5-2009

The Conformational Gymnastics of the Escherichia Coli SecA Molecular Machine and its Interactions with Signal Sequences

Jenny Lynn Maki
University of Massachusetts Amherst, makijen@gmail.com

Follow this and additional works at: https://scholarworks.umass.edu/open_access_dissertations
Part of the Biological and Chemical Physics Commons, and the Cell Biology Commons

Recommended Citation
https://scholarworks.umass.edu/open_access_dissertations/77

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
THE CONFORMATIONAL GYMNASTICS OF THE E. COLI SECA MOLECULAR MACHINE AND ITS INTERACTIONS WITH SIGNAL SEQUENCES

A Dissertation Presented

by

JENNY LYNN MAKI

Submitted to the Graduate School of the University of Massachusetts Amherst in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2009

Molecular and Cellular Biology
THE CONFORMATIONAL GYMNASTICS OF THE E. COLI SECA MOLECULAR MACHINE AND ITS INTERACTIONS WITH SIGNAL SEQUENCES

A Dissertation Presented

by

JENNY LYNN MAKI

Approved as to style and content by:

Lila M. Gierasch, Chair

Danny J. Schnell, Member

Scott C. Garman, Member

Jeanne A. Hardy, Member

David J. Gross, Interim Director
Molecular and Cellular Biology Program
DEDICATION

This dissertation is dedicated to my grandparents David (Butch) & Harriette Hulstrom and Viljo (Val) & Taimi Maki. Although they are no longer with us, their love and support is continuously with me and enables me to achieve any goal that I set for myself.
ACKNOWLEDGMENTS

Although this dissertation is in my name, I never could have completed this process without the help from a tremendous number of people. First, I need to express my gratitude to my advisor Lila Gierasch. She is an extraordinary scientist and amazing mentor. She has a gift of knowing when to let me struggle and when to give a gentle ‘nudge’ in the right direction. I want you to know that it has been a great privilege to work with you over the five and a half years. I have learned so much from you, probably much more than you realize. This work would have been impossible without you.

I also need to thank my other dissertation committee members, Danny Schnell, Scott Garman, and Jeanne Hardy. I could not have put together a better committee than these hard working professors. During my first committee meeting, they helped change the direction of my research and this made the project much better scientifically as well as more interesting. In addition, my committee always had great ideas and comments to help push this work to be the best. Thank you for all of your help.

I have also had the privilege to work with a great collection of lab mates. Joanna Swain was instrumental during my early graduate career. I am grateful for every question that answered for me as well as the advice she gave to me in dealing with people when problems arose. Beena Krishnan has been always been there to discuss experiments and results (as well as teach me about her culture and food). She has also had the privilege (or problem) of having her desk right next to mine and has had to put up with all of my singing in lab. We have laughed a lot during the last four years and that has helped me keep my sanity during this process. Aneta Szymanska helped me tremendously with the cross-linking experiments as well as teaching me how to fix the HPLC. Eugenia Clerico
has been my ‘control Queen’ because she always knows the right controls that are needed for every experiment. Thank you for also being my fellow member of the Signal Sequence subgroup even when you move on to the Molecular Chaperone subgroup. 

*Annie Marcelino*, my fellow grad student, always gave me copies her prospectus, committee meeting updates, data defense, and dissertation for formatting purposes. She is also one of the sweetest people that I have ever had the opportunity to know as well as a great friend. *Zoya Ignatova* is a wonderful scientist and friend. I thank her for all of her insight and her motivation to be a better scientist. I thank *Linda Rotondi* who taught me how to use and run the HPLC even though at times it has been the bane of my existence. I also thank my other lab mates for all of their support: current members – *Ellen Kalt, Qinghua Wang, Jiang Hong, Hare Sahoo, Ivan Budyak, Anastasia Zhuravleva, Rob Smock, T.J.Brunette, Mylene Ferrolino, Mangai Periasamy, Elliot Ethridge, Roy Jung, Dennis Kim, Dave Paquette, Ervin Pejo, and Dima Steblovsky* as well as previous lab members *Ken Rotondi, Marc Vogt, Elena Falkovskaia, Gizem Dinler, Virginie Sjoelund, Mathias Stotz, Linnea Freeman, Rob McCloughlin, and Jeff Bombadier* as well as many others. I thank you for making the lab a great place to work.

In addition to the people with whom I worked with on a daily basis, I also need to thank other labs on the UMass campus. I thank *Steve Eyles* from the UMass Mass Spectrometry Center and *Jennifer Normanly* and *Dan Kita* for help with 2-D gels. I also thank the *Weis Lab* with whom we shared the 6th floor and equipment. In addition to specific people, I also want to acknowledge everyone from the Molecular & Cellular Biology Program, the Chemistry, Molecular Biology & Biochemistry Departments, and the Chemistry and Biology Interface Program for enriching my scientific experience.
Thanks to Don Oliver from Wesleyan University who generously gave the lab the plasmid containing the SecA gene and the region-specific SecA antibodies. I need to thank Babis Kalodimos and his lab at Rutgers University for great SecA discussions and for hosting the ‘mini’ SecA meeting in the summer 2006. Giving a talk in front of some of the biggest names in the SecA field was the one of the most intimidating and nerve-racking experiences during graduate school. It was also one of the best because it gave me the confidence to be able to speak about my project in front of anyone.

Finally, I need to thank my family for all of their love and support during graduate school. Even though my parents, Tom and Bonnie Maki, do not understand the science, they are always more excited than I am after each milestone has been completed. My brother Todd Maki always invited me to visit him to get away from the lab and my niece Kate Maki for making me laugh. I thank my beautiful cat Nosey for just being his adorable self. In addition to my biological family, I also have my ‘East Coast’ family. The Froio family has taken me in as another family member. I would like to especially thank Skip and Ronnie Froio for all of the food, fun, and laughter and also for opening up their home to me. Michael Moore has helped me move several times and has been exceedingly generous with dinners at nice restaurants and trips to different places. Lastly, I need to thank Richard Froio. There really are not sufficient words to thank you. I don’t know if I would have survived graduate school without your emotional (and sometimes financial) support. No one else has the same sense of humor as we do (ie Now listen here Billy Bob and listen good…) so thank you for all of your craziness that made me feel sane during the insanity of graduate school.
ABSTRACT

THE CONFORMATIONAL GYMNASTICS OF THE E. COLI SECA MOLECULAR MACHINE AND ITS INTERACTIONS WITH SIGNAL SEQUENCES

MAY 2009

JENNY LYNN MAKI

B.S., UNIVERSITY OF MINNESOTA DULUTH

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Lila M. Gierasch

Protein secretion is a selective and regulated process that is essential in all organisms. In bacteria the preprotein translocase SecA, either free in the cytosol or associated with the SecYEG translocon, recognizes and binds most post-translational secretory proteins containing an N-terminal signal sequence. In Gram-negative bacteria, the molecular chaperone SecB binds many of the preproteins to keep them in a translocation-competent state. Subsequently, SecB delivers the preproteins to the translocon-associated SecA, which binds the signal sequence and also interacts with mature regions of the preprotein. After the preprotein/SecA/SecYEG complex has formed, the energy derived from ATP hydrolysis by SecA coupled with the proton motive force drives the insertion of the preprotein through the translocon pore. During the translocation reaction, the conformation of SecA dramatically changes from an inactive closed form (c-SecA) to one more active and open states. The various crystal structures of SecA have provided many structural details about c-SecA. The recent low resolution
crystal structure of a fragment of SecA bound to SecYEG (Zimmer et al., 2008) has provided a starting point for structural analysis of the active and open conformation of SecA. Previous work in our laboratory demonstrated that an N-terminal proteolytic fragment of SecA, SecA64, is an activated form of SecA that with higher affinity signal peptides better than c-SecA (Triplett et al., 2001). To correlate the SecA64 results with full-length SecA, we determined that SecA in the presence of low concentrations of urea has an enhanced ATPase activity similar to translocation level, which is comparable to what was observed with SecA64. Analysis by CD and Trp fluorescence indicates the presence of an intermediate at 2.2 M urea at 22°C (termed u-SecA). Using limited proteolysis, we determined that u-SecA is in an protease-sensitive conformation that mimics the translocation-active form of SecA. These structural rearrangements occur primarily in the C-terminal one-third of the protein. Next, we sought to understand the signal sequence interactions with c-SecA and translocation-active u-SecA. Using a photoactivatable cross-linking approach along with limited proteolysis, two-dimensional gels, and domain mapping with region-specific antibodies, the signal sequence-binding site was mapped to the interface of NBF II, PPXD, and HSD. The site is the same in both forms of SecA but in our data suggests u-SecA that the binding groove as expanded.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1. General overview of protein secretion</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Signal sequences</td>
<td>2</td>
</tr>
<tr>
<td>1.3. Protein translation and chaperones</td>
<td>10</td>
</tr>
<tr>
<td>1.4. Secretory pathways in prokaryotes</td>
<td>15</td>
</tr>
<tr>
<td>1.4.1. Type I, III, and IV secretion systems</td>
<td>15</td>
</tr>
<tr>
<td>1.4.2. Type II secretion systems</td>
<td>17</td>
</tr>
<tr>
<td>1.4.2.1. Bacterial Sec pathway</td>
<td>17</td>
</tr>
<tr>
<td>1.4.2.2. Tat pathway</td>
<td>18</td>
</tr>
<tr>
<td>1.4.2.3. YidC</td>
<td>21</td>
</tr>
<tr>
<td>1.5. Secretory pathways in eukaryotes</td>
<td>22</td>
</tr>
<tr>
<td>1.5.1. Eukaryotic Sec pathway</td>
<td>22</td>
</tr>
<tr>
<td>1.5.2. Protein import into mitochondria and chloroplasts</td>
<td>23</td>
</tr>
<tr>
<td>1.6. Protein targeting and translocation in bacteria</td>
<td>27</td>
</tr>
<tr>
<td>1.6.1 SRP</td>
<td>27</td>
</tr>
<tr>
<td>1.6.2. SecA</td>
<td>28</td>
</tr>
<tr>
<td>1.6.3. SecYEG Translocon</td>
<td>29</td>
</tr>
<tr>
<td>1.7. Preprotein translocase SecA</td>
<td>33</td>
</tr>
<tr>
<td>1.7.1. SecA interaction partners</td>
<td>34</td>
</tr>
<tr>
<td>1.7.1.1. SecA dimerization</td>
<td>34</td>
</tr>
<tr>
<td>1.7.1.2. ATP binding and hydrolysis</td>
<td>34</td>
</tr>
<tr>
<td>1.7.1.3. SecB interactions</td>
<td>38</td>
</tr>
<tr>
<td>1.7.1.4. SecYEG binding</td>
<td>38</td>
</tr>
</tbody>
</table>
1.7.1.5. Interactions with phospholipids..........................40
1.7.1.6. Preprotein binding..............................................41

1.7.2. SecA crystal structures........................................42
1.7.3. Different conformations of SecA................................46
1.7.4. Signal sequence interactions with SecA........................46

1.8. Statement of Dissertation ........................................47

2. CHARACTERIZATION OF TRANSLOCATION-ACTIVE SECA ....49

2.1. Introduction ..................................................................49
2.2. Results ........................................................................57

2.2.1. Activation of SecA in low concentrations of urea...........57
2.2.2. Structural analysis of intermediate .............................57

2.2.2.1. Tryptophan Fluorescence ....................................57
2.2.2.2. Far-UV CD ..........................................................62
2.2.2.3. Limited proteolysis ..............................................67
2.2.2.4. Mass Spectrometry of proteolytic fragments ............74

2.2.2.5. Oligomeric state of u-SecA

2.3. Discussion...................................................................80
2.4. Materials and Methods ................................................88

2.4.1. Recombinant DNA experiments................................88

2.4.1.1. Cloning of SecA gene..........................................88
2.4.1.2. Site-directed mutagenesis.....................................88

2.4.2. Protein expression and purification ............................89
2.4.3. Enzyme-coupled ATPase assay ................................90
2.4.4. Far-UV CD ..............................................................91
2.4.5. ATPase assay in urea ..............................................92
2.4.6. Equilibrium fluorescence measurements ........................92
2.4.7. CD measurements in urea .......................................92
2.4.8. Limited proteolysis ................................................93
2.4.9. Antibody detection ................................................93
2.4.10. Mass spectrometry ..............................................94
2.4.11. Velocity analytical ultracentrifugation .......................95
2.4.12. Gluteraldehyde cross-linking ..................................96

3. SIGNAL SEQUENCE INTERACTIONS WITH CYTOSOLIC SECA AND
TRANSLOCATION-ACTIVE SECA...........................................97
### 3.1. Introduction

3.2. Experimental strategy

3.3. Results

- 3.3.1. Establishing the validity of the SecA-signal peptide interaction
- 3.3.2. Analysis of SecA fragments cross-linked to signal peptide

3.4. Discussion

3.5. Materials and Methods

- 3.5.1. SecA titration assay
- 3.5.2. Biotin detection
- 3.5.3. Protein staining and visualization
- 3.5.4. Biotin signal quantification and normalization
- 3.5.5. Signal peptide competition assay
- 3.5.6. Labeling of signal peptide with maleimido-benzophenone
- 3.5.7. BPM cross-linking experiments
- 3.5.8. Protease digestion and TCA precipitation
- 3.5.9. 1-D gels
- 3.5.10. 2-D gels
- 3.5.11. 2-D gel data analysis

4. SIGNIFICANT FINDINGS AND FUTURE DIRECTIONS

- 4.1. Significant findings
- 4.2. Future directions

  - 4.2.1. Structure of translocation-active SecA bound to preprotein
  - 4.2.2. Preprotein interactions after release from SecB
  - 4.2.3. SecA/preprotein interactions in the cytosol

Bibliography
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The Sec secretion pathways</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>General features of signal sequences</td>
<td>7</td>
</tr>
<tr>
<td>1.3</td>
<td>Structure and function of trigger factor</td>
<td>14</td>
</tr>
<tr>
<td>1.4</td>
<td>The Tat and YidC secretion pathways</td>
<td>20</td>
</tr>
<tr>
<td>1.5</td>
<td>Protein import pathways in mitochondria and chloroplast</td>
<td>26</td>
</tr>
<tr>
<td>1.6</td>
<td>The SecYEG translocon</td>
<td>30</td>
</tr>
<tr>
<td>1.7</td>
<td>SecA interacts with several ligands</td>
<td>35</td>
</tr>
<tr>
<td>1.8</td>
<td>Different structures of SecA</td>
<td>44</td>
</tr>
<tr>
<td>2.1</td>
<td>Different conformations of SecA</td>
<td>52</td>
</tr>
<tr>
<td>2.2</td>
<td>Seven Trp residues in SecA</td>
<td>55</td>
</tr>
<tr>
<td>2.3</td>
<td>Low concentration of urea stimulate the ATPase activity of SecA</td>
<td>59</td>
</tr>
<tr>
<td>2.4</td>
<td>Changes in SecA tertiary structure</td>
<td>61</td>
</tr>
<tr>
<td>2.5</td>
<td>Secondary structural changes in SecA in urea</td>
<td>64</td>
</tr>
<tr>
<td>2.6</td>
<td>Activity and structural changes in SecA in urea</td>
<td>66</td>
</tr>
<tr>
<td>2.7</td>
<td>Limited proteolysis demonstrates u-SecA is in an open conformation</td>
<td>69</td>
</tr>
<tr>
<td>2.8</td>
<td>SecA region-specific antibodies</td>
<td>71</td>
</tr>
<tr>
<td>2.9</td>
<td>Different domain proteolysis in c-SecA and u-SecA</td>
<td>73</td>
</tr>
<tr>
<td>2.10</td>
<td>Mass spectrometry determined protease cleavage sites</td>
<td>76</td>
</tr>
<tr>
<td>2.11</td>
<td>AUC of c-SecA and u-SecA</td>
<td>79</td>
</tr>
<tr>
<td>2.12</td>
<td>Cross-linking demonstrates u-SecA is a monomer</td>
<td>82</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2.13</td>
<td>Model of the conformation change in SecA upon binding SecYEG</td>
<td>87</td>
</tr>
<tr>
<td>3.1</td>
<td>Model for SecA-mediated translocation</td>
<td>99</td>
</tr>
<tr>
<td>3.2</td>
<td>Proposed signal sequence binding regions</td>
<td>102</td>
</tr>
<tr>
<td>3.3</td>
<td>Experimental design and data analysis</td>
<td>110</td>
</tr>
<tr>
<td>3.4</td>
<td>Map of chymotrypsin cleavage sites</td>
<td>113</td>
</tr>
<tr>
<td>3.5</td>
<td>Signal peptide binding assay</td>
<td>116</td>
</tr>
<tr>
<td>3.6</td>
<td>Signal peptide competition assay</td>
<td>120</td>
</tr>
<tr>
<td>3.7</td>
<td>Limited protease digestion of signal peptide cross-linked to c-SecA</td>
<td>124</td>
</tr>
<tr>
<td>3.8</td>
<td>Limited protease digestion of signal peptide cross-linked to u-SecA</td>
<td>126</td>
</tr>
<tr>
<td>3.9</td>
<td>Bio-KRRLamB19C-MBP cross-linking to both c-SecA and u-SecA</td>
<td>128</td>
</tr>
<tr>
<td>3.10</td>
<td>2-D gel biotin detection of signal peptide cross-linked to c- and u- SecA</td>
<td>131</td>
</tr>
<tr>
<td>3.11</td>
<td>Analysis similar and different fragments in c-SecA and u-SecA</td>
<td>133</td>
</tr>
<tr>
<td>3.12</td>
<td>Antibody detection of 2-D gel of c-SecA cross-linked to signal peptide</td>
<td>136</td>
</tr>
<tr>
<td>3.13</td>
<td>Antibody detection of 2-D gel of u-SecA cross-linked to signal peptide</td>
<td>138</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

Portions of this chapter are part of a review paper written in collaboration with Eugenia M. Clerico and Lila M. Gierasch. (Clerico E.M., Maki, J.L., Gierasch, L.M. Use of synthetic signal sequences to explore the protein export machinery. Biopolymers (Peptide Science). 2008, 90 (3): 307-319.)

1.1 General overview of protein secretion

Secretion is a highly selective and regulated process that is essential for the survival of any organism. In all kingdoms of life, the genes that code for proteins are translated predominantly by cytosolic ribosomes, yet nascent polypeptide chains are targeted to a variety of locations throughout the cell. As newly synthesized proteins emerge from the ribosome, they are greeted by a host of chaperones and targeting factors that effectively partition the proteins that are cytosolic from those that need to be translocated across a membrane (Ullers et al., 2007). In general, proteins that are destined for secretion or integration into a membrane contain a zip code called a signal sequence, which targets the ribosome nascent chain complex (RNC) or the post-translational proteins to the secretory apparatus (von Heijne, 1998). These polypeptides bind to secretory components through specific interactions and are translocated across the membrane, which is typically the plasma membrane in prokaryotic organisms and the endoplasmic reticulum (ER) membrane in eukaryotic organisms. In bacteria, secretory proteins are either translocated across another membrane such as the outer membrane or are folded into their final native state in the periplasm. In eukaryotes secretory proteins are correctly folded in the ER and ultimately trafficked to their final destination such as the plasma membrane or an organelle (Clerico et al., 2008).
1.2 Signal sequences

The targeting of nascent polypeptides to their final non-cytoplasmic destination is a crucial process in both prokaryotic and eukaryotic organisms. The targeting of these proteins is usually mediated by an N-terminal signal sequence, which directs the post-translational protein or the RNC to the inner membrane in prokaryotes and the endoplasmic reticulum in eukaryotes (Figure 1.1) (Schnell and Hebert, 2003). Despite their universal functions in all realms of life, signal sequences lack primary sequence homology, even in proteins that are closely related. Although without primary structure homology, signal sequences share a common distribution of residue types and are typically 15 to 30 residues in length (Figure 1.2, A). The N-terminus portion of the signal sequence, the so-called n-region, has an overall positive charge and is variable in length (Gierasch, 1989). It has been suggested that the requirement for the positive charge at the N-terminus can be compensated by the addition of a longer hydrophobic core (Hikita and Mizushima, 1992). The middle part of a signal sequence called the h-region, is enriched several Leu, Ala, Met, Val, Ile, or Phe residues and therefore, is highly hydrophobic. This hydrophobic region is the hallmark of a signal sequence and consists of 7 to 13 residues. The hydrophobic segment is shorter than a membrane spanning helix but is longer than a typical hydrophobic segment found in globular proteins (von Heijne, 1985). The third region of the signal sequence, the c-region, is uncharged, polar, and contains the signal peptidase cleavage site, A-X-A where X is any amino acid (Gierasch, 1989). The signal sequence is typically cleaved during the translocation of the preprotein (von Heijne, 1990). These general features of signal sequences have likely been evolutionarily
**Figure 1.1 The Sec secretion pathways**

Secretory and integral membrane proteins are targeted to the endoplasmic reticulum (ER) membrane in eukaryotes and the plasma membrane in prokaryotes. SRP-mediated translocation: The membrane protein (pink) emerges from the ribosome (gray) and SRP (light blue) recognizes and binds the signal sequence (red). The nascent chain complex (RNC) and SRP are targeted to SRP receptor (dark blue) at the membrane. Further translation of the protein is arrested or retarded at this step in eukaryotes and the nascent chain is maintained in an unfolded conformation needed for translocation. The interaction between the SRP and the SRP receptor delivers the RNC to the Sec translocon (green) and the nascent protein crosses or is inserted into the membrane. SRP is released from it receptor to start another targeting cycle. SecA-mediated targeting: After translation by the ribosome, the mature regions of the secretory protein interact with the molecular chaperone SecB (orange) to prevent protein folding. Through signal sequence recognition, the SecB-preprotein complex associates with SecA (purple), which is either in the cytosol or associated with the plasma membrane. This complex binds to the SecYEG translocon and induces the release of SecB. SecA undergoes conformational gymnastics to facilitate preprotein translocation across the inner membrane. Adapted from Clerico et al. (Clerico et al., 2008).
optimized for the particular passenger protein and the role of the passenger protein in the cell (Gierasch, 1989).

