RNA Biology Special Focus: Type I toxin-antitoxin systems
Review

Divergently overlapping *cis*-encoded antisense RNA regulating toxin-antitoxin systems from *E. coli*: *hok*/*sok*, *ldr*/*rdl*, *symE*/*symR*

Mitsuoki Kawano*
Department of Applied Life Sciences, Niigata University of Pharmacy and Applied Life Sciences, Niigata, Japan
*Corresponding to: Mitsuoki Kawano; Department of Applied Life Sciences, Niigata University of Pharmacy and Applied Life Sciences, 265-1, Higashijima Akiha-ku, Niigata-shi, Niigata, 956-8603, Japan
Tel.: +81.250.25.5124; FAX: +81.250.25.5021
Email: mkawano@nupals.ac.jp

Key words: small RNA, noncoding RNA, small protein, post-transcriptional regulation, SOS response repetitive sequence

Abstract
Toxin-antitoxin (TA) systems are classified according to how the antitoxin prevents the toxin. In type I TA systems, small antisense RNAs (RNA antitoxins) inhibits translation of small toxic proteins by binding to the corresponding mRNAs. Those type I TA systems were originally identified as plasmid stabilization modules rendering a post-segregational killing (PSK) effect on the host cells. The type I TA loci also exist on the *Escherichia coli* chromosome but their biological functions are less clear. Genetic organization of regulatory elements and the toxin genes from *hok*/*sok* and *ldr*/*rdl* families are closely related structurally, in which the toxins are predicted to contain a transmembrane domain, but share no detectable sequence identity. This review will give an overview of *E. coli* type I TA modules, especially *hok*/*sok*, *ldr*/*rdl*, and SOS-inducible *symE*/*symR* systems which are regulated by divergently overlapping *cis*-encoded antisense RNAs, and discuss their functions and benefit to the cell.

Introduction
In the last ten years, small regulatory RNAs (sRNAs) have become the focus of a broad range of studies due to their potential involvement in various levels of biological systems.[Mattick 2005, Storz et al 2005] These sRNAs play an important role for regulating gene expression via a base-pairing mechanism with target mRNAs.[2-50, Beisel and Storz, 2010] In *E. coli*, most of the sRNAs bind to an RNA chaperon, Hfq that stimulates duplex formation by two complementary RNAs and stabilizes sRNAs.[Aiba 2007] Some sRNAs, like microRNAs discovered in eukaryotes and viruses, have been characterized to control gene expression via *trans*-acting means. Such sRNAs are partially complementary to their target RNAs; therefore, they often have multiple targets.[Gottesman and Storz, 2010] In contrast to the *trans*-encoded sRNAs encoded on the bacterial

chromosome, *cis*-encoded sRNAs are encoded on mainly plasmids, phages, and transposons. They are encoded in the same DNA locus and are therefore completely complementary to their targets.[Wagner et al 2002, Brantl 2007] Originally the *cis*-encoded sRNAs were identified as a regulatory RNA to control the initiation of plasmid replication and its copy numbers by changing the formation of a targeting mRNA secondary structure.[Tomizawa et al 1981] These RNAs also act on transcriptional attenuation, inhibition of translation, promotion of RNA degradation. One of the RNA encoded on plasmids, Sok antitoxin RNAs, repress the synthesis of small, hydrophobic proteins (Hok) that cause damage in bacterial cell membranes.[1-5] Because the protein toxin leads to cell death in which the plasmid is lost, these toxin-antitoxin (TA) modules have been termed post-segregational-killing (PSK) systems[1-12] or addiction modules.[3-3]

In addition to the *hok*/*sok* systems for plasmid addiction, a number of chromosomally encoded TA modules have been recently identified and characterized.[1-27, 1-32, 1-2, 2-23, 2-30] There are three types of TA systems, classified as type I, II, and III, which are distinguished by the nature of the antitoxins and the composition of TA gene systems.[Yamaguchi et al 2010, Hayes and Melderen 2011] This review will focus on the type I TA system, whose toxin translation is inhibited by antisense sRNA antitoxin (antitoxin of type II TA system is a small, unstable protein; type III antitoxin is RNA that inhibits toxin activity), especially *hok*/*sok*, *ldr*/*rdl*, and *symE*/*symR* systems in *E. coli* with an emphasis on their common aspects and merit to the host cells, and point out future issues in the field.

Plasmid-encoded type I TA modules

The first TA systems were identified on plasmids. The *hok*/*sok* system of the *E. coli* plasmid R1 was originally discovered in a screen for a locus that mediates efficient plasmid stabilization by killing plasmid-free cells,[2-18, 1-12] and perhaps it is one of the most characterized system at a molecular regulatory level among all TA systems so far.[1-5] *hok*/*sok* homologues are also found in other low-copy plasmids. The F plasmid carries two *hok*/*sok* homologus loci ([1-26] *fim*[2-34, 5-33] and *srnB*[2-37, 5-67]). There is one *hok*/*sok* homologus locus (*pndA/pndB)* on plasmids R483.[2-38, 2-36] The locus codes for three small genes that denoted *hok* (host killing), *sok* (suppression of killing), and *mok* (modulation of killing) and all of the genes act on in a coordinated manner. The *hok* gene encodes a highly toxic inner membrane-associated protein of 52 amino acids (aa) that irreversibly damages the cell membrane, and is thus lethal to host cells. The *mok* reading frame overlaps extensively with *hok*, and it is required for expression and regulation of *hok* translation. The *sok* gene specifies a small antisense RNA *in-cis* of 64 nucleotides (nt) which is complementary to the *hok* mRNA leader region. Sok RNA is unstable (half-life of ~ 30 seconds), but is constitutively expressed from a relatively strong promoter. In contrast, *hok* mRNA is very stable (half-life of ~ 20 minutes) and is constitutively expressed from a relatively weak promoter. Genetic analyses showed that Sok RNA inhibits translation of the *mok* reading frame by blocking and that translation of *hok* is coupled to the translation of *mok*. Consequently, Sok RNA indirectly inhibits translation of *hok* by preventing *mok* translation. Because Sok RNA is very unstable and is quickly degraded when the R1 plasmid is lost from the cell, the more stable *hok* mRNA is translated and cells increase plasmid maintenance by reducing the plasmid-free cells are killed.

This plasmid stabilization system is as a result of a phenotype called post-segregational killing (PSK).

