Stepwise evolution of the Sec machinery in Proteobacteria

van der Sluis, EO; Driessen, AJM; Sluis, Eli O. van der

Published in:
Trends in Microbiology

DOI:
10.1016/j.tim.2006.01.009

Genome Analysis

Stepwise evolution of the Sec machinery in Proteobacteria

Eli O. van der Sluis and Arnold J.M. Driessen

Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9751 NN Haren, The Netherlands

The Sec machinery facilitates the translocation of proteins across and into biological membranes. In several of the Proteobacteria, this machinery contains accessory features that are not present in any other bacterial division. The genomic distribution of these features in the context of bacterial phylogeny suggests that the Sec machinery has evolved in discrete steps. The canonical Sec machinery was initially supplemented with SecB; subsequently, SecE was extended with two transmembrane segments and, finally, SecM was introduced. The Sec machinery of *Escherichia coli* and other Enterobacteriales represents the end product of this stepwise evolution.

The protein-conducting channel of the Sec machinery is universally conserved

The Sec machinery is an ensemble of proteins that facilitates the translocation of proteins and pre-proteins into and across biological membranes [1]. The protein-conducting channel of the Sec machinery is universally conserved but the additional components that are required for protein translocation differ between Bacteria, Archaea and Eukarya. The evolutionary implications of the similarities and differences between Sec machineries of different domains have been reviewed previously [2] but variations within domains have not, thus far, been analyzed in an evolutionary context. We have examined the genomic distribution of three components that distinguish the model organism *Escherichia coli* from the vast majority of other bacteria. In the context of established phylogenetic relationships between bacteria, the distributions suggest that the Sec machinery of Proteobacteria has evolved in a stepwise fashion.

The canonical Sec machinery of Bacteria and differences observed in *Escherichia coli*

Most Gram-positive bacteria contain the canonical Sec machinery that comprises eight different proteins. Embedded within the cytoplasmic membrane are the trimeric SecYEG complex that forms the protein-conducting channel, the membrane protein integrase YidC and the trimeric complex with unknown function SecDFYajC. The motor protein SecA cycles between the cytoplasm and the membrane, where it binds to SecYEG to provide a driving force for the translocation reaction [3].

Corresponding author: Driessen, A.J.M. (a.j.m.driessen@rug.nl).

The Sec machinery of *E. coli* is more complex: it contains one protein that is substantially different and two additional proteins. Unlike the single-spanning SecE from Gram-positives, SecE from *E. coli* consists of three transmembrane segments (TMSs). The additional proteins of the *E. coli* Sec machinery are SecB and SecM. None of these additional features is essential for viability [4–6] but all contribute in different ways to protein translocation. SecB prevents the premature folding of secretory proteins [7] and targets them to SecA, which is bound to the membrane at SecYEG [8]. SecM is a small exported protein encoded from a gene that is directly upstream of *secA* and can induce the arrest of translation elongation at the *secM–secA* mRNA in order to regulate and localize the expression of SecA [9,10]. Removal of the additional TMSs of SecE renders cells cold-sensitive for growth but no specific function has yet been assigned to this region [4].

Reconstructing the evolution of multicomponent systems

Information on the evolutionary history of multicomponent biological systems can be obtained by analyzing their composition in organisms with known phylogenetic relationships [11]. So far, however, this method has only been sparsely applied to reconstruct how eukaryotic systems have evolved from a prokaryotic ancestor [12–14]. Because the phylogeny of bacteria is becoming increasingly clear owing to the availability of numerous completely sequenced genomes, a similar approach can be followed for bacterial multicomponent systems. To reveal how the bacterial Sec machinery might have evolved, we have analyzed all the completely sequenced bacterial genomes of free-living bacteria deposited in the NCBI database (207 genomes in July 2005) for the presence of genes encoding SecB, SecM and SecE with three TMSs (termed 'SecE3' here). Genomes of obligate endosymbionts were excluded from the analysis because these are known to be extremely reduced in size.

