Multilayered Regulation of the Type I Toxin ZorO in EHEC

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Jia Wen
August 2016

Copyright © 2016 by Jia Wen
All rights reserved.

DEDICATION

To Grandma, Mom, Dad, and Z.

ACKNOWLEDGEMENTS

To start, I want to thank my advisor, Dr. Elizabeth Fozo, for everything she has done for me. When I started graduate school, I barely had any practical knowledge of microbiology. Dr. Fozo allowed me to join her lab and trained me by hand to do science. She showed me the correct attitude to pursue science and accompanied me through all the challenges I encountered. Her input and support over the years have been vital to my success, and I could not have done it without her.

To everyone on my committee, thank you for your invaluable investment in my research and you all have made me a better scientist. Thank you Dr. Albrecht von Arnim for introducing me to the RNA world in the molecular biology class. Thank you Dr. Alison Buchan for all the critical advice you provided during my prelim. Thank you Dr. Erik Zinser for showing me how wonderful genetics could be in the 680 class – I felt like a kid entering the candy store for the first time.

Special thanks to Dr. Jeffrey Becker, who understood my fear and supported me when I was debating about my future career plan. I also want to thank Dr. Sarah Lebeis and Dr. Todd Reynolds, for all the help and encouragement they have given to me during my job hunting process.

To my lab mates, thank you for the laugh, the tears, the good, and the not so good days you all have shared with me. I have been very lucky to have a group of terrific people to work with. I want to specially thank John Harp, for his kindness and patience in reading and correcting almost everything I have written.

To all my friends, whom it is impossible to list here, thank you for fighting together with me. It is your support and love that keep me strong. Special thanks to Liyin Huang, who has been as good a friend, and as as nice a person, as I have ever met.

To my family, thank you for allowing me to freely pursue my dream. Everything I have today I owe to you.

ABSTRACT

A bacterial type I toxin-antitoxin system contains two genes: one encodes a small toxic protein and the second, a small regulatory RNA (sRNA) that inhibits toxin production. To date, very few type I loci have been described thoroughly in regards to the regulation of toxin and the function of the toxin at endogenous levels. In this study, I demonstrated that the *zor-orz* locus of *Escherichia coli* O157:H7 is composed of two highly homologous type I toxin-antitoxin systems: *zorO-orzO* and *zorP-orzP*. The *zor* genes encode the toxins and the *orz* genes encode the antitoxin sRNAs. Overexpression of *zorO* is toxic to *E. coli* and causes bacterial growth stasis or cell death; however, co-expression of *orzO* neutralizes this toxicity and restores normal bacterial growth. Rapid membrane depolarization was observed upon ZorO overproduction, suggesting that ZorO targets the membrane. Replacement of two charged amino acids in ZorO, glutamic acid at the 16th position and arginine at the 23rd position, can impair ZorO toxicity.

Given its inherent toxicity, production of ZorO is tightly controlled by both its 5' untranslated region (UTR) and the antitoxin OrzO. The *zorO* 5' UTR harbors two distinct regions that modulate *zorO* translation. One is a putative ribosome standby site that is exposed only upon processing of the 5' UTR and likely promotes translation by facilitating ribosome preloading onto the mRNA. The other region spans from +35 to +50 of *zorO* and is required for optimal translation of *zorO*, although the underlying mechanism remains unclear. The OrzO sRNA inhibits ZorO production by reducing both the stability and the translation of the *zorO* mRNA. Specifically, OrzO base pairs to the putative standby site of *zorO* and impedes translation, potentially through competition with ribosome for this site. Successful base pairing of *zorO* by OrzO requires at least 15 nucleotides of perfect sequence complementarity. Once paired, the RNA duplex can be degraded by RNase III, rendering the *zorO* mRNA untranslatable. Collectively, this multilayered control of ZorO production limits its toxicity. Hence, my work provides new insights into regulation of type I toxins by their antitoxin sRNAs and beyond the sRNAs.

TABLE OF CONTENTS

CHAPTER I Introduction 1
Publication note 2
I. An overview of bacterial toxin-antitoxin systems 3
II. Repression of the type I toxin by the antitoxin sRNA 4
III. Biological roles of type I gene pairs 8
IV. Perspectives 15
V. Research objectives 16
References 18
Appendix I: chapter I tables and figures 25
CHAPTER II The ZorO-OrzO type I toxin-antitoxin locus: repression by the OrzO antitoxin 32
Publication note 33
Abstract 34
I. Introduction 35
II. Materials and Methods 37
III. Results 39
IV. Discussion 46
References 53
Appendix II: chapter II tables and figures 58
CHAPTER III 5' UTR inhibits and enhances translation of the type I toxin ZorO 87
Publication note 88
Abstract 89
I. Introduction 90
II. Materials and Methods 91
III. Results 96
IV. Discussion 101
References 106
Appendix III: chapter III tables and figures 110
CHAPTER IV Characterization of the type I toxin ZorO of *E. coli* O157:H7 130

Publication note 131
Abstract 132
I. Introduction 133
II. Materials and Methods 134
III. Results 135
IV. Discussion 137
References 140
Appendix IV: chapter IV tables and figures 142
CHAPTER V Conclusions and future directions 149
References 154
VITA 156

LIST OF TABLES

Table 1.1. Features of decribed type I toxin-antitoxin loci 25
Table 2.1. Strains and plasmids used in this study 58
Table 2.2. Oligonucleotides used in this study 59
Table 2.3. Summary of the mutants generated and their repressive abilities 61
Table 3.1. Strains and plasmids used in this study 110
Table 3.2. Oligonucleotides used in this study 111
Table 4.1. Strains and plasmids used in this study 142
Table 4.2. Oligonucleotides used in this study 143
Table 4.3 Topology predictions of the ZorO toxin 144

LIST OF FIGURES

Figure 1.1. Repression of the BsrG toxin by the SR4 antitoxin 27
Figure 1.2. Repression of the IbsC toxin production by the SibC antitoxin..... 28
Figure 1.3. Repression of the Hok toxin production by the Sok antitoxin 29
Figure 1.4. Repression of the TisB toxin production by the IstR-1 antitoxin..... 30
Figure 1.5. Repression of the TxpA toxin production by the RatA antitoxin..... 31
Figure 2.1. The *zorO*-*orzO* gene locus is a true type I toxin-antitoxin system..... 64
Figure 2.2. The RNA chaperone Hfq is not required for OrzO to repress *zorO*-induced toxicity 66
Figure 2.3. RNase III is critical for OrzO repression of *zorO* 67
Figure 2.4. Effects of an *rnc* deletion on OrzO expression and processing of *zorO* 68
Figure 2.5. The 5' end of OrzO is sufficient for inhibition of *zorO*-mediated toxicity. ... 70
Figure 2.6. The V1 region is responsible for target discrimination by OrzO 71
Figure 2.7. OrzP expression levels are equivalent to OrzO expression levels 72
Figure 2.8. Restoration of the base pairing interactions between OrzO V1 to *zorO* prevents *zorO* toxicity 73
Figure 2.9. A specific residue within the V1 region of OrzO is necessary for *zorO* repression 74
Figure 2.10. Expression comparison of OrzO mutants to wild type OrzO..... 76
Figure 2.11. The ability of OrzO 9(1)8 to repress *zorO*-induced toxicity is variable..... 78
Figure 2.12. The minimal level of OrzO needed for repression and the minimal amount of *zorO* that can cause toxicity..... 80
Figure 2.13. Only 15 nucleotides of continuous base pairing are required for OrzO inhibition of *zorO*..... 82
Figure 2.14. Repression of *zorO* requires 17 nucleotides of discontinuous base pairing. 84
Figure 2.15. Predicted secondary structures of *zorO* mRNA and OrzO indicate the V1 regions are mostly single-stranded..... 86
Figure 3.1. The 5' processed form of *zorO* (*zorO* Δ28) shows enhanced translation..... 112
Figure 3.2. The 5' UTR-mediated translation effects can be transferred to *gfp* 114
Figure 3.3. The 5' processing of *zorO* does not alter RBS structure but exposes a single-

stranded region (EAP region) 116
Figure 3.4. Disruption of the EAP region of *zorO* leads to decreased translation..... 121
Figure 3.5. Exposing the RBS rescues the translation defect in *zorO* Δ82..... 123
Figure 3.6. The 5' UTR contains another region that impacts *zorO* translation..... 124
Figure 3.7. OrzO sRNA inhibits translation of *zorO* transcripts 127
Figure 3.8. The +35 to +50 region in *zorO* full-length and zorO Δ28..... 128
Figure 4.1. Toxicity caused by *zorO* overexpression is due to ZorO protein activity.... 145
Figure 4.2. Overexpression of *zorO* leads to rapid membrane depolarization 146
Figure 4.3. Neutralizing the charges at 16th and/or 23rd positions of ZorO impairs the toxicity of ZorO 147
Figure 4.4. Reversing the charges at 16th and/or 23rd positions of ZorO impairs the toxicity of ZorO 148

CHAPTER I

Introduction

Publication note

A version of the section III (III. Biological roles of type I gene pairs), the table, and all figures of this chapter was originally published by Jia Wen and Elizabeth M. Fozo:

Wen, J.; Fozo, E.M. sRNA Antitoxins: More than One Way to Repress a Toxin. *Toxins* **2014**, *6*, 2310-2335.

My contribution to this paper was the literature review and shared writing with Dr. Elizabeth Fozo.

I. An overview of bacterial toxin-antitoxin systems

Toxin-antitoxin (TA) systems have been identified on plasmids and chromosomes of numerous bacteria and archaea. In bacteria, these systems are involved in stress response [reviewed in (1)]. In response to different environmental stimuli, the TA systems can be activated and participate in diverse cellular processes including DNA replication, macromolecular synthesis, membrane disruption, phage infections, and cytoskeletal polymerization, therefore facilitating bacteria to cope with the ever changing environment [reviewed in (2,3)]. Each TA locus is composed of a set of two closely linked genes: one encodes a toxic protein whose overproduction leads to cell growth stasis or cell death, and the other encodes an antitoxin that counteracts the effects of the toxin. Depending on the nature of the antitoxin as well as the modes of action by which the antitoxin represses the toxin, the TA systems can be classified into five types.

The type I TA system contains a stable toxin protein and a labile antitoxin small RNA (sRNA). The antitoxin sRNA possesses sequence complementarity to the toxin mRNA that allows it to base pair with the toxin mRNA. Base pairing by the sRNA can lead to degradation of the mRNA and/or reduced translation of the mRNA, hence inhibiting toxin production.

In the type II system, both the toxin and the antitoxin are proteins. Most type II loci identified to date are arranged in an operon with the antitoxin gene located upstream of the toxin gene. Typically, the type II toxin functions as an endonuclease and the antitoxin neutralizes its activity by binding with the toxin protein [reviewed in (4)]. In addition, transcription of the TA operon can be autoregulated upon binding of the antitoxin or the antitoxin-toxin complex to the promoter, adding another layer of regulation (5-7).

Like the type I system, the type III TA system also is comprised of a protein toxin and a RNA antitoxin. However, repression by the antitoxin occurs through a distinct mechanism. The type III antitoxin gene is composed of a stretch of short tandem repeats and separated from the toxin gene by a weak terminator. Instead of base pairing to the toxin mRNA, the antitoxin RNA forms pseudoknots and directly binds to the toxin protein, sequestering the protein thereby blocking its activity (8). All type III TA systems identified to date are involved in phage resistance by inducing premature bacterial cell death upon phage infection [reviewed in (9)].

The type IV TA system is composed of a protein toxin and a protein antitoxin. Unlike type II loci, the toxin and the antitoxin of the type IV system do not directly interact. Rather, the antitoxin counteracts the effects of the toxin by acting on the targets of the toxin. The *cbtA/cbeA*

gene pair of *Escherichia coli* is the sole member of this type. The CbtA toxin inhibits the polymerization of the bacterial cytoskeletal proteins MreB and FtsZ. On the contrary, the antitoxin CbeA protein directly interacts with MreB and FtsZ and promotes bundling of FtsZ and MreB protofilaments, therefore enhancing polymerization (10).

To date, only one gene pair of the type V TA system has been identified. This type V gene pair *ghoT/ghoS* was discovered in the chromosome of *E. coli* (11). The GhoT toxin is a membrane lytic peptide whose overproduction causes ghost cell formation, whereas the antitoxin GhoS protein is a RNase that specifically cleaves the *ghoT* mRNA and renders it untranslatable.

Bacteria can harbor multiple different types of TA loci. For instance, at least 33 TA loci have been identified in lab strain *E. coli* K12 [reviewed in (3)]. Further, the pathogen *Mycobacterium tuberculosis* has more than 60 TA loci, with many of the loci duplicated pairs [reviewed in (12)]. Duplication of TA gene pairs is common, but it is not clear why bacteria may possess redundant loci. Homologs of bacterial TA systems have also been discovered through genome mining in many archaea (13). Additionally, BLAST searches against fungal genomes have identified some toxin homologues (14). However, whether these homologues are functional and whether their antitoxin counterparts are present have not been examined. Given the prevalence of TA systems, it would be interesting to know what potential advantages the TA systems can provide to organisms. This is particularly relevant with respect to evolution of the organisms across the three domains of life.

II. Repression of the type I toxin by the antitoxin sRNA

The type I TA systems were initially identified on plasmids where they confer plasmid maintenance through a post-segregational killing mechanism, as exemplified by the *hok/sok* of the R1 plasmid of *E. coli* and the *fst*/RNAII of the pAD1 plasmid of *Enterococcus faecalis* (15-17). Homologs of plasmid encoded type I TA systems as well as new type I loci have since been discovered in numerous bacterial chromosomes. Most type I toxins identified to date are small proteins that are less than 60 amino acids and contain one putative transmembrane domain. Owing to their hydrophobic nature, these toxins are proposed to form pores in the bacterial cell membrane and impair ATP synthesis, which eventually leads to cell growth stasis or cell death, similar to phage holins [reviewed in (18)]. To avoid such detrimental effects caused by type I toxins, their

levels must be carefully controlled. One mechanism to modulate the production of a type I toxin is through base pairing by its corresponding antitoxin sRNA.

Unlike the *trans*-encoded sRNAs that are located distal from their targets and act through limited base pairing, the type I antitoxin sRNAs are encoded antisense to the toxin genes and share extensive sequence complementarity with the toxin mRNAs [reviewed in (19)]. The genetic organizations of the type I TA loci discussed in this section are listed in Table 1.1. Due to their greater base pairing potential, most type I sRNAs can repress their cognate toxin mRNAs without the protein Hfq, a RNA chaperone required for the function of many identified *trans*-encoded sRNAs.

For some pairs, base pairing between the antitoxin sRNA and the toxin mRNA happens in a stepwise fashion: there is an initial contact between single-stranded regions of each RNA, then extends to the flanking sequences of the complimentary region or stimulates binding through other distal portions of the RNAs [reviewed in (9)]. This has been demonstrated by the *ibsC/sibC* type I locus in *E. coli* (20). The SibC sRNA first interacts with the *ibsC* mRNA through two isolated complimentary regions, TRD1 and TRD2. The base pairing then extends to the flanking region and leading to pairing of over 70 nt. Another more complicated example is the type I locus *fst*/RNAII. In this case, the RNAII sRNA possesses three distinct regions of complementarity to the *fst* mRNA: a 3' termination loop region that can base pair to the 3' end of *fst*, and two 5' direct repeats, DRa and DRb, that can base pair to the 5' end of the *fst* (21). Base pairing initiates via interactions between the complementary regions in the 3' terminator loops of both RNAs, resulting in the formation of a "kissing complex". This brings the 5' ends of the RNAII sRNA and the *fst* mRNA close to each other, which subsequently promotes a second base pairing event through the two direct repeats on RNAII and the corresponding sites at the 5' *fst* (22). Afterwards, the base pairing progresses to the complementary region in between the DRa and DRb sites, forming a stable *fst*-RNAII duplex.

Given the extensive base pairing between the sRNA and the toxin mRNA, one would imagine that base pairing will result in structural changes of both RNAs, as the previously structured complimentary sequences are now "occupied". Depending on whether the ribosome binding site (RBS) structure is involved, the effects of such structural changes may vary. In the type I pair *bsrG/sr4* of *Bacillus subtilis*, binding by SR4 sRNA to the 3' end of the *bsrG* mRNA leads to structural rearrangement of the *bsrG* RBS (Figure 1.1) (23). More specifically, the *bsrG*

RBS is located in a 4 nt-long stem and the upstream sequences of this stem are originally paired with part of the *bsrG* complementary region. Upon SR4 base pairing, the complementary region on *bsrG* is occupied, therefore releasing the sequences upstream of the RBS-containing stem. This leads to a structural rearrangement and pairing such that the length of the RBS-containing stem extends to 8 nt. The increased stem length was shown to impede translation, therefore preventing BsrG toxin production. To date, SR4 sRNA is the only type I antitoxin that base pairs to the 3' end of the toxin mRNA but inhibits toxin translation through impacting the distal RBS structure.

Aside from affecting RNA structure at a distal region, a more common outcome of base pairing seen in type I pairs is the interruption of ribosome binding to the "local" mRNA. This typically leads to impaired translation of the toxin mRNA; however, the detailed mechanisms can vary depending on where base pairing occurs. Some type I antitoxin sRNAs base pair over the translation initiation region (TIR) of the toxin mRNAs and block translation, as demonstrated by the SymR sRNA of the *symE/symR* pair (24,25), the SibC sRNA of the *ibsC/sibC* pair (Figure 1.2) (26), and the RalA sRNA of the *ralR/ralA* pair of *E. coli* (27). On the contrary, some type I antitoxin do not interfere with the TIR of the toxin, but instead the TIR of a leader peptide. One classic example is the *hok/sok* locus of the R1 plasmid (Figure 1.3) (15). The 5' untranslated region (UTR) of *hok* harbors a small open reading frame (ORF) termed *mok* whose translation is coupled with *hok* (16). The Sok sRNA base pairs to the *hok* 5' UTR but does not bind with its RBS. Rather, it overlaps with the TIR as well as part of the coding region of *mok*, leading to translation inhibition of both *mok* and *hok*. Similar to that of *hok/sok*, base pairing by the IstR-1 sRNA of the *tisB/istR-1* locus in *E. coli* also occurs at the 5' UTR of *tisB* distal to the *tisB* RBS (28). However, regulation of the TisB production mediated by the IstR-1 sRNA is conducted through a different mechanism in that the sRNA blocks a ribosome standby site on the mRNA (Figure 1.4) (29). More specifically, the *tisB* RBS is sequestered in a stem structure that prevents the ribosome from directly accessing the RBS. Instead, the ribosome first preloads onto a standby site upstream of the RBS at the *tisB* 5' UTR and then moves to the RBS when the RBS opens. This standby site is required for translation and overlaps the region of base pairing for the IstR-1 sRNA. Therefore, base pairing of IstR-1 to the standby site interferes with ribosome preloading, leading to translation inhibition of *tisB* mRNA.

Regardless of where the base pairing occurs, it often induces degradation of the mRNA. Typically, upon base pairing, the double-stranded specific RNase III cleaves the sRNA-mRNA

duplex, rendering the mRNA untranslatable. For example, base pairing by the IstR-1 sRNA results in a processed form of *tisAB* (+106 *tisAB*) that was proved to be translationally inactive (29). Further, RNase III cleaves the *bsrG* toxin mRNA at the complementary region after *bsrG* is paired with the SR4 antitoxin sRNA (30). In some cases, stimulating degradation of the toxin mRNA is the main mechanism used by the antitoxin sRNA to inhibit toxin production, particularly when base pairing by the sRNA solely occurs at the 3' end of the toxin mRNA. One such example is the *txpA/ratA* pair of *B. subtilis*, in which the RatA sRNA base pairs to the *txpA* mRNA at its 3' end resulting in a series of cleaved *txpA* transcripts that are subsequently degraded by PNPase (Figure 1.5) (31,32). In addition to *txpA/ratA*, three other type I loci in *B. subtilis*, the *bsrE/SR5* and the *bsrH/as-bsrH*, use the same strategy to modulate production of the toxin (32-34). It has been shown that the loss of cell viability seen in *B. subtilis* upon RNase III encoding gene deletion is caused by de-repression of two type I toxins including TxpA (32), which further reveals the essential role of RNase III in regulating toxin production of these type I loci.

Although in many described type I loci the sRNA can stimulate toxin mRNA degradation, two exceptions have been observed. One is the *fst*/RNAII pair in which the formation of the RNA duplex greatly increases the stability of both RNAs: the half-life of *fst* mRNA showed a 1.6- to 2.2-fold increase and the half-life of RNAII increased from 4 min to 16 min (35). The other is the pair *ralR/ralA* discovered on the *rac* prophage of *E. coli*. In this case, there was no substantial reduction in *ralR* mRNA levels following interaction with RalA, as detected by qRT-PCR (27). Interestingly, in both cases, the small RNA is located antisense to the 3' end of toxin mRNA but shares additional base pairing complementarity with the 5' end of the toxin mRNA. As described above, complete base paring between RNAII and *fst* occurs at both the 5' end and the 3' end of the mRNA, whereas base pairing in most other type I systems occurs only at one end. Perhaps the more extensive base pairing of *fst*/RNAII accounts for the resistance of the *fst* transcript to degradation; however, more experimental evidence is needed to make such assumption. Another fact worth noting is that the studies of *fst*/RNAII and *ralR/ralA* did not compare the stability of the mRNA in a wild type versus a RNase III and/or other double-stranded specific RNases deficient background. Therefore, it is possible that the RNase III dependent degradation still occurred, but its effect was hidden by a greater increase of mRNA stability caused by formation of the RNA duplex. It would be interesting to see how the half-life of *fst* or *ralR* changes upon RNase III deletion.

As mentioned in the previous section, for those instances where it has been examined, the toxin mRNA is relatively stable whereas the antitoxin sRNA usually has a very short half-life. For instance, the half-life of the *hok* mRNA of the *hok/sok* pair is ~ 30 minutes whereas the half-life of Sok is ~ 30 seconds (36). This adds an additional layer to toxin repression given that despite successful base pairing, the time window during which one sRNA is "alive" may be too short for the sRNA to exert its inhibitory effects. So, how does antitoxin repress its toxin given this shortcoming? it has been suggested that the antitoxin sRNA is in molar excess over the corresponding toxin mRNA. In the aforementioned *hok/sok* pair encoded on the R1 plasmid, both genes are constitutively expressed; however, the Sok is transcribed from a stronger promoter than *hok* (37). Hence, *hok* is outnumbered by Sok as long as the R1 plasmid is present to generate Sok. This compensates for the short half-life of Sok and allows successful repression of the *hok* mRNA. Similar observations were shown in the chromosomally encoded type I pairs *bsrG/sr4* and *bsrE/sr5* in *B. subtilis*, in which the antitoxin promoters are far stronger than the toxin promoters (30,38). Particularly, the stronger promoter of *sr5* was shown to lead to a 7-8 folds higher levels of SR5 sRNA than the *bsrE* mRNA under normal growth conditions. In other cases, the difference in the amount of the sRNA and the mRNA is due to transcriptional repression of the toxin, as is seen in the chromosomally encoded loci *tisB/istR-1* and *symE/symR* in *E. coli*. Both loci are involved in SOS response and each toxin gene is subject to LexA repression (28,39). On the contrary, the *istR-1* or *symR* is constitutively transcribed, leading to repression of the corresponding toxin mRNA (24).

To summarize, type I antitoxins repress the toxins production via base pairing, but the direct outcomes of base pairing can vary. Base pairing can lead to translation inhibition of the toxin mRNA and/or degradation of the toxin mRNA through several distinct mechanisms. Additional regulation is conferred through differences in RNA stability and abundance throughout growth conditions: although the sRNA is not as stable compared to the toxin mRNA, it is usually expressed in excessive amount over the toxin mRNA, thereby rendering successful repression.

III. Biological roles of type I gene pairs

To date, numerous examples of type I loci have been described in bacterial chromosomes. However, whether these chromosomal copies have a true biological function is still questioned. Given that the first loci were described on plasmids and serve to stabilize plasmids, many assumed that these

loci are "selfish" DNA elements. Yet, bioinformatic analyses indicate that many toxin-antitoxin pairs have no clear homology to mobile genetic elements nor is there evidence for horizontal gene transfer (40). Also, many toxin gene sequences are highly conserved with strong predicted ribosome binding sites and little evidence of sequence degeneration, suggesting that bacteria are maintaining these genes. This would support the argument that, at least for some chromosomally encoded loci, they do indeed possess a true biological function.

Plasmid remnants? The *hok*/*sok* locus of plasmid R1 was the first type I toxin-antitoxin gene pair to be described. Upon cell division, if a daughter cell does not inherit a copy of the plasmid, the unstable Sok antitoxin will be degraded, allowing the stable *hok* mRNA to be translated, and kill the plasmid-less cell. Specifically, the Hok toxin oligomerizes in the inner membrane, causing pore formation, and loss of viability (16). This is particularly evident with the appearance of "ghost" cells upon Hok overproduction. Hok though appears to act internally as application of exogenous Hok to cells could not induce cell killing; in fact, synthetic Hok could only kill cells if electroporated into the cells (41). Combined, these data supported that the function of the plasmid encoded *hok*/*sok* locus is to maintain the plasmid within the bacterial population.

With the advent of genomic sequencing, *hok/sok* loci were found within the genomes of numerous *E. coli* strains and other Gram-negative bacteria; for some *E. coli* strains examined, there are as many as 12 copies of the locus (40,42). In some instances, there is evidence of sequence insertions or sequence degeneration, and it is likely that these loci no longer produce the toxin or antitoxin. However, a large number appear to be intact and are potentially expressed (42). So, what are these chromosomal copies doing? Are they simple remnants of past plasmids? Do they serve to maintain plasmids within the population or maintain chromosomal stability in the population? Given that there are numerous copies within a single species and some loci appear degenerate, these questions are not trivial to address. Future work will hopefully resolve whether or not chromosomally encoded *hok*/*sok* gene pairs have a true biological function.

Impairment of chromosomal structure As described above, the *fst*/RNA II locus of the pAD1 plasmid of *E. faecalis* confers plasmid addiction through post-segregational killing (17). Like Hok and many other type I toxins, the Fst toxin also contains a putative transmembrane domain. When overexpressed, it can cause cell membrane permeabilization and cessation of macromolecular

synthesis (43). However, unlike Hok, overproduction of Fst does not result in the formation of ghost cells [reviewed in (44)]. A recent study revealed that the Fst-resistant *E. faecalis* M7 strain harbors a mutation in the *rpoC* gene encoding the β subunit of RNA polymerase (45). Normally, an Fst-sensitive *E. faecalis* strain will induce the expression of a variety of membrane transporters upon *fst* overexpression; however, the M7 strain did not. This suggested that Fst induces membrane transporters and that induction may deplete the cell of energy. To validate this, the authors treated wild type *E. faecalis* with reserpine, an inhibitor of transporters, and then induced *fst* in the treated strain. Upon treatment with reserpine, wild *E. faecalis* survived Fst toxicity far better than without (45). Thus, the induction of membrane transporters is a major cause of Fst-induced toxicity.

Intriguingly, Fst induced membrane damage occurs late upon Fst overproduction (46). This indicated that the loss of membrane integrity may not be the primary effect of Fst toxin. Studies have revealed that the primary effect of Fst overexpression is condensation of the nucleoid. The effects on nucleoid structure have been observed not only in *E. faecalis*, but also *Staphylococcus aureus*, *B. subtilis*, and *E. coli*; although how Fst disrupts nucleoid structure remains unclear (43,46,47). Interestingly, bioinformatic analysis concluded that both Fst and Ldr, a chromosomally encoded type I toxin of *E. coli* (26), are within the same toxin superfamily (40). Overexpression of LdrD led to a similar nucleoid condensation in *E. coli*, suggesting that this is a conserved feature for this family (26).

Like the *hok*/*sok* locus, numerous homologs to *fst*/RNAII have been identified in many bacterial chromosomes (40,47). Given the above effects on membrane transportation upon Fst overproduction, it is interesting to note that several chromosomal *fst*/RNAII loci are flanked by transporters, specifically, carbohydrate transporters (47). Also, for those examined, both the toxin and antitoxin are transcribed, indicative that they are likely functional. What their biological role may be is not clear. For a homolog identified in *Streptococcus mutans*, described below, there are links to persister cell formation (see below) (48). For a homolog found on the chromosome of *E. faecalis* V583, deletion of the antitoxin component (RNAII) led to increased virulence in a mouse model, though how *fst* could contribute to virulence is not known (49). It is important to note that even very low levels of Fst production can lead to the formation of suppressors, so further examination of this RNAII deletion strain is needed. Interestingly, a homolog of *fst*/RNAII in *S. aureus* indicates that expression of the toxic peptide is induced upon cellular stress, suggesting that it may have a role in responding to changing environments (50).

Persister formation Although a bacterial population maybe genetically identical, cellular growth and gene expression are not equivalent across the population of cells. A small portion of cells within a population are non-growing or slow-growers; these "dormant" cells are termed persisters (51,52). Given that most antibiotics target actively growing cells, persisters are often highly antibiotic resistant (53). The *tisB/istR-1* locus has been linked with persister formation. Specifically, the antibiotic ciprofloxacin targets bacterial DNA gyrase, leading to DNA damage and induction of the SOS response (54). The type I toxin gene *tisB* is induced by SOS damage (28). Deletion of the entire *tisB*/*istR1* locus in *E. coli* led to a decrease in persisters tolerant to ciprofloxacin. Yet a strain deleted for only the antitoxin IstR-1 exhibited a 10- to 100-fold increase in the level of persisters (55,56). This indicated that TisB, when expressed at its endogenous level, plays an essential role in persister formation.

How then could a small hydrophobic protein lead to persister formation? A previous study demonstrated that an epitope tagged version of TisB localized to the inner membrane of *E. coli* (57). Additionally, work with synthetic TisB demonstrated that TisB monomers can rapidly and spontaneously bind to membranes (58). It is thought that these monomers within the membrane may form dimers that would allow protons to pass or alternatively that TisB monomers aggregate to form pores that are selective for anions (59). Regardless, production of TisB and its insertion in the membrane would cause a loss of proton motive force and a decline of intracellular ATP. This was verified in vivo when, overproduction of TisB led to a dramatic loss of ATP within *E. coli* within 5 minutes (57). Along with this, ciprofloxacin requires ATP for its function, thus expression of TisB would cause a decrease in ATP levels leading to both an inhibition of macromolecule synthesis and an inhibition of the antibiotic's activity. To summarize, insertion of TisB into the membrane would cause a loss in proton motive force, depletion of ATP levels for ciprofloxacin activity and macromolecule synthesis, and the consequential increase in the number of persisters within the population (55,56,60).

This consequential decrease in cellular metabolism caused by TisB is similar to what has been reported for type II toxins. Type II toxin proteins are often "larger" (~100 amino acids) than type I toxins and are not particularly hydrophobic. They have defined biochemical activities with most examples functioning as either ribonucleases or inhibitors of DNA gyrase (61). It was hypothesized that by induction of type II toxin activity (specifically ribonuclease activity), cellular translation would be halted, giving rise to the "dormant" persister cell phenotype. This is supported

by the observation that deletion of multiple type II toxin-antitoxin loci in *E. coli* reduced the level of persisters (62). Further, a recent study utilizing chemical inhibitors of macromolecular synthesis provided additional evidence: by decreasing macromolecular synthesis chemically, the frequency of persister cell formation was far higher over control populations that were not treated (63). Thus, type II toxins and TisB appear to induce persister formation by reducing cellular metabolism through the inhibition of macromolecular synthesis. It is critical to note, however, that most identified type I toxins have not been examined to conclude whether or not they participate in persister formation, but such analysis would be of great interest.

The studies with TisB illustrate another important point: the critical balance of toxin levels within a cell. Overproduction of TisB clearly causes cell death (28,57,64), yet in some instances, it can lead to persister formation (55). In particular, analysis suggests that the expression of *tisB* is induced 1000-fold under SOS conditions yet this does not induce cell killing (55,57). Thus, the fine-tuning of toxin and antitoxin levels is critical for persister formation. When the toxin level is below that of the antitoxin, the antitoxin successfully represses toxin expression, and the cell is unharmed. If the toxin level is higher than that of the antitoxin, the toxin escapes repression and impairs cell growth. Depending upon how excessive toxin levels are to the levels of the antitoxin will determine whether the cell becomes dormant (*i.e.*, persister) or is killed. When the levels of the toxin and the antitoxin are close, some cells will have an imbalance in the ratio of toxin: antitoxin. Owing to the stochastic nature of bacterial populations, this small portion of the population will become persisters (65). The exact amounts of TisB needed to form ciprofloxacin-induced persisters is currently not known but will be of great interest in understanding the balance between the benefits and costs of possessing this toxic gene.

In the Gram-positive bacterium *S. mutans*, a homolog to the *fst*/RNAII locus of pAD1 has been linked to persister cell formation (48). Deletion of this locus, known as Fst-Sm/srSm, did not impact persister formation. However, the number of persisters was greatly reduced upon treatment with cell-wall damaging antibiotics in a strain harboring the entire locus on a multicopy plasmid (48). Unlike *tisB*, this observation appears to be due to increasing the number of Fst-Sm/srSm copies within the cell; analysis has not revealed whether or not this observation can be attributed either to the toxin, the antitoxin, or just increased copies of the entire locus. How then, would increasing the number of copies of this locus lead to decreased persister cell formation? This could be due again to levels of the toxin *versus* the antitoxin; by increasing the number of loci within the

cell, there could be an imbalance in the ratio of toxin: antitoxin, leading to heighten cell killing and decrease in persister formation. A detailed analysis of toxin and antitoxin RNA and toxin protein levels across the cellular population will be needed to further conclude how this gene pair impacts persister cell formation.

Chromosomal stabilization and recombination In *B. subtilis*, many of its predicted chromosomally encoded type I loci are located within integrated mobile elements. For instance, the *txpA*/*ratA* and *bsrH*/*as-bsrH* loci are located within the *skin* element (31,66,67), while *bsrG*/*sr4* and *yonT*/*as-yonT* are within the SPβ prophage (30,40). Additionally, the newly identified *bsrE/sr5* was found in the prophage-like region P6 (33). Overexpression of these toxin genes slows cell growth or causes cell lysis, while deletion of these genes does not lead to noticeable consequences.

The presence of these gene pairs on mobile genetic elements suggests that they may maintain these elements within the chromosome. This maintenance may then provide the cell with a selective advantage under specific environmental conditions. For example, studies of the SPβ prophage suggest that it contains genes that are beneficial to *B. subtilis*. The gene *sspC* encodes a small acid-soluble protein that provides high UV light resistance to spores and is found on the SPβ prophage (68). The gene *nonA*, examined in the SPβ prophage region of the *B. subtilis* Marburg strain, encodes a protein that protects cells against infection from the bacteriophage SP10 (69). Perhaps the two toxin-antitoxin loci then contribute to maintaining this element within the population. Interestingly, the BsrG type I toxic protein that is encoded on the SPβ prophage is temperature-sensitive. Rapid degradation of *bsrG* mRNA was observed at 48°C, indicating a potential role of this toxin in response to changing temperatures (30).

However, some prophage regions may simply act as selfish elements in the chromosome. One possible example for this theory is the *skin* element of *B. subtilis*. This large element (48 kb) is located within the *sigK* gene encoding the RNA polymerase sigma factor σ^K (70). During sporulation, *skin* is excised from the chromosome, creating an intact *sigK* gene. This excision only occurs from the chromosome of the mother cell, and not from the chromosome of the forespore. An engineered strain with the *skin* element deleted was able to grow and sporulate normally, suggesting that this region does not play an essential role in viability or sporulation (70). Thus, is *skin* simply a piece of selfish DNA? Within the skin are two different type I toxin-antitoxin loci: *txpA*/*ratA* and *bsrH*/*as-bsrH*. Thus, if the *skin* is lost, expression of those toxins could be lethal for

the cell, helping to maintain the element within the population. However, a study did identify genes conferring arsenate and arsenite resistance within the *skin* element (71), implying that there are genes of value to *B. subtilis* within this region. Perhaps a more detailed analysis of the genes encoded with the *skin* element could determine if the function for *txpA*/*ratA* and *bsrH*/*as-bsrH* is to maintain a "useful" piece of mobile DNA.

It is also possible that these type I modules of *B. subtilis* have functions other than stabilizing the chromosomal regions. several type I toxin encoding genes, such as *brsG*, *brsE* and *brsH*, have putative ResD response regulator binding sites upstream, indicating that their toxin products may participate in response to oxygen limitation (34). Durand et al noted that the described toxins and antitoxins are all under control of the vegetative sigma factor σ^A; thus all could be quickly induced (34). Further analysis of the balance between the expression of the toxins and antitoxins will provide much-needed evidence for their physiological roles.

In *E. coli*, the newly described *dinQ/agrB* locus may also play a role in chromosomal stability (72). Transcription of the *dinQ toxin*, like *tisB* and *symE*, is also regulated by LexA, and is induced upon DNA damage. Deletion of the antitoxin *agrB* led to elevated DinQ levels; however, this deletion did not directly impact SOS activities. Instead, a strain deleted for *agrB*, had a 400-fold reduction in recombination frequency, suggesting that DinQ interferes with recombination (72). The small toxin is found in the membrane, and its overproduction can lead to decreased ATP levels, like TisB, as well as increased nucleoid condensation, like Fst and LdrD. Further work to elucidate how these phenotypes are linked to DinQ biochemical activity will be of great interest.

Inhibition of competitors The *txpA*/*ratA* homologous locus, *sprG1*/*sprF1*, in *S. aureus* was recently described (73). Like the original *txpA*/*ratA* locus, the locus characterized was found on a mobile genetic element, ΦSa3 PI, a phage integrated within the chromosome; however, additional copies were predicted within the core genome (31,32,73). The *sprG1* gene encodes two toxic peptides from two in-frame initiation codons, and the *sprF1* gene encodes an antitoxin sRNA. Interestingly, we note that even the original member of this family, *txpA* of *B. subtilis*, may produce two peptides [see (31)]. For *S. aureus*, overproduction of either the long or short SprG1 peptide inhibited cell growth and caused cell death; however, co-expression of the SprF1 antitoxin could repress toxicity induced by either toxic peptide (73). Both SprG1 peptides contain a putative transmembrane domain and both proteins were detected in the supernatant. Given this, the authors

investigated whether or not external addition of either peptide could be toxic to cells. Application of synthetic forms of SprG1 inhibited the growth of *E. coli*, *Pseudomonas aeruginosa* and *S. aureus*, as well induced lysis of eukaryotic cells (73). Interestingly, the longer synthetic peptide was more effective against human erythrocytes and the shorter variant more active against bacterial isolates. While this is the first type I toxin shown to be able to induce toxicity when applied externally to cells, more analysis will be needed to confirm that sufficient SprG1 is produced and released from cells to cause killing under native, endogenous conditions.

Nucleic acid cleavage Almost all type I toxins described to date have a putative transmembrane domain that is thought to contribute to membrane damage upon their production. However, both SymE and RalR do not follow this paradigm. Overproduction of both SymE (SymE/SymR) and RalR (RalR/RalA) toxins were toxic to *E. coli*, yet co-expression of their cognate antitoxin (SymR or RalA) could prevent this toxicity (25,27). SymE, is a rather "large" protein for a type I toxin at 113 amino acids in length. SymE acts as a ribonuclease and its overproduction triggers mRNA degradation [for a more thorough discussion of the unusual homology of SymE see (25,74)]. Interestingly, SymE, like TisB and DinQ, is also induced by DNA damage and under the control of the LexA promoter. Thus, three type I toxins of *E. coli* are triggered by the SOS response, signifying the importance of DNA stability and repair for cell survival.

The RalR toxin is encoded on the *rac* prophage in the genome of *E. coli* along with its cognate antitoxin-encoding gene *ralA* (27). RalR was shown to cleave DNA in an *in vitro* assay, but not RNA, suggesting it functions as a DNase. Deletion of either *ralR* alone or *ralR*/*ralA* locus in *E. coli* resulted in greater sensitivity towards fosfomycin, an antibiotic that inhibits peptidoglycan biosynthesis (27). It will be interesting to see how the DNase activity of RalR could lead to increased resistance to a cell wall synthesis inhibitor.

IV. Perspectives

Despite our growing understandings of the bacteria type I TA systems, many questions remain to be answered. For example, the detailed mechanisms underlying the versatile roles of the type I toxins are largely unclear. Additionally, it is not known whether an antitoxin sRNA can regulate other target(s) besides the cognate toxin gene. This is particularly interesting given the observations that many *trans*-encoded sRNAs use different regions of the sequences to regulate

expression of multiple genes [reviewed in (75)]. Moreover, the evolution of the type I loci as well as potential functional connections between type I and other types of TA systems are fascinating areas to be examined.

V. Research objectives

A previous bioinformatics search had identified two highly homologous gene pairs in *E. coli* O157:H7 (EHEC) likely belonging to a type I TA system (40). We renamed these two loci as *zorO-orzO* and *zorP-orzP*. The *zorO* and *zorP* genes encode for small proteins of 29 amino acids, while the *orzO* and *orzP* genes encode for sRNAs of ~80 nts. Since these are newly identified gene pairs, it is unclear if they are true type I TA loci and how the putative toxin genes are regulated. The overarching goal of this dissertation was to characterize the gene pair *zorO-orzO* and to understand the regulation of *zorO* gene in EHEC.

When beginning this project, I noted that the OrzO sRNA shares extensive sequence complementarity to the *zorO* mRNA, indicating that these two RNAs can base pair. Base pairing of the mRNA by the sRNA is a key feature of a type I TA locus. Therefore, I first confirmed that the *zorO-orzO* gene pair is a bona fide type I system. Since there are 18 nts of perfect base pairing potential between the OrzO sRNA and the *zorO* mRNA, I wanted to know if all 18 nts of base pairing are needed for OrzO to repress *zorO*. Using mutational studies, I demonstrated that successful recognition of the *zorO* mRNA by the OrzO sRNA requires a minimal of 15 nts of perfect base pairing potential. Further, I showed that base pairing by the OrzO sRNA can stimulate degradation of the *zorO* mRNA in a RNase III-dependent fashion, leading to reduced ZorO production. Hence, the *zorO* gene can be regulated at the post-transcriptional level by the OrzO sRNA. These results are presented in chapter 2.

Along with this, I wanted to know if repression of *zorO* by OrzO solely relies on mRNA degradation. A close examination of the *zorO* mRNA indicated that it possesses a long 5' UTR of 174 nts where the OrzO sRNA base pairs. Specifically, the base pairing occurs at a region located ~60 nts upstream of the *zorO* RBS in this long 5' UTR. Given this, I asked why the *zorO* 5' UTR was so long and whether base pairing by OrzO could impact *zorO* translation despite the fact that the base pairing site is far from the translation initiation region of *zorO*. To address these questions, I determined the roles of *zorO* 5' UTR using a combination of *in vivo* and *in vitro* assays. I showed that the 5' UTR could modulate *zorO* translation through two independent regions, both of which

are required for *zorO* to achieve optimal translational efficiency. Interestingly, one of the region overlaps the base pairing site of *zorO*. Hence, binding by the OrzO sRNA at this site interferes with *zorO* translation. These findings expanded the work shown in chapter 2 and indicated that the *zorO* gene can also be regulated at the translational level by both its own 5' UTR and the OrzO sRNA. These results are presented in chapter 3.

Given my findings that the *zorO* gene is regulated at multiple levels, I was curious as to why so many layers of regulation are needed to control ZorO production. Although *zorO* overexpression causes bacterial growth stasis or cell death, the mechanism underlying the toxic nature of ZorO remained elusive. Predictive tools have suggested that ZorO can target the cytoplasmic membrane; however, this had never been experimentally tested. Indeed, I found that overproduction of ZorO leads to rapid membrane depolarization, confirming that it functions at the cell membrane. Interestingly, despite its hydrophobicity, the ZorO protein contains several charged amino acids and two of them are predicted in the putative transmembrane domain of ZorO. I thus focused on the roles of these two charged amino acids and demonstrated that they are critical for ZorO's toxicity. These findings provided much-needed information on the biochemical properties of the ZorO toxin and shed light on elucidating its biological function. These results are presented in chapter 4.

Taken together, the work presented in this dissertation expands the current knowledge of bacterial type I toxin-antitoxin systems and provides new insights into the regulatory mechanisms controlling the production of the type I toxin ZorO in EHEC.

