Regulation of the mobile genetic element ICE*Bs1* by a conserved repressor and anti-repressor

by

Baundauna Bose

B.A. Biology
Cornell University, College of Arts & Sciences, 2000

Submitted to the Department of Biology in Partial Fulfillment
of the Requirements for the Degree of

Doctor of Philosophy in Biology
at the
Massachusetts Institute of Technology

February 2010

Signature of author: _____
Department of Biology
December 11, 2009

Certified by: _____
Alan D. Grossman
Professor of Biology

Accepted by: _____
Stephen P. Bell
Professor of Biology
Co-chair, Biology Graduate Committee

Regulation of the mobile genetic element ICE*Bs1* by a conserved repressor and anti-repressor

by
Baundauna Bose

Submitted to the Department of Biology on December 11, 2009 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biology

Abstract

The mobile genetic element ICE*Bs1* is an integrative and conjugative element (a conjugative transposon) found in the *Bacillus subtilis* chromosome. The SOS response and the RapI-PhrI sensory system activate ICE*Bs1* gene expression, excision, and transfer by inactivating the ICE*Bs1* repressor protein ImmR. Although ImmR is similar to many characterized phage repressors, we found that, unlike these repressors, inactivation of ImmR requires an ICE*Bs1*-encoded anti-repressor ImmA (YdcM). Under ICE*Bs1*-inducing conditions, ImmA cleaves ImmR at a specific site to induce the element.

We found that changing the amount or the specific activity of ImmA can cause derepression of ICE*Bs1* without activation by RecA or RapI. We isolated and characterized mutations in *immA* (*immA*^{h}) that cause derepression of ICE*Bs1* gene expression in the absence of inducing signals. However, we also found that ImmA levels did not significantly change during activation by RapI, indicating that RapI-mediated induction is likely due to increased activity of ImmA. Therefore, we propose that RapI and RecA induce ICE*Bs1* by increasing its specific activity.

Along with earlier observations, some ImmA^{h} mutants highlighted the importance of ImmA's C-terminal sequence for regulation of ImmA protein levels. We demonstrated that GFP tagged with C-terminal residues of ImmA was less abundant in vivo than untagged GFP. We screened cells with mutations of ATP-dependent proteases for effects on ICE*Bs1* expression, and found that ClpXP might play a role in regulating ImmA stability and ICE*Bs1* gene expression.

To learn more about the repressor, ImmR, we isolated and characterized mutants of *immR* (*immR(ind-)*) that attenuate induction of ICE*Bs1* gene expression under the normally inducing conditions of treatment with DNA damaging reagent and overproduction of RapI. All four identified *immR(ind-)* mutations fall within a stretch of 10 residues flanking the cleavage site, emphasizing the importance of this sequence for ImmR proteolysis and ICE*Bs1* induction. To further characterize the C-terminal portion of ImmR, we demonstrated that it interacts with ImmA and with itself in yeast two-hybrid assays, indicating that this part of the protein likely functions in ImmR oligomerization and recognition of ImmR by ImmA.

Homologs of ImmA and ImmR are found in many mobile genetic elements, so the mode of regulation by ImmA and ImmR may be conserved in various systems.

Thesis Advisor: Alan D. Grossman

Title: Professor of Biology

Acknowledgements

I 'd like to thank my thesis advisor Alan Grossman for giving me the opportunity to work in his lab and for his efforts in mentoring me over the past several years. The other folks who have worked with ICE*Bs1* are fantastic colleagues, and I would like to especially acknowledge Jennifer Auchtung, Melanie Berkmen, and C. Lee for being excellent role models, labmates, and friends. Jenny deserves immense credit for her work with ICE*Bs1*, and I feel very lucky to have worked with and learned from her.

I can't thank C. Lee enough for all that she has contributed to my work-life during my stint in the Grossman lab. If I ever become half as great a scientist as she is, I'll feel extremely proud of myself. I don't dare hope of ever becoming as great a human being.

I feel deeply indebted to Bob Sauer for his mentorship and support, from the time I joined his lab as a technician to the present. He's been a great source of inspiration, advice, and humor. I will always admire his diabolical genius and be grateful for his consistent encouragement of scientific investigation that is motivated by genuine curiosity.

I'd like to thank the other members of my thesis committee- Mike Laub, Graham Walker, and Linc Sonenshein for their advice about my project. I'd chiefly like to acknowledge Mike for always taking the time to talk through various dilemmas related to both my project and the plotlines on Lost.

The members of the Grossman lab and the other labs in this neck of Building 68 (Baker, Sauer, Kaiser, Walker, Laub, and Endy) have made this an enviable work environment, and I thank them for their camaraderie and their generosity in sharing supplies and information. In particular, Jim Butler, Tahmeena Chowdhury, Joey Davis, Jen Hou, Andreas Martin, Anne Meyer, Mihaela Pruteanu, Caterina Schweidenback, Carolyn Sevier, Christos Tsokos, and Kevin Wang have positively and markedly impacted my life and my work here. I sincerely thank my special science + coffee buddies over the years- Aliaa Abdelhakim, Susan Cohen, Alexi Goranov, and Jade Wang for their roles as colleagues and enablers of my caffeine addiction..

I have a terrific urban family, and without them I couldn't function: Aliaa Abdelhakim, Melanie Arzt, Melanie Berkmen, Susan Cohen, Meaghan Donovan, Tracie Goldman, Meagan Hanna, Chris Hayes, Restu Ismail, Chaim Katz, Asya Khonina, Linda Kim, Shirley Li, Kathleen McGinness, Randi Rotjan, Mukund Thattai, David Wah, and Chia-yung Wu.

Finally and above all, I'm grateful to my mother for her patience, support, and love.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction
Horizontal Gene Transfer in bacteria

Outline:

(introduction)

Horizontal transfer of DNA in bacteria
- Natural transformation
- Transduction
- Conjugative plasmids
- ICEs
  - (introduction)
  - Host range
  - Attachment sites
  - Excision and integration
  - Nicking, Replication, and DNA transfer
  - Generation of diversity

Mobile element-encoded functions

Signals that regulate HGT
- Rationale
- DNA damage
- Cell-cell signaling
- Antibiotics
- Other signals

Mechanisms of regulation of HGT
- Phages
  - Autocleaving repressors
  - Phage anti-repressors
- ICEs
  - ICE*Bs1*
  - Other DNA damage-induced ICEs
  - Tetracycline-induced ICEs
  - pSAM2

Conclusion & thesis outline

There are three main types of horizontal gene transfer (HGT) in bacteria- natural transformation, transduction, and conjugation. Natural transformation is the active acquisition and incorporation of extracellular DNA. Transduction is phage-mediated transfer of DNA between bacteria. Conjugation, or bacterial mating, involves the transfer of conjugative plasmids or integrative and conjugative elements (ICEs, also known as conjugative transposons; Table 1) from cell to cell by direct contact.

Table 1. Partial list of ICEs

Element	Genus or species	Size (kb)	Characterized functions[a]	Ref.
CFB group				
CTnDOT[c]	*Bacteroides fragilis*	65	Tc^r Em^r	(Cheng *et al.*, 2000)
CTnERL[b, c]	*Bacteroides fragilis*	52	Tc^r	(Salyers *et al.*, 1995; Shoemaker *et al.*, 1989)
CTnGERM1	*Bacteroides ovatus* DH3716	75	Em^r	(Wang *et al.*, 2003)
α-Proteobacteria				
Symbiosis[c] island	*Mesorhizobium loti* MAFF303099	611	Symbiosis—type III secretion	(Sullivan *et al.*, 2002)
Symbiosis[c] island	*Mesorhizobium loti* R7A	502	Symbiosis—type IV secretion	(Sullivan *et al.*, 2002)
β-Proteobacteria				
Tn*4371*	*Ralstonia* sp. A5	55	Biphenyl degradation	(Toussaint *et al.*, 2003)
γ-Proteobacteria				
bph-sal	*Pseudomonas putida*	90	Biphenyl and salicylate degradation	(Nishi *et al.*, 2000)
clc[c]	*Pseudomonas* sp. B13	105	Chlorocatechol degradation	(Ravatn *et al.*, 1998)
CTn*scr94*[c]	*Salmonella enterica* Senftenberg	100	Sucrose utilization	(Hochhut *et al.*, 1997)
ICE*Ec1*	*Escherichia coli*	69	yersiniabactin siderophore system (virulence determinant)	(Schubert *et al.*, 2004)
LpPI-1	*Legionella pneumophila*	65	Various putative virulence-associated determinants	(Brassinga *et al.*, 2003)
p1056	*Haemophilus influenzae*	ND	Tc^r Ap^r	(Pickard *et al.*, 2003)
pJY1	*Vibrio cholerae*	ND	Su^r Cm^r Sm^r	(Yokota and Kuwahara, 1977)
pMERPH	*Shewanella putrefaciens*	ND	Hg^r	(Boltner *et al.*, 2002; Pembroke *et al.*, 2002; Peters *et al.*, 1991)

Element	Genus or species	Size (kb)	Characterized functions[a]	Ref.
R391	*Providencia rettgeri*	89	Kn^r Hg^r	(Boltner *et al.*, 2002; Pembroke *et al.*, 2002)
R997	*Proteus mirabilis*	85	Ap^r Sm^r Su^r	(Boltner *et al.*, 2002; Murphy and Pembroke, 1999; Pembroke *et al.*, 2002)
SPI-7	*Salmonella enterica* Typhi	134	Antigen Vi (capsule)	(Dimopoulou *et al.*, 2002)
SXT [c]	*Vibrio cholerae*	99.5	Su^r Tm^r Cm^r Sm^r	(Beaber *et al.*, 2002; Waldor *et al.*, 1996)
High G+ C Gram-positive				
pIJ110 [d]	*Streptomyces parvulus*	13.6	None identified	(Hopwood *et al.*, 1984)
pIJ408 [d]	*Streptomyces glaucescens*	15	None identified	(Hopwood *et al.*, 1984; Sosio *et al.*, 1989)
pMEA100 [d]	*Amycolatopsis mediterranei*	23.7	None identified	(Madon *et al.*, 1987; Moretti *et al.*, 1985)
pMEA300 [c, d]	*Amycolatopsis methanolica*	13.3	Mutator—stimulation of transformation	(Vrijbloed *et al.*, 1994)
pMR2 [d, e]	*Micromonospora rosaria*	11.2	None identified	(Hosted *et al.*, 2005)
pSA1	*Streptomyces cyaneus*	9.1	None identified	(Miyoshi *et al.*, 1986)
pSAM2 [c, d]	*Streptomyces ambofaciens*	10.9	None identified	(Pernodet *et al.*, 1984)
pSE101 [d, e]	*Saccharopolyspora erythraea*	10.9	None identified	(Brown *et al.*, 1988)
pSE211 [d]	*Saccharopolyspora erythraea*	18.1	None identified	(Brown *et al.*, 1990)
pSG1 [d, e]	*Streptomyces griseus*	16.9	None identified	(Cohen *et al.*, 1985)
SLP1 [d]	*Streptomyces coelicolor*	17.2	None identified	(Vogtli and Cohen, 1992)
Low G +C Gram-positive				
CdiA1	*Clostridium difficile*	30.6	None identified	(Burrus *et al.*, 2002)
CdiA2	*Clostridium difficile*	ND	None identified	(Burrus *et al.*, 2002)
CdiB3	*Clostridium difficile*	28.2	None identified	(Burrus *et al.*, 2002)
CdiB4	*Clostridium difficile*	ND	None identified	(Burrus *et al.*, 2002)

Element	Genus or species	Size (kb)	Characterized functions[a]	Ref.
CW459 *Tet*(M)	*Clostridium perfringens*	ND	Tc^r	(Roberts *et al.*, 2001)
EfaC1	*Enterococcus faecalis*	25.3	None identified	(Burrus *et al.*, 2002)
EfaC2	*Enterococcus faecalis*	32.7	None identified	(Burrus *et al.*, 2002)
EfaD2	*Enterococcus faecalis*	ND	None identified	(Burrus *et al.*, 2002)
ICE*Bs1*[c]	*Bacillus subtilis* 168	20.5	None identified	(Burrus *et al.*, 2002)
ICEF	*Mycoplasma fermentans* PG18	23	None identified	(Calcutt *et al.*, 2002)
ICE*Lm1*	*Listeria monocytogenes*	21.3	Putative Cd2+ resistance	(Burrus *et al.*, 2002)
ICE*St1*[c]	*Streptococcus thermophilus*	34.7	Type II restriction-modification *Sth*368I	(Burrus *et al.*, 2002)
ICE*St3*[c, e]	*Streptococcus thermophilus*	28.1	None identified	(Pavlovic *et al.*, 2004)
pRS01/sex factor	*Lactococcus lactis*	48.4	Tellurium resistance	(Gasson *et al.*, 1995)
SmuE	*Streptococcus mutans*	20.5	None identified	(Burrus *et al.*, 2002)
Tn*916*[c]	*Enterococcus faecalis*	18.4	Tc^r	(Clewell *et al.*, 1995; Gawron-Burke and Clewell, 1982)
Tn*916*S[c, e]	*Streptococcus intermedius*	18	Tc^r	(Lancaster *et al.*, 2004)
Tn*1545*[c, e]	*Streptococcus pneumoniae*	25.3	Tc^r Em^r Kn^r	(Courvalin and Carlier, 1986)
Tn*1549*	*Enterococcus* sp.	34	Vm^r	(Garnier *et al.*, 2000)
Tn*5251*[c, f]	*Streptococcus pneumoniae*	18	Tc^r	(Ayoubi *et al.*, 1991)
Tn*5252*[c, f]	*Streptococcus pneumoniae*	47	Cm^r UV^r	(Ayoubi *et al.*, 1991; Munoz-Najar and Vijayakumar, 1999; Vijayakumar and Ayalew, 1993)
Tn*5253*[c, e, f]	*Streptococcus pneumoniae*	65.5	Tc^r Cm^r UV^r	(Ayoubi *et al.*, 1991)
Tn*5276*[c]	*Lactococcus lactis*	70	Sucrose utilization and nisin synthesis	(Rauch and De Vos, 1992)
Tn*5386*[c, e]	*Enterococcus faecium*	29.5	putative: lantibiotic resistance	(Rice *et al.*, 2007)
Tn*5397*[c]	*Clostridium difficile*	21	Tc^r	(Wang *et al.*, 2000)
Tn*5801*	*Staphylococcus aureus*	25.8	Tc^r	(Kuroda *et al.*, 2001)
Tn*6000*[c, e, g]	*Enterococcus faecium*	ND	Tc^r	(Roberts *et al.*, 2006)

Table modified from Table 1 in Burrus & Waldor, 2004.

Ap, ampicillin; Cm, chloramphenicol; Em, erythromycin; Hg, mercury; Kn, kanamycin; ND, not determined; Sm, streptomycin; Su, sulfamethoxazole; Tc, tetracycline; Tm, trimethoprim; UV, ultra-violet light; Vm, vancomycin.

a Functions encoded which are not involved in integration/excision, DNA transfer or regulation of mobility.
b formerly Tc^{r}ERL
c mentioned in the text
d included in review (te Poele *et al.*, 2008)
e not present in original table by Burrus & Waldor, 2004.
f Tn*5253* contains Tn*5251* and Tn*5252*
g formerly EfcTn

HGT enables the rapid transfer of traits between bacteria. The various functions that can be transferred, and the many bacteria that are capable of transferring and receiving genes in this way make the significance of HGT clear. One area in which HGT has obvious consequences is the spread of antibiotic resistance among infectious bacteria. Resistance genes have been acquired via all types of HGT; these traits are carried on phages, conjugative plasmids, ICEs, and DNA taken up by naturally competent bacteria (reviewed in Barlow, 2009; reviewed in Salyers *et al.*, 1995; Salyers and Shoemaker, 1996; Shoemaker *et al.*, 2001; Whittle *et al.*, 2002). The pervasive spread of antibiotic resistance by HGT suggests that for antibiotics to remain effective, we may have to develop drugs to inhibit HGT (Barlow, 2009).

HGT is not only significant for its role in spreading antibiotic resistance and virulence traits among pathogens. Phages influence cycling of organic matter in the oceans (Canchaya *et al.*, 2003). The *clc* element, an ICE, is induced by and can metabolize aromatic chlorinated compounds; this may have environmental and industrial applications (Sentchilo *et al.*, 2003). This element resides in and transfers itself among bacteria in diverse environments, such as membrane reactors, groundwater, and wastewater-treatment plants (Springael *et al.*, 2002; Zhou and Tiedje, 1995). Phages and ICEs affect bacteria in ways that are of critical importance to the dairy industry (Canchaya *et al.*, 2003). Finally, HGT may critically influence the evolution of bacterial genomes (Canchaya *et al.*, 2003).

One fundamental aspect of HGT that is well worth understanding is its regulation. Regulation of HGT seems to be important for mobile elements to optimize maintenance in their current hosts and transfer to new hosts (Oppenheim *et al.*, 2005; Salyers *et al.*, 1995).

The work presented in this thesis focuses on the regulation of ICE*Bs1* (Fig. 1). ICE*Bs1* is

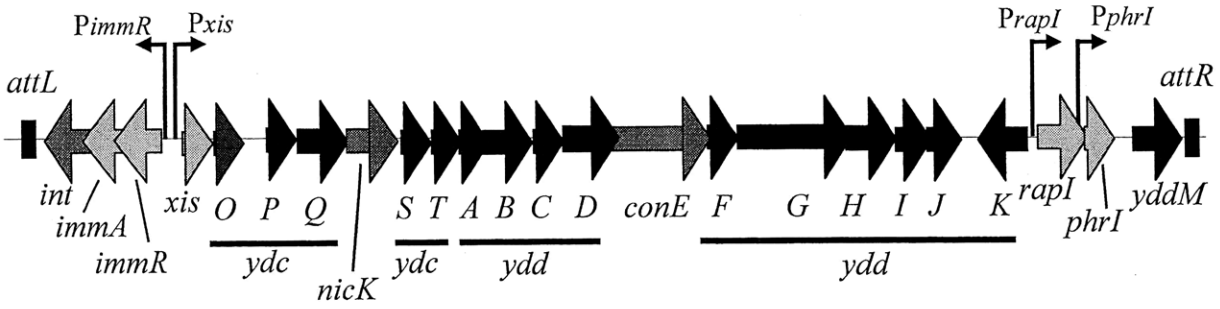


Figure 1. Organization of ICE*Bs*1. The 24 orfs and four promoters are shown as arrows. The name of each gene is indicated below its arrow. Boxes at the left and right represent attachment sites. Genes encoding characterized products are grey. *immA* encodes the anti-repressor and *immR* encodes the repressor for the element. *int* and *xis* encode the integrase and excisionase, respectively. *nicK* encodes a DNA relaxase that nicks the DNA at an origin of transfer (*oriT*) sequence within the *nicK* gene. *conE* encodes a protein that likely forms part of the mating apparatus. (This figure is identical to Figure 1 in Chapter 3.)

an approximately 20 kbp integrative and conjugative element inserted in a tRNA gene in *Bacillus subtilis* (Auchtung *et al.*, 2005; Burrus *et al.*, 2002). Genes at the left end of ICE*Bs1* are part of a regulatory module that resembles those found in many bacteriophages (Auchtung *et al.*, 2005; Burrus *et al.*, 2002). This module includes *immR* and *immA*, encoding the element's repressor and anti-repressor, respectively (Auchtung *et al.*, 2007; Bose* *et al.*, 2008). ImmR represses transcription of genes required for excision and transfer and both activates and represses its own expression (Auchtung *et al.*, 2007) (Fig. 2). ICE*Bs1* gene expression is derepressed in vivo during the RecA-dependent SOS response, or when the ICE*Bs1*-encoded cell-cell signaling regulator RapI is present and active (Auchtung *et al.*, 2005). In both cases, derepression requires the anti-repressor ImmA (Bose* *et al.*, 2008). ImmA is a site-specific protease that cleaves ImmR, thereby causing derepression of ICE*Bs1* gene expression (Bose* *et al.*, 2008). It is not known how ImmA-mediated proteolysis of ImmR is stimulated by RecA or RapI.

Horizontal transfer of DNA in bacteria

Natural transformation

Natural transformation is the active uptake and heritable integration of extracellular DNA (reviewed in Johnsborg *et al.*, 2007). The advantages to cells that become naturally competent are currently being explored and debated (reviewed in Claverys *et al.*, 2006; Johnsborg *et al.*, 2007). The DNA that is taken up could provide new genetic material, with which a cell can repair its own damaged DNA. Or this DNA could be a source from which cells acquire new traits and generate diversity. Alternatively, the DNA could provide nourishment for the cell.

The genes governing competence development have been classified into two groups for Gram-positive bacteria (reviewed in Johnsborg *et al.*, 2007). "Early genes" determine whether conditions are suitable for cells to become competent, and "late genes" encode DNA uptake and

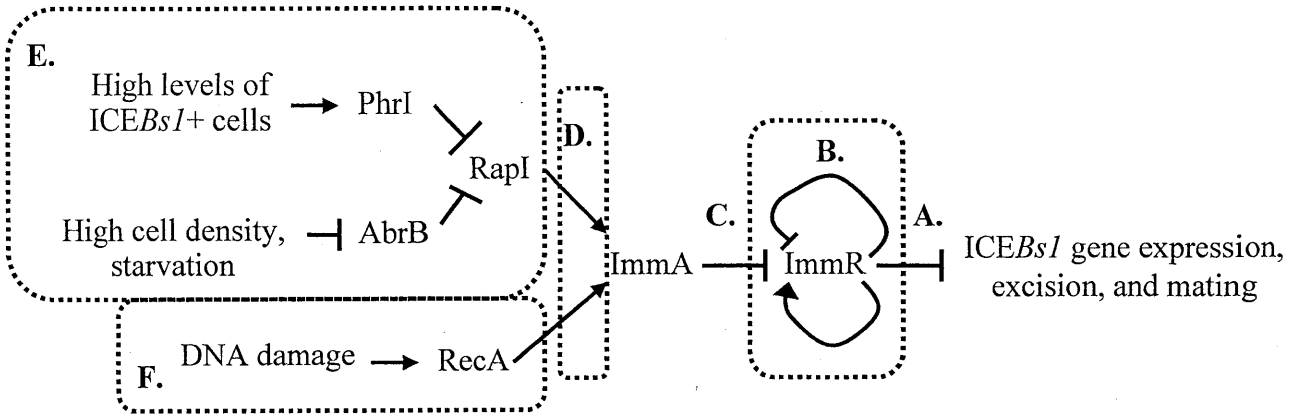


Figure 2. Regulation of ICE*Bs*1.

A. ICE*Bs1* gene expression, excision, and mating are repressed by ImmR.

B. ImmR also regulates its own synthesis. Low ImmR protein levels activate *immR* gene expression, while high ImmR levels repress *immR* expression.

C. ImmR's activity is antagonized by the anti-repressor ImmA, which catalyzes a site-specific cleavage of ImmR, promoting its further degradation.

D. ImmA-mediated cleavage of ImmR is activated by either of two types of signals.

E. RapI activates ImmA-mediated cleavage of ImmR, when RapI is synthesized (when high cell density relieves repression of *rapI* by AbrB) and active (when neighboring levels are not producing and secreting PhrI peptide).

F. RecA protein activated during DNA damage activates ImmA-mediated cleavage of ImmR

(Figure redrawn from originals by J.M. Auchtung & A.D. Grossman)

species, but they usually involve cell-cell communication via specific peptides (Johnsborg *et al.*, 2007). Late genes of Gram-positive bacteria are highly conserved (Lapidus *et al.*, 2002; Martin *et al.*, 2006). The majority of all naturally transformable bacteria use the same basic machinery to acquire DNA (reviewed in Claverys *et al.*, 2006). Twitching motility and DNA uptake are closely related (Johnsborg *et al.*, 2007). Usually, competent cells require type IV pili or pseudopili for internalization of DNA (Johnsborg *et al.*, 2007).

For some types of bacteria, mechanisms by which DNA is made available for uptake may also be regulated. Gonococcal cells acquire DNA that is released by autolysis of other cells or exported from live cells through a type IV secretion system (Dillard and Seifert, 2001; Hamilton *et al.*, 2001; Hamilton *et al.*, 2005). In *Streptococcus pneumoniae*, specific cell-cell signaling during competence development can lead to lysis of non-competent cells, sometimes referred to as allolysis or fratricide (reviewed in Claverys *et al.*, 2006; Johnsborg *et al.*, 2007). It has been suggested that these may be active mechanisms by which bacteria ensure that DNA is available for transformation.

Transduction

Transduction is the transfer of bacterial DNA from one cell to another by a bacteriophage. There are two different types of transduction- generalized and specialized (reviewed in Snyder and Champness, 1997). In generalized transduction, while phages are packaging their own DNA into new phage heads, they mistakenly package some host bacterial DNA as well. The phage containing this mistakenly-packaged DNA can then transfer it to another bacterial cell upon infection. DNA at any position in the bacterial chromosome may be transferred by generalized transduction. Specialized transduction, described below, is very different and only occurs in cells containing phage lysogens.

Temperate bacteriophages, exemplified by lambda, can grow lytically or create lysogens (reviewed in Arber, 1983; and Campbell, 1994; and Davis and Waldor, 2002; and Roberts and Devoret, 1983). During lytic growth, a phage actively replicates and packages its DNA and attempts to infect new hosts (Campbell, 1994). To become a lysogen, a phage incorporates its genome into the host's chromosome, thus becoming a prophage (Arber, 1983; Roberts and Devoret, 1983). Incorporation into the bacterial chromosome usually occurs by site-specific recombination mediated by a phage-encoded integrase (Davis and Waldor, 2002). For lambda, this integrase is encoded by the *int* gene (Arber, 1983). Recombination occurs between a site in the chromosome (*attB*) and a site on the phage (*attP*). This forms two sites on each extreme end of the prophage (*attL* and *attR*). Lambda integrates at one highly preferred *attB* in the *Escherichia coli* chromosome, but it can integrate into any of several secondary attachment sites if its preferred site is deleted (Arber, 1983). For most prophages, regulatory mechanisms ensure that they stay dormant until one or more signals cause them to excise and resume lytic growth. For lambda, excision from the chromosome requires Int and a phage-encoded excisionase (Xis) (Arber, 1983).

At some low frequency, prophage excision can occur by a compromised exchange, resulting in hybrid structures of the phage genome and adjacent segments of the bacterial chromosome (Arber, 1983). Transfer of this bacterial DNA to new hosts by the phage is specialized transduction (Snyder and Champness, 1997). This process differs markedly from generalized transduction, in that only DNA adjacent to the prophage attachment site may be transferred by specialized transduction.

Many temperate bacteriophages carry genes suggesting that they have transferred and/or are continuing to transfer bacterial DNA by specialized transduction. Virulence genes carried by

phages are frequently at one end of the prophage, near *attL* or *attR*, suggesting they may originally have been acquired by imprecise excision from a host (reviewed in Davis and Waldor, 2002). Similarly, many prophages have complete tRNA genes near *attL* or *attR* (Davis and Waldor, 2002). Phages frequently integrate at *attB* sites within tRNA genes; the presence of complete tRNA genes within a phage genome may indicate that it excised imprecisely at one time, taking a complete tRNA sequence in the process (Davis and Waldor, 2002).

Temperate phages also contribute to horizontal gene transfer by virtue of the diverse functions encoded on their genomes. Phage-encoded products include toxins and factors that help pathogenic hosts evade host defenses (reviewed in Davis and Waldor, 2002). Sequence diversity of phages appears to be facilitated by the formation of hybrids of different phages and the moblization of new DNA into phage genomes by transposons (reviewed in Canchaya *et al.*, 2003; reviewed in Davis and Waldor, 2002).

Conjugative plasmids

Conjugative plasmids and ICEs move directly from cell to cell by mating (Burrus *et al.*, 2002; Grohmann *et al.*, 2003). They generally encode the mating machinery required for them to transfer (Burrus *et al.*, 2002; Grohmann *et al.*, 2003). Following excision of an ICE from the chromosome, the mechanisms of transfer of ICEs and conjugative plasmids are very similar. Transfer of conjugative plasmids usually initiates at a specific site, the origin of transfer (*oriT*) (Byrd and Matson, 1997; Lanka and Wilkins, 1995). A relaxase, usually encoded by the plasmid, nicks the *oriT* and covalently attaches itself to the DNA (Byrd and Matson, 1997; Lanka and Wilkins, 1995). The relaxase attached to the DNA can interact with a coupling protein in the bacterial membrane to target the DNA to a transmembrane pore (Lee and Grossman, 2007). The relaxase-DNA complex can be transferred from the donor to the recipient, where the relaxase can

rejoin the plasmid ends to form single-stranded, circular DNA (Lee and Grossman, 2007). A complementary strand is synthesized in the recipient (Parker *et al.*, 2002). Some virulence plasmids contain incomplete DNA transfer systems, suggesting they may have been fully conjugative at one point (reviewed in Davis and Waldor, 2002). In some cases, these can be mobilized by conjugative elements that encode functional mating machinery (reviewed in Davis and Waldor, 2002).

ICEs

ICEs, or conjugative transposons, have characteristics of both phages and conjugative plasmids. Like prophages, they are integrated into the host cell's chromosome, passively replicated along with host cell DNA, and excise in order to transfer (Fig. 3). Like plasmids, ICEs encode conjugation machinery and move directly from cell to cell by mating.

The first identified conjugative transposon was Tn*916*. It was discovered in the late 1970s, when tetracycline resistance was found to transfer from one *Enterococcus faecalis* strain to another without any detectable plasmids (Franke and Clewell, 1981). The transferable resistance was associated with an 18kb DNA region (Clewell *et al.*, 1995). Since it could transpose within and between cells, it was dubbed a "conjugative transposon". Since many similar elements occupy a single chromosomal site, but do not transpose intracellularly, and move in and out of chromosomes by mechanisms distinct from those of other transposons, they have been renamed as ICEs (Burrus *et al.*, 2002).

ICEs range in size from less than 20kb to more than 100 kb (Churchward, 2002; Waldor *et al.*, 1996) (Table 1). Characterized ICEs include elements at each extreme of the size range. Tn*916* is only 18.4 kb, while *clc* is 105 kb (Flannagan *et al.*, 1994; Gawron-Burke and Clewell, 1982; Sentchilo *et al.*, 2003).

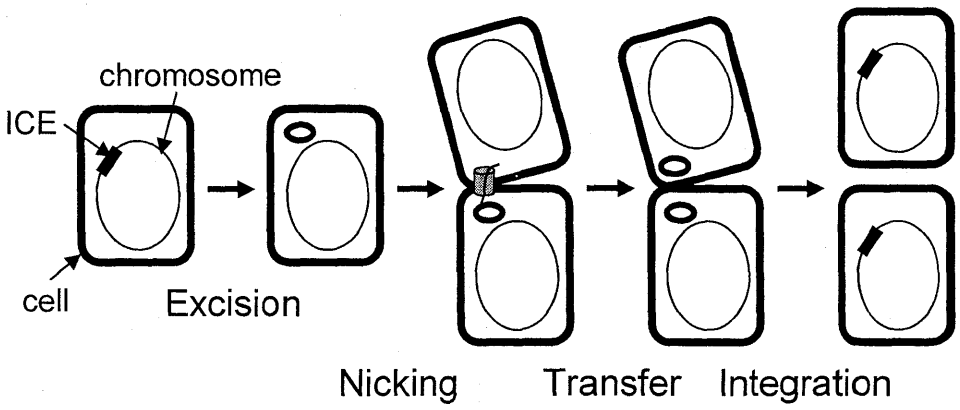


Figure 3. Model of ICE conjugation. An ICE integrated in the chromosome of its host bacterium can excise from the chromosome, and following nicking of its DNA, the ICE can be transferred from the host cell into a recipient by conjugation. Once inside the recipient. the ICE can integrate into that cell's chromosome to be stably maintained. (redrawn from a figure by J.M. Auchtung & A.D. Grossman)

Prior to mating, an ICE excises from the host cell's chromosome to form a circular intermediate (Fig. 3). ICE DNA is then transferred to recipient cells through a conjugative pore. After transfer, the ICE is incorporated into the recipient cell's chromosome. This process is described in more detail below.

Host range

The host ranges of different ICEs seem to vary. Tn*916* has a very broad host range (Scott, 2002). Tn*916*-family elements occur naturally in or have been transferred into more than 52 different species, including some Gram-positive and Gram-negative bacteria (Clewell *et al.*, 1995). ICE*Bs1* can transfer by mating from *Bacillus subtilis* to *Bacillus anthracis*, *Bacillus licheniformis*, or *Listeria monocytogenes* (Auchtung *et al.*, 2007). Transfer of ICE*Bs1* into Gram-negatives has not yet been observed.

In most cases, when an ICE fails to move into a bacterial species, the reasons why it can't mate are unknown. However, for Tn*5397*, mating into *E. coli* was made possible by cloning the ICE's *attB* from *Clostridium difficile* into the *E. coli* chromosome (Wang *et al.*, 2000). Thus in this case, the recombinase's target sequence specificity seemed to be the only barrier to mating into *E. coli*.

Attachment sites

Tn*916* was first identified as a conjugative transposon in *E. faecalis*, because of its ability to move intracellularly to different sites within a chromosome and intercellularly by mating. The intracellular transposition is made possible by site-specific excision and low specificity of integration (reviewed in Burrus *et al.*, 2002). Tn*916 attB* sites in *E. faecalis* are highly variable, but tend to be A-T rich (Scott *et al.*, 1994). Further study of Tn*916* transposition showed that it transposes by a different mechanism than that of type I and type II transposons (Burrus *et al.*,

2002). Furthermore, the attachment sites of Tn*916* depend on the host background (Roberts and Mullany, 2009). Similar variations in specificity of integration in various hosts have been observed for other ICEs (Brown *et al.*, 1988; Brown *et al.*, 1994; Wang *et al.*, 2000; Wang *et al.*, 2006).

ICE*Bs1* has one *attB*, within *trnS-leu2*, in *B. subtilis* (Auchtung *et al.*, 2005; Burrus *et al.*, 2002). In strains lacking this site, it can integrate into any of several secondary attachment sites (Lee *et al.*, 2007). Many ICEs and prophages have attachment sites in highly conserved genes, such as those for tRNAs; this may help these mobile genetic elements to be disseminated into new species and evolve (Williams, 2002).

Excision and integration

During conjugation, an ICE excises from the host cell's chromosome and forms a circular intermediate (Churchward, 2002; Scott *et al.*, 1988). This involves an element-encoded integrase (Int) that catalyzes site-specific recombination for excision and integration (Roberts and Mullany, 2009) Int can mediate recombination between sites at the ends of an integrated element (*attL*, *attR*) to excise the element from the donor's chromosome, or it can mediate recombination between a site on the ICE (*attI* or *attICE*) and one or more sites in the bacterial chromosome (*attB*) to integrate the element into the recipient's chromosome (Burrus and Waldor, 2004). Many ICEs also encode an excisionase (Xis), a recombination directionality factor that promotes excision of the element rather than integration (Burrus and Waldor, 2004; Burrus *et al.*, 2006; Roberts and Mullany, 2009). The integration-excision systems of Tn*916*-related elements are structurally and functionally similar to those of lambdoid phages (Clewell *et al.*, 1995; Poyart-Salmeron *et al.*, 1989, 1990; Scott and Churchward, 1995). Most ICEs encode tyrosine recombinases (Roberts and Mullany, 2009; Wang *et al.*, 2000). However, some ICEs encode

serine recombinases. These recombinases work by a different mechanism and without the help of an excisionase (Wang *et al.*, 2000). One such example is Tn*5397* of *C. difficile*, which apart from its recombinase, closely resembles Tn*916* (Redfield, 1988). In contrast to integrases, excisionases are generally not homologous to each other (Lewis and Hatfull, 2001).

Like other ICEs, to mate, ICE*Bs1* first excises from the chromosome to form a circular intermediate (Auchtung *et al.*, 2007). It encodes its own Int and Xis. Int is required for excision in the donor and site-specific integration in the recipient (Lee *et al.*, 2007). Int must be produced in the recipient for integration (Lee *et al.*, 2007). In contrast, *Tn916*'s Int may be transferred from donor to recipient during mating (Bringel *et al.*, 1992).

Some ICEs are found both integrated in the chromosome and as independently replicating, extrachromosomal elements. Several of these ICEs were originally classified as plasmids, because the plasmid state was detected first. Most AICEs (Actinomycete ICEs) can assume both integrated and freely replicating forms (te Poele *et al.*, 2008). pSAM2 of *Streptomyces ambofaciens* can exist as a single integrated copy or as one integrated copy and 5-10 extrachromosomal copies (Pernodet *et al.*, 1984). In a couple of cases, as described below, the different states of pSAM2 could be attributed to variations in transcription levels of the *pra* gene (Sezonov *et al.*, 1995). For AICEs present in both integrated and freely replicating forms, the copy number of the extrachromosomal versions is typically low (reviewed in te Poele *et al.*, 2008). For example, one plasmid copy of the AICE pMEA300 was detected for every 5-10 host cell chromosomes (te Poele *et al.*, 2008). Another group of elements that can exist in both integrated and free forms is the IncJ plasmids (reviewed in Churchward, 2002). These elements tend to be maintained in the free state only when the cell has one or more integrated copies

(Pembroke and Murphy, 2000). This group includes elements in *Proteus*, *Pseudomonas*, and *Vibrio* species (Coetzee *et al.*, 1972).

Nicking, replication, and DNA transfer

Transfer of bacterial ICEs, like that of conjugative plasmids, generally requires a DNA relaxase (Roberts and Mullany, 2009). The relaxase nicks the element's DNA at a specific sequence, the origin of transfer (*oriT*), to begin DNA transfer (Roberts and Mullany, 2009). ICE*Bs1* encodes its own relaxase, NicK, that specifically nicks the *oriT* at a site within the *nicK* gene (Lee and Grossman, 2007). Excision of the element is not required for nicking (Lee and Grossman, 2007). Relaxases of conjugative plasmids are thought to be covalently bound to DNA after nicking, such that they can mediate contacts between the DNA and the mating machinery, remain bound during passage through the mating pore, and re-join the elements' ends once inside the recipient (Lee and Grossman, 2007& refs therein). It seems likely that ICE relaxases act similarly.

Relaxases of conjugative plasmids and ICEs are thought to interact with coupling proteins associated with the mating pore (Lee and Grossman, 2007& refs therein). For ICE*Bs1*, *ydcQ* appears to encode an FtsK/SpoIIIE homolog, which might be the coupling protein (Lee and Grossman, 2007). FtsK/SpoIIIE-family proteins are DNA translocases with distant homology to coupling proteins of conjugative plasmids (Lee and Grossman, 2007& refs therein). Proteins with FtsK/SpoIIIE domains also participate in transfer of AICEs (te Poele *et al.*, 2008).

ICE*Bs1* can replicate autonomously by a plasmid-like rolling circle mechanism (Lee *et al.*, 2009). Replication is required for maintenance of the element, but not for conjugation (Lee *et al.*, 2009). This replication initiates at the *oriT* (Lee *et al.*, 2009). Replication requires the ICE-encoded relaxase, NicK, and the host-encoded proteins PolC, DnaN, and PcrA (Lee *et al.*, 2009).

Of these, only NicK and PcrA helicase are required for ICE*Bs1* mating (Lee *et al.*, 2009). PcrA might be needed to unwind the DNA enabling transfer of a single strand (Lee *et al.*, 2009).

Characterization of the mating proteins of ICE*Bs1* is in progress. One protein that probably forms part of the conjugation machinery is the element-encoded ConE (Berkmen *et al.*, 2009). ConE is required for mating, and during induction of ICE*Bs1*, ConE localizes to the cell poles along with ICE*Bs1* DNA (Berkmen *et al.*, 2009). The ATPase domains of ConE are required for mating, but not for localization (Berkmen *et al.*, 2009). At least one other ICE*Bs1* protein is required for the polar localization of ConE (Berkmen *et al.*, 2009).

For most ICEs and conjugative plasmids, mating most likely involves transfer of a single strand of DNA (Lanka and Wilkins, 1995; Llosa *et al.*, 2002; Scott *et al.*, 1994; Waters and Guiney, 1993). However, for the AICE pSAM2, mating experiments have indicated that double stranded DNA is transferred (Possoz *et al.*, 2001). For ICEs where ssDNA is transferred, this DNA may be replicated prior to integration in the recipient. For ICE*Bs1*, synthesis of the complementary strand is most likely required for *int* to be expressed, so that Int can mediate site-specific integration into the chromosome (Lee and Grossman, 2007). The strand of ICE*Bs1* that is probably transferred does not encode Int, and most promoters are only active as double-stranded DNA (Lee and Grossman, 2007; Masai and Arai, 1997).

Some ICEs may have mechanisms to protect their DNA following transfer. Tn*916*-like ICEs encode putative anti-restriction functions and have few restriction sites (Flannagan *et al.*, 1994).

Generation of diversity

Since ICEs move by site-specific excision and integration, it may not be immediately apparent how they could help to generate diversity. However, some ICEs might increase genetic

diversity by transferring part of their host chromosomes. ICEs, such as CTnERL and SXT, can transfer chromosomal markers by an Hfr-type mechanism (Hochhut *et al.*, 2000; Whittle *et al.*, 2006). Several AICEs can mobilize chromosomal loci (te Poele *et al.*, 2008). This occurs when the element is nicked and transferred without excising from the chromosome.

The mutability of ICE-carried sequences may be demonstrated by ICEs that are amalgamates of multiple mobile elements. Tn*5253*, found in *S. pneumoniae*, is the composite of two different ICES- Tn*5251* and Tn*5252* (Ayoubi *et al.*, 1991; Buu-Hoi and Horodniceanu, 1980; Provvedi *et al.*, 1996). Some Tn*916*-like ICEs harbor non-conjugative transposons, such as Tn*917*, and macrolide efflux genetic assembly (MEGA) elements (Roberts and Mullany, 2009& refs therein). In these cases, sequence diversity of the elements and their hosts is being promoted.

Mobile element-encoded functions

The significance of HGT is underscored by the variety of functions that have been found to be mobilizable from cell to cell. One prominent, significant type of function that is propagated by HGT is antibiotic resistance. Increased understanding of regulation and mechanisms of transfer seems most urgent when considering the spread of antibiotic resistance genes among pathogens. For example, SXT, an ICE of *Vibrio cholerae*, and its close relatives encode resistances to antibiotics such as sulfamethoxazole and trimethoprim that were previously used to treat cholera (Waldor *et al.*, 1996). The ICE Tn*916* was first identified because of its ability to mobilize tetracycline resistance (Clewell *et al.*, 1995; Flannagan *et al.*, 1994; Gawron-Burke and Clewell, 1982). Other Tn*916* family members confer resistance to other drugs, such as kanamycin and macrolides (Courvalin and Carlier, 1986). One such family member, Tn*1545*, was found to encode resistance to tetracycline, kanamycin, and erythromycin in *S. pneumoniae* (Courvalin and Carlier, 1986; Gawron-Burke and Clewell, 1982). Tn*5386* contains an operon

with homology to genes encoding lantibiotic immunity (Rice *et al.*, 2007). Some relatives of Tn*916*, such as Tn*6000* and Tn*916S*, carry a different tetracycline resistance gene than the one found in Tn*916* (Lancaster *et al.*, 2004; Roberts *et al.*, 2006). Transfer of antibiotic resistance determinants from one cell to another is not limited to ICEs, or to mobile elements. *Streptococcus pyogenes* has acquired fluoroquinolone resistance from *Streptococcus dysgalactiae* by natural transformation (Pletz *et al.*, 2006).

Some mobile elements carry genes for antibiotic synthesis. For example, Tn*5276* and similar ICEs enable cells to make the lantibiotic nisin (Rauch and De Vos, 1992; Rauch *et al.*, 1994; Rauch and de Vos, 1994).

The study of horizontal gene transfer also overlaps with attempts to characterize the causative agents of disease, because many virulence factors are encoded by plasmids and phages (reviewed in Davis and Waldor, 2002). Plasmids can encode an assortment of virulence factors, such as toxins, pili, and typeIII secretion systems (Davis and Waldor, 2002). Phage-encoded virulence factors include toxins, proteins to help bacteria escape host defenses, and proteins that alter the bacterial surface so as to interfere with detection by the host organism (Davis and Waldor, 2002). The first discovery of a phage-encoded virulence factor was documented in 1951, when diphtheria toxin was found to be encoded by beta-phage of *Corynebacterium diphtheriae* (Freeman, 1951). In some cases, the toxin genes within phage genomes are expressed during lysogeny and regulated independently of the phage life cycle (Davis and Waldor, 2002). In others, toxin production is coupled to the phage life cycle. For example, in Shiga toxin (Stx)-encoding phages of enterohaemorrhagic *E. coli* (EHEC), Stx genes are transcribed from a late-phage promoter (Neely and Friedman, 1998; Plunkett *et al.*, 1999; Wagner *et al.*, 2001). Lysogens of these phages are induced by antibiotics that stall DNA synthesis, such as quinolones

(Muhldorfer *et al.*, 1996; Zhang *et al.*, 2000). Understanding the regulation of mobile genetic elements may be particularly important in instances such as this, where virulence factors are produced when horizontal transfer is stimulated, and agents used to treat infection may trigger increases in virulence.

Mobile elements can help their hosts survive adverse conditions other than the presence of antibiotics and host cell defenses. Tn*5252* can help its host Streptococci, survive UV damage (Munoz-Najar and Vijayakumar, 1999). The best studied host, *S. pneumoniae* does not have a canonical SOS response (Gasc *et al.*, 1980). However, after developing natural competence *S. pneumoniae* does induce some genes homologous to those that are part of the SOS response in other bacteria (Prudhomme *et al.*, 2006). Tn*5252* helps its host survive severe damage, because it encodes an error-prone DNA repair system, homologous to UmuDC of *E. coli* (Munoz-Najar and Vijayakumar, 1999). Some IncJ plasmids confer mercury resistance to their hosts (Coetzee *et al.*, 1972). pNP40, a conjugative plasmid of *Lactococcus lactis*, endows its host with a veritable powerhouse of assets. In addition to nisin resistance, pNP40 also encodes cadmium resistance, putative cold-shock proteins, components of DNA repair systems, and phage resistance mechanisms (O'Driscoll *et al.*, 2006 & refs therein).

Some mobile elements change the conditions under which their hosts can grow, by allowing them to metabolize available compounds or live symbiotically with other organisms. Several strains of *E. coli* and *Salmonella* have sucrose fermentation genes on conjugative plasmids (Hochhut *et al.*, 1997; Schmid *et al.*, 1988; Smith and Parsell, 1975; Wohlhieter *et al.*, 1975). Some ICEs carry genes for a phosphotransferase-dependent sucrose fermentation pathway. Such systems are found on Ctn*scr94* of *Salmonella senftenberg* 5494-57 and Tn*5276* of *L. lactis* (Hochhut *et al.*, 1997; Rauch and De Vos, 1992; Thompson *et al.*, 1991). Some ICEs

confer complex degradation pathways for certain chemicals (Nishi *et al.*, 2000; Ravatn *et al.*, 1998; Ravatn *et al.*, 1998; Toussaint *et al.*, 2003). For example, products encoded by the *clc* element enable host cells to metabolize a number of toxic aromatic compounds (Ravatn *et al.*, 1998). Other ICEs allow their hosts to fix nitrogen (Sullivan and Ronson, 1998). The symbiosis island of *Mesorhizobium loti* enables its symbiotic growth with plant roots by encoding products involved in nitrogen fixation and symbiosis (Sullivan *et al.*, 2002).

It is not yet known what advantage, if any, ICE*Bs1* provides to its hosts.

Signals that regulate HGT

Rationale

In most cases, regulation of HGT permits transfer under certain conditions. Presumably, these systems evolved to balance between constitutive transfer and no transfer, because either extreme would be suboptimal for both mobile elements and their hosts. Mobile elements benefit from transfer in that they spread their DNA, and hosts often benefit by virtue of some function that is encoded by the acquired DNA. At the other extreme, unchecked expression of transfer functions has deleterious effects on host cells in various systems (Auchtung *et al.*, 2007; Beaber *et al.*, 2002; reviewed in Holcik and Iyer, 1997). In some cases, genes encoding transfer functions were originally dubbed *kil* genes, because their unregulated expression resulted in host cell death (reviewed in Holcik and Iyer, 1997). If an element were to kill its host by continually attempting to transfer, when the odds of success were very low, both movement to new cells and propagation within host progeny could fail, and the element could be lost.

The types of signals that induce HGT seem to reduce the burden placed on the host while allowing for optimal transfer of the element. Stressful conditions, like DNA damage, might induce a mobile element so that it could escape an unstable environment. In some organisms,

DNA damage or other stresses, can induce competence development (reviewed in Claverys *et al.*, 2006). This allows a cell to take up DNA, which could be used as a repair template, as a source of genetic diversity, or for food (reviewed in Claverys *et al.*, 2006; reviewed in Johnsborg *et al.*, 2007). Where the presence of an antibiotic induces an element encoding resistance to that antibiotic, or presence of a molecule induces an element encoding products that metabolize that molecule, cells benefit from the function encoded by the element and the element benefits by spreading its own DNA.

DNA damage

DNA damage induces horizontal gene transfer in various systems. DNA-damaging treatments such as UV-irradiation or addition of mitomycinC (MMC) induce the SOS response in different types of bacteria (Little and Mount, 1982; Marrero and Yasbin, 1988). DNA damage increases the amount of single-stranded DNA in the cell. By binding to single-stranded DNA, RecA protein is activated to promote the autoproteolytic cleavage of LexA repressor and the consequent transcription of SOS genes (Little and Mount, 1982). Activated RecA also promotes the cleavage of many prophage repressors, including the canonical example of phage lambda's CI repressor (Little, 1984; Roberts and Devoret, 1983; Susskind and Youderian, 1983). In some instances, such as with cyanophage and Stx-encoding phages of *E. coli*, UV and MMC induce lysogens, but the mechanism is unknown (Muhldorfer *et al.*, 1996; Williamson *et al.*, 2002; Zhang *et al.*, 2000). DNA damage, caused by treatment with MMC or ciprofloxacin induces SXT, an ICE from *V. cholerae,* by a similar mechanism to that of lambda phage (Beaber *et al.*, 2004).

ICE*Bs1* gene expression, excision, and mating are also stimulated upon MMC treatment, by a mechanism that is RecA-dependent and involves cleavage of the ICE*Bs1* repressor, ImmR

(Auchtung *et al.*, 2005; Auchtung *et al.*, 2007). However, ICE*Bs1* differs from the aforementioned systems in that ImmR does not cleave itself but is cleaved by the antirepressor, ImmA (Bose* *et al.*, 2008) (Fig. 2). Induction of ICE*Bs1* by DNA damage is independent of regulation resulting from cell-cell signaling. It does not require RapI or PhrI (Auchtung *et al.*, 2005).

Another noteworthy example of HGT induction by DNA-damaging agents is the development of natural competence in *S. pneumoniae*. *S. pneumoniae* is thought not to have an SOS system like those well characterized in other bacteria, as no SOS regulator with LexA's autocleavage signature has been found (Claverys *et al.*, 2006; Gasc *et al.*, 1980). Rather, *S. pneumoniae*'s response to DNA damage is thought to be regulated by the same factors that govern development of natural competence. Treatment of *S. pneumoniae* with MMC or certain antibiotics that induce the SOS response in other bacteria results in induction of *ssb* expression and high levels of genetic transformation (reviewed in Claverys *et al.*, 2006; Prudhomme *et al.*, 2006). MMC increases *recA* expression in *S. pneumoniae* in a manner that depends on intact competence regulation (Prudhomme *et al.*, 2006). In contrast, induction of the *com* regulon does not depend on *recA* (Claverys *et al.*, 2006).