Genomic distribution of SecB, SecE3 and SecM

Generalities

The genomic distribution of SecB, SecE3 and SecM can be summarized as follows: SecB is present in all α-, β-, and γ-proteobacteria, SecE3 is found in most β-proteobacteria and all γ-proteobacteria and SecM is found in all members of the γ-proteobacterial order Enterobacteriales (see Supplementary Material). Within the subdivision of

β-proteobacteria, three genomes contain a single-spanning SecE whereas the others encode SecE3. A phylogenetic analysis of all available SecE sequences based on the conserved C-terminal region indicates that within the β-proteobacterial subdivision, the single-spanning SecEs were the earliest to diverge (Figure 1a). These single-spanning β-proteobacterial SecE proteins might have either lost their extension or SecE might have been extended after this divergence had occurred.

Exceptions

The only two exceptions to the distribution mentioned earlier are a *secB* gene in *Desulfotalea psychrophila* and a *secE3* gene in *Bdellovibrio bacteriovorus*. These species both belong to the δ-proteobacteria. Phylogenetic analyses indicate that *D. psychrophila* SecB is most closely related to β-proteobacterial SecBs (data not shown), which suggests that it might have originated from a rare horizontal gene transfer (HGT) event. It seems unlikely that *B. bacteriovorus* SecE3 also originates from an HGT event because its C-terminal region is most closely related to that of other δ-proteobacteria and its N-terminal extension does not have any significant sequence similarity to that of β- and γ-proteobacterial SecE3s (data not shown). Hence, *B. bacteriovorus* SecE could have been extended independently.

SecM-like ORFs upstream of SecA in Pasteurellales

Inspection of the region upstream of *secA* in the genomes where *secM* was not found by BLAST searches revealed that the five sequenced members of the γ-proteobacterial order Pasteurellales encode a small open reading frame (ORF) that is remarkably similar to *secM*. Although these ORFs are substantially shorter than *secM* and the overall sequence similarity is low, the proteins they encode contain a similar atypical signal sequence and their C termini are nearly identical to the SecM motif that causes arrest of translation elongation (Figure 1b). Based on these similarities, it seems likely that these ORFs upstream of *secA* will have a similar function and, therefore, they are referred to here as SecM^{P}.

SecB paralogs in Gluconobacter oxydans *and* Francisella tularensis

An interesting phenomenon revealed by the analysis is the occurrence of two *secB* genes in the genomes of the

