport pathway for membrane or secretory proteins into or across the bacterial plasma membrane. It exists as both monomers and homo-dimers with two different orientations. Whether it functions as a monomer or dimer during transport remains controversial. Here we reexamined an *E. coli* secYE mutant whose translocons were reported to exist solely as monomers based on disulfide crosslinking. While the mutant showed comparable translocation efficiencies compared to wild-type, tellingly its translocons had similar levels of both dimers based on SecA footprinting and site-specific photo-crosslinking studies. Our work clarifies the nature of this important mutant and serves as a cautionary tale for use of disulfide crosslinking when assessing dynamic protein-protein interactions.
2. INTRODUCTION

The conserved Sec-dependent pathway is the major protein transport system in prokaryotes and eukaryotes. It inserts integral membrane proteins into the lipid bilayer or translocates secretory proteins across the membrane into the periplasm (prokaryotes) or lumen of the endoplasmic reticulum (eukaryotes). In bacteria, co-translational translocation starts by recognition of a highly hydrophobic signal-anchor sequence of generally a nascent membrane protein by signal recognition particle, which recruits this complex to its SecY-associated membrane receptor FtsY, whereby the energy of GTP binding and hydrolysis transfers the nascent chain-ribosome complex to the SecYEG translocon (for a review see [233]). It has been proposed that the energy of translation drives the polypeptide chain into the SecYEG channel for subsequent integration of the integral membrane protein into the lipid bilayer. On the other hand, post-translational translocation starts by recognition of less hydrophobic signal sequences or unfolded sequences within the mature region of a pre-protein by a variety of chaperones, which include the ribosome-bound trigger factor, SecA or SecB, which keeps the fully translated pre-protein in a translocation-competent, loosely-folded state before its transfer to the SecYEG-bound, SecA ATPase motor protein [3]. The energy of ATP binding and hydrolysis drives the translocation of the pre-protein through the SecYEG channel by an as yet poorly defined mechanism. The co-crystal structure of the SecA-bound SecYEG complex has provided a model for substrate protein transport through this complex [91]. ATP binding to SecA allows its pre-
protein-binding domain (PPXD) to capture a distal portion of the substrate polypeptide chain in a clamp formed by PPXD, the helical scaffold domain (HSD) and nucleotide-binding domain two (NBD2), while the proximal portion of the substrate forms a hairpin structure encompassing the signal peptide and early mature region of the pre-protein that is associated with the SecA two-helix finger sub-domain (THF) which is positioned at the mouth of the channel. In this location the THF could act as an ATP-dependent piston to drive the forward movement of the polypeptide chain across the channel. While a number of studies now support this model of SecA action [125, 190, 234, 235], other studies have casted doubt on this proposed mechanism and suggest that the ATPase cycle of SecA may allow for cycles of channel opening and closure to create a Brownian Ratchet mechanism with biased net forward movement of the substrate polypeptide chain [189, 197]

The universally conserved SecYEG channel complex (or Sec61αβγ in the case of Archaea and Eukarya) resembles an hourglass with larger openings on the exterior sides of the membrane and a narrower constriction ring lined with aliphatic residues located toward its center that would facilitate channel gating [136]. Indeed, the crystal structure of a SecA-SecYEG complex engaged with an artificial substrate showed tight interactions between the pore ring residues and the protein substrate [141]. Its main component, SecY protein, contains ten transmembrane helices that are arranged in a pseudo-symmetrical clamp-like structure with a SecE “hinge” at its back and a lateral gate at its front, the latter of which allows for the exit of integral membrane proteins
into the lipid bilayer. A movable plug domain that is located on the exterior side of the channel could facilitate channel closure during inactive periods, although this domain appears to be involved in channel activation as well [30].

Whether SecYEG protein functions as a monomer or dimer has significant impact on our understanding of the translocation mechanism. In the simplest case, a monomer may suffice for ribosome or SecA and substrate protein binding, translocon activation and the subsequent transport steps. By contrast, if the dimer is the functional unit, then it is likely to require more regulatory signals for the timely and correct assembly of the relevant ternary complex as well as collaboration between each protomer to perform its specific function during the transport process. Further complicating this issue is the finding that the SecYEG dimer exists in two distinct forms: both front-to-front (FtF) and back-to-back (BtB) states have been observed (see references contained in [179]). In the former case the lateral gate regions of SecY face one another at the dimer interface, while in the latter case they are at opposite ends of the dimer distal to the dimer interface. Thus there remains the possibility that two protomers of the FtF dimer could form a composite channel with larger dimensions beyond that of a single channel: a potentially useful solution for transporting larger, partially-folded, substrate proteins. By contrast, having the lateral gates distal to the dimer interface in the case of the BtB dimer would be important to accommodate its interaction with the large membrane protein chaperone YidC [170], which binds
transmembrane segments of integral membrane proteins as they exit the lateral gate in order to assist in their folding.

Whether the SecYEG channel complex functions as monomer or dimer during protein transport has been an ongoing debate [236]. SecYEG structures containing pre-activated SecA with or without bound substrate or bound to a ribosome exist, but they generally contain a single SecYEG protomer [37, 91, 138, 141, 237, 238], suggesting that the monomer may be the functional unit. In support of this possibility, when fluorescently labeled SecYEG protein was functionally reconstituted into giant unilamellar vesicles at physiological concentration and SecA dependent transport of an OmpA-dihydrofolate reductase substrate forming a stable translocation intermediate was studied, no SecYEG oligomers were detected [177]. Furthermore, an in vivo study showed that jamming of the translocon with an OmpA-GFP substrate in order to mimic a translocation intermediate state caused the dissociation of SecYEG dimers based on a loss of disulfide crosslinking, indicating that SecYEG monomerizes during active protein transport [179]. However this latter study has been disputed based on maintenance of the dimer during OmpA-GFP jamming as assessed by in vivo photo-crosslinking or disulfide crosslinking as well as in vivo SecA foot-printing [239].

Other biochemical studies are consistent with SecYEG dimer being the functional translocon unit. Interactions of SecYEG protein with SecA, substrate proteins, ATP and acidic phospholipids, all of which are required for protein
translocation, have been shown to promote SecYEG dimer formation [44, 180]. Estimates of the translocon channel diameter utilizing the \textit{in vitro} translocation of a protein substrate conjugated to rigid aromatic ring systems of different sizes gave a limit of \~\%22-24 \text{ Å} [137], which is difficult to reconcile with estimates of the maximal expandable size of a single channel of \~\%16 \text{ Å} based on molecular dynamics simulations [25]. This result would be more consistent with the existence of a FtF dimer with a fused composite channel. An \textit{in vivo} photo-crosslinking study mapping the interface between SecA and SecY favored the FtF dimer, since only this orientation was consistent with the footprint of the N- and C-domains of SecA on distinct SecY protomers [182]. Furthermore, genetic studies of a covalently-linked version of SecY dimer reveal the presence of intragenic complementation: a result most easily explained by functional division between the two SecYEG protomers making up the dimer [214, 240]. For example, a functional covalent dimer was formed from a SecY protomer defective in SecA binding attached to another one that was defective in its transport function. Finally, single molecule experiments comparing liposomes reconstituted with single, multiple, or BtB-crosslinked SecYEG copies found that while all three states were sufficient for SecA and substrate protein binding to the proteoliposomes, only the latter two states containing SecYEG dimers were able to promote protein translocation [241].

A recent study has complicated this matter further with the finding that SecYEG monomers or dimers may be required depending on the substrate protein of
interest. For example, by using proteins that contained jamming-dependent C-terminal loops as substrates to titrate SecYEG functional units during \textit{in vitro} protein translocation, it was found that the galactose-binding protein (GBP) precursor required twice the number of translocon units as pro-OmpA did for complete functional jamming to occur [184]. The simplest explanation for this observation is that SecYEG dimer is required for the translocation of GBP, while the monomer transports pro-OmpA. This brings up the possibility that both SecYEG monomer and dimer states may be functional: i.e. they may transport distinct substrates as suggested by the study above, or alternatively, the dimer may provide enhanced function over the monomer as was suggested previously [214].

Previously we have utilized both \textit{in vivo} photo-crosslinking and di-sulfide crosslinking to better define the FtF and BtB SecYEG dimer interfaces as well as the functional state of these dimers. In contradiction to a previously published study [179], we found that both dimer forms were maintained throughout translocon jamming, and they could be isolated when the jammed complex was purified, suggesting that SecYEG dimers were active in protein transport. Furthermore, by utilizing an \textit{in vivo} SecA foot-printing technique, we found that both protomers of the SecY dimer were in contact with the SecA-attached jamming substrate, indicative of a dimer-based transport mechanism [239]. However, a major limitation of our study was that we couldn’t exclude the possibility that a fraction of the observed transport occurred exclusively via a SecYEG monomer, or alternatively, that the monomer was part of a
monomer-dimer transport cycle that was not captured by our translocon jamming and crosslinking methodology.

In order to directly address the functional state of the SecYEG monomer in vivo, we constructed and characterized a strain that was previously described to be defective in formation of both the FtF and BtB dimers and to produce only SecYEG monomers [179]. The growth and transport properties of this strain for a number of different substrate proteins are described below, along with more rigorous methods to evaluate the extent of its SecYEG oligomerization defect.

