Alignment of the protein substrate hairpin along the SecA two-helix finger primes protein transport in *Escherichia coli*

Qi Zhang[a], Sudipta Lahiri[a], Tithi Banerjee[a], Zhongmou Sun[a], Donald Oliver[a], and Ishita Mukerji[a,1]

[a]Molecular Biophysics Program, Department of Molecular Biology and Biochemistry, Wesleyan University, Middletown, CT 06459

Edited by Linda L. Randall, University of Missouri, Columbia, MO, and approved July 11, 2017 (received for review February 8, 2017)

A conserved hairpin-like structure comprised of a signal peptide and early mature region initiates protein transport across the SecY or Sec61α channel in Bacteria or Archaea and Eukarya, respectively. When and how this initiator substrate hairpin forms remains a mystery. Here, we have used the bacterial SecA ATPase motor protein and SecYEG channel complex to address this question. Engineering of a functional miniprotein substrate onto the end of SecA allowed us to efficiently form ternary complexes with SecYEG for spectroscopic studies. Förster resonance energy transfer mapping of key residues within this ternary complex demonstrates that the protein substrate adopts a hairpin-like structure immediately adjacent to the SecA two-helix finger subdomain before channel entry. Comparison of ADP and ATP-γS–bound states shows that the signal peptide partially inserts into the SecY channel in the latter state. Our study defines a unique preinsertion intermediate state where the SecA two-helix finger appears to play a role in both templating the substrate hairpin at the channel entrance and promoting its subsequent ATP-dependent insertion.

protein transport | FRET mapping | Sec system

Protein transport into and across the plasma membrane of Archaea and Bacteria or the endoplasmic reticular membrane of Eukarya occurs through a universally conserved protein-conducting channel termed the SecY or Sec61 complex, respectively [reviewed in Park and Rapoport (1)]. These channels display remarkably similar hourglass-shaped structures that are doubly gated: they open vertically to allow protein transport across the membrane or open laterally to allow insertion of integral membrane proteins into the lipid bilayer (2–10). The SecYEG channel consists of 15 transmembrane helices with a short helical region that blocks the channel, referred to as the plug domain. Vertical opening is accomplished at least in part by rearrangement of the plug domain of SecY/Sec61α to open the channel interior, while lateral opening requires transverse movement of SecY/Sec61α helices that reside at the lateral gate to create a path for exit from the channel interior into the lipid bilayer.

Substrate proteins can initiate their transport either cotranslationally or posttranslationally depending on a given substrate and organism. For cotranslational transport, the signal recognition particle and its receptor promote the targeting of appropriate nascent chain-bearing polysomes to the channel complex, while for posttranslational transport species-specific proteins are responsible for such targeting. In many cases, a signal sequence located at the N terminus of the substrate protein contains information regarding the ultimate location of the protein and is needed for translocation. The degree of hydrophobicity of this signal sequence appears to control which of these two pathways is used (11, 12).

In Bacteria, SecA protein recognizes both signal peptides and SecY protein, and serves to target secretory preproteins to the SecY complex (13). SecA is a multidomain protein with two ATP binding domains (NBD-1, NBD-2), a preprotein cross-linking domain (PPXD) named for its ability to cross-link to preprotein substrates (14), central helix (CH) and two-helix finger (THF) subdomains, as well as helical wing domain (HWD) and carboxyl-terminal linker (CTL) domains (Fig. 1). As an ATPase motor protein, SecA uses ATP-driven hydrolytic cycles to promote domain movements required for substrate protein insertion into the SecY channel as well as subsequent processive protein transport through the channel. Several models of SecA action have been proposed [summarized in Kusters and Driessen (15)]; however, the precise mechanism of SecA-driven protein transport has yet to be fully elucidated. Recent structural and biochemical analysis have suggested that the THF subdomain located in the central region of SecA potentially acts as a molecular ratchet driving substrate proteins into the SecY channel (16, 17) (Fig. 1).

Protein transport occurs via a loop model in which the signal peptide and early mature region of the preprotein form a hairpin-like structure within the membrane to initiate the transport process (18, 19). In this topology, the signal peptide remains relatively fixed, while the polypeptide region that follows is processively threaded across the membrane to accommodate increasing larger translocation intermediates. A recent X-ray structure has visualized this substrate protein hairpin within the SecY complex and found that the signal peptide resides in a groove immediately outside of the lateral gate, while the early mature region resides within the channel proper (10), augmenting earlier electron cryomicroscopy studies where only the signal sequence was visible (5–7).

Previously, we used Förster resonance energy transfer (FRET)-based methods to map the location of the SecA signal peptide-binding site using chimeras containing SecA and the alkaline phosphatase (PhoA) or lambda receptor variant (KRRLamB) signal peptides (20, 21). We found that the signal peptide interacts

Significance

Specialized protein complexes transport and integrate peptides into membranes in all living cells. Protein transport within the universally conserved Sec system requires the formation of an initiator protein substrate hairpin comprised of the signal peptide and adjacent region. Fundamental questions remain regarding when and how this hairpin structure forms. Here, we show that the SecA two-helix finger templates the hairpin within the preformed SecYEG-bound SecA complex prior to its insertion into the SecY channel. In addition to capturing a novel preinsertion intermediate state, our study expands the role of the SecA two-helix finger, which has previously been suggested to be an ATP-powered ratchet that drives cycles of protein transport.

Author contributions: Q.Z., D.O., and I.M. designed research; Q.Z., S.L., T.B., and Z.S. performed research; Q.Z. contributed new reagents/analytic tools; Q.Z., S.L., D.O., and I.M. analyzed data; and Q.Z., S.L., D.O., and I.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: imukerji@wesleyan.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1702201114/-/DCSupplemental.

