# Supporting Information

## Short et al. 10.1073/pnas.1216039110

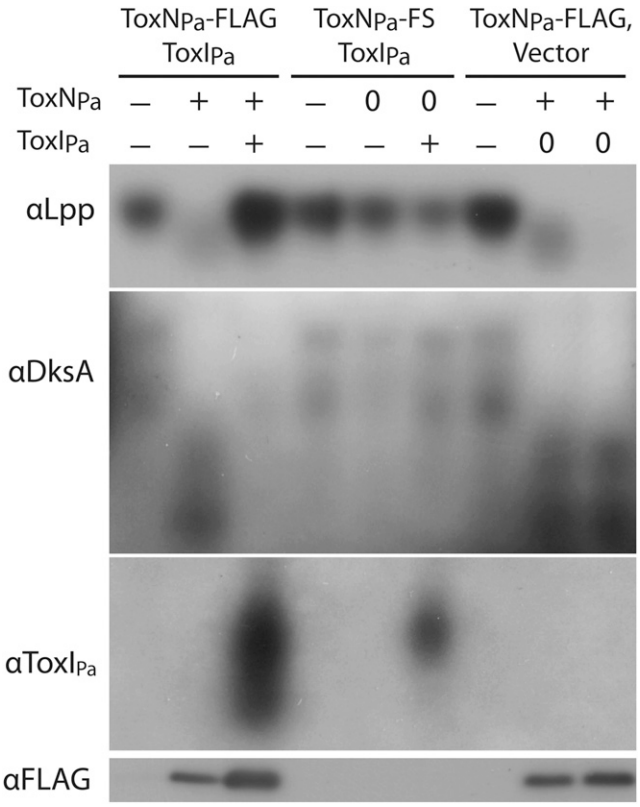


**Fig. S1.** $ToxN_{Pa}$ inhibition by $ToxI_{Pa}$. $ToxN_{Pa}$ degrades lpp and DksA RNAs and is inhibited by $ToxI_{Pa}$ in vivo. *Escherichia coli* cells containing separately inducible $ToxN_{Pa}$-FLAG and $ToxI_{Pa}$ plasmids were grown to log phase, and the effect of expression of $ToxN_{Pa}$-FLAG and subsequent coexpression of $ToxI_{Pa}$ on *lpp* and *dksA* transcript levels was analyzed by Northern blot (*Top Two Panels*). Expression of $ToxI_{Pa}$ (*Third Panel*) and $ToxN_{Pa}$-FLAG (*Bottom Panel*) also were assessed by Northern and Western blot, respectively. Negative controls for $ToxN_{Pa}$ (the frame shifted $ToxN_{Pa}$-FS) and $ToxI_{Pa}$ (vector) are as indicated.

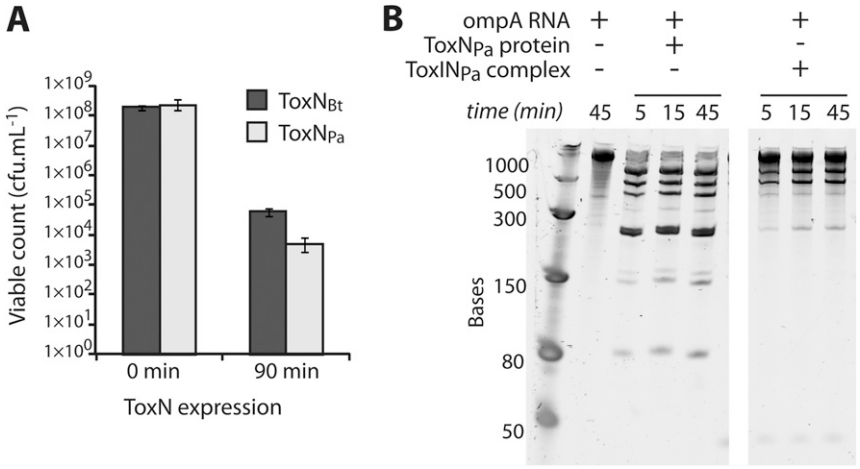


**Fig. S2.** (*A*) $ToxN_{Bt}$ has lower toxicity than $ToxN_{Pa}$ in *E. coli* in vivo. Viable counts of *E. coli* DH5α cultures carrying either $ToxN_{Bt}$ or $ToxN_{Pa}$ on pBAD30 are shown before and 90 min after induction of ToxN expression. Results shown are mean ± SD for three biological replicates. (*B*) $ToxN_{Pa}$ has higher activity than $ToxIN_{Pa}$ in vitro. Reactions (6 pmol *ompA* + 6 pmol $ToxN_{Pa}$) were incubated at 37 °C, and samples were taken at the times indicated. $ToxN_{Pa}$ protein and $ToxIN_{Pa}$ complex were purified by FPLC. It was not possible to match the concentration of $ToxN_{Pa}$ in the monomer and complex samples precisely because the monomeric form was less stable in solution.