3. EXPERIMENTAL PROCEDURES

Chemicals, Media, Strains and Plasmids. LB (Miller) broth, arabinose, ampicillin, kanamycin, chloramphenicol, streptomycin, and anhydrotetracycline were purchased from Fisher Scientific. pBpA and maltose were purchased from Bachem and Difco, respectively. Galactose, biotin, spectinomycin, kasugamycin, and the Protease Inhibitor mixture were obtained from Sigma Aldrich. Quik-Change™ and Wizard Plus SV Miniprep DNA Purification System kits were obtained from Stratagene and Promega, respectively. Restriction enzymes were obtained from New England Biolabs, while the WesternBright™ Sirius enhanced chemiluminescence kit was obtained from Advansta. Most other common laboratory chemicals were obtained from Fisher Scientific or Sigma Aldrich and were laboratory grade or better. Antisera to MBP, GBP or OmpA were prepared from purified proteins by hyper-immunizing rabbits, and the later two antisera were the kind gift of Linda Randall. Chicken anti-
GFP antibody and HRP-conjugated goat anti-rabbit IgG antibody were obtained from Abcam. HRP-conjugated goat anti-chicken IgG antibody was obtained from Jackson ImmunoResearch Laboratories. Peptide affinity-purified SecY antisera was prepared by hyper-immunizing rabbits to a peptide identical to the carboxyl terminus of SecY (CYESALKKANLKYGR) conjugated to keyhole limpet hemocyanin by maleimide chemistry utilizing Tana Laboratories, LC (Houstin, TX) for peptide synthesis, protein carrier conjugation, immunization, and peptide affinity purification of the antisera. Streptavidin HRP conjugate was obtained from ThermoFisher. The *E. coli* secYE chromosomal null mutant EP66 (*secY::hph secE::aadA* pTet2-SecYEG), a kind gift of Tom Rapoport [179], was constructed by replacing the *secY* and *secE* genes with hygromycin (*hph*) and streptomycin (*aadA*) antibiotic resistant genes, respectively, using strain BW25113 according to the method of Datsenko and Wanner [242]. It contains a p15A-derived plasmid with the *secYEG* genes under the control of the Tet-inducible promoter and a selectable kanamycin resistance gene. A derivative of EP66, EP66.1 (*secY::hph secE::aadA* pACYC-SecYEG) that contains the p15A-derived plasmid with the *secYEG* genes under the constitutive *rplN* operon promoter and a selectable chloramphenicol resistance gene was made by plasmid shuffling through transformation of EP66 with the relevant plasmid DNA and selection for chloramphenicol resistance. The *secY*, *secE*, and *secYE* mutant strains described in our study were made by QuikChange™ mutagenesis of pTet2-SecYEG followed by plasmid shuffling into EP66.1 with selection for kanamycin resistance. The mutants were verified by DNA sequencing of the entire *secY* and *secE* genes on pTet2-SecYEG
to confirm the presence of only the relevant mutations and by PCR analysis of chromosomal DNA using primers flanking the secY and secE loci to confirm the retention of the secY::hph and secE::aadA null alleles. The cloDF-derived plasmid pSup-pBpARS6TRN-cloDF encodes a Methanococcus jannaschii amber suppressor tRNA-tRNA synthetase pair that efficiently incorporates pBpA at the amber stop codon [201]. It was made by using QuikChange™ mutagenesis to insert BglII and MluI restriction sites flanking the origins of DNA replication of the cloDF-derived plasmid pCDFT7SecYEG [182] and the p15A-derived plasmid pSup-pBpARS6TRN [201]. This allowed replacement of the replication origin of pBpARS6TRN with that of pCDFT7SecYEG by swapping the relevant BglII-MluI restriction fragments by standard recombinant DNA methods. The pBAD-OmpA-GFP plasmid producing the OmpA-GFP translocon jamming chimera under the control of the arabinose inducible araBAD promoter was obtained from Tom Rapoport [179]. The pBAD-SecA-OmpA-GFP plasmid that produces the OmpA-GFP chimera fused to the carboxyl-terminus of SecA under the control of the arabinose inducible araBAD promoter has been described previously [239]. The pBAD-lacZ plasmid encoding the lacZ gene under the control of the araBAD promoter with an ampicillin resistance selectable marker was obtained from Invitrogen. The pGJ78-I plasmid producing the MalF-PSBT-I chimera was obtained from Jon Beckwith [243]. All site-directed mutations and plasmids constructed by recombinant DNA methods were verified by DNA sequence analysis utilizing the University of Pennsylvania DNA Sequencing Facility.
**Cell growth.** Cells were grown for analysis in two different ways depending on whether *secYEG* expression was being titrated or not. In general, a freshly struck out single colony of the appropriate strain was inoculated into 10 ml LB media supplemented with antibiotics (30 µg/ml kanamycin and 50 ng/ml anhydrotetracycline for pTet2-SecYEG plasmids; 25 µg/ml chloramphenicol for pACYC-SecYEG or pGJ78-I encoding MalF-PSBT-I plasmids; 100 µg/ml spectinomycin and 1 mM pBpA for pSup-pBpARS6TRN-cloDF; 100 µg/ml ampicillin for pBAD-SecA-OmpA-GFP or pBAD-lacZ plasmids) and grown overnight at 37°C with shaking at 250 rpm. The overnight culture was diluted 1:100 or 1:200 into LB supplemented with appropriate sugars (0.2% maltose or galactose for induction of MBP or GBP, respectively, where indicated) and appropriate antibiotics and grown at 37°C with shaking at 250 rpm to an \(A_{600}\) of 0.3, when the indicated \(A_{600}\) cell equivalents were harvested by sedimentation in a microcentrifuge at 14,000 rpm for 5 min, washed with 5 ml of PBS (10 mM sodium phosphate, pH 7.5, 140 mM NaCl), and resuspended in the appropriate buffer for a given experiment. In experiments where *secYEG* expression was titrated, after sub-culturing of the overnight culture, the culture was immediately divided into ten equal sub-cultures, each of which was supplemented with a different concentration of anhydrotetracycline (5, 8, 10, 13, 15, 18, 20, 25, 30, 50 ng/ml), and grown and harvested as described above.

**Cysteine crosslinking.** 9 \(A_{600}\) cell equivalent of each culture was harvested and resuspended in 6 ml of protoplast buffer (100 mM sodium phosphate, pH 7.5, 5 mM EDTA, 10 mM phenanthroline). A 4 ml sample was treated with copper
phenanthroline (180 mM phenanthroline, 60 mM CuSO₄, 50 mM NaH₂PO₄) at a final concentration of 300 μM [244] for 10 min at 30°C with shaking at 250 rpm, while the remaining 2 ml served as the untreated control. The reaction was quenched by addition of N-ethylmaleimide and EDTA at final concentrations of 1 mg/ml and 5 μM, respectively, for 10 min. For treatment with reducing agent, half of the copper phenanthroline-treated sample received DTT at a final concentration of 60 mM for 10 min. Whole cell analysis was done by resuspending each sample in 100 μl 2x non-reducing sample buffer (40 mM Tris–HCl, pH 7.8, 16 mM NaH₂PO₄, 2% (w/v) SDS, 50 μg/ml bromophenol blue, 20 mM EDTA, 2 mg/ml N-ethylmaleimide, 0.10 g/ml sucrose), followed by a 5 min incubation at 37°C before debris was removed by sedimentation in a microfuge at 14,000 rpm for 10 min. Samples were run on a 15% SDS-PAGE gel at 100 V at 4°C until the dye front reached the bottom. Western transfers were performed at 100 V for 1 h, and nitrocellulose membranes were blocked overnight in 10 ml TBS buffer (20 mM Tris-HCl, pH 7.5, 140 mM NaCl, 0.25% Tween 20) supplemented with 10% nonfat dry milk. Membranes were incubated with gentle shaking for 1 h at room temperature with rabbit anti-SecY peptide antibody, washed briefly with TBS, and incubated with gentle shaking for 1 h at room temperature with HRP-conjugated goat anti-rabbit IgG antibody, both at 1:5000 dilutions, followed by ECL visualization according to the manufacturer’s protocol.

**Translocation Efficiency of MBP, GBP and OmpA.** Duplicate samples of 3 A₆₀₀ cell equivalent of each culture were harvested, washed in PBS, and resuspended appropriately for assessment of either MBP or GBP secretion by cell fractionation or
SecY protein levels by western blotting as described previously [239]. For cell fractionation, cells were resuspended in an ice cold solution of 200 µl 20% sucrose buffered in 0.03 M Tris-HCl, pH 8, and half the sample was saved as a total cell control, while the other half was subjected to spheroplast formation by addition of 10 µl of 1 mg/ml lysozyme in 0.1M EDTA, pH 8, followed by incubation on ice for 15 min. Lysozyme-treated cells were sedimented in the microfuge at 5,500 rpm for 5 min, and the supernatant (periplasmic fraction) was removed, while the cell pellet (cytoplasm-membrane fraction) was washed with 600 µl of the above buffered sucrose solution and resuspended in 100 µl buffer A (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM MgSO₄). One-third volume of 4× sample buffer (8% SDS, 500 mM Tris-HCl, pH 6.8, 20% 2-mercaptoethanol, 60% glycerol, 0.02% bromophenol blue) was added to each fraction, followed by incubation at 100°C for 5 min. For analysis of SecY levels, cells were resuspended in 100 µl 2× sample buffer and incubated at 37°C for 5 min, when insoluble material was removed by sedimentation in a microfuge at 14,000 rpm for 10 min. 3 µl, 15 µl, or 20 µl samples for analysis of MBP, GBP, or SecY, respectively, were run on 15% SDS-PAGE gels and western blotted onto nitrocellulose membranes as described above. Membranes were probed with rabbit anti-MBP, anti-GBP or anti-SecY peptide antibodies followed by HRP-conjugated goat anti-rabbit IgG antibody, all at a 1:5000 dilution, followed by ECL visualization. Analysis of OmpA secretion was performed similarly as described above for the analysis of SecY levels except that samples were incubated at 100°C for 5 min, when they were run on a 20 cm long 15% SDS-PAGE gel at 72 Volt for 48 hr. Only the relevant portion of
the gel was excised in this case and transferred to the nitrocellulose membrane, which was probed with rabbit anti-pro-OmpA antibody and HRP-conjugated goat anti-rabbit IgG antibody, both at a 1:20,000 dilution prior to ECL visualization according to the manufacturer’s protocol.