References

1. Bukowski, M., Rojowska, A. and Wladyka, B. (2011) Prokaryotic toxin-antitoxin systems--the role in bacterial physiology and application in molecular biology. *Acta Biochim Pol*, **58**, 1-9.
2. Wen, Y., Behiels, E. and Devreese, B. (2014) Toxin-Antitoxin systems: their role in persistence, biofilm formation, and pathogenicity. *Pathog Dis*, **70**, 240-249.
3. Yamaguchi, Y., Park, J.H. and Inouye, M. (2011) Toxin-antitoxin systems in bacteria and archaea. *Annu Rev Genet*, **45**, 61-79.
4. Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Dreze, P. and Van Melderen, L. (2011) Diversity of bacterial type II toxin-antitoxin systems: a comprehensive search and functional analysis of novel families. *Nucleic Acids Res*, **39**, 5513-5525.
5. Kedzierska, B., Lian, L.Y. and Hayes, F. (2007) Toxin-antitoxin regulation: bimodal interaction of YefM-YoeB with paired DNA palindromes exerts transcriptional autorepression. *Nucleic Acids Res*, **35**, 325-339.
6. Li, G.Y., Zhang, Y., Inouye, M. and Ikura, M. (2008) Structural mechanism of transcriptional autorepression of the Escherichia coli RelB/RelE antitoxin/toxin module. *J Mol Biol*, **380**, 107-119.
7. Overgaard, M., Borch, J. and Gerdes, K. (2009) RelB and RelE of Escherichia coli form a tight complex that represses transcription via the ribbon-helix-helix motif in RelB. *J Mol Biol*, **394**, 183-196.
8. Blower, T.R., Short, F.L., Rao, F., Mizuguchi, K., Pei, X.Y., Fineran, P.C., Luisi, B.F. and Salmond, G.P. (2012) Identification and classification of bacterial Type III toxin-antitoxin systems encoded in chromosomal and plasmid genomes. *Nucleic Acids Res*, **40**, 6158-6173.
9. Brantl, S. and Jahn, N. (2015) sRNAs in bacterial type I and type III toxin-antitoxin systems. *FEMS Microbiol Rev*, **39**, 413-427.
10. Masuda, H., Tan, Q., Awano, N., Wu, K.P. and Inouye, M. (2012) YeeU enhances the bundling of cytoskeletal polymers of MreB and FtsZ, antagonizing the CbtA (YeeV) toxicity in Escherichia coli. *Mol Microbiol*, **84**, 979-989.
11. Wang, X., Lord, D.M., Cheng, H.Y., Osbourne, D.O., Hong, S.H., Sanchez-Torres, V., Quiroga, C., Zheng, K., Herrmann, T., Peti, W. *et al.* (2012) A new type V toxin-antitoxin

system where mRNA for toxin GhoT is cleaved by antitoxin GhoS. *Nat Chem Biol*, **8**, 855-861.

12. Sala, A., Bordes, P. and Genevaux, P. (2014) Multiple toxin-antitoxin systems in Mycobacterium tuberculosis. *Toxins (Basel)*, **6**, 1002-1020.
13. Sberro, H., Leavitt, A., Kiro, R., Koh, E., Peleg, Y., Qimron, U. and Sorek, R. (2013) Discovery of functional toxin/antitoxin systems in bacteria by shotgun cloning. *Mol Cell*, **50**, 136-148.
14. Fico, S. and Mahillon, J. (2006) TasA-tasB, a new putative toxin-antitoxin (TA) system from Bacillus thuringiensis pGI1 plasmid is a widely distributed composite mazE-doc TA system. *BMC Genomics*, **7**, 259.
15. Gerdes, K., Larsen, J.E. and Molin, S. (1985) Stable inheritance of plasmid R1 requires two different loci. *J Bacteriol*, **161**, 292-298.
16. Gerdes, K., Bech, F.W., Jorgensen, S.T., Lobner-Olesen, A., Rasmussen, P.B., Atlung, T., Boe, L., Karlstrom, O., Molin, S. and von Meyenburg, K. (1986) Mechanism of postsegregational killing by the hok gene product of the parB system of plasmid R1 and its homology with the relF gene product of the E. coli relB operon. *EMBO J*, **5**, 2023-2029.
17. Weaver, K.E. (1995) Enterococcus faecalis plasmid pAD1 replication and maintenance. *Dev Biol Stand*, **85**, 89-98.
18. Brantl, S. (2012) Bacterial type I toxin-antitoxin systems. *RNA Biol*, **9**, 1488-1490.
19. Thomason, M.K. and Storz, G. (2010) Bacterial antisense RNAs: how many are there, and what are they doing? *Annu Rev Genet*, **44**, 167-188.
20. Han, K., Kim, K.S., Bak, G., Park, H. and Lee, Y. (2010) Recognition and discrimination of target mRNAs by Sib RNAs, a cis-encoded sRNA family. *Nucleic Acids Res*, **38**, 5851-5866.
21. Greenfield, T.J. and Weaver, K.E. (2000) Antisense RNA regulation of the pAD1 par post-segregational killing system requires interaction at the 5' and 3' ends of the RNAs. *Mol Microbiol*, **37**, 661-670.
22. Greenfield, T.J., Franch, T., Gerdes, K. and Weaver, K.E. (2001) Antisense RNA regulation of the par post-segregational killing system: structural analysis and mechanism of binding of the antisense RNA, RNAII and its target, RNAI. *Mol Microbiol*, **42**, 527-537.

23. Jahn, N. and Brantl, S. (2013) One antitoxin--two functions: SR4 controls toxin mRNA decay and translation. *Nucleic Acids Res*, **41**, 9870-9880.

24. Kawano, M., Reynolds, A.A., Miranda-Rios, J. and Storz, G. (2005) Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in Escherichia coli. *Nucleic Acids Res*, **33**, 1040-1050.

25. Kawano, M., Aravind, L. and Storz, G. (2007) An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Mol Microbiol*, **64**, 738-754.

26. Kawano, M., Oshima, T., Kasai, H. and Mori, H. (2002) 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*, **45**, 333-349.

27. Guo, Y., Quiroga, C., Chen, Q., McAnulty, M.J., Benedik, M.J., Wood, T.K. and Wang, X. (2014) RalR (a DNase) and RalA (a small RNA) form a type I toxin-antitoxin system in Escherichia coli. *Nucleic Acids Res*, **42**, 6448-6462.

28. Vogel, J., Argaman, L., Wagner, E.G. and Altuvia, S. (2004) The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Curr Biol*, **14**, 2271-2276.

29. Darfeuille, F., Unoson, C., Vogel, J. and Wagner, E.G. (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Mol Cell*, **26**, 381-392.

30. Jahn, N., Preis, H., Wiedemann, C. and Brantl, S. (2012) BsrG/SR4 from Bacillus subtilis--the first temperature-dependent type I toxin-antitoxin system. *Mol Microbiol*, **83**, 579-598.

31. Silvaggi, J.M., Perkins, J.B. and Losick, R. (2005) Small untranslated RNA antitoxin in Bacillus subtilis. *J Bacteriol*, **187**, 6641-6650.

32. Durand, S., Gilet, L. and Condon, C. (2012) The essential function of B. subtilis RNase III is to silence foreign toxin genes. *PLoS Genet*, **8**, e1003181.

33. Meissner, C., Jahn, N. and Brantl, S. (2016) In Vitro Characterization of the Type I Toxin-Antitoxin System bsrE/SR5 from Bacillus subtilis. *J Biol Chem*, **291**, 560-571.

34. Durand, S., Jahn, N., Condon, C. and Brantl, S. (2012) Type I toxin-antitoxin systems in Bacillus subtilis. *RNA Biol*, **9**, 1491-1497.

35. Weaver, K.E., Ehli, E.A., Nelson, J.S. and Patel, S. (2004) Antisense RNA regulation by stable complex formation in the Enterococcus faecalis plasmid pAD1 par addiction system. *J Bacteriol*, **186**, 6400-6408.

36. Faridani, O.R., Nikravesh, A., Pandey, D.P., Gerdes, K. and Good, L. (2006) Competitive inhibition of natural antisense Sok-RNA interactions activates Hok-mediated cell killing in Escherichia coli. *Nucleic Acids Res*, **34**, 5915-5922.

37. Franch, T., Gultyaev, A.P. and Gerdes, K. (1997) Programmed cell death by hok/sok of plasmid R1: processing at the hok mRNA 3'-end triggers structural rearrangements that allow translation and antisense RNA binding. *J Mol Biol*, **273**, 38-51.

38. Muller, P., Jahn, N., Ring, C., Maiwald, C., Neubert, R., Meissner, C. and Brantl, S. (2016) A multistress responsive type I toxin-antitoxin system: bsrE/SR5 from the B. subtilis chromosome. *RNA Biol*, **13**, 511-523.

39. Fernandez De Henestrosa, A.R., Ogi, T., Aoyagi, S., Chafin, D., Hayes, J.J., Ohmori, H. and Woodgate, R. (2000) Identification of additional genes belonging to the LexA regulon in Escherichia coli. *Mol Microbiol*, **35**, 1560-1572.

40. Fozo, E.M., Makarova, K.S., Shabalina, S.A., Yutin, N., Koonin, E.V. and Storz, G. (2010) Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families. *Nucleic Acids Res*, **38**, 3743-3759.

41. Pecota, D.C., Osapay, G., Selsted, M.E. and Wood, T.K. (2003) Antimicrobial properties of the Escherichia coli R1 plasmid host killing peptide. *J Biotechnol*, **100**, 1-12.

42. Pedersen, K. and Gerdes, K. (1999) Multiple hok genes on the chromosome of Escherichia coli. *Mol Microbiol*, **32**, 1090-1102.

43. Weaver, K.E., Weaver, D.M., Wells, C.L., Waters, C.M., Gardner, M.E. and Ehli, E.A. (2003) Enterococcus faecalis plasmid pAD1-encoded Fst toxin affects membrane permeability and alters cellular responses to lantibiotics. *J Bacteriol*, **185**, 2169-2177.

44. Weaver, K.E. (2012) The par toxin-antitoxin system from Enterococcus faecalis plasmid pAD1 and its chromosomal homologs. *RNA Biol*, **9**, 1498-1503.

45. Brinkman, C.L., Bumgarner, R., Kittichotirat, W., Dunman, P.M., Kuechenmeister, L.J. and Weaver, K.E. (2013) Characterization of the effects of an rpoC mutation that confers resistance to the Fst peptide toxin-antitoxin system toxin. *J Bacteriol*, **195**, 156-166.

46. Patel, S. and Weaver, K.E. (2006) Addiction toxin Fst has unique effects on chromosome segregation and cell division in Enterococcus faecalis and Bacillus subtilis. *J Bacteriol*, **188**, 5374-5384.

47. Weaver, K.E., Reddy, S.G., Brinkman, C.L., Patel, S., Bayles, K.W. and Endres, J.L. (2009) Identification and characterization of a family of toxin-antitoxin systems related to the Enterococcus faecalis plasmid pAD1 par addiction module. *Microbiology*, **155**, 2930-2940.

48. Koyanagi, S. and Levesque, C.M. (2013) Characterization of a Streptococcus mutans intergenic region containing a small toxic peptide and its cis-encoded antisense small RNA antitoxin. *PLoS One*, **8**, e54291.

49. Michaux, C., Hartke, A., Martini, C., Reiss, S., Albrecht, D., Budin-Verneuil, A., Sanguinetti, M., Engelmann, S., Hain, T., Verneuil, N. *et al.* (2014) Involvement of Enterococcus faecalis small RNAs in stress response and virulence. *Infect Immun*, **82**, 3599-3611.

50. Sayed, N., Nonin-Lecomte, S., Rety, S. and Felden, B. (2012) Functional and structural insights of a Staphylococcus aureus apoptotic-like membrane peptide from a toxin-antitoxin module. *J Biol Chem*, **287**, 43454-43463.

51. Lewis, K. (2007) Persister cells, dormancy and infectious disease. *Nat Rev Microbiol*, **5**, 48-56.

52. Lewis, K. (2008) Multidrug tolerance of biofilms and persister cells. *Curr Top Microbiol Immunol*, **322**, 107-131.

53. Gefen, O. and Balaban, N.Q. (2009) The importance of being persistent: heterogeneity of bacterial populations under antibiotic stress. *FEMS Microbiol Rev*, **33**, 704-717.

54. Hooper, D.C. (2001) Mechanisms of action of antimicrobials: focus on fluoroquinolones. *Clin Infect Dis*, **32 Suppl 1**, S9-S15.

55. Dorr, T., Vulic, M. and Lewis, K. (2010) Ciprofloxacin causes persister formation by inducing the TisB toxin in Escherichia coli. *PLoS Biol*, **8**, e1000317.

56. Wagner, E.G. and Unoson, C. (2012) The toxin-antitoxin system tisB-istR1: Expression, regulation, and biological role in persister phenotypes. *RNA Biol*, **9**, 1513-1519.

57. Unoson, C. and Wagner, E.G. (2008) A small SOS-induced toxin is targeted against the inner membrane in Escherichia coli. *Mol Microbiol*, **70**, 258-270.

58. Steinbrecher, T., Prock, S., Reichert, J., Wadhwani, P., Zimpfer, B., Burck, J., Berditsch, M., Elstner, M. and Ulrich, A.S. (2012) Peptide-lipid interactions of the stress-response peptide TisB that induces bacterial persistence. *Biophys J*, **103**, 1460-1469.

59. Gurnev, P.A., Ortenberg, R., Dorr, T., Lewis, K. and Bezrukov, S.M. (2012) Persister-promoting bacterial toxin TisB produces anion-selective pores in planar lipid bilayers. *FEBS Lett*, **586**, 2529-2534.

60. Dorr, T., Lewis, K. and Vulic, M. (2009) SOS response induces persistence to fluoroquinolones in Escherichia coli. *PLoS Genet*, **5**, e1000760.

61. Goeders, N. and Van Melderen, L. (2014) Toxin-antitoxin systems as multilevel interaction systems. *Toxins (Basel)*, **6**, 304-324.

62. Maisonneuve, E., Shakespeare, L.J., Jorgensen, M.G. and Gerdes, K. (2011) Bacterial persistence by RNA endonucleases. *Proc Natl Acad Sci U S A*, **108**, 13206-13211.

63. Kwan, B.W., Valenta, J.A., Benedik, M.J. and Wood, T.K. (2013) Arrested protein synthesis increases persister-like cell formation. *Antimicrob Agents Chemother*, **57**, 1468-1473.

64. Fozo, E.M., Kawano, M., Fontaine, F., Kaya, Y., Mendieta, K.S., Jones, K.L., Ocampo, A., Rudd, K.E. and Storz, G. (2008) Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. *Mol Microbiol*, **70**, 1076-1093.

65. Rotem, E., Loinger, A., Ronin, I., Levin-Reisman, I., Gabay, C., Shoresh, N., Biham, O. and Balaban, N.Q. (2010) Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. *Proc Natl Acad Sci U S A*, **107**, 12541-12546.

66. Irnov, I., Sharma, C.M., Vogel, J. and Winkler, W.C. (2010) Identification of regulatory RNAs in Bacillus subtilis. *Nucleic Acids Res*, **38**, 6637-6651.

67. Saito, S., Kakeshita, H. and Nakamura, K. (2009) Novel small RNA-encoding genes in the intergenic regions of Bacillus subtilis. *Gene*, **428**, 2-8.

68. Tovar-Rojo, F. and Setlow, P. (1991) Effects of mutant small, acid-soluble spore proteins from Bacillus subtilis on DNA in vivo and in vitro. *J Bacteriol*, **173**, 4827-4835.

69. Yamamoto, T., Obana, N., Yee, L.M., Asai, K., Nomura, N. and Nakamura, K. (2014) SP10 infectivity is aborted after bacteriophage SP10 infection induces nonA transcription on the prophage SPbeta region of the Bacillus subtilis genome. *J Bacteriol*, **196**, 693-706.

70. Kunkel, B., Losick, R. and Stragier, P. (1990) The Bacillus subtilis gene for the development transcription factor sigma K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene. *Genes Dev*, **4**, 525-535.

71. Sato, T. and Kobayashi, Y. (1998) The ars operon in the skin element of Bacillus subtilis confers resistance to arsenate and arsenite. *J Bacteriol*, **180**, 1655-1661.

72. Weel-Sneve, R., Kristiansen, K.I., Odsbu, I., Dalhus, B., Booth, J., Rognes, T., Skarstad, K. and Bjoras, M. (2013) Single transmembrane peptide DinQ modulates membrane-dependent activities. *PLoS Genet*, **9**, e1003260.

73. Pinel-Marie, M.L., Brielle, R. and Felden, B. (2014) Dual toxic-peptide-coding Staphylococcus aureus RNA under antisense regulation targets host cells and bacterial rivals unequally. *Cell Rep*, **7**, 424-435.

74. Kawano, M. (2012) Divergently overlapping cis-encoded antisense RNA regulating toxin-antitoxin systems from E. coli: hok/sok, ldr/rdl, symE/symR. *RNA Biol*, **9**, 1520-1527.

75. Storz, G., Vogel, J. and Wassarman, K.M. (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell*, **43**, 880-891.

Appendix I: chapter I tables and figures

Table 1.1. Features of described type I toxin-antitoxin loci

Locus	Founding member (Plasmid or bacterium)[a]	Genetic organization[b]	Mode of antitoxin action
hok/sok	Plasmid R1 (16)		Inhibit protein synthesis Stimulate mRNA degradation
fst/RNAII	Plasmid pAD1 (21)		Inhibit protein synthesis
bsrG/sr4	*Bacillus subtilis* (30)		Inhibit protein synthesis Stimulate mRNA degradation
symE/symR	*Escherichia coli* (25)		Inhibit protein synthesis Stimulate mRNA degradation
ibs/sib	*Escherichia coli* (64)		Inhibit protein synthesis Stimulate mRNA degradation
ralR/ralA	*Escherichia coli* (27)		Inhibit protein synthesis
tisB/istR-1	*Escherichia coli* (28)		Inhibit protein synthesis Stimulate mRNA degradation

Table 1.1. (continued)

Locus	Founding member (Plasmid or bacterium) [a]	Genetic organization[b]	Mode of antitoxin action
txpA/ratA	*Bacillus subtilis* (21)		Stimulate mRNA degradation
bsrE/sr5	*Bacillus subtilis* (33)		Stimulate mRNA degradation
bsrH/as-bsrH	*Bacillus subtilis* (66)		Stimulate mRNA degradation

[a] The founding member (first description) of each toxin-antitoxin locus is indicated. Note that many homologs to these systems have been identified and characterized; the text gives details of those; [b] The toxin mRNA is represented by the blue arrow with the coding region shown as the blue box. The antitoxin RNA is represented by the red arrow.

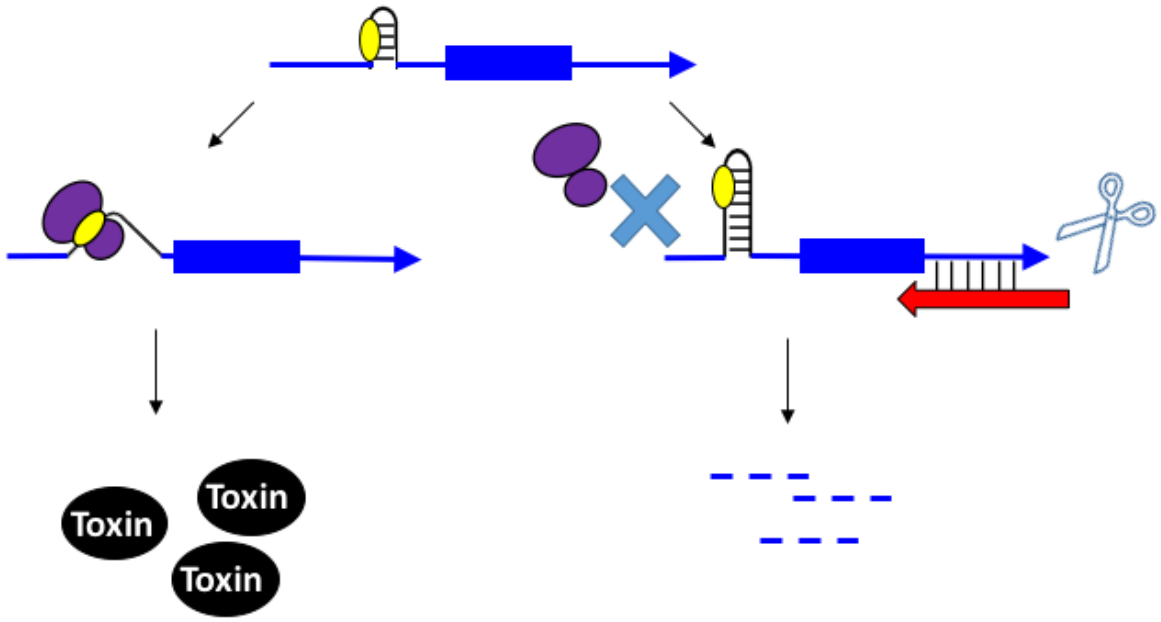


Figure 1.1. Repression of the BsrG toxin production by the SR4 antitoxin. The ribosome binding site (yellow) of the *bsrG* mRNA is normally located in a 4-nt-long stem. Binding by the SR4 antitoxin sRNA (red) causes structural changes of this stem and results in extension of the stem to 8-nt-long, which leads to reduced translation of the *bsrG* mRNA. This interaction with the antitoxin sRNA can also trigger toxin mRNA degradation.

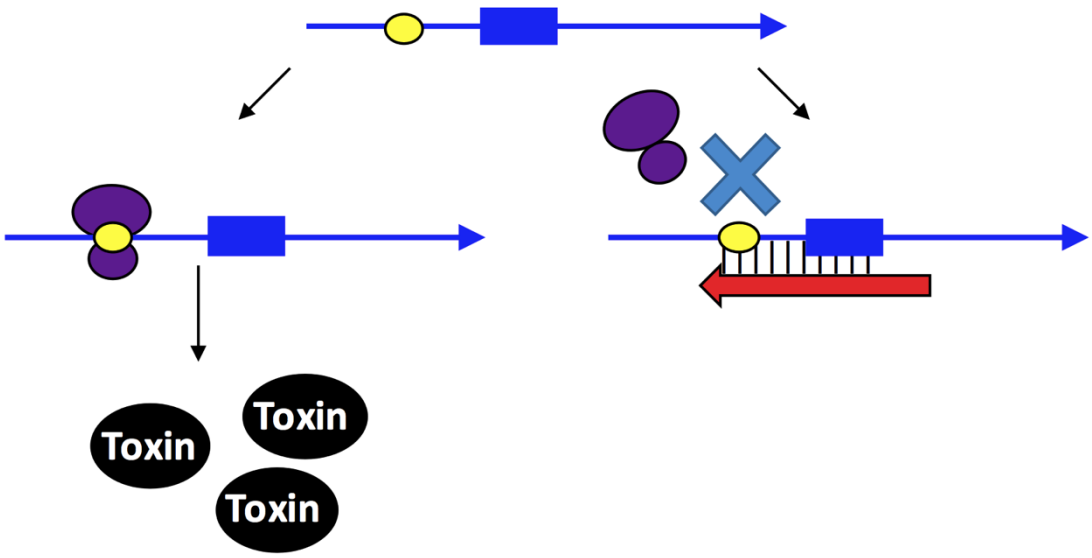


Figure 1.2. Repression of the IbsC toxin production by the SibC antitoxin. The *ibs* mRNA (blue) is translated by the ribosome (purple) when the Sib antitoxin (red) is not present. Binding of the SibC antitoxin RNA (red) blocks translation as it overlaps the ribosome binding site (yellow) and the translation initiation codon and then can potentially extend over the entire coding region.

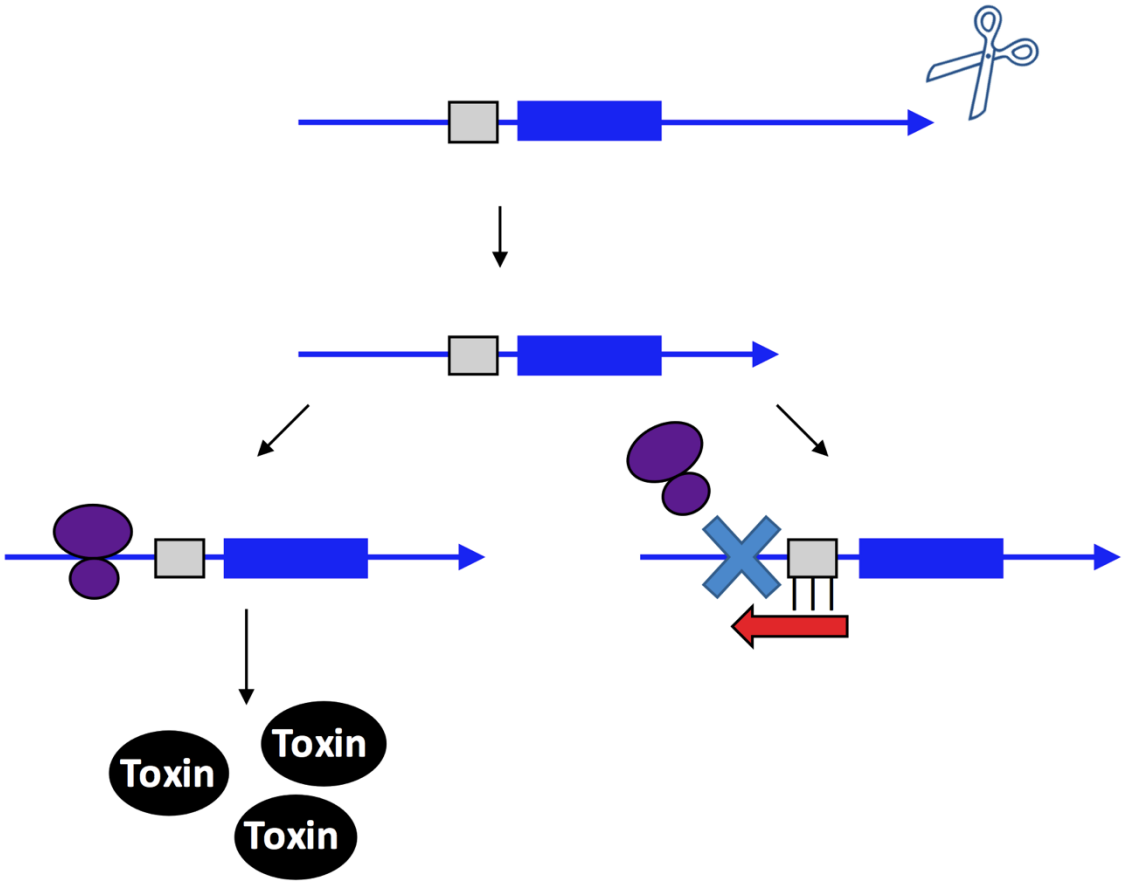


Figure 1.3. Repression of the Hok toxin production by the Sok antitoxin. The *hok* mRNA (blue) must first be processed at its 3' end prior to interaction with either the ribosome (purple) or the Sok sRNA (red). Translation of the *hok* mRNA is dependent upon translation of an upstream open reading frame *mok* (grey) by the ribosome. Binding of the Sok antitoxin RNA (red) blocks translation of the leader peptide, thus preventing toxin expression.

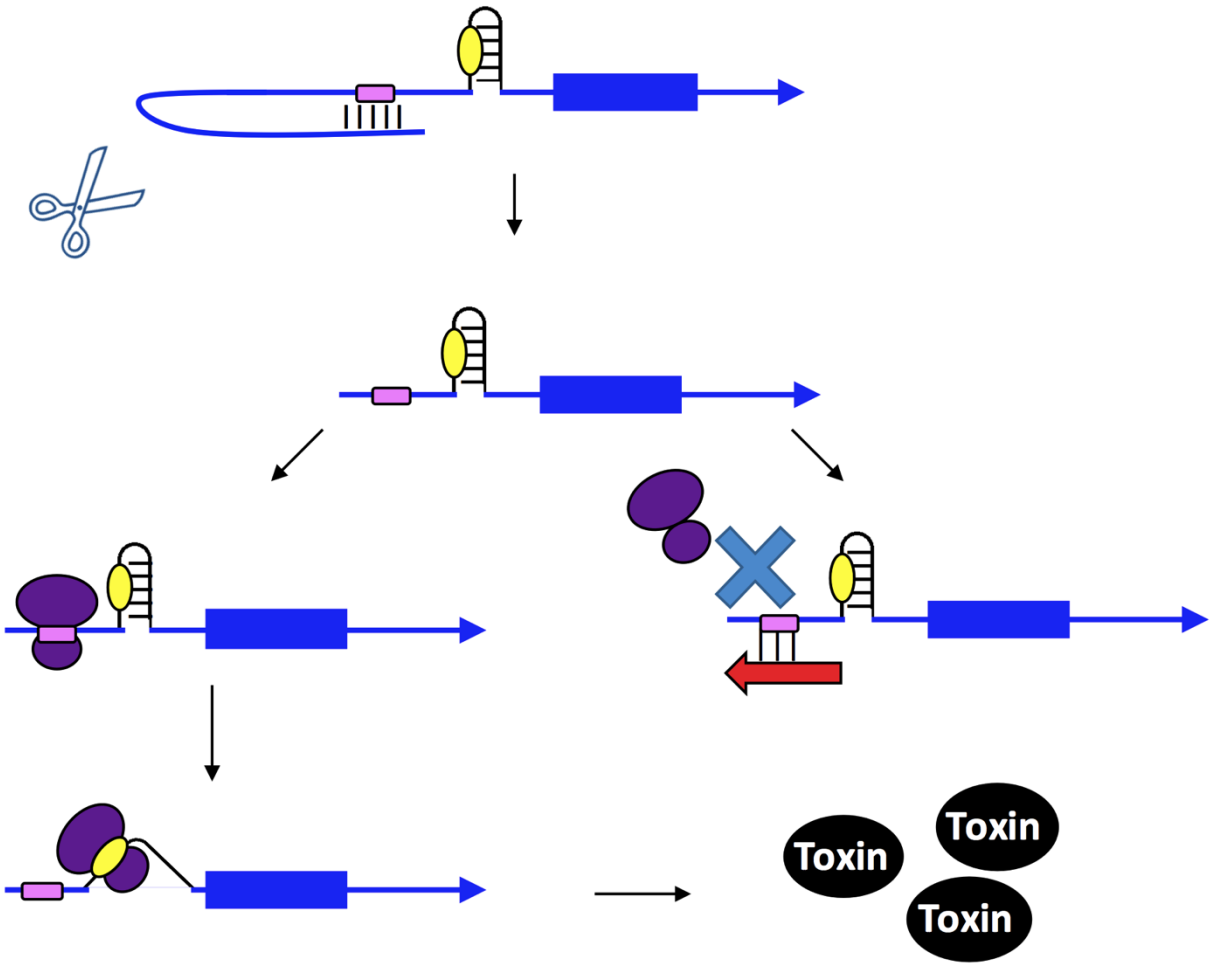


Figure 1.4. Repression of the TisB toxin production by the IstR-1 antitoxin. Translation of the *tisB* mRNA is dependent on both a processing event and binding of the ribosome (purple) to a standby site (pink) as the real toxin ribosome binding site (yellow) is sequestered in a secondary structure. First, the full-length *tisB* mRNA (blue) must be processed at its 5' end in order for the IstR-1 sRNA (red) or the ribosome to bind. Following processing, the standby site (pink) is accessible to the ribosome (purple) and the antitoxin (red). Binding by the IstR-1 antitoxin will block binding of the ribosome to the standby site. In the absence of the antitoxin, the ribosome can bind to the standby site and then onto the true ribosome binding site (yellow).

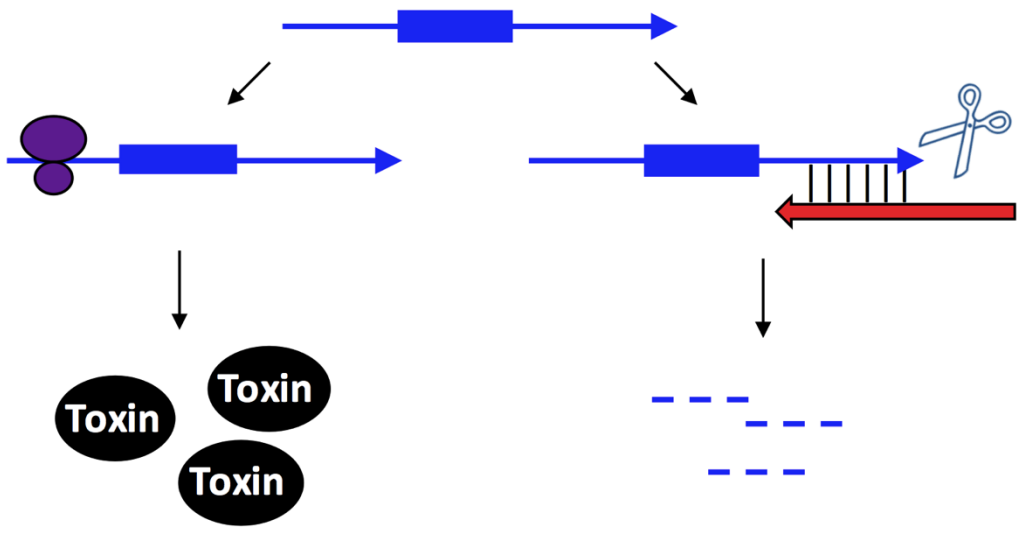


Figure 1.5. Repression of the TxpA toxin production by the RatA antitoxin. The *txpA* toxin mRNA (blue) can be translated by the ribosome (purple) to produce the toxic protein. However, if the antitoxin RatA (red) is present and interacts with the toxin, RNase III can cleave the double-stranded complex, thereby initiating mRNA degradation which eventually prevents toxin translation.

CHAPTER II

The ZorO-OrzO type I toxin-antitoxin locus: repression by the OrzO antitoxin

Publication note

A version of this chapter was originally published by Jia Wen, Daniel Won, and Elizabeth M. Fozo:

Jia Wen, Daniel Won, and Elizabeth M. Fozo. "The ZorO-OrzO type I toxin-antitoxin locus: repression by the OrzO antitoxin." *Nucleic Acids Research* (2014) 42 (3):1930-1946.

My contribution to this paper was mutant constructions, rescue experiments, RNA isolation, northern analyses, quantification of RNA levels, and most of the literature review and writing. Daniel Won assisted generating data presented in Figure 2.3, where he performed the rescue experiment with wild type *Escherichia coli* O157: H7 and Δ*rnc* mutant as well as the corresponding northern analyses. Dr. Elizabeth Fozo performed the primer extension assay represented in Figure 2.4B. I performed all other experiments.

Abstract

Type I toxin-antitoxin loci consist of two genes: a small, hydrophobic, potentially toxic protein, and a small RNA (sRNA) antitoxin. The sRNA represses toxin gene expression by base pairing to the toxin mRNA. A previous bioinformatics search predicted a duplicated type I locus within *Escherichia coli* O157:H7 (EHEC), which we have named the gene pairs *zorO-orzO* and *zorP-orzP*. We show that overproduction of the *zorO* gene is toxic to *E. coli*; co-expression of the sRNA OrzO can neutralize this toxicity, confirming that the *zorO-orzO* pair is a true type I toxin-antitoxin locus. However, OrzO is unable to repress *zorO* in a strain deleted for RNase III, indicating that cleavage of the target mRNA is critical for repression. Sequence analysis and mutagenesis studies have elucidated a nucleotide sequence region (V1) that allows differential recognition of the *zorO* mRNA by OrzO and not OrzP, and a specific, single nucleotide within the V1 of OrzO that is critical for repression of *zorO*. Although there are 18 nucleotides of complementarity between the OrzO sRNA and the *zorO* mRNA, not all base pairing interactions are needed for repression; however, the amount needed is dependent upon whether there is continuous or discontinuous complementarity to the target mRNA.

I. Introduction

As a newly recognized class of gene expression regulators, small, non-coding RNAs have been found in all kingdoms of life. In bacteria, these RNAs are referred to as sRNAs (small RNAs) and are often 50-300 nt in length. Some sRNAs bind to proteins and modulate their functions, as exemplified by the CsrB RNA of *Escherichia coli*, which directly binds and sequesters the CsrA protein, thus affecting carbon metabolism in *E. coli* (1). However, the majority of characterized sRNAs function by base pairing to target mRNAs, either regulating the translation or altering the stabilities of the targets.

These base pairing sRNAs fall into two categories: *trans*-encoded sRNAs and antisense sRNAs. *Trans*-encoded sRNAs are located at chromosomal positions distal from the genes they regulate. These sRNAs typically interact with their targets by limited base pairing and often require the RNA chaperone Hfq to function [reviewed in (2)]. True antisense sRNAs are encoded on the opposite strand of DNA to their targets. They share extensive base pairing potential (usually 60 nt or more) to their target mRNAs [reviewed in (3,4)]. Although antisense sRNAs were initially discovered on mobile genetic elements, several pairs have been described on bacterial chromosomes [reviewed in (5)].

A subset of antisense sRNAs in bacteria can repress the expression of small proteins under 60 amino acids in length. These proteins are usually very hydrophobic and toxic when overproduced. The sRNA (antitoxin) base pairs with the toxin mRNA and affects the translation and/or the stability of the mRNA. Gene pairs consisting of a toxic protein-encoding gene and a corresponding antitoxin sRNA-encoding gene are referred to as type I toxin-antitoxin systems. The first type I pair to be identified was the *hok/sok* locus of the R1 plasmid in *E. coli*, whereby Hok is the toxin, and Sok the antitoxin (6). The *hok* mRNA has a very long half-life while the Sok sRNA has a short half-life. When cell division occurs, a daughter cell that has not inherited the R1 plasmid will die due to the quick degradation of the Sok sRNA and the translation of the more stable *hok* mRNA. Besides the *hok/sok* locus, several type I toxin-antitoxin gene pairs have been described on other plasmids in gram negative bacteria and were also shown to play a role in plasmid maintenance [reviewed in (7)]. Additionally, the RNAI-RNAII locus of the pAD1 plasmid of *Enterococcus faecalis* clearly showed that similar loci were present in gram positive bacteria (8-10). Homologs of these plasmid loci were later found within bacterial chromosomes (11-14).

In recent years, several type I toxin-antitoxin systems that do not have homology to plasmid sequences have been identified and characterized in a number of bacterial chromosomes. Although both the toxin and antitoxin-encoding genes are located at the same locus, their relative positioning varies among the chromosomally encoded type I pairs. Some antitoxins overlap with the 5' untranslated region (UTR) of their targets, such as the Rdl sRNAs of the Ldr-Rdl family in *E. coli* (15), or the 3' UTR, like RatA and SR4 of the TxpA-RatA and BsrG/SR4 loci of *Bacillus subtilis* (16,17). Others directly overlap with the coding sequences of the target mRNAs, as exemplified by the Ibs-Sib family of *E. coli* in which Sib sRNA completely overlaps with the *ibs* mRNA coding region (18).

Along with these, there are a few unconventional type I loci in which the antitoxins are encoded divergent from the toxin-encoding genes [reviewed in (19)]. Unlike the traditional type I loci that can have 60 nt or more base paring potential, the unconventional type I loci have limited base pairing, often only 18-21 nt. These loci include the TisAB-IstR and the ShoB-OhsC pairs in *E. coli* (18,20).

A previous bioinformatics search identified two highly homologous loci encoded in tandem in the chromosome of *E. coli* O157:H7 EDL933 (EHEC) (14). Each locus includes two genes: one encodes a small hydrophobic protein, and the other encodes an sRNA that is located divergently from the protein-encoding gene. This genetic organization is very similar to that of the type I toxin-antitoxin pairs ShoB-OhsC and TisAB-IstR. In the two newly identified loci, the small protein-encoding genes (29 amino acids) were previously annotated as *z3289* and *z3290*. We have since renamed *z3289* and *z3290* as *zorO* (Z protein often repeated) and *zorP* respectively. Additionally, we have denoted the sRNA gene divergent from *zorO* as *orzO* (overexpression represses *zor* toxicity) and the sRNA gene divergent from *zorP* as *orzP* (Figure 2.1A; all tables and figures are located in the appendix). The sRNAs contain regions of complementarity (18-19 nt) to the mRNAs, suggesting that they could base pair. In addition, the *zor-orz* loci are highly conserved in most commensal and pathogenic *E. coli* and *Shigella* strains but not in the common lab strains, like *E. coli* MG1655 (14).

Overexpression of ZorO was shown to be toxic; however, it was never demonstrated that the sRNA OrzO could repress *zorO*-induced toxicity (14). Thus, is *zorO-orzO* a true type I toxin-antitoxin locus? Given the high homology between the *zorO-orzO* and *zorP-orzP* loci, is there "cross-talk" in the regulation? In this study, we confirmed that the *zorO-orzO* locus is indeed a

true type I toxin-antitoxin system. Utilizing the inherent toxicity associated with *zorO* overproduction, we performed mutational analyses to further investigate the requirements for successful repression by the OrzO antitoxin. Through these analyses, we revealed that the 5' end of the OrzO sRNA can solely repress *zorO* and the V1 region within 5' end of the Orz sRNA determines the specific recognition of the *zor* target.

II. Materials and Methods

Bacteria strains and plasmids All bacterial strains and plasmids utilized in this work are listed in Table 2.1. The sequences of all oligonucleotides are listed in Table 2.2.

Growth conditions *E. coli* strains were grown at 37°C in LB medium with shaking. Antibiotics were added when necessary at the following concentration: 100 μg/mL ampicillin, 25 μg/mL chloramphenicol, 25 μg/mL kanamycin. Arabinose was added to a final concentration of 0.2 or 0.002% as indicated. IPTG, when used, was at a final concentration of either 0.1 mM or 1 mM, as indicated.

Generation of strains for rescue experiments All strains used in the described rescue experiments are derived from DJ624 which contains *ΔlacX74* and does not possess the *zor-orz* locus (21). To better control expression from the P_{BAD} promoter across the population of cells, the *araE* gene was placed under control of the P_{CP18} constitutive promoter through P1 transduction of the *kan*-P_{CP18}-*araE* allele (22). The kanamycin cassette was then removed using pCP20, resulting in strain UTK007 (23). Generation of the *hfq* and *rnc* deletions was performed using the mini-λ-Red recombination system (24-26). In each case, pKD4 served as a PCR template for amplifying the kanamycin cassette using primers containing 20 nt of homology to the kanamycin cassette, flanked by approximately 40 nt of homology to the chromosomal region (25).

Plasmid construction For overexpression of the sRNAs OrzO and OrzP, the genes were amplified from genomic DNA (*E. coli* O157:H7 EDL933) and ligated into the AatII and EcoRI sites of pBR-plac generating pBR-plac-*orzO* and pBR-plac-*orzP* (21). The *zorO* gene was ligated into the PstI and HindIII sites of pEF21 from the mapped +1 of its transcription, generating pEF21-*zorO* (pBAD-*zorO* in the text) (14,18). Mutations of individual nucleotide residues for all constructs

were performed by site-directed mutagenesis of either pBR-plac-*orzO*, or pBR-plac-*orzP*, or pEF21-*zorO*. For generation of pBR-pLac-*orzO* 6(1)10(1) and pBR-plac-*orzO* (1)5(1)11, pBR-plac-*orzO* 6(1) 11 served as the template. The two chimeric constructs, pBR-plac-*orzO-ohsC* and pBR-plac-*orzO-istR*, were generated using SOE PCR (27,28). PCR product A amplified the first 18 nucleotides of *orzO* from pBR-plac-*orzO* along with 100 nt upstream of vector sequence (including the AatII cut site); the 3' end of the product contained 15 nt of PCR product B. PCR product B amplified either *ohsC* or *istR* from *E. coli* O157:H7 EDL933 genomic DNA; this product began immediately following the region of base pairing that these RNAs would have for their respective targets (18,20). The 5' end of PCR B contained 15 nt of homology to the 5' end of *orzO*. The products were then spliced together via PCR, digested with AatII and EcoRI, and ligated into pBR-plac. Plasmid DNA was isolated using the Qiagen Mini Plasmid Kit; PCR products were isolated using the Qiagen PCR Purification Kit.

Rescue experiments Rescue experiments were as described previously (15,18). The pBR-plac plasmid carrying the *orz* wild type/ mutant gene under an IPTG-induced $P_{LlacO-1}$ promoter was initially transformed into UTK007 or one of the derivatives described above, through electroporation (21). Afterwards, the pEF21 plasmid containing the *zor* wild type/mutant gene under an arabinose-induced P_{BAD} promoter was transformed into the same cells (29). The resulting transformants harboring two plasmids were grown overnight and then diluted to OD_{600} of 0.01. When the OD_{600} reached ≈ 0.1, the cultures were split and IPTG was added to half the culture to a final concentration of either 0.1 mM or 1 mM for the overexpression of the *orz* gene. Thirty minutes after IPTG induction, arabinose was added to a final concentration of either 0.002% or 0.2% for the overproduction of the *zor* gene. OD_{600} was taken and recorded every 30 minutes. Shown are averages ± standard deviations for a minimum of three independent experiments.

