Reciprocal regulation between the *prgQ* and *prgX* operons of pCF10, a conjugative plasmid of *Enterococcus faecalis*.

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Christopher Mark Johnson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Gary Dunny

July, 2011

Acknowledgements

More people have contributed to this work, directly or indirectly, than I can name, but there are a number who deserve special recognition for making this a fun and enriching experience. Thank you Gary Dunny for your mentorship and advice, for consistently suggesting the most prudent course of action but being patient when I chose to do other things, and for admitting me to your lab when prudence might have suggested against it.

Thank you to the members of the Dunny lab for the insightful discussions, critical questions and great collaborations, you made an interesting project and become a great time: Dawn Manias, Heather Haemig, Thinh Le, Laura Cook, Aaron Barnes, Kristi Frank, Chris Kristich, Katie Ballering, Olivia Chuang-Smith, Suzanne Grindle, Frank Barr, Nick Dillon and Rachel Leibman.

Thank you to the collaborators who have contributed so much to this work: Keith Weaver, Anushree Chatterjee, Sonia Shokeen, Wei-shou Hu and Tina Henkin.

Thank you to my committee members and my other mentors at the University of Minnesota: Matt Mescher, Jeff Gralnick, Yinduo Ji and Pat Cleary.

Thank you to my parents, Mark and Sue Johnson, for your love, support and encouragement during my graduate studies and throughout my life.

Thank you to my in-laws, Bill and Sue Palmer for your sense of humor and for being part of “the rest of my life” when time out of the lab was needed.

Thank you to my wife and children, Kelly, Andrew and Max, for your patience, love, support and for providing the greater context for this research.

Dedication

This dissertation is dedicated to Kelly, Andrew and Max.

Abstract

Regulation of conjugation of pCF10, a pheromone-response plasmid of *Enterococcus faecalis*, is a well characterized process that serves as a model for the control of gene expression in bacteria by intercellular signaling. The genes encoded in the pCF10 *prgQ* operon mediate conjugative transfer of the plasmid in response to a peptide pheromone secreted by plasmid-free *E. faecalis* cells. The products of the *prgX* operon, the small RNA Anti-Q and PrgX, the transcriptional repressor of the *prgQ* promoter, negatively regulate expression of the *prgQ* operon. Transcription of *prgX* is initiated at a promoter within the *prgQ* operon, but oriented in the opposite direction, making transcription of the two operons overlapping and convergent. Each operon encodes a small RNA within its 5' terminus, which is complementary to 5' sequences of the opposing operon. The orientation of the operons and RNA species derived from the 5' terminus of each operon suggested that, in addition to regulation of the *prgQ* operon by Anti-Q, there may be unrecognized regulatory interactions between the two operons. This led me to undertake a focused study of potential transcriptional and posttranscriptional regulatory interactions between the *prgQ* and *prgX* operons. I found that the operons encode mechanisms to negatively regulate each other. Both use a small RNA to reciprocally regulate the expression of downstream genes from the other operon. Notably, each RNA acts via a different mechanism; Qs, derived from the 5' terminus of the *prgQ* operon, directs posttranscriptional processing of the *prgX* mRNA by the host factor RNase III. Anti-Q, an RNA derived from the 5' terminus of the *prgX* operon, negatively regulates transcription elongation of the *prgQ* operon without the assistance of host-encoded proteins. Additionally, I gained understanding of cis-acting mechanisms that regulate expression of the *prgX* operon. Specifically, transcription from the *prgQ* promoter represses the activity of the *prgX* promoter via a mechanism termed transcriptional interference, and Anti-Q sequences act as an intrinsic terminator, attenuating the expression of *prgX* mRNA.

In addition to elucidating some of the complexity of regulation between the *prgQ* and *prgX* operons, this work revealed the high degree of functional efficiency in the RNA sequences of both Anti-Q and Qs. Anti-Q directs its own genesis by acting as an unusual factor-independent termination signal to RNA polymerase and the mature RNA regulates gene expression from the *prgQ* operon. Qs, for its part, fills at least three roles in the regulation of conjugation, acting as a leader sequence that can attenuate downstream gene expression, a mRNA for the *prgQ* gene, and a regulatory RNA that directs endonucleolytic processing of its target. Despite the fact that the overlapping sequences between *prgQ* and *prgQ* are shared by both operons, the sequences necessary for Anti-Q and Qs to mediate each of their activities are distinct. This raises the possibility that these regulatory functions evolved separately and that the combination of two ancestral regulatory pathways gave rise to the pheromone-response circuit of pCF10.

This work has a number of broader implications. It deepens our understanding of how bacteria use RNAs to regulate gene expression. Additionally, the mobile genetic elements of *E. faecalis* contribute to the evolution of multi-drug resistant *E. faecalis* and other pathogens. Understanding the molecular mechanisms used by conjugative plasmids may allow the development of novel interventions to prevent this evolution or target bacteria carrying these plasmids.

Table of Contents

List of Tables

List of Figures

Introduction

Enterococcus faecalis is a Gram positive bacterium that is typically found in the gastrointestinal tracts of animals, including humans, as a member of the commensal flora[1]. In recent decades, *E. faecalis* has also gained recognition as a leading cause of hospital acquired infections. It has several characteristics that allow it to exploit this niche. It is an intrinsically hardy organism, tolerating wide ranges of temperature, pH and osmolarity as well as exposure to high levels of bile salts and some antibiotics[2]. It is able to form robust biofilms on a variety of surfaces, including indwelling medical devices and human tissues[3]. *E. faecalis* is also associated with a wide variety of mobile genetic elements, such as phages, integrative conjugating elements (conjugative transposons) and conjugative as well as non-conjugative plasmids[4-6]. These can transmit genes between organisms, granting them new behaviors and allowing them to adapt to different conditions.

The conjugative mobile genetic elements of *E. faecalis* can transmit virulence factors and antibiotic resistance genes to other *E. faecalis* strains or to other bacterial species, such as *Staphylococcus aureus,* leading to the evolution of multi-drug resistant pathogens that can cause life-threatening infections [4,7]. The pheromone-responsive plasmid pCF10 is one such element. pCF10 was originally identified as a tetracycline-resistance element and has since been shown to also encode aggregation substance, a well-characterized Enterococcal virulence factor[8,9] (and references therein). This plasmid is a member of a family of Enterococcal plasmids that enable their hosts to detect a signal from nearby, plasmid-free *E. faecalis* cells and transfer a copy of the plasmid to these cells via conjugation[10]. Understanding how these plasmids function is an important part of understanding the spread of antibiotic resistance and virulence factors between bacteria. This understanding may allow us to develop novel interventions to prevent horizontal gene transfer between bacteria in clinical settings.

Regulation of the pheromone response of pCF10

Control of the conjugative response has been studied in pCF10 and related pheromone-response plasmids and serves as a model for gene regulation by cell-cell communication in bacteria (see [11] and references therein). Many aspects of this regulation are well understood at the cellular and molecular levels (Figure 1A). The signal that ultimately leads to pCF10 conjugation is a peptide pheromone, cCF10 (LVTLVFV), which is derived from a portion of the N-terminal sequence of CcfA, a chromosomally-encoded protein about which little is known. The pheromone is proteolytically cleaved from the protein during maturation and is secreted by plasmid-free cells[12]. Cells bearing pCF10 express PrgY, which limits secretion of cCF10, preventing autoinduction of conjugation[13]. In this way, plasmid-bearing cells can monitor the environment for the presence of a genetically distinct organism. This is different than quorum sensing pathways, in which a genetically homogenous population of bacteria uses cell-cell communication to measure their own abundance.

To the best of our knowledge, transcription of most pCF10 genes encoding conjugation factors is initiated at the *prgQ* promoter, P_Q (Figure 1B)[14,15]. Previous work has demonstrated that in the absence of pheromone, the plasmid-encoded protein PrgX represses transcription from P_Q. This repression is leaky, however, and generates Qs, an RNA that is the correct length to have its 3' terminus located at an inverted repeat sequence, IRS1[16]. Qs encodes *prgQ*, the peptide product of which is processed to generate iCF10 (AITLIFI), an inhibitory pheromone that interacts with PrgX to maintain repression[17]. Extracellular pheromone is recognized by the plasmid-encoded protein PrgZ, which is homologous to OppA peptide binding proteins of oligopepetide permease (Opp) systems. PrgZ recruits the *E. faecalis* Opp system to import pheromone[18]. Inside the host cell, cCF10 interacts with PrgX, alleviating its repression of P_Q[19].

Transcription of Qs increases; additionally some *prgQ* transcripts are extended several kilobases past IRS1 into genes that encode conjugation factors.

Regulation of the *prgQ* operon

Expression of genes from the *prgQ* operon is negatively regulated by products of the *prgX* operon. These include PrgX, a protein that acts as a repressor of P_Q, and Anti-Q a small RNA that also negatively regulates gene expression from the *prgQ* operon via a mechanism that is independent of PrgX [20,21].

A great deal is known about PrgX regulation of P_Q, resulting from several genetic, in vivo, in vitro and structural studies. PrgX is capable of binding to two distinct regions of DNA. The primary binding site is an inverted repeat sequence approximately 110 bp upstream of the P_Q transcription initiation site (Figure 2). The secondary binding site overlaps the P_Q -35 region and contains half of the inverted repeat, so it is more weakly bound by PrgX[22]. When PrgX binds to iCF10, it forms a homotetramer that occupies both binding sites, preventing RNAP from accessing the promoter. When PrgX binds to cCF10, there is a conformational change in the region of PrgX responsible for multimerization and PrgX forms dimers that occupy the primary, but not the secondary binding site, allowing RNAP to bind P_Q and initiate transcription[19]. The increase in transcription from P_Q is 5-10 fold. However, the increase in expression of downstream genes encoding conjugation factors is at least 100-fold[23]. It was hypothesized that Anti-Q plays a critical role in this augmentation, acting to cause termination of transcription in uninduced cells via a “countertranscript-driven attenuation mechanism” which is described below[21].

Regulation of the *prgX* operon

The *prgX* promoter is located downstream of P_Q, but oriented in the opposite direction, making transcription of the *prgQ* and *prgX* operons convergent for 223 nt (Figure 1B, Figure 2). Regulation of the *prgX* operon is not as well understood

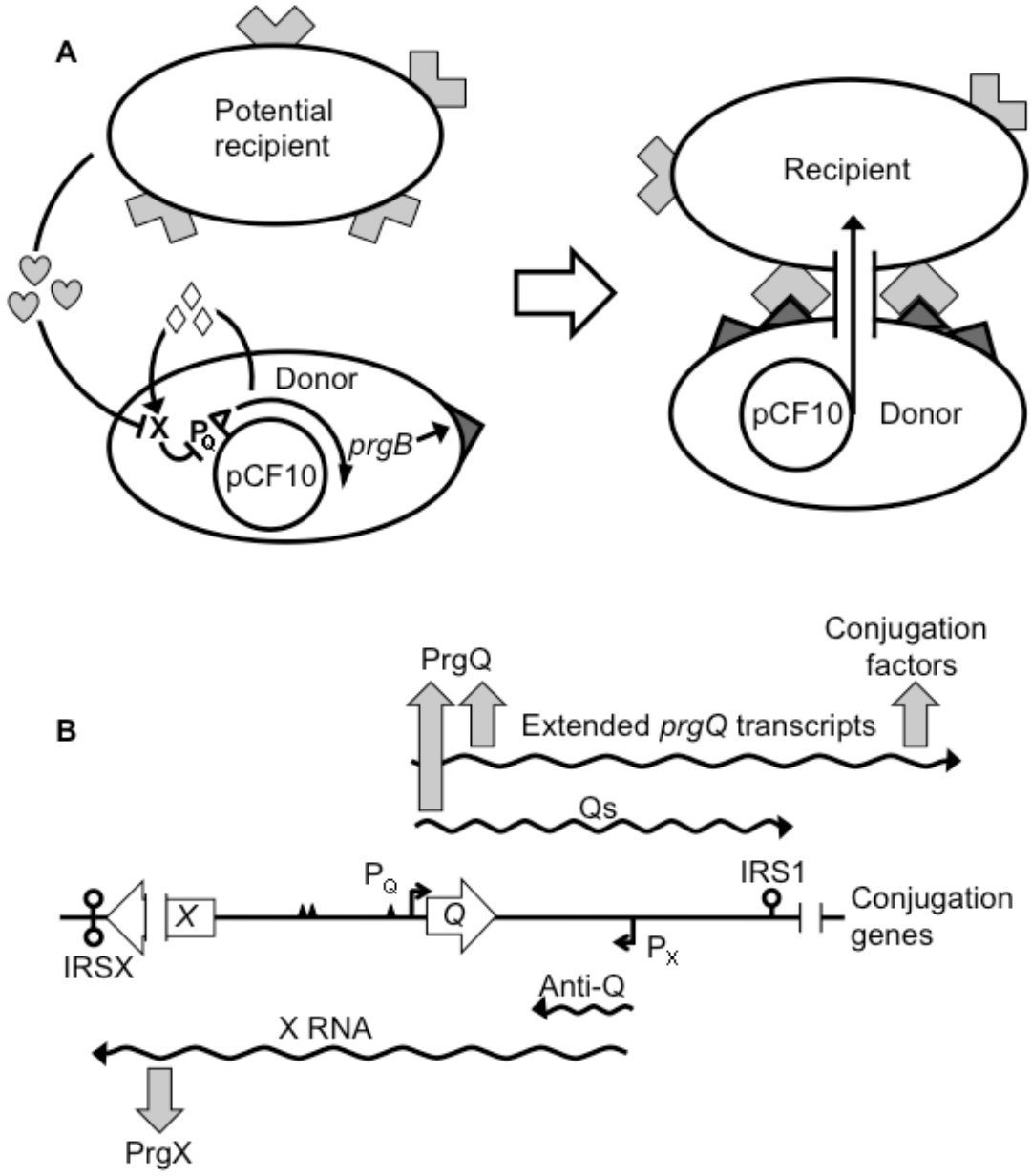

A
Potential recipient
Donor
IX
P_Q
pCF10
prgB
Recipient
pCF10
Donor
B
PrgQ
Conjugation factors
Extended prgQ transcripts
Qs
P_Q
IRS1
X
Q
Conjugation genes
IRSX
P_X
Anti-Q
X RNA
PrgX

Figure 1

Figure 1: Cellular and molecular factors that regulate the pheromone response of pCF10. **A**. Cell-cell signaling induces the pheromone response. Potential recipient and donor cells are shown as ovals. cCF10 (pheromone) is indicated by hearts and iCF10 (inhibitory pheromone) by diamonds, PrgX by an X. pCF10 is indicated by a circle. The *prgQ* promoter (P_Q) is shown as a flag and transcription of downstream genes is shown as a curved arrow. Downstream genes include *prgB*, which encodes Aggregation substance, shown as a triangle, and the type 4 secretion system, indicated by a channel between two cells. Aggregation substance binds to Enterococcal binding substance on the recipient cell, indicated by right angles. **B**. A map of the *prgQ* and *prgX* region, and products of the two operons. pCF10 DNA is represented by a thin line with features. These include the *prgQ* and *prgX* open reading frames, indicated by open arrows labeled *Q* and *X*, respectively. Transcriptional start sites for the *prgQ* (P_Q) and *prgX* (P_X) promoters are shown as bent arrows. Terminators are shown as lollipops. The PrgX primary and secondary binding sites are shown as double and single triangles, respectively. Transcripts of the operons are shown as wavy lines, *prgQ* above and *prgX* below the DNA map. Large arrows indicate translation into the indicated peptide products.

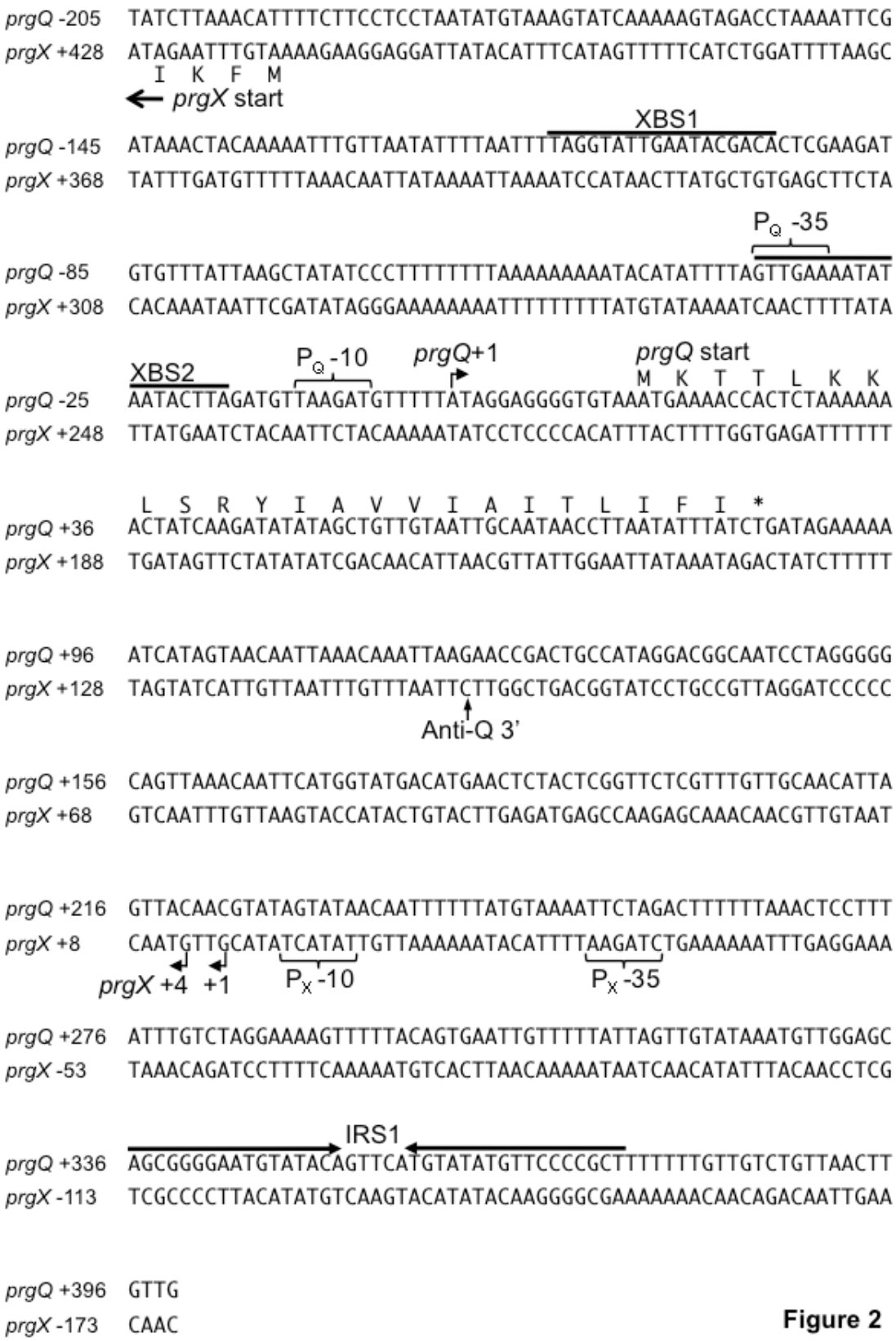

prgQ -205 TATCTTAAACATTTTCTTCCTCCTAATATGTAAAGTATCAAAAAGTAGACCTAAAATTCG
prgX +428 ATAGAATTTGTAAAAGAAGGAGGATTATACATTTCATAGTTTTTCATCTGGATTTTAAGC
I K F M
prgX start
XBS1
prgQ -145 ATAAACTACAAAAATTTGTTAATATTTTAATTTTAGGTATTGAATACGACACTCGAAGAT
prgX +368 TATTTGATGTTTTTAAACAATTATAAAATTAAAATCCATAACTTATGCTGTGAGCTTCTA
PQ -35
prgQ -85 GTGTTTATTAAGCTATATCCCTTTTTTTTAAAAAAAAATACATATTTTAGTTGAAAATAT
prgX +308 CACAAATAATTCGATATAGGGAAAAAAAATTTTTTTTTATGTATAAAATCAACTTTTATA
XBS2
PQ -10
prgQ+1
prgQ start
M K T T L K K
prgQ -25 AATACTTAGATGTTAAGATGTTTTTATAGGAGGGGTGTAAATGAAAACCACTCTAAAAAA
prgX +248 TTATGAATCTACAATTCTACAAAAATATCCTCCCCACATTTACTTTTGGTGAGATTTTTT
L S R Y I A V V I A I T L I F I *
prgQ +36 ACTATCAAGATATATAGCTGTTGTAATTGCAATAACCTTAATATTTATCTGATAGAAAAA
prgX +188 TGATAGTTCTATATATCGACAACATTAACGTTATTGGAATTATAAATAGACTATCTTTTT
prgQ +96 ATCATAGTAACAATTAAACAAATTAAGAACCGACTGCCATAGGACGGCAATCCTAGGGGG
prgX +128 TAGTATCATTGTTAATTTGTTTAATTCTTGGCTGACGGTATCCTGCCGTTAGGATCCCCC
Anti-Q 3'
prgQ +156 CAGTTAAACAATTCATGGTATGACATGAACTCTACTCGGTTCTCGTTTGTTGCAACATTA
prgX +68 GTCAATTTGTTAAGTACCATACTGTACTTGAGATGAGCCAAGAGCAAACAACGTTGTAAT
prgQ +216 GTTACAACGTATAGTATAACAATTTTTTATGTAAAATTCTAGACTTTTTTAAACTCCTTT
prgX +8 CAATGTTGCATATCATATTGTTAAAAAATACATTTTAAGATCTGAAAAAATTTGAGGAAA
prgX +4 +1
PX -10
PX -35
prgQ +276 ATTTGTCTAGGAAAAGTTTTTACAGTGAATTGTTTTTATTAGTTGTATAAATGTTGGAGC
prgX -53 TAAACAGATCCTTTTCAAAAATGTCACTTAACAAAAATAATCAACATATTTACAACCTCG
IRS1
prgQ +336 AGCGGGGAATGTATACAGTTCATGTATATGTTCCCCGCTTTTTTTGTTGTCTGTTAACTT
prgX -113 TCGCCCCTTACATATGTCAAGTACATATACAAGGGGCGAAAAAAACAACAGACAATTGAA
prgQ +396 GTTG
prgX -173 CAAC

Figure 2

Figure 2: Sequence of pCF10 in the P_Q/P_X region. The top strand is the *prgQ* sense strand, the bottom strand is the *prgX* sense strand. The -10 and -35 regions of P_Q and P_X are shown with labeled brackets. The Qs and X RNA transcription start sites are shown with bent arrows. The PrgX binding sites are indicated by black bars above the sequence labeled XBS1 (primary) and XBS2 (secondary). The 3' terminus of Anti-Q is indicated. The IRS1 terminator is shown as two arrows above the sequence that forms the terminator stem. The *prgQ* open reading frame is shown. The beginning of the *prgX* open reading frame is shown. The remainder of the *prgX* open reading frame and IRSX terminator have been omitted.

as regulation of the *prgQ* operon, though it has been shown that PrgX also plays a role in regulation of its own expression and that of Anti-Q. Unlike the *prgQ* operon, PrgX positively regulates products of the *prgX* operon. When the *prgX* operon is cloned into a reporter plasmid, mutations to the *prgX* open reading frame (ORF) that inactivate PrgX prevent expression of the *prgX* message, X RNA, and greatly reduce the amount of Anti-Q that is expressed. This is not caused by PrgX regulation of P_X, as PrgX does not effect expression of a reporter transcriptionally fused immediately downstream of P_X, nor does PrgX bind to DNA in the P_X region[20,22]. Instead, experiments suggested that this activity takes place at the level of posttranscriptional regulation of *prgX* transcripts. Deletion studies of the *prgX* operon demonstrate that sequences between the *prgQ* transcriptional start site (P_X +223) and the *prgX* ORF (P_X +417) negatively regulate expression of Anti-Q and X RNA when provided in cis, suggesting that they may encode a sequence that marks the RNA for degradation. Providing PrgX can overcome this negative regulation, however pheromone induction of PrgX restores the negative regulatory activity[20,22].

Research has uncovered substantial evidence that understanding RNA regulatory elements of the *prgQ* and *prgX* operons is critical to understanding the pheromone response of pCF10. This led us to undertake a focused study of the roles played by different cis and trans acting RNA elements in regulating gene expression of both the *prgQ* and *prgX* operons.

Regulation of RNA in bacteria

Over the past decades there has been a growing interest in different roles played by RNA in bacteria in addition to its place as mRNA, tRNA and in ribonucleoprotein complexes such as ribosomes. It is now recognized that RNA has catalyitic activity, not only in ribosomes, but in some RNases, such as RNase P and in Group I & II introns and other ribozymes[24-26]. RNAs can also regulate protein activity; an example is 6S, which sequesters RNA polymerase by

mimicking DNA[27]. Additionally, RNAs can interact via base-pairing with complementary sequences in target RNAs to influence the target's fate at almost every step in the life of the RNA, including transcription, translation and degradation [28-30]. These RNA-RNA interactions often have the effect of influencing the secondary structure of the target RNA, either by directly generating a duplex between the two RNAs, or by sequestering sequences within the target, preventing them from interacting with other sequences and thereby influencing the structure of distal regions of the RNA. The secondary structure of RNA is critical to its function. Many endoribonucleases specifically cleave single stranded RNA while others target double stranded RNA[31]. This cleavage can lead to a change in the function or degradation of the target. The ribonuclease complement of *E. faecalis* is not completely understood, but is similar to that of *Bacillus subtilis*[32]. Structure contributes to the function of RNA in more subtle ways as well. Loops of RNA at the ends of helical stems and other single-stranded regions often interact with other RNAs. The ability of these so-called "kissing loops" to interact with targets can be enhanced by various hairpin motifs, sequences within the loop that cause the sugar-phosphate backbone to twist such that the loop nucleotides are presented for optimal interaction[33-35].

Transcription of RNA is a tightly regulated process. Initiation of transcription is modulated by the promoter, which recruits a particular σ subunit of RNA polymerase (RNAP), and by transcription factors, which can enhance RNAP binding to a promoter or, as is the case with PrgX, compete with RNAP for access to the promoter[22]. Transcription can also be affected by the activity of other promoters in cis, a phenomenon known as convergent transcription or transcriptional interference[36]. This occurs when two promoters are oriented in a face-to-face configuration, so that RNAP enzymes initiating transcription at these promoters will proceed toward each other. RNAPs transcribing from one promoter actively repress transcription from the other promoter. It is thought that the RNAP enzymes collide, forcing one or both from the template, or that RNAP

transcribing from one promoter prevents RNAP from binding to the other promoter, either by physically blocking access to the second promoter or by dislodging a necessary transcription factor or RNAP that is in the process of initiating transcription[37,38,36]. Once transcription is initiated, different signals can act to terminate transcription upstream of a gene, attenuating its expression. Some of these signals are carried within the RNA itself. These are referred to as intrinsic terminators, factor-independent terminators or rho-independent terminators. Factor independent terminators typically consist of an inverted repeat sequence encoded in the DNA, followed by a tyrosine-rich region. Once transcribed, this is able to form a high G + C stem-loop structure followed by a polyuridine tract. The precise mechanism by which this signals RNAP to terminate transcription is unknown, but current data favors a model in which an actively transcribing RNAP enzyme pauses along the polyuridine tract. While RNAP is paused, the stem-loop structure forms, extending into the RNA exit channel of RNAP. In the active elongation complex, this channel can accommodate approximately 9 nucleotides of single-stranded RNA. The double-stranded helix within the channel forces RNAP to transition to a less stable conformation, which disengages from the DNA template and releases the RNA, terminating transcription[39].

The ability of a factor-independent terminator to function can be regulated by mechanisms that prevent formation of the stem-loop structure while RNAP is paused along the polyuridine tract. This can be accomplished by RNA-binding proteins such as TRAP of *B. subtilis*, small molecules or even other RNAs[28,40,41]. The countertranscript-driven attenuation mechanism, which regulates the copy number of some plasmids in Gram positive bacteria is an example. In such systems, a factor-independent terminator is coded in the 5' UTR of a gene that drives plasmid replication. Formation of the terminator prevents expression of the gene; however, the 5' leader can also form a competing secondary structure, an antiterminator, which precludes formation of the terminator and allows

transcription of the gene. Formation of the antiterminator is prevented by interactions between sequences in the 5' UTR of the nascent RNA and a sRNA transcribed from the opposing DNA strand in the leader region, termed a countertranscript[28,42,43]. This interaction stabilizes the nascent RNA in a conformation that allows terminator formation, preventing transcription of the downstream gene. The countertranscript is complementary to its targets along the entire length of the sRNA but does not need to form a complete duplex with the target to exert its influence. Rather, binding between complementary kissing loops is sufficient to disrupt the antiterminator in the regulated RNA, making the terminator energetically favorable[44,45].

After a gene has been transcribed, the mature mRNA may be posttranscriptionally regulated by a number of mechanisms. The secondary structure of the transcript may be influenced so that the ribosome-binding site (RBS) is single stranded and accessible to ribosomes, or occluded within a double-stranded RNA helix, preventing translation. Ribsowitches are cis-acting RNA elements, encoded in the 5' UTR of genes, that often function this way. In the absence of a small molecule the 5' UTR adopts a secondary structure that permits translation of the downstream message. When the small molecule is present it interacts with structures within the RNA, changing its conformation such that the RBS of the gene is sequestered within an RNA helix, preventing translation[29,46]. An RNA may also be targeted for degradation. There are several well-documented examples of regulatory RNAs interacting with and directing enzymatic cleavage of their targets. RNase III is an endoribonuclease that can mediate this activity by cleaving double stranded RNA. Homologs have been identified in all three domains of life and are well studied in both eukaryotes, where RNase III family members Drosha and Dicer are involved in small RNA maturation, and in bacteria, where RNase III has been found to play several roles, including maturation of tRNA and rRNAs, as well as RNA-directed gene silencing via degradation of regulatory RNA-mRNA pairs. Bacterial RNase III

recognizes and cleaves double stranded RNA that is at least 20 bp in length. Such duplexes can be formed by interactions between complementary regions of two RNAs, or within the stem of a single RNA. Unlike ribonucleases such as RNase E and RNase J1, RNase III is indifferent to triphosphate groups at the 5' terminus of substrate RNAs. This allows RNase III to initiate RNA decay by cleaving its target and generating a 5' terminus with a monophosphate, which can then serve as an efficient docking site for other RNases, leading to degradation of the RNA.

Methods

Bacterial strains, plasmids and reagents. *E. coli* DH5α was used to propagate the plasmids in this study and was grown in LB broth, YENB broth containing 0.075% yeast extract and 0.08% Nutrient Broth[47] or on LB plates with 1.5% agarose. Antibiotics were used at the following concentrations for *E. coli*: chloramphenicol, 20 μg/ml; erythromycin, 200 μg/ml. *E. faecalis* was grown in THB broth, M9-YE + glucose (M9-YEG) broth or on THB plates with 1.5% agarose. Antibiotics were used at the following concentrations for *E. faecalis*: spectinomycin, 1000 μg/ml; chloramphenicol, 20 μg/ml; erythromycin, 10 μg/ml. Strains and plasmids are listed in table 1.

Cloning of OG1RFΔ3097. OG1RFΔ3097 was generated by using a previously described allelic exchange system[48] to replace the EF3097 ORF in OG1RF with the sequence: ATGAAACAAGGTGAAATAAAAAAAAGTATTCCTCAGTAA. The sequence of the mutation was verified by Sanger sequencing.

Plasmid construction. Most of the plasmids used in this study were generated by modifying the plasmids pBK1 and pBK2, constructed by Brianna Kozlowics[49], or pDM4 and pDM5, constructed by Dawn Manias[23,50] (Figure 5). The approaches and reagents used to generate specific plasmids are listed in table 2. Oligonucleotides used in this study are listed in table 3.

The plasmid pCJ9 was constructed as follows (Figure 3). The vector pVE6007 was digested with HpaII and KpnI, the cohesive ends blunted with Klenow fragment and dephosphorylated with calf intestinal alkaline phosphatase (CIP). The *E. faecalis aad9* gene was digested with EcoRI and NcoI out of a pGEM T Easy clone into which it had been inserted, the cohesive ends blunted with Klenow fragment and ligated into the digested pVE6007 to generate pVE6007Spec. The temperature sensitive mutation in the RepR gene was

Table 1: Strains and plasmids used in this study

Strain or plasmid	Characteristics	Source
E. coli DH5a	Used for cloning	Lab stock
E. coli EC1000	Provides RepA in trans to pCJK47	[51]
E. faecalis OG1Sp	Derived from OG1, resistant to Spectinomycin	[48]
E. faecalis 100-5	Derived from OG1Sp, carries *prgX* in chromosome under control of constitutive P_{23} promoter	[52]
E. faecalis OG1RF	Derived from OG1, resistant to Rifampicin and Fusidic acid	Lab Stock
E. faecalis OG1RFD3097	Derived from OG1RF, carries in-frame deletion of *rnc* (EF3097)	This work
E. faecalis OG1SSpD*rnjB*	Derived from OG1SSp, carries in-frame deletion of *rnjB*. Made by Keith Weaver as described [53]	This work
E. faecalis CK111	Derived from OG1Sp, conjugatively delivers pCJK47	[48]
pAT18	pAMb1 based Gram-positive, *E. coli* shuttle vector	[54]
pBK1	-172 through +390 of *prgQ* inserted upstream of a *lacZ* transcriptional fusion in pCl3340, a Gram-positive *E. coli* shuttle vector	[49]
pBK1-1	pBK1 derivative carrying the antiterminator deletion, *prgQ* Δ120 to 253	[50]
pBK1-13	pBK1 derivative with *lacZ* transcriptionally fused to *prgQ* +354	[50]
pBK1-14	pBK1-13 derivative carrying the antiterminator deletion, *prgQ* Δ120 to 253	[50]
pBK1-17	pBK1 derivative carrying a deletion between IRS1 and the *lacZ* junction, *prgQ* Δ385-394	[50]
pBK1-18	pBK1 derivative carrying a deletion of the 3' half of IRS1, *prgQ* Δ354 to 376	[50]
pBK1-19	pBK1-18 derivative carrying the antiterminator deletion, *prgQ* Δ120 to 253	[50]
pBK1-21	pBK1 derivative carrying complementary mutations to 3 nt in each stem half of the terminator as well as the compensatory mutation within the antiterminator: *prgQ* +339/41GGG:CCC, +369/71CCC:GGG, +197/99CUC:GGG	[50]
pBK1-22	pBK1 derivative carrying the substitution *prgQ* T196C, Mutation 13 from [21]	This work
pBK1-23	pBK1 derivative carrying the substitution *prgQ* G138A, Mutation 33 from [21]	This work
pBK1-24	pBK1 derivative carrying a mutation in Qs complementary to the U in the Anti-Q YUNR loop (A215G)	[50]
pBK1-25	pBK1 derivative carrying a point mutation that inactivates P_X, *prgX* T(-5)G	[50]
pBK1-26	pBK1 derivative carrying a point mutation that inactivates P_Q, *prgQ* T(-12)A, A(-11)C	[23]
pBK1-27	pBK1 derivative carrying the substitution *prgQ* G173C	This work
pBK1-28	pBK1 derivative carrying the substitution *prgQ* G173A	This work

pBK1-29	pBK1 derivative carrying the substitution *prgQ* U174A	This work
pBK1-30	pBK1 derivative carrying the substitution *prgQ* A175U	This work
pBK1-31	pBK1 derivative carrying the substitution *prgQ* U176A	This work
pBK1-32	pBK1 derivative carrying the substitution *prgQ* G177C	This work
pBK1-33	pBK1 derivative carrying the substitution *prgQ* A178G	This work
pBK1-35	pBK1 derivative carrying the substitution *prgQ* A144U	This work
pBK1-36	pBK1 derivative carrying the substitution *prgQ* C143G	This work
pBK1-37	pBK1 derivative carrying the substitution *prgQ* G142C	This work
pBK1-38	pBK1 derivative carrying the substitution *prgQ* G141C	This work
pBK1-4	pBK1 derivative carrying complementary mutations to 3 nt in each stem half of the terminator: *prgQ* +339/41GGG:CCC, +369/71CCC:GGG	[50]
pBK1-57	pBK1-25 derivative carring the mutation *prgQ* Δ14 to 114	This work
pBK1-60	pBK1 derivative carrying the substitution *prgQ* U171C	This work
pBK1-61	pBK1 derivative carrying the substitutions *prgQ* G138A and C147U	This work
pBK1-72	pBK1 derivative carrying the substitutions *prgQ* U171C and A180G	This work
pBK1-9	pBK1 derivative carrying the mutation *prgQ* Δ10 to 110	This work
pBK2	pCI3340 with pCF10 fragment from IRSX to IRS1. Carries complete *prgX* operon and *prgQ* sequences through IRS1. *lacZ* transcriptionally fused downstream of IRS1	[49]
pBK2-126	pBK2-26 derivative carrying missense mutations in codons 3 and 4 of *prgX* ORF	This work
pBK2-146	pBK2-126 derivative containing mutations to a potential stem structure	This work
pBK2-24	pBK2 derivative carrying a mutation to the Anti-Q YUNR loop U (*prgQ* A215G)	[55]
pBK2-26	pBK2 derivative arrying a point mutation that inactivates P_Q, *prgQ* T(-12)A, A(-11)C	This work
pBK2-46	pBK2-26 derivative containing mutations to a potential stem structure	This work
pCJ1	transcriptional fusion of lacZ to P_X +221 in pCI3340 backbone	[23]
pCJ1-2	pCJ1 carrying mutation to P_X: *prgX* A(-9)G	This work
pCJ1-25	pCJ1 carrying mutation to P_X: *prgX* T(-5)G	This work
pCJ1-5	pCJ1 carrying mutation to P_X: *prgX* T(-10)A	This work
pCJ1-6	pCJ1 carrying mutations to P_X: *prgX* T(-10)A, A(-9)G, T(-5)G	This work
pCJ11	pCJ5 derivative carrying the mutation *prgX* D233 to 1446, includes IRSX	This work
pCJ12	pCJ5 derivative carrying the mutation *prgX* D215 to 1446, includes IRSX	This work
pCJ13	pCJ5 derivative carrying the mutation *prgX* D205 to 1446, includes IRSX	This work

pCJ14	pCJ13 derivative carrying IRS1 cloned upstream of IRSX	This work
pCJ15	pCJ12 derivative carrying IRS1 cloned upstream of IRSX	This work
pCJ16	pCJ11 derivative carrying IRS1 cloned upstream of IRSX	This work
pCJ18	pCJ6 derivative carrying mutation to *prgX* occluding sequence	This work
pCJ19	pCI3304 based transcriptional fusion of X RNA (~+220) to P_Q (prgQ-176 to +3ish) for IVT mapping	This work
pCJ30	pCJ6 with hinge deleted	This work
pCJ31	pCJ6 with hinge replaced by *prgQ* 112 to 206	This work
pCJ32	pCJ6 with hinge replaced by *prgQ* 30 to 112	This work
pCJ33	pCJ6 with hinge replaced by *prgQ* 30 to 206	This work
pCJ5	pBK1-25 derivative carrying IRSX inserted at XhoI site and a modified IRSX inserted at SalI site	This work
pCJ6	pBK2-126 derivative with P_X Δ6 to 210	This work
pCJ9	*E.coli/E.faecalis* shuttle and expression vector	This work
pCJ9:His-EF3097	Derived from pCJ9, expresses N-terminal His-tagged RNase III	This work
pCJK47	Allelic exchange suicide vector	This work
pDM4	pAT18 derivative containing *prgQ* +1 through +390, produces Anti-Q.	[50]
pDM4-22	pBK1 derivative carrying the substitution *prgX* A28G, Mutation 13 from [21]	This work
pDM4-23	pBK1 derivative carrying the substitution *prgX* C86U, Mutation 33 from [21]	This work
pDM4-27	pDM4 carrying the substitution *prgX* C51G	This work
pDM4-28	pDM4 carrying the substitution *prgX* C51U	This work
pDM4-29	pDM4 carrying the substitution *prgX* A50U	This work
pDM4-3	pDM4 derivative carrying a point mutation within the proposed YUNR loop of Anti-Q (T9C)	[50]
pDM4-30	pDM4 carrying the substitution *prgX* U49A	This work
pDM4-31	pDM4 carrying the substitution *prgX* A48U	This work
pDM4-32	pDM4 carrying the substitution *prgX* C47G	This work
pDM4-33	pDM4 carrying the substitution *prgX* U46C	This work
pDM4-35	pDM4 carrying the substitution *prgX* U80A	This work
pDM4-36	pDM4 carrying the substitution *prgX* G81C	This work
pDM4-37	pDM4 carrying the substitution *prgX* C82G	This work
pDM4-38	pDM4 carrying the substitution *prgX* C83G	This work
pDM4-60	pDM4 carrying the substitution *prgX* A53G	This work

pDM5-25	Contains pCF10 fragment expressing Qs with P_X inactivated	This work
pDM5-26	Derived from pDM5, carrys mutation inactivating P_Q: *prgQ* T(-12)A, A(-11)C	This work
pDM5-40	Derived from pDM5-25, with *prgQ* Δ10 to 110	This work
pDM5-48	Derived from pDM5-25, with mutation to a potential stem structure	This work
pDM5-49	Derived from pDM5-48, with mutation to correct potential stem structure	This work
pDM5-70	Derived from pDM5-25, with *prgQ* Δ4 to 36	This work
pDM5-73	Δ120 to the end of the pCF10 sequence inserted into the backbone	This work
pDM5-76	Derived from pDM5-25, with *prgQ* Δ34 to 335	This work
pDM5-77	Derived from pDM5-25, with *prgQ* Δ4 to 36 and 120 to 335	This work
pDM5-78	Derived from pDM5-25, with *prgQ* Δ4 to 80	This work
pDM5-79	Derived from pDM5-25, with *prgQ* Δ4 to 36 and 80 to 120	This work
TB13	Carries pCF10 sequences with *prgQ* T196C, mutation 13 from referenced paper	[21]
TB33	Carries pCF10 sequences with *prgQ* G138A, mutation 33 from referenced paper	[21]