**Membrane insertion efficiency of MalF-PSBT-I.** Duplicate samples of 3 $A_{600}$ cell equivalent of each culture were harvested, washed in PBS, and resuspended in 100 µl 2X sample buffer. Samples were either heated at 100°C for 5 min for analysis of MalF-PSBT-I protein, or they were incubated at 37°C for 5 min when insoluble material was removed by sedimentation in the microfuge at 14,000 rpm for 10 min for analysis of SecY protein levels. 20 µl samples were analyzed on 15% SDS-PAGE gels followed by western blotting. Membranes were probed with Streptavidin-HRP for detection of biotinylated MalF-PSBT-I protein or rabbit anti-SecY peptide antibody followed by HRP-conjugated goat anti-rabbit IgG antibody, all at a 1:5,000 dilution, for detection of SecY protein prior to ECL visualization according to the manufacturer’s protocol.

**Photo-crosslinking.** The procedure described previously by Zheng et al. [239] was employed. In particular, 6 $A_{600}$ cell equivalent of each culture was harvested, washed and resuspended in 4 ml of PBS. 2 ml samples were irradiated with UV light at 360 nm on ice for 20 min using a Rayonet 2000 UV crosslinker (Southern New England Ultraviolet Company), while the remaining 2 ml samples were saved as the non-irradiated control. Each sample was sedimented in the microfuge at 14,000 rpm for 2 min, resuspended in 1 ml of breakage buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA,
25 mM PMSF, 1 mM DTT, 100 μg/ml of RNase, 100 μg/ml DNase, 60 μg/ml of lysozyme, 1× Protease inhibitor mixture) and disrupted with 30 sec bursts of a cup horn sonicator (Heat Systems) in a polycarbonate tube until near clarity. Unbroken cells were removed by sedimentation in a microfuge at 14,000 rpm for 5 min. Supernatant was isolated and membranes were sedimented in a Sorvall S120 AT2 rotor at 82,000 rpm for 30 min at 4°C. Each membrane pellet was solubilized in 60 μl ABB buffer (5% SDS, 10 mM Tris-HCl, pH 8, 1 mM EDTA) with constant stirring for 45 min at 37°C, when an equal volume of 2× sample buffer was added and stirring continued for an additional 15 min. 20 μl samples were analyzed on 15% SDS-PAGE gels followed by western blotting. Membranes were probed with rabbit anti-SecY peptide antibody followed by HRP-conjugated goat anti-rabbit IgG antibody, both at a 1:5000 dilution, prior to ECL visualization according to the manufacturer’s protocol.

**SecA Foot-printing.** The procedure described previously by Zheng et al. [239] was employed. Strains were grown as described above in “Cell growth” except that 30 ng/ml of anhydrotetracycline was employed for modest secYEG over-expression, the Sec-OmpA-GFP trimers was induced at an $A_{600}$ of 0.15 by addition of arabinose to 0.2%, and cells were grown for an additional 50 min prior to harvesting. 18 $A_{600}$ cell equivalent of each culture was harvested, washed and resuspended in 12 ml of PBS. The sample was divided into 6 equal portions for the appropriate treatment regime: UV irradiated at 360 nm or non-irradiated, treated with copper phenanthroline or copper phenanthroline plus DTT, or treated with copper phenanthroline and UV
irradiation followed by DTT treatment or not. Samples were processed identically as described above for “Photo-crosslinking” except that non-reducing sample buffer was used for membrane solubilization and analysis of the copper phenanthroline-treated samples. 5 µl or 20 µl samples were run on 7.5% or 11.3% SDS-PAGE gels, respectively, for analysis of SecA-OmpA-GFP or SecY proteins. Following western blotting, membranes were probed with chicken anti-GFP antibody followed by HRP-conjugated goat anti-chicken IgG antibody, both at a 1:10,000 dilution, for detection of the SecA-OmpA-GFP trimera, or alternatively with rabbit anti-SecY peptide antibody followed by HRP-conjugated goat anti-rabbit IgG antibody, both at a 1:5000 dilution, for detection of SecY protein prior to ECL visualization according to the manufacturer’s protocol.

4. RESULTS

Disruption of SecYEG dimer. In order to address the functional state of the SecYEG monomer, we constructed and characterized a secYE mutant that has been described previously to produce solely SecYEG monomers in vivo [179]. Mutations were introduced into the secY and secE genes by site-directed mutagenesis of plasmid pTet2-SecYEG carrying the secYEG genes under the inducible tetracycline promoter that allows for regulated secYEG gene expression. This plasmid was then transformed into a strain deleted for the normally essential secY and secE chromosomal genes utilizing a plasmid shuffling scheme that allows for displacement of the resident secYEG-containing plasmid by the incoming mutated plasmid of the same
compatibility group (Figure 3.1A). In our case both the mutated and resident plasmids were of the p15A compatibility group and contained different antibiotic resistance markers (kanamycin or chloramphenicol resistance, respectively) allowing for easy plasmid exchange, since the relevant secY and secE mutations have been reported to maintain essential gene function [179].

Charged and bulky arginine substitutions were engineered at both the FtF and BtB dimer interfaces in order to disrupt both types of SecYEG dimers (Figure 3.1B). Arginine substitutions within SecY TM3 (at the FtF interface) or SecE TM3 (at the BtB interface) were utilized for this purpose (Figure 3.1C). In the former case we also replaced several residues flanking the arginine residues with more hydrophobic ones in order to prevent a possible SecY TM3 membrane integration defect as described previously [179]. Finally, we engineered appropriate cysteine substitutions within SecY in order to employ disulfide crosslinking to monitor dimerization capability. We utilized residues within the periplasmic loops of SecY that strongly crosslink either the FtF or BtB dimers for this purpose [179]. These genetic manipulations resulted in the creation of three plasmids that contained the aforementioned mutations in secY, secE, or in both genes. Upon successful shuffling of the relevant plasmids into the secYE null strain (Figure 3.1D), we obtained secY or secE mutant haploid strains that potentially would be defective in the formation of FtF or BtB dimers, respectively, as well as the secYE mutant haploid strain that potentially would be defective for assembly of both dimer states. We also constructed an isogenic wild-type strain by this
method for strict comparison purposes. The four strains were verified by PCR analysis to show that they still carried the relevant drug cartridges at the deleted \textit{secY} and \textit{secE} chromosomal loci, and by DNA sequencing of the relevant pTet2-SecYEG plasmids to show that the plasmid-borne \textit{secY} and \textit{secE} genes contained only the engineered mutations of interest (data not shown).

![Diagram](image.png)

**FIGURE 3.1** Construction of \textit{secY}, \textit{secE}, and \textit{secYE} mutants affecting SecYEG dimerization. (A) Genetic scheme depicting the construction of the relevant mutants utilizing plasmid shuffling. The incoming plasmid shown in green contains the relevant \textit{secY} and/or \textit{secE} mutations. \textit{hph} and \textit{aadA} indicate hygromycin resistant or
streptomycin resistant gene cartridges that replaced the chromosomal secY or secE genes, respectively. (B) Side view of T.maritima SecYEG docked in FtF (left) or BtB (right) conformations according to the previously known sites of cysteine crosslinking. SecY, SecE, and SecG are colored in green, red and cyan, respectively, with the two lateral gate helices TM2b and TM7 colored in orange. Blue spheres depict residues mutated to arginine to disrupt the dimer interface, and the black spheres are cysteine reporter pairs used to assess dimer disruption. (C) SecY or SecE TM3 amino acid sequence is shown with arginine substitutions for disruption of the FtF or BtB interface, respectively, labeled in red font and other flanking mutations to increase the hydrophobicity of the SecY TM3 helix labeled in green font. (D) The wild-type (Y or E) and secY:hphA (Y0) and secE:aadA (E0) null alleles were verified by agarose gel analysis of PCR reactions using primers that annealed to sequences flanking the chromosomal secY or secE genes. (E) in vivo disulfide crosslinking was performed on the wild-type (WT) and secYE mutant (YE) using the indicated reporter cysteine to assess formation of the FtF or BtB dimer as described in “Materials and Methods”. Western blots were probed with anti-SecY peptide antibody. CuPh and DTT indicate treatment with copper phenanthroline and dithiothreitol, respectively. The SecY monomer (Y) and crosslinked dimer (YY) bands are indicated along with two unidentified bands that are marked with stars. Appropriate molecular weight markers are also indicated.

Disulfide crosslinking was performed on the newly constructed secYE mutant in order to assess the extent of disruption of the individual FtF and BtB dimers. Upon treatment with the oxidizer copper phenanthroline, the wild-type strain formed strong SecY dimer bands with both cysteine reporters, which were reversed by treatment with the reductant DTT, indicative of authentic cysteine-driven dimer formation for both the FtF and BtB dimer states (Figure 3.1E). Analysis of the secYE mutant, by contrast, showed that both the FtF and BtB SecY dimer bands were largely undetectable, consistent with the previous finding that these arginine substitutions largely monomerized the SecYEG population in this case [179]. We also observed two unidentified minor bands (marked with stars) that may be due to the presence of cross-reactive antibodies within our anti-SecY peptide antisera; these species appeared to be
irrelevant to our study based on their unresponsiveness to dithiothreitol (DTT)-induced reduction following copper phenanthroline treatment.