Genetic studies in *E. coli* have demonstrated that defects in signal sequences can be compensated by suppressor mutations in another part of the signal sequence. Early genetic work on secretory proteins in *E. coli* performed by Emr and Silhavy (Emr and Silhavy, 1980) demonstrated that most point mutations in the signal sequence did not inhibit translocation but some drastic point mutations could block export. Many of the mutations that change translocation of a preprotein occur in the hydrophobic core. For example, if a residue in the h-region of the λ phage receptor (LamB) signal sequence was mutated to a negative residue export was inhibited. Export of LamB can be inhibited by the deletion of four residues from the h-region. The fact that the inhibition can be reversed by mutation of either Pro9 to Leu or Gly17 to Cys (Figure 1.2, B) (Emr and Silhavy, 1980). The importance of both the n- and h-regions has also been assessed using the alkaline phosphatase signal sequence. Mutant signal sequences with a negative charge at the N-terminus and a highly hydrophobic h-region can be secreted rapidly but a less hydrophobic h-region cannot be exported. Therefore, the n- and h-regions are both involved in for recognition by the transport machinery (Izard et al., 1996). These results demonstrated that the primary structure is not critical for targeting by the signal sequence but the positive n-region and even more so the hydrophobic core are very important in the correct targeting of preproteins (Gierasch, 1989).

Even though signal sequences lack primary sequence homology, they are interchangeable between proteins even among vastly different organisms. Proteins carrying signal sequences from a prokaryotic protein can be secreted in a eukaryotic cell
Figure 1.2  General features of signal sequences

A. The three regions of a signal peptide. Adapted from Chou. (Chou, 2003). B. The *E. coli* λ-receptor (LamB) series of signal sequences. The relative position from the start of the signal sequence is indicated by the numerals above the amino acid sequence. Deleted amino acids are denoted by the dash, inserted amino acids are indicated by bold letters, the position of amino acid substitutions are shown by single letters. Adapted from Triplett et al. (Triplett et al., 2001).
A.

<table>
<thead>
<tr>
<th>n-region</th>
<th>h-region</th>
<th>c-region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polar and positively charged residues (&lt;-10 a.a.)</td>
<td>Hydrophobic, high helical propensity, rich in Ala and Leu (-7-15 a.a.)</td>
<td>Polar and neutral residues. A-A box for cleavage site</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Helix-breaking proline or glycine</td>
</tr>
</tbody>
</table>

B. LamB series (*E. coli* λ-receptor):

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>LamBWT</td>
<td>M M I T L R K L P L A V A V A A G V M S A Q A M A</td>
</tr>
<tr>
<td>LamBΔ78</td>
<td>- - - -</td>
</tr>
<tr>
<td>LamBΔ78r1</td>
<td>- - - -</td>
</tr>
<tr>
<td>LamBΔ78r2</td>
<td>L - - - -</td>
</tr>
<tr>
<td>KRR-LamBWT</td>
<td>M M I T L R K R K L P L A V A V A A G V M S A Q A M A</td>
</tr>
<tr>
<td>KRR-LamB19C</td>
<td></td>
</tr>
</tbody>
</table>

C
and vice versa (Muller et al., 1982; Talmadge et al., 1980). This ability of signal sequences from different organisms to still direct protein secretion implies that the recognition mechanisms of secretory machinery have not changed greatly during evolution (Briggs and Gierasch, 1986).

Recent studies in mammalian cells have shown that signal sequences likely have a role in the efficiency of preprotein translocation. One example of this is the cell surface glycoprotein, prion protein (PrP). Incorrect targeting of this protein to the ER is caused by a slight inefficiency of the PrP signal sequence. This inefficiency leads to the generation and build-up of both a cytosolic form and an incorrect topologic form of PrP, leading to neurodegeneration (Hegde and Bernstein, 2006). If the PrP signal sequence is replaced with a more efficient signal sequence such as the one from prolactin, there is an increase in the amount of PrP in the correct topological form (Ott and Lingappa, 2004). These results suggest that signal sequences not only act as targeting sequences but also affect the efficiency of translocation.

Genetic work in a mutant E. coli strain demonstrated that the removal of four hydrophobic residues from the h-region of the LamB signal sequence caused a defect in export (Emr and Silhavy, 1983). In two pseudorevertant strains, LamB was successfully exported and genetic examination of these strains demonstrated that each had a secondary point mutation within the signal sequence. In one case, the Gly17 was replaced by Cys, LamBΔ78r1. In the other case, the Pro9 was replaced by Leu, LamBΔ78r2 (Figure 1.2, B). Therefore, the loss of four hydrophobic residues was somehow compensated by single residue changes. Since both Gly and Pro can act as helix breakers, it was proposed that the signal sequence must adopt α-helical conformation to function.
Our group tested the secondary structure hypothesis by using synthetic signal peptides in solution as well as in phospholipids monolayers and vesicles (Gierasch et al., 1985). In solution the LamB signal peptide adopted a largely random conformation as monitored by far-ultraviolet (UV) circular dichroism (CD) analysis. In the presence of polyfluorinated alcohols, which can mimic the membrane interior, the signal peptide become partially helical (Briggs and Gierasch, 1986). This work was extended further to examine the deletion and pseudorevertants mutants discovered by Emr and Silhavy (Emr and Silhavy, 1983). The deletion mutant remained predominately in a random coil conformation in buffer, sodium dodecyl sulfate (SDS) micelles, trifluoroethanol, and phospholipid vesicles. The wild type LamB signal peptide as well as the two pseudorevertants had random coil conformations in buffer but adopted a much more helical structure in the presence of an apolar environment (Briggs and Gierasch, 1986). These results are consistent with the hypothesis that signal sequences must adopt an $\alpha$-helical structure for function.

Further studies showed that some non-functional signal sequences adopt $\alpha$-helical structures in apolar environments meaning that the structural conformation itself is not sufficient for signal sequence function (McKnight et al., 1989). Therefore, our group used surface tensiometry experiments using phospholipids monomers to assess the interactions of signal peptides with phospholipids monolayers. The functional LamB signal peptides were able to insert into the monolayer but the non-functional signal peptides were only able to interact with the phospholipids head groups of the lipids. The non-functional signal peptides that can adopt $\alpha$-helical conformation could only insert into the monolayers at high peptide concentrations (Briggs et al., 1985; McKnight et al., 1989).
The same biophysical characteristics of the LamB signal peptide are observed in other functional and non-functional synthetic signal peptides such as the outer membrane protein OmpA (Hoyt and Gierasch, 1991a; Hoyt and Gierasch, 1991b; Rizo et al., 1993). At least in the case of the LamB signal sequence, the adjacent mature region of the preprotein does not affect the secondary structure or the membrane-binding properties of the signal sequence (McKnight et al., 1991). This implies that functional signal sequences not only need to be able to adopt an α-helical conformation but also need to be able to insert into the phospholipids membrane.

In spite of their diversity and lack of primary sequence homology, signal sequences have very distinct features and properties that specifically direct preproteins to the secretory apparatus. Signal sequences are interchangeable between species and are tolerant of many mutations. Recently, it has come to light that signal sequences are not just simple interchangeable domains that serve as simple targeting motifs but have specific properties that have been optimized for the proper biosynthesis of its substrate protein (Hegde and Bernstein, 2006).

1.3. **Protein translation and chaperones**

All proteins are synthesized by the decoding of the messenger ribonucleic acids (mRNA) by the ribosome. In *E. coli* the ribosome is made of 55 components and is formed by two subunits, 30S and 50S, which together form the 70S particle. This 70S ribosome consists of about two-thirds ribosomal RNA (rRNA) and one-third protein. The larger 50S subunit is made from two rRNAs, the 23S and the 5S, and 33 proteins. The
smaller 30S subunit is composed of the 16S rRNA and 21 ribosomal proteins (Kaczanowska and Ryden-Aulin, 2007).

The translation process can be broken down into three stages: initiation, elongation, and termination, and the initiation stage is thought to be the rate limiting in the translation process. The messenger RNA (mRNA) is recognized by the 30S subunit and initiation factors aid in the recognition of the Shine-Dalgrano sequence in the mRNA. This allows for the proper orientation of the start codon and then the 50S subunit assembles forming the 70S particle. During elongation aminoacylated transfer RNA (tRNA) corresponding to the codon in the mRNA binds the ribosome at the A site. The aminoacyl group is released by the elongation factor EF-Tu and the tRNA is moved into the P site. Another aminoaceylated tRNA corresponding to the next codon in the mRNA binds to the free A site. This positions the tRNAs so that peptide bond formation occurs. The deacylated tRNA is then repositioned from the P site to the E site and is eventually pushed into the exit tunnel of the ribosome. When a stop codon enters the A site, release factors bind and induce release of the peptide chain from the P site (Kaczanowska and Ryden-Aulin, 2007).

As nascent polypeptide chains emerge from the bacterial ribosome, they are greeted by two ribosome-bound chaperones, the signal recognition particle (SRP) and trigger factor (TF), and both bind the polypeptide chain co-translationally (Hoffmann et al., 2006). Proteins lacking signal sequences are recognized by TF, which can deliver the protein to another chaperone system such as GroEL (Kandror et al., 1995) or Hsp70 to assist in proper folding (Deuerling et al., 1999). Proteins containing signal sequences with long hydrophobic cores and transmembrane segments interact specifically with SRP
while TF binds less hydrophobic signal sequences to prevent SRP binding allowing for targeting the SecA post-translationally (Beck et al., 2000). A more in-depth description of SRP and SecA is presented in section 1.6.1. and 1.7., respectively. TF binds to stretches of eight amino acids that are enriched with basic and aromatic amino acids. The binding motif of TF is similar to the binding motif of another molecular chaperone, DnaK. Since these chaperones share a similar binding motif but are in different locations in the cell, they may be able to cooperate in protein folding (Deuerling and Bukau, 2004). The structure of TF revealed an unusual extended conformation that is made up of three domains (Figure 1.3, A). The ribosome-docking region is found in the N-terminal domain while the peptidyl-prolyl isomerase domain is at the other end of the structure. The C-domain is sandwiched between the other two domains and forms the “arms” of TF (Hoffmann et al., 2006). When bound to the ribosome, TF is hunched over the polypeptide exit tunnel, which shields the nascent polypeptide and allows for TF to capture the chain in its hydrophobic cradle (Figure 1.3, B). The binding of TF may delay folding to allow for more of the chain to emerge from the ribosome so that the correct amount of sequence is present for folding (Deuerling and Bukau, 2004).

In *E. coli* and many other Gram-negative bacteria, there is a dedicated secretion chaperone called SecB. There are several other molecular chaperones in bacteria such as GroEL, DnaJ, and DnaK but it is still unclear if these chaperones have any role in protein secretion. SecB can rapidly bind nascent chains still attached to the ribosome or to partially folded preproteins in the cytosol (van Wely et al., 2001). SecB associates and dissociates with polypeptide substrates without the use of ATP, which is in contrast to other chaperone such as GroEL and DnaK (Danese and Silhavy, 1998). The structural
Figure 1.3 Structure and function of trigger factor

A. Crystal structure (1W26) of soluble trigger factor color-coded for the different domains: red, the N-terminal ribosome-docking region, yellow, the peptidyl-prolyl isomerase domain, and green, the C-terminal domain. B. Schematic representation of trigger factor bound at the exit tunnel of the ribosome. Adapted from Hoffmann et al. (Hoffmann et al., 2006).
organization of SecB is a homotetramer that is a dimer of dimers. There is a long channel located on both sides of the tetramer that is proposed to be involved in the binding of the polypeptide chain. This groove has two subsites with one deep cleft lined with aromatic residues and one shallow groove, which is hydrophobic in nature. The orientation of the two grooves allows the polypeptide to wrap around the SecB tetramer. The basis for how SecB differentiates between cytosolic and secretory proteins is still unclear since the binding motif is about nine amino acid in length and enriched in both aromatic and basic residues, which is very similar to other molecular chaperones. SecB keeps the preprotein in an unfolded state for translocation and targets the complex to the preprotein translocase SecA (Driessen and Nouwen, 2008).

1.4. Secretory pathways in prokaryotes

The secretion of proteins in bacteria is an intricate dance that requires specific recognition and delivery of the correct substrates to the proper translocation machinery. Bacterial cells have several different types of secretion systems including outer membrane systems (Saier, 2006). The development of distinct systems for general secretion, type II represented by the Sec and Tat pathways, enable cell survival while the more specialized systems, type I, III, and IV, are involved in pathogenesis.

1.4.1. Type I, III, and IV secretion systems

Gram-negative bacteria have developed specialized membrane secretion systems. The type I, type III, and type IV secretory pathways can export proteins across both the inner and outer membranes in one energy-coupled step. The type II secretory pathways cross only the cytoplasmic membrane. The type I ATP-binding cassette (ABC)
macromolecular system generally consists of two integral membrane domains and two cytoplasmic domains. This ABC transporter can associate with the main terminal branch (MTB) and outer membrane factor (OMF) proteins to enable secretion across both membranes. This system can transport ions, drugs, and proteins of various sizes including pathogens outside of the cell (Saier, 2006).

Gram-negative pathogens have also developed two direct transport systems for the delivery of virulent proteins into the host organism to establish infection. The type III pathway assembles a needle complex and shares common ancestral relationship to the flagellar secretion system. The needle consists of a basal body structure that is assembled using the Sec pathway. This structure is followed by the needle filament assembly, which can interact with the host cell membrane and compromise its integrity. Once the needle has made contact with the host cell the virulent proteins are secreted into the host cell and begin to manipulate the host’s cellular processes (Cambronne and Roy, 2006).

The type IV secretion system has some similarities to the type III system but is ancestrally related to the conjugation systems in bacteria. This secretion system also delivers into the cytosol of the host cell various virulent proteins, which modify the host cell’s physiology. The type IV system is made up of multiple subunits that span both the inner and outer membranes in Gram-negative bacteria. It is composed of a secretion channel and a surface filament or proteins (Christie et al., 2005). It is still unclear how the type IV system compromises the host cell membrane but it is likely that a small conduit could support the transport of the unfolded substrates (Cambronne and Roy, 2006).
1.4.2. Type II secretion systems

1.4.2.1. Bacterial Sec pathway

The general secretory (Sec) pathway is one of the main secretory pathways in prokaryotes (Gierasch, 1989). Signal sequences target proteins to this pathway by binding either co-translationally to the signal recognition particle (SRP) or post-translationally to SecA (Vrontou and Economou, 2004). Both of these pathways consist of four steps: recognizing the preprotein, maintaining the unfolded state, targeting to the translocon, and translocating across the membrane (Figure 1.1) (Clerico et al., 2008). The Sec pathway translocates proteins in the unfolded state, thus, allowing for protein insertion into the cytoplasmic membrane, folding in the periplasmic space, and secretion across or integration into the outer membrane. The majority of the components of the Sec system were identified by genetic studies in *E. coli* and this system contains three major components: a protein targeting unit, a motor protein, and a membrane embedded protein channel (Natale et al., 2008). In the SRP pathway, the signal sequence is recognized co-translationally by SRP and the RNC complex is targeted to the SRP receptor (SR). The SRP/SR interaction delivers the RNC complex to the SecYEG translocon and continued translation of the polypeptide pushes it through the translocon. SecA, on the other hand, recognizes the signal sequence of the preprotein post-translationally. The molecular chaperone SecB keeps the preprotein in an unfolded conformation and brings it to SecA. SecA docks at the translocon and uses the energy generated from ATP hydrolysis to push the preprotein across the SecYEG translocon (Figure 1.1) (Clerico et al., 2008). In bacteria, the Sec pathway is responsible for the secretion of a majority of the extracellular proteins, which have a variety of functions needed for cell survival (Natale et al., 2008).
1.4.2.2. Tat pathway

Another major secretory pathway in bacteria is the Twin-arginine translocation pathway, often referred to as the Tat pathway (Natale et al., 2008). The biggest difference between the Sec and the Tat pathways is that the Tat-pathway translocates properly folded secretory proteins across the cytoplasmic membrane (Figure 1.4, A). The secretory proteins are targeted to the Tat pathway by N-terminal signal sequences that are similar to Sec signal sequences. Tat signal sequence are composed of both n- and h-regions and in some cases a c-region that is cleavable. The recognition element of a Tat signal sequence is the almost invariant twin-arginine motif at the n-and h-region interface. In general the motif pattern is Z-R-R-x-φ-φ where Z is any polar residue and φ is any hydrophobic residue. Additionally, a Tat signal sequence often contains a positive residue in the c-region to ensure that the substrate does not target to the Sec pathway. Many of the Tat substrates have cofactors added before translocation. Therefore, the Tat system is utilized to translocate proteins of redox pathways, anaerobic respiration, and the global nitrogen cycle (Natale et al., 2008) as well as for the establishment of infections in both plants and animals (Lee et al., 2006). Some Tat substrates do not contain cofactors and may have specific folding reasons for utilizing this pathway. The Tat translocase consists of TatA and TatC or TatA, TatB, and TatC, which are all integrated membrane proteins (Natale et al., 2008). Protein that are transported by the Tat pathway are involved in several cellular functions that are essential to the survival of the cell (Lee et al., 2006).
Figure 1.4 The Tat and YidC secretion pathways

A. Model of the targeting and transport in the Tat pathway. Embedded in the membrane is the SecYEG translocon in green, TatA in red, TatB in blue, and TatC in yellow. (1) The nascent polypeptide chain emerging from the ribosome avoids targeting to the Sec pathway by the presence of the twin Arg (RR) residues. (2) The preprotein folds into its final conformation and associates with any needed cofactors and/or additional subunits before targeting to the TatBC receptor complex (3). (4) The formation of the complete translocase is driven by the proton motive force. The folded preprotein is transported through the pore, which is formed predominately by TatA. (5) The signal sequence is removed and the mature protein is released into the periplasm. Adapted from Lee et al. (Lee et al., 2006).

B. Model for YidC interaction with TM segments. The transmembrane (TM) segment of the protein leaves the SecYEG channel and associates with YidC. The first TM helix is released into the lipid membrane and the next TM segment binds YidC. YidC enables the integration of the TM segments during the translocation process. Adapted from Xie and Dalbey (Xie and Dalbey, 2008).
1.4.2.3. **YidC**

YidC is another translocase that is found in bacteria (Scotti et al., 2000) and organelles of bacterial origin such as the mitochondria (Hell et al., 2001). This protein can act as an independent insertase to insert proteins into the cytoplasmic membrane (Scotti et al., 2000). In addition to its own independent function, YidC can also cooperate with the Sec pathway machinery to aid in membrane protein insertion (Veenendaal et al., 2004). The independent function of YidC was discovered by its depletion from the membrane. Once YidC is no longer present, the insertion of Sec-independent proteins was inhibited (Samuelson et al., 2000). As an insertase, YidC plays a crucial role with the F$_{1}$F$_{0}$ ATPase (van der Laan et al., 2004) and cytochrome bo$_{3}$ oxidase (Stenberg et al., 2007). The role of YidC in these membrane-bound enzymes is likely the insertion of one of the subunits for the F$_{1}$F$_{0}$ ATPase or insertion of part of the enzyme into the membrane for cytochrome bo$_{3}$ oxidase (Figure 1.4, B) (Xie and Dalbey, 2008). Recently, a medium-level resolution map of YidC was determined by electron cryomicroscopy and revealed that YidC can form dimers in membranes (Lotz et al., 2008). It was proposed from this structure that the nature of the dimer interface could indicate the pathway, either Sec-dependent or Sec-independent, for insertion of membrane proteins. YidC may act as a monomer in the complex with SecYEG but act as a dimer for independent insertion of proteins.
1.5. **Secretory pathways in eukaryotes**

Although prokaryotes and eukaryotes are very different, they still share a similar co-translationally preprotein targeting system, SRP but have evolved different post-translational systems. Organelles of prokaryotic origin have import machinery that is different from the plasma membrane export machinery in prokaryotes indicating that these organelles have evolved their translocation machinery.

1.5.1. **Eukaryotic Sec pathway**

In eukaryotic cells, protein secretion and membrane integration of proteins is more complicated than in bacteria due to the presence of membrane-bound organelles such as: nucleus, endoplasmic reticulum (ER), golgi apparatus, lysosome, endosome, peroxisome, mitochondria, and chloroplast. Most of the polypeptides are initially targeted to the ER for modification and folding before being sorted to their final destination (Alder and Johnson, 2004). The preprotein must translocate across the ER membrane from the cytosol where it was translated by the ribosome. In co-translational translocation, the RNC complex is recognized by the SRP and is similar to the system found in prokaryotes. The SRP/RNC complex is targeted to the SRP receptor (SR) at the ER membrane. This directs an interaction between the ribosome and the Sec61 translocon embedded in the ER membrane. The elongating nascent polypeptide chain associates with the translocon and the continued translation of the polypeptide translocates the preprotein through the channel (Figure 1.1) (Rapoport, 2008).

Another subset of proteins is translocated post-translationally in eukaryotes but this is different from the bacterial SecA system. This post-translational translocation
involves the Sec61 translocon associating with the Sec62/63 protein complex. Initial recognition of the preprotein is mediated by Sec62/63, which subsequently transfers the preprotein to Sec61 (Schnell and Hebert, 2003). The translocation of the preprotein, at least in yeast, occurs through a ratcheting mechanism by interactions between the preprotein, Sec62/63, and BiP, an ER luminal Hsp70 chaperone. After a transient interaction with Sec63, BiP hydrolyzes ATP and closes its peptide-binding pocket on the translocation substrate. This prevents any backward sliding into the cytosol by the preprotein and enables forward movement into the ER lumen. After additional segments of preprotein enter the ER, another BiP molecule binds and the process is repeated until the entire preprotein has been translocated. Once ADP is exchanged for ATP, the BiP molecule is released from the translocated substrate (Rapoport, 2008).

1.5.2. Protein import into mitochondria and chloroplasts

Both mitochondria and chloroplast contain multiple membranes, which likely arose from an endosymbiotic relationship with prokaryotes. During the evolution, almost all of the genes from the endosymbiotic prokaryote were transferred to the nucleus of the host cell (von Braun and Schleiff, 2007). As a consequence N-terminal presequences or transit peptides are needed to target these proteins back to the correct organelle. Therefore, mitochondria and chloroplasts possess numerous translocons to translocate preproteins to the outer membrane, inner membrane, intermembrane space, and the internal soluble compartments (Schnell and Hebert, 2003).

