Subsystem: General secretory pathway (sec-SRP) complex (TC 3.A.5.1.1)

Matthew Cohoon, Department of Computer Science, University of Chicago, Chicago, IL

The list of subsystem functional roles

ftsX	Cell division protein ftsX
ftsE	Cell division transporter, ATP-binding protein ftsE (TC 3.A.5.1.1)
ftsY	Cell division protein ftsY (TC 3.A.5.1.1)
secY	Preprotein translocase secY subunit (TC 3.A.5.1.1)
ffh	Signal recognition particle, subunit Ffh SRP54 (TC 3.A.5.1.1)
secA	Protein export cytoplasm protein SecA ATPase RNA helicase (TC 3.A.5.1.1)
secE	Preprotein translocase subunit SecE (TC 3.A.5.1.1)
secD	Protein-export membrane protein secD (TC 3.A.5.1.1)
secF	Protein-export membrane protein secF (TC 3.A.5.1.1)
YajC	Preprotein translocase subunit YajC (TC 3.A.5.1.1)
secG	Preprotein translocase subunit SecG (TC 3.A.5.1.1)
RNPA	Ribonuclease P protein component (EC 3.1.26.5)

Introduction

Approximately 20% of bacterial proteins are located partially or completely outside of the cytoplasm. Most reach their destination via the general secretory pathway (GSP). The first step of GSP is insertion and translocation of a preprotein-chaperone complex across the cytoplasmic membrane. Fully translocated proteins are released into externaml milieu (in Gram(+)s) or periplasm (in Gram(-)s). Some are transported across outer membrane by one of several terminal branches of the GSP used by gram-negative bacteria.

Subsystem analysis

The Sec translocase, catalizing the first step, is comprised of the protein-conducting SecYEG channel and the heterotrimeric SecDFYajC complex that facilitates translocation. Interestingly, the Sec complex is also associated with YidC. For translocation of large domains of membrane proteins, SecA is also needed. Typically, membrane proteins are targeted to the Sec translocase by the SRP pathway.

In addition to functional roles listed above, two more roles have included in this subsystem: FtsX and SecB. The *ftsX* homologs often cluster on the chromosome with those of *ftsE* to form an ABC transporter. SecB is directly involved with Sec-SRP-mediated transport of proteins as it acts as a chaperone, maintaining the proteins targeted for secretion in unfolded states.

Subsystem spreadsheet (a fragment)

Organism	ftsX	ftsE	ftsY	secY	ffh	secA	secE	secD	secF	YajC	secG	RNPA	secB
Escherichia coli K12 [B]	3395	3396	3397	3234	2577	99	3893	405	406	404	3118	3640	3545
Photorhabdus luminescens subsp. laumondii TTO1 [B]	3917	3918	3919	4491	1191	3477	410	3726	3725	3727	4319	4682	4616
Xylella fastidiosa Temecula1 [B]	1969	1970	840	430	50	1783	1930	169	170	168	227	2035	1017
Yersinia pestis KIM [B]	408	407	406	3954	877	3565	468	975	976	974	674	4057	72
Haemophilus influenzae R2866 [B]	221	220	219	249	1329	368	103	1436	1435	1437	1637	455	196
Mycobacterium tuberculosis CDC1551 [B]	3300	3301	3093	767	3087	19193457	669	2754	2753	2755	1526	4185	
Salmonella enterica serovar Choleraesuis SC-B67 [B]	3498	3499	3500	3354	2678	133	4031	449	450	448	3232	3758	3624

Notes on subsystem encoding in SEED

- First, uniform assignments are generated for all PEGs (=**p**rotein **e**ncoding **g**enes) in SEED potentially encoding a specific functional role (e.g. SecY)
- PEGs with annotations matching a functional role in a subsystem are automatically connected to subsystem spreadsheet
- Analysis of subsystem spreadsheet efficiently identifies all the genomes lacking properly annotated SecY orthologs
- Specialized tools in SEED allow user to limit the spreadsheet display to a specific phylogenetic group, or to sort organisms “by phylogeny” facilitating comparisons between related organisms
- The “show clusters” option can be activated to highlight ORFs located on the chromosome in close proximity of each other. Such clustering on the chromosome is one of the many strategies used in SEED to identify gene candidates for a “missing” functional role
- “Show missing with matches” tool can be used to automatically identify gene candidates for all functional roles in a subsystem based on sequence similarity. This computationally intensive search can also be limited to a single genome or a single functional role.
- “Annotate via translation” tool can be used to annotate groups of ORFs as a batch
- Using the SEED tools and search engine the Transporter Classification (TC) numbers were applied with the same consistency and regularity that is now prevalent with EC numbers