Cell-cell signaling

ICE*Bs1* is regulated in response to a cell-cell signaling system that can activate transfer of the element when host cells are crowded by potential recipients (Auchtung *et al.*, 2005). This mode of regulation is independent of induction by DNA damage and does not require *recA* (Auchtung *et al.*, 2005) (Fig. 2). The intercellular signaling system is comprised of RapI and PhrI, which are encoded by the element (Auchtung *et al.*, 2005) (Fig. 1). *B. subtilis* has several *rap*s and *phr*s. Phrs are one type of signaling peptide produced by *B. subtilis* (Lazazzera, 2001).

Phrs are secreted by cells, and imported through the Opp permease after accumulating extracellularly (Auchtung *et al.*, 2005) (Fig. 4). Upon entering the cell, Phr peptides directly inhibit the activities of cognate Rap proteins (Auchtung *et al.*, 2005). Characterized Raps directly or indirectly inhibit the activities of transcription factors that affect sporulation, competence, degradative enzyme production, and antibiotic synthesis (Auchtung *et al.*, 2005).

Several Bacillus plasmids and phages contain *rap-phr* cassettes (Auchtung *et al.*, 2005). Rap60 and Phr60 of pTA1060 are involved in degradative enzyme production (Koetje *et al.*, 2003). The functions of other mobile element-encoded Raps and Phrs are unknown, with the exception of RapI and PhrI of ICE*Bs1* (Auchtung *et al.*, 2005).

Transcription of *rapI* and *phrI* is repressed by AbrB, directly or indirectly (Auchtung *et al.*, 2005) (Fig. 2). AbrB represses transcription of many target genes in *B. subtilis* during exponential growth (Phillips and Strauch, 2002). AbrB-mediated repression is relieved under nutrient-limiting conditions and at high cell density (Phillips and Strauch, 2002). Under these conditions, *rapI* is transcribed. RapI activates ICE*Bs1* gene expression, excision, and mating, unless it is antagonized by PhrI (Auchtung *et al.*, 2005). Thus RapI can activate ICE*Bs1* mating when host cells are crowded by cells that do not contain the element and consequently do not produce PhrI (Auchtung *et al.*, 2005).

Cell-cell signaling regulates transfer of mobile genetic elements in many other systems. The mobilization of Ti plasmids in *Agrobacterium tumefaciens* is controlled by a plasmid-encoded signal received from plant cells after those cells have been transformed with plasmid-derived T-DNA (Fuqua and Winans, 1994; Winans *et al.*, 1999; Zhu *et al.*, 2000). Chromosome-encoded signaling peptides of *E. faecalis* stimulate transfer of specific conjugative plasmids (Chandler and Dunny, 2004; Clewell, 1993; Dunny, 2007). Plasmid-encoded peptides made by

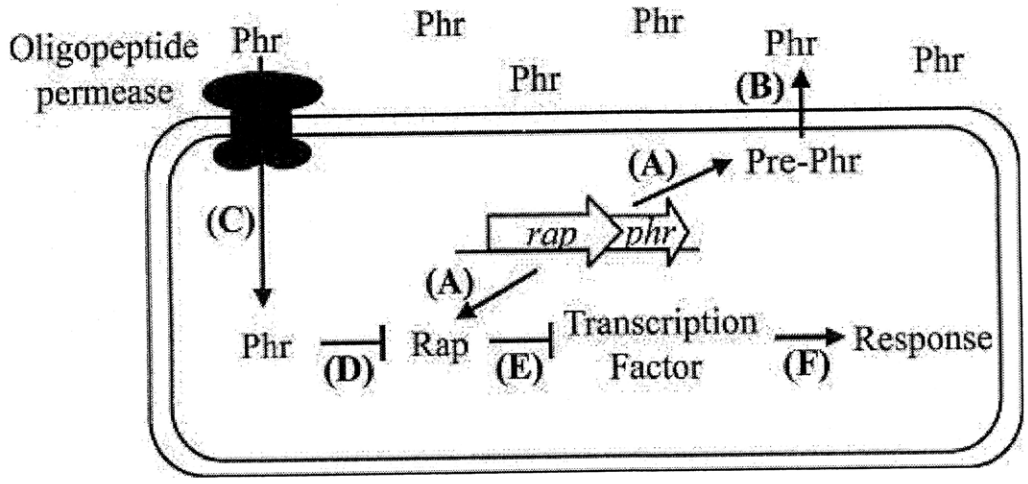


Figure 4. Phr peptide signaling in *B. subtilis*. (A) *rap* and *phr* genes are transcribed and translated; (B) pre-Phr peptides are secreted and processed ; (C) mature Phr peptides are transported into the cell by the Opp; (D) once inside the cell, Phr peptides inhibit the activities of regulators known as Rap proteins; (E) each characterized Rap protein inhibits the activity of a transcription factor, either directly or indirectly; (F) and inhibition of transcription factors lead to cellular responses.(Figure & legend from Auchtung et al., 2005)

E. faecalis host cells inhibit transfer of new plasmids (Chandler and Dunny, 2004; Clewell, 1999).

Intercellular communication also affects horizontal gene transfer in some bacteria by affecting the development of natural competence. Competence-stimulating peptides (CSPs) control competence development in *S. pneumoniae* (reviewed in Claverys *et al.*, 2006; reviewed in Johnsborg *et al.*, 2007). In *B. subtilis*, competence development is regulated in response to the secreted peptide, ComX (reviewed in Claverys *et al.*, 2006; Hamoen *et al.*, 2003; Magnuson *et al.*, 1994; Okada *et al.*, 2005). In both *S. pneumoniae* and *B. subtilis*, the peptide pheromone is detected by a dedicated two-component system that transmits the signal so as to lead to transcriptional activation of the competence (*com*) regulon . Despite the parallels between the two systems, there are some obvious differences. ComX is isoprenylated before export, while CSP is not modified (reviewed in Claverys *et al.*, 2006). Also, the response regulators in each of the two component systems are not homologous; they have distinct DNA-binding domains (reviewed in Claverys *et al.*, 2006). In *B. subtilis*, this process is also influenced by PhrC and PhrF, which may communicate information about growth phase and external pH (Bongiorni *et al.*, 2005; Pottathil and Lazazzera, 2003; Solomon *et al.*, 1996).

Antibiotics

Antibiotics induce horizontal gene transfer in several systems. In some cases, an element encoding resistance to an antibiotic is induced to transfer in the presence of that antibiotic. For example, tetracycline increases transfer frequencies of the tetracycline-resistance-conferring ICEs CTnDOT, Tn*916*, Tn*1545*, and other Tn*916*-like elements (Cheng *et al.*, 2000; Cheng *et al.*, 2001; Doucet-Populaire *et al.*, 1991; Rice *et al.*, 1992; Salyers *et al.*, 1995; Shoemaker and Salyers, 1988; Showsh and Andrews, 1992; Stevens *et al.*, 1990; Torres *et al.*, 1991). In other

cases, antibiotics promote HGT more generally; the transferred DNA may enhance host cell survival but does not provide specific resistance to that antibiotic. For example, aminoglycoside antibiotics and fluoroquinolones stimulate competence development in *S. pneumoniae* (Prudhomme *et al.*, 2006).

Other signals

A wide variety of other factors can impact horizontal gene transfer. Competence development in *S. pneumoniae* is influenced by a two-component system that monitors cell wall integrity and a stress-related global regulator, in addition to cell-cell signaling (Guenzi *et al.*, 1994; Saskova *et al.*, 2007). Cyanophage lysogens are induced by copper or cadmium in marine environments and by copper in freshwater (Lee *et al.*, 2006; Williamson *et al.*, 2002). Integrase expression of *clc*, an ICE, is stimulated by 3-chlorobenzoate, which is metabolized by enzymes encoded on *clc* (Sentchilo *et al.*, 2003). Replication and presumably conjugation of pMEA300, an AICE of *A. methanolica*, is activated by autoclaved sucrose or fructose (Vrijbloed *et al.*, 1995).

In some cases, general conditions that activate transfer of mobile genetic elements have been described, but the exact signals are not yet known. Some prophages, such as CTXϕ of *V. cholerae* and Stx-encoding phages of EHEC, are induced when the bacteria infect a host (Acheson *et al.*, 1998; Kimsey and Waldor, 1998; Lawrence and Ochman, 1998). Marine prophage lysogens are induced by pollutants and seasonal changes (Jiang and Paul, 1998; Williamson *et al.*, 2002; Wilson *et al.*, 1993).

Mechanisms of regulation of HGT

Phages

Autocleaving repressors

Induction of many prophage lysogens is mediated by autocleavage of the primary repressor protein in response to DNA damage. The canonical example of this is CI repressor of phage lambda. Phages that are similarly regulated include 434 and P22 (Roberts and Devoret, 1983).

Lambda lysogens are induced by agents that perturb DNA replication by producing DNA lesions or by preventing the replication fork from functioning normally (reviewed in Roberts and Devoret, 1983). Under these conditions, RecA protein becomes activated by binding single-stranded DNA (Roberts and Devoret, 1983). ATP-binding, but not hydrolysis are also required for RecA activity (Roberts and Devoret, 1983). Activated RecA stimulates the autocleavage of LexA repressor and lambda's CI repressor (Little, 1984; reviewed in Little, 1993; reviewed in Roberts and Devoret, 1983). Autoproteolysis of LexA relieves repression of SOS genes, including *recA* (Little and Mount, 1982). Autoproteolysis of CI begins a lambda prophage's transition to lytic growth (reviewed in Kim and Little, 1993; Little, 1984; reviewed in Roberts and Devoret, 1983).

CI repressor maintains lysogeny and confers immunity against superinfection (reviewed in Arber, 1983). CI of lambda has two domains. Its N-terminal domain binds DNA, and its C-terminal domain allows oligomerization and autocleavage (reviewed in Kim and Little, 1993). The cleavage site is in a linker between the two domains (Kim and Little, 1993). In lambda CI, the site is between Ala and Gly (Kim and Little, 1993). In homologous repressors, it is either between Ala and Gly or between Cys and Gly (reviewed in Roberts and Devoret, 1983). The

residues responsible for autocleavage are also conserved. In lambda CI, they are S149 and K192 (Lobocka *et al.*, 1996).

Phage anti-repressors

Induction of many phage lysogens and of ICE*Bs1* is regulated by a repressor and an anti-repressor. However, characterized phage anti-repressors function differently than the ICE*Bs1* anti-repressor ImmA. In most characterized phage systems, anti-repressors are thought to somehow interfere with the ability of the repressor to bind to operator sites. It seems likely that in most of these systems, the anti-repressor functions by directly binding to the repressor.

Coliphage 186 is regulated by a repressor (CI) and anti-repressor (Tum). Expression of *tum* is repressed by LexA (Brumby *et al.*, 1996; Lamont *et al.*, 1989). During the SOS response, LexA autocleavage allows expression of *tum* (Brumby *et al.*, 1996; Lamont *et al.*, 1989). Tum then interferes with CI to allow lytic growth. Tum does not cleave or cause autocleavage of CI, since Tum's antagonism of CI is reversible (Shearwin *et al.*, 1998). Tum is thought to bind CI and affect its activity so as to allow lytic growth of coliphage 186 (Shearwin *et al.*, 1998). *Salmonella enterica* phage Fels-2 harbors a Tum homolog and is regulated similarly to coliphage 186 (Bunny *et al.*, 2002).

Linear phage N15 seems to be regulated nearly identically to coliphage 186. In N15, SOS relieves LexA repression of *antC*, allowing production of the AntC anti-repressor (Mardanov and Ravin, 2007). AntC then counteracts the activity of N15's primary repressor CB (Lobocka *et al.*, 1996; Mardanov and Ravin, 2007). CB resembles lambda CI repressor in some respects, but it lacks the conserved cleavage site and catalytic residues (Lobocka *et al.*, 1996). AntC and CB interact in a bacterial two-hybrid test (Mardanov and Ravin, 2007). It has been proposed that

AntC antagonizes CB, because together they form complexes that can't bind DNA (Mardanov and Ravin, 2007).

The same model has been put forward for several other phage systems. The satellite phage P4 can induce lytic growth of P2 (Geisselsoder *et al.*, 1981; Liu *et al.*, 1997). It has been suggested that protein E of P4 and the repressor protein C of P2 form multisubunit complexes, such that C can no longer bind to operator sequences (Liu *et al.*, 1998).

Phage P22 has an anti-repressor, Ant, that seems to function by a similar mechanism to other phage anti-repressors. The primary repressor, C2, of phage P22 can be reversibly antagonized by Ant (Susskind and Botstein, 1975; Susskind and Youderian, 1983). However, this anti-repressor is not involved in prophage induction in response to DNA damage. Mutating *ant* in P22 does not prevent prophage induction by DNA-damaging agents (Susskind and Youderian, 1983). Induction by DNA-damaging agents involves RecA-stimulated cleavage of C2 (Susskind and Youderian, 1983). Phage P22's Ant can interfere with the ability of lambda phage's CI to bind DNA (Susskind and Youderian, 1983). And P22 Ant antagonizes lambda CI mutants that are insensitive to RecA-mediated induction, demonstrating that Ant and RecA affect phage repressors in different ways (Susskind and Youderian, 1983).

ICEs

ICEBs1

ICE*Bs1* is regulated by cell-cell signaling and DNA damage. Either activated RecA or RapI, when produced and not antagonized by PhrI, somehow activate ImmA-catalyzed cleavage of ImmR. Proteolysis of ImmR relieves repression of the *xis* promoter, allowing expression of *xis* and many other genes involved in ICE*Bs1* transfer.

P*xis* reflects ICE*Bs1* induction; low basal levels of transcription from this promoter are highly induced upon RapI overproduction or MMC treatment (Auchtung *et al.*, 2007). P*xis* is repressed by the ICE*Bs1* repressor, ImmR (Auchtung *et al.*, 2007). ImmR limits the ability of cells in which it is produced to acquire ICE*Bs1* (Auchtung *et al.*, 2007). Deletion of *immR* causes ICE*Bs1* to be lost at a relatively high frequency and it causes increased lysis of host cells, perhaps owing to unchecked production of mating pore components (Auchtung *et al.*, 2007).

ImmR is homologous to CI repressor of lambda phage (Auchtung *et al.*, 2005; Burrus *et al.*, 2002). Like CI, ImmR regulates its own transcription and is cleaved under conditions that activate RecA (Auchtung *et al.*, 2007; Dodd *et al.*, 2005; Little, 1984). However, ImmR does not contain the conserved cleavage site characteristic of CI-like, autocleaving repressors. Instead ImmR is cleaved by the anti-repressor, ImmA (Bose* *et al.*, 2008). Proteolysis of ImmR by ImmA is activated, directly or indirectly, by RapI or activated RecA (Bose* *et al.*, 2008). The mechanism of this activation is unknown.

Other DNA damage-induced ICEs

The ICE SXT of *V. cholerae* is also induced by SOS, but in a similar manner to lambda prophages. The SXT repressor, SetR, is homologous to CI of lambda (Beaber *et al.*, 2004; Beaber and Waldor, 2004). SetR regulates its own transcription and that of *setC* and *setD*, which are oriented divergently from *setR* (Beaber *et al.*, 2004; Beaber and Waldor, 2004). SetC and SetD activate transcription of the *int* and *tra* operons that encode proteins required for SXT transfer. Although autocleavage of SetR has not been observed, mutating the conserved Ala-Gly cleavage sequence in SetR renders SXT uninducible by SOS (Beaber *et al.*, 2004).

Three ICEs of *Streptococcus thermophilus* are induced by DNA damage and appear to contain regulatory components similar to those of lambda phage and ICE*Bs1* (Bellanger *et al.*,

2007). The *arp1* gene of each *S. thermophilus* ICE encodes a CI-homologue with the conserved DNA-binding and autocleavage domains (Bellanger *et al.*, 2007). MMC treatment results in increased excision of these ICEs (Bellanger *et al.*, 2007). In addition to Arp1, each of the *S. thermophilus* ICEs encodes two proteins that appear to resemble ImmR and ImmA of ICE*Bs1* (Bellanger *et al.*, 2007). Arp2 is a CI-like repressor that lacks the autoproteolytic domain, and OrfQ contains the conserved domain of unknown function and the HEXXH, putative-zinc-binding motif of ImmA (Bellanger *et al.*, 2007). The functions of Arp2 and OrfQ in the *S. thermophilus* ICEs have not been demonstrated, but their inclusion in an element with a functional lamba CI-like repressor is intriguing. Perhaps, having two different regulatory systems enables a single ICE to respond to complex environmental stimuli in a more sophisticated manner than an ICE with either single regulatory system.

Tetracycline-induced ICEs

The well-characterized ICEs CtnDOT and Tn*916* are both induced by tetracycline, and they both harbor tetracycline resistance genes. However, the mechanisms of regulation of these two ICEs appear to be very different from each other. CtnDOT is likely controlled by translational attenuation, while Tn*916* regulation involves transcriptional attenuation and requires circularization of the element.

Stimulation of CTnDOT excision and transfer is mediated by the products of an operon containing the tetracycline-resistance gene *tetQ*, *rteA*, and *rteB* (Shoemaker *et al.*, 1989; Stevens *et al.*, 1990; Stevens *et al.*, 1993). The transcription of this operon does not change in the presence of tetracycline (Wang *et al.*, 2004). Rather, hairpins in the mRNA leader sequence ahead of *tetQ* prevent downstream translation (Wang *et al.*, 2004). Tetracycline-dependent ribosomal stalling may relieve this translational blockage by allowing alternative mRNA

structures to form (Wang *et al.*, 2004). Destabilizing the most readily formed hairpins by mutating the *tetQ* leader does enable the formation of alternative structures (Wang *et al.*, 2005). The products of the genes downstream of *tetQ*, *rteA* and *rteB*, resemble a two component system (Cheng *et al.*, 2001; Stevens *et al.*, 1990; Stevens *et al.*, 1992; Stevens *et al.*, 1993; Whittle *et al.*, 2002). Together, RteA and RteB activate production of RteC, which activates genes required for excision and transfer (Moon *et al.*, 2005). Interruption of *rteA*, *rteB*, or *rteC* prevents CTnDOT excision (Cheng *et al.*, 2001; Stevens *et al.*, 1993).

Transfer of Tn*916* requires excision and circularization of the element, but tetracycline does not stimulate excision (Celli *et al.*, 1997; Celli and Trieu-Cuot, 1998). Instead, tetracycline appears to affect transcription of *tetM*, the gene for the resistance determinant, and transfer genes by relieving transcriptional attenuation (Celli and Trieu-Cuot, 1998; Su *et al.*, 1992). The presence of tetracycline allows transcription to read through a point where it would otherwise terminate and thus strongly increases expression of *orf7* and *orf8* (Celli and Trieu-Cuot, 1998). The products of *orf7* and *orf8* activate their own transcription and that of downstream genes (Celli and Trieu-Cuot, 1998). The downstream genes include *xis*, *int*, and chromosomal genes if the element is integrated or transfer genes if the element has excised and circularized (Celli and Trieu-Cuot, 1998).

pSAM2

The AICE pSAM2 encodes KorSA, a GntR-family repressor that represses its own transcription and that of *pra* (Sezonov *et al.*, 1998; Sezonov *et al.*, 2000). Pra activates the operon containing *rep*, *xis*, and *int* (Sezonov *et al.*, 1998; Sezonov *et al.*, 2000). Levels of Pra in pSAM2 govern the state of pSAM2. pSAM2 can be present in *S. ambofaciens* as a single integrated copy, as is $pSAM2_{B2}$ (Pernodet *et al.*, 1984). Or the element can exist as one

integrated copy and 5-10 extrachromosomal, autonomously replicating copies, as are $pSAM2_{B3}$ and $pSAM2_{B4}$ (Pernodet *et al.*, 1984). A point mutation in the *pra* promoter allows for increased production of the Pra activator in $pSAM2_{B3}$ than in $pSAM2_{B2}$ (Sezonov *et al.*, 1995). This mutation does not appear to increase *pra* transcription by interfering with KorSA repression (Sezonov *et al.*, 2000). Artificial overexpression of *pra* can cause $pSAM2_{B2}$ to be maintained in both the integrated and plasmid state or as a plasmid with no integrated copy (Sezonov *et al.*, 1995).

Conclusion & thesis outline

HGT is an important aspect of bacterial biology. Work aimed at understanding the regulation of HGT has elucidated some conserved signals and mechanisms and promises to reveal new and different ways of controlling HGT in the future.

The regulatory system of ICE*Bs1* resembles previously characterized regulatory systems of mobile genetic elements in some ways, but in many respects, it is different and exciting. This thesis focuses on the role of the ICE*Bs1* anti-repressor, ImmA, in regulation of the element. ImmA is particularly interesting because of the abundance of as-yet-uncharacterized ImmA homologs, many of which are found in known or putative mobile genetic elements.

Chapter 2 describes how ImmA cleaves ImmR to derepress gene expression, excision, and mating of the ICE*Bs1*. This chapter was published in *Molecular Microbiology* in 2008, with Jennifer M. Auchtung as co-first author. Information from Chapter 2 is cited in other parts of this thesis with references to the published paper. Appendix A contains data supporting the idea that ImmA is a metal-dependent protease; these data have not yet been published. Appendix B summarizes the results of a recent search for ImmA and ImmR homologs, providing a more current version of the information in Table 1 of Chapter 2.

Chapter 3 details ways in which ImmA can be activated to cleave ImmR. Overproduction of ImmA derepresses P*xis*, but RapI causes derepression of P*xis* without increasing ImmA levels. Mutations that render ImmA hyperactive reveal that the protein can be altered so as to increase its abundance in vivo or its specific activity. Chapter 3 is currently being prepared for submission.

Appendix C summarizes experiments suggesting that ImmA may be readily degraded by one or more ATP-dependent proteases, and that stabilizing ImmA by deleting these proteases has only minor effects on ICE*Bs1* gene expression.

Appendix D details the isolation and characterization of four uninducible mutants of ImmR (ImmR(ind-)). Four mutations were found to confer this phenotype. These mutations flank the site at which ImmA cleaves ImmR, highlighting the importance of this sequence for proteolysis of ImmR.

Chapter 4 summarizes findings presented in this thesis and considers what questions might next be posed with respect to the regulation of ICE*Bs1*.

a note about redundancy: Chapters 2 and 3 were prepared for publication outside of this thesis. Because of this, the introduction and discussion sections of those chapters are somewhat redundant with parts of the thesis introduction (this chapter) and the thesis discussion (Chapter 5). In addition, Appendices C and D include introduction sections that briefly summarize relevant information about ICE*Bs1* that may be described in other sections of the thesis.

References

Acheson, D.W., Reidl, J., Zhang, X., Keusch, G.T., Mekalanos, J.J., and Waldor, M.K. (1998) In vivo transduction with shiga toxin 1-encoding phage. *Infect Immun* **66**: 4496-4498.

Arber, W. (1983) A beginner's guide to lambda biology. In *Lambda II.* Hendrix, R.W. (ed). Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory, pp. 381-394.

Auchtung, J.M., Lee, C.A., Monson, R.E., Lehman, A.P., and Grossman, A.D. (2005) Regulation of a Bacillus subtilis mobile genetic element by intercellular signaling and the global DNA damage response. *Proc Natl Acad Sci U S A* **102**: 12554-12559.

Auchtung, J.M., Lee, C.A., Garrison, K.L., and Grossman, A.D. (2007) Identification and characterization of the immunity repressor (ImmR) that controls the mobile genetic element ICEBs1 of Bacillus subtilis. *Mol Microbiol* **64**: 1515-1528.

Ayoubi, P., Kilic, A.O., and Vijayakumar, M.N. (1991) Tn5253, the pneumococcal omega (cat tet) BM6001 element, is a composite structure of two conjugative transposons, Tn5251 and Tn5252. *J Bacteriol* **173**: 1617-1622.

Barlow, M. (2009) What antimicrobial resistance has taught us about horizontal gene transfer. In *Horizontal gene transfer: genomes in flux*. Gogarten, M.B., Gogarten, J.P. and Olendzenski, L. (eds). New York: Humana Press, pp. 397-411.

Beaber, J.W., Hochhut, B., and Waldor, M.K. (2002) Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from Vibrio cholerae. *J Bacteriol* **184**: 4259-4269.

Beaber, J.W., Hochhut, B., and Waldor, M.K. (2004) SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature* **427**: 72-74.

Beaber, J.W., and Waldor, M.K. (2004) Identification of operators and promoters that control SXT conjugative transfer. *J Bacteriol* **186**: 5945-5949.

Bellanger, X., Morel, C., Decaris, B., and Guedon, G. (2007) Derepression of excision of integrative and potentially conjugative elements from Streptococcus thermophilus by DNA damage response: implication of a cI-related repressor. *J Bacteriol* **189**: 1478-1481.

Berkmen, M.B., Lee, C.A., Loveday, E.K., and Grossman, A.D. (2009) Polar positioning of a conjugation protein from the integrative and conjugative element ICEBs1 of Bacillus subtilis. *J Bacteriol.*

Boltner, D., MacMahon, C., Pembroke, J.T., Strike, P., and Osborn, A.M. (2002) R391: a conjugative integrating mosaic comprised of phage, plasmid, and transposon elements. *J Bacteriol* **184**: 5158-5169.

Bongiorni, C., Ishikawa, S., Stephenson, S., Ogasawara, N., and Perego, M. (2005) Synergistic regulation of competence development in Bacillus subtilis by two Rap-Phr systems. *J Bacteriol* **187**: 4353-4361.

Bose*, B., Auchtung*, J.M., Lee, C.A., and Grossman, A.D. (2008) A conserved anti-repressor controls horizontal gene transfer by proteolysis. *Mol Microbiol* **70**: 570-582. * co-first authors.

Brassinga, A.K., Hiltz, M.F., Sisson, G.R., Morash, M.G., Hill, N., Garduno, E., Edelstein, P.H., Garduno, R.A., and Hoffman, P.S. (2003) A 65-kilobase pathogenicity island is unique to Philadelphia-1 strains of Legionella pneumophila. *J Bacteriol* **185**: 4630-4637.

Bringel, F., Van Alstine, G.L., and Scott, J.R. (1992) Conjugative transposition of Tn916: the transposon int gene is required only in the donor. *J Bacteriol* **174**: 4036-4041.

Brown, D.P., Chiang, S.J., Tuan, J.S., and Katz, L. (1988) Site-specific integration in Saccharopolyspora erythraea and multisite integration in Streptomyces lividans of actinomycete plasmid pSE101. *J Bacteriol* **170**: 2287-2295.

Brown, D.P., Idler, K.B., and Katz, L. (1990) Characterization of the genetic elements required for site-specific integration of plasmid pSE211 in Saccharopolyspora erythraea. *J Bacteriol* **172**: 1877-1888.

Brown, D.P., Idler, K.B., Backer, D.M., Donadio, S., and Katz, L. (1994) Characterization of the genes and attachment sites for site-specific integration of plasmid pSE101 in Saccharopolyspora erythraea and Streptomyces lividans. *Mol Gen Genet* **242**: 185-193.

Brumby, A.M., Lamont, I., Dodd, I.B., and Egan, J.B. (1996) Defining the SOS operon of coliphage 186. *Virology* **219**: 105-114.

Bunny, K., Liu, J., and Roth, J. (2002) Phenotypes of lexA mutations in Salmonella enterica: evidence for a lethal lexA null phenotype due to the Fels-2 prophage. *J Bacteriol* **184**: 6235-6249.

Burrus, V., Pavlovic, G., Decaris, B., and Guedon, G. (2002) Conjugative transposons: the tip of the iceberg. *Mol Microbiol* **46**: 601-610.

Burrus, V., Pavlovic, G., Decaris, B., and Guedon, G. (2002) The ICESt1 element of Streptococcus thermophilus belongs to a large family of integrative and conjugative elements that exchange modules and change their specificity of integration. *Plasmid* **48**: 77-97.

Burrus, V., and Waldor, M.K. (2004) Shaping bacterial genomes with integrative and conjugative elements. *Res Microbiol* **155**: 376-386.

Burrus, V., Marrero, J., and Waldor, M.K. (2006) The current ICE age: biology and evolution of SXT-related integrating conjugative elements. *Plasmid* **55**: 173-183.

Buu-Hoi, A., and Horodniceanu, T. (1980) Conjugative transfer of multiple antibiotic resistance markers in Streptococcus pneumoniae. *J Bacteriol* **143**: 313-320.

Byrd, D.R., and Matson, S.W. (1997) Nicking by transesterification: the reaction catalysed by a relaxase. *Mol Microbiol* **25**: 1011-1022.

Calcutt, M.J., Lewis, M.S., and Wise, K.S. (2002) Molecular genetic analysis of ICEF, an integrative conjugal element that is present as a repetitive sequence in the chromosome of Mycoplasma fermentans PG18. *J Bacteriol* **184**: 6929-6941.

Campbell, A. (1994) Comparative molecular biology of lambdoid phages. *Annu Rev Microbiol* **48**: 193-222.

Canchaya, C., Fournous, G., Chibani-Chennoufi, S., Dillmann, M.L., and Brussow, H. (2003) Phage as agents of lateral gene transfer. *Curr Opin Microbiol* **6**: 417-424.

Celli, J., Poyart, C., and Trieu-Cuot, P. (1997) Use of an excision reporter plasmid to study the intracellular mobility of the conjugative transposon Tn916 in gram-positive bacteria. *Microbiology* **143 (Pt 4)**: 1253-1261.

Celli, J., and Trieu-Cuot, P. (1998) Circularization of Tn916 is required for expression of the transposon-encoded transfer functions: characterization of long tetracycline-inducible transcripts reading through the attachment site. *Mol Microbiol* **28**: 103-117.

Chandler, J.R., and Dunny, G.M. (2004) Enterococcal peptide sex pheromones: synthesis and control of biological activity. *Peptides* **25**: 1377-1388.

Cheng, Q., Paszkiet, B.J., Shoemaker, N.B., Gardner, J.F., and Salyers, A.A. (2000) Integration and excision of a Bacteroides conjugative transposon, CTnDOT. *J Bacteriol* **182**: 4035-4043.

Cheng, Q., Sutanto, Y., Shoemaker, N.B., Gardner, J.F., and Salyers, A.A. (2001) Identification of genes required for excision of CTnDOT, a Bacteroides conjugative transposon. *Mol Microbiol* **41**: 625-632.

Churchward, G. (2002) Conjugative transposons and related mobile elements. In *Mobile DNA II.* Craig, N.L. (ed). Washington, D.C.: ASM Press, pp. 177-191.

Claverys, J.P., Prudhomme, M., and Martin, B. (2006) Induction of competence regulons as a general response to stress in gram-positive bacteria. *Annu Rev Microbiol* **60**: 451-475.

Clewell, D.B. (1993) Bacterial sex pheromone-induced plasmid transfer. *Cell* **73**: 9-12.

Clewell, D.B., Flannagan, S.E., and Jaworski, D.D. (1995) Unconstrained bacterial promiscuity: the Tn916-Tn1545 family of conjugative transposons. *Trends Microbiol* **3**: 229-236.

Clewell, D.B. (1999) Sex pheromone systems in *Enterococci.* In *Cell-Cell Signaling in Bacteria.* Dunny, G.M. and Winans, S.C. (eds). Washington, D. C.: ASM Press, pp. 47-65.

Coetzee, J.N., Datta, N., and Hedges, R.W. (1972) R factors from Proteus rettgeri. *J Gen Microbiol* **72**: 543-552.

Cohen, A., Bar-Nir, D., Goedeke, M.E., and Parag, Y. (1985) The integrated and free states of Streptomyces griseus plasmid pSG1. *Plasmid* **13**: 41-50.

Courvalin, P., and Carlier, C. (1986) Transposable multiple antibiotic resistance in Streptococcus pneumoniae. *Mol Gen Genet* **205**: 291-297.

Davis, B.M., and Waldor, M.K. (2002) Mobile genetic elements and bacterial pathogenesis. In *Mobile DNA II.* Craig, N.L. (ed). Washington, D.C.: ASM Press, pp. 1040-1059.

Dillard, J.P., and Seifert, H.S. (2001) A variable genetic island specific for Neisseria gonorrhoeae is involved in providing DNA for natural transformation and is found more often in disseminated infection isolates. *Mol Microbiol* **41**: 263-277.

Dimopoulou, I.D., Russell, J.E., Mohd-Zain, Z., Herbert, R., and Crook, D.W. (2002) Site-specific recombination with the chromosomal tRNA(Leu) gene by the large conjugative Haemophilus resistance plasmid. *Antimicrob Agents Chemother* **46**: 1602-1603.

Dodd, I.B., Shearwin, K.E., and Egan, J.B. (2005) Revisited gene regulation in bacteriophage lambda. *Curr Opin Genet Dev* **15**: 145-152.

Doucet-Populaire, F., Trieu-Cuot, P., Dosbaa, I., Andremont, A., and Courvalin, P. (1991) Inducible transfer of conjugative transposon Tn1545 from Enterococcus faecalis to Listeria monocytogenes in the digestive tracts of gnotobiotic mice. *Antimicrob Agents Chemother* **35**: 185-187.

Dunny, G.M. (2007) 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.

Flannagan, S.E., Zitzow, L.A., Su, Y.A., and Clewell, D.B. (1994) Nucleotide sequence of the 18-kb conjugative transposon Tn916 from Enterococcus faecalis. *Plasmid* **32**: 350-354.

Franke, A.E., and Clewell, D.B. (1981) 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.

Freeman, V.J. (1951) Studies on the virulence of bacteriophage-infected strains of Corynebacterium diphtheriae. *J Bacteriol* **61**: 675-688.

Fuqua, W.C., and Winans, S.C. (1994) A LuxR-LuxI type regulatory system activates Agrobacterium Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. *J Bacteriol* **176**: 2796-2806.

Garnier, F., Taourit, S., Glaser, P., Courvalin, P., and Galimand, M. (2000) Characterization of transposon Tn1549, conferring VanB-type resistance in Enterococcus spp. *Microbiology* **146 (Pt 6)**: 1481-1489.

Gasc, A.M., Sicard, N., Claverys, J.P., and Sicard, A.M. (1980) Lack of SOS repair in Streptococcus pneumoniae. *Mutat Res* **70**: 157-165.

Gasson, M.J., Godon, J.J., Pillidge, C.J., Eaton, T.J., Jury, K., and Shearman, C.A. (1995) Characterization and exploitation of conjugation in *Lactococcus lactis*. *Int. Dairy J.* **5**: 757-762.

Gawron-Burke, C., and Clewell, D.B. (1982) A transposon in Streptococcus faecalis with fertility properties. *Nature* **300**: 281-284.

Geisselsoder, J., Youdarian, P., Deho, G., Chidambaram, M., Goldstein, R., and Ljungquist, E. (1981) Mutants of satellite virus P4 that cannot derepress their bacteriophage P2 helper. *J Mol Biol* **148**: 1-19.

Grohmann, E., Muth, G., and Espinosa, M. (2003) Conjugative plasmid transfer in gram-positive bacteria. *Microbiol Mol Biol Rev* **67**: 277-301, table of contents.

Guenzi, E., Gasc, A.M., Sicard, M.A., and Hakenbeck, R. (1994) A two-component signal-transducing system is involved in competence and penicillin susceptibility in laboratory mutants of Streptococcus pneumoniae. *Mol Microbiol* **12**: 505-515.

Hamilton, H.L., Schwartz, K.J., and Dillard, J.P. (2001) Insertion-duplication mutagenesis of neisseria: use in characterization of DNA transfer genes in the gonococcal genetic island. *J Bacteriol* **183**: 4718-4726.

Hamilton, H.L., Dominguez, N.M., Schwartz, K.J., Hackett, K.T., and Dillard, J.P. (2005) Neisseria gonorrhoeae secretes chromosomal DNA via a novel type IV secretion system. *Mol Microbiol* **55**: 1704-1721.

Hamoen, L.W., Venema, G., and Kuipers, O.P. (2003) Controlling competence in Bacillus subtilis: shared use of regulators. *Microbiology* **149**: 9-17.

Hochhut, B., Jahreis, K., Lengeler, J.W., and Schmid, K. (1997) CTnscr94, a conjugative transposon found in enterobacteria. *J Bacteriol* **179**: 2097-2102.

Hochhut, B., Marrero, J., and Waldor, M.K. (2000) Mobilization of plasmids and chromosomal DNA mediated by the SXT element, a constin found in Vibrio cholerae O139. *J Bacteriol* **182**: 2043-2047.

Holcik, M., and Iyer, V.N. (1997) Conditionally lethal genes associated with bacterial plasmids. *Microbiology* **143 (Pt 11)**: 3403-3416.

Hopwood, D.A., Hintermann, G., Kieser, T., and Wright, H.M. (1984) Integrated DNA sequences in three streptomycetes form related autonomous plasmids after transfer to Streptomyces lividans. *Plasmid* **11**: 1-16.

Hosted, T.J., Jr., Wang, T., and Horan, A.C. (2005) Characterization of the Micromonospora rosaria pMR2 plasmid and development of a high G+C codon optimized integrase for site-specific integration. *Plasmid* **54**: 249-258.

Jiang, S.C., and Paul, J.H. (1998) Significance of Lysogeny in the Marine Environment: Studies with Isolates and a Model of Lysogenic Phage Production. *Microb Ecol* **35**: 235-243.

Johnsborg, O., Eldholm, V., and Havarstein, L.S. (2007) Natural genetic transformation: prevalence, mechanisms and function. *Res Microbiol* **158**: 767-778.

Kim, B., and Little, J.W. (1993) LexA and lambda Cl repressors as enzymes: specific cleavage in an intermolecular reaction. *Cell* **73**: 1165-1173.

Kimsey, H.H., and Waldor, M.K. (1998) CTXphi immunity: application in the development of cholera vaccines. *Proc Natl Acad Sci U S A* **95**: 7035-7039.

Koetje, E.J., Hajdo-Milasinovic, A., Kiewiet, R., Bron, S., and Tjalsma, H. (2003) A plasmid-borne Rap-Phr system of Bacillus subtilis can mediate cell-density controlled production of extracellular proteases. *Microbiology* **149**: 19-28.

Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A., Mizutani-Ui, Y., Takahashi, N.K., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H., and Hiramatsu, K. (2001) Whole genome sequencing of meticillin-resistant Staphylococcus aureus. *Lancet* **357**: 1225-1240.

Lamont, I., Brumby, A.M., and Egan, J.B. (1989) UV induction of coliphage 186: prophage induction as an SOS function. *Proc Natl Acad Sci U S A* **86**: 5492-5496.

Lancaster, H., Roberts, A.P., Bedi, R., Wilson, M., and Mullany, P. (2004) Characterization of Tn916S, a Tn916-like element containing the tetracycline resistance determinant tet(S). *J Bacteriol* **186**: 4395-4398.

Lanka, E., and Wilkins, B.M. (1995) DNA processing reactions in bacterial conjugation. *Annu Rev Biochem* **64**: 141-169.

Lapidus, A., Galleron, N., Andersen, J.T., Jorgensen, P.L., Ehrlich, S.D., and Sorokin, A. (2002) Co-linear scaffold of the Bacillus licheniformis and Bacillus subtilis genomes and its use to compare their competence genes. *FEMS Microbiol Lett* **209**: 23-30.

Lawrence, J.G., and Ochman, H. (1998) Molecular archaeology of the Escherichia coli genome. *Proc Natl Acad Sci U S A* **95**: 9413-9417.

Lazazzera, B.A. (2001) The intracellular function of extracellular signaling peptides. *Peptides* **22**: 1519-1527.

Lee, C.A., Auchtung, J.M., Monson, R.E., and Grossman, A.D. (2007) Identification and characterization of int (integrase), xis (excisionase) and chromosomal attachment sites of the integrative and conjugative element ICEBs1 of Bacillus subtilis. *Mol Microbiol* **66**: 1356-1369.

Lee, C.A., and Grossman, A.D. (2007) Identification of the origin of transfer (oriT) and DNA relaxase required for conjugation of the integrative and conjugative element ICEBs1 of Bacillus subtilis. *J Bacteriol* **189**: 7254-7261.

Lee, C.A., Babic, A., and Grossman, A.D. (2009) Autonomous plasmid-like replication of a conjugative transposon. *Mol Microbiol.*

Lee, L.H., Lui, D., Platner, P.J., Hsu, S.F., Chu, T.C., Gaynor, J.J., Vega, Q.C., and Lustigman, B.K. (2006) Induction of temperate cyanophage AS-1 by heavy metal--copper. *BMC Microbiol* **6**: 17.

Lewis, J.A., and Hatfull, G.F. (2001) Control of directionality in integrase-mediated recombination: examination of recombination directionality factors (RDFs) including Xis and Cox proteins. *Nucleic Acids Res* **29**: 2205-2216.

Little, J.W., and Mount, D.W. (1982) The SOS regulatory system of Escherichia coli. *Cell* **29**: 11-22.

Little, J.W. (1984) Autodigestion of lexA and phage lambda repressors. *Proc Natl Acad Sci U S A* **81**: 1375-1379.

Little, J.W. (1993) LexA cleavage and other self-processing reactions. *J Bacteriol* **175**: 4943-4950.

Liu, T., Renberg, S.K., and Haggard-Ljungquist, E. (1997) Derepression of prophage P2 by satellite phage P4: cloning of the P4 epsilon gene and identification of its product. *J Virol* **71**: 4502-4508.

Liu, T., Renberg, S.K., and Haggard-Ljungquist, E. (1998) The E protein of satellite phage P4 acts as an anti-repressor by binding to the C protein of helper phage P2. *Mol Microbiol* **30**: 1041-1050.

Llosa, M., Gomis-Ruth, F.X., Coll, M., and de la Cruz Fd, F. (2002) Bacterial conjugation: a two-step mechanism for DNA transport. *Mol Microbiol* **45**: 1-8.

Lobocka, M.B., Svarchevsky, A.N., Rybchin, V.N., and Yarmolinsky, M.B. (1996) Characterization of the primary immunity region of the Escherichia coli linear plasmid prophage N15. *J Bacteriol* **178**: 2902-2910.

Madon, J., Moretti, P., and Hutter, R. (1987) Site-specific integration and excision of pMEA100 in Nocardia mediterranei. *Mol Gen Genet* **209**: 257-264.

Magnuson, R., Solomon, J., and Grossman, A.D. (1994) Biochemical and genetic characterization of a competence pheromone from B. subtilis. *Cell* **77**: 207-216.

Mardanov, A.V., and Ravin, N.V. (2007) The antirepressor needed for induction of linear plasmid-prophage N15 belongs to the SOS regulon. *J Bacteriol* **189**: 6333-6338.

Marrero, R., and Yasbin, R.E. (1988) Cloning of the Bacillus subtilis recE+ gene and functional expression of recE+ in B. subtilis. *J Bacteriol* **170**: 335-344.

Martin, B., Quentin, Y., Fichant, G., and Claverys, J.P. (2006) Independent evolution of competence regulatory cascades in streptococci? *Trends Microbiol* **14**: 339-345.

Masai, H., and Arai, K. (1997) Frpo: a novel single-stranded DNA promoter for transcription and for primer RNA synthesis of DNA replication. *Cell* **89**: 897-907.

Miyoshi, Y.K., Ogata, S., and Hayashida, S. (1986) Multicopy derivative of pock-forming plasmid pSA1 in Streptomyces azureus. *J Bacteriol* **168**: 452-454.

Moon, K., Shoemaker, N.B., Gardner, J.F., and Salyers, A.A. (2005) Regulation of excision genes of the Bacteroides conjugative transposon CTnDOT. *J Bacteriol* **187**: 5732-5741.

Moretti, P., Hintermann, G., and Hutter, R. (1985) Isolation and characterization of an extrachromosomal element from Nocardia mediterranei. *Plasmid* **14**: 126-133.

Muhldorfer, I., Hacker, J., Keusch, G.T., Acheson, D.W., Tschape, H., Kane, A.V., Ritter, A., Olschlager, T., and Donohue-Rolfe, A. (1996) Regulation of the Shiga-like toxin II operon in Escherichia coli. *Infect Immun* **64**: 495-502.

Munoz-Najar, U., and Vijayakumar, M.N. (1999) An operon that confers UV resistance by evoking the SOS mutagenic response in streptococcal conjugative transposon Tn5252. *J Bacteriol* **181**: 2782-2788.

Murphy, D.B., and Pembroke, J.T. (1999) Monitoring of chromosomal insertions of the IncJ elements R391 and R997 in Escherichia coli K-12. *FEMS Microbiol Lett* **174**: 355-361.

Neely, M.N., and Friedman, D.I. (1998) Arrangement and functional identification of genes in the regulatory region of lambdoid phage H-19B, a carrier of a Shiga-like toxin. *Gene* **223**: 105-113.

Nishi, A., Tominaga, K., and Furukawa, K. (2000) A 90-kilobase conjugative chromosomal element coding for biphenyl and salicylate catabolism in Pseudomonas putida KF715. *J Bacteriol* **182**: 1949-1955.

O'Driscoll, J., Glynn, F., Fitzgerald, G.F., and van Sinderen, D. (2006) Sequence analysis of the lactococcal plasmid pNP40: a mobile replicon for coping with environmental hazards. *J Bacteriol* **188**: 6629-6639.

Okada, M., Sato, I., Cho, S.J., Iwata, H., Nishio, T., Dubnau, D., and Sakagami, Y. (2005) Structure of the Bacillus subtilis quorum-sensing peptide pheromone ComX. *Nat Chem Biol* **1**: 23-24.

Oppenheim, A.B., Kobiler, O., Stavans, J., Court, D.L., and Adhya, S. (2005) Switches in bacteriophage lambda development. *Annu Rev Genet* **39**: 409-429.

Parker, C., Zhang, X.L., Henderson, D., Becker, E., and Meyer, R. (2002) Conjugative DNA synthesis: R1162 and the question of rolling-circle replication. *Plasmid* **48**: 186-192.

Pavlovic, G., Burrus, V., Gintz, B., Decaris, B., and Guedon, G. (2004) Evolution of genomic islands by deletion and tandem accretion by site-specific recombination: ICESt1-related elements from Streptococcus thermophilus. *Microbiology* **150**: 759-774.

Pembroke, J.T., and Murphy, D.B. (2000) Isolation and analysis of a circular form of the IncJ conjugative transposon-like elements, R391 and R997: implications for IncJ incompatibility. *FEMS Microbiol Lett* **187**: 133-138.

Pembroke, J.T., MacMahon, C., and McGrath, B. (2002) The role of conjugative transposons in the Enterobacteriaceae. *Cell Mol Life Sci* **59**: 2055-2064.

Pernodet, J.L., Simonet, J.M., and Guerineau, M. (1984) Plasmids in different strains of Streptomyces ambofaciens: free and integrated form of plasmid pSAM2. *Mol Gen Genet* **198**: 35-41.

Peters, S.E., Hobman, J.L., Strike, P., and Ritchie, D.A. (1991) Novel mercury resistance determinants carried by IncJ plasmids pMERPH and R391. *Mol Gen Genet* **228**: 294-299.

Phillips, Z.E., and Strauch, M.A. (2002) Bacillus subtilis sporulation and stationary phase gene expression. *Cell Mol Life Sci* **59**: 392-402.

Pickard, D., Wain, J., Baker, S., Line, A., Chohan, S., Fookes, M., Barron, A., Gaora, P.O., Chabalgoity, J.A., Thanky, N., Scholes, C., Thomson, N., Quail, M., Parkhill, J., and Dougan, G. (2003) Composition, acquisition, and distribution of the Vi exopolysaccharide-encoding Salmonella enterica pathogenicity island SPI-7. *J Bacteriol* **185**: 5055-5065.

Pletz, M.W., McGee, L., Van Beneden, C.A., Petit, S., Bardsley, M., Barlow, M., and Klugman, K.P. (2006) Fluoroquinolone resistance in invasive Streptococcus pyogenes isolates due to spontaneous mutation and horizontal gene transfer. *Antimicrob Agents Chemother* **50**: 943-948.

Plunkett, G., 3rd, Rose, D.J., Durfee, T.J., and Blattner, F.R. (1999) Sequence of Shiga toxin 2 phage 933W from Escherichia coli O157:H7: Shiga toxin as a phage late-gene product. *J Bacteriol* **181**: 1767-1778.

Possoz, C., Ribard, C., Gagnat, J., Pernodet, J.L., and Guerineau, M. (2001) The integrative element pSAM2 from Streptomyces: kinetics and mode of conjugal transfer. *Mol Microbiol* **42**: 159-166.

Pottathil, M., and Lazazzera, B.A. (2003) The extracellular Phr peptide-Rap phosphatase signaling circuit of Bacillus subtilis. *Front Biosci* **8**: d32-45.

Poyart-Salmeron, C., Trieu-Cuot, P., Carlier, C., and Courvalin, P. (1989) Molecular characterization of two proteins involved in the excision of the conjugative transposon Tn1545: homologies with other site-specific recombinases. *Embo J* **8**: 2425-2433.

Poyart-Salmeron, C., Trieu-Cuot, P., Carlier, C., and Courvalin, P. (1990) The integration-excision system of the conjugative transposon Tn 1545 is structurally and functionally related to those of lambdoid phages. *Mol Microbiol* **4**: 1513-1521.

Provvedi, R., Manganelli, R., and Pozzi, G. (1996) Characterization of conjugative transposon Tn5251 of Streptococcus pneumoniae. *FEMS Microbiol Lett* **135**: 231-236.

Prudhomme, M., Attaiech, L., Sanchez, G., Martin, B., and Claverys, J.P. (2006) Antibiotic stress induces genetic transformability in the human pathogen Streptococcus pneumoniae. *Science* **313**: 89-92.

Rauch, P.J., and De Vos, W.M. (1992) Characterization of the novel nisin-sucrose conjugative transposon Tn5276 and its insertion in Lactococcus lactis. *J Bacteriol* **174**: 1280-1287.

Rauch, P.J., Beerthuyzen, M.M., and de Vos, W.M. (1994) Distribution and evolution of nisin-sucrose elements in Lactococcus lactis. *Appl Environ Microbiol* **60**: 1798-1804.

Rauch, P.J., and de Vos, W.M. (1994) Identification and characterization of genes involved in excision of the Lactococcus lactis conjugative transposon Tn5276. *J Bacteriol* **176**: 2165-2171.

Ravatn, R., Studer, S., Springael, D., Zehnder, A.J., and van der Meer, J.R. (1998) Chromosomal integration, tandem amplification, and deamplification in Pseudomonas putida F1 of a 105-kilobase genetic element containing the chlorocatechol degradative genes from Pseudomonas sp. Strain B13. *J Bacteriol* **180**: 4360-4369.

Ravatn, R., Studer, S., Zehnder, A.J., and van der Meer, J.R. (1998) Int-B13, an unusual site-specific recombinase of the bacteriophage P4 integrase family, is responsible for chromosomal insertion of the 105-kilobase clc element of Pseudomonas sp. Strain B13. *J Bacteriol* **180**: 5505-5514.

Redfield, R.J. (1988) Evolution of bacterial transformation: is sex with dead cells ever better than no sex at all? *Genetics* **119**: 213-221.

Rice, L.B., Marshall, S.H., and Carias, L.L. (1992) Tn5381, a conjugative transposon identifiable as a circular form in Enterococcus faecalis. *J Bacteriol* **174**: 7308-7315.

Rice, L.B., Carias, L.L., Marshall, S.H., Hutton-Thomas, R., and Rudin, S. (2007) Characterization of Tn5386, a Tn916-related mobile element. *Plasmid* **58**: 61-67.