The presequence of a mitochondrial protein directs it to the translocon of the mitochondrial outer membrane (TOM) (Figure 1.5, A). Tom20 and Tom22 act as the
receptor domains while Tom40 forms the channel. Tom40 is not only a channel but also recognizes the preproteins and aids in the sorting of preproteins to the correct compartment. The TOM complex is used for the import of all preproteins across the outer membrane (Kutik et al., 2007). Preproteins destined for the mitochondrial matrix or inner membrane are directed to the translocon of the mitochondrial inner membrane (TIM). There are two TIM translocons in mitochondria. One, Tim23, is predominately responsible for the translocation of matrix proteins and a subset of integral membrane proteins containing a single membrane-spanning domain (Schnell and Hebert, 2003). This translocation process needs the assistance of the presequence translocase-associated motor (PAM). Small proteins destined for the intermembrane space (IMS) are imported into the IMS by the mitochondrial intermembrane space assembly machinery (MIA). The outer membrane β-barrel preproteins are imported by TOM and transferred by chaperones to the sorting and assembly machinery (SAM). The proteins of the inner membrane translocate through TOM and are transferred to the other mitochondrial translocon, Tim22, for insertion into the membrane (Kutik et al., 2007).

The translocon of the outer membrane of the chloroplast (TOC) has three core components, Toc159, Toc34, and Toc75 (Figure 1.5, B). Toc159 and Toc34 are GTP-binding proteins located at the chloroplast surface and interact with the preprotein at the beginning of the import process while Toc75, which likely forms a β-barrel structure, constitutes the channel protein. Toc 159 is the primary receptor for preproteins and works with Toc75 to translocate preproteins across the chloroplast outer membrane. Most of the outer membrane proteins actually lack a transit peptide and are targeted by signals within their membrane anchor. Preprotein transport across the outer membrane requires both
Figure 1.5 Protein import pathways in mitochondria and chloroplast

A. Most mitochondrial proteins are synthesized in the cytosol by ribosomes and imported through the TOM complex. Matrix proteins containing presequences are transported by the TIM23 complex in conjunction with the motor protein PAM into the matrix. The mitochondrial processing peptidase (MPP) removes the presequence. The IMS small proteins are imported using the mitochondrial intermembrane space assembly machinery (MIA). Outer membrane (OM) precursors are transferred from TOM to SAM by Tim9-Tim10 chaperone. Inner membrane (IM) precursors are transferred to the TIM22 complex by Tim9-Tim10. TIM22 drives the insertion of these IM proteins into the inner membrane. Figure adapted from Kutik et al. (Kutik et al., 2007).

B. Core components of the chloroplast import pathway are shown in color. After translation by the ribosomes, preproteins are transported across the TOC complex, which consists of Toc159 and Toc34 as the receptors and Toc75 as the channel. Some preproteins are guided to the TOC complex by 14-3-3 protein and a Hsp70 chaperone. After translocation across the outer membrane, the preproteins interact with Hsp70 in the intermembrane space, which guides the preprotein to the TIC complex with assistance from Tic22. The TOC and TIC complexes associate and the preprotein crosses both membranes. The channel has been proposed to be formed by Tic110 and Tic20. The chaperones Cpn60 and Hsp93 help Tic40 fold the imported proteins. The transit peptide is cleaved from the preprotein by the stromal processing peptidase (SPP). Figure adapted from Kessler and Schnell (Kessler and Schnell, 2006).
ATP and GTP. Therefore, a Hsp70 has been proposed to function as the driving force for translocation while the Toc GTPases act as the regulators and receptors during initiation of translocation. The inner membrane translocon complex, the Tic components, has not been isolated in stoichiometric amounts. The Toc and Tic components begin to physically interact with each other during the early stages of translocation. This suggests that the inner membrane translocon is assembled in response to early interactions with Toc translocon. Two proteins form the proposed channel components for the inner membrane translocon, Tic110 and Tic20. It has been suggested that Tic110 is the docking site for the preproteins while Tic20 likely forms the membrane channel and is distantly related to channels of the mitochondrial inner membrane. Tic110 also associates with the molecular chaperones in the stroma and these chaperones provide the driving force the preprotein translocation across the inner membrane (Kessler and Schnell, 2006).

1.6. Protein targeting and translocation in bacteria

1.6.1. SRP

The signal recognition particle is a ribonucleoprotein, which is found in all three kingdoms of life. SRP was first identified in mammalian cells and later in prokaryotic cells. The \textit{E. coli} SRP is one of the simplest of the SRPs because it consists of only one protein, called Ffh for fifty-four homolog due to its homology to the mammalian SRP54 protein, and a 4.5S RNA (Luirink et al., 2005). The Ffh component of the \textit{E. coli} SRP is responsible for the binding of signal sequences. Ffh is made of three domains, the N domain, which interacts with the ribosome, the G domain, which has a GTP binding site and is also involved in binding SR, and the M domain, which has a high Met content and
interacts with the RNA (Clerico et al., 2008). The signal sequence likely binds to the M domain but also shows interaction with the NG domain (Cleverly and Giersch, 2002). Unlike its mammalian counterparts, which target both secretory, and membrane proteins to the ER, *E. coli* SRP is predominantly involved with the translocation of inner membrane proteins. SRP actually interacts with the ribosome near the exit tunnel and binds to the signal sequences of integral membrane proteins (Figure 1.1). The SRP targets the RNC complex to the SR in the plasma membrane. The interaction between SRP and SR mutually stimulates each other’s GTPase activity and the RNC complex is transferred to the adjacent SecYEG translocon. The GDP-bound SRP is released from the complex so it can start another cycle (Clerico et al., 2008).

1.6.2. **SecA**

SecA is a preprotein translocase found primarily in bacteria but is also found in organelles of prokaryotic origin such as the chloroplast. Not only do bacterial SecA recognize and many different secretory proteins but it also acts as a processive motor to translocate preproteins across the inner membrane (Clerico et al., 2008). After translation and release from the ribosome, secretory proteins are held in the unfolded state by the molecular chaperone SecB. This new complex of the preprotein/SecB is targeted to SecA either in the cytosol or at the membrane. SecA recognizes and binds the signal sequence, a mature region of the preprotein, and SecB. SecA also interacts with both the SecYEG translocon through specific interactions with SecY and the inner membrane. These interactions induce a dramatic conformational change in SecA, which results in the release of the chaperone SecB. Using the energy from ATP hydrolysis as well as the
proton motive force, SecA undergoes further conformational gymnastics to translocate the preprotein across the inner membrane through multiple rounds of binding and release of the mature portions of the preprotein. Once the preprotein has been completely translocated, SecA is released and can begin another translocation cycle (Driessen et al., 2001). A more in-depth description of SecA is presented in section 1.7.

1.6.3. SecYEG Translocon

The bacterial translocon is a heterotrimeric membrane channel composed of three proteins, SecYEG, while the eukaryotic counterpart to the channel is called Sec61 and is made from the α, β, and γ subunits. In prokaryotes the largest of the subunits, SecY, is the core component and has ten transmembrane regions (TM1-TM10), six cytoplasmic regions (C1-C6), and five periplasmic/luminal loops (P1-P5). SecE is a small membrane protein and has three transmembrane domains in which only one makes extensive contacts with SecY. SecG has two transmembrane domains (Collinson, 2005). The first crystal structure of the channel was solved at 3.2 Å and represents the closed channel. This SecYEβ structure was from the archaeabacterium Methanococcus jannaschii and showed that SecY has two domains TM1-TM5 and TM6-TM10 that have an inverted pseudo-symmetry (Figure 1.6) (Van den Berg et al., 2004). This structure suggests that one copy of SecY forms the channel pore and this is supported by mutations that compensate for defective signal mapping to the center of the SecY complex. The central pore is blocked on the extracytoplasmic side by the presence of a small helix called the plug, which is part of SecY transmembrane helix 2. In order for substrates to translocate across the membrane the plug is required to move out of the way. The shape of the
Figure 1.6 The SecYEG translocon

A. Side view of the SecY complex. The plug shown in green from helix 2a is held in the ‘plug-hole’ in the closed complex. Adapted from Collinson (Collinson, 2005). B. The crystal structure of *M. jannaschii* SecY complex (1RH5) seen from the cytoplasm. The plug shown in green blocks the pore in the closed conformation. Transmembrane (TM) 1-5 are in blue and TM 6-10 are in red. The bright blue helix is TM2b and SecE and Secβ are shown in gray. Adapted from Osborne et al. (Osborne et al., 2005).
channel is a hourglass lined with hydrophilic residues on both sides of the constriction point while at the point of constriction is several hydrophobic residues. This would provide a seal so that ions and other small molecules could not diffuse across the membrane during protein translocation. The pore must also widen to enable the passage of an unfolded polypeptide chain (Osborne et al., 2005).

In another study using single particle cryo-electron microscopy, the structure of a translocating ribosome bound to SecYEG was solved and contained two SecYEG complexes. In this structure, only structural features of 10 Å were observed. The orientation of the two SecYEG complexes was front-to-front where the lateral gates from the protomers face each other but do not form a larger channel (Mitra et al., 2005). Several lines of evidence suggest that SecYEG can exist in different oligomeric states but this oligomerization does not appear to significantly change the conformation of the heterotrimer. Another possibility is that translocation partners, such as SecA and the ribosomes, could modulate the oligomeric state of SecYEG (Rusch and Kendall, 2007).

Recently, two new crystal structures of the translocon were solved (Tsukazaki et al., 2008; Zimmer et al., 2008). SecYE from Thermus thermophilus bound to an anti-SecY Fab fragment was resolved to 3.2 Å resolution (Tsukazaki et al., 2008). Overall, this structural architecture was similar to the original M. jannaschii SecYEβ structure (Van den Berg et al., 2004) but the TM6-TM10 region has been altered creating a small opening and has been termed the ‘pre-open’ state. In a low-resolution structure of SecYE without the Fab fragment, the structure was similar to the closed form like the SecYEβ structure. The anti-SecY Fab fragment binds to a motif in C5 that is the same region where SecA has been demonstrated to bind. These structures suggest that the closed form
of SecYE is energetically favored when no interacting components are present but the ‘pre-open’ form is a conformation state induced by the binding of a ligand such as SecA (Tsukazaki et al., 2008).

The second crystal structure solved was of Thermotoga maritima SecYEG bound to Thermotoga maritima SecA, without its carboxyl-terminal domain. This four-protein structure was fitted to a 4.5 Å model. The N-terminal residues of SecY, SecE, and SecG, C-terminal residues of SecY and SecG, as well as the periplasmic loop between TM1 and TM2a of SecY were not resolved. The SecA-SecYEG complex was strongest in the presence of ADP-BeF₃, AlF₃, which mimics the intermediate states during ATP hydrolysis (Zimmer et al., 2008). Both of these structures likely represent the early stages of preprotein translocation.

1.7. Preprotein translocase SecA

SecA was discovered in the early 1980’s through genetic work. This protein plays a central role in the bacterial Sec pathway. SecA is only found in prokaryotes and organelles of prokaryotic origin such as the chloroplast. Like SRP, SecA selectively binds signal sequences of preproteins but recognizes them post-translationally and most of the preproteins are destined for the periplasm, the outer membrane, or extracellular secretion. In addition to signal sequence recognition, SecA also acts as a processive motor to translocate proteins across the inner membrane (Clerico et al., 2008).
1.7.1. SecA interaction partners

1.7.1.1. SecA dimerization

SecA is a large protein that interacts with several different ligands during the preprotein translocation reaction (Figure 1.7). This preprotein translocase is a 102 kD protein and in solution forms homodimers with a monomer-dimer equilibrium of 0.25 – 0.5 µM depending on protein concentration, salt concentration, temperature, and ligands (Woodbury et al., 2002). The bacterial cellular concentration of SecA is 5 µM suggesting that it is mostly in the dimeric form. However, studies have shown that SecA may function as a monomer during the translocation cycle (Or et al., 2005; Or and Rapoport, 2007; Duong, 2003) while other studies indicate that the SecA dimer stay intact during translocation (Jilaveanu et al., 2005; de Keyzer et al., 2005; Wang et al., 2008). A recent crystal structure of SecA bound to the SecYEG translocon indicates that SecA likely acts as a monomer during the translocation reaction (Zimmer et al., 2008).

1.7.1.2. ATP binding and hydrolysis

SecA is a processive motor enzyme that uses ATP hydrolysis to push preproteins through the SecYEG channel. Initial studies suggested two distinct ATP binding sites were available on SecA (Mitchell and Oliver, 1993; van der Wolk et al., 1995; den Blaauwen et al., 1996; van der Wolk et al., 1997; Miller et al., 2002). The first crystal structure of SecA showed that only one ATP binding site was present on SecA. The ADP or ATP binds SecA at the interface between nucleotide binding fold I (NBF I) and nucleotide binding fold II (NBF II) and stabilizes the SecA molecule in a compact conformation. The nucleotide binding folds share structural homology with DEAD-box
Figure 1.7 SecA interacts with several ligands

SecA crystal structure (1M6N) color-coded for the different domains: blue, NBF I; purple, NBF II; yellow, PPXD; green, HSD; orange, HWD; and red, CTL (see Figure 1.8 for more structural details). The nucleotide binding site has been precisely mapped but the other ligand binding sites are less clear. SecAβ indicates the second SecA protomer. Adapted from Vrontou and Economou (Vrontou and Economou, 2004).
helicases. The nucleotide binding site is formed by the helicase motifs. The Walker A consensus sequence binds the $\alpha$- and $\beta$-phosphates of the nucleotide by forming a P-loop helix-capping structure. The Walker B consensus sequence is a hydrophobic $\beta$-strand ending in an Asp residue, which contacts the Mg$^{2+}$ cofactor. The Glu from the DEAD-box is the likely candidate for the catalytic base in the ATP hydrolysis reaction (Hunt et al., 2002).

A subsequent study using NMR, thermodynamic, and biochemical techniques demonstrated that the regions lining the nucleotide binding cleft undergo transitions between disordered and ordered states that correspond to the functional catalytic states (Keramisanou et al., 2006). This disordered region was also previously observed in our lab by NMR (Chou et al., 2002). Furthermore, Arg 574 and Arg 509, both found in the helicase motifs, are important in enhancing the catalytic activity of SecA (Keramisanou et al., 2006).

During the translocation cycle, SecA goes from a low basal ATPase rate to eight-fold higher translocation ATPase rate (Vrontou et al., 2004). Truncation of the C-terminal one-third of SecA causes an unregulated, hyperactive ATPase rate (Karamanou et al., 1999; Triplett et al., 2001). Deletion of the intramolecular regulator of ATP hydrolysis 1 (IRA1) also causes the unregulated ATPase activity of SecA. In the closed and compact cytosolic form of SecA, IRA1 suppresses the ATPase rate of SecA but upon interaction with the membrane or the translocon, IRA1 is released and the ATPase rate of SecA is dramatically increased to help facilitate the translocation of the preprotein across the inner membrane (Karamanou et al., 1999).
1.7.1.3. SecB Interactions

In order for preproteins to be translocated across the inner membrane post-translationally by SecA in Gram-negative bacteria, the molecular chaperone SecB recognizes and binds preproteins to keep them in an unfolded state. SecB can then target them to the preprotein translocase SecA. SecB has a high affinity for the membrane-bound form of SecA ($K_D = 10^{-30} \text{nM}$) (Vrontou and Economou, 2004). SecB is a homotetramer and has three potential areas of interaction with SecA. One of the contact regions is between the negatively charge surface formed by each dimer of SecB and the zinc-binding domain at the extreme C-terminus of SecA. A second area involves the C-terminus of SecB and the interfacial region of the dimer (Patel et al., 2006) of SecA from the $B. subtilis$ structure (Hunt et al., 2002). The third area of interaction was discovered by truncation of both SecA and SecB, which eliminated the other binding interactions. This third interaction may not be stable enough to maintain association on its own. Even though SecB is a tetramer and SecA is a dimer, the complex formed is asymmetric which may have implication for preprotein translocation (Patel et al., 2006).

1.7.1.4. SecYEG Binding

Several years ago it was shown that SecA interacts with the SecYEG channel, and SecA was proposed to insert into the channel with the preprotein (Economou and Wickner, 1994). Only recently have the molecular details of the SecA-SecY interaction begun to be understood. Several indirect methods have been used to assess the SecA-SecYEG interactions (Snyders et al., 1997; van der Sluis et al., 2006; Karamanou et al., 2008) but they have not provided conclusive evidence.
Using an in-vivo site-directed cross-linking technique, Mori and Ito (Mori and Ito, 2006a) studied the SecA-SecY interaction by the incorporation of the photoactivatable cross-linker in the cytoplasmic loops of SecY. Cross-linking was observed with SecA to the C2, C4, C5, and C6 loops of SecY. The presence of NaN$_3$, a known powerful inhibitor of SecA translocation activity, enhanced cross-linking to C6 suggesting that this interaction is likely to be dynamic during the translocation reaction. The interactions with C2, C4, and C5 are more static in nature and the C5 loop appears to be involved in the binding of the N-terminal two-thirds of SecA (Mori and Ito, 2006a). The C5 residue Arg 357 was subsequently shown to be essential for the initiation of the translocation cycle by SecA (de Keyzer et al., 2007).

The interaction sites proposed by Mori and Ito (Mori and Ito, 2006a) were used to design disulfide cross-linking experiments with SecA. Cys substitutions were made in the C6 loop of SecY and the NBF I of SecA. The cross-links mapped one face of the NBF I to the tip of the C6 loop of SecY. In another experiment, two linked SecY proteins were used for cross-linking with SecA and a preprotein. The results indicated that one SecY channel interacted with the preprotein while the second one bound the NBFI of SecA. This suggests that SecYEG can function as a dimer with one active translocon for preproteins and a second translocon for binding SecA (Osborne and Rapoport, 2007).

The structure of SecA bound to the SecYEG translocon has recently been solved at 4.5 Å. This structure demonstrates that both SecA and the SecY complex undergo large conformational changes upon binding. In this structure only one SecA molecule is bound to one copy of the SecY channel. The majority of the interactions are made by the preprotein cross-linking domain (PPXD) of SecA and the loops connection TMs 8-9 and
TMs 6-7 of SecY. Additionally, the α-helical scaffold domain (HSD) makes extensive interactions with SecY. Other minor interactions occur as well but these have small contact surfaces and likely could not sustain the interaction alone. To correlate this structure (Zimmer et al., 2008) with the previous disulfide cross-linking results (Osborne and Rapoport, 2007), it was proposed that this new structure represents the active channel through which preproteins are translocated. The inactive channel that interacts with the NBF I was lost during detergent solubilization (Zimmer et al., 2008). Clearly, the SecA-SecY interaction is important in the translocation of preproteins and we are just beginning to understand some of the molecular details of this dynamic interaction.

1.7.1.5. Interactions with phospholipids

In addition to its interactions with other components of the secretory pathway, SecA can also interact and insert into the inner membrane. Protease protection assays in the presence of inverted membrane vesicles, ATP, and preproteins, demonstrated that a 30 kD fragment of SecA is protected from the protease. N-terminal sequencing revealed that this protected fragment started at residue 610 indicating that the 30 kD fragment is composed of the C-terminal third of the protein (Price et al., 1996).

Other experiments have demonstrated that long-chain phospholipids analogues cause SecA to monomerize. In addition to the change in oligomeric state, the phospholipids also cause domain dissociation in SecA (Benach et al., 2003). Acidic phospholipids are required for efficient translocation of preproteins by SecA. Fluorescence experiments with SecA and labeled phospholipids revealed that SecA can penetrate deep into the phospholipids bilayer. This also induces a partial unfolding or
domain dissociation of SecA that could be important for the translocation cycle (Ulbrandt et al., 1992).

1.7.1.6. Preprotein binding

The major function of SecA is the translocation of preproteins across the inner membrane. The preprotein consists of two parts, the signal sequence and the mature region and SecA interacts with both of these parts. Signal sequences can bind to SecA in solution as well as in the presence of liposomes (Vrontou and Economou, 2004). Several different regions in the N-terminal two-thirds of SecA have been proposed as the signal sequence-binding site in soluble SecA (Kimura et al., 1991; Triplett et al., 2001; Hunt et al., 2002; Baud et al., 2002; Chou and Gierasch, 2005; Papanikou et al., 2005; Gelis et al., 2007) as well as SecA bound to liposomes (Musial-Siwek et al., 2005; Musial-Siwek et al., 2007). Most of these observation or predictions have been made on the soluble form of SecA, which may not be the physiologically relevant form of SecA to study signal sequence interactions (Clerico et al., 2008). Another study indicated that the nucleotide state of SecA can modulate the affinity of the signal sequence (Shin et al., 2006). The recognition and binding of the mature portion of the preprotein is still poorly understood. The binding site for the mature regions of the preprotein should have some fundamental characteristics such as: binding to 20-30 amino acids at a time, no sequence specificity, and a small bound versus unbound energy barrier (Vrontou and Economou, 2004). One study demonstrated that the mature portion of a preprotein could interact with the PPXD of SecA (Papanikou et al., 2005). Despite several years of research effort and
the central importance of this question, preprotein interactions with SecA are still not clearly understood.

1.7.2. SecA crystal structures

The first crystal structure of dimeric SecA from *Bacillus subtilis* revealed the structural architecture of the monomer, which consists of five different domains (Figure 1.8, A). The catalytic activity of ATP hydrolysis is carried out by the first nucleotide binding fold (NBF I) in blue and the second nucleotide binding fold (NBF II) in purple. These two domains have a ‘RecA-like’ fold which creates the nucleotide binding site at the interface between the domains. Additionally, NBF I and II share structural homology with the superfamilies I and II ATP-dependent helicases. A unique domain in SecA is the preprotein cross-linking domain (PPXD) in yellow, which interrupts NBF I (Hunt et al., 2002) and has been demonstrated to bind preproteins (Vrontou and Economou, 2004; Papanikou et al., 2005). The α-helical scaffold domain (HSD) in green follows the NBF II and has been implicated as the conformational regulator of SecA (Mori and Ito, 2006b). The α-helical wing domain (HWD) in orange interrupts the HSD and can loosely pack against the PPXD but the function of this domain has not been determined. The carboxyl-terminal linker (CTL) in red at the C-terminus of the protein contains the zinc binding motif and has also been shown to interact with phospholipids as well as SecB. The last 40 residues at the extreme C-terminus were not resolved in the *B. subtilis* structure (Hunt et al., 2002) or any of the subsequent crystal structures.