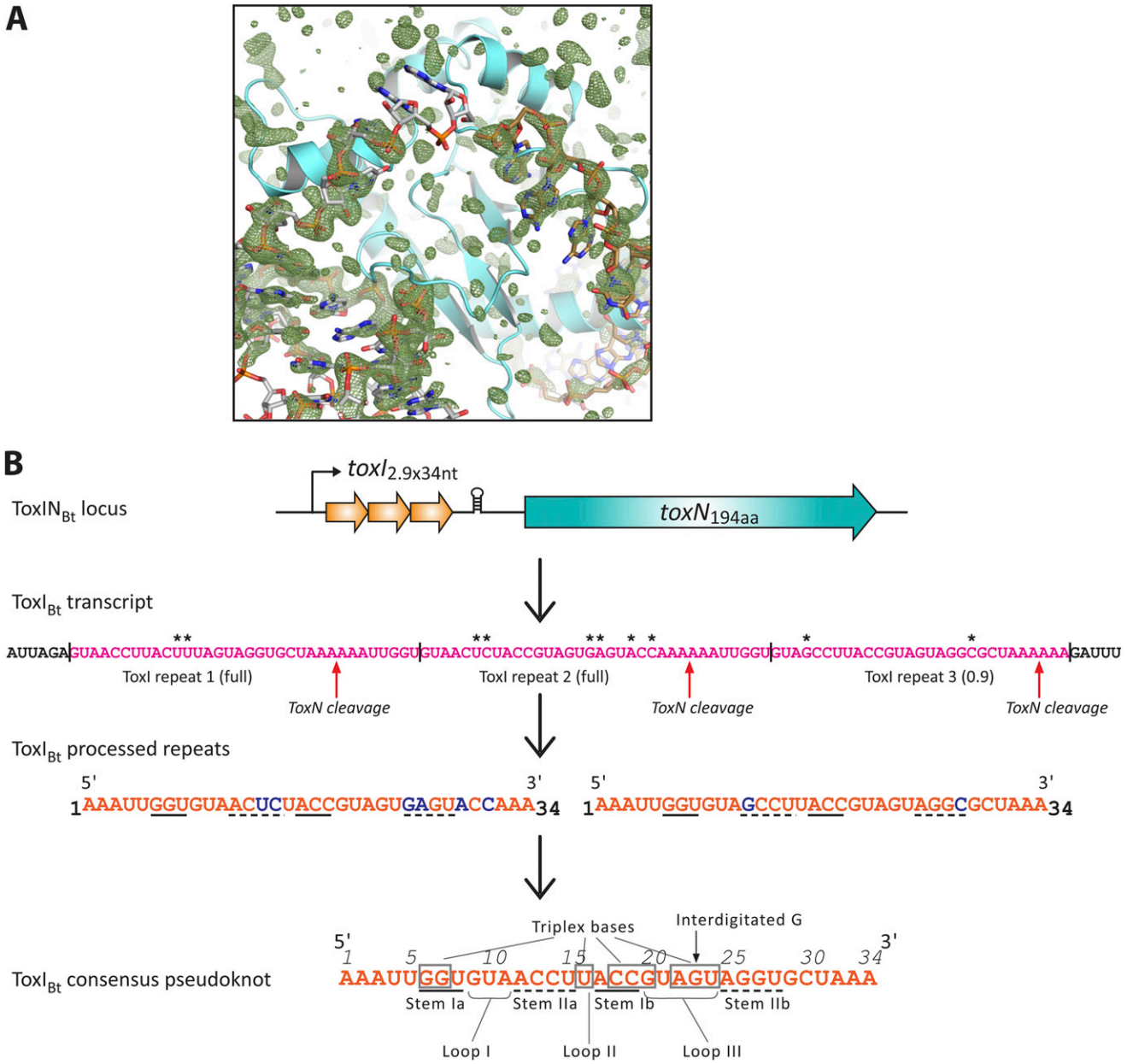


**Fig. S3.** (*A*) Initial Fo-Fc map following molecular replacement with a modified ToxN$_{Pa}$ protomer as the search model (see *Materials and Methods*). The map is contoured at 2.5 σ. ToxN is shown as cyan in a cartoon. The superimposed ToxI$_{Bt}$ chains are shown as silver sticks for the RNA contained in the asymmetric unit and as sand-colored sticks for the symmetry-related molecule. (*B*) ToxI$_{Bt}$ consensus RNA sequence. The ToxIN$_{Bt}$ locus consists of a single promoter, a series of 2.9 nearly perfect 34-nt repeats, a transcriptional terminator hairpin, and the gene for ToxN$_{Bt}$. The ToxI$_{Bt}$ repeats are transcribed as a continuous series and then are cleaved by ToxN$_{Bt}$ at AAA↓AAA sequences to produce two 34-nt processed ToxI$_{Bt}$ RNAs. Each transcript encodes only two final ToxI$_{Bt}$ pseudoknot sequences; therefore the three ToxI$_{Bt}$ RNAs observed in the complex structure must be generated from more than one transcript. Because the two final ToxI$_{Bt}$ RNAs are not identical in sequence, the consensus repeat sequence defined by tandem repeat finder (1) was used to solve the structure and as the wild-type sequence for ToxI$_{Bt}$ mutagenesis and antitoxicity assays. The two processed ToxI$_{Bt}$ RNAs generated from each transcript differ from the consensus ToxI$_{Bt}$ RNA sequence in that they contain