Roberts, A.P., Johanesen, P.A., Lyras, D., Mullany, P., and Rood, J.I. (2001) Comparison of Tn5397 from Clostridium difficile, Tn916 from Enterococcus faecalis and the CW459tet(M) element from Clostridium perfringens shows that they have similar conjugation regions but different insertion and excision modules. *Microbiology* **147**: 1243-1251.

Roberts, A.P., Davis, I.J., Seville, L., Villedieu, A., and Mullany, P. (2006) Characterization of the ends and target site of a novel tetracycline resistance-encoding conjugative transposon from Enterococcus faecium 664.1H1. *J Bacteriol* **188**: 4356-4361.

Roberts, A.P., and Mullany, P. (2009) A modular master on the move: the Tn916 family of mobile genetic elements. *Trends Microbiol* **17**: 251-258.

Roberts, J.W., and Devoret, R. (1983) Lysogenic induction. In *Lambda II.* Hendrix, R.W. (ed). Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory, pp. 123-144.

Salyers, A.A., Shoemaker, N.B., Stevens, A.M., and Li, L.Y. (1995) Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. *Microbiol Rev* **59**: 579-590.

Salyers, A.A., and Shoemaker, N.B. (1996) Resistance gene transfer in anaerobes: new insights, new problems. *Clin Infect Dis* **23 Suppl 1**: S36-43.

Saskova, L., Novakova, L., Basler, M., and Branny, P. (2007) Eukaryotic-type serine/threonine protein kinase StkP is a global regulator of gene expression in Streptococcus pneumoniae. *J Bacteriol* **189**: 4168-4179.

Schmid, K., Ebner, R., Altenbuchner, J., Schmitt, R., and Lengeler, J.W. (1988) Plasmid-mediated sucrose metabolism in Escherichia coli K12: mapping of the scr genes of pUR400. *Mol Microbiol* **2**: 1-8.

Schubert, S., Dufke, S., Sorsa, J., and Heesemann, J. (2004) A novel integrative and conjugative element (ICE) of Escherichia coli: the putative progenitor of the Yersinia high-pathogenicity island. *Mol Microbiol* **51**: 837-848.

Scott, J.R., Kirchman, P.A., and Caparon, M.G. (1988) An intermediate in transposition of the conjugative transposon Tn916. *Proc Natl Acad Sci U S A* **85**: 4809-4813.

Scott, J.R., Bringel, F., Marra, D., Van Alstine, G., and Rudy, C.K. (1994) Conjugative transposition of Tn916: preferred targets and evidence for conjugative transfer of a single strand and for a double-stranded circular intermediate. *Mol Microbiol* **11**: 1099-1108.

Scott, J.R., and Churchward, G.G. (1995) Conjugative transposition. *Annu Rev Microbiol* **49**: 367-397.

Scott, K.P. (2002) The role of conjugative transposons in spreading antibiotic resistance between bacteria that inhabit the gastrointestinal tract. *Cell Mol Life Sci* **59**: 2071-2082.

Sentchilo, V., Ravatn, R., Werlen, C., Zehnder, A.J., and van der Meer, J.R. (2003) Unusual integrase gene expression on the clc genomic island in Pseudomonas sp. strain B13. *J Bacteriol* **185**: 4530-4538.

Sezonov, G., Hagege, J., Pernodet, J.L., Friedmann, A., and Guerineau, M. (1995) Characterization of pra, a gene for replication control in pSAM2, the integrating element of Streptomyces ambofaciens. *Mol Microbiol* **17**: 533-544.

Sezonov, G., Duchene, A.M., Friedmann, A., Guerineau, M., and Pernodet, J.L. (1998) Replicase, excisionase, and integrase genes of the Streptomyces element pSAM2 constitute an operon positively regulated by the pra gene. *J Bacteriol* **180**: 3056-3061.

Sezonov, G., Possoz, C., Friedmann, A., Pernodet, J.L., and Guerineau, M. (2000) KorSA from the Streptomyces integrative element pSAM2 is a central transcriptional repressor: target genes and binding sites. *J Bacteriol* **182**: 1243-1250.

Shearwin, K.E., Brumby, A.M., and Egan, J.B. (1998) The Tum protein of coliphage 186 is an antirepressor. *J Biol Chem* **273**: 5708-5715.

Shoemaker, N.B., and Salyers, A.A. (1988) Tetracycline-dependent appearance of plasmidlike forms in Bacteroides uniformis 0061 mediated by conjugal Bacteroides tetracycline resistance elements. *J Bacteriol* **170**: 1651-1657.

Shoemaker, N.B., Barber, R.D., and Salyers, A.A. (1989) Cloning and characterization of a Bacteroides conjugal tetracycline-erythromycin resistance element by using a shuttle cosmid vector. *J Bacteriol* **171**: 1294-1302.

Shoemaker, N.B., Vlamakis, H., Hayes, K., and Salyers, A.A. (2001) Evidence for extensive resistance gene transfer among Bacteroides spp. and among Bacteroides and other genera in the human colon. *Appl Environ Microbiol* **67**: 561-568.

Showsh, S.A., and Andrews, R.E., Jr. (1992) Tetracycline enhances Tn916-mediated conjugal transfer. *Plasmid* **28**: 213-224.

Smith, H.W., and Parsell, Z. (1975) Transmissible substrate-utilizing ability in enterobacteria. *J Gen Microbiol* **87**: 129-140.

Snyder, L., and Champness, W. (1997) Bacteriophages. In *Molecular genetics of bacteria.* Washington, D.C.: ASM Press, pp. p. 161-193.

Solomon, J.M., Lazazzera, B.A., and Grossman, A.D. (1996) Purification and characterization of an extracellular peptide factor that affects two different developmental pathways in Bacillus subtilis. *Genes Dev* **10**: 2014-2024.

Sosio, M., Madon, J., and Hutter, R. (1989) Excision of pIJ408 from the chromosome of Streptomyces glaucescens and its transfer into Streptomyces lividans. *Mol Gen Genet* **218**: 169-176.

Springael, D., Peys, K., Ryngaert, A., Van Roy, S., Hooyberghs, L., Ravatn, R., Heyndrickx, M., van der Meer, J.R., Vandecasteele, C., Mergeay, M., and Diels, L. (2002) Community shifts in a seeded 3-chlorobenzoate degrading membrane biofilm reactor: indications for involvement of in situ horizontal transfer of the clc-element from inoculum to contaminant bacteria. *Environ Microbiol* **4**: 70-80.

Stevens, A.M., Shoemaker, N.B., and Salyers, A.A. (1990) The region of a Bacteroides conjugal chromosomal tetracycline resistance element which is responsible for production of plasmidlike forms from unlinked chromosomal DNA might also be involved in transfer of the element. *J Bacteriol* **172**: 4271-4279.

Stevens, A.M., Sanders, J.M., Shoemaker, N.B., and Salyers, A.A. (1992) Genes involved in production of plasmidlike forms by a Bacteroides conjugal chromosomal element share amino acid homology with two-component regulatory systems. *J Bacteriol* **174**: 2935-2942.

Stevens, A.M., Shoemaker, N.B., Li, L.Y., and Salyers, A.A. (1993) Tetracycline regulation of genes on Bacteroides conjugative transposons. *J Bacteriol* **175**: 6134-6141.

Su, Y.A., He, P., and Clewell, D.B. (1992) Characterization of the tet(M) determinant of Tn916: evidence for regulation by transcription attenuation. *Antimicrob Agents Chemother* **36**: 769-778.

Sullivan, J.T., and Ronson, C.W. (1998) Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *Proc Natl Acad Sci U S A* **95**: 5145-5149.

Sullivan, J.T., Trzebiatowski, J.R., Cruickshank, R.W., Gouzy, J., Brown, S.D., Elliot, R.M., Fleetwood, D.J., McCallum, N.G., Rossbach, U., Stuart, G.S., Weaver, J.E., Webby, R.J., De Bruijn, F.J., and Ronson, C.W. (2002) Comparative sequence analysis of the symbiosis island of Mesorhizobium loti strain R7A. *J Bacteriol* **184**: 3086-3095.

Susskind, M.M., and Botstein, D. (1975) Mechanism of action of Salmonella phage P22 antirepressor. *J Mol Biol* **98**: 413-424.

Susskind, M.M., and Youderian, P. (1983) Bacteriophage P22 antirepressor and its control. In *Lambda II.* Hendrix, R.W. (ed). Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory, pp. 347-363.

te Poele, E.M., Bolhuis, H., and Dijkhuizen, L. (2008) Actinomycete integrative and conjugative elements. *Antonie Van Leeuwenhoek* **94**: 127-143.

te Poele, E.M., Samborskyy, M., Oliynyk, M., Leadlay, P.F., Bolhuis, H., and Dijkhuizen, L. (2008) Actinomycete integrative and conjugative pMEA-like elements of Amycolatopsis and Saccharopolyspora decoded. *Plasmid* **59**: 202-216.

Thompson, J., Nguyen, N.Y., Sackett, D.L., and Donkersloot, J.A. (1991) Transposon-encoded sucrose metabolism in Lactococcus lactis. Purification of sucrose-6-phosphate hydrolase

and genetic linkage to N5-(L-1-carboxyethyl)-L-ornithine synthase in strain K1. *J Biol Chem* **266**: 14573-14579.

Torres, O.R., Korman, R.Z., Zahler, S.A., and Dunny, G.M. (1991) The conjugative transposon Tn925: enhancement of conjugal transfer by tetracycline in Enterococcus faecalis and mobilization of chromosomal genes in Bacillus subtilis and E. faecalis. *Mol Gen Genet* **225**: 395-400.

Toussaint, A., Merlin, C., Monchy, S., Benotmane, M.A., Leplae, R., Mergeay, M., and Springael, D. (2003) The biphenyl- and 4-chlorobiphenyl-catabolic transposon Tn4371, a member of a new family of genomic islands related to IncP and Ti plasmids. *Appl Environ Microbiol* **69**: 4837-4845.

Vijayakumar, M.N., and Ayalew, S. (1993) Nucleotide sequence analysis of the termini and chromosomal locus involved in site-specific integration of the streptococcal conjugative transposon Tn5252. *J Bacteriol* **175**: 2713-2719.

Vogtli, M., and Cohen, S.N. (1992) The chromosomal integration site for the Streptomyces plasmid SLP1 is a functional tRNA(Tyr) gene essential for cell viability. *Mol Microbiol* **6**: 3041-3050.

Vrijbloed, J.W., Madon, J., and Dijkhuizen, L. (1994) A plasmid from the methylotrophic actinomycete Amycolatopsis methanolica capable of site-specific integration. *J Bacteriol* **176**: 7087-7090.

Vrijbloed, J.W., Jelinkova, M., Hessels, G.I., and Dijkhuizen, L. (1995) Identification of the minimal replicon of plasmid pMEA300 of the methylotrophic actinomycete Amycolatopsis methanolica. *Mol Microbiol* **18**: 21-31.

Wagner, P.L., Neely, M.N., Zhang, X., Acheson, D.W., Waldor, M.K., and Friedman, D.I. (2001) Role for a phage promoter in Shiga toxin 2 expression from a pathogenic Escherichia coli strain. *J Bacteriol* **183**: 2081-2085.

Waldor, M.K., Tschape, H., and Mekalanos, J.J. (1996) A new type of conjugative transposon encodes resistance to sulfamethoxazole, trimethoprim, and streptomycin in Vibrio cholerae O139. *J Bacteriol* **178**: 4157-4165.

Wang, H., Roberts, A.P., Lyras, D., Rood, J.I., Wilks, M., and Mullany, P. (2000) Characterization of the ends and target sites of the novel conjugative transposon Tn5397 from Clostridium difficile: excision and circularization is mediated by the large resolvase, TndX. *J Bacteriol* **182**: 3775-3783.

Wang, H., Smith, M.C., and Mullany, P. (2006) The conjugative transposon Tn5397 has a strong preference for integration into its Clostridium difficile target site. *J Bacteriol* **188**: 4871-4878.

Wang, Y., Wang, G.R., Shelby, A., Shoemaker, N.B., and Salyers, A.A. (2003) A newly discovered Bacteroides conjugative transposon, CTnGERM1, contains genes also found in gram-positive bacteria. *Appl Environ Microbiol* **69**: 4595-4603.

Wang, Y., Shoemaker, N.B., and Salyers, A.A. (2004) Regulation of a Bacteroides operon that controls excision and transfer of the conjugative transposon CTnDOT. *J Bacteriol* **186**: 2548-2557.

Wang, Y., Rotman, E.R., Shoemaker, N.B., and Salyers, A.A. (2005) Translational control of tetracycline resistance and conjugation in the Bacteroides conjugative transposon CTnDOT. *J Bacteriol* **187**: 2673-2680.

Waters, V.L., and Guiney, D.G. (1993) Processes at the nick region link conjugation, T-DNA transfer and rolling circle replication. *Mol Microbiol* **9**: 1123-1130.

Whittle, G., Shoemaker, N.B., and Salyers, A.A. (2002) The role of Bacteroides conjugative transposons in the dissemination of antibiotic resistance genes. *Cell Mol Life Sci* **59**: 2044-2054.

Whittle, G., Shoemaker, N.B., and Salyers, A.A. (2002) Characterization of genes involved in modulation of conjugal transfer of the Bacteroides conjugative transposon CTnDOT. *J Bacteriol* **184**: 3839-3847.

Whittle, G., Hamburger, N., Shoemaker, N.B., and Salyers, A.A. (2006) A bacteroides conjugative transposon, CTnERL, can transfer a portion of itself by conjugation without excising from the chromosome. *J Bacteriol* **188**: 1169-1174.

Williams, K.P. (2002) Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. *Nucleic Acids Res* **30**: 866-875.

Williamson, S.J., Houchin, L.A., McDaniel, L., and Paul, J.H. (2002) Seasonal variation in lysogeny as depicted by prophage induction in Tampa Bay, Florida. *Appl Environ Microbiol* **68**: 4307-4314.

Wilson, W.H., Joint, I.R., Carr, N.G., and Mann, N.H. (1993) Isolation and Molecular Characterization of Five Marine Cyanophages Propagated on Synechococcus sp. Strain WH7803. *Appl Environ Microbiol* **59**: 3736-3743.

Winans, S.C., Zhu, J., and More, M.I. (1999) Cell Density-Dependent Gene Expression by *Agrobacterium tumefaciens* during Colonization of crown Gall Tumors. In *Cell-Cell Signaling in Bacteria*. Dunny, G.M. and Winans, S.C. (eds). Washington, D.C.: ASM Press, pp. 117-128.

Wohlhieter, J.A., Lazere, J.R., Snellings, N.J., Johnson, E.M., Synenki, R.M., and Baron, L.S. (1975) Characterization of transmissible genetic elements from sucrose-fermenting Salmonella strains. *J Bacteriol* **122**: 401-406.

Yokota, T., and Kuwahara, S. (1977) Temperature-sensitive R plasmid obtained from naturally isolated drug-resistant Vibrio cholerae (biotype El Tor). *Antimicrob Agents Chemother* **11**: 13-20.

Zhang, X., McDaniel, A.D., Wolf, L.E., Keusch, G.T., Waldor, M.K., and Acheson, D.W. (2000) Quinolone antibiotics induce Shiga toxin-encoding bacteriophages, toxin production, and death in mice. *J Infect Dis* **181**: 664-670.

Zhou, J.Z., and Tiedje, J.M. (1995) Gene transfer from a bacterium injected into an aquifer to an indigenous bacterium. *Mol Ecol* **4**: 613-618.

Zhu, J., Oger, P.M., Schrammeijer, B., Hooykaas, P.J., Farrand, S.K., and Winans, S.C. (2000) The bases of crown gall tumorigenesis. *J Bacteriol* **182**: 3885-3895.

Chapter 2: A conserved anti-repressor controls horizontal gene transfer by proteolysis

This work was published in *Molecular Microbiology* (2008) 70:554-6

by Baundauna Bose*, Jennifer M. Auchtung*, Catherine A. Lee, and Alan D. Grossman

* co-first authors

Fig. 5 shows BB's results. Fig. 2, Fig. 4(B-E), and Table 1 show JMA's results. For Fig. 3, BB made the strains and made the initial observations, and JMA did the experiment shown in the paper. CAL did the experiments in Fig. 4A.

Abstract

The mobile genetic element ICE*Bs1* is an integrative and conjugative element (a conjugative transposon) found in the *Bacillus subtilis* chromosome. The SOS response and the RapI-PhrI sensory system activate ICE*Bs1* gene expression, excision, and transfer by inactivating the ICE*Bs1* repressor protein ImmR. Although ImmR is similar to many characterized phage repressors, we found that, unlike these repressors, inactivation of ImmR requires an ICE*Bs1*-encoded anti-repressor ImmA (YdcM). ImmA was needed for the degradation of ImmR in *B. subtilis*. Co-expression of ImmA and ImmR in *E. coli* or co-incubation of purified ImmA and ImmR resulted in site-specific cleavage of ImmR. Homologs of *immR* and *immA* are found in many mobile genetic elements. We found that the ImmA homolog encoded by *B. subtilis* phage ø105 is required for inactivation of the ø105 repressor (an ImmR homolog). ImmA-dependent proteolysis of ImmR repressors may be a conserved mechanism for regulating horizontal gene transfer.

Introduction

Mobile genetic elements play significant roles in genome plasticity, the spread of antibiotic resistance, and acquisition of new traits (Dobrindt *et al.*, 2004; Frost *et al.*, 2005)}. Mobile genetic elements include bacteriophages, conjugative plasmids, and conjugative transposons (Churchward, 2002; Whittle *et al.*, 2002), also known as integrative and conjugative elements (ICEs) (Burrus *et al.*, 2002; van der Meer and Sentchilo, 2003). Mobile elements typically have regulatory mechanisms to reduce the burden placed on the host cells while maintaining the potential for the element to spread under specific conditions. A tremendous amount is known about phages and plasmids and mechanisms controlling their transfer. In contrast, much less is known about regulation of ICEs and mechanisms controlling their stability and transfer.

ICEs and putative ICEs are present in many bacteria (Burrus and Waldor, 2004) and are important agents of horizontal gene transfer (Beaber *et al.*, 2002; Rice, 2002; Scott, 2002; Whittle *et al.*, 2002). They reside integrated in the chromosome of a host cell and can excise and transfer to recipients through conjugation (mating) where they then integrate into the chromosome of the recipient (Churchward, 2002; Whittle *et al.*, 2002). ICEs typically encode proteins needed for their regulation, integration, excision, and transfer.

ICE*Bs1* (Fig. 1) is an integrative and conjugative element found in the chromosome of *Bacillus subtilis* (Auchtung *et al.*, 2005; Burrus *et al.*, 2002). During exponential growth, the ICE*Bs1* gene product ImmR, the immunity repressor, represses transcription of genes needed for excision and transfer (Auchtung *et al.*, 2007). Inactivation of ImmR and derepression of ICE*Bs1* gene expression, excision, and mating occurs either when cells induce the RapI-dependent sensory response due to crowding by neighboring cells that lack a copy of ICE*Bs1*, or when cells induce the RecA-dependent SOS response to DNA damage (Auchtung *et al.*, 2005).

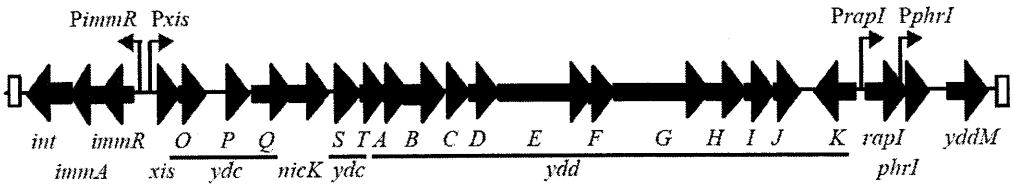


Figure 1. Map of ICE*Bs1*. Each of the 24 genes encoded by ICE*Bs1* is indicated by the thick black arrows oriented in the direction of transcription. Gene names are indicated under each arrow. Single letter designations are given for the genes of unknown function (*ydcO-ydcT*; *yddA-yddK; yddM*). Thin arrows indicate the positions of the characterized promoters. White boxes denote the ends of the element.

Many mobile genetic elements are induced by the highly conserved RecA-mediated response to DNA damage, the SOS response {reviewed in (Walker, 1996)}. The best characterized responses involve induction of lysogenic phages. After DNA damage, the host protein RecA is activated by binding to single-stranded DNA. In the paradigmatic example of lambda (and lambdoid phages), activated RecA directly facilitates the autocleavage of the phage repressor cI (Little, 1984; Roberts *et al.*, 1978). An analogous mechanism of autocleavage inactivates the cellular repressor LexA (Little, 1984; Miller *et al.*, 1996), thereby inducing expression of the many genes normally repressed by LexA (Au *et al.*, 2005; Courcelle *et al.*, 2001; Fernandez De Henestrosa *et al.*, 2000; Goranov *et al.*, 2006).

We investigated the mechanism by which ImmR is inactivated under conditions that promote ICE*Bs1* excision and transfer. Although inactivation of ImmR by DNA damage requires RecA, ImmR does not contain some of the key residues known to be required for auto-proteolysis and activated RecA does not directly facilitate the autocleavage of ImmR. Instead, we found that inactivation of ImmR requires the ICE*Bs1*-encoded protein, ImmA (ImmR Antagonist, formerly YdcM), an anti-repressor. ImmR was rapidly degraded in vivo when antagonized by ImmA. ImmA and ImmR interacted directly in a yeast two-hybrid assay. We purified ImmA and found that in vitro, it caused cleavage of purified ImmR. We conclude that ImmA is an anti-repressor that inactivates ImmR, likely through site-specific cleavage that results in ImmR degradation and derepression of ICE*Bs1* in vivo.

Homologs of *immA* and *immR* are found in many other mobile genetic elements. We found that the ImmA homolog of *B. subtilis* phage ø105 {(ø105)ImmA, formerly Orf2, a.k.a., ø105_33 in NCBI} was required for inactivation of the phage repressor cø105 (an ImmR homolog).

ImmA-dependent proteolysis of ImmR likely represents a conserved strategy for regulating transfer of many different mobile genetic elements.

Results

Identification of the anti-repressor gene *immA*

ImmR is the only ICE*Bs1* gene both necessary and sufficient to repress transcription from P*xis*, the promoter upstream of the gene encoding excisionase (Auchtung *et al.*, 2007; Lee *et al.*, 2007). We found that ImmR was not sufficient for induction of P*xis-lacZ* by RapI or the RecA-dependent DNA damage response. In cells cured of ICE*Bs1* (ICE*Bs1*^{0}) that express *immR* from its own promoter at an ectopic site, there was no detectable induction of P*xis-lacZ*, either after addition of mitomycin C (MMC) to induce the *recA*-dependent DNA damage response (Fig. 2A), or after overproduction of the cell-cell sensory regulator RapI (Fig. 2B). In contrast, in cells containing ICE*Bs1*, expression of P*xis-lacZ* was induced after addition of MMC (Fig. 2A) or overproduction of RapI (Fig. 2B), as observed previously (Auchtung *et al.*, 2007). These results indicate that at least one additional ICE*Bs1* gene is required to derepress expression from P*xis*.

Previously we found that a deletion-derivative of ICE*Bs1* containing only four genes, *int, immA* (*ydcM*), *immR,* and *xis*, was stably maintained in the chromosome and could be induced to excise by either the DNA damage response or by overexpression of *rapI* (Lee *et al.*, 2007). Since *int* and *xis* encode the integrase and excisionase respectively (Lee *et al.*, 2007), we suspected that *immA* might encode a regulatory protein required for derepression. Analysis of ImmA using the conserved domain architecture retrieval tool {C-DART (Geer *et al.*, 2002)} revealed that it has a conserved domain of unknown function (COG2856) that contains the characteristic HEXXH

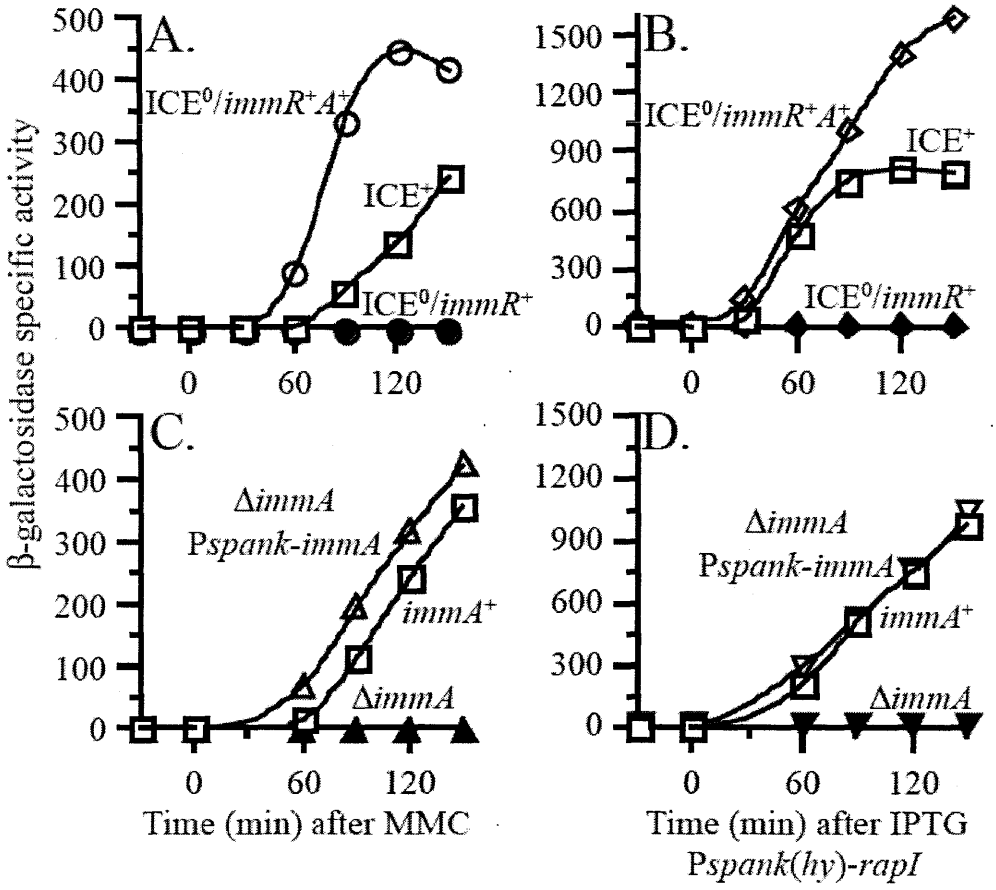

A.
B.
C.
D.
β-galactosidase specific activity
ICE⁰/immR⁺A⁺
ICE⁺
ICE⁰/immR⁺
ΔimmA
Pspank-immA
immA⁺
Time (min) after MMC
Time (min) after IPTG
Pspank(hy)-rapI

Figure 2. ImmA is required for de-repression of P*xis-lacZ*. Expression of P*xis-lacZ* was monitored in cells grown in defined minimal medium and treated at time 0 with either MMC to induce the SOS response (A & C), or with IPTG to induce expression of P*spank(hy)-rapI* (B & D). β-galactosidase specific activity is plotted as a function of time after the indicated treatment.

A. Effects of MMC on expression of P*xis-lacZ*. ICE*Bs1*$^+$ cells (JMA201, □, ICE$^+$); ICE*Bs1*^{0} cells expressing only *immR* from its own promoter (JMA421, ●, ICE^0/*immR*$^+$); ICE*Bs1*^{0} cells co-expressing *immR* and *immA* from the *immR* promoter (JMA436, ○, ICE^0/*immR*$^+$*A*$^+$). Data for JMA201 were published previously (Auchtung *et al.*, 2007) and are re-plotted for comparison.

B. Effects of overproducing RapI on expression of P*xis-lacZ*. ICE*Bs1*$^+$ (KLG126, □, ICE$^+$); ICE*Bs1*^{0} cells expressing *immR* (JMA444, ♦, ICE^0/*immR*$^+$); ICE*Bs1*^{0} cells expressing both *immR* and *immA* (JMA446, ◊, ICE^0/*immR*$^+$*A*$^+$).

C-D. Effects of ImmA on expression of P*xis-lacZ* induced with MMC (C) or RapI (D). All cells contained the Δ*int*::*cat* mutation and the indicated *immA* allele. **C.** *immA*$^+$ (CAL16, □); Δ*immA* (JMA726, σ); Δ*immA* P*spank-immA* (JMA840, △). IPTG was present throughout growth of JMA840. **D.** *immA*$^+$ (JMA836, □); Δ*immA* (JMA838, τ); Δ*immA* P*spank-immA* (JMA842, ▽).

motif found in many zinc-dependent metalloproteases (reviewed in Miyoshi and Shinoda, 2000; and Rawlings and Barrett, 1995).

We found that *immA*, the gene immediately downstream from *immR* (Fig. 1), was needed for derepression of P*xis-lacZ*. An *immA* null mutant was unable to derepress expression of P*xis-lacZ* in response to MMC (Fig. 2C) or overproduction of RapI (Fig. 2D). This defect in induction was rescued by providing wild type *immA* from a heterologous promoter (P*spank-immA*) at an ectopic site in the chromosome (Fig. 2C, D), indicating that the defect in induction was due to loss of *immA* and not to unintended secondary effects.

We also found that expression of *immA* and *immR* together from their normal promoter, in the absence of any other ICE*Bs1* genes, was sufficient for normal repression and derepression of P*xis-lacZ* (Fig. 2A, B). Transcription of P*xis-lacZ* was strongly derepressed following addition of MMC (Fig. 2A) or overproduction of RapI (Fig. 2B). This is in marked contrast to the lack of derepression in cells expressing only *immR* and not *immA* (Fig. 2A, B). These results indicate that of all the ICE*Bs1* genes, *immA* and *immR* together are sufficient for ICE*Bs1* induction in response to DNA damage or overproduction of RapI (also an ICE*Bs1* gene product).

ImmA contains an HEXXH motif that is required for proteolytic activity of many zinc-dependent metalloproteases. We constructed an *immAH75A* mutation (altering the first histidine in this motif) and found that it caused an *immA* null phenotype. That is, in cells containing wild type *immR* and the *immAH75A* allele, there was no derepression of P*xis-lacZ* by either MMC or overproduction of RapI (data not shown).

Taken together, all the genetic data indicate that ImmA is an anti-repressor required for derepression of ICE*Bs1* in response to DNA damage or RapI and the first histidine in the

HEXXH motif found in many metalloproteases is somehow important for ImmA function (see below).

ImmR and ImmA homologs encoded by other mobile genetic elements

immR and *immA* are encoded with the integrase gene in a so-called lysogeny module at the left end of ICE*Bs1* (Fig. 1). Lysogeny modules are found in many temperate (lysogenic) phage and encode genes that promote integration and suppress lytic phage growth. Several phages of low G+C-content gram- positive bacteria contain modules with homologs of *int*, *immR* and *immA* (Lucchini *et al.*, 1999).

Homologs of ImmA and ImmR are readily identifiable through comparative sequence analyses and over 80% of these homologs are encoded in known or putative mobile genetic elements (Table 1). The genes most similar to *immR* and *immA* from ICE*Bs1* are found in a putative mobile element in the genome of a vancomycin-resistant *Enterococcus faecalis* (Paulsen *et al.*, 2003). ImmA and ImmR homologs are also present in many phages that are induced by the global DNA damage response. It is generally assumed that induction is caused by RecA-stimulated auto-proteolysis of the phage repressor, although the mechanism has not been characterized for elements containing ImmA homologs. Some of these mobile genetic elements (ø105, *skin*, and λBa04) also contain homologs of RapI and PhrI (Auchtung *et al.*, 2005). Based on the conservation of ImmA and ImmR, the conservation of RecA, and the prevalence of lysogenic phages and other mobile elements that are induced by the RecA-dependent DNA damage response, we suspect that the function of the anti-repressor ImmA represents a conserved mechanism for inactivating repressor proteins and stimulating horizontal gene transfer.

Table 1. ImmR and ImmA homologs in mobile and putative mobile genetic element.

Organism	Element	ImmR[1] (%Id/%Sim)	ImmA[2] (%Id/%Sim)	Din[3]	Reference[4]
Known mobile elements					
Bacillus clarkii	BCJA1c phage	Gp5 (29/48)	Gp4 (39/59)	nd	(Kropinski *et al.*, 2005)
B. subtilis	ø105 phage	ø105 repressor (24/50)	ImmA(ϕ105) (32/53)	Yes	(McVeigh and Yasbin, 1996)
B. subtilis	PBSX phage	Xre (34/50)	XkdA (N.S.)	Yes	(McDonnell *et al.*, 1994)
B. subtilis	*skin* element[5]	YqaE (34/52)	YqaB (N.S.)	No	(Krogh *et al.*, 1996)
B. thuringiensis	MZTP02 phage	AAX62112.1 (N.S.)	AAX62113.1 (34/50)	Yes	(Liao *et al.*, 2007)
Listeria monocytogenes	A118 phage	Gp36 (33/62)	Gp35 (N.S.)	Yes	(Loessner *et al.*, 2000)
Geobacillus	GBSV1 phage	Gp48 (31/52)	Gp47 (36/51)	nd	--
Geobacillus thermodenitrificans	defective prophage	GTNG_2860 (23/44)	GTNG_2861 (38/53)	nd	(Feng *et al.*, 2007)
Streptococcus pneumoniae	MM1 phage	CI (25/52)	Orf2 (N.S.)	Yes	(Obregon *et al.*, 2003)
S. thermophilus	ϕO1205 phage	Orf4 (35/43)	Orf3 (N.S.)	Yes	(Stanley *et al.*, 1997)
Putative mobile elements[6]					
Actinobacillus pleuropneumonia serovar 1	--	Aple02002129 (35/54)	Aple02002131 (25/43)	nd	--
Bacillus anthracis	λBa04	BA3829 (30/52)	BA3830 (33/52)	nd	(Read *et al.*, 2003)
B. cereus subsp. cytotoxis NVH391-98	--	Bcer98DRAFT_3608 (33/60)	Bcer98DRAFT 3609 (28/53)	nd	--
B. halodurans C-125	--	BH3549 (22/42)	BH3550 (38/57)	nd	(Canchaya *et al.*, 2003)
B. weihensteph-anensis KBAB4	--	BcerKBAB4DRAFT 1206 (32/54)	BcerKBAB4DRAFT 1207 (34/51)	nd	--
B. weihensteph-anensis KBAB4	--	BcerKBAB4DRAFT 3005 (31/53)	BcerKBAB4DRAFT _3006 (32/53)	nd	--
C. perfringens ATCC13124	--	CPF_1021 (30/48)	CPF_1020 (33/53)	nd	(Myers *et al.*, 2006)
C. perfringens ATCC 13124	--	CPF_1604 (34/52)	CPF1607 (33/58)	nd	--
Clostridium sp. OhILAs	--	ClosDRAFT_0334 (22/49)	ClosDRAFT_0333 (30/52)	nd	--
Desulfitobacterium	--	DhafDRAFT_2630	DhafDRAFT_2631	nd	--

hafniense DCB-2		(22/40)	(N.S.)		
Desulfotomaculum reducens MI-1	--	Dred_0020 (30/48)	Dred_0019 (31/50)	nd	--
Enterococcus faecalis V583	--	EF2544 (45/68)	EF2545 (42/64)	nd	(Paulsen *et al.*, 2003)
E. faecium DO ctg653	--	EfaeDRAFT_2195 (29/47)	EfaeDRAFT_2196 (N.S.)	nd	--
Lactobacillus reuteri F275	--	Lreu_1144 (33/61)	Lreu_1145 (26/45)	nd	--
Listeria innocua CLIP 11262	--	Lin1762 (31/52)	Lin1763 (28/47)	nd	(Glaser *et al.*, 2001)
L. innocua CLIP 11262	--	Lin1234 (32/56)	Lin1233 (35/55)	nd	(Glaser *et al.*, 2001)
L. monocytogenes 1/2a F6854	--	LMOf6854_2699 (36/55)	LMOf6854_2699.1 (29/52)	nd	(Rasko *et al.*, 2004)
Staphylococcus haemolyticus JSC1435	--	SH1805 (26/55)	SH1806 (N.S.)	nd	--
Thermoanaerobacter tencongensis	--	TTE2125 (26/49)	TTE2126 (26/45)	nd	(Bao *et al.*, 2002)

[1] All ImmR-like proteins contain a predicted phage repressor helix-turn-helix motif identified by C-DART (Geer *et al.*, 2002). For those proteins that share significant sequence identity with ImmR, the % amino acid identity and similarity is reported. The protein that does not share significant sequence identity with ImmR (not significant, N.S.) was identified due to the presence of a protein that shares sequence identity with ImmA.

[2] All ImmA-like proteins contain predicted Zinc metalloprotease motifs identified by C-DART (Geer *et al.*, 2002). For those proteins that share significant sequence identity with ImmA, the % amino acid identity and similarity is reported. Proteins that do not share significant sequence identity with ImmA (not significant, N.S.) were identified due to the presence of a protein that shares sequence identity with ImmR.

[3] Din indicates damage inducible. Yes, no, or not determined (nd).

[4] Reference is given for the DNA sequence. When no reference is indicated (--), data were retrieved from unpublished sequences deposited in NCBI.

[5] *skin* is a defective prophage.

[6] Putative mobile genetic elements were identified based on the presence of multiple genes predicted to encode proteins homologous to those found in bacteriophage, transposons, or conjugative elements.

The ImmA-like protein from the *B. subtilis* phage ø105 is required for inactivation of the phage repressor in response to DNA damage

We investigated the roles of the *immR*- and *immA*-like genes encoded by the *B. subtilis* bacteriophage ø105 (*cø*105 and *orf2*, respectively) in regulating phage gene expression. The organization of these ø105 genes is similar to that of their counterparts in ICE*Bs1*. *cø*105, *orf2*, and *int* are in a putative operon that is transcribed divergently from other phage genes (Fig. 3A) (Van Kaer *et al.*, 1987). cø105 is the phage repressor, and it regulates transcription of itself and the divergent *orf4* (Van Kaer *et al.*, 1987). ø105 lysogens are induced by the *recA*-dependent DNA damage response (Love and Yasbin, 1984).

We constructed a fusion of the promoter for *orf4* to *lacZ* (P*orf4-lacZ*) and analyzed repression and derepression during DNA damage induced by MMC. Consistent with previous findings (Van Kaer *et al.*, 1987), P*orf4-lacZ* was expressed at very low levels in ø105 lysogens (Fig. 3B). Expression was significantly induced shortly after treatment of cells with MMC (Fig. 3B). Expression of P*orf4-lacZ* was high in non-lysogenic cells lacking ø105 and was not further induced by addition of MMC (Fig 3B).

The high level of expression of P*orf4-lacZ* in non-lysogenic cells was significantly reduced by expression of *cø*105 (Fig. 3B), indicating that the ø105 repressor, in the absence of any other phage gene products, is sufficient to repress transcription from P*orf4*. However, in cells in which *cø*105 was the only phage gene expressed, P*orf4-lacZ* was not derepressed in response to DNA damage induced by addition of MMC (Fig. 3B), indicating that at least one other ø105 gene is needed for induction by DNA damage.

We found that derepression of P*orf4-lacZ* during the DNA damage response required the ImmA homolog Orf2. Concomitant expression of *cø*105 and *orf2* {(ø105)*immA*} allowed

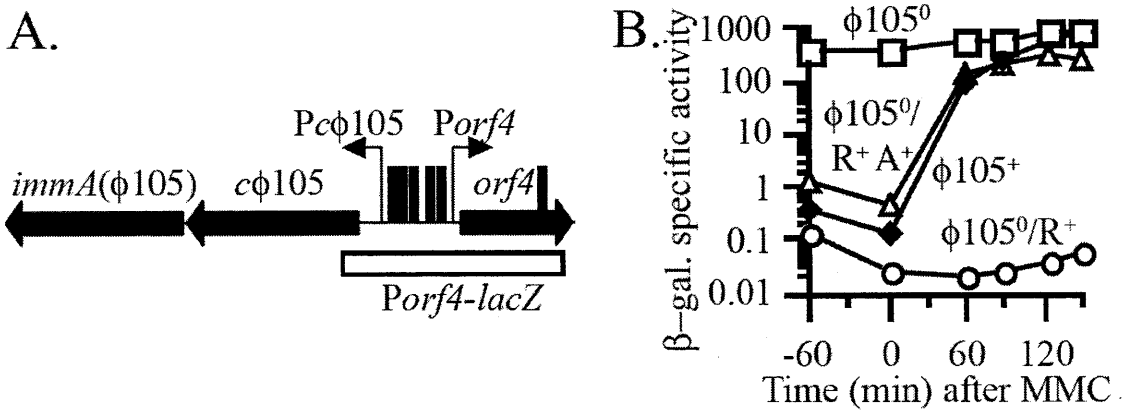

A.
Pcϕ105
Porf4
immA(ϕ105)
cϕ105
orf4
Porf4-lacZ
B.
β–gal. specific activity
1000
100
10
1
0.1
0.01
ϕ105⁰
ϕ105⁰/
R⁺A⁺
ϕ105⁺
ϕ105⁰/R⁺
-60
0
60
120
Time (min) after MMC

Figure 3. The bacteriophage ø105 homolog of ImmA, (ø105)ImmA, is needed for de-repression of phage gene expression in response to DNA damage.

A. Diagram of the region of ø105 that contains *immA*(ø105) {aka *orf2*}, *c*ø105, and *orf4*. Genes are indicated by thick black arrows, with the name of each gene indicated above. Thin arrows indicate the positions of the *orf4* and *c*ø105 promoters (Van Kaer *et al.*, 1987). Black rectangles indicate the positions of the six ø105 repressor bindings sites, three upstream from and adjacent to the leftward promoter P*c*ø105, two upstream from and adjacent to the rightward promoter P*orf4*, and one internal to *orf4* (Van Kaer *et al.*, 1989). The white box underneath the map indicates the region of the *orf4* promoter cloned upstream of *lacZ* that was used to assay gene expression in B.

B. P*orf4-lacZ* expression was monitored in a ø105 lysogen (BOSE447, ♦, ø105^{+}), and in cells lacking ø105 that were otherwise wild-type (BOSE446, □, ø105^{0}), expressed the ø105 repressor from its native promoter (BOSE451, ○, ø105^{0}/R^{+}), or expressed the ø105 repressor from its native promoter and also expressed ImmA(ø105) from the IPTG-inducible promoter P*spank* (BOSE567, △, ø105^{0}/R^{+}A^{+}). Cells were grown in minimal medium containing IPTG and were treated with mitomycin C at OD600 ~ 0.5. Samples were collected at the times indicated and β-galactosidase specific activity was determined.

efficient derepression of P*orf4-lacZ* after addition of MMC (Fig. 3B). Based on these results, we conclude that the ø105 repressor and (ø105)ImmA function analogously to ImmR and ImmA of ICE*Bs1*. It is likely that other ImmR and ImmA homologs similarly regulate the activities of their respective mobile elements.

ImmA is required for degradation of ImmR

To explore the mechanism of ImmA-mediated inactivation of ICE*Bs1* ImmR in vivo, we monitored the fate of ImmR under conditions that induce ICE*Bs1* gene expression, excision, and mating. We found that under these conditions, ImmR is proteolyzed in vivo in an ImmA-dependent manner.

We measured the stability of ImmR in vivo in pulse-chase experiments in ICE*Bs1*^{0} cells that expressed *immR* and *immA* ectopically and expressed *rapI* under control of a xylose-inducible promoter (P*xyl-rapI*). We monitored the amount of radioactively labeled ImmR by immunoprecipitation at various times before and after production of RapI. ImmR was stable in the absence of induction of *rapI* (Fig. 4A; lanes 1-4). However, 10-20 minutes after xylose was added to cultures to induce RapI overproduction, pulse-labeled ImmR was degraded (Fig. 4A, lanes 5-8). These results demonstrate that *rapI* overexpression promotes degradation of ImmR in cells that also express ImmA.

We also assessed ImmR levels using Western blots. We found that one hour after treatment with MMC (Fig. 4B) or overexpression of *rapI* (Fig. 4C), the level of ImmR protein was greatly reduced. This decrease occurred in cells that contained ICE*Bs1* as well as in ICE*Bs1*^{0} cells that expressed *immR* and *immA* (Fig. 4B, C). In contrast, in the absence of *immA*, ImmR was stable, even under inducing conditions (Fig. 4B, C), indicating that the reduced level of ImmR was dependent on the presence of ImmA.

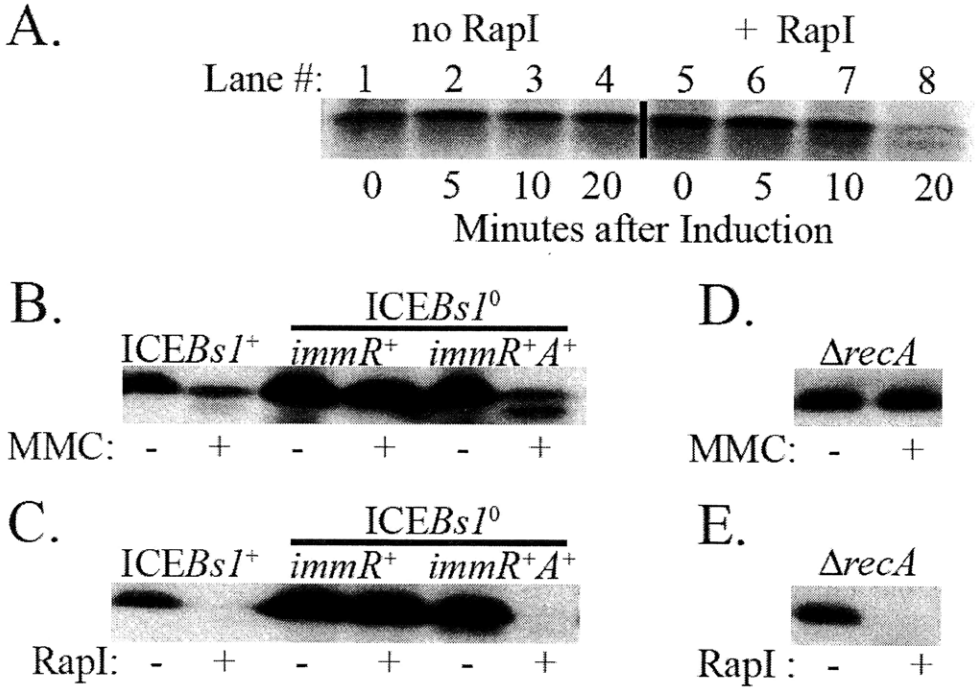

A.
no RapI
+ RapI
Lane #: 1 2 3 4 5 6 7 8
0 5 10 20 0 5 10 20
Minutes after Induction
B.
ICEBs1⁰
ICEBs1⁺ immR⁺ immR⁺A⁺
MMC: - + - + - +
C.
ICEBs1⁰
ICEBs1⁺ immR⁺ immR⁺A⁺
RapI: - + - + - +
D.
ΔrecA
MMC: - +
E.
ΔrecA
RapI : - +

Figure 4. ImmA promotes degradation of ImmR in vivo.

A. ImmR stability was monitored through pulse-chase experiments. ICE*Bs1*^{0} cells that co-expressed *immR* and *immA* from their native promoter and were otherwise wild-type (JMA436, Lanes 1-4) or expressed *rapI* from a xylose-inducible promoter (CAL746, Lanes 5-8) were grown in defined minimal medium lacking xylose. At OD600 ~ 0.5, ^{35}S-met was added to the cultures. One minute later >1000-fold excess unlabeled methionine was added along with xylose to induce expression of *rapI*. Samples were collected just prior to addition of unlabeled methionine and xylose, and 5, 10, and 20 minutes after addition. Samples were immunoprecipitated with anti-ImmR antibodies and analyzed by polyacrylamide gel electrophoresis followed by phosphorimaging.

B-E. ImmR levels were monitored with Western blots using anti-ImmR antibodies. Cultures were grown in minimal medium to an OD600 ~ 0.5 and split in two. Cells were untreated (-) (B-E), or treated with MMC to induce the SOS response (+) (B, D), or treated with IPTG to overproduce RapI (+) (C, E). Samples were collected 60 minutes after the indicated treatment.

B. ICE*Bs1*$^+$ (JMA201); ICE*Bs1*^{0} P*immR-immR* (JMA421); ICE*Bs1*^{0} P*immR-immR immA* (JMA436)

C. All three strains contained P*spank(hy)-rapI*. ICE*Bs1*$^+$ (KLG126); ICE*Bs1*^{0} P*immR-immR* (JMA444); ICE*Bs1*^{0} P*immR-immR immA* (JMA446)

D. ICE*Bs1*$^+$ Δ*recA* (IRN444)

E. ICE*Bs1*$^+$ P*spank(hy)-rapI* Δ*recA* (CAL92).

Whereas *immA* was required for degradation of ImmR under both conditions tested, *recA* was only required for ImmA-dependent ImmR degradation during the SOS response (Fig. 4D) and not after overproduction of RapI (Fig. 4E). *rapI* was not required for ImmA-dependent degradation of ImmR during the SOS response (Fig. 4B). Similarly, previous studies showed that either activated RecA or RapI induce ICE*Bs1* gene expression, excision, and transfer (Auchtung *et al.*, 2005). Taken together, our results demonstrate that RecA and RapI function independently to cause ImmA-dependent degradation of ImmR and subsequent derepression of ICE*Bs1* gene expression, excision, and transfer.

ImmA interacts directly with ImmR

We hypothesized that ImmA-dependent degradation of ImmR might involve direct interaction between the two proteins. Using a yeast two-hybrid system, we found that ImmA and ImmR interact. We constructed fusion proteins of ImmR to the Gal4 activation domain and ImmA to the Gal4 DNA binding domain. We introduced these fusions into *Saccharomyces cerevisiae* cells that had *ADE2*, a gene required for adenine synthesis, under control of a Gal4-activated promoter (James *et al.*, 1996). Growth of these cells on medium lacking adenine requires an interaction between ImmR and ImmA to unite the two domains of Gal4 and to activate transcription of *ADE2*. Cells that contained both the ImmR and ImmA fusion proteins were able to grow on medium lacking adenine (data not shown). This result indicates that ImmR and ImmA can interact, and based on the phenotypes of *immR* and *immA* mutants, we suspect that these proteins interact in *B. subtilis*.

Several phage repressors act as dimers or multimers to regulate gene expression {e.g., (Little, 1984; Walker, 1996)}. We detected self-interaction between ImmR fusion proteins (but not

ImmA), indicating that ImmR likely forms a dimer or higher order multimer to regulate gene expression.