The only structure of the extreme C-terminus of SecA is a solution NMR structure of the last 22 residues that includes the zinc binding motif (Dempsey et al., 2004). The
Figure 1.8 Different structures of SecA

Three different crystal structures of SecA demonstrate that the overall domain organization is similar but the preprotein cross-linking domain has three different orientations. A. The *B. subtilis* structure (1M6N) has the PPXD in a closed form with intramolecular interactions with the HWD. The last 40 residues are missing from this structure (Hunt et al., 2002). B. The NMR structure of *E. coli* SecA (2VDA) shows that the PPXD has rotated away from the HWD and is in the open conformation. This structure is missing the most of the CTL (Gelis et al., 2007). C. SecA structure from *T. maritima* bound to the SecYEG translocon (3DIN). The translocon is not shown for clarity. The PPXD has moved even further away from the HWD and has some intramolecular interactions with NBF II. The CTL of SecA is missing from this structure (Zimmer et al., 2008). D. Linear representation of SecA color-coded for the different domains. The sequence numbering beneath the cartoon is for *E. coli* SecA.
zinc is coordinated by three Cys and one His residues in the CXCXSGX₈CH sequence. The conserved Ser forms a strong hydrogen bond with the third Cys. This interaction likely plays an important role in the stability of the structure. This zinc binding domain is a key player in the interaction between SecA and SecB.

Other dimeric crystal structures of SecA from various prokaryotes species have been reported (Sharma et al., 2003; Zimmer et al., 2006; Vassylyev et al., 2006; Papanikolau et al., 2007) as well as a monomeric B. subtilis SecA structure (Osborne et al., 2004) All of these structures have similar monomer architecture except for the interface of the PPXD and HWD. In some structures the PPXD loosely packs against the HWD in the so called ‘closed’ state (Figure 1.8, A) (Hunt et al., 2002; Sharma et al., 2003; Zimmer et al., 2006; Vassylyev et al., 2006) while in other structures the PPXD rotates about 60° and is no longer in contact with the HWD in a so called ‘open’ state (Figure 1.8, B) (Papanikolau et al., 2007; Osborne et al., 2004). A recent elegant NMR structure of full-length Escherichia coli SecA demonstrated that SecA visits both the ‘open’ and ‘closed’ conformations in solution but the major species is the ‘open’ state (Gelis et al., 2007). The previously mentioned structure of SecA bound to the SecYEG translocon shows that SecA undergoes another conformational change upon binding to the translocon (Zimmer et al., 2008). The overall architecture remains the same but the PPXD has rotated even further away from the HWD and interacts with both NBF II and SecY. Another difference is the relative positions of the two NBFs. This structure was solved with ADP-BeFₓ, which mimics transition state during ATP hydrolysis and shows that NBF II has a rigid body rotation of about 15° towards the plane of the membrane.
This positions Arg 574 (E. coli numbering) towards the γ-phosphate of the ATP to trigger the hydrolysis reaction. Although these structures have started to provide atomic resolution snapshots of SecA but none of these structures provide full mechanistic details about SecA in its activated state prior to and during translocation.

1.7.3. Different conformations of SecA

In order for SecA to perform its required function as a preprotein translocase, it must perform conformational gymnastics. In the cytosol, SecA exists in ‘closed’ dimeric form and is protease resistant. Intramolecular interactions between the different domains help stabilize this compact conformation. Binding of different ligands such as SecB, phospholipids, and SecYEG trigger domain dissociation in SecA by weakening the intramolecular interactions to different extents. This domain dissociation results in a more open structure that is protease sensitive (Clerico et al., 2008). The conformational changes of SecA will be further described in the introduction to Chapter 2.

1.7.4. Signal sequence interactions with SecA

In order to test the secondary structure hypothesis of signal sequences (Emr and Silhavy, 1983) with SecA, our group performed line broadening and transfer Nuclear Overhauser Effect (trNOE) NMR experiments using synthetic signal peptides and soluble SecA (Chou and Gierasch, 2005). The observed trNOEs demonstrated that signal peptide bound to SecA adopted an α-helical structure in the h- and c-regions. Differential line broadening experiments indicated that one side of the helix form from residues in the h-region was more strongly bound to SecA. Additionally, the positive n-region of the signal sequence also contributed to the binding interaction. Therefore, the electrostatic and
hydrophobic characteristics as well as the secondary structure are needed for signal sequence-binding to SecA.

A considerable amount of effort has been expended in trying to determine the signal sequence-binding site on SecA. Initial studies demonstrated that signal sequences interact with the N-terminal two-thirds of SecA (Kimura et al., 1991; Triplett et al., 2001) but further biochemical and biophysical experiments do not converge on one signal site (Figure 3.3 and for further description and discussion see the introduction to Chapter 3). Major conformational changes are induced in SecA after initial interaction with preprotein and SecB. Very few signal sequence-binding studies have been performed on this more open translocation-active form, which is likely to be more physiologically relevant.

1.8. Statement of Dissertation

Protein secretion across or insertion into biological membranes occurs in all three kingdoms of life. Approximately, one-third of a cell’s proteome is secretory and integral membrane proteins (Papanikou et al., 2007). Over the last three decades considerable work has been performed to understand the targeting of the secretory proteins to their subcellular compartments and their active transport across membranes. Though significant progress has been made, many of the mechanistic details are still unresolved. In particular, the recognition mechanism of diverse signal sequences still remains elusive.

The studies of a truncated version of SecA, SecA64 (Triplett et al., 2001) and of the ATPase enhancement of SecA in denaturant (Song and Kim, 1997) lay the groundwork for this study. The SecA64 study demonstrated that the ATPase activated
form of SecA bound signal peptides better than the full-length protein (Triplett et al., 2001) while the denaturant study showed that low concentrations of denaturant have a similar effect on the activation of the ATPase activity as SecA64 but in the context of the full-length protein (Song and Kim, 1997). In this current study, full-length SecA in the presence of low concentrations of urea was characterized for structural changes to understand the conformational rearrangements of the activated translocation-competent SecA. Furthermore, this work describes the interactions of signal peptides with the cytosolic (solution) form of SecA as well as the translocation-active conformation of SecA.
CHAPTER 2

CHARACTERIZATION OF TRANSLOCATION-ACTIVE SECA

2.1 Introduction

SecA is a dynamic protein that visits different conformations during the preprotein translocation cycle. In the cytosol, SecA is in a closed form and has low ATPase activity (Schmidt et al., 2000). SecA can bind preproteins either in the cytosol or at the membrane. In the case of Gram-negative bacteria, the preprotein/SecB complex interacts with SecA already associated with the SecYEG translocon. These binding interactions with other components of the secretory pathway induce a major conformational change in SecA resulting in an open domain dissociated form (Clerico et al., 2008). This form of SecA possesses a high ATPase activity (Lill et al., 1990), which is used to help translocate the preprotein through the SecYEG channel.

Oliver and Beckwith (Oliver and Beckwith, 1982) first described SecA as a peripheral inner membrane protein but a subsequent study demonstrated that SecA partitioned between the cytosol and the inner membrane (Cabelli et al., 1991), and further experiments indicated that SecA could insert into membrane vesicles (Ulbrandt et al., 1992). Through protease-protection assays in membranes, Economou and Wickner (Economou and Wickner, 1994) showed that a 30 kD fragment corresponding to the C-terminal one-third of SecA was protected upon interaction with membranes, and they concluded it was inserted into the membrane. The cytosolic or solution form of SecA has been studied by FRET analysis (Ding et al., 2003a) and cryo-electron microscopy, which demonstrated that soluble SecA is in a closed dimeric conformation (Chen et al., 2008).
Taken together these results clearly illustrate that SecA has at least two distinct conformations, the closed cytosolic form and open membrane-bound form (Figure 2.1).

Studies performed on cytosolic SecA (c-SecA) have demonstrated that temperatures slightly higher than physiological temperature induced an endothermic conformational change (Schmidt et al., 2000; Fak et al., 2004). Using SecA mutants discovered through genetic analysis (Fikes and Bassford, 1989) (Fortin et al., 1990) (Oliver et al., 1990), Schmidt et al. (Schmidt et al., 2000) discovered at low temperatures these mutants are in the same open conformation as is wild-type SecA at elevated temperatures. These results suggest that the conformational changes observed at higher temperatures or because of physiologically relevant mutations change the conformation of SecA into a form similar to the membrane-bound form.

Driessen and coworkers (den Blaauwen et al., 1996) studied the effects of different nucleotides on the conformation of cytosolic SecA (c-SecA). Trp fluorescence revealed that ADP induced a closed form of SecA and the binding of an ATP analog, AMP-PNP enabled SecA to adopt a more open conformation. This open conformation was caused by the loosening of interactions between the C-terminal one-third of SecA and the N-terminal region. These results suggest that a domain dissociation of the C-terminal region of c-SecA upon AMP-PNP binding could aid in the priming of SecA for interaction with the membrane and/or SecYEG.

In another study using fluorescence resonance energy transfer (FRET) (Ding et al., 2003b), the close proximity of the PPXD and the C-terminal one-third of SecA was examined. This study demonstrated that ADP promotes a tight association between the PPXD and HWD while higher temperature and binding to model membranes caused a
Figure 2.1 Different conformations of SecA

SecA exists in at least two different conformations. In the cytosol SecA is in a closed ADP-bound form with low ATPase activity. Various ligands such as lipids, SecB, preprotein, and translocation as well as high temperature, denaturants, and truncation of the C-terminus force SecA to adopt the open conformation. Adapted from Fak et al. (Fak et al., 2004).
dissociation of the two domains. Furthermore, the hyperactive azide-resistant mutant of SecA (Oliver et al., 1990) is in an even more open conformation in the cytosolic form (Ding et al., 2003b). The recent NMR structure of SecA (Gelis et al., 2007) argued that c-SecA visits both the open and closed state but is predominantly in the open state (Figure 1.8, B). The NMR data on SecA was collected in the nucleotide-free form, which may have influenced the promotion of the open state. The recent structure of truncated SecA bound to the SecYEG translocon solved by Rapoport and colleagues (Zimmer et al., 2008) was determined in the presence of ADP BeF$_3^-$, a transition-state analog. As discussed in Chapter 1 section 1.7.2., this translocon-bound structure of SecA has the PPXD in an even more open conformation (Figure 1.8, C). All of these results suggest that the nucleotide bound to SecA influences PPXD and HWD dissociation and reorientation.

To try to understand the conformational changes induced in SecA upon binding to model membranes, Oliver and co-workers (Ding et al., 2001) employed fluorescence experiments using SecA mutants containing single Trp substitutions to identify the residues involved in phospholipid binding. SecA contains seven Trp residues (349, 519, 541, 622, 701, 723, and 775) (Figure 2.2) but Trp701, Trp723, and Trp775 located in the C-terminal portion of SecA are the major contributors to the overall fluorescence. In solution Trp723 and to a lesser extent Trp701 are solvent accessible while Trp775 is buried. In model membranes, Trp723 is less solvent exposed, Trp775 becomes more solvent exposed and Trp701 did not change in solvent accessibility. These results suggest that the C-terminal portion of SecA is flexible and therefore, may be crucial for the preprotein translocation cycle.
Figure 2.2 Seven Trp residues in SecA

The NMR structure from *E. coli* (Gelis et al., 2007) (2VDA) illustrates the seven Trp residues shown in the cyan spheres in SecA. The linear cartoon below the structure is colored for the different domains in SecA and show the positions of each of the Trp residues by the cyan boxes. *Indicate the Trp residues that contribute most to the overall fluorescence in SecA (Ding et al., 2001).*
Natale et al. (Natale et al., 2005) utilized fluorescence spectral changes in single Trp mutants of SecA along with acrylamide quenching to probe the structural rearrangements of SecA upon binding to the SecYEG translocon. The binding of SecA to the translocon induces SecA changes in the PPXD, NBF II and HWD. These results demonstrate that SecA bound to SecYEG adopts a conformation that is different from the cytosolic form and are observed in the recent truncated SecA/SecYEG structure (Zimmer et al., 2008).

In another set of experiments using proteolysis and fluorescence of a fluorescein-labeled SecY, SecA was observed to bind SecYEG in a protease-sensitive conformation independent of nucleotide. The fluorescence experiments indicated that additional conformational changes occur in the SecA/SecYEG complex upon ATP analog binding. This change in conformation is transmitted from the NBFs of SecA to the pore region of SecY. The studies suggest that there are several different conformational changes in the SecA/SecYEG complex during the preprotein translocation cycle (Robson et al., 2007).

Despite many biochemical and biophysical studies, the conformational changes in SecA during the translocation cycle are still unclear due to the lack of a good stable model of the active state. Many of the previous studies have utilized only one technique such as fluorescence and have used modified or mutated SecA protein. This chapter describes the activation and conformational change of SecA in adopting the translocation-active form. We use low concentrations of urea to activate SecA by domain dissociation. Through various biophysical experiments, limited proteolysis, and domain mapping, we have characterized a physiologically relevant, soluble form of SecA that
reflects the functional conformation prior to preprotein translocation across the inner membrane.

2.2 Results

2.2.1. Activation of SecA in low concentrations of urea

To generate an \textit{in vitro} form of full-length SecA that represents the conformational change during the early stages of translocation, we followed up on the work of Song and Kim (Song and Kim, 1997), which demonstrated that the ATPase activity of SecA becomes uncoupled in low concentration of denaturant. We performed urea titrations to examine the ATPase activity of SecA at 22°C. As the concentration of urea increased, the ATPase activity of SecA was enhanced to a peak rate at 2.2 M urea (Figure 2.3). This is an eight-fold higher activity than SecA in 0 M urea. After this peak in activity, the ATPase activity drops sharply and by 2.8 M urea no ATPase activity is observed. This eight-fold activation of the ATPase activity of SecA in 2.2 M urea is similar to the activation during preprotein translocation (Papanikou et al., 2004).

2.2.2. Structural analysis of intermediate

2.2.2.1. Tryptophan Fluorescence

The seven intrinsic Trp residues in SecA (Figure 2.2) provide us a tool to probe the tertiary structure of SecA during the course of a urea melt. The emission maximum of native SecA is 340 nm. Upon full denaturation in 8 M urea, the emission maximum red shifts to 350 nm and decreases in intensity (Figure 2.4, A). During a urea melt at 22°C, SecA Trp fluorescence at 340 nm begins decreasing at 1.4 M urea and at 2.2 M urea
Figure 2.3 Low concentration of urea stimulate the ATPase activity of SecA

The ATPase activity of SecA was measured as a function of urea concentration at 22°C. A peak in activity is observed at 2.2 M urea followed by a sharp decrease to no activity. SecA in the absence of urea was set to 100% activity. The error bars represent the standard deviation from three separate experiments.
Figure 2.4 Changes in SecA tertiary structure

A. The Trp fluorescence change in SecA upon unfolding in urea at 22°C. The red circles show the Trp fluorescence of native SecA excited at 295 nm. Native SecA has an emission peak around 340 nm. The blue circles show unfolding SecA in 8 M urea. Upon unfolding the fluorescence signal decreases and red-shifts to about 350nm. B. The Trp fluorescence was normalized to SecA in the absence of urea. The urea melt of SecA at 22°C shows an intermediate from 2.5 to 3.4 M urea and SecA is fully unfolded by 4 M urea.
A. 

![Fluorescence Intensity vs Wavelength](image)

- Native SecA
- 8 M Urea SecA

B. 

![Relative Fluorescence Intensity vs Urea Concentration](image)
about 50% of Trp fluorescence is lost. The maximal ATPase activity of SecA is in 2.2 M urea. The fluorescence data suggests the presence of an intermediate in 2.2 M urea region of the melt. The fluorescence remains nearly constant from 2.5 M urea to 3.4 M urea and SecA completely unfolds around 4 M urea. Trp 701, 723, and 775 have been shown to contribute most to the overall Trp fluorescence of SecA (Ding et al., 2001). Both Trp 701 and 723 are located in the HWD while Trp 775 is located in the C-terminal HSD. Trp775 is the most solvent inaccessible of the three Trp residues so the intramolecular interactions holding the C-terminal HSD to the other domains has been disrupted has been disrupted in 2.2 M urea. Taken together these results suggest that the increased ATPase activity of SecA is occurring due to a conformational change/unfolding event in the C-terminus of SecA.

2.2.2.2. Far-UV CD

In order to gain further understanding of the structural changes that occur in SecA, we examined the far-UV CD spectra in several urea concentrations (Figure 2.5). Due to the presence of urea, spectra could only be recorded to 215 nm. The helical content of SecA decreases as the urea concentration is increased. SecA in 2.2 M urea has lost about 25% of its helical content (Figure 2.6, green diamonds). Additionally, a structured intermediate is likely since there is very little change in the helical content from 2.2 M urea to 3.0 M urea. This loss of helical structure is likely attributed to unfolding or domain rearrangement in the C-terminal portion of the protein, in particular part of the HSD and the HWD.
Figure 2.5 Secondary structural changes in SecA in urea

The far-UV CD of SecA in different concentrations of urea at 22°C shows a decrease in the helical content of SecA as the concentration of urea is increased. SecA in the cluster of the spectra from 2.2 M to 3.0 M urea indicate the presence of an intermediate.
Low concentrations of urea activate the ATPase activity SecA. The red circles represent the percent of ATPase activity of SecA from Figure 2.3. The blue squares show the Trp fluorescence of SecA at 340 nm as shown in Figure 2.4. The green diamonds show the far-UV CD signal of SecA at 222 nm. The Trp and CD signal were normalized to SecA in the absence of urea. SecA in 2.2 M urea has an eight-fold enhancement in activity over SecA in 0 M urea and both the Trp fluorescence and far-UV CD indicate the presence of an intermediate at this urea concentration.
When the ATPase activity, Trp fluorescence, and far-UV CD are plotted together (Figure 2.6), it is clear that the peak ATPase activity coincides with a structured intermediate at 2.2 M urea. We term this intermediate form of SecA u-SecA.

2.2.2.3. Limited proteolysis

To gain further insight into how u-SecA differs structurally from the c-SecA structure, we performed a time course of limited proteolysis on both forms of SecA using $\alpha$-chymotrypsin (Figure 2.7). Strikingly, u-SecA is much more protease sensitive than c-SecA indicating that u-SecA is in a domain dissociated conformation similar to what has been observed for SecA in the presence of phospholipids (Ulbrandt et al., 1992). Quantitation of the amount of full-length SecA at each point in the digestion shows that in c-SecA 40% of full-length SecA remains after 15 minutes of digestion while full-length SecA in u-SecA is completely digested by five minutes (Figure 2.7, B). To further understand into the conformational change in u-SecA, we performed the chymotryptic digestion followed by detection with the region-specific antibodies developed by Ramamurthy and Oliver (Ramamurthy and Oliver, 1997) (Figure 2.8). These antibodies were generated before the first crystal structure of SecA was solved but the antibodies still roughly match the six different domains in SecA.

The region-specific antibody detection shows that a portion of NBF I, the C-terminal region including the HWD, the C-terminal HSD, and the CTL are the regions of the protein cleaved in both forms (Figure 2.9). These results are consistent with the fluorescence and near-UV CD suggesting that the C-terminal portion of the protein is
Figure 2.7 Limited proteolysis demonstrates u-SecA is in an open conformation

A. A time course of limited proteolysis of c-SecA and u-SecA was analyzed by an 8% SDS-PAGE and stained with Coomassie blue. The lanes are as follows: (1) SecA no digestion, (2) c-SecA 10 sec, (3) u-SecA 10 sec, (4) c-SecA 1 min, (5) u-SecA 1 min, (6) c-SecA 2 min, (7) u-SecA 2 min, (8) c-SecA 5 min, (9) u-SecA 5 min, (10) c-SecA 10 min, (11) u-SecA 10 min, (12) c-SecA 15 min, (13) u-SecA 15 min. B. Quantification of the amount of full-length SecA remaining at each time point and is the average of two separate experiments. The intensity of the bands were normalized to undigested SecA from lane 1.
The region-specific antibodies were generated as described by Ramamurthy and Oliver (Ramamurthy and Oliver, 1997). The cartoon representation of *E. coli* SecA is shown on the top and is color-coded for the different domains. The residues recognized by the six different antibodies (A1-A6) are indicated above the cartoon and the corresponding regions in the NMR structure from *E. coli* (Gelis et al., 2007) (2VDA) are show below. Most of the CTL domain was not determined in the NMR structure so this region is indicated by the red line in the A6 structure.
Figure 2.9 Different domain proteolysis in c-SecA and u-SecA

**A.** The region-specific antibodies against SecA were used to examine the digestion pattern of c-SecA. The 15-minute time point is shown with the molecular weight ladder indicated on the right side of the gel. The labels are as follows: C–Coomassie blue stained gel, A1–antibody against residues 1-209, A2–antibody against residues 211-350, A3–antibody against residues 351-509, A4–antibody against residues 519-664, A5–antibody against residues 665-820, A6–antibody against residues 822-901. Red boxes indicate proteolytic fragments that are similar in both c-SecA and u-SecA while the blue boxes demonstrate different proteolytic fragments. **B.** The region-specific antibodies against SecA were used to examine the digestion pattern of u-SecA. The 5-minute time point is shown with the molecular weight ladder indicated on the right side of the gel. The labels are the same as in A.
undergoing a conformational change in 2.2 M urea, and that this region of the protein could be experiencing similar changes during preprotein recognition at the membrane.

### 2.2.2.4. Mass Spectrometry of proteolytic fragments

The predicted chymotrypsin cleavage sites from ExPASy (Gasteiger et al., 2005) based on the linear SecA sequence do not completely explain all of the proteolytic fragments. To further identify these chymotryptic fragments in u-SecA, we utilized matrix assisted laser desorption ionization mass spectrometry (MALDI-MS). Four of the u-SecA fragments were analyzed but due to the large size of SecA (901 amino acids) and the large number of chymotrypsin cleavage sites, none the u-SecA fragments contained a homogenous SecA fragment. Based on the MS peak intensities (see Materials and Methods), predicted chymotrypsin cleavage sites, and antibody mapping, we determined the most abundant proteolytic species in the u-SecA fragments (Figure 2.10).