Table 2: Cloning of plasmids used in this study

Plasmid	Technique	Template	Backbone	Oligonucleotides
pBK1-1	Digestion/ligation[A]	pBK1	pBK1	pBR322_3265, pCF10_8274R_BsaI
pBK1-13	Digestion/ligation[A]	pBK1	pBK1	pCF10_8007_XhoI, pCF10_8510R_BamHI
pBK1-14	Digestion/ligation[A]	pBK1-1	pBK1	pCF10_8510R_BamHI, pCF10_8007_XhoI
pBK1-17	Digestion/ligation[A]	pBK1	pBK1	pCF10_8007_XhoI, pBK1NewLacZ_Jn
pBK1-18	Digestion/ligation[A]	pBK1	pBK1	pCF10_8007_XhoI, pBK1D354-376
pBK1-19	Digestion/ligation[A]	pBK1-1	pBK1	pCF10_8007_XhoI, pBK1D354-376
pBK1-21	Site-directed mutagenesis[D]	pBK1-4	pBK1-4	SuperTermCompF, SuperTermCompR
pBK1-22	Digestion/ligation[A]	TB13	pBK1	pCF10_8007_XhoI, pCF10_8544R_BamHI
pBK1-23	Digestion/ligation[A]	TB33	pBK1	pCF10_8007_XhoI, pCF10_8544R_BamHI
pBK1-25	Mutagenic primer[B]	pBK1	pBK1	pBR322_3265, Qa-10Mut4
pBK1-26	Digestion/ligation[A]	pBK1	pBK1	Qs-10Mut1_DnF, LacZ5R_Seq, Qs-10Mut1_UpR, pBR322_3265
pBK1-27	Mutagenic primer[B]	pBK1	pBK1	pBK1-27_F, pCF10_8544R_BamHI
pBK1-28	Mutagenic primer[B]	pBK1	pBK1	pBK1-28_F, pCF10_8544R_BamHI
pBK1-29	Mutagenic primer[B]	pBK1	pBK1	pBK1-29_F, pCF10_8544R_BamHI
pBK1-30	Mutagenic primer[B]	pBK1	pBK1	pBK1-30_F, pCF10_8544R_BamHI
pBK1-31	Mutagenic primer[B]	pBK1	pBK1	pBK1-31_F, pCF10_8544R_BamHI
pBK1-32	Mutagenic primer[B]	pBK1	pBK1	pBK1-32_F, pCF10_8544R_BamHI
pBK1-33	Mutagenic primer[B]	pBK1	pBK1	pBK1-33_F, pCF10_8544R_BamHI
pBK1-35	Mutagenic primer[B]	pBK1	pBK1	pBK1-35_R, pCF10_8007_XhoI
pBK1-36	Mutagenic primer[B]	pBK1	pBK1	pBK1-36_R, pCF10_8007_XhoI
pBK1-37	Mutagenic primer[B]	pBK1	pBK1	pBK1-37_R, pCF10_8007_XhoI
pBK1-38	Mutagenic primer[B]	pBK1	pBK1	pBK1-38_R, pCF10_8007_XhoI
pBK1-4	Site-directed mutagenesis[D]	pBK1	pBK1	prgQ_antiTerMut1F, prgQ_antiTerMut1R
pBK1-57	Digestion/ligation[A]	pBK1, pDM5-25	pBK1	pCF10_8007_XhoI, pCF10_8161_R_BsaI, pCF10_8291, Qa-10_Mut4
pBK1-60	Mutagenic primer[B]	pBK1	pBK1	Mut60F, pCF10_8544R_BamHI
pBK1-61	Site-directed mutagenesis[D]	pBK1	pBK1	pCF10_Qs1_QR, Mut61_SitDir
pBK1-72	Mutagenic primer[B]	pBK1	pBK1	Mut72F, pCF10_8544R_BamHI

pBK1-9	Digestion/ligation[A]	pBK1	pBK1	pCF10_8007_XhoI, BsaIRevd+10_+110, BsaIFd+10_+110, LacZ5'_Seq
pBK2-126	Mutagenic primer[B]	pBK2-26		pBK2_100_R, pBR322_3265
pBK2-146	Overlap extension PCR[C]	pBK2-126	pBK1-126	pCF10_8213, pCF10_8392R
pBK2-26	Digestion/ligation[A]	pBK1-26	pBK2	pCF10_8007_XhoI, pCF10_8544R_BamHI
pBK2-46	Mutagenic primer[B]	pBK2-146	pBK2-26	pCF10_8007_XhoI, pBK1_Qs_QR
pCJ1	Digestion/ligation[A]	pBK1	pBK1	pCF10_8176_BglII, LacZ5R_Seq
pCJ1-2	Mutagenic primer[B]	pCJ1	pCJ1	LacZ5R_Seq, Qa-10Mut2
pCJ1-25	Mutagenic primer[B]	pCJ1	pCJ1	LacZ5R_Seq, Qa-10Mut4
pCJ1-5	Mutagenic primer[B]	pCJ1	pCJ1	LacZ5R_Seq, Qa-10Mut5
pCJ1-6	Mutagenic primer[B]	pCJ1	pCJ1	LacZ5R_Seq, Qa-10Mut6
pCJ11	Digestion/ligation[A]	pBK1	pCJ5	pCJ11F, pBK1_Qs_QR
pCJ12	Digestion/ligation[A]	pBK1	pCJ5	pCJ12F, pBK1_Qs_QR
pCJ13	Digestion/ligation[A]	pBK1	pCJ5	pCJ13F, pBK1_Qs_QR
pCJ14	Digestion/ligation[A]	pBK1	pCJ5	pCJ14F, pBK1_Qs_QR
pCJ15	Digestion/ligation[A]	pBK1	pCJ14	pCJ12F, pBK1_Qs_QR
pCJ16	Digestion/ligation[A]	pBK1	pCJ14	pCJ11F, pBK1_Qs_QR
pCJ18	Mutagenic primer[B]	pCJ6	pBK2-26	pCF10_8007_XhoI, pCJ18R
pCJ19	Digestion/ligation[A]	pBK1	pBK1	pCJ19_UpR, pCF10_8007_XhoI, pCJ19_DnF, pCF10_8176_BglII
pCJ30	Digestion/ligation[A]	pBK2, pCJ6	pBK1	pCJ6_JnR_BsaI, pBK1_Qs_QR, pBR322_3265, pCF10_7996R_BsaI
pCJ31	Digestion/ligation[A]	pBK2, pCJ6	pBK1	pCJ6_JnR_BsaI, pBK1_Qs_QR, pCF10_8382R_BsaI, pCF10_8288_BsaI, pBR322_3265, pCF10_7996R_BsaI
pCJ32	Digestion/ligation[A]	pBK2, pCJ6	pBK1	pCJ6_JnR_BsaI, pBK1_Qs_QR, pCF10_8287R_BsaI, pCF10_8206_BsaI, pBR322_3265, pCF10_7996R_BsaI
pCJ33	Digestion/ligation[A]	pBK2, pCJ6	pBK1	pCJ6_JnR_BsaI, pBK1_Qs_QR, pCF10_8382R_BsaI, pCF10_8206_BsaI, pBR322_3265, pCF10_7996R_BsaI
pCJ6	Mutagenic primer[B]	pBK1-26	pBK2-26	pCF10_8007_XhoI, prgX/Px_fusion
pCJ9	see text			
pCJ9:His-EF3097	Digestion/ligation[A]	OG1RF	pCJ9	RNaseIII_His_NcoI_F, RNaseIII_SphI_R
pDM4-22	Digestion/ligation[A]	TB13	pDM4	pCF10_8544R_HpaI, pCF10_8177_XhoI
pDM4-23	Digestion/ligation[A]	TB33	pDM4	pCF10_8544R_HpaI, pCF10_8177_XhoI
pDM4-27	Mutagenic primer[B]	pDM4	pDM4	pBK1-27_F, pDM4_4795R

pDM4-28	Mutagenic primer[B]	pDM4	pDM4	pBK1-28_F, pDM4_4795R
pDM4-29	Mutagenic primer[B]	pDM4	pDM4	pBK1-29_F, pDM4_4795R
pDM4-3	Digestion/ligation[A]	pBK2-24	pDM4	pCF10_8177_XhoI, pCF10_8510R
pDM4-30	Mutagenic primer[B]	pDM4	pDM4	pBK1-30_F, pDM4_4795R
pDM4-31	Mutagenic primer[B]	pDM4	pDM4	pBK1-31_F, pDM4_4795R
pDM4-32	Mutagenic primer[B]	pDM4	pDM4	pBK1-32_F, pDM4_4795R
pDM4-33	Mutagenic primer[B]	pDM4	pDM4	pBK1-33_F, pDM4_4795R
pDM4-35	Mutagenic primer[B]	pDM4	pDM4	pBK1-35_R, pCF10_8177_XhoI
pDM4-36	Mutagenic primer[B]	pDM4	pDM4	pBK1-36_R, pCF10_8177_XhoI
pDM4-37	Mutagenic primer[B]	pDM4	pDM4	pBK1-37_R, pCF10_8177_XhoI
pDM4-38	Mutagenic primer[B]	pDM4	pDM4	pBK1-38_R, pCF10_8177_XhoI
pDM4-60	Mutagenic primer[B]	pDM5-25	pDM4	Mut60F, pDM4_4795R
pDM5-25	Mutagenic primer[B]	pDM5	pDM5-25	Manias-48/XhoI, Qa-10Mut4
pDM5-26	Digestion/ligation[A]	pBK1-26	pDM5	Manias-48/XhoI, pCF10_8510R
pDM5-40	Digestion/ligation[A]	pBK1-9	pDM5-25	Manias-48/XhoI, Qa-10Mut4
pDM5-48	Overlap extension PCR[C]	pDM5-25	pDM5-25	Extended_XhoI-48, StemA_SitDirMut, pCF10_8213, pCF10_8392R
pDM5-49	Overlap extension PCR[C]	pDM5-25	pDM5-25	Extended_XhoI-48, StemA_SitDirMut, pCF10_8213, pCF10_8392R
pDM5-70	Overlap extension PCR[C]	pDM5-25	pDM5-25	pCF10_8072, pCF10_8180R, pCF10_8213, pBK1_Qs_QR
pDM5-73	Digestion/ligation[A]	pDM5-25	pDM5-25	Extended_XhoI-48, pCF10_8274R_BsaI
pDM5-76	Overlap extension PCR[C]	pDM5-25	pDM5-25	pCF10_8201R, pDM4_8276, Mut76F, pDM4_4795R
pDM5-77	Overlap extension PCR[C]	pDM5-70	pDM5-25	pCF10_Qs4_QR, pDM4_8276, Mut77F, pDM4_4795R
pDM5-78	Overlap extension PCR[C]	pDM5-25	pDM5-25	pCF10_8180R, pDM4_8276, Mut78F, pDM4_4795R
pDM5-79	Overlap extension PCR[C]	pDM5-70	pDM5-25	Mut79R, pDM4_8276, Qa_RNA_R, pDM4_4795R

A. PCR fragments are digested and ligated to form a mutant fragment, which is ligated into the vector

B. Primers specifying a mutation are used to generate a mutant fragment that is ligated into the vector

C. Overlap extension PCR is used to generate mutant fragments that are ligated into the vector

D. Mutation-encoding primers are used to amplify the whole plasmid as per Invitrogen & Qiagen systems

Table 3: oligonucleotides used in this study

Name	Sequence
3'RACE_oligo	(PO_3)TCGTATGCCGTCTTCTGTTG
3'RACE_RT	CAACAGAAGACGGCATACGA
3535NisR	ATCAGATCTAGTCTTATAACTATAC
BsaIFd+10_+110	CTAGATGGTCTCGGAGGGAAACAAATTAAGAACCGACTGC
BsaIRevd+10_+110	CTAGATGGTCTCCCCTCCTATAAAAACATCTTAACATCTAAGT
E.fa_5S_Probe	AACAGGTGTATCCTTCTCGCTAT
EF3097_DnF	ATTGGTCTCCTTTTATTTACACCTTGTTTCATTATACT
EF3097_DnR	TCAGCATGCAATACTAAATTACTTGATGCGATTG
EF3097_DnSeq	GCACAGGTTCAGATCCTTC
EF3097_UpF	TTTTCTAGACCCTGAATCCATAAATAAATCTTG
EF3097_UpR	GCGGGTCTCGAAAAGTATTCCTCAGTAAGG
EF3097_UpSeq	CGACGGTCTTCAGGTTTAC
Extended_XhoI-48	GTGCCTCGAGATACATATTTTAGTTGAAAATATAATACTTAGATGTTAAGATGTTTTTA
gyrB_F	CAAGCCAAAACAGGTCGCC
gyrB_R	ACCAACACCGTGCAAGCC
LacZ5R_Seq	GCTGGCGAAAGGGGGATG
Manias-48/XhoI	GTGCCTCGAGATACATATTTTAGTTGAAAATATAATAC
Mut60F	ATCCTAGGGGGCAGTTAAACAATTCACGGTATGAC
Mut61_SitDir	TTAAACAAATTAAGAACCGACTGCCATAGAACGGCAATTCTAGG
Mut72F	ATCCTAGGGGGCAGTTAAACAATTCACGGTATGACGTGAAC
Mut76F	GTTTTTATAGGAGGGGTGTAAATGAAAACCACTCTAAAAAGCGGGGAATGTATACAGTTCATG
Mut77F	TCTGATAGAAAAAATCATAGTAACAATTAAACAAATTAGCGGGGAATGTATACAGTTCATG
Mut78F	TATAATACTTAGATGTTAAGATGTTTTTATAGTATCTGATAGAAAAAATCATAGTA
Mut79R	CTAGGATTGCCGTCCTATGGCAGTCGGTTCTTAATATTAAGGTTATTGCAATTACAAC
Old_Anti-Q_RNA_T7	TAATACGACTCACTATAGTTGTAACTAATGTTGCAACAAA
p043lacZ_391R	GGGATAGGTTACGTTGGTGTAG
P1	GGTATTGCGGTACCCTTGTAC
pBK1_lacZ_QF	CGCTTTAATGATGATTTCAGCC

pBK1_lacZ_QR	TGCCATAAAGAAACTGTTACCC
pBK1_lacZ_QR_SP6	CGGATTTAGGTGACACTATAGAATGCCATAAAGAAACTGTTACCC
pBK1_Qs_QR	TCCTAGACAAATAAAGGAGTTTAAAAAAG
pBK1-27_F	ATCCTAGGGGGCAGTTAAACAATTCATGCTATGACATG
pBK1-28_F	ATCCTAGGGGGCAGTTAAACAATTCATGATATGACATG
pBK1-29_F	ATCCTAGGGGGCAGTTAAACAATTCATGGAATGACATG
pBK1-30_F	ATCCTAGGGGGCAGTTAAACAATTCATGGTTTGACATG
pBK1-31_F	ATCCTAGGGGGCAGTTAAACAATTCATGGTAAGACATG
pBK1-32_F	ATCCTAGGGGGCAGTTAAACAATTCATGGTATCACATG
pBK1-33_F	ATCCTAGGGGGCAGTTAAACAATTCATGGTATGGCATG
pBK1-35_R	CCCCTAGGATAGCCGTCCTATGG
pBK1-36_R	CCCCTAGGATTCCCGTCCTATGG
pBK1-37_R	CCCCTAGGATTGGCGTCCTATGG
pBK1-38_R	CCCCTAGGATTGCGGTCCTATGG
pBK1D354-376	TTTGGATCCCCCAACAGACAACAAAACTGTATACATTCCCCGCTGC
pBK1NewLacZ_Jn	TTTGGATCCAACAAAAAAAGCGGGGAACATATAC
pBK2-100R	ATACTCGAGATATTAGGAGGAAGAAAATGTTTTAGTAAGGTTC
pBR322_3265	GTATATATGAGTAAACTTGGTCTGACAG
pCF10_7002_T7	TAATACGACTCACTATAGGGAGAAATAAAAGAGCAGTCATGAAAATA
pCF10_7149R	TATTCACAGATCTTTAGTTGAGG
pCF10_7845	ACTCTTCTACAGAAATCGGTC
pCF10_7922	TCAATCTGATGGTAATTTAACTCTTGCCTTATTTG
pCF10_7983R	ATGTTTAAGATAGGTTCTGTCC
pCF10_7996R_BsaI	NNNGGTCTCGTAATTAGGAGGAAGAAAATGTTTAAGATAG
pCF10_8007_XhoI	AGCTCGAGTATCAAAAAGTAGACCTAAAATTCG
pCF10_8072	TGAATACGACACTCGAAGATG
pCF10_8137	TTTAGTTGAAAATATAATACTTAGATGTTAAGATG
pCF10_8161R_BsaI	NNNGGTCTCTTTTTACACCCCTCCTATAAAAACATC
pCF10_8176_BglII	NNAGATCTATAGGAGGGGTGTAAATGAAAACC
pCF10_8177_T7	TAATACGACTCACTATAGGGAGATAGGAGGGGTGTAAATGAAAAC

pCF10_8177_XhoI	TTCTCGAGATAGGAGGGGTGTAAATGAAAAC
pCF10_8177F	ATAGGAGGGGTGTAAATGAAAAC
pCF10_8180R	CTATAAAAACATCTTAACATCTAAGT
pCF10_8201R	TGGTTTTCATTTACACCCCTCC
pCF10_8206_BsaI	NNNGGTCTCATAATAAAAAAACTATCAAGATATATAGCTGT
pCF10_8213	CTATCAAGATATATAGCTGTTGTAATTGC
pCF10_8274R_BsaI	AAGGTCTCTCTAGAATTTGTTTAATTGTTACTATG
pCF10_8287R_BsaI	NNNGGTCTCCATTATAATTGTTACTATGATTTTTTCTATCAG
pCF10_8288_BsaI	NNNGGTCTCATAATAACAAATTAAGAACCGACTGCC
pCF10_8291_BsaI	NNNGGTCTCAAAAATTAAGAACCGACTGCCATAG
pCF10_8298F_SP6	CGGATTTAGGTGACACTATAGAACCGACTGCCATAGGACG
pCF10_8382R_BsaI	NNNGGTCTCCATTAAACAAACGAGAACCGAGTAG
pCF10_8392R	CTAATGTTGCAACAAACGAGAACCGAGTAGAG
pCF10_8395R	GTAACTAATGTTGCAACAAACG
pCF10_8437R	AAGTCTAGAATTTTACATAAAAAATTGTTATACTATACG
pCF10_8510R	GTATACATTCCCCGCTGC
pCF10_8510R_BamHI	NNNGGATCCTGTATACATTCCCCGCTGC
pCF10_8544R_BamHI	AAGTTAACAGACAACAAAAAAAGCGGGG
pCF10_8544R_HpaI	TTGGATCCCCCAACAGACAACAAAAAAAGCGGGG
pCF10_8559R_T7	TAATACGACTCACTATAGGGAGACAACAAAAAAAGCGGGGAAC
pCF10_Qs1_QR	TTCTTAATTTGTTTAATTGTTAC
pCF10_Qs4_QR	TTAATTGTTACTATGATTTTTTCTATCAGATA
pCJ11F	NNNCTCGAGAAAAAATCATAGTAACAATTAAACAAA
pCJ12F	NNNCTCGAGAATTAAACAAATTAAGAACCGACTG
pCJ13F	NNNCTCGAGATTAAGAACCGACTGCCATAGG
pCJ14F	NNNGGTCTCTTCGAACAAAAAAAGCGGGGAACATATACATGAACTGTATACATTCCCCGCTCGAGATTAAGA ACCGACTGCCATAGG
pCJ18R	NNNTCTAGAATTTTACATAAAAAATTGTTATACTATACGTTGTAACACCCGAGATATAAAAACATCTTAA
pCJ19_DnF	NNNGGTCTCATAGTAACTAATGTTGCAACAAACGAG
pCJ19_UpR	NNNGGTCTCAACTATAAAAACATCTTAACATCTAAGTATTA
pCJ6_JnR_BsaI	NNNGGTCTCCATTATAGGAGGGGTGTTACAACG

pCJ9_P23_R	NNNCCATGGTGAGTGCCTCCTTATAATTTGTCGACGGCCGAATTCTAAGACTATTTTATC
pDM4_4795R	TCAGGCGCCATTCGCCATTC
pDM4_8276	GCGGATAACAATTTCACACAGG
pMSP3535_7856-78	GAAGAGCGCCCAATACGCAAAC
prgQ_antiTerMut1F	GAATGTATACAGTTCATGTATATGTTCGGGGCTTTTTTTGTTGTC
prgQ_antiTerMut1R	GAACATATACATGAACTGTATACATTCGGGGCTGCTCCAACT
prgX/Px_fusion	NNNTCTAGAATTTTACATAAAAAATTGTTATACTATACGTTGTAACACCCCTCCTATAAAAACATC
pUC_F	CCCAGTCACGACGTTGTAAAACG
pUC_R	GGAAACAGCTATGACCATGATTAC
pWV01_RepR_WTF	AAATGGAAAGCACTATAAAAAACCACACTATC
pWV01_RepR_WTR	TTTTATAGTGCTTTCCATTTCGTATAACATCACTACTATTCCAT
Qa_RNA_R	GAACCGACTGCCATAGGACG
Qa-10Mut2	AAGTCTAGAATTTTACATAAAAAATTGTTGTACTATACG
Qa-10Mut4	AAGTCTAGAATTTTACATAAAAAATTGTTATACGATACG
Qa-10Mut5	AAGTCTAGAATTTTACATAAAAAATTGTAATACTATACG
Qa-10Mut6	AAGTCTAGAATTTTACATAAAAAATTGTAGTACGATACG
Qs_RNA_R	AAAAAAAGCGGGGAACATATACATG
Qs-10Mut1_DnF	AGATGTACAGATGTTTTTATAGGAGGGGTG
Qs-10Mut1_UpR	ATCTGTACATCTAAGTATTATATTTTCAACTAAAA
RNA_Oligo	rCUAGUACUCCGGUAUUGCGGUACCCUUGUACGCCUGUUUUAUA
RNaseIII_His_NcoI_F	NNNCCATGGGTCACCATCACCACCATCACGGATCCAAACAAGGTGTAAATAAATGTAAAGG
RNaseIII_SphI_R	NNNGCATGCTTACTGAGGAATACTTTTCAGTG
StemA_SitDirMut	ATATCTTGATAGTTTTTTTCCTCCCCATTTCATTTACACCCCTCCTA
SuperTermCompF	CATGGTATGACATGAACTCTACTCGGTT
SuperTermCompR	AACCGAGTAGAGTTCATGTCATACCATG
XhoI_IRS1	CTGTATACATTCCCCGCTCGAGA

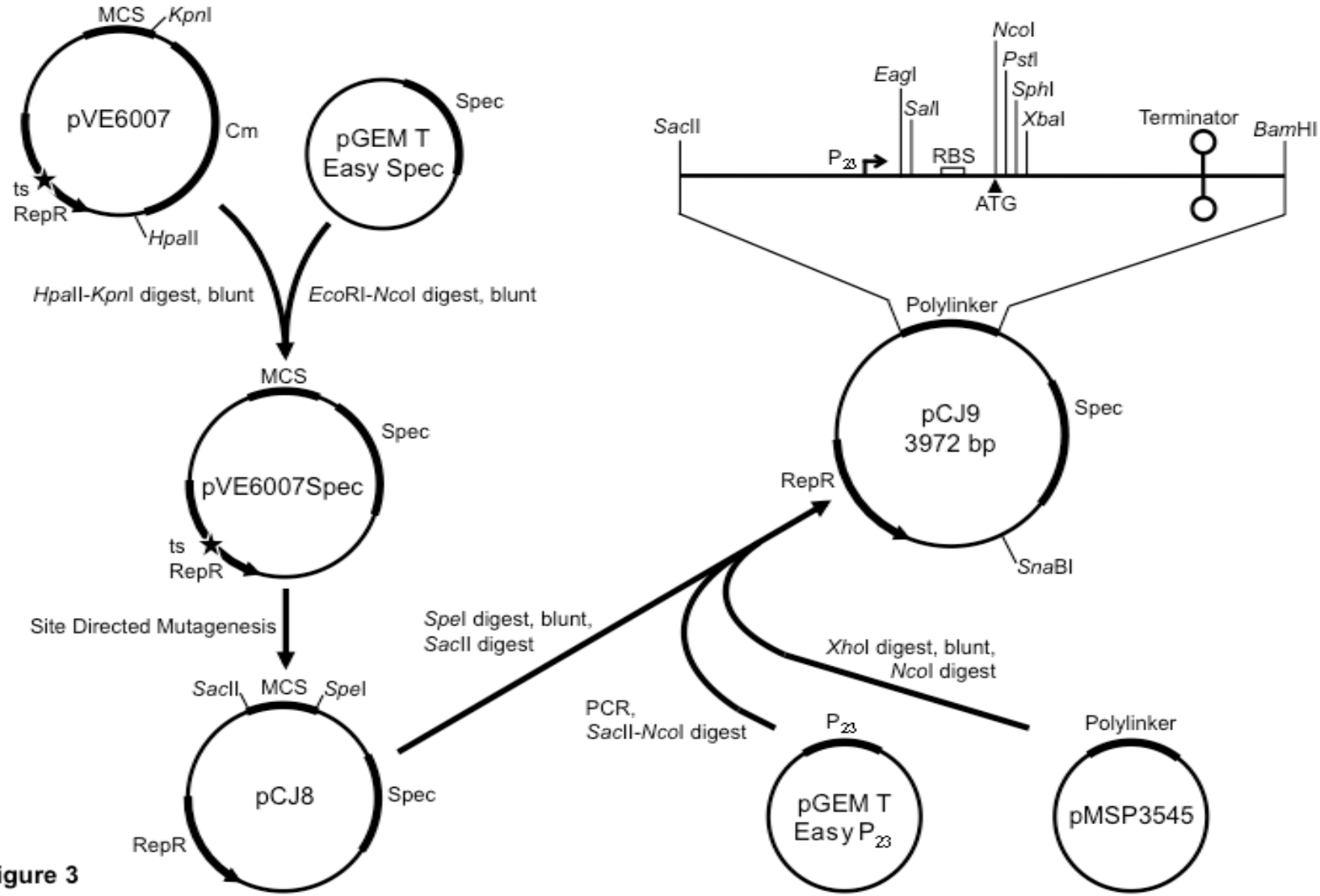

MCS
KpnI
pVE6007
Cm
ts
RepR
HpaII
Spec
pGEM T
Easy Spec
HpaII-KpnI digest, blunt
EcoRI-NcoI digest, blunt
MCS
Spec
pVE6007Spec
ts
RepR
Site Directed Mutagenesis
SacII
MCS
SpeI
pCJ8
Spec
RepR
SpeI digest, blunt,
SacII digest
NcoI
PstI
EagI
SphI
SalI
XbaI
SacII
Terminator
BamHI
P23
RBS
ATG
Polylinker
pCJ9
3972 bp
Spec
RepR
SnaBI
XhoI digest, blunt,
NcoI digest
PCR,
SacII-NcoI digest
P23
pGEM T
Easy P23
Polylinker
pMSP3545

Figure 3

Figure 3: Construction of the expression vector pCJ9 as described in the text. Plasmids are shown as circles. Regions of interest are thick lines. Restriction sites used in construction are labeled. The star indicates the mutation in RepR that causes temperature sensitivity[56]. The polylinker region is magnified. A bent arrow indicates the P_{23} promoter; a box indicates the ribosome binding site; an arrowhead indicates the translation start codon; unique restriction sites within the multi-cloning site are labeled; a lollipop indicates the terminator containing stop codons in all three forward reading frames. Transcriptional fusions can be made using the restriction sites *Eag*I or *Sal*I; translational fusions can be made using the restriction sites *Nco*I, *Pst*I, *Sph*I or *Xba*I.

corrected using site-directed mutagenesis with primers pWV01_RepR_WTF and pWV01_RepR_WTR to generate pCJ8. The multi-cloning site was then replaced with the P_{23} promoter[57] ligated to the polylinker from pMSP3545 [58]. The P_{23} promoter was PCR amplified from a pGEM-T Easy insert using primers pUC_F and pCJ9_P23_R, then digested with SacII and NcoI. The polylinker was PCR amplified from pMSP3545 with primers pMSP3535_7856-78 and 3535NisR, then digested with XhoI, blunted with Klenow fragment and digested with NcoI. The two fragments were ligated into pCJ8 that had been digested with SpeI, the cohesive ends blunted with Klenow fragment, then digested with SacII. The net result of this cloning strategy is pCJ9, an expression vector with the wide host range pWV01 replicon, bearing the Spec gene, which confers spectinomycin resistance. The polylinker contains the P_{23} promoter, followed by a RBS and initiating codon with a terminator downstream. Insertions into the polylinker can be used to generate either transcriptional or translational fusions. Each module, the replication machinery, the antibiotic resistance cassette and the polylinker is flanked by unique restriction sites, making replacement of these modules straight forward.

Plasmid DNA was isolated from *E. coli* using a Qiagen miniprep kit following the manufacturer's instructions. For isolation of plasmid DNA from *E. faecalis*, 3 ml cultures were grown overnight in M9-YEG, then diluted 1:2 in M9-YEG + 3% glycine and incubated for 90 min. Cells were pelleted and resuspended in 200 μl of lysis buffer (10 mM Tris-HCl, pH 8, 50 mM NaCl, 10 mM EDTA and 30 mg/ml lysozyme) and incubated 15 min at 37°C. DNA was then prepared using a Qiagen miniprep kit. Plasmids were electroporated into competent *E. coli* prepared using YENB[47] or *E. faecalis* prepared using lysozyme[20].

Quantitative reverse-transcription PCR (qRT-PCR). *E. faecalis* strains were grown overnight in M9-YE broth at 37°C with selective antibiotics. In the morning, cultures were diluted 1:5 in fresh medium and grown for 90 minutes at

37°C. Then, 600 μl of each culture was added to 1200 μl of Bacteria RNA protect reagent (Qiagen), prepared as per reagent instructions, snap-frozen in ethanol and dry ice and stored at -80°C. RNA was prepared using an RNeasy kit (Qiagen) following a modified enzymatic lysis procedure[59]. RNA (2-5 μg) was treated with Turbo DNase (Ambion) as per the manufacturers' instructions. RNA (20 ng) was then reverse transcribed using a Superscript III first strand synthesis kit (Invitrogen) and gene-specific primers (GSPs) following the manufacturer's instructions. Each reaction contained GSPs for *lacZ* and *gyrB*. Approximately 1 ng cDNA was then used for quantitative PCR using IQ™ SYBR Green Supermix (Bio-Rad) on an IQ5 real-time PCR system (Bio-Rad). The amplification efficiency of each primer set was assessed by amplifying 10-fold serial dilutions of a DNA control consisting of linearized pBK1 (*lacZ*) or *E. faecalis* genomic DNA (*gyrB*). Transcript levels were then determined relative to pBK1 using the $\Delta\Delta C_T$ method with *gyrB* serving as the reference gene.

RNA stability assay and Northern blots. *E. faecalis* strains were cultured as described for the β-galactosidase assays. After the 90 min subculture, rifampicin (350 μg/ml) was added to 5 ml of each culture. At T= 0, 5, 10, 20, 40, 80 min, 600 μl aliquots were removed from each culture and added to 1200 μl Bacteria RNA protect reagent, prepared as per reagent instructions, snap-frozen in ethanol and dry ice and stored at -80°C. Total RNA was then prepared with an RNeasy kit following a modified enzymatic lysis procedure as above. RNA (1 μg) from each sample was electrophoresed on a 1.2% denaturing agarose gel in MOPS buffer as previously described (Bae 2000 JMB). RNA was transferred to a positively charged nylon membrane (Roche) by passive elution with 5X SSC + 10 mM NaOH for 2 h (http://www.ambion.com/techlib/tb/tb_169.html). Membranes were cross-linked using a Stratalinker (Stratagene).

RNA probes were generated as follows. Primers were designed to amplify the region of interest and append an SP6 or T7 promoter to one end of the PCR product such that transcription produced a probe complementary to the region of interest. The PCR product was then used as a template for SP6 RNA polymerase (RNAP) to generate a digoxygenin-UTP (DIG-UTP; Roche) labeled RNA probe as per the manufacturer's instructions. The primers pCF10_7002_T7 and pCF10_7983R were used to generate the probe specific for the *prgX* ORF. Primers pCF10_8298F_SP6 and pCF10_8395R were used to generate the probe specific for the 5' terminus of X RNA. Primers pCF10_8599R_T7 and pCF10_8177F were used to generate the probe specific for Qs. Primers pBK1_lacZ_QF and pBK1_lacZ_QR_SP6 were used to generate the probe specific for *lacZ*.

Membranes were probed with the DIG-labeled RNA probes as per the manufacturers instructions and probe hybridization was detected using DIG-specific Fab fragments and CDP-star reagent (Roche). Membranes were then probed with a DIG end-labeled DNA probe specific for *E. faecalis* 5S rRNA[60] and detected as above. Blots were scanned using a Scanjet 4200c flatbed scanner (Hewlett-Packard) and analyzed using ImageJ software (NIH). RNA half-life was determined using regression analysis.

In vitro transcription (IVT) assays. A region of *prgQ* from –176 to +512 relative to the transcription start-site was PCR amplified using pBK1-25 as template. The resulting DNA fragment was used as a linear template for IVT assays. This template contains P_Q and IRS1 with a point mutation in the P_X -10 region that inactivates P_X. The template (2.5nM) was mixed with *E. faecalis* RNAP (24, 27, 34) in a 35 μl reaction containing 20 mM Tris-HCl, pH 7.9, 20 mM NaCl, 20 mM $MgCl_2$, 0.1 mM EDTA, 11 mM β-mercaptoethanol, 286 μm each ATP, GTP, CTP, 28.6 μM UTP, 2.5 μg of the initiating dinucleotide: ApU for P_Q and GpU for P_X (Ribomed) and 6 μCi α^{32}P-UTP (MP biomedicals). Exogenous RNA was added

at 150 nM where indicated. Reactions were incubated at 37°C for 15 min and stopped by adding an equal volume of IAA:phenol:chloroform, pH 5.2. The aqueous fractions were recovered, resolved on a denaturing polyacrylamide gel and scanned on a phosphorimager (Molecular Dynamics). Gels were analyzed using Imagequant software.

Preparation of RNA for in vitro assays. Exogenous RNAs were prepared using a T7 MEGAshortscript kit (Ambion) following the kit instructions. The transcripts were resolved on a denaturing polyacrylamide gel, the bands of interest were excised and the RNA was eluted as per the kit instructions. The RNA was then precipitated using ethanol and resuspended in DMPC-treated H_2O.

The template for Anti-Q was made by PCR amplification of the Anti-Q region from pBK1 using primers that positioned a T7 RNAP promoter in place of P_X. Human 18S rRNA was also produced (using the template included with the kit) for use as a negative control in the antitermination assays. Qs was prepared using primers pCF10_8177_T7 and Qs_RNA_R. The X RNA 5' fragment was prepared using primers Old_Anti-Q_RNA_T7 and pCF10_7845.

β-Galactosidase assays. *E. faecalis* cultures were grown overnight in M9-YEG broth at 37°C with selective antibiotics. In the morning, cultures were diluted 1:5 in fresh medium and grown for 90 minutes at 37°C. A modified Miller assay was then performed as follows. 1 ml of each culture was centrifuged, the supernatant discarded and the pellet resuspended in Z buffer[61]. The suspended cells (0.1 or 0.2 ml) were transferred to duplicate tubes, Z buffer was added for a total volume of 0.2 ml and 40 μl toluene was added to each tube. The tubes were placed in a 28°C heat block and allowed to equilibrate for 10 min. Z buffer (0.2 ml) containing 4 mg/ml ONPG was added to each tube. The tubes were incubated for 6-8 min and the reaction was stopped by adding 0.5 ml 1 M Na_2CO_3 and

placing the tubes on ice. The debris in each tube was pelleted and the supernatant was removed for analysis as described by Miller[61].

5' 3' RACE. The 5' and 3' termini of single X RNA transcripts were determined following a method similar to one already described[62]. Cellular RNA, 6μg, was treated with the Turbo DNA free kit (Applied Biosystems) as per the manufacturers instructions. RNA was then precipitated by adding 0.1 volumes 3M Sodium acetate pH5.2 and 2.5 volumes Ethanol. chilling at -20℃ for 2h and centrifuging for 30' at 13000 rpm in a 4°C microcent rifuge. The supernatant was aspirated and the pellet washed with 70% ethanol and centrifuged for 15' at 13000 rpm in a 4°microcentrifuge. The supernatant was aspirated, the pellet dried and resuspended in 42μl water. Samples were split into two equal aliquots, and treated with 5 units Tobacco acid pyrophosphatase (TAP, Epicenter) or mock-treated with water. Reactions were performed in 1X TAP buffer (Epicenter) supplemented with 20 U RNasin Plus (Promega) in a volume of 50μl and incubated for 1 hour at 37°C. Reactions were then bro ught to 100μl with water and RNA was extracted with an equal volume of phenol:chloroform. The aqueous phase was recovered and the RNA precipitated with Sodium acetate and ethanol as described above. The pellet was resuspended in a final volume of 300μl 1x T4 RNA ligase buffer (New England Biolabs) supplemented with 10% Dimethyl sulfoxide, 20 Units RNasin Plus (Promega) and 40 units T4 RNA ligase (New England Biolabs) and incubated for 2 hours at 37°C to self-ligate the RNAs. RNA was extracted with phenol:chloroform, pH 4.5, precipitated with Sodium acetate and Ethanol as described above and resuspended in 8μl H_2O. This was used as template for a reverse transcription reaction using Superscript III and primed with 3.2 pmol oligonucleotide pCF10_8072 as per the manufacturers instructions. The cDNA was then PCR amplified using pCF10_8072 and pCF10_7149R. The reactions were ligated into pGEM-T Easy, which was used to transform *E. coli*

DH5α. Vector insertions were PCR amplified from individual colonies using pUC_F and pUC_R and the *prgX* 5'-3' junctions were sequenced.