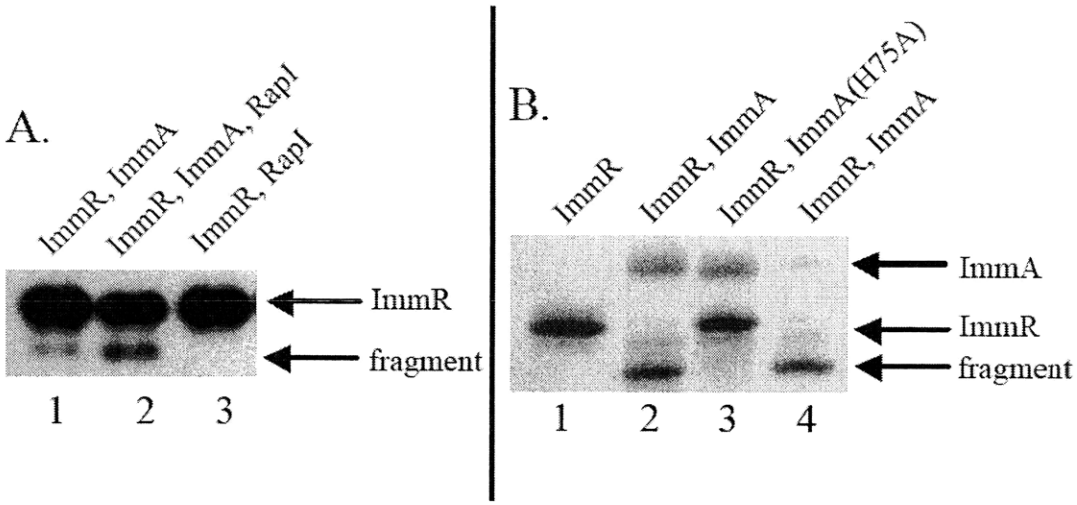
ImmA causes degradation of ImmR in *E. coli*

To further explore the action of ImmA outside of *B. subtilis,* we assayed ImmR degradation in *E. coli*. Using Western blots, we observed a shortened form of ImmR when N-terminally His-tagged ImmR (His6-ImmR) and ImmA were co-expressed in *E. coli* (Fig. 5A). This fragment was not detected in the absence of ImmA, and the relative abundance of the fragment was increased by concurrent expression of RapI (Fig. 5A). Mass analysis of the His-tagged ImmR fragment (Fig. 5C) indicated that in *E. coli*, ImmR is cleaved between F95 and M96 (Fig. 5D) in an ImmA-dependent fashion.

Purified ImmA causes cleavage of ImmR in vitro

To determine its mechanism of action, we purified ImmA and tested its activity in vitro. We overproduced and purified His6-ImmR and untagged ImmA from *E. coli*. Incubation of purified ImmA with His6-ImmR led to cleavage of ImmR (Fig. 5B). There was no detectable cleavage of His6-ImmR in the absence of ImmA (Fig. 5B). To be sure that the proteolytic activity observed in vitro was due to ImmA and not a contaminant, we also purified and tested ImmA(H75A) mutant protein. His6-ImmR was not cleaved in vitro during incubation with ImmA(H75A). Previously we had purified a C-terminally His-tagged ImmR (Auchtung *et al.*, 2007). This protein was also cleaved in vitro when incubated with ImmA, but remained stable during incubation alone or with ImmA(H75A) (data not shown).

Although we did not detect autocleavage of ImmR in vitro in incubations containing ImmR alone (above), repressors from phages λ, 434, and P22 cleave themselves in vitro when incubated with activated RecA or without RecA at elevated pH. We conducted additional tests



C.

system	ImmR	ImmA	RapI	ImmR molecular mass expected (Da)	ImmR molecular mass observed (Da)
in vitro	His6-ImmR	ImmA	-	12961.3 3738.3	12960.9 3741.6
in vitro	His6-ImmR	ImmA(H75A)	-	16681.6	16682.2
in vitro	ImmR-His6	ImmA	-	11060.4 4803.4	11060.3 4806.7
E. coli	His6-ImmR	ImmA	RapI	12961.3	12961.8

D.

```
  1  (Helix-turn-Helix domain)
 62 Leu Gly Lys Asp Glu Val Ser Lys Lys Asn
 72 Glu Thr Asp Leu Leu Asn Lys Thr Ile Asn
 82 Glu Ala Ile Gln Glu Leu Lys Asp Glu Asp
 92 Thr Leu Leu Phe/Met Asn Asp Gly Glu Phe
102 Asp Glu Glu Thr Ala Arg Leu Val Lys Lys
112 Ala Leu Lys Asn Gly Ile Lys Phe Ile Asp
122 Leu Lys Lys Lys Glu
```

Figure 5. ImmA-mediated cleavage of ImmR in *E. coli* and in vitro.

A. ICE*Bs1* proteins were overexpressed in *E. coli* and lysates were analyzed by Western blot using anti-ImmR antibodies. Lysates were collected two hours after induction from cells that overexpressed His6-ImmR with ImmA (BOSE799, lane 1), with ImmA and RapI (BOSE817, lane 2), or with RapI (BOSE819, lane 3).

B. His6-ImmR (17μM) was incubated in vitro overnight at 37°C: alone (lane 1); with 22μM ImmA (lane 2); with 18μM ImmA(H75A) (lane 3); or with 11μM ImmA (lane 4). Reactions were run on SDS-PAGE and stained with Coomassie brilliant blue. Bands representing ImmA, His6-ImmR, and the ImmR fragment are indicated.

C. Products of ImmR cleavage assays in vitro and in *E. coli* were analyzed by MALDI-TOF. The form of the ImmA, ImmR, and RapI present in each reaction is indicated. Experimentally determined masses are shown alongside expected masses for the cleavage occurring between F95 and M96 of ImmR. The smaller ImmR fragment was detected from the in vitro reactions by not from *E. coli*.

D. The cleavage site in ImmR. The helix-turn-helix domain, contained in the first 61 N-terminal residues, is the presumed DNA binding domain that is similar to that of other phage-like repressors. The amino acid sequence from residue 62 to the C-terminus is shown with the cleavage site between F95 and M96 marked with a slash.

for autocleavage of ImmR by incubating ImmR-His6 over a pH range from 4 to 11 at 30°, 37°, or 42° for ~19 hrs and ~188 hrs. We observed no evidence of specific autocleavage (data not shown).

We determined the masses of ImmR fragments from the ImmA-dependent cleavage reactions by mass spectrometry (MALDI-TOF). In reactions with His6-ImmR and ImmA, we detected masses corresponding to two fragments of ImmR (Fig. 5C), consistent with site-specific cleavage between F95 and M96 of ImmR (Fig. 5D). The mass of the larger fragment matched that of the His6-ImmR fragment from ImmA-dependent cleavage in *E. coli* (Fig. 5B, C). In the presence of ImmA(H75A), we only detected full length His6-ImmR (Fig. 5C). In reactions with C-terminally tagged ImmR-His6 and ImmA, we also detected mass peaks consistent with site-specific cleavage between F95 and M96 (Fig. 5C). Taken together, our results indicate that ImmR is cut in an ImmA-dependent manner at a single site, between F95 and M96 (Fig. 5D).

Discussion

Using in vivo and in vitro analyses, we found that ImmA is an anti-repressor that most likely functions as a site-specific protease to cleave and inactivate the repressor ImmR thereby causing derepression of the mobile genetic element ICE*Bs1*. ImmA mediates ICE*Bs1* induction in response to cues of DNA damage and cell-cell signaling. ImmA homologs are widely conserved in mobile genetic elements and the mechanism of action of members of this anti-repressor family appears to be quite different from that of other characterized anti-repressors involved in horizontal gene transfer.

SOS-induced auto-proteolysis of repressors

A well understood mechanism of de-repression of mobile genetic elements is the RecA-stimulated autocleavage of phage repressors in response to DNA damage. RecA bound to single-stranded DNA stimulates autocleavage of the repressors of phages λ, 434, and P22 {e.g., (Little, 1984; Roberts and Roberts, 1975; Roberts *et al.*, 1978; Sauer *et al.*, 1982)}. RecA also appears to facilitate autocleavage of repressors of two conjugative transposons SXT and ICE*St1* (Beaber *et al.*, 2004; Bellanger *et al.*, 2007). In addition, RecA facilitates autocleavage of the cellular SOS repressor, LexA (Little, 1984; Miller *et al.*, 1996). In all cases, autocleavage depends on the presence of conserved catalytic residues in the C terminus of the repressor (Slilaty and Little, 1987).

Proteolysis mediated by the anti-repressor ImmA

Destruction of the ICE*Bs1* repressor ImmR does not resemble auto-proteolysis of previously characterized repressors. ImmR does not contain the conserved residues necessary for autocleavage of LexA and phage repressors. The cleavage site in ImmR is between a phenylalanine and methionine, whereas the autocleavage site of LexA and similar repressors is typically between an alanine and glycine. We were unable to detect specific autocleavage of purified ImmR under conditions suitable for in vitro autocleavage of characterized phage repressors (Little, 1984; Slilaty *et al.*, 1986).

Unlike previously characterized phage repressors, cleavage of ImmR in vivo and in vitro requires the anti-repressor ImmA and there is no evidence to suggest that ImmR is capable of cleaving itself. The simplest conclusion from our data is that the anti-repressor ImmA is a site-specific protease required for inactivation of the repressor ImmR.

The mechanism(s) by which RapI and RecA stimulate proteolysis of ImmR by ImmA is unknown. RapI and RecA may increase synthesis, stability, or specific activity of ImmA. Expression of *immA* from the IPTG-inducible promoter P*spank* did not derepress P*xis-lacZ* unless RapI was overproduced or DNA damage was induced (Fig. 2C,D). Therefore, RapI and RecA likely affect the anti-repressor ImmA at a level other than transcription initiation.

Anti-repressors and horizontal gene transfer

Several anti-repressors from mobile genetic elements have been characterized, but none are known to cause degradation of the target repressor. Many anti-repressors function by direct interaction with their cognate repressor. For example, Tum, Coi, E, and Ant, antirepressors from bacteriophage 186, P1, P4, and P22, respectively, all form complexes with their respective repressors and inhibit the ability of these repressors to bind DNA (Heinzel *et al.*, 1992; Liu *et al.*, 1998; Shearwin *et al.*, 1998; Susskind and Youderian, 1983). The satellite phage RS1 encodes RstC, an antirepressor that promotes aggregation of the RstR repressor of the CTX phage (Davis *et al.*, 2002). In contrast, the λ Cro protein binds directly to DNA and competes with the λ CI repressor for its binding sites (Ptashne, 1992), and references therein).

Conservation of ImmA and ImmR

ImmR and ImmA homologs reside in other known or putative mobile genetic elements (Table 1). We found that the ImmR and ImmA homologs encoded by the *B. subtilis* bacteriophage ø105 (cø105 and (ø105)ImmA, respectively) function as a repressor and anti-repressor to regulate expression of genes in ø105. Thus, we propose that these proteins represent a conserved strategy for regulating the activity of mobile genetic elements.

ICE*Bs1* differs from previously characterized DNA damage-inducible mobile genetic elements in that it requires the anti-repressor ImmA for inactivation of the repressor ImmR, and

in that excision and mating occur in response to cues from cell-cell signaling as well as cues of DNA damage. RapI and RecA function independently of each other to mediate cues from these different pathways and stimulate transfer of ICE*Bs1* by stimulating ImmA-mediated proteolysis of ImmR. ImmA homologs in other mobile genetic elements may be capable of responding to more than one signal, thereby allowing multiple conditions to affect horizontal gene transfer.

Experimental Procedures

Media and growth conditions

B. subtilis cells were grown at 37°C with aeration in S7 minimal salts medium {(Vasantha and Freese, 1980), except that 50 mM MOPS was used instead of 100} supplemented with 1% glucose or arabinose, 0.1% glutamate, 40 μg/ml tryptophan and phenylalanine, and 120 μg/ml threonine. Arabinose was used when *rapI* was expressed from P*xyl*. When appropriate, antibiotics were used at the following concentrations: chloramphenicol (5 μg/ml); kanamycin (5 μg/ml); tetracycline (10 μg/ml); spectinomycin (100 μg/ml); and erythromycin (0.5 μg/ml) and lincomycin (12.5 μg/ml) together to select for macrolide-lincosamide-streptogramin B (MLS) resistance. IPTG and MMC (Sigma) were used at final concentrations of 1 mM and 1 μg/ml respectively.

S. cerevisiae cells were grown at 30°C in yeast peptone dextrose medium or synthetic complete medium lacking uracil and leucine or lacking uracil, leucine, and adenine (Sambrook and Russell, 2001).

E. coli cells were grown in LB at 30°C or 37°C. When appropriate, ampicillin (100-200 μg/ml) and or chloramphenicol (15 μg/ml) were added. To induce gene expression in BL21-AI cells (Invitrogen), growth medium was supplemented with 1-4mM IPTG and 0.2% arabinose.

Strains and alleles

B. subtilis strains used in this study are listed in Table 2. Standard techniques were used for cloning and strain construction (Harwood and Cutting, 1990; Sambrook and Russell, 2001). The ICE*Bs1*^{0} strain, and the Δ*int*205::*cat*, *amyE*::{P*spank(hy)-rapI*) *spc*}, *cgeD*::{(P*spank(hy)-rapI*) *kan*}, *cgeD*::{(P*immR-immR*) *kan*}, *thrC*::{(P*xis-lacZ*Ω343) *mls*}, Δ*recA260*::*cat-mls*, and Δ*(rapI phrI)*::*kan* alleles were previously described (Auchtung *et al.*, 2005; Auchtung *et al.*, 2007). (Note: *mls* is the same as the previously used *erm*).

The promoter for *orf4* from *B. subtilis* phage ø105 was fused to *lacZ* (P*orf4-lacZ*) by cloning from 298 bp upstream of the *orf4* start codon to 252 bp into *orf4* upstream of the promoterless *lacZ* in the vector pDG793 (Guerout-Fleury *et al.*, 1996). This promoter fragment contains all six immunity repressor binding sites known to control expression of *orf4* (Van Kaer *et al.*, 1989). This fusion was integrated into *thrC* by homologous recombination.

immR and *immA* from ICE*Bs1* were cloned together under control of their own promoter into the integration vector pMMB124 (Auchtung *et al.*, 2007). Sequences extended from 268 bp upstream of the *immR* start codon to 2 bp downstream of the *immA* stop codon. *c*ø105 was cloned with its own promoter into pMMB124. Sequences extended from 293 bp upstream of the *c*ø105 start codon to 24 bp downstream of the stop codon. Each clone was introduced into the *B. subtilis* chromosome at *cgeD* by homologous recombination.

LacI-repressible, IPTG-inducible copies of *immA* and *immA*(ϕ105) were generated by cloning from 24 bp upstream of *immA* to 2 bp downstream of its stop codon and from 28 bp upstream of *immA*(ϕ105) to 29 bp downstream of its stop codon downstream of P*spank* in pDR110 (Rokop *et al.*, 2004), followed by integration of these plasmids at *amyE* by homologous

recombination. P*spank-immAH75A* was constructed using Quikchange (Invitrogen) site-directed mutagenesis of a plasmid containing P*spank-immA*. The xylose-inducible copy of *rapI* was constructed by M. Berkmen and was created by cloning *rapI* downstream of P*xyl* in vector pDR160, (from D. Rudner) followed by integration of the plasmid at *amyE* by homologous recombination.

Δ*immA* Δ*int*::*cat* was generated through a combination of splicing by overlap extension (SOE) and long-flanking homology PCR. This construct creates a deletion of *immA* (first 3 codons at 5' end joined to last two codons at 3' end) linked to a replacement of +53 to +1097 of *int* with the chloramphenicol resistance gene from pGEM*cat* (Youngman *et al.*, 1989). This *int* deletion removes the same sequence as Δ*int205*::*cat* (Auchtung *et al.*, 2005).

The state of ImmR in *E. coli* was monitored in strains BOSE799 (BL21-AI pBOSE794), BOSE817 (BL21-AI pBOSE794, pBOSE801), and BOSE819 (BL21-AI pBOSE792, pBOSE801). pBOSE792 expresses immR with the N-terminal tag MGSSH$_6$SSGLVPRGSH from the T7 promoter in pET14b (Novagen). pBOSE794 expresses the aforementioned N-terminally tagged ImmR and untagged ImmA from the T7 promoter in pET14b. pBOSE801 expresses untagged RapI from pBAD33 (Guzman *et al.*, 1995), with an exogenous rbs incorporated during cloning.

Proteins for in vitro assays were purified from the following strains. N-terminally His-tagged ImmR was purified from BOSE798 (BL21-AI pBOSE792); pBOSE792 is described above. C-terminally His-tagged ImmR was purified from JMA622 (BL21-AI pJMA605); pJMA605 was previously described (Auchtung *et al.*, 2007). ImmA was purified from BOSE848 (BL21-AI pBOSE831). pBOSE831 expresses untagged immA from the T7 promoter in pSA27 (Duncan *et al.*, 1996). In pBOSE831, the codon for V14 was mutated (gtg to gta) by Quikchange

(Stratagene) to prevent translation from starting at this site. ImmA(H75A) was purified from BOSE847 (BL21-AI pBOSE841). pBOSE841 was made by Quikchange site-directed mutagenesis of H75 to A in the immA encoded by pBOSE831.

Yeast two-hybrid assays

Plasmids encoding ImmR and ImmA fused to the Gal4 DNA binding domain (Gal4-BD) or Gal4 activation domain (Gal4-AD) were generated by cloning the coding sequence of *immR* or *immA* in the same reading frame as the upstream Gal4-AD coding sequence in plasmid pGAD-c1 (James *et al.*, 1996) or the upstream Gal4-BD coding sequence in plasmid pGBDu-c3 (James *et al.*, 1996). Plasmids containing the ImmR and ImmA Gal4 binding and activation domain fusions were transformed into *S. cerevisae* strain PJ69-4A (*trp-901 leu2-3 ura3-52 his3-200 gal4Δ gal80Δ LYS2*::(*GALl-HIS3*) *GAL2-ADE2 met2*::(*GAL7-lacZ*) (James *et al.*, 1996) selecting for growth on synthetic complete medium lacking uracil and leucine. Six transformants were purified to single colonies on medium without uracil and leucine, then re-purified and tested on the same medium without and with adenine. A pair of proteins were designated as interacting if all 6 transformants grew on medium lacking adenine. Similar results were observed with 6 transformants from each of three independent transformation experiments.

Pulse-chase experiments

Cells were grown in defined minimal medium containing arabinose to an OD600 ~ 0.5, at which point ^{35}S-labeled methionine (1,100 mCi/mmol) was added to the culture to a final concentration of between 30 - 60 μCi/ml. Cells were incubated with label for 1 min and an aliquot was removed (T=0 sample). Unlabeled methionine (50 μM final concentration) and xylose (2% final concentration to induce expression from P*xyl*) were added. Samples were

collected at 5, 10, and 20 minutes as indicated. All samples were frozen immediately on dry ice and stored at -80°C.

Samples were thawed and centrifuged to recover the cells. Cells were resuspended in buffer (10 mM Tris 1 mM EDTA) containing 20 mg/ml lysozyme and 50 μg/ml PMSF and incubated on ice for 30 min. Cell pellets were resuspended in a volume of buffer proportionate to the OD600 at which they were recovered in order to normalize the concentration of proteins in the sample. SDS was added to 1%, then cells were lysed by heating at 100°C for 10 min. Twenty volumes of KI buffer (50 mM Tris 150 mM NaCL 1 mM EDTA 0.5% Triton X-100) containing 25 μg/ml PMSF were mixed with the cell lysates to reduce the SDS concentration to 0.05%. Samples were then centrifuged at 4°C for 15 min. The supernatant was removed and incubated overnight in the presence of rabbit polyclonal anti-ImmR sera (Covance) at 4°C (~1:300 dilution). Bound protein was separated from the supernatant by affinity purification with protein A sepharose. Bound protein was washed six times with KI Buffer containing PMSF. Proteins were eluted at 70°C for 15 min. in 2X SDS Buffer (0.2 M DTT 4% SDS 100 mM Tris-HCl pH 6.8 20% glycerol 0.005% bromophenol blue). The immunoprecipitated samples were analyzed by SDS-PAGE on 15% gels followed by electro-blotting to PVDF membranes. Radioactivity was detected by phosphorimaging using the Typhoon imager 9400 (Amersham Biosciences).

Western Blots

B. subtilis cells were grown in defined minimal glucose medium to OD600 ~ 0.5. Cultures were split and inducer (either MMC or IPTG) was added to half of the cultures. Samples were collected 60 minutes after addition of inducer from induced and uninduced cultures. Cells were pelleted by centrifugation, washed with TN Buffer (50 mM Tris 300 mM NaCl, pH8), and stored at –20°C.

Cell pellets were thawed on ice, resuspended in an appropriate volume of Buffer (10 mM Tris 10 mM EDTA, pH 8) containing 0.05 mg/ml lysozyme and 1 mM AEBSF, and incubated at 37°C for 20 min. Pellets were resuspended in a volume of buffer proportionate to the OD600 at which they were recovered in order to normalize the concentration of proteins in the sample. SDS-Loading Buffer was then added to samples, which were heated at 100°C for 10 min. followed by centrifugation to remove insoluble material.

Proteins were separated by SDS-PAGE on 15% gels and transferred to PolyScreen PVDF membrane (Perkin Elmer) using the Trans-blot semi-dry electro-blot transfer apparatus (BioRad). Membranes were blocked in 0.2% I-Block (Tropix) in TBST (50 mM Tris 200 mM NaCl 0.05% Tween-20 pH 8) either at room temperature for 1 hr. or overnight at 4°C. Membranes were incubated in 1:10,000 anti-ImmR rabbit polyclonal antisera in 0.2% I-Block TBST for 1 hr. at room temperature, washed several times in TBST, incubated in 1:3,000 goat anti-rabbit IgG-HRP conjugate (BioRad) for 1 hour at room temperature, and washed several times in TBST. Signals were detected using Western Lightning chemiluminescence reagents (Perkin-Elmer) followed by exposure to Kodak Biomax Light film.

The amount of ImmR was consistently higher in cells that expressed *immR* ectopically from its native promoter than in ICE*Bs1*$^+$ cells, and this made the pulse-chase experiments technically easier. We suspect that this increased production of ImmR is due to the loss of a potential ImmR binding site upstream of the promoter that might be involved in auto-repression (Auchtung *et al.*, 2007). Increased production of ImmR is likely not due to the absence of another protein in ICE*Bs1* because levels of ImmR protein in cells lacking most of the genes in ICE*Bs1* (Δ*xis-yddM*) and expressing *immR, immA*, and *int* from its native promoter were similar to ICE*Bs1*$^+$ cells (data not shown).

β-galactosidase assays

β-galactosidase specific activity was assayed as described (Jaacks *et al.*, 1989). Specific activity was calculated relative to the optical density at 600 nm of the samples and is plotted relative to the time of treatment as indicated. Results shown are from a single experiment and are representative of results obtained in at least two independent experiments.

Comparative sequence analysis

Potential homologs of ImmR and ImmA were identified by a combination of protein BLAST (Altschul *et al.*, 1997) and C-DART (Geer *et al.*, 2002). More specifically, protein-protein BLAST was used to identify potential homologs of ImmR or ImmA that shared significant sequence identity and similarity. The adjacent or nearby genes were then analyzed using BLAST 2 (Tatusova and Madden, 1999) or C-DART to identify those proteins that had either significant sequence identity and similarity to ImmR or ImmA or contained the conserved domains found in ImmR (XRE family helix-turn-helix) or ImmA (COG2856). When pairs of *immR*- and *immA*-like genes were identified, the genes in the surrounding area were analyzed to determine if they encoded proteins likely indicative of a mobile genetic element or remnant (e.g., putative transposases, integrases, or other phage-related proteins). Many proteins were found to be ImmR- and ImmA-like and over 80% of these appear to be in putative mobile elements or remnants of mobile elements (Table 1).

ImmR cleavage in *E. coli*

ImmR in *E. coli* was monitored by Western blots of samples from BOSE799, BOSE817, and BOSE819. Strains were grown in shaking LB plus antibiotic(s) at 37°C to an OD600 of 0.45-0.65 and then induced with 4mM IPTG and 0.2% arabinose. Growth was continued for 2 hours, cells were pelleted, and pellets were stored at -20°C. For Western blots, cells were resuspended

into TN buffer, mixed with 5x SDS sample buffer, boiled, and run on SDS-PAGE. Western blots to detect ImmR were performed as described above. His6-ImmR and the His-tagged ImmR fragment were purified from BOSE817 pellets using the lysis and chromatography procedures for purification of His6-ImmR from BOSE798, described below.

Protein purification

N-terminally His-tagged ImmR was purified from BOSE798. Cells were grown with shaking in LB at 37°C to OD600~ 0.5 and induced with 1mM IPTG and 0.2% arabinose. Approximately 4 hours after induction, cells were pelleted, and pellets were stored at -20°C. Cells were lysed by inversion at room temperature in a lysis buffer (43.4mM Tris, 263mM NaCl, 1x Cellytic, 196μg/ml lysozyme, 4.9μg/ml DnaseI, 9.8mM imidazole pH8). The lysate was centrifuged at 17,000 rcf, and His-tagged protein was purified from the supernatant by Ni-NTA (Qiagen) affinity chromatography using TN buffer supplemented with increasing concentrations of imidazole (10mM-1M). Eluted protein was dialyzed against TN buffer, then mixed with an equal volume of glycerol, and stored at 4°C or -80°C.

ImmA and ImmA(H75A) were purified from BOSE848 and BOSE847, respectively. Cells were grown shaking at 30°C overnight in LB with 100μg/ml Amp, 4mM IPTG, and 0.2% arabinose. Cells were pelleted by centrifugation, and pellets were stored at -20°C. Pellets were resuspended in 50mM Tris, 50mM NaCl, 10mM EDTA, 20mM DTT, pH8 and lysed by sonication on ice. Upon centrifugation, ImmA was found in inclusion bodies in the pellet. Inclusion bodies were isolated (Martin and Schmid, 2003), and protein solubilized in 100mM Tris, 6M GdmCl, 1mM EDTA, 50mM reduced glutathione, pH8.5 was stored at 4°C. ImmA was refolded immediately before use in assays by 11-fold dilution on ice into refolding buffer (50mM Tris, 10mM DTT, 5%glycerol, 100μM $ZnCl_2$, 0.8M L-arginine, pH7). Diluted samples were

centrifuged at 16,100 rcf at 4°C for 20 min. Protein in the supernatant was used for in vitro assays.

In vitro assays

Cleavage of ImmR was assayed in vitro by incubating 17μM His6-ImmR in vitro overnight at 37°C alone, with 11μM or 22μM ImmA, or with 18μM ImmA(H75A), as indicated. The final reaction buffer was 43.8mM Tris, 6.44mM DTT, 19.22% glycerol, 48mM NaCl, 64.4 μM $ZnCl_2$, 0.516M L-arginine, 0.22M GdmCl, 36.4 μM EDTA, 1.82mM GSH, pH~7.3. Reactions were mixed with 5x sample buffer, boiled, and run on SDS-PAGE. Gels were stained with Coomassie brilliant blue.

Mass analysis

In vitro reactions and His-tagged protein purified from BOSE817 were analyzed by mass spectometry (MALDI-TOF) by the MIT Biopolymers Laboratory. Expected masses were calculated using the "compute pI/MW tool" on the Expasy proteomics server {http://us.expasy.org/tools/pi_tool.html}.

Table 2. *B. subtilis* strains.

Strain	Relevant genotype
1L11	ø105 lysogen
BOSE447	ø105 lysogen (ø105^{+}) *thrC*::{(P*orf4-lacZ*) *mls*}
BOSE446	*thrC*::{(P*orf4-lacZ*) *mls*}
BOSE451	*cgeD*::{(Pcø105-cø105) *kan*} *thrC*::{(P*orf4-lacZ*) *mls*}
BOSE567	*cgeD*::{(Pcø105-cø105) *kan*} *amyE*::{(P*spank-orf2*) *spc*} *thrC*::{(P*orf4-lacZ*) *mls*}
CAL16	Δ*int*205::*cat thrC*::{(P*xis-lacZ*Ω343) *mls*}
CAL92	*amyE*::{(P*spank*(*hy*)-*rapI*) *spc*} Δ(*rapI phrI*)::*kan* Δ*recA260*::*cat mls*
CAL746	ICE*Bs1*^{0} *cgeD*::{(P*immR-immR immA*) *kan*} *amyE*::{(P*xyl-rapI*) *spc*} *thrC*::{(P*xis-lacZ*Ω343) *mls*}
IRN342	Δ(*rapI phrI*)::*kan*
IRN444	Δ*recA260*::*cat mls*
JMA168	*amyE*::{(P*spank*(*hy*)-*rapI*) *spc*} Δ(*rapI phrI*)::*kan*
JMA201	*thrC*::{(P*xis-lacZ*Ω343) *mls*}
JMA421	ICE*Bs1*^{0} *cgeD*::{(P*immR-immR*) *kan*} *thrC*::{(P*xis-lacZ*Ω343) *mls*}
JMA436	ICE*Bs1*^{0} *cgeD*::{(P*immR-immR immA*) *kan*} *thrC*::{(P*xis-lacZ*Ω343) *mls*}
JMA444	ICE*Bs1*^{0} *cgeD*::{(P*immR-immR*) *kan*} *amyE*::{(P*spank*(*hy*)-*rapI*) *spc*} *thrC*::{(P*xis-lacZ*Ω343) *mls*}
JMA446	ICE*Bs1*^{0} *cgeD*::{(P*immR-immR immA*) *kan*} *amyE*::{(P*spank*(*hy*)-*rapI*) *spc*} *thrC*::{(P*xis-lacZ*Ω343) *mls*}
JMA726	(Δ*immA*720 Δ*int*::*cat*) *thrC*::{(P*xis-lacZ*Ω343) *mls*}
JMA836	Δ*int*205::*cat cgeD*::{(P*spank*(*hy*)-*rapI*) *kan*} *thrC*::{(P*xis-lacZ*Ω343) *mls*}
JMA838	(Δ*immA*720 Δ*int*::*cat*) *cgeD*::{(P*spank*(*hy*)-*rapI*) *kan*} *thrC*::{(P*xis-lacZ*Ω343) *mls*}
JMA840	(Δ*immA*720 Δ*int*::*cat*) *amyE*::{(P*spank-immA*Ω218) *spc*} *thrC*::{(P*xis-lacZ*Ω343) *mls*}
JMA842	(Δ*immA*720 Δ*int*::*cat*) *cgeD*::{(P*spank*(*hy*)-*rapI*) *kan*} *amyE*::{(P*spank-immA*Ω218) *spc*} *thrC*::{(P*xis-lacZ*Ω343) *mls*}
JMA895	ICE*Bs1*^{0} *lacA*::{(P*immR-immR immA*) *tet*} *thrC*::{(P*xis-lacZ*Ω343) *mls*}
JMA900	ICE*Bs1*^{0} *lacA*::{(P*immR-immR immAH75A*) *tet*} *thrC*::{(P*xis-lacZ*Ω343) *mls*}
KLG126	*amyE*::{(P*spank*(*hy*)-*rapI*) *spc*} *thrC*::{(P*xis-lacZ*Ω343) *mls*}

*All *B. subtilis* strains are derived from AG174 (JH642) and contain *trpC2* and *pheA1* except for 1L11 and BOSE447, which are derivatives of *B. subtilis* 168 and contain *trpC2*. 1L11 was obtained from the *Bacillus* Genetic Stock Center.

Acknowledgements

We thank Melanie B. Berkmen for constructing the P*xyl-rapI* fusion and Andreas Martin, Alexi Goranov, and Bob Sauer for helpful discussions and comments on the manuscript. This work was supported in part by an NSF graduate fellowship to BB and Public Health Service grant GM50895 from the NIH to ADG.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-3402.

Au, N., Kuester-Schoeck, E., Mandava, V., Bothwell, L.E., Canny, S.P., Chachu, K., Colavito, S.A., Fuller, S.N., Groban, E.S., Hensley, L.A., O'Brien, T.C., Shah, A., Tierney, J.T., Tomm, L.L., O'Gara, T.M., Goranov, A.I., Grossman, A.D., and Lovett, C.M. (2005) Genetic composition of the Bacillus subtilis SOS system. *J Bacteriol* **187**: 7655-7666.

Auchtung, J.M., Lee, C.A., Monson, R.E., Lehman, A.P., and Grossman, A.D. (2005) Regulation of a Bacillus subtilis mobile genetic element by intercellular signaling and the global DNA damage response. *Proc Natl Acad Sci U S A* **102**: 12554-12559.

Auchtung, J.M., Lee, C.A., Garrison, K.L., and Grossman, A.D. (2007) Identification and characterization of the immunity repressor (ImmR) that controls the mobile genetic element ICEBs1 of Bacillus subtilis. *Mol Microbiol* **64**: 1515-1528.

Bao, Q., Tian, Y., Li, W., Xu, Z., Xuan, Z., Hu, S., Dong, W., Yang, J., Chen, Y., Xue, Y., Xu, Y., Lai, X., Huang, L., Dong, X., Ma, Y., Ling, L., Tan, H., Chen, R., Wang, J., Yu, J., and Yang, H. (2002) A complete sequence of the *T. tengcongensis* genome. *Genome Res* **12**: 689-700.

Beaber, J.W., Hochhut, B., and Waldor, M.K. (2002) Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from Vibrio cholerae. *J Bacteriol* **184**: 4259-4269.

Beaber, J.W., Hochhut, B., and Waldor, M.K. (2004) SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature* **427**: 72-74.

Bellanger, X., Morel, C., Decaris, B., and Guedon, G. (2007) Derepression of excision of integrative and potentially conjugative elements from Streptococcus thermophilus by DNA damage response: implication of a cI-related repressor. *J Bacteriol* **189**: 1478-1481.

Burrus, V., Pavlovic, G., Decaris, B., and Guedon, G. (2002) Conjugative transposons: the tip of the iceberg. *Mol Microbiol* **46**: 601-610.

Burrus, V., Pavlovic, G., Decaris, B., and Guedon, G. (2002) The ICESt1 element of Streptococcus thermophilus belongs to a large family of integrative and conjugative elements that exchange modules and change their specificity of integration. *Plasmid* **48**: 77-97.

Burrus, V., and Waldor, M.K. (2004) Shaping bacterial genomes with integrative and conjugative elements. *Res Microbiol* **155**: 376-386.

Canchaya, C., Proux, C., Fournous, G., Bruttin, A., and Brussow, H. (2003) Prophage genomics. *Microbiol Mol Biol Rev* **67**: 238-276.

Churchward, G. (2002) Conjugative transposons and related mobile elements. In *Mobile DNA II.* Craig, N.L. (ed). Washington, D.C.: ASM Press, pp. 177-191.

Churchward, G.G. (2002) Conjugative Transposons and Related Mobile Elements. In *Mobile DNA II.* Craig, N.L., Craigie, R., Gellert, M. and Lambowitz, A.M. (eds). Washington, D.C.: ASM Press, pp. 177-191.

Courcelle, J., Khodursky, A., Peter, B., Brown, P.O., and Hanawalt, P.C. (2001) Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient Escherichia coli. *Genetics* **158**: 41-64.

Davis, B.M., Kimsey, H.H., Kane, A.V., and Waldor, M.K. (2002) A satellite phage-encoded antirepressor induces repressor aggregation and cholera toxin gene transfer. *Embo J* **21**: 4240-4249.

Dobrindt, U., Hochhut, B., Hentschel, U., and Hacker, J. (2004) Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol* **2**: 414-424.

Duncan, L., Alper, S., and Losick, R. (1996) SpoIIAA governs the release of the cell-type specific transcription factor sigma F from its anti-sigma factor SpoIIAB. *J Mol Biol* **260**: 147-164.

Feng, L., Wang, W., Cheng, J., Ren, Y., Zhao, G., Gao, C., Tang, Y., Liu, X., Han, W., Peng, X., Liu, R., and Wang, L. (2007) Genome and proteome of long-chain alkane degrading Geobacillus thermodenitrificans NG80-2 isolated from a deep-subsurface oil reservoir. *Proc Natl Acad Sci U S A* **104**: 5602-5607.

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

Frost, L.S., Leplae, R., Summers, A.O., and Toussaint, A. (2005) Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* **3**: 722-732.

Geer, L.Y., Domrachev, M., Lipman, D.J., and Bryant, S.H. (2002) CDART: protein homology by domain architecture. *Genome Res* **12**: 1619-1623.

Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couve, E., de Daruvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K.D., Fsihi, H., Garcia-del Portillo, F., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L.M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vicente, J.M., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., de Pablos, B., Perez-Diaz, J.C., Purcell, R., Remmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J.A., Voss, H., Wehland, J., and Cossart, P. (2001) Comparative genomics of *Listeria* species. *Science* **294**: 849-852.

Goranov, A.I., Kuester-Schoeck, E., Wang, J.D., and Grossman, A.D. (2006) Characterization of the Global Transcriptional Responses to Different Types of DNA Damage and Disruption of Replication in Bacillus subtilis. *J Bacteriol* **188**: 5595-5605.

Guerout-Fleury, A.M., Frandsen, N., and Stragier, P. (1996) Plasmids for ectopic integration in *Bacillus subtilis*. *Gene* **180**: 57-61.

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

Harwood, C.R., and Cutting, S.M. (1990) *Molecular biological methods for Bacillus.* Chichester; New York: Wiley.

Heinzel, T., Velleman, M., and Schuster, H. (1992) C1 repressor of phage P1 is inactivated by noncovalent binding of P1 Coi protein. *J Biol Chem* **267**: 4183-4188.

Jaacks, K.J., Healy, J., Losick, R., and Grossman, A.D. (1989) Identification and characterization of genes controlled by the sporulation-regulatory gene spo0H in Bacillus subtilis. *J Bacteriol* **171**: 4121-4129.

James, P., Halladay, J., and Craig, E.A. (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**: 1425-1436.

Krogh, S., O'Reilly, M., Nolan, N., and Devine, K.M. (1996) The phage-like element PBSX and part of the skin element, which are resident at different locations on the Bacillus subtilis chromosome, are highly homologous. *Microbiology* **142**: 2031-2040.

Kropinski, A.M., Hayward, M., Agnew, M.D., and Jarrell, K.F. (2005) The genome of BCJA1c: a bacteriophage active against the alkaliphilic bacterium, *Bacillus clarkii*. *Extremophiles* **9**: 99-109.

Lee, C.A., Auchtung, J.M., Monson, R.E., and Grossman, A.D. (2007) Identification and characterization of *int* (integrase), *xis* (excisionase), and chromosomal attachment sites of the integrative and conjugative element ICE*Bs1* of *Bacillus subtilis*. *Mol Microbiol* **accepted for publication**.

Liao, W., Sun, F., Song, S.Y., Shi, W., and Pang, Y. (2007) [Biology of two lysogenic phages from Bacillus thuringiensis MZ1]. *Wei Sheng Wu Xue Bao* **47**: 92-97.

Little, J.W. (1984) Autodigestion of lexA and phage lambda repressors. *Proc Natl Acad Sci U S A* **81**: 1375-1379.

Liu, T., Renberg, S.K., and Haggard-Ljungquist, E. (1998) The E protein of satellite phage P4 acts as an anti-repressor by binding to the C protein of helper phage P2. *Mol Microbiol* **30**: 1041-1050.

Loessner, M.J., Inman, R.B., Lauer, P., and Calendar, R. (2000) Complete nucleotide sequence, molecular analysis and genome structure of bacteriophage A118 of *Listeria monocytogenes*: implications for phage evolution. *Mol Microbiol* **35**: 324-340.

Love, P.E., and Yasbin, R.E. (1984) Genetic characterization of the inducible SOS-like system of Bacillus subtilis. *J Bacteriol* **160**: 910-920.

Lucchini, S., Desiere, F., and Brussow, H. (1999) Similarly organized lysogeny modules in temperate Siphoviridae from low GC content gram-positive bacteria. *Virology* **263**: 427-435.

Martin, A., and Schmid, F.X. (2003) Evolutionary stabilization of the gene-3-protein of phage fd reveals the principles that govern the thermodynamic stability of two-domain proteins. *J Mol Biol* **328**: 863-875.

McDonnell, G.E., Wood, H., Devine, K.M., and McConnell, D.J. (1994) Genetic control of bacterial suicide: regulation of the induction of PBSX in *Bacillus subtilis*. *J Bacteriol* **176**: 5820-5830.

McVeigh, R.R., and Yasbin, R.E. (1996) Phenotypic differentiation of "smart" versus "naive" bacteriophages of *Bacillus subtilis*. *J Bacteriol* **178**: 3399-3401.

Miller, M.C., Resnick, J.B., Smith, B.T., and Lovett, C.M., Jr. (1996) The Bacillus subtilis *dinR* gene codes for the analogue of *Escherichia coli* LexA. Purification and characterization of the DinR protein. *J Biol Chem* **271**: 33502-33508.

Miyoshi, S., and Shinoda, S. (2000) Microbial metalloproteases and pathogenesis. *Microbes Infect* **2**: 91-98.

Myers, G.S., Rasko, D.A., Cheung, J.K., Ravel, J., Seshadri, R., DeBoy, R.T., Ren, Q., Varga, J., Awad, M.M., Brinkac, L.M., Daugherty, S.C., Haft, D.H., Dodson, R.J., Madupu, R., Nelson, W.C., Rosovitz, M.J., Sullivan, S.A., Khouri, H., Dimitrov, G.I., Watkins, K.L., Mulligan, S., Benton, J., Radune, D., Fisher, D.J., Atkins, H.S., Hiscox, T., Jost, B.H., Billington, S.J., Songer, J.G., McClane, B.A., Titball, R.W., Rood, J.I., Melville, S.B., and Paulsen, I.T. (2006) Skewed genomic variability in strains of the toxigenic bacterial pathogen, Clostridium perfringens. *Genome Res* **16**: 1031-1040.

Obregon, V., Garcia, J.L., Garcia, E., Lopez, R., and Garcia, P. (2003) Genome organization and molecular analysis of the temperate bacteriophage MM1 of *Streptococcus pneumoniae*. *J Bacteriol* **185**: 2362-2368.

Paulsen, I.T., Banerjei, L., Myers, G.S., Nelson, K.E., Seshadri, R., Read, T.D., Fouts, D.E., Eisen, J.A., Gill, S.R., Heidelberg, J.F., Tettelin, H., Dodson, R.J., Umayam, L., Brinkac, L., Beanan, M., Daugherty, S., DeBoy, R.T., Durkin, S., Kolonay, J., Madupu, R., Nelson, W., Vamathevan, J., Tran, B., Upton, J., Hansen, T., Shetty, J., Khouri, H., Utterback, T., Radune, D., Ketchum, K.A., Dougherty, B.A., and Fraser, C.M. (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* **299**: 2071-2074.

Ptashne, M. (1992) *A Genetic Switch.* Cambridge, MA U.S.A: Blackwell Scientific Publications and Cell Press.

Rasko, D.A., Ravel, J., Okstad, O.A., Helgason, E., Cer, R.Z., Jiang, L., Shores, K.A., Fouts, D.E., Tourasse, N.J., Angiuoli, S.V., Kolonay, J., Nelson, W.C., Kolsto, A.B., Fraser, C.M., and Read, T.D. (2004) The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic adaptations and a large plasmid related to *Bacillus anthracis* pXO1. *Nucleic Acids Res* **32**: 977-988.

Rawlings, N.D., and Barrett, A.J. (1995) Evolutionary families of metallopeptidases. *Methods Enzymol* **248**: 183-228.

Read, T.D., Peterson, S.N., Tourasse, N., Baillie, L.W., Paulsen, I.T., Nelson, K.E., Tettelin, H., Fouts, D.E., Eisen, J.A., Gill, S.R., Holtzapple, E.K., Okstad, O.A., Helgason, E., Rilstone, J., Wu, M., Kolonay, J.F., Beanan, M.J., Dodson, R.J., Brinkac, L.M., Gwinn, M., DeBoy, R.T., Madpu, R., Daugherty, S.C., Durkin, A.S., Haft, D.H., Nelson, W.C., Peterson, J.D., Pop, M., Khouri, H.M., Radune, D., Benton, J.L., Mahamoud, Y., Jiang, L., Hance, I.R., Weidman, J.F., Berry, K.J., Plaut, R.D., Wolf, A.M., Watkins, K.L., Nierman, W.C., Hazen, A., Cline, R., Redmond, C., Thwaite, J.E., White, O., Salzberg, S.L., Thomason, B., Friedlander, A.M., Koehler, T.M., Hanna, P.C., Kolsto, A.B., and Fraser, C.M. (2003) The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. *Nature* **423**: 81-86.

Rice, L.B. (2002) Association of different mobile elements to generate novel integrative elements. *Cell Mol Life Sci* **59**: 2023-2032.

Roberts, J.W., and Roberts, C.W. (1975) Proteolytic cleavage of bacteriophage lambda repressor in induction. *Proc Natl Acad Sci U S A* **72**: 147-151.

Roberts, J.W., Roberts, C.W., and Craig, N.L. (1978) Escherichia coli recA gene product inactivates phage lambda repressor. *Proc Natl Acad Sci U S A* **75**: 4714-4718.

Rokop, M.E., Auchtung, J.M., and Grossman, A.D. (2004) Control of DNA replication initiation by recruitment of an essential initiation protein to the membrane of *Bacillus subtilis*. *Mol Microbiol* **52**: 1757-1767.

Sambrook, J., and Russell, D.W. (2001) *Molecular cloning: a laboratory manual.* Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Sauer, R.T., Ross, M.J., and Ptashne, M. (1982) Cleavage of the lambda and P22 repressors by recA protein. *J Biol Chem* **257**: 4458-4462.

Scott, K.P. (2002) The role of conjugative transposons in spreading antibiotic resistance between bacteria that inhabit the gastrointestinal tract. *Cell Mol Life Sci* **59**: 2071-2082.

Shearwin, K.E., Brumby, A.M., and Egan, J.B. (1998) The Tum protein of coliphage 186 is an antirepressor. *J Biol Chem* **273**: 5708-5715.

Slilaty, S.N., Rupley, J.A., and Little, J.W. (1986) Intramolecular cleavage of LexA and phage lambda repressors: dependence of kinetics on repressor concentration, pH, temperature, and solvent. *Biochemistry* **25**: 6866-6875.

Slilaty, S.N., and Little, J.W. (1987) Lysine-156 and serine-119 are required for LexA repressor cleavage: a possible mechanism. *Proc Natl Acad Sci U S A* **84**: 3987-3991.

Stanley, E., Fitzgerald, G.F., Le Marrec, C., Fayard, B., and van Sinderen, D. (1997) Sequence analysis and characterization of phi O1205, a temperate bacteriophage infecting *Streptococcus thermophilus* CNRZ1205. *Microbiology* **143**: 3417-3429.

Susskind, M.M., and Youderian, P. (1983) Bacteriophage P22 Antirepressor and Its Control. In *Lambda II.* Hendrix, R.W., Roberts, J.W., Stahl, F.W. and Weisberg, R.A. (eds). Cold Spring Harbor: Cold Spring Harbor Laboratory, pp. 347-363.

Tatusova, T.A., and Madden, T.L. (1999) BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol Lett* **174**: 247-250.

van der Meer, J.R., and Sentchilo, V. (2003) Genomic islands and the evolution of catabolic pathways in bacteria. *Curr Opin Biotechnol* **14**: 248-254.

Van Kaer, L., Van Montagu, M., and Dhaese, P. (1987) Transcriptional control in the EcoRI-F immunity region of Bacillus subtilis phage phi 105. Identification and unusual structure of the operator. *J Mol Biol* **197**: 55-67.

Van Kaer, L., Van Montagu, M., and Dhaese, P. (1989) Purification and in vitro DNA-binding specificity of the *Bacillus subtilis* phage phi 105 repressor. *J Biol Chem* **264**: 14784-14791.

Vasantha, N., and Freese, E. (1980) Enzyme changes during Bacillus subtilis sporulation caused by deprivation of guanine nucleotides. *J Bacteriol* **144**: 1119-1125.

Walker, G.C. (1996) The SOS Response of *Escherichia coli.* In *Escherichia coli and Salmonella: cellular and molecular biology.* Neidhardt, F.C., Curtiss III, R., Ingraham, J.L., Lin, E.C., Low, K.B., Magasanik, B., Reznikoff, W.S., Riley, M., Schaechter, M. and Umbarger, H.E. (eds). Washington, D.C.: American Society for Microbiology, pp. 1400-1416.

Whittle, G., Shoemaker, N.B., and Salyers, A.A. (2002) The role of *Bacteroides* conjugative transposons in the dissemination of antibiotic resistance genes. *Cell Mol Life Sci* **59**: 2044-2054.

Youngman, P., Poth, H., Green, B., York, K., Olmedo, G., and Smith, K. (1989) Methods for Genetic Manipulation, Cloning, and Functional Analysis of Sporulation Genes in *Bacillus subtilis*. In *Regulation of Procaryotic Development.* Smith, I., Slepecky, R.A. and Setlow, P. (eds). Washington, D.C.: ASM Press, pp. 65-87.

Appendix A: Proteolysis of ImmR by ImmA is metal-dependent

The ICE*Bs1* anti-repressor ImmA cleaves the repressor ImmR to allow ICE*Bs1* gene expression, excision, and mating. ImmA harbors an HEXXH motif that is conserved amongst many zinc-dependent metalloproteases (reviewed in Miyoshi and Shinoda, 2000; and Rawlings and Barrett, 1995). We mutated the first histidine in this motif to alanine to make ImmA(H75A), which is non-functional both in vivo and in vitro (Bose* *et al.*, 2008). In cells containing *immAH75A*, ICE*Bs1* gene expression was not derepressed by either MMC or RapI overproduction. In vitro, ImmA(H75A) did not cleave N- or C-terminally His-tagged ImmR under conditions where wild type ImmA did. These data indicate that the first residue in the putative Zinc-binding motif is important for ImmA function, and this supports the idea that ImmA is a metalloprotease.

To explore the metal-dependence of ImmA's proteolytic activity, we tested the effect of adding the metal-chelating agent EDTA to in vitro reactions with N-terminally His-tagged ImmR and untagged ImmA. Proteolysis of ImmR was assayed under conditions where an ImmR fragment was observed following incubation with ImmA, but ImmR was stable in the absence of ImmA or in the presence of ImmA(H75A) (Fig. 1). We found that EDTA prevented in vitro cleavage of ImmR, indicating that one or more metals are required for ImmA to proteolyze ImmR (Fig. 1). We tested whether adding a surplus of certain metals to a reaction mixture containing EDTA could restore proteolysis of ImmR. We found that adding excess $ZnCl_2$ or $MnCl_2$ allowed cleavage of ImmR, whereas $MgCl_2$ had no effect (Fig.1). Adding excess $MgSO_4$, $Ca(NO3)_2$, or $FeSO_4$ also had no effect (data not shown). Proteolysis of ImmR was more efficient in EDTA-treated reactions that were restored by addition of $ZnCl_2$ or $MnCl_2$, than in the reaction that was not treated with EDTA or metal. This probably indicates that the amount of Zn or Mn in the untreated reaction is limiting.

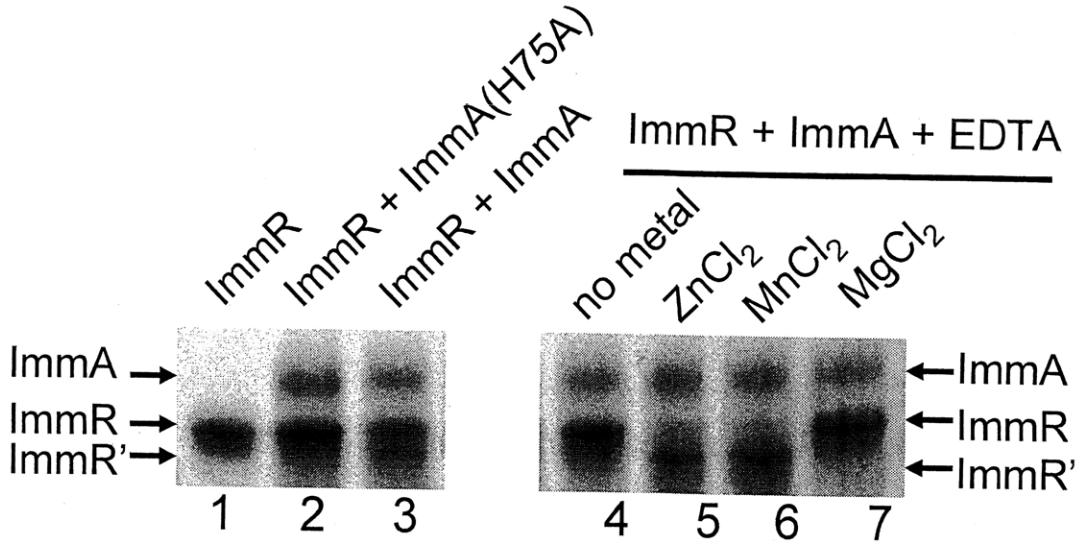


Figure 1. Effects of metals on the proteolysis of ImmR by ImmA. Cleavage of His6-ImmR by ImmA was assayed in vitro. Reactions were conducted under conditions where His6-ImmR was cleaved by wild type ImmA (lane 3), but was stable on its own (lane 1) or in the presence of ImmA(H75A) (lane 2). Adding the metal-chelating agent EDTA abolished proteolysis of His6-ImmR (lane 4). Adding excess ZnCl2 or MnCl2 in addition to EDTA restored proteolysis of His6-ImmR (lanes 5-6), but adding excess MgCl2 did not (lane 7). ImmA, ImmR, and the N-terminal fragment of ImmR (ImmR') are indicated.