The U50 fragment (Figure 2.10, C) contained peptides covering almost the entire SecA sequence. To determine the abundant 50 kD proteolytic fragment(s) in this sample, we analyzed the high-intensity peaks (Figure 2.10, C, gray boxes) and discovered two SecA chymotryptic fragments were the abundant species. The first fragment contains NBF I, PPXD, and half of NBF II, while the second fragment consists of the C-terminal portion of the PPXD, NBF II, N-terminal HSD, and the HWD. These two fragments differ only by 1 kD hence they cannot be resolved by SDS-PAGE. The U30 sample was analyzed using the same method (Figure 2.10, D). This most abundant fragment is 32 kD and is located at the C-terminal end of SecA. Two other minor species of 33 kD and 32 kD are present in the U30 sample and are in the N-terminal two-thirds of SecA. The
Figure 2.10 Mass spectrometry determined protease cleavage sites

**A.** A four-minute digestion of u-SecA followed by Coomassie blue staining (C) and region-specific antibody detection. The boxes indicate the chymotryptic u-SecA fragment that was analyzed by MALDI-MS/MS. The molecular weight standard is indicated on the left of the gel.  
**B.** The notation used to present the MS data in C and D. Manual analysis used high-intensity peaks to determine the abundant species in each u-SecA fragment analyzed by MALDI-MS/MS (see Materials and Methods).  
**C.** MALDI-MS analysis of the U50 fragment shows peptide coverage across almost the entire sequence. High-intensity peak data determined that U50 contained two predominant fragments of 51 kD and 52 kD.  
**D.** MALDI-MS analysis of the U30 fragment shows peptide coverage across almost the entire sequence. High-intensity peak data determined that U30 contained one major C-terminal fragment of 32 kD and two other minor N-terminal fragments of 33 kD.  
**E.** Combining the MS and antibody data, a map of the chymotrypsin cleavage sites in u-SecA, shown below the cartoon, was determined. The large, medium, and small arrows indicate the relative cleavage at each of the sites. The chymotrypsin cleavage sites in c-SecA, shown above the cartoon, were determined based on the probability of cleavage at a specific residue and surface accessibility (see text for details).
smaller U27 and U20 fragments were also analyzed but due to numerous MS identified peptides, an abundant fragment(s) could not be determined. Using the peptide identified by MS experiments and the antibody data, we were able to determine the relative cleavage of the chymotrypsin sites in u-SecA (Figure 2.10, E). The surface accessibility calculation of the side chain residues of the *E. coli* NMR structure (2VDA) using VADAR software (Willard et al., 2003) identified the surface exposed aromatic residues. This data along with the predicted cleavage sites from ExPASy (Gasteiger et al., 2005) enabled us to determine the chymotryptic site in c-SecA. Taken together this data demonstrates that the C-terminal portion of both c-SecA and u-SecA are proteolytically labile. A stable 30 kD fragment is generated if cleavage occurs at Trp622 prior to digestion of other parts of HWD, HSD, and CTL. This C-terminal fragment is consistent with the 30 kD membrane-inserted domain characterized by Wickner et al. (Price et al., 1996).

### 2.2.2.5. Oligomeric state of u-SecA

One of the many unresolved questions about SecA’s molecular mechanism during the translocation cycle is whether it functions at the membrane as a dimer or a monomer (see Chapter 1 section 1.7.1.1. for further discussion). In the case of u-SecA, we also needed to determine the oligomeric state of the protein. Typical experiments such as size-exclusion chromatography yielded ambiguous results so we performed velocity analytical ultracentrifugation (AUC) experiments. C-SecA was used as a control since it has previously been shown by this method to be a dimer (Woodbury et al., 2002). At 2 μM, 5 μM, and 10 μM concentrations, c-SecA exists in a dimeric form (Figure 2.11, A) but
Figure 2.11 AUC of c-SecA and u-SecA

A. C-SecA at 5 μM was analyzed by velocity analytical ultracentrifugation. The top of the graph shows absorbance scans of the cell with the black symbols indicating the scans at the beginning of the centrifugation run and red symbols representing the scans at the end of the run. The scans were fit to a continuous c(s) distribution model using SEDFIT software (Schuck et al., 2002). The sedimentation coefficient translated into a molecular weight of 203 kD, which is 1 kD off from the calculated 204 kD molecular weight of dimeric SecA. B. The same analysis as in A using 5 μM u-SecA. U-SecA sediments much faster than c-SecA and application of the continuous distribution model indicates large molecular weight species. This indicates that u-SecA may be aggregating.
u-SecA was found to be in several higher molecular weight species (Figure 2.11, B). These larger species are likely due to protein aggregation. To determine if u-SecA was aggregating, we used a gluteraldehyde cross-linking approach. Glutaraldehyde is a non-specific amine-amine cross-linker so residues in close proximity to each other will be cross-linked. We used this method to determine the oligomeric state of both c-SecA and u-SecA at several different protein concentrations (Figure 2.12). C-SecA is mostly in a dimeric state as expected but u-SecA is found in a monomeric form. As the protein concentration of u-SecA is increased, some u-SecA is in the dimeric form as well as in an aggregated form. This experiment explains the higher molecular species found in the AUC experiments with u-SecA. This result demonstrates that u-SecA is mostly in a monomeric form at low protein concentration and that SecA unfolds as a monomer not through a dimeric intermediate as previously reported (Doyle et al., 2000)

2.3. Discussion

SecA in the cytosol is in a compact state with low basal level of ATPase activity and the atomic details have been described by the available structures (Hunt et al., 2002; Sharma et al., 2003; Osborne et al., 2004; Zimmer et al., 2006; Vassylyev et al., 2006; Papanikolau et al., 2007). On the other hand, little is actually understood about the highly active membrane/translocon bound form of SecA. Therefore, we generated and characterized u-SecA: a full-length, soluble urea intermediate that is in an open and active physiologically relevant conformation mimicking the translocation-active form.

Various studies have demonstrated that when SecA interacts with either the phospholipids membrane or the SecYEG translocon, massive structural rearrangements
Figure 2.12 Cross-linking demonstrates u-SecA is a monomer

C-SecA and u-SecA oligomeric state was determined by gluteraldehyde cross-linking and run on a 6% tricine SDS-PAGE gel. The molecular weight standard is indicated on the left side and the different oligomeric forms of SecA are indicated on the right side. The lanes are as follows: (1) 0.5 µM cross-linked c-SecA, (2) 1.0 µM cross-linked c-SecA, (3) 2.5 µM cross-linked c-SecA, (4) 5.0 µM cross-linked c-SecA, (5) 1.0 µM non cross-linked SecA, (6) 0.5 µM cross-linked u-SecA, (7) 1.0 µM cross-linked u-SecA, (8) 2.5 µM cross-linked u-SecA (9) 5.0 µM cross-linked u-SecA. C-SecA is predominately in a dimeric form while u-SecA is in a monomeric form. The diffuse nature of cross-linked band is due to intramolecular cross-linking.
occur (Economou and Wickner, 1994; Ding et al., 2001; Natale et al., 2005; Robson et al., 2007). This conformational change causes the cytosolic protease-resistant SecA to adopt an open, domain-dissociated, and protease-sensitive form. Through limited chymotrypsin proteolysis, we determined that NBF I is destabilized in u-SecA (Figure 2.9, B, A1 lane). The unfolded nature of this domain in u-SecA suggests that it is missing an interaction partner. Supporting our observation is a study by Osborne and Rapoport (Osborne and Rapoport, 2007), which demonstrated that NBF I of SecA is involved in the binding interaction with SecY.

In our limited protease digestion studies, the PPXD in u-SecA can be subdivided into two regions where one region is more protease sensitive than the other (Figure 2.9, B, A2 lane). This stabilization of one portion of the PPXD is consistent with the crystal structure of SecA /SecYEG complex (Zimmer et al., 2008), which shows that the PPXD has rotated far away from the HWD and makes molecular interactions with NBF II. This crystal structure also supports our observation that portions of the PPXD, NBF II, and the long helix from the HSD form a stable core in the translocation-active form of SecA.

Previous data has suggested that in the presence of anionic membranes SecA partially unfolds and this unfolding event increases membrane insertion (Ulbrandt et al., 1992). Other evidence showed that about one third of SecA can insert into the membrane (Price et al., 1996). Since Trp 701, 723, and 775 contribute most to the overall Trp fluorescence (Ding et al., 2001), the C-terminal third of SecA can interact with the inner membrane. Our biophysical, proteolysis, and MS data demonstrates that the HWD, C-terminal HSD, and CTL of u-SecA have dissociated from the rest of the molecule. The proteolysis and MS results indicate a stable 30 kD fragment from the C-terminal one-
third of SecA is generated (Figure 2.9, B, A5 and A6 lanes and Figure 2.10, D). This C-terminal region is highly protease sensitive in u-SecA but if chymotrypsin cleavage first occurs at Trp 623, the released C-terminal region can refold into a more stable fragment. This 30 kD fragment is similar to protected membrane-inserted fragment identified by Price et al. (Price et al., 1996). The domain dissociation in activated SecA causes the release of IRA 1 located in the C-terminal HSD increasing the ATPase activity required for preprotein translocation (Karamanou et al., 1999). This conformational change in SecA primes the molecule for preprotein translocation and also could reveal a preprotein interaction sites that are not accessible in the cytosolic form of SecA.

In our studies, we determined that u-SecA is in a mostly monomeric form, which is consistent with the crystal structure of SecA bound to SecYEG (Zimmer et al., 2008) and other translocation experiments (Duong, 2003; Osborne et al., 2004; Osborne and Rapoport, 2007). This dissociation of the dimer for translocation likely reveals residues needed for specific SecY binding interactions. Additionally, we demonstrated that u-SecA begins to nonspecifically aggregate at higher protein concentrations. The dimer dissociation into the activated translocation state reveals hydrophobic surfaces for binding interactions with SecY, preprotein, and the inner membrane. In u-SecA these interaction region are unsatisfied leading to aggregation. These observations suggest that the monomeric form of SecA is important in the translocation cycle.

Using our results in conjunction with previous studies, we propose a model of SecA interactions with the inner membrane and the SecYEG translocon. Upon interaction with the translocon, the SecA dimer dissociates into a monomer and undergoes massive structural rearrangements. This conformational change breaks PPXD interactions with the
HWD and promotes NBF I binding to the translocon (Figure 2.13). The loss of the PPXD interactions drive the HWD, C-terminal HSD, and CTL insertion into the membrane, which in Gram-negative bacteria induces SecB dissociation. These new interactions create the proper positioning of SecA for the productive translocation of preproteins.
Figure 2.13 Model of the conformation change in SecA upon binding SecYEG

The different domains of SecA have been color-coded: blue, NBF I; yellow, PPXD; purple, NBF II; green, N-terminal HSD; orange, HWD; dark green, C-terminal HSD; and red, CTL. (1) Cytosolic SecA is in a closed antiparallel dimeric form and has low ATPase activity. (2) Upon interaction with SecY the SecA dimer dissociates. (3) The HWD, C-terminal HSD, and CTL dissociate from the rest of the SecA monomer and insert into the membrane. This new conformation positions SecA on the SecYEG translocon for productive translocation of preproteins.
2.4. Materials and Methods

2.4.1. Recombinant DNA experiments

2.4.1.1. Cloning of SecA gene

The SecA gene was amplified by PCR from the pT7-SecA2 plasmid (D. Oliver, Wesleyan University) using forward primer (5′-GGAATTCATATGCCATGCTTAATCAAATTTGAAC-3′; introduced restriction enzyme sites are underlined) and reverse primer (5′-CCGCTCGAGTTATTGCAGGCGGCCATGGC-3′) using Taq DNA Polymerase (New England Biolabs, Ipswich, MA). The 2.7-kb PCR fragment was subcloned into the pGEM®-T vector (Promega, Madison, WI). The resulting plasmid was digested with NdeI/XhoI restriction enzymes (New England Biolabs, Ipswich, MA) and ligated into the same sites in pET17b (Novagen, Madison, WI) creating the pET-17b-FL SecA plasmid. DNA sequencing (Davis Sequencing, Davis, CA) verified the correct sequence of the SecA gene.

2.4.1.2. Site-directed mutagenesis

Sequencing results from the pET-17b-FL SecA revealed the presence of two point mutations. One mutation was silent while the other one changed an amino acid T371A. The mutagenesis primer (5′-CCAGAACGAAAAACCCACGCTGGCTTCGATCACC-3′) and its complement (the single nucleotide change underlined) were used to change the Ala back to Thr at position 371. Pfu Turbo DNA polymerase (Stratagene, La Jolla, CA) was used with 50 ng of pET-17b-FL SecA 371A and 125 ng of each primer to
incorporate the single nucleotide change following the manufacture’s protocol. DNA sequencing (Davis Sequencing, Davis, CA) verified the incorporation of the mutation in the SecA gene.

2.4.2. Protein expression and purification

The pET-17b-FL SecA plasmid was transformed into the BL21 (DE3) bacterial strain. Cells were grown at 37°C to OD$_{600}$ of 0.5 in LB supplemented with LinA salts and 50 µg/ml ampicillin. Expression of SecA protein was induced by the addition of IPTG to a final concentration of 0.75 mM and the culture was grown for another 2.5 hours at 37°C. The bacterial cells were pelleted, resuspended in lysis buffer (50 mM Tris pH 7.6, 100 mM NaCl, 1 mM DTT), frozen in liquid nitrogen, and stored at -80°C until purification.

SecA protein was purified as described previously (Fak et al., 2004) with some modifications. Bacterial cells were treated with a cocktail of protease inhibitors (150 µM PMSF, 250 µg leupeptin, 10 µg pepstatin, 150 µM ABESF) followed by lysis. Cell lysis was performed either by lysozyme treatment and sonication or by microfluidics system. For lysozyme treatment, cells were thawed and 20 mg of lysozyme was added (per liter of original culture). The cells were incubated on ice for 20 minutes and then 4 mg of DNase I was added (per liter of original culture) and incubated on ice for 10 minutes. The cells were sonicated for 30 seconds followed by two minutes on ice for a total of five cycles. For lysis by the microfluidics system, the cells were thawed and the total volume was increase to 30 ml with lysis buffer. The cells were passed through the microfluidizer® processor (Microfluidics, Newton, MA) at a pressure of 16,000 PSI two times.
After a 10,000 rpm spin at 4°C, the supernatant was passed through a 0.22 µm filter and loaded onto a DEAE Sepharose CL-6B (Sigma, St. Louis, MO) column equilibrated with 50 mM Tris pH 7.6, 1 mM DTT, 0.025% NaN₃. The protein was purified on an AKTA purification system (GE Healthcare, Piscataway, NJ) using a 100 ml salt gradient from 0 to 1 M NaCl. Fractions were run on 8% Tricine SDS-PAGE. Relevant fractions containing the SecA protein were pooled and loaded onto a SP Sepharose Fast Flow (Sigma, St. Louis, MO) column equilibrated with 50 mM HEPES pH 7.5, 1 mM DTT, 0.025% NaN₃. The loaded protein was purified using a 100 ml salt gradient from 0 to 1 M NaCl. Fractions were run on 8% SDS-PAGE to determine which fractions contained SecA protein and the relevant fractions were concentrated to a final volume of 2.5 - 5 ml. The concentrated SecA protein was injected onto a Superdex 200 prep grade (GE Healthcare, Piscataway, NJ) gel filtration column equilibrated with 25 mM Tris pH 7.6, 150 mM KCl, 1 mM MgCl₂. The relevant fractions containing pure SecA were pooled, concentrated, snap-frozen in liquid nitrogen, and stored at -80°C. Protein concentration was determined by absorbance at 275 nm (the peak absorbance for SecA in an absorbance scan) after dilution into 6 M guanidinium-HCl, 20 mM sodium phosphate pH 6.5, assuming an extinction coefficient of 0.73 OD per mg mL⁻¹. Each batch of protein was checked for ATPase activity by the enzyme-coupled ATPase assay and secondary structure by far-UV CD.

2.4.3. Enzyme-coupled ATPase assay

The ATPase activity of purified SecA was measured by the steady-state enzyme-coupled ATPase assay. This assay uses pyruvate kinase and lactate dehydrogenase to
couple the ATP hydrolysis to the conversion of NADH to NAD$^+$ (Norby, 1988). To pre-equilibrate the sample, 400 nM SecA was added to 50 mM Tris pH 8.0, 30 mM KCl, 30 mM NH$_4$Cl, 1 mM DTT, 5 mM Mg(OAc)$_2$, 3 mM phospho(enol)pyruvate, 10 U of pyruvate kinase, and 14.4 U of lactate dehydrogenase at 37°C for 15 minutes. After pre-equilibration, 0.5 mM ATP and 0.25 mM NADH was added to the reaction mixture and the change in transmittance at 340 nm was monitored using the J715 spectropolarimeter (JASCO) instrument at 37°C for 16 minutes. The rate of hydrolysis was determined by changing the percent transmittance to absorbance. The absorbance at 340 nm over time was plotted excluding the first 60 seconds in which the reaction was achieving steady-state and the slope was determined by linear least squares curve fitting. The change in absorbance is directly proportional to the mole of NADH hydrolyzed using the $\varepsilon_{340}$ of NADH as 6.22 x10$^3$ cm$^{-1}$ M$^{-1}$. The correspondence between NADH hydrolyzed and ATP hydrolysis is 1:1, therefore, the rate of ATP hydrolysis is equal to the rate measured for NADH depletion. The ATPase activity was expressed as pmol ATP hydrolyzed/µg SecA minute and was the average of three independent measurements.

2.4.4. Far-UV CD

CD spectra of proteins were measure using a J715 spectropolarimeter (JASCO). Spectra were taken in the far-UV region from 195-250 nm using a 1-mm pathlength cell. Scan speed was 20 nm/min and five scans were averaged. The samples were in 300 µl of 10 mM KHPO$_4$ pH 7.6 and 1 mM MgCl$_2$ at a concentration of 0.5 or 1 µM. The spectra were buffer corrected.
2.4.5. ATPase assay in urea

The ATPase activity of SecA in various urea concentrations was measured using the malachite green-ammonium molybdate ATPase assay (Lanzetta et al., 1979) with o-phosphoric acid (Sigma, St. Louis, MO) as a standard curve. After equilibration at 22°C for 12 – 14 hours of 1 µM SecA in 50 mM Tris pH 8.0, 30 mM KCl, 30 mM NH₄Cl, 5 mM Mg(OAc)₂, 1 mM DTT with various concentrations of urea, 5 mM ATP was added and the reaction was incubated for one hour at 22°C. The absorbance was measured at 660 nm using a Genesys 10 UV scanning spectrophotometer (Thermo Electron Corp).

2.4.6. Equilibrium fluorescence measurements

Samples containing 1 µM SecA were incubated at 22°C with various concentrations of urea in 50 mM Tris pH 8.0, 30 mM KCl, 30 mM NH₄Cl, 5 mM Mg(OAc)₂, 1 mM DTT for 12-14 hours. Trp fluorescence of SecA was measured using a Photon Technology International QM-1 spectrofluorometer at 22°C. The fluorescence excitation wavelength was 295 nm and the emission spectra were measured from 300 to 380 nm. The excitation and emission bandwidths were both set to 2 nm. The fluorescence signal at 340 nm was plotted against the urea concentration and corrected for background fluorescence using urea buffer.

2.4.7. CD measurements in urea

Far-UV CD measurements of 1 µM SecA in 20 mM HEPES pH 7.5, 50 mM KCl, 1 mM MgCl₂, 1 mM DTT with various concentrations of urea incubated at 22°C overnight were recorded on a J715 spectropolarimeter (JASCO) from 215-250 nm using
a 1-mm pathlength cell. Scan speed was 20 nm/min and five scans of each sample were averaged. Spectra were taken of urea buffer to substrate any background signal.

2.4.8. Limited proteolysis

C-SecA in 20 mM HEPES pH 7.5, 50 mM KCl, 1 mM MgCl$_2$, 1 mM DTT and u-SecA at 2.5 µM in 20 mM HEPES pH 7.5, 50 mM KCl, 1 mM MgCl$_2$, 1 mM DTT, 2.2 M urea were generated by incubation at 22°C for four hours. The samples were subjected to a time course of limited α-chymotrypsin (1:150 w/w) digestion from 0 to 60 minutes. At each time point, adding 3X SDS-PAGE sample buffer and flash freezing in liquid nitrogen quenched the digestion. Samples were either boiled and subjected to SDS-PAGE or stored at -80°C until further use. The samples were run in duplicate on two 8% or 10% SDS-PAGE gels. The gels were either stained with Coomassie blue or transferred to PVDF membrane for further analysis.

2.4.9. Antibody Detection

Digested c-SecA and u-SecA protein transferred to PVDF membrane was detected with region-specific SecA antiserum (Ramamurthy and Oliver, 1997) to determine the identity of the proteolytic bands. The blots were blocked overnight at 4°C or at room temperature for one hour in 5% non-fat milk in Tris-Buffered Saline with 0.5% Tween 20 (TBS-T). After washing with TBS-T, the blot was probed with one of the region-specific antibodies at 1:5000 dilution in TBS-T for one hour at room temperature. Next, a secondary goat anti-rabbit IgG alkaline phosphatase (AP) conjugated antibody (Sigma) was used according to the manufacturer’s protocol. Antibodies were detected by AP Lumino (Pierce, Rockford, IL) using a G:Box gel documentation unit (Syngene,
Frederick, MD). Blots were stripped in 50 mM Tris pH 6.8, 2% SDS, and 100 mM β-mercaptoethanol buffer at 55°C for one hour. The blots were then reprobed with the secondary antibody, and if no background signal was detected, the blot was blocked and probed with the other region-specific antibodies following the above protocol.

2.4.10. Mass spectrometry

Four u-SecA fragments, U50, U30, U27, and U20 numbered for their apparent molecular weight, were analyzed by MALDI-MS/MS. The fragments were excised from 8% tricine SDS-PAGE gel using a clean razor and analyzed by the Center for Advanced Proteomics Research at the New Jersey Medical School of the University of Medicine and Dentistry of New Jersey. The gel pieces were washed with 30% acetonitrile in 50 mM ammonium bicarbonate followed by reduction with DTT and alkylation with iodacetamide. The samples were further digested with trypsin at 37°C overnight. The resulting trypsin digested peptides were extracted with 1% trifluoracetic acid two time followed by 80% acetonitrile. The samples were concentrated using a speed vac and desalted using C18 ZipTips. The SecA peptides were mixed in a 1:1 ration with 7 mg/ml α-cyano-4-hydroxy-cinnamic acid and spotted onto the MALDI plate. The data was analyzed using the local Mascot search against the *E. coli* SecA sequence. The variable modifications employed were oxidized methionine and carbamidomethyl labeled-cystenine and the error of the MS was within 50 ppm.