compensatory mutations in stem II, and the first repeat also contains two nucleotide substitutions in the tail following stem IIb. Sequences are colored as follows: ToxI$_{Bt}$ DNA repeats, orange; ToxN$_{Bt}$ gene, teal; ToxI$_{Bt}$ unprocessed transcript, pink with nonconsensus nucleotides indicated by an asterisk; ToxI$_{Bt}$ processed repeats, orange with nonconsensus nucleotides in blue and base-pairing regions underlined; ToxI$_{Bt}$ consensus pseudoknot, orange with structural features and numbering as indicated.

1. Benson G (1999) Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res* 27(2):573–580.

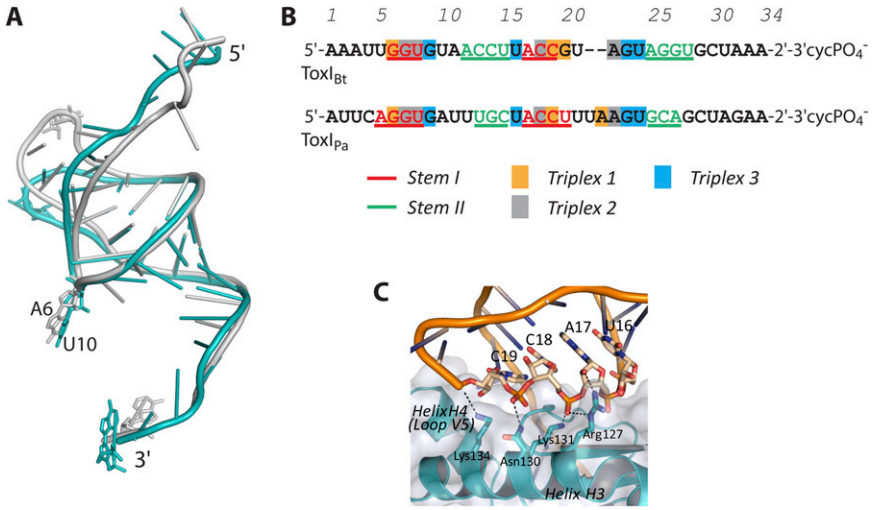


**Fig. S4.** Comparison of $ToxI_{Bt}$ and $ToxI_{Pa}$ structures. (*A*) Least-squares superimposition of the $ToxI_{Bt}$ (teal) and $ToxI_{Pa}$ (silver) (PDB ID code 2XD0) RNA pseudoknot structures, both represented as cartoons. (*B*) Structure-based alignment of $ToxI_{Bt}$ and $ToxI_{Pa}$ sequences. Base-pairing regions and nucleotides forming triplexes are indicated. Numbers correspond to nucleotides of $ToxI_{Bt}$. (*C*) Detail of $ToxN_{Bt}$ interactions with $ToxI_{Bt}$ backbone at interface 2, shown as view from behind helix H3. $ToxI_{Bt}$ is shown as a cartoon with key nucleotides as pale pink sticks; $ToxN_{Bt}$ is shown as a cartoon in teal with key residues shown as sticks. Black dashed lines indicate hydrogen bonds.