These findings provide additional evidence that ImmA is a metal-dependent protease. The results are consistent with ImmA binding Zn or Mn to cleave ImmR. Further work is needed to learn which metal(s) are involved in ImmA-mediated proteolysis in vivo. In addition, for characterized Zinc-dependent metalloproteases, one or more residues assist the two histidines in the HEXXH motif in coordinating Zinc. Alignments of ImmA with the known classes of HEXXH-containing metalloproteases did not reveal which residue(s) in ImmA might additionally be required for metal coordination. Perhaps ImmA and its homologs comprise a new class of HEXXH-containing metalloproteases, with an as-yet uncharacterized metal-binding signature.

Experimental Procedures

Proteins were purified (His6-ImmR from BOSE798, ImmA from BOSE848, and ImmA(H75A) from BOSE847) as described previously (Bose* *et al.*, 2008). In vitro reactions were incubated for ~17 hours at 37°C. Reactions contained a final concentration of 20μM of each protein indicated. The final reaction buffer was 34.3mM Tris, 57.6mM NaCl, 4.11mM DTT, 21.3% glycerol, 41.1 μM $ZnCl_2$, 0.329M L-arginine, 0.247M guanidinium chloride, 41.1 μM EDTA, 2.05mM reduced glutathione, pH~7.5. When present, EDTA was added to each reaction mixture on ice to a concentration of 1.11mM. Five minutes after treatment with EDTA, water or metal-containing solution was added so that the final concentrations in each reaction of EDTA, $ZnCl_2$, $MnCl_2$, and $MgCl_2$ were 1.00mM, and reactions were transferred to 37°C. Reactions were also conducted in which addition of EDTA to 10mM was followed by addition of FeSO4 to 100 μM, $Ca(NO_3)_2$ to 100mM, or $MgSO_4$ to 100mM; no ImmR proteolysis was observed in these reactions (data not shown).

References

Bose*, B., Auchtung*, J.M., Lee, C.A., and Grossman, A.D. (2008) A conserved anti-repressor controls horizontal gene transfer by proteolysis. *Mol Microbiol* **70**: 570-582. * co-first authors.

Miyoshi, S., and Shinoda, S. (2000) Microbial metalloproteases and pathogenesis. *Microbes Infect* **2**: 91-98.

Rawlings, N.D., and Barrett, A.J. (1995) Evolutionary families of metallopeptidases. *Methods Enzymol* **248**: 183-228.

Appendix B: Pairs of proteins that are homologous to

ICE*Bs1* ImmA and ImmR are encoded in many other systems

Table1 in Chapter 2 lists homologs of ImmA and ImmR that were found during searches prior to March, 2008. Table 1 in this appendix lists proteins with homology to ImmA and ImmR that were found during a recent search, in January, 2010. I identified the proteins listed in Table 1 of this appendix by a combination of approaches.

I first conducted BLAST searches at the NCBI website to identify proteins that were similar to ImmA and ImmR (Altschul *et al.*, 1997). During these searches conserved domains within ImmA and ImmR were identified as well (Marchler-Bauer and Bryant, 2004; Marchler-Bauer *et al.*, 2009). ImmR contains an N-terminal helix-turn-helix domain (HTH XRE) that is characteristic of transcriptional repressors of bacteriophages. ImmA contains two conserved domains- a domain of unknown function (DUF955) and COG2856. DUF955-containing proteins are bacterial and viral proteins harboring an HEXXH motif that suggests these proteins may have some catalytic activity. COG2856 is found in predicted zinc peptidases that may be involved in amino acid transport and metabolism. Searching for COG2856-containing proteins at the NCBI website retrieved 361 listings; 4706 were found for DUF955.

For proteins with similarity to ImmA identified by BLAST searching, and for some of the proteins with either conserved domain in ImmA, I scanned the annotated genome sequences for adjacent or nearby, upstream proteins that contained a predicted HTH XRE domain, signifying some homology to ImmR. For ImmR-like proteins that were identified by BLAST search, I looked for a protein just downstream that contained either conserved domain found in ImmA. When a pair of proteins was found in this manner, I checked for indications that they were in a mobile genetic element. For pairs that were not in known mobile elements, those that are nearby known or predicted transposase, phage proteins, or conjugation proteins are listed as being in a putative mobile element (Table 1).

Table 1. Partial list of paired proteins with homology to ImmR and ImmA

Organism	Element	ImmR[1]		ImmA[2]		Reference[3]
		homolog	score	homolog	score	
Known mobile elements						
Bacillus clarkii	BCJA1c phage	Gp5	124	Gp4	115	(Kropinski *et al.*, 2005)
Bacillus licheniformis ATCC 14580 (DSM 13)	PBSX prophage	Xre	138	XkdA	16	(Veith *et al.*, 2004)
Bacillus subtilis	ø105 phage[4]	cø105	134	ImmA(ϕ105)	188	(McVeigh and Yasbin, 1996)
Bacillus subtilis	PBSX phage[4]	Xre	128	XkdA	26	(McDonnell *et al.*, 1994)
Bacillus subtilis	*skin* element[5]	YqaE	151	YqaB	43	(Krogh *et al.*, 1996; Mizuno *et al.*, 1996)
Bacillus weihenstephanensis KBAB4	phage-like element PBSX	BcerKBAB4_3618[7]	175	BcerKBAB4_3621[7]	n.s.	(Lapidus *et al.*, 2008)
Clostridium difficile	phage phiCD27	Gp72[8]	34	Gp72[8]	n.s.	(Mayer *et al.*, 2008)
Enterococcus faecalis	phage phiFL2A	Gp03	72	Gp02	39	(Yasmin *et al.*)
Enterococcus faecalis	phage phiFL4A	Gp04	152	Gp03	n.s.	(Yasmin *et al.*)
Geobacillus	GBSV1 phage	Gp48	175	Gp47	202	--
Geobacillus thermodenitrificans NG80-2	defective prophage	GTNG_2860	60	GTNG_2861	179	(Feng *et al.*, 2007)
Lactobacillus	phage phig1e	phig1ep05	159	phig1ep06	33	(Kodaira *et al.*, 1997)
Lactobacillus delbrueckii subsp. *lactis*	JCL1032 phage	Orf120	163	Orf131	23	(Riipinen *et al.*, 2007)
Lactobacillus gasseri	phage phiadh	Rad	55	ORF2	n.s.	(Altermann *et al.*, 1999; Engel *et al.*, 1998)
Lactobacillus plantarum JDM1	prophage Lp2	JDM1_0475	88	JDM1_0474	58	(Zhang *et al.*, 2009)

Organism	Element	ImmR[1]		ImmA[2]		Reference[3]
		homolog	score	homolog	score	
Lactococcus lactis	phage TP901-1	Orf4	41	Orf3	n.s.	(Brondsted *et al.*, 2001; Madsen and Hammer, 1998)
Lactococcus lactis IL1403	phage bIL285	Orf4	48	Orf3	n.s.	(Chopin *et al.*, 2001)
Listeria	B054 phage	Gp41	129	Gp40	123	(Dorscht *et al.*, 2009)
Staphylococcus aureus	phage 187	Gp37	94	Gp36	n.s.	(Kwan *et al.*, 2005)
Streptococcus dysgalactiae subsp. *equisimilis* NS3396	phage phi3396	phi3396_04	58	phi3396_03	8	(Davies *et al.*, 2007)
Streptococcus equi	phage P9	Gp4	127	Gp3	n.s.	--
Streptococcus mitis SF100	phage SM1[4]	Gp4	89	Gp3	33	(Bensing *et al.*, 2001)
Streptococcus oralis	phage P10	Gp4	93	Gp3	16	--
Streptococcus pneumoniae	MM1 phage[4]	CI	83	Orf2	n.s.	(Obregon *et al.*, 2003)
Streptococcus pyogenes MGAS315	phage 315.1	SpyM3_0684	85	SpyM3_0683	n.s.	(Beres *et al.*, 2002)
Streptococcus pyogenes MGAS315	phage 315.5	SpyM3_1350	n.s.	SpyM3_1351	n.s.	(Beres *et al.*, 2002)
Streptococcus pyogenes MGAS315	phage 315.6	SpyM3_1455	142	SpyM3_1456	n.s.	(Beres *et al.*, 2002)
Streptococcus pyogenes MGAS5005	phage 5005.2	M5005_Spy1219	78	M5005_Spy1220	n.s.	(Sumby *et al.*, 2005)
Streptococcus pyogenes MGAS9429	phage 9428.3	Spy1466	113	Spy1467	n.s.	(Beres *et al.*, 2006)
Streptococcus pyogenes MGAS10270	phage 10270.3	Spy1358	128	Spy1359	18	(Beres *et al.*, 2006)
Streptococcus thermophilus	ICE*St3*	Arp2[7]	203	OrfQ[7]	32	(Pavlovic *et al.*, 2004)
Streptococcus thermophilus	phage TP-J34	Orf121	78	Orf122	n.s.	(Neve *et al.*, 1998)

Organism	Element	ImmR[1]		ImmA[2]		Reference[3]
		homolog	score	homolog	score	
Streptococcus thermophilus	ϕO1205 phage[4]	Orf4	50	Orf3	n.s.	(Stanley *et al.*, 1997)
Streptococcus thermophilus CSK939	phage 5093	st5093phage_43	16	st5093phage_42	n.s.	--
Putative mobile elements[6]						
Actinobacillus pleuropneumoniae serovar 1 str. 4074	--	Aple02002129	115	Aple02002131	65	--
Aeromicrobium marinum DSM 15272	--	AmarD1_1200	6	AmarD1_1195	20	--
Alicyclobacillus acidocaldarius subsp. *acidocaldarius* DSM 446	--	Aaci_0465	162	Aaci_0464	68	--
Alicyclobacillus acidocaldarius subsp. *acidocaldarius* DSM 446	--	Aaci_1064	89	Aaci_1063	130	--
Alkaliphilus oremlandii OhILAs	--	Clos_1048	183	Clos_1047	22	--
Alkaliphilus oremlandii OhILAs	--	Clos_1286	96	Clos_1285	136	--
Ammonifex degensii KC4	--	Adeg_0697	193	Adeg_0696	61	--
Anaerocellum thermophilum DSM 6725	--	Athe_0094	50	Athe_0095	n.s.	(Kataeva *et al.*, 2009)
Anaerocellum thermophilum DSM 6725	--	Athe_2467	148	Athe_2466	8	(Kataeva *et al.*, 2009)
Anaerocellum thermophilum DSM 6725	--	Athe_2672	163	Athe_2671	31	(Kataeva *et al.*, 2009)
Anaerococcus prevotii DSM 20548	--	Apre_0223	77	Apre_0222	99	--
Bacillus anthracis str. Ames	λBa04	BA_3829	130	BA_3830	145	(Read *et al.*, 2003)
Bacillus cellulosilyticus DSM 2522	--	BcellDRAFT_2906[8]	85	BcellDRAFT_2906[8]	30	--
Bacillus cereus AH187	--	BCAH187_A0630	209	BCAH187_A0629	189	--
Bacillus cereus AH820	--	BCAH820_4414	309	BCAH820_4415	195	--
Bacillus cereus AH820	--	BCAH820_0555	39	BCAH820_0553	184	--
Bacillus cereus AH1273	--	bcere0030_10090	100	bcere0030_10080	142	--
Bacillus cereus ATCC 10876	--	bcere0002_36230	215	bcere0002_36240	205	--
Bacillus cereus BGSC 6E1	--	bcere0004_33280	184	bcere0004_33290	88	--
Bacillus cereus E33L	--	BCZK3459	n.s.	BCZK3460	145	(Han *et al.*, 2006)
Bacillus cereus G9842	--	B1349	195	B1348	226	--
Bacillus cereus MM3	--	bcere0006_55030	43	bcere0006_55020	190	--

Organism	Element	ImmR[1]		ImmA[2]		Reference[3]
		homolog	score	homolog	score	
Bacillus cereus m1293	--	bcere0001_51600	209	bcere0001_51610	183	--
Bacillus cereus NVH0597-99	--	BCO59799_3567	184	BCO59799_3568	158	--
Bacillus cereus Rock1-15	--	bcere0018_34160	130	bcere0018_34170	158	--
Bacillus cereus subsp. *cytotoxis* NVH 391-98	--	Bcer98_2969	167	Bcer98_2970	144	(Lapidus *et al.*, 2008)
Bacillus cereus 95/8201	--	bcere0016_4670	309	bcere0016_4660	133	--
Bacillus halodurans C-125	--	BH3549	56	BH3550	203	(Canchaya *et al.*, 2003)
Bacillus licheniformis ATCC 14580 (DSM 13)	phage NZP1	BL00664[7]	149	BL01374[7]	14	(Rey *et al.*, 2004)
Bacillus mycoides Rock1-4	--	bmyco0002_54560	21	bmyco0002_54550	152	--
Bacillus mycoides Rock3-17	--	bmyco0003_31350	206	bmyco0003_31360	174	--
Bacillus sp. B14905	--	BB14905_02315	88	BB14905_02320 (XkdA)	n.s.	--
Bacillus sp. B14905	--	BB14905_21323	129	BB14905_21318	23	--
Bacillus thuringiensis IBL 200	--	bthur0013_61880	176	bthur0013_61890	102	--
Bacillus thuringiensis IBL 4222	--	bthur0014_57190[8]	70	bthur0014_57190[8]	25	--
Bacillus thuringiensis IBL 4222	--	bthur0014_62300	220	bthur0014_62290	196	--
Bacillus thuringiensis serovar *andalousiensis* BGSC 4AW1	--	bthur0009_35660	225	bthur0009_35670	203	--
Bacillus thuringiensis serovar *berliner* ATCC 10792	--	bthur0008_35250	225	bthur0008_35260	213	--
Bacillus thuringiensis serovar *berliner* ATCC 10792	--	bthur0008_52890	112	bthur0008_52880	117	--
Bacillus thuringiensis serovar *berliner* ATCC 10792	--	bthur0008_58320[8]	19	bthur0008_58320[8]	25	--
Bacillus thuringiensis serovar *berliner* ATCC 10792	--	bthur0008_59200	181	bthur0008_59190	164	--
Bacillus thuringiensis serovar *huazhongensis* BGSC 4BD1	--	bthur0011_55960	180	bthur0011_55950	162	--
Bacillus thuringiensis serovar *huazhongensis* BGSC 4BD1	--	bthur0011_56080	331	bthur0011_56070	211	--
Bacillus thuringiensis serovar *huazhongensis* BGSC 4BD1	--	bthur0011_56740	113	bthur0011_56730	61	--

Organism	Element	ImmR[1]		ImmA[2]		Reference[3]
		homolog	score	homolog	score	
Bacillus thuringiensis serovar *monterrey* BGSC 4AJ1	--	bthur0007_58320	179	bthur0007_58330	150	--
Bacillus thuringiensis serovar *pakistani* str. T13001	--	bthur0005_62300	167	bthur0005_62290	59	--
Bacillus thuringiensis serovar *pulsiensis* BGSC 4CC1	--	bthur0012_55730	180	bthur0012_55720	149	--
Bacillus thuringiensis serovar *thuringiensis* str. T01001	--	bthur0003_54580	183	bthur0003_54570	164	--
Bacillus weihenstephanensis KBAB4	--	BcerKBAB4_3465	178	BcerKBAB4_3466	167	--
Bacillus weihenstephanensis KBAB4	--	BcerKBAB4_5811	164	BcerKBAB4_5810	146	--
Bacteroides capillosus ATCC 29799	--	BACCAP_02975	119	BACCAP_02976	99	--
Blautia hydrogenotrophica DSM 10507	--	RUMHYD_01734	38	RUMHYD_01733	171	--
Brachyspira murdochii DSM 12563	--	BmurDRAFT_27570[8]	14	BmurDRAFT_27570[8]	42	--
Bradyrhizobium sp. ORS278	--	BRADO6390[8]	44	BRADO6390[8]	n.s.	(Giraud *et al.*, 2007)
Brevundimonas sp. BAL3	--	BBAL3_2589[8]	26	BBAL3_2589[8]	7	--
Burkholderia ambifaria IOP40-10	--	BamIOP4010DRAFT_3742	35	BamIOP4010DRAFT_3743	43	--
Burkholderia ambifaria MC40-6	--	BamMC406_6747	129	BamMC406_6746	n.s.	--
Burkholderia pseudomallei BCC215	--	BpseBC_05323	72	BpseBC_05318	n.s.	--
Caldicellulosiruptor saccharolyticus DSM 8903	--	Csac_2536[8]	34	Csac_2536[8]	n.s.	(van de Werken *et al.*, 2008)
Caldicellulosiruptor saccharolyticus DSM 8903	--	Csac_2642	152	Csac_2641	20	(van de Werken *et al.*, 2008)
Candidatus Koribacter versatilis Ellin345	--	Acid345_0227	28	Acid345_0226	54	(Ward *et al.*, 2009)
Carnobacterium sp. AT7	--	CAT7_05074	51	CAT7_05079	47	--
Carnobacterium sp. AT7	--	CAT7_07613	106	CAT7_07618	34	--
Caulobacter sp. K31	--	Caul_1155	n.s.	Caul_1156	44	--
Cellvibrio japonicus Ueda107	--	CJA_1208	58	CJA_1207	48	(DeBoy *et al.*, 2008)
Clostridium bartlettii DSM 16795	--	CLOBAR_01007	131	CLOBAR_01008	68	--
Clostridium botulinum A2 str. Kyoto	--	CLM_2575	42	CLM_2576	234	--
Clostridium botulinum B str. Eklund 17B	--	CLL_A0920	110	CLL_A0919	71	--

Organism	Element	ImmR[1]		ImmA[2]		Reference[3]
		homolog	score	homolog	score	
Clostridium botulinum Ba4 str. 657	--	CLJ_B2578	95	CLJ_B2579	239	(Smith *et al.*, 2007)
Clostridium botulinum Ba4 str. 657	--	CLJ_B1361	138	CLJ_B1360	n.s.	(Smith *et al.*, 2007)
Clostridium botulinum Ba4 str. 657	--	CLJ_B1766	204	CLJ_B1765	202	(Smith *et al.*, 2007)
Clostridium botulinum Bf	--	CBB_0515	81	CBB_0514	124	--
Clostridium botulinum NCTC 2916	--	CBN_3015	62	CBN_3016	145	--
Clostridium cellulolyticum H10	--	Ccel_3076	94	Ccel_3077	n.s.	--
Clostridium hathewayi DSM 13479	--	ChatD1_10967	82	ChatD1_10972	145	--
Clostridium hathewayi DSM 13479	--	ChatD1_11689	58	ChatD1_11684	n.s.	--
Clostridium hathewayi DSM 13479	--	ChatD1_17589	52	ChatD1_17594	n.s.	--
Clostridium hiranonis DSM 13275	--	CLOHIR_01411	190	CLOHIR_01410	250	--
Clostridium kluyveri NBRC 12016	--	CKR_1693[8]	80	CKR_1693[8]	14	--
Clostridium nexile DSM 1787	--	CLONEX_01042	136	CLONEX_01041	88	--
Clostridium perfringens ATCC13124	--	CPF_1021	94	CPF_1020	195	(Myers *et al.*, 2006)
Clostridium perfringens CPE str. F4969	--	AC5_1039	91	AC5_1038	179	--
Clostridium phytofermentans ISDg	--	Cphy_0782	211	Cphy_0781	180	--
Clostridium sordellii	--	BAC57552.1	211	BAC57551.1	180	--
Clostridium sp. L2-50	--	CLOL250_01157	101	CLOL250_01158	143	--
Clostridium sp. M62/1	--	ClM62_09276	136	ClM62_09281	78	--
Clostridium sp. M62/1	--	ClM62_08456	159	ClM62_08451	101	--
Clostridium sp. SS2/1	--	CLOSS21_02459	158	CLOSS21_02458	175	--
Clostridium sporogenes ATCC 15579	--	CLOSPO_00711[8]	64	CLOSPO_00711[8]	n.s.	--
Clostridium sporogenes ATCC 15579	--	CLOSPO_01870	129	CLOSPO_01869	123	--
Cupriavidus taiwanensis pRALTA	--	pRALTA_0123	113	pRALTA_0124	22	(Amadou *et al.*, 2008)
Cyanothece sp. PCC 7424	--	PCC7424_1462	57	PCC7424_1463	n.s.	--
Desulfitobacterium hafniense DCB-2	--	Dhaf_4801	82	Dhaf_4800	n.s.	--
Desulfotomaculum acetoxidans DSM 771	--	Dtox_1454	74	Dtox_1455	24	--
Desulfotomaculum acetoxidans DSM 771	--	Dtox_2224	71	Dtox_2224	30	--

Organism	Element	ImmR[1]		ImmA[2]		Reference[3]
		homolog	score	homolog	score	
Desulfotomaculum acetoxidans DSM 771	--	Dtox_2248	72	Dtox_2247	51	--
Desulfotomaculum reducens MI-1	--	Dred_0020	78	Dred_0019	124	--
Dialister invisus DSM 15470	--	GCWU000321_01350	118	GCWU000321_01349	141	--
Diaphorobacter sp. TPSY	--	Dtpsy_1331[8]	n.s.	Dtpsy_1331[8]	n.s.	--
Diaphorobacter sp. TPSY	--	Dtpsy_3236	22	Dtpsy_3235	n.s.	--
Dickeya dadantii Ech703	--	Dd703_3835[8]	n.s.	Dd703_3835[8]	n.s.	--
Dictyoglomus turgidum DSM 6724	--	Dtur_1292	149	Dtur_1293	21	--
Dinoroseobacter shibae DFL 12	--	Dshi_0407	57	Dshi_0408	n.s.	(Wagner-Dobler *et al.*)
Dorea formicigenerans ATCC 27755	--	DORFOR_02707	96	DORFOR_02708	93	--
Enterococcus faecalis ATCC 4200	--	EFDG_01604	168	EFDG_01603	67	--
Enterococcus faecalis V583	--	EF2544 (45/68)	321	EF2545 (42/64)	310	(Paulsen *et al.*, 2003)
Enterococcus faecium Com15	--	EFWG_01456	86	EFWG_01457	84	--
Enterococcus faecium 1,230,933	--	EFPG_1537	96	EFPG_1536	159	--
Enterococcus faecium C68	--	EFXG_02685	116	EFXG_02686	139	--
Escherichia coli 83972	--	HMPREF0358_4676[8]	28	HMPREF0358_4676[8]	21	--
Eubacterium siraeum DSM 15702	--	EUBSIR_02070[8]	83	EUBSIR_02070[8]	n.s.	--
Faecalibacterium prausnitzii M21/2	--	FAEPRAM212_00671[8]	49	FAEPRAM212_00671[8]	15	--
Fibrobacter succinogenes subsp. *succinogenes* S85	--	Fisuc_2447	45	Fisuc_2446	n.s.	--
Fibrobacter succinogenes subsp. *succinogenes* S85	--	Fisuc_2499	44	Fisuc_2498	n.s.	--
Frankia sp. EAN1pec	--	Franean1_0062[8]	25	Franean1_0062[8]	9	--
Frankia sp. EAN1pec	--	Franean1_1588[8]	16	Franean1_1588[8]	10	--
Fusobacterium periodonticum ATCC 33693	--	FperA3_10171	123	FperA3_10176	62	--
Fusobacterium sp. D11	--	PrD11_07466	69	PrD11_07461	65	--
Fusobacterium sp. D12	--	FuD12_02584	56	FuD12_02589	93	--
Fusobacterium sp. 3_1_36A2	--	HMPREF0946_01108	59	HMPREF0946_01107	149	--
Fusobacterium sp. 3_1_5R	--	F3_00752	57	F3_00747	100	--

Organism	Element	ImmR[1]		ImmA[2]		Reference[3]
		homolog	score	homolog	score	
Fusobacterium sp. 3_1_5R	--	F3_08316[7]	136	F3_08306[7]	80	--
Geobacillus sp. G11MC16	--	G11MC16DRAFT_3519	n.s.	G11MC16DRAFT_3520	192	--
Geobacillus sp. Y412MC61	--	GYMC61_2168	51	GYMC61_2168	36	--
Geobacillus sp. Y412MC10	--	GYMC10_5405	56	GYMC10_5406	136	--
Haemophilus influenzae PittHH	--	CGSHiHH_00608	37	CGSHiHH_00613	33	--
Jonesia denitrificans DSM 20603	--	Jden_2129	n.s.	Jden_2128	n.s.	--
Jonesia denitrificans DSM 20603	--	Jden_2271	39	Jden_2270	55	--
Kosmotoga olearia TBF 19.5.1	--	Kole_0647[8]	104	Kole_0647[8]	17	--
Lactobacillus paracasei subsp. *paracasei* 8700:2	--	LBPG_00198	184	LBPG_00197	n.s.	--
Lactobacillus plantarum subsp. *plantarum* ATCC 14917	--	HMPREF0531_1071	141	HMPREF0531_1070	76	--
Lactobacillus reuteri DSM 20016	--	Lreu_1144	215	Lreu_1145	105	--
Lactobacillus reuteri 100-23	--	Lreu23DRAFT_3079	115	Lreu23DRAFT_3078	96	--
Lactobacillus rhamnosus LMS2-1	--	HMPREF0539_0397[8]	77	HMPREF0539_0397[8]	57	--
Lactobacillus salivarius UCC118	--	LSL_0742	107	LSL_0741	80	(Claesson *et al.*, 2006)
Lactobacillus salivarius UCC118	--	LSL_1835[8]	66	LSL_1835[8]	61	(Claesson *et al.*, 2006)
Lactobacillus ultunensis DSM 16047	--	HMPREF0548_1649[8]	90	HMPREF0548_1649[8]	42	--
Lactobacillus vaginalis ATCC 49540	--	HMPREF0549_0135	n.s.	HMPREF0549_0134	63	--
Listeria innocua CLIP 11262	--	Lin1762	171	Lin1763	126	(Glaser *et al.*, 2001)
Listeria innocua CLIP 11262	--	Lin1234	123	Lin1233	108	(Glaser *et al.*, 2001)
Listeria monocytogenes HPB2262	--	LmonH_01860	237	LmonH_01855	167	--
Listeria monocytogenes str. 1/2a F6854	--	LMOf6854_2699	244	LMOf6854_2700	169	(Nelson *et al.*, 2004)
Methanobrevibacter smithii ATCC 35061	--	Msm_1528	101	Msm_1529	40	(Samuel *et al.*, 2007)
Methanobrevibacter smithii DSM 2374	--	MsmiD23_020100005676	113	MsmiD23_020100005671	56	--

Organism	Element	ImmR[1]		ImmA[2]		Reference[3]
		homolog	score	homolog	score	
Micromonospora aurantiaca ATCC 27029	--	MicauDRAFT_2931	40	MicauDRAFT_2930	n.s.	--
Moorella thermoacetica ATCC 39073	--	Moth_0278	203	Moth_0277	n.s.	(Pierce *et al.*, 2008)
Nakamurella multipartita DSM 44233	--	Namu_2694[8]	12	Namu_2694[8]	n.s.	--
Nakamurella multipartita DSM 44233	--	Namu_4637	15	Namu_4638	29	--
Nitrobacter hamburgensis X14	--	Nham_3355[8]	36	Nham_3355[8]	25	--
Nitrobacter sp. Nb-311A	--	NB311A_05900	72	NB311A_05895	20	--
Nitrobacter sp. Nb-311A	--	NB311A_13716	63	NB311A_13711	n.s.	--
Nitrobacter sp. Nb-311A	--	NB311A_16529	45	NB311A_16534	n.s.	--
Nitrosomonas europaea ATCC 19718	--	NE2537	49	NE2538	21	(Chain *et al.*, 2003)
Nostoc punctiforme PCC 73102	--	Npun_F5445[8]	76	Npun_F5445[8]	57	--
Novosphingobium aromaticivorans DSM 12444	--	Saro_2204	56	Saro_2203	n.s.	--
Paenibacillus larvae subsp. *larvae* BRL-230010	--	Plarl_13349	118	Plarl_13344	119	--
Paenibacillus sp. oral taxon 786 str. D14	--	POTG_03770	199	POTG_03769	n.s.	--
Parvibaculum lavamentivorans DS-1	--	Plav_1824	39	Plav_1823	33	--
Pelodictyon phaeoclathratiforme BU-1	--	Ppha_2629[8]	71	Ppha_2629[8]	n.s.	--
Pelotomaculum thermopropionicum SI	--	PTH_0725[8]	77	PTH_0725[8]	n.s.	(Kosaka *et al.*, 2006; Kosaka *et al.*, 2008)
Polaromonas naphthalenivorans CJ2 pPNAP05	--	Pnap_4792[8]	63	Pnap_4792[8]	37	(Yagi *et al.*, 2009)
Pseudomonas putida GB-1	--	PputGB1_1625[8]	34	PputGB1_1625[8]	n.s.	--
Ralstonia pickettii 12D	--	Rpic12D_4677	23	Rpic12D_4678	n.s.	--
Rhizobium leguminosarum bv. *trifolii* WSM2304	--	Rleg2_6153[8]	84	Rleg2_6153[8]	n.s.	--
Rhodobacter sphaeroides 2.4.1	--	RSP_1622[8]	35	RSP_1622[8]	n.s.	--

Organism	Element	ImmR[1]		ImmA[2]		Reference[3]
		homolog	score	homolog	score	
Rhodobacterales bacterium HTCC2654	--	RB2654_02659	64	RB2654_02654	n.s.	--
Rhodopseudomonas palustris CGA009	--	RPA4129	110	RPA4130	17	(Larimer *et al.*, 2004)
Rhodopseudomonas palustris DX-1	--	Rpdx1DRAFT_0436	52	Rpdx1DRAFT_0435	n.s.	--
Rhodopseudomonas palustris DX-1	--	Rpdx1DRAFT_0653	50	Rpdx1DRAFT_0652	38	--
Roseovarius sp. TM1035	--	RTM1035_05255	144	RTM1035_05260	29	--
Ruminococcus flavefaciens FD-1	--	RflaF_16036	45	RflaF_16041	122	(Berg Miller *et al.*, 2009)
Ruminococcus obeum ATCC 29174	--	RUMOBE_00961	125	RUMOBE_00962	131	--
Ruminococcus torques ATCC 27756	--	RUMTOR_02056	116	RUMTOR_02055	167	--
Sebaldella termitidis ATCC 33386	--	Sterm_0822	22	Sterm_0823	150	--
Sebaldella termitidis ATCC 33386	--	Sterm_1389	53	Sterm_1388	164	--
Sebaldella termitidis ATCC 33386	--	Sterm_3926	113	Sterm_3927	88	--
Selenomonas noxia ATCC 43541	--	SnoxA4_08410	138	SnoxA4_08405	76	--
Shewanella sp. ANA-3	--	Shewana3_4356	57	Shewana3_4357	28	--
Sphaerobacter thermophilus DSM 20745	--	Sthe_2842[8]	42	Sthe_2842[8]	n.s.	--
Staphylococcus aureus subsp. *aureus* ED98	--	SAAV_2062	81	SAAV_2063	n.s.	--
Staphylococcus aureus subsp. *aureus* str. CF-Marseille	--	SauraC_12199	94	SauraC_12194	n.s.	--
Staphylococcus haemolyticus JSC1435	--	SH1805	103	SH1806	48	(Takeuchi *et al.*, 2005)
Stenotrophomonas maltophilia K279a	--	Smlt1906[8]	44	Smlt1906[8]	n.s.	(Crossman *et al.*, 2008)
Streptococcus equi subsp. *equi* 4047	--	SEQ_0790	121	SEQ_0789	n.s.	(Holden *et al.*, 2009)
Streptococcus equi subsp. *equi* 4047	--	SEQ_2086	78	SEQ_2087	n.s.	(Holden *et al.*, 2009)
Streptococcus pneumoniae CCRI 1974M2	--	SpneC19_04965[7]	71	SpneC19_04975[7]	n.s.	(Feng *et al.*, 2009)
Streptococcus pneumoniae CDC1873-00	--	SP187300_0114[7]	121	SP187300_0111[7]	n.s.	--

Organism	Element	ImmR[1]		ImmA[2]		Reference[3]
		homolog	score	homolog	score	
Streptococcus pneumoniae SP18-BS74	--	CGSSp18BS74_06382	71	CGSSp18BS74_06377	n.s.	--
Streptococcus pneumoniae SP19-BS75	--	CGSSp19BS75_00906	n.s.	CGSSp19BS75_00911	n.s.	--
Streptococcus pneumoniae SP23-BS72	--	CGSSp23BS72_00845	44	CGSSp23BS72_00840	n.s.	--
Streptococcus suis 89/1591	--	SsuiDRAFT_2808[8]	46	SsuiDRAFT_2808[8]	72	--
Streptococcus suis 89/1591	--	SsuiDRAFT_4039	84	SsuiDRAFT_4040	9	--
Streptomyces sp. AA4	--	StAA4_26329[8]	64	StAA4_26329[8]	n.s.	--
Subdoligranulum variabile DSM 15176	--	SUBVAR_04734	119	SUBVAR_04733	101	--
Thermoanaerobacterium thermosaccharolyticum DSM 571	--	TtheDRAFT_1054	60	TtheDRAFT_1053	45	--
Thermoanaerobacterium thermosaccharolyticum DSM 571	--	TtheDRAFT_2606	n.s.	TtheDRAFT_2605	n.s.	--
Thermoanaerobacter tencongensis MB4	--	TTE2125	122	TTE2126	108	(Bao *et al.*, 2002)
uncultured bacterium	--	ORFT45-17	163	ORFT45-16	101	(Kazimierczak *et al.*, 2008)
Verrucomicrobiae bacterium DG1235	--	VDG1235_3181[8]	22	VDG1235_3181[8]	n.s.	--
Vibrio mimicus VM223	--	VMA_000497[8]	61	VMA_000497[8]	n.s.	--
Victivallis vadensis ATCC BAA-548	--	Vvad_PD2341	n.s.	Vvad_PD2342	n.s.	--
Victivallis vadensis ATCC BAA-548	--	Vvad_PD3969	60	Vvad_PD3970	n.s.	--
Xanthobacter autotrophicus Py2	--	Xaut_0552	90	Xaut_0551	40	--
Xanthobacter autotrophicus Py2	--	Xaut_3976[8]	32	Xaut_3976[8]	n.s.	--
No apparent mobile element						
Acinetobacter sp. ATCC 27244	--	HMPREF0023_0553[8]	13	HMPREF0023_0553[8]	28	--
Actinosynnema mirum DSM 43827	--	Amir_3160[8]	18	Amir_3160[8]	n.s.	--
Anaerococcus tetradius ATCC 35098	--	HMPREF0077_0478[8]	72	HMPREF0077_0478[8]	n.s.	--
Anaerofustis stercorihominis DSM 17244	--	ANASTE_00175	41	ANASTE_00176	n.s.	--

Organism	Element	ImmR[1]		ImmA[2]		Reference[3]
		homolog	score	homolog	score	
Arthrospira maxima CS-328	--	AmaxDRAFT_1690	63	AmaxDRAFT_1691	n.s.	--
Bacillus sp. NRRL B-14911	--	B14911_01154	43	B14911_01159	10	--
Bacteroides intestinalis DSM 17393	--	BACINT_01474	104	BACINT_01473	67	--
Candidatus Accumulibacter phosphatis clade IIA str. UW-1	--	CAP2UW1_3866	68	CAP2UW1_3865	n.s.	--
Catenulispora acidiphila DSM 44928	--	Caci_6930	12	Caci_6931	10	--
Catenulispora acidiphila DSM 44928	--	Caci_7309[8]	57	Caci_7309[8]	16	--
Chlorobium phaeobacteroides BS1	--	Cphamn1_2492[8]	n.s.	Cphamn1_2492[8]	n.s.	--
Clostridium sporogenes ATCC 15579	--	CLOSPO_01445	74	CLOSPO_01444	26	--
Cyanothece sp. CCY0110	--	CY0110_24286	81	CY0110_24291	n.s.	--
Cyanothece sp. PCC 8802	--	Cyan8802_1489[8]	89	Cyan8802_1489[8]	22	--
Desulfatibacillum alkenivorans AK-01	--	Dalk_3082[8]	94	Dalk_3082[8]	15	--
Dictyoglomus turgidum DSM 6724	--	Dtur_0679	95	Dtur_0678	138	--
Enterococcus faecalis HH22	--	HMPREF0346_2373[8]	69	HMPREF0346_2373[8]	56	--
Enterococcus faecalis TX1322	--	HMPREF0349_0407[8]	77	HMPREF0349_0407[8]	n.s.	--
Enterococcus faecalis TX1322	--	HMPREF0349_2537[8]	69	HMPREF0349_2537[8]	56	--
Eubacterium eligens ATCC 27750	--	EUBELI_20262	48	EUBELI_20261	94	(Mahowald *et al.*, 2009)
Eubacterium siraeum DSM 15702	--	EUBSIR_02284	135	EUBSIR_02285	52	--
Frankia sp. EuI1c	--	FraEuI1cDRAFT_0226[8]	n.s.	FraEuI1cDRAFT_0226[8]	n.s.	--
Fusobacterium periodonticum ATCC 33693	--	FperA3_02249	58	FperA3_02254	49	--
Fusobacterium sp. D11	--	PrD11_12015	63	PrD11_12010	52	--
Hahella chejuensis KCTC 2396	--	HCH_06281	82	HCH_06282	30	(Jeong *et al.*, 2005)
Kribbella flavida DSM 17836	--	Kfla_4651[8]	28	Kfla_4651[8]	n.s.	--
Lactobacillus paracasei subsp. *paracasei* ATCC 25302	--	HMPREF0530_3024[8]	107	HMPREF0530_3024[8]	35	--
Lyngbya sp. PCC 8106	--	L8106_14995	81	L8106_14990	n.s.	--
marine gamma proteobacterium HTCC2080	--	MGP2080_02256	70	MGP2080_02261	37	--
Methylobacterium extorquens AM1	--	MexAM1_ META2p0468	44	MexAM1_ META2p0467	n.s.	(Vuilleumier *et al.*, 2009)

Organism	Element	ImmR[1]		ImmA[2]		Reference[3]
		homolog	score	homolog	score	
Paenibacillus larvae subsp. *larvae* BRL-230010	--	Plarl_20922	172	Plarl_20917	139	--
Pelobacter propionicus DSM 2379	--	Ppro_0104[8]	68	Ppro_0104[8]	n.s.	--
Pelodictyon phaeoclathratiforme BU-1	--	Ppha_2792[8]	n.s.	Ppha_2792[8]	n.s.	--
Pyramidobacter piscolens W5455	--	HMPREF7215_2390	56	HMPREF7215_2389	n.s.	--
Rhizobium leguminosarum bv. *trifolii* WSM1325	--	Rleg_5883[8]	30	Rleg_5883[8]	6	--
Rhodospirillum rubrum ATCC 11170	--	Rru_A0537[8]	n.s.	Rru_A0537[8]	n.s.	--
Ruminococcus obeum ATCC 29174	--	RUMOBE_02098	89	RUMOBE_02097	55	--
Staphylococcus aureus subsp. *aureus* MN8	--	HMPREF0769_0272[8]	76	HMPREF0769_0272[8]	66	--
Staphylococcus aureus subsp. *aureus* TCH60	--	HMPREF0772_0473[8]	76	HMPREF0772_0473[8]	65	--
Staphylococcus aureus subsp. *aureus* USA300_TCH959	--	HMPREF0776_0883[8]	67	HMPREF0776_0883[8]	66	--
Streptococcus pneumoniae TCH8431/19A	--	HMPREF0837_0961[8]	80	HMPREF0837_0961[8]	23	--
Syntrophobacter fumaroxidans MPOB	--	Sfum_1441[8]	14	Sfum_1441[8]	n.s.	--
Thermanaerovibrio acidaminovorans DSM 6589	--	Taci_1585	56	Taci_1586	n.s.	--
Veillonella parvula ATCC 17745	--	HMPREF1035_0096	90	HMPREF1035_0095	n.s.	--
Xanthobacter autotrophicus Py2	--	Xaut_4724[8]	36	Xaut_4724[8]	n.s.	--
Yersinia pseudotuberculosis PB1/+	--	YPTS_0261[8]	n.s.	YPTS_0261[8]	31	--

[1] All ImmR-like proteins contain a predicted phage repressor helix-turn-helix motif characteristic of phage repressors (HTH_XRE). The score for alignment of each protein with ImmR using LALIGN (global alignment, no end-gap penalty) is indicated. Higher scores indicate greater similarity with ImmR. A perfect score for ImmR (matched with itself) is 794. "n.s." indicates that the result for this alignment was not significant. Proteins that were not aligned well to ImmR by LALIGN were included, because of the presence of a neighboring ImmA-like protein.

[2] All ImmA-like proteins contain either of the conserved domains in ImmA- COG2856 or DUF955. The score for alignment of each protein with ImmR using LALIGN (global alignment, no end-gap penalty) is indicated. Higher scores indicate greater similarity with

ImmA. A perfect score for ImmA (matched with itself) is 1097. "n.s." indicates that the result for this alignment was not significant. Proteins that were not aligned well to ImmA by LALIGN were included, because of the presence of a neighboring ImmR-like protein.

[3] Indicated references include published sequence information and information about relevant mobile genetic elements . When no reference is indicated (--), data were retrieved from unpublished sequences deposited in NCBI.

[4] Element is damage-inducible.

[5] *skin* is a defective prophage.

[6] Putative mobile genetic elements were identified based on the presence of genes predicted to encode proteins homologous to those found in bacteriophage, transposons, or conjugative elements.

[7] The genes encoding the ImmA- and ImmR-like pair are separated by one or more orfs, according to the annotated sequences.

[8] The same protein contains an HTH XRE and either COG2856 or DUF955, indicating similarity to ImmR and ImmA.

For some ImmA-like proteins, there were multiple listings of products with identical sequences. Sometimes these were found in different organisms or mobile elements, and they may have ImmR-like partners that are not identical to each other. Because Table 1 is a partial list of homologs, only one of each set of identical ImmA-like and/or ImmR-like proteins is listed to avoid redundancy. Curiously, some proteins annotated as having a DUF955 domain don't have an HEXXH motif; where this was noticed, the protein was no long considered similar to ImmA.

The recent search revealed many more ImmR- and ImmA-homologs than were found previously. This is mostly due to the dramatic increase in the available sequence information during the time between the searches. However, some of the pairs identified during this search were found in sequences that were available prior to 2005. This is likely because investigating the proteins identified as having COG2856 or DUF955 reveals many more homologs than are found by BLAST searching alone. In addition, many of the newly identified proteins have both an HTH XRE domain and either COG2856 or DUF955. These may resemble fusions of ImmR and ImmA that were not documented during the previous search.

The similarity of each protein to ImmR or ImmA was gauged using LALIGN (hosted at ch.EMBnet.org) to do pairwise, global alignments between the homolog and the ICE*Bs1* protein (Huang and Miller, 1991). The program assigns a score to reflect the degree of similarity between two aligned proteins. Higher scores indicate greater similarity. Perfect matches for ImmR and ImmA generate scores of 794 and 1097, respectively. In some cases, LALIGN produced poor alignments between proteins. Usually, these appeared as end-to-end alignments, where a short stretch of amino acids at the C-terminal end of one protein was aligned with those at the N-terminal end of another. In some cases, the similarities between these proteins might be detected better by other alignment programs.

Some COG2856 and DUF955-containing proteins that didn't align well with ImmA in LALIGN were more usefully assessed in multiple sequence alignments from CLUSTALW (Larkin *et al.*, 2007). Several of these ImmA-like proteins from known mobile genetic elements are aligned with ImmA using CLUSTALW in Fig.1. In addition, many of the proteins with ImmR- and ImmA-like domains could not be aligned well with ImmA in LALIGN, but better alignments were produced using CLUSTALW (Fig. 2). This alignment highlights the conservation of several residues around the HEXXH motif in ImmA. The conserved glutamate residue indicated by an asterisk in Fig. 2 is a good candidate residue to test for its importance in metal coordination by ImmA. In some characterized HEXXH metalloproteases, a glutamate downstream of the conserved motif participates in metal-binding, along with the two conserved histidines (Rawlings and Barrett, 1995).

Along with further characterization of the metal-binding signature of ImmA, this homology search and future updates could inform characterization of the regulatory system of ICE*Bs1* in a variety of ways. Analysis of co-variation between these pairs of homologs could reveal which parts of the proteins are important for their interaction with each other. Proteins that resemble ImmR fused to ImmA could be useful tools for biochemical characterization of this system. Also, comparing the structure of ICE*Bs1* to that of other mobile elements with similar regulatory systems could shed light on the dissemination and evolution of these genes.

Figure 1. Alignment of ImmA with ImmA-like proteins from other mobile genetic elements. Sequences of some ImmA-like proteins from mobile genetic elements that did not align well with ImmA in LALIGN were aligned with ImmA using CLUSTALW. Asterisks indicate amino acid identity and dots indicate similarity.

```
ImmA                --ENLQHQQFVIAHELGHYFLHKRLN-TFKVVNCSKVLKDKLEHQASLFASELILTDKML 120
Amir_3160           IKRDYYRQRFDVAHELGHLVMHGD---AEPGG-------RIVEEQAHRFAAELLTPADQV 256
Ppro_0104           LGKTASRARFDAAHELGHLVLHEE---HTPGC-------SLTENEADRFAGAFLAPRDSF 242
HMPREF0077_0478     GKKGKARINFDMAHELGHILLHPW---TEDIETLSNEEFKERERQANKFASSFLLPRDTF 253
phiCD27_Gp72        NKTSASRIHFDIAHELGHICLHEW---SEDVEALEKQEFKDRESEANRFASAFLLPEETF 251
BcellDRAFT_2906     DSGSFYRQQFSIAHEIGHWVLHQS---INPQE-LDKDEYKTMEDEANKLASIFLLPKESF 246
BBAL3_2589          LKESLCRQRFDFAHECGHLIMHRG---LQTGD-------RATEDQAHRFASAFMFPRGAV 240
FraEul1c_0226       ATGNWFWENWSLAHELGHLVLGHH---DEPIT---EAERDPREAAANGFAADLLLPKDAM 229
Ppha_2792           KQASDSRQTFTLMHELAHLLLHKS---SSIDDNFDLDSSHGRERDANSFAGHLLVPDNFL 259
BRAD06390           GKDRSNGRVFTLLHEFCHLALRQSGV-SNMGGDRNDAPHPDVEKFCNAVAAAALMPRDWL 261
Rru_A0537           DRDARAARSFTLLHELAHLWLGQTGVSGAVETAEISSRVGVIERFCNDVAGEFLLPSAAF 281
Cyan8802_1489       YLDDLYRQNFSVAHEYCHALFDSS---LEQEMTYFKDNNNNREWRANSFAGCFLVPENAI 253
                            :   **  *  :                    *  .  .*.  : .   .
```

Figure 2. Alignment of ImmA with single proteins, each of which harbors domains found in ImmR and ImmA. Part of the alignment is shown for ImmA and several proteins that harbored HTH XRE domains and either COG2856 or DUF955. This alignment was generated by CLUSTALW; each of these proteins could not be aligned with ImmA using LALIGN. Asterisks indicate amino acid identity and dots indicate similarity.

References

Altermann, E., Klein, J.R., and Henrich, B. (1999) Primary structure and features of the genome of the Lactobacillus gasseri temperate bacteriophage (phi)adh. *Gene* **236**: 333-346.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-3402.

Amadou, C., Pascal, G., Mangenot, S., Glew, M., Bontemps, C., Capela, D., Carrere, S., Cruveiller, S., Dossat, C., Lajus, A., Marchetti, M., Poinsot, V., Rouy, Z., Servin, B., Saad, M., Schenowitz, C., Barbe, V., Batut, J., Medigue, C., and Masson-Boivin, C. (2008) Genome sequence of the beta-rhizobium Cupriavidus taiwanensis and comparative genomics of rhizobia. *Genome Res* **18**: 1472-1483.

Bao, Q., Tian, Y., Li, W., Xu, Z., Xuan, Z., Hu, S., Dong, W., Yang, J., Chen, Y., Xue, Y., Xu, Y., Lai, X., Huang, L., Dong, X., Ma, Y., Ling, L., Tan, H., Chen, R., Wang, J., Yu, J., and Yang, H. (2002) A complete sequence of the *T. tengcongensis* genome. *Genome Res* **12**: 689-700.

Bensing, B.A., Siboo, I.R., and Sullam, P.M. (2001) Proteins PblA and PblB of Streptococcus mitis, which promote binding to human platelets, are encoded within a lysogenic bacteriophage. *Infect Immun* **69**: 6186-6192.

Beres, S.B., Sylva, G.L., Barbian, K.D., Lei, B., Hoff, J.S., Mammarella, N.D., Liu, M.Y., Smoot, J.C., Porcella, S.F., Parkins, L.D., Campbell, D.S., Smith, T.M., McCormick, J.K., Leung, D.Y., Schlievert, P.M., and Musser, J.M. (2002) Genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence. *Proc Natl Acad Sci U S A* **99**: 10078-10083.

Beres, S.B., Richter, E.W., Nagiec, M.J., Sumby, P., Porcella, S.F., DeLeo, F.R., and Musser, J.M. (2006) Molecular genetic anatomy of inter- and intraserotype variation in the human bacterial pathogen group A Streptococcus. *Proc Natl Acad Sci U S A* **103**: 7059-7064.

Berg Miller, M.E., Antonopoulos, D.A., Rincon, M.T., Band, M., Bari, A., Akraiko, T., Hernandez, A., Thimmapuram, J., Henrissat, B., Coutinho, P.M., Borovok, I., Jindou, S., Lamed, R., Flint, H.J., Bayer, E.A., and White, B.A. (2009) Diversity and strain specificity of plant cell wall degrading enzymes revealed by the draft genome of Ruminococcus flavefaciens FD-1. *PLoS One* **4**: e6650.

Brondsted, L., Ostergaard, S., Pedersen, M., Hammer, K., and Vogensen, F.K. (2001) Analysis of the complete DNA sequence of the temperate bacteriophage TP901-1: evolution, structure, and genome organization of lactococcal bacteriophages. *Virology* **283**: 93-109.

Canchaya, C., Proux, C., Fournous, G., Bruttin, A., and Brussow, H. (2003) Prophage genomics. *Microbiol Mol Biol Rev* **67**: 238-276.

