Structural similarities and differences between two functionally distinct SecA proteins: the *Mycobacterium tuberculosis* SecA1 and SecA2

Stephanie Swanson[1]
Thomas R. Ioerger[2]
Nathan W. Rigel[3†,4]
Brittany K. Miller[4]
Miriam Braunstein[4]
James C. Sacchettini[1,*]

[1]Department of Biochemistry & Biophysics, Texas A&M University
[2]Department of Computer Science and Engineering, Texas A&M University
[3]Department of Biology, Hofstra University, Hempstead, New York
[4]Department of Microbiology and Immunology, University of North Carolina - Chapel Hill
†current affiliation
*corresponding author, Contact Information:
2138 ILSB
TAMUS 3474
College Station TX, 77843-3474 USA
phone: (979)862-7636
fax: (979)862-7638
email: sacchett@tamu.edu

ABSTRACT

While SecA(1) is the ATPase component of the major bacterial secretory (Sec) system, mycobacteria and some Gram-positive pathogens have a second paralog, SecA2. In bacteria with two SecA paralogs, each SecA is functionally distinct and they cannot compensate for one another. Compared to SecA1, SecA2 exports a distinct and smaller set of substrates, some of which have roles in virulence. In the mycobacterial system, some SecA2-dependent substrates lack a signal peptide while others contain a signal peptide but possess features in the mature protein that necessitate a role for SecA2 in their export. It is unclear how SecA2 functions in protein export, and one open question is whether SecA2 works with the canonical SecYEG channel to export proteins. In this study, we report the structure of *M. tuberculosis* SecA2, which is the first structure of any SecA2 protein. A high level of structural similarity is observed between SecA2 and SecA1. The major structural difference is the absence of the helical wing domain, which is likely to play a role in how *M. tuberculosis* SecA2 recognizes its unique substrates. Importantly, structural features critical to the interaction between SecA1 and SecYEG are preserved in SecA2. Further, suppressor mutations of a dominant-negative *secA2* mutant map to the surface of SecA2 and help identify functional regions of SecA2 that may promote interactions with SecYEG or the translocating polypeptide substrate. These results support a model in which the mycobacterial SecA2 works with SecYEG.

Importance

SecA2 is a paralog of SecA1, which is the ATPase of the canonical bacterial Sec secretion system. SecA2 has a non-redundant function with SecA1, and SecA2 exports a distinct and smaller set of substrates than SecA1. This work reports the crystal structure of SecA2 of *Mycobacterium tuberculosis* (the first SecA2 structure reported for any organism). Many of the structural features of SecA1 are conserved in the SecA2 structure, including putative contacts with the SecYEG channel. Several structural differences are also identified that could relate to the unique function and selectivity of SecA2. Suppressor mutations of a *secA2* mutant map to the surface of SecA2 and help identify functional regions of SecA2 that may promote interactions with SecYEG.

INTRODUCTION

SecA is the ATPase component of the bacterial Sec secretion pathway (1). SecA recognizes proteins destined for export from the cytoplasm and provides energy to translocate them across the cytoplasmic membrane by way of the SecYEG translocase channel. The proteins exported by SecA are synthesized as preproteins with N-terminal signal peptides. Following translocation, the signal peptide is cleaved to release the mature protein species. Both the signal peptide and features of the mature protein are recognized by SecA (2). Some Gram-positive and acid-fast bacteria, including mycobacteria, have a SecA paralog referred to as SecA2. SecA1, the canonical SecA in these organisms, is essential for growth and responsible for the majority of protein export

69 that occurs. In contrast, SecA2 is typically not essential and is required for the export of a
70 more limited subset of proteins (3, 4). Studies in mycobacteria show that even when over-
71 expressed, the two SecA proteins are unable to compensate for each other (5). Thus, each
72 SecA protein has distinct functions in protein export. In *Mycobacterium tuberculosis*
73 (*Mtb*), SecA2 is not essential for growth in culture but it is essential for *Mtb* virulence *in*
74 *vivo* (6) (7). Further, SecA2 is required for intracellular growth of *Mtb* in macrophages
75 (8). The role of SecA2 in promoting growth in macrophages is attributed to a role in
76 preventing phagosome maturation (9). In *Mycobacterium marinum*, export of protein
77 kinase G (PknG) by the SecA2 pathway is suggested to at least be partially responsible
78 for the SecA2 effect on phagosome maturation (10). In *Mtb*, the SecA2 pathway is
79 additionally required to restrict apoptosis of infected macrophages. A possible
80 explanation for this latter effect is the SecA2-dependent secretion of superoxide
81 dismutase, which may reduce ROS-mediated apoptosis (11, 12). An association between
82 SecA2 and the secretion of virulence factors extends to other bacterial pathogens, as well
83 (13-16). There is also an intriguing association between the SecA2 pathway and the
84 export of S-layer proteins by some Gram-positive bacteria, such as *Bacillus anthracis*
85 (17) and *Clostridium difficile* (18).
86
87 It is unclear how *Mtb* SecA2 carries out its unique function in protein export. In some
88 organisms with two SecAs, there is a SecY paralog (SecY2), with which SecA2 likely
89 interacts (19). In SecA2-SecY2 systems, SecY2 and several accessory Sec proteins (Asp)
90 are thought to form an accessory protein translocation channel in the cytoplasmic
91 membrane (4). Mycobacteria, however, are in a group of bacteria referred to as 'SecA2-

only' systems that lack a second SecY ortholog (3). Mycobacteria, as well as several Gram-positive species including *Listeria monocytogenes* (13), *Corynebacterium glutamicum* (20), and *C. difficile* (18) are in the 'SecA2-only' group. An important but unresolved question is whether SecA2 works with the canonical SecYEG channel to export proteins in these systems lacking a second SecY.

The mycobacterial proteins currently known to be exported by SecA2 include examples with typical Sec signal peptides, as well as proteins lacking signal peptides altogether (3). Superoxide dismutase (SodA) in *Mtb* and PknG in *Mtb* and in *M. marinum* are examples of proteins lacking signal peptides that are exported in a SecA2-dependent manner (7, 10, 21). Of the signal peptide-containing proteins exported by the SecA2 systems of *Mycobacterium smegmatis* (22), *M. marinum* (10) and *Mtb* (21), the most thoroughly studied proteins are the *M. smegmatis* Ms1704 and Ms1712 proteins (22). Studies of Ms1704 and Ms1712 demonstrate that they require their signal peptide for export, but it is a feature of the mature portions of these proteins that necessitates export via the SecA2-dependent pathway (23). Interestingly, when fused to a signal peptide for the Twin-arginine translocation (Tat) pathway the mature domain of Ms1704 is exported by the Tat pathway. This result suggests that the defining feature of SecA2 substrates may be a tendency to fold prior to export (23). This is because proteins that get translocated across the membrane by the Tat pathway must be folded in the cytoplasm prior to export (24). In contrast, preproteins exported by the canonical SecA must be unfolded (25), sometimes with the help of export chaperones (26, 27), due to the narrow diameter of the SecYEG central channel. Therefore, if SecA2 works with SecYEG, the role of SecA2

may be to facilitate the export of proteins that have a tendency to fold prior to export by either helping to maintain such proteins in an unfolded state or assisting in the recognition or export of such problematic substrates.

There is only 38% amino acid sequence identity between *Mtb* SecA1 and SecA2 proteins. Yet, SecA2, like SecA1, has a DEAD-box ATPase domain (28) and ATPase activity is required for SecA2 function (29). Further, SecA2 variants lacking ATPase activity due to an amino acid substitution in the Walker-box are dominant negative, and a *secA2* dominant negative mutant exhibits *secA2* mutant phenotypes (growth defect on rich agar and azide sensitivity) that are more severe than those exhibited by a Δ*secA2* null mutant (29). Extragenic suppressors of this dominant-negative *secA2* allele map to the *secY* promoter, and increased SecY levels suppress the *secA2* dominant-negative phenotype (30). These findings suggest that the SecA2 dominant-negative protein is locked in a non-productive interaction with the essential SecYEG channel, which inhibits SecYEG function but can be overcome by increased SecY production. This is consistent with SecA2 working with SecYEG. In a recent study of the SecA2-only system of *L. monocytogenes*, suppressors of a *secA2* mutation also mapped to *secY* (Durack et al., 2015). Furthermore, behavior of a dominant-negative SecA1 mutant in the *C. difficile* system is consistent with the SecYEG translocase used by SecA1 also being used by SecA2 (18). Thus, it seems likely that, in these SecA2-only systems, SecY is involved. However, a direct interaction between SecA2 and SecYEG has not been demonstrated in any system.

138 Previously, the crystal structure of the canonical SecA1 was solved in *Mtb* (31), as well
139 as several other organisms, including *Escherichia coli* (32), *Bacillus subtilis* (33),
140 *Thermotoga maritima* (34) and *Thermus thermophilus* (35). SecA structures contain five
141 canonical domains, organized roughly in the shape of a barbell: a core helical scaffold
142 domain (HSD, forming the "axis"), 2 nucleotide-binding domains (NBD1 and NBD2)
143 which together form a DEAD-box, RecA-like, or superfamily II helicase motor domain
144 on one end of the barbell, and a helical-wing domain (HWD) and preprotein cross-linking
145 domain (PPXD) on the other end of the barbell. In addition, a helix-loop-helix domain
146 called IRA1 (for "intramolecular regulator of ATPase") packs against the HSD, with
147 helices aligned in parallel. The loop connecting the helices of IRA1 is known as the two-
148 helix finger (2HF). The 2HF is shown to insert into the SecYEG pore and it is proposed
149 to promote forward movement of the preprotein through the channel (34, 36), although
150 the interaction between the 2HF and SecYEG could also serve an alternate role besides
151 pushing the translocating protein through the channel (37). During preprotein
152 translocation, SecA undergoes significant conformational changes, one of which involves
153 the orientation of the PPXD domain. According to one model (38), the PPXD likely
154 starts out oriented towards the HWD, forming a hydrophobic “cleft” for binding the
155 signal peptide of the preprotein (39, 40), and then rotates toward NBD2 to form a
156 “clamp” around the translocating polypeptide chain, which has been proposed to be
157 initiated by docking with SecYEG (41).
158
159 In order to better understand the unique function of SecA2, we solved the crystal
160 structure of *Mtb* SecA2, which is the first SecA2 structure to be determined in any

organism. The structure reveals that the HWD domain is completely absent in *Mtb* SecA2. The HWD could play a role in interacting with protein substrates, as it forms part of a cleft with the PPXD that is implicated in peptide binding (40). Though the residues that directly bind the signal peptide (based on NMR studies) are contributed by the PPXD and IRA1 domains (40), the HWD would likely be physically proximal to the untranslocated portion of protein substrates. Further, residues in the HWD of *E. coli* SecA (along with the PPXD and HSD) have been shown to cross-link with synthetic signal peptides in cysteine-substitution experiments (42). The lack of an HWD in SecA2 leads to a signal peptide binding cleft that is more highly solvent-exposed than in SecA1, which we propose could account for recognition of specific SecA2-dependent substrates and prevent export of the larger number of SecA1-dependent preproteins. The structure also reveals conservation in *Mtb* SecA2 of features critical to the interaction between SecA and SecYEG proteins. Finally, by mapping intragenic suppressor mutations onto the SecA2 structure, we show that the mutated residues appear in surface-exposed regions and map to three functional domains that are likely involved in mediating interactions with other protein partners such as SecYEG.