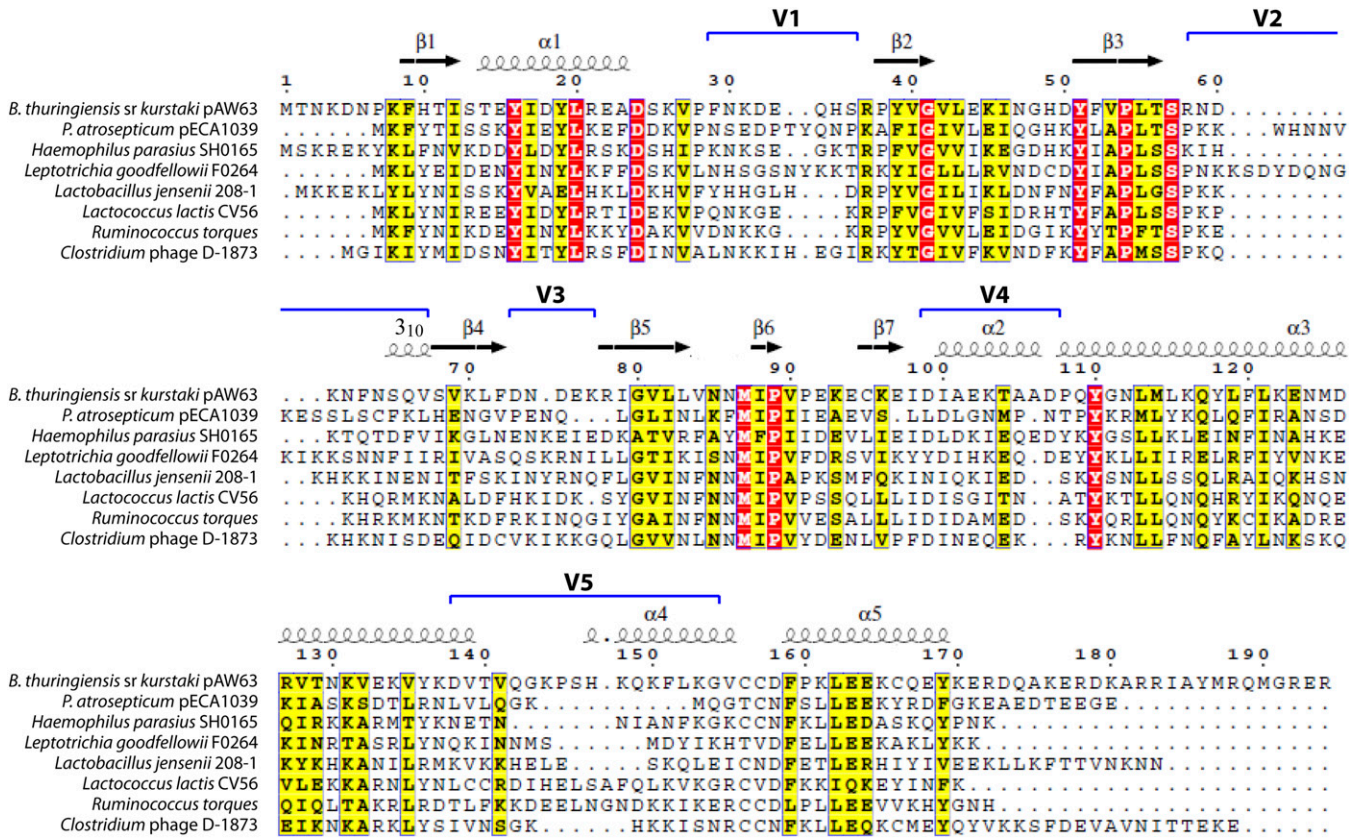


**Fig. S5.** Sequence alignment of eight diverse ToxN proteins selected from a phylogenetic tree of the ToxN protein family (1) and aligned using CLUSTALW (2). The secondary structure of $ToxN_{Bt}$ is indicated above the alignment. Note that $ToxN_{Bt}$ residues 174–194 were disordered in the crystal structure. The figure was generated using ESPript2 (3).


1. Blower TR, et al. (2012) Identification and classification of bacterial type III toxin-antitoxin systems encoded in chromosomal and plasmid genomes. *Nucleic Acids Res* 40(13):6158–6173.
2. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947–2948.
3. Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: Analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15(4):305–308.