Through the painstaking molecular studies by the Gerdes group, the detailed mechanisms leading to activation of *hok* translation in plasmid-free cells are well understood. The full-length *hok* mRNA which is translationally inactive and prevents binding of Sok RNA, is accumulated in plasmid-carrying cells. This primary mRNA is then activated by 3′ end processing by the RNase II (*rnb* gene product) and polyribonucleotide nucleotidyltransferase (PNPase: *pnp* gene product), leading to generation of the truncated mRNA. After refolding, this mRNA contains the Sok RNA target hairpin structure and therefore is active with respect to both translation and Sok RNA binding. In plasmid-carrying cells, however, the Sok RNA rapidly binds to the truncated *hok* mRNA and thereby inhibits its translation by preventing ribosome entry. This *hok* mRNA-Sok RNA duplex is cleaved by RNase III, thus truncated *hok* mRNA is not translated. In plasmid-free cells, the Sok RNA is rapidly degraded due to its instability. Then, the continued 3′ processing of full-length *hok* mRNA leads to accumulation of the stable truncated mRNA, resulting Hok protein synthesis and selective killing of the plasmid-free cells.

Chromosome-encoded type I TA modules

hok/*sok* gene family.

There are multiple classes of *hok*/*sok* toxin-antitoxin modules which are encoded not only on plasmids but on bacterial chromosomes as well. The chromosome of *E. coli* K-12 contains five *hok*/*sok* homologous loci (Fig. 1). All of the *hok*/*sok* genes, with the exception of the *hokB*/*sokB* locus, clearly appear to have degenerated with mutations and transposon insertions. Three *hok*/*sok* loci (*hokA*/*sokA, hokC/sokC, and hokE*) are inactivated by insertion sequence (IS) elements. It is revealed that about half of the 72 wild type *E. coli* strains of the ECOR collection encode a *hokA*/*sokA* and *hokC/sokC* systems without an IS element. The *hokE* locus does not encode the *sok* antisense RNA gene. The *hokD* gene, formerly known as *relF*, is encoded by the third gene of the *relBEF* operon. This gene does not contain upstream regulatory elements and lacks the *sok* gene. Thus, the *hokD* locus probably constitutes an evolutionary relic of an ancient *hok*-homologue. The *hokB*/*sokB* locus seems to contain all the regulatory elements as previously described for the *hok* system of plasmid R1 (Fig. 2). However, all the *hok*/*sok* homologues on the chromosome have lost PSK activity, probably through inactivation by IS elements, point mutations and a genetic rearrangement. The chromosome-encoded mRNA is insufficiently translated *in vitro*, further explaining the absence of the PSK phenotype. Therefore, biological function of the chromosomally encoded *hok*/*sok* loci is not elucidated yet. It is possible that induction of the chromosomal *hok* gene may be activated by as-yet unknown signals and is required in natural conditions rather than laboratory environments.

ldr/*rdl* gene family.

The *E. coli* K-12 chromosome also contains a family of multiple TA modules. Four copies of long direct repeat (LDR) sequences were detected upon completion of the *E. coli* genomic sequence. Three of the repeats (LDR-A, -B, -C), each approximately 500 bp in length, are located as tandem repeats at 27.4 min on the genetic map. Another single fourth copy (LDR-D), 450 bp in length and nearly identical to LDR-A, -B and -C,

is located at 79.7 min. It is interesting that the LDR sequences are symmetrically positioned on the chromosome (Fig. 1), but the reasons for this are unknown. In previous studies, another four statistically significant LDR sequences were identified with more than 187 bp matched to LDR-A near the LDR loci. However, these are probably remnant sequences due to their loss of sense and antisense genes. There are several similarities between LDRs (*ldr*/*rdl* modules) and *hok*/*sok* modules on the chromosome of *E. coli*: (i) encoding of a small toxic protein (LdrD: 35 aa; HokB: 49 aa) whose overexpression leads to rapid host cell killing; (ii) expression of a highly stable *ldr* mRNA repressed by an unstable small antisense Rdl RNA; (iii) formation of a stable mRNA secondary structure (*ldrD* mRNA: 374 nt in length, $\Delta G = -177.24$ kcal mol^{-1}; *hok* mRNA: 433 nt in length, $\Delta G = -197.02$ kcal mol^{-1}); (iv) presence of a putative *ldrD*-overlapping *mok*-like second open reading frame named as *ldrX* (the ribosome binding site and the initiation codon of *ldrD* have no capacity to bind to the RdlD RNA but it can base-pair with those regions of *ldrX*); and (v) presence on the chromosome as multiple copies in many enteric bacteria. Despite of their highly similar genetic organization and regulatory mode, there are no homologies between those two systems within the amino acid or DNA sequences and functionally the *ldr*/*rdl* module has no PSK activity. LDR-homologous sequences have not yet been reported in known plasmid sequences. These results suggest that the LDRs are genetic elements that are not used for stabilizing plasmid inheritance and probably do not show PSK-like activity in LDR-locus deleted cells.

***symE*/*symR* gene family.**

In a cloning-based screen for *E. coli* sRNAs, several *cis*-encoded antisense RNAs were identified. One such 77 nucleotide RNA was denoted *symR* (sym biotic RNA) based on its genomic position opposite to the 5′ end of the *symE* (SOS-induced *yjiW* gene with similarity to *Maz*E) mRNA (Fig. 1). The SymR promoter is embedded in the *symE* coding sequence and the SymR antisense RNA is transcribed from three nucleotides behind the start codon of *symE* (Fig. 2). The SymE synthesis is tightly repressed at multiple levels; by the LexA repressor at the level of transcription, by the SymR RNA at the level of mRNA stability and translation, and by the Lon protease at the level of protein stability (Fig. 3). This multilayer control system is probably an effective way to keep endogenous SymE toxin level low enough not to damage the cell until it is required. In contrast to the previously mentioned antitoxin RNAs and proteins which are rapidly degraded by ribonuclease and protease activity respectively, SymR antitoxin RNA is quite stable and surprisingly the SymE toxin is degraded by the Lon protease.

After SOS response induced by the DNA damaging agent mitomycin C, SymE synthesis is very slow; detection occurred 30 min after induction with a peak at 90 min. One interpretation of this is that the timing of SymE activity is dependent on the initial activity of DNA repair proteins (Fig. 4). This late expression of SymE after the SOS response could be controlled by SymR, the Lon protease, and other potential unknown factors. Without SOS-induction, i.e. in normally growing cells, SymE levels were slightly elevated when one of the symR or lon genes is disrupted, similar to levels observed in a *lexA* repressor mutant strain. However, SymE levels were significantly elevated in the double mutant and were induced even further upon deactivation of the LexA repressor with mitomycin C treatment. This indicates that the contributions of

LexA, SymR, and Lon to SymE repression are additive. RNase III is not required for repression. In addition, Hfq RNA chaperon protein is not necessary for SymR regulation of SymE synthesis although most *trans*-encoded antisense RNAs in *E. coli* require the regulatory activity of binding between sRNA and the target mRNAs.