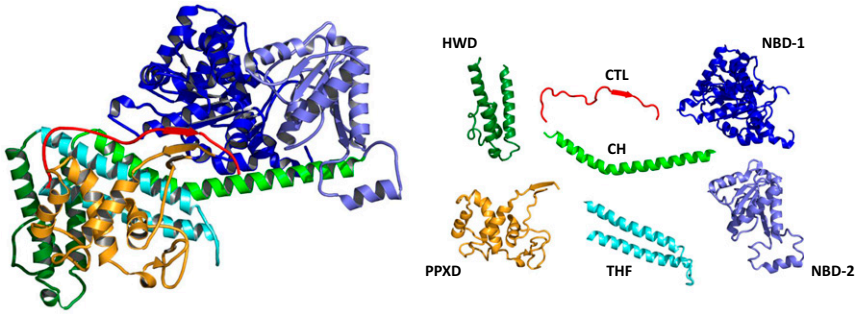

Fig. 1. Ribbon representation of the *B. subtilis* SecA protein colored by domain (*Left*) with the individual domains shown on the *Right* (PDB ID code 1M6N). They include (N-terminal to C-terminal on SecA): the nucleotide-binding domain-1 (NBD-1) (blue), the preprotein cross-linking domain (PPXD) (gold), the nucleotide-binding domain-2 (NBD-2) (light blue), the central helix subdomain (CH) (green), the helical wing domain (HWD) (dark green), the two helix-finger subdomain (THF) (cyan), and the carboxyl-terminal linker (CTL). The CTL is depicted in red and serves as a model of PhoA signal peptide bound to *B. subtilis* SecA based on the FRET mapping study of Zhang et al. (21).

with both the PPXD and THF subdomain and adopts a parallel binding orientation along the long axis of THF mimicking the location of the carboxyl-terminal tail of SecA in the *Bacillus subtilis* X-ray crystal structure (shown in red, Fig. 1). We also demonstrated that the lambda receptor signal peptide, KRRLamB, bound to SecA in the same location and orientation as its PhoA counterpart, indicative of a common binding site and orientation (21).

In this report, we have used a similar approach to address a critical and unresolved question concerning the initiation of protein transport in the SecA–SecYEG complex: Does the protein substrate hairpin-like structure that initiates protein transport form before or after insertion of the substrate into the membrane? We now report that the hairpin-like structure forms and binds to the SecA THF subdomain before entry into the SecY channel. Given the importance of the substrate hairpin loop in channel activation, templating of the hairpin before channel insertion may be a conserved mechanism for the cotranslational protein transport pathway as well.

Results

To determine the location of the signal peptide and early mature region of the protein substrate within the purified SecYEG-bound SecA complex, we used our genetically engineered SecA–PhoA chimeras and FRET mapping approach as reported previously (21). We have demonstrated that the SecA and PhoA signal peptide portions of the chimera were functional in vivo and in vitro, and that the attached PhoA signal peptide bound specifically to SecA at the previously mapped binding site (20). In the present study, a longer chimera was created in a similar manner by genetically fusing the PhoA substrate to SecA in which we removed the dispensable carboxyl-terminal 67-residue linker domain of SecA and replaced it with a short glycine–serine linker followed by the first 68 residues of PhoA. The PhoA portion contained the 21-residue signal peptide and an additional 47 residues of the early mature region terminated by a hexahistidine tag (Table S1). To achieve differential labeling between SecA and PhoA, dye labels were introduced into the PhoA portion of the chimera by in vivo incorporation of azidophenylalanine at engineered amber codons followed by protein purification and dye incorporation at these sites using click chemistry (22). In contrast, labeling of selected residues within SecA or SecY was achieved by genetically engineering unique cysteine residues and labeling them using maleimide chemistry. Efficient assembly of the relevant PhoA substrate-bound SecA–SecYEG complex was improved by this approach, as it ensures a one-to-one stoichiometry of bound substrate to SecA–SecYEG complex, and avoids solubility issues associated with high concentrations of free substrate peptides or their premature folding or aggregation in the case of larger protein substrates (20).

To capture an early intermediate in the transport process, we chose to use *n*-dodecyl-β-D-maltopyranoside (DDM)-solubilized SecYEG protein for our study, since SecA binds this form of SecYEG with high affinity and the binding stimulates SecA ATPase activity (23, 24). In addition, this system is comparable to that used in the *Thermotoga maritima* SecA–SecYEG X-ray cocrystal structure (16), which we use for modeling our results. Robust site-specific in vivo photo–cross-linking has demonstrated that the *Escherichia coli* SecA–SecY complex is structurally similar to its *T. maritima* counterpart (25). By avoiding the use of a SecYEG-containing proteoliposome system, we further simplified our approach and mitigated concerns over a mixed topology of SecYEG protein within the bilayer, its potential for dimerization, or that a fraction of SecA might bind solely to phospholipids using its known lipid-binding activity.