RESULTS

Crystal Structure of *Mtb* SecA2

Mtb SecA2 (Rv1821) was crystallized in space group $P2_1$, and the structure was solved by single-wavelength anomalous diffraction (43) to a resolution of 2.8 Å. The asymmetric unit of the crystal contains a single monomer, and there is no indication of a

184 higher-order oligomer in the crystal lattice. A total of 705 out of 778 residues of the apo-
185 protein were visible in the electron density and could be built. The crystallographic
186 statistics are shown in Table 1.
187
188 **Broad structural similarity between *Mtb* SecA1 and SecA2**
189 The tertiary structure of SecA2 is very similar overall to *Mtb* SecA1 and other orthologs
190 in the SecA family (Figure 1). SecA2 has a long 65 Å (45-amino acid) helix scaffold
191 (HSD), which interconnects four other domains, including two nucleotide-binding
192 domains (NBD1 and NBD2), the IRA1 domain, and the PPXD domain. NBD1 and
193 NBD2 pack together to form a DEAD-box motor domain with an ATP-binding site
194 between them. Catalytically-important residues, such as K115 and R545 are conserved
195 (see Figure S1), consistent with demonstrated ATPase activity of SecA2 (28). As in
196 other SecA structures, the IRA1 domain consists of a pair of alpha-helices packed in
197 parallel to the HSD (forming a 3-helix bundle) and connected by a 9-amino acid loop
198 (known as the 2-helix finger, 2HF). SecA2 lacks the ~70-amino acid C-terminal domain
199 (CTD) which is present in SecA1 orthologs. However, the short linker to this domain,
200 called the C-terminal linker (CTL, residues 734-778), is retained in the SecA2 sequence.
201 The CTL is largely disordered in the crystal structure. However, as observed in previous
202 SecA structures (33), part of the CTL of SecA2 (residues 749-759, shown in yellow in
203 Figure 1) forms a third β-strand along the outside of the preprotein binding site. Note
204 that this region is preceded by a disordered loop (residues 734-748), which appears as a
205 discontinuity between IRA1 and CTL in the figure, and followed by only 19 residues to
206 the C-terminus, which are also disordered. During model-building, sequence assignment

in this strand was aided by the location of SeMet757 and the density of bulky side chains, which helped to rule out the possibility of a bound preprotein substrate.

Differences between the structures of *Mtb* SecA1 and SecA2

Despite the overall similarity between the structures of SecA1 and SecA2, there are several notable differences. One structural difference between SecA1 and SecA2 is found in the nucleotide-binding region. SecA2 lacks the VAR domain (44), which in other SecA orthologs consists of a pair of helices that reach out from NBD2 and cover over the ATP-binding site (Figure 2). Consequently, the ATP binding site is more solvent-exposed in SecA2. The VAR domain is present in some SecA orthologs, including *Mtb* SecA1 (31) and *E. coli* SecA (32), but it is absent in others, such as *B. subtilis* (33) and *T. maritima* SecA (39). The functional significant of the absence of the VAR domain in SecA2 is unknown.

A second structural difference involves the orientation of the PPXD domain. As in other SecA structures, the PPXD domain consists of an α+β fold that is attached to the NBD1 motor domain by a pair of anti-parallel β-strands that cross over the HSD. The PPXD of *Mtb* SecA2 occupies a distinct orientation compared to previous SecA structures, as illustrated in Figure 3. The PPXD in previous SecA structures been observed in several different orientations ranging from contact with the HWD (to form a "signal peptide binding-cleft closed" conformation, as observed in 1nl3) to contact with NBD2 (to form a "preprotein clamp closed" conformation, as observed in 3din) (38, 45) produced by a

229 rigid-body rotation relative to the rest of the protein (38, 39). The PPXD in SecA2
230 occupies an intermediate position between these two extremes.
231
232 The most striking structural difference in SecA2 is that the HWD is missing (Figure 2)
233 due to deletion of 70 amino acids that form a helical domain at the end of the HSD, as
234 anticipated from the sequence alignment (Figure S2). In SecA2, the remaining 23
235 residues connect the HSD directly to IRA1, bypassing the helical wing domain. In other
236 SecA structures, including *Mtb* SecA1, the body of the HWD forms a deep hydrophobic
237 cleft with PPXD, which can open or close against it (39), with the signal peptide binding
238 site at the base (formed by residues from PPXD and IRA1) (39, 40). The absence of the
239 HWD in SecA2 makes the cleft significantly more open and solvent-exposed (illustrated
240 in Figure S3), which could help SecA2 recognize its unique substrates that are
241 distinguished by features of their mature domains, possibly a tendency to fold prior to
242 export (23).
243
244 The functionally-important two-helix finger (2HF), which is a 9-residue loop connecting
245 two helices in the IRA1 domain that inserts into the SecYEG pore, is conserved in the
246 *Mtb* SecA2 structure (residues 695-703, Table 2). However, the 2HF loop in *Mtb* SecA2
247 adopts a different three-dimensional conformation compared to previous structures. In
248 the *Mtb* SecA2 structure the 2HF is observed to close down approximately 10Å onto the
249 HSD, like a jaw-hinge (Figure 4), due to differences in how the ends of the helices
250 unwind (even though the 2HF amino acid sequence itself is highly conserved, as shown
251 in Table 2). This orientation contrasts with the conformation observed in most other SecA

structures, in which the loop is more flipped out into solvent (Figure 4); however, the conformation of the 2HF loop is also quite variable among SecA crystal structures (Figure S4). Fluorescence studies also suggest that the 2HF loop is flexible and can adopt different conformations in solution (37).

Similarities between SecYEG binding regions of *Mtb* SecA1 and SecA2

The conservation of the overall structure of SecA2 is consistent with a model in which SecA2 works with SecYEG to translocate SecA2-dependent proteins across the membrane. Further, the key regions of SecA2 that would interact with the SecYEG pore are conserved, including the 2HF. The helix-terminating proline in the 2HF is present in SecA2 (Pro703), as it is in all SecA homologs (see Table 2). Tyr794 in *E. coli* SecA is another critical residue in the 2HF (36). Although it is substituted by Leu698 in *Mtb* SecA2, this tyrosine is substituted by large hydrophobic residues in 20% of SecA homologs (methionine in *Mtb* SecA1). Further, structural data from the *Tm* SecA-SecYEG complex supports that hydrophobic substitutions, such as leucine, can be accommodated at this position, as the side-chain sits in a hydrophobic pocket in SecY (34).

Structural superposition of *Mtb* SecA2 onto *Tm* SecA in the *Tm* SecA-SecYEG complex (3din; (34); Figure 5 and Figure S5) further indicates that SecA2 preserves many of the structural features of SecA implicated in binding to SecYEG. This includes amino acids in the *Mtb* SecA2 2HF and immediately adjacent regions of IRA1 that contact SecY in

the *Tm* SecA-SecYEG complex (amino acids 687-715 in SecA2) (Table 3 and Figure 5). There are also regions of NBD2 and the HSD that are structurally conserved in the SecA2 structure and positioned for contact with SecY (Table 3 and Figure 5). These residues in NBD2, IRA1, and the HSD are clustered at the interface with SecY. In addition, although the PPXD of SecA2 is rotated away and does not appear to make direct contact with SecY in the superposition, if it were rotated into an orientation similar to that observed in the *Tm* SecA in the complex, it would place additional SecA2 residues (listed in Table 3) in contact with SecY, as shown in Figure 5. It is notable that *Mtb* SecA2 D607 (in the HSD) corresponds to one of the residues in *E. coli* SecA (640) that can be cross-linked with SecY using photo-activatable unnatural amino acids (46).

Mapping of suppressor mutations on the SecA2 structure

Prior studies indicate that a SecA2 dominant negative protein with an amino acid substitution in the ATP binding Walker-box, making it unable to bind ATP, is locked in a non-functional complex, likely with SecYEG, at the membrane. In order to identify important residues in SecA2, we identified intragenic suppressor mutations that could overcome the *secA2* dominant negative phenotypes (30) with the rationale being that such mutations might map to sites of protein interactions in SecA2 complexes. For convenience, these experiments were performed with the *M. smegmatis* ortholog of SecA2, which has 83% amino acid identity to *Mtb* SecA2 and is able to substitute for the *Mtb* SecA2 in cross-species complementation experiments (29). An *M. smegmatis* (*Ms*) strain expressing the dominant negative *Ms* SecA2^{K129R}, which has an amino acid substitution in the Walker box (equivalent to K115 in *Mtb* SecA2) was used. All

298 suppressors identified reversed the severe dominant negative phenotypes caused by
299 SecA2^{K129R}, as assessed by azide sensitivity assays (5) and colony size on rich agar (30)
300 (Figure 6 and data not shown), but they still exhibited a phenotype similar to that of a
301 Δ*secA2* null mutant.
302
303 Eight independent suppressors with mutations in the coding sequence of *secA2*^{K129} were
304 identified by sequencing, mapping to four different domains: NBD1, NBD2, PPXD, and
305 IRA1 (Table 4). All eight suppressor mutants produced full-length SecA2 protein at
306 normal levels, as confirmed by western blot analysis. Each mutation was validated to be
307 responsible for the suppression by retesting the phenotype of individual mutations when
308 introduced into a fresh *secA2*^{K129R} mutant background (Figure 6).
309
310 When mapped to the SecA2 structure, all of the suppressor mutations were located on the
311 surface of the protein (Figure 7). For simplicity, below we will refer to the suppressors
312 using amino acid numbering that corresponds to *Mtb* SecA2 (Table 4). There were three
313 categories of suppressors. The first set of suppressor mutations affected the same surface
314 loop of NBD1. There were two suppressors derived from independent cultures with
315 identical mutations in NBD1 and a third suppressor with a different mutational alteration
316 that mapped to the same site in NBD1. These NBD1 suppressors involve a 4-residue
317 loop ^{168}STPD^{172} in *Mtb* connecting a β-strand and an α-helix; this loop was deleted in
318 one mutant and duplicated in another. It is currently unknown what role these residues
319 play, but it is striking that three out of eight suppressor mutations involved this surface-

localized loop of the nucleotide binding domain, suggesting it is a functionally important point of contact for SecA2.