Chain, P., Lamerdin, J., Larimer, F., Regala, W., Lao, V., Land, M., Hauser, L., Hooper, A., Klotz, M., Norton, J., Sayavedra-Soto, L., Arciero, D., Hommes, N., Whittaker, M., and Arp, D. (2003) Complete genome sequence of the ammonia-oxidizing bacterium and obligate chemolithoautotroph Nitrosomonas europaea. *J Bacteriol* **185**: 2759-2773.

Chopin, A., Bolotin, A., Sorokin, A., Ehrlich, S.D., and Chopin, M. (2001) Analysis of six prophages in Lactococcus lactis IL1403: different genetic structure of temperate and virulent phage populations. *Nucleic Acids Res* **29**: 644-651.

Claesson, M.J., Li, Y., Leahy, S., Canchaya, C., van Pijkeren, J.P., Cerdeno-Tarraga, A.M., Parkhill, J., Flynn, S., O'Sullivan, G.C., Collins, J.K., Higgins, D., Shanahan, F.,

Fitzgerald, G.F., van Sinderen, D., and O'Toole, P.W. (2006) Multireplicon genome architecture of Lactobacillus salivarius. *Proc Natl Acad Sci U S A* **103**: 6718-6723.

Crossman, L.C., Gould, V.C., Dow, J.M., Vernikos, G.S., Okazaki, A., Sebaihia, M., Saunders, D., Arrowsmith, C., Carver, T., Peters, N., Adlem, E., Kerhornou, A., Lord, A., Murphy, L., Seeger, K., Squares, R., Rutter, S., Quail, M.A., Rajandream, M.A., Harris, D., Churcher, C., Bentley, S.D., Parkhill, J., Thomson, N.R., and Avison, M.B. (2008) The complete genome, comparative and functional analysis of Stenotrophomonas maltophilia reveals an organism heavily shielded by drug resistance determinants. *Genome Biol* **9**: R74.

Davies, M.R., McMillan, D.J., Van Domselaar, G.H., Jones, M.K., and Sriprakash, K.S. (2007) Phage 3396 from a Streptococcus dysgalactiae subsp. equisimilis pathovar may have its origins in streptococcus pyogenes. *J Bacteriol* **189**: 2646-2652.

DeBoy, R.T., Mongodin, E.F., Fouts, D.E., Tailford, L.E., Khouri, H., Emerson, J.B., Mohamoud, Y., Watkins, K., Henrissat, B., Gilbert, H.J., and Nelson, K.E. (2008) Insights into plant cell wall degradation from the genome sequence of the soil bacterium Cellvibrio japonicus. *J Bacteriol* **190**: 5455-5463.

Dorscht, J., Klumpp, J., Bielmann, R., Schmelcher, M., Born, Y., Zimmer, M., Calendar, R., and Loessner, M.J. (2009) Comparative genome analysis of Listeria bacteriophages reveals extensive mosaicism, programmed translational frameshifting, and a novel prophage insertion site. *J Bacteriol* **191**: 7206-7215.

Engel, G., Altermann, E., Klein, J.R., and Henrich, B. (1998) Structure of a genome region of the Lactobacillus gasseri temperate phage phiadh covering a repressor gene and cognate promoters. *Gene* **210**: 61-70.

Feng, J., Lupien, A., Gingras, H., Wasserscheid, J., Dewar, K., Legare, D., and Ouellette, M. (2009) Genome sequencing of linezolid-resistant Streptococcus pneumoniae mutants reveals novel mechanisms of resistance. *Genome Res* **19**: 1214-1223.

Feng, L., Wang, W., Cheng, J., Ren, Y., Zhao, G., Gao, C., Tang, Y., Liu, X., Han, W., Peng, X., Liu, R., and Wang, L. (2007) Genome and proteome of long-chain alkane degrading Geobacillus thermodenitrificans NG80-2 isolated from a deep-subsurface oil reservoir. *Proc Natl Acad Sci U S A* **104**: 5602-5607.

Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.C., Jaubert, M., Simon, D., Cartieaux, F., Prin, Y., Bena, G., Hannibal, L., Fardoux, J., Kojadinovic, M., Vuillet, L., Lajus, A., Cruveiller, S., Rouy, Z., Mangenot, S., Segurens, B., Dossat, C., Franck, W.L., Chang, W.S., Saunders, E., Bruce, D., Richardson, P., Normand, P., Dreyfus, B., Pignol, D., Stacey, G., Emerich, D., Vermeglio, A., Medigue, C., and Sadowsky, M. (2007) Legumes symbioses: absence of Nod genes in photosynthetic bradyrhizobia. *Science* **316**: 1307-1312.

Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couve, E., de Daruvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K.D., Fsihi, H., Garcia-del Portillo, F., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L.M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vicente, J.M., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., de Pablos, B., Perez-Diaz, J.C., Purcell, R., Remmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vazquez-

Boland, J.A., Voss, H., Wehland, J., and Cossart, P. (2001) Comparative genomics of *Listeria* species. *Science* **294**: 849-852.

Han, C.S., Xie, G., Challacombe, J.F., Altherr, M.R., Bhotika, S.S., Brown, N., Bruce, D., Campbell, C.S., Campbell, M.L., Chen, J., Chertkov, O., Cleland, C., Dimitrijevic, M., Doggett, N.A., Fawcett, J.J., Glavina, T., Goodwin, L.A., Green, L.D., Hill, K.K., Hitchcock, P., Jackson, P.J., Keim, P., Kewalramani, A.R., Longmire, J., Lucas, S., Malfatti, S., McMurry, K., Meincke, L.J., Misra, M., Moseman, B.L., Mundt, M., Munk, A.C., Okinaka, R.T., Parson-Quintana, B., Reilly, L.P., Richardson, P., Robinson, D.L., Rubin, E., Saunders, E., Tapia, R., Tesmer, J.G., Thayer, N., Thompson, L.S., Tice, H., Ticknor, L.O., Wills, P.L., Brettin, T.S., and Gilna, P. (2006) Pathogenomic sequence analysis of Bacillus cereus and Bacillus thuringiensis isolates closely related to Bacillus anthracis. *J Bacteriol* **188**: 3382-3390.

Holden, M.T., Heather, Z., Paillot, R., Steward, K.F., Webb, K., Ainslie, F., Jourdan, T., Bason, N.C., Holroyd, N.E., Mungall, K., Quail, M.A., Sanders, M., Simmonds, M., Willey, D., Brooks, K., Aanensen, D.M., Spratt, B.G., Jolley, K.A., Maiden, M.C., Kehoe, M., Chanter, N., Bentley, S.D., Robinson, C., Maskell, D.J., Parkhill, J., and Waller, A.S. (2009) Genomic evidence for the evolution of Streptococcus equi: host restriction, increased virulence, and genetic exchange with human pathogens. *PLoS Pathog* **5**: e1000346.

Huang, X., and Miller, W. (1991) A time-efficient, linear-space local similarity algorithm. *Advances in Applied Mathematics* **12**: 337-357.

Jeong, H., Yim, J.H., Lee, C., Choi, S.H., Park, Y.K., Yoon, S.H., Hur, C.G., Kang, H.Y., Kim, D., Lee, H.H., Park, K.H., Park, S.H., Park, H.S., Lee, H.K., Oh, T.K., and Kim, J.F. (2005) Genomic blueprint of Hahella chejuensis, a marine microbe producing an algicidal agent. *Nucleic Acids Res* **33**: 7066-7073.

Kataeva, I.A., Yang, S.J., Dam, P., Poole, F.L., 2nd, Yin, Y., Zhou, F., Chou, W.C., Xu, Y., Goodwin, L., Sims, D.R., Detter, J.C., Hauser, L.J., Westpheling, J., and Adams, M.W. (2009) Genome sequence of the anaerobic, thermophilic, and cellulolytic bacterium "Anaerocellum thermophilum" DSM 6725. *J Bacteriol* **191**: 3760-3761.

Kazimierczak, K.A., Rincon, M.T., Patterson, A.J., Martin, J.C., Young, P., Flint, H.J., and Scott, K.P. (2008) A new tetracycline efflux gene, tet(40), is located in tandem with tet(O/32/O) in a human gut firmicute bacterium and in metagenomic library clones. *Antimicrob Agents Chemother* **52**: 4001-4009.

Kodaira, K.I., Oki, M., Kakikawa, M., Watanabe, N., Hirakawa, M., Yamada, K., and Taketo, A. (1997) Genome structure of the Lactobacillus temperate phage phi g1e: the whole genome sequence and the putative promoter/repressor system. *Gene* **187**: 45-53.

Kosaka, T., Uchiyama, T., Ishii, S., Enoki, M., Imachi, H., Kamagata, Y., Ohashi, A., Harada, H., Ikenaga, H., and Watanabe, K. (2006) Reconstruction and regulation of the central catabolic pathway in the thermophilic propionate-oxidizing syntroph Pelotomaculum thermopropionicum. *J Bacteriol* **188**: 202-210.

Kosaka, T., Kato, S., Shimoyama, T., Ishii, S., Abe, T., and Watanabe, K. (2008) The genome of Pelotomaculum thermopropionicum reveals niche-associated evolution in anaerobic microbiota. *Genome Res* **18**: 442-448.

Krogh, S., O'Reilly, M., Nolan, N., and Devine, K.M. (1996) The phage-like element PBSX and part of the skin element, which are resident at different locations on the Bacillus subtilis chromosome, are highly homologous. *Microbiology* **142**: 2031-2040.

Kropinski, A.M., Hayward, M., Agnew, M.D., and Jarrell, K.F. (2005) The genome of BCJA1c: a bacteriophage active against the alkaliphilic bacterium, *Bacillus clarkii*. *Extremophiles* **9**: 99-109.

Kwan, T., Liu, J., DuBow, M., Gros, P., and Pelletier, J. (2005) The complete genomes and proteomes of 27 Staphylococcus aureus bacteriophages. *Proc Natl Acad Sci U S A* **102**: 5174-5179.

Lapidus, A., Goltsman, E., Auger, S., Galleron, N., Segurens, B., Dossat, C., Land, M.L., Broussolle, V., Brillard, J., Guinebretiere, M.H., Sanchis, V., Nguen-The, C., Lereclus, D., Richardson, P., Wincker, P., Weissenbach, J., Ehrlich, S.D., and Sorokin, A. (2008) Extending the Bacillus cereus group genomics to putative food-borne pathogens of different toxicity. *Chem Biol Interact* **171**: 236-249.

Larimer, F.W., Chain, P., Hauser, L., Lamerdin, J., Malfatti, S., Do, L., Land, M.L., Pelletier, D.A., Beatty, J.T., Lang, A.S., Tabita, F.R., Gibson, J.L., Hanson, T.E., Bobst, C., Torres, J.L., Peres, C., Harrison, F.H., Gibson, J., and Harwood, C.S. (2004) Complete genome sequence of the metabolically versatile photosynthetic bacterium Rhodopseudomonas palustris. *Nat Biotechnol* **22**: 55-61.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., and Higgins, D.G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947-2948.

Madsen, P.L., and Hammer, K. (1998) Temporal transcription of the lactococcal temperate phage TP901-1 and DNA sequence of the early promoter region. *Microbiology* **144 (Pt 8)**: 2203-2215.

Mahowald, M.A., Rey, F.E., Seedorf, H., Turnbaugh, P.J., Fulton, R.S., Wollam, A., Shah, N., Wang, C., Magrini, V., Wilson, R.K., Cantarel, B.L., Coutinho, P.M., Henrissat, B., Crock, L.W., Russell, A., Verberkmoes, N.C., Hettich, R.L., and Gordon, J.I. (2009) Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. *Proc Natl Acad Sci U S A* **106**: 5859-5864.

Marchler-Bauer, A., and Bryant, S.H. (2004) CD-Search: protein domain annotations on the fly. *Nucleic Acids Res* **32**: W327-331.

Marchler-Bauer, A., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Lu, S., Marchler, G.H., Mullokandov, M., Song, J.S., Tasneem, A., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., and Bryant, S.H. (2009) CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res* **37**: D205-210.

Mayer, M.J., Narbad, A., and Gasson, M.J. (2008) Molecular characterization of a Clostridium difficile bacteriophage and its cloned biologically active endolysin. *J Bacteriol* **190**: 6734-6740.

McDonnell, G.E., Wood, H., Devine, K.M., and McConnell, D.J. (1994) Genetic control of bacterial suicide: regulation of the induction of PBSX in *Bacillus subtilis*. *J Bacteriol* **176**: 5820-5830.

McVeigh, R.R., and Yasbin, R.E. (1996) Phenotypic differentiation of "smart" versus "naive" bacteriophages of *Bacillus subtilis*. *J Bacteriol* **178**: 3399-3401.

Mizuno, M., Masuda, S., Takemaru, K., Hosono, S., Sato, T., Takeuchi, M., and Kobayashi, Y. (1996) Systematic sequencing of the 283 kb 210 degrees-232 degrees region of the

Bacillus subtilis genome containing the skin element and many sporulation genes. *Microbiology* **142 (Pt 11)**: 3103-3111.

Myers, G.S., Rasko, D.A., Cheung, J.K., Ravel, J., Seshadri, R., DeBoy, R.T., Ren, Q., Varga, J., Awad, M.M., Brinkac, L.M., Daugherty, S.C., Haft, D.H., Dodson, R.J., Madupu, R., Nelson, W.C., Rosovitz, M.J., Sullivan, S.A., Khouri, H., Dimitrov, G.I., Watkins, K.L., Mulligan, S., Benton, J., Radune, D., Fisher, D.J., Atkins, H.S., Hiscox, T., Jost, B.H., Billington, S.J., Songer, J.G., McClane, B.A., Titball, R.W., Rood, J.I., Melville, S.B., and Paulsen, I.T. (2006) Skewed genomic variability in strains of the toxigenic bacterial pathogen, Clostridium perfringens. *Genome Res* **16**: 1031-1040.

Nelson, K.E., Fouts, D.E., Mongodin, E.F., Ravel, J., DeBoy, R.T., Kolonay, J.F., Rasko, D.A., Angiuoli, S.V., Gill, S.R., Paulsen, I.T., Peterson, J., White, O., Nelson, W.C., Nierman, W., Beanan, M.J., Brinkac, L.M., Daugherty, S.C., Dodson, R.J., Durkin, A.S., Madupu, R., Haft, D.H., Selengut, J., Van Aken, S., Khouri, H., Fedorova, N., Forberger, H., Tran, B., Kathariou, S., Wonderling, L.D., Uhlich, G.A., Bayles, D.O., Luchansky, J.B., and Fraser, C.M. (2004) Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen Listeria monocytogenes reveal new insights into the core genome components of this species. *Nucleic Acids Res* **32**: 2386-2395.

Neve, H., Zenz, K.I., Desiere, F., Koch, A., Heller, K.J., and Brussow, H. (1998) Comparison of the lysogeny modules from the temperate Streptococcus thermophilus bacteriophages TP-J34 and Sfi21: implications for the modular theory of phage evolution. *Virology* **241**: 61-72.

Obregon, V., Garcia, J.L., Garcia, E., Lopez, R., and Garcia, P. (2003) Genome organization and molecular analysis of the temperate bacteriophage MM1 of *Streptococcus pneumoniae*. *J Bacteriol* **185**: 2362-2368.

Paulsen, I.T., Banerjei, L., Myers, G.S., Nelson, K.E., Seshadri, R., Read, T.D., Fouts, D.E., Eisen, J.A., Gill, S.R., Heidelberg, J.F., Tettelin, H., Dodson, R.J., Umayam, L., Brinkac, L., Beanan, M., Daugherty, S., DeBoy, R.T., Durkin, S., Kolonay, J., Madupu, R., Nelson, W., Vamathevan, J., Tran, B., Upton, J., Hansen, T., Shetty, J., Khouri, H., Utterback, T., Radune, D., Ketchum, K.A., Dougherty, B.A., and Fraser, C.M. (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* **299**: 2071-2074.

Pavlovic, G., Burrus, V., Gintz, B., Decaris, B., and Guedon, G. (2004) Evolution of genomic islands by deletion and tandem accretion by site-specific recombination: ICESt1-related elements from Streptococcus thermophilus. *Microbiology* **150**: 759-774.

Pierce, E., Xie, G., Barabote, R.D., Saunders, E., Han, C.S., Detter, J.C., Richardson, P., Brettin, T.S., Das, A., Ljungdahl, L.G., and Ragsdale, S.W. (2008) The complete genome sequence of Moorella thermoacetica (f. Clostridium thermoaceticum). *Environ Microbiol* **10**: 2550-2573.

Rawlings, N.D., and Barrett, A.J. (1995) Evolutionary families of metallopeptidases. *Methods Enzymol* **248**: 183-228.

Read, T.D., Peterson, S.N., Tourasse, N., Baillie, L.W., Paulsen, I.T., Nelson, K.E., Tettelin, H., Fouts, D.E., Eisen, J.A., Gill, S.R., Holtzapple, E.K., Okstad, O.A., Helgason, E., Rilstone, J., Wu, M., Kolonay, J.F., Beanan, M.J., Dodson, R.J., Brinkac, L.M., Gwinn, M., DeBoy, R.T., Madpu, R., Daugherty, S.C., Durkin, A.S., Haft, D.H., Nelson, W.C., Peterson, J.D., Pop, M., Khouri, H.M., Radune, D., Benton, J.L., Mahamoud, Y., Jiang, L., Hance, I.R., Weidman, J.F., Berry, K.J., Plaut, R.D., Wolf, A.M., Watkins, K.L.,

Nierman, W.C., Hazen, A., Cline, R., Redmond, C., Thwaite, J.E., White, O., Salzberg, S.L., Thomason, B., Friedlander, A.M., Koehler, T.M., Hanna, P.C., Kolsto, A.B., and Fraser, C.M. (2003) The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. *Nature* **423**: 81-86.

Rey, M.W., Ramaiya, P., Nelson, B.A., Brody-Karpin, S.D., Zaretsky, E.J., Tang, M., Lopez de Leon, A., Xiang, H., Gusti, V., Clausen, I.G., Olsen, P.B., Rasmussen, M.D., Andersen, J.T., Jorgensen, P.L., Larsen, T.S., Sorokin, A., Bolotin, A., Lapidus, A., Galleron, N., Ehrlich, S.D., and Berka, R.M. (2004) Complete genome sequence of the industrial bacterium Bacillus licheniformis and comparisons with closely related Bacillus species. *Genome Biol* **5**: R77.

Riipinen, K.A., Raisanen, L., and Alatossava, T. (2007) Integration of the group c phage JCL1032 of Lactobacillus delbrueckii subsp. lactis and complex phage resistance of the host. *J Appl Microbiol* **103**: 2465-2475.

Samuel, B.S., Hansen, E.E., Manchester, J.K., Coutinho, P.M., Henrissat, B., Fulton, R., Latreille, P., Kim, K., Wilson, R.K., and Gordon, J.I. (2007) Genomic and metabolic adaptations of Methanobrevibacter smithii to the human gut. *Proc Natl Acad Sci U S A* **104**: 10643-10648.

Smith, T.J., Hill, K.K., Foley, B.T., Detter, J.C., Munk, A.C., Bruce, D.C., Doggett, N.A., Smith, L.A., Marks, J.D., Xie, G., and Brettin, T.S. (2007) Analysis of the neurotoxin complex genes in Clostridium botulinum A1-A4 and B1 strains: BoNT/A3, /Ba4 and /B1 clusters are located within plasmids. *PLoS One* **2**: e1271.

Stanley, E., Fitzgerald, G.F., Le Marrec, C., Fayard, B., and van Sinderen, D. (1997) Sequence analysis and characterization of phi O1205, a temperate bacteriophage infecting *Streptococcus thermophilus* CNRZ1205. *Microbiology* **143**: 3417-3429.

Sumby, P., Porcella, S.F., Madrigal, A.G., Barbian, K.D., Virtaneva, K., Ricklefs, S.M., Sturdevant, D.E., Graham, M.R., Vuopio-Varkila, J., Hoe, N.P., and Musser, J.M. (2005) Evolutionary origin and emergence of a highly successful clone of serotype M1 group a Streptococcus involved multiple horizontal gene transfer events. *J Infect Dis* **192**: 771-782.

Takeuchi, F., Watanabe, S., Baba, T., Yuzawa, H., Ito, T., Morimoto, Y., Kuroda, M., Cui, L., Takahashi, M., Ankai, A., Baba, S., Fukui, S., Lee, J.C., and Hiramatsu, K. (2005) Whole-genome sequencing of staphylococcus haemolyticus uncovers the extreme plasticity of its genome and the evolution of human-colonizing staphylococcal species. *J Bacteriol* **187**: 7292-7308.

van de Werken, H.J., Verhaart, M.R., VanFossen, A.L., Willquist, K., Lewis, D.L., Nichols, J.D., Goorissen, H.P., Mongodin, E.F., Nelson, K.E., van Niel, E.W., Stams, A.J., Ward, D.E., de Vos, W.M., van der Oost, J., Kelly, R.M., and Kengen, S.W. (2008) Hydrogenomics of the extremely thermophilic bacterium Caldicellulosiruptor saccharolyticus. *Appl Environ Microbiol* **74**: 6720-6729.

Veith, B., Herzberg, C., Steckel, S., Feesche, J., Maurer, K.H., Ehrenreich, P., Baumer, S., Henne, A., Liesegang, H., Merkl, R., Ehrenreich, A., and Gottschalk, G. (2004) The complete genome sequence of Bacillus licheniformis DSM13, an organism with great industrial potential. *J Mol Microbiol Biotechnol* **7**: 204-211.

Vuilleumier, S., Chistoserdova, L., Lee, M.C., Bringel, F., Lajus, A., Zhou, Y., Gourion, B., Barbe, V., Chang, J., Cruveiller, S., Dossat, C., Gillett, W., Gruffaz, C., Haugen, E., Hourcade, E., Levy, R., Mangenot, S., Muller, E., Nadalig, T., Pagni, M., Penny, C.,

Peyraud, R., Robinson, D.G., Roche, D., Rouy, Z., Saenampechek, C., Salvignol, G., Vallenet, D., Wu, Z., Marx, C.J., Vorholt, J.A., Olson, M.V., Kaul, R., Weissenbach, J., Medigue, C., and Lidstrom, M.E. (2009) Methylobacterium genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources. *PLoS One* **4**: e5584.

Wagner-Dobler, I., Ballhausen, B., Berger, M., Brinkhoff, T., Buchholz, I., Bunk, B., Cypionka, H., Daniel, R., Drepper, T., Gerdts, G., Hahnke, S., Han, C., Jahn, D., Kalhoefer, D., Kiss, H., Klenk, H.P., Kyrpides, N., Liebl, W., Liesegang, H., Meincke, L., Pati, A., Petersen, J., Piekarski, T., Pommerenke, C., Pradella, S., Pukall, R., Rabus, R., Stackebrandt, E., Thole, S., Thompson, L., Tielen, P., Tomasch, J., von Jan, M., Wanphrut, N., Wichels, A., Zech, H., and Simon, M. The complete genome sequence of the algal symbiont Dinoroseobacter shibae: a hitchhiker's guide to life in the sea. *Isme J* **4**: 61-77.

Ward, N.L., Challacombe, J.F., Janssen, P.H., Henrissat, B., Coutinho, P.M., Wu, M., Xie, G., Haft, D.H., Sait, M., Badger, J., Barabote, R.D., Bradley, B., Brettin, T.S., Brinkac, L.M., Bruce, D., Creasy, T., Daugherty, S.C., Davidsen, T.M., DeBoy, R.T., Detter, J.C., Dodson, R.J., Durkin, A.S., Ganapathy, A., Gwinn-Giglio, M., Han, C.S., Khouri, H., Kiss, H., Kothari, S.P., Madupu, R., Nelson, K.E., Nelson, W.C., Paulsen, I., Penn, K., Ren, Q., Rosovitz, M.J., Selengut, J.D., Shrivastava, S., Sullivan, S.A., Tapia, R., Thompson, L.S., Watkins, K.L., Yang, Q., Yu, C., Zafar, N., Zhou, L., and Kuske, C.R. (2009) Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. *Appl Environ Microbiol* **75**: 2046-2056.

Yagi, J.M., Sims, D., Brettin, T., Bruce, D., and Madsen, E.L. (2009) The genome of Polaromonas naphthalenivorans strain CJ2, isolated from coal tar-contaminated sediment, reveals physiological and metabolic versatility and evolution through extensive horizontal gene transfer. *Environ Microbiol* **11**: 2253-2270.

Yasmin, A., Kenny, J.G., Shankar, J., Darby, A.C., Hall, N., Edwards, C., and Horsburgh, M.J. Comparative Genomics and Transduction Potential of Enterococcus faecalis Temperate Bacteriophages. *J Bacteriol* **192**: 1122-1130.

Zhang, Z.Y., Liu, C., Zhu, Y.Z., Zhong, Y., Zhu, Y.Q., Zheng, H.J., Zhao, G.P., Wang, S.Y., and Guo, X.K. (2009) Complete genome sequence of Lactobacillus plantarum JDM1. *J Bacteriol* **191**: 5020-5021.

Chapter 3: Cleavage of the ICE*Bs1* repressor by the conserved anti-repressor ImmA can be activated by increasing ImmA protein levels or specific activity

Abstract

The mobile genetic element ICE*Bs1* is an integrative and conjugative element (a conjugative transposon) found in the *Bacillus subtilis* chromosome. The RecA-dependent SOS response and the RapI-PhrI sensory system activate ICE*Bs1* gene expression, excision, and mating by stimulating cleavage of the element's repressor ImmR by the anti-repressor ImmA. The mechanism by which RapI and RecA activate ImmA-catalyzed cleavage of ImmR is unknown. We found that changing the amount or the specific activity of ImmA can cause derepression of ICE*Bs1* without activation by RecA or RapI. We isolated and characterized mutations in *immA* that cause derepression of ICE*Bs1* gene expression in the absence of inducing signals (*immA*^{h}). We obtained two types of ImmA^{h} mutants. One type was more abundant in the cell than wild type, and the other cleaved ImmR faster than wild type ImmA in vitro. In addition, we analyzed the effects of overproducing wild type ImmA in vivo, and found that raising ImmA levels could derepress ICE*Bs1* gene expression. However, we also found that ImmA levels did not significantly change during activation by RapI, indicating that RapI-mediated induction is likely due to an increase in the activity of ImmA. We propose that RapI and RecA induce ICE*Bs1* by increasing ImmA's specific activity. Homologs of ImmA and ImmR are found in many mobile genetic elements, so the mechanisms that regulate ImmA-mediated cleavage of ImmR may be conserved in various systems.

Introduction

Integrative and conjugative elements (ICEs), also known as conjugative transposons, are mobile genetic elements found in a wide range of bacteria. These elements reside in the host chromosome and contribute to genome plasticity. They facilitate the acquisition of new traits, including antibiotic resistance, symbiosis, and virulence (reviewed in Burrus *et al.*, 2002; Burrus *et al.*, 2006; Churchward, 2002; Roberts and Mullany, 2009; Salyers *et al.*, 1995). ICEs can excise from the host chromosome and transfer to other cells by conjugation. Once inside the recipient, the ICE typically integrates into the host chromosome and is stably maintained and propagated by host replication and cell division. ICEs generally encode proteins that function in their regulation, integration, excision, and transfer (Churchward, 2002; Whittle *et al.*, 2002).

ICE*Bs1* is an approximately 20 kbp integrative and conjugative element (Fig. 1) inserted in a tRNA gene in *Bacillus subtilis* (Auchtung *et al.*, 2005; Burrus *et al.*, 2002). Genes at the left end of ICE*Bs1* are part of a regulatory module that resembles those found in many bacteriophages (Auchtung *et al.*, 2005; Burrus *et al.*, 2002). This module includes *immR* and *immA*, encoding the element's repressor and anti-repressor, respectively (Auchtung *et al.*, 2007; Bose* *et al.*, 2008). ImmR represses transcription of genes required for excision and transfer and both activates and represses its own expression (Auchtung *et al.*, 2007). ICE*Bs1* gene expression is derepressed in vivo during the RecA-dependent SOS response, or when the ICE*Bs1*-encoded cell-cell signaling regulator RapI is present and active (Auchtung *et al.*, 2005). In both cases, derepression requires the anti-repressor ImmA (Bose* *et al.*, 2008). ImmA is a site-specific protease that cleaves ImmR thereby causing derepression of ICE*Bs1* gene expression (Bose* *et al.*, 2008). It is not known how the protease activity of ImmA is stimulated by RecA or RapI.

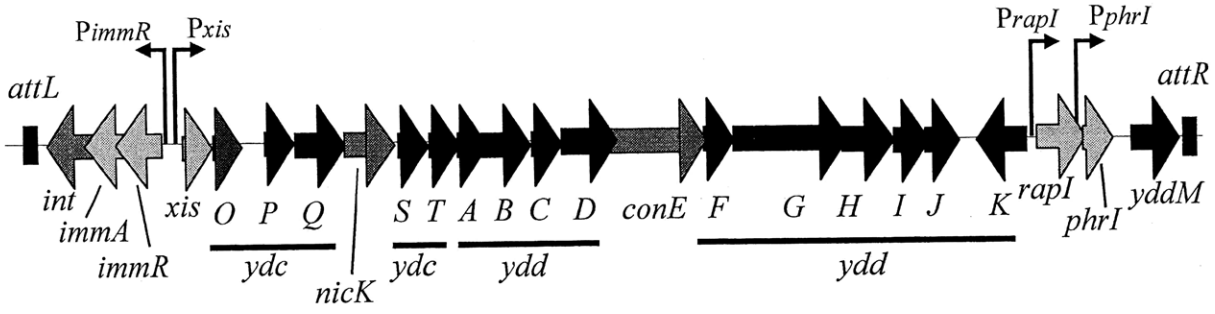


Figure 1. Organization of ICE*Bs*1. The 24 orfs and four promoters are shown as arrows. The name of each gene is indicated below its arrow. Boxes at the left and right represent attachment sites. Genes encoding characterized products are grey. *immA* encodes the anti-repressor and *immR* encodes the repressor for the element. *int* and *xis* encode the integrase and excisionase, respectively. *nicK* encodes a DNA relaxase that nicks the DNA at an origin of transfer (*oriT*) sequence within the *nicK* gene. *conE* encodes a protein that likely forms part of the mating apparatus.

We found that changing the amount or changing the specific activity of ImmA can cause derepression of ICE*Bs1*, even without activation by RecA or RapI. We isolated and characterized mutations in *immA* that cause derepression of ICE*Bs1* gene expression in the absence of exogenous inducing signals. We also analyzed the effects of artificially elevating the amount of ImmA in the cell. Our results indicate that there are at least two ways in which ImmA-mediated cleavage of ImmR can be activated: 1) by increasing the activity of ImmA, or 2) by increasing the cellular concentration of ImmA. We also found that ImmA levels did not significantly change during activation by RapI, indicating that RapI-mediated induction likely results from an increase in the activity of ImmA.

Results

Overexpression of *immA* can derepress P*xis-lacZ*

Previous experiments indicated that cells with C-terminally epitope-tagged ImmA had elevated ImmA protein levels and a small increase in ICE*Bs1* gene expression (J.M. Auchtung, ADG, unpublished results). Based on these results, we decided to directly test the relationship between levels of native (untagged) ImmA and expression of the major promoter in ICE*Bs1*. We used transcription of a P*xis-lacZ* fusion as an indicator of ICE*Bs1* gene expression. P*xis* is strongly repressed by ImmR (Auchtung *et al.*, 2007). Derepression requires ImmA and occurs after overexpression of *rapI* or during the RecA-mediated SOS response (Bose* *et al.*, 2008). This regulation occurs both in the presence and absence of all other ICE*Bs1* genes (Auchtung *et al.*, 2007; Bose* *et al.*, 2008). For simplicity, the experiments presented here were done is cells lacking ICE*Bs1* (ICE*Bs1*^{0}).

Increased expression of *immA* caused increased expression of P*xis-lacZ*. We modulated expression of *immA* using a fusion to the IPTG-inducible promoter P*spank* (P*spank-immA*) or to the stronger P*spank(hy)* (P*spank(hy)-immA*) and varying induction with different concentrations of IPTG. Cells expressed *immR* from its own promoter and had a copy of *rapI* fused to the xylose-inducible promoter P*xyl* (P*xyl-rapI*), but were grown in the absence of xylose. Expression of P*xis-lacZ* (Fig. 2A) was low in cells containing the weaker P*spank-immA* fusion and growing in 1 mM IPTG (fully-induced P*spank*).

Expression of P*xis-lacZ* was higher when cells containing P*spank(hy)-immA* were grown in 10 μM IPTG, and higher still when grown in 100 μM IPTG (Fig. 2A). Expression of P*xis-lacZ* was fully derepressed (~1000 specific activity units) when cells containing P*spank(hy)-immA* were grown in 1 mM IPTG (data not shown). We also measured levels of ImmA protein, in Western blots with anti-ImmA antibodies, under the conditions used for monitoring expression of P*xis-lacZ*. As expected, the relative amount of ImmA increased with higher concentrations of IPTG (Fig. 2B). Together, these results indicate that increased levels of ImmA cause increased transcription from P*xis*.

Induction of P*xis-lacZ* by overexpression of RapI does not cause an increase in the amount of ImmA

Although increasing the amount of ImmA caused increased expression of P*xis-lacZ*, the amount of ImmA did not increase when expression of P*xis-lacZ* was induced by overproduction of RapI. Addition of xylose to induce expression of *rapI* (from P*xyl-rapI*), in cells with P*spank-immA* grown in 1 mM IPTG, caused an increase in expression of P*xis-lacZ* such that after 90-120 min, cells had ~1000-fold more ß-galactosidase specific activity than before production of RapI (Fig. 2A). Under these conditions, there was little or no detectable change in the amount of

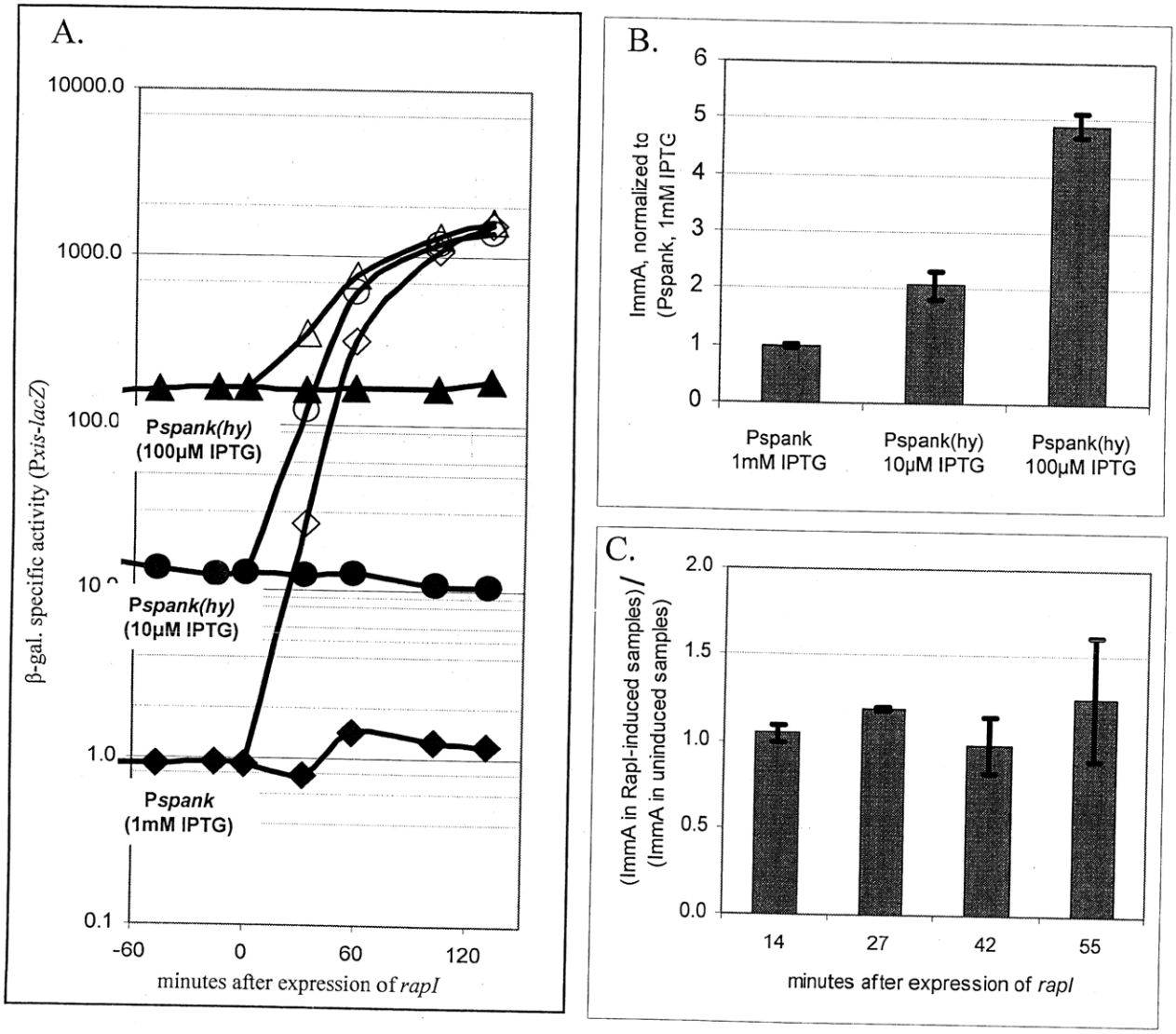

A.
10000.0
1000.0
100.0
10.0
1.0
0.1
β-gal. specific activity (Pxis-lacZ)
Pspank(hy)
(100μM IPTG)
Pspank(hy)
(10μM IPTG)
Pspank
(1mM IPTG)
-60
0
60
120
minutes after expression of rapI
B.
ImmA, normalized to
(Pspank, 1mM IPTG
6
5
4
3
2
1
0
Pspank
1mM IPTG
Pspank(hy)
10μM IPTG
Pspank(hy)
100μM IPTG
C.
(ImmA in RapI-induced samples) /
(ImmA in uninduced samples)
2.0
1.5
1.0
0.5
0.0
14
27
42
55
minutes after expression of rapI

Figure 2. Effects of ImmA levels on expression of P*xis-lacZ*. All strains were cured of ICE*Bs1* (ICE*Bs1*^{0}) and expressed *immR*, *immA*, and *rapI* at ectopic loci. Expression of *immA* from P*spank* or the stronger P*spank(hy)* was controlled with different concentrations of IPTG. Cells were grown in defined minimal medium with arabinose at 37°C. At t=0, cultures were split and xylose was added to one part to induce expression of P*xyl-rapI*. Effects on ICE*Bs1* gene expression were monitored using a P*xis-lacZ* fusion and levels of ImmA protein were determined by Western blotting with anti-ImmA antibodies.

A. β-galactosidase activity is plotted as a function of time relative to addition of xylose to induce expression of P*xyl-rapI*. closed symbols, untreated; open symbols, treated with xylose at time 0; diamonds, P*spank-immA* (BOSE534)with 1mM IPTG; circles, P*spank(hy)* (BOSE1095) with 10μM IPTG; triangles, P*spank-immA* (BOSE1095) with 100 μM IPTG.

B. The amount of ImmA relative to that in cells with P*spank-immA* (BOSE534) grown in 1mM IPTG is plotted. Results from cells with P*spank(hy)-immA* (BOSE1095) grown in 10μM and 100 μM IPTG are shown. For each time point, the average value for two replicates is plotted, and the error bar indicates one standard deviation above and below the average.

C. The ratio of ImmA in xylose-treated samples to that in untreated samples is plotted for cells with P*spank-immA* (BOSE534) grown in 1mM IPTG. The time at which samples were taken is indicated below each bar, as minutes after splitting of the culture for treatment with xylose. For each time point, the average value for at least two replicates is plotted, and the error bar indicates one standard deviation above and below the average.

ImmA (Fig. 2C). These results indicate that P*xis-lacZ* is derepressed by RapI even though *immA* is expressed from a heterologous promoter, and that derepression can occur without a significant increase in the amount of ImmA.

Isolation of hyperactive mutants of ImmA

To better understand the mechanisms of derepression of ICE*Bs1* gene expression, we isolated and characterized mutations in *immA*, *immA*^{h} that cause increased expression of P*xis-lacZ* without exogenous induction. The cells contained P*xis-lacZ*, expressed *immR* from its own promoter (to repress P*xis-lacZ*), and had a P*xyl-rapI* fusion. Under normal growth conditions (without xylose) cells containing wild type *immA* or an *immA* null mutation form white colonies on indicator plates containing X-gal (bromo-chloro-indolyl-galactopyranoside), indicating repression of P*xis-lacZ*. We introduced mutagenized *immA* under control of P*spank* into these cells and screened for mutants that formed blue colonies, indicating at least partial derepression of P*xis-lacZ*. The *immA* alleles from candidate mutants were backcrossed and cells retested to be sure the mutant phenotype was linked to *immA*. Mutant *immA* alleles were then sequenced. Since several alleles contained multiple mutations, we reconstructed each single mutation by site-directed mutagenesis of *immA* and tested each for effects on expression of P*xis-lacZ* (Experimental procedures). Each mutation is designated by the amino acid in wild type ImmA, its position in the sequence, and the amino acid to which it was changed. Changes to a stop codon are designated with a "*". Single mutations causing K13I, T23S, R85G, K90R, V92A, V92E, N93D, F144I, F156S, and I165* were remade (Fig. 3) and tested for effects on expression of P*xis-lacZ*. Most had no significant effect on expression of P*xis-lacZ* (data not shown) and were not characterized further.

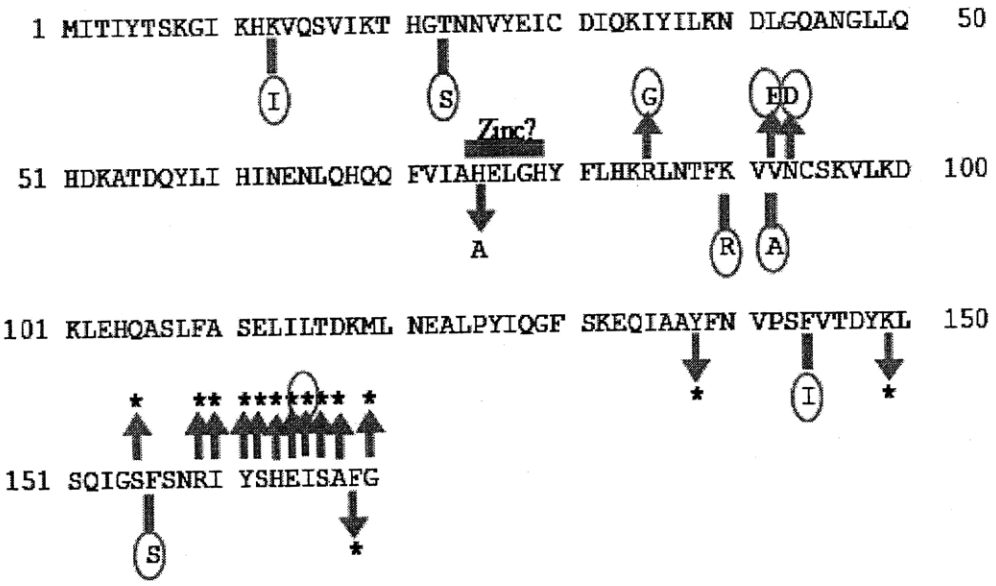


Figure 3. ImmA sequence and mutations. The complete amino acid sequence of ImmA is shown. The horizontal bar indicates the putative Zinc-binding motif, HELGH. Single amino acid substitutions are shown above and below the sequence. Circled mutations were identified in the mutant hunt, remade, and tested. Up arrows indicate mutations that cause a hyperactive phenotype. Down arrows indicate mutations that cause a null phenotype. Vertical bars indicate mutations that maintain the wild type phenotype. Stars indicate stop codons.

Four mutations in *immA*, R85G, V92E, N93D, and I165*, caused increased expression of P*xis-lacZ* on X-gal indicator plates, indicating that they somehow caused ImmA to be hyperactive. Three of these mutations, R85G, V92E, N93D, are in the central region of ImmA, just C-terminal to the predicted Zinc-binding motif, HEXXH (Fig. 3). The I165* mutation truncates the protein by five amino acids at its C terminus. We further explored the effect of truncating ImmA at its C terminus by targeted mutagenesis. Nonsense mutations at positions 155, 159 through 167, and 169 caused a hyperactive phenotype and nonsense mutations at positions 138, 149, and 168 caused a null phenotype (Fig. 3), as detected on X-gal indicator plates.

Effects of ImmA^h mutants on expression of P*xis-lacZ*

We further characterized the three central missense mutations (R85G, V92E, and N93D) and three of the C-terminal nonsense mutations (I160*, I165*, and G169*). Each caused an increase in expression of P*xis-lacZ* in cells grown in liquid medium (Fig. 4), although to different extents. R85G had the largest effect (Fig. 4A, C) and G169* the smallest (Fig. 4B, D). Expression of P*xis-lacZ* in all mutants was further induced after overexpression of *rapI*, and all strains reached the same maximum level of P*xis-lacZ* expression (Fig. 4A, B). However, in response to Mitomycin C (MMC) to induce a DNA damage response, strains with higher uninduced levels of P*xis-lacZ* expression had a higher induced level of ß-galactosidase specific activity than those with lower uninduced expression levels (Fig. 4C, D). This is consistent with typically lower induction of ICE*Bs1* in response to MMC than in response to RapI (Auchtung *et al.*, 2005). In addition, MMC caused a decrease in cell growth and viability as judged by a drop in OD600 ~2 hours after treatment, indicating that cells may have begun to die before all the strains could accumulate the maximum level of ß-galactosidase.

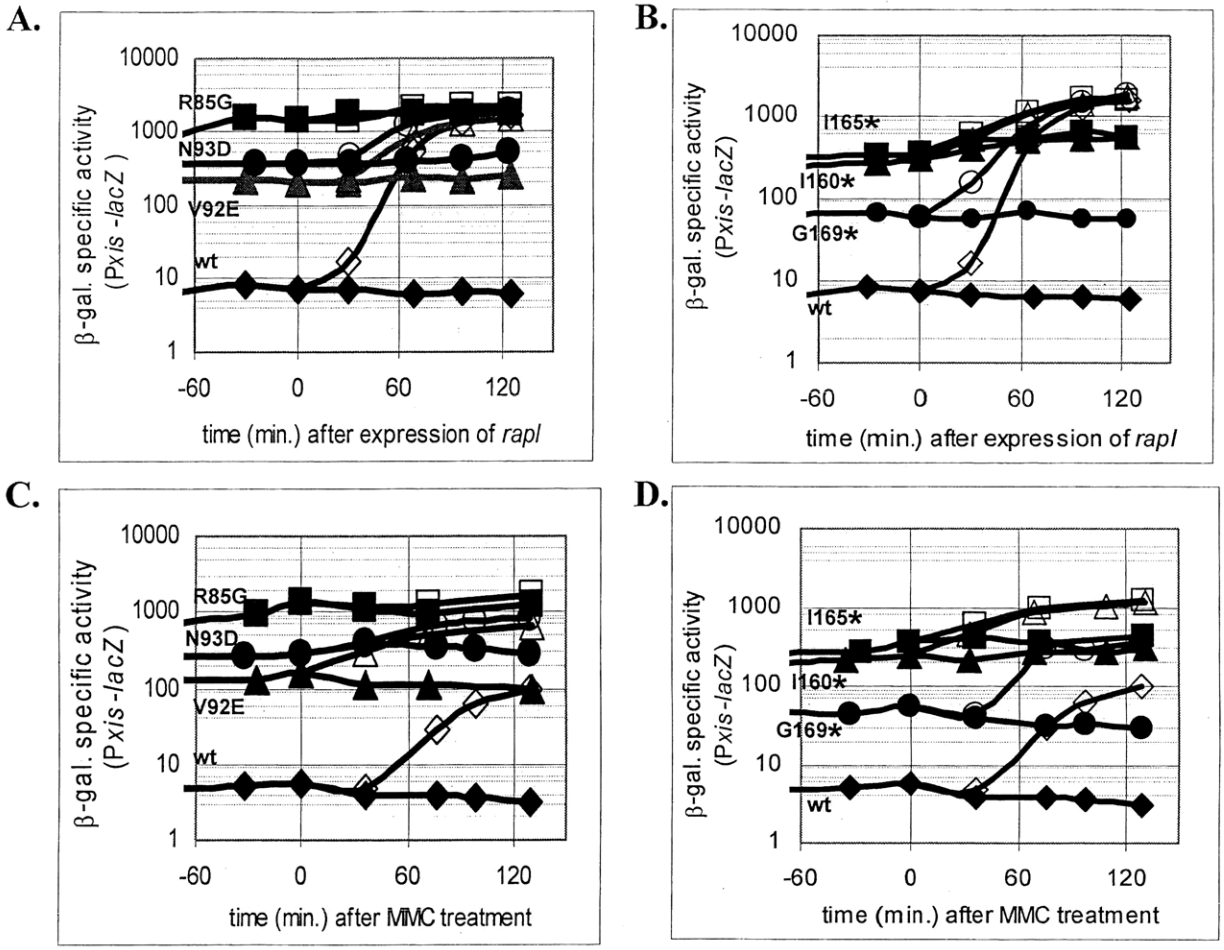


Figure 4. Effects of ImmA^h mutants on P*xis-lacZ*. All strains were cured of ICE*Bs1* (ICE*Bs1*^{0}) and expressed *immR*, *immA*, and *rapI* at ectopic loci. Strains harboring *immA*^{h} mutants are compared to wild type. Cultures were grown in defined minimal medium plus 1mM IPTG for continuous expression of *immA* variants from P*spank*. At t=0 cultures were split, and one part was treated with xylose to induce expression of P*xyl-rapI* (**A.** and **B.**, open symbols) or MMC to induce the SOS response (**C.** and **D.**, open symbols).The other part of each culture was left untreated (closed symbols). Effects on ICE*Bs1* gene expression were monitored using a P*xis-lacZ* fusion.

A. and C. *immA*^{h} variants with mutations near the center of ImmA's sequence. wild type, BOSE534, diamonds; R85G, BOSE569, squares; V92E, BOSE708, triangles; N93D, BOSE672, circles.

B. and D. *immA*^{h} variants with mutations near ImmA's C terminus. wild type, BOSE534, diamonds; I160*, BOSE912, triangles; I165*, BOSE713, squares; G169*, BOSE904, circles.

We constructed several alleles of *immA* that contained two mutations each of which causes a hyperactive phenotype. The double mutants V92E G169* (Fig. 5A), V92E I165* (Fig. 5B), V92E N93D (Fig. 5C), and R85G I165* (Fig. 5D), all caused higher expression of P*xis-lacZ* than either of the respective single mutations. A fifth combination, N93D, S155*, had higher expression of P*xis-lacZ* than one of the single mutants (S155*), but less than the other (N93D) (data not shown). These results indicate that for some of the combinations, the effects on expression of P*xis-lacZ* appear to be additive.

RapI and RecA are not required for ImmA^h phenotypes

The ImmA^h mutants respond to both RapI and DNA damage, and the strain background in which we isolated the *immA^h* mutants contained P*xyl-rapI* and was *recA+*. Therefore, we tested whether the ImmA^h mutants required endogenous RecA and perhaps low level (leaky) expression of P*xyl-rapI* under otherwise non-inducing conditions. We found that the *immA^h* mutations tested (R85G, V92E, I165*, G169*) all caused elevated expression of P*xis-lacZ* in a *recA* null mutant in the absence of *rapI* (Fig. 6A), although expression was lower than that in *recA$^+$* cells with P*xyl-rapI* (Fig. 6B). In cells producing wild type ImmA, expression of P*xis-lacZ* was also reduced in the *recA* mutant in the absence of *rapI*, indicating that much of this effect is independent of the *immA^h* mutations. These finding indicate that the ImmA^h mutants are hyperactive in the absence of any of the known inducers (RecA and RapI).