RNA extraction In order to examine RNA levels of *zorO* and OrzO, cells harboring the appropriate plasmids were grown to an $OD_{600} \approx 0.2$, and arabinose was added as indicated to induce *zorO* (time 0). Five minutes post-induction with arabinose, the culture was split and IPTG (0.1 mM or 1 mM) was added to half. RNA was extracted via direct lysis as previously described with some modifications (30). Cells (750 µL aliquots) from time 0, 5, 15, 30, and 60 minutes post-arabinose were harvested and incubated at 65°C with 500µL acid phenol: chloroform (preheated

to 65°C) and 102 µL of direct lysis solution (320 mM sodium acetate, 8% SDS, 16 mM EDTA) for 10 minutes. After incubation, the mixtures were centrifuged at 13000 rpm at room temperature for 10 minutes. Supernatants were transferred to tubes containing 500 µL of 65°C phenol: chloroform, mixed thoroughly and spun again for 10 minutes at room temperature. The supernatants were transferred and extracted twice with 400 µL phenol: chloroform. The RNA was then ethanol precipitated and resuspended in RNase free water. When needed, RNA was treated with Turbo DNaseI (Life Technologies) following manufacturer's instructions.

Northern analysis Total RNA (10 µg) isolated at indicated time points following arabinose induction was separated on a denatured 6% (for the separation of *zor* mRNA) or 8% (for the separation of *orz* sRNA) polyacrylamide-urea gel and transferred to a Zeta-Probe Genomic GT membrane (Bio-Rad). Specific probes were generated by end-labeling oligonucleotides with γ-^{32}P by T4 polynucleotide kinase. Incubation of the membrane and washes were performed as outlined by Opdyke *et al* (31). Northern analysis was performed from a minimum of three independent experiments for every construct examined.

Quantification of RNA levels The relative intensities of the resulting bands from northern analysis detected via autoradiography were quantified using ImageJ (32). The intensity of the band for WT at T_{15} (induced with the level of IPTG indicated) was set to a value of "1" as a standard for comparison. Relative intensities to WT at T_{15} for the other bands on the same northern blot were thus calculated. Averages and standard deviations were determined, and the statistical significance was calculated using Student's *t* test.

Primer extension Total RNA (5 µg) isolated thirty minutes post arabinose induction from rescue experiments was separated on a denatured 8% polyacrylamide-urea gel as described previously (18,31).

III. Results

***zorO-orzO* is a bona fide type I toxin-antitoxin locus** The *zorO* gene of *E. coli* O157:H7 EDL933, initially annotated as *z3289*, was predicted to be type I toxin (14). Overexpression of the small protein from a high-copy plasmid in *E. coli* MG1655 led to cell growth stasis as well as a decrease

in colony forming units, indicating that ZorO is toxic (14). In the same study, the sRNA OrzO (denoted at that time as sRNA-1) was detected via northern analysis. This sRNA, though encoded divergent from the toxin gene, could potentially base pair with the 5' UTR of the *zorO* mRNA (Figure 2.1A, 2.1B).

Despite these data, the crucial question remained whether the *zorO-orzO* pair is a true type I toxin-antitoxin locus. If so, then the OrzO sRNA should be able to repress the toxicity of ZorO overproduction. Rescue experiments were employed to test if OrzO can repress ZorO. We expressed *zorO* from an arabinose-inducible promoter (P_{BAD}) on one plasmid and *orzO* from an IPTG-inducible promoter ($P_{LlacO-1}$) on a second compatible plasmid (15,18). These experiments were performed in a strain derived from *E. coli* MG1655 that does not possess the *zor-orz* loci so as to eliminate any possible repression from a chromosomally encoded OrzO. While the induction of *zorO* led to cell growth stasis, indicating that ZorO is toxic, the co-expression of OrzO prevented this cell stasis, suggesting that OrzO serves as an antitoxin for ZorO (Figure 2.1B). Thus, we conclude that the *zorO-orzO* pair is a true type I toxin-antitoxin locus.

RNase III but not Hfq is critical for repression of *zorO*-induced toxicity The majority of *trans*-encoded sRNAs characterized to date require the RNA chaperone Hfq to function efficiently whereas the antisense sRNAs typically do not (2,7,33). However, given that there is limited complementarity between the *zorO* mRNA and the corresponding OrzO sRNA (18 nt perfect complementarity), we wanted to examine whether Hfq is required for the observed repression of the ZorO toxin. A similar rescue experiment was performed as outlined above in a strain deleted for *hfq*. The results showed that the OrzO sRNA can fully repress *zorO* in both the wild type and the *Δhfq* strain (Figure 2.2), indicating that Hfq is not needed for the OrzO sRNA to regulate *zorO*.

RNase III, a double-stranded ribonuclease, has been shown to be critical for cleaving paired toxin-antitoxin complexes including *hok/sok*, *tisB-istR, bsrG/sr4*, *txpA-ratA*, as well as for the sRNA GadY and its mRNA target *gadX-gadW* (16,20,34,35,38). To determine whether repression by OrzO was dependent upon expression of RNase III, we performed our rescue experiment in a strain deleted for *rnc*, the gene encoding RNase III. The results (Figure 2.3A) show that OrzO is incapable of rescuing *zorO*-induced toxicity in this mutant, even at low levels of inducing agent (0.002% arabinose, see below; data not shown). When *zorO* is expressed with OrzO in a wild type strain, we detect major transcripts of approximately 280 and 310 nt in length via northern analysis

(Figure 2.3B). When examining *zorO* in the *Δrnc* strain, we noted an accumulation of these full-length transcripts in comparison to expression in the wild type strain, suggesting that the full-length mRNAs are more stable in the deletion strain (Figure 2.3B). This implies that cleavage of the *zorO*-OrzO RNA duplex by RNase III is important for repression of *zorO*-induced toxicity. Surprisingly, OrzO levels (induced by 1 mM IPTG) appear reduced in the *Δrnc* strain (Figure 2.4A).

We attempted to map the cleavage sites of *zorO* when OrzO is expressed in a wild type and a *Δrnc* strain using primer extension. Similar to our northern analyses, we noted an accumulation of full-length *zorO* mRNA in the *Δrnc* strain even when OrzO is expressed (Figure 2.4B). However, we are able to map OrzO-specific cleavage sites within the *zorO* mRNA in the *Δrnc* strain (Figure 2.1A, 2.4B), similar to what has been reported for cleavage of *gadX-gadW* by GadY and cleavage of *cII-O* by the OPP RNA (35-37). This finding suggests that while RNase III may be critical for regulation of *zorO* repression, there may be additional factors involved (see Discussion). As summarized in Figure 1A and shown in Figure 2.4B, the major processing sites lie within the region of base pairing for both the wild type and *Δrnc* strain. We did note that in the *Δrnc* strain there were sites that appear to be shifted from the wild type cleavage sites, as well as decreased cleavage at some sites (Figure 2.1A, 2.4B). We also detected processing outside the region of base pairing for both strains (Figure 2.1A, 2.4B), but it is not clear at this point the origin of these products.

5' end of the OrzO sRNA can solely repress *zorO* A previous study has shown that the antitoxin IstR-1 of *E. coli* regulates its toxin target TisB by 5' end pairing (20). Similarly, the OhsC antitoxin is predicted to base pair to its target ShoB through its 5' end (18). In *zorO-orzO* pair, the longest stretch of potential base pairing between OrzO and the *zorO* mRNA also occurs in the immediate 5' end of OrzO. To test if just the 5' end of OrzO can repress *zorO*, we first constructed two OrzO mutants truncated at the 3' end. However, we were unable to detect the expression of these truncated sRNAs (data not shown). We hypothesized that the inability to detect these truncated RNAs was due to their instability; therefore, we generated chimeric sRNAs to the OhsC and IstR-1 sRNAs. The base pairing regions that OhsC and IstR-1 had to their own toxins were replaced with the first 18 nt of the 5' end of the OrzO sRNA (Figure 2.5A). Consequently, only the 5' end of the chimeras can base pair with the *zorO* mRNA. Rescue experiment results showed that

although the wild type OhsC and IstR sRNAs failed to repress *zorO*, both chimeras successfully repressed the expression of *zorO* (Figure 2.5B, 2.5C). This demonstrates that the first 18 nt of OrzO are sufficient for full repression.

Orz regulation is target specific The 5' end of OrzO contains a continuous stretch of 18 nt that can base pair to the *zorO* mRNA (Figure 2.1B). The OrzP sRNA, which is a homolog of OrzO, also shares a total of 15 nt complementarity to *zorO* (Figure 2.6B). Because of the sequence similarities between OrzO and OrzP, it is possible that OrzP is capable of regulating ZorO (Figure 2.6A).

Rescue experiments were used to determine whether the OrzP sRNA can repress the expression of *zorO*. When OrzO was co-expressed with *zorO*, the strain was able to grow. However, when OrzP was co-expressed with *zorO*, cell stasis occurred (Figure 2.6C), suggesting that only OrzO, but not OrzP, can repress *zorO*. We also confirmed that OrzP was being expressed at levels equivalent to or higher than OrzO by northern analysis (Figure 2.7), which indicates that the failure of OrzP to rescue is not due to expression differences. Thus, repression by the Orz sRNAs is target-specific.

V1 region determines OrzO specificity for *zorO* Although OrzO and OrzP share the same predicted base pairing region with *zorO*, only OrzO is able to repress *zorO*. To further investigate what dictates this specificity, we compared the sequences of the two sRNAs (Figure 2.6A). Despite the great sequence similarity between OrzO and OrzP, there are several differences. In particular, the nucleotide sequences from +9 to +11 compose a variable region of 3 nt between OrzO (GAA) and OrzP (ACG). Indeed, this variable region (referred to as V1) is the most prominent difference observed within the 5' regulatory domain of the two sRNAs.

To determine if this V1 region is responsible for specificity in the OrzO regulation of *zorO*, the three nucleotides in the V1 region of OrzO were changed from "GAA" to "ACG", the sequence found in OrzP (Table 2.3). We then tested the ability of the resulting mutant strain (15 nt of base pairing potential) to rescue cells from *zorO*-induced toxicity. Our results showed that the OrzO V1 mutant completely failed to rescue cells (Figure 2.6D), suggesting that the V1 region is critical for repression by the OrzO sRNA.

To confirm these results, we generated compensatory mutations in *zorO* (referred to as *zorO* V1) so as to restore complementarity to the OrzO V1 mutant. Indeed, restoring the base pairing allowed the OrzO V1 to fully repress ZorO V1 induced toxicity (Figure 2.8). These data implied that the V1 region is important for OrzO to recognize its target.

Furthermore, we engineered an OrzP V1 mutant, in which its V1 region was changed from "ACG" to "GAA", the sequence found in OrzO. Consequently, the OrzP V1 mutant has increased ability to base pair to *zorO*, as it could perfectly match *zorO* in the V1 region (Table 2.3). Unlike the OrzP WT, which cannot inhibit *zorO* mRNA expression, the OrzP V1 mutant fully represses *zorO* (Figure 2.6D). These data, in addition to the observations from the *zorO* WT-OrzO V1 pair and the *zorO* V1-OrzO V1 pair, further support the hypothesis that the V1 region is responsible for the Orz sRNA specificity to its *zor* target.

The first base of the V1 region is critical for OrzO regulation of *zorO* Since the V1 region of the 5' end of OrzO is composed of three nucleotides (GAA), we then asked whether all three nucleotides contribute equally to the ability of OrzO to regulate *zorO*. Each nucleotide of the V1 region was mutated individually and tested for the ability to rescue cells from *zorO*-toxicity. The resulting mutants are named OrzO 8(1)9, OrzO 9(1)8 and OrzO 10(1)7. These are named such that the number in the parentheses indicates the number of unpaired nucleotides, while the numbers on the two sides of the parentheses represent the number of the nucleotides within the 5' end of OrzO that are capable of base pairing with *zorO* (Table 2.3). For example, in OrzO 8(1)9, the first 8 nt of OrzO can pair to *zorO*, followed by a single unpaired nt, and then 9 nt of pairing.

Rescue experiments were used to assess the repression abilities of the three mutants. Interestingly, the repression ability of the single mutants was not equivalent. The point mutation in the first nucleotide of the V1 region of OrzO completely abolished the repression ability of OrzO sRNA (Figure 2.9A), as the OrzO 8(1)9 mutant failed to rescue cells even at a low induction level of *zorO* (0.002% arabinose, see below). Northern analysis further indicated that the OrzO 8(1)9 mutant showed levels equivalent to that of the OrzO wild type (WT) at T_{15} and T_{60}, suggesting that the failure of the OrzO 8(1)9 mutant to rescue is due to mutating the first nucleotide of the V1 region (Figure 2.9B, 2.10A). Mutation of the third nucleotide of the V1 region did not affect the repression ability of OrzO, as the OrzO 10(1)7 mutant still fully repressed *zorO* (Figure 2.9C).

The OrzO 9(1)8 mutant (second nucleotide of the V1 mutated) was highly variable in its ability to repress *zorO* (Figure 2.11A), despite generation of the mutant independently multiple times. To clarify whether this may be due to expression problems, we performed a more in-depth expression comparison to wild type OrzO. We first examined the basal level of OrzO induction required to repress *zorO*, and determined that this occurs at 0.1 mM IPTG induction (Figure 2.12A, 2.12B, 2.13C). We then compared expression of OrzO 9(1)8 when induced at 1 mM IPTG to wild type OrzO expressed at 0.1mM IPTG (Figure 2.11B, 2.11C). Indeed, expression of OrzO 9(1)8 is higher than the basal levels of wild type OrzO that is needed for *zorO* repression. This suggests that the attenuated repression ability observed in OrzO 9(1)8 is likely due to the nucleotide change rather than the expression difference. Regardless, our results imply that the first nucleotide of the V1 region is most critical for the OrzO sRNA to regulate *zorO*.

15 nt of continuous base pairing interactions are sufficient for OrzO regulation We have shown that the 5' base pairing domain of OrzO is critical for regulating *zorO* and the V1 region dictates the correct recognition of the cognate *zor* targets. However, it is not clear whether all nucleotides within the 5' end of OrzO besides the V1 region are needed for repression of *zorO*. To better refine the requirements for successful base pairing, we generated two deletion constructs in which the first 3 and 6 nt were deleted from the 5' end of OrzO while leaving the V1 intact. The OrzO Δ3 and Δ6 share a continuous stretch of 15 nt and 12 nt complementarity with *zorO* mRNA, respectively (Table 2.3).

Rescue experiments were used to assess the abilities of these two mutants to repress *zorO*. Neither the OrzO Δ3 nor Δ6 completely repressed *zorO* when we used 0.2% arabinose to induce expression of the toxin (Figure 2.13A, data not shown). We also inspected the levels of OrzO Δ3 and Δ6 and observed that their levels were decreased as compared to the level of full-length OrzO when all were induced with 1mM IPTG (data not shown).

Upon closer examination of the rescue experiment using OrzO Δ3, we noticed a "transient" rescue or a delay in growth stasis compared to expression of *zorO* alone when induced with 0.2% arabinose. When we reduced the arabinose concentration to 0.002%, we noted that overproduction of ZorO still caused cell growth stasis, in a pattern identical to the 0.2% arabinose induction group (Figure 2.12D, 2.12E). In addition, co-expression of OrzO Δ3 relieved the toxicity caused by induction of *zorO* at 0.002%, suggesting that the failure of this mutant to rescue at higher

concentration of arabinose was likely due to the reduced levels of this sRNA mutant (Figure 2.13A). However, OrzO Δ6 still could not repress toxicity when ZorO was induced by 0.002% arabinose (Figure 2.13C).

We observed that the RNA levels of OrzO Δ3 and Δ6 (induced by 1 mM IPTG) were higher than the basal levels of wild type OrzO (induced by 0.1 mM IPTG) at T_{15} (Figures 2.13B, 2.13D, 2.10B, 2.10C). The levels of the wild type sRNA continued to increase over time (T_{30} and T_{60}), whereas the levels of the truncated sRNAs remained relatively steady. Even though OrzO Δ3 did not accumulate to a level as high as the wild type sRNA, it was capable of repressing *zorO*. OrzO Δ6, however, could not even transiently rescue cells, despite being induced initially to levels higher than wild type OrzO (Figure 2.13D, 2.10C). Combined, these data imply that 15 nt, but not 12 nt, of continuous base pairing is sufficient for the OrzO sRNA to repress *zorO*.

A minimum of 17 nt of discontinuous base pairing interactions is required for OrzO repression of *zorO* As shown above, 15 nt of continuous complementarity to *zorO* is sufficient for repression. Our OrzP V1 mutant can also repress *zorO*, but it does not have 15 nt of continuous base pairing (Table 2.3). Instead, this mutant has 17 nt of discontinuous pairing. This suggests that the amount of complementary base pairs required for discontinuous base pairing interactions may be different from that for continuous base pairing interactions.

Given that the OrzP V1 (pairing: 5(1)12) can rescue cells from *zorO*-induced toxicity, we designed series of mutants with decreased base pairing potential as compared to OrzP V1 (Table 2.3). We constructed OrzO 5(2)11 with a total of 16 nt of base pairing, which is composed of a 5 nt stretch of complementarity, followed by a 2 nt gap and then another 11 nt stretch of complementarity. The OrzO 5(2)11 mutant could not repress ZorO toxicity when the mRNA was induced with either 0.2% or 0.002% arabinose (Figure 2.14A and data not shown). We also confirmed that the OrzO 5(2)11 mutant (induced by 1 mM IPTG) was being expressed at levels higher than the basal levels of OrzO WT (induced by 0.1 mM IPTG) at T_{15} and T_{30} (Figure 2.14B, 2.10D), suggesting that the failure of OrzO 5(2)11 to rescue is likely not due to insufficient RNA levels, but rather base pairing differences.

Our OrzO 5(2)11 possesses a total of 16 nt of pairing to *zorO*, whereas our OrzP V1 possesses 17 nt of pairing (5(1)12). It maybe that a minimum of 17 nt of discontinuous base pairing is required for repression. To explore this further, we generated several more constructs. The OrzO

6(1)11 mutant has a total of 17 nt of pairing, and indeed it can fully repress *zorO* (Figure 2.14C). We then mutated the last nucleotide of pairing of this construct, generating the OrzO 6(1)10(1) mutant, which possess a total of 16 nt of pairing. We also generated the OrzO (1)5(1)11 mutant, in which the first nucleotide of pairing was mutated, and in total, possesses only 16 nt of pairing. When either mutant was used in a rescue experiment, they were unable to repress toxicity, even if *zorO* was induced with 0.002% arabinose (Figure 2.14D). Furthermore, northern analysis indicated that these mutants (induced by 1 mM IPTG) were expressed at levels equivalent to or even higher than OrzO WT induced by 0.1 mM IPTG (Figure 2.14E, 2.14F, 2.10E, 2.10F). Thus, the failure to rescue is likely due to insufficient pairing and not due to insufficient production of the antitoxin.

IV. Discussion

In this study, we show that the *zorO-orzO* gene pair of *E. coli* O157:H7 is a true type I toxin-antitoxin locus. More specifically, our data are consistent with the properties of an unconventional type I system, as the OrzO antitoxin is encoded divergently from *zorO* and has limited base pairing potential with it. Overproduction of ZorO is toxic, however, co-expression of OrzO represses this toxicity and rescues the cells. We discovered that the 5' end of OrzO can solely repress *zorO*-induced toxicity, and the V1 region dictates the correct recognition of *zorO*. Furthermore, the first base of the V1 region is critical for the OrzO sRNA to regulate the expression of *zorO* mRNA. Although there is 18 nt complementarity with *zorO* at the 5' end of OrzO, not all base pairing interactions are needed for repression. For continuous complementarity, 15 nt of perfect base pairing is sufficient to function, while for discontinuous complementarity, a minimum of 17 nt of perfect base pairing is required to maintain the repression ability if a single non-pairing nucleotide is present within the base pairing region.

Role of RNase III in repression of *zorO* expression Ribonuclease III (RNase III), encoded by the *rnc* gene, cleaves double-stranded RNA, and has been linked to processing of several type I toxin-antitoxin pairs found in *E. coli* and *B. subtilis* (16,20,34). While RNase III is not essential for growth in *E. coli*, deletion of the gene in *B. subtilis* is lethal. Work by Durand, et al. has shown that the essentiality of RNase III in *B. subtilis* is because it cleaves the type I toxins *txpA* and *yonT* (38). Interestingly, while the RNA duplex of another type I toxin-antitoxin locus in *B. subtilis*, BsrG/SR4, is cleaved by RNAse III, repression of the toxin does not require this ribonuclease (16).

Thus, while RNase III plays a role in the processing of several described type I pairs in both *E. coli* and *B. subtilis*, its essentiality for repression can be variable between the different systems.

RNase III is also involved in the processing of other RNA duplexes that are not type I toxin-antitoxin systems. These include the interaction of the mRNA *gadX-gadW* with the sRNA GadY and the interaction of the mRNA *cII-O* with the OOP RNA of λ (35,37,39). For the *zorO-orzO* pair, deletion of RNase III led to accumulation of the full-length *zorO* mRNA, and in this strain background, *zorO*-induced toxicity could not be repressed by co-expression of OrzO. Yet, we still could detect OrzO-dependent processing of *zorO* in a *Δrnc* strain. Similar observations were seen with *gadX-gadW* and *cII-O*; cleavage still occurred without RNase III, although the processing sites were slightly shifted (35,37,39). We observed a similar shift in the processing of *zorO* in a *Δrnc* strain (Figure 2.1A, 2.4B). There are other ribonucleases within *E. coli* that may be contributing to this processing [reviewed in (40)]. For example, RNaseE is critical for the function of many Hfq-dependent base pairing RNAs [reviewed in (41)]. Further analysis is needed to determine what other enzymes may be contributing to processing of *zorO*-OrzO; however, based on the in depth analysis of the *gadX-gadW* processing by GadY, it is possible that an uncharacterized double-stranded ribonuclease may exist (35).

We were surprised to note that the levels of OrzO were reduced in a *Δrnc* strain; we have observed that the chromosomal expression levels of OrzO within an *E. coli* O157:H7 deleted for *rnc* are also decreased in comparison to a wild type strain (Wen and Fozo, unpublished observations). We had anticipated that the levels of OrzO would either be unaffected by the loss of RNase III or perhaps higher, as there would be decreased cleavage of the *zorO*-OrzO RNA duplex. It maybe that RNase III cleaves OrzO directly, and that this cleavage impacts its stability and/or activity. However, we have been unable to detect alternative forms of OrzO, and cannot conclude that this is occurring.

Regulation of ZorO by the 5' end of OrzO To determine whether the 5' end of OrzO was sufficient for regulating *zorO*, we generated two constructs expressing either the first 42 or 34 nt of OrzO. These truncated sRNAs did not rescue, nor were we ever able to detect their overexpression (data not shown). However, our chimeric sRNAs, OrzO-OhsC (75 nt) and OrzO-IstR (72 nt) were stable and able to rescue *E. coli* from ZorO toxicity. Thus, although the 5' end of OrzO is sufficient for repression, additional RNA sequence at the 3' end may be needed for

stability. Many type I antitoxins are predicted to have extensive secondary structure (7). The reason for such extensive structure among the type I antitoxins has not been thoroughly examined, but our data suggests that it may be required to prevent quick degradation of the sRNA.

Our data shows that the OrzO sRNA represses *zorO* expression through base pairing interactions between the 5' end of OrzO and the 5' UTR of *zorO*. The binding region in the 5' end appears to be conserved across many sRNAs. For example, a previous study that examined 18 well-characterized *E. coli/Salmonella* Hfq-dependent sRNAs found that more than one third used their 5' ends to repress their targets (42). In addition, the Hfq-independent sRNA IstR-1 also acts by 5' end pairing [reviewed in (19)]. Similar interactions have been predicted to occur with the *shoB-ohsC* pair, in which the 5' end of OhsC shares 19 nt of complementarity to the 5' UTR of *shoB* [(18) and reviewed in (19)]. Although the reason why so many sRNAs function through 5' pairing remains unclear, it has been proposed that similar to eukaryotic microRNAs, some bacterial sRNAs may have a 5' conserved "seed" region that facilitates the selection for correct mRNA targets (42). Thus, the OrzO sRNA likely uses the first 18 nt of its 5' end as the "seed" region. Further sequence analysis studies of a larger group of sRNAs and their targets will be helpful to evaluate whether the existence of the "seed" region can be applied as a general principle in the prediction of bacterial 5' pairing sRNAs and their target mRNAs.

Recognition of *zorO* by the V1 region of OrzO We have shown specifically that the V1 region of OrzO plays a crucial role in recognizing its target. Replacing the V1 region of OrzO with that of OrzP abrogates the ability of the sRNA to repress the *zorO* mRNA, whereas replacing the V1 region of OrzP with that of OrzO allows the mutated OrzP to repress *zorO*.

When inspecting the predicted structures of OrzO and *zorO* (Figure 2.15A, 2.15B), we find that both the V1 region of OrzO and the V1 target site in *zorO* are located in single-stranded structures. The relatively relaxed property of these single-stranded structures may allow for enhanced access of the single-stranded nucleotides to pair with their targets compared to those in a stem structure. Thus, it is likely that the pairing between the V1 region of OrzO sRNA and the corresponding site of *zorO* mRNA occurs prior to that of other complementary nucleotide regions.

In addition, when we mutated the bases of the V1 region, we noted that mutation of the first nucleotide of the V1 region caused a complete loss of the repression of *zorO* without affecting the levels of the sRNA. However, other single base mutants constructed in this study all

successfully repress *zorO* (Table 2.3). These results suggested that as compared to other nucleotides, pairing via the first nucleotide of the V1 region is most critical for the OrzO sRNA to regulate *zorO* mRNA expression.

These observations suggest that OrzO and *zorO* may interact through a "kissing complex". Generally, in a "kissing complex", the binding regions of the sRNA and the mRNA are located in loop structures (single-stranded). The initial base pairing occurs at the short exposed regions; thus, mutations of nucleotides involved the primary "kissing" region impact the repression outcome, as exemplified by the RNAII-RNAI pair in ColE1-derived plasmid and the CopA-CopT pair in R1 plasmid (43,44). Once the initial pairing is established, the kissing complex can extend and propagate the formation of double-stranded RNA in the rest of the sequence. Similarly, this has been reported for the interaction of the *ibs*-Sib type I toxin-antitoxin pair in *E. coli*; however, in this system, there are two recognition sites, in *zorO-orzO*, there appears to be only one recognition domain (45).

While these previously described systems seem to corroborate the results of this study, structure-probing and kinetic experiments need to be performed to validate the predictions and to define the order of occurrence of base pairing in different regions.

Multiple factors contribute to the repression ability of Orz Antitoxin sRNAs function by base pairing to their target mRNAs. To successfully repress the toxin mRNA, an adequate amount of functional antitoxin sRNA is required. Otherwise, even if the sRNA is capable of repressing its mRNA target, insufficient RNA levels will result in partial or complete loss of repression.

We observed this directly with our OrzO Δ3 mutant, the expression levels of which were lowered when compared to full-length OrzO. Because all of our Orz constructs are under the control of the same promoter (IPTG-inducible $P_{Llac\text{-}O1}$), the decreased levels of the OrzO Δ3 are likely due to accelerated degradation. When performing the rescue experiment using the OrzO Δ3 mutant, an initial repression of *zorO* was observed when we induced the toxin with 0.2% arabinose, but this repression was not long lasting (Figure 2.13A). Because the antitoxin sRNA is induced 30 minutes prior to that of the toxin mRNA, it is likely that enough OrzO Δ3 sRNA had accumulated to initially repress *zorO*; however, as the "pre-made" OrzO Δ3 was consumed through base pairing with *zorO*, the newly produced OrzO Δ3 was not sustained at a high enough level to maintain *zorO* repression. When the inducer concentration was decreased from 0.2% to 0.002% to reduce *zorO*

levels, the OrzO Δ3 mutant was able to repress *zorO*, implying that the levels of the sRNA were now sufficient for repression.

It is also likely that the structure and thus, specific nucleotide sequence, of OrzO contributes as well to its functional ability. For example, several of our mutants showed reduced expression levels at 1 mM IPTG when compared to the wild type at the same IPTG induction. As stated, all of our OrzO constructs were controlled by the same inducible promoter, so transcription should not be affected. Thus, it is more likely that these constructs are more unstable than wild type OrzO. As shown in Figure 2.15, much of the 5' end of OrzO (region of base pairing) is predicted to be in a stem structure. Disruption of this long stem, even by a single nucleotide, may impact overall stability of the sRNA. Further analysis of the effect of these mutations on the structure and degradation of OrzO is needed to resolve this.

Aside from RNA levels, the base pairing potential of the Orz sRNA also contributes to its ability to repress *zorO*. Our data show that 15 nt of continuous complementarity at 5' end of OrzO are sufficient for the repression of *zorO* mRNA, as represented by the OrzO Δ3 mutant. However, unlike continuous base pairing, in which the pairing region of the sRNA perfectly matches the target sequence, unpaired nucleotides would interrupt the formation of the sRNA-mRNA duplex, affecting repression of the mRNA. Thus, it is likely that more base pairing interactions are needed to compensate for the instability of the RNA duplex caused by the unpaired base(s). The OrzO 6(1)10(1) and 6(1)11 mutants support this theory. Both RNAs were produced at levels similar to wild type OrzO, but only Orzo 6(1)11 was able to successfully repress ZorO toxicity. The main difference between OrzO 6(1)10(1) and OrzO 6(1)11 comes from the total number of base pairing interactions, 16 versus 17, respectively. Additionally, the OrzP V1 mutant (pairing: 5(1)12) also successfully represses *zorO*. Combined, these data suggest that the total amount of pairing impacts the repression ability, rather than specific positioning of nucleotides (outside the V1 region).

How does binding to the 5' UTR repress *zorO* toxicity? The region of complementarity in the *zorO* mRNA is located in its long 5' UTR, approximately 73-91 nt upstream of the start codon. How could binding to this region prevent *zorO* expression? It is important to note that regulation of the type I toxins Hok and TisB by their respective antitoxins occurs through pairing far upstream of the start codons (46,47). Translation of *hok* requires the translation of a small peptide (encoded by *mok*) found within the 5' UTR of the *hok* mRNA. The antitoxin, Sok, is complementary to *mok*,

and hence binding of the sRNA prevents translation of *mok*, and consequently, *hok* (47). For *tisB*, its ribosome binding site is inaccessible due to its secondary structure. However, within its 5' UTR is a stand-by ribosome binding site; here the ribosome can bind, and as the structure breathes, move to the true ribosome binding site for translating *tisB*. The antitoxin, IstR-1, has sequence complementarity that overlaps the stand-by site; thus, formation of a *tisB*-IstR-1 RNA duplex would prevent ribosome binding (46).

The *zorO* mRNA has 5' UTR that is approximately 180 nt in length (Figure 2.1A); within this UTR are several potential small open reading frames and ribosome binding sites. It is possible that *zorO* translation is mediated in a manner similar to either *hok* or *tisB*, and that binding by OrzO may block either translation of an internal peptide or binding of the ribosome to a stand-by site. Ongoing experiments will help elucidate how binding of the antitoxin so far upstream from the ATG start site successfully represses *zorO* expression.

Role of the *zor-orz* locus? A major question regarding many type I toxin-antitoxin loci is what is their true biological function? This is particularly true for the chromosomally encoded loci. Work has implicated the role of the type I toxin TisB in halting cell division during the SOS-response, allowing the cell time to repair damaged DNA, as well as playing a role in persister cell formation (48,49). Several of the type I loci described in *B. subtilis* exist on prophages [reviewed in (50)]. In particular, *txpA-ratA* is located on the *skin* element, which is excised during sporulation (17). It has been suggested that TxpA helps maintain the *skin* element within *B. subtilis*. For the BsrG/SR4 locus, data has shown that degradation of the *bsrG* toxin mRNA is accelerated under high temperatures (16). How this may influence the function of the toxin has not yet been elucidated.

In the case of the *zor-orz* locus, it is not found beyond *E. coli* and *Shigella* species, and is absent from laboratory strains of *E. coli* like MG1655 (14). This implies that the locus may have been lost upon domestication of *E. coli* and perhaps the *zor-orz* locus is needed for growth in the normal habitat *of E. coli*, the intestine. Furthermore, previous work showed that there are differences in the chromosomal expression of *zorO* and *orzO* during growth with glucose as the sole carbon source (14). We have confirmed that the endogenous expression of the sRNA and the mRNA is sensitive to carbon source, implying a possible role for the locus in cellular metabolism (Wen, Fozo, unpublished observations). Thus, the *zor-orz* locus may play a role in controlling cellular growth in response to changing nutrient conditions in the native environment. Current

work examining the effects of deleting the locus on total cellular metabolism and gene expression are ongoing to decipher how these genes may impact overall cellular growth.

Overall, our data indicates that the 5' end of OrzO is sufficient for repressing *zorO*; that the V1 region dictates Orz sRNA specificity for its target; and that there appears to be a minimum length required for base pairing. Further investigations are aimed at understanding the precise regulatory mechanisms of this untraditional type I toxin-antitoxin system to better our understanding of the intricacies of RNA regulation.

References

1. Liu, M.Y., Gui, G., Wei, B., Preston, J.F., 3rd, Oakford, L., Yuksel, U., Giedroc, D.P. and Romeo, T. (1997) The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in *Escherichia coli*. *J. Biol. Chem.*, **272**, 17502-17510.
2. Waters, L.S. and Storz, G. (2009) Regulatory RNAs in bacteria. *Cell*, **136**, 615-628.
3. Georg, J. and Hess, W.R. (2011) *cis*-Antisense RNA, another level of gene regulation in bacteria. *Microbiol. Mol Biol. Rev.*, **75**, 286-300.
4. Thomason, M.K. and Storz, G. (2010) Bacterial antisense RNAs: how many are there, and what are they doing? *Ann. Rev. Gen.*, **44**, 167-188.
5. Brantl, S. (2007) Regulatory mechanisms employed by *cis*-encoded antisense RNAs. *Curr. Opin. Microbiol.*, **10**, 102-109.
6. Gerdes, K., Larsen, J.E.L. and Molin, S. (1985) Stable inheritance of plasmid R1 requires two different loci. *J. Bacteriol.*, **161**, 292-298.
7. Fozo, E.M., Hemm, M.R. and Storz, G. (2008) Small toxic proteins and the antisense RNAs that repress them. *Microbiol. Mol. Biol. Rev.*, **72**, 579-589.
8. Weaver, K.E., Jensen, K.D., Colwell, A. and Sriram, S.I. (1996) Functional analysis of the *Enterococcus faecalis* plasmid pAD1-encoded stability determinant *par*. *Mol. Microbiol.*, **20**, 53-63.
9. Weaver, K.E. and Tritle, D.J. (1994) Identification and characterization of an *Enterococcus faecalis* plasmid pAD1-encoded stability determinant which produces two small RNA molecules necessary for its function. *Plasmid*, **32**, 168-181.
10. Greenfield, T.J., Ehli, E., Kirshenmann, T., Franch, T., Gerdes, K. and Weaver, K.E. (2000) The antisense RNA of the *par* locus of pAD1 regulates the expression of a 33-amino-acid toxic peptide by an unusual mechanism. *Mol. Microbiol.*, **37**, 652-660.
11. Faridani, O.R., Nikravesh, A., Pandey, D.P., Gerdes, K. and Good, L. (2006) Competitive inhibition of natural antisense Sok-RNA interactions activates Hok-mediated cell killing in *Escherichia coli*. *Nucleic Acids Res.*, **34**, 5915-5922.
12. Pedersen, K. and Gerdes, K. (1999) Multiple *hok* genes on the chromosome of *Escherichia coli*. *Mol. Microbiol.*, **32**, 1090-1102.

13. Weaver, K.E., Reddy, S.G., Brinkman, C.L., Patel, S., Bayles, K.W. and Endres, J.L. (2009) Identification and characterization of a family of toxin-antitoxin systems related to the *Enterococcus faecalis* plasmid pAD1 *par* addiction module. *Microbiology*, **155**, 2930-2940.

14. Fozo, E.M., Makarova, K.S., Shabalina, S.A., Yutin, N., Koonin, E.V. and Storz, G. (2010) Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families. *Nucleic Acids Res.*, **38**, 3743-3759.

15. Kawano, M., Oshima, T., Kasai, H. and Mori, H. (2002) 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.*, **45**, 333-349.

16. Jahn, N., Preis, H., Wiedemann, C. and Brantl, S. (2012) BsrG/SR4 from *Bacillus subtilis*--the first temperature-dependent type I toxin-antitoxin system. *Mol. Microbiol.*, **83**, 579-598.

17. Silvaggi, J.M., Perkins, J.B. and Losick, R. (2005) Small untranslated RNA antitoxin in *Bacillus subtilis*. *J. Bacteriol.*, **187**, 6641-6650.

18. Fozo, E.M., Kawano, M., Fontaine, F., Kaya, Y., Mendieta, K.S., Jones, K.L., Ocampo, A., Rudd, K.E. and Storz, G. (2008) Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. *Mol. Microbiol.*, **70**, 1076-1093.

19. Fozo, E.M. (2012) New type I toxin-antitoxin families from "wild" and laboratory strains of *E. coli*: Ibs-Sib, ShoB-OhsC and Zor-Orz. *RNA Biol.*, **9**, 1504-1512.

20. Vogel, J., Argaman, L., Wagner, E.G. and Altuvia, S. (2004) The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Curr. Biol.,* **14**, 2271-2276.

21. Guillier, M. and Gottesman, S. (2006) Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Mol. Microbiol.*, **59**, 231-247.

22. Khlebnikov, A., Datsenko, K.A., Skaug, T., Wanner, B.L. and Keasling, J.D. (2001) Homogeneous expression of the $P_{(BAD)}$ promoter in *Escherichia coli* by constitutive expression of the low-affinity high-capacity AraE transporter. *Microbiology*, **147**, 3241-3247.

23. Cherepanov, P.P. and Wackernagel, W. (1995) Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene*, **158**, 9-14.

24. Court, D.L., Swaminathan, S., Yu, D., Wilson, H., Baker, T., Bubunenko, M., Sawitzke, J. and Sharan, S.K. (2003) Mini-lambda: a tractable system for chromosome and BAC engineering. *Gene*, **315**, 63-69.

25. Datsenko, K.A. and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA*, **97**, 6640-6645.

26. Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G. and Court, D.L. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*, **97**, 5978-5983.

27. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, **77**, 51-59.

28. Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K. and Pease, L.R. (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene*, **77**, 61-68.

29. Guzman, L.M., Belin, D., Carson, M.J. and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J. Bacteriol*, **177**, 4121-4130.

30. Masse, E., Escorcia, F.E. and Gottesman, S. (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev.*, **17**, 2374-2383.

31. Opdyke, J.A., Kang, J.G. and Storz, G. (2004) GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J. Bacteriol.*, **186**, 6698-6705.

32. Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, **9**, 671-675.

33. Zhang, A., Wassarman, K.M., Rosenow, C., Tjaden, B.C., Storz, G. and Gottesman, S. (2003) Global analysis of small RNA and mRNA targets of Hfq. *Mol. Microbiol.*, **50**, 1111-1124.

34. Gerdes, K., Nielsen, A., Thorsted, P. and Wagner, E.G. (1992) Mechanism of killer gene activation. Antisense RNA-dependent RNase III cleavage ensures rapid turn-over of the stable *hok*, *srnB* and *pndA* effector messenger RNAs. *J. Mol. Biol.*, **226**, 637-649.

35. Opdyke, J.A., Fozo, E.M., Hemm, M.R. and Storz, G. (2011) RNase III participates in GadY-dependent cleavage of the *gadX-gadW* mRNA. *J. Mol. Biol*, **406**, 29-43.

36. Krinke, L., Mahoney, M. and Wulff, D.L. (1991) The role of the OOP antisense RNA in coliphage lambda development. *Mol. Microbiol.*, **5**, 1265-1272.

37. Krinke, L. and Wulff, D.L. (1987) OOP RNA, produced from multicopy plasmids, inhibits lambda *cII* gene expression through an RNase III-dependent mechanism. *Genes Dev.*, **1**, 1005-1013.

38. Durand, S., Gilet, L. and Condon, C. (2012) The essential function of *B. subtilis* RNase III is to silence foreign toxin genes. *PLoS Genet.*, **8**, e1003181.

39. Krinke, L. and Wulff, D.L. (1990) RNase III-dependent hydrolysis of lambda *cII-O* gene mRNA mediated by lambda OOP antisense RNA. *Genes Dev.*, **4**, 2223-2233.

40. Arraiano, C.M., Andrade, J.M., Domingues, S., Guinote, I.B., Malecki, M., Matos, R.G., Moreira, R.N., Pobre, V., Reis, F.P., Saramago, M. *et al.* (2010) The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol. Rev.*, **34**, 883-923.

41. Lalaouna, D., Simoneau-Roy, M., Lafontaine, D. and Masse, E. (2013) Regulatory RNAs and target mRNA decay in prokaryotes. *Biochim. Biophys. Acta*, **1829**, 742-747.

42. Papenfort, K., Bouvier, M., Mika, F., Sharma, C.M. and Vogel, J. (2010) Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. *Proc. Natl. Acad. Sci. USA*, **107**, 20435-20440.

43. Hjalt, T. and Wagner, E.G. (1992) The effect of loop size in antisense and target RNAs on the efficiency of antisense RNA control. *Nucleic Acids Res.*, **20**, 6723-6732.

44. Lin-Chao, S. and Cohen, S.N. (1991) The rate of processing and degradation of antisense RNAI regulates the replication of ColE1-type plasmids in vivo. *Cell*, **65**, 1233-1242.

45. Han, K., Kim, K.S., Bak, G., Park, H. and Lee, Y. (2010) Recognition and discrimination of target mRNAs by Sib RNAs, a *cis*-encoded sRNA family. *Nucleic Acids Res.*, **38**, 5851-5866.

46. Darfeuille, F., Unoson, C., Vogel, J. and Wagner, E.G. (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Mol. Cell*, **26**, 381-392.

47. Thisted, T. and Gerdes, K. (1992) Mechanism of post-segregational killing by the *hok/sok* system of plasmid R1. Sok antisense RNA regulates *hok* gene expression indirectly through the overlapping *mok* gene. *J. Mol. Biol.*, **223**, 41-54.

48. Dorr, T., Vulic, M. and Lewis, K. (2010) Ciprofloxacin causes persister formation by inducing the TisB toxin *in Escherichia coli*. *PLoS Biol.*, **8**, e1000317.

49. Unoson, C. and Wagner, E.G. (2008) A small SOS-induced toxin is targeted against the inner membrane in *Escherichia coli*. *Mol. Microbiol.*, **70**, 258-270.

50. Durand, S., Jahn, N., Condon, C. and Brantl, S. (2012) Type I toxin-antitoxin systems in *Bacillus subtilis*. *RNA Biol.*, **9**, 1491-1497.

51. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406-3415.