The second group of suppressors (three in total) clustered in the SecA "polypeptide clamp" region made up of PPXD and NBD2 domains. Two suppressor mutations mapped to the SecA2 PPXD domain: a non-synonymous substitution D316H and an insertion of a second glutamate at E354. These amino acids are in separate loops in the PPXD domain, but they are proximal in the three-dimensional structure, approximately 7Å apart (Figure 7a). The PPXD is positioned far from the NBD2 domain in the SecA2 structure (distance between closest residues of the two domains is 23 Å, representing a "clamp open" state). However, in the *Tm* SecA-SecYEG complex, the corresponding PPXD loops to which these suppressor mutations map come in contact with NBD2. Moreover, the *Tm* SecA residue corresponding to the D316H suppressor in the *Mtb* SecA2 PPXD is in direct contact with NBD2 in the *Tm* SecA-SecYEG complex (34) (illustrated in Figure 7b). It should be noted that this *Tm* SecA complex with SecYEG represents an extreme conformation (induced by ADP and BeFx in the crystallization buffer) in which the preprotein channel is entirely collapsed (a loop of the PPXD actually inserts into the preprotein binding channel). In a structure of SecA bound to a preprotein substrate (3jv2; (47)), the PPXD does not rotate quite as far toward NBD2 as in the SecA-SecYEG complex, but the residues corresponding to the suppressor mutations are still on the surface of the PPXD in a region that would be in position to interact with SecYEG or the lipid bilayer (similar to the red residues highlighted in Figure 5). Thus, these suppressor mutations could disrupt intramolecular interactions when the PPXD

343 rotates to form the "clamp" around the translocating polypeptide (41), or could lock it in
344 the extreme closed state such that the preprotein channel is collapsed altogether.
345 Strikingly, the NBD2 suppressor T449I in *Mtb* also maps to the SecA "preprotein clamp"
346 region, and is proximal (within 10 Å) to the two PPXD suppressor mutations when the
347 clamp is closed (based on the analogous residues in the *Tm* SecA-SecYEG docked
348 structure (34), Figure 7). Thus, these three suppressors in NBD2 and PPXD could
349 conceivably cause a defect in clamp closure during translocation. In light of past studies
350 suggesting that interactions between SecA2^{K129R} and SecYEG are responsible for the
351 dominant negative phenotype (30), these results suggest that a defect in clamp closure
352 may dislodge or prevent SecA2 interactions with SecYEG by disrupting interactions with
353 the polypeptide being translocated through the channel.
354
355 The final group of intragenic suppressors identified have deletions in IRA1. One
356 suppressor has a deletion of 714-721 in *Mtb* and another suppressor has a very similar,
357 yet distinct, deletion of residues 712-719 in *Mtb*. These deletions are in the middle of one
358 of the α-helices, just downstream from the 2HF that forms part of the interface with
359 SecYEG (Figures 5, 7 and Table 3), and similar mutations in IRA1 have previously been
360 shown to disrupt binding to SecYEG (48). Furthermore, one of the deleted residues in
361 both of the IRA1 suppressors is phenylalanine Phe715, which is a conserved residue
362 predicted to contact SecY (colored red in Figure 5) that is equivalent to the highly
363 conserved Phe798 (in *Tm* SecA). In the *Tm* SecA-SecYEG structure, Phe798 (in *Tm*
364 SecA) forms an aromatic stacking interaction with Tyr418 in the C-terminal tail of *Tm*
365 SecY (34). This interaction appears to be crucial to docking as the equivalent tyrosine

366 residue in *E. coli* SecY (Tyr429) is the location of a cold-sensitive mutation that
367 prevented insertion of SecA into the membrane channel (49). These interacting residues
368 are highly conserved in all Sec systems, including *Mtb* SecA2 (Phe715) and *Mtb* SecY
369 (Tyr436). The fact that this group of intragenic *secA2* suppressor mutants harbors
370 deletions in a structurally conserved and critical SecY-interacting region of IRA1 (Figure
371 5, Table 3) is consistent with their mode of suppression being avoidance of complex
372 formation between SecA2^{K129R} and SecYEG.
373
374 **Intragenic suppressors alter membrane localization of the dominant negative SecA2**
375 In wild-type *M. smegmatis* SecA2 is predominantly found in the soluble cytoplasmic-
376 containing fraction. In contrast, the localization of SecA2^{K129R} is almost exclusively in
377 the membrane-containing cell envelope pellet (29) (Figure 8). This is consistent with a
378 model for SecA2^{K129R} being locked in a protein complex with SecYEG at the membrane.
379 Since we predicted that some of the intragenic suppressors alleviate SecYEG interactions
380 we determined the membrane localization of SecA2^{K129R} in the intragenic suppressor
381 mutant background. Strains were lysed and then fractionated into cell envelope (pellet)
382 and soluble (cytoplasmic) fractions. Western blot analysis with anti-SecA2 antibodies on
383 fractions was then carried out to localize the protein. In each of the representative
384 intragenic suppressors analyzed, the distribution of SecA2^{K129R} shifted from the envelope,
385 as seen in the starting *secA2*^{K129R} strain, to the soluble cytoplasmic fraction (Figure 8).
386 Suppressor mutations in the "clamp" (PPXD and NBD2) and IRA1 domains had the most
387 dramatic effects, restoring partitioning of SecA2 between cell envelope and cytoplasm to
388 almost wild-type levels. This data supports a model in which the intragenic suppressor

mutations alleviate the dominant-negative phenotype by disrupting protein-protein interactions involving the SecYEG membrane complex and/or the translocating polypeptide.

Discussion

Over 30 years ago, SecA was identified as a critical component of the protein export system of bacteria (50). Since that time, there has been extensive genetic, molecular, biochemical, biophysical and structural studies to understand SecA function. Of the two SecAs in *Mtb*, SecA1 is the counterpart of the well-studied canonical SecA, while SecA2 has a distinct function from SecA1 and a non-overlapping substrate specificity profile. The structure of *Mtb* SecA2 we report is the first structure of any SecA2 protein. The broad structural similarity observed between the two solved *Mtb* SecA structures indicates that, even after decades of mechanistic studies, gaps in our understanding of SecA proteins remain.

The smaller size of SecA2 versus SecA1 and canonical SecA proteins appears to come from the absence of the HWD, the VAR domain and a C-terminal domain (CTD, though it still retains the CTL linker) reducing the overall size of the protein product from 949 aa to 778 aa (Figure S2). The lack of a HWD is the most striking structural difference in SecA2. Without the HWD, the signal peptide recognition site of SecA2 is more solvent-exposed and thus more accessible to protein substrates. This structural difference may help explain the ability of SecA2 to export substrates with distinctive

features of their mature domain, possibly a propensity to fold prior to export (23). The "open" nature of the cleft created by the absence of a HWD could provide a broad surface against which folded proteins could possibly dock and unfold for translocation through the SecYEG transmembrane pore. Several pieces of experimental evidence support the possibility that the HWD could interact with preproteins. While an NMR structure of SecA bound to a signal peptide did not identify any residues of the HWD that directly interact with the signal peptide (40), several residues of the HWD were found to form cysteine-based cross-links with a synthetic signal peptide (42), which may result from transient states (i.e. alternative conformations of the HWD) sampled dynamically in solution. In fact, the HWD is observed to rotate by up to 15º between different crystal structures, depending on oligomeric state (39). This suggests the HWD itself is mobile in solution, which is supported by fluourescence-based (FRET) studies (51). Furthermore, the mobility of the HWD appears to be influenced by the presence of a preprotein (52). Thus, the absence of the HWD in SecA2 could potentially affect substrate recognition. The 70-residue deletion of the HWD observed in *Mtb* SecA2 is a general feature among actinomycetes (including *Mycobacterium* and *Corynebacterium* species) (see Figure S2). It should be noted that other Gram-positive SecA2 proteins also appear to have a truncated version of this domain (deletions of 13-18 residues for *S. gordonii* and *L. monocytogenes*, respectively). Until structures of these other SecA2 orthologs are solved, the potential consequences of these HWD truncations remains unknown. It is possible that a reduced HWD could open up the signal-peptide binding cleft and/or increase the site of interaction with preproteins, as we propose for *Mtb* SecA2. To achieve a complete

434 picture of SecA2 function going forward, the consequences of a truncated or deleted
435 HWD will need to be explored in both mycobacterial and Gram-positive SecA2 proteins.

436 The significance of the absence of the VAR domain in the SecA2 NBD region is
437 less clear. The lack of the VAR domain leaves the nucleotide-binding site relatively
438 solvent-exposed. While other SecA2 orthologs also lack the VAR domain (Figure S2),
439 one third of bacterial SecA(1) proteins lack this domain as well (44). In *E. coli* SecA, the
440 VAR domain has been shown to regulate ATPase activity and ADP release, as *secA* Δ*var*
441 mutants display higher ATPase activity and faster ADP release rates (44). However, *Mtb*
442 SecA2 (28) was recently reported to release ADP more slowly (not more quickly) than
443 the VAR-containing *Mtb* SecA1 (53).

444 Mycobacterial SecA2 proteins, as well as SecA2s in many other organisms, lack
445 the C-terminal domain (CTD) (Figure S2). The CTD in SecA1 proteins consists of a tail
446 of ~70 amino acids that is disordered in all previous crystal structures (1). In most
447 bacteria, the CTD of SecA contains a Zn^{2+}-finger domain that binds to the protein export
448 chaperone SecB (54). Mycobacteria are an exception, in that the CTD of SecA1 does not
449 contain the conserved cysteines of a Zn^{2+}-finger motif. However, this may not be too
450 surprising because, like Gram positive bacteria (55), no SecB ortholog with a function in
451 protein export has so far been identified in mycobacteria. Thus, because of the lack of
452 Zn^{2+}-finger motif in the CTD of SecA1 and lack of a SecB ortholog, the absence of a
453 CTD in *Mtb* SecA2 seems unlikely to be a significant contributing factor to the unique
454 function of SecA2.

455 In comparison to all prior SecA structures, the SecA2 structure also revealed new
456 orientations of the PPXD and the 2HF loop. However, these differences probably reflect

457 the conformational plasticity of these two structural elements. Given the mobility of the
458 PPXD domain already established for canonical SecA proteins, it seems likely that the
459 PPXD orientation observed in SecA2 represents a previously unobserved structural
460 intermediate in the transition of the preprotein binding clamp from the open to closed
461 position (38). The unique orientation of the 2HF loop observed in SecA2, which occurs
462 at a key point of interaction with the translocation channel and varies considerably among
463 SecA structures, is probably a consequence of the flexibility of this loop in solution.