Due to the heterogeneous nature of the original α-chymotrypsin fragments further manual analysis was performed. Taking into consideration all of the peptides that were detected by the MS analysis, the size of the original chymotrypsin U50 fragment would
correspond to an 83 kD fragment. In the mass spectrum, a single peak of high intensity does not correlate to the most abundant peptide but if most of the intense peaks cluster to one region of the protein, that region is likely the most abundant species in the original u-SecA chymotryptic fragment. A threshold of 10% intensity was set to determine all of the high intensity peaks. All of the high-intensity peaks corresponding to the four original fragments were mapped onto the SecA sequence. Based on the predicted chymotrypsin cleavage sites along with the intensity analysis, the abundant species in the original U50 and U30 SecA fragments were identified. For the smaller U27 and U20 fragments, the clustering of the MS intensities no longer applies due to the large number of proteolytic fragments of similar sizes. The sequencing data from the MS/MS was used to determine if a particular chymotrypsin cleavage site was used. If any of the aromatic residues appeared in the sequenced peptide of all four u-SecA fragments more than two times, this residue was not considered to be cleaved in the original chymotryptic digestion. From this analysis the cleavage sites in u-SecA were assigned.

2.4.11. Velocity analytical ultracentrifugation

C-SecA in 20 mM HEPES pH 7.5, 50 mM KCl, 1 mM MgCl₂, 1 mM TCEP and u-SecA in 20 mM HEPES pH 7.5, 50 mM KCl, 1 mM MgCl₂, 1 mM TCEP, 2.2 M urea, at 2 µM, 5 µM, and 10 µM were equilibrated overnight at 22°C. SecA samples and buffer were loaded into cells containing two-sector centerpieces. The An-60 Ti rotor was equilibrated to 22°C prior to the AUC run. The three separate cells containing samples were centrifuged at 30,000 rpm until sedimentation of the protein was complete. Radial scans measured the absorbance profile of each of the three samples at four-minute
intervals. The sedimentation coefficient was determined using the SEDFIT software (Schuck et al., 2002) using the continuous c(s) distribution model. The partial –specific volume, density, and viscosity of both samples were calculated using the freeware program SEDNTERP (http://www.jphilo.mailway.com/). The partial-specific volume used for both samples is 0.7335. The density and viscosity values used are 1.00071 and 0.01002 for c-SecA and 1.03508 and 0.010997 for u-SecA.

2.4.12. Gluteraldehyde Cross-linking

C-SecA in 20 mM HEPES pH 7.5, 50 mM KCl, 1 mM MgCl₂, 1 mM DTT and u-SecA in 20 mM HEPES pH 7.5, 50 mM KCl, 1 mM MgCl₂, 1 mM DTT, 2.2 M urea were equilibrated for four hours at 22°C. Gluteraldehyde was added to a final concentration of 0.1% and incubated at room temperature for two minutes. The cross-linking was quenched by the addition of 100 mM Tris pH 8.0. The samples were boiled for five minutes and loaded onto a 6% Tricine SDS-PAGE. The proteins were visualized by staining the gel with Coomassie Blue.
CHAPTER 3

SIGNAL SEQUENCE INTERACTIONS WITH CYTOSOLIC SECA AND TRANSLOCATION-ACTIVE SECA

3.1. Introduction

A critical step in the post-translational secretion pathway in bacteria is the recognition and binding of the signal sequence by SecA. After the preprotein has been translated by the ribosome, either SecB interacts with the mature region of the preprotein to maintain it in an unfolded state or the preprotein is recognized by cytosolic SecA. The preprotein/SecB complex is targeted to SecA, which binds the signal sequence and mature region of the preprotein and SecB. Upon interactions with the preprotein/SecB complex as well as with the SecYEG translocon, SecA undergoes a dramatic conformational change into the translocation-active form (Figure 3.1). This structural rearrangement likely reveals a higher affinity signal sequence-binding site that is used as a proofreading step prior to insertion of the precursor into the SecYEG translocon. SecA correctly positions the preprotein and uses ATP hydrolysis for translocation through the SecYEG channel (Clerico, 2008). Despite the vast amount of biochemical and biophysical experiments performed to try to identify the signal sequence-binding site, its exact location in both cytosolic and translocation-active SecA is still poorly defined.

Various biochemical methods have been used to try to determine the region where signal sequences bind SecA. Kimura et al. (Kimura et al., 1991) first tried to try to identify the signal sequence-binding site using the zero-length cross-linker, 1-ethyl-3-(3-dimethylaminoproyl) carbodiimide (EDC), which couples carboxyl groups to primary
Figure 3.1 Model for SecA-mediated translocation

The preprotein (pink) is translated by the ribosome nascent chain complex (RNC). Tetrameric SecB (tan hexagons) binds the preprotein keeping it in an unfolded state. SecA (green) interacts with the SecB/preprotein complex and binds to the SecYEG translocon and undergoes a major conformational change. SecA binds the signal sequence (red box) of the preprotein and positions the preprotein for translocation through SecYEG.
amines. Reconstituted SecA fragments were cross-linked to radio-labeled *in vitro* translated preproteins. These experiments indicated that the preprotein cross-linked to the N-terminal portion of SecA and demonstrated that residues 267-340 encompassed the binding site.

Initial studies performed by our group, compared the binding of signal peptides to full-length SecA and an activated form of SecA, SecA64, which is a 64 kD N-terminal chymotryptic fragment generated in the presence of ATP. This experiment showed that signal peptides bound more strongly to SecA64 than to full-length SecA (Triplett et al., 2001). These results suggest that either the signal sequence-binding site is more accessible in SecA64 or that the activated form of SecA has a different binding site.

Efforts to identify the signal sequence-binding site began in earnest when the first crystal structure of SecA was solved by Hunt el al. (Hunt et al., 2002), a computer program VOIDOO (Kleywegt and Jones, 1994) was used (Hunt et al., 2002) to detect unknown grooves or cavities on the surface of the structure. Three regions were identified as possible preprotein interactions sites: the CTL groove, the single-stranded nucleic acids (SS-NA) groove, and the methionine canyon (Figure 3.2, A). Each of these cavities is located at an interface between two or more domains. Binding of the signal sequence of a preprotein to the CTL groove would require the displacement of the CTL of SecA (Hunt et al., 2002). Previous studies have shown that SecB binds to the CTL of SecA, and this likely disengages the CTL from the rest of the molecule (Dempsey et al., 2004). Our group previously reported (Chou and Gierasch, 2005) the structure of the signal peptide bound to SecA and proposed a binding site in the same region as the CTL groove (Figure 3.2, B, red and green), which has both electrostatic and hydrophobic
Figure 3.2 Proposed signal sequence binding regions

A. Surface and cartoon representation of the *B. subtilis* SecA crystal structure (Hunt et al., 2002) (1M6N) color-coded for the different domain: blue, first nucleotide binding fold (NBFI); yellow, preprotein cross-linking domain (PPXD); purple, second nucleotide binding fold (NBFII); green, α-helical scaffold domain (HSD); orange, α-helical wing domain (HWD); and red, carboxyl-terminal linker (CTL). The black boxes indicate the location of possible ligand-binding sites, CTL groove, single-stranded nucleic acids (SS-NA) groove, and methionine canyon. B. Monomeric *B.subtilis* SecA crystal structure (Osborne et al., 2004) (1TF5) surface and cartoon representation color-coded for proposed signal sequence binding regions: red, Chou and Gierasch, 2005 (Chou and Gierasch, 2005); blue, Papanikou et al., 2005 (Papanikou et al., 2005); green, overlap between the latter two sites; cyan, Musial-Siwek et al., 2007 (Musial-Siwek et al., 2007); yellow, Osborne et al., 2004 (Osborne et al., 2004). C. Cartoon and surface representation of the *E. coli* SecA NMR structure (Gelis et al., 2007) (2VDA) color-coded for domains as described above. The modeled signal peptide is shown in cyan. D. Structure of SecA from *T. maritima* represented by surface and cartoon bound to SecYEG (not shown) (Zimmer et al., 2008) (3DIN). The arrow represents the clamp region proposed to be a potential signal sequence binding site.
character. SecA has structural homology to helicases, and the proposed SS-NA groove is analogous to the site where nucleic acids interact with helicases. The methionine canyon at the interface between NBF II and the N-terminal HSD has weakly packed Met residues (Hunt et al., 2002). In SRP the signal sequence has been shown to interact with the Met-rich M-domain (Zopf et al., 1993), which has a hydrophobic groove lined with Met residues (Bernstein et al., 1989). All of these potential interaction sites are located within the first two-thirds of the SecA molecule, which is consistent with the finding that signal peptides interact with SecA64 (Triplett et al., 2001).

In another study, Economou and co-workers (Baud et al., 2002) used chemical cross-linking with dithiobis(succinimidyl propionate) to demonstrate that a synthetic alkaline phosphatase signal peptide interacts with N-terminal fragments of SecA containing the first 263 residues. They further showed by deletion analysis that residues 219-244 are essential for signal peptide binding but residues 234-263 may be needed for optimal signal sequence-binding (Baud et al., 2002). This proposed binding region is located in the preprotein cross-linking domain (PPXD), which was so named from the results presented in the original preprotein cross-linking study by Kimura et al. (Kimura et al., 1991). These results were further supported by surface plasmon resonance assays, which showed that several SecA modification including C-terminal truncation (residues 1-610), IRA1 deletion (Δ783-795) (for further discussion of IRA1, see Chapter 1), or mutation of Trp775 to Ala, increase binding to the immobilized signal peptide when compared to wild-type SecA (Vrontou and Economou, 2004). These results indicate that IRA1 may be acting as a molecular switch to regulate signal sequence binding (Baud et al., 2002). This group also studied the PPXD in isolation as well as various SecA deletion
mutants to determine if the PPXD could bind both the signal peptide and the mature region of the preprotein. They found that signal peptides can bind directly to the isolated PPXD and stabilize this domain in a distinct conformation (Papanikou et al., 2005).

The PPXD can be further broken down into two subdomains, the stem (residues 219-240 and 362-377) and the bulb (residues 234-361). Through protease protection assays using either a signal sequence or a mature region of a preprotein, signal sequences were shown to bind to the stem region (Figure 3.2, B, blue and green) while the mature region of the preprotein can bind to the bulb (Papanikou et al., 2005). This potential signal sequence binding to the stem region consistent with to the CTL groove proposed in by (Hunt et al., 2002) when the first crystal structure of SecA was solved.

In 2004 Rapoport and co-workers (Osborne et al., 2004) solved the crystal structure of a monomeric version of SecA. In this structure, the PPXD has rotated and moved away from the HWD creating a groove between these two domains. Based on this structure, this group proposed two alternative preprotein binding grooves. Groove 1 is located between the PPXD and HWD and is lined with several polar and charged residues (Figure 3.2, B, yellow). The second proposed site, groove 2, overlaps with the signal sequence-binding site proposed by several other groups (Figure 3.2, B, blue, green, and red) (Papanikou et al., 2005; Chou and Gierasch, 2005; Hunt et al., 2002).

In order to determine the region important for signal sequence-binding, Musial-Siwek et al. (Musial-Siwek et al., 2005) monitored signal peptide binding to SecA by fluorescence resonance energy transfer (FRET). In this case, the intrinsic Trp residues of SecA were used as the donor, and the alkaline phosphatase signal sequence synthesized with a C-terminal Cys, which was labeled with IAEDANS, served as the acceptor. Using
this system they were able calculated that c-SecA has one signal sequence-binding site with a $K_d$ of 2.4 µM but in lipid vesicles the binding fit best to a two-state model with a $K_{d1}$ of 0.9 µM and $K_{d2}$ of 11.2 µM. By Trp mutagenesis, they observed that Trp 349 appears to contribute the most to the FRET signal but Trp 775 also appears to play a role. They also observed that mutations to residues in the third helix of the PPXD also decreased FRET efficiency, indicating that this region may be important for signal sequence-binding. Subsequently, this same group performed cross-linking studies using a synthetic alkaline phosphatase signal peptide. At the end of the h-region of the signal peptide, benzyolphenylalanine was incorporated as a photoactivatable cross-linker and the C-terminal Cys labeled with biotin. At various regions throughout the SecA protein they incorporated Factor Xa cleavage sites to specifically cleave SecA into two parts. Through cross-linking of the signal peptide to the various SecA mutants, digesting with Factor Xa, and detecting of the biotin, they observed cross-linking to residues 269-322 (Figure 3.2, B, cyan), (Musial-Siwek et al., 2007) which is similar to the site identified by Kimura (Kimura et al., 1991).

Recently, an NMR structure of the SecA homodimer was determined using a domain-parsing strategy in conjunction with specific labeling of the methyl groups of Ile, Leu, Val, and Met (Gelis et al., 2007). To determine the structure of the complex between c-SecA and the signal sequence, Cys mutants of the KRRLamB signal peptide were labeled with a nitroxide spin label. Using methyl-TROSY experiments, the distance-dependent broadening of the NMR signals by the nitroxide-labeled signal peptide were obtained. Changes in the methyl resonance intensities of SecA in the presence of the signal peptide were converted to distances. From these distance constraints, models of
signal peptide bound to SecA were generated. The orientation of the signal peptide bound to SecA was determined by labeling the signal peptide with nitroxide at two different positions, before and after the hydrophobic core. The proposed binding site on SecA is the groove between the PPXD and HWD with the signal peptide interacting predominantly with the PPXD (Figure 3.2, C). The groove has both hydrophobic and electrostatic components (Gelis et al., 2007). This site is not accessible in the original crystal structure because the PPXD packs against the HWD (Hunt et al., 2002) (Figure 3.2, A). In the monomeric version of B. subtilis SecA (Osborne et al., 2004) and as well as the NMR structure (Gelis et al., 2007), the PPXD has rotated slightly and moved so there are no interactions with the HWD (Figure 3.2, C). This binding region modeled from the NMR distance restraints (Gelis et al., 2007) and the site determined by the FRET and cross-linking studies (Musial-Siwek et al., 2005; Musial-Siwek et al., 2007) are in similar regions on the PPXD. This region also has some overlap with the stem region (residues 219-240 and 362-377) (Papanikou et al., 2005) as well as groove 1 from the monomeric crystal structure (Osborne et al., 2004) (Figure 3.2).

All of the signal sequence-binding studies except for Musial-Siwek et al. (Musial-Siwek et al., 2007) have been performed on the cytoplasmic form of SecA, which is not the conformation involved in the translocation reaction (see the introduction to Chapter 2 for a detailed description). A new low resolution structure of truncated SecA bound to the SecYEG translocon has recently been solved (Zimmer et al., 2008) and may provide a framework for signal sequence binding to SecA. From this structure, a new preprotein binding site was proposed. In this SecA structure, the PPXD has rotated even further away from the HWD and makes contact with both SecY and NBF I. The PPXD, NBF II,
and portions of the HSD form a new ‘clamp’ region, which is lined with conserved residues some of which forms a hydrophobic patch (Figure 3.2, D). This region could be used to recognize the signal sequence and position the preprotein above the SecY pore (Zimmer et al., 2008). A previous study (Cooper et al., 2008) using site-directed spin labeling with electron paramagnetic resonance (EPR) spectroscopy with SecA, preproteins, and inverted membrane vesicles supports the preprotein interactions with the ‘clamp’ region.

Although a significant amount of work has been performed to identify the signal sequence-binding site on SecA, the location of the site is still unclear especially in the translocation-active form. Nearly all of the studies have been performed on the cytosolic form of SecA, which is likely not to be the physiologically relevant form of the protein needed for signal sequence binding. Identifying the signal sequence-binding site on the membrane- or translocon-bound form of SecA has been difficult due to the presence of liposomes and other components of the secretory pathway. This chapter describes signal sequence interactions with c-SecA and u-SecA, a soluble translocation-active form of SecA, as well as cross-linking studies to define the signal sequence-binding region on both forms. A possible signal sequence recognition mechanism is proposed based on the results and the recently published SecA/SecYEG crystal structure (Zimmer et al., 2008).

3.2. Experimental strategy

In Chapter 2, we established that low concentrations of urea convert SecA to a state (u-SecA) that mimics the translocation-active form of full-length SecA. U-SecA thus provides a system to determine the signal sequence-binding site on the translocation-
active form of SecA. To gain further insights into the translocation mechanism, the signal sequence-binding site on both the cytosolic and translocation-active forms of SecA must be determined. We have developed a powerful strategy using photoactivatable cross-linking in conjunction with limited proteolysis, two-dimensional gels, biotin and region-specific SecA antibody detection to map the signal peptide binding region on both c-SecA and u-SecA.

In our strategy, the N-terminal biotinylated modified LamB signal peptide, Bio-KRRLamB19C, was labeled with the photoactivatable cross-linker 4-maleimido-benzophenone. The biotin on the signal peptide allows for the specific detection of only the cross-linked products. As shown in Chapter 2, c-SecA is relatively resistant to protease digestion. In order to compare the cross-linking patterns between c-SecA and u-SecA, c-SecA was first cross-linked to signal peptide (Figure 3.3 (A1)) followed by incubation in urea in the same conditions that were used to generate u-SecA (A2) before proteolysis. This enables the cross-linking to occur in the c-SecA form but leads to the same conditions for limited chymotrypsin digestion (A5). U-SecA was generated as described in Chapter 2 (A3), cross-linked to signal peptide (A4), and subjected to limited protease digestion (A5).

Due to the large size of the SecA protein, numerous chymotryptic cleavage sites are anticipated. SecA has 40 potential digestion sites predicted from ExPASy (Gasteiger et al., 2005) with a probability threshold set at 80% based on enzyme specificity for the surrounding sequence. Chymotrypsin preferentially digests at aromatic residues, and the probability of cleavage is based upon the type of amino acid in the sequence before and after the preferred aromatic residue (Keil, 1992). Therefore, chymotrypsin digestion may
Figure 3.3 Experimental design and data analysis

Flow charts of the experiment design for the 2-D gels and for how the data from the 2-D gels was processed.
A. Experimental Design for 2-D Gels
1. Cross-link SP to c-SecA
2. Incubate c-SecA in urea
3. Generate u-SecA
4. Cross-link SP to u-SecA
5. Digest with chymotrypsin
6. Isoelectric focusing
7. SDS-PAGE
8. Transfer to PVDF membrane
9. Biotin Detection
10. Region-specific Ab Detection

B. Analysis of 2-D Gels
1. Manual identification of biotin-positive fragments
2. Comparison of c-SecA and u-SecA biotin fragments
3. ID of similar and different fragments in c-SecA and u-SecA
4. c-SecA biotin fragment ID by Ab mapping
5. u-SecA biotin fragment ID by Ab mapping
6. Compare IDs of similar biotin fragments
generate several different proteolytic SecA fragments of similar molecular weight. This complexity of the SecA digestion pattern makes antibody mapping by one-dimensional SDS-PAGE difficult to resolve.

To circumvent this problem, cross-linked SecA fragments were analyzed by two-dimensional gel electrophoresis. The first dimension is isoelectric focusing (IEF), which separates fragments based on their isoelectric point (Figure 3.3 (A6)), and the second dimension is molecular weight separation by SDS-PAGE (A7). This powerful technique with antibody mapping enabled the identification of different SecA domains involved in binding of signal sequences (A9, A10). The analysis of the two-dimensional gels for cross-linked proteins by biotin detection and antibody mapping is not trivial. Manual analysis of the biotin detection (Figure 3.3, (B1)) of the two-dimensional gels was performed to compare the pattern of cross-linked fragments between c-SecA and u-SecA (B2) and differences in the cross-linking pattern were identified (B3). The identities of the cross-linked SecA fragments were manually assigned using the SecA region-specific antibody data (B4) and the identifications of the cross-linked fragments from c-SecA and u-SecA were compared (B5). Using the mass spectrometry results of the chymotrypsin SecA fragments from Chapter 2, we were able to identify the cross-linked SecA proteolytic fragments (Figure 3.4, A). This powerful approach has enabled us to determine the signal sequence-binding region on both the soluble and translocation-active forms of SecA.
Figure 3.4 Map of chymotrypsin cleavage sites

Based on the surface accessibility of the aromatic residues and mass spectrometry results from Chapter 2, the chymotrypsin cleavage sites are shown. The linear sequence of SecA is color-coded for the different domains and the region-specific antibodies are indicated above the cartoon. The relative cleavage at each site is indicated by the thickness of the arrow. B. The cartoons illustrate the different proteolytic fragments of SecA cross-linked to signal peptide in the 2-D gels. The SecA cross-linked fragments are broken down into three subsets based on domains. The name of each fragment is indicated on top of the cartoon.
3.3. Results

3.3.1. Establishing the validity of the SecA-signal peptide interaction

Although signal peptide binding to c-SecA has previously been established (Gelis et al., 2007), we needed to confirm that signal peptides bind to u-SecA in a specific manner. Due to the complications of working with the highly hydrophobic signal peptide and working in urea solutions, typical binding assays were unsuccessful so we developed a qualitative gluteraldehyde cross-linking based binding assay. This method uses gluteraldehyde to cross-link a biotinylated KRRLamB19C signal peptide (Figure 1.2, B) to c-SecA and u-SecA. Therefore, if the signal peptide is bound to either conformation of SecA, the protein and signal peptide will be cross-linked upon addition of glutaraldehyde.