Appendix II: chapter II tables and figures

Table 2.1. Strains and plasmids used in this study

Name	Relevant genotype or description	Source
Escherichia coli EDL933	Wildtype EHEC, O157:H7	D. Friedman
DJ480	MG1655 *ΔlacX74* (wild type strain)	D. Jin, (20)
UTK007	DJ480 PCP18-*araE*	This study
UTK008	UTK007 Δ*hfq::kan*	This study
UTK011	UTK007 Δ*rnc::kan*	This study
Plasmids		
pKD4	Km^R	(24)
pBR-pLac	Amp^R; P_{LacO} promoter	(20)
pEF21	Cm^R; P_{BAD} promoter	(17)
pBR-pLac-*orzO*	AmpR	This study
pEF21-*zorO*	Cm^R	This study
pBR-pLac-*orzO*-*ohsC*	Amp^R	This study
pBR-pLac-*orzO*-*istR*	Amp^R	This study
pBR-pLac-*orzP*	Amp^R	This study
pBR-pLac-*orzO* V1	Amp^R	This study
pEF21-*zorO* V1	Cm^R	This study
pBR-pLac-*orzP* V1	Amp^R	This study
pBR-pLac-*orzO* 8(1)9	Amp^R	This study
pBR-pLac-*orzO* 10(1)7	Amp^R	This study
pBR-pLac-*orzO* 9(1)8	Amp^R	This study
pBR-pLac-*orzO*Δ3	Amp^R	This study
pBR-pLac-*orzO*Δ6	Amp^R	This study
pBR-pLac-*orzO* 5(2)11	Amp^R	This study
pBR-pLac-*orzO* 6(1)11	Amp^R	This study
pBR-pLac-*orzO* 6(1)10(1)	Amp^R	This study
pBR-pLac-*orzO* 1(5)1(11)	Amp^R	This study

Table 2.2. Oligonucleotides used in this study

Name	Sequence[a]	Use
EF506	CCTGCTGGTGAGGCAGGTTCTTTTATTT	OrzO Northern analysis
EF511	GAGCACGGTTAACTTTTGTGTCAGCGTGT	*zorO* Primer extension analysis
EF517	CTACAGTGACGCTAACTTGCTGATT	OrzO Northern analysis
EF524	GTTGGTACGAAACGTTGCTCTCCG	*zorO* Northern analysis
EF769	CAGAATCGAAAGGTTCAAAGTACAAATAAGCATATAAGGAAAAGAGAGAATGAT TGAACAAGATGGATTG	UTK008 PCR (Δ*hfq*)
EF770	CGCAGGATCGCTGGCTCCCCGTGTAAAAAAAACAGCCCGAAACCTTAGAAGAACT CGTCAAGAAG	UTK008 PCR (Δ*hfq*)
EF771	GTCTGTTTCGTGTGCTGAATTGTTGACGCATTTATTTATTGGTATCGCATGATTGAA CAAGATGGA	UTK011 PCR (Δ*rnc*)
EF772	GATGGCAATAAATCCGCAGTAACTTTTATCGATGCTCATTCCAGCTCCAGTCAGAA GAACTCGTCAAGAAG	UTK011 PCR (Δ*rnc*)
EF537	GTGGATATGGTTATAATGACGTCGTTGTTACGAA	pBR-pLac-*orzO* PCR
EF538	GATGTGGAATTCTATGGTGGGCTATTGGC	pBR-pLac-*orzO* PCR; pBR-pLac-*orzO-ohsC* PCR; pBR-pLac-*orzO-istR* PCR; pBR-pLac-*orzO*Δ3 PCR, pBR-pLac-*orzO*Δ6 PCR
EF531	GAGATTTATAATCCTGCAGGTTGGGACGTTGC	pEF21-*zorO* PCR
EF532	GAGTTAAAGCTTGAATTAATAAAAACAGTAATC	pEF21-*zorO* PCR
EF78	GACACGGAAATGTTGAATAC	pBR-pLac-orzO-ohsC PCR; pBR-pLac-orzO-istR PCR
EF790	CTTTTTAACTTTAATTTTGCAACATTTCGTACCAACGACGTC	pBR-pLac-*orzO-ohsC* PCR
EF791	GAAATGTTGCAAAATTAAAGTTAAAAAGTAAAACCCCCGTTCC	pBR-pLac-*orzO-ohsC* PCR
EF792	GCCACGGTAA GAATTC AAAGTTAAAAATAATACCG	pBR-pLac-*orzO-ohsC* PCR
EF793	GCGGCTGGTAACCGCAGCAACATTTCGTACCAACGACGTC	pBR-pLac-*orzO-istR* PCR
EF794	CGAAATGTTGCTGCGGTTACCAGCCGCGGGCGGCTGACG	pBR-pLac-*orzO-istR* PCR
EF795	CTGTACTGCAGAATTCAAAAAACCCCGCCGGAGCG	pBR-pLac-*orzO-istR* PCR

Table 2.2. (continued)

Name	Sequence[a]	Use
EF541	CCAGCCGTGGTTATAATGACGTCGTTGGAACACG	pBR-pLac-*orzP* PCR
EF542	GTTGTTGAATTCAAAAATATCTGGG	pBR-pLac-*orzP* PCR
EF700	CTGACGTCGTTGGTACACGATGTTGCACAATCAGC	pBR-pLac-*orzO* V1 PCR
EF701	GCTGATTGTGCAACATCGTGTACCAACGACGTCAG	pBR-pLac-*orzO* V1 PCR
EF725	GATCCGGAGAGCAACGTCGTGTACCAACATACGTAAAC	pEF21-*zorO* V1 PCR
EF726	GTTTACGTATGTTGGTACACGACGTTGCTCTCCGGATC	pEF21-*zorO* V1 PCR
EF702	CTGACGTCGTTGGAACGAAATGTTGCACAGGCTGTG	pBR-pLac-*orzP* V1 PCR
EF703	CACAGCCTGTGCAACATTTCGTTCCAACGACGTCAG	pBR-pLac-*orzP* V1 PCR
EF998	CTGACGTCGTTGGTACCAAATGTTGCACAATCAG	pBR-pLac-*orzO* 8(1)9 PCR
EF999	CTGATTGTGCAACATTTGGTACCAACGACGTCAG	pBR-pLac-*orzO* 8(1)9 PCR
EF976	GACGTCGTTGGTACGACATGTTGCACAATCAGCAAG	pBR-pLac-*orzO* 10(1)7 PCR
EF977	CTTGCTGATTGTGCAACATGTCGTACCAACGACGTC	pBR-pLac-*orzO* 10(1)7 PCR
EF745	CTGACGTCGTTGGTACGCAATGTTGCACAATCAGC	pBR-pLac-*orzO* 9(1)8 PCR
EF746	GCTGATTGTGCAACATTGCGTACCAACGACGTCAG	pBR-pLac-*orzO* 9(1)8 PCR
EF853	GGTTATAATTTA GACGTC GGTACGAAATGTTGCACAATC	pBR-pLac-*orzO*Δ3 PCR
EF875	GGTTATAATTTATTC GACGTC ACGAAATGTTGCACAATC	pBR-pLac-*orzO*Δ6 PCR
EF940	GATACTGACGTCGTTGGATCGAAATGTTGCACAATCAGC	pBR-pLac-*orzO* 5(2)11 PCR
EF941	GCTGATTGTGCAACATTTCGATCCAACGACGTCAGTATC	pBR-pLac-*orzO* 5(2)11 PCR
EF934	GATACTGACGTCGTTGGTTCGAAATGTTGCACAATCAGC	pBR-pLac-*orzO* 6(1)11 PCR
EF935	GCTGATTGTGCAACATTTCGAACCAACGACGTCAGTATC	pBR-pLac-*orzO* 6(1)11 PCR
EF972	CGTCGTTGGTTCGAAATGTTGACAATCAGCAAGTTAGCG	pBR-pLac-*orzO* 6(1)10(1) PCR
EF973	CGCTAACTTGCTGATTGTCAACATTTCGAACCAACGACG	pBR-pLac-*orzO* 6(1)10(1) PCR
EF944	GATACTGACGTCCTTGGTTCGAAATGTTGCACAATCAGC	pBR-pLac-*orzO* 1(5)1(11) PCR
EF945	GCTGATTGTGCAACATTTCGAACCAAGGACGTCAGTATC	pBR-pLac-*orzO* 1(5)1(11) PCR

[a] 5' - 3', restriction sites underlined, nucleotides altered via site directed mutagenesis in red.

Table 2.3. Summary of the mutants generated and their repressive abilities

sRNA	Target mRNA	Repression ability[a]	Total number of perfect base pairing	Base pairing pattern	Base pairing
OrzO WT	*zorO* WT	Full	18	18	**OrzO WT** 5′ GUUGGUACGAAAUGUUGC 3′ ||||||||||||:||||| ***zorO* WT** 3′ CAACCAUGCUUUGCAACG 5′
OrzO V1	*zorO* WT	No	16	8(2)8	**OrzO V1** 5′ GUUGGUACACGAUGUUGC 3′ |||||||| :|:||||| ***zorO* WT** 3′ CAACCAUGCUUUGCAACG 5′
OrzP WT	*zorO* WT	No	15	5(1)2(2)8	**OrzP WT** 5′ GUUGGAACACGAUGUUGC 3′ ||||| || :|:||||| ***zorO* WT** 3′ CAACCAUGCUUUGCAACG 5′
OrzP V1	*zorO* WT	Full	17	5(1)12	**OrzP V1** 5′ GUUGGAACGAAAUGUUGC 3′ ||||| ||||||:||||| ***zorO* WT** 3′ CAACCAUGCUUUGCAACG 5′
OrzO Δ3	*zorO* WT	Repress *zorO* induced by 0.002% arabinose	15	15	**OrzO Δ3** 5′ GGUACGAAAUGUUGC 3′ |||||||||:||||| ***zorO* WT** 3′ CAACCAUGCUUUGCAACG 5′

Table 2.3. (continued)

sRNA	Target mRNA	Repression ability[a]	Total number of perfect base pairing	Base pairing pattern	Base pairing
OrzO Δ6	*zorO* WT	No	12	12	OrzO Δ6 5′ ACGAAAUGUUGC 3′ ||||||:||||| *zorO* WT 3′ CAACCAUGCUUUGCAACG 5′
OrzO 6(1)11	*zorO* WT	Full	17	6(1)11	OrzO 6(1)11 5′ GUUGGUUCGAAAUGUUGC 3′ |||||| |||||:||||| *zorO* WT 3′ CAACCAUGCUUUGCAACG 5′
OrzO 5(2)11	*zorO* WT	No	16	5(2)11	OrzO 5(2)11 5′ GUUGGAUCGAAAUGUUGC 3′ ||||| |||||:||||| *zorO* WT 3′ CAACCAUGCUUUGCAACG 5′
OrzO (1)5(1)11	*zorO* WT	No	16	(1)5(1)11	OrzO (1)5(1)11 5′ CUUGGUUCGAAAUGUUGC 3′ ||||| |||||:|||||| *zorO* WT 3′ CAACCAUGCUUUGCAACG 5′
OrzO 6(1)10(1)	*zorO* WT	No	16	6(1)10(1)	OrzO 6(1)10(1) 5′ GUUGGUUCGAAAUGUUGA 3′ |||||| |||||:|||| *zorO* WT 3′ CAACCAUGCUUUGCAACG 5′

Table 2.3. (continued)

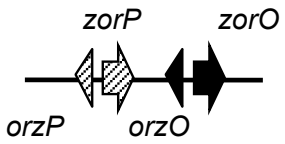
sRNA	Target mRNA	Repression ability[a]	Total number of perfect base pairing	Base pairing pattern	Base pairing
OrzO 10(1)7	*zorO* WT	Full	17	10(1)7	OrzO 10(1)7 5′ GUUGGUACGACAUGUUGC 3′ |||||||||| |:||||| *zorO* WT 3′ CAACCAUGCUUUGCAACG 5′
OrzO 9(1)8	*zorO* WT	ND[b]	17	9(1)8	OrzO 9(1)8 5′ GUUGGUACGCAAUGUUGC 3′ ||||||||| ||:||||| *zorO* WT 3′ CAACCAUGCUUUGCAACG 5′
OrzO 8(1)9	*zorO* WT	No	17	8(1)9	OrzO 8(1)9 5′ GUUGGUACCAAAUGUUGC 3′ |||||||| |||:||||| *zorO* WT 3′ CAACCAUGCUUUGCAACG 5′
OrzO V1	*zorO* V1	Full	18	18	OrzO V1 5′ GUUGGUACACGAUGUUGC 3′ ||||||||||||:||||| *zorO* V1 3′ CAACCAUGUGCUGCAACG 5′

[a] Repression ability when *zorO* is induced by 0.2% arabinose unless otherwise indicated.

[b] Despite generation of this mutant independently multiple times, we never obtained consistent results when analyzing its ability to repress *zorO* induced either with 0.2 or 0.002% arabinose.

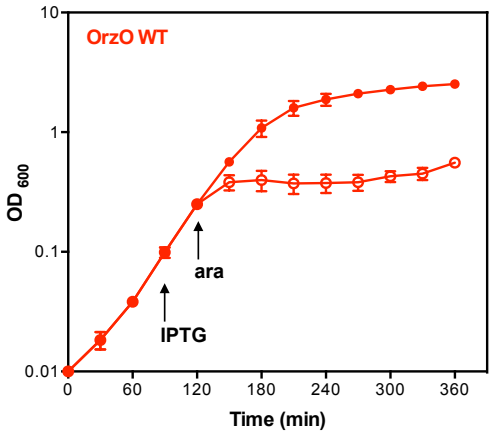
Figure 2.1. The *zorO-orzO* gene locus is a true type I toxin-antitoxin system. (**A**) Genetic organization of the duplicated *zor-orz* locus in *E. coli* EDL933. Sequence details of the *zorO-orzO* genes; region of base pairing is shaded, and previously mapped transcription starts are indicated in bold (14). Predicted -35 and -10 promoter elements are boxed. Black arrows indicate RNA processing products found in both the wild type and *Δrnc* strain; green arrows, wild type strain only; red arrows, *Δrnc* strain only. (**B**) Overexpression of OrzO can repress *zorO*-mediated toxicity. *E. coli* strain UTK007 was transformed with pBR-plac-o*rzO* (induced by IPTG) and pBAD-*zorO* (induced by arabinose). The strain was grown to $OD_{600} \approx 0.1$, and the culture was split into two. IPTG (1 mM final concentration) was added to one half of the culture. Arabinose (0.2% final concentration) was added 30 minutes later to each culture and OD_{600} was measured over time. Solid symbols indicate the addition of IPTG; open symbols, no IPTG. Shown are the mean values ± standard deviations for three independent cultures. Predicted region of base pairing is shown.

A.




```
1    TAAAAAAACCTGCTGGTGAGGCAGGTTCTTTTATTTACTACAGTGACGCTAACTTGCTGATTGTGCAACATTTCGTACCAACGAA  85
     ATTTTTTTGGACGACCACTCCGTCCAAGAAAATAAATGATGTCACTGCGATTGAACGACTAACACGTTGTAAAGCATGGTTGCTT
                                                                                  +1 orzO

                                 -35 zorO              -10 zorO      +1 zorO
86   TAAATTATAACCATATCCACACTCCATTGCAATTCAGTGAGTGAGATTTATAATCCTGCCCGTTGGGAGGTTGTGCCGGATCGGA  171
     ATTTAATATTGGTATAGGTGTGAGGTAACGTTAAGTCACTCACTCTAAATATTAGGACGGGCAACCCTCCAACACGGCCTAGCCT
       -10 orzO                -35 orzO

172  CCAATTTTAAGTCCTGGCTGCCGGACGGGTGGTGCCGCAGGCGGTGCCCTGATCCGGAGAGCAACGTTTCGTACCAACATACGTA  257
     GGTTAAAATTCAGGACCGACGGCCTGCCCACCACGGCGTCCGCCACGGGACTAGGCCTCTCGTTGCAAAGCATGGTTGTATGCAT

                                                                  zorO start codon
342  AACCAGTATCGACAGGCAGTGTAAGGGTAAGGTGCTGGTGTTGCGCTCCTTAATAAGGAGAGCGGATGGACACGCTGACACAAA   437
     TTGGTCATAGCTGTCCGTCACATTCCCATTCCACGACCACAACGCGAGGAATTATTCCTCTCGCCTACCTGTGCGACTGTGTTT

                                                                  zorO stop codon
438  AGTTAACCGTGCTCATTGCCGTACTGGAGTTATTAGTGGCTCTGTTACGGTTGATTGATTTGTTGAAGTAA
     TCAATTGGCACGAGTAACGGCATGACCTCAATAATCACCGAGACAATGCCAACTAACTAAACAACTTCATT
```

B.




```
OrzO 5′ GUUGGUACGAAAUGUUGC 3′
        ||||||||||||:|||||
zorO 3′ CAACCAUGCUUUGCAACG 5′
```

Figure 2.1. (continued)

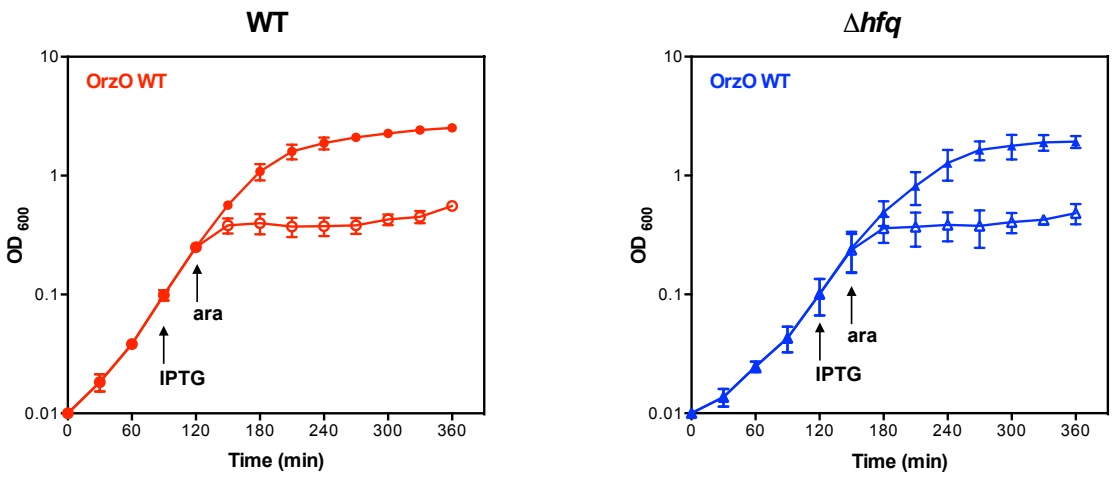


Figure 2.2. The RNA chaperone Hfq is not required for OrzO to repress *zorO*-induced toxicity. *E. coli* wild type strain UTK007 (red) and *E. coli* strain UTK008 (blue), a derivative of the wild type which is deleted for *hfq*, were transformed with pBR-plac-*orzO* and pBAD-*zorO*. The strains were grown to $OD_{600} \approx 0.1$, and split into two. IPTG (1 mM) was added to one half of the culture. Arabinose (0.2%) was added 30 minutes later to each culture and OD_{600} was measured over time. Solid symbols indicate the addition of IPTG; open symbols, no IPTG. Averages ± standard deviations for three independent cultures are shown.

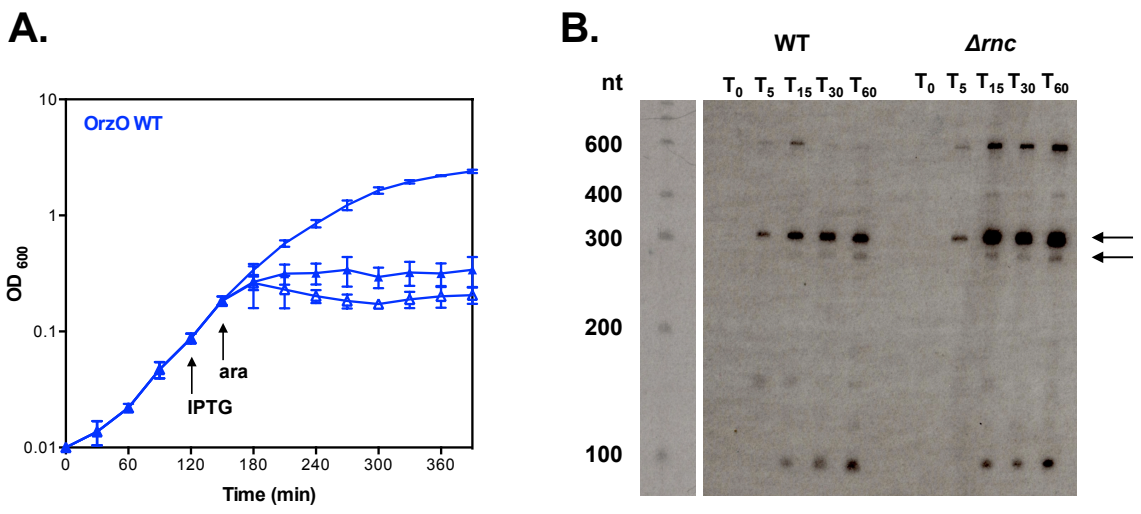


Figure 2.3. RNase III is critical for OrzO repression of *zorO*. (**A**) Overexpression of OrzO cannot repress *zorO*-mediated toxicity in a strain deleted for *rnc*. *E. coli* strain UTK011, which is deleted for *rnc,* was transformed with pBR-plac-*orzO* and pBAD-*zorO*. The strain was grown to $OD_{600} \approx 0.1$, and split into three. One of the three cultures received no supplementation. IPTG (1 mM) was added as indicated to one culture. After 30 minutes, arabinose (0.2%) was added to two of the cultures and cell growth was monitored by measuring the OD_{600} over time. Solid symbols indicate the addition of arabinose and IPTG; open symbols, arabinose and no IPTG; no symbol indicates unsupplemented culture. Shown are the averages ± standard deviations from three independent experiments. (**B**) Accumulation of full-length *zorO* mRNA in the *rnc* deletion strain. Total RNA was isolated from a wild type strain (UTK007) and an *rnc* deletion strain (UTK011) harboring pBAD-*zorO* 0, 5, 15, 30, 60 minutes after the addition of 0.2% arabinose. Arrows indicate the predicted full-length *zorO* transcripts.

Figure 2.4. Effects of an *rnc* deletion on OrzO expression and processing of *zorO*. (A) Expression of OrzO in a Δ*rnc* strain. Total RNA was isolated as described in Materials and Methods from a wild type strain (UTK007) and an *rnc* deletion strain (UTK011) harboring pBAD-*zorO* and pBR-plac-*orzO*. The antitoxin OrzO was induced with 1 mM IPTG, and RNA was isolated 0, 5, 15, 30, 60 minutes after the addition of 0.2% arabinose. Average intensities are shown as described in the Materials and Methods. **(B)** Processing of *zorO* in a wild type strain and a Δ*rnc* strain. RNA was isolated as described in the Materials and Methods 30 minutes after arabinose addition ± induction of OrzO (1 mM IPTG). The sequencing gel was loaded twice; the panel on the left was run longer, to allow better separation of the cleavage products closer to the start of transcription. Black arrows indicate products found in both the wild type and Δrnc strain; green arrows, wild type strain only; red arrows, Δ*rnc* strain only. The black brackets indicate the sequence region overlap of the two panels.

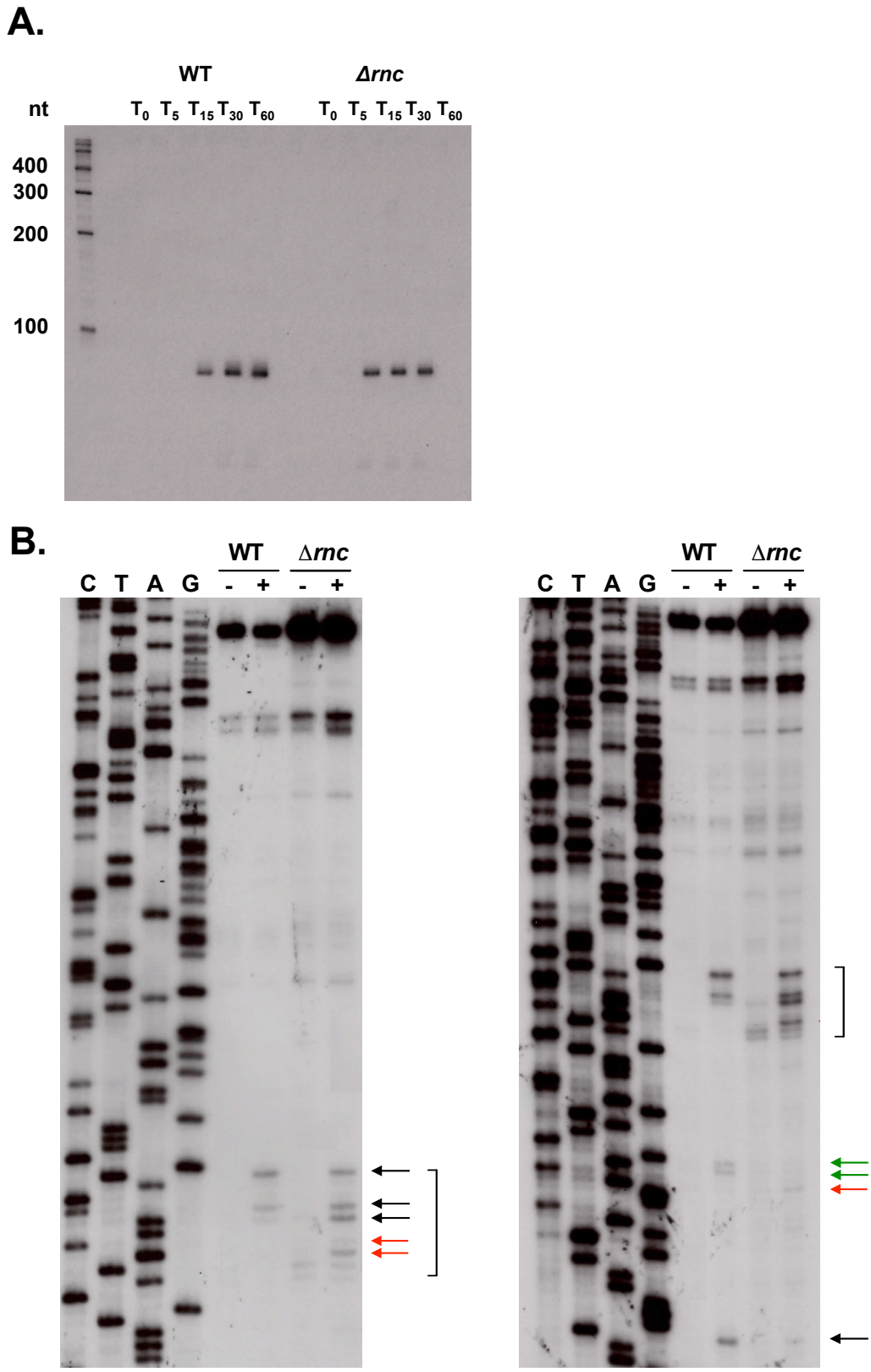


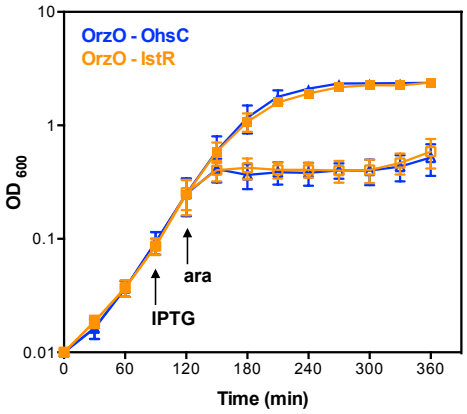
Figure 2.4. (continued)

A.

>OrzO-OhsC
GUUGGUACGAAAUGUUGCAAAAUUAAAGUUAAAAAGUAAAACCCCCGUUCCUUACCAGUUCGGGGGUUUUACUUU

>OrzO-IstR
GUUGGUACGAAAUGUUGCUGCGGUUACCAGCCGCGGGCGGCUGACGAAACCUCGCUCCGGCGGGGUUUUUU

B.



C.

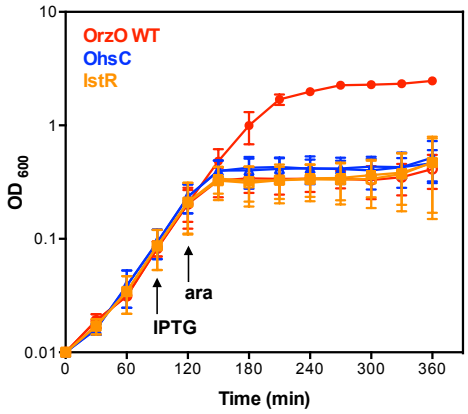


Figure 2.5. The 5' end of OrzO is sufficient for inhibition of *zorO*-mediated toxicity. (**A**) Generation of an OrzO-OhsC and OrzO-IstR chimeras. The 5' ends of the type I antitoxins OhsC and IstR were replaced with the 5' end of OrzO, the longest region of complementarity to *zorO*. Underlined sequences indicate the OrzO nucleotides. (**B**) Overexpression of both OrzO-OhsC and OrzO-IstR chimeras repress *zorO*. Rescue experiment was performed as outlined in Figure 2.1B with either OrzO-OhsC (blue triangles) or OrzO-IstR (orange squares) expressed from the $P_{LlacO-1}$ promoter of pBR-plac. Solid symbols indicate the addition of IPTG; open symbols, no IPTG. The averages of three independent cultures ± standard deviations are shown. (**C**) Neither wild type OhsC nor IstR can repress *zorO*-induced toxicity. Rescue experiment was performed with pBR-plac-*orzO* (red circles), pBR-plac-*ohsC* (blue triangles), or pBR-plac-*istR* (orange squares) as outlined in Figure 2.1B. Solid symbols indicate addition of IPTG; open symbols, no IPTG. Shown are the averages ± standard deviations for three independent cultures.

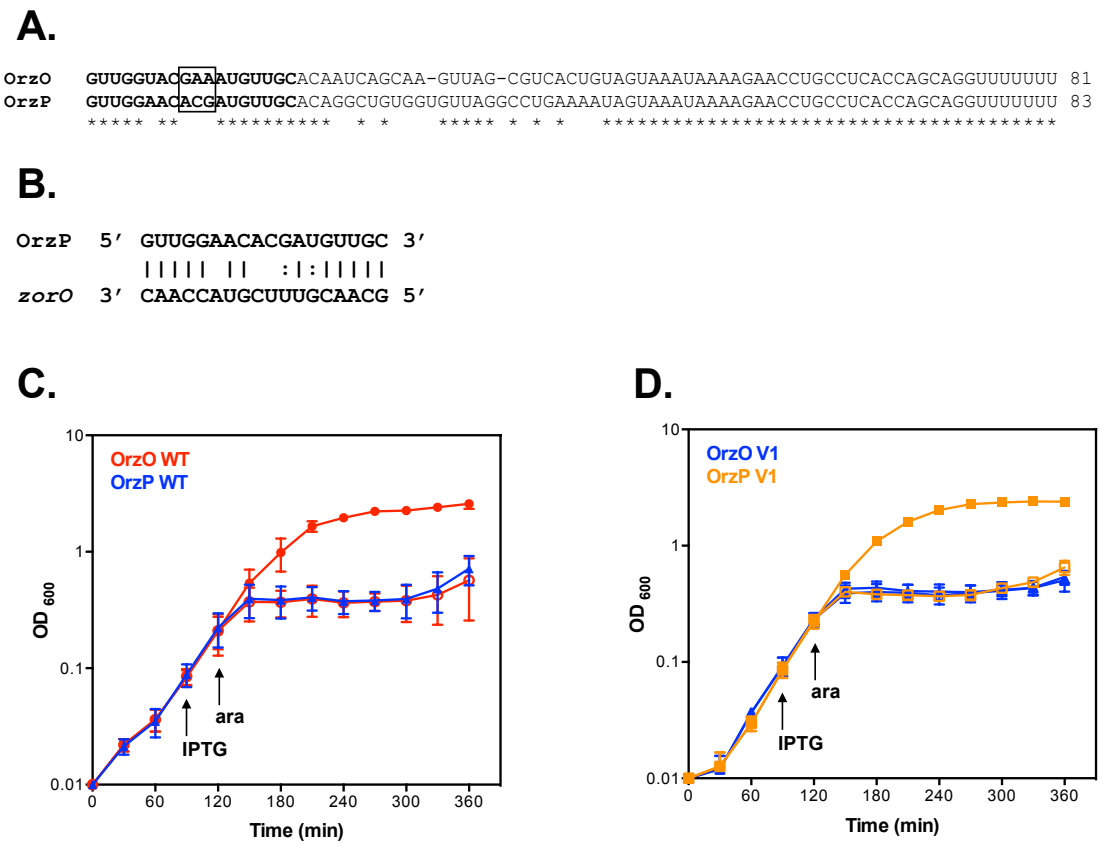


Figure 2.6. The V1 region is responsible for target discrimination by OrzO. (**A**) Alignment of OrzO and OrzP sRNAs. The V1 region is indicated in the black box; the entire base pairing region is in bold. (**B**) Potential base pairing interactions between OrzP and *zorO* in the predicted binding region. (**C**) OrzO, not OrzP, can repress *zorO*. Rescue experiment was performed as outlined in Figure 1B with either wild type OrzO (red circles) or OrzP (blue triangles) expressed from the $P_{LlacO-1}$ promoter of pBR-plac. Solid symbols indicate the addition of IPTG; open symbols, no IPTG. Averages ± standard deviations of three independent cultures are shown. (**D**) An OrzP V1 mutant can rescue cells from *zorO* overexpression. Rescue experiment was performed as outlined in Figure 2.1B using pBR-plac-*orzP* V1 (orange squares), in which its V1 region was mutated to match OrzO, and pBR-plac-*orzO* V1 (blue triangles), in which its V1 was mutated to match OrzP. Solid symbols indicate addition of IPTG; open symbols, no IPTG. Averages ± standard deviations for three independent cultures are shown.

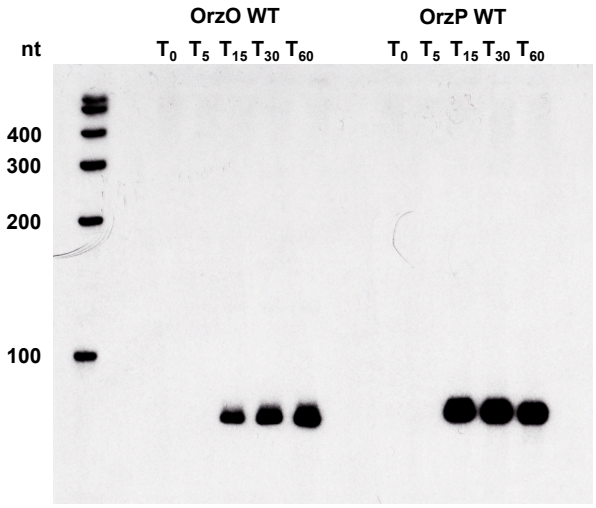
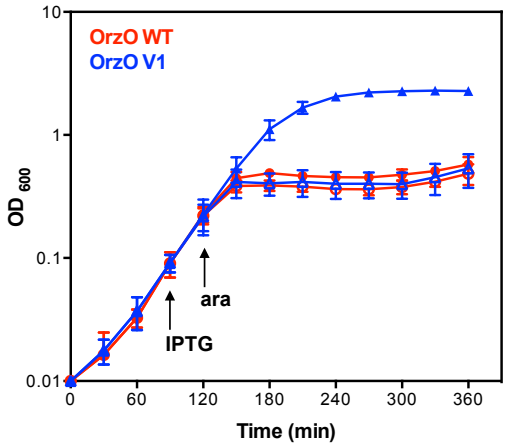


Figure 2.7. OrzP expression levels are equivalent to OrzO expression levels. Total RNA was isolated as described in Materials and Methods from UTK007 harboring pBAD-*zorO* and either pBR-plac-*orzO* or pBR-plac-*orzP* after 0, 5, 15, 30, 60 minutes of induction with 0.2% arabinose. Note that both OrzO and OrzP were induced by 1 mM IPTG.




```
OrzO V1 5′ GUUGGUACACGAUGUUGC 3′
           |||||||||||||:||||||
zorO V1 3′ CAACCAUGUGCUGCAACG 5′
```

Figure 2.8. Restoration of the base pairing interactions between OrzO V1 to zorO prevents zorO toxicity. The compensatory mutations were made in pBAD-*zorO* (Table 2.3) so that it could perfectly base pair with OrzO V1. The resulting construct, pBAD-*zorO* V1 was transformed into UTK007 harboring either pBR-plac-*orzO* (red circles) pBR-plac-*orzO* V1 (blue triangles) and a rescue experiment was performed as outlined in Figure 2.1B. Solid symbols indicate addition of IPTG; open symbols, no IPTG. Averages ± standard deviations for three independent cultures are shown. Predicted base pairing interactions between OrzO V1 and zorO V1 are shown on the right.

Figure 2.9. A specific residue within the V1 region of OrzO is necessary for *zorO* repression. (**A**) Mutation of the first nucleotide of the V1 region fails to rescue *E. coli* from ZorO toxicity. Rescue experiment was performed as outlined in Figure 2.1B except that arabinose was added to a final concentration of 0.002%. Shown are the averages ± standard deviations of pBR-plac-*orzO* (red circles) or pBR-plac-*orzO* 8(1)9 (blue triangles), which has the first nucleotide of the V1 region mutated. Solid symbols indicate addition of IPTG; open symbols, no IPTG. (**B**) Expression of OrzO 8(1)9 is similar to wild type OrzO. Total RNA was isolated from *E. coli* UTK007 overexpressing *zorO* (0.002% arabinose) and either OrzO (1 mM IPTG) or OrzO 8(1)9 (1 mM IPTG) as described in Materials and Methods. Quantification of RNA levels is described in the Materials and Methods, with OrzO WT at T_{15} set to a value of 1. Numbers below the panel indicate the average relative intensities. (**C**) Mutation of the third nucleotide of the V1 region has no impact on the ability of OrzO to repress *zorO* expression. The third nucleotide of the V1 region of OrzO was mutated to generate OrzO 10(1)7. A rescue experiment was performed as outlined in Figure 2.1B with wild type OrzO (red circles) or OrzO 10(1)7 (blue triangles) expressed from the $P_{LlacO-1}$ promoter of pBR-plac. Solid symbols indicate addition of IPTG; open symbols, no IPTG. Averages ± standard deviations for three independent cultures are shown.

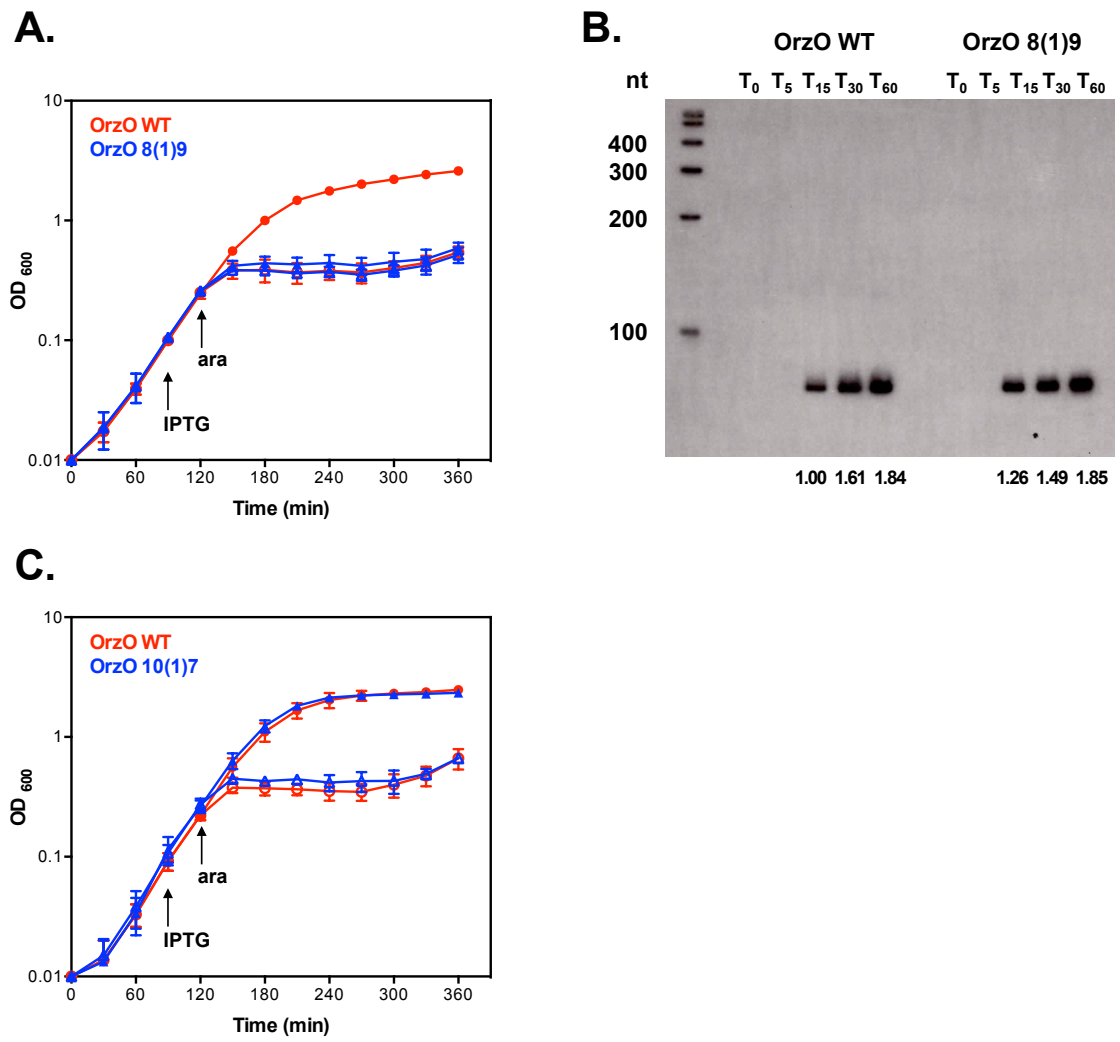


Figure 2.9. (continued)

Figure 2.10. Expression comparison of OrzO mutants to wild type OrzO. Quantifications of the northern analysis of OrzO WT to OrzO 8(1)9 **(A)**, OrzO Δ3 **(B)**, OrzO Δ6 **(C)**, OrzO 5(2)11 **(D)**, OrzO (1)5(1)11 **(E)**, or OrzO 6(1)10(1) **(F)**. Note that the mutant sRNAs were induced with 1 mM IPTG, whereas the wild type was induced with either 1 mM **(A)** or 0.1 mM IPTG **(B, C, D, E, and F)**. The intensity of the band representing OrzO WT at T_{15} in each northern was set to a value of 1. The relative intensities of the other bands in the same northern blot were thus calculated and plotted in bar graph. Shown are the averages ± standard deviations from the mean values of the three independent assays. *, $P < 0.05$ determined by Student's t test.

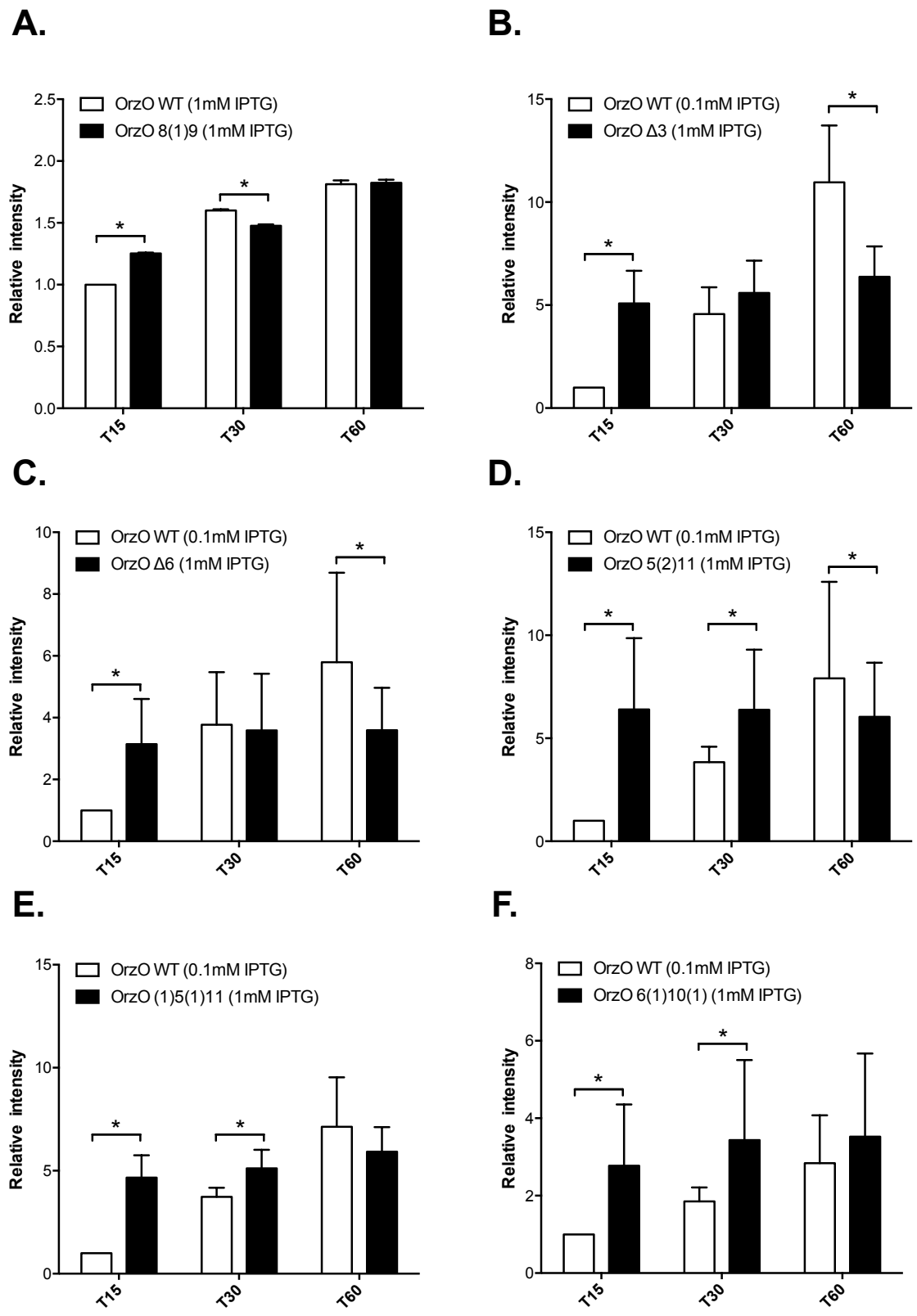


Figure 2.10. (continued)

Figure 2.11. The ability of OrzO 9(1)8 to repress *zorO*-induced toxicity is variable. (A) Rescue experiment was performed as outlined in Figure 2.1B using either pBR-plac-*orzO* or pBR-plac-*orzO* 9(1)8 as indicated. For co-expression of OrzO WT with *zorO*, 1 mM IPTG and 0.002% arabinose (red circles) was used; for co-expression of OrzO 9(1)8 and *zorO*, 1mM IPTG and 0.2% (orange squares) or 0.002% arabinose (blue triangles) were used. Solid symbols indicate addition of IPTG; open symbols, no IPTG. Averages ± standard deviations for three independent cultures are shown. **(B)** OrzO 9(1)8 is expressed at levels higher than the basal levels of wild type OrzO needed to repress *zorO*. Total RNA was isolated from *E. coli* UTK007 overexpressing *zorO* (0.002% arabinose) and either *orzO* WT (0.1 mM IPTG) or *orzO* 9(1)8 (1 mM IPTG) as described in Materials and Methods. Average intensities are shown as described in the Materials and Methods. **(C)** Quantification of the expression of OrzO 9(1)8 to wild type OrzO. Quantification of RNA levels is described in the Materials and Methods, with OrzO WT at T_{15} set to a value of 1. The relative intensities of the other bands in the same northern blot were thus calculated and plotted in the bar graph. Shown are the averages ± standard deviations from the mean values of the three independent assays. *, $P < 0.05$ determined by Student's t test.