464 Given that there is no corresponding SecY2 partner in the *M. tuberculosis*
465 genome, an important mechanistic question to be answered is whether SecA2 works with
466 the canonical SecYEG channel to export proteins. In prior studies, we described a
467 dominant negative *secA2* mutation that exhibits more severe phenotypes than a Δ*secA2*
468 deletion mutant (29). Such phenotypes often result from a dominant negative protein
469 being locked in a non-productive complex with its normal binding partners. Further, we
470 showed extragenic suppressors that overexpress SecY suppress the *secA2*^{K129R} dominant
471 negative phenotype, which argues for an interaction between SecA2^{K129R} and SecY (30).
472 Here, we identified intragenic suppressors of *secA2*^{K129R}, and all of them mapped to the
473 surface of the SecA2 structure. One group of suppressors mapped to the IRA1 domain of
474 SecA2 in regions where similar mutations disrupt *E. coli* SecA binding to SecYEG (48).
475 These IRA1 suppressors also restored cytoplasmic localization of SecA2^{K129R}. These
476 results can be explained by the IRA1 suppressor mutations preventing SecA2^{K129R}
477 interactions at the membrane SecYEG channel, and they support the model for SecA2
478 working with SecY to promote export of its specific substrates. The suppressors that
479 mapped to the "polypeptide clamp" region of SecA could similarly suppress the dominant

480 negative phenotype. However, in this case, the suppression would result from the
481 inability of SecA2 to trap the translocating polypeptide in the center of the SecYEG
482 channel, causing SecA2 to fail to engage SecYEG (without the substrate), or causing the
483 ternary system (SecYEG-SecA2-preprotein) to dissociate.

484 The SecA structure reported here is of a monomer. In other studies SecA proteins
485 have been crystallized as monomers (56) or dimers (39), and the issue of the oligomeric
486 state of SecA during protein translocation has remained controversial (37, 57-59). A
487 recent study demonstrated the ability of recombinant *Mtb* SecA1 and SecA2 to physically
488 interact *in vitro* (60). If SecA1-SecA2 heterodimers form, it is possible that interactions
489 between SecA1 and SecY might avoid the need of SecA2 to directly interact with SecY.
490 However, it is currently unclear if SecA1-SecA2 dimers exist and/or are functional in
491 mycobacteria. Further, the dominant negative SecA2 phenotypes and the intragenic
492 suppressors reported here, combined with structural conservation of SecA-SecY contact
493 sites in SecA2, argue for the ability of SecA2 and SecY to interact. Ultimately, to clarify
494 the mechanistic details of SecA2-dependent protein export it will be necessary to study
495 the pathway with an *in vitro* reconstitution system, as was used to dissect the mechanistic
496 details of the *E. coli* Sec pathway.

497 Since the SecYEG channel requires that proteins be unfolded for translocation
498 (25), the possibility of SecA2 working with the SecYEG channel is intriguing, in light of
499 experiments suggesting that SecA2 substrates are distinguished by a tendency to fold in
500 the cytoplasm (23). The role of SecA2 could be to promote recognition of proteins that
501 would normally be overlooked by the canonical SecA1-SecYEG translocase or to help
502 maintain proteins in an unfolded state prior to or during export. The regions of structural

503 difference and suppressor mutations identified in this study represent exciting new
504 directions for exploring the functional differences between SecA2 and SecA1 proteins.
505
506 **METHODS**
507 **Protein Expression and Purification**
508 The 778-residue open reading frame of *Mtb* SecA2 was cloned into expression vector
509 pNR14. Several genomic databases list *Mtb* SecA2 as having a total length of 808 amino
510 acids (e.g. NCBI accession NP_216337). However, the start site in this annotation is
511 likely to be incorrect, as the first 30 amino acids are not required for function and
512 represents an N-terminal extension that is not observed in other SecA orthologs (28, 61).
513 Therefore, we designate the GTG codon corresponding to residue 31 in the NCBI
514 annotation as the true start codon, yielding a total ORF length of 778 amino acids. The
515 expression construct, pNR14, produces a tag-less form of the protein (28).
516 Selenomethionyl protein was produced by transforming the *E. coli* methionine auxotroph
517 B834(DE3) (Novagen) with the pNR14 expression vector. 6 L of culture were grown
518 under standard conditions to mid-log phase. The cells were pelleted and used to
519 innoculate 12 L of M9 minimal media supplemented with 50 mg/L of L-
520 selenomethionine, 50 mg/L of standard L-amino acids (excluding methionine), 100 nm
521 vitamin B12, and trace elements (62). Expression was induced with 0.5 mM ITPG at
522 16°C for 12 hours. Cells were then harvested and resuspended in lysis buffer containing
523 50 mM tris pH 8.0, 50 mM NaCl, 1 mM DTT, 10 mM $MgCl_2$, 20 ug/ml DNase, and 1X
524 protease inhibitor cocktail V (EMD Biosciences). The cells were disrupted in a
525 BeadBeater (Biospec) using 0.1 mm glass beads. Cellular debris was cleared from the

526 lysate by spinning at 27,200 X *g* for 2 hours. The supernatant was then filtered and
527 loaded onto a Blue Sepharose column (GE Healthcare) that had been equilibrated in 50
528 mM tris pH 8.0, 50 mM NaCl, and 1 mM DTT. Protein collected from the flow-through
529 was further purified by anion exchange chromatography using a HiTrap Q HP column
530 (GE healthcare). The purified protein was dialyzed overnight against buffer containing 50
531 mM tris pH 8.0, 50 mM NaCl, and 1 mM DTT and was then concentrated to 10 mg/ml
532 using a Centriprep centrifugal concentrator (Milipore) and flash frozen until further use.
533
534 **Crystallization**
535
536 Purified protein was crystallized in 20% PEG 8K, 0.1 M tris pH 8.0, 0.2 M NaCl, 3%
537 ethylene glycol, and 8 mM 3-[(3-cholamidopropyl)dimethylammonio]-1-
538 propanesulfonate (CHAPS). Wells were set up using sitting-drop vapor diffusion at 21°C,
539 with drops consisting of 2 parts buffer and one part protein. Crystals grew to 100 µm
540 within 3-4 days. Perfluoropolyether (Hampton Research) was used as a cryo-protectant.
541 The protein crystallized in space group $P2_1$ with the unit cell parameters a = 39, b =
542 165, c = 67 Å and β = 97°. The corresponding unit cell volume can accommodate a single
543 molecule in the asymmetric unit.
544
545 **Crystal Dehydration**
546
547 A crystal dehydration method was developed that significantly improved the mosaic
548 spread and diffraction power of the crystals (63). Both the well and drop solution were
549 replaced with mother liquor that had a 3-5% increase in precipitant
550 concentration. Crystals were left to dehydrate for a minimum of 48 hours before making
551 another incremental increase in the precipitant. Successfully dehydrated crystals had a

552 reduced *b* unit cell parameter of up to 15 Å with the largest difference resulting in a
553 10.5% decrease in the unit cell volume. The crystal that produced the best diffraction data
554 and led to structure solution had only a 3 Å difference in the *b* unit cell parameter.
555
556 **Data Collection, Structure Determination, and Refinement**
557 The structure was solved by single-wavelength anomalous dispersion using a seleno-
558 methionine derivative (64). Anomalous diffraction data were collected at beamline 23-ID
559 of the GM/CA-CAT facilities of the Advanced Photon Source, Argonne National
560 Laboratory. Crystals were partitioned using the 10 µm mini-beam (65). This prevented
561 global-scale radiation exposure and allowed for more data to be collected from a single
562 crystal. The data was processed and reduced using the HKL2000 software package (66).
563 The location of 3 Se sites were found using SHELX C/D and were used as a starting
564 point for locating additional sites in autoSHARP (67, 68). The resulting experimental
565 phases extended to 3.8 Å resolution and produced an electron density map in which
566 approximately 60% of the backbone could be placed in NBD1, NBD2, and parts of the
567 HSD. Model building was performed in Coot (69). The phases from the partial model
568 were then combined with the experimental phases using SigmaA and used as a starting
569 point for progressive runs of density modification in DM (70, 71). This facilitated the
570 placement of the backbone in the PPXD as well as in other parts of the model. Initially,
571 sequence was assigned by the positions of the Se atoms and from the density of large side
572 chains. Then a real-space cross-validation procedure called "ping-pong" cross-validation
573 was used to complete the structure (56). Briefly, the model was split into two sets. Side
574 chains that could be identified in the first set of residues were used during phase

575 combination and density modification. The resulting map was used to place side chains
576 for the second set of residues, and the process continued in alternation. Structure
577 refinement was carried out in autoBuster (72). The structural coordinates have been
578 deposited in the Protein Data Bank with the identifier 4UAQ.
579
580 **Suppressor Screen and Reconstruction**
581 Spontaneous suppressors of the *secA2*^{K129R} strain were isolated by plating onto Mueller-
582 Hinton agar at 37 °C, as described previously (30). The *secA2*^{K129R} strain has the
583 chromosomal *secA2* gene deleted and carries a copy of the *secA2* gene encoding for
584 SecA2^{K129R} integrated at the chromosomal L5 *att* site. The *secA2*^{K129R} gene of the
585 suppressors was PCR amplified and sequenced to identify intragenic suppressor
586 mutations. To confirm that suppressor phenotypes were due to sequenced mutations in
587 *secA2*^{K129R}, the intragenic suppressors were recreated in a fresh strain background. PCR
588 amplified *secA2*^{K129R} gene products from the intragenic suppressors were subcloned into
589 pCR2.1 followed by cloning into pMV306. The resulting vectors were electroporated into
590 the Δ*secA2* mutant of *M. smegmatis* and transformants were tested for sensitivity to
591 sodium azide and SecA2 localization.
592
593 **Azide Sensitivity Assay**
594 Cultures were plated for sensitivity to sodium azide as previously described (29). In brief,
595 200 µL of saturated (OD_{600nm} = 2.0) *M. smegmatis* culture was mixed with 3.5 mL of
596 molten 7H9 top agar and then poured onto a 7H10 bottom agar plate lacking tween.
597 Sterile 6-mm filter discs were placed onto the surface of the cooled top agar. 10 µL of

0.15 M sodium azide was then added to the disc. The plates were incubated for 2 days at 37°C, and the resulting zones of growth inhibition were measured. Each strain was tested in triplicate.

Subcellular Fractionation and Immunoblotting

To determine the subcellular localization of SecA2 in *M. smegmatis*, we fractionated bacterial whole cell lysates as described previously (22, 29). Whole cell lysates were generated by five passages through a French pressure cell. The lysates were separated into cell envelope (100,000 x *g* pellet) and soluble (100,000 x *g* supernatant) fractions. Protein derived from the same amount of starting cells for each fraction was analyzed by SDS-PAGE and immunoblots using an anti-SecA2 antibody at a 1:20,000 dilution (73). For quantification, secondary antibody conjugated to alkaline phosphatase was used and detected using the ECF reagent (GE Healthcare). Fluorescence was quantified using a phosphorimager and ImageQuant 5.2 (Molecular Dynamics).

ACKNOWLEDGEMENTS

This work was supported by National Institutes of Health grant NIAID AI054540 (M.B.), NIH grant P01 AIO95208 (J.C.S), and grant A0015 from the Robert A. Welch Foundation (J.C.S).

REFERENCES

1. **Chatzi KE, Sardis MF, Economou A, Karamanou S.** 2014. SecA-mediated targeting and translocation of secretory proteins. Biochim Biophys Acta **1843:**1466-1474.