5' RACE. 5' termini of processed X RNA transcripts were determined as previously described[14] except no TAP-treated samples were prepared. The oligonucleotide ligated to 5' termini was RNA_oligo. The RT reaction was primed with pCF10_7922 and the PCR was primed with pCF10_8072 and P1. PCR products were ligated into pGEM-T Easy (Promega), which was used to transform electrocompetent *E. coli* DH5α. Transformants were spread on plates with selective antibiotics. Individual colonies were picked and the sequence inserted into the vector was PCR amplified using primers pUC_F and pUC_R, and sequenced.

3' RACE. To map the 3' terminus of Anti-Q, cellular RNA, 1μg, was treated with the Turbo DNA free kit (Applied Biosystems) as per the manufacturers instructions. RNA was precipitated with 0.3M Sodium acetate and ethanol as described above and resuspended in 26μl H_2O. The 5' termini of the RNA were dephosphorylated by adding 1 unit CIP (New England biolabs) and 3μl NEB buffer 3 and incubating for 60 minutes at 37°C. This prevents them from serving as a substrate for T4 RNA ligase. The RNA was extracted with phenol:chloroform, precipitated with Sodium acetate and ethanol and resuspended in 35μl H_2O. The 5' phosphorylated DNA oligo (3'RACE_oligo), 100 pmol, was then ligated to the 3' termini of the RNA by adding 5μl T4 RNA ligase buffer, 60 units T4 RNA ligase (New England Biolabs), 5μl dimethylsulfoxide and 20 Units RNasin Plus and incubating for 2 hours at 37°C . RNA was extracted with phenol:chloroform and precipitated with Sodium acetate and ethanol, then resuspended in 8μl H_2O. This was then used as a template for reverse transcription using Superscript III reverse transcriptase and primed with 2 pmol of the oligonucleotide 3'RACE_RT, which is complementary to the ligated

oligonucleotide, as per the manufacturers instructions. The cDNA, 1μl, was then used as template for a PCR reaction primed with 3'RACE_RT and pCF10_8392R. The products were electrophoresed on an agarose gel, bands of the correct size were excised, gel purified and sequenced.

Purification of RNase III. OG1RFΔ3097 cells containing pCJ9::His-EF3097 were cultured at 37°C overnight in 100ml of M9-YEG broth with selective antibiotics. This was diluted in 1L of fresh medium and grown to an OD600 of 1.2 – 1.4. Cells were pelleted by centrifugation and the pellet was stored at -80°C. The pellet was thawed on ice and resuspended in 10 ml of 10 mM $NaHPO_4$ pH 7.4, 50 mM NaCl containing 30 mg/ml lysozyme and 250 U/ml Mutanolysin (Sigma). The suspension was incubated at 37°C for 20 minutes, split into two fractions and centrifuged at 5000g for 10 minutes at 4°C. The supernatant was discarded and the pellets resuspended in 7.5 ml Y-Per mixed 1:1 with 2x WB (1x WB is 50 mM $NaHPO_4$ pH 7.4, 300 mM NaCl, 10 mM Imidazole) with one tablet Complete Mini, EDTA-free protease inhibitor (Roche). The cells were lysed by sonication. The lysate was centrifuged at 21500g for 20 min at 4°C and the supernatant was applied to a gravity-flow column containing a 1.5 ml bed volume of HisPur cobalt resin (Pierce) prepared as per the manufacturer's instructions. The flowthrough was collected and applied a second time. The column was washed and the bound proteins eluted as per the manufacturer's instructions. 500μl fractions were collected, analyzed by SDS-PAGE and visualized using GelCode Blue (Thermo Scientific) following the manufacturers instructions. Fractions containing highly enriched RNase III were pooled and applied to a Amicon-ultra centrifugal concentrator with a of 10 kDa molecular weight cut-off (Millipore). The sample was washed 3x with 60 mM Tris-Cl pH 8.0, 800 mM KCl, 0.2 mM EDTA. The final solution was diluted with two volumes of 80% glycerol and stored at -20°C.

RNase III digestion of X RNA and Qs. Qs and X RNAs were independently incubated at 60°C for 10 minutes in a dry heat block, which was then turned off and the RNAs allowed to cool to 37°C. The labeled (0 .5 pmol) and unlabeled (2.5 pmol) RNA were mixed in 10μl TMN (TMN contains 20mM Tris-acetate pH 7.5, 10mM Magnesium acetate, 100mM Sodium chloride, 1mM Dithiothreitol) and incubated at 37°C for 5 minutes. This was mixed with 12 or 30ng of RNase III in 10μl TMN and incubated at 37°for 1 minute. Reactions w ere stopped with phenol:chloroform. RNA was Ethanol precipitated from the aqueous phase, resuspended in 20μl Gel loading dye II (Ambion), electrophoresed on a 6% denaturing urea-polyacrylamide gel and examined using a phosphorimager.

^{32}P 5' labeling of RNA. RNAs were 5' labeled as described[63]. In vitro transcribed RNA, 10 pmol, was dephosphorylated by mixing with 1 unit CIP (New England Biolabs) in 1x NEB buffer #3 in a 10μl reaction and incubating for 30 minutes at 37°C. The RNA was extracted with phenol:chloroform, p recipitated with 0.3M Sodium acetate and Ethanol and resuspended in 5μl water. The RNA was then labeled in a 10 μl reaction containing with 10 units Polynucleotide kinase (New England Biolabs) and 20 μCi γ^{32}P-ATP in 1x PNK buffer. The reaction was incubated for 30 minutes at 37°C, mixed with loading dye and electrophoresed on a denaturing urea-polyacrylamide gel. Radioactive bands were detected using a phosphorscreen and phosphorimager, excised, mixed with 200μl probe elution buffer (probe elution buffer contains 0.5 M Ammonium acetate, 1mM EDTA and 0.2% SDS) and incubated at 37°for 2 hours. The eluti on was repeated, the RNA precipitated from the fractions with 2.5 volumes ethanol and resuspended in 20μl water.

^{32}P 5' labeling of oligonucleotides. Oligonucleotides used for primer extension were 5' labeled in a 10 μl reaction containing with 10 units Polynucleotide kinase (New England Biolabs) and 20 μCi γ^{32}P-ATP in 1x PNK buffer. The reaction was

incubated for 30 minutes at 37°C, mixed with loading dye and electrophoresed on a denaturing urea-polyacrylamide gel. Radioactive bands were detected using a phosphorscreen and phosphorimager, excised, mixed with 200μl probe elution buffer and incubated at 37°for 2 hours. The elution was repeated, the oilgonucleotide precipitated from the fractions with 2.5 volumes ethanol and resuspended in 20μl water.

Primer extension. Primer extension reactions were performed as described[64]. The RNA to be examined was resuspended in water, then 2 pmol was mixed with 0.6 pmol of 5' ^{32}P-labeled oligonucleotide pCF10_8137 in 10 μl of 1x SB (1x SB contains 10mM Tris-acetate pH 7.4, 60m Ammonium chloride and 6mM β-mercaptoethanol). The RNA mix was heated to 60°C for 20 min in a dry heat block. The heat block was turned off and the reaction allowed to cool to 42°C. The primed RNA was mixed with 2μl 1x SB containing 60μM Magnesium acetate. For the primer extension, 4μl of the primed RNA was mixed with 1x SB + 10mM Magnesium acetate and supplemented with all 4 dNTPs to a final concentration of 1.5mM each and incubated for 5 minutes at 42°C. Reactions for sequence ladders were supplemented with 1.5mM of the desired ddNTP at this time. Then 0.5 units of AMV reverse transcriptase was added to each tube and the reactions incubated for 15 minutes at 42°C. Reactions were stopped and cDNA extracted with phenol:chloroform. The cDNA was precipitated with 2.5M Ammonium acetate and Ethanol, then resuspended in 20μl Gel loading dye II (Ambion), heated to 95°C for 5 minutes and electrophoresed on a denaturing urea-polyacrylamide gel.

SDS/PAGE and Westerns. SDS/PAGE and western blotting of PrgX were performed as previously described[20]. Strains were grown overnight at 37°C without shaking in M9-YEG with selective antibiotics. Cultures were then diluted 1:5 in fresh medium and grown for 90 minutes. A spectrophotometer was used to

measure the optical density of the cultures at λ=600 nm (OD600) and $1*10^9$ cells were collected, assuming that 1 ml of culture at OD600=1.0 contains $1*10^9$ cells. Cells were then pelleted in a microcentrifuge. The pellets were resuspended in 100μl of a solution containing 10mM Tris-HCl pH8, 1mM EDTA, 25% Sucrose and 15 mg/ml Lysozyme. Cells were incubated for 15 minutes at 37°C, then resuspended in 50μl SDS/PAGE loading dye containing β-mercaptoethanol. Samples were then electrophoresed on a 12.5% polyacrylamide gel and transferred to a nitrocellulose membrane (Whatman, protran BA 83, 0.2μm pore) using electrophoresis in Towbin buffer. Membranes were blocked overnight at 4° in PBST + 10% dehydrated skim milk (PBST is 10mM Potassium phosphate buffered saline pH7.4 with 0.1% Tween 20). Membranes were incubated in PBST + 10% dehydrated skim milk containing a 1:1500 dilution of a polyclonal PrgX antibody (Bae 2000 JMB) for 2 hours at room temperature, then washed 3 times in PBST for 10 minutes. Membranes were then incubated for 1-2 hours in PBST + 10% dehydrated skim milk containing a 1:5000 dilution of HRP-Goat α-Rabbit IgG (Invitrogen) and washed 3 times in PBST. Antibody binding to the blots was detected with Super Signal Pico West Chemiluminescent substrate (Thermo scientific) as per the manufacturers instructions, and visualized with film or with a gel dock equipped with a chemiluminescent detector (Cell Biosciences).

Random mutagenesis of Anti-Q. A fragment of pCJ15 that contained Anti-Q was PCR amplified in a buffer known to induce misincorporation of nucleotides due to the presence of manganese salts and a high level of one of the nucleotides[65]. These error-prone PCR reactions contained 10mM Tris-HCl pH 8.7, 50mM KCl, 5μg/ml bovine serum albumin, 0.5μm each oligonucleotide primer p043lacZ_391R and pBR322_3265, 600 pmol DNA template, 4.2mM $MgCl_2$, 0.5mM $MnCl_2$, 2 U Taq DNA polymerase, 0.2mM each dATP, dGTP, dTTP and 3.4 mM dCTP. Reactions were run for 15, 20, 26 and 30 cycles, then

digested for 1 hour with 10 units of the restriction enzyme DpnI to remove the plasmid template.

Error-prone PCR reactions were then used as template for PCR reactions with Pfu-Ultra II high-fidelity DNA polymerase under conditions recommended by the manufacturer. Products of these reactions were ligated into pGEM-T Easy, which was used to transform *E. coli* DH5α. Individual inserts were sequenced to estimate frequency of misincorporated residues for each numbers of cycles tested. We thought it desirable to maximize the frequency of Anti-Q fragments with 1 substitution while minimizing the frequency of Anti-Q fragments with >1 substitution, because it may be difficult to attribute phenotypic effects to a single substitution within a fragment carrying multiple substitutions, so we calculated the frequency at which 0, 1 or >1 mutations should be found within the 115 bp Anti-Q fragment for different numbers of cycles of error-prone PCR. We decided that 15 and 20 cycles of error-prone PCR were optimal for our purposes. These reactions were then used as template for PCR reactions with Taq DNA polymerase and thermopol buffer (New England Biolabs) and primed with oligos pCF10_8437R and XhoI_IRS1. These reactions were digested with XhoI and XbaI and ligated into a similarly digested pCJ16 backbone, which was then used to transform *E. coli* DH5α. Transformants were streaked for single colonies, which were used as template for colony PCR to screen for inserts into the vector. Clones containing vectors with correctly sized and oriented were used as templates for colony PCR using Pfu Ultra II and primed with oligonucleotides pBR322_3265 and lacZ5'R_Seq. These PCR products were purified using a PCR purification kit (Qiagen) and used as templates for IVT reactions.

Results

The *prgQ* and *prgX* operons regulate each other using cis-acting and trans-acting mechanisms

The convergent orientation of the P_Q and P_X promoters raises questions about regulatory interactions between the two operons. Closely spaced promoters can be directly regulated by a single transcription factor, as is the case for pR and pL of the coliphage λ[66], but this was found not to be the case for P_Q and P_X. Anushree Chatterjee of the Hu lab recognized the potential for these operons to regulate each other via transcriptional interference. Previous work had revealed that Anti-Q, derived from the *prgX* operon, negatively regulates *prgQ* transcripts and that *prgX* transcripts were negatively regulated during pheromone induction by unknown mechanisms[20,21]. Because of this and the suggestive arrangement of the promoters we hypothesized that these promoters could regulate each other in cis via transcriptional interference and that RNAs from each operon could also regulate the opposing operon in trans.

In order to test this hypothesis and distinguish between regulation via transcriptional interference, in which the promoters must be in cis, and regulation in which RNAs can be provided in trans, we developed a system whereby the operons could be located in either configuration. To develop this system, we needed to be able to selectively inactivate each promoter. To do this we sought to introduce mutations into the -10 region for each promoter and test their effects. The transcription initiation site of the *prgQ* operon had been precisely identified via a 5'-RACE strategy[14]. We compared the sequence upstream of *prgQ* +1 to the *E. coli* consensus promoter and identified likely -10 and -35 sequences[67] (Figure 4). We introduced a two nucleotide substitution into the -10 region, generating a *Bsr*GI restriction site, and tested a plasmid bearing this mutation for

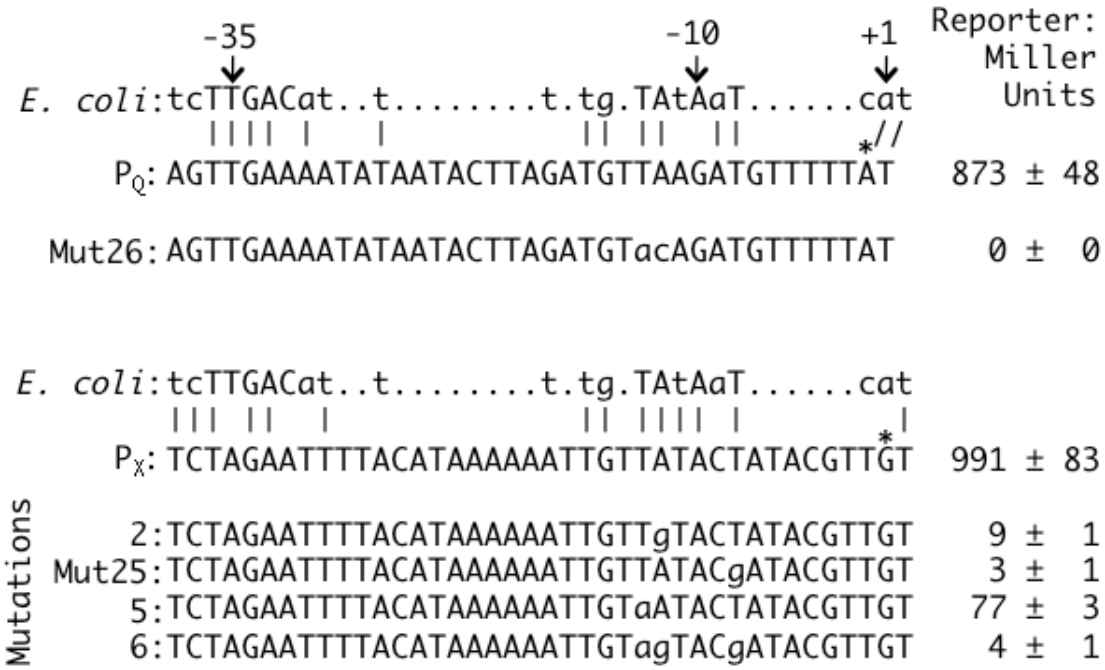


Figure 4: Mutations used to inactivate P_X and P_Q. The sequence of P_X and P_Q is shown, along with an alignment to the canonical *E. coli* σ^{70} promoter. In the *E. coli* promoter, uppercase letters indicate highly conserved residues, lowercase letters indicate poorly conserved residues, dots (.) indicate the presence of non-conserved residues. The numbers +1, -10 and -35 are positions in relation to the transcriptional start site (+1). In the *E. faecalis* promoters, uppercase letters indicate wild-type residues and lowercase letters indicate substitutions in the mutation being tested. The mutations indicated were cloned into plasmids upstream of a transcriptional *lacZ* fusion, used to transform *E. faecalis* OG1Sp and tested for β-galactosidase production using a Miller assay. Mutations to P_X were tested in pCJ1, which has *lacZ* fused at P_X +221, in *E. faecalis* strains bearing pCF10. The mutation to P_Q was tested in pBK1. Average β-galactosidase production and range for wild-type and mutated promoters is shown.

expression of a reporter fused downstream. This mutation caused expression of the reporter to drop below detectable levels.

The transcription initiation site of the *prgX* operon had been previously mapped to two closely spaced nucleotides using primer extension. Previous attempts to inactivate P_X based on the upstream start site had failed. Because of this we used a 5'-RACE strategy to reevaluate the *prgX* initiation site and identified the downstream nucleotide as the major transcription initiation site; this experiment was performed in collaboration with Thinh Le. Comparing sequence upstream of this nucleotide with *L. lactis* and *E. coli* consensus promoters revealed previously untargeted regions to be the likely -10 and -35 sequences. We selectively introduced substitutions into the -10 region and tested plasmids bearing these mutations for changes in expression of a reporter fused downstream (Figure 4). Changes in reporter expression suggested that we had identified the -10 region. We selected a mutation and designated it #25 for continued use.

In order to place the *prgQ* and *prgX* operons in a trans configuration, we cloned these promoter mutations onto pBK1 and pDM5, compatible plasmids that contain similar fragments of the *prgQ prgX* region of pCF10 (Figure 5). The plasmid copy number of pDM5 is approximately 5-fold greater than pBK1. By examining products from the promoter active on pBK1 we were able to test RNA expression of each operon when the other operon was absent, located in cis, or located in trans at a higher copy number. Neither plasmid contains the *prgX* ORF, allowing us to examine interactions between the operons in the absence of PrgX. We transformed OG1Sp with various combinations of the plasmids, prepared RNA from these strains and analyzed the RNA using northern blots. We found that plasmids containing a mutated promoter did not produce detectable RNA from that promoter (Figure 6, part A lane 7 and part B lane 7) and that pDM5 containing an active *prgQ* or *prgX* operon produced more

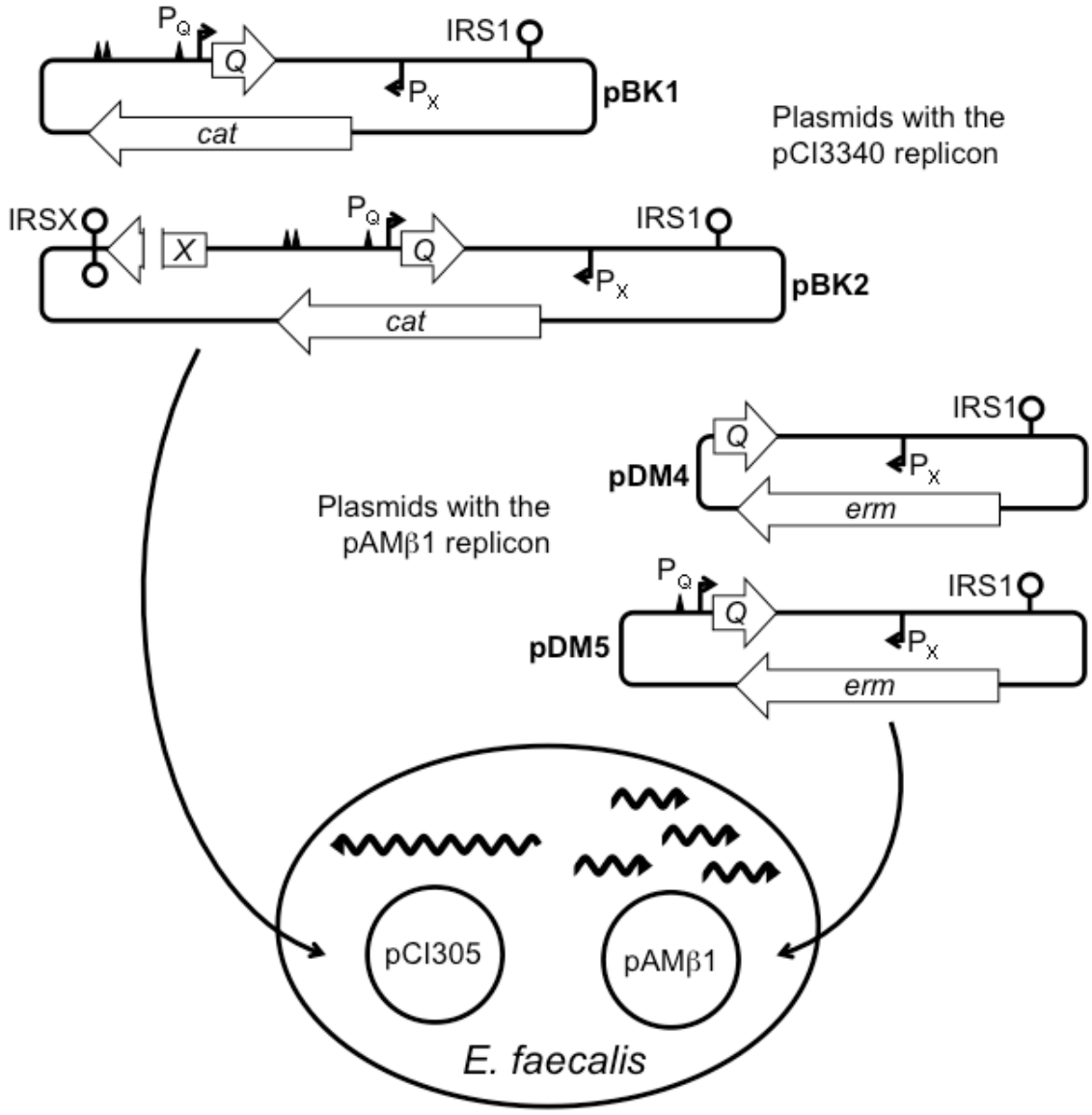
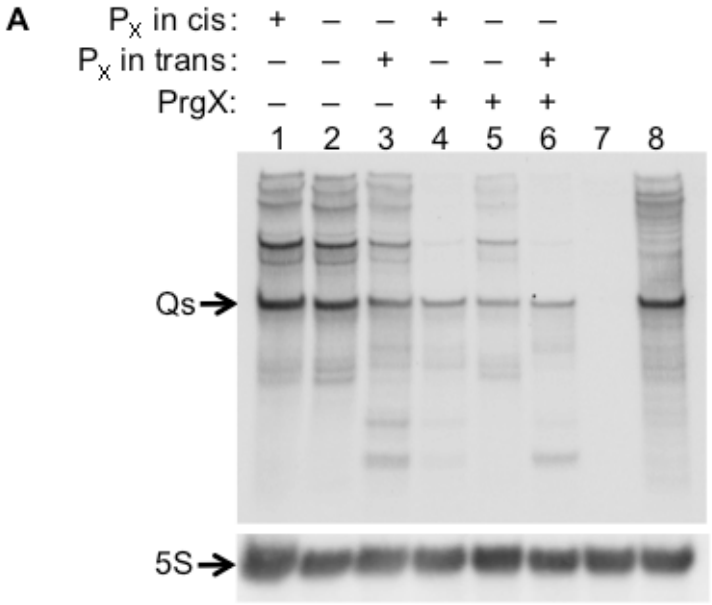

P_Q
Q
IRS1
P_X
pBK1
cat
Plasmids with the
pCI3340 replicon
IRSX
X
P_Q
Q
IRS1
P_X
pBK2
cat
IRS1
Q
P_X
pDM4
erm
Plasmids with the
pAMβ1 replicon
P_Q
Q
IRS1
P_X
pDM5
erm
pCI305
pAMβ1
E. faecalis

Figure 5

Figure 5: Schematic for providing *prgQ* and *prgX* products in trans. Slightly different regions of pCF10 were cloned onto compatible plasmids pCI305, bearing the pCI3340 replicon and pAT18, bearing the pAMβ1 replicon with different antibiotic selection markers. The *cat* gene confers chloramphenicol resistance. The *erm* gene confers erythromycin resistance. The name of each plasmid is indicated in bold, features of pCF10 cloned into each plasmid are shown. Bent arrows indicate transcription start sites. Open arrows indicate open reading frames, labeled Q and X for *prgQ* and *prgX*, respectively. Lollipops indicate factor-independent terminators. Triangles indicate PrgX binding sites. Each promoter on any plasmid can be selectively inactivated by a point mutation in its -10 region (see figure 3). pBK1 can provide Qs and/or Anti-Q and is responsive to PrgX. pBK2 can provide Qs and/or Anti-Q and X RNA and is responsive to PrgX. pDM4 can provide Anti-Q. pDM5 can provide Qs and/or Anti-Q and is not responsive to PrgX because the primary PrgX binding site is missing. These plasmids are then used to transform electrocompetent *E. faecalis* and both amplicons maintained by antibiotic selection. pCI305 has a copy number of 11 ± 4 and pAT18 has a copy number of 61 ± 25 in *E. faecalis*, as determined by Dawn Manias using real-time PCR to measure the abundance of the *prgX* gene from total cellular DNA after it had been cloned in single copy into different plasmids.


A
P_X in cis: + – – + – –
P_X in trans: – – + – – +
PrgX: – – – + + +
1 2 3 4 5 6 7 8
Qs
5S

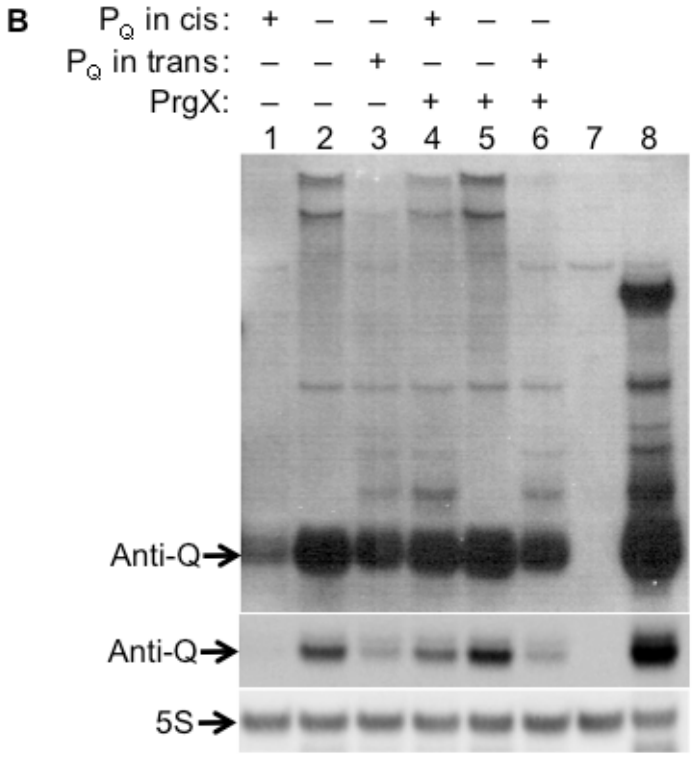

B
P_Q in cis: + – – + – –
P_Q in trans: – – + – – +
PrgX: – – – + + +
1 2 3 4 5 6 7 8
Anti-Q
Anti-Q
5S

Figure 6

Figure 6: *prgQ* and *prgX* regulate each other in cis and in trans. Combinations of pBK1 and pDM5 with and without mutations to inactivate P_Q and P_X were used to transform electrocompetent *E. faecalis* OG1Sp or 100-5, which carries *prgX* on the chromosome. RNA was prepared from these strains and 1μg was used for northern analysis. **A**. A blot hybridized to a probe specific for Qs. Lane 1, OG1Sp + pBK1; lane 2, OG1Sp + pBK1-25; lane 3, OG1Sp + pBK1-25 + pDM5-26; lane 4, 100-5 + pBK1; lane 5, 100-5 + pBK1-25; lane 6, 100-5 + pBK1-25 + pDM5-26; lane 7, OG1Sp + pBK1-26; lane 8, OG1Sp + pBK1-26 + pDM5-25. **B**. A blot hybridized to a probe specific for Anti-Q. Lane 1, OG1Sp + pBK1; lane 2, OG1Sp + pBK1-26; lane 3, OG1Sp + pBK1-26 + pDM5-25; lane 4, 100-5 + pBK1; lane 5, 100-5 + pBK1-26; lane 6, 100-5 + pBK1-26 + pDM5-25; lane 7, OG1Sp + pBK1-25; lane 8, OG1Sp + pBK1-25 + pDM5-26. The lower panel labeled Anti-Q is a shorter exposure of the same blot shown in the upper panel. *E. faecalis* 5S is used as a loading control. Promoters labeled “in cis” are located on pBK1. The RNA detected was produced from pBK1-derived plasmids. If the opposing promoter is present in cis, it is also on pBK1, if the opposing promoter is present in trans, it is on a pDM5-derived plasmid. PrgX, when present, is provided from a chromosomal copy under control of the ectopic P_{23} promoter. Plasmids designated -25 have P_X inactivated and plasmids designated -26 have P_Q inactivated.

transcripts from that operon than pBK1 with the same operon (Figure 6, parts A and B, compare lane 8 to lane 2)

To examine the effects of *prgX* on *prgQ* expression, we hybridized a blot with a probe specific for full-length Qs (Figure 6A). We found that when both operons were present in cis, pBK1 generated abundant Qs as well as longer RNAs (lane 1), indicating transcription extends past IRS1. Inactivating P_X caused an increase in *prgQ* transcripts (lane 2), but providing *prgX* transcripts in trans reversed this effect (lane 3), indicating that products of the *prgX* operon negatively regulate *prgQ* expression, and that the bulk of this regulation may be due to trans-acting mechanisms.

To examine the effects of *prgQ* products on *prgX* expression, we hybridized a blot with a probe specific for Anti-Q (Figure 6B). We found that when both operons were present in cis, pBK1 generated very little detectable *prgX* transcripts (lane 1), though a band corresponding to Anti-Q is faintly visible. Inactivating P_Q caused a substantial increase in expression of Anti-Q as well as expression of extended *prgX* transcripts (lane 2). Providing *prgQ* products in trans caused a partial reversal of this effect (lane 3), suggesting that *prgQ* products regulate *prgX* transcripts in trans, but that *prgQ* also represses *prgX* using mechanisms that depend on the promoters being located in cis.

Because P_Q exerts transcriptional interference against P_X we hypothesized that the cis acting sequences that negatively regulate Anti-Q and *prgX* expression in a PrgX-dependent manner are the *prgQ* promoter. In order to test this we used 100-5, a strain of *E. faecalis* in which PrgX is provided from the chromosome under control of P_{23}, a constitutive promoter. We transformed 100-5 with plasmids that provide *prgQ* and *prgX* transcripts in cis and in trans and compared RNA expression to that of OG1Sp.

When hybridized a blot with a probe specific for Anti-Q (Figure 6B), we found that, when both operons were present in cis, PrgX caused a substantial increase in *prgX* transcript expression (lane 4). Inactivating P_Q caused *prgX* transcript expression to increase to the same level as when P_Q was inactivated in the absence of PrgX (lane 5, compare to lane 2) . Providing Qs in trans caused a substantial reduction in *prgX* transcript expression (lane 6). When we examined a blot hybridized with a probe specific for Qs (Figure 6A), we found that PrgX reduced expression of *prgQ* transcripts, particularly those that extend past IRS1 (lane 4, compare with lane 1), in agreement with previous work. Inactivating P_X allowed an increase in *prgQ* expression (lane 5) but providing *prgX* transcripts in trans reversed this effect (lane 6). This is generally the same pattern as seen in the absence of PrgX. These data indicate that the cis acting sequences that negatively regulate Anti-Q and *prgX* expression are the *prgQ* promoter and that the major effects of PrgX in regulation of the *prgX* operon are due to repression of P_Q, which removes both cis-acting and trans-acting regulation of the *prgQ* operon on the *prgX* operon.

Anti-Q is generated via multiple pathways

Although the experiment described above elucidated the role of PrgX and P_Q in regulation of Anti-Q, it was unclear what factors directed generation of Anti-Q RNA. Because Anti-Q does not resemble either a factor-independent terminator or a rho-dependent terminator, we hypothesized that Anti-Q was processed from longer transcripts and that *prgX* sequences downstream of Anti-Q were necessary for generation of Anti-Q. The 3' terminus of Anti-Q was previously identified by S1 mapping[21]. We sought to confirm the location of this 3' terminus using a 3' RACE protocol. We determined the 3' terminus of Anti-Q prepared from *E. faecalis* OG1Sp with either pCF10 or pDM4. The 3' terminus that we identified was the same for both strains, but different than the 3' terminus identified by S1 mapping of Anti-Q from OG1RF (Figure 7A). This suggests that

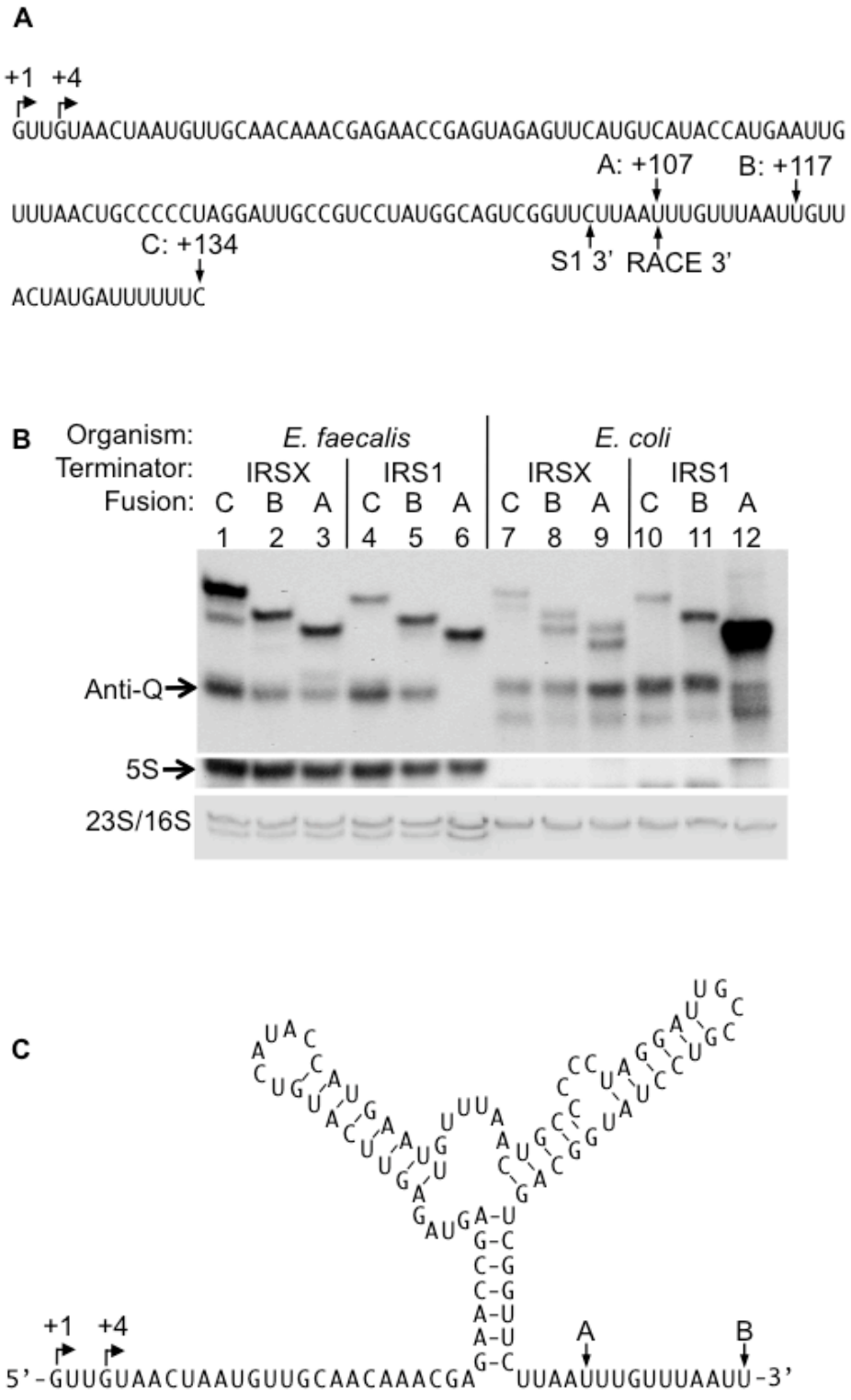

A
+1 +4
GUUGUAACUAAUGUUGCAACAAACGAGAACCGAGUAGAGUUCAUGUCAUACCAUGAAUUG
A: +107
B: +117
UUUAACUGCCCCCUAGGAUUGCCGUCCUAUGGCAGUCGGUUCUUAAUUUGUUUAAUUGUU
C: +134
S1 3'
RACE 3'
ACUAUGAUUUUUUC
B
Organism:
E. faecalis
E. coli
Terminator:
IRSX
IRS1
IRSX
IRS1
Fusion:
C B A C B A C B A C B A
1 2 3 4 5 6 7 8 9 10 11 12
Anti-Q
5S
23S/16S
C
+1 +4
A
B
5'-GUUGUAACUAAUGUUGCAACAAACGA
UUAAUUUGUUUAAUU-3'

Figure 7

Figure 7: Anti-Q production can be directed by *prgX* sequences proximal and distal to its 3' terminus. **A**. Sequence map of Anti-Q. The +1 and +4 *prgX* transcriptional start sites are shown. The 3' termini of Anti-Q are shown, as determined by S1 mapping and 3' RACE. For experiments to identify *prgX* sequences necessary for Anti-Q genesis, two different terminators were fused to three locations, labeled A (pCJ13, pCJ14), B (pCJ12, pCJ15) and C (pCJ11, pCJ16). The terminators were either the wild-type *prgX* terminator, IRSX (pCJ11, pCJ12, pCJ13) or the ectopic IRS1 terminator (pCJ14, pCJ15, pCJ16) **B**. Plasmids containing various deletions of *prgX* sequences were used to transform electrocompetent *E. faecalis* OG1Sp or *E. coli* DH5α. RNA was prepared from these strains and 1μg was used for northern analysis. The membrane was hybridized with a RNA probe specific for Anti-Q, *E. faecalis* 5S is shown as a loading control, as is a an Ethidium bromide stain of the 23S and 16S rRNA. Lanes are labeled with the species of bacterium from which the RNA was prepared, the terminator, IRS1 or IRSX that was fused to the 3' end of Anti-Q and the location of this fusion: A, B or C, as shown on the map in part A. **C**. The experimentally determined structure of Anti-Q is shown, with transcriptional start sites indicated by bent arrows and the location of fusions A and B indicated.

there may be some variability in the precise location of this terminus in vivo, possibly due to strain differences, or that the location identified by one of the protocols is incorrect.