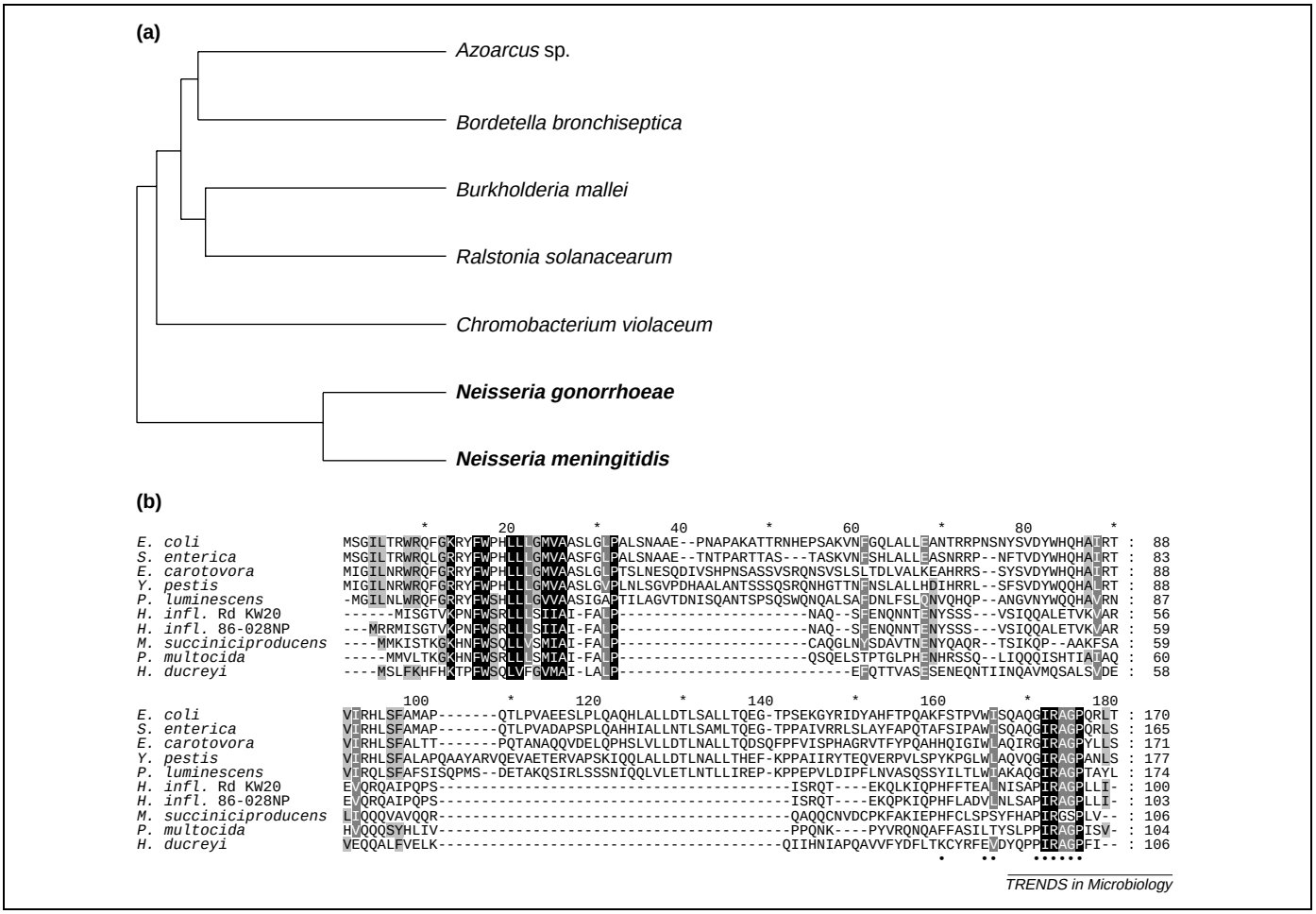

Figure 1. Sequence analysis of Sec machinery proteins. **(a)** Neighbor-joining tree of SecE based on the conserved C-terminal region that starts at the equivalent of Thr62 in *Escherichia coli*. Only unique β-proteobacterial SecE sequences are shown. Organisms that contain single-spanning SecE are in bold and those that contain SecE3 are in regular typeface. **(b)** Multiple sequence alignment of SecM sequences from Enterobacteriales and Pasteurellales. The top five sequences from Enterobacteriales are referred to as SecM (*E. coli*, *Salmonella enterica*, *Erwinia carotovora*, *Yersinia pestis* and *Photorhabdus luminescens*) and the bottom five sequences derived from Pasteurellales are referred to as SecM^{P} (*Haemophilus influenzae* Rd KW20, *H. influenzae* 86-028NP, *Mannheimia succiniciproducens*, *Pasteurella multocida* and *Haemophilus ducreyi*). Conserved residues are shaded with increasing gray intensities. Positions that are important for translation arrest in *E. coli* are indicated with a black dot.

α-proteobacterium *Gluconobacter oxydans* and the γ-proteobacterium *Francisella tularensis*. Paralogs of SecA, SecY, SecE and SecG are observed in Actinobacteria and Firmicutes and some of these are related to virulence [15,16]. Similarly, the SecB paralog of the human pathogen *F. tularensis* could be required for the secretion of pathogenicity-related proteins [17].

Stepwise evolution

To investigate whether the genomic distribution of the genes analyzed here contains information on the evolution of the Sec machinery, they were projected onto the relevant part of the bacterial phylogenetic tree. Because the details of branching patterns obtained by different tree-reconstruction methods are often dissimilar [18], only a simplified tree is shown at the lowest resolution required for interpretation of the results (Figure 2). This projection reveals that SecB was introduced in the last common ancestor (LCA) of the α-, β-, and γ-proteobacteria, SecE was extended in the LCA of the β- and γ-proteobacteria (see earlier) and SecM was introduced in the LCA of the γ-proteobacterial order Enterobacteriales. This implies that the canonical Sec machinery with a minimal *B. subtilis*-like composition has been successively modified in the following steps: first, the chaperone SecB was introduced, second, SecE was extended with two transmembrane segments and third, the regulatory protein SecM was introduced, resulting in the well studied *E. coli* Sec machinery.