**Table S1. Size exclusion chromatography of $ToxIN_{Pa}$ assembly reactions**

| Sample | Molecular mass (Da) | Elution volume (mL) | $K_{av}$* | Calculated molecular mass (Da)† |
|---|---|---|---|---|
| ToxIN samples | | | | |
| ToxIN complex | 9,4341 | 12.77 | 0.271 | 73,025 |
| ToxN protein | | 16.42 | 0.508 | 28,376 |
| ToxI monomer | | 16.48 | 0.512 | 28,052 |
| ToxI transcript | | 11.05 | 0.160 | 162,264 |
| ToxN + ToxI single peak 1 | | 12.7 | 0.267 | 74,906 |
| ToxN + ToxI single peak 2 | | 16.5 | 0.513 | 27,945 |
| ToxN + ToxI transcript peak 1 | | 12.65 | 0.263 | 76,300 |
| ToxN + ToxI transcript peak 2 | | 14.56 | 0.387 | 42,692 |
| ToxN + ToxI transcript peak 3 | | 16.37 | 0.505 | 28,651 |
| Standards | | | | |
| Blue dextran | Void | 8.59 | | |
| Lysozyme | 14,300 | 20.17 | 0.751 | 15,740 |
| Myoglobin | 17,000 | 16.95 | 0.543 | 25,711 |
| BSA monomer | 66,780 | 14.76 | 0.400 | 40,625 |
| BSA dimer | 133,560 | 12.35 | 0.244 | 85,651 |
| Thyroglobulin | 669,000 | 9.31 | 0.047 | 1,032,383 |

*$K_{av} = (V_e - V_o)/(V_t - V_o)$ where $V_e$ is the elution volume of the sample, $V_o$ is the void volume, and $V_t$ is the total column volume.
†The $K_{av}$ and molecular mass values for the five size standards were fitted to an exponential function to give the equation $K_{av} = 452.39(Da)^{-0.663}$, from which molecular mass values for ToxIN samples were calculated.

**Table S2. Data collection and refinement statistics**

| | |
|---|---|
| Data collection | |
| Space group | P6 |
| Cell dimensions | |
| *a*, *b*, *c* (Å) | 127.10, 127.10, 37.74 |
| $\alpha$, $\beta$, $\gamma$ (°) | 90.00, 90.00, 120.00 |
| Wavelength (Å) | 0.9795 |
| Resolution (Å) | 27.95–2.2 |
| $R_{merge}$(%)[#,1] | 5 (54.7) |
| $I/\sigma I$ | 34.8 (4.6) |
| Completeness (%)[#] | 100.0 (99.9) |
| Redundancy[#] | 11.0 (11.1) |
| Molecular replacement | |
| Search model | A: 2XDB, protein part only, modified<br>B: $ToxN_{Bt}$-$ToxI_{Bt}$ structure solved using A, with $ToxI_{Bt}$ built into omit map |
| Rotation and translation search score | A: Rotation function Z = 5.9; translation function Z = 8.8; packing clashes = 0; log likelihood gain = 52<br>B: Rotation function Z = 10.0; translation function Z = 21.0; packing clashes = 0; log likelihood gain = 102 |
| Refinement | |
| Resolution (Å) | 27.95–2.2 |
| No. Reflections | 34,617 |
| $R_{work}$ (%)/$R_{free}$ (%)[2] | 16.23/19.59 |
| No. atoms | |
| Protein | 1,382 |
| RNA | 726 |
| Water | 174 |
| *B*-factors (Å$^2$) | |
| Protein | 46.56 |
| RNA | 43.32 |
| Water | 48.98 |
| r.m.s. deviations | |
| Bond lengths (Å) | 0.01 |
| Bond angles (°) | 1.25 |
| Ramachandran plot | 97% favored, 3% allowed, 0% disallowed |
| Coordinate error (Luzzati plot) | $\sigma$ = 0.2429 |