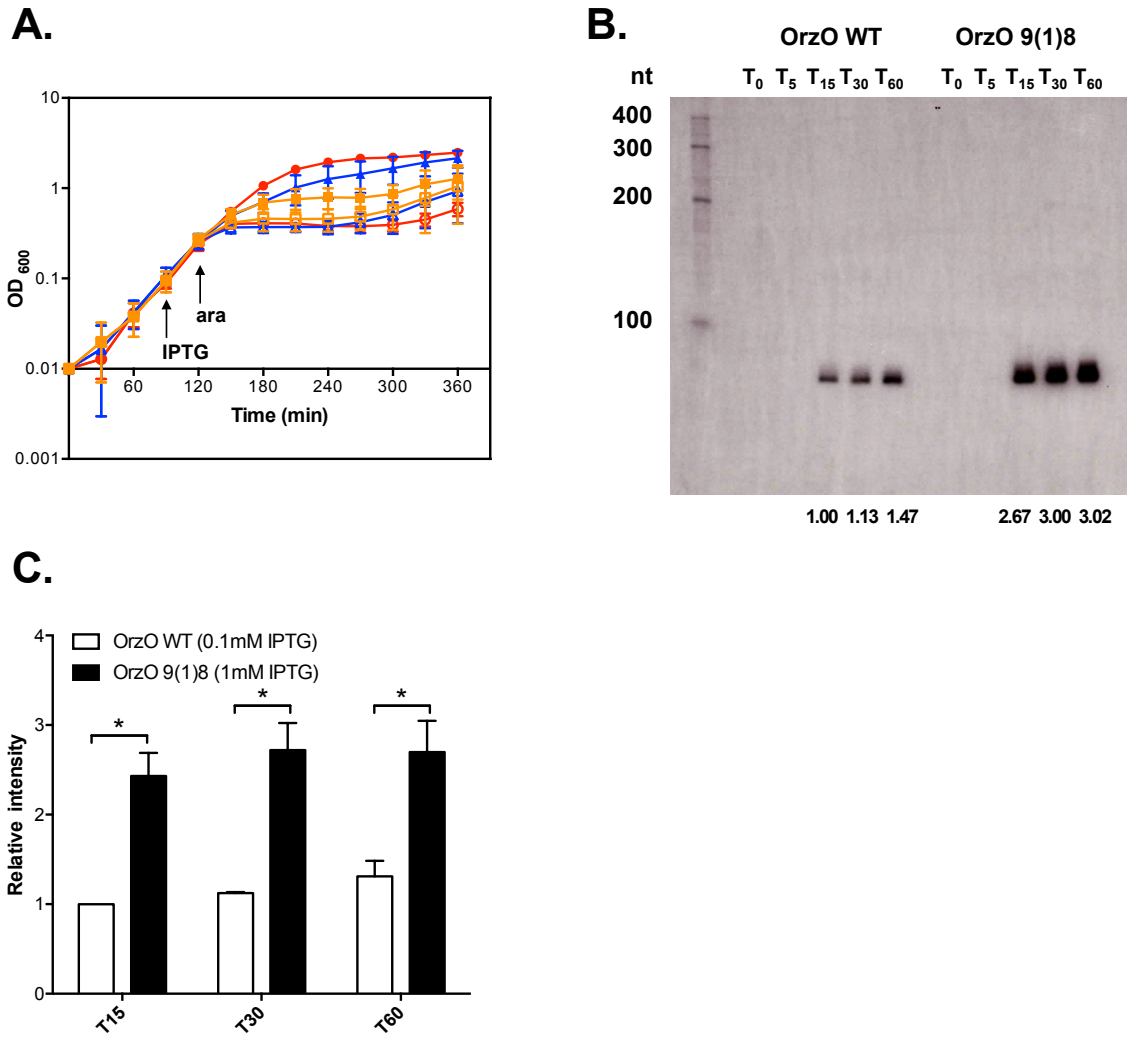


Figure 2.11. (continued)

Figure 2.12. The minimal level of OrzO needed for repression and the minimal amount of *zorO* that can cause toxicity. (A) Overexpression of OrzO with 0.1mM IPTG fully represses *zorO*-induced toxicity. Rescue experiment was performed as outlined in Figure 2.1B. Arabinose was added to a final concentration of 0.002%; IPTG, 1 mM (red circles) or 0.1 mM (blue triangles). Solid symbols indicate addition of IPTG; open symbols, no IPTG. Averages ± standard deviations for three independent cultures are shown. **(B)** Levels of OrzO induced by 0.1mM IPTG versus 1 mM IPTG. Total RNA was isolated as described in Materials and Methods from *E. coli* UTK007 harboring pBAD-*zorO* and pBR-plac-*orzO*. Shown are samples 0, 5, 15, 30, 60 minutes post induction with 0.002% arabinose and either 0.1mM or 1mM IPTG. Average intensities are shown as described in the Materials and Methods. **(C)** Expression comparison of wild type OrzO induced by 0.1 mM and 1 mM IPTG. Quantification of RNA levels is described in the Materials and Methods, with OrzO WT induced by 0.1 mM IPTG at T_{15} set to a value of 1. The relative intensities of the other bands in the same northern blot were thus calculated and plotted in bar graph. Shown are the averages ± standard deviations from the mean values of the three independent assays. *, $P < 0.05$ determined by Student's t test. **(D)** Overexpression of zorO with 0.002% arabinose causes cell stasis. *E. coli* strain UTK007 harboring pBAD-*zorO* was grown to $OD_{600} \approx 0.2$, and the culture was split into three. Arabinose, 0.2% (red circles), 0.002% (blue triangles) or no arabinose (no symbols) were added and OD_{600} was measured over time. Averages ± standard deviations for three independent cultures are shown. **(E)** Expression of *zorO* induced by either 0.2% or 0.002% arabinose. Total RNA was isolated as described in Materials and Methods from *E. coli* UTK007 harboring pBAD-*zorO* and pBR-plac-*orzO*. Shown are samples 0, 5, 15, 30, 60 minutes post induction with 1mM IPTG and either 0.2% or 0.002% arabinose.

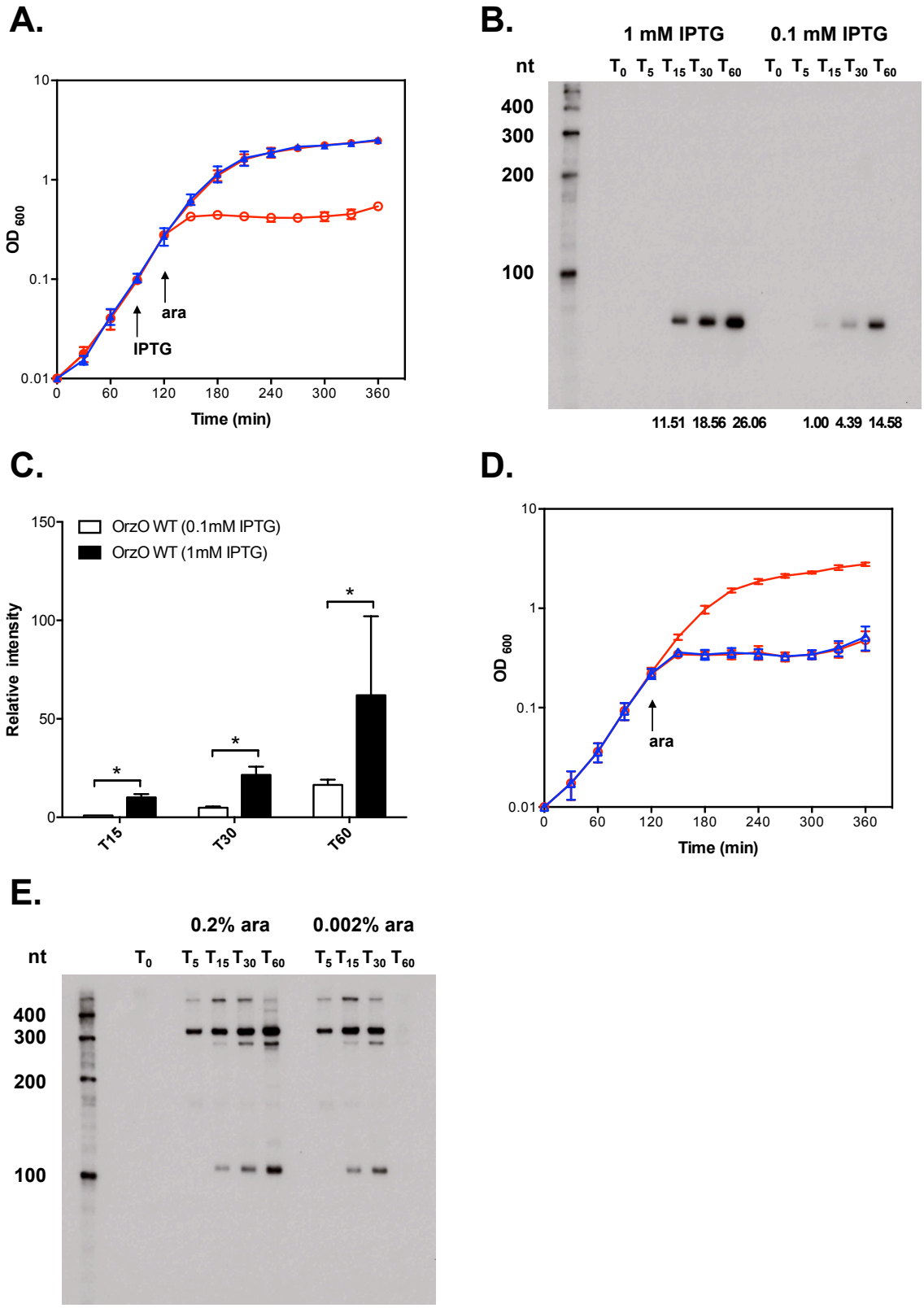


Figure 2.12. (continued)

Figure 2.13. Only 15 nucleotides of continuous base pairing are required for OrzO inhibition of *zorO*. (A) OrzO Δ3 prevents *zorO*-induced toxicity at low levels of arabinose. Rescue experiment was performed as outlined in Figure 2.1B with OrzO Δ3 expressed from the $P_{LlacO-1}$ promoter of pBR-plac. Note that *zorO* was induced by either 0.2% (blue triangles) or 0.002% (orange squares). Solid symbols indicate addition of IPTG; open symbols, no IPTG. Averages ± standard deviations for three independent cultures are shown. (**B**) Comparison of RNA levels between wild type OrzO induced by 0.1mM IPTG and OrzO Δ3 induced by 1mM IPTG. Total RNA was isolated from *E. coli* UTK007 overexpressing OrzO or OrzO Δ3 from pBR-plac as described in Materials and Methods. Average intensities are shown as described in the Materials and Methods. **(C)** OrzO Δ6 fails to repress *zorO*-induced toxicity even at low levels of arabinose induction. Rescue experiment was performed using either pBR-plac-*orzO* (red circles) or pBR-plac-*orzO* Δ6 (blue triangles) as indicated. Note that the arabinose concentration used was 0.002%. Solid symbols indicate addition of IPTG; open symbols, no IPTG. Averages ± standard deviations for three independent cultures are shown. **(D)** Comparison of RNA levels between the wild type OrzO induced by 0.1mM IPTG and the OrzO Δ6 mutant induced by 1mM IPTG. Total RNA was isolated from *E. coli* UTK007 overexpressing OrzO or OrzO Δ6 from pBR-plac as described in Materials and Methods. Average intensities are shown as described in the Materials and Methods.

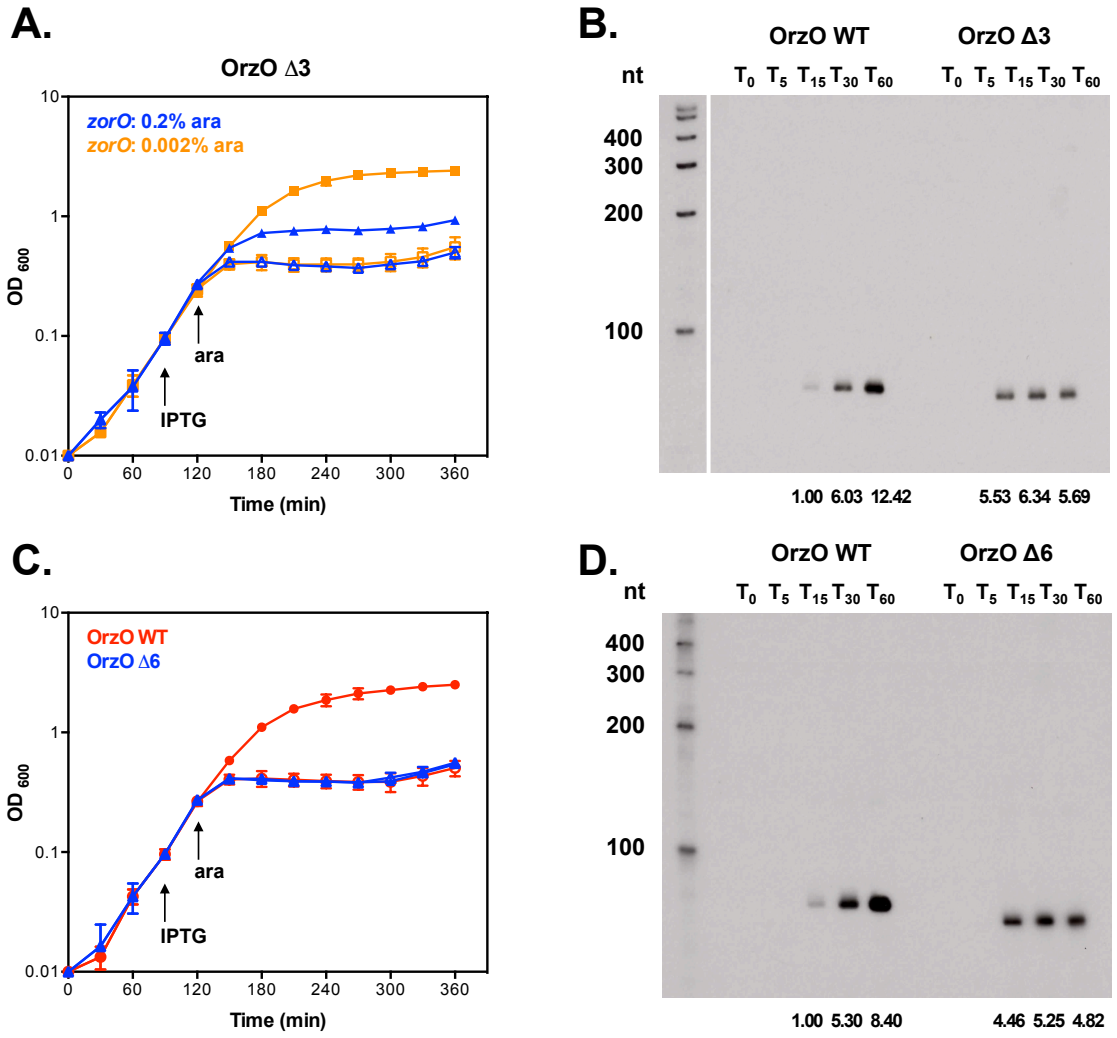


Figure 2.13. (continued)

Figure 2.14. Repression of *zorO* requires 17 nucleotides of discontinuous base pairing. (**A**) OrzO 5(2)11 fails to inhibit ZorO toxicity. Rescue experiment was performed as outlined in Figure 2.1B with either wild type OrzO (red circles) or OrzO 5(2)11 (blue triangles) expressed from the $P_{LlacO-1}$ promoter of pBR-plac. Solid symbols indicate addition of IPTG; open symbols, no IPTG. Averages ± standard deviations for three independent cultures are shown. (**B**) Comparison of RNA levels between the wild type OrzO induced by 0.1mM IPTG and the OrzO 5(2)11 mutant induced by 1mM IPTG. Average intensities are shown as described in the Materials and Methods. (**C**) OrzO 6(1)11 can repress *zorO*. Rescue experiment was performed as outlined in Figure 2.1B with either wild type OrzO (red circles) or OrzO 6(1)11 (blue triangles) expressed from the $P_{LlacO-1}$ promoter of pBR-plac. Solid symbols indicate addition of IPTG; open symbols, no IPTG. Averages ± standard deviations for three independent cultures are shown. (**D**) Neither OrzO 6(1)10(1) nor OrzO (1)5(1)11 can prevent *zorO*-induced toxicity. Rescue experiment was performed with pBR-plac-*orzO* (red circles), pBR-plac-*orzO* (1)5(1)11 (blue triangles), or pBR-plac-*orzO* 6(1)10(1) (orange squares). Note that IPTG was added to 1mM final concentration and arabinose was added to 0.002% final concentration. Solid symbols indicate addition of IPTG; open symbols, no IPTG. Averages ± standard deviations for three independent cultures are shown. (**E**) Comparison of RNA levels between the wild type OrzO induced by 0.1mM IPTG and the OrzO (1)5(1)11 mutant induced by 1mM IPTG. Average intensities are shown as described in the Materials and Methods. **(F)** Comparison of RNA levels between the wild type OrzO induced by 0.1mM IPTG and the OrzO 6(1)10(1) mutant induced by 1mM IPTG. Average intensities are shown as described in the Materials and Methods.

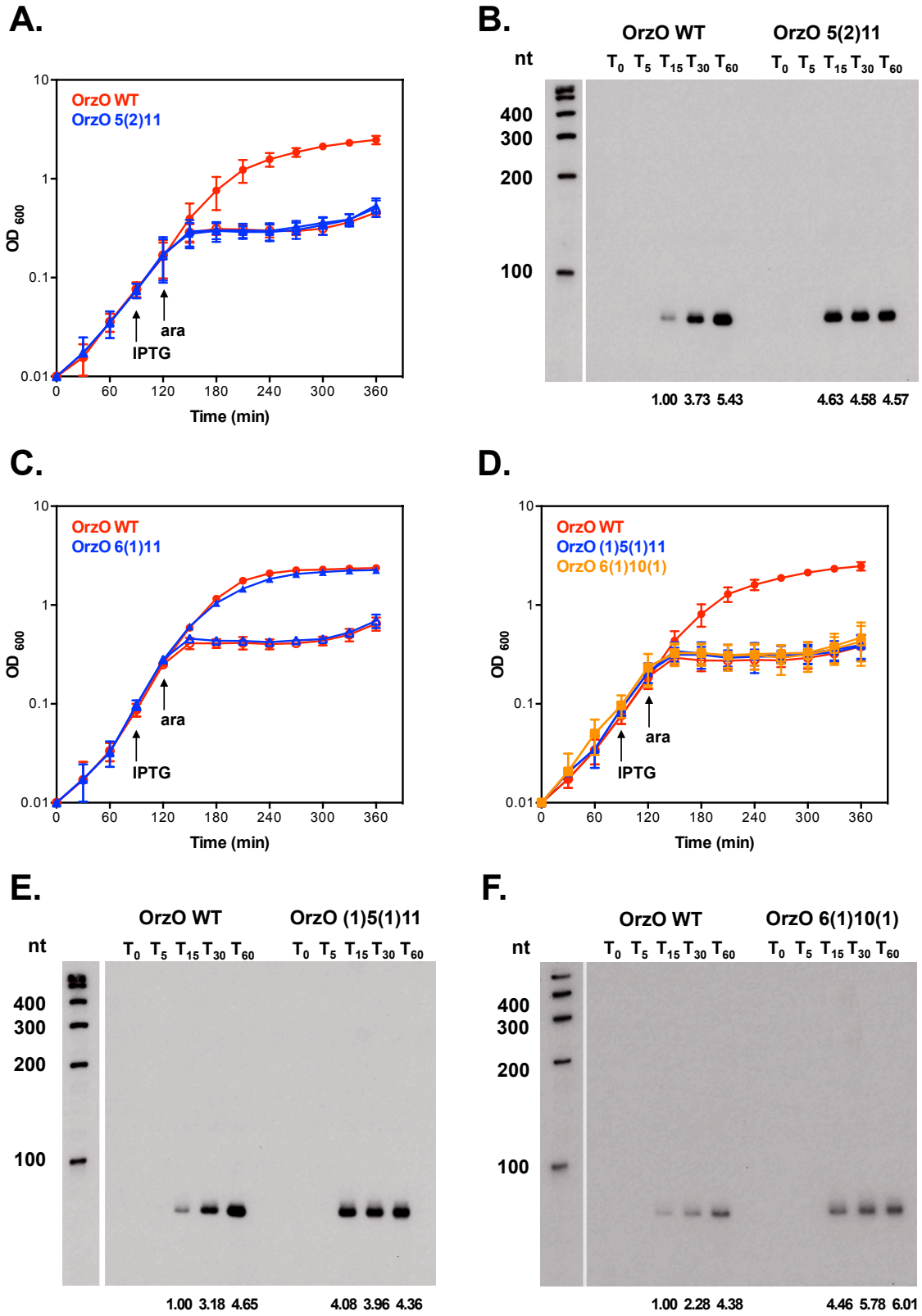


Figure 2.14. (continued)

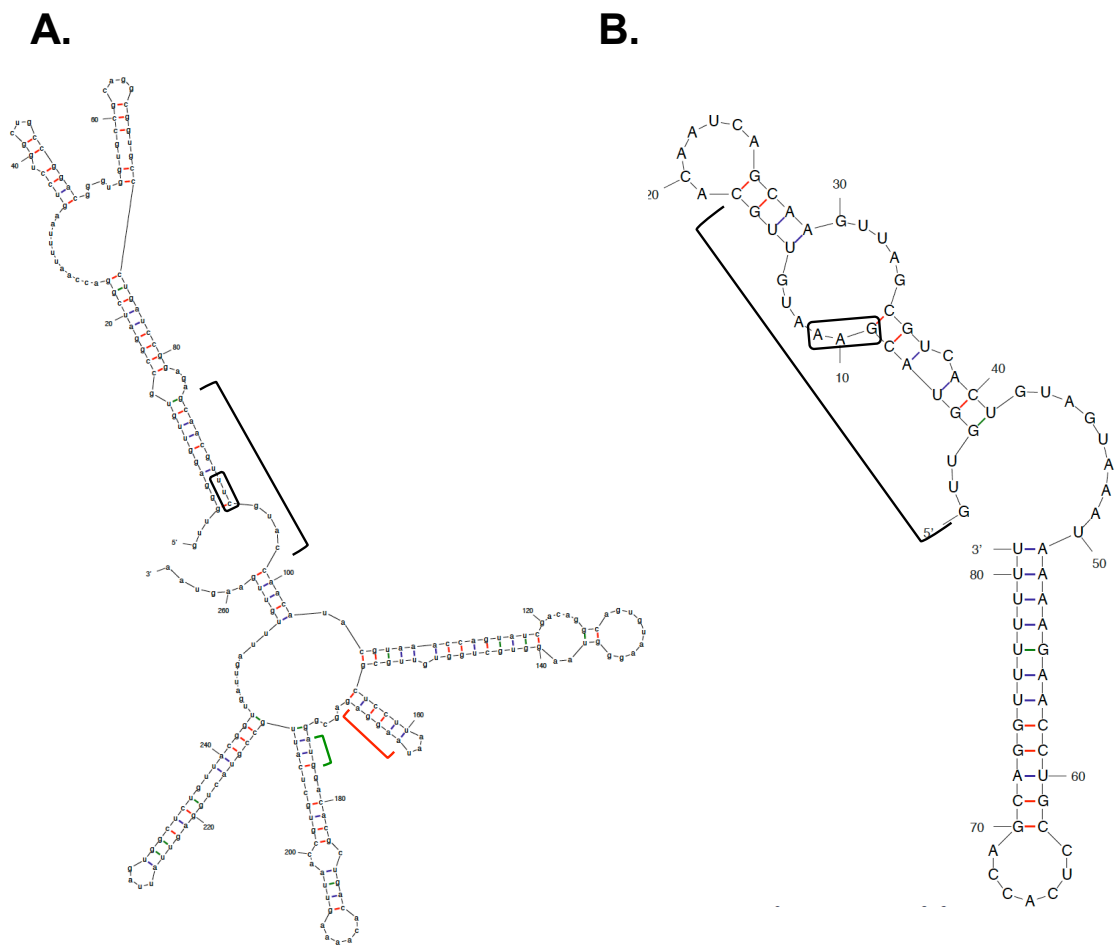


Figure 2.15. Predicted secondary structures of *zorO* mRNA and OrzO indicate the V1 regions are mostly single-stranded. (**A**) Putative RNA structure of *zorO* mRNA generated by M-fold (http://http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form) (51). The black bracket indicates the region of complementarity to OrzO; red brackets the putative ribosome binding site; green bracket, the start codon. The black box is the V1 region binding site. (**B**) Putative RNA structure of OrzO sRNA generated by M-fold (http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form) (51). The sequence indicated by the black bracket is the region of complementarity to the *zorO* mRNA; the black box is the V1 region.

CHAPTER III

5’ UTR inhibits and enhances translation of the type I toxin ZorO

Publication note

This chapter is in preparation for submission. The tentative title is:

Jia Wen, John R. Harp, and Elizabeth M. Fozo. The 5' UTR of the type I toxin ZorO can both inhibit and enhance translation [In preparation].

My contribution to this paper was mutant constructions, toxicity assays, rescue experiments, RNA isolation, northern analysis, *in vitro* transcription, structural probing, *in vitro* translation, literature review, and writing. John R. Harp assisted in generating data presented in Figure 3.2, where he performed the flow cytometry analyses. I performed all other experiments.

Abstract

Many bacterial type I toxin mRNAs possess a long 5' untranslated region (UTR) that can serve as the target site of the corresponding antitoxin sRNA. This is the case for the *zorO-orzO* type I system where the OrzO sRNA antitoxin base pairs to the 174-nucleotide *zorO* 5' UTR. Here, we demonstrate that the 5' UTR of the *zorO* type I toxin controls *zorO* translation independent of the sRNA such that the full-length 5' UTR hinders translation whereas a processed 5' UTR (*zorO* Δ28) promotes translation. The full-length *zorO* 5' UTR folds into an extensive secondary structure sequestering the ribosome binding site (RBS). Processing of the 5' UTR does not alter the RBS structure, but exposes a single-stranded region which we designate as the EAP (Exposure After Processing) region located upstream of the RBS, which likely serves as a ribosome standby site. Truncation of this EAP region severely impairs *zorO* translation, but this defect can be rescued by exposure of the RBS. In addition to the EAP region, another region spanning +35 to +50 of the *zorO* mRNA also facilitates translation and is needed for optimal translation of *zorO*. Importantly, the positive and negative effects on translation imparted by the 5' UTR can be transferred onto a reporter gene, indicative that the 5' UTR can solely drive regulation. Moreover, we show that the cognate OrzO sRNA can inhibit *zorO* translation via base pairing to the 3' end of the EAP region.

I. Introduction

Across bacterial species, numerous toxin-antitoxin loci have been identified that are believed to function in stress responses. These two-gene loci are categorized by both the nature of the antitoxin and the mode of its action to repress the toxin. For a type I toxin-antitoxin locus, one gene encodes a small protein that is toxic when overproduced and the second gene encodes a small RNA (sRNA) that represses toxin production through base pairing [reviewed in (1)]. Most type I toxins share common biochemical features: they usually contain a single putative transmembrane domain and are hypothesized to function at the cytoplasmic membrane. Overproduction of type I toxins is detrimental to bacterial cells and results in growth stasis or cell death. For instance, overexpression of either the *ibsC* or *shoB* toxin gene in *Escherichia coli* inhibited cell growth and led to rapid membrane depolarization (2). Further, TisB toxin overproduction in *E. coli* caused similar membrane damage along with a dramatic decrease in cellular ATP levels (3). Additionally, some type I toxins can impair nucleoid segregation and cell division prior to damaging the cell membrane, as exemplified by the Fst toxin of *Enterococcus faecalis* (4).

Given its inherent toxicity, the production of a type I toxin must be tightly controlled. This can be achieved through regulation by its corresponding antitoxin sRNA. More specifically, the antitoxin sRNA base pairs with the toxin mRNA and inhibits translation of the mRNA and/or stimulates the degradation of the mRNA, thereby preventing toxin production. This is exemplified by the *symE-symR* locus of *E. coli* (5) in that the SymR sRNA base pairs to the translation initiation region of the *symE* mRNA and blocks its translation, therefore repressing SymE production (6). Further, in the *txpA-ratA* pair of *Bacillus subtilis*, base pairing of the RatA sRNA to the *txpA* mRNA leads to RNase III-dependent degradation of *txpA*, rendering it untranslatable (7,8).

Aside from the sRNA antitoxin mediated regulation, production of some type I toxins can be modulated by the 5' untranslated regions (UTRs) of the toxin mRNAs themselves, particularly for those toxin mRNAs that possess long and structured 5' UTRs [reviewed in (9)]. One well studied example is the type I toxin gene *hok* of plasmid R1 in *E. coli* (10). The *hok* mRNA has a 177-nt-long 5' UTR, which harbors a small leader gene termed *mok* (11). Translation of the *hok* mRNA is dependent upon translation of *mok*; however, the 5' UTR of *hok* forms into a secondary structure that occludes the *mok* RBS, preventing not only its translation but also that of Hok (12,13). Although translation of the type I toxin TisB is also repressed by its long 5' UTR, the underlying mechanism is distinct from that of *hok*. In this case, the *tisB* RBS is structurally sequestered and

translation of the *tisB* mRNA requires a single-stranded standby site upstream of the RBS that allows ribosomes to preload onto the mRNA (14). This standby site, though, is occluded in a stem structure and only becomes accessible upon processing of the 5' UTR; thus, the full-length *tisB* mRNA is translationally inactive.

We previously reported that the *zorO-orzO* gene pair in *E. coli* O157:H7 (EHEC) is a true type I toxin-antitoxin locus, with *zorO* encoding the toxic small protein and *orzO* encoding the antitoxin sRNA (15,16). The *zorO* mRNA has a long 5' UTR of 174 nts with the region of base pairing to OrzO located 60 nts upstream of the *zorO* RBS. When examining the predicted structures of the *zorO* mRNA, we noted that the RBS is located in a stem structure (15), which could prevent ribosomal access. We hypothesized that translation of *zorO* may require ribosomal preloading onto a standby site to compensate for the closed RBS. Thus, disruption of this putative ribosome standby site of *zorO* would impair translation of the *zorO* mRNA and alleviate the toxic effects of *zorO* overexpression.

In this study, we confirm that the *zorO* RBS is occluded in stem structure and that processing of the 5' end is required for efficient translation. The 5' processing does not alter the structure at the RBS, but leads to the opening of a putative standby site (EAP region) that promotes translation. Optimal translation of *zorO* requires not only the EAP region, but also another region spanning from +35 to +50 of *zorO* 5' UTR. Moreover, we demonstrate that the OrzO sRNA affects both the stability and the translation of the *zorO* mRNA via base pairing at the EAP region of *zorO*, therefore inhibiting the production of ZorO toxin.

II. Materials and Methods

Bacterial strains and plasmids The strains and plasmids used in this study are listed in Table 3.1, and the sequences of all oligonucleotides are given in Table 3.2.

Growth conditions *E. coli* strains were grown in Luria–Bertani (LB) medium at 37°C with shaking. For strains carrying pAZ3-*zorO* Δ28 or pAZ3-*zorO* Δ34, 0.2% glucose was added to reduce leaky expression of the P_{BAD} promoter. Antibiotics were used at the following concentrations when needed: ampicillin, 100 μg/ml; chloramphenicol, 25 μg/ml; kanamycin, 25 μg/ml. Arabinose was added to a final concentration of 0.2%, 0.0001%, or 0.00001% as indicated. When used, IPTG was added to a final concentration of 1 mM.

Plasmid construction For overexpression of the *zorO* gene under the P_{BAD} promoter, the full 5' UTR of *zorO* along with the coding region was amplified from *E. coli* O157:H7 EDL933 genomic DNA, digested with EcoRI and HindIII, and cloned into the corresponding sites of pAZ3 (5), generating pAZ3-*zorO*. The pAZ3-*zorO* Δ28, pAZ3-*zorO* Δ34, pAZ3-*zorO* Δ50, and pAZ3-*zorO* Δ82 were constructed in the same fashion. The pEF21-*zorO* Δ28 was generated following the same scheme, but using a lower copy number vector pEF21 (pBAD33 derivative, pACYC origin) (2). The pAZ3-*zorO* Δ82 RBS was constructed through site-directed mutagenesis as previously described using pAZ3-*zorO* Δ82 as template (17).

For *in vitro* transcription assays, the full 5' UTR of *zorO* along with the coding region was amplified from *E. coli* O157:H7 EDL933 genomic DNA, digested with EcoRI and SmaI, and cloned into the corresponding sites of pGEM®-3Zf(+) Vector (Promega), generating pGEM-T7-*zorO*. Plasmids pGEM-T7-*zorO* Δ50, and pGEM-T7-*zorO* Δ82 were generated in the same manner. The pGEM-T7-*zorO* Δ82 RBS was generated through site-directed mutagenesis using pGEM-T7-*zorO* Δ82 as template.

The translational *gfp* fusion to *zorO* was generated using SOE PCR (18). PCR product A amplified the full length 5' UTR and the start codon of *zorO* from genomic DNA along with an additional 18 nts at the 3' end that overlapped with the PCR product B. PCR product B was the result of amplifying the coding sequence of *gfp* from vector pFA6a-GFP(S65T)-kanMX6 (19). The two products were spliced together using the external primers, digested with EcoRI and HindIII, and cloned into the corresponding sites of pAZ3, generating pAZ3-*zorO* UTR-*gfp*. The pAZ3-Δ28 UTR-*gfp* was constructed through amplifying the corresponding UTR-*gfp* region using pAZ3-*zorO* UTR-*gfp* as template via PCR, and cloned into pAZ3.

Overproduction of toxic proteins ZorO toxicity assays were performed as described previously (15). Briefly, the indicated pAZ3 plasmid was transformed into UTK007 via electroporation (20). The resulting transformants were grown overnight and then diluted to an OD_{600} of 0.01. When the OD_{600} reached 0.2 - 0.3, the culture was split and arabinose was added to half of the culture to a final concentration of 0.2%, 0.0001%, or 0.00001%. Prior to arabinose induction, cells harboring either pAZ3- *zorO* Δ28 or pAZ3- *zorO* Δ34 were washed twice with 1×phosphate-buffered saline (PBS) and resuspended in LB medium to ensure no carryover of glucose. OD_{600} was measured

every 30 min. Shown are averages ± standard deviations for a minimum of three independent experiments.

Rescue experiments Rescue experiments were conducted as described previously (15,21). The pBR-plac plasmid carrying the *orzO* gene under an IPTG-induced $P_{LlacO-1}$ promoter was transformed into UTK011 (Δ*rnc*) through electroporation. Afterwards, the pEF21 plasmid containing the full-length 5' UTR of *zorO* and its coding sequence (*zorO* full-length) or the processed 5' UTR of *zorO* and its coding sequence (*zorO* Δ28) under the P_{BAD} promoter was transformed into the same cells. The resulting transformants harboring the two plasmids were grown overnight and then diluted to OD_{600} of 0.01. When the OD_{600} reached ≈ 0.1, the cultures were split and IPTG was added to half the culture to a final concentration of 1 mM. Thirty minutes after IPTG induction, arabinose was added to a final concentration of either 0.0001% or 0.00002% for the overproduction of *zorO*. OD_{600} was measured every 30 min. Shown are averages ± standard deviations for a minimum of three independent experiments.

RNA extraction To compare the RNA levels of *zorO* wild type versus *zorO* mutants, cells carrying the appropriate plasmid were grown as indicated in the toxicity assays to an OD_{600} ≈ 0.2 – 0.3, and arabinose (0.2% or 0.0001%) was added (time 0). Cells were harvested at time 0, 5, 15, 30, and 60 min post-arabinose induction and total RNA was isolated via hot acid-phenol or direct lysis as described previously (21,22). When needed, RNA was treated with TURBO™ DNase (Life Technologies) following the manufacturer's instructions.

Northern analysis Total RNA (10 µg) was separated on a denatured 6% polyacrylamide-urea gel for detection of *zorO* wild type or mutant mRNA and transferred to a Zeta-Probe Genomic GT membrane (Bio-Rad). Specific oligonucleotide probes were 5' end-labeled with γ-^{32}P by T4 polynucleotide kinase (New England Biolabs). Hybridization and washes of the membrane were performed as described previously (23). Northern analysis was conducted from a minimum of three independent experiments for every construct examined.

Flow cytometry analysis Cells containing a translational *gfp* fusion to *zorO* were grown as previously indicated to an approximate OD_{600} ≈ 0.3 and split into two cultures. One culture was

induced with 0.2% arabinose and the other served as a control. After the cultures were split, aliquots were taken from the cultures at 0, 30, and 60 min. Due to cellular growth, 50 µl aliquots were taken at time 0 and 25 µl aliquots were taken at time 30 and 60. These aliquots were flooded with 4 ml of 1X PBS and washed extensively twice. Cells were resuspended in 1ml 1X PBS and analyzed by flow cytometry in a LSR II flow cytometer (Becton Dickinson) with a 488-nm laser. Samples were run at an event rate of 3,000 events per second. Green fluorescence was collected in the fluorescein isothiocyanate (FITC) channel. Differences in fluorescence intensity of the induction group versus the control group were obtained using geometric mean fluorescence for n = 3.

***In vitro* transcription** T7 *in vitro* transcription was conducted as described previously (24) with modifications. Briefly, pGEM-T7-*zorO* or its derivatives was linearized with SmaI, purified using the Qiagen PCR Purification Kit, and 1 µg of digested plasmid was used as the DNA template for T7 transcription. A DNA template for *zorO* Δ28 mRNA was amplified from *E. coli* O157:H7 EDL933 genomic DNA via PCR with primers containing the T7 promoter. The DNA template, ribonucleotides (each at 2 mM, New England Biolabs), 8 mM DTT (Invitrogen), and 100 U T7 RNA polymerase (50 U/µl, New England Biolabs) were mixed in a 50 µl volume and incubated at 37°C for 16 hrs. Afterwards, 2 U of TURBO™ DNase (2 U/µl, Life Technologies) was added to the transcription reaction, followed by incubation at 37°C for an additional 15 min. Transcribed RNAs were then precipitated and gel purified (25).

RNA radiolabeling and *in vitro* structure probing For 5' end labeling, the RNAs synthesized from T7 transcription were dephosphorylated and end-labeled with γ-^{32}P by T4 polynucleotide kinase (New England Biolabs). For 3' end labeling, RNAs were labeled with [5- γ ^{32}P] pCp and T4 RNA Ligase 1 (New England Biolabs).

In vitro structure probing was conducted in 10 µl reactions as described previously (26). Radiolabeled *zorO* RNA and its derivatives (0.2 pmol) were denatured at 95°C for 1 min and incubated on ice for 5 min, followed by the addition of 1 × Structure Buffer (Ambion) and 0.1 µg/µl yeast RNA. Unlabeled OrzO sRNA (250 nmol or 500 nmol) was added as indicated. The reaction mixtures were incubated at 37°C for 1 h. For enzymatic probing, 2 µl of RNase T1 (0.01 U/µl; Ambion) were added and the samples were incubated for 6 min at 37°C. For chemical

probing, 2 µl of fresh lead(II) acetate (25 mM) were added to corresponding reactions and incubated for 1.5 min at 37°C. Reactions were stopped with the addition of 20 µl of Inactivation/Precipitation buffer (Ambion) mixed by vortexing, and precipitated. The RNA pellets were suspended in 7 µl Loading Buffer II (95% formamide, 18 mM EDTA, 0.025% SDS, 0.2% xylene cyanol, 0.2% bromophenol blue; Ambion).

To generate the RNase T1 ladder, radiolabeled RNA (0.4 pmol) was mixed with 1 × Sequencing Buffer (Ambion), denatured at 95°C for 1 min, and chilled on ice for 5 min. The mixture was incubated with RNase T1 (1 µl, 0.1 U/µl) at 37°C for 5 min, and 12 µl of loading buffer II was added to stop the reaction. To generate the hydroxyl ladder, radiolabeled RNA (0.4 pmol) was mixed with 1 × Alkaline Hydrolysis Buffer (Ambion) and incubated at 90°C for 5 min. The reaction was stopped with the addition of 12 µl of loading buffer II.

All samples were denatured at 95°C for 3 min and 3 µl of each sample was resolved on a denatured 6% polyacrylamide-urea gel. Gels were dried and then exposed to BioMax XAR film (Kodak).

***In vitro* translation** *In vitro* translation was performed using *E. coli* S30 Extract System (Promega) as described previously with modifications (14). Briefly, 0.1 µM or 0.25 µM of *in vitro* transcribed *zorO* RNA was mixed with 0.2 mM ^{35}S-labelled methionine (>1000 Ci/mmol, PerkinElmer), 7.5 µl S30 extract, 10 µl S30 premix without amino acids, and 0.1 mM of each amino acid minus methionine. 5 µM of *orzO* RNA were added as indicated in the text. The reactions were incubated at 37°C for 30 min and chilled on ice for 5 min. For each reaction, 5 µl aliquots were mixed with 20 µl cold acetone and incubated on ice for an additional 15 min. Afterwards, the reaction mixtures were centrifuged for 10 min at 4°C at 12,000 g. The protein pellets were suspended in 15 µl of H_2O, mixed with 4X Bolt LDS sample buffer (Life Technologies), and resolved on NuPAGE Novex 4-12% Bis-Tris Plus Gels in MES Buffer (Life Technologies) alongside SeeBlue® Pre-Stained Protein Standard (Life Technologies). Gels were fixed in methanol/acetic acid solution (45% methanol, 10% acetic acid) for 30 min at room temperature and washed in 100 ml Amplify Fluorographic Reagent solution (GE Healthcare) for 60 min. Gels were subsequently dried and exposed to BioMax XAR film (Kodak).

III. Results

5' processed form of *zorO* exhibits more efficient translation Previous primer extension analysis of the *zorO* mRNA revealed a major processed form at the + 29 from the transcription start site (15). It has been shown that 5' processing of some type I toxin mRNAs increases translation of the mRNA [reviewed in (9)]. Thus, we hypothesized that the processed form of *zorO* may be translated more efficiently than the full-length *zorO*.

To test if processing of the *zorO* mRNA impacts its translation, we constructed a *zorO* Δ28 mutant under the control of the P_{BAD} promoter in the plasmid pAZ3 (pAZ3-*zorO* Δ28) in which the first 28 nts of the *zorO* 5' UTR were removed (Figure 3.1A). This mutant reflects the native processed form of the *zorO* transcript. We predicted that if 5' processing of the *zorO* mRNA promotes *zorO* translation, the *zorO* Δ28 mutant would have a higher translation efficiency and exhibit a more pronounced toxic phenotype than the *zorO* full-length under the same induction condition.

To compare the toxicity of the *zorO* full-length and *zorO* Δ28 plasmids, we first determined that 0.0001% was the lowest amount of arabinose needed to induce stasis in cultures overproducing the full-length *zorO* mRNA (Figure 3.1B). Decreased levels of arabinose (*i.e.* 0.00001%) did impact growth, but complete stasis was not observed (Figure 3.1B). On the contrary, *zorO* Δ28 still caused complete cell growth stasis at 0.00001% (Figure 3.1C). Given that the *zorO* coding sequence remains unchanged in *zorO* Δ28, this "hypertoxic" phenotype observed is indicative of increased production of ZorO toxin.