**Reduced growth rate of secY, secE, and secYE mutants.** We compared the growth rate of the secY, secE, and secYE mutants with the isogenic wild-type strain at varying secYEG expression levels. This parameter should be a good albeit indirect read-out of the different SecYEG functional states given the large number of essential proteins that utilize this channel for their biogenesis. Strains were grown at varying concentrations of the Tet promoter inducer anhydrotetracycline, and western blots were performed on SecY protein in order to account for any differential translocon stability in our mutants. All four strains grew comparably at the highest inducer concentration examined. However, they showed a clear hierarchy in their decline of growth rate as anhydrotetracycline concentration was lowered: the secYE mutant exhibited the slowest growth rate, followed by the secY mutant, while the secE mutant was the closest to the wild-type strain, which had the fastest growth rate (Figure 3.2A-F). In fact, the wild-type strain showed very little decline in growth rate over this dynamic range of secYEG expression. By contrast, the secYE double mutant barely grew at the lowest inducer concentration, although its SecY steady state level was only approximately half that of the wild-type strain at this inducer concentration (Figure 3.2A and G). Thus at least part of the difference in the observed growth rates between these strains was due to differential SecY protein levels at a given inducer concentration. The approximately two-fold reduction in SecY content was most likely due to a modest assembly defect for this mutant, since transcriptional and translational
differences were unlikely given the isogenic nature of our strains that differed only in the engineered point mutations within the plasmid-borne secY and secE genes. It has been shown previously that misfolded SecY protein is rapidly degraded by the ATP-dependent membrane proteases FtsH [245]. In order to normalize for this effect, we compared the growth rate curves of the different strains at approximately comparable SecY levels (Figure 3.2H and I). This analysis showed that the SecY-normalized growth rate in the ~30-45% chromosomal level was similar for the wild-type and secE mutant, while it was decreased significantly for the secY mutant, and decreased even more for the secYE mutant. Ignoring possible secondary effects caused by the engineered mutations, such results suggest that the FtF dimer may be the most active translocon species, with declining activity for the BtB dimer, followed by the least active monomer species. However, these data suggest that all translocon states appear to be sufficiently active once SecY levels reached ~60% of the chromosomal level, since any growth rate variation largely disappeared at the highest inducer concentration tested here (Figure 3.2F).
FIGURE 3.2 Analysis of growth rates of mutant strains. (A-F) Wild-type (blue), secY mutant (red), secE mutant (green) and secYE mutant (purple) strains were grown in duplicate in the presence of the indicated concentration of anhydrotetracycline and selective antibiotics as needed, and A_{600} readings were taken at the indicated time and plotted. Cultures were diluted each time they reached A_{600} of ~0.5 to insure maintenance of exponential growth and linearity of measurements, and dilution factors were taken into account during plotting. (G) Western blots of SecY protein level from the same cultures shown in panels A-E. Bands were quantified and expressed as percentage of the SecY chromosomal level (% Chr) for the secY wild-type strain BW25113 (Chr 100% loading). Numbers highlighted in red and green correspond to the amount of anhydrotetracycline that gave ~30 or ~45%
SecY chromosomal level, respectively. (H-I) Comparison of growth curves for duplicate strains at anhydrotetracycline levels that facilitated SecY levels of ~30% (H) or ~45% (I) SecY chromosomal level.

**Translocation efficiency of mutants for different substrates.** One plausible explanation for the existence of different SecYEG oligomeric states may have to do with different translocation needs of particular substrate proteins. Even if the SecYEG monomer is capable of translocating all of the substrate proteins within a cell, a particular SecYEG dimer state may do so much more rapidly for a sub-set of more “difficult” substrates for example. Such a difference in transport rates between these two different translocon states would become manifest by the accumulation of the relevant substrate proteins within the cytoplasm of the strain making solely the SecYEG monomer when secYEG expression was reduced below a critical level. Therefore, we compared the translocation efficiency of the secYE mutant with the isogenic wild-type strain for a number of substrate proteins at a variety of different secYEG expression levels, normalizing for the difference in SecY protein levels between the two strains. Given the laborious nature of this analysis, we only compared these two strains.

Our analysis utilized one of three different assays to assess the degree of substrate protein transport or membrane insertion: (i) the extent of signal peptide processing (for outer membrane protein OmpA), (ii) cell fractionation (for periplasmic proteins maltose-binding protein (MBP) and GBP) or (iii) the extent of cytosolic biotinylation (for membrane insertion of a MalF-PSBT chimera containing a the
amino-terminal portion of the maltose permease transmembrane subunit fused to a biotinylation domain that resides on the periplasmic side of the inner membrane in its membrane-inserted state). Ten different anhydrotetracycline concentrations were utilized for this purpose in order to vary SecYEG protein levels within the appropriate dynamic range. In addition, incremental titration of the inducer concentration was necessary in order to obtain a good comparison between the wild-type and secYE mutant, given different SecY protein levels at a given inducer concentration.

We first looked at the translocation of the major outer membrane protein OmpA in our strains, since this is one of the best characterized substrates for the Sec system. Signal peptide processing was used to measure transport in this case, since a significant portion of the substrate polypeptide chain must be transported across the plasma membrane in order to reach the signal peptidase II processing enzyme, which is on the periplasmic side of the plasma membrane [246]. Comparison of the decrease in pro-OmpA levels with increasing inducer concentrations revealed a modest difference between the wild-type and secYE mutant (Figure 3.3A). However, most of this difference can be attributed to a difference in SecY levels at a given inducer concentration, since comparison of SecY-normalized cultures yielded approximately similar levels of pro-OmpA accumulation between the two strains (Figure 3.3A and B, see different shaded arrows that compare similar SecY levels between the two strains).
It has been reported recently that the number of active translocons in a highly efficient *in vitro* translocation system is dependent on the substrate protein utilized, since GBP precursor required twice the number of SecYEG heterotrimeric units as pro-OmpA did for its transport [184]. An obvious rationale for this observation is that translocation of GBP required a SecYEG dimer, while OmpA translocation occurred via a monomer. Therefore, we were interested in measuring MBP and GBP translocation efficiencies in our system. In this case cultures grown at different anhydrotetracycline concentrations were fractionated into periplasmic and cytoplasm-membrane fractions in order to separate the secreted MBP or GBP from their non-secreted precursor forms. As expected, the extent of MBP precursor accumulation within the cytoplasm-membrane fraction

**FIGURE 3.3 Translocation efficiency of OmpA.** (A-B) The wild-type (WT) and secYE mutant (YE) were grown in the presence of the indicated concentration of anhydrotetracycline (Tet) and selective antibiotics as needed until an $A_{600}$ of 0.3, when cells were harvested, and (A) OmpA or (B) SecY proteins were analyzed by western blotting as described in “Materials and Methods”. %P/O is the percentage ratio of pro-OmpA divided by total OmpA (pro-OmpA plus OmpA). SecY levels were normalized.
using the value of the wild-type culture at 10 ng/ml anhydrotetracycline that was arbitrarily set to 100%. The filled and open arrows indicate wild-type and secYE mutant pairs that have similar SecY levels. The two images shown in panel B are from the same western blot.

was inversely related to the amount of anhydrotetracycline added, and MBP was found almost entirely within the periplasmic fraction at the higher inducer concentrations tested (particularly for the wild-type strain) (Figure 3.4A). When MBP accumulation was plotted as the ratio of MBP in the cytoplasm-membrane fraction over total MBP (C/T) versus inducer concentration, it was clear that the secYE mutant had a higher MBP accumulation ratio than the wild-type strain for a given inducer concentration (Figure 3.4B). However again as for OmpA substrate, when this data was corrected for SecY protein levels, both strains had similar C/T ratios, indicative of similar MBP secretion efficiencies (Figure 3.4B and C, see different shaded arrows that compare similar SecY
FIGURE 3.4 Translocation efficiency for MBP and GBP. The wild-type (WT) and secYE mutant (YE) were grown in the presence of the indicated concentration of anhydrotetracycline (Tet) and selective antibiotics as needed and (A-C) 0.2% maltose
or (D-F) 0.2% galactose until an $A_{600}$ of 0.3, when each culture was harvested and fractionated into total cell (T), cytoplasmic-membrane (C) and periplasmic (P) fractions, which were analyzed for (A) MBP or (D) GBP by western blotting as described in the “Materials and Methods”. The ratio of (B) MBP or (E) GBP in the cytoplasmic-membrane fraction to total cell fraction (C/T) was quantified and plotted against the anhydrotetracycline concentration. The filled, open, or hatched arrows indicate C/T ratios for the wild-type and secYE mutant with similar SecY levels. The same cultures used in (C) panel A or (F) panel D were harvested and analyzed for SecY levels by western blotting. For ease of comparison, band intensity was normalized using the (C) 10 ng/ml or (F) 13 ng/ml anhydrotetracycline concentration of the wild-type strain that was arbitrarily set at 100%. The two images shown in panel C or F are from the same western blot.

levels between the two strains). We performed a similar analysis for GBP translocation, and we obtained similar results in this case as well (Figure 3.4D-F). Thus we were unable to find a substantial difference in the translocation efficiency between the wild-type and secYE mutant strains over a dynamic range of SecY levels for three different secreted substrates: OmpA, MBP, and GBP.

Bacteria like *E. coli* generally utilize the co-translational pathway starting off with SRP and SRP receptor for the Sec-dependent insertion of integral membrane proteins into the plasma membrane, while the SecA and SecB-dependent post-translational pathway is generally utilized for the secretion of periplasmic and outer membrane proteins [33]. Because of the fundamental differences between these two pathways, we were interested in studying the insertion efficiency of an integral membrane protein substrate in our system. In order to do so, we made use of a previously described fusion between the first 712 residues of MalF permease and a biotinylatable domain from *Propionibacterium shermanii* (PSBT) (Figure 3.5A) [243]. During normal Sec-dependent assembly of this chimera into the plasma
membrane, the PBST domain is rapidly transferred to the periplasmic side of the membrane in its non-biotinylated state. By contrast, if the biogenesis of this chimera is delayed (e.g. in sec-defective mutants), then it remains within the cytoplasm long enough to become biotinylated by biotin ligase. Therefore, the extent of MalF-PSBT-I biotinylation can be used as a direct read out for a defect in its SecYEG-dependent membrane insertion efficiency.