In addition, the analysis provides an alternative view on the $SecM^P$ mentioned earlier. Rather than being a functional equivalent of SecM in the Pasteurellales, $SecM^P$ could represent an ancestral protein with a slightly different function, from which SecM is derived. Considering the presence of the translation-arrest motif, $SecM^P$ is expected to stall translation elongation in a similar manner to SecM but the exact regulatory mechanism could differ as a result of the shorter N-terminal region. Thus, the evolution of the Sec machinery might be most accurately described by a four-step process that includes the extension of $SecM^P$ to SecM.

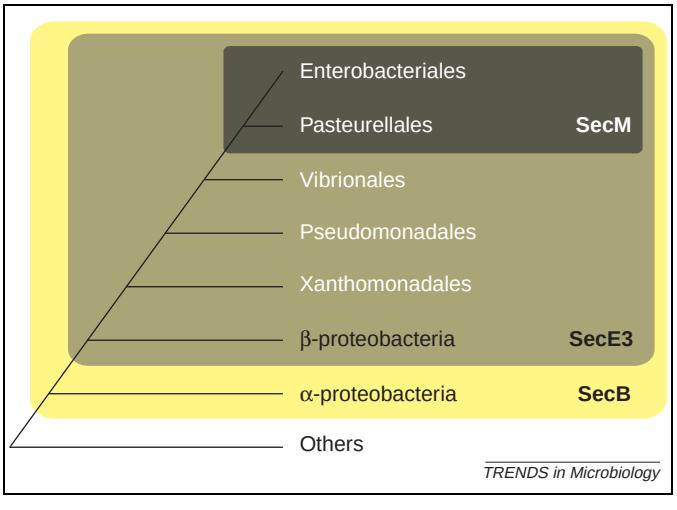

Figure 2. Genomic distribution of accessory features of the Sec machinery combined with bacterial phylogeny. The simplified phylogenetic tree is derived from Refs [21,22]. Only the relevant taxa are shown. The names of γ-proteobacterial orders are in white and the presence of SecB, SecE3 and SecM is indicated by the colored boxes.

Biological significance

The stepwise evolutionary process that is suggested by these results could be interpreted as a random series of independent improvements of the Sec machinery. However, such an interpretation does not take into account that each accessory feature was introduced in a different context or that the accessory features might be functionally interdependent. We hypothesize that SecE3 and SecM have specifically improved SecB-dependent protein translocation rather than protein translocation *per se*. By regulating and localizing the expression of SecA [9,10], SecM maximizes the amount of SecYEG-bound SecA that forms the receptor for preprotein-SecB complexes [8]. SecE3 might similarly enhance SecB-dependent targeting by increasing the affinity of SecA for SecYEG. Although this has not been demonstrated experimentally, SecA from *E. coli* is known to bind to SecYEG with a much higher affinity than that of *B. subtilis* SecA [19,20]. It is likely that the additional TMSs of *E. coli* SecE are at least partially responsible for the higher affinity. Moreover, a sufficiently high affinity of SecA for SecYEG is a prerequisite for achieving the regulatory effects of SecM. In other words, the accessory features of the Sec machinery might contribute to protein translocation synergistically rather than independently.

Concluding remarks and future perspectives

The genome analysis presented here suggests that the proteobacterial Sec machinery has evolved in discrete steps. Each step seems to have improved the efficiency of protein translocation in a different manner. Thus, this study reveals in detail how the Sec machinery might have undergone successive improvements by acquisition or modification of individual components. The possible synergism between SecB, SecE3 and SecM can now be investigated experimentally. In addition, it should be stressed that the phylogenetic position of *E. coli* provides an excellent opportunity to reveal how other bacterial model systems might have evolved.

Acknowledgements

This study was supported by the Council for Chemical Sciences (CW), which is subsidized by the Dutch Organization for the Advancement of Scientific Research (NWO). The authors thank Juke Lolkema for valuable discussions.