To determine the sequences necessary for Anti-Q generation we performed a deletion analysis of *prgX* in vivo. The plasmid pDM4 contains P_X +1 to +221, followed by the *prgX* terminator, IRSX, and is capable of producing Anti-Q. Using these sequences as a starting point we cloned P_X and various amounts of downstream sequences into pCI305 and placed either IRSX or an ectopic terminator, IRS1, immediately downstream (Figure 7A). This allowed us to test the contribution of proximal and distal *prgX* sequences to Anti-Q formation. We then transformed OG1Sp with these plasmids and used northern blots to test these strains for their ability to generate Anti-Q (Figure 7B). When *prgX* transcripts were terminated with IRSX, they produced Anti-Q, even when all proximal *prgX* sequences downstream of Anti-Q were deleted (Lanes 1-3). When *prgX* transcripts were terminated with IRS1 Anti-Q was generated when some downstream sequence was included (lanes 4 and 5) but disappeared when all *prgX* sequences downstream of Anti-Q were deleted (lane 6). These data indicated that *prgX* sequences both proximal and distal to the 3' terminus of Anti-Q contribute to Anti-Q generation, and that either are sufficient, but neither necessary, for Anti-Q generation.

Because RNAP enzymes from distantly related organisms operate in the same way, intrinsic terminators are broadly recognized by bacterial and phage RNAP enzymes[68,69]. RNA degradation machinery, on the other hand, is more nuanced. Some factors are well conserved whereas others vary between phylogenically distant bacteria. We hypothesized that if Anti-Q were generated by well conserved mechanisms, it should be formed from a *prgX* template in distantly related bacteria. We transformed *E. coli* with the plasmids used to study Anti-Q formation and repeated the *prgX* deletion analysis in that organism (Figure 7B

lanes 7-12). The results were the very similar to those obtained from *E. faecalis*. The only difference was that a small amount of Anti-Q was formed when all downstream *prgQ* sequences were deleted and the transcript terminated with IRS1. This suggests that the factors needed to form Anti-Q in vivo are conserved across great evolutionary distance, though it is unclear if these are conserved nuclease activities or recognition of a termination signal.

The 3' end of Anti-Q functions as a factor independent terminator

Analysis of sequences immediately downstream of Anti-Q revealed a series of thymidines between fusion points A and B (Figure 7), which would become a polyuridine tract when transcribed into RNA. Such tracts are characteristic of factor-independent terminators and are thought to stall RNAP long enough for a stable hairpin structure to form in the nascent RNA, signaling the polymerase to disengage from the DNA template. The secondary structure of Anti-Q RNA has been determined by lead-acetate digestion and such a hairpin structure does not exist[55]. Rather, there is a stem that is part of a branching structure upstream of the polyuridine tract. We wondered if the basal stem of this branching structure could serve the same role as the missing hairpin, signaling RNAP to terminate transcription within the thymidine-rich region (Figure 7C).

The mechanism by which factor-independent terminator function is not completely understood, but it is known that other host factors are not needed in order for termination to occur. We reasoned that if Anti-Q acts as a termination signal, we could reconstitute this activity in vitro. We hypothesized that the sequences immediately downstream of Anti-Q constitute a polyuridine tract and are necessary for termination in vitro and that termination should not be sensitive to which terminator, IRS1 or IRSX, was fused downstream. To test this we performed in vitro transcription (IVT) assays using some of the plasmids used in the *prgX* deletion study as templates (Figure 8A). We found that if the polyuridine tract was present, IVT reactions generated a band of the correct size to be Anti-Q

(lanes 1 and 3). If the polyuridine tract was absent, no Anti-Q was generated, regardless of which terminator was located downstream (lanes 2 and 4). This indicates that the polyuridine tract is necessary for formation of Anti-Q in vitro.

It is possible that the Anti-Q band in generated by the IVT assay is the result of transcriptional pausing or arrest, rather than termination. If the band is the result of termination, it should no longer be associated with RNAP. In order to distinguish between these possibilities we performed IVT reactions using His-tagged RNAP bound to nickel-agarose beads. We electrophoresed transcripts associated with the nickel-agarose beads separately from those that were released into the supernatant (Figure 8). We found that, when compared to total transcripts generated in a reaction (lane 1) Anti-Q was depleted in the bead-associated fraction (lane 2) and enriched in the supernatant (lane 3), confirming that formation of Anti-Q in vitro is the result of termination.

Because Anti-Q behaves like an intrinsic terminator, we reasoned that it must have structural features that allow it to act as such. We hypothesized that nucleotides contributing to the stability of the stems, particularly the basal stem, are more important for Anti-Q termination than nucleotides in the loops. To test this we performed an unbiased mutagenesis of Anti-Q and determined the sequence of mutants that impacted termination of Anti-Q.

In order to mutate the Anti-Q region we used error-prone PCR to amplify Anti-Q and ligated the mutated fragment into a plasmid upstream of the IRS1 and IRSX terminators (see methods). We used clones of this plasmid as a template for PCR reactions to generate linear templates for IVT reactions (Figure 9A). We performed the IVT reactions on templates containing mutated as well as wild-type Anti-Q sequences and determined the frequency of termination at Anti-Q relative to the frequency of termination at IRS1. We sequenced templates that

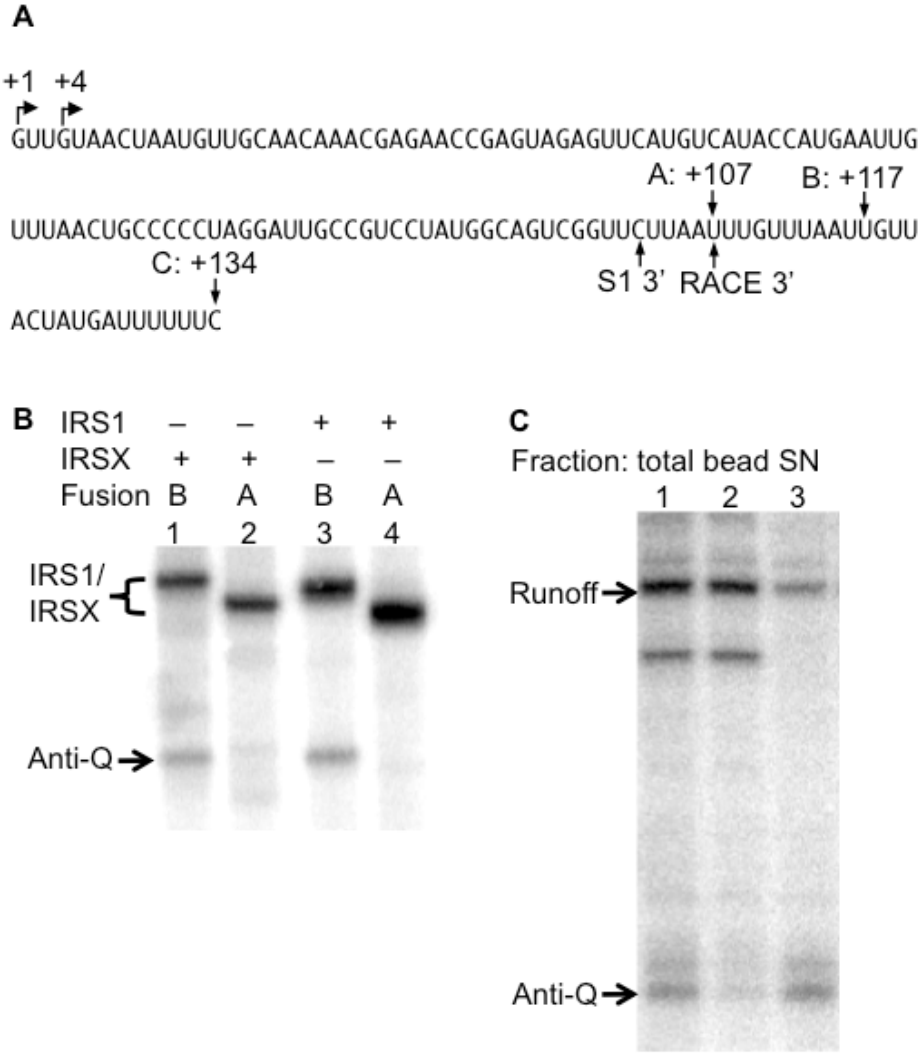

A
+1 +4
GUUGUAACUAAUGUUGCAACAAACGAGAACCGAGUAGAGUUCAUGUCAUACCAUGAAUUG
A: +107
B: +117
UUUAACUGCCCCCUAGGAUUGCCGUCCUAUGGCAGUCGGUUCUUAAUUUGUUUAAUUGUU
C: +134
S1 3'
RACE 3'
ACUAUGAUUUUUUC
B
IRS1 – – + +
IRSX + + – –
Fusion B A B A
1 2 3 4
IRS1/
IRSX
Anti-Q
C
Fraction: total bead SN
1 2 3
Runoff
Anti-Q

Figure 8

Figure 8: Anti-Q encodes a factor-independent terminator **A**. An IVT assay performed using plasmid templates: lane 1, pCJ12; lane 2, pCJ13; lane 3, pCJ15; lane 4; pCJ14. The band formed by termination at Anti-Q is indicated, as is the band formed by termination at IRSX (lanes 1 & 2) or IRS1 (lanes 3 & 4). **B**. An IVT performed using RNAP that was attached to Nickel-agarose beads. The template was a linear PCR product containing P_X -167 to +400 with P_Q inactivated by a mutation to its -10 region. Two identical reactions were performed in parallel. One reaction was stopped using phenol:chloroform (lane 1), which denatures the RNAP, releasing all transcripts. The other reaction was stopped by putting it on ice. The reaction was then centrifuged for 15 seconds in a microcentrifuge and 20μl of the supernatant, containing terminated transcripts, was removed and added to loading dye (lanes 3). The bead pellet was washed once with 100μl IVT buffer, then phenol:chloroform (35μl) and IVT buffer (20μl) were added to the pellet. The aqueous phase, containing RNAP-assiciated transcripts, was recovered and added to loading dye (lanes 2).
caused a 20% change in termination at Anti-Q compared to a wild-type template. Most of the mutant sequences contained a single nucleotide substitution. Several mutants were discarded because they contained multiple substitutions; however, two mutants with two substitutions each have been included in the analysis because in both cases one of the substitutions is shared with another mutant identified in the screen, allowing us to distinguish between the effects of the two substitutions.

A

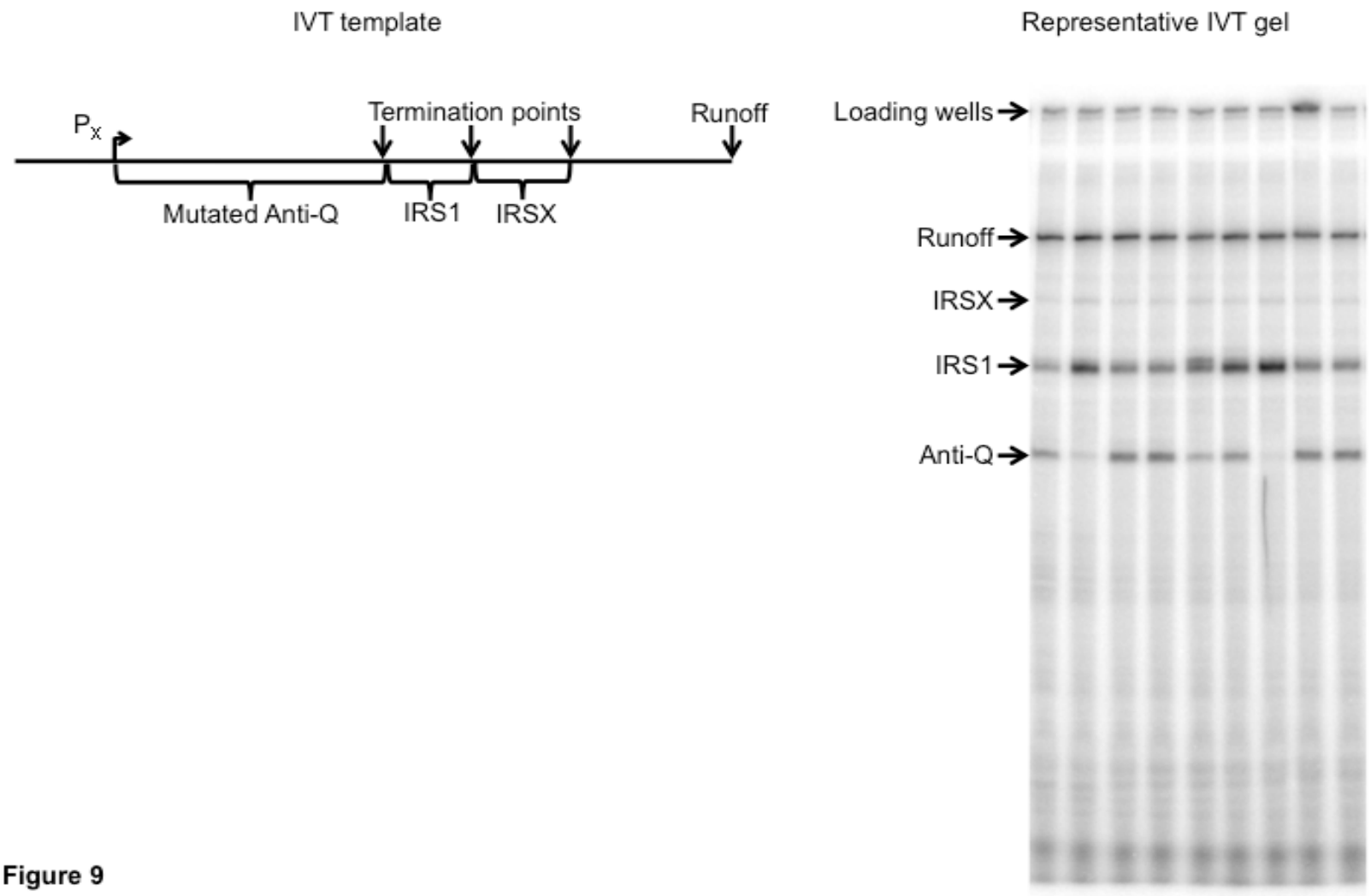
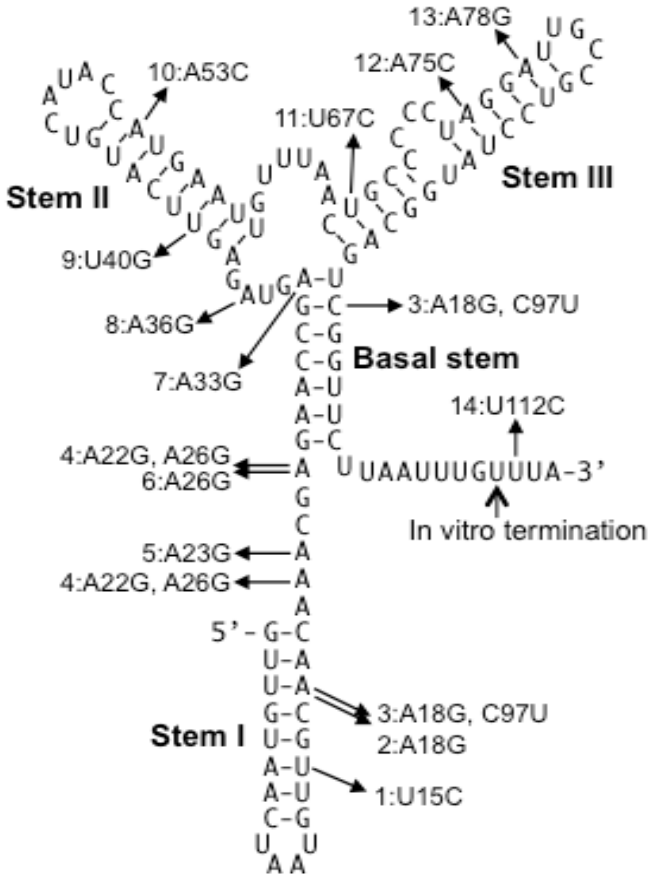
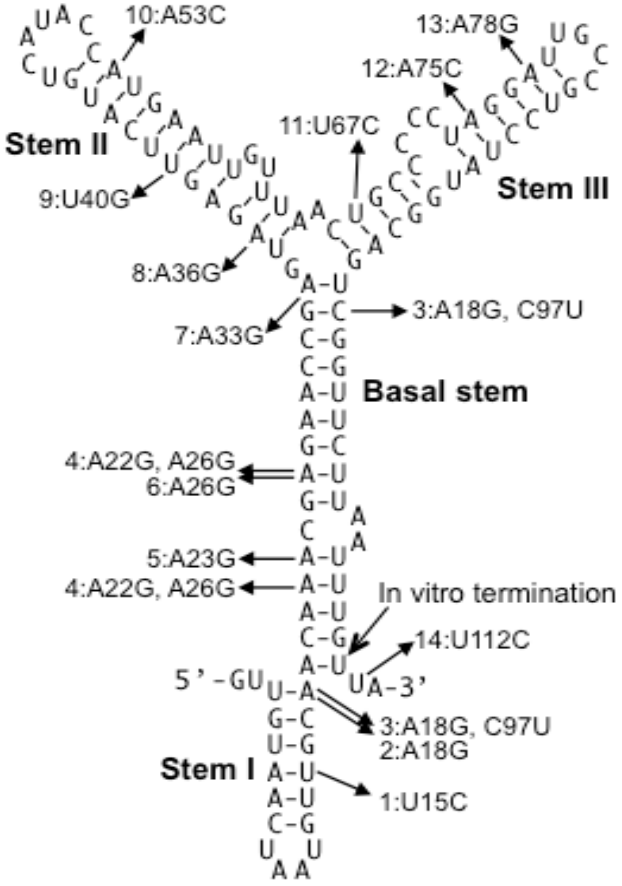

IVT template
Representative IVT gel
Px
Termination points
Runoff
Mutated Anti-Q
IRS1
IRSX
Loading wells
Runoff
IRSX
IRS1
Anti-Q

Figure 9



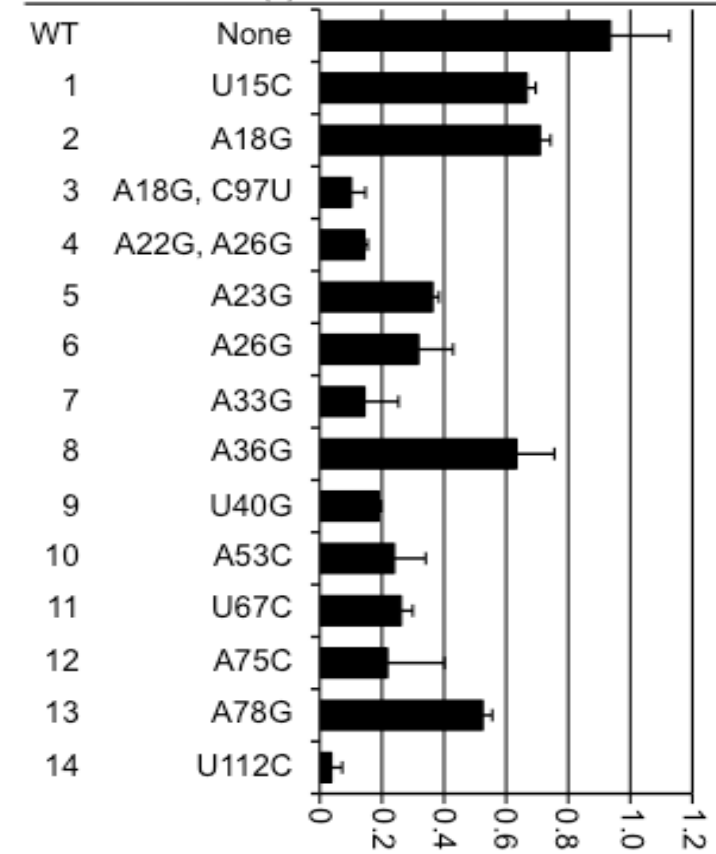
B. Experimentally determined structure



C. Proposed alternate structure

Figure 9

D



No.	Mutation(s)	Anti-Q/IRS1	ΔG (kcal/mole) of given structure A	B	Predicted structural effects
WT	None		-40.4	-37.2	
1	U15C		-37.5	-34.3	
2	A18G		-40.1	-37.1	
3	A18G, C97U		-39.0	-35.0	Shortens basal stem by 2 base pairs
4	A22G, A26G		-40.2	-36.5	
5	A23G		-40.4	-36.9	
6	A26G		-40.2	-36.4	
7	A33G		-40.5	-36.9	Shortens basal stem by 1 base pair
8	A36G		-40.4	-36.8	
9	U40G		-37.4	-35.4	Shortens stem II by 2 base pairs
10	A53C		-35.4	-32.2	Shortens stem II by 2 base pairs
11	U67C		-37.9	-35.2	Shortens stem III by 2 base pairs
12	A75C		-37.4	-33.9	Shortens stem III
13	A78G		-36.8	-33.6	
14	U112C		-40.4	-36.9	Disrupts polyuridine tract

Figure 9

Figure 9: An unbiased mutagenic study of the Anti-Q terminator. **A.** A cartoon of the linear DNA fragment containing random mutations used as an IVT template is shown on the left. Sequences carrying Anti-Q, IRS1 and IRSX are indicated by labeled brackets. The 3' end of transcripts formed by termination or polymerase reaching the end of the template (runoff) are indicated by arrows. The panel at the right shows a representative IVT assay with bands generated by termination or runoff indicated. **B**. Experimentally determined and **C**. Proposed alternate structures of Anti-Q, as described in the text. The different stems are labeled in bold. The in vitro termination point, as described in the text and shown in Figure 10 is labeled. Polymorphisms carried by mutants identified in the screen are labeled with an identifying number and sequence substitutions. **D**. Characteristics of the mutants identified in the screen. The table shows the identification of the mutant, polymorphisms within that mutant, the degree of termination at Anti-Q as determined by the intensity of the Anti-Q termination band divided by the intensity of the IRS1 termination band, the change in Gibb's free energy, a measure of structural stability, of each Anti-Q structure and any predicted effects of the substitutions not described by changes in overall stability.

We identified the location of the mutations within the experimentally determined structure of Anti-Q[55]. All of the mutations analyzed decreased the frequency of termination at IRS1 (Figure 9C). Most of the mutations identified are within stems and are predicted to increase the minimum free energy of Anti-Q, as determined by mfold, indicating that they decrease the stability of the Anti-Q structure (Figure 9A). The structure of Anti-Q was determined using the 3' terminus of the mature RNA. Analysis of Anti-Q sequences reveals that transcription of the poly-U tract might lead to additional base-pair interactions with the effect of extending the basal stem of Anti-Q. Additionally, the stem of loop II could also be extended through additional nucleotide pairings. We analyzed the effects of the mutations within the context of an altered Anti-Q structure that contains these expended pairings. With the exception of mutation #14, which will be discussed below, the mutations identified in the screen all fall within the stems of and destabilize the alternate Anti-Q structure (Figure 9B).

The mutations that reduced termination at Anti-Q the most were generally those that fell within the basal stem (Figure 9, mutants 3, 4, 5, 6 and 7). All of these weakened the structure of Anti-Q. Mutants 3 and 7 are also predicted to shorten the basal stem by 2 or 1 base pairs, respectively. Mutations that fell within stems I, II and III have varying effects on termination at Anti-Q. Mutations predicted to shorten the stem in which they occur (mutants 9, 10, 11 and 12) have a more dramatic effect than those that simply reduce the overall stability of the Anti-Q structure (mutants 1, 2, 8 and 13). These results suggest that the stems of the predicted Anti-Q structure are important for termination of Anti-Q. The stability of the structure appears to contribute to termination, but is secondary to the structure itself, i.e. base-pairing within the stems. It appears that the stability and size of the basal stem helix is critical for causing termination.

The mutation that had the most dramatic effect on termination of Anti-Q is # 14, U112C. Its effect could be explained by proposing that U112 pairs with A18 and

forms part of the basal stem. Subsequent experiments suggest that this nucleotide disrupts the ability of the polyuridine tract to pause RNAP at the proper location for termination instead.

In order to better understand the contribution of the polyuridine tract to termination at Anti-Q we mapped the site of termination in vitro. To do this we compared an IVT reaction of Anti-Q to a chain-termination sequencing ladder also made by in vitro transcription. P_X has two transcriptional start sites; we anticipated that this would confound interpretation of the ladder and the location of termination, so we cloned Anti-Q immediately downstream of P_Q, which has only one identified transcriptional start site and mapped the precise site of Anti-Q termination in this context. We found that Anti-Q directed termination downstream of this ectopic promoter, and that termination occurred at two residues within the polyuridine tract (Figure 10), which is not uncommon for factor independent terminators[70]. Both of these residues are U's with G's immediately upstream and another U immediately downstream. In order for an intrinsic terminator to function, the RNAP must pause at the point of termination. Mutation #14 is a U->C transition to the base immediately downstream of the first termination point. We suspect that this substitution disrupts the pause signal, allowing RNAP to transcribe past the termination point before the terminator has time to function.

We note that the DNA fragment studied in the random mutagenesis screen contains the upstream Anti-Q termination point, but not the downstream termination point. While it is not uncommon for intrinsic terminators to have multiple termination points, it does raise the possibility that some mutants identified in the study may have quantitatively different effects on termination of Anti-Q transcribed from a template with both termination points, particularly

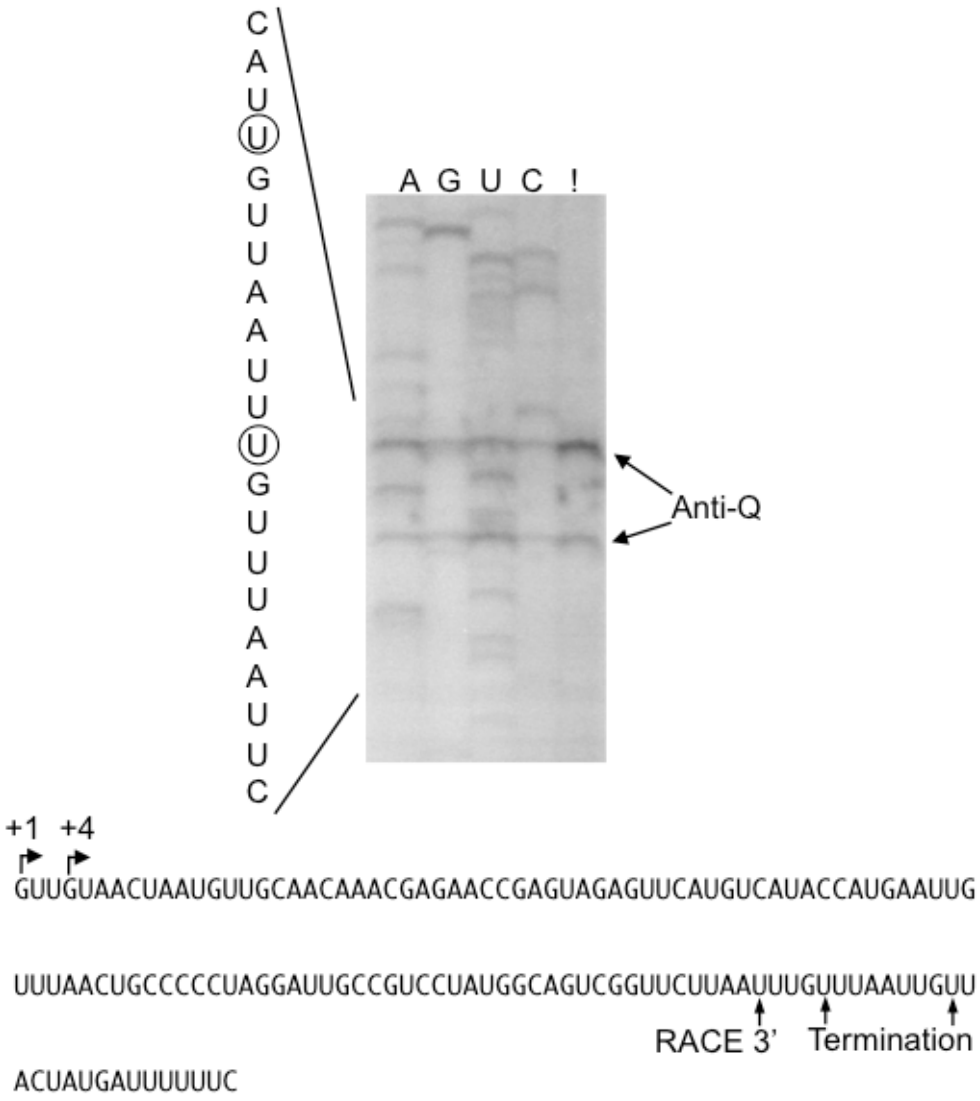


Figure 10: Determination of the Anti-Q termination point in vitro. A PCR amplification product of pCJ19 was used as a template for the IVT reactions shown. This contains Anti-Q sequences (*prgX* mRNA-encoding sequence +1 to +223) fused to P_Q +4. The lanes marked A, G, U, C are chain termination reactions using the indicated 3'-Methoxy-NTP. The lane marked with an ! is a standard IVT reaction showing terminated products. The RNA sequence of Anti-Q is shown below with arrows indicating the in vitro termination points and the 3' terminus of Anti-Q as identified by 3' RACE.

mutant 14. We also note that further work is needed to test the feasibility of the proposed alternate Anti-Q structure, as it is not yet experimentally supported. Despite these caveats, the results of the study support the view that Anti-Q is a factor independent terminator.

Nascent *prgQ* transcripts contain competing terminator and antiterminator structures.

Work by Taeok Bae revealed that Anti-Q negatively regulates the expression of genes in the *prgQ* operon downstream of IRS1[21]. Analysis of *prgQ* sequences suggested that IRS1 is a factor independent terminator, but that *prgQ* also contains sequences to form a competing antiterminator structure, and that Anti-Q could interact with nascent *prgQ* transcripts to favor formation of the terminator. In order to investigate the activities of the proposed terminator and antiterminator structures, we used a plasmid cloned by Brianna Kozlowicz that contains the pCF10 region from P_Q to IRS1 into a reporter plasmid with a *lacZ* transcriptional fusion downstream of the terminator (pBK1, Fig. 5). This allowed us to use β-galactosidase expression and *lacZ* transcript abundance to report extension of transcripts past IRS1. Although pBK1 retains P_X, it lacks the *prgX* structural gene and produces very little Anti-Q (Figure 6B, lane 1), allowing us to use pBK1 to examine the intrinsic termination behavior of nascent *prgQ* transcripts and mutant derivatives without the influence of these negative regulators. Initially, we used β-galactosidase assays to assess transcription past IRS1 in pBK1 and its derivatives (Table 4). We then used qRT-PCR to quantify *lacZ* mRNA levels directly.

In most strains β-galactosidase activity mirrored *lacZ* transcript levels as measured by qRT-PCR. However, some vectors with mutations to IRS1 showed decreased β-galactosidase expression relative to the amount of *lacZ* mRNA. We hypothesized that pBK1 and most of its derivatives contained pCF10 sequences between IRS1 and the *lacZ* junction that weakly interacted with the *lacZ* RBS,

blocking *lacZ* translation (Figure 11A). Deleting the 3' half of the terminator might increase the efficiency of this RBS occlusion, repressing translation of β-galactosidase. In order to test this possibility, we created derivatives of pBK1 and pBK1-18 which lack the potential RBS occluding sequence, generating pBK1-17 and pBK1-13, respectively. The β-galactosidase levels of these mutants was significantly increased while the *lacZ* transcript abundance remained unchanged (Figure 11B). This supports the hypothesis that the differences in β-galactosidase expression observed between these reporters was due to efficiency of translation, an artifact of plasmid construction, rather than to changes in termination at IRS1. For this reason, in experiments that involved mutations to IRS1 we used *lacZ* transcript abundance rather than β-galactosidase activity to assay terminator activity in vivo.

IRS1 encodes a G+C rich inverted repeat that can form a stable stem-loop structure followed by a polyuridine tract, both characteristic of factor-independent terminators[71,72]. We hypothesized that disrupting IRS1 would lead to a decrease in termination, which would result in an increase in transcription past IRS1. To test this we compared *lacZ* transcript abundance between pBK1 and a pBK1-derived construct with a deletion of the 3' side of the inverted repeat (Figure 12, pBK1-18), predicted to prevent formation of the terminator stem. This mutation had no effect on *lacZ* transcript levels, suggesting that IRS1 does not act as an intrinsic terminator in the context of our plasmid.

These results conflicted with previous interpretations of northern blots and sequence data suggesting that IRS1 is a factor-independent terminator. To reconcile these results, we hypothesized that, in the absence of Anti-Q RNA, formation of a previously postulated antiterminator is constitutive within nascent *prgQ* transcripts, precluding terminator formation and allowing extension into downstream sequences[21].

Table 4: β-galactosidase activity of OG1Sp strains bearing the indicated plasmids, mean and standard deviation expressed in Miller Units (M.U.)

Plasmids	β-gal activity	stdev
pBK1	1090	± 275
pBK1 + pAT18[1]	1107	± 152
pBK1 + pDM4	513	± 139
pBK1 + pDM4-3	279	± 61
pBK1-1 (pBK1 + AT deletion)	146	± 14
pBK1-4	274	± 48
pBK1-4 + pAT18[1]	275	± 46
pBK1-4 + pDM4	61	± 5
pBK1-13	1874	± 411
pBK1-13 + pAT18[1]	2149	± 105
pBK1-13 + pDM4	2129	± 350
pBK1-17	1340	± 360
pBK1-18	439	± 32
pBK1-18 + pAT18[1]	421	± 11
pBK1-18 + pDM4	307	± 15
pBK1-19 (pBK1-18 + AT deletion)	314	± 42
pBK1-21	723	± 12
pBK1-21 + pAT18[1]	736	± 49
pBK1-21 + pDM4	274	± 24
pBK1-24	718	± 32
pBK1-24 + pDM4	288	± 43
pBK1-24 + pDM4-3	258	± 61

[1] pAT18 is the vector control for pDM4

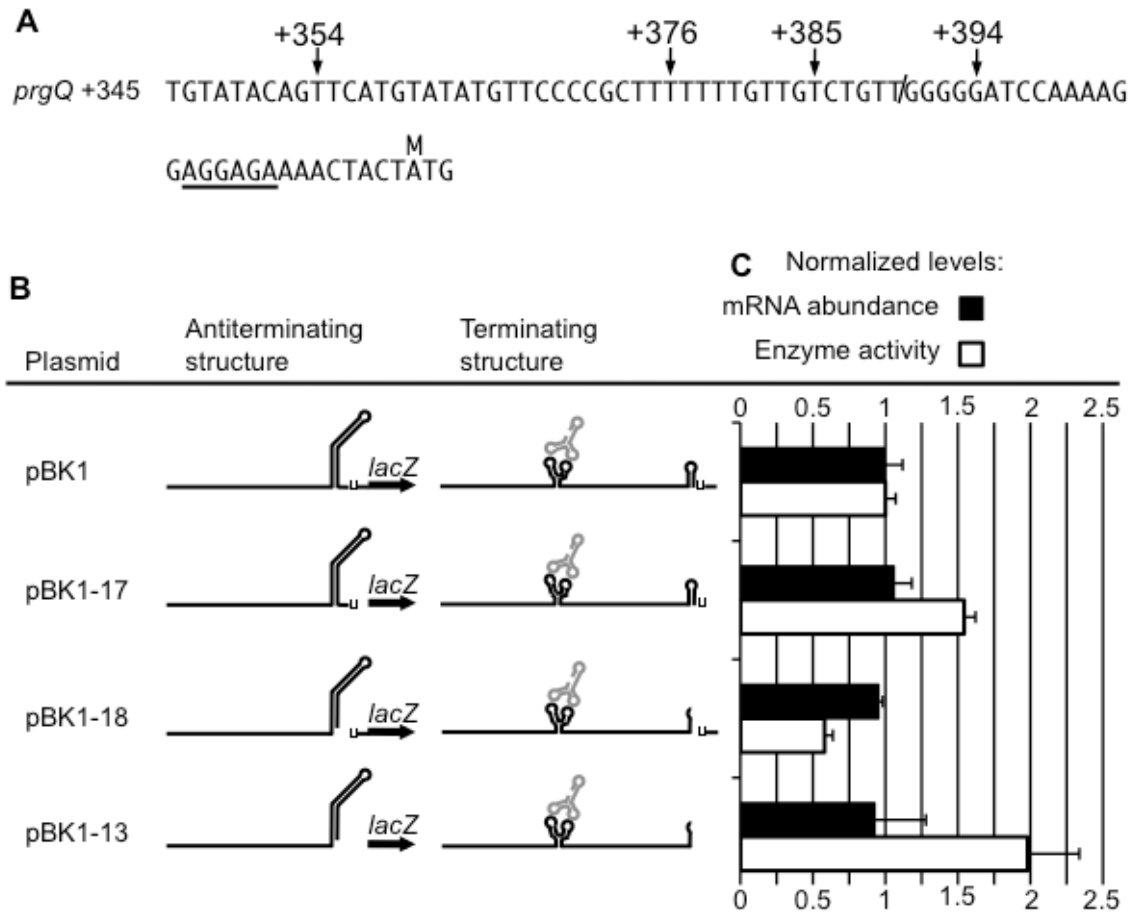

A
+354
+376
+385
+394
prgQ +345 TGTATACAGTTCATGTATATGTTCCCCGCTTTTTTTGTTGTCTGTTGGGGGGATCCAAAAG
M
GAGGAGAAAACTACTATG
B
C Normalized levels:
mRNA abundance
Enzyme activity
Plasmid
Antiterminating structure
Terminating structure
pBK1
pBK1-17
pBK1-18
pBK1-13
lacZ
0 0.5 1 1.5 2 2.5

Figure 11

Figure 11: Effects of altering the *prgQ/lacZ* junction on *lacZ* translational efficiency. **A.** Sequence of the *prgQ*/*lacZ* junction. pCF10 sequences are to the left of the backslash, *lacZ* sequences are to the right. Numbers given are relative to the *prgQ* transcriptional start site. Sites referenced by deletion constructs are indicated by arrows pointing toward the sequence. The *lacZ* RBS is underlined and the initiation codon is indicated. **B.** Reporter plasmids and *prgQ* transcripts with the proposed antiterminating and terminating structures. *prgQ*/*lacZ* is shown in black, Anti-Q is shown in gray. The location of the polyuridine tract is indicated by a 'U', deletions are indicated by gaps. pBK1-17 contains a deletion that disrupts the RBS occluding sequence (Δ385-394), pBK1-18 contains a deletion within IRS1, but has the RBS occluding sequence (Δ354-376), pBK1-13 contains a deletion of the second half of IRS1 and the RBS occluding sequence (Δ354-394). **C.** Transcription past IRS1 is shown as measured by qRT-PCR of *lacZ* transcript abundance and β-galactosidase activity, both normalized to pBK1. *E. faecalis* strains were cultured and tested as detailed in Materials and Methods. pBK1-17 expression of β-galactosidase from *lacZ* mRNA is more efficient than pBK1, indicating that removal of the *lacZ* RBS occlusion sequence derepresses translation. pBK1-13 expression of β-galactosidase from *lacZ* mRNA is more efficient than pBK1-18, also indicating that removal of the *lacZ* RBS occlusion sequence derepresses translation.