**Table S3. Plasmids used in this study**

| Name | Primers used* | Description | Source |
|---|---|---|---|
| pACYC184 | — | *E. coli* cloning vector, p15A origin, $Cm^R$ | (1) |
| pET21b+ | — | Expression vector, $Ap^R$, T7 promoter | Novagen |
| pBAD30 | — | Expression vector, $Ap^R$, Ara promoter induced by L-ara, repressed by D-glc | (2) |
| pBluScript KSII+ | — | Phagemid, T7 promoter, $Ap^R$ | Stratagene |
| pHCMC05 | — | *E. coli-Bacillus* shuttle vector, $Ap^R$ $Cm^R$ | (3) |
| pRBJ200 | — | *E. coli par*-deficient single-copy vector, $Ap^R$ | (4) |
| pFLS44 | FS49, PF196 | $ToxI_{Bt}$ promoter, repeats and terminator in pACYC184, $Cm^R$ | This study |
| pFLS49 | FS59, FS60 | *E. coli* K-12 *ompF* in pBluScript KSII+, $Ap^R$ | This study |
| pFLS50 | FS61, FS62 | *E. coli* K-12 *ompA* in pBluScript KSII+, $Ap^R$ | This study |
| pFLS51 | FS63, FS64 | *E. coli* K-12 *dksA* in pBluScript KSII+, $Ap^R$ | This study |
| pFLS52 | FS65, FS66 | *E. coli* K-12 *rpoD* in pBluScript KSII+, $Ap^R$ | This study |
| pFLS53 | FS67, FS68 | *E. coli* K-12 *lpp* in pBluScript KSII+, $Ap^R$ | This study |
| pFLS67 | FS45, FS77 | $ToxN_{Bt}$ C-terminal 6xHis in pET21b+, $Ap^R$ | This study |
| pFLS79 | FS105, [FS73, FS74], FS101 | $ToxIN_{Bt}$-frameshift locus in pHCMC05, $Ap^R$ $Cm^R$ | This study |
| pFLS80 | FS104, FS101 | $ToxIN_{Bt}$ locus in pHCMC05, $Ap^R$ $Cm^R$ | This study |
| pFLS84 | FS112, PF185 | $ToxI_{Bt}$ single consensus repeat in pTA100, $Sp^R$ | This study |
| pFLS88 | FS130, PF185 | $ToxI_{Bt}$ randomized sequence U4-U31 in pTA100, $Sp^R$ | This study |
| pFLS99 | FS146, PF185 | $ToxI_{Bt}$ G20C in pTA100, $Sp^R$ | This study |
| pFLS100 | FS147, PF185 | $ToxI_{Bt}$ G20U in pTA100, $Sp^R$ | This study |
| pFLS103 | PF197, [FS144, FS145], PF195 | $ToxN_{Bt}$ F29A in pBAD30, $Ap^R$ | This study |
| pFLS118 | FS49, FS168 | $ToxIN_{Bt}$ locus in pRBJ200, $Ap^R$ | This study |
| pFLS121 | FS169, FS183 | $ToxIN_{Pa}$ locus in pRBJ200, $Ap^R$ | This study |
| pSLO1 | PF197, [SO3, SO4], PF195 | $ToxN_{Bt}$ S57A in pBAD30, $Ap^R$ | This study |
| pSLO4 | PF197, [SO14, SO15], PF195 | $ToxN_{Bt}$ K31A in pBAD30, $Ap^R$ | This study |
| pSLO5 | PF197, [SO18, SO19], PF195 | $ToxN_{Bt}$ R58A in pBAD30, $Ap^R$ | This study |
| pSLO7 | PF197, [SO22, SO23], PF195 | $ToxN_{Bt}$ Y110F in pBAD30, $Ap^R$ | This study |
| pSLO8 | PF197, [SO24, SO25], PF195 | $ToxN_{Bt}$ K148A in pBAD30, $Ap^R$ | This study |
| pSLO10 | PF185, SO9 | $ToxI_{Bt}$ U8A in pTA100, $Sp^R$ | This study |
| pSLO11 | PF185, SO10 | $ToxI_{Bt}$ G9U in pTA100, $Sp^R$ | This study |
| pSLO12 | PF185, SO11 | $ToxI_{Bt}$ U10A in pTA100, $Sp^R$ | This study |
| pSLO13 | PF185, SO12 | $ToxI_{Bt}$ G20A in pTA100, $Sp^R$ | This study |
| pSLO14 | PF185, SO13 | $ToxI_{Bt}$ G23A in pTA100, $Sp^R$ | This study |
| pTA49 | — | $ToxN_{Pa}$ in pBAD30, $Ap^R$ | (5) |
| pTA50 | — | Frameshift $ToxN_{Pa}$ in pBAD30, $Ap^R$ | (5) |
| pTA76 | — | Full $ToxI_{Pa}$ array in pTA100, $Sp^R$ | (5) |
| pTA100 | — | pQE80-L derivative, $Sp^R$ | (5) |
| pTA110 | — | $ToxI_{Pa}$ in pBluScript KSII+ for antisense transcription, $Ap^R$ | This study |
| pTA111 | — | $ToxI_{Pa}$ in pBluScript KSII+ for sense transcription, $Ap^R$ | (6) |
| pTA115 | — | Full $ToxI_{Bt}$ array in pTA100, $Sp^R$ | (5) |
| pTA117 | — | $ToxN_{Bt}$ in pBAD30, $Ap^R$ | (5) |
| pTRB1 | — | $ToxN_{Pa}$-FLAG in pBAD30, $Ap^R$ | (5) |
| pTRB14 | — | $ToxN_{Pa}$ in pTYB1, $Ap^R$ | (6) |
| pTRB18 | — | $ToxI_{Pa}$ promoter, repeats and terminator in pACYC184, $Cm^R$ | (6) |

*Overlap PCR primers used to introduce mutations are shown in brackets.