Characteristics of the small toxic proteins

The Hok and Ldr toxin protein family members have some common features. They are very hydrophobic and contain one α–helical transmembrane domain and positively charged amino acids flanked by the domain (Fig. 5). All the toxins in the type I TA system of *E. coli* characterized so far have a transmembrane domain and are predicted to be localized in inner membrane and/or interact with other proteins in the cell membrane or periplasm. It is proposed that an oligomeric form of the toxins similar to phage holin proteins creates pore-like structure and permeabilize the membrane to impair ATP synthesis; consequently replication, transcription, and translation may be inhibited.

Overexpression of Hok protein from a multi-copy plasmid has been shown to lead to loss of the cell membrane potential, arrest of respiration, efflux/influx of small molecules, and change morphology to so-called “ghost-cells” which inhibit cell growth and reduce colony-forming ability. Therefore, Hok proteins are likely to kill the cells by mediating irreversible damage to the host cell membrane.

Ectopic expression of Ldr protein causes rapid growth inhibition, loss of cell viability, inhibition of global translation, and nucleoid condensation. The condensation of nucleoid structure is a quick reaction that is observed within 2 min of induction of *ldrD*. This speed, together with the failure to detect the *ldrD* gene product, makes it seem unlikely that this nucleoid condensation is caused by the accumulation of LdrD on the chromosome. The LdrD might interact with an unknown target that is important for maintaining normal nucleoid structure and cell growth. The physiological function of LdrD related to this phenotype is at present unknown; however, microarray analysis suggests that overexpression of LdrD leads to physiological alteration in purine metabolism and decreases cAMP levels in the cell. Thus, the identification of the specific molecular target(s) of the small protein will be an important next subject of investigation.

SymE overexpression also affects cell growth and global protein synthesis. Although SymE belongs to type I TA families, it is not a hydrophobic protein and does not show functional homology to other type I toxin proteins. SymE actually promotes RNA degradation of mRNAs and noncoding RNAs but not SymR RNA. This resembles the function of type II toxins such as MazF which can cleave mRNA independent of the ribosome (Fig. 6). However, SymE has homology to the AbrB-fold superfamily proteins such as MazE which act as transcriptional factors and antitoxins in various type II TA modules. Analyses of amino acid conservation and operon organization of the SymE-like gene family imply that SymE has evolved into an RNA cleavage protein with toxin-like properties from a transcription factor or antitoxin. It is interesting to note that many SymE family genes locate adjacent to an antitoxin-like gene encoding protein antitoxin. In case *symE* homologues solely exist on the genome, SymR-like antisense RNA might antagonize SymE toxicity; this inferred from an observation that nucleotide sequences around initiation codon of SymE-like gene are relatively conserved among the families.

The biological roles of the Hok, Ldr and SymE proteins expressed from the chromosome are unclear. None of the chromosome-encoded *hok*/*sok* and *ldr*/*rdl* systems has the PSK activity, and antitoxin-deficient strains do not show growth inhibition and killing effect to the cells. Thus, it has been proposed that they are beneficial to cell survival by being part of the global cellular response to environmental stress such as amino acid and/or carbon source starvation, rather than being cell-killing modules (not toxin actually). In the case of SymE, which is induced in response to DNA damage or other factors causing SOS response, it has been suggested that RNA cleavage property by SymE may be important for the recycling of RNAs damaged under the conditions. Additional experiments are required to reveal what conditions determine activity and to clarify functional roles in the cell.

Diversity, evolution, and merit of the type I TA Family

In *E. coli* K-12, multiple and polymorphic TA modules are present on the chromosome, encoding five *hok*/*sok* and four *ldr*/*rdl* (LDR) loci but one copy of *symE*/*symR* module. Those homologues of the type I toxin are also found in other enteric bacteria. The *hok* genes are found in a broader spectrum of enterobacteria and *ldr* and *symE* genes are conserved in only closely related enteric bacteria. *Hok* gene loci are repeated multiple times (e.g. *E. coli* O157 H7 EDL933 has thirteen copies) and are almost randomly scattered on the chromosome. In contrast, multiple *ldr* genes are located as a tandem repeat sequence. In contrast to type II TA system, type I systems seem unlikely to move via horizontal gene transfer but rather have evolved by lineage-specific gene duplication. It is also interesting to compare corresponding genetic loci of the toxin genes between bacteria species to speculate on their evolution. For example, one of the *ldr* genes in *Salmonella* exists between *tsx* and *yajD* genes at 9.1 min on the genetic map. In contrast, a REP-like sequence is present between *tsx* and *yajD* genes in *E. coli*. This may imply that REP-like sequences could be used as a target sequence for LDR-insertion or it exists as a remnant sequence after excision of an LDR sequence. Interestingly, *symE* genes tend to associate with mobile or selfish elements such as transposons, restriction-modification (RM) modules, and pathogenicity islands. The *symE*/*symR* module in *E. coli* K-12 is located between *mcrB* and *hsdS* RM-related genes. It is also found that other TA module, a *relBE* system in the gram positive bacteria *Streptococcus mutants*, is associated with RM system gene *hsdS*. These observations may suggest that the *symE*/*symR* module is distributed from one chromosome to another by utilizing of selective advantages of the mobile elements.

A clear understanding of the importance of type I TA modules has yet to be obtained despite more than 25 years of devoted study. Each TA cassette seems to not be essential because they can be deleted easily from the chromosome. So, what is the biological significance to bacteria of containing multiple TA systems in its genome, though the gene product is highly toxic to the cell when overexpressed? It is speculated that chromosomally encoded TA systems act as integral regulators of cellular activity as cell survival/persistence elements during starvation or antibiotic exposition, quality control elements by RNA recycling, beyond PSK-like function that stabilizing of mobile genetic elements (reviewed in ref. x-x). I favour the model that they have a benefit as a defense system against invasion factors like bacteriophages. Since bacteriophage infection can lead to rapid cessation of the host cell transcription, the unstable antitoxin RNAs can be depleted

much faster than stable toxin-encoding mRNAs, and thereby, the toxin protein is synthesized. In addition, it is reported that the *hok*/*sok* modules could reduce the efficiency of T4 plating and decrease the plaque size when carried on a high-copy plasmid. The toxin is actually not bactericidal to the cell but may play a role in interference with phage propagation by modulating the membrane or preventing mature particle formation. The SOS-induced SymE protein might also contribute to preventing the spread of phage infection if it targets phage mRNA preferentially to inhibit translation. In addition, any toxin that reduces host metabolism upon phage or infected-cell attack to the natural bacteria cell could cause abortive phage infection or immune response respectively, so this effect could be common for TA systems. It is worth mentioning the evidence that free-living prokaryotes including pathogenic bacteria which grow very slowly contain abundant TA loci whereas obligate host-associated intracellular organisms have no TA loci.[1-3] Consequently, the TA systems seem like non-redundant systems and their diversity and multiplicity may be explained by a natural consequence of pathogen-host conflict to increase the fitness and hence diversified to respond to various forms of stimuli.