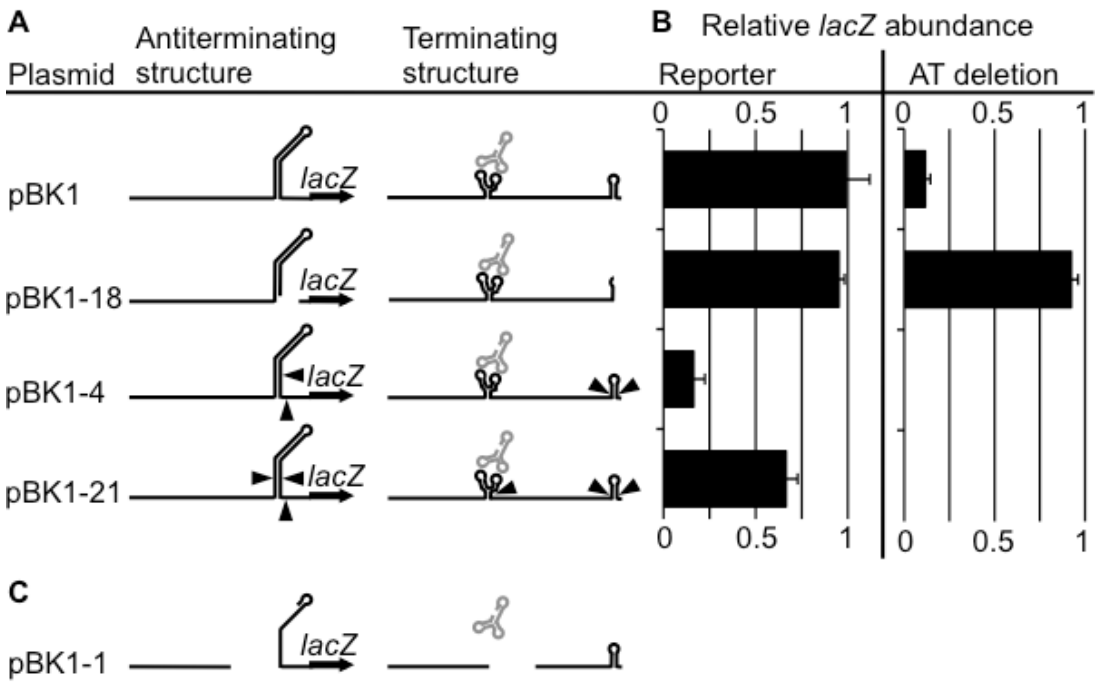


Figure 12: Effects of mutations within *prgQ* on expression of a downstream reporter. **A.** Reporter plasmids and *prgQ* transcripts with the proposed antiterminating and terminating structures. *prgQ*/*lacZ* mRNA is shown in black, Anti-Q is shown in gray. The location of the polyuridine tract is indicated by a 'U', deletions are indicated by gaps and point mutations are indicated by triangles. pBK1 is the wild-type reporter, pBK1-18 contains a deletion within IRS1 (⊗354-376), pBK1-4 contains mutations within IRS1 (339/41GGG:CCC, 369/71CCC:GGG), pBK1-21 contains mutations within the antiterminator and IRS1(197:99CUC:GGG, 339/41GGG:CCC, 369/71CCC:GGG). **B.** Transcription past IRS1 is shown as measured by qRT-PCR of *lacZ*, normalized to a representative pBK1 sample. Strains tested were *E. faecalis* OG1Sp bearing the indicated plasmid or a plasmid that is identical except containing the antiterminator deletion (AT deletion, ⊗120-253). Strains were cultured and tested as detailed in Materials and Methods. **C.** Diagram of proposed antiterminating and terminating structures for pBK1-1, a reporter identical to pBK1 except carrying the AT deletion.

To test this hypothesis, we deleted the 5' side of the antiterminator (Figure 12C, pBK1-1; data for plasmids carrying this deletion are shown in the right-most panel under "AT deletion"). We reasoned that in the absence of the competing antiterminator, the terminator should form readily. *lacZ* transcript abundance from a reporter carrying this mutation dropped to 12% of that from pBK1 (Figure 12B, AT deletion, Δ120-253), supporting the model that the antiterminator precludes termination at IRS1.

If disruption of the antiterminator decreases *lacZ* transcript abundance by allowing IRS1 to act as a terminator, then deletion of IRS1 should suppress the effect of this disruption. To test this, we introduced the antiterminator deletion into pBK1-18, which carries a deletion of the 3' side of IRS1. This construct displayed no change in *lacZ* transcript abundance relative to pBK1-18 (Figure 12, pBK1-18). Taken together, these data indicate that the antiterminator deletion decreases *lacZ* transcript abundance by allowing formation of the terminator.

To further examine the ability of IRS1 to terminate transcription, we swapped 3 nt between the 5' and 3' sides of the IRS1 helix. This mutation (pBK1-4) maintained the structure of the terminator but, because the 5' side of IRS1 participates in formation of the antiterminator, it was predicted to destabilize the antiterminator element. This mutation reduced *lacZ* transcript abundance to 16% of that of pBK1 (Figure 12), consistent with reduced antitermination. To ensure that this change in expression was due to the loss of complementarity within the antiterminator, we constructed plasmid pBK1-21, which is identical to pBK1-4 except that it carries an additional change in the 5' side of the antiterminator that restores complementarity. This construct exhibited an increase in *lacZ* transcript abundance by 290% compared to pBK1-4 (67**%** compared to pBK1), partially suppressing the phenotype displayed by pBK1-4. These data support our model in which IRS1 can act as a terminator, but this is precluded by formation of an upstream antiterminator within the nascent RNA.

Anti-Q enforces termination at IRS1.

We next wanted to examine the direct effects of Anti-Q on nascent *prgQ* transcript termination at IRS1 in the absence of any potential complicating factors such as PrgX control of P_Q. pDM4 contains a segment of pCF10 DNA including P_X and DNA encoding Anti-Q, but not P_Q. Cells containing pDM4 produce abundant Anti-Q, consistent with previous results, allowing us to provide Anti-Q in trans[22].

Using qRT-PCR, we found that providing Anti-Q in trans reduced *lacZ* transcript abundance produced by pBK1 to 28% compared to the level without Anti-Q (Table 5). This supports the argument that Anti-Q acts on nascent *prgQ* transcripts to cause terminator formation. The reporter bearing a deletion of the 3' side of IRS1 was not sensitive to Anti-Q (Table 5, pBK1-18) demonstrating that IRS1 is necessary for Anti-Q to decrease expression of downstream genes. Taken together, these data support the model that Anti-Q interacts with nascent *prgQ* transcripts to allow termination at IRS1.

We also analyzed the impact of Anti-Q on *prgQ* transcript termination when the antiterminator was destabilized by point mutations within the terminator (Table 5, pBK1-4). Because the antiterminator was already destabilized, reducing *lacZ* transcript abundance to 16% compared to pBK1, we expected that Anti-Q would have only a modest ability to further disrupt the antiterminator, making this mutant less sensitive to Anti-Q than the wild-type. Surprisingly, Anti-Q reduced *lacZ* transcript abundance from pBK1-4 to 10% of pBK1-4 alone, a statistically significant difference ($p=0.01$) from the response of pBK1 to Anti-Q, indicating that this mutant is more responsive to Anti-Q. Providing the compensatory mutation within the antiterminator appeared to partially alleviate this increased

Table 5: *lacZ* mRNA levels produced by strains bearing reporter plasmids in the absence and presence of pDM4, which provides Anti-Q.

	lacZ mRNA level[1]	
Plasmid	- pDM4	+ pDM4
pBK1	99 ± 13	27 ± 7
pBK1-18	95 ± 3	97 ± 9
pBK1-4	16 ± 6	2 ± 1
pBK1-21	67 ± 6	11 ± 4

[1] Values are normalized to a representative pBK1 sample

sensitivity to Anti-Q, as Anti-Q reduced *lacZ* transcript abundance produced by this construct to 16% compared to the construct alone, though this change in response was not statistically significant (Table 5, pBK1-21).

An alternate explanation for some of the above results could be that mutations within *prgQ* or interaction between nascent *prgQ* transcripts and Anti-Q cause rapid degradation of the fused *lacZ* transcript. In order to test this, we examined the half-life of *lacZ* transcripts from different strains during log-phase growth by halting transcription with rifampicin and examining the rate of transcript decay as detailed in Materials and Methods. We tested the stability of *lacZ* transcripts produced by the wild-type construct, pBK1, with and without Anti-Q produced in trans. We also tested the stability of *lacZ* transcripts produced by constructs carrying a deletion within the antiterminator (pBK1-1, Δ120-253) and within the 3' half of IRS1 through the *lacZ* fusion (pBK1-13). We found that production of Anti-Q in trans did not decrease the half-life of the *lacZ* transcripts, nor did the presence of the selected mutations (Figure 13). This indicates that the decrease in *lacZ* transcript abundance observed by qRT-PCR is caused by changes in the frequency of termination at IRS1, rather than a decrease in transcript stability.

The role of specific motifs that drive interaction between Anti-Q and nascent *prgQ* transcripts.

Loop I of Anti-Q contains a YUNR motif (sequence CUAA, Figure 14A). YUNR and other "hairpin" motifs have been shown to accelerate RNA/RNA interactions in other systems, and mutations to the "U" residue have been shown to disrupt the YUNR motif[33,73]. A mutation to the U residue of the Anti-Q YUNR loop, however, had no effect on Anti-Q interactions with Qs in vitro or when both RNAs are transcribed from the same template in vivo[55]. We sought to extend these results using our two plasmid system. To do this we cloned the YUNR mutation into both pBK1(generating pBK1-24) and into pDM4, (generating pDM4-3) and tested the ability of both the wild-type and mutant alleles of Anti-Q to reduce the

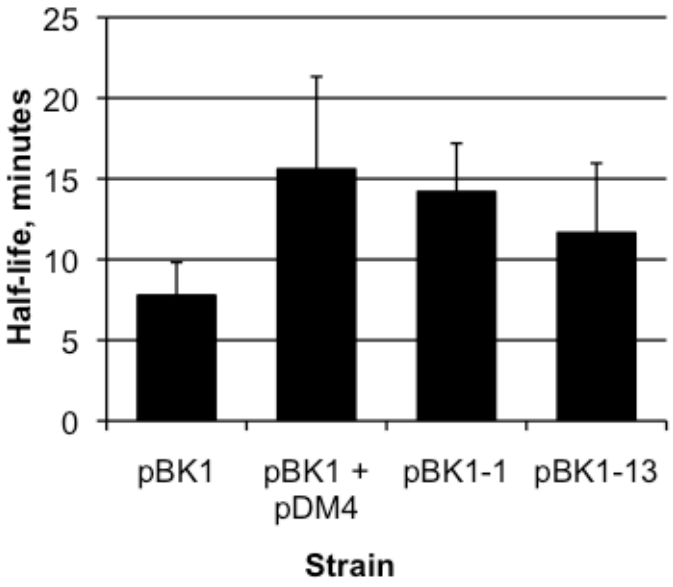


Figure 13: Half-life of *lacZ* transcripts in *E. faecalis* OG1Sp produced by different plasmids. Total RNA was isolated at various time points after transcriptional arrest with rifampicin. Samples were prepared as described in materials and methods to determine the rate of transcript decay. pBK1 produces *lacZ* transcripts fused to wild-type *prgQ* sequence. Production of Anti-Q in trans (pBK1 + pDM4) stabilizes *lacZ* transcripts (p=0.04). pBK1-1 carries the AT deletion (Δ120-253) within *prgQ*. This mutation stabilizes *lacZ* transcripts (p=0.01). pBK1-18 carries a deletion of the 3' half of IRS1 through the *lacZ* junction (Δ354-394) the observed change in stability compared to pBK1 *lacZ* transcripts is not statistically significant (p=0.15).

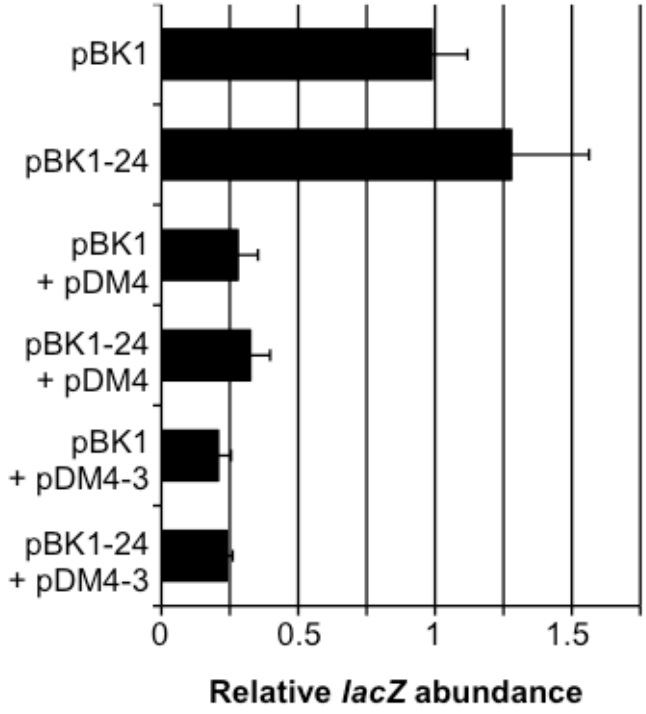


Figure 14: Effects of the YUNR mutation on Anti-Q-mediated attenuation. Transcription past IRS1 is shown as measured by qRT-PCR of *lacZ*, normalized to a representative pBK1 sample. Strains tested were *E. faecalis* OG1Sp bearing the indicated plasmids. pDM4-3 is an derivative of pDM4 with a T:C substitution at the "U" position in the YUNR motif. pBK1-24 is a derivative of pBK1 with a mutation (A215G) that complements the point mutation in pDM4-3. Strains were cultured and tested as detailed in Materials and Methods.

abundance of *lacZ* transcripts produced from pBK1 bearing both wild-type and mutant *prgQ* alleles. (Figure 14). Both the wild-type and mutated pBK1 were equally sensitive to both wild-type and mutated Anti-Q. These results confirm that this YUNR loop does not play a physiological role in Anti-Q-mediated termination at IRS1.

Unlike loop I, loops II and III of Anti-Q were found to be protected from lead-acetate cleavage by interaction with Qs, suggesting that these act as sites of interaction between the two RNAs[55]. We hypothesized that the region of Qs complementary to Anti-Q formed a binding platform, containing stem-loops complementary to loops II and III of Anti-Q and that Anti-Q mediated its effects by interacting with these loops.

To test this we mutated residues within loops II and III and the supporting stems of Anti-Q and the complementary nucleotides of Qs (Figure 15A and B). Mutations within the stems were chosen so as to disrupt the stem structure within Qs, but not within Anti-Q. We then used our trans-plasmid system to test the ability of the Anti-Q alleles to attenuate expression of the transcriptionally fused *lacZ* gene in pBK1 plasmids that were wild-type or carried complementary mutations to the binding platform (Figure 15C). We found that the mutations tested fell into four distinct classes: major pairing residues, minor pairing residues, structural residues and irrelevant residues. For major pairing residues, a mismatch between the RNAs prevented Anti-Q from attenuating β-galactosidase expression, regardless of which RNA carried the mutation (mutants 30, 31, 32, 36, 37). This defect could be corrected by introducing the mutation to both RNAs. For minor pairing residues a mismatch between the RNAs prevented Anti-Q from attenuating β-galactosidase expression if the mutation was within Anti-Q, but was tolerated when the mutation was within Qs (mutations 27, 28, 29, 35, 38). This defect could be corrected if both RNAs

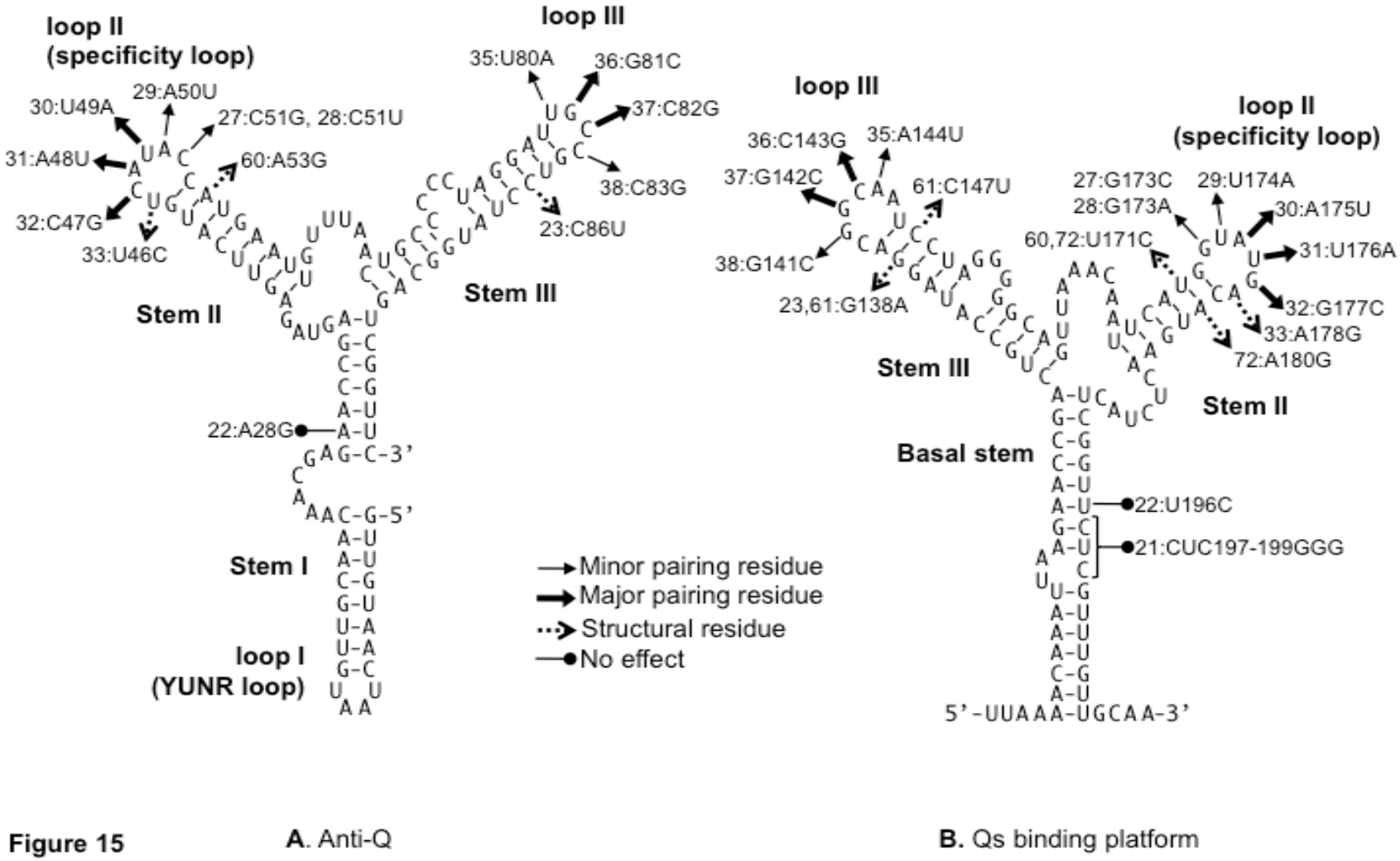


Figure 15 **A**. Anti-Q **B.** Qs binding platform

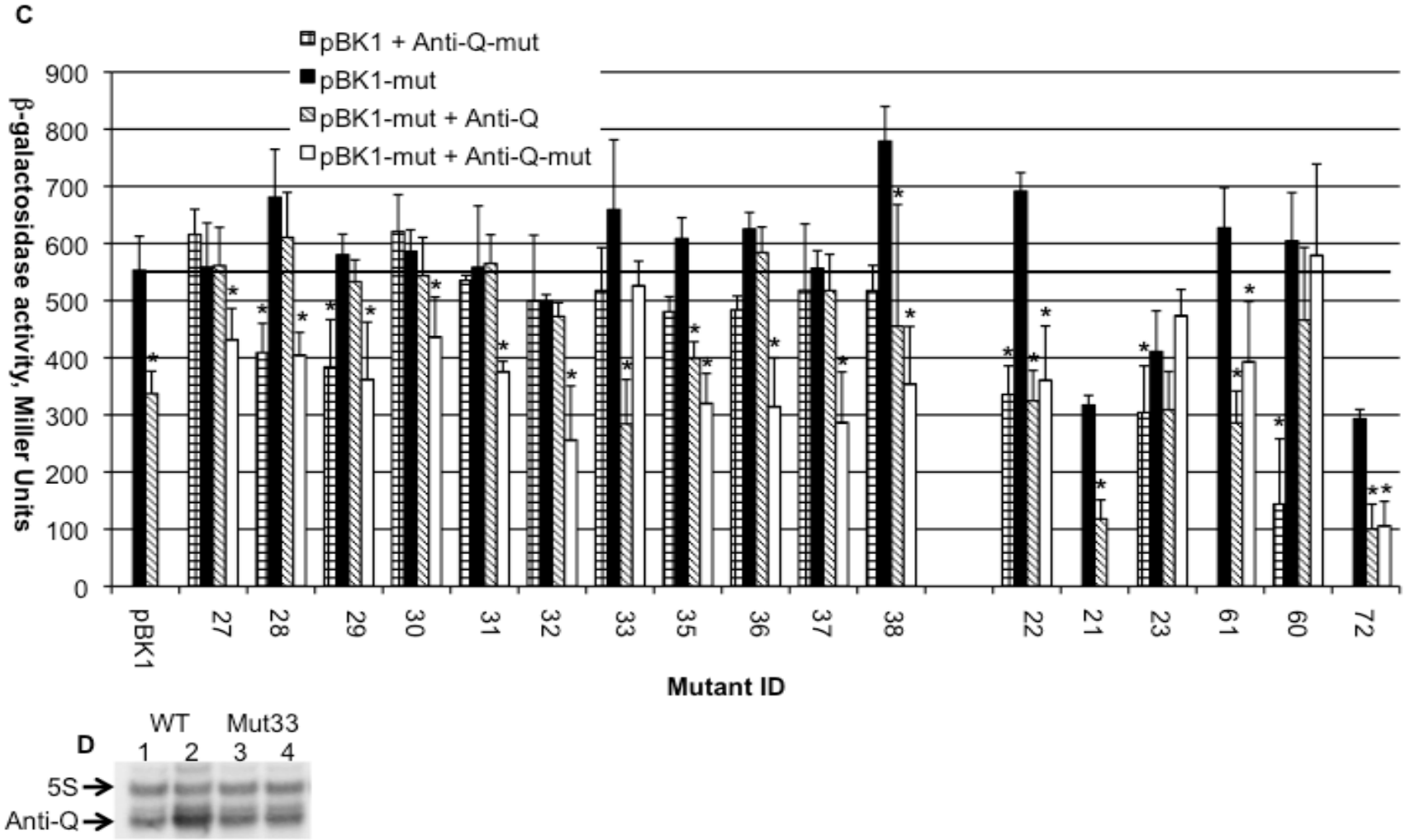

C
pBK1 + Anti-Q-mut
pBK1-mut
pBK1-mut + Anti-Q
pBK1-mut + Anti-Q-mut
β-galactosidase activity, Miller Units
0
100
200
300
400
500
600
700
800
900
pBK1
27
28
29
30
31
32
33
35
36
37
38
22
21
23
61
60
72
Mutant ID
D
WT
Mut33
1
2
3
4
5S
Anti-Q

Figure 15

Figure 15: Genetic analysis of interaction between kissing loops on Anti-Q and nascent *prgQ* transcripts. **A**. The structure of Anti-Q and **B**. the predicted structure of the *prgQ* binding platform. The location of the mutations tested in the experiment, as described in the text, are indicated by arrows pointing away from the sequence. Each mutation is identified by a number and the sequence substitution made. The arrows indicate how the residue contributes to Anti-Q effects on *prgQ* termination at IRS1. Thin arrows indicate that disrupting complementarity at that residue between the RNAs sometimes prevented Anti-Q from significantly increasing termination at IRS1; thick arrows indicate that disrupting complemetarity always prevented Anti-Q from significantly increasing termination at IRS1; dotted arrows indicate that the residue serves a structural role in the *prgQ* binding platform; arrows with dots for heads indicate that the mutation had no effect on interactions between *prgQ* and Anti-Q. **C**. β-galactosidase expression of *E. faecalis* OG1Sp transformed with pBK1 ± pDM4 carrying the indicated mutations. An asterisk indicates that the Anti-Q allele tested significantly decreases β-galactosidase expression from pBK1 carrying the indicated *prgQ* allele.

carried the mutation, indicating that the effect was due to the mismatch, rather than a general defect of the mutant Anti-Q. For structural residues, a mutation within one RNA prevented Anti-Q from attenuating gene expression, but this was not corrected when both RNAs carried the mutation (23, 33, 60). For structural residues within stems, the defect could be corrected by introducing a compensatory mutation to the other side of the proposed helix, restoring base-pairing within the stem (mutations 61, 72). For irrelevant residues, the mutations tested did not interfere with Anti-Q–driven attenuation of β-galactosidase expression (mutations 21, 22).

In general, major pairing residues occurred within the middle of loops, minor pairing residues occurred toward the edges of loops, adjacent to the stems, and structural residues occurred within stems II and III. Mutations to the basal stem in Qs were irrelevant (Figure 15). This suggests that stems II and III form a binding platform in nascent *prgQ* transcripts that is necessary for interaction with Anti-Q, but that the basal stem is not necessary for this function.

An exception to these generalizations is Anti-Q +46, tested by Anti-Q_{mut33} (*prgQ* A178G, Anti-Q U46C). This mutation prevented Anti-Q from attenuating gene expression in pBK1 even if the binding platform carried the complementary mutation. Furthermore, pBK1 carrying this mutation was responsive to wild-type Anti-Q. This suggests that this residue serves a structural function in Anti-Q. Northern blots show that Anti-Q_{mut33} is as abundant as wild-type Anti-Q (Figure 15D), indicating the effect is not due to a loss of stability. Anti-Q_{mut33} has the same secondary structure as wild-type Anti-Q, as determined by lead-acetate probing (Personal communication, Keith Weaver) indicating that the mutation does not cause unexpected changes in the overall structure of Anti-Q. The context of this residue, normally a “U”, suggests that it could be a part of a “UNR” motif. Such motifs have been shown to enhance interaction with complementary RNAs in the same way as YUNR motifs[34]. However, when compared to wild-type

Anti-Q using gel shift assays Anti-Q_{mut33} is not defective for interaction with either wild-type Qs or Qs bearing the complementary mutation (Keith Weaver, personal communication).

Countertranscript-driven attenuation functions in vitro.

To confirm our interpretations of the in vivo studies and resolve discrepancies between these and in the in vitro gel-shift experiments examining the possible UNR loop in Anti-Q we examined termination at IRS1 using in vitro transcription (IVT) assays (Figure 16). We used a supercoiled (plasmid) template containing an inactivated P_X to examine termination at IRS1 in the presence and absence of Anti-Q. We found that in the absence of Anti-Q, 45% of *prgQ* transcripts terminated at IRS1 (lane 1). Anti-Q increased the frequency of termination to 91% (lane 2) while an irrelevant RNA of similar size did not effect frequency of termination (lane 3). These data support the model that nascent *prgQ* transcripts contain a transcriptional terminator at IRS1, and Anti-Q interaction with these transcripts enforces termination. It has been reported that some terminators function less efficiently in vitro when transcribed from supercoiled (plasmid) templates than from linear templates[74]. We found that termination at IRS1 occurred with equal frequency using either supercoiled or linear templates generated via PCR and that Anti-Q was equally effective at enforcing termination of RNAs transcribed from either template (Compare lanes 1-3 with 4, 5, and 7). This demonstrates that this system is not sensitive to supercoiling of the DNA template. We also tested if disrupting the proposed UNR loop in Anti-Q (Mut33: U46C) impacts Anti-Q enforced termination in vitro. We found that wild-type and mutant Anti-Q were equally effective at enforcing termination of both wild type *prgQ* transcripts or those carrying the complementary mutation (compare lanes 4-7 with 8-11).

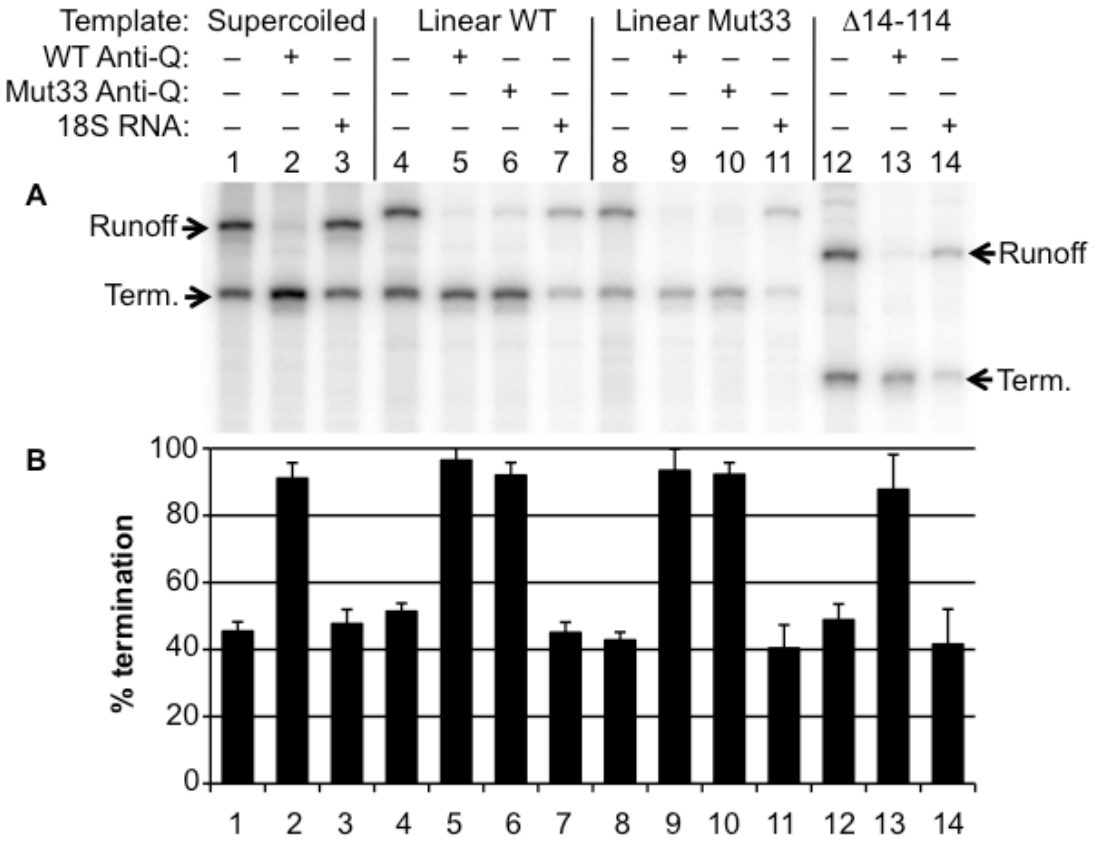


Figure 16: Termination and antitermination occur in vitro. **A.** IVT reactions performed using *E. faecalis* RNAP to generate *prgQ* transcripts from the indicated template. Supercoiled: pCJ5-25 plasmid; linear WT, a PCR product amplified from pBK1-25 carrying P_Q -176 to +512; linear Mut33, a PCR product identical to the linear WT except carrying Mut33 as described in the text and shown in Figure 14; Δ14-114, a PCR product identical to the linear WT except carrying a deletion from P_Q +1 to +114. Reactions were performed with or without the indicated RNA. Transcripts terminated at IRS1 are indicated (Term.) as are runoff (R.O.) transcripts that extend to the end of the DNA template, or to a downstream IRSX terminator, in the case of the supercoiled template. **B.** Termination at IRS1 as determined by dividing the intensity of the IRS1 band over the sum of the intensities of the IRS1 and the runoff bands. The results of three independent IVT assays were analyzed. Error bars indicate one standard deviation from the mean.

Structural analysis of *prgQ* sequences suggested that sequences upstream of the binding platform do not contribute to formation of the binding platform, antiterminator or terminator structures in nascent *prgQ* transcripts. We hypothesized that sequences upstream of the binding platform were dispensable for correct function of countertranscript-driven attenuation. In vivo experiments suggested that deletions to this region impacted the compatibility or stability of the plasmids used to provide Anti-Q and *prgQ* transcripts in trans so we pursued this question using IVT assays. We cloned a deletion of *prgQ* +14 to +114 in pBK1-25, which contains a mutated P_X, generating pBK1-57. We used this plasmid as a template for PCR to generate a linear template for IVT reactions. We found that termination of transcripts at IRS1 occurred with the same frequency from this template, as a wild-type template and that Anti-Q caused the same increase in termination frequency, whereas an irrelevant RNA did not (Figure 16, lanes 12-14). This indicates that sequences in *prgQ* upstream of the binding platform are dispensable for correct function of the countertranscript-driven attenuation system.

Qs regulation of the *prgX* operon.

Cells bearing pCF10 produce X RNA, a 1.4 kb transcript initiated at the *prgX* promoter that encodes PrgX. Upon pheromone induction, X RNA levels drop dramatically and PrgX protein levels drop slightly[20]. We have demonstrated that the *prgQ* operon negatively regulates *prgX* products in cis and in trans (Figure 6); however, in this experiment the *prgX* operon was truncated and did not include *prgX* sequences, nor did it distinguish between effects of Qs RNA and the PrgQ peptide. Additionally, it was not shown if the *prgX* target was a mature RNA, nascent transcript or the *prgX* promoter. We wanted to better understand the trans-acting mechanisms. To test if the target was P_X, we transcriptionally fused *lacZ* to P_X +16, generating pCJ3. We chose +16 because this should be sufficient distance from the transcriptional start site for RNAP to transition from an initiation complex to an elongation complex, but short enough to minimize RNA-RNA

interactions with Qs, ensuring that any effects are dependent on Qs action on the native promoter. We transformed OG1Sp with pCJ3, with and without pDM5-25, which provides Qs in trans, and tested these strains for β-galactosidase expression, using a Miller assay. β-galactosidase expression from our reporter without Qs was 738 ± 23 Miller units and with Qs was 777 ± 242 Miller units, indicating that Qs does not regulate initiation of transcription at P_X. We hypothesized that the RNA Qs negatively regulates the expression of *prgX* products via a mechanism involving interaction between complementary sequences on the RNAs.

To test if Qs regulates PrgX protein expression, we provided transcripts from the *prgX* and *prgQ* operons separately from compatible plasmids (Figure 5). We generated *E. faecalis* strains bearing the plasmid that transcribes the *prgX* operon, pBK2-26, with and without the plasmid that transcribes Qs, pDM5-25. We then performed a western blot using a PrgX-specific antibody on whole cell lysates from these strains (Figure 17A). We found that providing Qs in trans reduced PrgX expression to 77% ± 10% (compare lane 2 to lane 1) whereas the empty vector did not (lane 3)

In order to test Qs effects on X RNA and Anti-Q we prepared RNA from these strains and analyzed it using northern blots. When we hybridized blots using a probe specific for the *prgX* ORF we detected a single 1.4 kb band in the absence of Qs (Figure 17B). This band was the correct size to be a transcript that initiated at P_X and terminated at IRSX, the *prgX* terminator. When Qs was provided this band diminished in intensity and a second, ~1.2 kb band was detected. A probe specific for 99 nt of the 5' end of X RNA hybridized to the 1.4 kb band, but not the ~1.2 kb band (Figure 17C). To test if this phenomenon was dependent on PrgX,

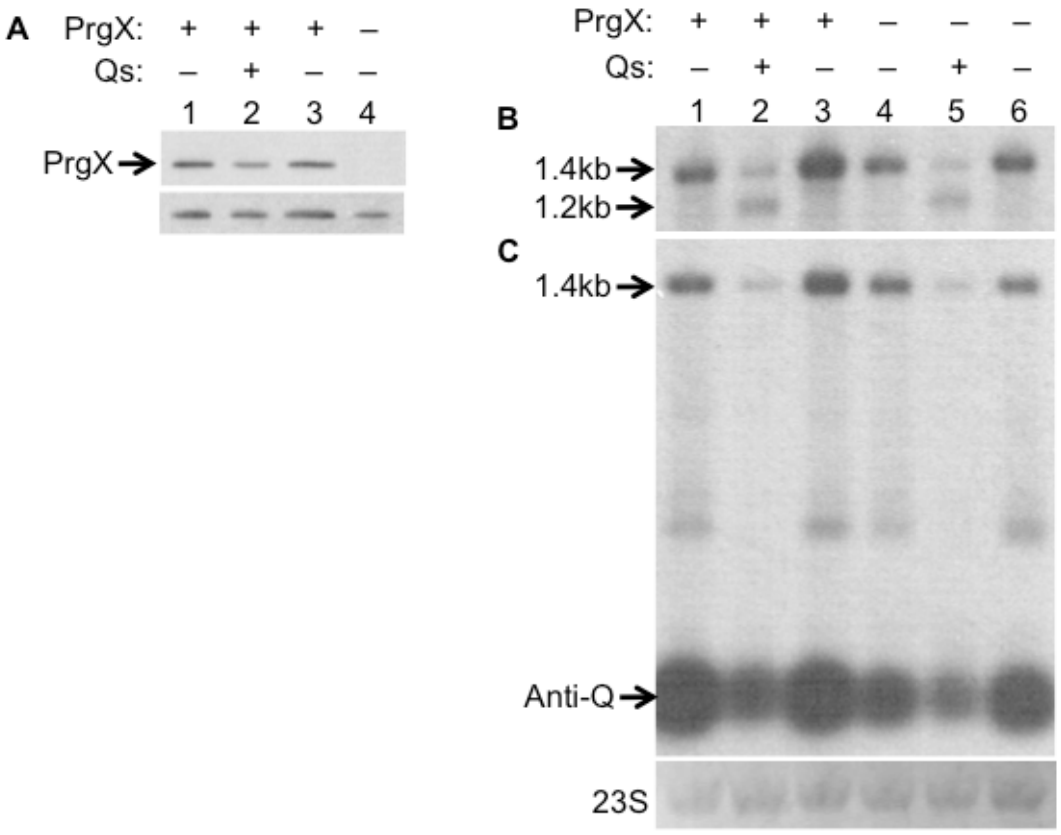


Figure 17: Qs negatively regulates expression of PrgX and *prgX* transcripts. Qs and *prgX* transcripts were provided in trans from compatible plasmids transformed into OG1Sp. Products of the following strains were prepared and examined by blot. Lane 1, pBK2-26; lane 2, pBK2-26 + pDM5-25; lane 3, pBK2-26 + pAT18; lane 4, pBK2-126; lane 5, pBK2-126 + pAT18; lane 6, pBK2-126 + pDM5-25. pBK2-26 produces X RNA with a wild type *prgX* open reading frame (ORF). pBK2-126 produces X RNA with missense mutations in codons 3 and 4 of the *prgX* ORF. pDM5-25 produces Qs. pAT18 is the empty vector control for pDM5-25. **A**. A western blot on total cell lysates using an antibody against PrgX. The bottom panel shows an unidentified, cross-reactive host protein that serves as a loading control. **B,C**. RNA was prepared and analyzed by Northern blots hybridized to a probe specific for the *prgX* ORF (**B**) or the first 99 nt of X RNA (**C**). The loading control is an ethidium bromide stain of 23S.

we introduced nonsense mutations into the third and fourth codons of the *prgX* gene and found this had no effect (Figure 17B). Further examination of this blot revealed that Qs also reduced the abundance of Anti-Q, whereas the empty vector did not. Taken together these data indicate that Qs causes differential expression of *prgX* transcripts, reducing Anti-Q and generating a second *prgX* message that is missing the 5' terminus of full-length X RNA, and that PrgX is dispensable for these phenomenon.