2. **Gouridis G, Karamanou S, Gelis I, Kalodimos CG, Economou A.** 2009. Signal peptides are allosteric activators of the protein translocase. Nature **462:**363-367.
3. **Feltcher ME, Braunstein M.** 2012. Emerging themes in SecA2-mediated protein export. Nat Rev Microbiol **10:**779-789.
4. **Bensing BA, Seepersaud R, Yen YT, Sullam PM.** 2014. Selective transport by SecA2: an expanding family of customized motor proteins. Biochim Biophys Acta **1843:**1674-1686.
5. **Braunstein M, Brown AM, Kurtz S, Jacobs WR, Jr.** 2001. Two nonredundant SecA homologues function in mycobacteria. J Bacteriol **183:**6979-6990.
6. **Sassetti CM, Rubin EJ.** 2003. Genetic requirements for mycobacterial survival during infection. Proc Natl Acad Sci U S A **100:**12989-12994.
7. **Braunstein M, Espinosa BJ, Chan J, Belisle JT, Jacobs WR, Jr.** 2003. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of Mycobacterium tuberculosis. Mol Microbiol **48:**453-464.
8. **Kurtz S, McKinnon KP, Runge MS, Ting JP, Braunstein M.** 2006. The SecA2 secretion factor of Mycobacterium tuberculosis promotes growth in macrophages and inhibits the host immune response. Infect Immun **74:**6855-6864.
9. **Sullivan JT, Young EF, McCann JR, Braunstein M.** 2012. The Mycobacterium tuberculosis SecA2 system subverts phagosome maturation to promote growth in macrophages. Infect Immun **80:**996-1006.
10. **van der Woude AD, Stoop EJ, Stiess M, Wang S, Ummels R, van Stempvoort G, Piersma SR, Cascioferro A, Jimenez CR, Houben EN, Luirink J, Pieters J, van der Sar AM, Bitter W.** 2014. Analysis of SecA2-dependent substrates in Mycobacterium marinum identifies protein kinase G (PknG) as a virulence effector. Cell Microbiol **16:**280-295.
11. **Break TJ, Jun S, Indramohan M, Carr KD, Sieve AN, Dory L, Berg RE.** 2012. Extracellular superoxide dismutase inhibits innate immune responses and clearance of an intracellular bacterial infection. J Immunol **188:**3342-3350.
12. **Hinchey J, Lee S, Jeon BY, Basaraba RJ, Venkataswamy MM, Chen B, Chan J, Braunstein M, Orme IM, Derrick SC, Morris SL, Jacobs WR, Jr., Porcelli SA.** 2007. Enhanced priming of adaptive immunity by a proapoptotic mutant of Mycobacterium tuberculosis. J Clin Invest **117:**2279-2288.
13. **Lenz LL, Mohammadi S, Geissler A, Portnoy DA.** 2003. SecA2-dependent secretion of autolytic enzymes promotes Listeria monocytogenes pathogenesis. Proc Natl Acad Sci U S A **100:**12432-12437.
14. **Siboo IR, Chambers HF, Sullam PM.** 2005. Role of SraP, a Serine-Rich Surface Protein of Staphylococcus aureus, in binding to human platelets. Infect Immun **73:**2273-2280.
15. **Wu H, Mintz KP, Ladha M, Fives-Taylor PM.** 1998. Isolation and characterization of Fap1, a fimbriae-associated adhesin of Streptococcus parasanguis FW213. Mol Microbiol **28:**487-500.
16. **Xiong YQ, Bensing BA, Bayer AS, Chambers HF, Sullam PM.** 2008. Role of the serine-rich surface glycoprotein GspB of Streptococcus gordonii in the pathogenesis of infective endocarditis. Microb Pathog **45:**297-301.

17. **Nguyen-Mau SM, Oh SY, Kern VJ, Missiakas DM, Schneewind O.** 2012. Secretion genes as determinants of Bacillus anthracis chain length. J Bacteriol **194:**3841-3850.
18. **Fagan RP, Fairweather NF.** 2011. Clostridium difficile has two parallel and essential Sec secretion systems. J Biol Chem **286:**27483-27493.
19. **Takamatsu D, Bensing BA, Sullam PM.** 2004. Four proteins encoded in the gspB-secY2A2 operon of Streptococcus gordonii mediate the intracellular glycosylation of the platelet-binding protein GspB. J Bacteriol **186:**7100-7111.
20. **Caspers M, Freudl R.** 2008. Corynebacterium glutamicum possesses two secA homologous genes that are essential for viability. Arch Microbiol **189:**605-610.
21. **Feltcher ME, Gunawardena HP, Zulauf KE, Malik S, Griffin JE, Sassetti CM, Chen X, Braunstein M.** 2015. Label-free Quantitative Proteomics Reveals a Role for the Mycobacterium tuberculosis SecA2 Pathway in Exporting Solute Binding Proteins and Mce Transporters to the Cell Wall. Mol Cell Proteomics **14:**1501-1516.
22. **Gibbons HS, Wolschendorf F, Abshire M, Niederweis M, Braunstein M.** 2007. Identification of two Mycobacterium smegmatis lipoproteins exported by a SecA2-dependent pathway. J Bacteriol **189:**5090-5100.
23. **Feltcher ME, Gibbons HS, Ligon LS, Braunstein M.** 2013. Protein export by the mycobacterial SecA2 system is determined by the preprotein mature domain. J Bacteriol **195:**672-681.
24. **DeLisa MP, Tullman D, Georgiou G.** 2003. Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. Proc Natl Acad Sci U S A **100:**6115-6120.
25. **Nouwen N, Berrelkamp G, Driessen AJ.** 2007. Bacterial sec-translocase unfolds and translocates a class of folded protein domains. J Mol Biol **372:**422-433.
26. **Weiss JB, Ray PH, Bassford PJ, Jr.** 1988. Purified secB protein of Escherichia coli retards folding and promotes membrane translocation of the maltose-binding protein in vitro. Proc Natl Acad Sci U S A **85:**8978-8982.
27. **Kusters I, Driessen AJ.** 2011. SecA, a remarkable nanomachine. Cell Mol Life Sci **68:**2053-2066.
28. **Hou JM, D'Lima NG, Rigel NW, Gibbons HS, McCann JR, Braunstein M, Teschke CM.** 2008. ATPase activity of Mycobacterium tuberculosis SecA1 and SecA2 proteins and its importance for SecA2 function in macrophages. J Bacteriol **190:**4880-4887.
29. **Rigel NW, Gibbons HS, McCann JR, McDonough JA, Kurtz S, Braunstein M.** 2009. The Accessory SecA2 System of Mycobacteria Requires ATP Binding and the Canonical SecA1. J Biol Chem **284:**9927-9936.
30. **Ligon LS, Rigel NW, Romanchuk A, Jones CD, Braunstein M.** 2013. Suppressor Analysis Reveals a Role for SecY in the SecA2-Dependent Protein Export Pathway of Mycobacteria. J Bacteriol **195:**4456-4465.
31. **Sharma V, Arockiasamy A, Ronning DR, Savva CG, Holzenburg A, Braunstein M, Jacobs WR, Jr., Sacchettini JC.** 2003. Crystal structure of Mycobacterium tuberculosis SecA, a preprotein translocating ATPase. Proc Natl Acad Sci U S A **100:**2243-2248.

32. **Papanikolau Y, Papadovasilaki M, Ravelli RB, McCarthy AA, Cusack S, Economou A, Petratos K.** 2007. Structure of dimeric SecA, the Escherichia coli preprotein translocase motor. J Mol Biol **366:**1545-1557.
33. **Hunt JF, Weinkauf S, Henry L, Fak JJ, McNicholas P, Oliver DB, Deisenhofer J.** 2002. Nucleotide control of interdomain interactions in the conformational reaction cycle of SecA. Science **297:**2018-2026.
34. **Zimmer J, Nam Y, Rapoport TA.** 2008. Structure of a complex of the ATPase SecA and the protein-translocation channel. Nature **455:**936-943.
35. **Vassylyev DG, Mori H, Vassylyeva MN, Tsukazaki T, Kimura Y, Tahirov TH, Ito K.** 2006. Crystal structure of the translocation ATPase SecA from Thermus thermophilus reveals a parallel, head-to-head dimer. J Mol Biol **364:**248-258.
36. **Erlandson KJ, Miller SB, Nam Y, Osborne AR, Zimmer J, Rapoport TA.** 2008. A role for the two-helix finger of the SecA ATPase in protein translocation. Nature **455:**984-987.
37. **Whitehouse S, Gold VA, Robson A, Allen WJ, Sessions RB, Collinson I.** 2012. Mobility of the SecA 2-helix-finger is not essential for polypeptide translocation via the SecYEG complex. J Cell Biol **199:**919-929.
38. **Chen Y, Bauer BW, Rapoport TA, Gumbart JC.** 2015. Conformational Changes of the Clamp of the Protein Translocation ATPase SecA. J Mol Biol **427:**2348-2359.
39. **Osborne AR, Clemons WM, Jr., Rapoport TA.** 2004. A large conformational change of the translocation ATPase SecA. Proc Natl Acad Sci U S A **101:**10937-10942.
40. **Gelis I, Bonvin AM, Keramisanou D, Koukaki M, Gouridis G, Karamanou S, Economou A, Kalodimos CG.** 2007. Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. Cell **131:**756-769.
41. **Bauer BW, Rapoport TA.** 2009. Mapping polypeptide interactions of the SecA ATPase during translocation. Proc Natl Acad Sci U S A **106:**20800-20805.
42. **Bhanu MK, Zhao P, Kendall DA.** 2013. Mapping of the SecA signal peptide binding site and dimeric interface by using the substituted cysteine accessibility method. J Bacteriol **195:**4709-4715.
43. **Dauter Z, Dauter M, Dodson E.** 2002. Jolly SAD. Acta Crystallogr D Biol Crystallogr **58:**494-506.
44. **Das S, Grady LM, Michtavy J, Zhou Y, Cohan FM, Hingorani MM, Oliver DB.** 2012. The variable subdomain of Escherichia coli SecA functions to regulate SecA ATPase activity and ADP release. J Bacteriol **194:**2205-2213.
45. **Gold VA, Whitehouse S, Robson A, Collinson I.** 2013. The dynamic action of SecA during the initiation of protein translocation. Biochem J **449:**695-705.
46. **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.
47. **Zimmer J, Rapoport TA.** 2009. Conformational flexibility and peptide interaction of the translocation ATPase SecA. J Mol Biol **394:**606-612.