We compared the membrane insertion efficiency of the MalF-PSBT-I chimera in the wild-type and secYE mutant at a range of different anhydrotetracycline concentrations. The extent of biotinylation of this substrate was normalized to the presumed constant level of biotinylation of the endogenous biotin carboxyl carrier protein (Figure 3.5B). The extent of biotinylation of the chimera clearly diminished for both strains when going from lower to higher inducer concentrations as expected. Comparison of the membrane insertion efficiency between the two
FIGURE 3.5 Membrane insertion efficiency of MalF-PSBT-I. (A) A membrane topology map of MalF protein redrawn from Jander et al. [243] is shown with the PSBT domain fused to MalF after isoleucine 712. (B-C) The wild-type (WT) and secYE mutant (YE) harboring the MalF-PSBT-I-producing plasmid were grown in the presence of 1.6 nM biotin and the indicated concentration of anhydrotetracycline (Tet) and selective antibiotics as needed until an $A_{600}$ of 0.3, when cells were harvested and analyzed for (B) biotinylation of MalF-PSBT-I or (C) SecY protein by western blotting as described in “Materials and Methods”. Biotinylation of MalF-PSBT-I was quantified and normalized to the endogenous biotin carboxyl carrier protein (BCCP) and expressed as a percentage ratio of the two proteins (%M/B). SecY levels were normalized using the 13 ng/ml anhydrotetracycline concentration of the wild-type strain that was arbitrarily set at 100%. The two images shown in panel C are from the same western blot. The colored arrows indicate %M/B for the wild-type and secYE mutant with similar SecY levels.

strains for similar SecY levels, however, revealed an interesting pattern: the wild-type strain displayed a somewhat reduced efficiency of MalF-PSBT-I membrane insertion (i.e. higher level of biotinylation) compared to the secYE mutant at low SecY levels,
but this pattern was reversed at higher SecY levels, where the wild-type strain was clearly more efficient in membrane insertion of the chimera than the mutant (Figure 3.5B and C, see different colored arrows that compare similar SecY levels between the two strains). It is possible that the apparent under-performance of the wild-type versus mutant at low SecY levels may simply be due to higher production of the MalF-PSBT-I protein, given its higher growth rate and presumed higher ribosomal content (which should be proportional to growth rate [247]). Since we were unable to obtain antibodies against MalF protein, we were unable to address this possibility. The clearer result here is that the secYE mutant displayed an approximately two-fold reduction in membrane insertion efficiency for this inner membrane substrate compared to the wild-type strain at more normal SecY levels where the growth difference between the two strains largely disappears (Figure 3.2F).

**SecA foot-printing analysis.** Since the results of our analysis showed that the wild-type and secYE mutant had similar translocation efficiencies per unit SecY protein for the three substrates examined, we felt that it was important to confirm that the translocons of the secYE mutant where indeed functioning as monomers by an alternative method besides disulfide crosslinking (Figure 3.1E) [179]. Therefore, we turned to our previously documented SecA foot-printing method for this purpose. Utilizing site-specific in vivo photo-crosslinking, we previously established that the SecA N-domain (NBD1 and NBD2) and C-domain (PPXD, HSD, HWD, and CTL) bind to distinct SecYEG protomers within the dimer [182]. In order to devise a SecA foot-printing assay for substrate-engaged translocons, we previously fused the
translocon jamming substrate OmpA-GFP to the carboxyl-terminus of SecA to make a SecA-OmpA-GFP trimera [179, 239]. This trimera allowed us to visualize the substrate-engaged, SecYEG protomer (through a disulfide crosslink between cysteines at the end of the channel-inserted OmpA signal peptide (Cys21) and the SecY plug domain (Cys68) on the periplasmic side of the membrane) along with the second SecYEG protomer that is bound to the SecA motor domain (through a photocrosslink between ρ-benzoyl-phenylalanine (pBpA) at SecA residue 59 with a C-H target on SecY protein) (Figure 3.6A). Thus we reasoned that if SecYEG functioned as a monomer during the translocation arrest in the secYE mutant, then we would see only the disulfide-induced crosslinked species where the trimera was covalently linked to a single SecY copy. By contrast, if SecYEG dimers were present during the translocation arrest in the secYE mutant, then we would see the double crosslinked species where the trimera was covalently linked to two SecY copies.

We compared the wild-type and secYE mutant strains producing the SecA-OmpA-GFP trimera by either photo-crosslinking, disulfide crosslinking, or double crosslinking. We found that both strains produced a UV or copper phenanthroline-dependent higher molecular weight species that corresponded to the SecY crosslinked trimera by western analysis. Furthermore, when the double crosslinking procedure was performed, we observed a super shifted species that corresponded to the trimera crosslinked to two copies of SecY protein (Figure 3.6B and C). This result clearly suggests that translocons are functioning as dimers in the secYE mutant, at least during
translocon jamming, which has been utilized to mimic the translocationally active state [179]. Since it remained possible that the SecA-OmpA-GFP trimera may have served to artificially recruit or tether normally separate SecYEG monomers into a dimeric state in the secYE mutant, we sought another independent method to assess the SecYEG oligomeric state.

FIGURE 3.6 SecA-foot-printing of wild-type and secYE mutant. (A) A cartoon of SecA foot-printing, where binding of the SecA N-domain with the non-translocating SecY protomer (SecY1) or the SecA C-domain with the translocating SecY protomer (SecY2) is depicted along with the substrate jamming OmpA-GFP chimera attached to the C-terminus of SecA and inserted into the SecY channel. Sites for photocrosslinking of the SecA N-domain (depicted by an X at residue 59) with SecY1 or for
disulfide crosslinking of the OmpA signal sequence (C21) to the plug domain (C68) of SecY2 are shown. (B-C) The wild-type (WT) and secYE mutant (YE) containing pTet2-SecYEG with the relevant secY and secE alleles including secY Cys68, pBAD-SecA-OmpA-GFP with secA Amber 59 and ompA Cys 21, and the pBpARS6TRN-cloDF plasmid were grown in the presence of 1 mM pBpA and 30 ng/ml anhydrotetracycline and selective antibiotics as needed until an A600 of 0.15, when arabinose was added to a final concentration of 0.2%, and cultures were grown for an additional 50 min. Cells were harvested and treated with UV, CuPh, and DTT as indicated, and analyzed by western blotting using (B) GFP antibody or (C) anti-SecY peptide antibody as described in “Materials and Methods”. The positions of the trimera singly crosslinked to SecY (SecA-OmpA-GFP-Y) or doubly crosslinked to SecY (SecA-OmpA-GFP-YY) are indicated by a solid square or triangle, respectively, and the trimera and SecY are indicated by the filled circle and diamond, respectively.

**In vivo photo-crosslinking analysis.** We have previously characterized the FtF and BtB interfaces of the two SecYEG dimers by in vivo photo-crosslinking [239]. We decided to utilize this approach again in order to test for the presence of dimers within the secYE mutant, since the more promiscuous pBpA-driven photo-chemistry would be more invariant than disulfide crosslinking to minor perturbations of the dimer interfaces arising from the arginine substitutions. Irradiation of the pBpA incorporated residue at 360 nm converts it into a diradical that covalently attaches the donor to a nearby C-H acceptor, which are relatively abundant within protein complexes [203]. By contrast, the precise positioning of engineered cysteine pairs used to detect particular SecYEG dimer states would be more readily perturbed by arginine substitutions at the dimer interfaces, leading to a potentially false-negative result.

We engineered secY amber mutations into the wild-type or secYE mutant at hot spots for detection of the FtF (residue 154 in P2 or residue 111 in TM3) or BtB (residue 20 in TM1 or residue 434 in C6) dimer or in regions distal from either interface as
negative controls (residues 369 and 376 in TM9) [239], and then performed an in vivo photo-crosslinking experiment. To our surprise, we found that the secYE mutant was as equally capable of forming both the FtF or BtB dimer as the wild-type strain (Figure 3.7A). Control residues away from either interface were negative for photo-crosslinking, consistent with the idea that the original interfaces were being used for dimerization. This result implies that the arginine substitutions within SecY or SecE proteins didn’t effectively monomerize the translocon as reported [179], but rather, they caused sufficient perturbation of the respective interfaces to render the cysteine reporters incapable of efficient disulfide crosslinking and thus dimer detection. To make absolutely certain of our results, we analyzed the same cultures used for photo-crosslinking to verify that the plasmid-borne secY and secE genes contained solely the correct mutations and that the chromosomal secY and secE null alleles were still present (data not shown), thus ruling out some sort of genetic mutation or rearrangement in the mutant during the course of this experiment.
**FIGURE 3.7** *In vivo* photo-crosslinking of wild-type (WT) and *secYE* mutant (YE). (A) Strains containing pACYC-SecYEG with the indicated *secY* and *secE* alleles and the pBpARS6TRN-cloDF plasmid were grown in the presence of 1mM pBpA and selective antibiotics as needed until an $A_{600}$ of 0.3, when cells were harvested and subjected to *in vivo* photo-crosslinking, followed by isolation of membranes, and analysis of SecY protein by western blotting as described in “Materials and Methods”. SecY monomer (Y) and dimer (YY) bands are indicated. The *secY* wild-type strain BW25113 was included to indicate the normal chromosomal level of SecY ($Y^{WT}$). (B) Similar to panel A except strains were grown to an $A_{600}$ of 0.2, when the cultures were split, and the portion shown received kasugamycin at a final concentration of 2 mg/ml (Treated with Ksg), and all cultures were incubated for an additional 10 min before being analyzed by *in vivo* photo-crosslinking. (C) The wild-type (WT) and *secYE* mutant (YE) containing pACYC-SecYEG with the indicated *secY* and *secE* alleles, pBpARS6TRN-cloDF, and pBAD-lacZ plasmid were grown to an $A_{600}$ of 0.2, when the cultures were split into four parts, and kasugamycin was added to three of them at a final concentration of 2 mg/ml at time 0. Arabinose was added to a final concentration of 0.2% at the indicated times, and 40 min later β-galactosidase activity was measured. The average Miller units [230] of β-galactosidase activity of duplicate samples are shown. The basal level of β-galactosidase activity in the untreated culture lacking both kasugamycin and arabinose was subtracted as background from all assays.
The above *in vivo* photo-crosslinking experiment indicates that both the FtF and BtB dimers were present at an approximately normal level in the *secYE* mutant, but it did not distinguish whether such dimers preexist in their translocationally inactive state or whether they form only upon engagement with substrate. In order to distinguish between these two possibilities, we made use of the translation initiation inhibitor kasugamycin, which can be used to halt the production of substrate proteins to the Sec pathway, thereby allowing us to examine only the translocationally-inactive SecYEG population. Kasugamycin inhibits translation initiation by perturbing the mRNA codon-tRNA anticodon interaction, thus preventing the binding of the tRNA\(^{\text{Fm}}\) to the P-site on the 30S ribosome subunit [248]. We showed previously that pre-treatment of *E. coli* strains with 2 mg/ml of kasugamycin for 2 min was sufficient to prevent >99% of subsequently induced β-galactosidase activity [239].