We hypothesized that the ~1.2 kb band was generated by nucleolytic removal of about 200 nt from the 5' terminus of the 1.4 kb band. In order to test this we used a coupled 5' 3' RACE protocol to identify both termini of single X RNA transcripts. We also used Tobacco Acid Pyrophosphatase (TAP) to distinguish 5' termini generated as a result of *de novo* RNA synthesis from those derived by RNA processing.

In the presence of Qs, most of the 5' termini clustered in two distinct regions (Figure 18). One was the *prgX* transcriptional start site, a pair of closely spaced G residues[20]. These 5' termini were detected only in the TAP-treated sample (δ). The other region was about 205 nucleotides downstream, within sequences complementary to Qs. Termini in this region were identified in TAP-treated as well as untreated samples (γ), indicating that they were generated as a result of processing, rather than transcript initiation. In the absence of Qs the most commonly detected 5' terminus was the *prgX* transcriptional start site, detectable only in the TAP-treated sample (β). In the sample not treated with TAP (α), most sequences tested did not contain a bona-fide *prgX* 5'-3' junction, but rather non-specifically amplified sequence. The 3' termini from all samples generally fell within IRSX, though recessed 3' termini were detected. Taken together, these data indicate that the smaller RNA is generated by removal of about 200 nt from

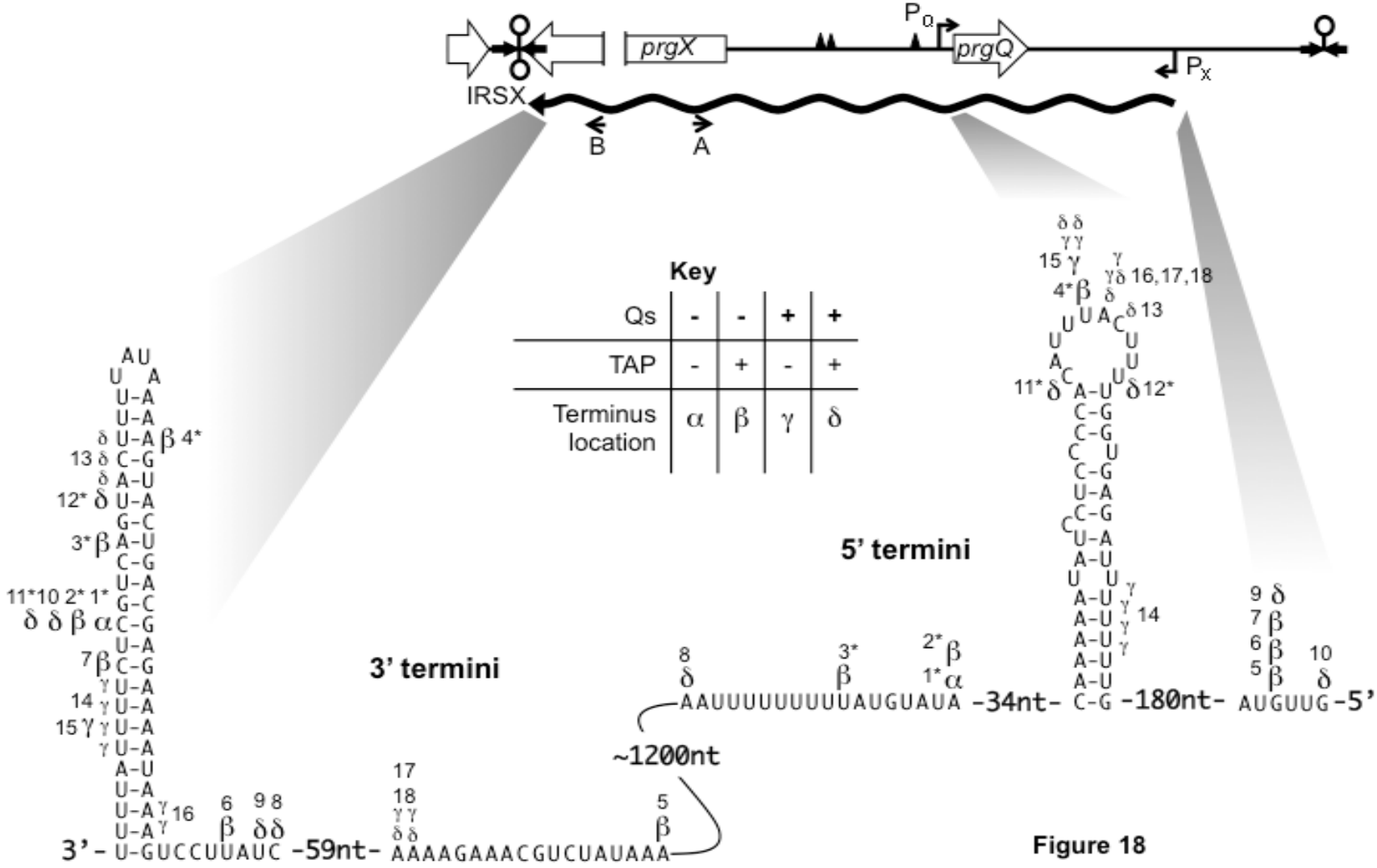

P$_Q$
prgX
prgQ
P$_X$
IRSX
B
A
Key
Qs - - + +
TAP - + - +
Terminus location α β γ δ
5' termini
3' termini
AAUUUUUUUUUUAUGUAUA -34nt- C-G -180nt- AUGUUG-5'
~1200nt
3'- U-GUCCUUAUC -59nt- AAAAGAAACGUCUAUAAA

Figure 18

Figure 18: Qs directs removal of ~200 nt from the 5' terminus of X RNA. The 5' and 3' termini of single X RNA transcripts were determined using a coupled 5'-3' RACE protocol. RNA was divided into aliquots to be treated with TAP to remove 5' triphosphate groups or mock treated, then circularized with T4 RNA ligase. cDNA was generated by reverse transcription using Superscript III primed with oligonucleotide A, then PCR amplified using oligonucleotides A & B. The reactions were ligated into a vector and transformed into *E. coli*. Vector insertions were PCR amplified from individual colonies and the *prgX* 5'-3' junctions were sequenced. A map of the region is shown, with the full length X RNA transcript drawn as a wavy line. Labeled arrows indicate the relative positions of the oligonucleotides used. Shaded triangles indicate the relative position of sequences along the X RNA transcript. A partial sequence and partial proposed structures of the 5' and 3' termini of X RNA are shown, along with the number of nucleotides in any gaps. The mapped 5' and 3' termini are indicated by letters that correspond to the RNA sample from which that transcript was derived (see key and text). Numbers next to these letters indicate corresponding 5' and 3' termini identified from single transcripts. Asterisks (*) next to numbers indicate that 1-6 nt of ectopic sequence were inserted at the junction between the 5' and 3' termini.

the 5' end of full-length X RNA, rather than the activity of a second promoter, and that there is no coordinated processing of the 3' terminus.

RNase III cleaves X RNA in a Qs-directed fashion.

Because X RNA is processed in a region that is complementary to Qs, we hypothesized that RNase III is responsible for the processing. *E. faecalis* V583 has one annotated gene that codes for RNase III: *rnc* (EF3097). The predicted protein sequence of this gene is 39% identical to *E. coli* RNase III and is conserved in *E. faecalis* OG1RF. BLAST searches of the OG1RF genome using the predicted protein sequence of EF3097 and the protein sequence of *E. coli* RNase III did not produce any other alignments with an E value below 1.5, indicating that these alignments are not biologically meaningful and that there is one copy of *rnc* encoded in the *E. faecalis* chromosome. We generated an in-frame deletion of *rnc* in *E. faecalis* OG1RF, generating the strain OG1RFΔ3097. Under the conditions tested, this strain was not defective for growth in broth culture (Figure 19).

To test if RNase III cleaved X RNA in a Qs-dependent fashion we transformed OG1RFΔ3097 with the X RNA producing plasmid with and without the Qs producing plasmid. We also provided RNase III in trans from a plasmid that transcribed *rnc* under control of P_{23}, a constitutive promoter. We found that in the absence of RNase III, Qs did not direct processing of X RNA (Figure 20A, lanes 3 and 4). When RNase III expression was complemented in trans, Qs–directed processing of X RNA was restored (lanes 5-8). X RNA was expressed at substantially higher levels under these conditions (lanes 5 and 6). The reason for this is unclear, thought it could indicate that the his-tagged RNase III enzyme used for complementation is less efficient at cleaving X RNA than the wild-type allele and is protecting X RNA from alternate degradation pathways.

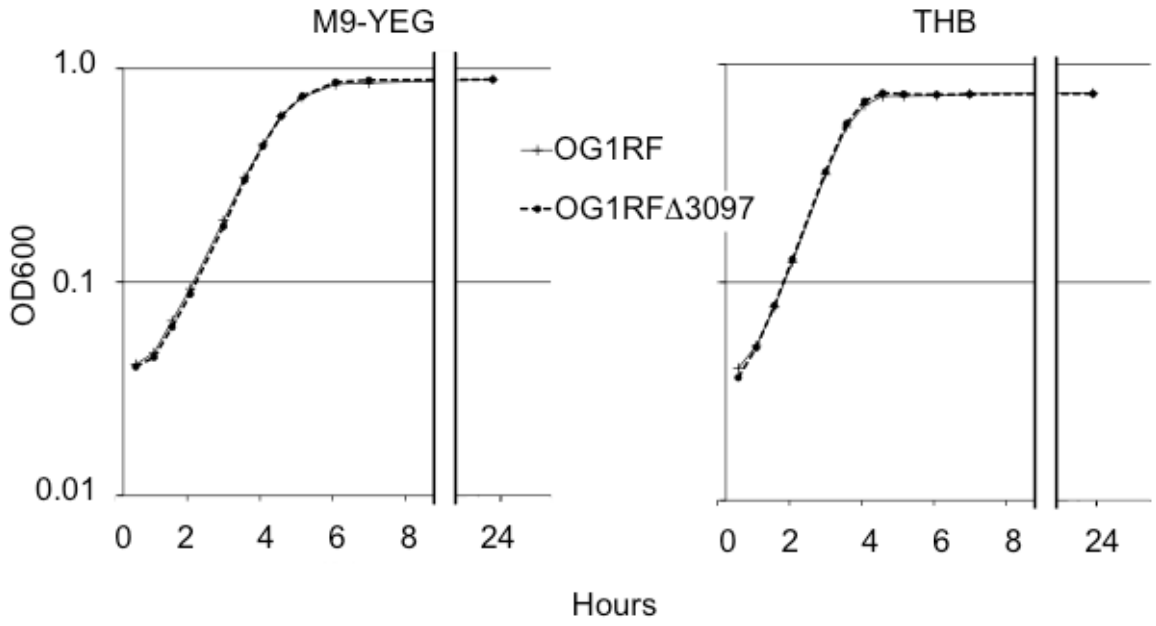


Figure 19: OG1RFΔ3097 is not defective for growth in broth culture. Growth curves for OG1RF and OG1RFΔ3097 were performed. Three cultures were grown for each strain in both M9-YEG (a minimal medium) and THB overnight at 37°C. Cultures were diluted in fresh medium to an OD600 of 0.05, 200μl from each culture was aliquoted into 3 wells of a 96 well plate and cultured at 37°. The OD600 was determined at various time points using a Modulus Microplate reader (Turner Biosystems). Average OD600 values of the cultures are shown with error bars representing standard deviation.

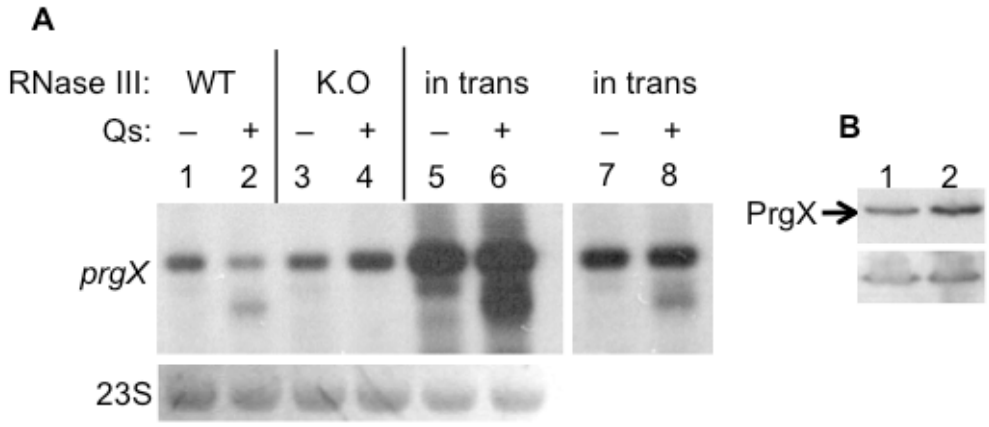


Figure 20: RNase III is necessary for Qs-directed cleavage of X RNA in vivo. **A.** RNA was prepared from the indicated strains of *E. faecalis* with pBK2-126, producing X RNA, with (+) or without (–) Qs from pDM5-25. RNase III is either present on the chromosome (WT), deleted from the chromosome (K.O.) or deleted from the chromosome, but provided from pCJ9:His-EF3097 (in trans). This RNA was analyzed by northern blot using a probe specific for the *prgX* open reading frame. Lanes 7 and 8 are identical to lanes 5 and 6 except the film was exposed for a shorter duration to allow better visualization of the bands. The loading control is an ethidium bromide stain of 23S. **B.** A western blot was performed on total cellular protein from the OG1RFΔ3097 bearing pBK2-26 providing X RNA with a wild type *prgX* ORF with and without pDM5-25 providing Qs. The bottom panel shows an unidentified, cross-reactive host protein that serves as a loading control.

We then tested if RNase III was necessary for Qs to repress PrgX expression. We performed a western blot on lysates from OG1RFΔ3097 bearing an X RNA-producing plasmid with a wild-type *prgX* ORF with and without a Qs producing plasmid. We found that in the absence of RNase III, Qs did not repress PrgX expression (Figure 20B). In fact, X RNA and PrgX expression appeared to increase in response to Qs (2.1 ± 1.2 and 1.6 ± 0.1 fold, respectively). It is possible that in the absence of RNase III, formation of a duplex with Qs protects X RNA from alternate degradation pathways. This conjecture was not explored further. These data demonstrate that RNase III is necessary for Qs–directed processing of X RNA and repression of PrgX.

To confirm that RNase III mediates cleavage of X RNA we fused a 6-histidine tag to the N-terminus of *E. faecalis* RNase III, similar to a method used to purify *S. aureus* RNase III[75] and purified His-tagged RNase III from *E. faecalis* cell lysates (Figure 21). We then tested this preparation for its ability to cleave both X RNA and Qs in the presence and absence of the other RNA in vitro. We found that RNase III did not cleave X RNA or Qs alone, but cleaved both RNAs at several locations in the presence of the complementary transcript (Figure 22A). We found that 5' labeled cleavage products were about 200 nt or less in length, indicating that cleavage took place within the region of complementarity between the RNAs. These data indicate that RNase III cleaves X RNA when directed by Qs.

We tested if RNase III cleaved X RNA in vitro at processing sites identified in vivo. To do this we used RNase III to cleave unlabeled X RNA in the presence and absence of Qs and used primer extension to map the cleavage sites (Figure 22B). In the absence of Qs, we could not detect RNase III-mediated cleavage of X RNA. When Qs was added without RNase III, primer extension generated numerous new bands within the region of complementarity, indicating that the

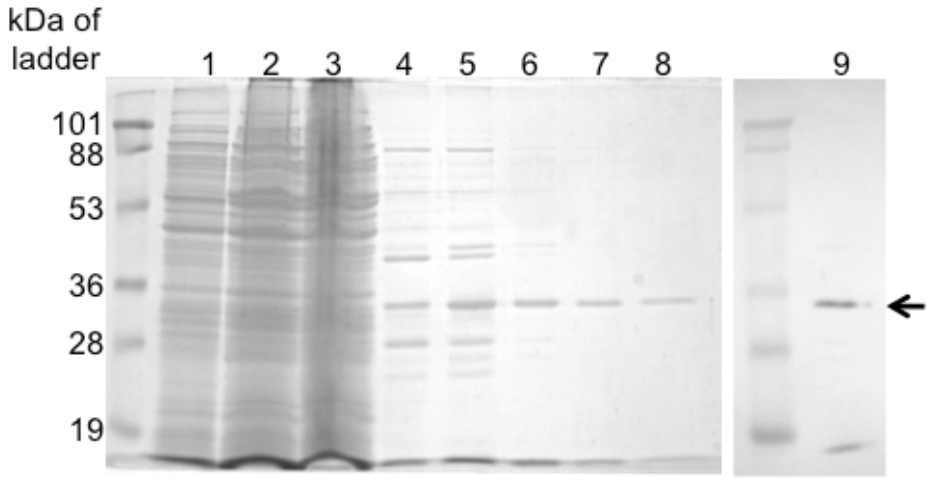


Figure 21: Preparation of RNase III. His-tagged RNase III was prepared from *E. faecalis* using a cobalt column as described in Materials and Methods. The contents of the lanes are: lane 1, pellet from cleared lysate; lane 2, cleared lysate; lane 3, cleared lysate after application to column; lanes 4-8, elution fractions. The fractions shown in lanes 6-8 were collected for concentration over an Amicon column with a 3 kDa MWCO. This preparation was analyzed by SDS-PAGE, as above (lane 9) and the identity of the band indicated by the arrow was confirmed to be RNase III by western blot using an antibody against the His-tag (not shown). The identity of the band that ran with the ion front is not known, though N-terminal sequencing suggested that it contained a mixture of peptides.

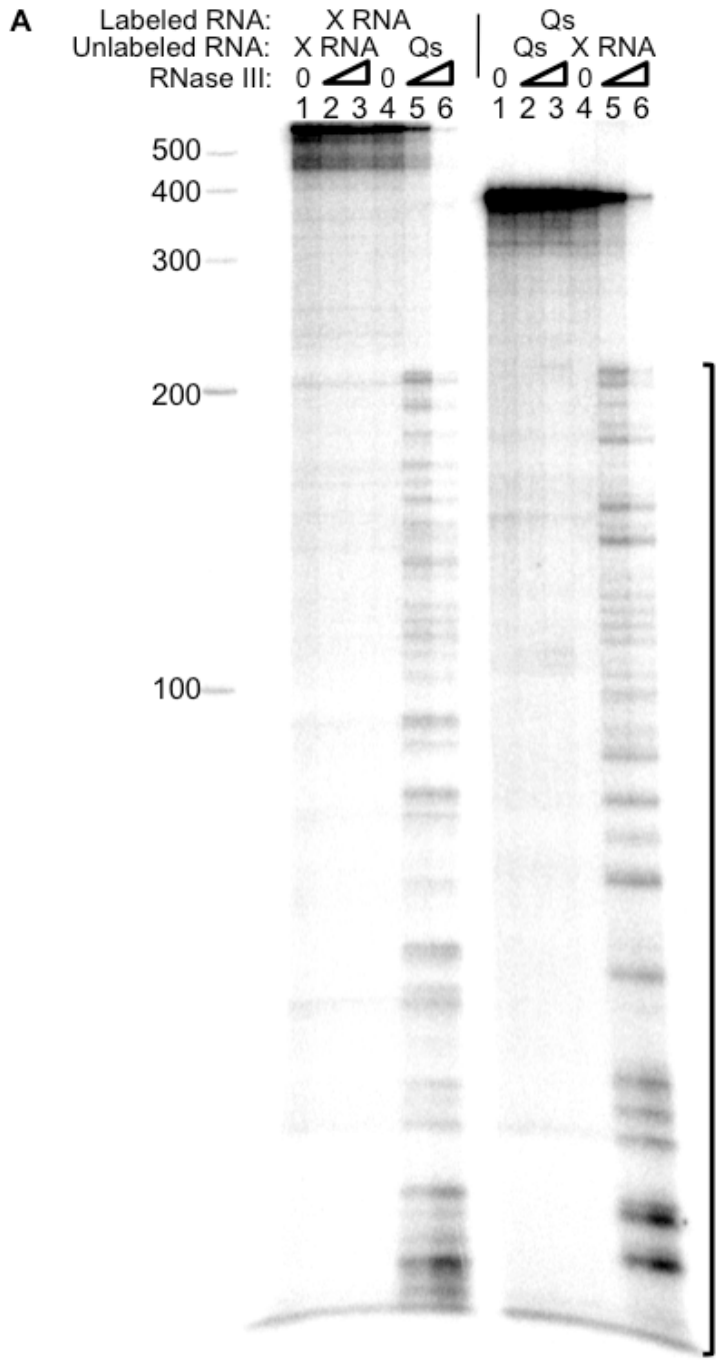

A
Labeled RNA:
X RNA
Qs
Unlabeled RNA:
X RNA
Qs
Qs
X RNA
RNase III:
0
0
0
0
1 2 3 4 5 6
1 2 3 4 5 6
500
400
300
200
100

Figure 22

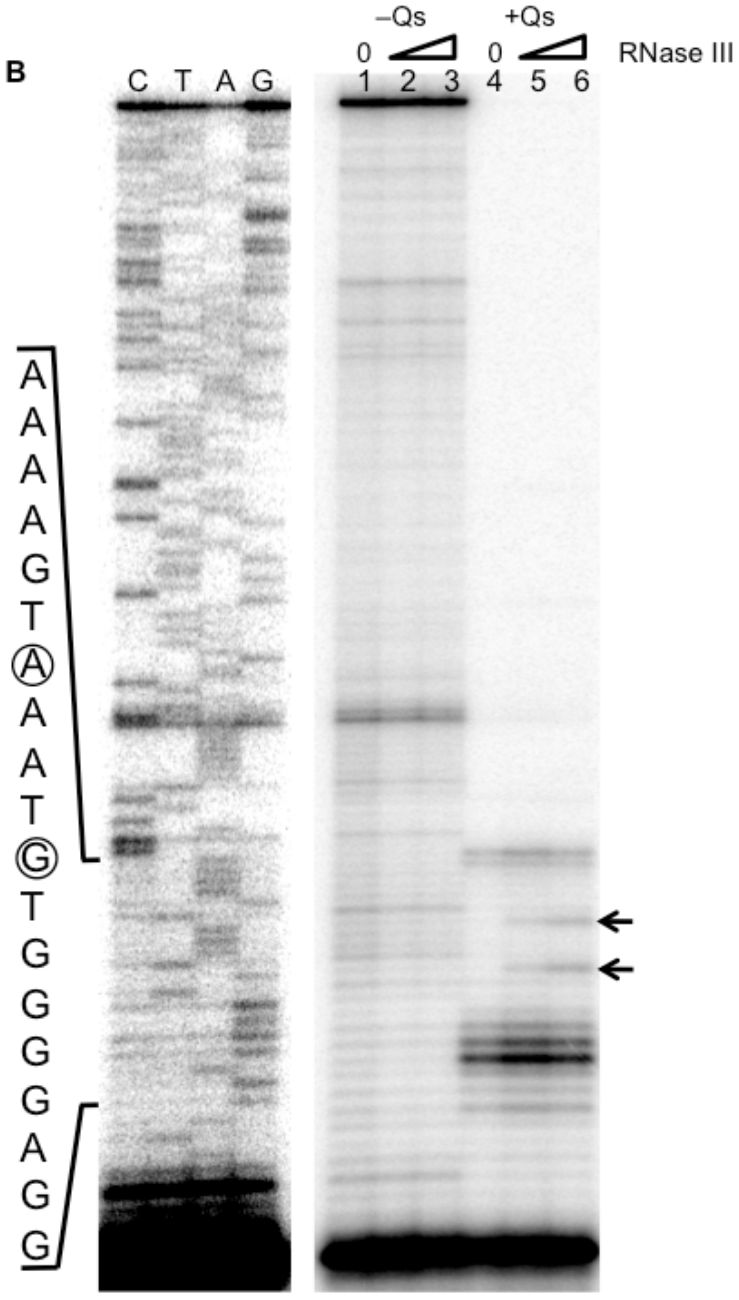


Figure 22

Figure 22: RNase III mediates Qs-directed cleavage of X RNA in vitro. **A.** Full length Qs and a 550 nt 5' fragment of X RNA were transcribed in vitro. One RNA was 5' labeled with ^{32}P and incubated with or without the unlabeled antisense RNA and with or without increasing levels of RNase III. The products were electrophoresed on a denaturing polyacrylamide-urea gel and examined using a phosphorimager. A square bracket indicates products digested in the overlapping region between the RNAs. **B.** Primer extension reaction done on an unlabeled 5' X RNA fragment incubated with and without Qs and with and without RNase III. The DNA oligonucleotide was 5' labeled with ^{32}P and used to prime a reverse transcription reaction. Products were electrophoresed on a denaturing polyacrylamide-urea gel and examined using a phosphorimager. A chain termination ladder is shown at the left. The ladder is from the same gel as the extensions, but with contrast and brightness adjusted to visualize the bands.

reverse transcriptase had difficulty extending through the RNA duplex. When RNase III was added two additional bands appeared, both mapped to 5' termini identified by in vivo experiments. These data indicate that RNase III is responsible for the processing of X RNA observed in vivo.

The 5' end of Qs directs X RNA processing.

To better understand how Qs directs processing of X RNA, we performed a deletion analysis of Qs. We cloned the deletions into pDM5-25 and transformed OG1Sp with these plasmids and a plasmid that produced X RNA. We then used northern blots to analyze processing of X RNA. Deleting the 5' third of Qs prevented processing as judged by the intensity of the 1.4 kb band (Figure 23, lane 3) but deleting the remainder did not prevent processing (lane 4). We found that Qs containing only the extreme 5' terminus of Qs, which is complementary to the region of processing, could direct processing of X RNA (lane 5). However, deletion of this region, which contains the *prgQ* RBS and initiation codon, did not prevent processing (lane 6), though the two X RNA bands appeared to be less well separated, suggesting that the processing site had shifted upstream. To test if there was a secondary site necessary to direct ectopic processing we made mutations that deleted the 5' region and different parts of the remaining Qs sequence. These mutants directed processing, though apparently at a variety of locations (lanes 7-9). Taken together, these data indicate that the only aspect of Qs necessary to direct processing of X RNA is a sufficient length of Qs between +1 and +110, which is entirely complementary to X RNA. Qs sequences that are complementary to X RNA but outside this region do not direct processing. Furthermore, translation of *prgQ* is not necessary for processing of X RNA.

We used sfold[76] (http://sfold.wadsworth.org) to analyze the region of X RNA processing and identified a possible stem-loop structure (Figure 18). To test if

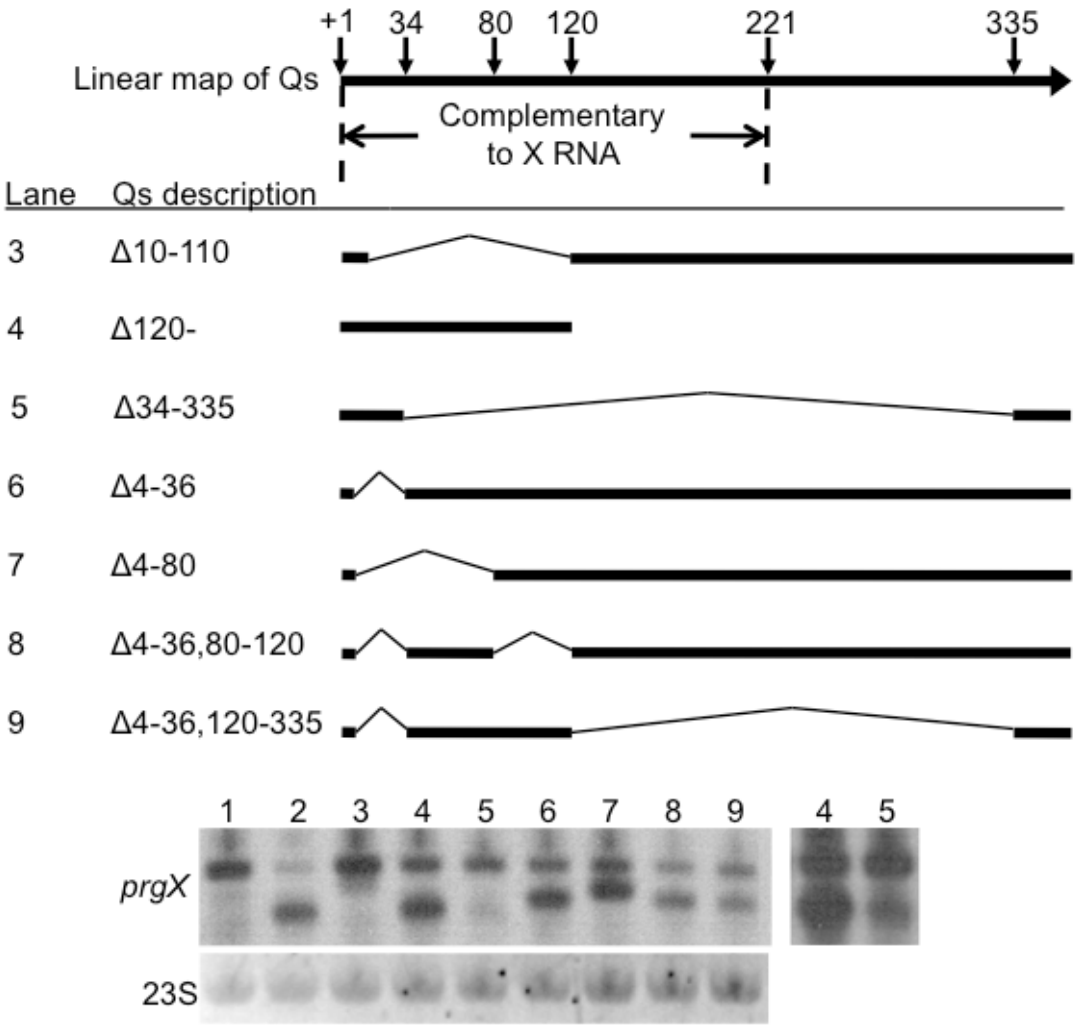


Figure 23: Sequences within the 5' end of Qs direct X RNA processing. A linear map of Qs is shown with arrows indicating points at various distances from the transcriptional start site. OG1Sp containing pBK2-126, producing X RNA, was transformed with versions of pDM5 containing different alleles of Qs. The chart indicates, pictorially and numerically, the region of Qs that has been deleted in each strain. RNA was prepared from these strains and analyzed by Northen blot using a probe specific for the *prgX* open reading frame. Lane 1 has pBK1-126 only with no pDM5, lane 2 has pDM5 producing wild-type Qs, lanes 3-9 have pDM5 bearing the deletion indicated in the chart. The loading control is an ethidium bromide stain of 23S.

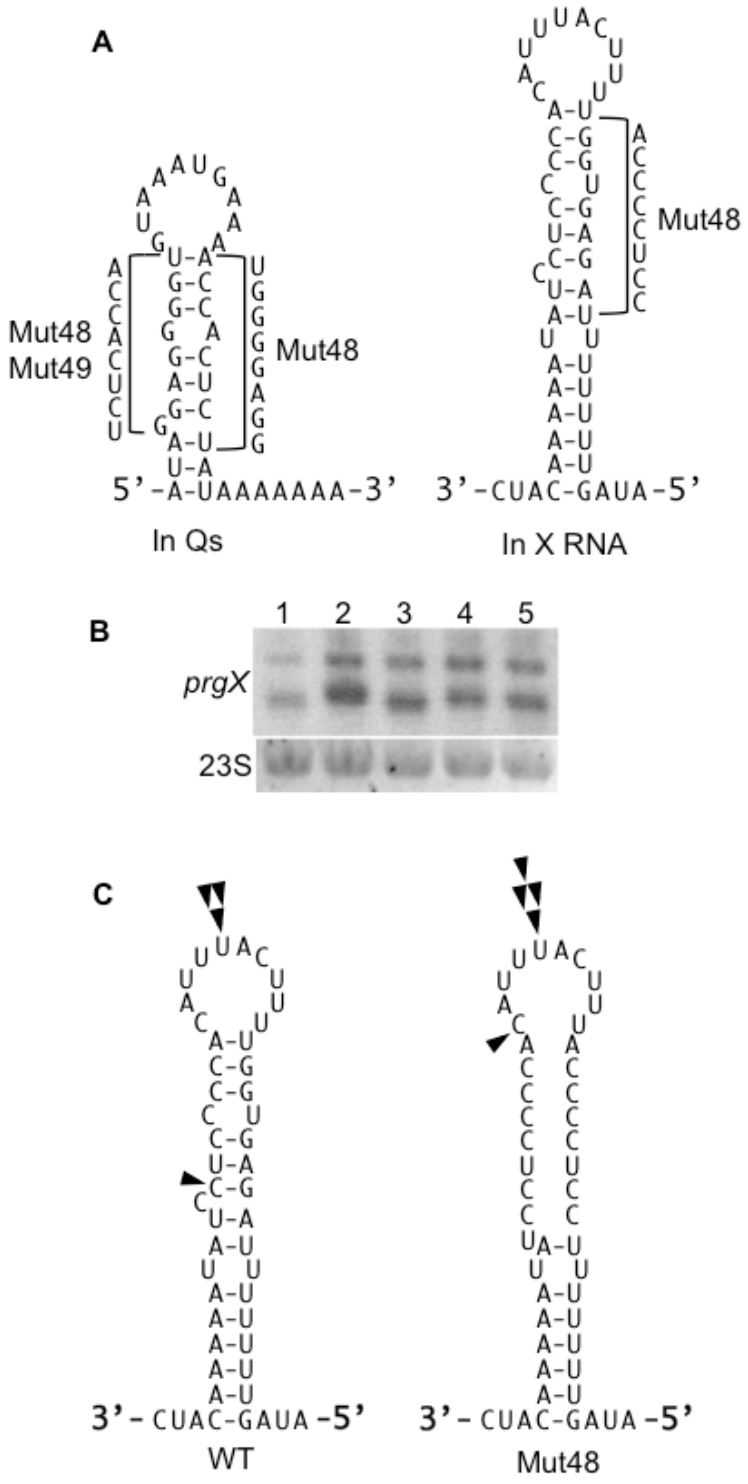

A
Mut48
Mut49
Mut48
Mut48
5'-A-UAAAAAAA-3'
In Qs
3'-CUAC-GAUA-5'
In X RNA
B
1 2 3 4 5
prgX
23S
C
3'- CUAC-GAUA -5'
WT
3'- CUAC-GAUA -5'
Mut48

Figure 24

Figure 24: A predicted stem-loop structure is not necessary to direct processing. A. This sequence and a predicted structure of Qs and X RNA in the region of processing with the RNA indicated below each diagram. Mutations tested in this experiment are shown. B. RNA was prepared from OG1Sp carrying plasmids that produce the X RNA and Qs alleles indicated, then analyzed by northern blot using a probe specific for the *prgX* open reading frame. Lane 1: WT X RNA, WT Qs. Lane 2: WT X RNA, Mut48 Qs. Lane 3: Mut 48 X RNA, Mut48 Qs. Lane 4: Mut48 X RNA, WT Qs. Lane 5: WT X RNA, Mut49 Qs. C. 5' RACE analysis was performed on RNA samples from lanes 1 and 3 of part B. A predicted structure of X RNA in the region of processing is shown. Arrowheads indicate individually sequenced 5' termini. The loading control is an ethidium bromide stain of 23S.

this predicted structure directed processing of X RNA to a precise location we generated a mutation that disrupted the stem structure in each RNA (Mut48) and a compensatory mutation that restored the stem structure in Qs (Mut49) (Figure 24A). We transformed OG1Sp with plasmids bearing wild-type or mutant X RNA and Qs alleles and assessed processing of X RNA using northern blots. Consistent with the deletion analysis, these mutations did not prevent processing (Figure 24B). When either Qs or X RNA carried the stem-disrupting mutation, the processing site was altered (compare lane 1 to lanes 2 and 4). This was not corrected if Qs was further mutated to restore the predicted stem structure (lane 5). However, if both RNAs carried the stem-disrupting mutation the site of processing shifted back to its wild-type location (lane 3).

We used a 5' RACE protocol to precisely identify the processing site when both RNAs were wild-type or both carried the stem-disrupting mutation (mut48). The primary location of processing was identical in both cases (Figure 24C). We conclude that, under the conditions tested, the predicted stem-loop structure is dispensable for directing processing to the wild-type location, and that the location of the processing site is instead dependent on complementarity between the two RNAs.

Translation of PrgX is suppressed from processed X RNA.

Processed X RNA still contains the *prgX* RBS and ORF, however, sequence analysis using sfold revealed that nucleotides within the 3' half of stem-loop A are complementary to and may interact with the *prgX* RBS, approximately 180nt downstream (Figure 25A and B). We hypothesized that cleavage of X RNA liberates these nucleotides from the helix, allowing them to pair with the *prgX* RBS and act as an occluding sequence, blocking ribosome binding to the RBS and preventing translation of PrgX from processed X RNA.