48. **Vrontou E, Karamanou S, Baud C, Sianidis G, Economou A.** 2004. Global co-ordination of protein translocation by the SecA IRA1 switch. J Biol Chem **279:**22490-22497.
49. **Duong F, Wickner W.** 1999. The PrlA and PrlG phenotypes are caused by a loosened association among the translocase SecYEG subunits. EMBO J **18:**3263-3270.
50. **Oliver DB, Beckwith J.** 1982. Regulation of a membrane component required for protein secretion in Escherichia coli. Cell **30:**311-319.
51. **Ding H, Hunt JF, Mukerji I, Oliver D.** 2003. Bacillus subtilis SecA ATPase exists as an antiparallel dimer in solution. Biochemistry **42:**8729-8738.
52. **Auclair SM, Oliver DB, Mukerji I.** 2013. Defining the solution state dimer structure of Escherichia coli SecA using Forster resonance energy transfer. Biochemistry **52:**2388-2401.
53. **D'Lima NG, Teschke CM.** 2013. ADP-dependent conformational changes distinguish Mycobacterium tuberculosis SecA2 from SecA1. J Biol Chem doi:M113.533323 [pii]
10.1074/jbc.M113.533323.
54. **Zhou J, Xu Z.** 2003. Structural determinants of SecB recognition by SecA in bacterial protein translocation. Nat Struct Biol **10:**942-947.
55. **Schneewind O, Missiakas DM.** 2012. Protein secretion and surface display in Gram-positive bacteria. Philos Trans R Soc Lond B Biol Sci **367:**1123-1139.
56. **Hunt JF, Deisenhofer J.** 2003. Ping-pong cross-validation in real space: a method for increasing the phasing power of a partial model without risk of model bias. Acta Crystallogr D Biol Crystallogr **59:**214-224.
57. **de Keyzer J, van der Sluis EO, Spelbrink RE, Nijstad N, de Kruijff B, Nouwen N, van der Does C, Driessen AJ.** 2005. Covalently dimerized SecA is functional in protein translocation. J Biol Chem **280:**35255-35260.
58. **Jilaveanu LB, Zito CR, Oliver D.** 2005. Dimeric SecA is essential for protein translocation. Proc Natl Acad Sci U S A **102:**7511-7516.
59. **Gouridis G, Karamanou S, Sardis MF, Scharer MA, Capitani G, Economou A.** 2013. Quaternary dynamics of the SecA motor drive translocase catalysis. Mol Cell **52:**655-666.
60. **Prabudiansyah I, Kusters I, Driessen AJ.** 2015. In Vitro Interaction of the Housekeeping SecA1 with the Accessory SecA2 Protein of Mycobacterium tuberculosis. PLoS One **10:**e0128788.
61. **DeJesus MA, Sacchettini JC, Ioerger TR.** 2013. Reannotation of translational start sites in the genome of Mycobacterium tuberculosis. Tuberculosis (Edinb) **93:**18-25.
62. **Studier FW.** 2005. Protein production by auto-induction in high density shaking cultures. Protein Expr Purif **41:**207-234.
63. **Heras B, Martin JL.** 2005. Post-crystallization treatments for improving diffraction quality of protein crystals. Acta Crystallogr D Biol Crystallogr **61:**1173-1180.
64. **Leahy DJ, Hendrickson WA, Aukhil I, Erickson HP.** 1992. Structure of a fibronectin type III domain from tenascin phased by MAD analysis of the selenomethionyl protein. Science **258:**987-991.

65. **Fischetti RF, Xu S, Yoder DW, Becker M, Nagarajan V, Sanishvili R, Hilgart MC, Stepanov S, Makarov O, Smith JL.** 2009. Mini-beam collimator enables microcrystallography experiments on standard beamlines. J Synchrotron Radiat **16:**217-225.

66. **Otwinowski Z, Minor W.** 1997. Processing of X-ray diffraction data collected in oscillation mode. Macromolecular Crystallography, Pt A **276:**307-326.

67. **Sheldrick GM.** 2010. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta Crystallogr D Biol Crystallogr **66:**479-485.

68. **Vonrhein C, Blanc E, Roversi P, Bricogne G.** 2007. Automated structure solution with autoSHARP. Methods Mol Biol **364:**215-230.

69. **Emsley P, Lohkamp B, Scott WG, Cowtan K.** 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr **66:**486-501.

70. **Read RJ.** 1986. Improved Fourier Coefficients for Maps Using Phases from Partial Structures with Errors. Acta Crystallographica Section A **42:**140-149.

71. **Cowtan KD, Main P.** 1996. Phase combination and cross validation in iterated density-modification calculations. Acta Crystallographica Section D-Biological Crystallography **52:**43-48.

72. **Bricogne G. BE, Brandl M., Flensburg C., Keller P., Paciorek W.,, Roversi P SA, Smart O.S., Vonrhein C., Womack T.O. .** 2011. BUSTER version 2.8.0. Cambridge, United Kingdom: Global Phasing Ltd.

73. **Guo XV, Monteleone M, Klotzsche M, Kamionka A, Hillen W, Braunstein M, Ehrt S, Schnappinger D.** 2007. Silencing Mycobacterium smegmatis by using tetracycline repressors. J Bacteriol **189:**4614-4623.

827 **Table 1 Data collection and refinement statistics for SAD (Semet) structure of SecA2**
828

	SecA2
Data collection	
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	39.60, 162.09, 67.31
α, β, γ (°)	90.00 95.87 90.00
Wavelength	0.97949
Resolution (Å)	35.64-2.8 (2.85-2.8)
R_{sym} or R_{merge}	0.096 (0.171)
$I / \sigma I$	24.3 (1.4)
Completeness (%)	0.72 (0.168)
Redundancy	4.1 (1.9)
Refinement	
Resolution (Å)	2.8
No. reflections	16255
R_{work} / R_{free}	0.2910 (0.2115)
No. atoms	
Protein	4894
Water	67
B-factors	
Protein	85.23
Water	58.28
R.m.s. deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1.26

829 *Values in parentheses are for highest-resolution shell.
830

Table 2: Conservation of the two-helix finger among SecA homologs.

E. coli	SecA1	LRGYAQKDP
T. maritima	SecA1	LRSYGQKDP
M. tuberculosis	SecA1	LRAMAQRDP
M. smegmatis	SecA1	LRAMAQRDP
M. tuberculosis	**SecA2**	**LRALGRQNP**
M. avium	SecA2	LRALGRQNP
M. smegmatis	SecA2	LRALGRQNP
S. aureus	SecA2	LRSYAQQNP
L. monocytogenes	SecA2	LRAYGQIDP
S. gordonii	SecA2	LRGYAQNNP
C. difficile	SecA2	LKSYAQKDP
C. glutamicum	SecA2	LRAIARETP

848 Table 3. Residues of *Mtb* SecA2 predicted to be in contact with SecY based on structural
849 superposition with *Tm* SecA in complex with SecYEG (3din).

850

Domain	***Mtb* SecA2 residues predicted to contact SecY**
NBD1	None
NBD2	E392, R395, Q396
HSD	V600, R604, D607, A610, R614
IRA1	Most residues spanning 687-715 (including residues of the 2HF and surrounding IRA1 helices)
PPXD	N270, H272, T274, E275, D289

851

852

853 Table 4. Suppressor mutations observed in *M. smegmatis* SecA2^{K129R}.

854

isolate	residue affected in *M. smegmatis* SecA2	domain	Corresponding residues in *Mtb* SecA2
6S, 9S	deletion of residues 182-185	NBD1	168-171 (STPD)
23S*	duplication of residues 182-185	NBD1	168-171
2S	Asp326 → His	PPXD	D316
25S*	Glu insertion at residue 364	PPXD	E354
34S*	Thr 459 → Ile	NBD2	T449
21B*	deletion of residues 734-741	IRA1	714-721
38S	deletion of residues 732-739	IRA1	712-719

855 * indicates suppressors were subcloned and retested in *M. smegmatis*.

856

857

858 **Figure Legends**

859 **Figure 1:** Domain architecture of Mtb SecA2, orange=NBD1, green=NBD2,
860 cyan=PPXD, magenta=IRA1, black=HSD, yellow=C-terminal linker (CTL).

861

862 **Figure 2.** Comparison of *Mtb* SecA2 with SecA1. Relative to SecA1 (gray backbone,
863 PDB: 1nl3), SecA2 (tan backbone) is smaller, lacking the HWD (red) and the VAR
864 domain (yellow). Also, the PPXD domain has undergone rotation (SecA1:blue →
865 SecA2:cyan)

866

867 **Figure 3.** Comparison of different orientations of the PPXD domain. *Mtb* SecA2 (cyan),
868 *Mtb* SecA1 (1nl3, dark blue), *B. subtilis* SecA (1tf2, yellow), *T. maritima* SecA-SecYEG
869 complex (3din, red). All four PPXD domains are superposed onto the body of *Mtb*
870 SecA2 (orange). At one extreme, in the *Mtb* SecA1 structure (right, dark blue, the PPXD
871 is packed against the HWD (missing in SecA2), representing the signal-peptide-
872 recognition-site-closed conformation. At the other extreme (left, red), the PPXD from
873 the *T. maritima* complex with SecYEG represents the "preprotein-clamp-closed"
874 configuration, where contact is made with NBD2 (orange, lower-left). The *Mtb* SecA2
875 PPXD occupies a unique intermediate position (cyan).

876

877 **Figure 4.** Conformation of the two-helix finger (2HF). The HSD of *Mtb* SecA2 (cyan) is
878 shown superimposed on the apo structure of *Mtb* SecA1 (purple). The loops connecting
879 the two helices are shown in orange (SecA1) and green (SecA2).

880

881 **Figure 5:** Interface residues (red) of Mtb SecA2 (tan) that would contact SecYEG (cyan),
882 based on superposition with Tm SecA (3din). Note that the red residues highlighted in the
883 PPXD correspond to residues of Tm SecA that contact SecYEG, as its PPXD is rotated
884 into contact with SecYEG.
885
886 **Figure 6.** Intragenic suppressors suppress the azide sensitivity phenotype of *secA2*^{K129R}.
887 Lawns of the indicated strains were plated and tested for sensitivity to 10 µL of 0.15 M
888 sodium azide (applied to a paper disk in the center of the plate) for 2 days at 37 ºC.
889 Average inhibition was calculated by measuring the diameter of the zone of azide
890 inhibition, and values are the means of three biological replicates. The Δ*secA2* mutant *M.*
891 *smegmatis* strain was transformed with plasmids containing either secA2, *secA2*^{K129R} or a
892 reconstructed intragenic suppressor with the secA2 ^{K129R} mutation in combination with an
893 intragenic suppressor mutation located in one of the following domains: NBD1, NBD2,
894 PPXD, and IRA1.
895
896 **Figure 7.** a) Suppressor mutations from *M. smegmatis* SecA2^{K129R} mapped onto *M.*
897 *tuberculosis* SecA2. Mutations are shown as yellow spheres. b) Suppressor mutations
898 mapped onto the *T. maritima* SecA complex with SecYEG (3din). In the complex with
899 SecYEG, SecA is in the "pre-protein clamp-closed" conformation, in which PPXD (cyan)
900 is swung down (black arrow) to make contact with NBD2 (green). In this conformation,
901 the residues affected by the suppressor mutations in the PPXD and NBD2 domains can be
902 seen to come into contact (circled in red).
903

Journal of Bacteriology

904 **Figure 8.** Subcellular localization of SecA2 is altered in intragenic suppressors of
905 *secA2*^{K129R} . Whole cell lysates of the indicated strains were fractionated into a
906 cytoplasmic-containing soluble fraction and membrane-containing cell envelope fraction.
907 Representative intragenic suppressors (PPXD, NBD2, IRA1, and NBD1) are indicated.
908 Protein derived from an equal number of cells was analyzed by SDS-PAGE and
909 quantitative immunoblot analysis with anti-SecA2 antibodies was performed. Percent
910 localization in a given fraction is plotted. Error bars represent the mean of three
911 independent replicates.
912

Figure 1: Domain architecture of Mtb SecA2. orange=NBD1. green=NBD2. cyan=PPXD magenta=IRA1, black=HSD, yellow=C-terminal linker (CTL).