Comparison of the kasugamycin treated and untreated control cultures showed that both the wild-type and *secYE* mutant contained similar levels of both the BtB and FtF dimers despite antibiotic treatment, which abolished >99% of translation based on β-galactosidase activity assays run on these cultures (Figure 3.7A-C). We conclude that the arginine substitutions within SecY or SecE did not perturb the ability of the *secYE* mutant to make approximately normal levels of resting FtF or BtB dimers, respectively. Thus regrettably, the previously described *secYE* mutant that is presumably identical to our strain does not effectively monomerize the SecYEG translocon based on two independent assays (*in vivo* photo-crosslinking and SecA
foot-printing), and thus additional studies are now required to attempt to make the requisite mutant in order to determine if the SecYEG monomer is active or not in \textit{in vivo} protein transport.

5. DISCUSSION

Earlier studies have been equivocal as to whether the SecYEG complex functions in protein transport as a monomer or dimer or alternatively as part of some sort of monomer-dimer translocation cycle. While most structural studies have provided us with a picture of a ribosome or SecA-bound monomer engaged or not with substrate that should be sufficient for protein transport or membrane protein bilayer insertion [37, 91, 138, 141, 237, 238], exceptions to this pattern exist where a ribosome-nascent chain complex bound to a SecYEG dimer has been described [38]. Furthermore for at least one x-ray structure, it was reported that the second SecYEG protomer was lost during purification of the relevant SecA-SecYEG complex [91]. Regrettably the purification requirements and conditions required for x-ray crystallography or cryo-EM studies severely hamper our ability to properly address this question. Likewise, biochemical studies have produced similar conflicting conclusions and are difficult to interpret as well, since the SecYEG monomer-dimer equilibrium appears to be dynamic in nature and subject to changes in response to interactions with SecA protein, adenosine nucleotides, substrate proteins, acidic phospholipids or the size of nanodiscs employed in such studies [44, 177, 180, 214, 240, 241, 249].
It is with this backdrop that we previously initiated an \textit{in vivo} approach with the hope of clarifying this unresolved major question in a more physiological manner. While a pioneering study employing \textit{in vivo} cysteine crosslinking suggested that the pool of FtF and BtB dimers dissociated once they engaged with an OmpA-GFP substrate used to create a stable translocation intermediate [179], we obtained an opposite result in this case. In particular, we found that both dimer pools were maintained during translocon jamming based on both \textit{in vivo} photo-crosslinking and disulfide crosslinking, where an OmpA-GFP associated SecY dimer was enriched in immunoprecipitates with GFP antibody compared to the substrate-associated monomer [239]. In addition, an \textit{in vivo} SecA foot-print of SecY protein demonstrated that the two halves of SecA engaged distinct protomers of the SecY dimer that was jammed with substrate. Thus while our prior study strongly supported an \textit{in vivo} role of the SecYEG dimer in ongoing protein transport, it did not address participation of the monomer in the transport process.

Thus the report of a strain defective in formation of both FtF and BtB dimers potentially opened the way to an \textit{in vivo} investigation of SecYEG monomer function [179]. The viability of such a mutant, if properly confirmed, would demonstrate that the monomer is indeed functional on its own, and such a strain would be invaluable for both \textit{in vivo} and \textit{in vitro} studies characterizing monomer activity. Indeed, utilizing arginine substitution at the FtF and BtB dimer interfaces along with appropriate cysteine reporters, we reconfirmed that both dimer states were essentially undetectable
in vivo by disulfide crosslinking (Figure 3.1). We next proceeded to study the growth and protein secretion physiology of the resulting secYE mutant compared to the isogenic wild-type strain under limiting secYEG expression levels (Figures 3.2-3.5). Our finding that the secYE mutant showed approximately similar translocation efficiencies for the three substrates studied, OmpA, MBP, and GBP, lead us to reexamine the extent of its dimerization defect by two alternative methods, namely in vivo SecA footprinting and in vivo site-specific photo-crosslinking (Figures 3.6 and 3.7). Both experiments clearly showed that the secYE mutant still contained similar levels of FtF and BtB dimers compared to the wild-type strain. Thus we conclude that both dimer interfaces are still relatively intact in this mutant, but perturbed sufficiently by the arginine substitutions in order to prevent disulfide crosslinking of the cysteine reporters utilized to detect each dimer state. Accordingly, the present study serves as a cautionary tale for the need to avoid overreliance on a single site-specific methodology with relatively fixed geometries like cysteine crosslinking when assessing dynamic protein-protein interactions even at relatively stable oligomeric interfaces. In that regard, the more promiscuous photo-crosslinking chemistry afforded by pBpA provides a more reliable readout of protein-protein interactions in such cases.

The perturbation of cysteine crosslinking by the engineered arginine substitutions is not unexpected in retrospect. In the case of SecY TM3, the two arginine substitutions are far enough from the membrane center so as to have a less energetic cost for bilayer occupancy, and the five neighboring aliphatic amino acid substitutions should restore the stability of this transmembrane helix within the bilayer (Figure
However, it is likely that the flexible arginine side chain with its charged guanidinium group would tend to snorkel into the polar region of the bilayer where it could multiply hydrogen bond with polar phospholipid head groups. Alternatively or additionally, the arginine substitutions are likely to cause deformation of the phospholipids and increased solvation resulting in membrane thinning. In either scenario the position of SecY TM3 within the bilayer could certainly be shifted, resulting in a concordant shift of the adjacent P3 domain containing the two cysteine reporters. Such shifts in transmembrane helices caused by even a single arginine substitution have been documented previously. In the case of the achondroplasia mutation in the transmembrane domain of fibroblast growth factor receptor 3, the most common form of human dwarfism, a neutron diffraction study showed that the relevant residue, G380R, was shifted from ~6 Å to ~11 Å from the bilayer center [251]. In the case of SecE TM3, even a slight shift in this helix due to the single arginine substitution could have significant long-range conformational effects, since this helix is transverse to the plane of the membrane and makes multiple contacts with SecY helices by forming the “hinge” at the back of the translocon complex. This view would explain why the relevant cysteine reporter residue within the P3 domain of SecY was negative for crosslinking in this case even though it was >20 Å from the arginine substitution within SecE TM3.

Given the present failure to adequately disrupt the two SecYEG dimer interfaces, it is worth asking whether a more systematic approach toward this end is likely to be successful or not. An extensive analysis of small molecule inhibitors of
protein-protein interactions over the past two decades indicates that relatively small hot spots of interaction, often near the center of the interface, occur that contribute most of the binding energy for protein complex stabilization (for a recent review, see [252]). Thus this literature suggests that by pursuing a more thermodynamically-targeted approach with an appropriate number of amino acid substitutions, it may well be feasible to disrupt the two dimer states while preserving the ability of SecYEG protein to properly fold and integrate into the plasma membrane. However, given the difficulty in locating key residues responsible for dimerization of membrane proteins (for a recent review, see [253]), alternative strategies should be considered as well. For example, it is conceivable that a small membrane protein or hydrophobic domain could be appropriately inserted into or adjacent to one of the SecY or SecE transmembrane segments at the relevant dimer interface that would sterically block dimerization. Such strategies are not without risk, given the potentially essential nature of one or both dimer forms. In that regard the genetic system utilized here that employs plasmid shuffling into a secYE chromosomal null background should be useful to determine whether one or both of the dimer states are indeed essential for cell viability or not. An alternative approach might be to engineer a cysteine residue(s) at the relevant dimer interface and to utilize a bulky sulfhydryl adduct that could be added to growing cells to sterically block dimerization in a conditional manner.