Concluding remarks

Our knowledge of TA systems is still limited but there is increasing evidence that a number of TA modules are harbored in a diverse range of bacteria and archaea. Future investigation of the TA systems will provide insights into fascinating questions why such dispensable multiple units are tolerated in evolution, as well as when and how TA loci are activated to play roles in natural bacterial cell populations. Two experiments should be useful to obtain one more piece of the whole picture; (i) detection of endogenously expressed toxin proteins in a variety of conditions through fusion to a reporter molecule at single-cell resolution by using a fluorescence microscopy since it seems nether all cells relieve from their tight repression for the toxin gene expression nor show a strong phenotypic effect to be observed by a population handling study. (ii) to investigate any differences between a wild type stain and its mutant strain where all known TA modules are deleted or disrupted under conditions in which they may act, for example, phage infection, sudden depletion of nutrient factor, and mixed culture with other type of cells.

In addition, despite the variation in the entire sequences and modes of action of the toxins across type I TA families, these toxin-antitoxin genes pairs seem like a conserved family consisted of a minimum unit of regulatory circuit by different but related genes encoded in the same locus. These so called symbiotic relationship systems probably require complementary features to ensure their retention under evolutionary of pressures; an intact toxin gene is necessary to maintain a functional antitoxin gene on the genome and an antitoxin gene is indispensable for maintaining cellular activity in the survival of the fittest. Thus, the TA system might have some advantages in survival, and being sophisticated and diverse enough to adapt to various environmental circumstances. Therefore, it can be expected that revealing the biological role even from one of the functional studies on a particular TA module might give hints revealing significance of existence on the genome and common property in all toxin-antitoxin systems. Continual research introducing novel ideas will give answers to those pressing questions.

Acknowledgements

The author is very much grateful to Gisela Storz for her comments on the manuscript and stimulating discussions to the *symE*/*symR* project, and Kenn Gerdes for his pioneering studies to the *hok/sok* system.

Legend

Figure 1. Localization of three different type I TA loci on the *E. coli* K-12 genome. *hok*/*sok* (pink), *ldr*/*rdl* (blue) and *symE*/*symR* (green) loci are shown. Asterisks show genes that are clearly degenerated or relics.

Figure 2. Genetic organization of *hokB*/*sokB*, *ldrD*/*rdlD*, and *symE*/*symR* type I TA modules of *E. coli* K-12. (A) The *hokB*/*sokB* locus is located between *cybB* and *trg* at 32.1 min (Fig. 1). This system contains all of the regulatory elements as described for *hok*/*sok* system in plasmid R1, such as *fbi* (foldback inhibition) element, *tac* (translational activation) element, *ucb* (upstream complementary box) promoter sequences, Shine-Dalgarno sequences, and an overlapping reading frame *mokB* (mediation of killing). The *mokB* reading frame is out-of-frame with *hokB* and terminates 38 nt upstream of *hokB*. (B) The *ldrD*/*rdlD* locus is located between *bcsG* and *yhjV* at 79.7 min (Fig. 1). A second open reading frame *ldrX*, noted by Gerdes and Wagner, that overlaps with *ldrD* as in-frame, thus they share the same translational termination codon. It is predicted that RdlD RNA regulates *ldrD* translation by regulating *ldrX* translation. (C) The *symE*/*symR* locus is located between restriction-modification related genes *mcrB* and *hsdS* at 98.7 min (Fig. 1). The *sym E* promoter has a LexA binding site and is strongly induced by DNA damaging agents. SymR is encoded opposite the 5' untranslated region (UTR) of *symE*, and base pairing can extend over the Shine-Dalgarno sequence as well as the initiation codon of *symE*.

Figure 3. Model for SymE synthesis. SOS-induced symE gene is repressed at three levels by (1) the LexA repressor (transcriptionally), (2) the SymR antisense RNA (post-transcriptionally and/or translationally), and (3) the Lon protease (post-translationally). Other as-yet unknown factors such as ribonucleases and chaperon proteins could be involved in the modulation of SymE synthesis. Endogenous levels of the SymE protein might play a role in degrading particular RNA damaged concomitantly with DNA.

Figure 4. Model for timing of SymE synthesis during the SOS response. The SOS genetic network consists of more than 40 genes in *E. coli* that carry out diverse functions in response to DNA damage, including nucleotide excision repair, homologous recombination, translesion DNA replication, and cell division arrest. The network is controlled by the LexA repressor, which downregulates itself and the expression of the other SOS genes but the peak timing of the induced protein levels seems to be different. SymE protein synthesis may occur at the late stage of the SOS response but before cell lysis. It is suggested that SymE promotion of RNA cleavage may be important for ribosome rescue by the recycling of RNAs damaged under SOS-inducing conditions.

Figure 5. Multiple amino acid sequence alignments. (A) Hok proteins from *E. coli* K-12, *E. coli* O157, and

plasmid R1. (B) Ldr proteins from *E. coli* K-12, *Salmonell typhimurium* LT2*, Salmonella typhi* CT18, and *Citrobacter freundii*. Identical amino acids are boxed, and similar amino acids are indicated by an asterisk at the bottom. The similarity of amino acids was determined by the following rules: L = I = M = V = F = W = A, K = R = H, D = E = Q = N, G = A = S, T = V, A = V and F = Y = H = W. + shows C-terminal positive charged residues. The black line above the aligned amino acids indicates a putative *trans*-membrane a-helical domain predicted by a computer program (SOSUI: http://bp.nuap.nagoya-u.ac.jp/sosui/).

Figure 6. Summary of the type I and II toxins. The targets, types of activity, and cellular process that are affected by the endogenous levels of expression of the toxins need to be more examined. Asterisk denotes paired antitoxin gene.

References

Mattick JS, Makunin IV. Small regulatory RNAs in mammals. Hum Mol Genet. 2005 Apr 15;14 Spec No 1:R121-32

Storz G, Altuvia S, Wassarman KM. An abundance of RNA regulators. Annu Rev Biochem. 2005;74:199–217.