Supplementary data

Supplementary data associated with this article can be found at doi:10.1016/j.tim.2006.01.009

References

1 Osborne, A.R. *et al*. (2005) Protein translocation by the Sec61/SecY channel. *Annu. Rev. Cell Dev. Biol.* 21, 529–550
2 Pohlschroder, M. *et al*. (1997) Protein translocation in the three domains of life: variations on a theme. *Cell* 91, 563–566
3 de Keyzer, J. *et al*. (2003) The bacterial translocase: a dynamic protein channel complex. *Cell. Mol. Life Sci.* 60, 2034–2052

4 Schatz, P.J. *et al*. (1991) One of three transmembrane stretches is sufficient for the functioning of the SecE protein, a membrane component of the *E. coli* secretion machinery. *EMBO J.* 10, 1749–1757
5 Kumamoto, C.A. and Beckwith, J. (1985) Evidence for specificity at an early step in protein export in *Escherichia coli*. *J. Bacteriol.* 163, 267–274
6 Rajapandi, T. *et al*. (1991) The first gene in the *Escherichia coli secA* operon, gene X, encodes a nonessential secretory protein. *J. Bacteriol.* 173, 7092–7097
7 Collier, D.N. *et al*. (1988) The antifolding activity of SecB promotes the export of the *E. coli* maltose-binding protein. *Cell* 53, 273–283
8 Hartl, F.U. *et al*. (1990) The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the *E. coli* plasma membrane. *Cell* 63, 269–279
9 Nakatogawa, H. and Ito, K. (2001) Secretion monitor, SecM, undergoes self-translation arrest in the cytosol. *Mol. Cell* 7, 185–192
10 Nakatogawa, H. *et al*. (2005) SecM facilitates translocase function of SecA by localizing its biosynthesis. *Genes Dev.* 19, 436–444
11 Huynen, M.A. *et al*. (2005) Variation and evolution of biomolecular systems: searching for functional relevance. *FEBS Lett.* 579, 1839–1845
12 Friedrich, T. and Weiss, H. (1997) Modular evolution of the respiratory NADH:ubiquinone oxidoreductase and the origin of its modules. *J. Theor. Biol.* 187, 529–540
13 Mathiesen, C. and Hagerhall, C. (2003) The 'antiporter module' of respiratory chain complex I includes the MrpC/NuoK subunit – a revision of the modular evolution scheme. *FEBS Lett.* 549, 7–13
14 Gabaldon, T. *et al*. (2005) Tracing the evolution of a large protein complex in the eukaryotes, NADH:ubiquinone oxidoreductase (complex I). *J. Mol. Biol.* 348, 857–870
15 Lenz, L.L. *et al*. (2003) SecA2-dependent secretion of autolytic enzymes promotes *Listeria monocytogenes* pathogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 100, 12432–12437
16 Braunstein, M. *et al*. (2003) SecA2 functions in the secretion of superoxide dismutase A and in the virulence of *Mycobacterium tuberculosis*. *Mol. Microbiol.* 48, 453–464
17 Larsson, P. *et al*. (2005) The complete genome sequence of *Francisella tularensis*, the causative agent of tularemia. *Nat. Genet.* 37, 153–159
18 Delsuc, F. *et al*. (2005) Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* 6, 361–375
19 Swaving, J. *et al*. (1999) Preprotein translocation by a hybrid translocase composed of *Escherichia coli* and *Bacillus subtilis* subunits. *J. Bacteriol.* 181, 7021–7027
20 van der Wolk, J. *et al*. (1993) Characterization of a *Bacillus subtilis* SecA mutant protein deficient in translocation ATPase and release from the membrane. *Mol. Microbiol.* 8, 31–42
21 Kunin, V. *et al*. (2005) The net of life: reconstructing the microbial phylogenetic network. *Genome Res.* 15, 954–959
22 Lerat, E. *et al*. (2003) From gene trees to organismal phylogeny in prokaryotes: the case of the γ-Proteobacteria. *PLoS Biol.* 1, E19

0966-842X/$ - see front matter
doi:10.1016/j.tim.2006.01.009