While our study hopefully clarifies the present state of the field somewhat more, we are still left trying to make sense of a conflicting set of literature on SecYEG monomer-dimer function. In that regard such studies would make sense on a collective
level if the monomer was the minimal functional unit of the translocon, while the dimer provided enhanced translocon function, perhaps necessary for sustaining the rapid \textit{in vivo} transport rates for the large variety of different protein substrates present in the modern bacterial secretome. Given the dimeric nature of the SecA footprint on SecY, where its N-domain (which contributes most to the initial SecY binding affinity [232]) interacts with the non-translocating SecYEG protomer, while its C-domain interacts with the adjoining translocating SecYEG copy, it seems likely that the dimer may have evolved to provide a sufficient platform for the various substrate protein targeting and transport factors like SecA as suggested previously [214]. In this context there might be significant flexibility between the two SecYEG protomers as long as they each are able to bind their appropriate factor(s) with the correct geometry. In the case of SecA, there is inherent flexibility between its N- and C-domains that could accommodate a modest shift in the alignment between two SecYEG protomers making up the dimer [254]. Such a model could help to explain why the arginine substitutions at the dimer interface were relatively tolerated on a functional level in our study. In fact, given all of the transport and assembly factors that need to be accommodated in the modern bacterial holo-translocon, it is reasonable to speculate that two dimer forms may have evolved in order to accommodate a wider diversity of accessory factors. These and other speculations on the functional oligomeric state of the bacterial translocon will have to await future studies aimed at addressing these important questions.
Chapter IV

Discussion

The SecYEG or Secαβγ translocon is at the heart of the Sec translocation pathway and is responsible for the biogenesis of over 95% of the secretome. Since the advent of the first high resolution crystal structure of Secαβγ from *M. jannaschii*, greater detail of its function has been revealed. However, whether it works as monomer or dimer, or alternates between the two states has been a great controversy. Most structural studies have suggested a monomer as the functional unit, but this has to be questioned because SecYEG can be artificially monomerized due to the unavoidable use of high detergent in those studies. Many biochemical, biophysical and genetic studies have resulted in a set of conflicting findings that support the monomer, dimer and even both as the functional unit.

Beyond these considerations, our work was triggered by the study of Park *et al* who showed that SecYEG dimer monomerized when it was arrested in a translocation-intermediate state by OmpA-GFP substrate, as evident by the loss of dimer-specific cysteine-crosslinking [179]. The experiment was done *in vivo*, thus its validity can’t be challenged as an *in vitro* artifact. However, one caveat is that SecYEG monomerization could also be explained by a SecA and substrate-induced conformational change in SecY that rendered those cysteine pairs incapable of forming a disulfide bond, and thus the dimer could still be involved in translocation. Due to this concern, chapter II of my thesis was dedicated to the study of the oligomeric state
of SecYEG in both translocating and resting states by using the less biased photo-crosslinking technique. In contrast to Park et al, my results indicated that the SecY dimer level was insensitive to global translocon jamming or translocon clearing. Our SecA-foot-printing and immunoprecipitation of the OmpA-GFP jammed SecY dimer complex were consistent with this finding and argues strongly that SecYEG dimer participates in active protein translocation. In fact, dimer was preferentially pulled down in the OmpA-GFP jammed complex, which is ironic because the DDM level used in solubilization of SecYEG prior to immunoprecipitation was similar to the level used in crystallization of the SecA-SecYEG complex by Zimmer et al., which resulted in SecYEG monomerization [91].

Chapter III of my thesis focused on the study of SecYEG monomer. This study was also inspired by Park et al who appeared to demonstrate successful creation of a monomer-biased secYE mutant, as also accessed by the loss of dimer specific cysteine-crosslinking [179]. This presented a good opportunity to study the function of SecYEG monomer in vivo. We created the identical secYE mutant in which arginines were engineered at the TM3 helix of SecY and SecE to disrupt the FtF and BtB dimers, respectively, and investigated the growth rate of the mutant and its translocation rate for secretory proteins MBP, GBP, RBP and OmpA, and the membrane protein MalF in comparison to the WT strain. Our study took into account the recent finding by Mao et al. who showed that GBP used twice the number of active translocon units compared to MBP, suggesting a mechanism in which substrate determines the usage of either monomer or dimer. We discovered that the secYE mutant was essentially as active as
the WT strain in translocating different substrates, when taking into account SecY expression level. This finding precipitated our re-investigation of the monomer-dimer status of the SecYEG translocon in the mutant strain. Further UV-crosslinking experiments showed that the secYE mutant formed approximately similar dimer levels compared to the WT strain in both active translocating and resting states. Additional reading through the literature suggested that introduction of positively-charged amino acids can cause more distortion in TM helices as opposed to dimer disruption. In fact, such substitutions can induce dimer formation because the presence of positive charges in the hydrophobic portion of the lipid bilayer is not stable, thus proteins tend to dimerize in order to hide the positive charges [255]. To truly create a monomer, we would have to scan through more mutations along the dimer interfaces, which is beyond the scope of this study. Our results proved once again that cysteine-crosslinking, because of its higher stereo-specificity, is not always the best method to study protein-protein interaction despite its robust crosslinking efficiency.

We also concluded that the attempted monomerization of SecY through the introduction of positively-charged amino acids is not an optimal choice as we have observed that the SecYE mutant protein wasn’t produced as well as the WT. This is less likely due to inefficient transcription and translation because the mutant was on the same plasmid, under the same Tet-driven promoter, and in the same strain background as the WT. The problem most likely came from the three extra arginines which hindered membrane integration and proper assembly. Even if some monomer-biased mutant was successfully constructed that inserted into the membrane and was
viable, there is a possibility that the difference in translocation function between the mutant and WT SecY is due to the mutations that compromised its function, rather than its monomeric state. If I had another chance to perform this study, I would probably set up *in vitro* protein translocation assay. For example, I would make four different types of proteoliposomes: the first one would contain an average of just one copy of SecYEG per liposome, the second one would contain two copies, the third one would contain a single copy of the covalently-linked SecY dimer, and the fourth one, which is the control, would contain WT SecYEG at the physiological concentration. These can be achieved by varying the lipid to protein ratio during the SecYEG reconstitution step. I would then perform *in vitro* protein translocation to see which one works better. To synchronize the reaction, I would add SecA and preprotein (such as proOmpA) first under the right buffer conditions and initiate the reaction with addition of ATP. The extent of translocation can be accessed by quantifying the amount of OmpA being protected from the subsequent protease digestion step (such as protease K). At the end, translocation activity will be normalized to the control. Comparing the first two proteoliposomes will enable us to determine whether having the extra copy of SecY which enables formation of dimer is important for translocation or not. Comparing the second and the third proteoliposomes will allow us to determine whether some kind of monomer-dimer equilibrium is essential for protein translocation or not.

There are other factors that may possibly influence whether SecYEG monomer or dimer gets used during protein translocation. For example, we could have also
investigated the changes in SecY oligomeric state at different stages of protein translocation, which probably include i) priming in which SecA or ribosome binding causes initial limited opening of the lateral gate; ii) translocation initiation in which the signal sequence or the signal-anchor sequence is inserted into the lateral gate, which opens up the channel even more; iii) translocation of the mature region into the periplasm or membrane; iv) completion of translocation in which protein substrates are released into their final destinations, followed by restoration of SecYEG back to its resting state. It is hard to arrest the channel at these distinct steps. Probably the best way to do it is to create an OmpA-GFP chimera with different lengths of mature sequence. One can then look at the formation of monomer and dimer when SecY is arrested at different stages of mature domain translocation.

The controversy surrounding the functional oligomeric state of SecYEG is not going to be resolved shortly. While my contribution to this ongoing debate is limited, my results did conclude one important point, that is, SecY dimer participates in a particular step of protein translocation. Due to the technical limitation of photo-crosslinking, or any other crosslinking technique, we were unable to discover anything about the dynamics of SecYEG monomer and dimer function during protein translocation. For this reason, we do not know whether the dimer initiates translocation and remains throughout the entire reaction, or if it monomerizes at a particular step. Frank Duong created a covalently linked SecYE dimer and showed that it could complement both the temperature sensitive and cold sensitive secY mutant E.coli strains [181]. This suggests that SecYEG dimer doesn’t need to monomerize anytime during the
translocation cycle, with the caveat that this covalent dimer could act as a tethered monomer. The same dimer was also proven to be functional during *in vitro* protein translocation. For example, the study by Osborn *et al* showed that one protomer of such a dimer forms the high affinity binding site for SecA, while the second protomer actually translocates protein substrate [214]. In fact, we also created the same covalently-linked dimer, and it was able to complement the growth of a *secYE* null strain. However, when we analyzed the dimer-encoding plasmid through restriction digestion, we found that its size was reduced to the size consistent with a plasmid that contains only one copy of SecY. This suggests that the SecY dimer-encoding plasmid was not stable. Further analysis confirmed that such a strain produced only SecY monomer by western blotting (result not shown). This result also suggests that the ability to monomerize and dimerize seems to be important for protein translocation and the SecY covalent dimer may not be acting as suggested by Duong.

The issue of SecYEG monomer and dimer function is challenging to address. While covalently-linked SecY is functional, one can argue that it functions as a tethered monomer at least during portions of the translocation cycle. Likewise, reports of monomer function could be explained by SecA bridging two monomers in order to utilize it as a dimer. There is no clear cut distinction between tethered monomers and a covalent dimer. Perhaps it is useful to distinguish three types of SecYEG dimers: the free dimer, the pre-translocation dimer and the translocation-engaged dimer. Free dimer forms as a result of collision between two SecYEG monomers in detergent or in a membrane bilayer. It forms in a SecYEG concentration-dependent manner. It may
be that a portion of such a dimer can have physiological function, for example, as a storage form. Free dimer can be exemplified by the FtF and BtB forms that we observed in the translocon-clearing state. When SecA, preprotein or other factors bind, it signals SecYEG to come together to form a pre-translocation dimer. The conformation of this dimer is not fixed and has a tendency to monomerize easily. The examples of this dimer are the ones that scientists have observed in vitro when partial translocation conditions are established. When SecA, protein substrates and ATP bind under favorable conditions (for example at 37°C in the presence of sufficient acid phospholipids), the translocation-engaged dimer forms in which two protomers make very specific contacts with one another, inducing conformational changes to facilitate protein translocation. The dimer that we captured using the SecA-foot-printing approach is likely belong to this type of dimer.