To track the location of the signal peptide and early mature region of the protein substrate in the SecA–SecYEG complex, we introduced dye labels periodically along its length (Fig. 2*A*). We selected residues at the beginning and end of the signal peptide (residues 2 and 22, respectively) as well as at two positions within the early mature region (residues 37 and 45). We avoided placing a dye within the hydrophobic core of the signal peptide, since we found previously that it perturbed SecA signal peptide binding (20). To accurately map the location of the protein substrate, we also introduced dye labels at three distinct locations in the SecA–SecYEG complex: two within SecA, located within the NBD-1 and the PPXD, and one within SecYEG (shown as spheres in Fig. 2 *D* and *E*). The residues selected for labeling the SecA–SecYEG complex needed to satisfy several criteria: (*i*) be structurally well distributed throughout the complex but within accurate range for FRET measurements, (*ii*) be surface accessible for good labeling efficiency, and (*iii*) reside near the ends of well-structured regions for greater accuracy in distance determinations with minimal structural and functional perturbations. As the three residues chosen (SecA37, SecA321, and SecY292) are positioned in distinct regions of the complex, the location of the signal peptide and early mature region on the SecA–SecYEG complex in the presence of either ADP or ATP-γS could be determined from the measured distances (Fig. 2).

To ensure that both SecA and SecYEG were in their monomeric forms and avoid interprotomer FRET, a buffer system containing 300 mM KCl and 0.1% DDM was used. At this salt concentration, less than 3% of SecA dimer was detected (26), while SecYEG dimer was undetectable at this detergent concentration (27). SecA affinity for exogenous PhoA signal peptide or an extended signal peptide that also contained the early mature region was ∼2 μM at this salt concentration similar to previous reports (20) (Fig. S1*A*). SecA also bound SecYEG with