Since both the *zorO* full-length and the *zorO* Δ28 constructs examined above share the same promoter and coding sequence, the more pronounced toxic phenotype caused by *zorO* Δ28 is likely not due to increased transcription or protein stability. Rather, it could be due to greater RNA stability and/or higher translation efficiency. If the *zorO* Δ28 RNA is more stable, we would expect to see increased RNA levels compared to that of the *zorO* full-length under the same induction conditions. To test this, we examined the mRNA levels of the *zorO* full-length as well as the *zorO* Δ28 by northern blot analyses. As shown in Figure 3.1D, *zorO* Δ28 was expressed at levels equivalent to or even lower than the *zorO* full-length under the same induction condition. This suggested that the elevated toxic phenotype of the *zorO* Δ28 mutant is likely due to increased translation of the mRNA and not due to improved mRNA stability. To directly compare the translation efficiency of *zorO* full-length versus *zorO* Δ28, an *in vitro* translation assay was

performed. Although both were translated, the *zorO* Δ28 mRNA was more robustly translated than the *zorO* full-length (Figure 3.1E). Therefore, the 5' processing of *zorO* mRNA facilitates the translation of *zorO*.

Enhanced translation by 5' processing can be applied to another system We next wanted to determine if the observed translational enhancement upon processing was due to changes in the 5' UTR. If changes in the UTR drove the effects on translation, then the transfer of the *zorO* 5' UTR onto another gene would mimic the above results. To test this, we generated two translational *gfp* fusions with either the full-length UTR (*zorO* UTR) or the processed UTR (Δ28 UTR); both fusions were under control of the P_{BAD} promoter. We compared the production of GFP following arabinose induction using flow cytometry. Without induction, these two constructs showed similar background fluorescence throughout all time points taken (Figure 3.2A). Thirty minutes post-induction, however, the Δ28 UTR-*gfp* showed a dramatic increase (~8 fold) in fluorescence intensity as compared to its uninduced control, whereas the fluorescence intensity of the full-length UTR-*gfp* was similar to the control. Further, after 60 minutes, the fluorescence produced by cells overexpressing the Δ28 UTR-*gfp* was 20 fold greater than its control, while the fluorescence produced by cells overexpressing the full-length UTR-*gfp* was only 1.5 fold greater than its control (Figure 3.2B and 3.2C). Additionally, we noted the full-length UTR-*gfp* was processed (data not shown) similarly as we have noted for *zorO* (15) indicating that processing can occur regardless of the downstream coding sequence. These results fully recapitulate the observed "hypertoxic" phenotype of *zorO* Δ28 and confirm our findings that processing of the *zorO* 5' UTR promotes translation.

Opening of the +73 to +102 region in *zorO* Δ28 Our data suggested that removal of the first 28 nts of the 5' UTR promotes *zorO* translation, yet the mechanism underlying this increased translation remained unclear. The RBS of the *zorO* mRNA is predicted to be in a stem structure that may impede ribosome binding (15,27). This raised the possibility that the 5' processing could lead to an overall rearrangement of the structure of the mRNA, rendering the RBS more accessible. We thus employed *in vitro* structure probing to compare the *zorO* full-length and *zorO* Δ28 structures. Surprisingly, no structural differences were detected at the RBS or the flanking regions of the *zorO* full-length and *zorO* Δ28 and both RBSs were in stems (Figure 3.3A and 3.3B). This

indicates that the increased translation efficiency of *zorO* Δ28 was not due to a more accessible RBS.

When comparing the two RNA structures, the only major difference we observed was in the region spanning from the +73 to +102 from the transcription start site (Figure 3.3C). This region was single-stranded in *zorO* Δ28 but was highly structured in the full-length *zorO* mRNA, matching the structural predictions (15,27). Thus, processing of the 5' end exposes this +73 to +102 region, which we refer to as the EAP (Exposure After Processing) region. Interestingly, the EAP region encompasses the *zorO* base pairing region by the antitoxin sRNA OrzO (Figure 3.3D) (15). This indicates that the EAP region may also be involved in sRNA dependent post-transcriptional regulation of *zorO* (see below).

Disruption of the EAP region impairs *zorO* translation Since the EAP region is located upstream (~ 100 nts) of the *zorO* RBS, we wanted to determine how exposure of this region impacted translation. Given that the RBS of *zorO* is structurally sequestered, direct ribosomal access to the RBS seems unlikely. Therefore, we hypothesized that optimal translation of *zorO* may require a ribosome standby site to facilitate ribosome preloading onto the mRNA. Data suggests that a standby site may not consist of a specific sequence, but instead is usually single-stranded (28). Owing to the single-stranded nature of the EAP region in *zorO* Δ28, we predicted that it may serve as a standby site. We thus generated a truncated EAP mutant by deleting the first 82 nts of the *zorO* 5' UTR (*zorO* Δ82) while maintaining the portion of the EAP region where the OrzO sRNA base pairs (Figure 3.1A). If the EAP region serves as a standby site, truncation of this region would interfere with ribosomal preloading and the subsequent translation, rendering the *zorO* Δ82 less toxic than the *zorO* full-length.

Indeed, the *zorO* Δ82 mutant was not toxic when induced by 0.0001% arabinose in contrast to full-length *zorO* (Figure 3.4A). Even when induced with saturating amounts of arabinose (0.2%), *zorO* Δ82 only showed impaired cell growth (Figure 3.4B). Further, within 15 minutes of induction of the full-length *zorO* with 0.2% arabinose, no viable colonies were detected via plating. However, viable colonies were still detectable after 2 hours of induction of the *zorO* Δ82 using the same arabinose concentration (0.2%; data not shown). Taken together, these data suggested that cells overexpressing *zorO* Δ82 produced less ZorO toxin as compared to those overexpressing full-length *zorO*.

To separate out the effect of RNA stability from that of translation efficiency, we determined the RNA levels of *zorO* Δ82 and *zorO* full-length by northern analysis. When both were induced with 0.2% arabinose, the levels of *zorO* Δ82 mRNA was lower than that of *zorO* full-length (data not shown), indicating that *zorO* Δ82 is less stable. We did note, though, that the levels of *zorO* Δ82 induced by 0.2% arabinose were equivalent to that of *zorO* full-length induced by 0.0001% arabinose (Figure 3.4C). However, despite the similar RNA levels, expression of only the *zorO* full-length mRNA and not *zorO* Δ82, conferred full cell growth stasis (Figure 3.4A and 3.4B). This suggested that in addition to decreased RNA stability, *zorO* Δ82 was not translated as efficiently.

To test whether the decreased toxicity was due in part to decreased translation of the *zorO* Δ82, we performed *in vitro* translation utilizing equal amounts of *zorO* full-length and *zorO* Δ82 mRNAs. As shown in Figure 3.4D, translation of the *zorO* Δ82 mRNA was barely detectable as compared to the *zorO* full-length. Together, these data suggest that disruption of the EAP region impairs translation efficiency.

Exposing the RBS rescues the translational defect of *zorO* Δ82 A standby site is thought to increase the concentration of ribosomes around the RBS (14,29). As a result, when the RBS is transitionally exposed, the preloaded ribosomes can rapidly initiate translation. In line with this, we hypothesized that increased accessibility of the *zorO* RBS would obviate the need for a standby site. We thus disrupted the RBS-containing stem in the *zorO* Δ82 mRNA by mutating the region that base pairs with the *zorO* RBS (Figure 3.1A). This resulted in a *zorO* mutant that harbors a constitutively exposed RBS (*zorO* Δ82 RBS). The *zorO* Δ82 RBS mutant was far more toxic than the parental *zorO* Δ82 and its toxicity mirrored that of the full-length *zorO* when induced by 0.0001% arabinose (Figure 3.5A). Additionally, northern blot analysis revealed that the *zorO* Δ82 RBS was expressed at much lower levels than *zorO* full-length (Figure 3.5B), suggesting that the restored toxicity seen in the *zorO* Δ82 RBS mutant was due to more efficient translation. This was further supported by the results of the *in vitro* translation assay shown in Figure 3.4D. Taken together, our data support the hypothesis that the EAP region of the *zorO* mRNA functions to compensate for a closed *zorO* RBS.

An additional element in *zorO* that affects *zorO* translation efficiency Our data indicate that the EAP region facilitates the translation of *zorO*. Since the *zorO* mRNA harbors a long (174 nts) and extensively structured 5' UTR, we next wanted to examine if other regions of the 5' UTR contribute to *zorO* translation. We thus constructed two additional *zorO* mutants in which either the first 34 nts (Δ34) or 50 nts (Δ50) from the transcription start site were deleted (Figure 3.1A). They were subjected to the same toxicity assays and northern blot analyses as described above. The *zorO* Δ34 mutant was similar to the aforementioned *zorO* Δ28 mutant in that it exhibited a hypertoxic phenotype and the RNA levels were lower than the full-length *zorO* under the same induction conditions (Figure 3.6A, 3.6B). This suggested that the translation efficiency of *zorO* Δ34 is also higher than that of *zorO* full-length.

Interestingly, the *zorO* Δ50 mutant showed a decreased toxic phenotype as compared to the full-length when induced by 0.0001% arabinose (Figure 3.6C). Northern blot analysis indicated that *zorO* Δ50 was expressed at levels similar or even higher to that of the full-length, suggesting that the decreased toxicity was not due to differences in RNA levels (Figure 3.6D). When compared to the full-length *zorO* mRNA, there was a moderate reduction in ZorO protein production from the *zorO* Δ50 mRNA *in vitro* (Figure 3.6E). These data suggested that the *zorO* Δ50 mutant is translated less efficiently, leading to its overall decreased toxicity. Surprisingly, our structural probing results revealed that the *zorO* Δ50 possesses the same open EAP region as the *zorO* Δ28 (Figure 3.6F) even though its translation efficiency is much lower than the *zorO* Δ28 (Figure 3.6E). Given that the *zorO* Δ34 is similar to that of the *zorO* Δ28, the reduced toxicity seen in *zorO* Δ50 is likely due to truncation of the sequence from +35 to +50. Therefore, these data implied that the region between +35 to +50 harbors an additional element that contributes to *zorO* translation (see discussion).

The OrzO sRNA causes both translation inhibition and degradation of the *zorO* mRNA Our previous study demonstrated that base pairing of the OrzO sRNA to the *zorO* mRNA triggers RNase III-dependent degradation (15). Whether this is the only mechanism by which the OrzO sRNA represses *zorO* expression is unclear. As mentioned above, the structural probing data revealed that the OrzO sRNA can base pair with both the *zorO* full-length and the *zorO* Δ28 mRNAs at the 3' end of the EAP region (Figure 3.3C). If the EAP region acts as a standby site, base pairing by OrzO would prevent the ribosome from preloading onto the mRNA thus inhibiting

translation of the *zorO* mRNA. Under this scenario, the OrzO sRNA would still be capable of repressing *zorO* even if RNase III is not present. To test this, we co-expressed the *orzO* sRNA gene along with either the *zorO* full-length or the *zorO* Δ28 gene in a RNase III deletion (Δ*rnc*) background to determine if OrzO could still restore bacterial growth from ZorO-induced toxicity. As shown in Figure 3.7A and 3.7B, bacterial growth stasis caused by overexpression of either the *zorO* full-length (0.0001% arabinose) or *zorO* Δ28 (0.00002% arabinose) was rescued by co-expression of *orzO*. This indicates that the OrzO sRNA can repress ZorO toxin production in the absence of RNase III, likely through inhibiting *zorO* mRNA translation.

To directly assess the influence of the OrzO sRNA on *zorO* mRNA translation, we examined the *in vitro* translation efficiency of either *zorO* full-length or *zorO* Δ28 mRNA in the presence or absence of OrzO. The ZorO protein levels for either mRNA were reduced greatly in the presence of OrzO (Figure 3.7C). These results are concordant with the observed rescue from cellular toxicity *in vivo* (Figure 3.7A and 3.7B) and indicate that OrzO can regulate *zorO* expression at the translational level. Interestingly, the OrzO sRNA exhibited a more pronounced inhibitory effect on translation of *zorO* Δ28 mRNA as compared to *zorO* full-length mRNA. This corroborates with our structural probing results that the *zorO* Δ28 mRNA has a more accessible EAP region for OrzO sRNA to bind. We also tested if the OrzO sRNA could impact the translation of *zorO* Δ82 RBS mRNA. Only a very mild reduction in ZorO protein levels was observed upon addition of OrzO sRNA (Figure 3.7D), which further supports the idea that an opened RBS would render the EAP region dispensable. Collectively, these data suggested that the OrzO sRNA represses ZorO production by affecting both translation and stability of the *zorO* mRNA.

IV. Discussion

In this study, we show that the 5' UTR of *zorO* regulates the production of ZorO toxin from three major aspects. First, the 5' UTR affects the stability of the *zorO* mRNA as almost all *zorO* truncation mutants tested exhibited decreased RNA levels as compared to the *zorO* full-length. Second, the 5' UTR modulates the translation of the *zorO* mRNA through two distinct regions. One region, EAP, is exposed only upon processing and is a putative ribosome standby site that likely facilitates ribosome preloading onto the mRNA, thereby promoting *zorO* translation. The other region spans from +35 to +50 of *zorO* and is required for optimal translation of *zorO*, although the underlying mechanism remains unclear (see below). Third, the *zorO* 5' UTR serves

as the target site of the antitoxin sRNA OrzO. The OrzO sRNA base pairs to the 3' of the EAP region of *zorO* and inhibits *zorO* translation, likely by competing with ribosomes for the EAP region. This base pairing also stimulates *zorO* degradation in an RNase III-dependent fashion (15), thereby preventing accumulation of the *zorO* mRNA.

Optimal translation of *zorO* requires a standby site The 5' UTR mediated translational regulation of ZorO toxin synthesis is reminiscent to that of the TisB toxin of the type I pair *tisAB-istR* of *E. coli*. Without IstR-1 sRNA base pairing, the *tisAB* mRNA possesses two forms: a full-length form (+1) and a processed form (+42) (14). The RBS of both forms of *tisAB* are folded in stems. Our work revealed that the *zorO* mRNA also exists in a full-length form and a processed form (Δ28) (15). Structural probing data demonstrated that the RBS is shielded in stems in both the *zorO* full-length and the Δ28 mRNAs. The processed *tisAB* undergoes structural changes at the 5' UTR as compared to full-length *tisAB* and opens a standby site that was initially sequestered in full-length transcript (14). Similar structural rearrangement also occurs in *zorO* upon processing of the *zorO* 5' UTR, which leads to exposure of a single-stranded region in the *zorO* Δ28 that spans from +73 to +102 (the EAP region). Additionally, the processed *tisAB* exhibited robust translation whereas the full-length *tisAB* is translationally inert (14). Our results from the toxicity and *in vitro* translation assays also showed that *zorO* Δ28 has increased translation relative to full-length *zorO*. Given these extensive similarities between *zorO* and *tisAB*, it is likely that the translation of *zorO* follows the standby model and the EAP region serves as a standby site.

Unlike *tisAB*, both the *zorO* full-length and the *zorO* Δ28 forms are translatable (Figure 3.1E). Hence, the putative standby region of *zorO* can promote translation, but is not required for translation. As proposed by de Smit *et al*, a standby site operates by increasing the local concentration of ribosomes given that the time window during which a folded RBS is transiently opened is too short to efficiently recruit ribosomes (29). Thus, the effect of a standby site is dependent on the opening of the closed RBS. Our data comparing the *zorO* Δ82 versus *zorO* Δ82 RBS mutant showed that despite the perturbation of the standby site, translation of the toxin mRNAs were recovered after permanently exposing the RBS. This suggests the standby site is dispensable when the RBS is exposed. Therefore, although *zorO* full-length may not harbor an accessible standby site, transitory opening of the RBS in *zorO* may provide an opportunity for translation to occur.

Potential requirement for ribosome to preload onto a standby site As mentioned above, the translational regulation of *zorO* by its 5' UTR likely follows the standby model as proposed in *tisAB,* though the actual lengths of the standby sites vary for each. The features of a standby site that allow ribosome preloading remain unclear. For normal translation initiation, the 30S subunit of the ribosome anchors onto the mRNA through base pairing interactions between the 3' terminus of 16S rRNA and the RBS of the mRNA. We do not recognize any obvious complementary sequence in *zorO* or *tisB* that could base pair with the 3' of 16S rRNA。 However, it has been shown that alternative sequence features of the mRNA, other than complementarity to the 16S rRNA, can permit 30S subunit preloading onto mRNA. Studies by Barendt *et al* revealed that translation initiation of non-Shine Dalgarno-led genes exhibit unique preferences on certain purines or pyrimidines at the 5' UTR (30,31). We note that both the *zorO* and the *tisB* mRNAs contain a stretch of 5-nt-long sequence (5' CAACA 3') at the 3' end of predicted standby site where the corresponding antitoxin sRNA targets. Whether this stretch of sequence plays a role in anchoring the ribosome is currently under investigation.

A more widely accepted feature of a standby site suggests that preloading of the ribosome onto a standby site is predominantly dependent on the single stranded nature of the site. Using fluorescence energy transfer experiments on designed mRNAs containing different secondary structures, Studer *et al* revealed that a single-stranded region in a structured mRNA allows the binding to the 30S subunit of the ribosome in a sequence independent manner (32). Another recent study also suggests that hairpin structures harboring single-stranded surface can function as standby sites regardless of their sequences (28).

Given that the association affinity between a standby site and the 30S subunit may not be as strong as that in the translation initiation complex, direct biochemical detection of a bound form of 30S with the standby site can be challenging. It is noteworthy that the sequence and the structure requirements of a standby site may not be mutually exclusive. Hence, it is possible that one standby site can harbor both specific sequence features as well as structure features.

Optimal translation of *zorO* requires an additional site Along with the EAP standby region, we identified another region spanning from +35 to +50 that contributes to *zorO* translation. We could not resolve the structure of this region due to poor RNase T1 cutting, but our lead(II) digestion indicated that the +35 to +50 region comprises a stretch of single nucleotides and a stem (Figure

3.8). A study by Borujeni *et al* suggested that any single hairpin structure surrounded by distal and proximal single-stranded binding sites could serve as a standby site and contribute to translation (28). Hence, perhaps the single-stranded sequences in the +35 to +50 region could also promote ribosome loading. However, a synthetic sRNA that can base pair with this region failed to rescue cells from *zorO* induced toxicity in a Δ*rnc* strain (data not shown), implying that sRNA pairing to this region does not interfere with *zorO* translation. Therefore, it is unlikely that this region serves as a site for ribosome loading.

Another possibility is that this region may impact the kinetics of RNA unfolding. RNA can adopt different structures before reaching its native fold (33). The folding process can occur in a sequential fashion in which the first stable structure formed directs subsequent folding (34). In line with this, one may imagine that there could be many alternative ways for a structured RNA to reverse to a "functional" unfolded state. Perhaps the presence of the +35 to +50 region confines the choices of possible unfolding pathways and streamlines this unfolding process, therefore allowing a more rapid translation of *zorO*. Additionally, it has been suggested that the folding of some RNAs occurs in a cooperative fashion that requires interactions of different portions of the RNA, as exemplified by the RNase P of *B. subtilis* (35). This raises the possibility that the +35 to +50 region may accelerate the folding/unfolding of the *zorO* RNA but not impact the final structure. This could explain why lack of this region results in reduced translation efficiency (Figure 3.6E) although the overall structure was not affected. Additional studies of *zorO* RNA unfolding could provide mechanistic insights on how this region affects *zorO* translation.

Alternative model for *zorO* translation Another mode of regulation of type I toxins is the translation coupling model represented by *mok-hok* (10,36). In this model, translation of the *hok* toxin mRNA is dependent on the translation of a leader peptide, Mok, encoded in the long 5' UTR of *hok*. The RBS of *mok* is sequestered in secondary structure and only exposed upon 3' processing of *hok* (12,13). In *zorO*, structural alteration is mediated by 5' processing instead of the 3' processing seen in *hok*. Further, no AUG start codon is found within the *zorO* 5' UTR. Although there are several GUG or UUG sequences that may serve as an alternative start codon, no canonical RBS is detected upstream of these sequences nor are they located within the EAP. Given that the translation coupling model requires the presence of a leader peptide, it is unlikely that the *zorO* mRNA uses this mechanism to initiate its translation.

Besides ZorO and TisB, 5' processing has been observed in two other type I toxins in *E. coli*, ShoB and DinQ, and appears linked to the translational efficiency of those toxin mRNAs. Specifically, two forms of *shoB* transcripts were detected: a full-length form (~320 nts) and a 5' processed form (~280 nts) (5), but only the processed form showed measurable β-galactosidase activities when fused translationally to *lacZ* (2). Further, *dinQ* possesses four processed forms yet only one transcript (+44) was translationally active (37). These observations suggest that optimal translation of *shoB* or *dinQ* requires 5' processing, likely through de-repressing the inhibitory effects caused by their respective 5' UTRs. This raises the possibility that similar to *zorO* and *tisB*, the *shoB* and *dinQ* may use standby sites to facilitate translation; furthermore, their respective antitoxins OhsC and AgrB may act by targeting the standby sites. Surprisingly, the *zorO-orzO*, *tisB-istR*, *shoB-ohsC*, and *dinQ-agrB* type I pairs share two distinct features: the toxin is encoded divergently from the corresponding antitoxin sRNA gene and base pairing occurs distal from the RBS at the toxin 5' UTR. Indeed, these four type I loci are the only ones identified to date that follow such genetic arrangement and base pairing pattern. It would be interesting to see whether the *shoB-ohsC* and *dinQ-agrB* pairs employ the standby mechanism to regulate toxin production, and if they do, the next fascinating question would be: can we predict the regulatory mechanisms of a type I pair based on its intrinsic genetic arrangement and the base pairing pattern?

References

1. Fozo, E.M., Hemm, M.R. and Storz, G. (2008) Small toxic proteins and the antisense RNAs that repress them. *Microbiol Mol Biol Rev*, **72**, 579-589.
2. Fozo, E.M., Kawano, M., Fontaine, F., Kaya, Y., Mendieta, K.S., Jones, K.L., Ocampo, A., Rudd, K.E. and Storz, G. (2008) Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. *Mol Microbiol*, **70**, 1076-1093.
3. Unoson, C. and Wagner, E.G. (2008) A small SOS-induced toxin is targeted against the inner membrane in *Escherichia coli*. *Mol Microbiol*, **70**, 258-270.
4. Patel, S. and Weaver, K.E. (2006) Addiction toxin Fst has unique effects on chromosome segregation and cell division in *Enterococcus faecalis* and *Bacillus subtilis*. *J Bacteriol*, **188**, 5374-5384.
5. Kawano, M., Reynolds, A.A., Miranda-Rios, J. and Storz, G. (2005) Detection of 5'- and 3'-UTR-derived small RNAs and *cis*-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Res*, **33**, 1040-1050.
6. Kawano, M., Aravind, L. and Storz, G. (2007) An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Mol Microbiol*, **64**, 738-754.
7. Silvaggi, J.M., Perkins, J.B. and Losick, R. (2005) Small untranslated RNA antitoxin in *Bacillus subtilis*. *J Bacteriol*, **187**, 6641-6650.
8. Durand, S., Gilet, L. and Condon, C. (2012) The essential function of *B. subtilis* RNase III is to silence foreign toxin genes. *PLoS Genet*, **8**, e1003181.
9. Brantl, S. and Jahn, N. (2015) sRNAs in bacterial type I and type III toxin-antitoxin systems. *FEMS Microbiol Rev*, **39**, 413-427.
10. Gerdes, K., Larsen, J.E. and Molin, S. (1985) Stable inheritance of plasmid R1 requires two different loci. *J Bacteriol*, **161**, 292-298.
11. Gerdes, K., Thisted, T. and Martinussen, J. (1990) Mechanism of post-segregational killing by the *hok/sok* system of plasmid R1: *sok* antisense RNA regulates formation of a *hok* mRNA species correlated with killing of plasmid-free cells. *Mol Microbiol*, **4**, 1807-1818.
12. Franch, T. and Gerdes, K. (1996) Programmed cell death in bacteria: translational repression by mRNA end-pairing. *Mol Microbiol*, **21**, 1049-1060.

13. Franch, T., Gultyaev, A.P. and Gerdes, K. (1997) Programmed cell death by *hok/sok* of plasmid R1: processing at the *hok* mRNA 3'-end triggers structural rearrangements that allow translation and antisense RNA binding. *J Mol Biol*, **273**, 38-51.

14. Darfeuille, F., Unoson, C., Vogel, J. and Wagner, E.G. (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Mol Cell*, **26**, 381-392.

15. Wen, J., Won, D. and Fozo, E.M. (2014) The ZorO-OrzO type I toxin-antitoxin locus: repression by the OrzO antitoxin. *Nucleic Acids Res*, **42**, 1930-1946.

16. Fozo, E.M., Makarova, K.S., Shabalina, S.A., Yutin, N., Koonin, E.V. and Storz, G. (2010) Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families. *Nucleic Acids Res*, **38**, 3743-3759.

17. Kunkel, T.A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc Natl Acad Sci U S A*, **82**, 488-492.

18. Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K. and Pease, L.R. (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene*, **77**, 61-68.

19. Bahler, J., Wu, J.Q., Longtine, M.S., Shah, N.G., McKenzie, A., 3rd, Steever, A.B., Wach, A., Philippsen, P. and Pringle, J.R. (1998) Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast*, **14**, 943-951.

20. Guzman, L.M., Belin, D., Carson, M.J. and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J Bacteriol*, **177**, 4121-4130.

21. Kawano, M., Oshima, T., Kasai, H. and Mori, H. (2002) 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*, **45**, 333-349.

22. Masse, E., Escorcia, F.E. and Gottesman, S. (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev*, **17**, 2374-2383.

23. Opdyke, J.A., Kang, J.G. and Storz, G. (2004) GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J Bacteriol*, **186**, 6698-6705.

24. Opdyke, J.A., Fozo, E.M., Hemm, M.R. and Storz, G. (2011) RNase III participates in GadY-dependent cleavage of the *gadX-gadW* mRNA. *J Mol Biol*, **406**, 29-43.

25. Nilsen, T.W. (2013) Gel purification of RNA. *Cold Spring Harb Protoc*, **2013**, 180-183.
26. Beisel, C.L., Updegrove, T.B., Janson, B.J. and Storz, G. (2012) Multiple factors dictate target selection by Hfq-binding small RNAs. *EMBO J*, **31**, 1961-1974.
27. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*, **31**, 3406-3415.
28. Espah Borujeni, A., Channarasappa, A.S. and Salis, H.M. (2014) Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res*, **42**, 2646-2659.
29. de Smit, M.H. and van Duin, J. (2003) Translational standby sites: how ribosomes may deal with the rapid folding kinetics of mRNA. *J Mol Biol*, **331**, 737-743.
30. Barendt, P.A., Shah, N.A., Barendt, G.A. and Sarkar, C.A. (2012) Broad-specificity mRNA-rRNA complementarity in efficient protein translation. *PLoS Genet*, **8**, e1002598.
31. Barendt, P.A., Shah, N.A., Barendt, G.A., Kothari, P.A. and Sarkar, C.A. (2013) Evidence for context-dependent complementarity of non-Shine-Dalgarno ribosome binding sites to *Escherichia coli* rRNA. *ACS Chem Biol*, **8**, 958-966.
32. Studer, S.M. and Joseph, S. (2006) Unfolding of mRNA secondary structure by the bacterial translation initiation complex. *Mol Cell*, **22**, 105-115.
33. Li, P.T., Vieregg, J. and Tinoco, I., Jr. (2008) How RNA unfolds and refolds. *Annu Rev Biochem*, **77**, 77-100.
34. Mahen, E.M., Watson, P.Y., Cottrell, J.W. and Fedor, M.J. (2010) mRNA secondary structures fold sequentially but exchange rapidly in vivo. *PLoS Biol*, **8**, e1000307.
35. Baird, N.J., Westhof, E., Qin, H., Pan, T. and Sosnick, T.R. (2005) Structure of a folding intermediate reveals the interplay between core and peripheral elements in RNA folding. *J Mol Biol*, **352**, 712-722.
36. Gerdes, K., Poulsen, L.K., Thisted, T., Nielsen, A.K., Martinussen, J. and Andreasen, P.H. (1990) The *hok* killer gene family in Gram-negative bacteria. *New Biol*, **2**, 946-956.
37. Weel-Sneve, R., Kristiansen, K.I., Odsbu, I., Dalhus, B., Booth, J., Rognes, T., Skarstad, K. and Bjoras, M. (2013) Single transmembrane peptide DinQ modulates membrane-dependent activities. *PLoS Genet*, **9**, e1003260.
38. Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*, **9**, 671-675.

39. Vogel, J., Argaman, L., Wagner, E.G. and Altuvia, S. (2004) The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Curr Biol*, **14**, 2271-2276.

Appendix III: chapter III tables and figures

Table 3.1. Strains and plasmids used in this study.

Name	Relevant genotype or description	Source
Escherichia coli EDL933	Wild type EHEC, O157:H7	D. Friedman
DJ480	MG1655 Δ*lacX74* (wild type strain)	D. Jin, (11)
UTK007	DJ480 PCP18-*araE*	(12)
UTK011	UTK007 Δ*rnc::kan*	(12)
Plasmids		
pFA6a-GFP(S65T)-kanMX6	Amp^R; Km^R	(24)
pEF21	Cm^R; P_{BAD} promoter; 10 - 12 copies per cell	(46)
pEF21-*zorO* full-length	Cm^R	(12)
pEF21-*zorO* Δ28	Cm^R	This study
pAZ3	Cm^R; P_{BAD} promoter; 15 - 20 copies per cell	(17)
pAZ3-*zorO* full-length	Cm^R	This study
pAZ3-*zorO* Δ28	Cm^R	This study
pAZ3-*zorO* Δ34	Cm^R	This study
pAZ3-*zorO* Δ50	Cm^R	This study
pAZ3-*zorO* Δ82	Cm^R	This study
pAZ3-*zorO* Δ82 RBS	Cm^R	This study
pAZ3-*zorO* UTR-*gfp*	Cm^R	This study
pAZ3-Δ28 UTR-*gfp*	Cm^R	This study
pBR-pLac	Amp^R; P_{LlacO} promoter	(11)
pBR-pLac-*orzO*	Amp^R	(12)
pGEM®-3Zf(+)	Amp^R; T7 promoter	Promega
pGEM-T7-*zorO* full-length	Amp^R	This study
pGEM-T7-*zorO* Δ50	Amp^R	This study
pGEM-T7-*zorO* Δ82	Amp^R	This study
pGEM-T7-*zorO* Δ82 RBS	Amp^R	This study

Table 3.2. Oligonucleotides used in this study

Name	Sequence[a]	Use
EF524	GTTGGTACGAAACGTTGCTCTCCG	Northern analysis of *zorO*, *zorO* UTR-*gfp* and their derivatives
EF1065	CAGTGAGTGAGATTTATAATCGAATTCGTTGGGAG	pAZ3-*zorO* full-length PCR; pAZ3-*zorO* UTR-*gfp* PCR; pGEM-T7-*zorO* full-length PCR
EF1068	ATACTCAAGCTTATTAAAGTCGCAGCACATGCAAC	pAZ3-*zorO* full-length PCR; pAZ3-*zorO* Δ28 PCR; pAZ3-*zorO* Δ50 PCR; pAZ3-*zorO* Δ82 PCR; pEF21-*zorO* Δ28 PCR
EF1141	GGTTGTGCCGGATCGGAATTCTTTTAAGTCCTGGCTGC	pAZ3-*zorO* Δ28 PCR; pAZ3-Δ28 UTR-*gfp* PCR
EF1127	GACCAAGAATTCGTCCTGGCTGCCGGACGGGTGGTGCCGC	pAZ3-*zorO* Δ34 PCR
EF1066	GATAATTAAGCTTGCAGCACATGCAACTTGAAG	pAZ3-*zorO* Δ34 PCR
EF1170	CAATTTTAAGTCCTGGCTGGAATTCGGGTGGTGCCGCAGGC	pAZ3-*zorO* Δ50 PCR; pGEM-T7-*zorO* Δ50 PCR
EF392	GCCCTGGAATTCAGAGCAACGTTTCGTACCAAC	pAZ3-*zorO* Δ82 PCR; pGEM-T7-*zorO* Δ82 PCR
EF1290	GTGTAAGGGTAAGGTGCTGGTGTTGCG**TTGGTA**AATAAGGAGA GCGGATGGACACGCTGAC	pAZ3-*zorO* Δ82 RBS PCR; pGEM-T7-*zorO* Δ82 RBS PCR
EF1291	GTCAGCGTGTCCATCCGCTCTCCTTATT**TACCAA**CGCAACACCA GCACCTTACCCTTACAC	pAZ3-*zorO* Δ82 RBS PCR; pGEM-T7-*zorO* Δ82 RBS PCR
EF1181	GAAAAGTTCTTCTCCTTTCATCCGCTCTCCTTATTAAGG	pAZ3-*zorO* UTR-*gfp* PCR
EF1182	CCTTAATAAGGAGAGCGGATGAAAGGAGAAGAACTTTTC	pAZ3-*zorO* UTR-*gfp* PCR
EF1183	GAAATTCGCTTATTTAGAAGCTTCGCGCCCTATTTGTATAG	pAZ3-*zorO* UTR-*gfp* PCR
EF1328	GGTTGTGCCGGATCGCTGCAGTTTTAAGTCCTGGCTGC	pEF21-*zorO* Δ28 PCR
EF1343	GAGTATAGCTCCCGGGAAGGGGGAAACGGTATTC	pGEM-T7 *zorO* full-length PCR; pGEM-T7-*zorO* Δ50 PCR; pGEM-T7-*zorO* Δ82 PCR; T7-*zorO* Δ28 PCR
EF1408	GAGTCTGCAGAAGCTTTAATACGACTCACTATAGGGCGAATTC TTTTAAGTCCTGGCTGCCGGACGGGTG	T7-*zorO* Δ28 PCR

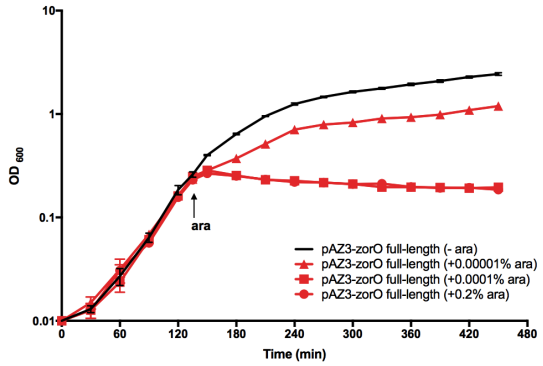
[a] 5' - 3', restriction sites underlined, nucleotides altered via site directed mutagenesis in red.

Figure 3.1. The 5' processed form of *zorO* (*zorO* Δ28) shows enhanced translation. **(A)** Sequence details of the *zorO* gene. Region of base pairing is shaded in grey. The predicted -35 and -10 promoter elements, the start codon, and the stop codon of *zorO* are indicated in bold. Transcription start site of the *zorO* full-length is highlighted in red. Transcription start site of each *zorO* truncation mutant is highlighted in color with the mutant name listed in the same color. The EAP region is labeled in red. The RBS of *zorO* is highlighted in yellow and the region base pairing with the RBS in the stem structure is underlined. **(B)** *E. coli* strain UTK007 harboring pAZ3-*zorO* was grown to mid-log, split, and arabinose added to the final concentration indicated by the arrow. Shown are the averages ± standard deviations for three independent cultures. **(C)** UTK007 harboring pAZ3-*zorO* or pAZ3-*zorO* Δ28 was grown in LB supplemented with 0.2% glucose to exponential phase, washed, and arabinose was added as indicated by the arrow. Shown are the mean values ± standard deviations for three independent cultures. **(D)** Total RNA was isolated from *E. coli* MG1655 harboring pAZ3-*zorO* or pAZ3-*zorO* Δ28 at the indicated times following the addition of 0.0001% arabinose to exponentially growing cultures (OD_{600} = 0.2 - 0.3). Shown is a representative of three independent northern blots. **(E)** *In vitro* translation assays with [^{35}S]-Met were performed as described in the Material and Methods with 0.25 μM of *zorO* full-length or *zorO* Δ28 mRNAs. No RNA was added to the negative control. Quantification of band intensities were determined using ImageJ (38). The intensity of the band representing the *zorO* full-length was set to 100% translational efficiency.

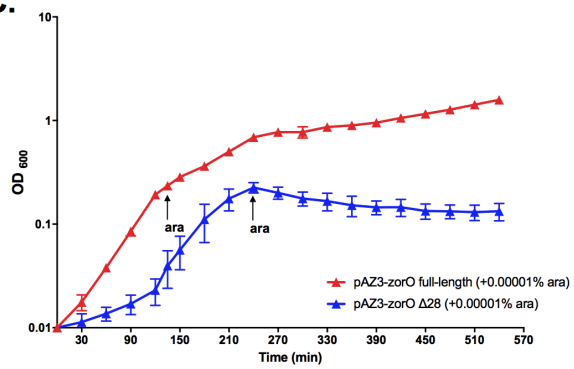
A.

```
                 -35                     -10          +1 zorO
  1  ATCCACACTCCATTGCAATTCAGTGAGTGAGATTTATAATCCTGCCCGTTGGGAGGTTGT
  1  TAGGTGTGAGGTAACGTTAAGTCACTCACTCTAAATATTAGGACGGGCAACCCTCCAACA

                 Δ28   Δ34             Δ50
 61  GCCGGATCGGACCAATTTTAAGTCCTGGCTGCCGGACGGGTGGTGCCGCAGGCGGTGCCC
 61  CGGCCTAGCCTGGTTAAAATTCAGGACCGACGGCCTGCCCACCACGGCGTCCGCCACGGG

     EAP     Δ82   base pairing
121  TGATCCGGAGAGCAACGTTTCGTACCAACATACGTAAACCAGTATCGACAGGCAGTGTAA
121  ACTAGGCCTCTCGTTGCAAAGCATGGTTGTATGCATTTGGTCATAGCTGTCCGTCACATT

                                RBS      zorO start codon
181  GGGTAAGGTGCTGGTGTTGCGCTCCTTAATAAGGAGAGCGGATGGACACGCTGACACAAA
181  CCCATTCCACGACCACAACGCGAGGAATTATTCCTCTCGCCTACCTGTGCGACTGTGTTT

241  AGTTAACCGTGCTCATTGCCGTACTGGAGTTATTAGTGGCTCTGTTACGGTTGATTGATT
241  TCAATTGGCACGAGTAACGGCATGACCTCAATAATCACCGAGACAATGCCAACTAACTAA

           zorO stop codon
301  TGTTGAAGTAACGGGCAAATTGAAACGCATCTAAGGGGGAATACCGTTTCCCCCTTTAAG
301  ACAACTTCATTGCCCGTTTAACTTTGCGTAGATTCCCCCTTATGGCAAAGGGGGAAATTC
```

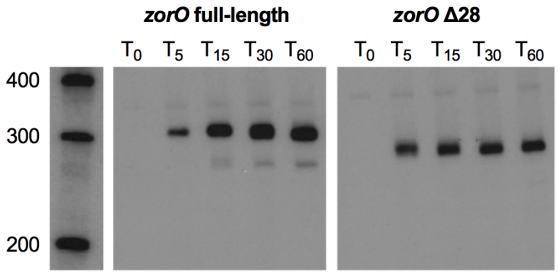
B.



C.



D.



E.

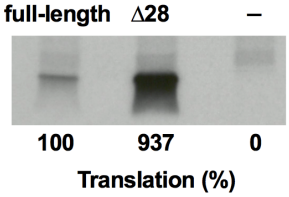


Figure 3.1. (continued)

Figure 3.2. The 5' UTR-mediated translation effects can be transferred to *gfp*. **(A)** Shown are histograms comparing fluorescence intensity of GFP in pAZ3- *zorO* UTR-*gfp* or pAZ3-Δ28 UTR-*gfp* without arabinose induction. Black line indicates 0 min post induction, red indicates 30 min post induction, and blue indicates 60 min post induction. Shown is a representative histogram of three independent experiments. **(B)** Shown are histograms comparing fluorescence intensity of GFP from cells harboring either pAZ3-*zorO* UTR-*gfp* or pAZ3-Δ28 UTR-*gfp* after 0.2% arabinose induction. The black line indicates 0 min post induction, red indicates 30 min post induction, and blue indicates 60 min post induction. Shown is the representative histogram of three independent experiments. **(C)** Fold difference in *gfp* expression as determined by geometric mean fluorescence and normalized to uninduced controls. Gray bars indicate pAZ3-*zorO* UTR-*gfp* and black bars indicate pAZ3-Δ28 UTR-*gfp*. Shown are the mean values ± standard deviations for three independent experiments.

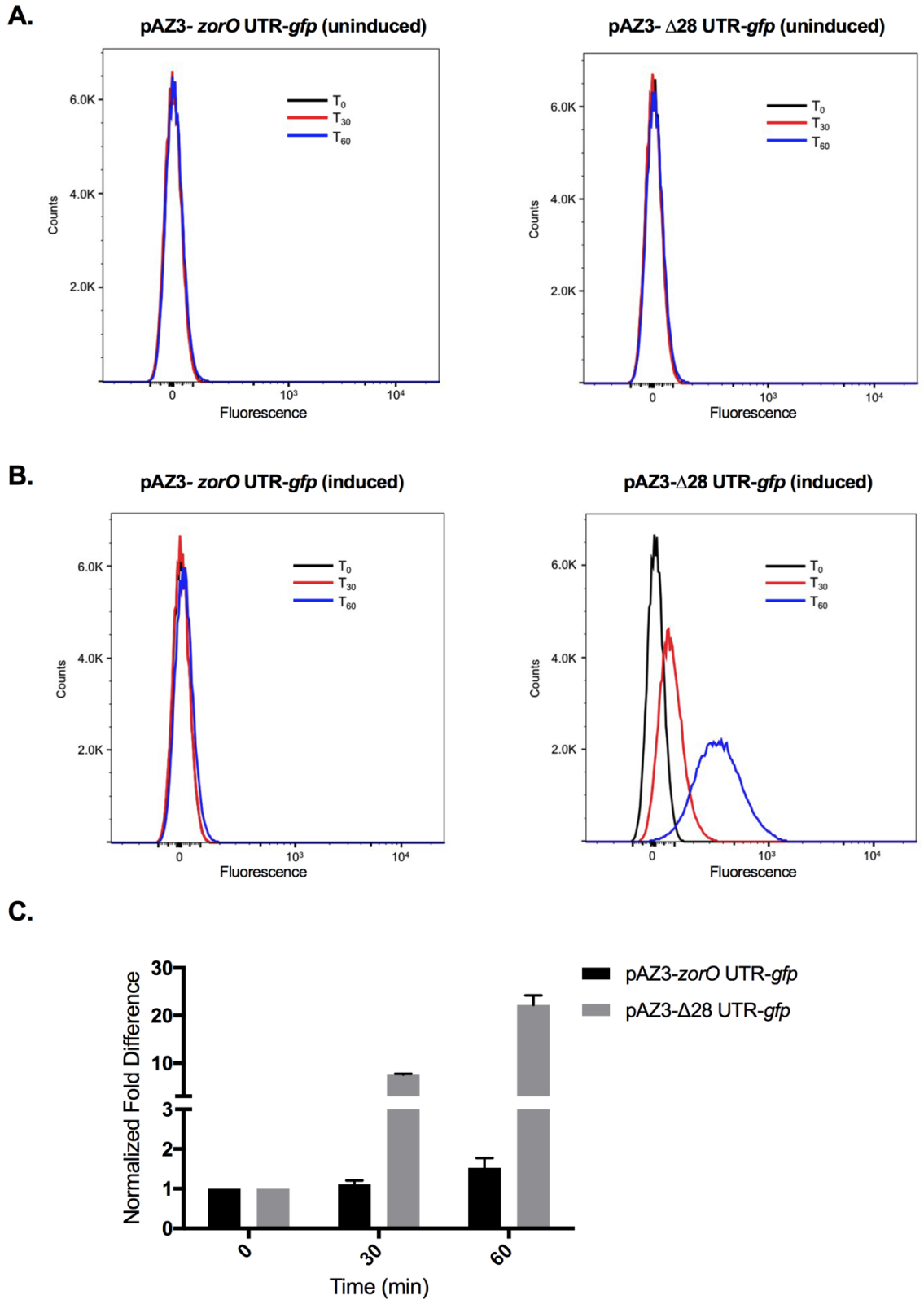


Figure 3.2. (continued)

Figure 3.3. The 5' processing of *zorO* does not alter RBS structure, but exposes a single-stranded region (EAP region). **(A)** The RBS of both *zorO* full-length and *zorO* Δ28 are in a stem structure. 5′-end-labeled *zorO* full-length or *zorO* Δ28 RNA was subjected to RNase T1 and lead(II) acetate cleavage as described in the Material and Methods. C, control RNA; Ladder/T1, RNase T1 ladder of indicated RNA; Ladder/OH, Alkaline ladder; Reaction/T1, RNase T1 digestion under denaturing condition; and Reaction/Lead, Lead(II) acetate digestion under denaturing condition. The position of selected cleaved G residues is given at the left of each T1 ladder. Yellow boxes indicate the RBS region of each *zorO* transcript. **(B)** *In vitro* structure probing of 3'-end-labeled *zorO* and its derivatives. 3′-end-labeled *zorO* full-length, *zorO* Δ28, *zorO* Δ50, or *zorO* Δ82 RNA (1.67 μM) was subjected to RNase T1 and lead(II) acetate cleavage as outlined in Figure 3A. The position of selected cleaved G residues is given at the left of the T1 ladder. Yellow boxes indicate the corresponding RBS of *zorO* full-length, *zorO* Δ28, *zorO* Δ50, and *zorO* Δ82. **(C)** *In vitro* structure probing was conducted as in (A). The position of selected cleaved G residues is given at the left of each T1 ladder. Red boxes indicate the corresponding EAP region (+73 to +102 from *zorO* full-length transcription start site) of *zorO* full-length and *zorO* Δ28. **(D)** The OrzO sRNA can base pair with both *zorO* full-length and *zorO* Δ28 at the 3' end of EAP region. Lead(II) acetate structure probing was conducted as in (A) with addition of 0 (-), 0.25 μM (+), and 0.5 μM (++) unlabeled OrzO sRNA. Red boxes indicate the corresponding EAP region of the *zorO* full-length and *zorO* Δ28. Black bars indicate the region protected by OrzO binding. Base pairing between *zorO* and OrzO is shown on the bottom.