1. Chang AC, Cohen SN (1978) Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* 134(3):1141–1156.
2. Guzman LM, Belin D, Carson MJ, Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* 177(14): 4121–4130.
3. Nguyen HD, et al. (2005) Construction of plasmid-based expression vectors for *Bacillus subtilis* exhibiting full structural stability. *Plasmid* 54(3):241–248.
4. Jensen RB, Grohmann E, Schwab H, Díaz-Orejas R, Gerdes K (1995) Comparison of *ccd* of F, *parDE* of RP4, and *parD* of R1 using a novel conditional replication control system of plasmid R1. *Mol Microbiol* 17(2):211–220.
5. Fineran PC, et al. (2009) The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proc Natl Acad Sci USA* 106(3):894–899.
6. Blower TR, et al. (2009) Mutagenesis and functional characterization of the RNA and protein components of the *toxIN* abortive infection and toxin-antitoxin locus of *Erwinia*. *J Bacteriol* 191(19):6029–6039.

**Table S4. Primers used in this study**

| Name | Sequence 5′–3′* | Restriction site |
|---|---|---|
| FS45 | GGTGGTCATATGACTAATAAAGATAATCCT | NdeI |
| FS49 | CCTTGGATCCGCAGAGAGAGATAAATAA | BamHI |
| FS59 | GGAGGAGAGCTCATGATGAAGCGCAATATT | SacI |
| FS60 | GGAGGAAAGCTTTTAGAACTGGTAAACGAT | HindIII |
| FS61 | GGAGGAGAGCTCATGAAAAAGACAGCTATC | SacI |
| FS62 | GGAGGAAAGCTTTTAAGCCTGCGGCTGAGT | HindIII |
| FS63 | GGAGGAGAGCTCATGCAAGAAGGGCAAAAC | SacI |
| FS64 | GGAGGAAAGCTTTTAGCCAGCCATCTGTTT | HindIII |
| FS65 | GGAGGAGAGCTCATGGAGCAAAACCCGCAG | SacI |
| FS66 | GGAGGAAAGCTTTTAATCGTCCAGGAAGCT | HindIII |
| FS67 | GGAGGAGAGCTCATGAAAGCTACTAAACTG | SacI |
| FS68 | GGAGGAAAGCTTTTACTTGCGGTATTTAGT | HindIII |
| FS73 | GAAAAAACAGCATAATAATCAGTATGGTAATTTGAT | — |
| FS74 | TTACCATACTGATTATTATGCTGTTTTTTCTGC | — |
| FS77 | GGTGAAGCTTAATGGTGATGGTGATGGTGCGCTCTCTCACGCCCCATTTG | HindIII |
| FS79 | CAAAGCCAACATACGGGT | — |
| FS80 | AGTGTCTGCACGCCATAC | — |
| FS81 | CAGGAGTGATCGCGTACT | — |
| FS82 | GGATGATGTCTTCGATCT | — |
| FS83 | CAGATTCCACGCTGGAAA | — |
| FS84 | CACGCATGTACATGCGTA | — |
| FS85 | TTCGCGAGCCAGTTCCGG | — |
| FS86 | TGCACTGCTCAACGCAGA | — |
| FS87 | CTTAGCGATAGAAATAAC | — |
| FS88 | ATCAACCGCTTTCATCAG | — |
| FS89 | CAACACCGCAGGATTCGC | — |
| FS90 | CTGCAGCGTTTGCAGTAC | — |
| FS91 | GAGTTTCCCCTTTAAAAC | — |
| FS92 | AAGACCCGCGAATGCCAG | — |
| FS93 | ATTAGTGATCGGCGTAGC | — |
| FS94 | GGTGTAAGCGATGGACGG | — |
| FS95 | TAGCGTTGCTGGAGCAAC | — |
| FS96 | AGTAGCCATGTTGTCCAG | — |
| FS97 | GTTGAAGCTTCAGATTCCACGCTGGAAA | HindIII |