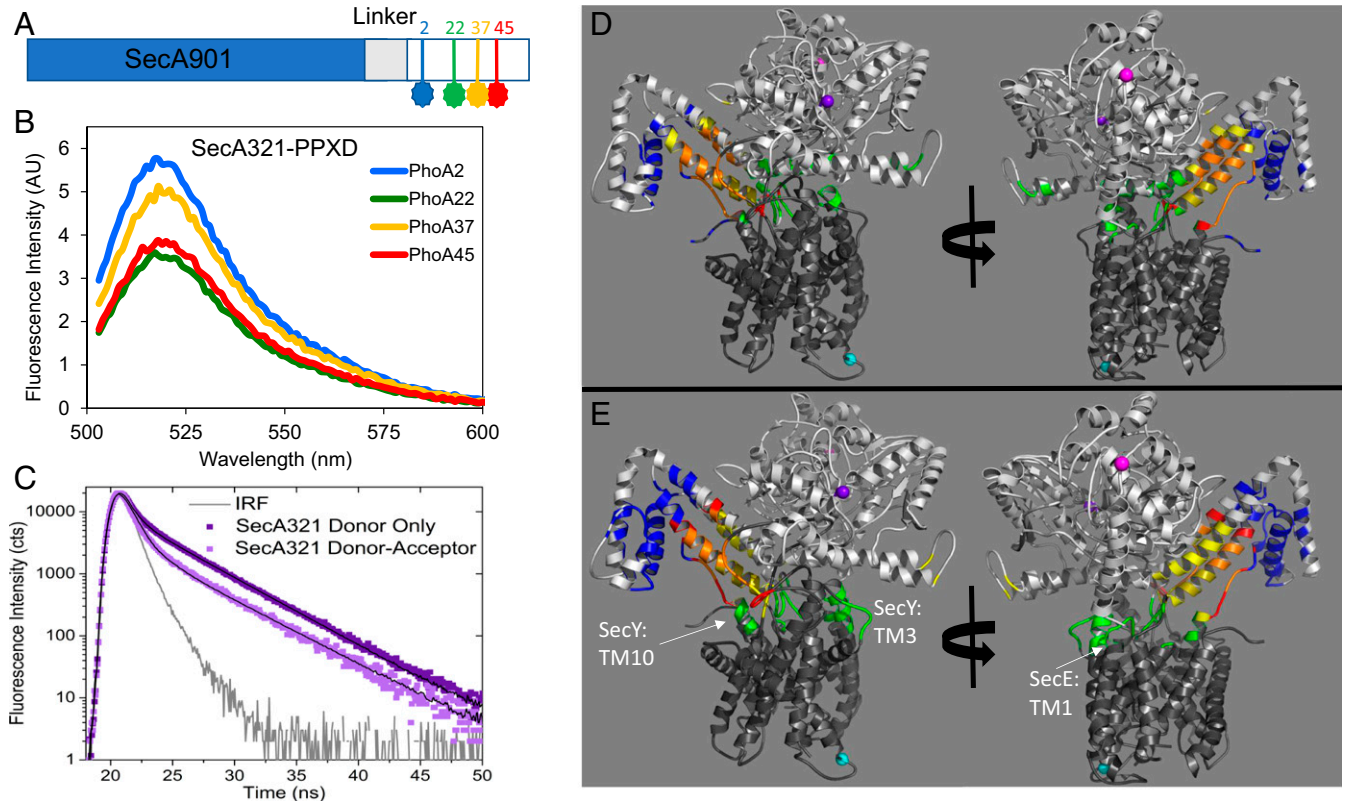

Fig. 2. Mapping of PhoA signal peptide and early mature region on to the SecA–SecYEG complex. (*A*) Schematic of the SecA–PhoA chimera construct in which the PhoA substrate peptide is genetically fused to SecA after a Gly–Ser linker (not drawn to scale). Cys residues were introduced for dye labeling at the indicated positions and are depicted in blue, green, yellow, and red. (*B*) Representative fluorescence spectra of the doubly labeled SecA–PhoA chimeras in the presence of SecYEG and ADP, with the donor dye positioned at different points within PhoA and the acceptor dye positioned at SecA residue 321. The spectrum generated with the donor dye at PhoA position 22 has the lowest intensity and highest transfer efficiency. Data were acquired and analyzed as described in *SI Materials and Methods*. (*C*) Time-resolved fluorescence decay spectra of the SecA–PhoA chimera labeled with the donor dye at PhoA position 22 and with the acceptor dye at SecA residue 321. A donor-only decay is shown in dark violet and a donor–acceptor decay is shown in light violet. The instrument response function (IRF) is given in gray. The donor–acceptor decay yields a shorter lifetime indicative of energy transfer. Data were acquired and analyzed as described in *SI Materials and Methods*. (*D* and *E*) The *T. maritima* SecA–SecYEG complex (PDB ID code 3DIN) is shown as a ribbon diagram with SecA and SecYEG in light and dark gray, respectively. The locations of *T. maritima* residues homologous to *E. coli* SecA37, SecA321, and SecY292 are shown by magenta-, violet-, or cyan-colored spheres, respectively. Mapped location of the PhoA substrate within the SecA–SecYEG complex in the presence of ADP (*D*) or ATP-γS (*E*). The region of overlap of the structure with the FRET data for PhoA residue 2, 22, 37, or 45 is shown in blue, green, yellow, or red, respectively. Overlap regions of PhoA residues 22, 37, and 45 are shown in olive, and overlap regions of 37 and 45 are shown in orange. Figures on the *Right* are rotated by ≈180°.

high affinity in this buffer system with a K_d of 6 nM (Fig. S1*B*) similar to previous measurements where the complex was shown to consist almost entirely of SecA monomers bound to SecYEG at this salt concentration (28). As expected, the ATPase activity of our various dye-labeled SecA–PhoA chimeras increased in the presence of SecYEG, exhibiting a range of activity from 30 to 160%. We have shown previously that the presence and position of dyes within SecA can affect ATPase activities without necessarily compromising function (20). In addition, we found that the various SecA–PhoA chimeras bound SecYEG with roughly similar affinities based on the similar strength of the FRET signals (see below as well as in *SI Materials and Methods*).

By using four positions on PhoA (Fig. 2*A*) and three positions on the SecA–SecYEG complex, we measured 12 different distances to position the PhoA substrate within the complex. To observe membrane-retracted or membrane-inserted states of SecA, spectra were collected in the presence of ADP or ATP-γS, respectively (Fig. S2) (29). For two sets of these measurements, the acceptor dye was located on the SecA portion of the chimera (residue 37 or 321 shown as violet or magenta spheres, respectively; Fig. 2 *D* and *E*), and the donor dye was located on the PhoA portion of the chimera. In the third set of measurements, the donor dye was on SecY adjacent to its plug domain (residue 292 shown as a cyan sphere, Fig. 2 *D* and *E*), and the acceptor dye was located on the PhoA portion of the chimera. Representative steady-state fluorescence spectra from a doubly labeled system are shown in Fig. 2*B* and demonstrate visually that the highest transfer efficiencies (lowest donor intensities) were observed for the dye located at PhoA residue 22 of the chimera, while the lowest transfer efficiencies (highest donor intensities) were observed for the dye located at PhoA residue 2. Fluorescence spectra of doubly labeled species from all three locations exhibited similar trends where the measured efficiencies were not linearly proportional to the distance of the label from the start of PhoA in the chimera, consistent with the formation of a hairpin loop (Fig. S2).

Energy transfer efficiencies were also determined using time-resolved fluorescence spectroscopy to verify the homogeneity of the transfer (Fig. 2*C*). Donor-only spectra were obtained in the presence of unlabeled proteins to control for any quenching associated with SecA–SecYEG complex formation. Analysis of these decays yielded one dominant lifetime, consistent with only one species participating in energy transfer for a given donor–acceptor pair. The presence of the acceptor reduced the donor lifetime, indicative of

energy transfer (Fig. 2*C* and Fig. S3), and the calculated efficiencies were in excellent agreement with steady-state measurements (Tables S2–S5). The steady-state FRET efficiencies (Tables S2–S4) are all within the linear efficiency range of the dye pairs used (0.2–0.8). The determined distances have a larger error than that obtained experimentally for the efficiencies alone to account for the error introduced by the uncertainty in dye position as determined from steady-state anisotropy values (21, 30, 31).

Using this information, we could position the signal peptide and early mature region of PhoA within the SecA–SecYEG complex. To do this, we considered each of the three distances measured between a given PhoA residue and their corresponding FRET partner on either SecA or SecY as a radius for a spherical shell (Fig. S4). The intersecting region of the three shells defined the position of a given PhoA residue within the complex, where the width of each shell corresponded to our uncertainty in the distance (Tables S2–S4). Thus, in Fig. 2 *D* and *E*, the mapped location of the second residue of the PhoA signal peptide (PhoA2) is shown in blue. Similarly, the mapped regions for PhoA residues 22 (PhoA22), 37 (PhoA37), and 45 (PhoA45) are shown in green, yellow, and red, respectively. There is considerable overlap of the PhoA22, PhoA37, and PhoA45 regions, which is shown in olive, while overlapping regions of PhoA37 and PhoA45 are shown in orange. In the ADP-bound form (Fig. 2*D*), the region mapped to PhoA2 is located mainly on the HWD and lies at the end of the THF farthest from the mouth of the channel. In contrast, the regions mapped solely by PhoA22, shown in green, are primarily found positioned on the loop of THF and lie directly over the mouth of the channel, poised for translocation. The PhoA22 region also maps to one of the helices of the THF along with the regions defined by PhoA37 and PhoA45. This finding supports a model in which the PhoA37 and PhoA45 residues are close to each other in space and are sandwiched in between PhoA22 and PhoA2. This spatial distribution strongly points to the PhoA signal peptide and early mature region forming a hairpin-like structure immediately adjacent to and paralleling the SecA THF subdomain (Fig. 2 *D* and *E*). The hairpin loop corresponding to the junction between the end of the signal peptide (PhoA22) and the beginning of the early mature region is formed at the mouth of the channel while the ends of the hairpin (PhoA2 and PhoA45) lie at the end of the THF near to the wing domain, far from the channel mouth. The relatively large area defined by the PhoA2 FRET measurements suggests the N-terminal end of the signal peptide is fairly flexible.