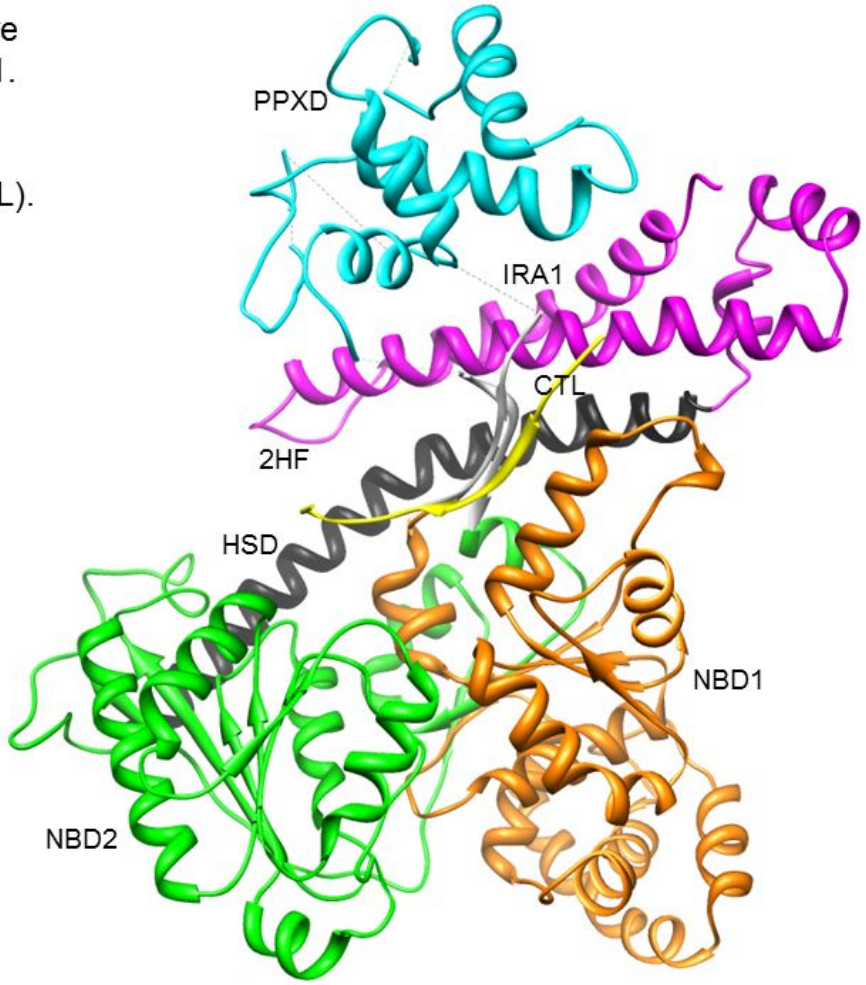
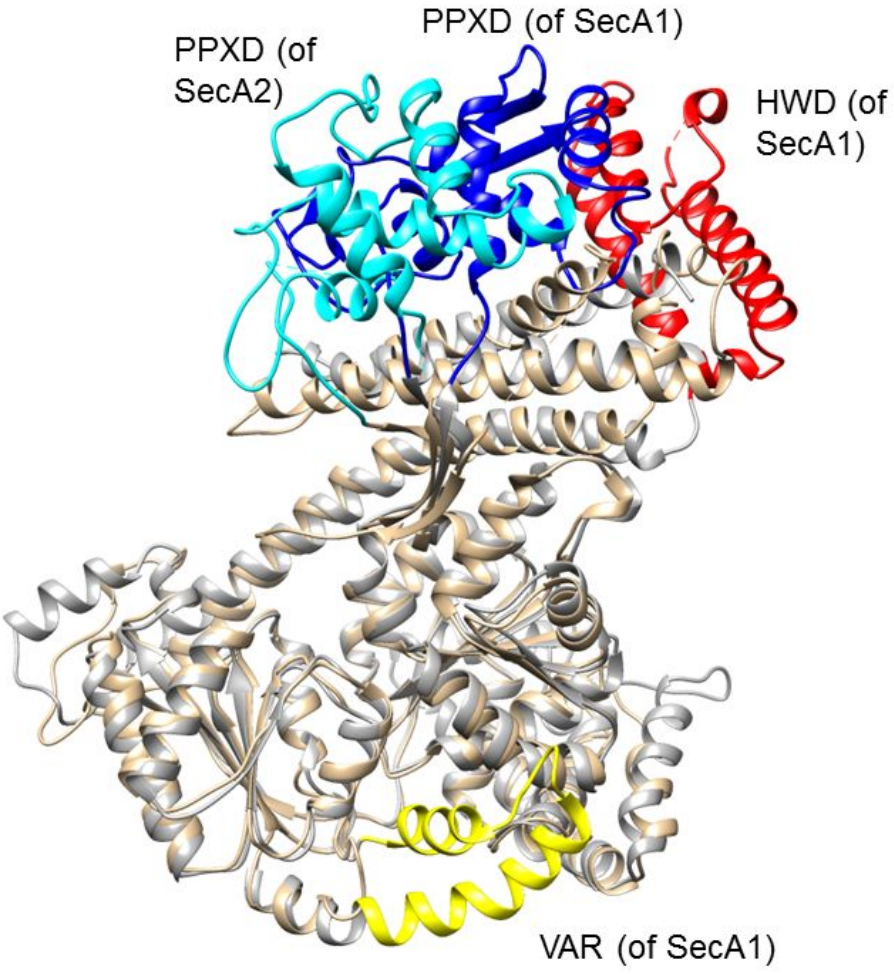


Figure 2. Comparison of *Mtb* SecA2 with SecA1. Relative to SecA1 (gray backbone, PDB: 1nl3), SecA2 (tan backbone) is smaller, lacking the HWD (red) and the VAR domain (yellow). Also, the PPXD domain has undergone rotation (SecA1:blue → SecA2:cyan)



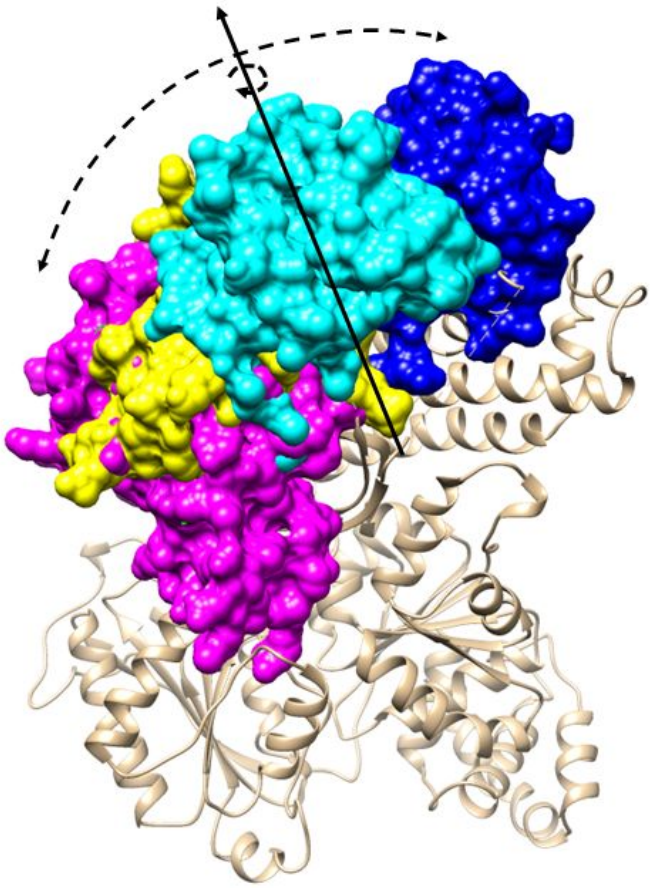
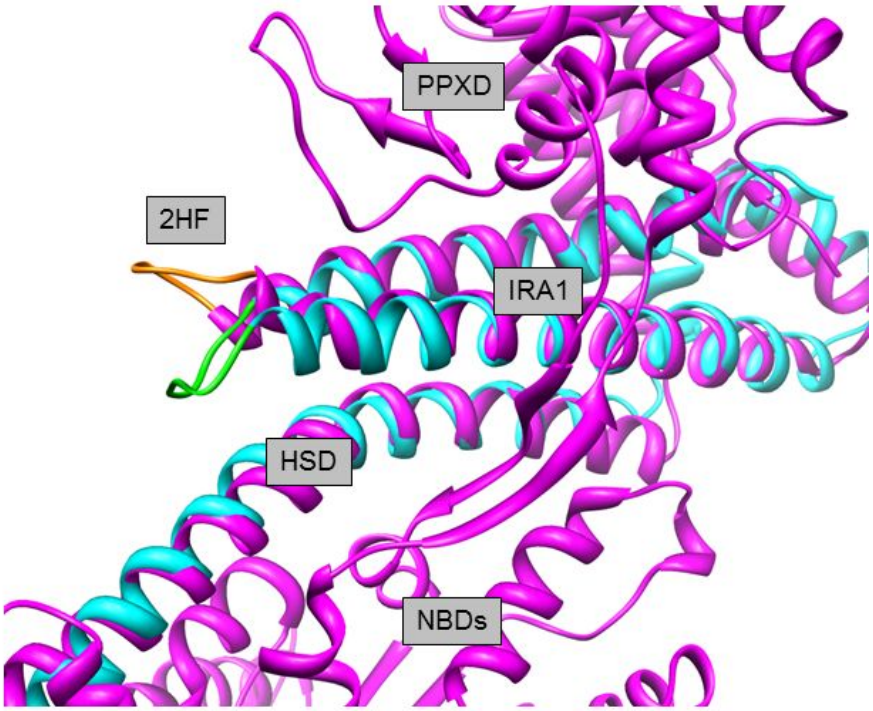


Figure 3. Comparison of different orientations of the PPXD domain. *Mtb* SecA2 (cyan), *Mtb* SecA1 (1nl3, dark blue), *B. subtilis* SecA (1tf2, yellow), *T. maritima* SecA-SecYEG complex (3din, magenta). All four PPXD domains are superposed onto the body of *Mtb* SecA2 (ribbon). At one extreme, in the *Mtb* SecA1 structure (right, dark blue), the PPXD is packed against the HWD (not shown, since it is missing in SecA2), representing the signal-peptide-recognition-site-closed conformation. At the other extreme (left, magenta), the PPXD from the *T. maritima* complex with SecYEG represents the "preprotein-clamp-closed" configuration, where contact is made with. The *Mtb* SecA2 PPXD occupies a unique intermediate position (cyan). Dashed arrows indicate rotational degrees of freedom among alternative conformations.