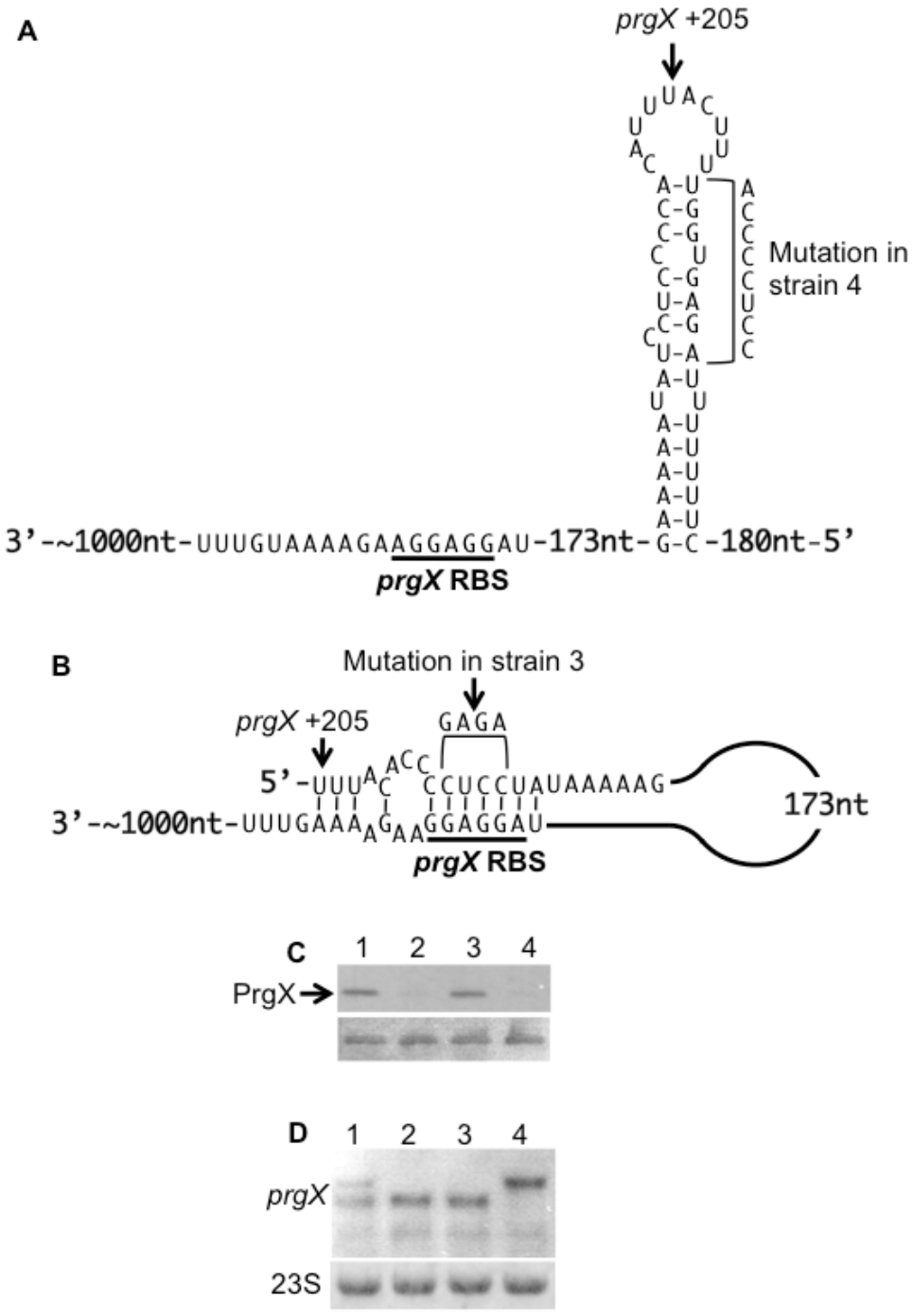

A
prgX +205
Mutation in strain 4
3'-~1000nt-UUUGUAAAAGAAGGAGGAU-173nt-G-C-180nt-5'
prgX RBS
B
Mutation in strain 3
prgX +205
GAGA
5'-UUUAACC CCUCCUAUAAAAAG
3'-~1000nt-UUUGAAAAGAAGGAGGAU
173nt
prgX RBS
C
1 2 3 4
PrgX
D
1 2 3 4
prgX
23S

Figure 25

Figure 25: Processed X RNA suppresses translation of PrgX. **A, B**. The sequence and proposed structure of X RNA before (**A**) and after (**B**) RNase III processing. The *prgX* RBS is underlined; *prgX* +205, the site of RNase III processing, is shown; the approximate number of nucleotides omitted in gaps is shown, mutations used in the experiment are indicated by brackets. **C, D**. OG1Sp was transformed with plasmids to test translational suppression as follows: lane 1, pBK2-26 + pDM5-25; lane 2, pCJ6; lane 3, pCJ18; lane 4, pBK2-46. These strains were used to prepare samples for **C.** A western blot probed with a PrgX specific antibody. The bottom panel show an unidentified, cross reactive host protein that serves as a loading control. It is the same blot as shown in the lower panel, but exposed for a longer duration. **D** A northern blot hybridized to a probe specific for the *prgX* ORF.

To test this, we deleted *prgX* +4 to +208 from pBK2-26, generating pCJ6. This fuses the P_X transcriptional start site to the region of X RNA processing so that the transcribed RNA simulates processing (Figure 25B). We found that when this plasmid was electroporated into OG1Sp, the strain produced abundant X RNA of the correct size (Figure 25D lane 2), but very little PrgX (Figure 25C, lane 2) compared to a strain bearing the plasmid that produces wild-type X RNA and Qs (Lane 1). Mutating 4 of the 7 nucleotides in the occluding sequence restored abundant translation of PrgX (lane 3). We hypothesized that PrgX can be translated from full-length X RNA because the occluding sequence is sequestered within the 3' half of a stem-loop helix. To test this, we mutated the 5' half of the stem-loop helix in our X RNA-producing plasmid, freeing the occluding sequence to interact with the *prgX* RBS. This mutation suppressed PrgX translation while producing abundant X RNA (Figure 25C, lane 4). Taken together, these data support a model in which translation of PrgX is suppressed from processed X RNA by nucleotides that occlude to the *prgX* RBS, but that in unprocessed X RNA translation is allowed because these nucleotides are sequestered within the helix of a stem-loop structure.

Analysis of the sequences involved in translational repression of PrgX indicated that the helix formed by the occluding sequence and the *prgX* RBS is only 7 base pairs in length. Considering that there are almost 200 nucleotides separating these sequences, this is a very short region of interaction. This suggested to us that the intervening region is important for translational repression, acting as a hinge to bring the occluding sequence and RBS into close proximity. We hypothesized that replacing the hinge with ectopic sequences would disrupt translational suppression. In order to test this we deleted the hinge region from pCJ6, which produces "processed" X RNA, and replaced it with sequences of various length and structural complexity. All of the sequences chosen were predicted not to directly interact with either the RBS or the occluding sequence. We used these plasmids to transform OG1Sp and tested the strains for

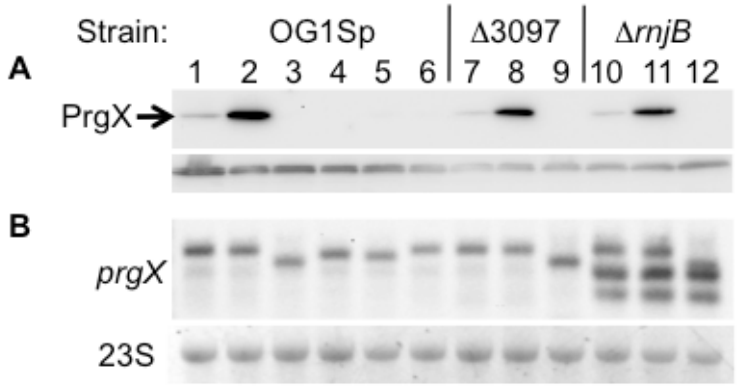


Figure 26: The X RNA hinge region regulates suppression of PrgX translation and this is independent of RNase III and RNase J2. Plasmids carrying the *prgX* Δ +6 to +210 allele and substitutions to the *prgX* hinge region, *prgX* +223-403 were cloned and used to transform OG1Sp (lanes 1-6), OG1RFΔ3097 (lanes 7-9) or OG1SSpΔ*rnjB* (lanes 10-12) as follows: lanes 1, 7, 10, pCJ6 (wild-type hinge); lanes 2, 8, 11, pCJ18 (wild-type hinge with a mutated occluding sequence, see figure 26); lane 3, pCJ30 (hinge is replaced with the sequence TAAT); lane 4, pCJ31 (hinge is replaced with *prgQ* +112 to +206); lane 5, pCJ32 (hinge is replaced with *prgQ* +30 to +112); lane 6, pCJ33 (hinge is replaced with *prgQ* +30 to +206). These strains were used to prepare samples for: **A.** A western blot probed with a PrgX specific antibody. The bottom panel shows an unidentified, cross reactive host protein that serves as a loading control. **B**. A Northern blot hybridized to a probe specific for the *prgX* ORF, the lower panel shows an ethidium bromide stain of 23S, used as a loading control.

production of PrgX and transcripts containing the *prgX* ORF. We found that in all cases suppression of PrgX translation was more effective with the ectopic hinges than with the wild type hinge (Figure 26A). This was not due to a decrease in the abundance of *prgX* mRNA (Figure 26B). This suggested that, contrary to our hypothesis, the wild type hinge specifically permits a low level of translation to occur from processed X RNA.

These results suggested that unidentified factors may interact with the hinge region to maintain a correct level of translational suppression. We hypothesized that such a factor would need to interact with processed X RNA in order to function. We also hypothesized that if a factor were missing, translation of processed X RNA would decrease, due to an inability to alleviate suppression, as seen in the hinge mutants. To test this we transformed strains containing deletions of candidate genes with pCJ6, producing processed X RNA, pCJ18, which does not suppress translation due to mutations in the occluding sequence, and pCJ30, which does not alleviate translational suppression. We chose to test strains bearing deletions to RNase III and RNase J2 in this manner, based on availability. We then compared the pattern of PrgX production from these plasmids in the different genetic backgrounds to the pattern seen in OG1Sp (Figure 26A). We found that the pattern of PrgX expression from the three plasmids was identical in the wild type and mutant genetic backgrounds. A northern blot showed that the amount of *prgX* mRNA produced by the plasmids is the same within each strain, although the strain containing the deletion of RNase J2 shows accumulation of shorter RNAs (Figure 26B). These data indicate that neither RNase J2 nor RNase III is necessary for the correct function of translational suppression from processed X RNA, but RNase J2 may be involved in subsequent steps in X RNA degradation.

Discussion

The *prgQ* and *prgX* operons encode mechanisms to reciprocally regulate each other.

Transcription of the *prgQ* and *prgX* operons is convergent for 223 nt. One consequence of this arrangement is that when P_Q is derepressed, transcription of the *prgQ* operon represses transcription from P_X via transcriptional interference. Another consequence is that RNAs from each operon transcribed within the overlapping region are perfectly complementary to transcripts from the other operon, giving rise to at least two RNA–RNA regulatory interactions. These interactions allow one operon to become dominant at the expense of the other; pheromone-sensitive repression of the *prgQ* promoter by PrgX ultimately determines the relative strength of transcription of the two operons.

From the experiments presented in this work, we conclude that the *prgQ* and *prgX* operons use RNAs to reciprocally regulate downstream gene expression from the other operon, each using a distinct mechanism. Anti-Q impacts *prgQ* at the level of transcription elongation, inducing a conformation change in nascent transcripts that signals RNAP to terminate transcription, preventing expression of downstream genes. Qs targets mature X RNA for processing, which leads to translational silencing of the *prgX* message. This processing may have other effects on X RNA, which will be discussed below. Qs also negatively regulates expression of Anti-Q, though the mechanism by which it does this is unknown. These regulatory effects are embedded within a larger network of interactions between the *prgQ* and *prgX* operons.

Anti-Q enforces termination at IRS1.

Previous work on pCF10 revealed that control of extension of *prgQ* transcripts past IRS1 is an important checkpoint in the pheromone response[77]. In uninduced

conditions, extended transcripts are not detectable by northern blot, whereas they become readily detectable after pheromone induction.

The results presented in this work support a model in which nascent *prgQ* transcripts can form one of two mutually exclusive structures (Figure 27A). One structure contains a terminator encoded by IRS1 and an upstream binding platform (Figure 27B). The other structure contains a large antiterminator, which includes sequences from both the terminator and binding platform, marked 1 and 2 (Figure 27C). In the absence of Anti-Q this antiterminator forms and allows transcription to continue past IRS1 and into downstream conjugation genes. When Anti-Q is at a sufficient concentration, it interacts with nascent *prgQ* transcripts by direct pairing of nucleotides in loops II and III with their complementary sequences in the Qs binding platform. This stabilizes the binding platform, prevents formation of the antiterminator and allows the terminator to form, halting transcription and attenuating the expression of downstream genes.

A mode of direct interaction between Anti-Q and nascent *prgQ* transcripts is supported by: 1) the success of Anti-Q in enhancing termination in vitro when RNAP is the only protein factor present; 2) in vitro studies that demonstrate that Qs and Anti-Q can interact in the absence of additional factors[55], and 3) In vivo studies demonstrating that mismatches between loop residues of Anti-Q and nascent *prgQ* transcripts can disrupt this interaction. This is consistent with results from in vitro studies of other cis-encoded antisense RNA/RNA interactions[63,78], but different than studies of systems in which the regulatory RNA is encoded in a different genomic location than the target and the two share limited complementarity. In these trans-coded systems, a host protein factor is needed for productive interactions between the regulatory RNA and the target. In eukaryotes, Argonaute proteins support interaction between mature miRNAs and their targets[79], and in bacteria, trans-coded RNAs typically require Hfq for productive interactions with their targets[80]. *E. faecalis* does not have an identified Hfq homolog. It is possible, however, that an unidentified host factor is needed

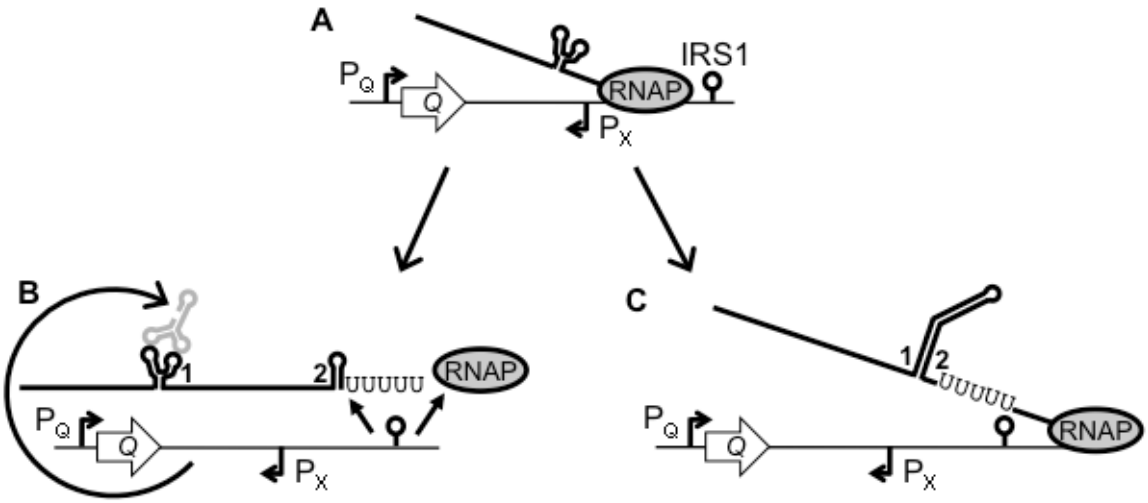


Figure 27: A model of how Anti-Q enforces termination at IRS1 as described in the text. A map of the *prgQ*/*prgX* region of pCF10 is shown as a thin black line with features. Nascent *prgQ* transcripts are indicated by a thick black line being extruded from RNAP (the shaded oval); terminated Qs transcripts are free of RNAP; Anti-Q is shown in gray.

for optimal interactions between these cis-coded RNAs in vivo. This is suggested by discrepancies between the results of our in vivo and in vitro experiments. In vivo results suggest that termination does not occur at IRS1 in the absence of Anti-Q, while our in vitro results demonstrate a basal level of termination. Additionally the in vivo results suggest that there is a structural UNR motif in Loop II of Anti-Q, whereas in vitro experiments do not support this. These observations could be due to differences in the biochemical conditions within the cell compared to the IVT reaction or to the existence of a host factor that mediates optimal RNA/RNA interaction in vivo, but which is absent from the IVT assays.

Recently, members of the Weaver lab determined the secondary structure of Anti-Q in vitro and characterized its interactions with Qs. They found that Anti-Q contains three stem-loop structures, including one with a YUNR motif; however, we found no evidence that this loop is involved in Anti-Q interactions with nascent *prgQ* transcripts. Subsequent analysis of *prgQ* sequences suggested that a loop complementary to Anti-Q loop I probably never forms. Instead, loops within *prgQ* that are complementary to Anti-Q loops II and III do form, and the complementary loops between the RNAs interact during Anti-Q driven attenuation at IRS1. This is supported by in vitro studies demonstrating that Qs protects Anti-Q sequences within these loops from lead acetate cleavage[55]. This is also supported by the genetic analysis demonstrating that mismatches between the loops of the two RNAs prevented Anti-Q from performing its normal function. This genetic study identified a potential UNR motif in a different loop of Anti-Q; however, this was not supported by other, independent means of investigation.

Genetic analysis of *prgQ* Anti-Q interactions further reveal that loops II and III are supported by stem structures within nascent *prgQ* transcripts, forming a RNA binding platform. Disruption of the stems of the binding platform impeded

interaction with Anti-Q, but mismatches between the RNAs at positions within these stems were tolerated. This suggests that the stems support the interacting loop sequences for interaction between the RNAs and that pairing of stem sequences between the RNAs is not necessary for productive interaction. This is supported by structural analysis of Qs–Anti-Q complexes, performed by the Weaver lab, which demonstrate that full duplexes do not form. Similar observations have been made in countertranscript-driven attenuation systems that control the copy number of plasmids, such as pIP501 and pAMβ1. Kinetic studies in these systems reveals that productive interaction occurs more rapidly than full duplex formation. This feature is necessary because transcription elongation is a rapid process. There is a finite window of time between transcription of the binding platform and transcription of terminator sequences. If Anti-Q induced the terminating conformation in nascent *prgQ* transcripts too slowly, the terminator helix would not form while RNAP was paused at the polyuridine tract and transcription would continue past IRS1.

Qs directs RNase III to cleave X RNA.

The results of experiments presented in this work suggest a model for Qs regulation of X RNA presented in Figure 28. X RNA is transcribed from the *prgX* operon. When not interacting with Qs, X RNA does not present any suitable target sites for RNase III. Qs is transcribed from the opposite strand of DNA as X RNA. The 5' region of Qs interacts with complementary sequences in the 5' region of X RNA, forming a duplex over 100 bp in length. This double stranded heteroduplex is cleaved in multiple locations by RNase III; the most 3' cleavage site on X RNA generating the new 5' termini detected in vivo. Processed X RNA may serve as a substrate for subsequent degradation, as discussed below.

RNase III mediates the effects of many antisense RNAs, both cis-encoded, such as the R plasmid *hok-sok* system, and trans-encoded, such as RNAIII-*spa*[30,75]. RNase III is able to cleave double-stranded RNA helices greater than 20 base

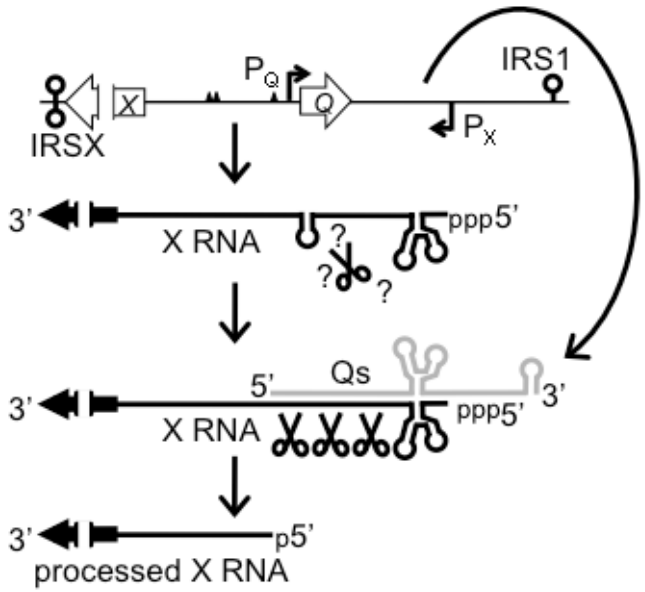


Figure 28. A model of how Qs directs cleavage of X RNA as described in the text. A map of the *prgQ*/*prgX* region of pCF10 is shown as a thin black line with features. RNA is indicated by thick black (X RNA) or gray (Qs) lines, “ppp” and “p” indicate that an RNA has a triphosphate or monophosphate group, respectively, at the 5’ terminus. Scissors represent RNase III.

pairs, although some nucleotide combinations act as “anti-determinants” to prevent cleavage at certain sites[81]. These features are apparent for *E. faecalis* RNaseIII from analysis of our in vitro reconstituted system (Figure 4). The labeled RNAs show a laddering pattern that suggests cleavage occurs at several distinct locations along the duplex, RNase III recognizing the entire duplex, but guided to certain sites by the sequences within the duplex. In vivo results, however, suggest that a complete duplex between complementary sequences of Qs and X RNA may not form. Qs +120 to +223 is complementary to X RNA +1 to +103 but does not direct RNase III cleavage. These complementary RNA regions are both highly structured. It is possible that kissing loops within this region of X RNA and complementary structures within Qs interact, imposing torsional constraints on the RNAs that prevent formation of a full duplex and RNase III target sites within this region.

Because RNase III cleavage of X RNA is upstream of the *prgX* RBS and ORF, and because this is not coordinated with processing of the 3’ terminus, processed X RNA still has all of the components necessary for PrgX translation. However, pCF10 has evolved a mechanism to repress translation of PrgX from this transcript. Genetic data indicates that cleavage of X RNA liberates sequences that are complementary to the *prgX* RBS, allowing them to occlude the RBS and suppress translation. In the unprocessed transcript, the occluding sequences are sequestered within a stem-loop structure, allowing translation.

Given the distance between the *prgX* RBS and the occluding sequences, we hypothesized that the intervening region acted as a hinge, facilitating effective translational suppression. We explored this conjecture by replacing the hinge region with ectopic sequences that varied in length and degree of structure, expecting that these sequences would allow greater translation of PrgX. The ectopic sequences had the opposite effect, repressing translation of PrgX completely. It may be that the hinge region recruits a factor that would permit

translation from processed X RNA to begin after a delay, and that failure to recruit this factor results in permanent suppression. The biological rationale for this hypothesis is that, after delivering pCF10 to a recipient cell, pheromone induced cells need to reestablish repression of the conjugation genes. The combination of transcriptional interference and Qs-directed processing of X RNA would prevent translatable *prgX* message from being expressed, unless processed X RNA could be recruited for this purpose after a delay sufficient to allow expression of conjugation genes. Ribonculeases are obvious candidates for factors that would restart translation from processed X RNA. These could act by directly removing the *prgX* RBS occluding sequence. We tested strains bearing in frame deletions of RNase III and RNase J2 and found that the pattern of expression between processed X RNA with a wild type hinge did not change relative to processed X RNA with an ectopic hinge or an mutated occluding sequence, indicating that these enzymes are not required to restart translation of PrgX.

Processing may initiate decay of X RNA. The 5' terminus of unprocessed X RNA has a triphosphate group and extensive secondary structure, both of which protect RNA from degradation[82]. RNase III cleavage generates a 5' monophosphate and may allow X RNA to serve as a substrate for docking and activation of other ribonucleases, such as RNase J1. RNase III cleavage has been shown to initiate decay of several mRNAs in *E. coli*[83,84]. However, the abundance of processed X RNA indicates that it does not decay rapidly, suggesting that it may serve a biological function. If this is the case, Qs-directed cleavage of X RNA may be part of the normal degradation pathway for X RNA because Qs is expressed from pCF10 in both induced and uninduced *E. faecalis* cells. The accumulation of short *prgX* transcripts in the RNase J2 mutant suggests that this enzyme plays a role in X RNA degradation.

Upon pheromone induction, levels of Qs rise while those of X RNA drop rapidly, suggesting that this mechanism allows the cell to rapidly downregulate PrgX production[20]. PrgX protein is modestly repressed during induction. Because the molecular mechanism by which cCF10 interaction with PrgX alleviates repression of P_Q has not been fully elucidated, it is unclear how this fluctuation in PrgX levels impacts repression of P_Q. It is also unclear how repression of P_Q is re-established. With multiple mechanisms acting to prevent expression of new PrgX, the *prgQ* operon may be expected to become continuously active following induction. The continuous expression of genes involved in conjugation would be energetically expensive and disadvantageous to the host cell, suggesting that there is a molecular timing device that switches conjugation off.

Anti-Q acts as a factor-independent terminator.

A surprising conclusion of this work is that the 3' end of Anti-Q acts as a factor-independent terminator. Factor-independent terminators are encoded in bacterial chromosomes as well as extrachromosomal elements and are generally comprised of a stem-loop structure with a high G + C content, followed by a polyuridine tract. Some terminators have imperfect stems, containing a bulge or a mismatch within the helix. There is also variation within the polyuridine tract; different RNAP enzymes have different sensitivities to the length and homogeneity of the polyuridine tract[85]. An extreme example is the phage ϕC31, which encodes terminators without a polyuridine tract, which are recognized by the RNAP encoded by the phage[86]. Anti-Q carries a polyuridine tract in which termination occurs, but in place of the simple stem-loop is a branched structure, which forms a high G + C helix at its base. This basal stem could serve the same purpose as the stem structure of a more typical terminator, extending into the RNA exit channel of RNAP and forcing its dissociation.

Regardless of the exact mechanism of termination, the fact that this unusual structure acts as a terminator has important consequences for how DNA

sequences are analyzed and annotated. Currently, bioinformatic-based analysis of genomic and metagenomic data only recognize terminators containing simple stem-loops. This information is used for the annotation of genes and operons, as well as the identification of riboswitches and other cis-acting RNA elements that contain potential terminators within 5' untranslated regions. While it is not yet clear how frequently Anti-Q–like terminators occur, it is possible that they are relatively abundant, acting to attenuate the expression of downstream genes within operons in many settings, and have escaped recognition because they are inefficient terminators that do not have a dramatic effect on downstream transcript levels.

It is evident from our experiments that the Anti-Q terminator functions at low efficiency. Like IRS1, it may be thought of as a transcriptional attenuator, which allows pCF10 to transcribe different amounts of Anti-Q RNA, encoded upstream of the termination point, and the *prgX* gene, encoded downstream. It is also evident that pCF10 is able to produce Anti-Q via a second pathway that is distinct from termination of Anti-Q. In the absence of the polyuridine tract, *prgX* transcripts that terminate in IRSX, but not IRS1, can produce Anti-Q. This phenomenon occurs in vivo, but not in vitro, suggesting that IRSX recruits a host factor that processes the longer transcript, removing sequences downstream of the Anti-Q 3' terminus. This occurs both in *E. faecalis* as well as *E. coli*, suggesting that the factor is conserved between the two bacteria. A candidate factor is PNPase. This enzyme is an exonuclease that degrades RNA from the 3' terminus and has homologs in both bacteria. By itself it is ineffective at degrading RNA with structured 3' termini, such as a terminator, but can recruit helicases to assist with this task. It is evident that the sequences sufficient to generate Anti-Q in *E. coli* and *E. faecalis* are slightly different, suggesting that the host factor that generates Anti-Q in *E. coli* has broader substrate recognition than its *E. faecalis* counterpart.

At this point it is unknown why there are two routes to generate Anti-Q. It may be that this redundancy is used to generate different amounts of Anti-Q at different points in the induction cycle. The strain that produces Anti-Q by termination only generated more Anti-Q than the strain that produced Anti-Q by processing only (Figure 7, compare lane 4 to lane 3). Production of abundant Anti-Q by termination would ensure that, when conjugation is repressed, the cell has sufficient Anti-Q to terminate all *prgQ* transcripts at IRS1. When the cell becomes induced, transcriptional interference from P_Q prevents transcription of new Anti-Q; however, the cell can still generate Anti-Q by processing already extant X RNA transcripts. This would provide a low level of Anti-Q that may help to reestablish the repression of conjugation genes once conjugation is complete.

Another important question for future study is the mechanism by which Qs regulates Anti-Q expression in trans. Several possibilities suggest themselves. It could be that Qs interacts with nascent *prgX* sequences, preventing formation of the Anti-Q terminator. There is no obvious antiterminator for Anti-Q, but Qs could directly sequester sequences important for termination. A second possibility is that Qs could interact with mature Anti-Q and recruit a host ribonuclease to degrade Anti-Q, much like Qs interaction with X RNA recruits RNase III. A third possibility is that Qs could interact with longer transcripts that serve as precursors for Anti-Q and prevent them from being processed into Anti-Q. This possibility fits well with a model in which PNPase acts to generate Anti-Q from longer transcripts because a duplex between Qs and *prgX* transcripts could block PNPase activity. These models are not mutually exclusive, and will require some effort to untangle, particularly in light of the multiple pathways that can generate Anti-Q.

P_Q exerts transcriptional interference against P_X.

Previous studies revealed that PrgX does not directly regulate P_X, suggesting that transcription from this promoter may be constitutive. In vivo experiments

revealed that products of P_X are negatively regulated by P_Q transcription in cis, a phenomenon referred to as transcriptional interference. Further work by Anushree Chatterjee has revealed that when P_Q is repressed by PrgX, it has minimal inhibitory effect on P_X transcription. When P_Q is derepressed (i.e. pheromone-induced) it represses P_X approximately 10-fold. She has also shown that collision between RNA polymerases transcribing from the two promoters may mediate transcriptional interference. Interestingly, transcription from P_X does not appear to exert transcriptional interference against P_Q to any appreciable degree. This could be because transcription from P_X is low enough that its effect is negligible, or that the mechanisms used by P_Q against P_X act in an asymmetric fashion, only affecting P_X. Whatever the reason, PrgX regulates transcription from P_X indirectly by modulating the strength of P_Q.

Our current model of regulation between the *prgQ* and *prgX* operons is depicted in Figure 29. In the absence of cCF10, the products of the *prgX* operon are dominant and prevent conjugation. Transcription from P_Q is repressed by PrgX. Repression is leaky, allowing transcription of the *prgQ* gene. Anti-Q attenuates the expression of downstream conjugation genes by interacting with nascent *prgQ* transcripts to favor formation of a factor-independent terminator within these transcripts. This terminated transcript is Qs, which encodes iCF10, an inhibitory peptide that interacts with PrgX to maintain repression[17,19]. Plasmid-free *E. faecalis* cells secrete the pheromone cCF10. When this is imported into the host cell, products of the *prgQ* operon become dominant. cCF10 interacts with PrgX, alleviating repression of P_Q. The increase in transcription from P_Q directly represses transcription from P_X, reducing expression of products of the *prgX* operon. Additionally, substantially more Qs is produced, which directs processing of X RNA, decreasing PrgX expression. The increase in transcription from P_Q eventually titrates Anti-Q levels, allowing transcription to extend past IRS1 into genes functionally involved in conjugation. This regulatory pathway

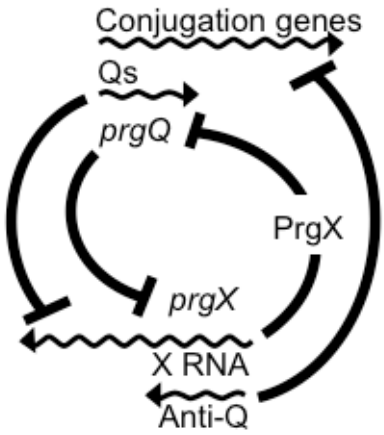


Figure 29. A model for mechanisms of reciprocal regulation between the *prgQ* and *prgX* operons, which control of the pheromone response, as described in the text. *prgQ* and *prgX* indicate transcription initiated at the respective promoters, RNAs are indicated by wavy lines, inhibitory relationships are indicated by barred arrows, the PrgX protein is indicated.

allows sensitive detection of and a robust and bistable response to changes in the ratio of cCF10 to iCF10[15,23].

The regulatory RNA segments of the *prgQ* and *prgX* operons are modular and non overlapping.

Although *prgQ* and *prgX* transcripts are perfectly complementary within the overlapping region, these studies have revealed that different sequences this region are used by Qs and Anti-Q to mediate their effects against X RNA and nascent *prgQ* transcripts, respectively. Qs requires sequences within its 5' region to direct RNase III cleavage of X RNA, whereas sequences encoding the binding platform, antiterminator and terminator are dispensable and do not direct processing of X RNA. The region of Qs required for this activity includes the *prgQ* ORF, suggesting that the Qs-X RNA interaction evolved alongside the iCF10-PrgX interaction. Anti-Q, on the other hand, interacts with the Qs binding platform, approximately Qs +120 to 220, and requires the antiterminator and terminator sequences in order to function. An IVT template that has *prgQ* Δ +14 to 114 responds to Anti-Q in identical fashion as the full length template, demonstrating that +14 to +114 are dispensable for this interaction. This modular organization suggests that the two regulatory units have distinct evolutionary origins.

Comparison to other regulatory modules.

Anti-Q driven attenuation of *prgQ* transcripts bears a striking resemblance to countertranscript-driven attenuation systems that control the copy number of several plasmids of Gram positive bacteria, including pAMβ1 from *E. faecalis* and the well studied Streptococcal plasmid, pIP501 (Figure 28A). Replication of pIP501 is controlled by the abundance of the plasmid-encoded RepR protein[87,88]. The *repR* message is RNAII, transcription of which is initiated at the pII promoter, which is negatively regulated by the CopR repressor[89]. Within the 5' UTR of RNAII is a competing terminator, encoded by IR2, and antiterminator.

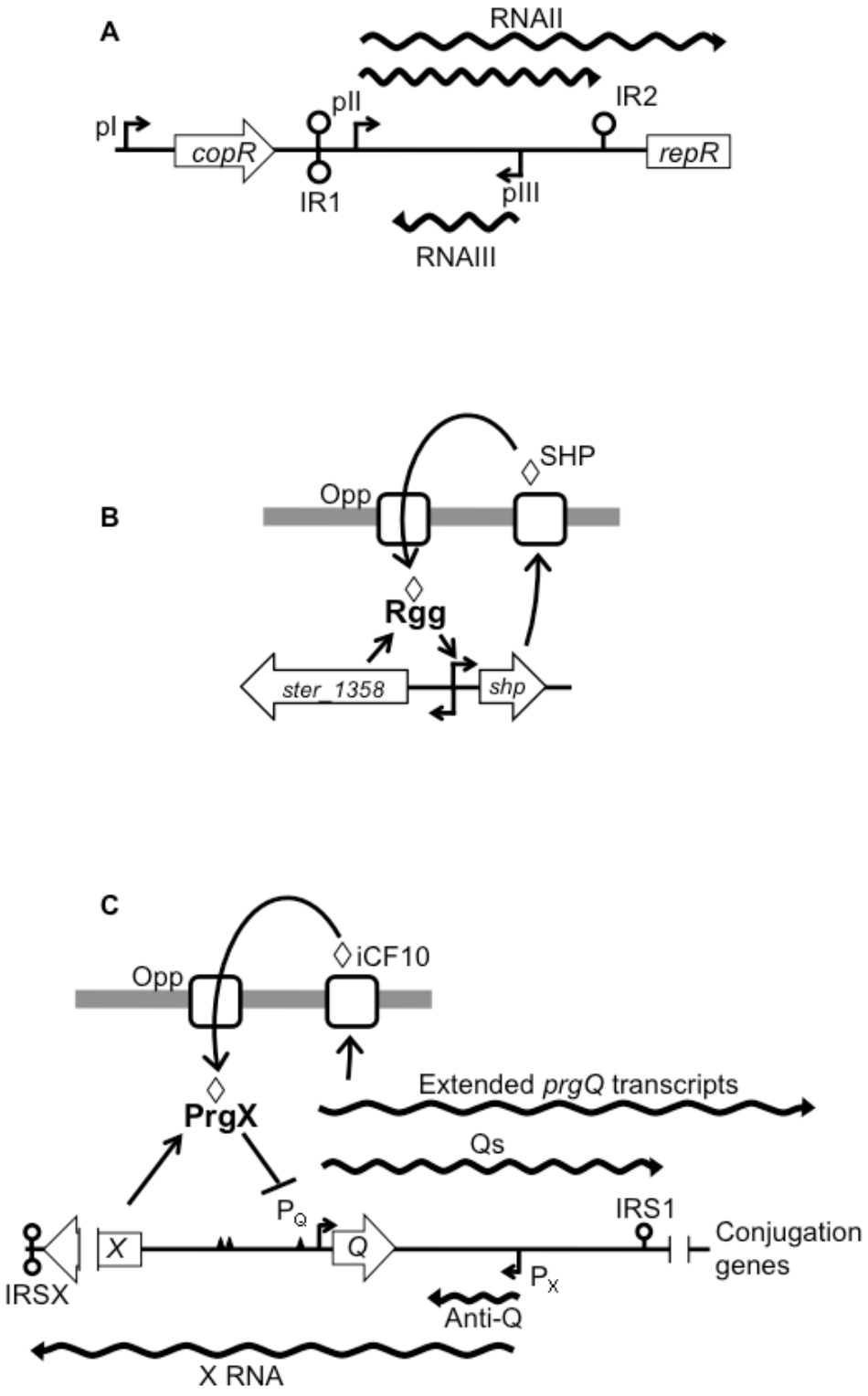

A
RNAII
pII
IR2
pI
copR
repR
IR1
pIII
RNAIII
B
SHP
Opp
Rgg
ster_1358
shp
C
iCF10
Opp
Extended prgQ transcripts
PrgX
Qs
P_Q
IRS1
Conjugation genes
X
Q
IRSX
P_X
Anti-Q
X RNA

Figure 30

Figure 30: Similarities between control of pCF10 conjugation and known regulatory pathways in Streptococcus and Enterococcus. **A**. Countertranscript-driven attenuation in pIP501. A map of the DNA is shown as a thin black line. Bent arrows indicate promoters; lollipops indicate terminators; open arrows indicate open reading frames; wavy lines indicate RNAs. **B**. A schematic depicting the function of the Rgg regulator encoded by *ster_1358* and SHP, its cognate peptide. A map of the DNA is shown as a thin black line; open arrows indicate open reading frames; bent arrows indicate putative promoters. The thick gray line indicates the cell membrane and wall. The Opp system is indicated. **C**. A map of the *prgQ* and *prgX* region, and products of the two operons. pCF10 DNA is represented by a thin line with features. These include the *prgQ* and *prgX* open reading frames, indicated by open arrows labeled *Q* and *X*, respectively. Transcriptional start sites for P_Q and P_X are shown as bent arrows. Terminators are shown as lollipops. The PrgX primary and secondary binding sites are shown as double and single triangles, respectively. Transcripts of the operons are shown as wavy lines, *prgQ* above and *prgX* below the DNA map. The PrgX protein is indicated; iCF10 is shown as a diamond. The thick gray line indicates the cell membrane and wall; the Opp system is indicated.

Termination is favored by interaction with RNAIII a, small RNA transcribed from the opposite DNA strand as RNAII, but within the 5' UTR. Loop structures within RNAIII interact with complementary loops within RNAII that are upstream of the terminator to induce formation of the terminator and attenuate transcription of *repR*[42].

Despite the parallels, the regulatory circuits of pIP501 copy number and pCF10 conjugation have different behavior. The pIP501 copy number control system is tuned to respond to small changes in plasmid abundance to keep plasmid copy number at an optimum level. If the copy number is too low, the plasmid risks not segregating to both daughters during cell division. If the copy number is too high the plasmid imposes a burden on the host. This is distinct from control of conjugation in pCF10, which, in the absence of potential recipients, needs to repress expression of energetically expensive conjugation genes, but also needs to be sensitive to the presence of potential recipients and commit to conjugation when these are present.

The primary difference between the two countertranscript-driven attenuation systems is the activity and regulation of the repressor protein. The activity of pII is controlled by the abundance of the CopR repressor, which is produced constitutively from a nearby locus. The level of CopR is dependent on *copR* gene dosage, i.e. the plasmid copy number. A drop in CopR alleviates repression of pII allowing increased transcription from pII to overcome the attenuating effects of RNAIII and express RepR. CopR is not known to respond to peptide signals. PrgX, on the other hand, is responsive to signaling peptides, which control the whether or not it represses P_Q. Expression of *prgX* is also a highly regulated process, being responsive to transcriptional interference from P_Q as well as posttranscriptional regulation directed by Qs. This suggests that differences in the behavior and expression of the repressor govern overall behavior of the two regulatory circuits.

PrgX is a member of a family of transcriptional regulators found in *Streptococcus* spp. and closely related bacteria that includes Rgg proteins. Work on Ster_1358, an Rgg family transcriptional regulator in *Streptococcus thermophilus* has revealed several parallels with PrgX activity and regulation[90] (Figure 28B). Like PrgX, the ability of Rgg to regulate transcription is responsive to a short, hydrophobic peptide pheromone (SHP). This peptide is processed from a precursor, which is encoded upstream of Rgg but on the opposite strand of DNA. In order to function, translated SHP needs to be exported, at which time it is proteolycially matured, and then reimported via the Opp system. In the cell, SHP interacts with Rgg relieving repression of the *shp* promoter and other promoters. This generates a positive feedback loop in which SHP positively regulates its own expression and that of other genes in a cell-density dependent fashion, forming a quorum-sensing system. Bioinformatic analysis of several bacteria closely related to *Streptococcus thermophilus*, including *E. faecalis* suggests that several short, hydrophobic peptides are encoded within the genomes of these bacteria[91]. Our own unpublished analysis indicates that, in *E. faecalis*, several of these are encoded upstream of transcriptional regulators with the ORFs in a divergent configuration. This raises the possibility that this mechanism of gene regulation may be relatively common in these bacteria.