The cross-linked samples were run on tricine SDS-PAGE gels and transferred to a PVDF membrane. The biotin signal from the signal peptide was detected by streptavidin-horseradish peroxidase (HRP). Only cross-linked SecA protein, which migrated around 100 kD for the monomer or about 200 kD for the dimer are detected by streptavidin-HRP. In this set of experiments, the amount of Bio-KRRLamB19C signal peptide was held constant while the concentration of c-SecA or u-SecA was increased. For both forms of SecA, the biotin signal increased as the concentration of protein was increased (Figure 3.5). As shown in Chapter 2, c-SecA is predominantly in a dimeric form but in this signal peptide binding assay the small amount of monomeric c-SecA also binds signal peptide (Figure 3.5, A, B). This is likely to be a non-specific binding event since the signal of the monomer does not increase as the protein concentration is increased.
Figure 3.5 Signal peptide binding assay

Glutaraldehyde cross-linking was used to assess if either form of SecA can bind a biotinylated signal peptide. A. Biotin detection (top) and Coomassie blue stained SDS-PAGE gel (bottom) of increasing concentrations of c-SecA cross-linked to 1.0 µM Bio-KRRLamB19C signal peptide. The c-SecA concentrations in the lanes are as follows: (1) 0.13 µM, (2) 0.21 µM, (3) 0.44 µM, (4) 0.49 µM, (5) 0.61 µM, (6) 0.87 µM, (7) 1.0 µM, (8) 1.41 µM, (9) 1.81 µM. Dimeric SecA migrates at 204 kD while monomeric SecA migrates at 102 kD. B. Quantitation of the biotin signal from two different experiments. Due to variation in the protein concentrations only six concentrations are shown. Dimeric SecA is shown by the red bars while monomeric SecA is shown by the blue bars. C. Biotin detection (top) and Coomassie blue stained SDS-PAGE gel (bottom) of increasing concentrations of u-SecA cross-linked to Bio-KRRLamB19C signal peptide. The u-SecA concentrations in the lanes are as follows: (1) 0.01 µM, (2) 0.14 µM, (3) 0.41 µM, (4) 0.68 µM, (5) 1.0 µM, (6) 1.25 µM, (7) 1.29 µM, (8) 1.47 µM, (9) 1.63 µM. Dimeric SecA migrates at 204 kD while monomeric SecA migrates at 102 kD. D. Quantitation of the biotin signal from two different experiments. Due to variation in the protein concentrations only six concentrations are shown. Dimeric SecA is shown by the red bars while monomeric SecA is shown by the blue bars.
In u-SecA most of the signal peptide binds to the monomer but as the total protein concentration increases more dimer is present (Figure 3.5, C, D). The titrations cannot be performed at higher protein concentrations because c-SecA forms higher oligomers (Shin et al., 2006) and u-SecA begins to aggregate non-specifically. In both cases, the functional form, dimer for c-SecA and monomer for u-SecA, appears to saturate around a protein concentration of one µM. This qualitative value is comparable to binding affinities previously reported. (Musial-Siwek et al., 2005; Gelis et al., 2007). We concluded that signal peptide binds to both c-SecA and u-SecA with similar affinities. This result was surprising at first because previous work showed that an activated form of SecA, SecA64, which lacks the C-terminal one-third of the protein, bound signal peptides better than full-length c-SecA (Triplett et al., 2001). Therefore, u-SecA in which the C-terminal portion has undergone domain dissociation might also have been expected to bind signal peptides with a higher affinity. However, u-SecA is generated by low concentrations of urea, and besides pushing SecA into the activated form, the urea could also weaken signal peptide binding. Urea weakens non-covalent interactions and since signal sequence recognition is mediated through hydrophobic interactions, the binding affinity to u-SecA is likely decreased. In the case of c-SecA, part of the signal peptide adopts a helical conformation upon binding (Chou and Gierasch, 2005; Gelis et al., 2007) and in u-SecA the helical structure may be disrupted resulting in a decreased binding affinity.

To test the specificity of c-SecA and u-SecA signal peptide interaction, a qualitative competition assay similar to the assay as described above was performed. The
concentration of both forms of SecA and the Bio-KRRLamB19C peptide was held constant at 1 µM and increasing amounts of the non-biotinylated signal peptide, KRRLamB WT (Figure 1.2, B), were added as a competitor. If the binding of the biotinylated signal peptide is specific, it should be displaced from SecA with the addition of the non-biotinylated signal peptide and result in a decrease in the biotin signal.

For c-SecA, the non-biotinylated signal peptide starts to compete with the biotinylated peptide to a small degree with equimolar amounts. When 20-fold excess KRRLamB WT is added, a little less than three-quarters of the biotinylated peptide is displaced (Figure 3.6, A, B). The concentration of signal peptide cannot be increased beyond 20-fold excess because it starts to induce higher oligomers/aggregation of c-SecA. The competing peptide was equilibrated with the c-SecA for one minute because longer equilibration times showed little differences in the biotin signal. This suggests that the biotinylated form of the signal peptide may have a slightly higher affinity for c-SecA than the non-biotinylated form as well as has a relatively fast on/off rate.

In the case of u-SecA, the equimolar amount of non-biotinylated signal peptide competes off about one-quarter of the biotinylated signal peptide. By 15-fold excess of the non-biotinylated signal peptide, more than three-quarters of the biotin signal has disappeared indicating that a majority of the biotinylated signal peptide has been displaced (Figure 3.6, C, D). If higher concentrations of signal peptides are used, u-SecA is pushed into the dimeric and aggregated forms. This dimerization/aggregation phenomenon with monomeric u-SecA is similar to SecA that was monomerized by phospholipid analogs. Monomeric SecA associated to the phospholipids is also dimerized/polymerized with the addition of high concentrations of signal peptide.
Figure 3.6 Signal peptide competition assay

Competition between biotinylated signal peptide and non-biotinylated signal peptide was monitored by gluteraldehyde cross-linking. **A.** Biotin detection (top) and Coomassie blue stained SDS-PAGE gel (bottom) of increasing concentrations of the competing KRR LamB WT signal peptide cross-linked to 1.0 µM each c-SecA and Bio-KRR LamB19C signal peptide. The KRR LamB WT concentrations in the lanes are as follows: (1) 0 µM, (2) 1.0 µM, (3) 2.5 µM, (4) 5.0 µM, (5) 10.0 µM, (6) 12.5 µM, (7) 15.0 µM, (8) 17.5 µM, (9) 20.0 µM. Dimeric SecA migrates at 204 kD while monomeric SecA migrates at 102 kD. **B.** Quantitation of the biotin signal in A based on three separate experiments. Dimeric SecA is shown by the red bars while monomeric SecA is shown by the blue bars. **C.** Biotin detection (top) and Coomassie blue stained SDS-PAGE gel (bottom) of increasing concentrations of the competing KRR LamB WT signal peptide cross-linked to 1.0 µM each u-SecA and Bio-KRR LamB19C signal peptide. The KRR LamB WT concentrations in the lanes are as follows: (1) 0 µM, (2) 1.0 µM, (3) 2.5 µM, (4) 5.0 µM, (5) 10.0 µM, (6) 12.5 µM, (7) 15.0 µM. Dimeric SecA migrates at 204 kD while monomeric SecA migrates at 102 kD. **D.** Quantitation of the biotin signal in C based on two separate experiments. Dimeric SecA is shown by the red bars while monomeric SecA is shown by the blue bars.
(Benach et al., 2003). For u-SecA the competing signal peptide needed to be equilibrated with u-SecA for one hour prior to cross-linking otherwise little competition was observed. This observation suggests that the signal peptide has a slow off-rate from u-SecA.

Due to the highly hydrophobic nature of the signal peptide, it is not surprising that all of the biotinylated signal peptide is not displaced in the competition assay. There is likely to be some level of non-specific interaction between SecA and the signal peptide that cannot be competed off of SecA in the gluteraldehyde cross-linking assay. Overall, these results demonstrate that c-SecA and u-SecA directly and specifically bind signal peptide.

3.3.2. Analysis of SecA fragments cross-linked to signal peptide

In order to determine the signal sequence-binding region on c-SecA and u-SecA, we employed a photoactivatable cross-linking approach. The Bio-KRRLamB19C signal peptide was labeled with 4-maleimido-benzophenone. Irradiation with UV light leads to a benzophenone diradical that has the ability to abstract a hydrogen atom from a nearby electron-rich σ-bond and upon recombination, forms a covalent linkage. If no suitable bond is present the diradical reacts with water making benzophenone (BP) a highly specific cross-linking reagent (Dorman and Prestwich, 1994).

To discern if the signal sequence-binding site is different in c-SecA and u-SecA, we developed a strategy to compare the cross-linked proteolytic fragments of both forms of SecA (Figure 3.3, A (1-5)). This strategy enables us to determine differences in the cross-linking pattern between c-SecA and u-SecA. Any differences can be attributed to
alternative cross-linking and not differences in protease digestion between the two different conformations of SecA.

In our initial studies, cross-linked proteins, c-SecA (Figure 3.7) and u-SecA (Figure 3.8), were subjected to limited chymotrypsin digestion over a time course ranging from 10 sec to 10 min. The digests were separated on tricine SDS-PAGE gels and analyzed by biotin and SecA region-specific antibody detection. The overall pattern of digestion, biotin, and antibody detection between c-SecA and u-SecA is similar but there are differences in the lower molecular weight bands in the biotin detection. The two minute time point of digestion (Figure 3.9) shows the biotin detection of SecA fragments around 66 kD corresponds to the A1, A2, A3, and A4 antibodies (see Chapter 2, Figure 2.8), which represents the N-terminal two-thirds of the protein as previously demonstrated with SecA64. A smaller proteolytic fragment of about 45 kD is recognized by the A1 and A2 antibodies. These cross-linked SecA fragments are present in both c-SecA and u-SecA. Smaller molecular weight bands, approximately 30 and 20 kD are detected only in u-SecA. The antibodies that correspond to these bands are A5 and A6 for the 30 kD band and A3 and A4 for the 20 kD band (Figure 3.9, B red boxes). SecA has 40 potential chymotrypsin digestion sites (calculated from ExPASy with a threshold set at 75% (Gasteiger et al., 2005)) so several different proteolytic fragments of similar molecular weight can be generated. Therefore, mapping the identity of the cross-linking fragments by 1-D gel antibody detection is difficult.

SecA is an acidic protein with a theoretical isoelectric point (pI) of 5.43 (calculated from ExPASy) (Gasteiger et al., 2005). To further separate and identify the cross-linked SecA fragments, we chose to use two-dimensional gels. Isoelectric focusing
Cross-linking was performed as c-SecA but then incubated in 2.2 M urea for four hours to have a similar chymotrypsin pattern as u-SecA. The lanes are as follows: (1) no digestion, (2) 10 sec, (3) 1 min, (4) 2 min, (5) 3 min, (6) 4 min, (7) 5 min, (8) 10 min. The labels of the different panels correspond to the different antibodies (A1 – A6), Coomassie blue stain (C) and Biotin detection (B). The molecular weight standard is indicated to the left of the Coomassie blue stained gel. A control of chymotrypsin alone was also included in the experiment but did not react with any of the detection methods so it is not shown.
Figure 3.8 Limited protease digestion of signal peptide cross-linked to u-SecA

After incubated in 2.2 M urea for four hours to generate u-SecA, cross-linking was performed followed by protease digestion. The lanes are as follows: (1) no digestion, (2) 10 sec, (3) 1 min, (4) 2 min, (5) 3 min, (6) 4 min, (7) 5 min, (8) 10 min. The labels of the different panels correspond to the different antibodies (A1 – A6), Coomassie blue stain (C) and Biotin detection (B). The molecular weight standard is indicated to the left of the Coomassie blue stained gel. A control of chymotrypsin alone was also included in the experiment but did not react with any of the detection methods so it is not shown.
Figure 3.9 Bio-KRRLamB19C-MBP cross-linking to both c-SecA and u-SecA

A. The two minute time point from c-SecA limited protease digestion. The molecular weight marker is shown on the left and each lane is labeled with the corresponding detection, A1 – A6 antibodies, biotin detection (B), and Coomassie blue stain (C). B. The two minute time point from u-SecA limited protease digestion. The molecular weight marker is shown on the left and each lane is labeled with the corresponding detection as described above. The red boxes indicate biotin detected fragments that are only found in u-SecA.
(IEF) was performed using pH 4-7 IEF strips (Figure 3.3, (A6)). After IEF the SecA fragments were further separated by tricine SDS-PAGE (A7). The 2-D gels were transferred to polyvinylidene fluoride (PVDF) membrane (A8) for biotin (A9) and region-specific antibody (A10) detection.

We analyzed the 2-D biotin detection of c-SecA and u-SecA blots (Figure 3.10) for cross-linked SecA fragments (Figure 3.3 (B1)) as described in the Materials and Methods. Only SecA fragments smaller than 50 kD were analyzed. Out of 21 potential fragments found on each blot, 12 fragments in c-SecA (Figure 3.11 A) and 18 fragments in u-SecA (Figure 3.11 C) were calculated to be above background signal (see Materials and Methods for description). The intensity of each spot was quantified (Figure 3.11 B, D) using ImageJ software (Abramoff et al., 2004) normalizing the data to the free signal peptide found at the bottom of both blots.

C-SecA (Figure 3.11 A) and u-SecA (Figure 3.11 B) biotin detection blots were aligned using the Flicker software (Lemkin et al., 2005). Since focusing differences can arise during IEF, the Flicker software enables the alignment of two different gels as described in the Material and Methods. Flickering between the two gel images allows for visualization of overlapping cross-linked fragments in c-SecA and u-SecA. Several SecA fragments are the same in both biotin detections (S1-S13) but some of these fragments (S8-S11 and S13) were observed in c-SecA but were not above background signal to be defined as a fragment. In addition to having similar proteolytic fragments, three unique fragments in c-SecA, C1-C3, and five unique fragments in u-SecA, U1-U5 were found (Figure 3.11).
Figure 3.10 2-D gel biotin detection of signal peptide cross-linked to c- and u-SecA

Bio-KRRLamB19C-MBP signal peptide was cross-linked to both c-SecA and u-SecA followed by digestion with chymotrypsin for two minutes. The digestions were separated by isoelectric focusing, 4-7 pH range, followed by molecular weight separation with 8% SDS-PAGE gels. **A.** Biotin detection of c-SecA proteolytic fragments. Yellow circles indicate potential spots (see materials and methods for spot determination). **B.** Same as in A but with u-SecA and orange circles indicating potential spots.
Figure 3.11 Analysis similar and different fragments in c- and u-SecA

A. The biotin detection of 2-D gel for cross-linked c-SecA. Spots that are similar to u-SecA are S1-S12 and spots that are different are C1-C3. Not all of the similar spots in c-SecA are above background. B. Quantitation of the spots shown in A using the ImageJ software. Each spot was measured for biotin signal three times. C. The biotin detection of 2-D gel for cross-linked u-SecA. Spots that are similar to c-SecA are S1-S13 and spots that are different are U1-U5. D. Quantitation of spots shown in C. Each spot was measured for biotin signal three times.
After biotin detection, each blot was probed with the different region-specific SecA antibodies. This enables the putative mapping of SecA fragments that are cross-linked to signal peptide. Both the biotin and antibody probes were performed on the same blot, thus, allowing for the biotin probed gel to be aligned separately with each of the antibody detections using the Flicker software (Lemkin et al., 2005). The flickering between the two images enabled visualization of overlaying fragments between the biotin and antibody images. From this method, the SecA domain corresponding to the proteolytic fragment was determined. One striking feature that is apparent when comparing the biotin detection to the antibody detection is that the cross-linking is amazingly specific. Only a small subset of the all SecA proteolytic fragments were cross-linked to the signal peptide.

All the similar cross-linked fragments in c-SecA (Figure 3.12) and u-SecA (Figure 3.13) are recognized by exactly the same antibodies with the exception of the S12. This fragment was only recognized by the A4 antibody in c-SecA but was identified by the A3 and A4 antibodies in u-SecA. Using the mass spectrometry data from Chapter 2 and the antibody detection data, we were able to determine the SecA chymotryptic fragments that were cross-linked to signal peptide (Figure 3.4, B). One subset of fragments represents cross-linking to the NBF I and PPXD domains of SecA. The S1-S5 and S7-S8 fragments are recognized by the A1 and A2 antibodies. S1, S2, and S3 fragments, about 40 kD in size, have slightly different isoelectric points (Figures 3.12 and 3.13). These SecA proteolytic fragments represent residues 11-405 (Figure 3.4 B) but have different cleavage sites a few residues away at one or both of the termini. The S4
Figure 3.12 Antibody detection of 2-D gels of c-SecA cross-linked to signal peptide

2-D gels were detected for biotin (B) and with the region-specific antibodies. The digested 1-D sample stained with Coomassie blue (C) with the molecular weights is shown to the left of the biotin image. The pH range for the IEF is shown on the top of the images. The lane on the left side of the image is cross-linked c-SecA sample that was not subjected to IEF. The yellow circles indicate the positions of the different SecA proteolytic fragments identified in the biotin image.
Figure 3.13 Antibody detection of 2-D gels of u-SecA cross-linked to signal peptide

2-D gel were detected for biotin (B) and with the region-specific antibodies. The digested 1-D sample stained with Coomassie blue (C) with the molecular weights is shown to the left of the biotin image. The pH range for the IEF is shown on the top of the images. The lane on the left side of the image is cross-linked u-SecA sample that was not subjected to IEF. The yellow circles indicate the positions of the different SecA proteolytic fragments identified in the biotin image.
and S5 fragments are approximately 24 kD in size and represent smaller proteolytic fragments of S1, S2, or S3. S4 and S5 include amino acids 80-299 and one of the fragments has a slightly different cleavage site resulting in a different isoelectric point. S7, about 20 kD in size, and S8, approximately 17 kD, are two further proteolytic cleavages of the S1-S3 fragments (Figure 3.4, B). S7 is only detected by the A1 antibody while S8 is exclusively detected by A2.

Another subset of similar SecA proteolytic fragments cross-linked to signal peptide in both c-SecA and u-SecA maps to NBF II. The approximately 19 kD S12 fragment is recognized by the A3 and A4 antibodies and includes residues 406-598. A second fragment, S13, is eight kD and has been further proteolyzed from the S12 fragment. One other cross-linked fragment is the same in c-SecA and u-SecA belongs to the third subset, the HSD and HWD. S6 is about 20 kD and is recognized only by the A5 antibody. This fragment includes residues 623-794 (Figure 3.4, B)

Several biotin-detected fragments are specific for c-SecA and u-SecA. The three unique fragments found in c-SecA (Figure 3.12) belong to the NBF II subset. The larger C2 and C3 fragments, approximately 43 kD, are recognized by the A2, A3, and A4 antibodies and correspond to residues 300-691. C2 and C3 are similar in molecular weight but differ slightly in isoelectric point indicating a small difference in digestion. A smaller unique c-SecA cross-linked fragment, C1 about 19 kD, represents a proteolytic fragment from C2 or C3. C1 is very similar to S12 but has a different isoelectric point due to a different chymotrypsin cleavage site.

In u-SecA five unique cross-linked fragments are observed (Figure 3.13) but only three of the five fragments were identified by antibody mapping. The two unidentified
fragments have low biotin signal (Figure 3.11, D) and probably do not contain enough SecA protein to be detected by the antibodies. The U1 fragment is in the NBF I and PPXD cross-linking subset. U1 is similar to S1, S2, S3 in molecular weight, isoelectric point, and antibody detection. Therefore, U1 is similar to the S1, S2, and S3 fragments but has a slightly different chymotrypsin digestion site. The U2 fragment is a member of the NBF II subset and is roughly 15 kD. This fragment is exclusively recognized by the A3 antibody and represents amino acids 406-541. U2 is very similar to S12 and S13 fragments and is a further proteolytic fragment of S12. The third u-SecA specific cross-linked fragment is part of the HSD and HWD subset. U4 is about 15 kD, recognized by A5 antibody, and corresponds to residues 623-762. This fragment is a smaller proteolytic piece from the S6 fragment.

These results demonstrate that c-SecA and u-SecA share the same signal-sequence binding site. Signal peptide cross-linking is observed to three regions of the protein indicating that the binding region is at the interface between NBFI, NBFI, PPXD, and N-terminal HSD.

3.4. Discussion

The mechanistic details of preprotein translocation through the SecYEG translocon by SecA are poorly understood. Activation of SecA by the addition of low concentrations of urea has provided us with a soluble translocation-active form with which to examine a critical step in the translocation process, signal sequence recognition. Several previous studies have probed signal sequence binding to the cytosolic form of SecA (Kimura et al., 1991; Triplett et al., 2001; Hunt et al., 2002; Baud et al., 2002; Chou
and Gierasch, 2005; Papanikou et al., 2005; Gelis et al., 2007) but only one study by Musial-Siwek et al. (Musial-Siwek et al., 2007) has tried to identify the signal sequence-binding site on an activated form of SecA. This group performed cross-linking studies in lipid vesicles but based on its ATPase rate lipid-bound SecA is not in the same conformation as translocation-active SecA (Lill et al., 1990).

These data argue that the signal sequence-binding site is located at the interface between NBF I, NBF II, PPXD, and HSD domains, which maps to an area similar to the ‘clamp’ region proposed by Rapoport et al. (Zimmer et al., 2008) in their recent crystal structure of truncated SecA bound to SecYEG. This result is in conflict with previously determined signal sequence-binding sites (Musial-Siwek et al., 2007; Gelis et al., 2007).

We used benzophenone-mediated (BPM) cross-linking combined with limited chymotrypsin digestion in conjunction with biotin and antibody detection to map the signal sequence-binding site. Due to the heterogeneous nature of the different proteolytic fragments of SecA, 2-D gels enabled us to examine SecA fragments of similar molecular weight since fragments are separated by IEF as well as molecular weight.

The length of the maleimido-benzophenone cross-linker from the backbone of the residue to which it is attached to the end of the benzophenone (BP) moiety is 12 Å. BP can cross-link to electron-rich σ bonds in the side chains Leu, Val, Arg, and Lys (Dorman and Prestwich, 1994) but is most likely to cross-link to Met (Wittelsberger et al., 2006). To correlate our BPM cross-linking data to the NMR model of signal peptide bound to SecA, all of the favorable cross-linking residues within 20 Å of Met 19 in the NMR structure were examined. From this analysis, cross-linked SecA fragments should be recognized predominantly by the A1, A2, and A5 antibodies. In the 2-D gel analysis,
proteolytic fragments are strongly recognized by the A1, A2, A3, and A4 antibodies. Two fragments, S6 and U4, are recognized by the A5 antibody and both fragments have a low biotin signal meaning that this was not a predominant cross-linking site. One possibility for the lack of cross-linking to the HWD (A5 antibody recognition) is that the orientation of the signal peptide does not allow for cross-linking to this region but if the NMR model is correct then we should only have cross-linking to SecA fragments recognized by the A1 and A2 antibodies. An alternative explanation is that the fragments recognized by the A3 and A4 antibodies are rich in surface-exposed Met residues, a favored cross-linker for BP. All of the surface-accessible side chains of the Met from residues 351 to 664 in the NMR structure were examined and only two surface-exposed Met were found. Both of these Met are a great distance away from the region that was modeled to be the signal sequence-binding site. Therefore, our data clearly suggest that the signal peptide is binding in a different region than the one model by NMR or that there are actually two signal sequence-binding sites.