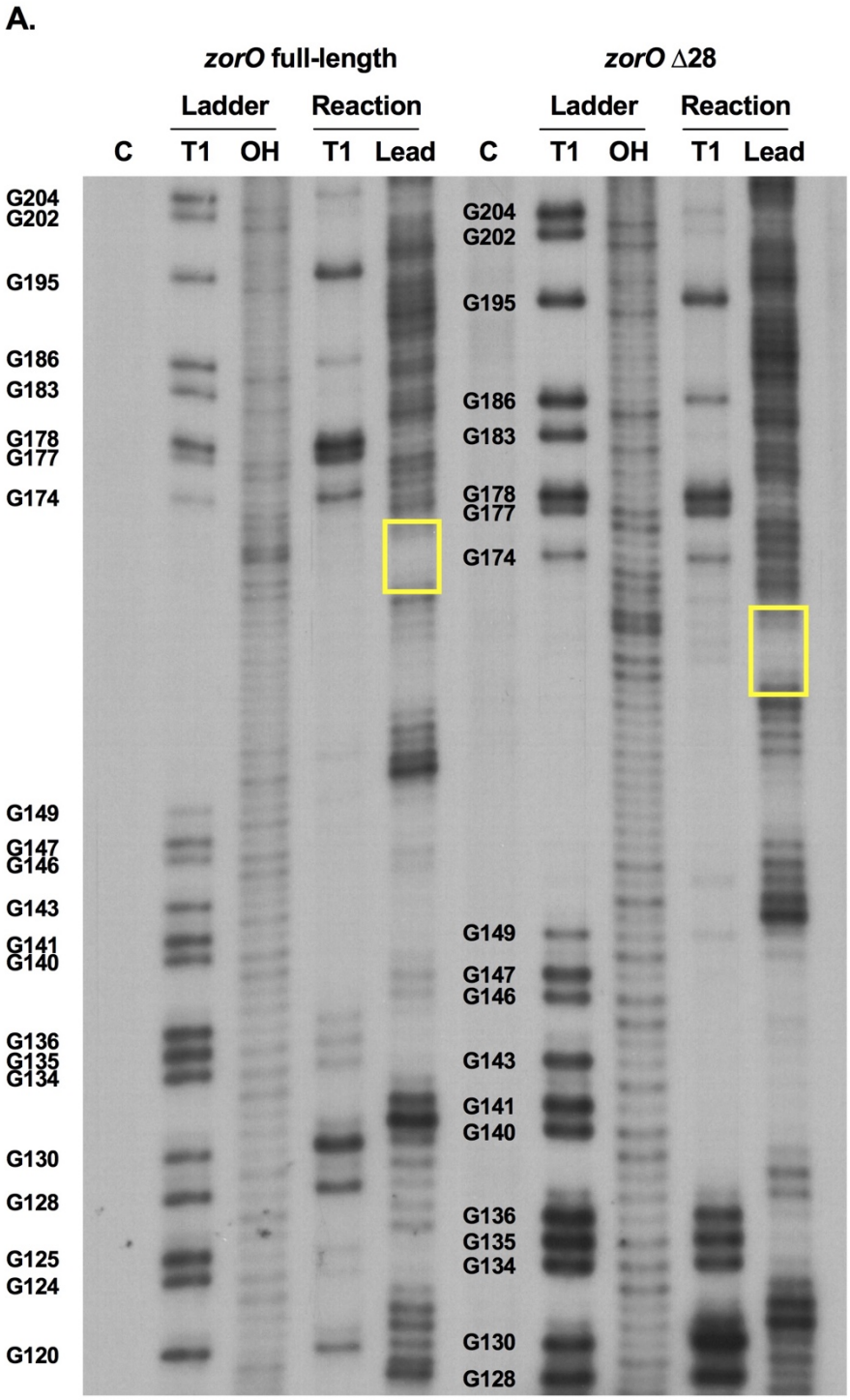


Figure 3.3. (continued)

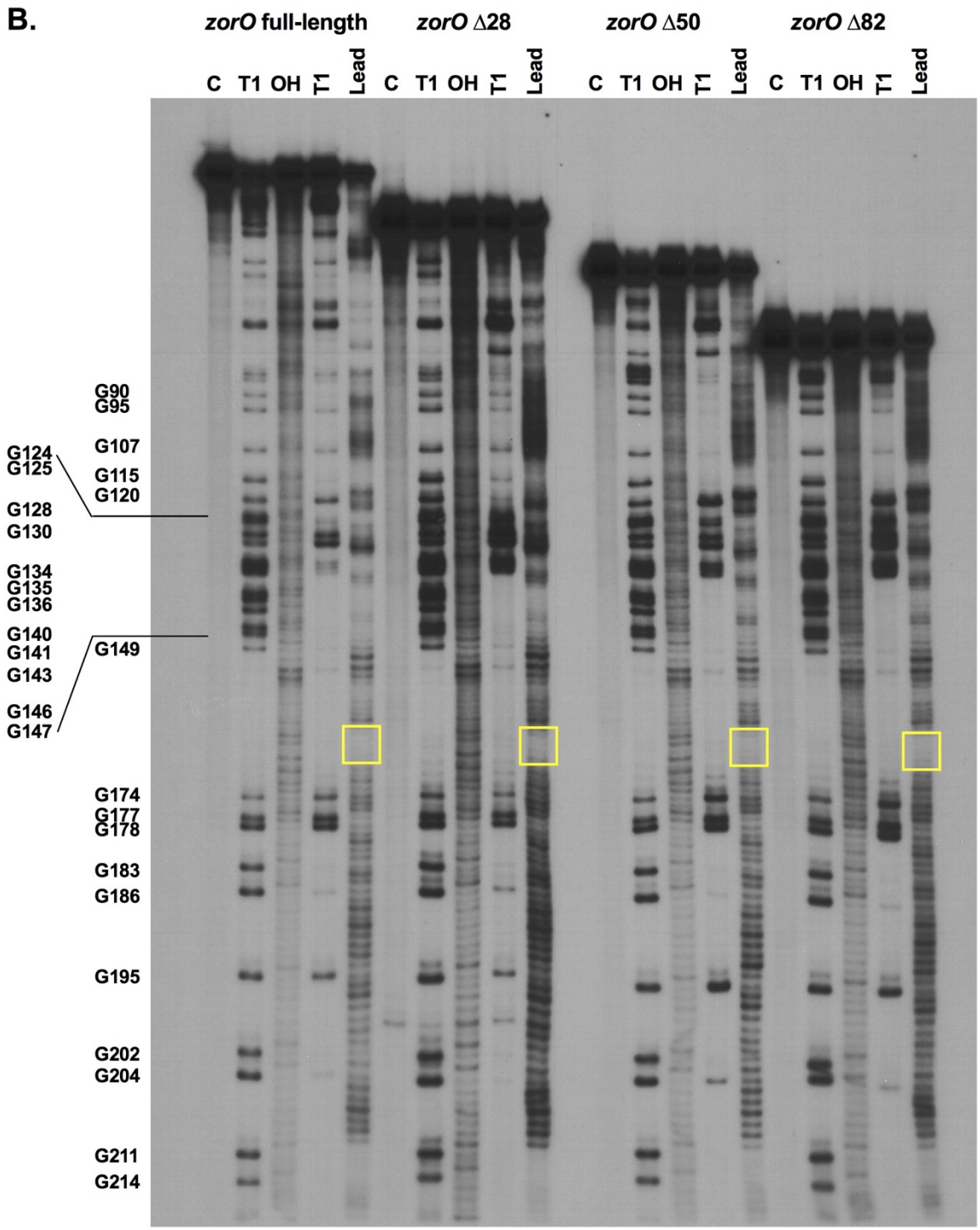


Figure 3.3. (continued)

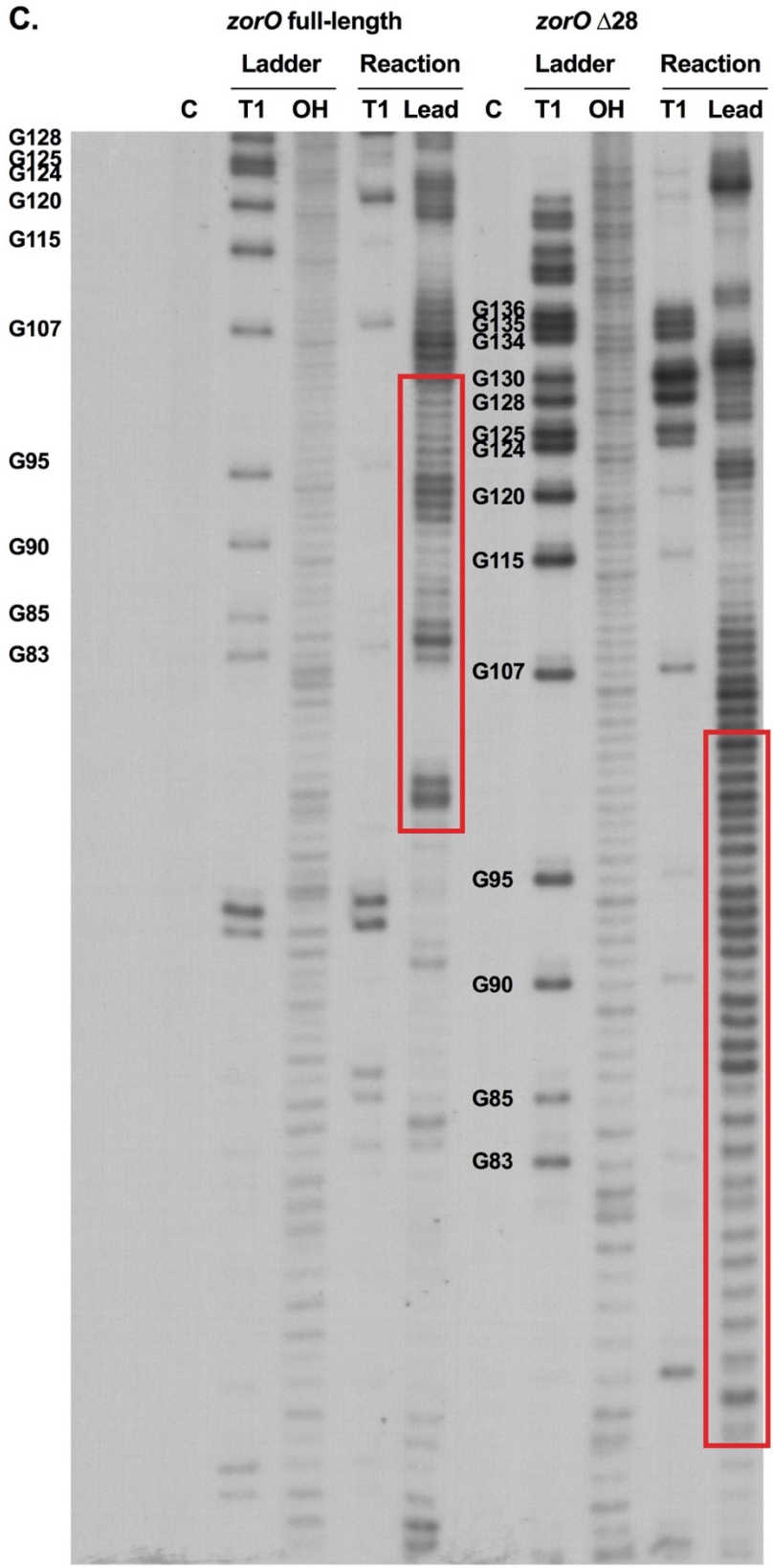


Figure 3.3. (continued)

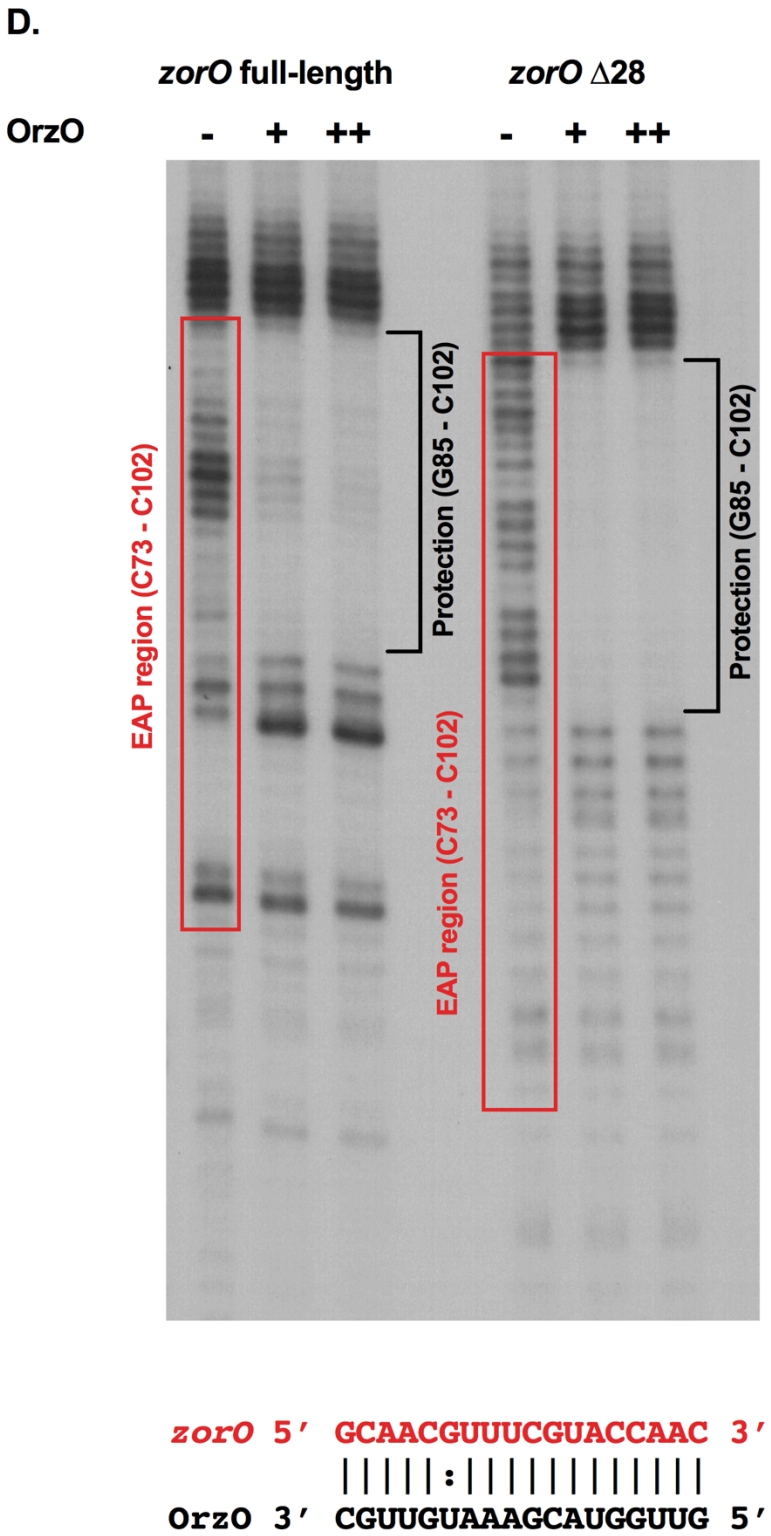


Figure 3.3. (continued)

Figure 3.4. Disruption of the EAP region of *zorO* leads to decreased translation. **(A)** Toxicity assay was performed as in Figure 3.1A with either *zorO* full-length or *zorO* Δ82 induced by 0.0001% arabinose. Shown are the mean values ± standard deviations for three independent cultures. **(B)**. Toxicity assay was performed as in Figure 3.1A with either *zorO* full-length or *zorO* Δ82 induced by 0 (uninduced control) or 0.2% arabinose. Shown are the mean values ± standard deviations for three independent cultures. **(C)** Expression of *zorO* Δ82 induced by 0.2% arabinose is similar to or higher than *zorO* full-length induced by 0.0001% arabinose. Total RNA was isolated from *E. coli* MG1655 harboring pAZ3-*zorO* or pAZ3-*zorO* Δ82 at 0, 5, 15, 30, and 60 min following the addition of either 0.0001% (to induce *zorO* full-length) or 0.2% (to induce *zorO* Δ82) arabinose to exponentially growing cultures. Shown is a representative of three independent northern blots. **(D)** *In vitro* translation assays with [^{35}S]-Met were performed as described in the Material and Methods with 0.25 μM of *zorO* full-length, *zorO* Δ82, or *zorO* Δ82 RBS mRNAs. No mRNA was added to the negative control. Quantification of band intensities were determined using ImageJ (38). The intensity of the band representing the *zorO* full-length was set to 100% translational efficiency.

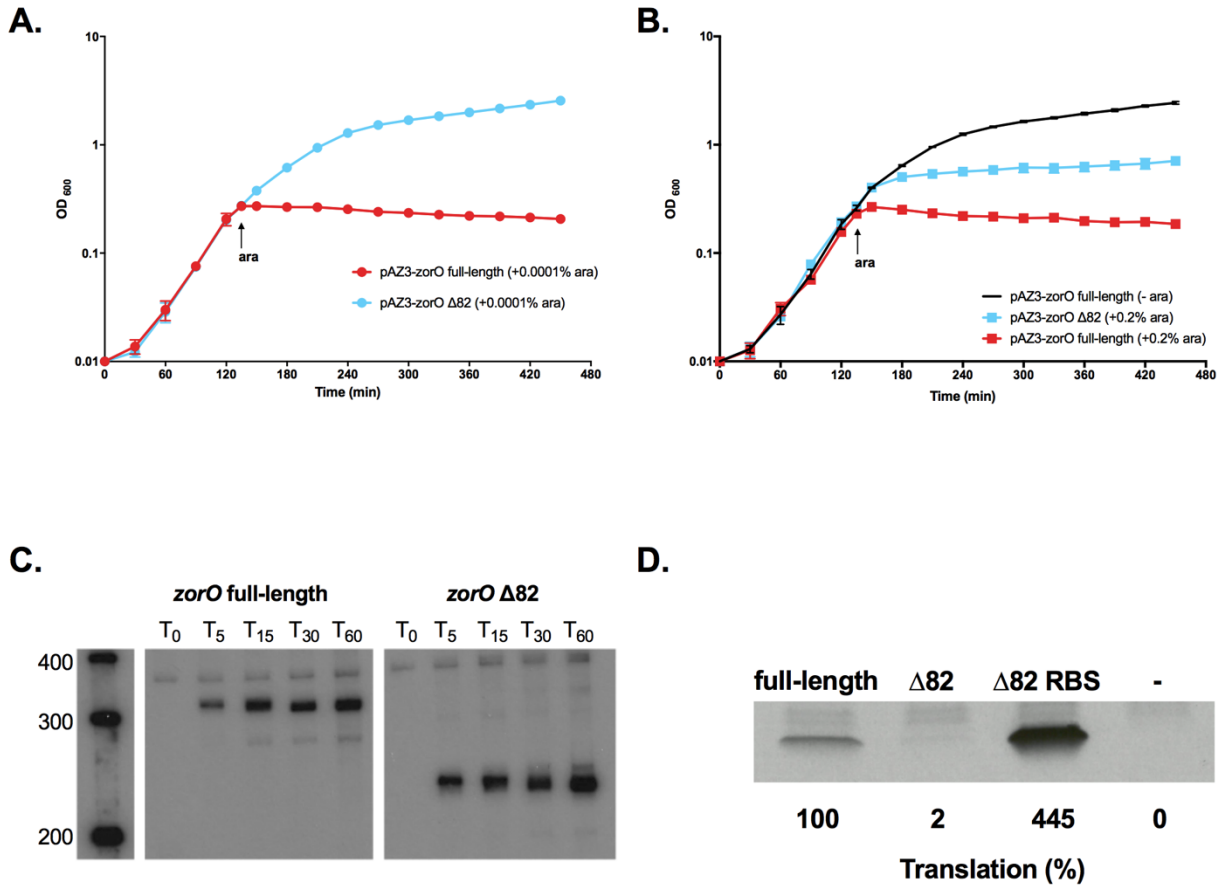


Figure 3.4. (continued)

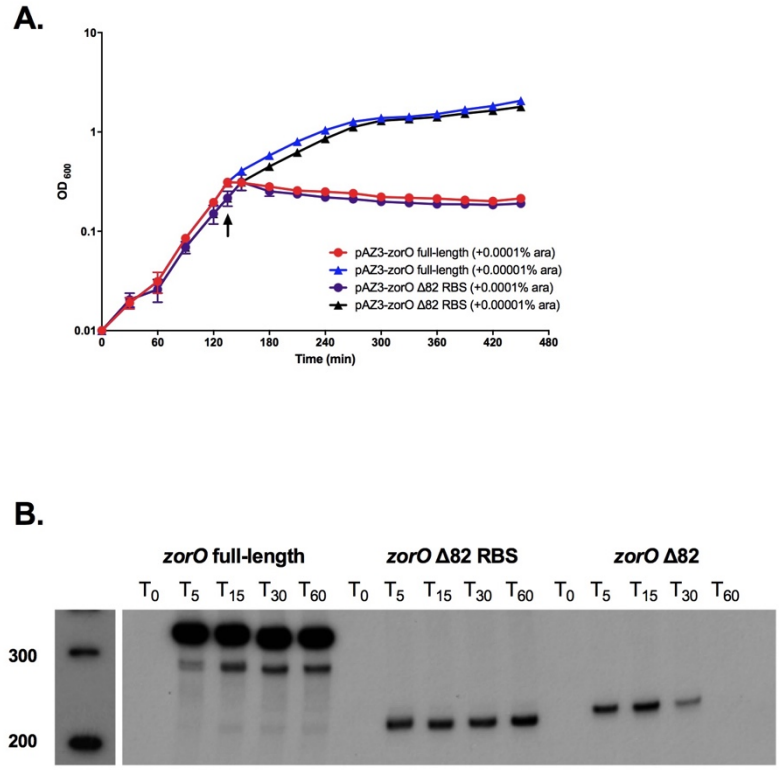


Figure 3.5. Exposing the RBS rescues the translational defect of *zorO* Δ82. **(A)** Toxicity assay was performed as in Figure 3.1A with either *zorO* full-length or *zorO* Δ82 RBS by either 0.0001% or 0.00001% arabinose. Shown are the mean values ± standard deviations for three independent cultures. **(B)** Total RNA was isolated from *E. coli* MG1655 harboring pAZ3-*zorO*, pAZ3-*zorO* Δ82, or pAZ3-*zorO* Δ82 RBS at 0, 5, 15, 30, and 60 min following the addition of 0.0001% arabinose to exponentially growing cultures. Shown is a representative of three independent northern blots.

Figure 3.6. The 5' UTR contains an additional region that impacts *zorO* translation. **(A)**. Toxicity assay was performed as in Figure 3.1B with either *zorO* full-length or *zorO* Δ34 induced by 0.00001% arabinose. Shown are the mean values ± standard deviations for three independent cultures. **(B)**. Total RNA was isolated from *E. coli* MG1655 harboring pAZ3-*zorO* full-length or pAZ3-*zorO* Δ34 at 0, 5, 15, 30, and 60 min following the addition of 0.0001% arabinose to exponentially growing cultures. Shown is a representative of three independent northern blots. **(C)** Toxicity assay was performed as in Figure 3.1A with either *zorO* full-length or *zorO* Δ50 induced by 0.0001% arabinose. Shown are the mean values ± standard deviations for three independent cultures. **(D)** Total RNA was isolated from *E. coli* MG1655 harboring pAZ3-*zorO* or pAZ3-*zorO* Δ50 at 0, 5, 15, 30, and 60 min following the addition of 0.0001% arabinose to exponentially growing cultures. Shown is a representative of three independent northern blots. **(E)** *In vitro* translation assays with [^{35}S]-Met were performed as described in the Material and Methods with 0.25 μM of *zorO* full-length, *zorO* Δ28, or *zorO* Δ50 mRNAs. No mRNA was added to the negative control. Quantification of band intensities were determined using ImageJ (38). The intensity of the band representing the *zorO* full-length was set to 100% translational efficiency. **(F)** *In vitro* structure probing was conducted as in Figure 3.3A. Red boxes indicate the corresponding EAP region of *zorO* Δ28 and *zorO* Δ50.

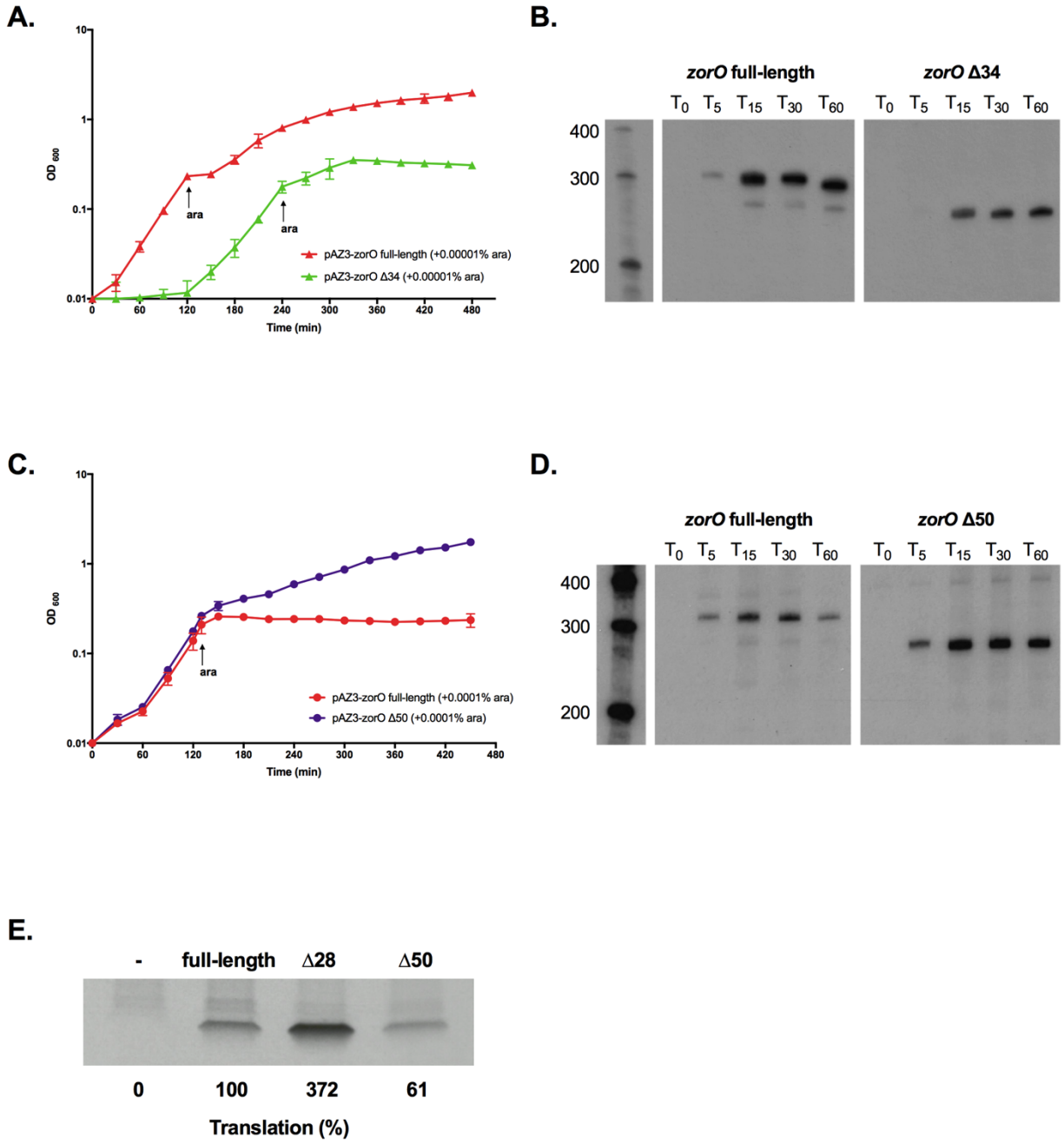


Figure 3.6. (continued)

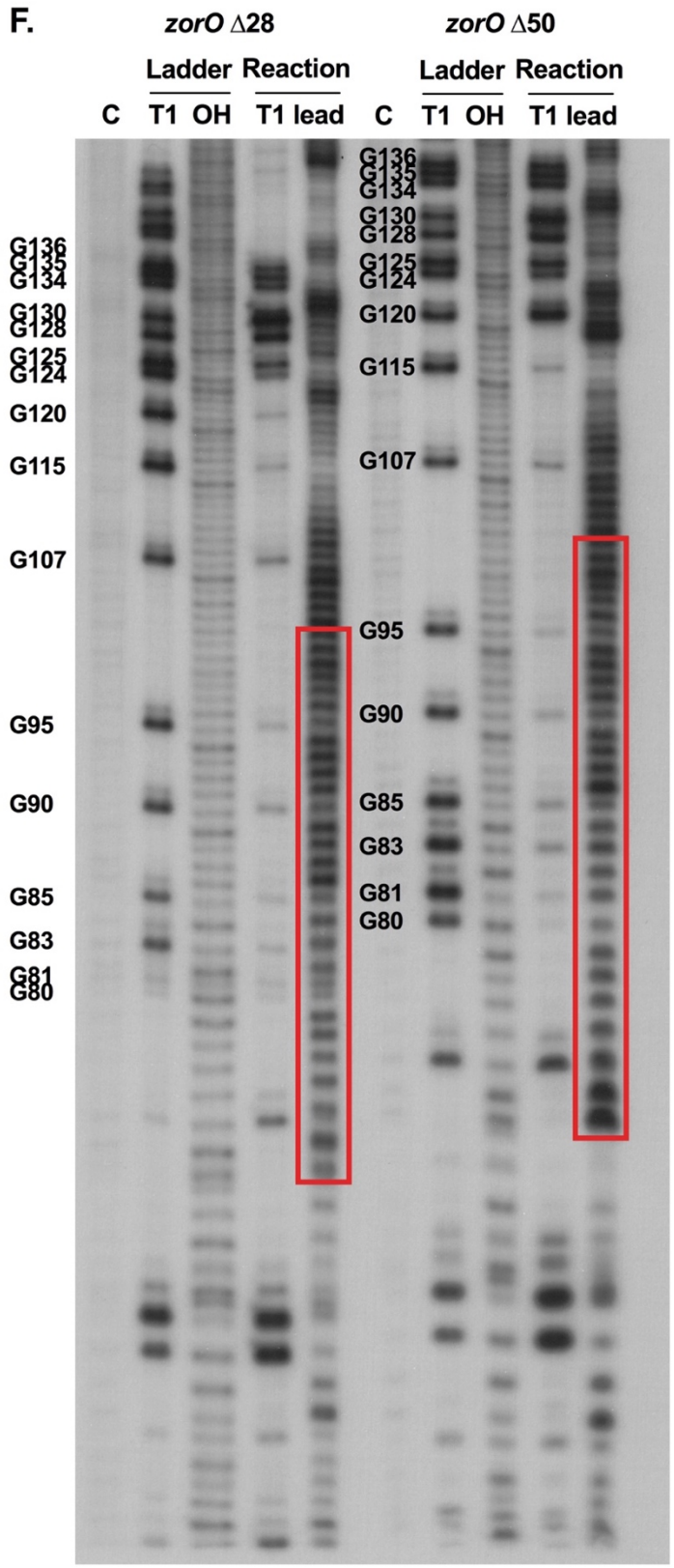


Figure 3.6. (continued)

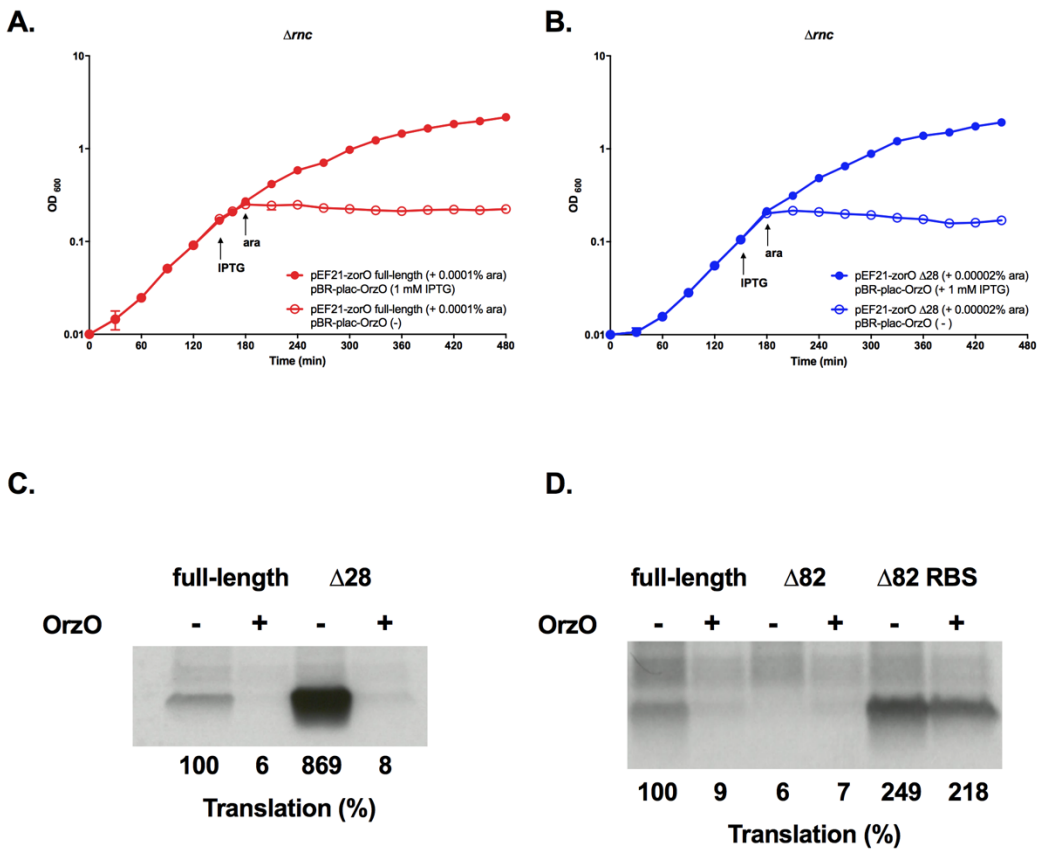


Figure 3.7. OrzO sRNA inhibits translation of *zorO* transcripts. **(A)** *E. coli* strain UTK011 carrying pBR-plac-*orzO* and pEF21-*zorO* was induced as indicated with arabinose (0.0001%) and/or IPTG (1mM). Shown are the mean values ± standard deviations for three independent cultures. **(B)** Rescue experiment was performed as in (A) with *zorO* Δ28 expressed induced by 0.00002% arabinose. Shown are the mean values ± standard deviations for three independent cultures. **(C)** *In vitro* translation assays with [^{35}S]-Met were performed as described in the Material and Methods with 0.1 μM of *zorO* full-length or *zorO* Δ28 mRNAs. 5 μM of OrzO RNA were added as indicated. Quantification of band intensities were determined using ImageJ (38). The intensity of the band representing the *zorO* full-length without the addition of OrzO was set to 100% translational efficiency. **(D)** *In vitro* translation assays with [^{35}S]-Met were performed as described in the Material and Methods with 0.1 μM of *zorO* full-length, *zorO* Δ82, or *zorO* Δ82 RBS mRNAs. 5 μM of OrzO RNA were added as indicated. Quantification of band intensities were determined using ImageJ (38). The intensity of the band representing the *zorO* full-length without the addition of OrzO was set to 100% translational efficiency.

Figure 3.8. The +35 to +50 region in *zorO* full-length and *zorO* Δ28. *In vitro* structure probing was conducted as in Figure 3.3 A. Purple bars indicate the corresponding +35 to +50 region of *zorO* full-length and *zorO* Δ50.

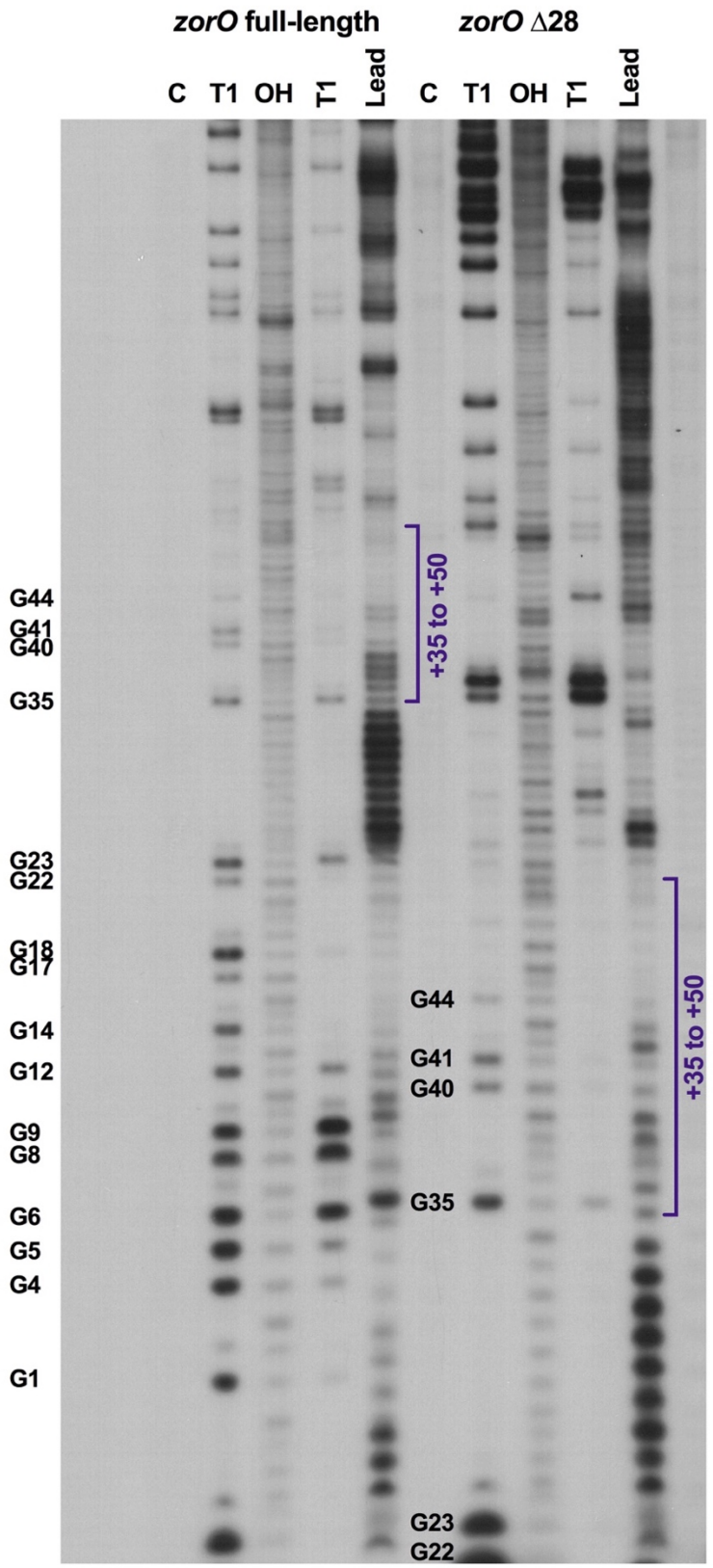


Figure 3.8. (continued)

CHAPTER IV

Characterization of the type I toxin ZorO of *E. coli* O157:H7

Publication note

This chapter is in preparation for submission. The tentative title is:

Jia Wen, Huston T. Heatherly, John R. Harp, and Elizabeth M. Fozo. Characterization of the type I toxin ZorO of *E. coli* O157:H7 [In preparation].

My contribution to this chapter was mutant constructions, toxicity assays, the literature review, and the writing. John R. Harp assisted in generating data presented in Figure 4.2, where he performed the flow cytometry analyses. Huston T. Heatherly assisted in generating data presented in Figure 4.4, where he performed the toxicity assays with pAZ3-*zorO* E16R, pAZ3-*zorO* R23E, and pAZ3-*zorO* E16R/R23E. I performed all other experiments.

Abstract

Most type I toxins are small hydrophobic proteins that possess a putative transmembrane domain. We previously showed that the *zorO-orzO* gene pair in *Escherichia coli* O157:H7 is a true type I toxin-antitoxin locus, with *zorO* encoding the toxic protein and *orzO* encoding the antitoxin sRNA. Overproduction of *zorO* is toxic and can cause cell growth stasis or cell death. This toxic phenotype is solely due to the function of the ZorO protein and not due to the *zorO* mRNA. ZorO overproduction leads to rapid membrane depolarization, suggesting that it targets the inner membrane. Six charged amino acids are found in the ZorO sequence, arranged in an alternating charge pattern. Two residues, glutamic acid at the 16th position (E16) and arginine at the 23rd position (R23), are predicted to reside in the putative transmembrane domain of ZorO. Mutational studies show that replacement of the E16 and/or R23 with amino acids carrying no charge or opposite charge impairs the toxicity of ZorO. However, the reduction in toxicity varies between different mutants, indicative of roles of both the charge and the amino acid position in determining ZorO toxicity.

I. Introduction

A bacterial type I toxin-antitoxin (TA) system is composed of a toxic protein whose overproduction causes cell growth stasis or cell death and an antitoxin small RNA (sRNA) that can inhibit production of the toxin by base pairing with its mRNA. The first described type I system, *hok-sok*, is located on the R1 plasmid of *Escherichia coli* where it functions to maintain the plasmid in a bacterial population through a post-segregational killing mechanism (1). With increasing numbers of bacterial genomes sequenced, homologs of *hok-sok* and novel type I loci have since been identified in bacterial chromosomes.

Most identified type I toxins share common biochemical features. They are small hydrophobic proteins (< 60 amino acids) with a single putative transmembrane domain and are hypothesized to function at the cytoplasmic membrane (2). Indeed, overproduction studies have confirmed that some type I toxins localize in the inner membrane and disrupt membrane integrity, as was verified for a Hok homolog (3). Overexpression of the *hok* gene in *E. coli* caused membrane damage and "ghost cells" were clearly visible (1). Similarly, overproduction of the type I toxin ShoB or IbsC led to rapid membrane depolarization and an immediate change in membrane potential shortly after induction (4). The TisB toxin also localized in the inner membrane of *E. coli* and its overproduction dramatically reduced intracellular ATP concentration (5). *In vitro* studies further revealed that the TisB proteins form narrow pores in a planar lipid bilayer causing membrane depolarization (6). Such membrane damage was suggested to reduce the proton motive force and impair ATP synthesis, thereby leading to a decline in the overall cellular energy.