The relationship between Rgg and SHP is similar to the relationship between PrgX and iCF10. In each case the transcriptional regulator, PrgX or Rgg, is encoded in a divergent configuration with a cognate peptide, iCF10 or SHP, respectively. The peptides require an extracellular phase for maturation and are reimported via the Opp system. Once in the cell the mature peptides directly interact with their transcriptional regulators to modulate expression of the peptides themselves. There are two notable differences between the systems. Although both PrgX and Rgg appear to act as transcriptional repressors, interaction between PrgX and iCF10 increases repression whereas interaction between Rgg and SHP decreases repression. PrgX interaction with cCF10

alleviates repression in much the same way as Rgg interaction with SHP, suggesting that this difference is not the result of fundamental differences between the proteins, but rather between the functions that the systems have evolved to fulfill. It is noteworthy that other Rgg-family regulators act as transcriptional activators. This appears to be a function of where the protein binding sites are located on the DNA template relative to the promoter being regulated[22,92,93]. Another difference between *prgX* and *ster_1358* is the location of the two promoters. Transcription of *prgX* message is convergent with transcription of the *prgQ* message. This allows the *prgQ* mRNA, Qs, to regulate expression of PrgX. Transcription of *ster_1358* and *shp* is not known to be convergent, however, identification of the promoters is putative, based on sequence analysis, and has not been confirmed genetically or biochemically. It is possible that transcription of these genes is convergent and overlapping, allowing for additional levels of regulation.

The modularity of the RNAs involved in reciprocal regulation between the *prgQ* and *prgX* operons suggests that the functions of the two RNAs may have evolved independently. The identification of separate homologous systems in closely related Streptococcal and Enterococcal bacteria further support this conjecture. It is interesting to think that regulation of conjugation of pCF10 and the other pheromone-responsive Enterococcal plasmids evolved from the merger of a peptide responsive transcriptional regulator and a countertranscript-driven attenuation system. Most of these plasmids have very specific host requirements and are unable to replicate outside of *E. faecalis*, suggesting that this event probably occurred in *E. faecalis*. However, the pheromone-response plasmid pAM373 [94] propagates in and responds to a pheromone produced by several Gram positive bacteria, suggesting that the pheromone response system could have evolved in elsewhere and later conjugated into *Enterococcus faecalis*, where the plasmids on which it is borne have since flourished.

Bibliography

1. Devriese, L., Baele, M. & Butaye, P. The Genus *Enterococcus*. *The Prokaryotes* 163-174 (2006).at <http://www.springerlink.com.ezp2.lib.umn.edu/content/k865515187274gw1>
2. Gilmore, M.S. The *enterococc*i: pathogenesis, molecular biology, and antibiotic resistance. (ASM Press: 2002).
3. Moellering, R.C., Jr Emergence of *Enterococcus* as a significant pathogen. *Clin. Infect. Dis* **14**, 1173-1176 (1992).
4. Dunny, G.M. & Clewell, D.B. Transmissible toxin (hemolysin) plasmid in *Streptococcus faecalis* and its mobilization of a noninfectious drug resistance plasmid. *J. Bacteriol* **124**, 784-790 (1975).
5. Franke, A.E. & Clewell, D.B. Evidence for a chromosome-borne resistance transposon (Tn916) in *Streptococcus faecalis* that is capable of “conjugal” transfer in the absence of a conjugative plasmid. *J. Bacteriol* **145**, 494-502 (1981).
6. Paulsen, I.T. et al. Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* **299**, 2071-2074 (2003).
7. Weigel, L.M. et al. Genetic analysis of a high-level vancomycin-resistant isolate of *Staphylococcus aureus*. *Science* **302**, 1569-1571 (2003).
8. Yagi, Y. et al. Plasmid content of *Streptococcus faecalis* strain 39-5 and identification of a pheromone (cPD1)-induced surface antigen. *J. Gen. Microbiol* **129**, 1207-1215 (1983).
9. Chuang, O.N. et al. Multiple functional domains of *Enterococcus faecalis* aggregation substance Asc10 contribute to endocarditis virulence. *Infect. Immun* **77**, 539-548 (2009).
10. Dunny, G.M., Brown, B.L. & Clewell, D.B. Induced cell aggregation and mating in *Streptococcus faecalis*: evidence for a bacterial sex pheromone. *Proc. Natl. Acad. Sci. U.S.A* **75**, 3479-3483 (1978).
11. Dunny, G.M. The peptide pheromone-inducible conjugation system of *Enterococcus faecalis* plasmid pCF10: cell-cell signalling, gene transfer, complexity and evolution. *Philos. Trans. R. Soc. Lond., B, Biol. Sci* **362**, 1185-1193 (2007).
12. Antiporta, M.H. & Dunny, G.M. *ccfA*, the genetic determinant for the cCF10 peptide pheromone in *Enterococcus faecalis* OG1RF. *J. Bacteriol* **184**, 1155-1162 (2002).
13. Buttaro, B.A., Antiporta, M.H. & Dunny, G.M. Cell-associated pheromone peptide (cCF10) production and pheromone inhibition in *Enterococcus faecalis*. *J. Bacteriol* **182**, 4926-4933 (2000).
14. Bensing, B.A., Meyer, B.J. & Dunny, G.M. Sensitive detection of bacterial transcription initiation sites and differentiation from RNA processing sites in the pheromone-induced plasmid transfer system of *Enterococcus faecalis*. *Proc. Natl. Acad. Sci. U.S.A* **93**, 7794-7799 (1996).

15. Hirt, H. et al. Characterization of the pheromone response of the *Enterococcus faecalis* conjugative plasmid pCF10: complete sequence and comparative analysis of the transcriptional and phenotypic responses of pCF10-containing cells to pheromone induction. *J. Bacteriol* **187**, 1044-1054 (2005).
16. Chung, J.W. & Dunny, G.M. Transcriptional analysis of a region of the *Enterococcus faecalis* plasmid pCF10 involved in positive regulation of conjugative transfer functions. *J. Bacteriol* **177**, 2118-2124 (1995).
17. Nakayama, J., Ruhfel, R.E., Dunny, G.M., Isogai, A. & Suzuki, A. The *prgQ* gene of the *Enterococcus faecalis* tetracycline resistance plasmid pCF10 encodes a peptide inhibitor, iCF10. *J. Bacteriol* **176**, 7405-7408 (1994).
18. Leonard, B.A., Podbielski, A., Hedberg, P.J. & Dunny, G.M. *Enterococcus faecalis* pheromone binding protein, PrgZ, recruits a chromosomal oligopeptide permease system to import sex pheromone cCF10 for induction of conjugation. *Proc. Natl. Acad. Sci. U.S.A* **93**, 260-264 (1996).
19. Kozlowicz, B.K. et al. Molecular basis for control of conjugation by bacterial pheromone and inhibitor peptides. *Mol. Microbiol* **62**, 958-969 (2006).
20. Bae, T., Clerc-Bardin, S. & Dunny, G.M. Analysis of expression of *prgX*, a key negative regulator of the transfer of the *Enterococcus faecalis* pheromone-inducible plasmid pCF10. *J. Mol. Biol* **297**, 861-875 (2000).
21. Bae, T., Kozlowicz, B.K. & Dunny, G.M. Characterization of cis-acting *prgQ* mutants: evidence for two distinct repression mechanisms by Qa RNA and PrgX protein in pheromone-inducible enterococcal plasmid pCF10. *Mol. Microbiol* **51**, 271-281 (2004).
22. Bae, T., Kozlowicz, B. & Dunny, G.M. Two targets in pCF10 DNA for PrgX binding: their role in production of Qa and *prgX* mRNA and in regulation of pheromone-inducible conjugation. *J. Mol. Biol* **315**, 995-1007 (2002).
23. Chatterjee, A. et al. Convergent transcription confers a bistable switch in *Enterococcus faecalis* conjugation. *Proc Natl Acad Sci U S A* (2011).doi:10.1073/pnas.1101569108
24. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849-857 (1983).
25. Kruger, K. et al. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* **31**, 147-157 (1982).
26. Prody, G.A., Bakos, J.T., Buzayan, J.M., Schneider, I.R. & Bruening, G. Autolytic processing of dimeric plant virus satellite RNA. *Science* **231**, 1577-1580 (1986).
27. Wassarman, K.M. & Storz, G. 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* **101**, 613-623 (2000).
28. Novick, R.P., Iordanescu, S., Projan, S.J., Kornblum, J. & Edelman, I. pT181 plasmid replication is regulated by a countertranscript-driven transcriptional attenuator. *Cell* **59**, 395-404 (1989).

29. Winkler, W., Nahvi, A. & Breaker, R.R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952-956 (2002).
30. Gerdes, K., Nielsen, A., Thorsted, P. & Wagner, E.G. 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 (1992).
31. Arraiano, C.M. et al. The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol. Rev* **34**, 883-923 (2010).
32. Condon, C. RNA processing and degradation in *Bacillus subtilis*. *Microbiol. Mol. Biol. Rev* **67**, 157-174, table of contents (2003).
33. Franch, T., Petersen, M., Wagner, E.G., Jacobsen, J.P. & Gerdes, K. Antisense RNA regulation in prokaryotes: rapid RNA/RNA interaction facilitated by a general U-turn loop structure. *J. Mol. Biol* **294**, 1115-1125 (1999).
34. Gutell, R.R., Cannone, J.J., Konings, D. & Gautheret, D. Predicting U-turns in ribosomal RNA with comparative sequence analysis. *J. Mol. Biol* **300**, 791-803 (2000).
35. Tomizawa, J. Control of ColE1 plasmid replication: the process of binding of RNA I to the primer transcript. *Cell* **38**, 861-870 (1984).
36. Palmer, A.C., Ahlgren-Berg, A., Egan, J.B., Dodd, I.B. & Shearwin, K.E. Potent transcriptional interference by pausing of RNA polymerases over a downstream promoter. *Mol. Cell* **34**, 545-555 (2009).
37. Stougaard, P., Light, J. & Molin, S. Convergent transcription interferes with expression of the copy number control gene, copA, from plasmid R1. *EMBO J* **1**, 323-328 (1982).
38. Brantl, S. & Wagner, E.G. Dual function of the copR gene product of plasmid pIP501. *J. Bacteriol* **179**, 7016-7024 (1997).
39. Vassylyev, D.G., Vassylyeva, M.N., Perederina, A., Tahirov, T.H. & Artsimovitch, I. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**, 157-162 (2007).
40. Babitzke, P. Regulation of transcription attenuation and translation initiation by allosteric control of an RNA-binding protein: the *Bacillus subtilis* TRAP protein. *Curr. Opin. Microbiol* **7**, 132-139 (2004).
41. Grundy, F.J. & Henkin, T.M. The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in gram-positive bacteria. *Mol. Microbiol* **30**, 737-749 (1998).
42. Brantl, S., Birch-Hirschfeld, E. & Behnke, D. RepR protein expression on plasmid pIP501 is controlled by an antisense RNA-mediated transcription attenuation mechanism. *J. Bacteriol* **175**, 4052-4061 (1993).
43. Le Chatelier, E., Ehrlich, S.D. & Jannière, L. Countertranscript-driven attenuation system of the pAM beta 1 repE gene. *Mol. Microbiol* **20**, 1099-1112 (1996).

44. Brantl, S. Regulatory mechanisms employed by cis-encoded antisense RNAs. *Curr. Opin. Microbiol* **10**, 102-109 (2007).
45. Heidrich, N. & Brantl, S. Antisense RNA-mediated transcriptional attenuation in plasmid pIP501: the simultaneous interaction between two complementary loop pairs is required for efficient inhibition by the antisense RNA. *Microbiology (Reading, Engl.)* **153**, 420-427 (2007).
46. Fuchs, R.T., Grundy, F.J. & Henkin, T.M. The S(MK) box is a new SAM-binding RNA for translational regulation of SAM synthetase. *Nat. Struct. Mol. Biol* **13**, 226-233 (2006).
47. Sharma, R.C. & Schimke, R.T. Preparation of electrocompetent E. coli using salt-free growth medium. *BioTechniques* **20**, 42-44 (1996).
48. Kristich, C.J., Chandler, J.R. & Dunny, G.M. Development of a host-genotype-independent counterselectable marker and a high-frequency conjugative delivery system and their use in genetic analysis of *Enterococcus faecalis. Plasmid* **57**, 131-144 (2007).
49. Kozlowicz, B.K. The molecular mechanism and peptide signaling response of PrgX used to control pheromone-induced conjugative transfer of pCF10. (2005).
50. Johnson, C.M. et al. Direct evidence for control of the pheromone-inducible *prgQ* operon of *Enterococcus faecalis* plasmid pCF10 by a countertranscript-driven attenuation mechanism. *J. Bacteriol* **192**, 1634-1642 (2010).
51. Leenhouts, K. et al. A general system for generating unlabelled gene replacements in bacterial chromosomes. *Mol. Gen. Genet* **253**, 217-224 (1996).
52. Fixen, K.R. et al. Analysis of the amino acid sequence specificity determinants of the enterococcal cCF10 sex pheromone in interactions with the pheromone-sensing machinery. *J. Bacteriol* **189**, 1399-1406 (2007).
53. Gao, P. et al. *Enterococcus faecalis rnjB* is required for pilin gene expression and biofilm formation. *J. Bacteriol* **192**, 5489-5498 (2010).
54. Trieu-Cuot, P., Carlier, C., Poyart-Salmeron, C. & Courvalin, P. An integrative vector exploiting the transposition properties of Tn1545 for insertional mutagenesis and cloning of genes from gram-positive bacteria. *Gene* **106**, 21-27 (1991).
55. Shokeen, S. et al. Structural analysis of the Anti-Q-Qs interaction: RNA-mediated regulation of *E. faecalis* plasmid pCF10 conjugation. *Plasmid* **64**, 26-35 (2010).
56. Maguin, E., Duwat, P., Hege, T., Ehrlich, D. & Gruss, A. New thermosensitive plasmid for gram-positive bacteria. *J. Bacteriol* **174**, 5633-5638 (1992).
57. van der Vossen, J.M., van der Lelie, D. & Venema, G. Isolation and characterization of *Streptococcus cremoris* Wg2-specific promoters. *Appl. Environ. Microbiol* **53**, 2452-2457 (1987).
58. Bryan, E.M., Bae, T., Kleerebezem, M. & Dunny, G.M. Improved vectors for nisin-controlled expression in gram-positive bacteria. *Plasmid* **44**, 183-190 (2000).

59. Ballering, K.S. et al. Functional genomics of *Enterococcus faecalis*: multiple novel genetic determinants for biofilm formation in the core genome. *J. Bacteriol* **191**, 2806-2814 (2009).
60. Weaver, K.E., Ehli, E.A., Nelson, J.S. & Patel, S. Antisense RNA regulation by stable complex formation in the *Enterococcus faecalis* plasmid pAD1 par addiction system. *J. Bacteriol* **186**, 6400-6408 (2004).
61. Miller, J. *Experiments in molecular genetics*. (Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y., 1972).
62. Toledo-Arana, A. et al. The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* **459**, 950-956 (2009).
63. Greenfield, T.J., Franch, T., Gerdes, K. & Weaver, K.E. 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 (2001).
64. Hartz, D., McPheeters, D.S., Traut, R. & Gold, L. Extension inhibition analysis of translation initiation complexes. *Meth. Enzymol* **164**, 419-425 (1988).
65. Fromant, M., Blanquet, S. & Plateau, P. Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal. Biochem* **224**, 347-353 (1995).
66. Oppenheim, A.B., Kobiler, O., Stavans, J., Court, D.L. & Adhya, S. Switches in bacteriophage lambda development. *Annu. Rev. Genet* **39**, 409-429 (2005).
67. Hawley, D.K. & McClure, W.R. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res* **11**, 2237-2255 (1983).
68. Jeng, S.T., Gardner, J.F. & Gumport, R.I. Transcription termination by bacteriophage T7 RNA polymerase at rho-independent terminators. *J. Biol. Chem* **265**, 3823-3830 (1990).
69. Macdonald, L.E., Zhou, Y. & McAllister, W.T. Termination and slippage by bacteriophage T7 RNA polymerase. *J. Mol. Biol* **232**, 1030-1047 (1993).
70. Dunn, J.J. & Studier, F.W. The transcription termination site at the end of the early region of bacteriophage T7 DNA. *Nucleic Acids Res* **8**, 2119-2132 (1980).
71. Artsimovitch, I. & Landick, R. Interaction of a nascent RNA structure with RNA polymerase is required for hairpin-dependent transcriptional pausing but not for transcript release. *Genes Dev* **12**, 3110-3122 (1998).
72. Yarnell, W.S. & Roberts, J.W. Mechanism of intrinsic transcription termination and antitermination. *Science* **284**, 611-615 (1999).
73. Heidrich, N. & Brantl, S. Antisense-RNA mediated transcriptional attenuation: importance of a U-turn loop structure in the target RNA of plasmid pIP501 for efficient inhibition by the antisense RNA. *J. Mol. Biol* **333**, 917-929 (2003).
74. Reynolds, R., Bermúdez-Cruz, R.M. & Chamberlin, M.J. Parameters affecting transcription termination by *Escherichia coli* RNA polymerase. I. Analysis of 13 rho-independent terminators. *J. Mol. Biol* **224**, 31-51 (1992).

75. Huntzinger, E. et al. *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate *spa* gene expression. *EMBO J* **24**, 824-835 (2005).
76. Ding, Y. & Lawrence, C.E. A statistical sampling algorithm for RNA secondary structure prediction. *Nucleic Acids Res* **31**, 7280-7301 (2003).
77. Bensing, B.A. & Dunny, G.M. Pheromone-inducible expression of an aggregation protein in *Enterococcus faecalis* requires interaction of a plasmid-encoded RNA with components of the ribosome. *Mol. Microbiol* **24**, 295-308 (1997).
78. Brantl, S. & Wagner, E.G. Antisense RNA-mediated transcriptional attenuation occurs faster than stable antisense/target RNA pairing: an in vitro study of plasmid pIP501. *EMBO J* **13**, 3599-3607 (1994).
79. Carthew, R.W. & Sontheimer, E.J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**, 642-655 (2009).
80. Waters, L.S. & Storz, G. Regulatory RNAs in bacteria. *Cell* **136**, 615-628 (2009).
81. Zhang, K. & Nicholson, A.W. Regulation of ribonuclease III processing by double-helical sequence antideterminants. *Proc. Natl. Acad. Sci. U.S.A* **94**, 13437-13441 (1997).
82. Belasco, J.G. All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay. *Nat. Rev. Mol. Cell Biol* **11**, 467-478 (2010).
83. Portier, C., Dondon, L., Grunberg-Manago, M. & Régnier, P. The first step in the functional inactivation of the *Escherichia coli* polynucleotide phosphorylase messenger is a ribonuclease III processing at the 5' end. *EMBO J* **6**, 2165-2170 (1987).
84. Régnier, P. & Grunberg-Manago, M. RNase III cleavages in non-coding leaders of *Escherichia coli* transcripts control mRNA stability and genetic expression. *Biochimie* **72**, 825-834 (1990).
85. Artsimovitch, I., Svetlov, V., Anthony, L., Burgess, R.R. & Landick, R. RNA polymerases from *Bacillus subtilis* and *Escherichia coli* differ in recognition of regulatory signals in vitro. *J. Bacteriol* **182**, 6027-6035 (2000).
86. Ingham, C.J., Hunter, I.S. & Smith, M.C. Rho-independent terminators without 3' poly-U tails from the early region of actinophage øC31. *Nucleic Acids Res* **23**, 370-376 (1995).
87. Brantl, S. & Behnke, D. The amount of RepR protein determines the copy number of plasmid pIP501 in *Bacillus subtilis*. *J. Bacteriol* **174**, 5475-5478 (1992).
88. Brantl, S., Behnke, D. & Alonso, J.C. Molecular analysis of the replication region of the conjugative *Streptococcus agalactiae* plasmid pIP501 in *Bacillus subtilis*. Comparison with plasmids pAM beta 1 and pSM19035. *Nucleic Acids Res* **18**, 4783-4790 (1990).
89. Brantl, S. The *copR* gene product of plasmid pIP501 acts as a transcriptional repressor at the essential *repR* promoter. *Mol. Microbiol* **14**, 473-483 (1994).

90. Ibrahim, M. et al. Control of the transcription of a short gene encoding a cyclic peptide in *Streptococcus thermophilus*: a new quorum-sensing system? *J. Bacteriol* **189**, 8844-8854 (2007).
91. Ibrahim, M. et al. A genome-wide survey of short coding sequences in streptococci. *Microbiology (Reading, Engl.)* **153**, 3631-3644 (2007).
92. Fleuchot, B. et al. Rgg proteins associated with internalized small hydrophobic peptides: a new quorum-sensing mechanism in streptococci. *Mol Microbiol* **80**, 1102-1119 (2011).
93. Rawlinson, E.L.A., Nes, I.F. & Skaugen, M. Identification of the DNA-binding site of the Rgg-like regulator LasX within the lactocin S promoter region. *Microbiology (Reading, Engl.)* **151**, 813-823 (2005).
94. Clewell, D.B., An, F.Y., White, B.A. & Gawron-Burke, C. *Streptococcus faecalis* sex pheromone (cAM373) also produced by *Staphylococcus aureus* and identification of a conjugative transposon (Tn918). *J. Bacteriol* **162**, 1212-1220 (1985).

Appendix A: Summary of studies of PrgZ and chromosomal *opp* genes

Deletion of *prgZ* decreases sensitivity to pheromone.

We generated an in frame deletion of *prgZ* in pCF10, replacing the open reading frame with the sequence ATGAAGAAGTACAAGGTTTTAATGCATTCAATAGGA GCAAAATATGACTACAAAAAAATGAGAATAGAGAAGTAA, which fuses the first 5 codons to the last 20 codons of *prgZ*. We then tested the ability of *E. faecalis* CK104 (OG1RFΔ*upp2*) bearing either wild type pCF10 or pCF10Δ*prgZ* to form clumps in response to pheromone. Clumping is caused by expression of aggregation substance, Asc10, which is encoded by the *prgB* gene. Asc10 expression is dependent on pheromone induction, making clumping a simple way to assess if a strain is responding to pheromone. The strains were grown overnight in M9-YEG or THB. The next morning we performed clumping assays on these cultures (Table 1). Briefly: a 96 well microtiter plate was prepared with 2 fold serial dilutions of cCF10. The overnight culture was diluted 1:10 into these wells using the same medium that they were cultured in overnight, and shaken at 250 rpm for 2 h at 37°C. The cultures were then visua lly assessed for clumping. The starting cCF10 concentration for each experiment is indicated along with the highest dilution at which clumping was observed in all replicates. The pheromone concentration of this dilution has been calculated (Table 6).

Strains bearing the pCF10 are sensitive to lower concentrations of cCF10 than strains bearing pCF10Δ*prgZ*, supporting the hypothesis that PrgZ increases the uptake of pheromone. Also, the the response of the Δ*prgZ* strain is sensitive to the medium in which it is grown, needing a higher concentration of pheromone when grown in THB than in M9. This may be due to CodY dependent repression of the ABC transporter responsible for pheromone uptake in rich media.

Table 6: Pheromone induced clumping response of pCF10 strains

Medium	pCF10 strain	Starting cCF10 conc. (pg/ml)	Clumping dilution	Minimum cCF10 conc. for clumping (pg/ml)
THB	pCF10	500	8	63
THB	pCF10	250	8	31
THB	pCF10Δ*prgZ*	200000	8	25000
THB	pCF10Δ*prgZ*	200000	8	25000
M9	pCF10	500	8	63
M9	pCF10	1000	32	31
M9	pCF10	1000	16	63
M9	pCF10Δ*prgZ*	5000	16	313
M9	pCF10Δ*prgZ*	500	2	250
M9	pCF10Δ*prgZ*	1000	8	125
M9	pCF10Δ*prgZ*	1000	4	250

iCF10 may interfere with PrgZ-specific import of cCF10

We next wanted to test if PrgZ might be responsible for iCF10 import as well. In order to do this we tested the ability of *E. faecalis* bearing pCF10 or pCF10Δ*prgZ* to transfer a copy of the plasmid to a recipient strain in response to a combination of cCF10 and iCF10. *E. faecalis* CK104 bearing pCF10 or pCF10Δ*prgZ* (donor strain) and OG1SSp (recipient strain) were cultured overnight in the indicated medium. Cultures were diluted 1:10 in medium containing pheromone and inhibitor peptides (donors) or fresh medium (recipients). The cultures were induced for 90 minutes at 37°C. Then donors and recipients were mixed 1:9 and conjugative transfer of the plasmid was allowed to occur for 15 min at 37°C. The matings were diluted with 1-9 volumes of cold KPBS (pH 7.4) and the cultures were plated on THB agar with antibiotics selective for donors, recipients or transconjugates (TCs). The number of TCs per donor for each culture was determined and the percent induction of mating was calculated by dividing the number of TC/donor for each treatment by the number of TC/donor for the pheromone-only induction in that experiment.

The results of a mating assay are shown in Figure 31. Both pCF10 and pCF10Δ*prgZ* are inhibited by similar levels of iCF10, relative to cCF10, suggesting that both iCF10 and cCF10 are recruited for import by PrgZ. If PrgZ did not recruit iCF10, then we would expect pCF10Δ*prgZ* to be much more sensitive to iCF10 than wild type pCF10, since the inhibitor uptake mechanism would still be intact.

When the mating assay is performed in M9 with 500 pg/ml of pheromone, conjugation of pCF10 responds to lower levels of iCF10 than pCF10Δ*prgZ*. The kinetics of inhibition also appear to be different. In the presence of PrgZ, iCF10 causes a precipitous decline in mating behavior that may level off as mating

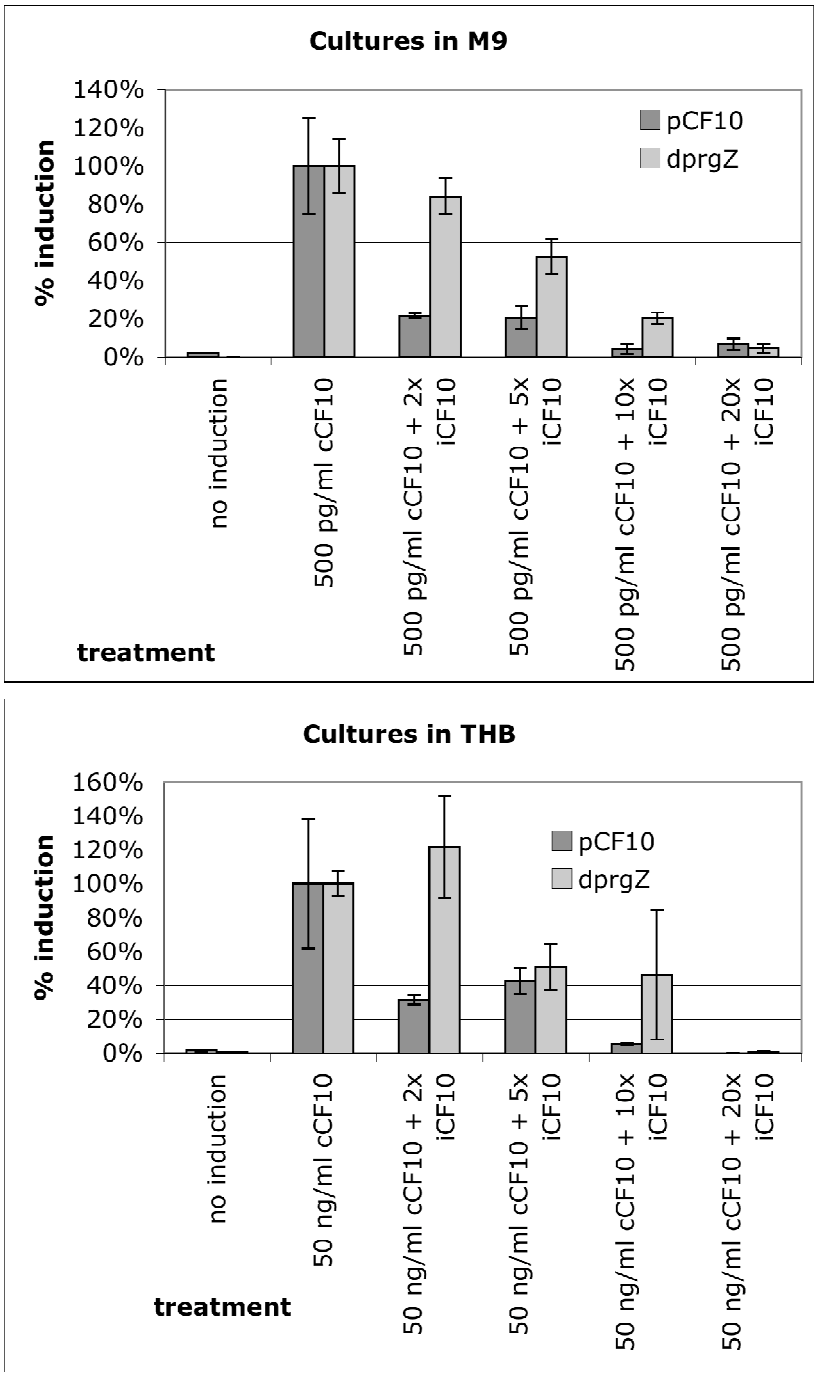


Figure 31. Conjugative transfer of pCF10 variants between *E. faecalis* strains. Mating experiments were performed and results calculated as described in the text. Dark gray bars indicate frequency of transfer of wild-type pCF10, light gray bars indicate frequency of transfer of pCF10Δ*prgZ*.

approaches uninduced levels. In the absence of PrgZ, the reduction in mating is more graded. These data suggest that at these concentrations, interactions between PrgZ and iCF10 are physiologically significant for determining whether or not a cell commits to conjugation. It may be that iCF10 is inhibiting cCF10 uptake via occupancy of PrgZ or that the intracellular concentration of iCF10 is being increased more quickly than that of cCF10 and inhibiting mating behavior through interactions with PrgX.

Confounding this interpretation of the data is the unknown contribution of endogenously produced iCF10 by organisms bearing the different plasmids. One of the first responses of pheromone induction is to increase transcription of the *prgQ* ORF, which encodes the iCF10 precursor. The two plasmids may therefore be generating different amounts of endogenous iCF10 in response to pheromone induction.

Expression of Opp locus genes in rich and minimal media

PrgZ is homologous to OppA, the substrate binding protein of Opp ABC transporter systems. Earlier work suggested that PrgZ cooperates with an *E. faecalis* Opp system to import cCF10 [18]. We have measured the expression of candidate genes from the two *opp* loci in *E. faecalis* using reverse transcription real time PCR. Briefly, overnight cultures of OG1RF ± pCF10 were grown in either THB or M9-YEG. The next morning cultures were diluted with four volumes fresh medium ± 50 ng/ml cCF10 and incubated for 90 minutes at 37°C. Total RNA was then prepared from the cultures as described in materials and methods. To prepare cDNA we used Superscript III reverse transcriptase primed with random hexamer oligonucleotides as per manufacturer instructions. We then used optimized primer sets to examine expression of the genes shown in Figure 32, using *gyrB* as an internal standard. Olionucleotides used to measure specific genes are given in Table 7.

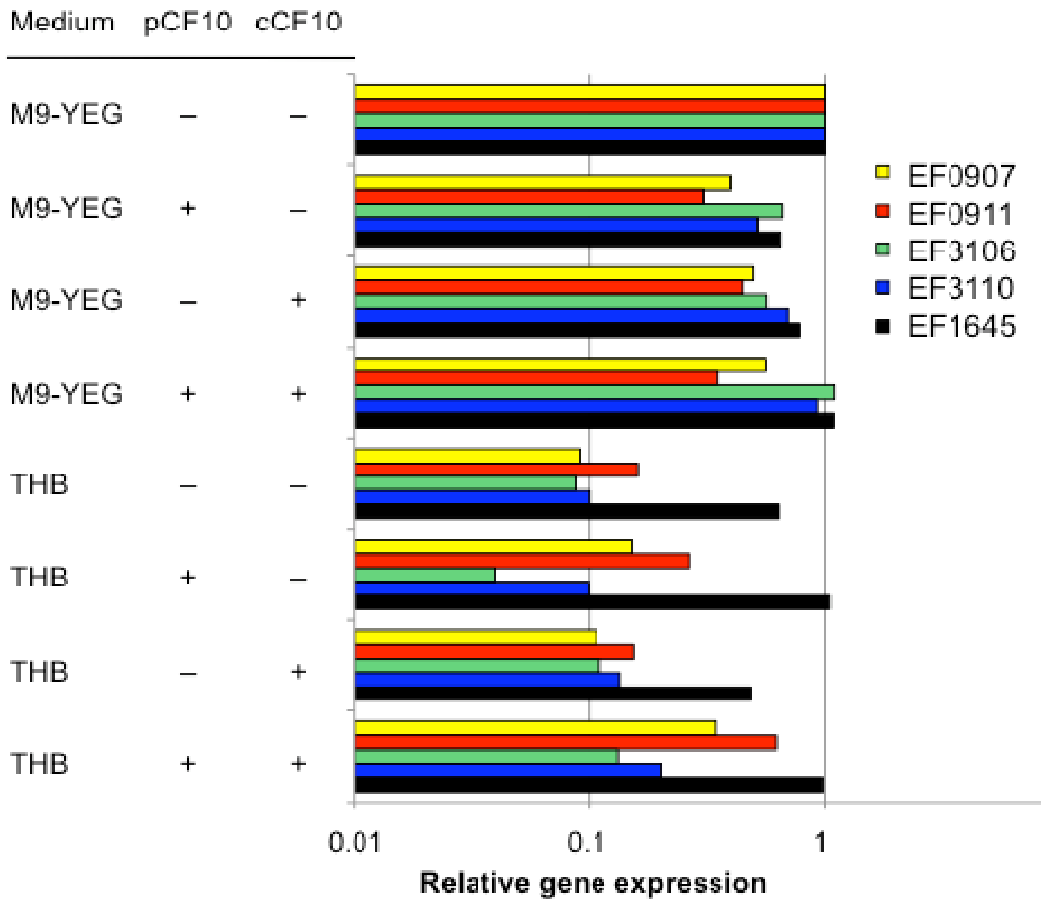


Figure 32. Relative *opp* gene expression in OG1RF cultures under the given conditions. Expression levels of each gene are normalized to the expression level in cells without pCF10 cultured in M9 medium without pheromone. The genes tested were: EF0907, homologous to *oppA* and *prgZ*, a substrate binding protein; EF0911, *oppD* homolog, an ATP binding protein; EF3106, homologous to *dtpA*, a substrate-binding protein with closer homology to dipeptide permeases than oligopeptide permeases; EF3110, *oppD* homolog, ATP binding protein. EF1645, *codY* homolog. PCR amplification cycles for all genes except EF0907 were done in two steps: a 95°melt cycle for 15 seconds followed by a 62° annealing, extension and data collection cycle for 30 seconds. For EF0907, each cycle was performed in three stems: a 95°melt cycl e for 15 seconds, a 58° annealing and extension cycle for 30 seconds and a 76°data collection cycle for 30 seconds.

Table 7: Oligonucleotide primers used to measure gene expression in Figure 32

Gene	Forward primer	Reverse primer
EF0907	TCTGGTGAATTAGCCCAACA	TGGTGATTTTTCATCACGTTG
EF0911	GGTGGACGTCTGGTTGAAGT	ATCGTTGGCATTGACCCTAA
EF3106	AGCGATGAAAGCGAAACAAT	CGGGACGTCCTAAGATTTGA
EF3110	GATCCGTTGTCTGCACTGAA	TGTTTTTGTTGTGCGGTCAT
EF1645	GCCGTGGATATGCTCAAAGT	CCAAACGGATACAGCTCTCG

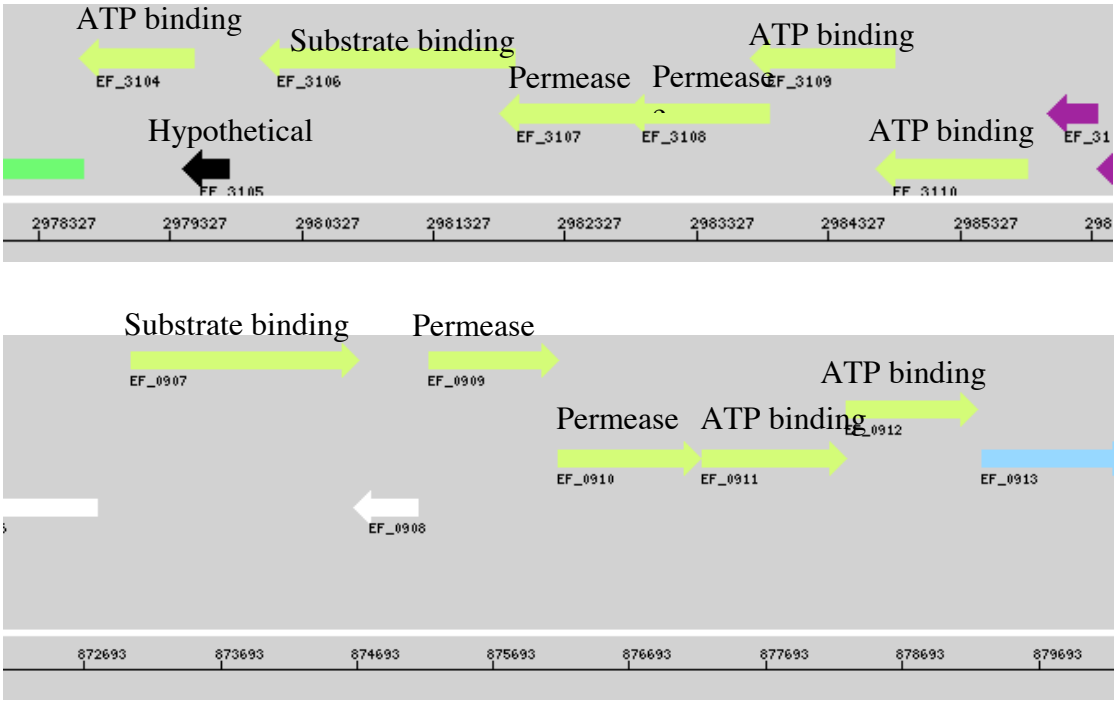


Figure 33: A region map of the genes encoding potential Opp transporters in *E. faecalis* V583 as generated by the NCBI genome viewer web application. The figure is enhanced with text describing the potential function of the genes.

opp gene expression is reduced in cells grown in THB medium compared to M9-YEG, explaining why cells grown in THB are less sensitive to exogenous pheromone. The expression of both sets of *opp* genes is affected by the growth medium, though it appears that genes in the 3104-3110 group show a greater sensitivity than those in 0907-0912. It appears that pCF10 and excess cCF10 do not effect the expression of *opp* genes.

The regions of the *E. faecalis* V583 genome encoding the two candidate Opp transporters are shown in Figure 33, with some annotation. We have used allelic exchange to generate markerless, in-frame deletions of the permease genes in both operons within OG1RF. Specifically, in OG1RFΔ3107-3108 the two ORFs were replaced by the sequence ATGTGGAAAACAATTCAACGATTAGGCTAA, which fused the first 5 codons of EF3108 to the last 5 codons of EF3107. In OG1RFΔ0909-0910 the two ORFs were replaced by the sequence GGTGTCATT ATTGAATAGTTTTGGAAAGATGAAAGAGTAA, which fused the first 8 codons of EF0909 to the last 5 codons of EF0910. These strains have not been phenotypically characterized.