[2-50] Storz, G., and S. Gottesman. 2006. Versatile roles of small RNA regulators in bacteria, p. 567–594. In R. F. Gesteland, T. R. Cech, and J. F. Atkins (ed.), The RNA world, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Beisel CL, Storz G. Base pairing small RNAs and their roles in global regulatory networks. FEMS Microbiol Rev. 2010;34:866–882.

Gottesman S, Storz G. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Biol. 2010 3:a003798.

[1-10] Wagner EG, Altuvia S, Romby P: Antisense RNAs in bacteria and their genetic elements. Adv Genet 2002, 46:361-398.

Brantl S. Regulatory mechanisms employed by cis-encoded antisense RNAs. Curr Opin Microbiol. 2007 Apr;10(2):102-9.

Tomizawa J, Itoh T, Selzer G, Som T. Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. Proc Natl Acad Sci U S A. 1981
Mar;78(3):1421-5.

[1-5] Gerdes K, Gultyaev AP, Franch T, Pedersen K, Mikkelsen ND: Antisense RNA-regulated programmed cell death. Annu Rev Genet 1997, 31:1-31.

[1-12]. Gerdes K, Rasmussen PB, Molin S: Unique type of plasmid maintenance function: postsegregational killing of plasmidfree cells. Proc Natl Acad Sci USA 1986, 83:3116-3120.

[3-3] Yarmolinsky, M.B. (1995) Programmed cell death in bacterial populations. Science 267, 836–837

[1-27] Pedersen K, Gerdes K: Multiple hok genes on the chromosome of Escherichia coli. Mol Microbiol 1999, 32:1090-1102.

[1-32] Kawano M, Oshima T, Kasai H, Mori H: Molecular characterization of long direct repeat (LDR) sequences expressing a stable mRNA encoding for a 35-amino-acid cell-killing peptide and a cis-encoded small antisense RNA in Escherichia coli. Mol Microbiol 2002, 45:333-349.

[1-2] Faridani OR, Nikravesh A, Pandey DP, Gerdes K, Good L: Competitive inhibition of natural antisense Sok–RNA interactions activates Hok-mediated cell killing in Escherichia coli. Nucleic Acids Res 2006, 34:5915-5922.

[2-23] Gerdes, K., and E. G. H. Wagner. 2007. RNA antitoxins. Curr. Opin. Microbiol. 10:117–124.

[2-30] Kawano, M., L. Aravind, and G. Storz. 2007. An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. Mol. Microbiol. 64:738–754.

Yamaguchi Y, Park JH, Inouye M. Toxin-antitoxin systems in bacteria and archaea. Annu Rev Genet. 2011;45:61-79.

Hayes F, Van Melderen L. Toxins-antitoxins: diversity, evolution and function. Crit Rev Biochem Mol Biol. 2011 Oct;46(5):386-408.

[2-18] Gerdes, K., J. E. Larsen, and S. Molin. 1985. Stable inheritance of plasmid R1 requires two different loci. J. Bacteriol. 161:292–298.

[1-12] Gerdes K, Rasmussen PB, Molin S: Unique type of plasmid maintenance function: postsegregational killing of plasmidfree cells. Proc Natl Acad Sci USA 1986, 83:3116-3120.

[1-26] Gerdes K, Poulsen LK, Thisted T, Nielsen AK, Martinussen J, Andreasen PH: The hok killer gene family in Gram-negative bacteria. New Biol 1990, 2:946-956.

[2-34] Loh, S. M., D. S. Cram, and R. A. Skurray. 1988. Nucleotide sequence and transcriptional analysis of a third function (Flm) involved in F-plasmid maintenance.
Gene 66:259–268.

[5-33] Golub EI, Panzer HA. 1988. The F factor of Escherichia coli carries a locus of stable plasmid inheritance stm, similar to the parB locus of plasmid RI. Mol. Gen. Genet. 214:353–57

[2-37] Ohnishi, Y. 1975. F factor promotes turnover of stable RNA in Escherichia coli. Science 187:257–258.

[5-67] Ohnishi Y, Iguma H, Ono T, Nagaishi H, Clark AJ. 1977. Genetic mapping of the F plasmid gene that promotes degradation of stable ribonucleic acid in Escherichia coli. J. Bacteriol. 132:784–89

[2-38] Ohnishi, Y., and S. Akimoto. 1980. I-like R plasmids promote degradation of stable ribonucleic acid in Escherichia coli. J. Bacteriol. 144:833–835.

[2-36] Nielsen, A. K., P. Thorsted, T. Thisted, E. G. H. Wagner, and K. Gerdes. 1991. The rifampicin-inducible genes srnB from F and pnd from R483 are regulated by antisense RNAs and mediate plasmid maintenance by killing of plasmid-free segregants. Mol. Microbiol. 5:1961–1973.

[1-3] Pandey DP, Gerdes K: Toxin – antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. Nucleic Acids Res 2005, 33:966-976.

Figure 1

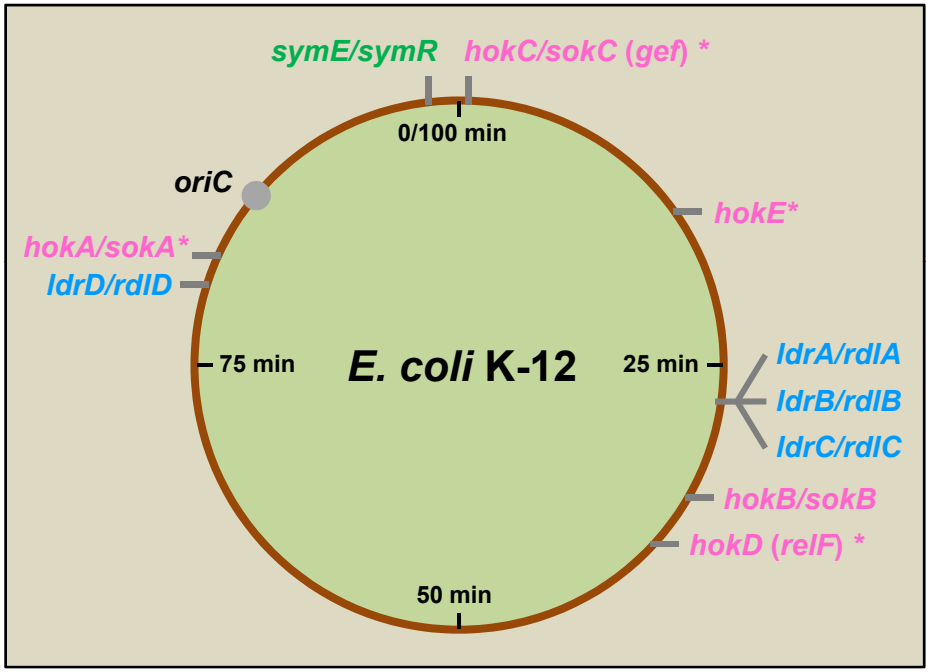

symE/symR
hokC/sokC (gef) *
0/100 min
oriC
hokE*
hokA/sokA*
ldrD/rdlD
75 min
E. coli K-12
25 min
ldrA/rdlA
ldrB/rdlB
ldrC/rdlC
hokB/sokB
hokD (relF) *
50 min