The ImmA C-terminal truncation mutants have increased protein levels in vivo

Because elevated levels of ImmA caused increased expression of P*xis-lacZ* (Fig. 2), we determined whether any of the *immA^h* mutations caused an increase in the amount of ImmA in the cell. We measured relative amounts of wild type ImmA and six ImmA^h mutants using Western blots with anti-ImmA antibodies. Each of the three central ImmA^h mutants (R85G,

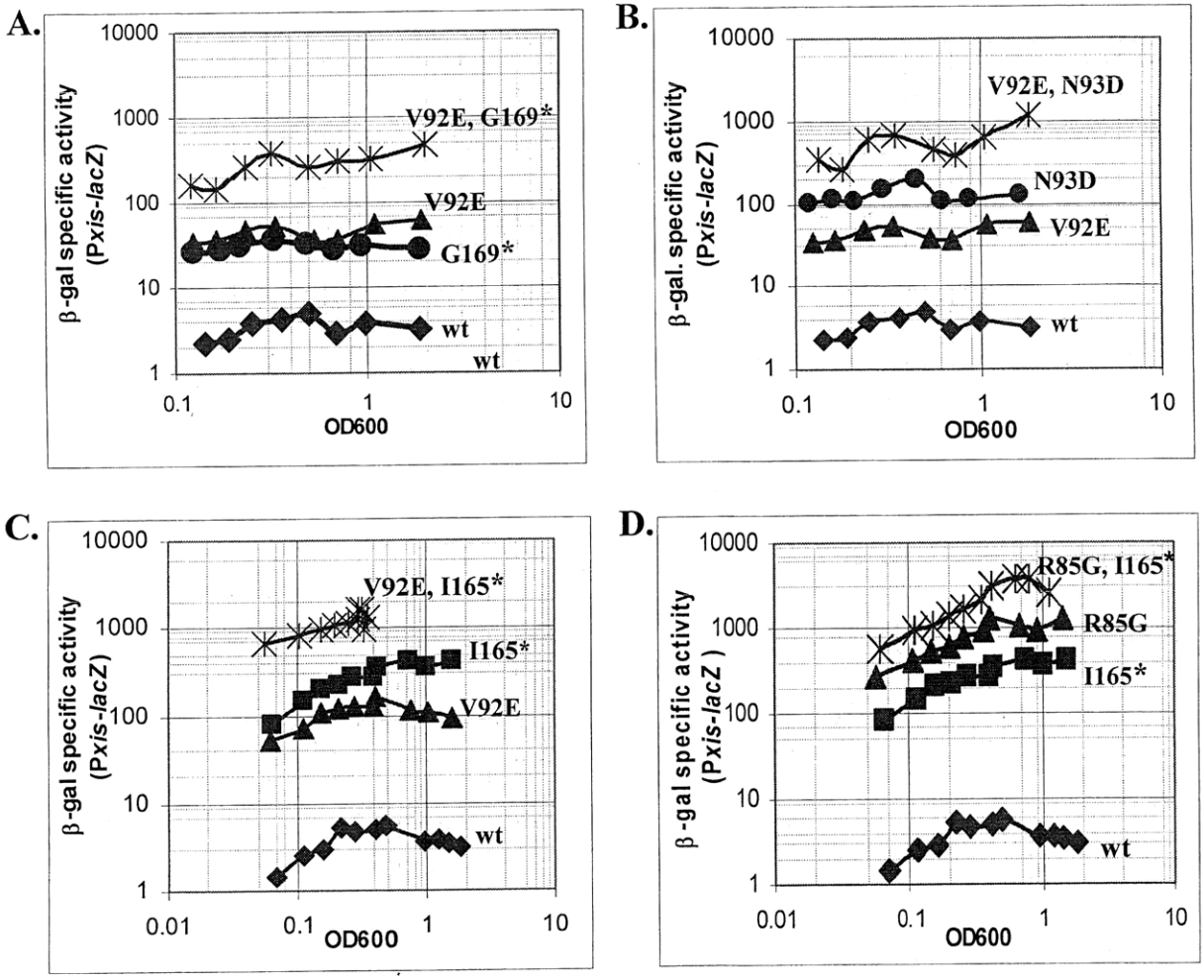

A.
β-gal specific activity (Pxis-lacZ)
10000
1000
100
10
1
0.1
1
10
OD600
V92E, G169*
V92E
G169*
wt
wt
B.
β-gal. specific activity (Pxis-lacZ)
V92E, N93D
N93D
V92E
wt
C.
β-gal specific activity (Pxis-lacZ)
0.01
V92E, I165*
I165*
V92E
wt
D.
β-gal specific activity (Pxis-lacZ)
R85G, I165*
R85G
I165*
wt

Figure 5. Effects on P*xis-lacZ* expression of *immA* alleles that contain two mutations, either of which causes a hyperactive phenotype. All strains were cured of ICE*Bs1* (ICE*Bs1*^{0}) and expressed *immR* and *immA* at ectopic loci. Cultures were grown in defined minimal medium with 1mM IPTG for expression of *immA* from P*spank*. Effects on ICE*Bs1* gene expression were monitored using a P*xis-lacZ* fusion. β-galactosidase activity is plotted as a function of OD600 of the culture.

A. wt, BOSE534, diamonds; V92E, BOSE708, triangles; G169*, BOSE904, circles; V92E G169*, BOSE1237, asterisks

B. wt, BOSE534, diamonds; V92E, BOSE708, triangles; N93D, BOSE672, circles; V92E N93D, BOSE1257, asterisks

C. wt, BOSE534, diamonds; V92E, BOSE708, triangles; I165*, BOSE713, squares; V92E I165*, BOSE1111, asterisks

B. wt, BOSE534, diamonds; R85G, BOSE569, triangles; I165*, BOSE713, squares; R85G I165*, BOSE1110, asterisks

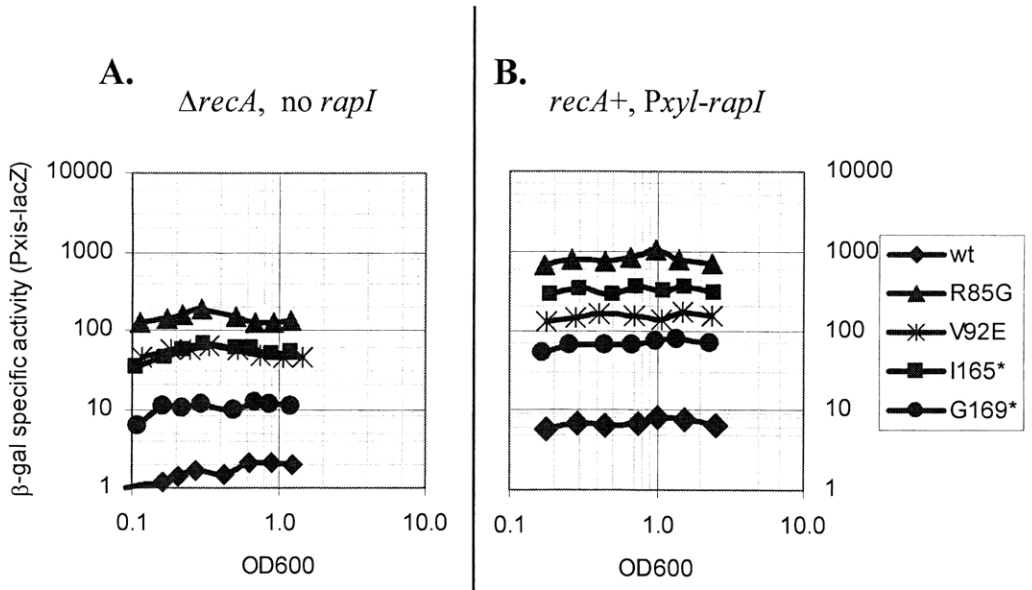


Figure 6. Effects of *rapI* and *recA* on P*xis-lacZ* expression. All strains were cured of ICE*Bs1* (ICE*Bs1*^{0}) and expressed *immR* and *immA* at ectopic loci. Cells harbored wild type *immA* or an *immA*^{h} mutant. Cultures were grown in defined minimal medium with 1% arabinose and 1mM IPTG for expression of *immA* from P*spank*. Effects on ICE*Bs1* gene expression were monitored using a P*xis-lacZ* fusion. β-galactosidase activity is plotted as a function of OD600 of the culture.

A. Strains lack the P*xyl-rapI* construct, harbor a null mutation of *recA*, and express the indicated allele of *immA*. wild type, BOSE1256, diamonds; R85G, BOSE1071, triangles; V92E, BOSE1073, asterisks; I165*, BOSE1074, squares; G169*, BOSE1076, circles.

B. Strains have the ectopic P*xyl-rapI* construct and wild type *recA* and express the indicated allele of *immA*. Expression from P*xyl* was not induced. wild type, BOSE534, diamonds; R85G, BOSE569, triangles; V92E, BOSE708, asterisks; I165*, BOSE713, squares; G169*, BOSE904, circles.

V92E, and N93D) was present in the cell at a level similar to that of wild type ImmA (Fig. 7). In contrast, the three C-terminal mutants examined (I160*, I165* and G169*) had significantly higher protein levels in vivo compared to wild type ImmA (Fig. 7). ImmAG169*, while more abundant than wild type ImmA, was less abundant than the other two truncation mutants. These relative levels are consistent with the effect of each of the truncation mutations on expression of P*xis-lacZ* (Fig. 4). It seemed most likely that the truncation mutations were affecting stability of ImmA, indicating that the wild type protein is normally unstable. Initial results screening known cellular protease mutants indicated that ImmA levels were significantly higher in a *clpP* mutant (BB and ADG, unpublished results). We conclude that the truncation mutations likely cause increased expression of P*xis-lacZ* due to increased stability of the mutant ImmA protein and that the central mutations likely cause increased expression of P*xis-lacZ* by some other mechanism.

Central *immA*^{h} mutations increase activity of ImmA in vitro

The ability of the central ImmA^h mutant proteins to cleave ImmR in vitro was greater than that of wild type ImmA or the truncation mutants. We purified the three central mutants (R85G, V92E, N93D) and three C-terminal truncations ((I160*, I165*, G169*), and compared the abilities of these ImmA mutants to cleave ImmR (His6-ImmR) in vitro. The C-terminal truncations all cleaved ImmR (His6-ImmR) at a rate comparable to that of wild type ImmA (Fig. 8). In contrast, all three central mutants cleaved ImmR significantly faster than did wild type ImmA (Fig. 8). The R85G mutant was the most active in these assays, consistent with the in vivo effects on expression of P*xis-lacZ* (Fig. 4). Together, our results indicate that the C-terminal truncation mutations affect the amount of ImmA protein in vivo, but do not have a significant effect on ImmA specific activity, and that the central mutations primarily affect the specific activity of ImmA, and do not significantly affect ImmA levels in vivo.

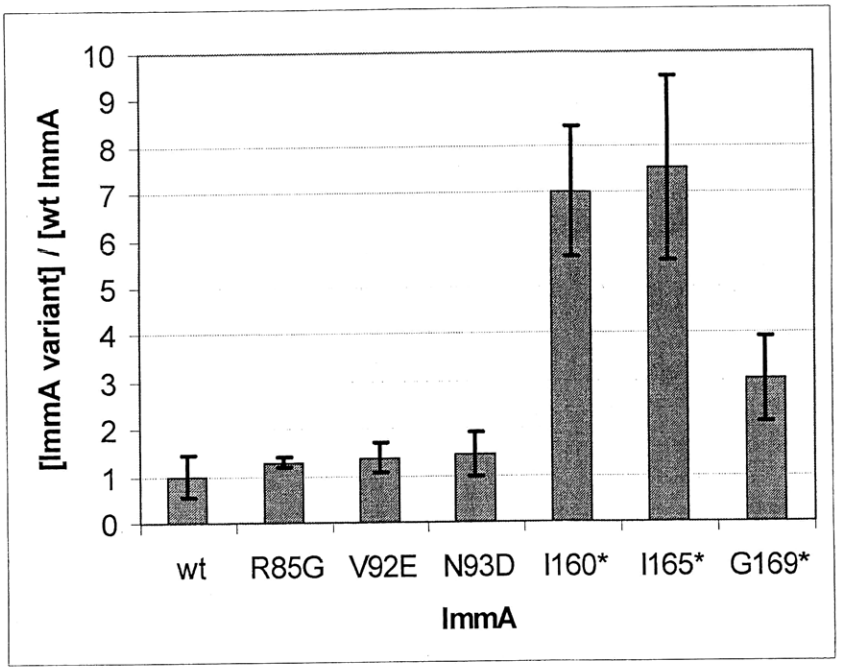


Figure 7. Cellular levels of ImmA^h mutant proteins. Levels of ImmA protein in strains expressing different alleles of *immA* were determined by Western blotting with anti-ImmA antibodies. For each strain, ImmA levels were quantified, and the average of at least three samples was calculated. The average for each type of ImmA was then normalized to the average for wild type ImmA and plotted. Error bars indicate one standard deviation above and below each average. Samples were taken from cells grown in defined minimal medium with 1mM IPTG for expression of *immA* from P*spank*. All strains expressed *immR* from its own promoter. wt, BOSE534; R85G, BOSE569; V92E, BOSE708; N93D, BOSE672; I160*, BOSE912; I165*, BOSE713; G169*, BOSE904

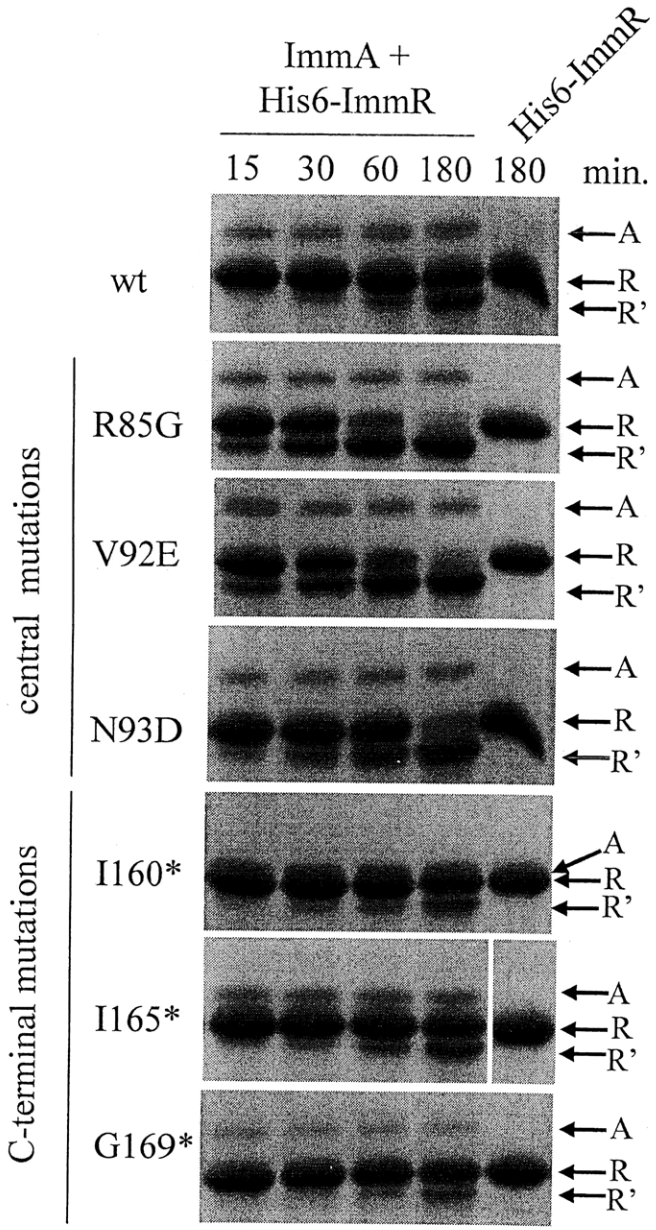

ImmA +
His6-ImmR
His6-ImmR
15 30 60 180 180 min.
wt
R85G
V92E
N93D
I160*
I165*
G169*
central mutations
C-terminal mutations
A
R
R'

Figure 8. In vitro proteolysis of ImmR by ImmA^{h} mutants. 180μM His6-ImmR and 12μM ImmA were incubated together at 37°C for the times indicated above each column. Reaction products were visualized by Coomassie-stained SDS-PAGE. Each row shows the result of reactions containing a single ImmA variant (or no ImmA, for the last column) at successive time points. Wild type ImmA and six ImmAh mutants were tested. ImmA (**A**), intact His6-ImmR (**R**), and the N-terminal fragment of His6-ImmR (**R'**) are marked to the right of each row. ImmA I160* migrates to a position barely distinguishable from that of intact His6-ImmR, so the ImmR fragment can be used to gauge the progress of these reactions.

Discussion

ImmA is required for derepression of ICE*Bs1* during the SOS response or after production of the activator RapI (Bose* *et al.*, 2008). When activated, ImmA cleaves the ICE*Bs1* repressor, ImmR, thereby inducing ICE*Bs1* gene expression (Bose* *et al.*, 2008). We found that P*xis-lacZ* could be derepressed by artificially increasing ImmA production in vivo. However, derepression of P*xis-lacZ* in response to RapI occurred without any significant change in ImmA levels. Isolation and characterization of mutations in *immA* (*immA*^{h}) that cause a hyperactive phenotype, i.e., increased expression of P*xis-lacZ* in the absence of exogenous signals to induce ICE*Bs1* gene expression, identified two classes of mutations. In one class, ImmA was truncated at the C terminus causing an increase in the amount of the mutant protein relative to that of the wild type protein. The second class of mutations were in the central region of ImmA and did not cause an increase in the amount of the mutant protein in vivo. Rather, these mutations caused ImmA to cleave ImmR more rapidly than either wild type ImmA or the truncation mutants. Taken together, our results indicate that there are at least two ways to increase ImmA-mediated proteolysis of ImmR: 1) increase the amount of ImmA, and 2) increase the specific activity of ImmA. Our findings support the idea that the primary mechanism of derepressing ICE*Bs1* in response to RapI is by increasing the specific activity of ImmA and not by increasing the amount of ImmA in the cell.

ImmA levels and derepression of P*xis*

We found that increasing the amount of wild type ImmA, by increasing its production from a heterologous promoter, caused increased expression of P*xis-lacZ*, a reporter for ICE*Bs1* gene expression. In addition, truncations of ImmA that remove several C-terminal amino acids also cause an increase in the amount of ImmA and an increase in expression of P*xis-lacZ*. This

suggests that ImmA is normally unstable, and that this instability is influenced by C-terminal residues of ImmA. Preliminary measurements indicate that ImmA is significantly stabilized in a *clpP* protease mutant (BB and ADG, Appendix C), consistent with the notion that ImmA is normally unstable. There are many examples of regulatory proteins that are constitutively degraded, allowing for rapid changes in their intracellular concentrations when they are stabilized against proteolysis.

Although increasing the amount of ImmA in the cell causes derepression of ICE*Bs1* gene expression, this does not seem to be the primary mechanism for derepression in response to RapI. The amount of ImmA did not detectably change when P*xis-lacZ* expression was induced by production of RapI. Under these conditions, there is an approximately 1,000-fold increase in ß-galactosidase specific activity from P*xis-lacZ*. When the amount of wild type ImmA is increased approximately 4-5-fold, there was an approximately 100-fold increase in ß-galactosidase specific activity from P*xis-lacZ*, significantly less than the increase in expression caused by RapI. In addition, the C-terminal ImmA mutants that cause increased expression from P*xis-lacZ* were still inducible by expression of RapI. Together, these findings indicate that something other than an increase in ImmA levels is causing ImmA-dependent derepression of ICE*Bs1* gene expression. We propose that the activation of ImmA by RapI to cleave ImmR and induce ICE*Bs1* involves an increase in the specific activity of ImmA and not an increase in its intracellular concentration.

Mutations that increase the specific activity of ImmA

In addition to the C-terminal truncation mutants affecting the levels of ImmA, we identified mutations affecting residues near the putative Zinc-binding motif in ImmA. Accumulation of these mutant proteins was indistinguishable from that of wild type ImmA. In vitro, these mutant

proteins were more efficient than wild type at cleaving ImmR, and these effects were generally consistent with the effects on P*xis-lacZ* expression in vivo. These findings indicate that the mutations cause an increase in the specific activity of ImmA. That the phenotype of the double mutant V92E N93D is greater than either single mutant is consistent with the mutants being partially activated through the same mechanism, or partially activated by different mechanisms.

There are several mechanisms by which the *immA* mutations could increase the activity of ImmA, and these could be related to the mechanisms by which RapI and RecA increase activity of ImmA (see below). They might have altered conformations that enhance their solubility or that enable them to recognize and/or cleave ImmR more efficiently than wild type. It's also possible that the ImmA mutants' structures enhance incorporation or retention of metal (presumably zinc) into the protein.

Activation of ICE*Bs1* gene expression by RapI and RecA

ICE*Bs1* gene expression is normally induced by production of active RapI or during the RecA-dependent SOS response (Auchtung *et al.*, 2005). The two activation mechanisms are independent of each other in that: 1) SOS stimulates induction of ICE*Bs1* in a *rapI* null mutant, and 2) that RapI production induces ICE*Bs1* in a *recA* null mutant (Auchtung *et al.*, 2005). Both RapI and RecA stimulate ImmA-dependent cleavage of ImmR, and RapI-mediated stimulation appears to be direct as it occurs in the heterologous host *E. coli* (Bose* *et al.*, 2008). For simplicity, we assume that RecA-stimulated cleavage is also direct, although this has not yet been determined. It is also not clear whether RapI and RecA stimulate ImmA-mediated cleavage of ImmR by the same mechanism. For discussion purposes, and simplicity, we will assume that the RapI and RecA work similarly, although this clearly need not be the case.

There are several possible mechanisms by which RapI and RecA could activate ImmA-mediated cleavage of ImmR. Three general possibilities include: 1) activation of ImmA, 2) activation of ImmR, making it a better substrate for ImmA, and 3) activation by bringing ImmA and ImmR together.

There are several possible mechanisms by which ImmA might be activated to cleave ImmR. RapI and RecA might bind to ImmA to make it favor a conformation that cleaves ImmR faster. Or they might chaperone ImmA to help it fold properly and to prevent its forming insoluble aggregates. In either of these cases, further characterization of the ImmA^{h} mutants that cleave ImmR more efficiently than wild type ImmA in vitro might aid in understanding how its function is regulated. We favor the idea that the central ImmA^{h} mutations alter the protein in a way that mimics what happens to ImmA under ICE*Bs1*-inducing conditions over the model that these mutations cause ImmA to become more active in some entirely novel way.

ImmA and ImmR in other mobile genetic elements

Homologs of ImmA and ImmR are found in many other mobile elements and putative mobile elements (Bose* *et al.*, 2008). The roles of ImmA and ImmR homologs in phage ø105 are similar to those of ImmA and ImmR in regulating ICE*Bs1* (Bose* *et al.*, 2008). The ImmR homolog of phage ø105 represses transcription to maintain lysogeny, and the ImmA homolog is required for induction of lytic growth when the SOS response is induced. Thus, the mode of regulation involving ImmA and ImmR is likely conserved in a variety of other systems. Further characterization of ImmA and the mechanisms by which it is regulated may therefore be relevant to understanding the regulation of many agents of horizontal gene transfer.

Experimental Procedures

Media and growth conditions. *B. subtilis* cells were grown at 37°C with aeration in S7 minimal salts medium (Vasantha and Freese, 1980), except that 50 mM MOPS was used instead of 100} supplemented with 1% arabinose, 0.1% glutamate, 40 μg/ml tryptophan and phenylalanine, and 120 μg/ml threonine. 1% xylose was used to express *rapI* from P*xyl*. When appropriate, antibiotics were used at the following concentrations: chloramphenicol (5 μg/ml); kanamycin (5 μg/ml); tetracycline (10 μg/ml); spectinomycin (100 μg/ml); and erythromycin (0.5 μg/ml) and lincomycin (12.5 μg/ml) together to select for macrolide-lincosamide-streptogramin B (MLS) resistance. MMC (Roche) was used at a final concentrations of 1 μg/ml. X-gal in LB agar plates was 120μg/ml. Strains with Δ*recA260*::*cat-mls* were kept in the dark whenever possible.

E. coli cells were grown in LB at 30°C or 37°C. When appropriate, ampicillin (100-200 μg/ml) and or chloramphenicol (15 μg/ml) were added. To induce gene expression in BL21-AI cells (Invitrogen), growth medium was supplemented with 1-4mM IPTG and 0.2% arabinose.

Strains and alleles. *B. subtilis* strains used in this study are listed in Table 1. Standard techniques were used for cloning and strain construction (Harwood and Cutting, 1990; Sambrook and Russell, 2001). The ICE*Bs1*^{0} strain, and the *cgeD*::{P*immR-immR*) *kan*}, *thrC*::{(P*xis-lacZ*Ω343) *mls*}, Δ*recA260*::*cat-mls*, and the *amyE*::{(P*spank-immA*Ω218) *spc*} alleles were previously described (Auchtung *et al.*, 2005; Auchtung *et al.*, 2007; Bose* *et al.*, 2008). (Note: *mls* is the same as the previously used *erm*).

P*xyl* and *xylR* from pDR160 (Rudner *et al.*, 1998) were cloned into pMMB752 (Berkmen *et al.*, 2009) to generate pBOSE508, a vector for introduction of a xylose-inducible gene at *lacA*. A

sequence extending from 33 bp upstream of the *rapI* start codon to 2 bp downstream of the stop codon was cloned into pBOSE508 to generate pBOSE525, which was integrated at *lacA* by homologous recombination.

Mutations in P*spank-immA* were made by Quikchange (Stratagene) site-directed mutagenesis of pBOSE540, a plasmid containing P*spank-immA*. The sequence from 24 bp upstream of *immA* to 2 bp downstream of its stop codon was cloned downstream of P*spank* in pDR110 (Rokop *et al.*, 2004) to generate pBOSE540. The P*spank* in pBOSE540 was converted to P*spank(hy)* by Quikchange mutagenesis, generating pBOSE1069. This allele was introduced into JH642 by a double crossover at *amyE*.

Proteins for in vitro assays were purified from the following strains. N-terminally His-tagged ImmR was purified from BOSE798 and untagged ImmA was purified from BOSE848 (Bose* *et al.*, 2008). Constructs for overproduction of ImmA mutants in *E. coli* were made by Quikchange mutagenesis of pBOSE831 (Bose* *et al.*, 2008). The resulting plasmids were introduced in BL21-AI *E. coli* to produce BOSE843 (R85G), BOSE844 (V92E), BOSE845 (N93D), BOSE846 (I165*), BOSE1133 (I160*), and BOSE1134 (G169*). ImmA^h mutants were overproduced in and purified from these strains.

ImmA^h mutant hunt

immA was amplified from JH642 DNA by mutagenic PCR with $MnCl_2$. Products were digested by restriction enzymes (NEB) and ligated into the vector pDR110 (gift from D. Rudner, Harvard Medical School). Ligation mixtures were used to transform competent *E. coli* DH5α cells. Transformants were selected by plating on LB agar containing ampicillin. Many colonies from a single PCR reaction were pooled, and plasmid DNA was miniprepped from each pool. Plasmid DNA was then used to transform BOSE533, and transformants were grown on LB

plates containing spectinomycin (to select transformants), IPTG (to induce P*spank-immA*), and X-gal (to visualize P*xis-lacZ* expression). On these plates, colonies with wild type *immA* or no *immA* appear white. Blue colonies were picked and re-streaked to purify single colonies. Each candidate colony was grown in liquid LB, and genomic DNA was obtained from these cells by phenol-chloroform extraction. This DNA was back-crossed to verify linkage of the blue-color phenotype with the *spc* marker associated with *immA*, as follows. DNA was used to transform BOSE533, and cells were plated on LB agar containing spectinomycin, IPTG, and X-gal. If genomic DNA from a candidate produced blue transformants, *immA* was amplified from this DNA by high fidelity PCR using Platinum Taq DNA polymerase (Invitrogen), and the PCR products were sequenced (MIT Biopolymers lab).

Mutations identified in the candidates were reconstructed as follows. The sequence of *immA* in pBOSE540 (pDR110-*immA*) was altered by Quikchange site-directed mutagenesis (Stratagene), and the resulting plasmids were introduced into JH642 by double crossovers at *amyE*. Genomic DNA from these strains was used to introduce each $immA^h$ allele into indicator strains like BOSE533.

Western Blots. Samples collected from *B. subtilis* cultures were flash-frozen on liquid nitrogen and stored at -20C before being thawed and pelleted, or they were pelleted immediately. Cells were pelleted by centrifugation, washed with TN Buffer (50 mM Tris 300 mM NaCl, pH8), and stored at –20°C. Cell pellets were thawed, resuspended in an appropriate volume of Buffer (10 mM Tris, 10 mM EDTA, pH 7) containing 0.1 mg/ml lysozyme and 1 mM AEBSF, and incubated at 37°C for 30 min. Pellets were resuspended in a volume of buffer proportionate to the OD600 at which they were recovered in order to normalize the concentration of proteins in

the sample. SDS-Loading Buffer was then added to samples, which were heated at 100°C for 10 min. followed by centrifugation to remove insoluble material.

Proteins were separated by SDS-PAGE on 15% or 18% gels and transferred to PolyScreen PVDF membrane (Perkin Elmer) using the Trans-blot semi-dry electro-blot transfer apparatus (BioRad). Membranes were blocked in 0.2% I-Block (Tropix) in TBST (50 mM Tris 200 mM NaCl 0.05% Tween-20 pH 8) either at room temperature for 1 hr. or overnight at 4°C. Membranes were incubated in 1:5,000 anti-ImmA rabbit polyclonal antisera in 0.2% I-Block TBST for 1 hr. at room temperature, washed several times in TBST, incubated in 1:3,000 goat anti-rabbit IgG-HRP conjugate (BioRad) for 1 hour at room temperature, and washed several times in TBST. Signals were detected using Western Lightning chemiluminescence reagents (Perkin-Elmer) followed by exposure to Kodak Biomax Light film.

β-galactosidase assays. β-galactosidase specific activity was assayed as described [Jaacks, 1989, spo0h]. Specific activity was calculated relative to the optical density at 600 nm of the samples. Results shown are from a single experiment and are representative of results obtained in at least three independent experiments, except for Figure 2A, which shows the results of one experiment.

Protein purification and in vitro assays. N-terminally His-tagged ImmR, untagged ImmA, and untagged mutants of ImmA were purified as previously described (Bose* *et al.*, 2008). In vitro assays were conducted with these proteins, also as previously described (Bose* *et al.*, 2008). Because different proportions of proteins were mixed in the assays presented here, the final reaction buffer is 44mM Tris pH~7.2, 38.8mM NaCl, 77mM GdmCl, 12.8μM EDTA, 640μM GSH, 7.3mM DTT, 73μM $ZnCl_2$, 583mM L-arg, 16.6% glycerol.

Table 1. *B. subtilis* strains

BOSE533	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE534	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE569	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(R85G)*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE672	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(N93D)* Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE708	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(V92E)*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE713	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(I165*)*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE904	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(G169*)*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE912	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(I160*)*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE1071	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(R85G)*Ω218) spc] Δ*recA260*::*cat-mls*
BOSE1073	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(V92E)*Ω218) spc] Δ*recA260*::*cat-mls*
BOSE1074	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(I165*)*Ω218) spc] Δ*recA260*::*cat-mls*
BOSE1076	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(G169*)*Ω218) spc] Δ*recA260*::*cat-mls*
BOSE1095	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank(hy)-immA*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE1110	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(R85G, I165*)*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]

BOSE1111	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(V92E, I165*)*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE1237	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(V92E, G169*)*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE1257	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA(V92E, N93D)*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]

Acknowledgements

We thank J.M. Auchtung , C.A. Lee, A.I. Goranov, M.B. Berkmen, and S.E. Cohen for helpful discussions and C.A. Lee for comments on the manuscript. This work was supported in part by an NSF graduate fellowship to BB and Public Health Service grant GM50895 from the NIH to ADG.

References

Auchtung, J.M., Lee, C.A., Monson, R.E., Lehman, A.P., and Grossman, A.D. (2005) Regulation of a Bacillus subtilis mobile genetic element by intercellular signaling and the global DNA damage response. *Proc Natl Acad Sci U S A* **102**: 12554-12559.

Auchtung, J.M., Lee, C.A., Garrison, K.L., and Grossman, A.D. (2007) Identification and characterization of the immunity repressor (ImmR) that controls the mobile genetic element ICEBs1 of Bacillus subtilis. *Mol Microbiol* **64**: 1515-1528.

Berkmen, M.B., Lee, C.A., Loveday, E.K., and Grossman, A.D. (2009) Polar positioning of a conjugation protein from the integrative and conjugative element ICEBs1 of Bacillus subtilis. *J Bacteriol.*

Bose*, B., Auchtung*, J.M., Lee, C.A., and Grossman, A.D. (2008) A conserved anti-repressor controls horizontal gene transfer by proteolysis. *Mol Microbiol* **70**: 570-582. * co-first authors.

Burrus, V., Pavlovic, G., Decaris, B., and Guedon, G. (2002) Conjugative transposons: the tip of the iceberg. *Mol Microbiol* **46**: 601-610.

Burrus, V., Pavlovic, G., Decaris, B., and Guedon, G. (2002) The ICESt1 element of Streptococcus thermophilus belongs to a large family of integrative and conjugative elements that exchange modules and change their specificity of integration. *Plasmid* **48**: 77-97.

Burrus, V., Marrero, J., and Waldor, M.K. (2006) The current ICE age: biology and evolution of SXT-related integrating conjugative elements. *Plasmid* **55**: 173-183.

Churchward, G. (2002) Conjugative transposons and related mobile elements. In *Mobile DNA II.* Craig, N.L. (ed). Washington, D.C.: ASM Press, pp. 177-191.

Harwood, C.R., and Cutting, S.M. (1990) *Molecular biological methods for Bacillus.* Chichester; New York: Wiley.

Roberts, A.P., and Mullany, P. (2009) A modular master on the move: the Tn916 family of mobile genetic elements. *Trends Microbiol* **17**: 251-258.

Rokop, M.E., Auchtung, J.M., and Grossman, A.D. (2004) Control of DNA replication initiation by recruitment of an essential initiation protein to the membrane of Bacillus subtilis. *Mol Microbiol* **52**: 1757-1767.

Rudner, D.Z., Breger, K.S., and Rio, D.C. (1998) Molecular genetic analysis of the heterodimeric splicing factor U2AF: the RS domain on either the large or small Drosophila subunit is dispensable in vivo. *Genes Dev* **12**: 1010-1021.

Salyers, A.A., Shoemaker, N.B., Stevens, A.M., and Li, L.Y. (1995) Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. *Microbiol Rev* **59**: 579-590.

Sambrook, J., and Russell, D.W. (2001) *Molecular cloning: a laboratory manual.* Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Vasantha, N., and Freese, E. (1980) Enzyme changes during Bacillus subtilis sporulation caused by deprivation of guanine nucleotides. *J Bacteriol* **144**: 1119-1125.

Whittle, G., Shoemaker, N.B., and Salyers, A.A. (2002) The role of Bacteroides conjugative transposons in the dissemination of antibiotic resistance genes. *Cell Mol Life Sci* **59**: 2044-2054.

Appendix C: ImmA protein levels are affected by its C-terminal sequence and by host proteases

Abstract

The mobile genetic element ICE*Bs1* is an integrative and conjugative element (a conjugative transposon) found in the *Bacillus subtilis* chromosome. The RecA-dependent SOS response and the RapI-PhrI sensory system activate ICE*Bs1* gene expression, excision, and mating by stimulating cleavage of the element's repressor ImmR by the anti-repressor ImmA. Earlier work showed that the normally low levels of ImmA in *B. subtilis* cells could be increased by changing the C-terminal sequence of ImmA, and that increased levels of ImmA derepressed ICE*Bs1* gene expression. We investigated whether ImmA is degraded by any of *B. subtilis*'s energy-dependent proteases by screening strains with mutated proteases for elevated ICE*Bs1* gene expression. We found that gene expression was partially derepressed in cells with null mutations of *clpX* or *clpP*. We verified that levels of ImmA were higher in the *clpP*-deleted cells than in wild type. We also demonstrated that the increase in ICE*Bs1* gene expression in *clpP* cells was moderate compared with derepression following induction by DNA-damaging treatment or overproduction of *rapI*, and that expression could be fully induced in *clpP*-deleted cells by either of those ICE*Bs1*-inducing conditions, with no apparent change in the levels of ImmA. These results support previous findings that ImmA levels are not affected by RapI or by DNA damage. However, they also indicate that ImmA is likely degraded by ClpXP. In addition, we demonstrate that GFP tagged with C-terminal sequences of ImmA is less abundant in vivo than untagged GFP, perhaps indicating that the C-terminal sequence of ImmA targets it for degradation. Future work may elucidate the role of ClpXP in degradation of ImmA and regulation of ICE*Bs1* and better define how the C-terminal sequence of ImmA limits protein levels.

Introduction

ICE*Bs1* is an integrative and conjugative element found in the chromosome of *Bacillus subtilis* (Auchtung *et al.*, 2005; Burrus *et al.*, 2002). At its left end, ICE*Bs1* harbors a regulatory module that resembles those of several bacteriophages (Auchtung *et al.*, 2005; Burrus *et al.*, 2002). This module includes the gene for a phage-like repressor, *immR*. ImmR represses transcription of genes required for excision and transfer (Auchtung *et al.*, 2007). The module also encodes the conserved anti-repressor, ImmA, which cleaves ImmR in response to cues from cell-cell signaling or DNA damage (Bose* *et al.*, 2008). Inactivation of ImmR by degradation leads to derepression of ICE*Bs1* gene expression, excision, and mating (Bose* *et al.*, 2008).

Cellular levels of ImmA are normally low, but can be increased by altering ImmA's C terminus (JMA and ADG unpublished observations; Chapter 3). The higher levels of C-terminally altered ImmA could result from increased synthesis or decreased degradation. C-terminal sequences are known to target many bacterial proteins for degradation by conserved proteases (Flynn *et al.*, 2003). We tested whether ImmA's C-terminal sequence affected the levels of the protein to which it was attached, by tagging an otherwise abundant protein with ImmA's C terminus. To identify factors that might help degrade ImmA, we screened strains in which one or more components of the cellular proteolytic machinery had been deleted.

To rapidly ascertain whether ImmA levels might be affected in a given strain, we used the P*xis-lacZ* reporter instead of directly detecting ImmA protein. Transcription from the promoter for the *xis* gene (P*xis*) in ICE*Bs1* is indicative of ICE*Bs1* excision and mating (Auchtung *et al.*, 2007). P*xis* is strongly repressed, but can be derepressed in an ImmA-dependent manner by overexpression of RapI or during the RecA-mediated SOS response (Bose* *et al.*, 2008). P*xis* can also be derepressed in the absence of ICE*Bs1*-inducing signals by overproduction of ImmA

(Chapter 3). Thus, we reasoned that deleting a cellular protease or other factor involved in degradation of ImmA might lead to ImmA accumulation and consequent derepression of P*xis-lacZ*.

We found that P*xis-lacZ* expression was elevated in cells with null mutations of *clpX* or *clpP*. One or both of the *lon* proteases might affect expression as well. Quantification of ImmA protein revealed that ImmA levels were significantly higher in Δ*clpP* cells than in cells with wild type *clpP*. We also found that tagging GFP with C-terminal sequences from ImmA decreases the amount of GFP protein, supporting the model that the C-terminal residues of ImmA might target it for degradation.

Results

Some cellular proteases affect P*xis-lacZ* expression

Transcription from the promoter of the *xis* gene (P*xis*) in ICE*Bs1* is indicative of ICE*Bs1* excision and mating (Auchtung *et al.*, 2007). We used a fusion of this promoter to *lacZ* (P*xis-lacZ*) to monitor regulation of ICE*Bs1*. P*xis* is strongly repressed by ImmR (Auchtung *et al.*, 2007). Derepression requires ImmA and occurs after overexpression of RapI or during the RecA-mediated SOS response (Bose* *et al.*, 2008). This regulation occurs both in the presence and absence of all other ICE*Bs1* genes (Bose* *et al.*, 2008). For simplicity, the experiments presented here were done in cells lacking ICE*Bs1* (ICE*Bs1*^{0}).

Raising levels of ImmA derepresses P*xis* (Chapter 3). A null mutation of any protein involved in the degradation of ImmA might therefore increase transcription from P*xis*. We assayed P*xis-lacZ* expression in strains bearing one or more null mutations in components of the cellular proteolytic machinery. Strains with deletions of *clpP* and/or *clpX* also harbored a

disruption of the *spx* gene, because this disruption lessens the severity of the growth defect caused by deletion of *clpP* (Nakano *et al.*, 2001; Nakano *et al.*, 2002; Nakano *et al.*, 2002). To account for the effects of disrupting *spx*, we also tested P*xis-lacZ* in a strain with the *spx* mutation only.

The strains in Table 1 were assayed by patching onto indicator plates containing x-gal (bromo-chloro-indolyl-galactopyranoside). On these plates, patches are white if the strain expresses *immA* and *immR*, and patches are blue if the strain lacks *immR*. Most of the mutations had little or no effect on P*xis-lacZ* expression as judged by the color of the patch. Slight blue coloration was observed for strains in which *clpX* or *clpP* were deleted. Deletion of *lonA* and *lonB* appeared to increase the blue color of *clpX*-deleted colonies, but the *lonB* deletion alone seemed to have no effect. Curiously, deletion of *clpX*, *lonB*, and *lonA* had a more drastic effect on P*xis-lacZ* than did deletion of *clpX*, *lonB*, *lonA*, and *clpP*. Effects of *clpP* were further evaluated.

To assess the effect of deletion of *clpP* on P*xis-lacZ*, we assayed strains grown in liquid culture (Fig. 1). Disruption of *spx* slightly increased P*xis-lacZ* expression, relative to wild type. Deletion of *clpP* in addition to the *spx* disruption increased P*xis-lacZ* expression roughly ten-fold. In all cases, P*xis-lacZ* expression was induced by overproduction of RapI or addition of mitomycinC to induce the SOS response. It was particularly important that P*xis-lacZ* was induced to the same maximum level in wild type and *clpP*-deleted cells, because this demonstrated that induction of the system had not been dampened by one or more of the many pleiotropic effects of *clpP* deletion in *B. subtilis* (Gerth *et al.*, 1998; Msadek *et al.*, 1998).

Table 1. Strains in which P*xis-lacZ* expression was assayed

Strain	Relevant genotype
BOSE917	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} Δ*ftsH*::*spc*
BOSE918	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} Δ*clpX*::*spc*
BOSE919	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} Δ*clpYQ*::*spc*
BOSE920	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} Δ*lonB*::*neo*
BOSE921	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} Δ*clpE*::*neo*
BOSE922	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} *spx*::*neo*
BOSE923	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} Δ*ssrA*::*cat*
BOSE924	ICE*Bs1*^{0} *thrC*::{(P*xis-lacZ*Ω343) *mls*} *cgeD*::{(P*immR-immR immA*) *kan*} Δ*ftsH*::*spc*
BOSE925	ICE*Bs1*^{0} *thrC*::{(P*xis-lacZ*Ω343) *mls*} *cgeD*::{(P*immR-immR immA*) *kan*} Δ*clpC*::*tet*
BOSE926	ICE*Bs1*^{0} *thrC*::{(P*xis-lacZ*Ω343) *mls*} *cgeD*::{(P*immR-immR immA*) *kan*} Δ*ssrA*::*cat*
BOSE927	ICE*Bs1*^{0} *thrC*::{(P*xis-lacZ*Ω343) *mls*} *cgeD*::{(P*immR-immR*) *kan*} *amyE*::{(P*spank-immA*Ω218) *spc*} Δ*ssrA*::*cat*
BOSE928	ICE*Bs1*^{0} *thrC*::{(P*xis-lacZ*Ω343) *mls*} *cgeD*::{(P*immR-immR*) *kan*} *amyE*::{(P*spank-immA*Ω218) *spc*} Δ*clpC*::*tet*
BOSE929	ICE*Bs1*^{0} *thrC*::{(P*xis-lacZ*Ω343) *mls*} *cgeD*::{(P*immR-immR*) *kan*} *lacA*::{(P*spank-immA*Ω218) *tet*} Δ*ftsH*::*spc*
BOSE930	ICE*Bs1*^{0} *thrC*::{(P*xis-lacZ*Ω343) *mls*} *cgeD*::{(P*immR-immR*) *kan*} *lacA*::{(P*spank-immA*Ω218) *tet*} Δ*clpX*::*spc*
BOSE931	ICE*Bs1*^{0} *thrC*::{(P*xis-lacZ*Ω343) *mls*} *cgeD*::{(P*immR-immR*) *kan*} *lacA*::{(P*spank-immA*Ω218) *tet*} Δ*clpYQ*::*spc*
BOSE932	ICE*Bs1*^{0} *thrC*::{(P*xis-lacZ*Ω343) *mls*} *cgeD*::{(P*immR-immR*) *kan*} *lacA*::{(P*spank-immA*Ω218) *tet*} Δ*ssrA*::*cat*
BOSE942	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} Δ*lonB*::*neo* Δ*ftsH*::*spc*
BOSE943	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} Δ*lonB*::*neo* Δ*clpX*::*spc*
BOSE944	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} Δ*lonB*::*neo* Δ*clpYQ*::*spc*
BOSE945	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} Δ*clpE*::*neo* Δ*ftsH*::*spc*

BOSE946	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} Δ*clpE*::*neo* Δ*clpX*::*spc*
BOSE947	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} Δ*clpE*::*neo* Δ*clpYQ*::*spc*
BOSE964	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} *spx*::*neo* Δ*ftsH*::*spc*
BOSE965	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} *spx*::*neo* Δ*clpX*::*spc*
BOSE966	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} *spx*::*neo* Δ*clpYQ*::*spc*
BOSE1008	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} Δ*clpP*::*cat*
BOSE1019	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} *spx*::*neo* Δ*clpP*::*cat*
BOSE1020	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} *spx*::*neo* Δ*clpX*::*spc* Δ*clpP*::*cat*
BOSE1046	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} *amyE*::{(P*xyl-rapI*) *spc*} *spx*::*neo*
BOSE1059	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} *amyE*::{(P*xyl-rapI*) *spc*} *spx*::*neo* Δ*clpP*::*cat*
BOSE1088	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} *spx*::*neo* Δ(*clpX lonB lonA)*::*spc*
BOSE1089	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} *amyE*::{(P*xyl-rapI*) *cat*} *spx*::*neo* Δ(*clpX lonB lonA)*::*spc*
BOSE1090	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} *spx*::*neo* Δ(*clpX lonB lonA)*::*spc* Δ*clpP*::*cat*
BOSE1262	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*} *amyE*::{(P*xyl-rapI*) *spc*}
JMA436	ICE*Bs1*^{0} *thrC*::{(P*xis-lacZ*Ω343) *mls*} *cgeD*::{(P*immR-immR immA*) *kan*}
JMA895	ICE*Bs1*^{0} *thrC*::{P*xis-lacZ*Ω343) *mls*} *lacA*::{(P*immR-immR immA*) *tet*}

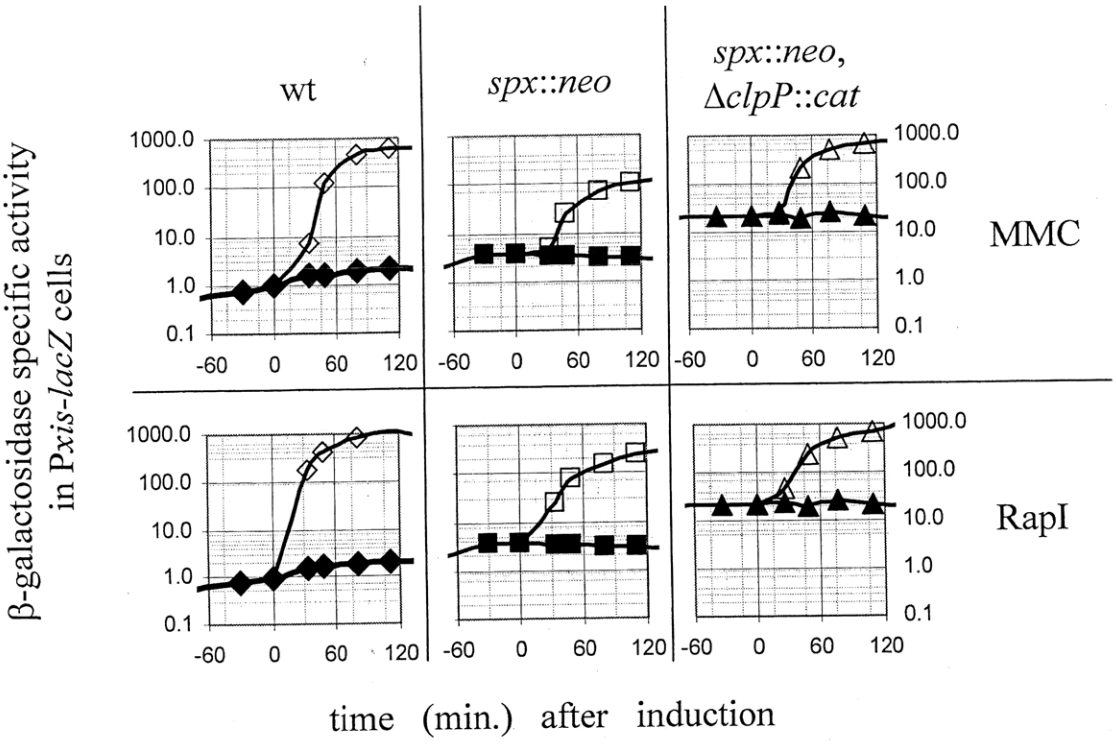


Figure 1. P*xis-lacZ* expression in Δ*clpP* cells. P*xis-lacZ* expression was monitored in cells with wild type *spx* and *clpP* (diamonds, BOSE1262), an *spx* disruption and wild type *clpP* (squares, BOSE1046), or an *spx* disruption and a *clpP* deletion (triangles, BOSE1059). Cultures were split at time 0 on each graph. Part of each culture was left untreated (closed symbols), and the other part was treated with MMC to induce the SOS response (top graphs, open symbols) or xylose to induce overproduction of RapI (bottom graphs, open symbols).

ClpP affects ImmA levels

We searched for mutants of proteolytic machinery that caused increases in P*xis-lacZ* expression, because eliminating any protein that helps to degrade ImmA should elevate P*xis-lacZ* expression. To verify that ImmA levels were elevated in a *clpP*-deleted strain, and to assess the magnitude of any change, we quantified ImmA using Western blots with antibodies agains ImmA. ImmA levels were significantly higher in *spx*::*neo* Δ*clpP*::*cat* cells than in cells with wild type *spx* and *clpP* (Fig. 2A). ImmA levels in *spx*::*neo* cells were comparable to those in cells with wild type *spx*. Taken together, these data indicate that deleting *clpP* increases ImmA levels, likely because ImmA is degraded by the ClpP protease.

We also compared ImmA levels in samples where P*xis-lacZ* was induced by MMC treatment or RapI overproduction to ImmA levels without exogenous induction. This comparison was made at a time point where P*xis-lacZ* expression had substantially risen and was continuing to climb, indicating that ImmA was active. Induction by MMC treatment or RapI overproduction did not significantly elevate ImmA levels (Fig. 2B). The most drastic change in ImmA levels was a roughly 50% decrease in *spx*::*neo* cells that were overproducing RapI, compared to *spx*::*neo* cells that were uninduced.