In the ATP-γS–bound state (Fig. 2*E*), the regions defined solely by PhoA22 (shown in green) primarily map to SecYEG and the top of channel. Exclusive mapping is observed at the top (cytosolic side) of SecY on TM3 and TM10 as well as TM1 of SecE. The loop or tip of the finger is also mapped to PhoA22 only. These results are in excellent agreement with increased insertion of substrates into the channel in the ATP-bound state (32, 33) and are consistent with the highest transfer efficiencies observed between the PhoA22 and SecY292 FRET pair in the presence of ATP-γS (Fig. 2*C* and Fig. S2*B*). The regions mapped by PhoA37 and PhoA45 remain along the THF and the CH subdomains as well as the C terminus of SecY (often termed the C6 cytosolic domain of SecY). The relative progression and orientation of the signal peptide and early mature region within the SecA–SecYEG complex can be visualized along the THF, in which the color change from green to red (green, olive, yellow, orange, red) signifies mapping of the regions from PhoA22 to PhoA45. The more orderly structure of the substrate in its ATP-γS–bound versus ADP-bound state suggests that it becomes more templated for translocation upon ATP binding. Although there is still considerable flexibility at the substrate ends as shown by their relatively large mapped areas, we note that these two regions, PhoA2 and PhoA45 (blue and red, respectively), are proximal to one another and far from the channel entry. Taken together, these findings strongly support a model where formation of the initial substrate hairpin is templated by the THF subdomain before its insertion into the channel.

Comparison of our result with the recent X-ray structure of the SecA–SecYEG complex with substrate inserted into the channel (10) (Fig. S5) clearly suggests that our complex represents a preinsertion intermediate state, while the latter structure represents a more mature, postinsertion, translocation-intermediate state. Remarkably, the signal peptide and early mature regions of the substrate adopt similar hairpin-like structures in both cases despite the fact that they are different substrates and were fused to SecA very differently: the OmpA signal peptide and early mature region were inserted at the end of the THF finger in the Li et al. (10) X-ray structure, while in our study the PhoA peptide was fused onto the C-terminal end of SecA. The similar nature of the results strongly implies that the initial substrate hairpin-like structure preexists outside of the SecYEG channel, is nucleated and/or stabilized by the assembled SecA–SecYEG complex, and is a conserved and fundamental unit for initiating transported protein substrates.

As shown in Fig. 3, we have used the OmpA signal peptide and early mature region hairpin (depicted in pink) from the Li et al. structure to model the location of its PhoA counterpart (depicted in cyan) in the preinsertion state. To generate this structure, the structurally unaltered OmpA hairpin was modeled into our FRET-identified locations. The relative accuracy of the proposed structure can be assessed by comparing specific locations along the OmpA hairpin with their mapped PhoA counterparts. As shown in Fig. 3*B*, the residues at positions 2 (Lys), 22 (Tyr), and 37 (Gly) of the OmpA hairpin are shown in blue, green, and yellow, respectively, and they match up exactly with the regions identified by our FRET measurements. Although the OmpA peptide is truncated at position 41, if it were extended, the additional segment would completely agree with our predicted position for PhoA45. In fact, the unstructured C-terminal end of SecY (which lies parallel to the unstructured region of the OmpA hairpin) also provides a model demonstrating how the early mature region of the peptide substrate could be binding within this region. The modeled-in hairpin structure further illustrates how the regions defined by PhoA2 and PhoA45 lie proximal to one another and distal from the channel opening. The excellent agreement observed between the structure of the OmpA signal peptide and early mature region and our FRET-based predictions further supports a model in which the initial substrate hairpin is templated along the THF before insertion into the channel.

Discussion

The major goal of our study was to determine the location of the amino-terminal portion of the substrate protein within the SecA–SecYEG complex in a preinsertion state. Previously, it was unclear (*i*) whether SecA contains two distinct signal peptide-binding sites: one for initial recognition by SecA in solution and another for later use within the SecA–SecYEG complex, (*ii*) whether SecA also contains a region for binding the early mature region of the substrate that is distinct from its signal peptide-binding site(s), and (*iii*) whether the initial substrate hairpin-like structure that is requisite for the initiation of protein transport forms before or concurrent with substrate entry into the channel. Our study satisfactorily addresses all three questions. In particular, we have visualized a unique preinsertion intermediate state before the deeper penetration of substrate into the channel proper. In this unique form, the signal peptide and early mature region of the substrate form a hairpin-like conformation lying along the SecA THF subdomain with their two structures in approximate register. The agreement between the PhoA and OmpA (shown in Fig. 3) substrate topologies is striking and attests to the presumably universal nature of the hairpin conformation for substrate entry into the Sec-dependent protein transport pathway. Our results expand the proposed function of