Figure 2

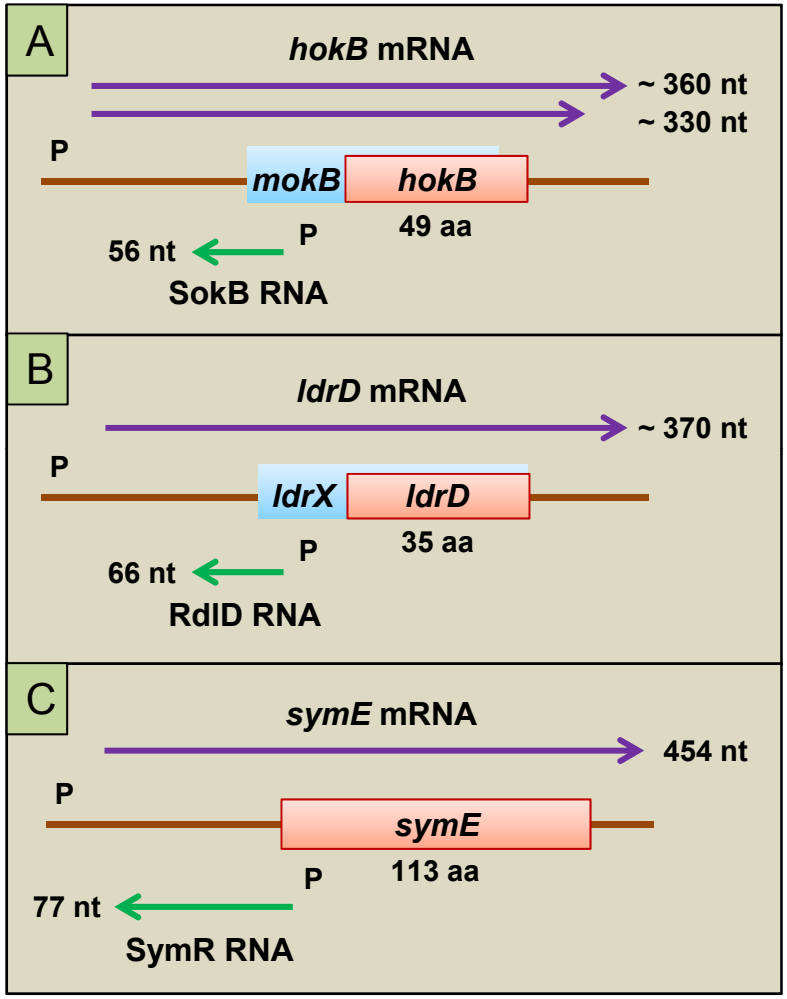

A
hokB mRNA
~ 360 nt
~ 330 nt
P
mokB
hokB
49 aa
P
56 nt
SokB RNA
B
ldrD mRNA
~ 370 nt
P
ldrX
ldrD
35 aa
P
66 nt
RdlD RNA
C
symE mRNA
454 nt
P
symE
113 aa
P
77 nt
SymR RNA