Another reason why this problem is challenging to address is that we don’t really know much about the SecYEG dimer structure other than the existence of the putative FtF or BtB dimers (e.g. there are no dimer x-ray structures available). However, these dimers can be found in the translocon-clearing state, and especially the BtB dimer, can also be found engaged with OmpA-GFP protein substrate in vivo, suggesting that it participates in active protein translocation. No one has ever tried to determine whether the covalently-linked, in vivo functional dimer constructed by Duong [181] forms the FtF or BtB dimer states. With no linker in between the two protomers, it would be hard for it to gain enough flexibility to form either dimer. This suggests that the covalent dimer may indeed function by association with the chromosomally-produced
SecYEG, to produce a non-covalent dimer with an extraneous covalently attached monomer. This would explain why the covalent dimer-encoded plasmid was not tolerated in the secYE null mutant. The FtF dimer could be advantageous when extra expansion of the channel is needed in order to accommodate translocation of larger, partially-folded substrates since the lateral gates can potentially fuse to form a composite dimer. One can imagine that opening of the FtF dimer would be tightly regulated so that it opens just enough to accommodate such substrates. In such a scenario, the pore ring residues of the two protomers would have to coordinate with each other so that they still form a tight seal around the translocating polypeptide. The BtB dimer appears to be the ideal form for translocation of membrane proteins since the lateral gate of each protomer is facing outwards towards the phospholipid bilayer. The extra copy of SecY could provide the high affinity binding site for SecA and enhance its translocation ATPase activity. There is limited evidence in support of either the FtF or BtB dimer as the translocation dimer. One study showed that nanodiscs that contained predominantly BtB dimer supported SecA translocation ATPase (a presumed requirement for translocation), suggesting that this particular form might be the translocation dimer [240]. Another important observation regarding the BtB dimer came from our study described in chapter II of this thesis, when OmpA-GFP jamming was shown to result in the pull down of the SecY dimer-substrate complex that was fixed in the BtB orientation by cysteine-crosslinking. In general, we observed a much stronger crosslinking signal for the BtB dimer compared to the FtF dimer, therefore, at least in our hands, the BtB dimer is the more prevalent form. The
number of crosslinking residues examined for each form suggests that this conclusion is not simply due to local effects of crosslinking efficiency.

If I had another chance, I would probably characterize Duong’s covalently-linked dimer further. The first thing that I would do is to make a recA- version of the secYE null strain described in Chapter III, and reintroduce the SecY dimer-encoding plasmid to see if it is stable and complements the null strain. Different substrates could be investigated just like we did in Chapter III. I probably would also study the function of the proposed FtF and BtB dimers. Since they can’t be made through in vivo synthesis, I would have to purify them first and artificially link them in vitro through previously established cysteine disulfide crosslinking. Then they could be reconstituted into proteoliposomes for subsequent in vitro protein translocation assays.

SecYEG dimer formation is influenced by factors such as SecA and preprotein binding [180]. In particular, SecA-enhanced formation of SecYEG dimer adds another layer of complexity to the story. This is because the SecA oligomeric state is as complicated and confused as the oligomeric state of SecYEG; it is not known whether SecA dimerization impacts SecYEG dimerization or not, and vice versa. SecA in the cytosol most likely exists as a dimer as previously mentioned. Binding of signal peptide was shown to monomerize SecA [110] or had no effect [95]. So whether substrate targeting occurs through SecA dimer or monomer is unclear. If monomer is enough to target substrates to the translocon, then the second SecA copy would bind at a later stage (if SecA works as dimer), which can potentially have affects on the SecYEG oligomeric
state. So, the oligomeric state of SecYEG during protein translocation is probably also influenced by how SecA targets substrates. The observation that translocation occurs in a step-wise manner suggests that SecA dissociates after ATP hydrolysis and rebinding of SecA permits translocation of a 2.5kDa segment of polypeptide in an ATP-independent manner [27]. Such dissociation and rebinding is likely to modulate the SecY oligomeric state as well. However, recently it was suggested that SecA undergoes multiple rounds of ATP hydrolysis before dissociation from the SecYEG channel and rebinding [190]; therefore modulation of SecYEG oligomeric state due to SecA rebinding may be limited. It is well known from crystal structures that SecA binding caused a marked opening of the SecY lateral gate. Therefore, if anything, binding of SecA to SecY during the initial targeting of substrate and rebinding to SecYEG later in translocation will likely cause certain conformational changes of the FtF dimer, since this dimer involves extensive contacts at the lateral gate between the two protomers.

Recently Banerjee et al. showed that parts of SecA inserted deeply into the SecY channel, with major contacts being made around the lateral gate and the pore ring. This is contradictory to the popular belief ever since the publication of the SecA-SecY crystal structure where the authors disputed the likelihood of SecA channel insertion [91]. Insertion of SecA into SecYEG will surely introduce major structural rearrangements in SecY which may not be thermodynamically favorable. However, the insertion of SecA, in particular at the lateral gate, can potentially provide extra helices required for channel expansion in order to accommodate translocation of
longer, potentially folded substrates. The portion of SecA that inserts and may form a temporal part of the channel is probably not the 2HF, since the 2HF has been heavily implicated in the power stroke model of protein translocation. Other regions of SecA can fulfill this job, for example, the HSD domain, which is a part of the original C-terminal 30kDa domain that was shown to insert into the membrane in a translocation dependent manner [28]. If this is true, then probably there is no need to form the FtF dimer to accommodate larger substrates.

There are also many other unknowns in the field that are less pertinent to the issue of SecYEG oligomeric state. For example, a polytopic integral membrane protein with a large periplasmic domain requires SecA for translocation; it is not known how SecA coordinates with the ribosome to translocate such proteins. Therefore, the pathways for co-translational and post-translational translocation shown in Figure 3 of the introduction are probably not mutually exclusive. SecA and the ribosome bind to the same cytoplasmic loop regions on SecY, therefore they can’t bind at the same time. However, SecA is also known to bind near the peptide exit tunnel of the RNC complex. So perhaps when SecYEG-bound RNC translates the periplasmic domain, some signals in the polypeptide, perhaps the mature targeting sequence will recruit SecA to bind near the peptide exit tunnel. In this state, SecA probably can also bind to the substrate through its patch A binding pocket. The original tight binding between the ribosome and SecY must be loosened somehow to allow room for SecA binding. After translocation of the periplasmic domain, the hydrophobicity of the next signal anchor
sequence would presumably disengage SecA and replace it with SRP and translation-coupled translocation would resume.

Other unknowns regarding the co-translational translocation mode include the mechanism by which the GTPase activity of SRP and its receptor FtsY causes transfer of the RNC complex to SecYEG translocon. FtsY is in contact with SecY, which makes it convenient for such transfer to occur. But how the GTPase activity is coupled to dissociation of RNC from SRP, subsequent binding to SecY, and insertion of the signal-anchor sequence into the lateral gate is unknown. Furthermore, how folding of both membrane protein and secretory protein is coordinated with the translocation process is largely known: does protein folding starts once enough polypeptide segments have been translocated, or is folding delayed by the action of accessory proteins such as YidC or TRAM, so that once translocation is completed, folding can occur robustly? How do YidC and TRAM coordinate with SecY to facilitate membrane protein folding? With regards to the post-translational translocation route, although several studies have pointed to the SecA dimer described by Hunt et al. [101] as the physiological dimer, more studies are needed to confirm this viewpoint. It is also not completely established how SecB or TF delivers preprotein to SecA. It is known that the preprotein-bound SecB has higher affinity for SecA [81], and the basis of SecA-SecB binding, in part, is the electrostatic interaction between the positively charged amino acids on the zinc-containing domain at the extreme C-terminus of SecA and the negatively charged amino acids on the solvent-exposed side of β-sheet of SecB [75]. SecA-SecB interaction was also found to be promoted by the extreme C-terminal
α-helix of SecB and the N-terminal interfacial region of the SecA dimer [256]. But the exact mechanism of preprotein transfer from SecB to SecA is not known. A crystal structure of the SecB-preprotein-SecA complex could address this question. The mechanism of SecA-driven protein translocation is also still largely unknown. Among the many models discussed previously, the Power Stroke and Brownian Ratchet models have been the most well established. Other models require coordinated actions of SecA and/or SecYEG dimers, both of which are still controversial in their own right. The role of other translocon subunits is also not well known. Recent crystal structures led scientists to propose that SecDF binds to the translocating polypeptide from the trans-side of membrane and pulls it towards the periplasm [161], possibly providing a net forward movement since backward sliding can occur even in the power stroke model (e.g. when ATP is hydrolysed to ADP). More studies are needed to establish and confirm the model of SecDF-enhanced translocation. Last but not least, very little is known about the function of YajC as it is the smallest and non-essential subunit of the Sec translocon. Its binding to, and regulation of SecDF production, suggest an interesting and potentially novel regulatory paradigm.
REFERENCES


28. Economou, A., et al., *SecA membrane cycling at SecYEG is driven by distinct ATP binding and hydrolysis events and is regulated by SecD and SecF*. Cell, 1995. 83: p. 1171-1181.


129. Economou, A., et al., SecA membrane cycling at SecYEG is driven by distinct ATP binding and hydrolysis events and is regulated by SecD and SecF. Cell, 1995. 83(7): p. 1171-1181.
132. Kaufmann, A., et al., Cysteine-directed cross-linking demonstrates that helix 3 of SecE is close to helix 2 of SecY and helix 3 of a neighboring SecE. Biochemistry, 1999. 38: p. 9115-9125.
152. Murphy, C.K. and J. Beckwith, Residues essential for the function of SecE, a membrane component of the Escherichia coli secretion apparatus, are located


181. Dalal, K., et al., Two copies of the SecY channel and acidic lipids are necessary to activate the SecA translocation ATPase at the dimeric SecYEG translocase. Embo J, 2003. 22: p. 4375-4384.


197. Allen, W., et al., *Two-way communication between SecY and SecA suggest a Brownian ratchet mechanism for protein translocation.* Elife, 2016. 5: p. 15598.