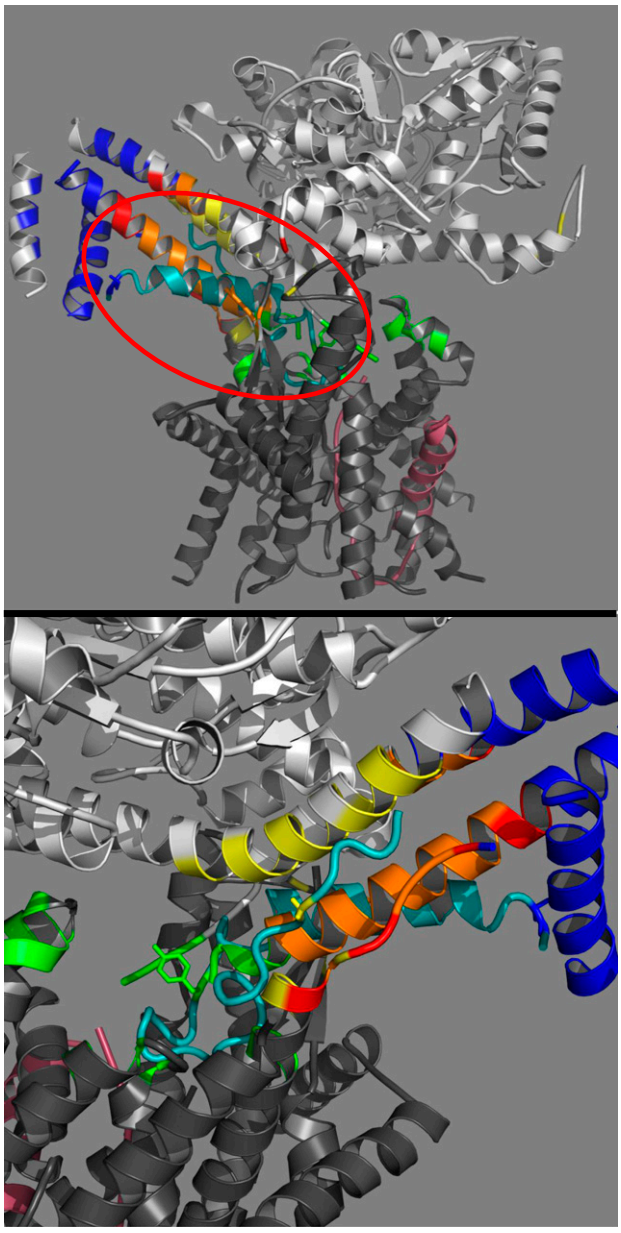

Fig. 3. (*A*) FRET-mapped regions projected on the *B. subtilis* SecA–*Geobacillus thermodenitrificans* SecYE cocrystal structure (PDB ID code 5EUL). SecA is shown in light gray, SecYE is in dark gray, and the OmpA peptide substrate inserted at the end of the THF is shown in pink. For clarity, the nanobody crystallized with the complex has been omitted (10). Generation of FRET-mapped regions and their associated colors in the presence of ATP-γS was done as described in Fig. 2. Circled in red is the peptide substrate (residues 749–791) (shown in cyan) excised from the original 5EUL PDB structure and modeled into the mapped regions without any alteration of the original structure. (*B*) Enlarged view of the modeled peptide (cyan) and mapped locations. Residues 2 (Lys), 22 (Tyr), and 37 (Gly) of the OmpA peptide are shown in a stick representation in blue, green, and yellow, respectively, and exhibit excellent agreement with the PhoA-mapped locations. Note the adjacent C-terminal portion of SecY discussed in the text.

the THF subdomain as a molecular ratchet (16, 17), suggesting that it first serves to template the substrate into a hairpin for subsequent entry into the channel (Fig. S6).

Our results suggest an obvious model where the THF and adjacent regions form two conjoined substrate-binding sites: one for the signal peptide and another for the early mature region that lie on opposite sides of the finger (shown in Fig. 3 *A* and *B*, respectively). We note that the first site coincides with our previously mapped solution state SecA-signal peptide-binding site (20) (Fig. 1), indicating this portion of the substrate does not reposition itself after SecYEG association. We also note that both sites have potential alternative binding partners in the absence of substrate, namely the C-terminal end of SecA (shown in red in Fig. 1) for the signal peptide-binding site and the C-terminal end of SecY (shown in FRET-mapped colors in Fig. 3*B*) for the early mature region-binding site. It appears likely that these alternative partners occupy these sites in the absence of substrate and potentially play regulatory roles in controlling substrate binding or channel insertion based on the existing literature (21, 34, 35).

The identification of an early mature region-binding site within SecA provides a structural handle to potentially understand how substrates with defective or missing signal peptides are accommodated for transport in certain Sec (Prl) mutants or why substitutions of multiple positively charged amino acid residues within the early mature region of substrates (so-called Sec-avoidance sequences) strongly inhibit their transport (36, 37). Regarding the proofreading activity of the translocon, we note that the existence of a preinsertion intermediate state that is linked to the SecA ATP binding and hydrolysis cycle allows the SecA–SecYEG complex to "scan" the substrate during the templating process and potentially "reject" it in a more readily reversible fashion than later on when the substrate has inserted into the channel proper. Our results imply that both formation of the proper SecA–SecYEG complex and correct placement of the substrate loop within it would be required for activation of SecA ATPase and SecA–SecYEG proofreading activities. The ADP and ATP-γS–bound preinsertion states depicted here could represent good working models of such proofreading steps, which have remained elusive.

Our data also point to the importance of the CH of SecA in binding and positioning the tip of the hairpin (PhoA22, green, Fig. 2*D*) adjacent to the channel opening, particularly in the ADP-bound state. Previous work has demonstrated the importance of this region of SecA for coupling its ATPase activity to substrate translocation (38). In this context, substrate contact with one end of the remarkably long CH subdomain, which is also in contact with NBD-1, NBD-2, and SecY protein, could provide the appropriate signaling for activation of SecA translocation ATPase activity and its coupling to substrate translocation. The role of the CH subdomain in potentially integrating these different events now deserves further study.