The BPM cross-linking data from the 1-D gels showed differences in signal peptide cross-linking to c-SecA and u-SecA. U-SecA has smaller molecular weight biotin positive bands (Figure 3.9) despite both c-SecA and u-SecA being digested in the same condition. Although unique c-SecA and u-SecA cross-linked fragments were present in the biotin detection, further analysis shows that the unique fragments are similar to the other identified S-fragments (Figure 3.4). This result demonstrates that c-SecA and u-SecA share the same signal sequence-binding site.

Overall, u-SecA contains a greater number of unique biotin-positive fragments and this could result from a difference in chymotrypsin digestion or more favorable
cross-linking residues exposed to the signal peptide in the translocation-active form. A difference in chymotryptic digestion can be ruled out since the digestion pattern and relative band intensities in the Coomassie blue stained gels are identical for u-SecA and c-SecA (Figure 3.9). The greater number of cross-links to u-SecA is likely due to more favorable residues closer to the signal peptide. This suggests that the signal-sequence binding groove in u-SecA has expanded. This would explain why some SecA natural substrates such as LamB with the native signal sequence bind to cytosolic SecA with low affinity but bind with high affinity to the activated SecA (Triplett et al., 2001).

Hunt et al. (Hunt et al., 2002) proposed the single-stranded nucleic acids (SS-NA) groove (Figure 3.1 A) as a potential site for preprotein binding. The ‘SS-NA groove’ occupies the interface of NBF I, NBF II, PPXD, and HSD. This is consistent with our data regarding the binding site for signal peptide in both c-SecA and u-SecA but in u-SecA the domain dissociation causes the site to become extended. The ‘SS-NA groove’ in SecA is similar to other ssDNA grooves found in DNA helicases (Korolev et al., 1997; Velankar et al., 1999), which have structural homology with the NBFs of SecA. Signal sequence-binding to the SS-NA groove imply a molecular operate through a mechanism similar to the helicases, such that the motor domains power the translocation of preproteins in SecA as they unwinding of DNA in the helicases.

Based on our data, we propose a model where signal sequences bind to the SS-NA region originally proposed by Hunt et al. (Hunt et al., 2002). The binding region is the same in both the cytosolic and translocon-bound forms but in the latter form, the binding region has been expanded which could allow for better signal sequence-binding. This binding region is also in close proximity to the nucleotide-binding site. The signal
sequence binding affinity can therefore be modulated by the presence of different nucleotides in the binding pocket (Shin et al., 2006). Additionally, the signal sequence can suppress the rate of ATPase activity as seen in the case of SecA64 (Triplett et al., 2001). This SS-NA binding region is in a region similar to the clamp region proposed by Zimmer et al. (Zimmer et al., 2008) based on the truncated SecA/SecYEG crystal structure. Therefore, we propose a mechanism where the signal sequence binds to the SS-NA groove and the mature portion of the preprotein binds to the PPXD. After the proofreading step of signal sequence binding, ATP hydrolysis releases the signal sequence and pushes it into the translocon.

During the course of this study, several experimental problems had to be resolved. A critical experiment was verifying that u-SecA interacts with signal peptide in a specific manner. This experiment was problematic do to the highly hydrophobic nature of the signal peptide and the propensity of u-SecA aggregate at protein concentrations above 2.5 µM. Typical methods to assess binding such as size exclusion chromatography, isothermal titration calorimetry, and equilibrium dialysis were unsuccessful. To overcome this problem, we developed the qualitative gluteraldehyde cross-linking binding assay. Although this assay does not provide a binding constant, we showed that u-SecA indeed binds to signal peptide in a specific manner to further confirmed that u-SecA is a translocation-active form of SecA.

Another major obstacle we overcame was determining the region on c-SecA and u-SecA that was cross-linked to signal peptide. As shown in Chapter 2, c-SecA is less proteolytically labile than u-SecA and therefore, comparison of the cross-linked chymotryptic fragments was difficult. We developed a strategy (Figure 3.3, A (1-5)) that
enabled us to cross-link signal peptide to c-SecA and digest this cross-linked product in the same manner as u-SecA. One-dimensional SDS-PAGE revealed that the digestion pattern was identical but as demonstrated in MS experiments in Chapter 2, the heterogeneous nature of the chymotryptic SecA fragments generates different proteolytic fragments of the same size. This heterogeneous population made the antibody mapping difficult to interpret that prompted us to used two-dimensional gels to resolve this problem since the SecA proteolytic fragments of similar size have different isoelectric points (Figure 3.3, A (5-10)). This method helped us identify the SecA domains that were cross-linked to signal peptide. Unfortunately, MS experiments to identify the residues in SecA cross-linked to signal peptide were unsuccessful. Due to the high specificity of BP, the cross-linking efficiency is low, in our case less than 10%. We tried to purify the cross-linked SecA products using a monomeric streptavidin column but the presence of urea decreased biotin affinity for the column and therefore, cross-linked u-SecA could not purified. In the MS experiments several potential cross-linked SecA fragments were identified but we were unsuccessful in obtaining sequence information by MS/MS. The masses of the potential cross-linked fragments were mapped to the SecA sequence but these fragments either did not contain a favorable cross-linking residue or the size of the fragment was similar to two or more SecA proteolytic fragments. Thus, we relied on our domain mapping strategy using region-specific antibodies to identify sites of cross-linking, which provided us a powerful method to determine the signal sequence-binding site on c-SecA and the translocation-active u-SecA.

3.5. Materials and Methods
3.5.1. **SecA titration Assay**

C-SecA at 2.5 μM protein concentration in 20 mM HEPES pH 7.5, 50 mM KCl, 1 mM MgCl$_2$, 1 mM DTT and u-SecA at 2.5 μM in 20 mM HEPES pH 7.5, 50 mM KCl, 1 mM MgCl$_2$, 1 mM DTT, 2.2 M urea were generated by incubated at 22°C for four hours. The biotinylated signal peptide, Bio-KRRLamB19C (Bio-MMITLRKRRKLPLAVAVAAGVCSAQAMA), was held constant at one μM while the concentration of c-SecA and u-SecA from the previous equilibration step was titrated from 0.1 μM to 2 μM in a reaction volume of 40 μl. The samples were equilibrated for two hours at 22°C. After equilibration the samples were cross-linked by the addition of gluteraldehyde to a final concentration of 0.1% and incubated at room temperature for two minutes. Cross-linking reaction was quenched by the addition of 100 mM Tris pH 8.0 and each sample was run in duplicate on two different 6% tricine SDS-PAGE gels. One gel was transferred to PVDF membrane while the other gel was stained with Coomassie Blue or with SYPRO Ruby (BioRad).

3.5.2. **Biotin detection**

Membranes were blocked overnight at 4°C in 5% non-fat milk in phosphate buffered saline with 0.1% tween-20 (PBS-T). The next day the membranes washed, and probed with a streptavidin-HRP conjugate (GE Healthcare) at a 1:1000 dilution in PBS-T for two hour at room temperature for gluteraldehyde cross-linking or overnight at 4°C for 1-D and 2-D gels. The cross-linked bands were detected using the SuperSignal West Pico kit (Pierce, Rockford IL) following the manufacture’s instructions. The signal was
visualized and recorded using a G:Box gel documentation unit (Syngene, Frederick, MD).

3.5.3. Protein staining and visualization

Gels were stained with either Coomassie Blue or SYPRO ruby protein stain (BioRad) overnight at room temperature with gentle shaking. The next day the gels were destained according to the manufacture’s protocol. For visualization the SYPRO ruby dye was excited with 254 nm light and recorded using a G:Box gel documentation unit (Syngene, Frederick, MD). Coomassie Blue stain was recording using white light in the G:Box gel documentation unit.

3.5.4. Biotin signal quantification and normalization

The signal from the biotin detection was quantified using ImageJ software (http://rsb.info.nih.gov/ij/docs/index.html) (Abramoff et al., 2004). The images were background subtracted and inverted so that the background was black and the fragments were gray/white. The rectangle tool was used to measure the area and the mean-gray value (pixels) in the cross-linked protein band (the size of the rectangle depended on the size of the band). For each measured band a corresponding rectangle of the same area was taken of the local background. The ImageJ software output is the area, the mean-gray value, and the absolute intensity (AI), which is calculated by the program as a product of the area times and mean-gray value. The relative intensity of each band was calculated by dividing the corrected absolute intensity by the absolute intensity of one µM Bio-KRRLamB19C cross-linked to one µM SecA. The data was normalized for protein
concentration by quantification of the SYPRO Ruby stained or Coomassie Blue stained gel following a similar procedure outline above for biotin quantification.

3.5.5. Signal peptide competition assay

One µM of both c-SecA and u-SecA (in the buffers described above) were incubated at 22°C for four hours. Bio-KRRlamB19C signal peptide at one µM was added to both c-SecA and u-SecA and incubated at 22°C for 15 min. The samples were added to tubes containing KRRlamB WT signal peptide (MMITLRKRRKLPLAVAVAAGVMSAQAMA) from 0 – 20 µM. c-SecA was incubated with the non-biotinylated signal peptide for one min before addition of 0.1% gluteraldehyde. Cross-linking occurred at room temperature for two minutes before quenching with 100 mM Tris pH 8.0. In the case of u-SecA, the samples were incubated with the non-biotinylated signal peptide for one hour before cross-linking and quenching. Each sample was run in duplicate on two different 6% tricine SDS-PAGE gels. One gel was transferred to PVDF membrane while the other gel was stained with Coomassie Blue. Quantification of the biotin bands was performed as described above and the data were normalized for protein concentration from the Coomassie blue stained gels.

3.5.6. Labeling of signal peptide with maleimido-benzophenone

The Cys at position 19 of one mM Bio-KRRlamB19C signal peptide (GL Biochem (Shanghai) Ltd) in 20 mM HEPES pH 7.5, 1.2 mM TCEP buffer was alkylated by the addition of 1.5 mM 4-male imido-benzophenone (MBP) (Sigma). The reaction was stirred for 30 minutes at room temperature and the reaction was quenched by the addition of TFA to lower the pH to ~2. The MBP-labeled peptide was purified by preparative RP-
HPLC using a phenyl (25x250 mm, 10 mm, 300 Å, Vydac) column using an appropriate water/acetonitrile gradient. The purified MBP-labeled peptide was checked by analytical RP-HPLC with a phenyl (4.6x150 mm, 5 mm, 300 Å, Vydac) column using an appropriate water/acetonitrile gradient demonstrating that the purity of the MBP-labeled peptide was greater than 95%. The identity and purity of the peptide was confirmed by mass spectrometry analysis using an Esquire-LC Ion trap mass spectrometer (Bruker Daltonics, Billerica, MA).

3.5.7. BPM cross-linking experiments

In order to determine if there are differences in the signal peptide binding region between c-SecA and u-SecA, Bio-KRRLamB19C-MBP signal peptide was cross-linked to both forms. In the absence of light, 2.5 µM Bio-KRRLamB19-MBP was added to 2.5 µM c-SecA in cross-linking buffer (20 mM HEPES pH 7.5, 50 mM KCl, 1 mM MgCl$_2$, 1 mM DTT) and incubated for 15 min at room temperature to allow for complex formation. The sample was then illuminated with 365 nm light at ~5 cm distance for one minute to induce cross-linking. To obtain a chymotrypsin digestion pattern similar to u-SecA, urea was added to the cross-linked c-SecA sample to a final concentration of 2.2 M and incubated at 22°C for four hours. Prior to cross-linking, u-SecA was generated by incubation of 2.5 µM of SecA in 2.2 M urea cross-linking buffer for four hours at 22°C. Cross-linking of u-SecA was performed as described above followed by the addition of 2.2 M urea buffer to obtain the same final volume as the c-SecA sample.

3.5.8. Protease digestion and TCA precipitation
Samples were digested with chymotrypsin (1:150 w/w) at room temperature for either a time course of 10 sec to 10 min for 1-D gel analysis or for two minutes for 2-D gel analysis. Digestion was quenched by the addition of 5 mM of the protease inhibitor ABESF. For 2-D gels an aliquot of each sample was taken, run on an 8% tricine SDS-PAGE gel, and stained with Coomassie Blue to ensure proper digestion. The samples were precipitated by the addition of one volume 100% trichloroacetic acid (TCA) to four volumes of sample and then incubated overnight at -20°C. The precipitated proteins were pelleted by a spin at 15,800 x g in a microcentrifuge. The pellet was washed with cold acetone two times and dried at 95°C for 10 minutes. The dried protein pellets were resuspended in 3X SDS-PAGE buffer for 1-D gel analysis or in rehydration buffer for 2-D gel analysis.

3.5.9. 1-D gels

After resuspension in the 3X SDS-PAGE buffer, the TCA precipitated samples were boiled in a water bath for five minutes. Each sample was run in duplicate on two different 6% tricine SDS-PAGE gels. One gel was transferred to PVDF membrane while the other gel was stained with Coomassie Blue. The membranes were probed streptavidin-HRP, as described above, and the different region-specific SecA antibodies as described in the previous chapter.

3.5.10. 2-D gels

TCA pellets for both c-SecA and u-SecA were resuspended in rehydration buffer (8 M urea, 2% CHAPS, 0.4% Biolytes). Isoelectric focusing (IEF) strips (7 cm) pH range 4-7 were rehydrated with c-SecA and u-SecA samples according to manufacture’s
protocol (Method 1 Passive rehydration, BioRad) and incubated at room temperature for 14-16 hours. The strips were focused using a linear voltage gradient according to the manufacture’s program (BioRad). After focusing the strips were frozen at -80°C until further use.

Prior to running the second dimension, the strips were thawed at room temperature for 10 min. The strips were incubated in equilibration buffer I (6 M urea, 0.375 M Tris pH 8.8, 2% SDS, 20% glycerol, 2% (w/v) DTT) for 10 minutes at room temperature followed by incubation in equilibration buffer II (6 M urea, 0.375 M Tris pH 8.8, 2% SDS, 20% glycerol, 2.5% (w/v) iodacetamide) for 10 minutes. The polymerizing stacking gel was added to the top of an 8% Tricine SDS-PAGE gel and the strips were slide into the polymerizing gel to ensure complete contact between the IEF strips and the SDS-PAGE gel. A small well was created in order to run the non-IEF sample for comparison to the IEF sample. After separation by SDS-PAGE, the gels were transferred to PVDF membrane and probed with streptavidin-HRP, as described above, and the different region-specific SecA antibodies as described in the previous chapter.

3.5.11. 2-D gel data analysis

To compare the cross-linking pattern between c-SecA and u-SecA, the Biotin detection of both proteins was analyzed using Flicker program (http://open2dprot.sourceforge.net/Flicker/) (Lemkin et al., 2005). Due to various distortions and other variations between the two samples, the Flicker program morphs and aligns two different images. In brief, three SecA fragments that are similar in both images were selected as landmarks. These landmarks were used to affine warp the
images resulting in the alignment of these landmarks. Flickering, which allows for rapid flickering between the two overlaid images was enabled and similar and different cross-linked fragments in c-SecA and u-SecA were visualized.

The SecA proteolytic fragments from the biotin detection were manually picked and analyzed using the ImageJ software (Abramoff et al., 2004). Fragments that have a molecular weight lower than 50 kD were analyzed as described earlier except the circle tool was used to define the fragment. Two criteria were used to define a fragment. The biotin signal intensity must have fit completely inside the area of a circle with a minimum radius of 5.0 pixels and a maximum radius of 10.5 pixels. This definition was used to eliminate small background speckles from the chemiluminescence as well as streaks of unfocused protein. The second criteria was that the mean-gray value of the fragment must be two times greater than the mean-gray value of the local background. Once a fragment has been defined, the AI of the local background was subtracted from the AI of the fragment. Finally, the relative intensity of each fragment was calculated by normalizing the data to the free signal peptide signal at the bottom of the gel. The average relative intensity of each fragment was determined by take three different measurements of the fragment using circles of different radii. If any of the mean-gray values of the fragments were not two times higher than the local background signal the fragment was eliminated from further analysis.

The different region-specific antibodies were used to identify the cross-linked fragments. The Flicker program was used to align and identify the cross-linked fragments (biotin detection) with the corresponding antibody detection. Since the same membrane was used for all of the detections, the full-length SecA protein from the non-IEF sample
(left side of the gel) was used to align the different images. Once the images were aligned, flickering between the biotin image and antibody image was enabled which allowed for the identification of the cross-linked fragments. Any fragment that was not identified by the antibodies was eliminated.
CHAPTER 4

SIGNIFICANT FINDINGS AND FUTURE DIRECTIONS

4.1. Significant findings

The translocation of preproteins across membranes is an essential step in the secretory pathway in all organisms. The bacterial secretory system requires three components to accomplish the translocation process: the signal sequence, SecA, and the membrane-embedded translocon. Although a significant amount of work has been performed on each of these components, the molecular details of the interaction between the signal sequence, SecA, the SecYEG translocon is still poorly defined. In this study, we were able to define a signal sequence-binding site on the cytosolic and the translocation-active conformations of SecA. Our work showed that both conformations share the same binding site but this site is likely expanded in the activated form.

Our work demonstrated that the translocation-active form of SecA undergoes dramatic domain dissociation. Dissociation of the C-terminal domains HWD, HSD, and CLT creating the activated form of SecA is consistent with the highly active truncated SecA64, which is missing the C-terminal one-third of the protein. Previous studies have suggested that this C-terminal region of SecA is protease protected in translocation assays. The dissociation of the PPXD and HWD is needed for the binding of SecA to the SecYEG translocon as demonstrated by the recent crystal structure of truncated SecA bound to the SecYEG translocon (Zimmer et al., 2008). The C-terminal domains, HWD, C-terminal HSD, and CTL can form a 30 kD stable fragment upon protease digestion. This fragment is similar to the membrane inserted portion of SecA previously described.
by Price et al. (Price et al., 1996). This observation suggests that the C-terminal domain association with the membrane may help stimulate SecA’s translocation activity. When SecA is pushed into the activated conformation, NBF I is also in a more proteolytic labile conformation. This observation is consistent with the binding of NBF I to SecYEG (Osborne and Rapoport, 2007) and that in our activated SecA NBF I is missing its interaction partner and therefore, exists in a more molten globular state.

Although the exact interactions of full-length SecA with the translocon are still poorly understood, we propose a model where NBF I, NBF II, and PPXD interact directly with the SecY while the C-terminal one-third of SecA disengages from the rest of the molecule and either interact with other components of the translocation system or the phospholipid membrane. These interactions of the C-terminus cause it to adopt a protease protected state while also releasing the suppression of the ATPase activity of SecA.

We have identified a signal-sequence binding site on SecA that is similar to the one proposed by Zimmer et al. (Zimmer et al., 2008) which is located at the interface between NBF II, PPXD, and N-terminal HSD. This result is in conflict with previously reported signal-sequence binding sites (see Chapter 3 Introduction for a full description of the other binding sites). All of these results suggest three different possibilities: (1) there are multiple signal-sequence binding sites on the cytosolic and activated forms of SecA, (2) one site is the actually signal-sequence binding site and the other sites are mature preprotein binding sites, and (3) all of the sites are mature preprotein binding sites. Although there is evidence that SecA interacts directly with the signal sequence, it
is also possible that the ‘reading’ of the signal sequence is performed by the translocon and not SecA.

4.2. Future directions

The development of a soluble translocation-active form of SecA has been useful in determining the conformational changes in the activated form of SecA as well as in mapping the signal sequence-binding site. Several questions regarding the mechanistic details of SecA translocation remain unanswered. How do the preprotein, SecA, and translocon work together? How does the preprotein remain in an unfolded conformation after release from SecB? Why does SecA interact with preproteins in the cytosolic and translocon-bound forms? These issue are explored and discussed in the sections below.

4.2.1. Structure of translocation-active SecA bound to preprotein

The crystal structure of truncated SecA/SecYEG (Zimmer et al., 2008) and the work presented in this study have provided details about the conformation of translocation-active SecA and interactions with signal sequences but this does not provide the entire picture for how these three molecules interact in-vivo. To further understand the nature of the interactions between preprotein, SecA, and SecYEG, more structural details are needed. Cryo-electron microscopy would be an excellent option to capture the SecA machine working to translocate the preprotein across the membrane. This kind of experiment would provide insight into the SecA mechanism and may also be applicable to other translocation machines.
4.2.2. **Preprotein interactions after release from SecB**

At the onset of the SecA translocation cycle, the molecular chaperone SecB is released from the preprotein/SecA/SecYEG complex. An intriguing question is how is the non-translocated portion of the preprotein maintained in an unfolded state. There must exist a sort of competition between the folding and translocation of the preprotein. Single-molecule studies using fluorescence could monitor the folding of the non-translocated portion of preprotein during the translocation process. One possibility is that the pushing of the preprotein through the translocon provides enough force to keep the preprotein in the unfolded state. This question could also be addressed using by using single molecule techniques.

4.2.3. **SecA /protein interactions in the cytosol**

One rather puzzling question is why do preproteins interact with SecA in the cytosol since the productive nature of the interactions is for translocation. A recent study by Eser and Ehrmann (Eser and Ehrmann, 2003) suggests that SecA might actually function in a quality-control manner to assist cytosolic proteins to fold by excluding them from the secretory pathway. The partitioning of SecA between the membrane and the cytosol indicates SecA could indeed have dual roles yet SecA can also recognize preproteins in the cytosol. Further studies using single-molecule fluorescence may be able to aid in the understanding of how SecA is able to differentiate between a secretory and a cytosolic protein. These kinds of studies could also provide insight into how other chaperones discriminate between substrates.
BIBLIOGRAPHY


Chou, Y. T. (2003) The characterization of SecA and its interactions with signal sequences in the bacterial secretory pathway, PhD, University of Massachusetts Amherst, Amherst, MA.


Duong, F. (2003). Binding, activation and dissociation of the dimeric SecA ATPase at the dimeric SecYEG translocase. Embo J 22, 4375-4384.


Saier, M. H., Jr. (2006). Protein secretion and membrane insertion systems in gram-