Figure 3

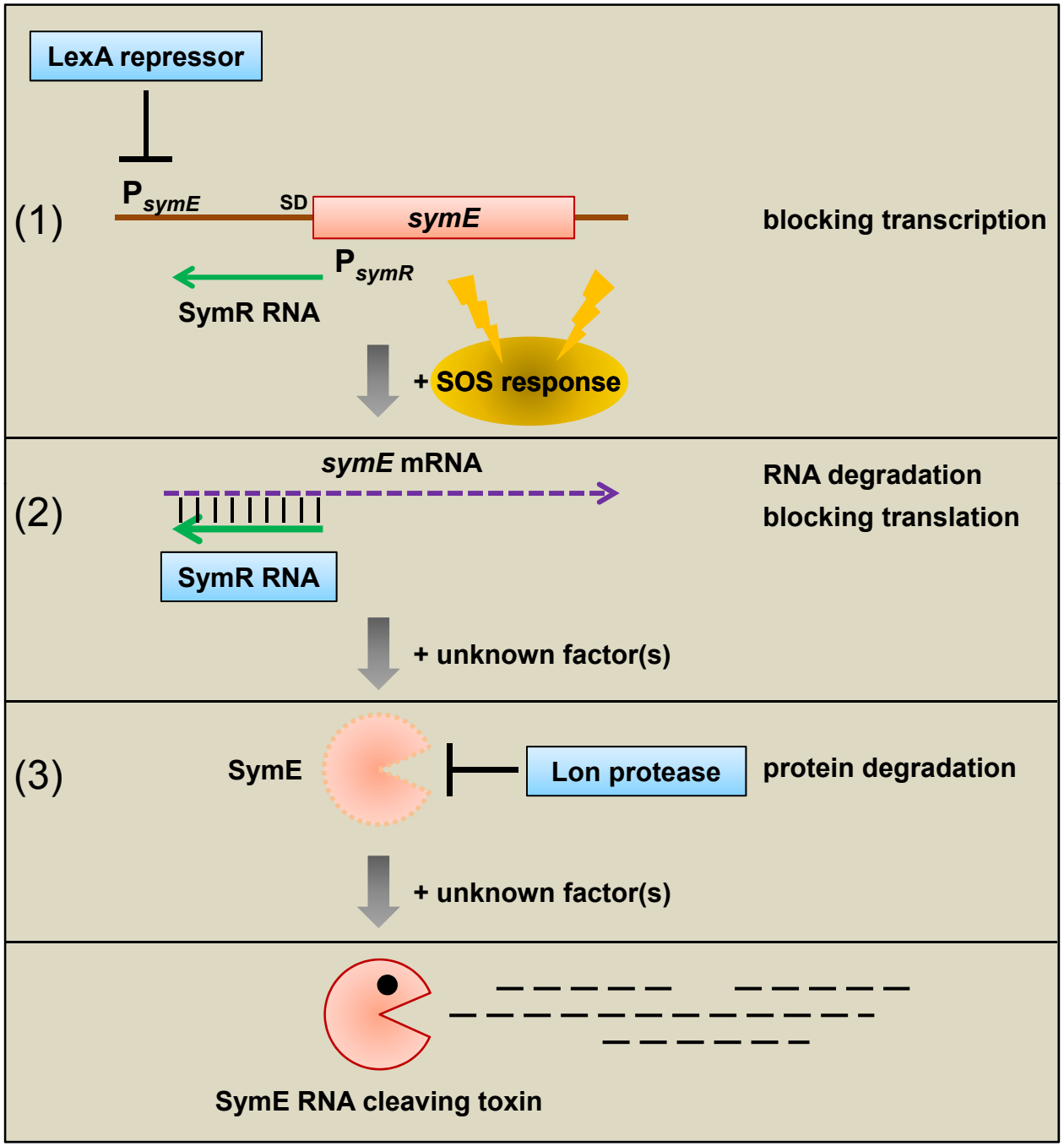

LexA repressor
(1)
P symE
SD
symE
blocking transcription
P symR
SymR RNA
+ SOS response
symE mRNA
RNA degradation
blocking translation
(2)
SymR RNA
+ unknown factor(s)
(3)
SymE
Lon protease
protein degradation
+ unknown factor(s)
SymE RNA cleaving toxin

Figure 4

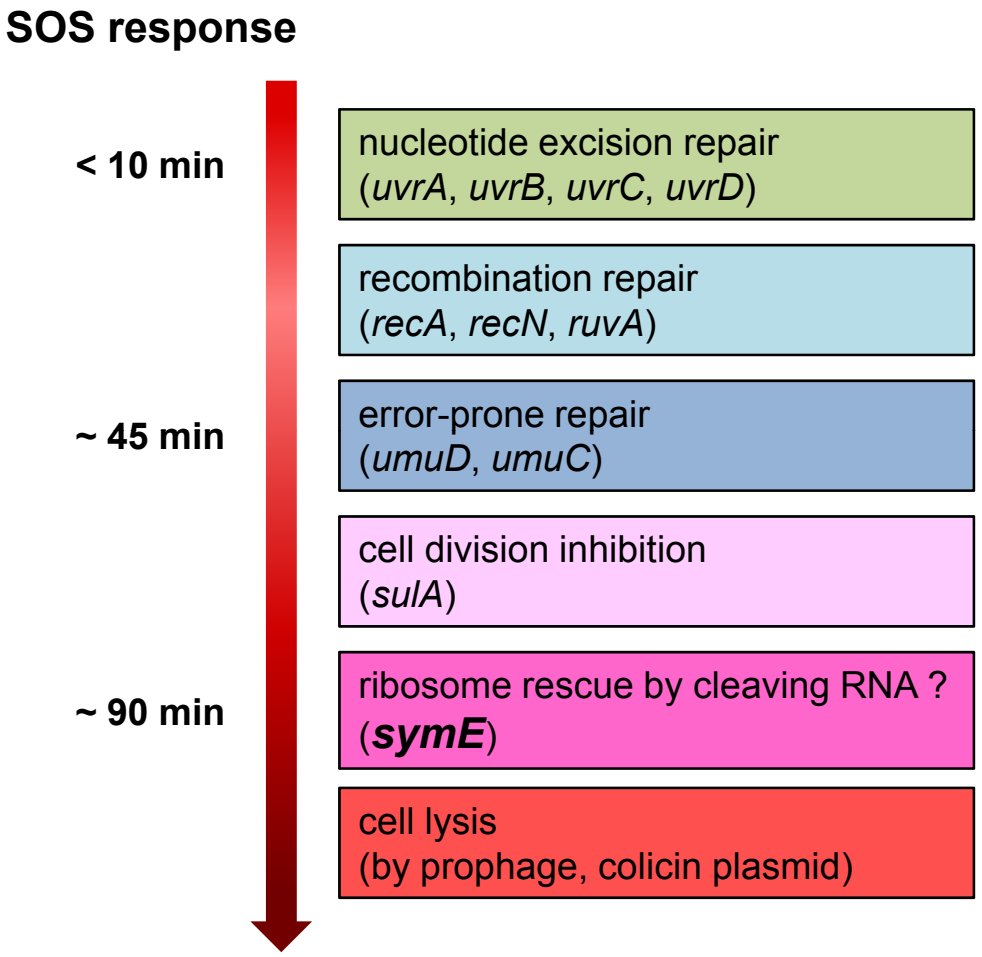

SOS response
< 10 min
~ 45 min
~ 90 min
nucleotide excision repair
(uvrA, uvrB, uvrC, uvrD)
recombination repair
(recA, recN, ruvA)
error-prone repair
(umuD, umuC)
cell division inhibition
(sulA)
ribosome rescue by cleaving RNA ?
(symE)
cell lysis
(by prophage, colicin plasmid)