Figure 4. Conformation of the two-helix finger (2HF). The HSD of *Mtb* SecA2 (cyan) is shown superimposed on the apo structure of *Mtb* SecA1 (purple). The loops connecting the two helices are shown in orange (SecA1) and green (SecA2).



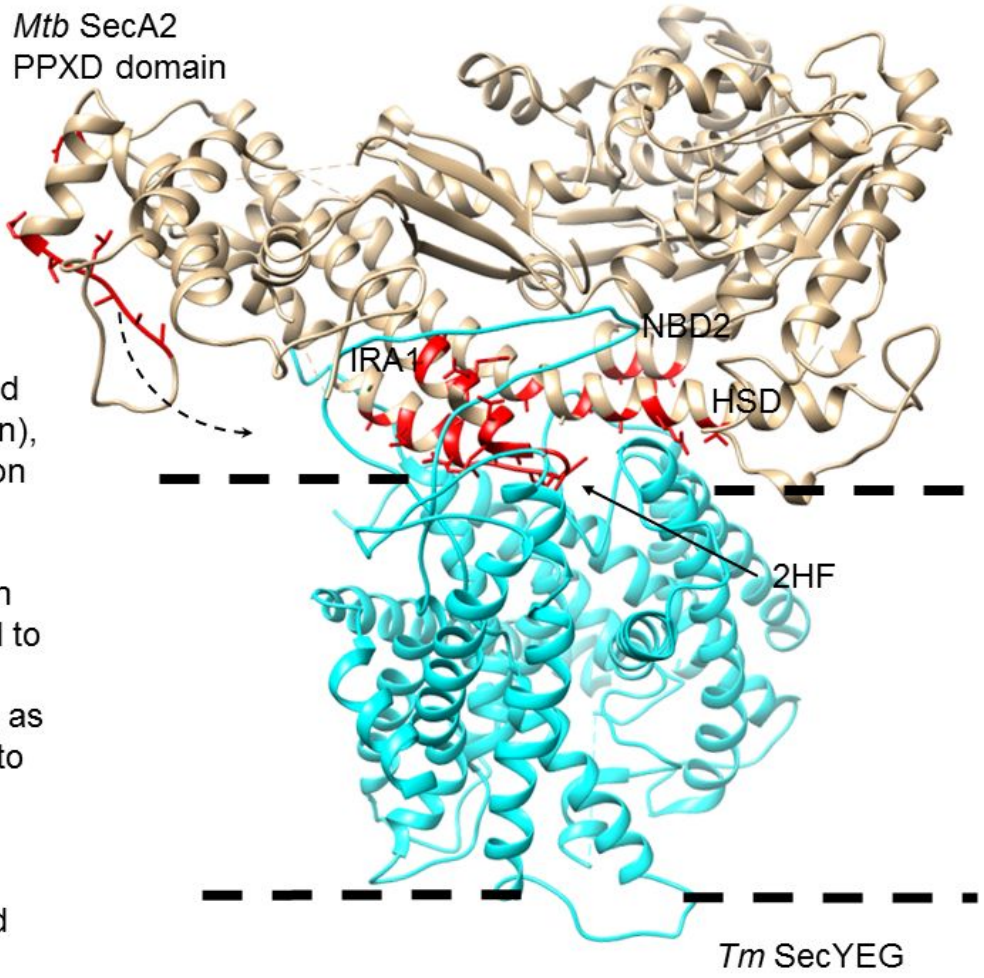


Figure 5: Interface residues (red) of *Mtb* SecA2 (tan) that would contact SecYEG (cyan), based on superposition with *Tm* SecA (3din). Note that the red residues highlighted in the PPXD correspond to residues of *Tm* SecA that contact SecYEG, as its PPXD is rotated into contact with SecYEG. Thick dashed lines indicate approximate boundaries of the lipid bilayer.

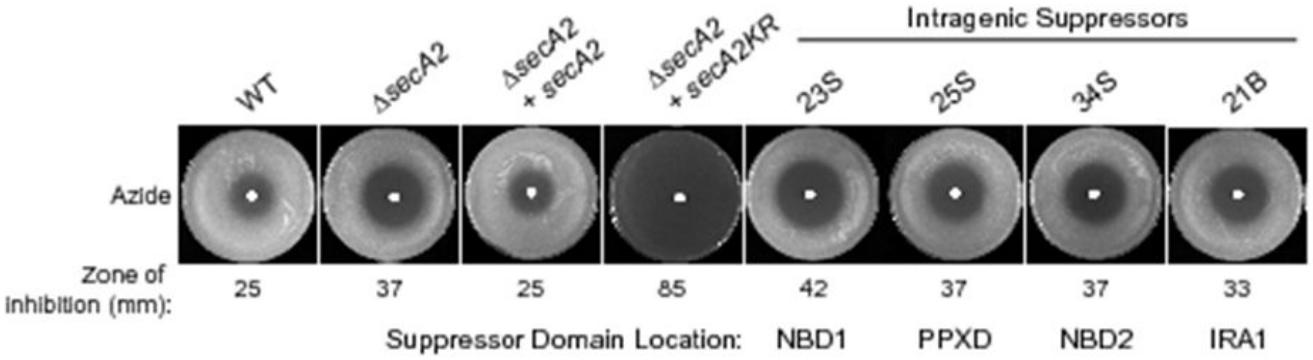


Figure 6. Intragenic suppressors suppress the azide sensitivity phenotype of *secA2*^{K129R}. Lawns of the indicated strains were plated and tested for sensitivity to 10 µL of 0.15 M sodium azide (applied to a paper disk in the center of the plate) for 2 days at 37 ºC. Average inhibition was calculated by measuring the diameter of the zone of azide inhibition, and values are the means of three biological replicates. The Δ*secA2* mutant *M. smegmatis* strain was transformed with plasmids containing either secA2, *secA2*^{K129R} or a reconstructed intragenic suppressor with the secA2 ^{K129R} mutation in combination with an intragenic suppressor mutation located in one of the following domains: NBD1, NBD2, PPXD, and IRA1.

Figure 7. a) Suppressor mutations from *M. smegmatis* SecA2^{K129R} mapped onto *M. tuberculosis* SecA2. Mutations are shown as yellow spheres. b) Suppressor mutations mapped onto the *T. maritima* SecA complex with SecYEG (3din). In the complex with SecYEG, SecA is in the "pre-protein clamp-closed" conformation, in which PPXD (cyan) is swung down (black arrow) to make contact with NBD2 (green). In this conformation, the residues affected by the suppressor mutations in the PPXD and NBD2 domains can be seen to come into contact (circled in red).

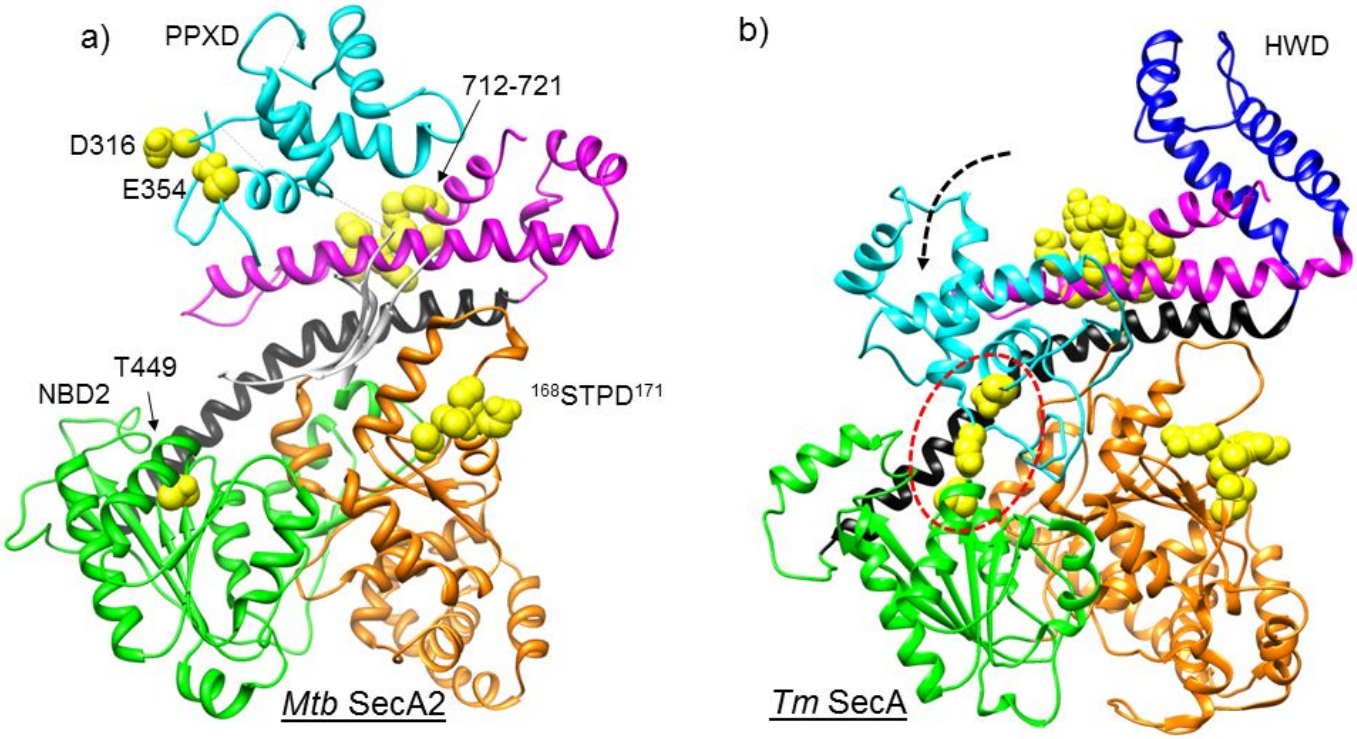


Figure 8. Subcellular localization of SecA2 is altered in intragenic suppressors of *secA2*^{K129R}. Whole cell lysates of the indicated strains were fractionated into a cytoplasmic-containing soluble fraction and membrane-containing cell envelope fraction. Representative intragenic suppressors (PPXD, NBD2, IRA1, and NBD1) are indicated. Protein derived from an equal number of cells was analyzed by SDS-PAGE and quantitative immunoblot analysis with anti-SecA2 antibodies was performed. Percent localization in a given fraction is plotted. Error bars represent the mean of three independent replicates.