Despite the detrimental effects caused by type I toxins, it is believed that bacteria employ these toxins to reduce both their metabolic and growth rates as means to increase stress tolerance and promote recovery. It has been reported that when challenged with the DNA-damaging antibiotic ciprofloxacin, the type I toxin TisB induces persister formation of *E. coli* cells by shutting down cellular metabolism (7). Similarly, the chromosomally-encoded *hok* homolog known as *hokB* is transcribed in *E. coli* in response to nutrient starvation, leading to dormancy (8). Aside from their roles in stress response, many type I toxins can provide bacteria fitness advantages in their natural habitats [reviewed in (9,10)]. For instance, deletion of either the type I toxin gene *hok*, *ldr*, or *tisB* was shown to reduce the survival rate of the intracellular bacterial pathogen *Salmonella enterica* serovar Typhimurium inside eukaryotic cells (11).

Our previous studies demonstrated that the *zorO-orzO* gene pair in the chromosome of *E. coli* O157:H7 (EHEC) is a true type I TA locus, with *zorO* encoding the toxin and *orzO* encoding the antitoxin sRNA (12). The ZorO toxin is 29-amino-acid in size and possesses one potential transmembrane domain. This suggests that the ZorO protein may employ a similar strategy as seen with TisB to disrupt the membrane (6). Surprisingly, 20% of the ZorO sequence is made of charged amino acids. Charged amino acids are usually less commonly found in the transmembrane segment of a membrane protein (13), and in many cases they are essential for the function of the membrane-bound protein. For example, in *E. coli*, a single, membrane-embedded, negatively charged amino acid (E26) determines the substrate specificity of the multidrug resistance protein MdfA (14). Further, charged residues have been reported to promote protein-protein interactions, as exemplified by the ionic interactions between the amphiphilic helix domain and the densely charged region of the TatA protein (15). Moreover, the positively charged residue R72 within the transmembrane segment one of the YidC2 protein is essential for proper insertion of YidC dependent membrane proteins in *Bacillus subtilis* (16). In this study, we confirm that overproduction of the ZorO toxin causes membrane damage and identify two charged residues that appear to be critical for the function of ZorO. The putative mode of action of ZorO within the cell membrane is also discussed.

II. Material and methods

Bacteria strains and plasmids All bacterial strains and plasmids used in this study are listed in Table 4.1. The sequences of all oligonucleotides are listed in Table 4.2.

Growth conditions *E. coli* strains were grown at 37°C in Luria–Bertani (LB) medium with shaking. Chloramphenicol was added at a final concentration of 25 μg/ml when necessary. Arabinose was added at a final concentration between 0.0001% and 0.2% as indicated.

Plasmid construction Mutations of individual nucleotides were performed through site-directed mutagenesis as described previously (17).

Overproduction of ZorO toxin ZorO toxicity assays were performed as described previously (12). Briefly, the pAZ3 plasmid harboring the *zorO* wild type or mutant gene under the P_{BAD}

arabinose inducible promoter was transformed into *E. coli* strain UTK007 through electroporation (18). The resulting transformants were grown overnight and then diluted to an OD_{600} of 0.01. When the OD_{600} reached ≈ 0.3, the cultures were split and arabinose was added to half the culture as indicated to overexpress the gene. OD_{600} was recorded every 30 min. Shown are averages ± standard deviations for a minimum of three independent experiments.

Flow cytometry analysis *E. coli* cells carrying the pAZ3-*zorO* wild type plasmid were grown as previously indicated to an OD_{600} ≈ 0.3 and split into two cultures. One culture was induced with 0.2% arabinose and the other served as the control (time 0). At 0, 5, 15, 30, or 60 min post-induction, 50 μl aliquots were taken, flooded with 4 ml of 1X PBS and centrifuged. Cell pellets were resuspended in 1 ml of 1X phosphate-buffered saline (PBS) and stained with 10 μg/ml DiBAC4(3) [bis-(1,3-dibarbituric acid)-trimethine oxanol] (Life Technologies). Cells were incubated at room temperature for 30 min in the dark, washed with 4 ml of 1X PBS extensively twice, and then resuspended in 0.5 ml 1X PBS for cells harvested at 0, 5, or 15 min, or in 1 ml 1X PBS for cells harvested at 30 and 60 min, and analyzed by flow cytometry in a LSR II flow cytometer (Becton Dickinson) with a 488-nm laser. Samples were run at an event rate of 3,000 events per second. DiBAC4(3) fluorescence was collected in the fluorescein isothiocyanate (FITC) channel. Differences in induction versus control were obtained using geometric mean fluorescence for n = 3.

III. Results

ZorO confers toxicity as a protein and not as a small RNA In addition to their protein function, the mRNAs of certain small proteins possess a regulatory role (19). For example, the 43-amino-acid SgrT protein of *E. coli* is produced under glucose-phosphate stress and inhibits glucose transport by modulating the activity of the glucose transporter PtsG. Moreover, the RNA that encodes SgrT, known as the SgrS sRNA, can base pair and destabilize the mRNA of the same glucose transporter (20,21). Hence, the response towards glucose-phosphate stress is attributed to both the effects of the SgrT protein and that of the SgrS sRNA. With this in mind, we first wanted to know whether the *zorO* mRNA plays a role in the toxic phenotype resulting from *zorO* overexpression.

To test this, we constructed a *zorO* AAG mutant under an arabinose inducible P_{BAD} promoter in which the start codon of *zorO* is mutated to AAG to disrupt translation. If the toxic phenotype seen in *zorO* overexpression is dependent upon its function as a protein, we would expect to see a loss of the phenotype. Indeed, as shown in Figure 4.1, the *zorO* AAG mutant failed to confer cell growth stasis under saturating levels of the inducing agent arabinose (0.2% arabinose). This suggests that the toxicity conferred by *zorO* relies entirely on the biochemical function of the ZorO protein.

Overproduction of ZorO causes membrane depolarization Since the ZorO protein is very hydrophobic and is predicted to contain a transmembrane domain, we examined whether it can cause cell membrane damage similar to the overproduction of type I toxins ShoB and IbsC (4). *E. coli* harboring pAZ3-*zorO* wild type was induced with 0.2% arabinose for ZorO overproduction; cells were subsequently incubated with DiBAC4(3) to monitor membrane depolarization (22). As shown in Figure 4.2A and 4.2B, overproduction of ZorO caused significant membrane damage whereas there was no damage in the control cells. More specifically, within 5 min after induction, 57% of the cells overproducing ZorO were positive for membrane depolarization; this increased to over 97% of the population within 15 min of induction. Along with membrane depolarization, we also noted a dramatic loss in viable cells within 15 min post induction (data shown in Chapter 3). Moreover, when reduced amounts of arabinose was used for induction (0.0001%), the kinetics of membrane depolarization were slower (data not shown), but still resulted in significant membrane depolarization over time. These data suggest that the ZorO toxin targets the cell membrane, leading to membrane depolarization and eventual cell death, when expressed at high enough levels.

Specific charged amino acids are important for ZorO toxicity When examining the protein sequence of ZorO, we noted that its predicted transmembrane domain possesses two charged amino acids: glutamic acid (E) at the 16th position and arginine (R) at the 23rd position (Table 4.3). As mentioned above, charged residues in membrane associated proteins are shown to be essential for their functions. To test if either E16 or R23 in ZorO contributed to its function, we first substituted these charged residues with neutral amino acids and evaluated the toxicity of the resultant mutants. As shown in Figure 4.3A, replacement of the negatively charged glutamic acid

with glutamine (Q) at position 16 caused a mild reduction in ZorO toxicity. When induced by 0.0001% arabinose, only the *zorO* wild type, but not the *zorO* E16Q mutant, caused complete cell stasis. However, when induced with 0.0005% arabinose, the *zorO* E16Q exhibited the same cell growth repression as the wild type *zorO*. This suggested that the ZorO E16Q was still toxic, but not as potent as ZorO wild type. On the contrary, substituting the positively charged arginine at the 23rd position of ZorO with a leucine (L) completely abolished the toxicity of ZorO (Figure 4.3B). The ZorO R23L mutant failed to confer cell growth stasis even if it was induced by the maximum amount of arabinose (0.2% final concentration). In addition, the *zorO* E16Q/R23L double mutant showed the same nontoxic phenotype as the *zorO* R23L mutant (Figure 4.3C). Taken together, neutralization of the charge at the 16th or the 23rd position of ZorO impairs the toxicity of ZorO.

We then tested if the specific positive or negative charges are required at the aforementioned positions to maintain full toxicity of ZorO. We constructed additional mutants in which the glutamic acid at the 16th position and/or the arginine at the 23rd position were substituted with amino acids with the opposite charge. The *zorO* E16R, *zorO* R23E, and *zorO* E16R/R23E, were then subjected to toxicity assays. As shown in Figure 4.4, all three mutants showed a moderate reduction of the toxicity of ZorO and could only confer full growth stasis when the concentration of the inducing agent arabinose was 0.005% or above 0.005%. These data suggest that the nature of the charge at these two positions also plays a role in determining ZorO toxicity.

IV. Discussion

In this study, we have begun to address the key features of ZorO that are critical for toxicity. Upon overproduction, ZorO causes rapid membrane depolarization, resulting in cell growth stasis or cell death. Such toxicity is solely dependent on the biochemical function of the ZorO protein as a *zorO* mutant without a functional start codon failed to confer cell stasis. Additionally, we also investigated the role of the two charged amino acids, glutamic acid at the 16th position and arginine at the 23rd position, and showed that disruption of the native charge at these positions impairs ZorO toxicity.

Our data clearly showed that the ZorO toxin targets the cell membrane. However, it is still unclear how exactly the toxin affects the membrane. It has been suggested that the type I toxin TisB disrupts membrane integrity by dimerizing and forming pores (15). Walther *et al* proposed

that the opposite charged amino acids on two monomers of the TisB protein form salt bridges, allowing the TisB monomers to assemble into an antiparallel dimer in a "charge zipper" fashion. Interestingly, we noted that similar to TisB, the ZorO protein contains six charged amino acids and those charged amino acids are arranged in an alternating pattern. Additionally, two charged amino acids, glutamic acid at the 16th position and arginine at the 23rd position, are localized on the same side of the α-helix as predicted by the helical wheel projection program (data not shown). Such positioning may allow for potential electrostatic interactions with a nearby protein containing opposite charges. Therefore, it is possible that ZorO dimerizes in the membrane and forms "charge zippers" as predicted for TisB (15) and that disruption of those amino acid charges would impair the toxicity of ZorO.

In agreement with this, our mutational studies demonstrated that ZorO E16Q, ZorO R23L, ZorO E16Q/R23L, ZorO E16R, and ZorO R23E all have decreased toxicity. These results confirmed the importance of the appropriate charges at these positions for ZorO function. However, we found that the extent of toxicity reduction varied in response to the residue mutated. Having a charge at the 23rd position seems more essential than at the 16th position as R23L ZorO is non-toxic but E16Q still maintained toxicity, albeit reduced compared to the wild type ZorO. At this point, we are unclear why the 23rd position is more important. Additional mutational studies and modeling/simulations would provide more insights into the connection between the location of the charges and the impact on ZorO toxicity.

To our surprise, the ZorO E16R/R23E mutant also exhibited reduced toxicity. Since the charges at the 16th and the 23rd positions are both reversed in this mutant, the electrostatic interactions could still be maintained if ZorO forms antiparallel dimers. However, given that the charged amino acids in ZorO are not evenly spaced, the antiparallel charged zipper formed by two ZorO proteins may not be perfectly aligned. Therefore, reversing the charge may affect the overall positioning of the two ZorO monomers and lead to inefficient dimerization. Topology analysis of ZorO can shed light on the protein-protein interactions of both ZorO WT and/or ZorO mutants.

When we predicted the topology of the ZorO protein using different programs, we obtained contradictory results (Table 4.3). Some small proteins of *E. coli* possess dual topology, which could result in ambiguous topology predictions, as exemplified by the YnhF, YoaK, and YohP proteins (23). Hence, it is possible that the ZorO protein may also exhibit dual topology, which could permit the assembly of antiparallel dimers. Fusions of ZorO with GFP or PhoA would aid

in the elucidation of its topology as GFP only functions in the cytoplasm whereas PhoA functions only in the periplasmic space. Ongoing studies will focus on using structural modeling and biochemical approaches to characterize the topology and *in vivo* activity of ZorO protein.

References

1. Gerdes, K., Rasmussen, P.B. and Molin, S. (1986) Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proc Natl Acad Sci U S A*, **83**, 3116-3120.
2. Fozo, E.M., Hemm, M.R. and Storz, G. (2008) Small toxic proteins and the antisense RNAs that repress them. *Microbiol Mol Biol Rev*, **72**, 579-589.
3. Poulsen, L.K., Refn, A., Molin, S. and Andersson, P. (1991) Topographic analysis of the toxic Gef protein from *Escherichia coli*. *Mol Microbiol*, **5**, 1627-1637.
4. Fozo, E.M., Kawano, M., Fontaine, F., Kaya, Y., Mendieta, K.S., Jones, K.L., Ocampo, A., Rudd, K.E. and Storz, G. (2008) Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. *Mol Microbiol*, **70**, 1076-1093.
5. Unoson, C. and Wagner, E.G. (2008) A small SOS-induced toxin is targeted against the inner membrane in *Escherichia coli*. *Mol Microbiol*, **70**, 258-270.
6. Gurnev, P.A., Ortenberg, R., Dorr, T., Lewis, K. and Bezrukov, S.M. (2012) Persister-promoting bacterial toxin TisB produces anion-selective pores in planar lipid bilayers. *FEBS Lett*, **586**, 2529-2534.
7. Dorr, T., Vulic, M. and Lewis, K. (2010) Ciprofloxacin causes persister formation by inducing the TisB toxin in *Escherichia coli*. *PLoS Biol*, **8**, e1000317.
8. Verstraeten, N., Knapen, W.J., Kint, C.I., Liebens, V., Van den Bergh, B., Dewachter, L., Michiels, J.E., Fu, Q., David, C.C., Fierro, A.C. *et al.* (2015) Obg and membrane depolarization are part of a microbial bet-hedging strategy that leads to antibiotic tolerance. *Mol Cell*, **59**, 9-21.
9. Yamaguchi, Y., Park, J.H. and Inouye, M. (2011) Toxin-antitoxin systems in bacteria and archaea. *Annu Rev Genet*, **45**, 61-79.
10. Brielle, R., Pinel-Marie, M.L. and Felden, B. (2016) Linking bacterial type I toxins with their actions. *Curr Opin Microbiol*, **30**, 114-121.
11. Lobato-Marquez, D., Moreno-Cordoba, I., Figueroa, V., Diaz-Orejas, R. and Garcia-del Portillo, F. (2015) Distinct type I and type II toxin-antitoxin modules control *Salmonella* lifestyle inside eukaryotic cells. *Sci Rep*, **5**, 9374.
12. Wen, J., Won, D. and Fozo, E.M. (2014) The ZorO-OrzO type I toxin-antitoxin locus: repression by the OrzO antitoxin. *Nucleic Acids Res*, **42**, 1930-1946.

13. Ulmschneider, M.B. and Sansom, M.S. (2001) Amino acid distributions in integral membrane protein structures. *Biochim Biophys Acta*, **1512**, 1-14.
14. Edgar, R. and Bibi, E. (1999) A single membrane-embedded negative charge is critical for recognizing positively charged drugs by the *Escherichia coli* multidrug resistance protein MdfA. *EMBO J*, **18**, 822-832.
15. Walther, T.H., Gottselig, C., Grage, S.L., Wolf, M., Vargiu, A.V., Klein, M.J., Vollmer, S., Prock, S., Hartmann, M., Afonin, S. *et al.* (2013) Folding and self-assembly of the TatA translocation pore based on a charge zipper mechanism. *Cell*, **152**, 316-326.
16. Chen, Y., Soman, R., Shanmugam, S.K., Kuhn, A. and Dalbey, R.E. (2014) The role of the strictly conserved positively charged residue differs among the Gram-positive, Gram-negative, and chloroplast YidC homologs. *J Biol Chem*, **289**, 35656-35667.
17. Kunkel, T.A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc Natl Acad Sci U S A*, **82**, 488-492.
18. Guzman, L.M., Belin, D., Carson, M.J. and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J Bacteriol*, **177**, 4121-4130.
19. Ulveling, D., Francastel, C. and Hube, F. (2011) When one is better than two: RNA with dual functions. *Biochimie*, **93**, 633-644.
20. Vanderpool, C.K. and Gottesman, S. (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol Microbiol*, **54**, 1076-1089.
21. Wadler, C.S. and Vanderpool, C.K. (2007) A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proc Natl Acad Sci U S A*, **104**, 20454-20459.
22. Wickens, H.J., Pinney, R.J., Mason, D.J. and Gant, V.A. (2000) Flow cytometric investigation of filamentation, membrane patency, and membrane potential in *Escherichia coli* following ciprofloxacin exposure. *Antimicrob Agents Chemother*, **44**, 682-687.
23. Fontaine, F., Fuchs, R.T. and Storz, G. (2011) Membrane localization of small proteins in *Escherichia coli*. *J Biol Chem*, **286**, 32464-32474.
24. Kawano, M., Aravind, L. and Storz, G. (2007) An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Mol Microbiol*, **64**, 738-754.

Appendix IV: chapter IV tables and figures

Table 4.1. Strains and plasmids used in this study

Name	Relevant genotype or description	Source
UTK007	DJ480 PCP18-*araE*	(12)
Plasmids		
pAZ3	Cm^R; P_{BAD} promoter	(24)
pAZ3-*zorO* wild type	Cm^R	This study
pAZ3-*zorO* AAG	Cm^R	This study
pAZ3-*zorO* E16Q	Cm^R	This study
pAZ3-*zorO* R23L	Cm^R	This study
pAZ3-*zorO* E16Q/R23L	Cm^R	This study
pAZ3-*zorO* E16R	Cm^R	This study
pAZ3-*zorO* R23E	Cm^R	This study
pAZ3-*zorO* E16R/R23E	Cm^R	This study

Table 4.2. Oligonucleotides used in this study

Name	Sequence[a]	Use
EF1184	GCTCATTGCCGTACTG**C**AGTTATTAGTGGCTCTG	pAZ3-*zorO* E16Q PCR forward
EF1185	CAGAGCCACTAATAACT**G**CAGTACGGCAATGAGC	pAZ3-*zorO* E16Q PCR reverse
EF1186	GTTATTAGTGGCTCTGTTAC**T**GTTGATTGATTTGTTGAAG	pAZ3-*zorO* R23L PCR forward
EF1187	CTTCAACAAATCAATCAAC**A**GTAACAGAGCCACTAATAAC	pAZ3-*zorO* R23L PCR reverse
EF1188	CATTGCCGTACTG**C**AGTTATTAGTGGCTCTGTTAC**T**GTTGATTG	pAZ3-*zorO* E16Q/R23L PCR forward
EF1189	CAATCAAC**A**GTAACAGAGCCACTAATAACT**G**CAGTACGGCAATG	pAZ3-*zorO* E16Q/R23L PCR reverse
EF1215	CGTGCTCATTGCCGTACTG**CG**GTTATTAGTGGCTCTGTTACG	pAZ3-*zorO* E16R PCR forward
EF1216	CGTAACAGAGCCACTAATAAC**CG**CAGTACGGCAATGAGCACG	pAZ3-*zorO* E16R PCR reverse
EF1217	GAGTTATTAGTGGCTCTGTTA**GA**GTTGATTGATTTGTTGAAGTAAC	pAZ3-*zorO* R23E PCR forward
EF1218	GTTACTTCAACAAATCAATCAAC**TC**TAACAGAGCCACTAATAACTC	pAZ3-*zorO* R23E PCR reverse
EF1219	CATTGCCGTACTG**CG**GTTATTAGTGGCTCTGTTA**GA**GTTGATTGATTTG	pAZ3-*zorO* E16R/R23E PCR forward
EF1220	CAAATCAATCAAC**TC**TAACAGAGCCACTAATAAC**CG**CAGTACGGCAATG	pAZ3-*zorO* E16R/R23E PCR reverse
EF1325	GCTCCTTAATAAGGAGAGCGGA**A**GGACACGCTGACACAAAAGTTAAC	pAZ3-*zorO* AAG PCR forward
EF1326	GTTAACTTTTGTGTCAGCGTGTCC**T**TCCGCTCTCCTTATTAAGGAGC	pAZ3-*zorO* AAG PCR reverse

[a] 5' - 3', restriction sites underlined, nucleotides altered via site directed mutagenesis in red.

Table 4.3. Topology predictions of the ZorO toxin

Software	Predicted transmembrane domain[a]	Predicted orientation
HMMTOP	LTQKLTVLIAVLELLVALLRLIDLL	N_{out}
Memsat	LTQKLTVLIAVLELLVALLRLI	N_{out}
Octopus	LTQKLTVLIAVLELLVALLRLIDLL	N_{in}
Philius	LTQKLTVLIAVLELLVALLRLIDL	N_{out}
Phobius	QKLTVLIAVLELLVALLRLI	N_{out}
Pro	LTQKLTVLIAVLELLVALLRLI	N_{out}
Prodiv	LTQKLTVLIAVLELLVALLRLI	N_{out}
Scampi	LTVLIAVLELLVALLRLIDL	N_{in}
ScampiMsa	LTVLIAVLELLVALLRLIDL	N_{in}
TMHMM	KLTVLIAVLELLVALLRLI	N_{in}
CCTOP	LTVLIAVLELLVALLRLID	N_{out}

[a] The full-length ZorO toxin sequence is "MDTLTQKLTVLIAVLELLVALLRLIDLLK". Positively charged amino acids are shown in red; negatively charged amino acids are shown in blue.

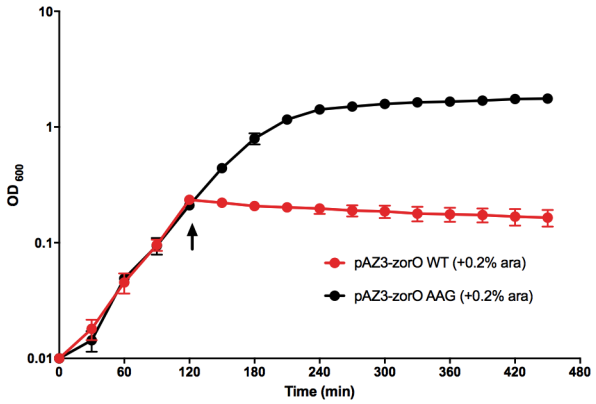


Figure 4.1. Toxicity caused by *zorO* overexpression is due to ZorO protein activity. *E. coli* strain UTK007 harboring pAZ3-*zorO* wild type or pAZ3-*zorO* AAG was grown to mid-log, and arabinose was added to a final concentration of 0.2%. Shown are the averages ± standard deviations for three independent cultures.

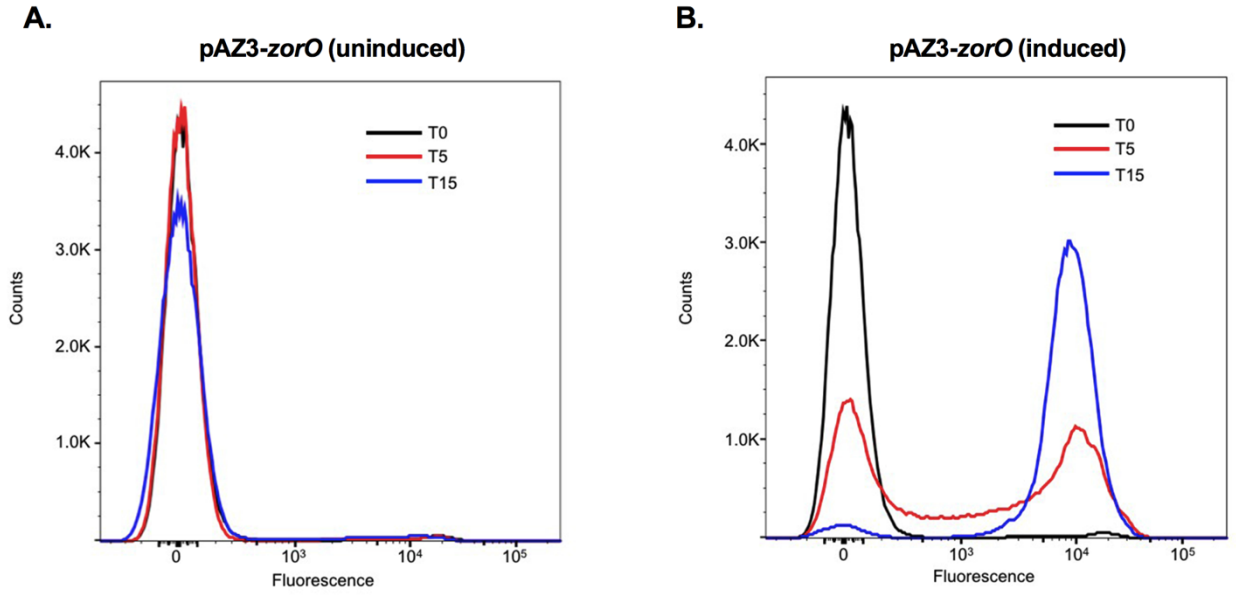


Figure 4.2. Overexpression of *zorO* leads to rapid membrane depolarization. Cells containing the pAZ3-*zorO* wild type plasmid were grown to an approximate $OD_{600} = 0.3$ and split into two cultures. One culture was left uninduced **(A)** and the other was induced with 0.2% arabinose **(B)** for 0 (black line), 5 (red), and 15 (blue) min.

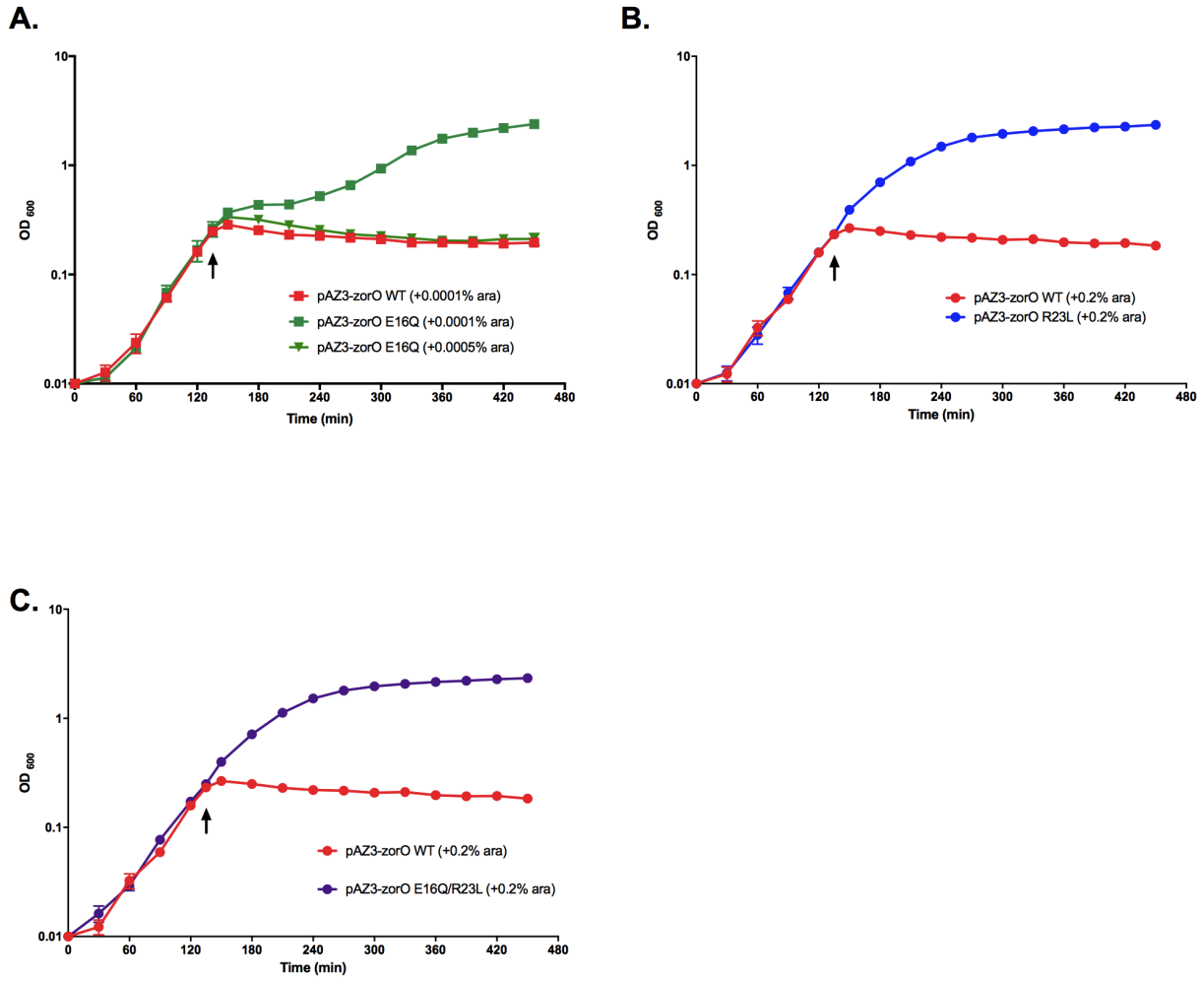


Figure 4.3. Neutralizing the charges at 16th and/or 23rd positions of ZorO impairs the toxicity of ZorO. (A) Toxicity assay was performed as in Figure 4.1 with either *zorO* wild type or *zorO* E16Q induced by 0.0001% or 0.0005% arabinose. Shown are the mean values ± standard deviations for three independent cultures. **(B)** Toxicity assay was performed as in Figure 4.1 with either *zorO* wild type or *zorO* R23L induced by 0.2% arabinose. Shown are the mean values ± standard deviations for three independent cultures. **(C)** Toxicity assay was performed as in Figure 4.1 with either *zorO* wild type or *zorO* E16Q/R23L induced by 0.2% arabinose. Shown are the mean values ± standard deviations for three independent cultures.

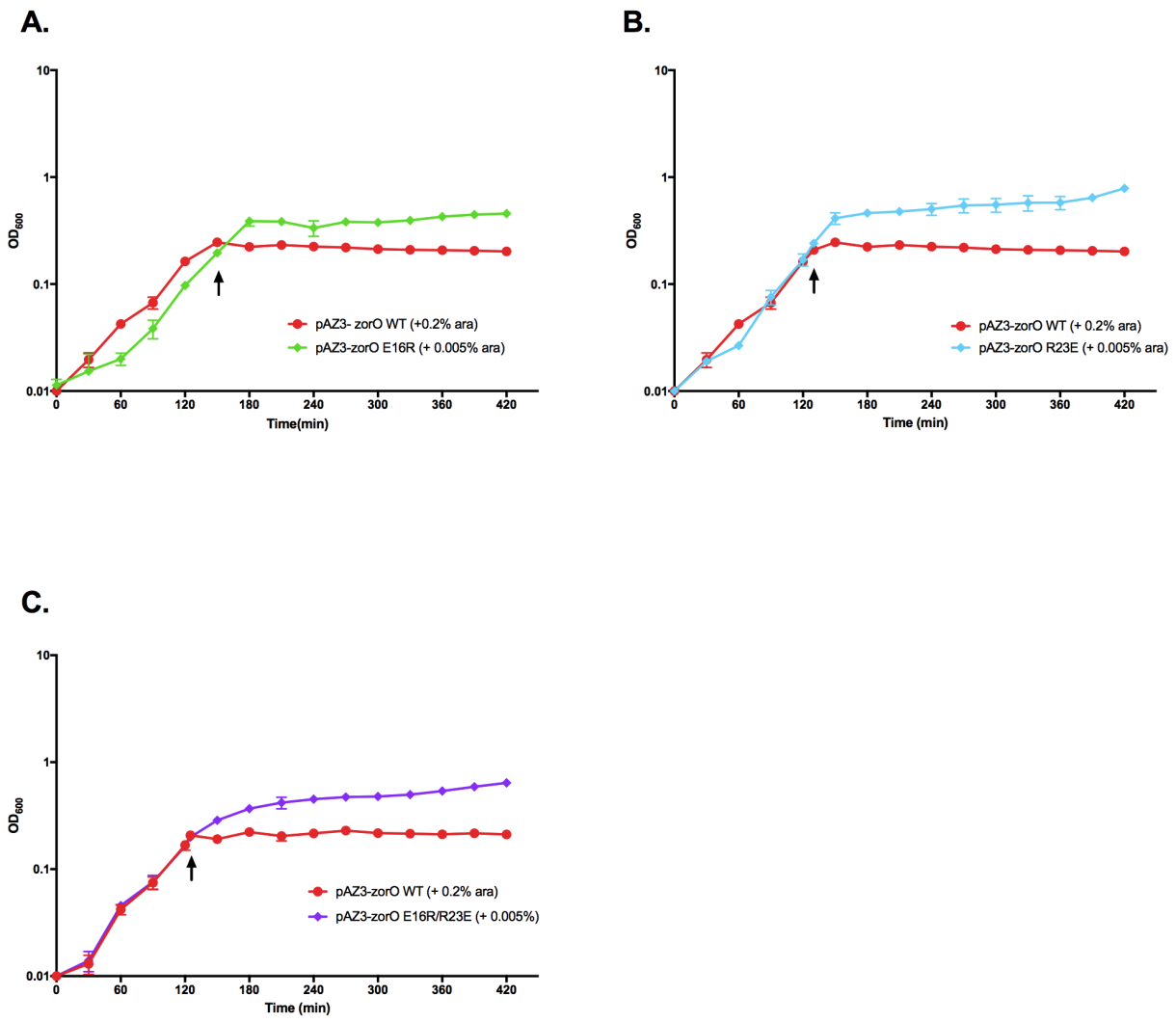


Figure 4.4. Reversing the charges at 16th and/or 23rd positions of ZorO impairs the toxicity of ZorO. (A) Toxicity assay was performed as in Figure 4.1 with either *zorO* wild type or *zorO* E16R induced by 0.2% or 0.005% arabinose. Shown are the mean values ± standard deviations for three independent cultures. **(B)** Toxicity assay was performed as in Figure 4.1 with either *zorO* wild type or *zorO* R23E induced by 0.2% or 0.005% arabinose. Shown are the mean values ± standard deviations for three independent cultures. **(C)** Toxicity assay was performed as in Figure 4.1 with either *zorO* wild type or *zorO* E16R/R23L induced by 0.2% or 0.005% arabinose. Shown are the mean values ± standard deviations for three independent cultures.

CHAPTER V

Conclusions and future directions

The studies presented in this dissertation demonstrate the multiple regulatory mechanisms aimed at controlling the production of the type I toxin ZorO in *E. coli* O157: H7 (EHEC). More specifically, the *zorO* gene is regulated post-transcriptionally via its 5' UTR and its corresponding antitoxin sRNA OrzO.

The 5' UTR of *zorO* folds into an extensive secondary structure that sequesters the RBS limiting *zorO* translation. Processing of the 5' UTR exposes a single-stranded EAP region upstream of the RBS that promotes *zorO* translation, likely due to allowing ribosome preloading on the *zorO* mRNA. Another region spanning +35 to +50 of the 5' UTR also contributes to *zorO* translation and is needed for maximal translation of *zorO*. While we have not yet elucidated how the +35 to +50 region promotes translation, we hypothesize that it may increase *zorO* unfolding or stabilize the unfolded structure. Our work has not only expanded on the standby model [reviewed in (1)] but also provides new insights in 5' UTR mediated regulation: we present the first example of a 5' UTR using two independent regions, a standby site and an upstream sequential region, to modulate translation of a gene. It would be interesting to examine if this additional region is needed for other genes using the standby model for translation regulation.

Aside from 5' UTR mediated regulation, the *zorO* mRNA can be regulated by the sRNA OrzO through base pairing. The OrzO sRNA base pairs to the 3' of the EAP region of *zorO* and inhibits the *zorO* translation, likely by interfering with ribosomal interaction with the standby site. It also stimulates *zorO* degradation in an RNase III-dependent fashion, hence preventing accumulation of *zorO* transcripts. A minimum of 15-nucleotide (nt) of consecutive sequence complementarity or 17-nt intermittent sequence complementarity is required for successful repression of the *zorO* mRNA by the OrzO sRNA. Such extensive base pairing potentials assure formation of the RNA duplex, which is particularly important given that the RNA chaperone protein Hfq is not required for the repression of *zorO* mRNA by the OrzO sRNA.

Why then are so many layers of regulation used to control *zorO* expression? This may be partially attributed to the extensive toxicity of ZorO. Our work demonstrated that the ZorO toxin targets the cell membrane and leads to rapid membrane depolarization. Such toxic effects are at least in part dependent on the two charged amino acids (Glu at the 16th position and Arg at the 23rd position) given that mutations of these residues reduce ZorO toxicity. Additionally, we have conducted metabolomics analyses to determine the potential impacts of *zor-orz* deletion on EHEC metabolism (unpublished data). Deletion of the *zor-orz* locus leads to an overall increase in the

levels of metabolites detected, particularly of those involved in amino acid or nucleotide biosynthesis. Given the detrimental effects of the Zor toxins, a delicate control of its production may be required to assure that ZorO can exert proper function(s) without killing the cells.

The toxic nature of ZorO raises questions about its biological role(s) in EHEC. It is thought that the type I toxins are involved in bacterial stress responses. In the case of the *zorO-orzO* locus, the stress stimulus that induces ZorO production has not yet been identified. Many stress stimuli that the type I toxin-antitoxin (TA) systems respond to were discovered through identification of transcriptional factor binding sites around the promoter region of the toxin gene. For example, both the *tisB* gene (*tisB-istR)* and *dinQ* gene (*dinQ-agrB)* are repressed by the transcriptional repressor LexA and are induced in SOS response (2,3). In line with this, we examined the promoter and its flanking regions of *zorO* and found several putative transcription factor binding sites. One is the cAMP receptor protein (CRP) binding site located 67 nts upstream of the *zorO* transcription start site. Expression of *zorO* was decreased in a Δ*crp* strain, indicative of a role of CRP in the regulation of *zorO* (unpublished data). Yet, no significant growth difference was observed between the EHEC wild type and the Δ*zor-orz* strain when grown with different carbon sources. Perhaps growth measurements are not sensitive enough to detect subtle changes caused by altered ZorO production, especially when taking into consideration the stochastic expression seen in many TA systems (4). Moreover, it is possible that ZorO protein levels are limited by other unidentified factors. To examine these possibilities, we are currently constructing *zorO* transcriptional and translational fusions using GFP as a reporter. This would allow us to assess the changes of both *zorO* mRNA and ZorO protein levels in a more accurate fashion.

An alternative hypothesis is that the type I TA systems possess other functions in addition to stress response. One interesting observation is that many bacteria harbor multiple different type I TA loci in the chromosome and some loci can respond to the same stress. For example, the aforementioned *tisB-istR* and the *dinQ-agrB* loci of *E. coli* are induced upon DNA damage and both toxins exert their function through targeting the inner membrane, and are seemingly functionally redundant (3,5). This raises the possibility that each type I locus may have its own unique function. Indeed, although most type I toxins contain a transmembrane domain and likely localize within the membrane, loss of membrane integrity is not always the primary effect of the toxin. For example, the primary effects of the type I toxin Fst or LdrD are on nucleoid segregation and cell division (6,7); membrane disruption is a secondary effect caused by defects of nucleoid

segregation. Work by Göbl *et al* also suggests that the main function of Fst is to interact with a specific membrane-bound target rather than disrupting the membrane itself (8). However, the target of Fst has not yet been identified. We have performed RNA-seq analysis on wild type EHEC and Δ*zor-orz* strain to assess the global transcriptional effects of *zor-orz* (unpublished data). Comparison of the two transcriptome profiles revealed a series of genes that were differentially expressed between the two strains. Among those genes, we observed that several involved in copper homeostasis were upregulated in the Δ*zor-orz* strain. Interestingly, the most highly upregulated gene *copA* encodes an ATPase that is known to localize at the inner membrane, leading to the possibility that ZorO and CopA may interact. A recombinant ZorO protein (1X-FLAG fused to the C-terminus) will be used to help elucidate not only the localization of ZorO but also potential binding partners, shedding light on the function of the *zorO-orzO* locus in EHEC.

In addition to the *zorO-orzO*, EHEC also harbors another type I gene pair *zorP-orzP* that is highly homologous to *zorO-orzO* (9). Our previous study revealed that there is no cross-talk between the regulation of these two Zor toxins by the antitoxin Orz sRNAs (10). More specifically, each Zor toxin is regulated only by its cognate Orz sRNA. Such independent regulation may imply that these loci participate in distinct cellular processes. One interesting feature of the Zor proteins is that they differ by one amino acid at the 3rd position (threonine in ZorO and serine in ZorP). If the type I toxins have additional targets besides the cell membrane as suggested by studies of Fst (6,8), this different amino acid in the Zor toxins may determine their target specificity. Although much of the work in this dissertation focuses only on the *zorO-orzO* pair, it remains a fascinating area to uncover the function(s) of *zorP-orzP* in EHEC.

Many other *E. coli* strains and *Shigella* species possess *zor-orz* loci, but some only contain one gene pair (9). This is not uncommon for type I systems as many exist in various duplicated copies in bacterial chromosome. For example, *E. coli* K12 strain has five copies of type I loci *ibs-sib* whereas EHEC harbors six (11). Computational analysis suggested that the type I systems evolve through lineage-specific duplication or deletion (9). Gene duplication and the subsequent diversification of the duplicated gene are known to be an important mechanism for acquisition of novel functions in evolution (12). This is well aligned with the notion that different type I loci may have specific yet undiscovered functions. Phylogenetic analysis of bacterial strains containing various copies of the same type I pair may provide useful hints on the function and evolution of a type I system.

Although it has been over three decades since the discovery of the first type I TA pair, we are just scratching the surface of characterizing these genetic loci. In addition to the urgent needs on understanding the functions of these type I systems, questions such as the evolution of the type I loci remain to be answered.

References

1. Unoson, C. and Wagner, E.G. (2007) Dealing with stable structures at ribosome binding sites: bacterial translation and ribosome standby. *RNA Biol*, **4**, 113-117.
2. Vogel, J., Argaman, L., Wagner, E.G. and Altuvia, S. (2004) The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Curr Biol*, **14**, 2271-2276.
3. Weel-Sneve, R., Kristiansen, K.I., Odsbu, I., Dalhus, B., Booth, J., Rognes, T., Skarstad, K. and Bjoras, M. (2013) Single transmembrane peptide DinQ modulates membrane-dependent activities. *PLoS Genet*, **9**, e1003260.
4. Maisonneuve, E., Castro-Camargo, M. and Gerdes, K. (2013) (p)ppGpp controls bacterial persistence by stochastic induction of toxin-antitoxin activity. *Cell*, **154**, 1140-1150.
5. Unoson, C. and Wagner, E.G. (2008) A small SOS-induced toxin is targeted against the inner membrane in *Escherichia coli*. *Mol Microbiol*, **70**, 258-270.
6. Patel, S. and Weaver, K.E. (2006) Addiction toxin Fst has unique effects on chromosome segregation and cell division in *Enterococcus faecalis* and *Bacillus subtilis*. *J Bacteriol*, **188**, 5374-5384.
7. Kawano, M., Oshima, T., Kasai, H. and Mori, H. (2002) 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*, **45**, 333-349.
8. Gobl, C., Kosol, S., Stockner, T., Ruckert, H.M. and Zangger, K. (2010) Solution structure and membrane binding of the toxin Fst of the *par* addiction module. *Biochemistry*, **49**, 6567-6575.
9. Fozo, E.M., Makarova, K.S., Shabalina, S.A., Yutin, N., Koonin, E.V. and Storz, G. (2010) Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families. *Nucleic Acids Res*, **38**, 3743-3759.
10. Wen, J., Won, D. and Fozo, E.M. (2014) The ZorO-OrzO type I toxin-antitoxin locus: repression by the OrzO antitoxin. *Nucleic Acids Res*, **42**, 1930-1946.
11. Fozo, E.M., Kawano, M., Fontaine, F., Kaya, Y., Mendieta, K.S., Jones, K.L., Ocampo, A., Rudd, K.E. and Storz, G. (2008) Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. *Mol Microbiol*, **70**, 1076-1093.

12. Jordan, I.K., Makarova, K.S., Spouge, J.L., Wolf, Y.I. and Koonin, E.V. (2001) Lineage-specific gene expansions in bacterial and archaeal genomes. *Genome Res*, **11**, 555-565.

VITA

Jia Wen was born in Wuhan, China on March, 1986, the daughter of Zhenhuan Wen and Shuxian Zhang. After completing her degree at Wuhan No. 3 High School, Wuhan, China, in 2003, she entered the Wuhan University, receiving her M.D. degree in early 2010 with specialty in Pulmonology. She joined the University of Tennessee in August 2010 to pursue a Ph.D. degree in the Department of Microbiology. Jia dreams to become a light resistant vampire so that she can enjoy sunshine on the beach without worrying about time passing by. If you are actually reading this, it means Jia received her Ph.D. in August 2016.