Figure 5

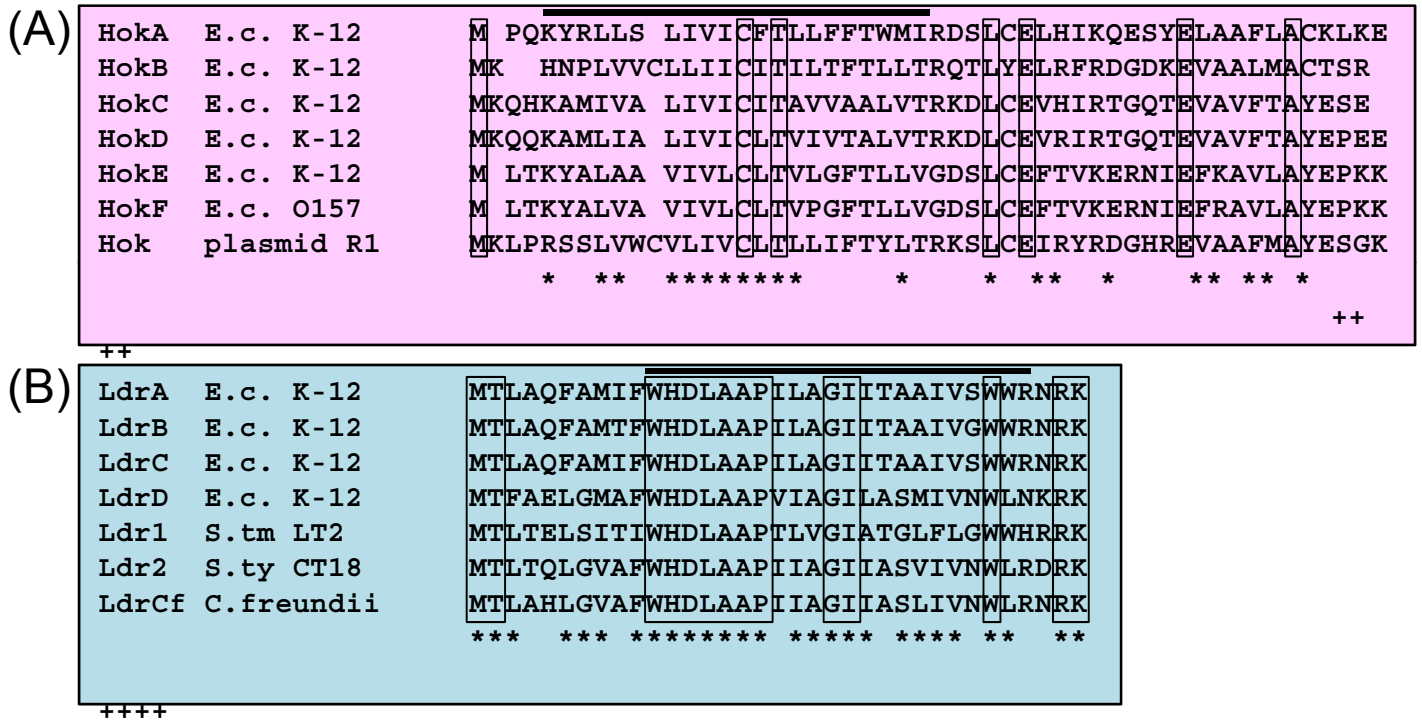
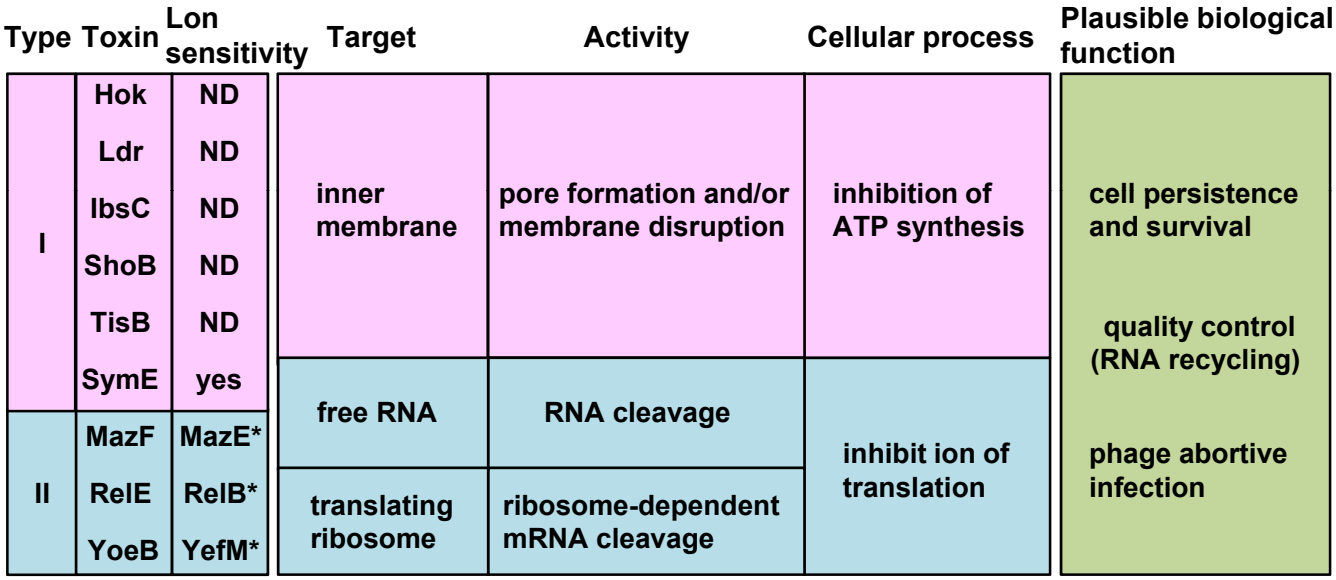

(A)
HokA E.c. K-12 M PQKYRLLS LIVICFTLLFFTWMIRDSLCELHIKQESYELAAFLACKLKE
HokB E.c. K-12 MK HNPLVVCLLIICITILTFTLLTRQTLYELRFRDGDKEVAALMACTSR
HokC E.c. K-12 MKQHKAMIVA LIVICITAVVAALVTRKDLCEVHIRTGQTEVAVFTAYESE
HokD E.c. K-12 MKQQKAMLIA LIVICLTVIVTALVTRKDLCEVRIRTGQTEVAVFTAYEPEE
HokE E.c. K-12 M LTKYALAA VIVLCLTVLGFTLLVGDSLCEFTVKERNIEFKAVLAYEPKK
HokF E.c. O157 M LTKYALVA VIVLCLTVPGFTLLVGDSLCEFTVKERNIEFRAVLAYEPKK
Hok plasmid R1 MKLPRSSLVWCVLIVCLTLLIFTYLTRKSLCEIRYRDGHREVAAFMAYESGK
++
(B)
LdrA E.c. K-12 MTLAQFAMIFWHDLAAPILAGIITAAIVSWWRNRK
LdrB E.c. K-12 MTLAQFAMTFWHDLAAPILAGIITAAIVGWWRNRK
LdrC E.c. K-12 MTLAQFAMIFWHDLAAPILAGIITAAIVSWWRNRK
LdrD E.c. K-12 MTFAELGMAFWHDLAAPVIAGILASMIVNWLNKRK
Ldr1 S.tm LT2 MTLTELSITIWHDLAAPTLVGIATGLFLGWWHRRK
Ldr2 S.ty CT18 MTLTQLGVAFWHDLAAPIIAGIIASVIVNWLRDRK
LdrCf C.freundii MTLAHLGVAFWHDLAAPIIAGIIASLIVNWLRNRK
++++

Figure 6


Type
Toxin
Lon sensitivity
Target
Activity
Cellular process
Plausible biological function
I
Hok
ND
Ldr
ND
IbsC
ND
ShoB
ND
TisB
ND
SymE
yes
inner membrane
pore formation and/or membrane disruption
inhibition of ATP synthesis
cell persistence and survival
quality control (RNA recycling)
II
MazF
MazE*
RelE
RelB*
YoeB
YefM*
free RNA
RNA cleavage
translating ribosome
ribosome-dependent mRNA cleavage
inhibit ion of translation
phage abortive infection