The preinsertion complex visualized in our study extends the SecA power stroke or ratchet model that relies on the THF subdomain for positioning and pushing substrate into and across the SecY channel in a processive manner powered by the SecA DEAD ATPase motor (discussed with references in ref. 15). The ADP and ATP-γS–bound states characterized here indeed depict modest movement of the substrate into the channel, but are clearly insufficient to explain the observed ∼20-aa step size for substrate transport during a single ATP turnover cycle. However, step size may differ significantly for the SecA–SecYEG complex in its preinsertion versus postinsertion states. In addition, protein translocation in this system appears to use both SecA-dependent pushing and Brownian motion-dependent sliding to achieve efficient transport (33). Since the mobility and role of the THF subdomain in protein transport remain controversial (39, 40), additional studies will be required to resolve this matter. Further studies can now address the biochemical and structural requirements needed to transition from the preinitiation to postinitiation states. In that regard, our FRET-based mapping methodology that makes use of functional protein chimeras coupled with site specific dye labeling provides a compelling approach to address this problem.

Materials and Methods

Construction, Expression, and Purification of SecA, SecA–PhoA Chimeras, and SecYEG Proteins. A series of *E. coli* SecA mutants or SecA–PhoA chimeras were constructed as described in Table S1. DH5α [F^- ϕ80*lacZΔ M15Δ (lacZYA-argF)U169 recA1 endA1 hsdR17*(r_K- m_K+) *phoA supE44 thi-1 gyrA96 relA1*] was used for all plasmid construction and purification, and DNA sequence analysis was performed at the University of Pennsylvania DNA-Sequencing Facility. *E. coli* BL21.19 [*secA13(Am) supF(Ts) trp(Am) zch*::*Tn10 recA*::*CAT clpA*::*KAN*] is derived from BL21(λDE3) (41) and was used as the host for all SecA- or SecA–PhoA-containing plasmids. Expression of SecA or SecA–PhoA mutants requiring incorporation of H-4-azido-Phe-OH at amber codons used the pEVOL-pAzF plasmid along with the appropriate pT7SecA or pT7SecA–PhoA plasmid (22) in *E. coli* BLR(λDE3). C43(DE3) [*ompT hsdSB* (r_B- m_B-) *gal dsm* (λDE3)] was used as the host for all SecYEG-containing plasmids. Expression and purification of SecA, SecA–PhoA chimeras, and SecYEG proteins are described in the *SI Materials and Methods* and are shown in Fig. S7.

Fluorescence Measurements. Steady-state fluorescence spectra were collected on a Fluoromax 4 (Horiba) spectrofluorometer. Time-resolved fluorescence decays were collected by the time-correlated single-photon counting method (PTI Timemaster) using a 490-nm LED laser (rep rate, 180 kHz). Acquisition and analysis details are given in *SI Materials and Methods*.

SecA ATPase Activity. SecA ATPase activity was measured for all of the chimeras and was determined by the Malachite green method with the modifications described previously (20). In addition to using conventional inverted membrane vesicles, we used purified SecYEG protein and high-salt TKM buffer with 0.1% DDM in our assay. The PhoA portion of the SecA–PhoA chimera served as the substrate protein in lieu of additional proOmpA for measurements performed with the chimera.

ACKNOWLEDGMENTS. We thank T. Rapoport (Harvard Medical School) and Peter Schultz (Scripps) for provision of the SecYEG-producing and azidophenylalanine–incorporating plasmids, respectively. Q.Z. thanks his committee members Manju Hingorani and Richard Olson for their helpful suggestions. This work was supported by National Institutes of Health Grant GM110552 (to D.O.) and National Science Foundation Grant MCB-0843656 (to I.M.).