| FS98 | GTTGAAGCTTTGCACTGCTCAACGCAGA | HindIII |
| FS99 | GTTGAAGCTTATCAACCGCTTTCATCAG | HindIII |
| FS101 | GGTGCCCGGGTTAATGGTGATGGTGATGGTGTCTCTCACGCCCCATTTG | SmaI |
| FS104 | GGTGCCCGGGTTATCTCTCACGCCCCATTTG | |
| FS105 | CCTTGGTACCGCAGAGAGAGATAAATAA | KpnI |
| FS112 | TTTAAGCTTtttagcacctactacggtaaggttacaccaatttTTGAATCTATTATAATTGTTATCCG | HindIII |
| FS130 | TTTAAGCTTtttaccatgactagaataacgagcctcgacttttTTGAATCTATTATAATTGTTATCCG | HindIII |
| FS144 | AGTAAAGTACCTGCCAATAAAGATGAACAGCATAGC | — |
| FS145 | CTGTTCATCTTTATTGGCAGGTACTTTACTATCCGCTTC | — |
| FS146 | TTTAAGCTTtttagcacctactagggtaaggttacaccaatttTTGAATCTATTATAATTGTTATCCG | HindIII |
| FS147 | TTTAAGCTTtttagcacctactaaggtaaggttacaccaatttTTGAATCTATTATAATTGTTATCCG | HindIII |
| FS168 | TTTCTCGAGATCTCTCACGCCCC | XhoI |
| FS169 | TTTTGGATCCGTTTTATCGACATTGTGAACC | BamHI |
| FS183 | TTTTCTCGAGCTATTACTCGCCTTCTTCC | XhoI |
| M13 Fwd -20 | GTTTTCCCAGTCACGAC | — |
| MJ12 | TTTTAAGCTTCATATTTTTCTCGTAAAAAAGGCGACTATG | HindIII |
| PF185 | AAACAAATAGGGGTTCCG | — |
| PF195 | TTTAAGCTTATCTCTCACGCCCC | HindIII |
| PF196 | TTTAAGCTTCAACTTTCTTCTCCCCC | HindIII |
| PF197 | TTTGAATTCGGAGAAGAAAGTTGACTAATAAAG | EcoRI |
| polyC Race | CGTATCGATGTCGACCCCCCCCCCCCCCCD | SalI, ClaI |
| SO3 | GTACCTTTAACAGCCCGTAACGATAAAAATTTTAAC | — |
| SO4 | TTTATCGTTACGGGCTGTTAAAGGTACAAAATAATCATG | — |
| SO9 | TTTAAGCTTtttagcacctactacggtaaggttactccaatttTTGAATCTATTATAATTGTTATCCG | HindIII |
| SO10 | TTTAAGCTTtttagcacctactacggtaaggttaaaccaatttTTGAATCTATTATAATTGTTATCCG | HindIII |
| SO11 | TTTAAGCTTtttagcacctactacggtaaggtttcaccaatttTTGAATCTATTATAATTGTTATCCG | HindIII |
| SO12 | TTTAAGCTTtttagcacctactatggtaaggttacaccaatttTTGAATCTATTATAATTGTTATCCG | HindIII |
| SO13 | TTTAAGCTTtttagcacctattacggtaaggttacaccaatttTTGAATCTATTATAATTGTTATCCG | HindIII |
| SO14 | AAAGTACCTTTTAATGCCGATGAACAGCATAGCAGA | — |

**Table S4. Cont.**

| Name | Sequence 5′–3′* | Restriction site |
|---|---|---|
| SO15 | ATGCTGTTCATCGGCATTAAAAGGTACTTTACTATCCGC | — |
| SO18 | GTACCTTTAACATCAGCCAACGATAAAAATTTTAACAGT | — |
| SO19 | ATTTTTATCGTTGGCTGATGTTAAAGGTACAAAATAATC | — |
| SO22 | GCAGACCCTCAGTTTGGTAATTTGATGTTAAAACAG | — |
| SO23 | TAACATCAAATTACCAAACTGAGGGTCTGCTGCTGT | — |
| SO24 | GGAAAGCCTTCACATGCCCAAAAATTCTTAAAAGGAGTT | — |
| SO25 | TAAGAATTTTTGGGCATGTGAAGGCTTTCCTTG | — |
| TRB57 | TTTGAGCTCAAGGTGATTTGCTACCTTTAAG | SacI |
| TRB230 | GCGTAATACGACTCACTATAGGGCGATCAGTTGAACGCCCTGAG | — |

*For $ToxI_{Bt}$ primers, the sequence corresponding to the $ToxI_{Bt}$ repeat is shown in lowercase.