1. Park E, Rapoport TA (2012) Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annu Rev Biophys* 41:21–40.
2. Van den Berg B, et al. (2004) X-ray structure of a protein-conducting channel. *Nature* 427:36–44.
3. Tsukazaki T, et al. (2008) Conformational transition of Sec machinery inferred from bacterial SecYE structures. *Nature* 455:988–991.
4. Egea PF, Stroud RM (2010) Lateral opening of a translocon upon entry of protein suggests the mechanism of insertion into membranes. *Proc Natl Acad Sci USA* 107: 17182–17187.
5. Frauenfeld J, et al. (2011) Cryo-EM structure of the ribosome-SecYE complex in the membrane environment. *Nat Struct Mol Biol* 18:614–621.
6. Hizlan D, et al. (2012) Structure of the SecY complex unlocked by a preprotein mimic. *Cell Rep* 1:21–28.
7. Park E, et al. (2014) Structure of the SecY channel during initiation of protein translocation. *Nature* 506:102–106.
8. Tanaka Y, et al. (2015) Crystal structures of SecYEG in lipidic cubic phase elucidate a precise resting and a peptide-bound state. *Cell Rep* 13:1561–1568.
9. Voorhees RM, Hegde RS (2016) Structure of the Sec61 channel opened by a signal sequence. *Science* 351:88–91.
10. Li L, et al. (2016) Crystal structure of a substrate-engaged SecY protein-translocation channel. *Nature* 531:395–399.
11. Ng DT, Brown JD, Walter P (1996) Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J Cell Biol* 134:269–278.
12. Lee HC, Bernstein HD (2001) The targeting pathway of *Escherichia coli* presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. *Proc Natl Acad Sci USA* 98:3471–3476.
13. Lycklama A, Nijeholt J, Driessen A (2012) The bacterial Sec-translocase: Structure and mechanism. *Philos Trans R Soc Lond B Biol Sci* 367:1016–1028.
14. Kimura E, Akita M, Matsuyama S, Mizushima S (1991) Determination of a region in SecA that interacts with presecretory proteins in *Escherichia coli*. *J Biol Chem* 266: 6600–6606.
15. Kusters I, Driessen AJ (2011) SecA, a remarkable nanomachine. *Cell Mol Life Sci* 68: 2053–2066.
16. Zimmer J, Nam Y, Rapoport TA (2008) Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature* 455:936–943.
17. Erlandson KJ, et al. (2008) A role for the two-helix finger of the SecA ATPase in protein translocation. *Nature* 455:984–987.
18. Duffaud GD, Lehnhardt SK, March PE, Inouye M (1985) Structure and function of the signal peptide. *Current Topics in Membranes and Transport*, ed Cook J (Academic, New York), Vol 24, pp 65–104.
19. Plath K, Mothes W, Wilkinson BM, Stirling CJ, Rapoport TA (1998) Signal sequence recognition in posttranslational protein transport across the yeast ER membrane. *Cell* 94:795–807.
20. Auclair SM, et al. (2010) Mapping of the signal peptide-binding domain of *Escherichia coli* SecA using Förster resonance energy transfer. *Biochemistry* 49:782–792.
21. Zhang Q, Li Y, Olson R, Mukerji I, Oliver D (2016) Conserved SecA signal peptide-binding site revealed by engineered protein chimeras and Förster resonance energy transfer. *Biochemistry* 55:1291–1300.
22. Chin JW, et al. (2002) Addition of *p*-azido-L-phenylalanine to the genetic code of *Escherichia coli*. *J Am Chem Soc* 124:9026–9027.
23. Duong F (2003) Binding, activation and dissociation of the dimeric SecA ATPase at the dimeric SecYEG translocase. *EMBO J* 22:4375–4384.
24. Robson A, Booth AE, Gold VA, Clarke AR, Collinson I (2007) A large conformational change couples the ATP binding site of SecA to the SecY protein channel. *J Mol Biol* 374:965–976.
25. Das S, Oliver DB (2011) Mapping of the SecA·SecY and SecA·SecG interfaces by site-directed in vivo photocross-linking. *J Biol Chem* 286:12371–12380.
26. Das S, Stivison E, Folta-Stogniew E, Oliver D (2008) Reexamination of the role of the amino terminus of SecA in promoting its dimerization and functional state. *J Bacteriol* 190:7302–7307.
27. Bessonneau P, Besson V, Collinson I, Duong F (2002) The SecYEG preprotein translocation channel is a conformationally dynamic and dimeric structure. *EMBO J* 21: 995–1003.
28. Kusters I, et al. (2011) Quaternary structure of SecA in solution and bound to SecYEG probed at the single molecule level. *Structure* 19:430–439.
29. Economou A, Wickner W (1994) SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. *Cell* 78:835–843.
30. Ivanov V, Li M, Mizuuchi K (2009) Impact of emission anisotropy on fluorescence spectroscopy and FRET distance measurements. *Biophys J* 97:922–929.
31. Auclair SM, Oliver DB, Mukerji I (2013) Defining the solution state dimer structure of *Escherichia coli* SecA using Förster resonance energy transfer. *Biochemistry* 52: 2388–2401.
32. Schiebel E, Driessen AJM, Hartl F-U, Wickner W (1991) Delta mu H+ and ATP function at different steps of the catalytic cycle of preprotein translocase. *Cell* 64:927–939.
33. Bauer BW, Shemesh T, Chen Y, Rapoport TA (2014) A "push and slide" mechanism allows sequence-insensitive translocation of secretory proteins by the SecA ATPase. *Cell* 157:1416–1429.
34. Gelis I, et al. (2007) Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* 131:756–769.
35. Chiba K, Mori H, Ito K (2002) Roles of the C-terminal end of SecY in protein translocation and viability of *Escherichia coli*. *J Bacteriol* 184:2243–2250.
36. Derman AI, Puziss JW, Bassford PJJ, Jr, Beckwith J (1993) A signal sequence is not required for protein export in prlA mutants of *Escherichia coli*. *EMBO J* 12:879–888.
37. Li P, Beckwith J, Inouye H (1988) Alteration of the amino terminus of the mature sequence of a periplasmic protein can severely affect protein export in *Escherichia coli*. *Proc Natl Acad Sci USA* 85:7685–7689.
38. Mori H, Ito K (2006) The long α-helix of SecA is important for the ATPase coupling of translocation. *J Biol Chem* 281:36249–36256.
39. Whitehouse S, et al. (2012) Mobility of the SecA 2-helix-finger is not essential for polypeptide translocation via the SecYEG complex. *J Cell Biol* 199:919–929.
40. Allen W, et al. (2016) Two-way communication between SecY and SecA suggest a Brownian ratchet mechanism for protein translocation. *eLife* 5:15598.
41. Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* 185:60–89.
42. Musial-Siwek M, Rusch SL, Kendall DA (2005) Probing the affinity of SecA for signal peptide in different environments. *Biochemistry* 44:13987–13996.
43. Lakowicz JR (2006) *Principles of Fluorescence Spectroscopy* (Springer, New York), 3rd Ed.
44. Jilaveanu LB, Zito CR, Oliver D (2005) Dimeric SecA is essential for protein translocation. *Proc Natl Acad Sci USA* 102:7511–7516.