Tagging GFP with ImmA's C-terminal sequence lowers protein levels

C-terminally truncated or epitope-tagged variants of ImmA are more abundant in vivo than wild type ImmA (Chapter 3; JMA and ADG, unpub obs). Thus, it seems likely that some aspect of the C-terminal sequence of ImmA targets it for degradation. We investigated whether appending C-terminal sequences from ImmA to an otherwise abundant protein would lower its levels. We tested GFP tagged with 5 (GFP-ImmA165-169), 15 (GFP-ImmA155-169), or 21 (GFP-ImmA149-169) residues from the C terminus of ImmA. Genes encoding these proteins

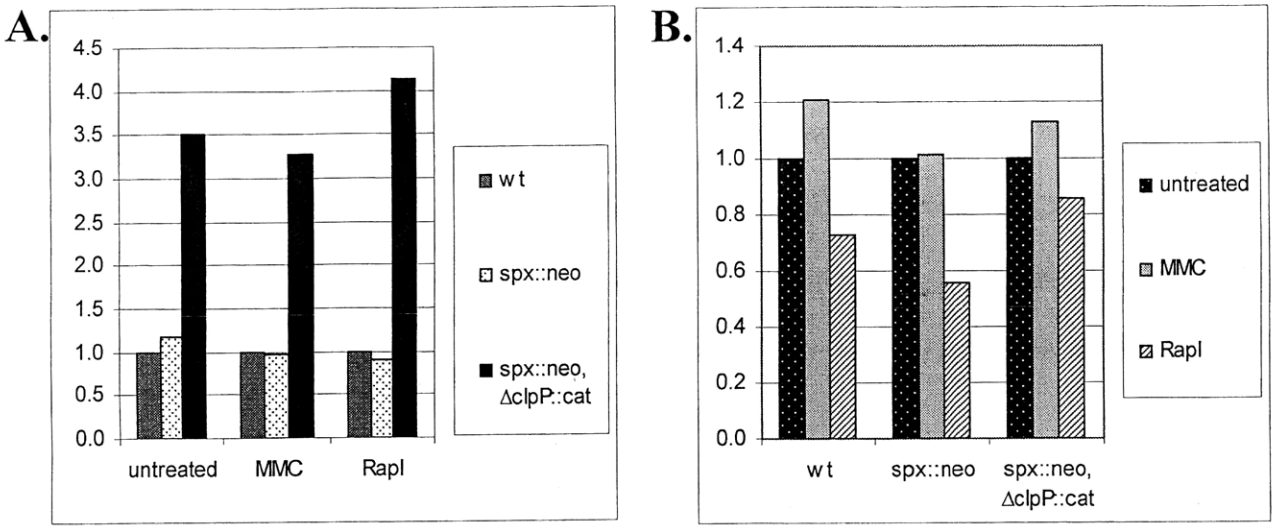


Figure 2. ImmA levels in Δ*clpP* cells. Levels of ImmA protein were determined for cells with wild type *spx* and *clpP* (BOSE1262), an *spx* disruption and wild type *clpP* (BOSE1046), or an *spx* disruption and a *clpP* deletion (BOSE1059). Cultures were grown to mid-exponential phase and split into three parts. One part was left untreated, one was treated with MMC to induce the SOS response, and one was treated with xylose to induce overproduction of RapI. Samples were taken 44-48 minutes after splitting cultures, and levels of ImmA were quantified by Western blots with antibodies against ImmA. ImmA levels in each sample were normalized to wild type for each treatment (**A.**) or to the untreated sample for each strain (**B.**)

were constitutively expressed at an ectopic locus in *B. subtilis*. We assessed protein levels using Western blots with antibodies against GFP. GFP-ImmA155-169 and GFP-ImmA149-169 were less abundant in vivo than untagged GFP (data not shown). GFP-ImmA149-169 exhibited the more dramatic decrease in protein levels and was analyzed further (Fig. 3). Repeated assays of GFP-ImmA149-169 revealed two bands; one migrates at the same position as untagged GFP, while the other is slightly larger, as would be expected for GFP-ImmA149-169. The larger of the two detected proteins is ~ one-tenth as abundant as untagged GFP, and the smaller is ~ one-fourth as abundant as untagged GFP. Further work is needed to identify the two species detected for GFP-ImmA149-169.

Discussion

We found that transcription from P*xis* was elevated in cells with null mutations of *clpX* or *clpP*, compared to cells with wild type alleles of these proteins. Deletion of both *lon* proteases exacerbated the effect of the *clpX* mutant, perhaps indicating that either or both play a role in the absence of ClpX. We expected that the increase in P*xis-lacZ* expression to result from accumulation of ImmA protein. Measurements of ImmA protein levels confirmed that ImmA was much more abundant in in Δ*clpP* cells than in cells with wild type *clpP*. In addition, GFP that was C-terminally tagged with sequences from ImmA's C terminus was less abundant in vivo than untagged GFP. This supports the idea that the C-terminal residues of ImmA play a role in limiting its levels in vivo, perhaps because ImmA's C terminus targets it for degradation by ClpXP.

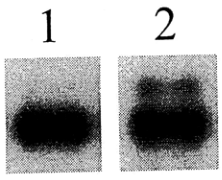


Figure 3. Cellular levels of GFP and GFP-ImmA149-169. GFP in extracts from cells expressing untagged GFP (lane 1, KG873) or GFP-ImmA149-169 (lane2, BOSE973) were detected using Western blots with antibodies against GFP. Lane 2 contains four times as much extract as lane 1.

Roles of ATP-dependent proteases in regulation of ICE*Bs1*

We found that P*xis-lacZ* expression was higher in cells with deletions of *clpX* or *clpP* than in wild type cells. P*xis-lacZ* expression in Δ(*clpX-lonB-lonA*) cells was higher than in Δ*clpX* cells, suggesting that one or both of the *lon* proteases might play a role in regulation of P*xis-lacZ* expression. However, P*xis-lacZ* expression in Δ(*clpX-lonB-lonA*) Δ*clpP* cells was lower than in Δ(*clpX-lonB-lonA*) cells, suggesting that the roles of these proteases in affecting P*xis* might be quite complex. Further investigation is needed to understand how they affect regulation of ICE*Bs1*.

Increasing levels of ImmA by overproducing it or by deleting *clpP* de-represses P*xis-lacZ*. However, no increase in ImmA levels was observed during induction of P*xis-lacZ* by overproduction of RapI (Chapter 3). Here, we assayed ImmA protein levels following addition of MMC and after inducing *rapI* expression. Only slight changes in ImmA levels were observed after both ICE*Bs1*-inducing treatments in all strains tested. This supports the idea that the SOS response and RapI stimulate ImmA-mediated cleavage of ImmR by a mechanism other than stabilizing ImmA protein. If P*xis* expression and ImmA levels are being controlled by ATP-dependent cellular proteases perhaps regulation of ICE*Bs1* by proteolysis of ImmA functions under different conditions than those that have been tested thus far.

Effect of ImmA's C-terminal sequence on protein levels

We tested the effect of tagging GFP with C-terminal sequences from ImmA and found that GFP tagged with 15 or 21 amino acids from ImmA were less abundant in vivo than untagged GFP. We observed two products for the variant that was least abundant in vivo, GFP-ImmA149-169. Further work is needed to identify these two proteins and to determine how they are

produced. More generally, additional investigation is needed to understand how the C terminus of ImmA affects its levels in vivo.

Changing the C terminus of ImmA by epitope tagging or truncation causes an increase in ImmA levels in vivo. Since ImmA levels were also affected in *clpP* mutants and perhaps in *clpX* levels, we wonder whether the C terminus of ImmA might encode a signal for degradation by ClpXP. C-terminal signals for degradation by ClpXP have been extensively characterized in *E. coli* (Flynn *et al.*, 2003). The C terminus of ImmA does not appear to resemble any of the C-terminal sequences that are recognized by *E. coli* ClpXP. However, less is known about sequences that target proteins to *B. subtilis* ClpXP, and some C-terminal sequences are known to be treated differently by the *B. subtilis* and *E. coli* proteases. If ImmA is a ClpXP substrate, characterizing the signals that target it for degradation could increase our understanding of these proteases in *B. subtilis*.

Experimental procedures

Media and growth conditions. *B. subtilis* cells were grown at 37°C with aeration in S7 minimal salts medium (Vasantha and Freese, 1980), except that 50 mM MOPS was used instead of 100} supplemented with 1% arabinose, 0.1% glutamate, 40 μg/ml tryptophan and phenylalanine, and 120 μg/ml threonine. 1% xylose was used to express *rapI* from P*xyl*. When appropriate, antibiotics were used at the following concentrations: chloramphenicol (5 μg/ml); kanamycin (5 μg/ml); tetracycline (10 μg/ml); spectinomycin (100 μg/ml); and erythromycin (0.5 μg/ml) and lincomycin (12.5 μg/ml) together to select for macrolide-lincosamide-streptogramin B (MLS) resistance. MMC (Roche) was used at a final concentrations of 1 μg/ml. X-gal in LB agar plates was 120μg/ml.

Strains and alleles. *B. subtilis* strains in which P*xis-lacZ* expression was assayed on plates are listed in Table 1. Alleles found in those strains are listed in Table 2. Standard techniques were used for cloning and strain construction (Harwood and Cutting, 1990; Sambrook and Russell, 2001). The ICE*Bs1*^{0} strain was previously described (Auchtung *et al.*, 2005). (Note: *mls* is the same as the previously used *erm*).

Δ*clpP*::*cat* was constructed by replacing *clpP*, from 62 bp upstream of its start codon to its penultimate codon, with the chloramphenicol-resistance gene from pGEM-*cat* (Youngman *et al.*, 1989). Δ(*clpX lonB lonA)*::*spc* was made by replacing the segment from 21 bp upstream of the start codon for clpX to 69 bp upstream of the stop codon for *lonA* with the spectinomycin-resistance gene from pJL73 (LeDeaux *et al.*, 1995).

Untagged *gfp* was expressed in KG873 (JH642 *thrC*::(P_c-*gfp erm*)). Construction of this allele was previously described (Griffith and Grossman, 2008). Tagged versions of *gfp* were made in the same way as the untagged construct, except that downstream primers included sequences from *immA* between the last codon in *gfp* and the stop codon.

Western Blots. Cells were pelleted by centrifugation, washed with TN Buffer (50 mM Tris 300 mM NaCl, pH8), and stored at –20°C. Cell pellets were thawed, resuspended in an appropriate volume of Buffer (10 mM Tris, 10 mM EDTA, pH 7) containing 0.1 mg/ml lysozyme and 1 mM AEBSF, and incubated at 37°C for 30 min. Pellets were resuspended in a volume of buffer proportionate to the OD600 at which they were recovered in order to normalize the concentration of proteins in the sample. SDS-Loading Buffer was then added to samples, which were heated at 100°C for 5 min. followed by centrifugation to remove insoluble material. Proteins were separated by SDS-PAGE on 15% or 18% gels and transferred to PolyScreen PVDF membrane (Perkin Elmer) using the Trans-blot semi-dry electro-blot transfer apparatus

(BioRad). Membranes were blocked in 0.2% I-Block (Tropix) in TBST (50 mM Tris 200 mM NaCl 0.05% Tween-20 pH 8) either at room temperature for 1 hr. or overnight at 4°C. Membranes were incubated in 1:5,000 anti-ImmA or 1:10,000 anti-GFP rabbit polyclonal antisera in 0.2% I-Block TBST for 1 hr. at room temperature. They were then washed several times in TBST, incubated in 1:3,000 goat anti-rabbit IgG-HRP conjugate (BioRad) for 1 hour at room temperature, and washed several times in TBST. Signals were detected using Western Lightning chemiluminescence reagents (Perkin-Elmer) followed by exposure to Kodak Biomax Light film.

β-galactosidase assays. β-galactosidase specific activity was assayed as described (Jaacks *et al.*, 1989). Specific activity was calculated relative to the optical density at 600 nm of the samples. Results shown are from a single experiment and are representative of results obtained in at least two independent experiments.

Table 2. Alleles used to construct strains for this study

Allele	Reference or source
amyE::{(P*spank-immA*Ω218) *spc*}	(Bose* *et al.*, 2008)
amyE::{(P*xyl-rapI*) *spc*}	(Bose* *et al.*, 2008)
cgeD::{(P*immR-immR*) *kan*}	(Auchtung *et al.*, 2007)
lacA::{(P*immR-immR immA*) *tet*}	(Bose* *et al.*, 2008)
thrC::{P*xis-lacZ*Ω343) *mls*}	(Auchtung *et al.*, 2007)
Δ*clpC*::*tet*	(Ruvolo *et al.*, 2006)
Δ*clpE*::*neo*	(Ruvolo *et al.*, 2006)
Δ*clpX*::*spc*	(Liu *et al.*, 1999)
Δ(*clpX lonB lonA*)::*spc*	this work
Δ*clpP*::*cat*	this work
Δ*clpYQ*::*spc*	(Ruvolo *et al.*, 2006)
Δ*ftsH*::*spc*	(Lysenko *et al.*, 1997)
Δ*lonB*::*neo*	(Ruvolo *et al.*, 2006)
spx::*neo*	(Nakano *et al.*, 2001)
Δ*ssrA*::*cat*	(Wiegert and Schumann, 2001)

References

Auchtung, J.M., Lee, C.A., Monson, R.E., Lehman, A.P., and Grossman, A.D. (2005) Regulation of a Bacillus subtilis mobile genetic element by intercellular signaling and the global DNA damage response. *Proc Natl Acad Sci U S A* **102**: 12554-12559.

Auchtung, J.M., Lee, C.A., Garrison, K.L., and Grossman, A.D. (2007) Identification and characterization of the immunity repressor (ImmR) that controls the mobile genetic element ICEBs1 of Bacillus subtilis. *Mol Microbiol* **64**: 1515-1528.

Bose*, B., Auchtung*, J.M., Lee, C.A., and Grossman, A.D. (2008) A conserved anti-repressor controls horizontal gene transfer by proteolysis. *Mol Microbiol* **70**: 570-582. * co-first authors.

Burrus, V., Pavlovic, G., Decaris, B., and Guedon, G. (2002) The ICESt1 element of Streptococcus thermophilus belongs to a large family of integrative and conjugative elements that exchange modules and change their specificity of integration. *Plasmid* **48**: 77-97.

Flynn, J.M., Neher, S.B., Kim, Y.I., Sauer, R.T., and Baker, T.A. (2003) Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol Cell* **11**: 671-683.

Gerth, U., Kruger, E., Derre, I., Msadek, T., and Hecker, M. (1998) Stress induction of the Bacillus subtilis clpP gene encoding a homologue of the proteolytic component of the Clp protease and the involvement of ClpP and ClpX in stress tolerance. *Mol Microbiol* **28**: 787-802.

Griffith, K.L., and Grossman, A.D. (2008) Inducible protein degradation in Bacillus subtilis using heterologous peptide tags and adaptor proteins to target substrates to the protease ClpXP. *Mol Microbiol* **70**: 1012-1025.

Harwood, C.R., and Cutting, S.M. (1990) *Molecular biological methods for Bacillus.* Chichester; New York: Wiley.

Jaacks, K.J., Healy, J., Losick, R., and Grossman, A.D. (1989) Identification and characterization of genes controlled by the sporulation-regulatory gene spo0H in Bacillus subtilis. *J Bacteriol* **171**: 4121-4129.

LeDeaux, J.R., Yu, N., and Grossman, A.D. (1995) Different roles for KinA, KinB, and KinC in the initiation of sporulation in Bacillus subtilis. *J Bacteriol* **177**: 861-863.

Liu, J., Cosby, W.M., and Zuber, P. (1999) Role of lon and ClpX in the post-translational regulation of a sigma subunit of RNA polymerase required for cellular differentiation in Bacillus subtilis. *Mol Microbiol* **33**: 415-428.

Lysenko, E., Ogura, T., and Cutting, S.M. (1997) Characterization of the ftsH gene of Bacillus subtilis. *Microbiology* **143 (Pt 3)**: 971-978.

Msadek, T., Dartois, V., Kunst, F., Herbaud, M.L., Denizot, F., and Rapoport, G. (1998) ClpP of Bacillus subtilis is required for competence development, motility, degradative enzyme synthesis, growth at high temperature and sporulation. *Mol Microbiol* **27**: 899-914.

Nakano, M.M., Hajarizadeh, F., Zhu, Y., and Zuber, P. (2001) Loss-of-function mutations in yjbD result in ClpX- and ClpP-independent competence development of Bacillus subtilis. *Mol Microbiol* **42**: 383-394.

Nakano, M.M., Nakano, S., and Zuber, P. (2002) Spx (YjbD), a negative effector of competence in Bacillus subtilis, enhances ClpC-MecA-ComK interaction. *Mol Microbiol* **44**: 1341-1349.

Nakano, S., Zheng, G., Nakano, M.M., and Zuber, P. (2002) Multiple pathways of Spx (YjbD) proteolysis in Bacillus subtilis. *J Bacteriol* **184**: 3664-3670.

Ruvolo, M.V., Mach, K.E., and Burkholder, W.F. (2006) Proteolysis of the replication checkpoint protein Sda is necessary for the efficient initiation of sporulation after transient replication stress in Bacillus subtilis. *Mol Microbiol* **60**: 1490-1508.

Sambrook, J., and Russell, D.W. (2001) *Molecular cloning: a laboratory manual.* Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Vasantha, N., and Freese, E. (1980) Enzyme changes during Bacillus subtilis sporulation caused by deprivation of guanine nucleotides. *J Bacteriol* **144**: 1119-1125.

Wiegert, T., and Schumann, W. (2001) SsrA-mediated tagging in Bacillus subtilis. *J Bacteriol* **183**: 3885-3889.

Youngman, P., Poth, H., Green, B., York, K., Olmedo, G., and Smith, K. (1989) Methods for Genetic Manipulation, Cloning, and Functional Analysis of Sporulation Genes in *Bacillus subtilis*. In *Regulation of procaryotic development: structural and functional analysis of bacterial sporulation and germination.* Smith, I., Slepecky, R. and Setlow, P. (eds). Washington, D.C.: American Society for Microbiology, pp. 65-87.

Appendix D: Mutations near the cleavage site in ImmR attenuate induction of ICE*Bs1* gene expression

Abstract

The mobile genetic element ICE*Bs1* is an integrative and conjugative element (a conjugative transposon) found in the *Bacillus subtilis* chromosome. The RecA-dependent SOS response and the RapI-PhrI sensory system activate ICE*Bs1* gene expression, excision, and mating by stimulating cleavage of the element's repressor ImmR by the anti-repressor ImmA. ImmR is predicted to contain an N-terminal helix-turn-helix domain. The single site at which ImmA cleaves ImmR is C-terminal to this domain, in a portion of ImmR that has no predicted architecture. In this work, we isolated and characterized mutants of *immR* (*immR(ind-)*) that attenuate induction of ICE*Bs1* gene expression under the normally inducing conditions of treatment with DNA damaging reagent and overproduction of RapI. The mutations in *immR* lessen or prevent cleavage of ImmR protein in vivo. All four *immR(ind-)* mutations fall within a stretch of 10 residues flanking the cleavage site, emphasizing the importance of this sequence for ImmR proteolysis and ICE*Bs1* induction. To further characterize the C-terminal portion of ImmR, we demonstrated that it interacts with ImmA and with itself in yeast two-hybrid assays, indicating that this part of the protein likely functions in ImmR oligomerization and recognition of ImmR by ImmA.

Introduction

ICE*Bs1* is an integrative and conjugative element found in the chromosome of *Bacillus subtilis* (Auchtung *et al.*, 2005; Burrus *et al.*, 2002). At its left end, ICE*Bs1* contains a regulatory module that resembles those of several bacteriophages (Auchtung *et al.*, 2005; Burrus *et al.*, 2002). This module includes the gene for a phage-like repressor, *immR*. ImmR represses transcription of genes required for excision and transfer of ICE*Bs1*, by repressing transcription from the promoter for the *xis* gene (P*xis*) (Auchtung *et al.*, 2007). Repression of P*xis* ceases when ImmR is inactivated by degradation (Bose* *et al.*, 2008). This occurs in response to signals that result from DNA damage or intercellular communication (Auchtung *et al.*, 2005; Bose* *et al.*, 2008).

ICE*Bs1* resembles many other mobile genetic elements in that it is induced in a RecA-dependent manner in response to DNA damage and involves the proteolysis of a CI-like repressor (Auchtung *et al.*, 2005; Bose* *et al.*, 2008; Roberts and Devoret, 1983). However, unlike paradigmatic phage repressors, inactivation of the ICE*Bs1* repressor ImmR additionally requires the element-encoded anti-repressor ImmA (Bose* *et al.*, 2008). ImmR resembles lambda repressor in that it has an N-terminal helix-turn-helix (HTH) domain that is thought to bind DNA (Auchtung *et al.*, 2007). However, ImmR differs from lambda CI in that it lacks CI's conserved catalytic residues and cleavage site, which are found in CI's C-terminal domain and the linker between the N- and C-terminal domains, respectively (Bose* *et al.*, 2008; Little, 1984; Pabo *et al.*, 1979; Roberts and Devoret, 1983). The C-terminal domain of CI mediates dimerization in addition to catalyzing autoproteolysis (Pabo *et al.*, 1979; Roberts and Devoret, 1983). The structure and function of ImmR's C terminus have not been investigated. The anti-

repressor ImmA cleaves ImmR at a single site, which is situated in the C-terminal portion of the protein (Fig. 1) (Bose* *et al.*, 2008).

To learn more about ImmR, we isolated and characterized mutations in *immR* (*immR(ind-)*) that prevent or abrogate induction of ICE*Bs1*. We used transcription of a P*xis-lacZ* fusion as an indicator of ICE*Bs1* gene expression. We identified four mutations in *immR*, flanking the cleavage site by no more than five residues on either side, that lower or abolish induction of P*xis-lacZ* in response to RapI overproduction or DNA damage induced by MMC treatment. To get a better idea of the role of different parts of ImmR, we also assayed the ability of truncated versions of ImmR to interact with ImmA and with themselves. We found that a C-terminal portion of ImmR could interact with ImmA and with itself in yeast-two hybrid assays.

Results

Isolation of ImmR mutants that attenuate induction of transcription from P*xis*

To better understand the ICE*Bs1* repressor, ImmR, we isolated and characterized mutations in *immR*, *immR(ind-)*, that prevent or lessen induction of P*xis-lacZ* expression in cells that are overproducing RapI. ImmR mutants were screened in a strain that was cured of ICE*Bs1*, contained P*xis-lacZ*, expressed *immA* from the IPTG-inducible promoter P*spank*, and had a P*xyl-rapI* fusion. Under growth conditions that allow expression of *immA* and *rapI*, ImmA-mediated cleavage of ImmR is stimulated, and P*xis-lacZ* is de-repressed. Thus, cells containing wild type *immR* or an *immR* null mutation form blue colonies on indicator plates containing x-gal (bromo-chloro-indolyl-galactopyranoside), 1mM IPTG (to induce P*spank-immA*), and 1% xylose (to induce P*xyl-rapI*), indicating de-repression or no repression of P*xis-lacZ*. We introduced PCR-mutagenized P*immR-immR* into these cells and screened for mutants that formed white or light

blue colonies, indicating at least partially abrogated induction of P*xis-lacZ*. The *immR* alleles from candidate mutants were backcrossed and cells retested to be sure the mutant phenotype was linked to *immR*. Mutant *immR* alleles were then sequenced. Since many alleles contained multiple mutations, we reconstructed several single mutations by site-directed mutagenesis of *immR* and tested each for effects on expression of P*xis-lacZ*. Each mutation is designated by the amino acid in wild type ImmR, its position in the sequence, and the amino acid to which it was changed. Single mutations D91E, D91G, L93H, N97S, E100G were remade (Fig. 1) and tested for effects on expression of P*xis-lacZ*. D91E had no significant effect on expression of P*xis-lacZ* (data not shown) and was not characterized further. Some candidate *immR(ind-)* mutations have yet to be remade and tested. These mutations are K13E, N26S, D65V, S68G, T79A, N97D, D98G, E103V, and F119L.

Four mutations in *immR*, D91G, L93H, N97S, and E100G, caused decreased expression of P*xis-lacZ* on x-gal, IPTG, xylose indicator plates, indicating that these mutations somehow caused ImmR to be resistant to antagonism. The four mutations flank ImmR's cleavage site, between F95 and M96, by fewer than five residues on each side.

Effects of ImmR(ind-) mutants on expression of P*xis-lacZ*

The effect of each ImmR(ind-) mutant on P*xis-lacZ* induction was determined by measuring β-galactosidase activity in strains with different alleles of *immR*. Cultures were grown to mid-exponential phase and split. One part was treated with mitomycinC (MMC) to induce the SOS response, and the other was left untreated. In untreated cells, all four of the strains carrying *immR(ind-)* alleles exhibited lower levels of P*xis-lacZ* expression than the strain carrying wild type ImmA (Fig. 2). Following treatment with MMC, β-galactosidase activity in cells with wild type ImmR rose sharply, increasing more than 100-fold to ~ 400 units within 2 hours. In

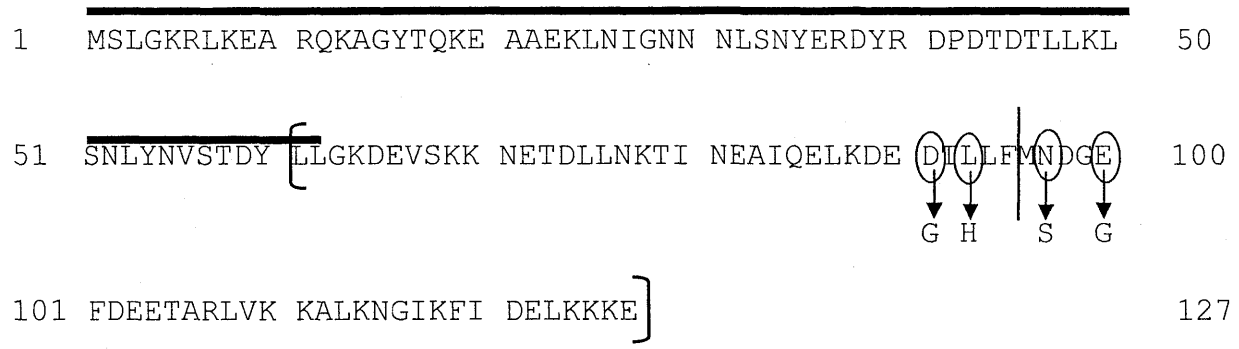


Figure 1. ImmR sequence. The full amino acid sequence of ImmR is shown. Circles indicate residues that can be mutated singly to produce ImmR(ind-) mutants. D91G and L93H are uninducible, while N97S and E100G are partially inducible in *B. subtilis*. The vertical line indicates the single site of cleavage by ImmA. The horizontal line indicates the predicted helix-turn-helix domain. The bracketed portion interacts with itself and with ImmA in yeast-two-hybrid assays.

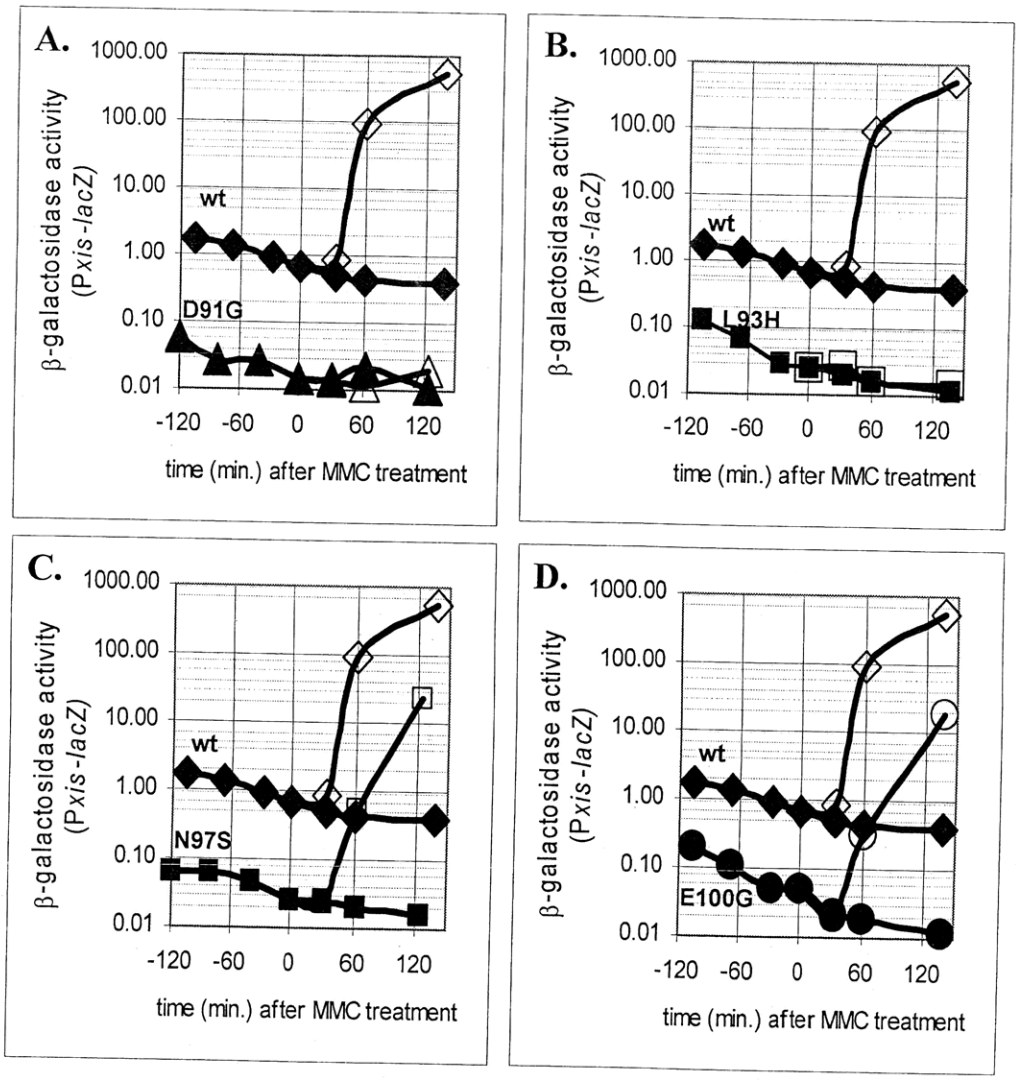


Figure 2. Effects of ImmR(ind-) mutants on P*xis-lacZ* expression. Expression of P*xis-lacZ* was monitored in *B. subtilis* strains with different alleles of *immR*. Cultures were grown to mid-exponential phase and split. Half of each culture was treated to have a final concentration of 1μg/ml MMC (open symbols), and the other half was left untreated (closed symbols).β-galactosidase specific activity is plotted relative to the time each culture was split. Values for the strain with wild type *immR* (diamonds, BOSE534) are plotted in all panels.

A. D91G, triangles, BOSE719

B. L93H, squares, BOSE714

C. N97S, squares, BOSE720

D. E100G, circles, BOSE716

contrast, cells with D91G or L93H mutants of ImmR displayed no de-repression of P*xis-lacZ* following MMC treatment, and cells with N97S or E100G mutants of ImmR showed partial de-repression of P*xis-lacZ* to ~20 units within 2 hours. The strains bearing the N97S and E100G mutants actually exhibited a larger fold-change in β-galactosidase activity than the strain with wild type ImmR, but had lower absolute levels because they started from a lower baseline. The fold-change of β-galactosidase activity in cells with wild type ImmR may have been limited as transcription approached its maximum level from P*xis*.

ImmR(ind-) mutants are cleaved less efficiently than wild type ImmR in vivo

Under ICE*Bs1*-inducing conditions, ImmR is degraded (Bose* *et al.*, 2008). We used Western blots with antibodies against ImmR to visualize degradation of wild type ImmR and the four *ind-* mutants following MMC treatment. For wild type ImmR, following MMC treatment, a smaller fragment appears and the overall level of protein decreases (Fig. 3). For the D91G and L93H alleles, no difference in ImmR was observed following induction with MMC. For the N97S and E100G alleles, a cleavage fragment is observed, but the protein level appears not to decrease as much as it does for wild type ImmR. These findings agree well with the behavior of the P*xis-lacZ* reporter in each of these strains, with the amount of repressor having an inverse relationship to transcription from the reporter.

The C terminus of ImmR interacts with itself and with ImmA

To better understand ImmR, we explored the possibility that its C-terminal portion was involved in oligomerization. The N-terminal part of ImmR contains a predicted helix-turn-helix motif that likely binds operator sites (Fig. 1). No conserved domains were found in the C-terminal part of ImmR. Previously, ImmR was found to interact with ImmA and with itself in yeast two-hybrid assays (Bose* *et al.*, 2008). We used the same type of assay to test the ability of

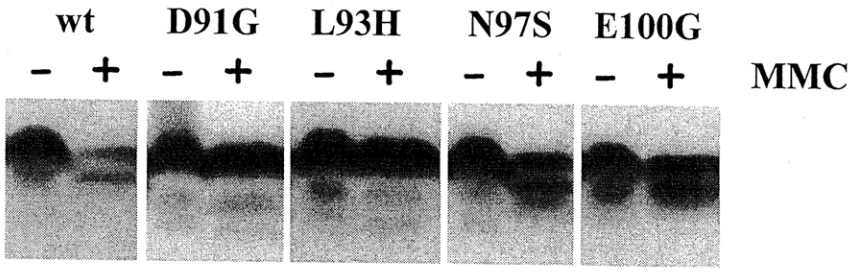


Figure 3. Proteolysis of ImmR(ind-) mutants in vivo. Samples from cells with wt or *ind-* alleles of *immR* were analyzed using Western blots with antibodies against ImmR. Samples were taken from cultures that had been grown to mid-exponential phase, split, and treated with MMC (+) or left untreated (-).The allele of *immR* is indicated above each pair of samples. wt, BOSE534; D91G, BOSE719; L93H, BOSE714; N97S, BOSE720; E100G, BOSE716.

N-terminal (ImmR1-68) and C-terminal (ImmR59-127) fragments of ImmR to interact with themselves and with ImmA. We found that ImmR59-127 could interact with ImmA and with itself. No interactions were observed for the N-terminal fragment, as expected.

Discussion

To learn more about ImmR, we isolated and characterized mutations in *immR* (*immR(ind-)*) that prevent or abrogate induction of ICE*Bs1*. We identified four mutations in *immR* that lower or abolish induction of P*xis-lacZ* in response to RapI overproduction or DNA damage induced by MMC treatment. Each mutation is no more than five amino acids from either side of the site of cleavage by ImmA. To explore the roles of different parts of ImmR, we also assayed the ability of truncated versions of ImmR to interact with ImmA and with themselves. We found that a C-terminal portion of ImmR could interact with ImmA and with itself in yeast-two hybrid assays.

The hunt for *immR(ind-)* mutants was not exhaustive. The screen could be continued, and the remaining isolated mutations could be reconstructed and tested. It would be interesting to know how many *ind-* mutations map to within a few residues of the cleavage site. Since the four mutations that we isolated fall within a stretch of 10 residues around the cleavage site, perhaps only a small sequence is required for ImmA to recognize and cut ImmR. Alternatively, screening for addtional mutants might highlight other parts of the protein that are important for cleavage by ImmA, or that prevent induction of P*xis* in some other way.

Even in the absence of ICE*Bs1*-inducing signals (MMC or RapI overproduction), strains containing ImmR(ind-) mutants had lower levels of P*xis-lacZ* expression than the strain with wild type ImmR (Fig. 2). This is consistent with previous results that cells with wild type *recA* and P*xyl-rapI* have elevated P*xis-lacZ* expression compared to cells lacking *recA* and P*xyl-rapI*,

even when the cells are not treated so as to activate RecA or induce expression of *rapI* (Chapter 3). Leaky expression of *rapI* and low levels of activated RecA could cause this difference. If the ImmR(ind-) mutants are resistant to cleavage resulting from the action of RapI and RecA, P*xis-lacZ* expression could drop to that in cells lacking *recA* and P*xyl-rapI*. This hypothesis could be tested by directly comparing these strains.

We obtained some general information about the C-terminal part of ImmR by showing that it binds itself and ImmA in yeast two-hybrid assays. Additional experiments are needed to verify these interactions and test their significance. Future work could focus on the architecture and function of this poorly understood section of ImmR.

Experimental procedures

Media and growth conditions. *B. subtilis* cells were grown at 37°C with aeration in S7 minimal salts medium (Vasantha and Freese, 1980), except that 50 mM MOPS was used instead of 100} supplemented with 1% arabinose, 0.1% glutamate, 40 μg/ml tryptophan and phenylalanine, and 120 μg/ml threonine. 1% xylose was used to express *rapI* from P*xyl*. When appropriate, antibiotics were used at the following concentrations: kanamycin (5 μg/ml); tetracycline (10 μg/ml); spectinomycin (100 μg/ml); and erythromycin (0.5 μg/ml) and lincomycin (12.5 μg/ml) together to select for macrolide-lincosamide-streptogramin B (MLS) resistance. MMC (Roche) was used at a final concentrations of 1 μg/ml. X-gal in LB agar plates was 120μg/ml.

S. cerevisiae cells were grown at 30°C in yeast peptone dextrose medium or synthetic complete medium lacking uracil and leucine or lacking uracil, leucine, and adenine (Sambrook and Russell, 2001).

Strains and alleles. *B. subtilis* strains used in this study are listed in Table 1. Standard techniques were used for cloning and strain construction (Harwood and Cutting, 1990; Sambrook and Russell, 2001). The ICE*Bs1*^{0} strain, and the *cgeD*::{(P*immR-immR*) *kan*}, *thrC*::{(P*xis-lacZ*Ω343) *mls*}, and the *amyE*::{(P*spank-immA*Ω218) *spc*} alleles were previously described (Auchtung *et al.*, 2005; Auchtung *et al.*, 2007; Bose* *et al.*, 2008). (Note: *mls* is the same as the previously used *erm*). The *lacA*::[(P*xyl-rapI*) *tet*] allele is described in Chapter 3.

ImmR(ind-) mutant hunt

P*immR-immR* was amplified from JH642 DNA by mutagenic PCR with $MnCl_2$. Products were digested by restriction enzymes and ligated into the vector pMMB124 (Auchtung *et al.*, 2007). Ligation mixtures were used to transform competent *E. coli* DH5α cells. Transformants were selected by plating on LB agar containing ampicillin. Many colonies from a single PCR reaction were pooled, and plasmid DNA was miniprepped from each pool. Plasmid DNA was then used to transform BOSE535, and transformants were grown on LB plates containing kanamycin (to select transformants), IPTG (to induce P*spank-immA*), xylose (to induce P*xyl-rapI*), and X-gal (to visualize P*xis-lacZ* expression). On these plates, colonies with wild type *immR* or no *immR* appear blue. White and light blue colonies were picked and re-streaked to purify single colonies. Each candidate colony was grown in liquid LB, and genomic DNA was obtained from these cells by phenol-chloroform extraction.

The DNA from each candidate was back-crossed to verify linkage of the white- or light blue-color phenotype with the *kan* marker associated with P*immR-immR*, as follows. DNA was used to transform BOSE535, and cells were plated on LB agar containing kanamycin, IPTG, xylose, and X-gal. If genomic DNA from a candidate produced white or light blue transformants,

P*immR-immR* was amplified from this DNA by high fidelity PCR using Platinum Taq DNA polymerase (Invitrogen), and the PCR products were sequenced (MIT Biopolymers lab).

Coding mutations identified in the candidates were reconstructed as follows. The sequence of *immR* in pJMA404 (pMMB124-P*immR-immR*), the plasmid used to construct *cgeD*::{(P*immR-immR*) *kan*}(Auchtung *et al.*, 2007) was altered by Quikchange site-directed mutagenesis (Stratagene), and the resulting plasmids were introduced into JH642 by double crossovers at *cgeD*. Genomic DNA from these strains was used to introduce each *immR(ind-)* allele into indicator strains (e.g. BOSE535).

Western Blots. Cells were pelleted by centrifugation, washed with TN Buffer (50 mM Tris 300 mM NaCl, pH8), and stored at –20°C. Cell pellets were thawed, resuspended in an appropriate volume of Buffer (10 mM Tris, 10 mM EDTA, pH 7) containing 0.1 mg/ml lysozyme and 1 mM AEBSF, and incubated at 37°C for 30 min. Pellets were resuspended in a volume of buffer proportionate to the OD600 at which they were recovered in order to normalize the concentration of proteins in the sample. SDS-Loading Buffer was then added to samples, which were heated at 100°C for 5 min. followed by centrifugation to remove insoluble material.

Proteins were separated by SDS-PAGE on 15% or 18% gels and transferred to PolyScreen PVDF membrane (Perkin Elmer) using the Trans-blot semi-dry electro-blot transfer apparatus (BioRad). Membranes were blocked in 0.2% I-Block (Tropix) in TBST (50 mM Tris 200 mM NaCl 0.05% Tween-20 pH 8) either at room temperature for 1 hr. or overnight at 4°C. Membranes were incubated in 1:10,000 anti-ImmR rabbit polyclonal antisera in 0.2% I-Block TBST for 1 hr. at room temperature, washed several times in TBST, incubated in 1:3,000 goat anti-rabbit IgG-HRP conjugate (BioRad) for 1 hour at room temperature, and washed several

times in TBST. Signals were detected using Western Lightning chemiluminescence reagents (Perkin-Elmer) followed by exposure to Kodak Biomax Light film.

β-galactosidase assays. β-galactosidase specific activity was assayed as described (Jaacks *et al.*, 1989). Specific activity was calculated relative to the optical density at 600 nm of the samples. Results shown are from a single experiment.

Yeast two-hybrid assays. Plasmids encoding ImmR1-68 and ImmR59-127 fused to the Gal4 DNA binding domain (Gal4-BD) or Gal4 activation domain (Gal4-AD) were generated by cloning the coding sequence of *immR* or *immA* in the same reading frame as the upstream Gal4-AD coding sequence in plasmid pGAD-c1 (James *et al.*, 1996) or the upstream Gal4-BD coding sequence in plasmid pGBDu-c3 (James *et al.*, 1996). Construction of a plasmid encoding ImmA fused to Gal4-BD was previously described (Bose* *et al.*, 2008). Plasmids containing the fusions were transformed into *S. cerevisae* strain PJ69-4A (*trp-901 leu2-3 ura3-52 his3-200 gal4Δ gal80Δ LYS2*::(*GALl-HIS3*) *GAL2-ADE2 met2*::(*GAL7-lacZ*) (James *et al.*, 1996) selecting for growth on synthetic complete medium lacking uracil and leucine. 6-10 transformants were purified to single colonies on medium without uracil and leucine, then re-purified and tested on the same medium without and with adenine. A pair of proteins were designated as interacting if more than half the transformants grew on medium lacking adenine. Similar results were observed in each of two independent transformation experiments.

Table 1. *B. subtilis* strains

BOSE534	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR*) *kan*] *amyE*::[(P*spank-immA*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE535	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *amyE*::[(P*spank-immA*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE714	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR L93H*) *kan*] *amyE*::[(P*spank-immA*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE716	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR E100G*) *kan*] *amyE*::[(P*spank-immA*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE719	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR D91G*) *kan*] *amyE*::[(P*spank-immA*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]
BOSE720	ICE*Bs1*^{0} *thrC*::[(P*xis-lacZ*Ω343) *mls*] *cgeD*::[(P*immR-immR N97S*) *kan*] *amyE*::[(P*spank-immA*Ω218) spc] *lacA*::[(P*xyl-rapI*) *tet*]

References

Auchtung, J.M., Lee, C.A., Monson, R.E., Lehman, A.P., and Grossman, A.D. (2005) Regulation of a Bacillus subtilis mobile genetic element by intercellular signaling and the global DNA damage response. *Proc Natl Acad Sci U S A* 102: 12554-12559.

Auchtung, J.M., Lee, C.A., Garrison, K.L., and Grossman, A.D. (2007) Identification and characterization of the immunity repressor (ImmR) that controls the mobile genetic element ICEBs1 of Bacillus subtilis. *Mol Microbiol* 64: 1515-1528.

Bose*, B., Auchtung*, J.M., Lee, C.A., and Grossman, A.D. (2008) A conserved anti-repressor controls horizontal gene transfer by proteolysis. *Mol Microbiol* 70: 570-582. * co-first authors.

Burrus, V., Pavlovic, G., Decaris, B., and Guedon, G. (2002) The ICESt1 element of Streptococcus thermophilus belongs to a large family of integrative and conjugative elements that exchange modules and change their specificity of integration. *Plasmid* 48: 77-97.

Harwood, C.R., and Cutting, S.M. (1990) *Molecular biological methods for Bacillus.* Chichester; New York: Wiley.

Jaacks, K.J., Healy, J., Losick, R., and Grossman, A.D. (1989) Identification and characterization of genes controlled by the sporulation-regulatory gene spo0H in Bacillus subtilis. *J Bacteriol* 171: 4121-4129.

James, P., Halladay, J., and Craig, E.A. (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* 144: 1425-1436.

Little, J.W. (1984) Autodigestion of lexA and phage lambda repressors. *Proc Natl Acad Sci U S A* 81: 1375-1379.

Pabo, C.O., Sauer, R.T., Sturtevant, J.M., and Ptashne, M. (1979) The lambda repressor contains two domains. *Proc Natl Acad Sci U S A* 76: 1608-1612.

Roberts, J.W., and Devoret, R. (1983) Lysogenic induction. In *Lambda II.* Hendrix, R.W. (ed). Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory, pp. 123-144.

Sambrook, J., and Russell, D.W. (2001) *Molecular cloning: a laboratory manual.* Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Vasantha, N., and Freese, E. (1980) Enzyme changes during Bacillus subtilis sporulation caused by deprivation of guanine nucleotides. *J Bacteriol* 144: 1119-1125.

Chapter 4: Discussion

Summary

In response to RapI or the SOS response, ICE*Bs1* gene expression, excision, and mating are induced (Auchtung *et al.*, 2005). The ICE*Bs1*-inducing signals somehow activate cleavage of the repressor ImmR by the anti-repressor ImmA (Auchtung *et al.*, 2007; Bose* *et al.*, 2008). Destruction of ImmR allows transcription from the *xis* promoter, which controls expression of many genes encoding products involved in ICE*Bs1* mating (Auchtung *et al.*, 2007; Bose* *et al.*, 2008).

In this thesis, I demonstrated that ImmA cuts ImmR at a single site and mapped this cleavage site. I have shown that cleavage of ImmR by ImmA is metal-dependent, supporting the idea that ImmA could be a Zinc-dependent protease (Appendix A). I isolated four mutants of ImmR that attenuate induction of P*xis* in response to ICE*Bs1*-inducing signals. The single amino acid substitution in each of these four mutants flanks the cleavage site by no more than five residues on either side. The mutations abrogate or abolish cleavage of ImmR by ImmA in vivo.

Jennifer Auchtung had found that C-terminally tagged variants of ImmA were more abundant than wild type ImmA in vivo, indicating that ImmA might be unstable (JMA and ADG, unpublished observations). I explored whether RapI might activate ImmA by increasing its levels, possibly by increasing ImmA stability. I found that P*xis* could be derepressed by overexpression of *immA*, when this resulted in significant increases in ImmA protein levels. However, I also found that during induction of transcription from P*xis* by RapI, ImmA levels did not significantly increase.

To better understand ImmA, I isolated and characterized hyperactive mutants of ImmA (ImmA^{h}). I found two classes of ImmA^{h} mutants. Those with missense mutations

near the center of ImmA's sequence cleaved ImmR more efficiently in vitro than did wild type ImmA, but were approximately as abundant as wild type ImmA in vivo. In contrast, $ImmA^h$ mutants that were C-terminally truncated were significantly more abundant than wild type ImmA in vivo, but cleaved ImmR in vitro with the same efficiency as wild type ImmA.

Although activation of ImmA by RapI or activated RecA (RecA*) does not seem to involve stabilization of ImmA, the C terminus of ImmA might target it for degradation by host cell proteases. I screened strains bearing one or more null mutations of energy-dependent proteases for the increase in P*xis-lacZ* expression that should result from accumulation of ImmA. I found that deletion of *clpX*, *clpP*, *lonA*, or *lonB* might stabilize ImmA. I also found that appending residues from the C terminus of ImmA to the C terminus of the otherwise abundant protein GFP caused a significant decrease in its levels.

The findings in this thesis have increased our understanding of the regulatory proteins of ICE*Bs1* and revealed or clarified some interesting potential directions for future research, discussed below.

Activation of ImmA by RapI and RecA

How RapI and RecA activate cleavage of ImmR remains unknown. RapI and RecA could affect ICE*Bs1* by different mechanisms. Three broad ways in which they could work are (1) by activating ImmA to cleave ImmR, (2) by making ImmR a better substrate for cleavage, or (3) by bringing ImmA and ImmR together.

In addition, their effect could be direct or indirect. Because RapI can activate ImmA-dependent cleavage of ImmR in *E. coli*, the effect of RapI is likely direct. Alternatively, it

is possible that some other factor is required for RapI's effect, and this factor is conserved between *B. subtilis* and *E. coli*. Current work is aimed at determining whether cleavage of ImmR can be stimulated by the SOS response in *E. coli* by *E. coli* RecA and/or *B. subtilis* RecA.

Thus far, only yeast-two hybrid assays have been attempted to investigate whether RapI and RecA directly interact with ImmA. Of the ImmA variants tested (wild type, R85G, V92E, N93D, I165*), only the C-terminally truncated ImmA^{h} mutant I165* gave a positive yeast two-hybrid result (BB and ADG, unpublished observations). It interacted with both RecA and RapI in these tests. Further work is needed to investigate the possible interactions of RapI and RecA with ImmA. Such interactions may stimulate the ability of ImmA to cleave ImmR.

We have explored one possible way in which RapI could stimulate cleavage of ImmR- by stabilizing ImmA against degradation. Our findings do not support this model. Several possible mechanisms remain. RapI could promote proper folding of ImmA and keep it from forming insoluble aggregates. Or, RapI could change ImmA's localization, bringing it into contact with ImmR. RapI could covalently modify ImmA, or RapI could alter ImmA's conformation by binding to it. Currently, attempts are being made to purify RapI and RecA and to see whether they can enhance cleavage of ImmR in vitro.

ImmA, the protease

ImmA was predicted to be a metalloprotease, because it contains an HEXXH motif characteristic of many Zinc-dependent proteases (reviewed in Miyoshi and Shinoda, 2000; and Rawlings and Barrett, 1995). However, proteases with this motif usually have one or more residues C-terminal to this sequence that participate in coordination of the

Zinc. Based on alignments with the known groups of HEXXH-containing proteases, it is not clear which other residues in ImmA might be involved in metal coordination (BB and ADG, unpublished observations). Other than this motif, ImmA consists of one large domain of unknown function. Future work might explore the characteristics of ImmA. In addition to determining whether additional residues in ImmA are required for metal-binding, it might be worthwhile to determine its structure and oligomeric state. Any of these characteristics might be changed upon activation by RapI or RecA and/or be altered in the central $ImmA^h$ mutants.

ImmA, the substrate

ImmA is likely targeted, by its C-terminal sequence, for degradation by one or more ATP-dependent proteases. Future work could explore this. Although it seems likely that degradation of ImmA is not involved in induction by RapI and RecA, turnover of ImmA could play some other role in regulation of ICE*Bs1*. The rates of synthesis and degradation of ImmA could be measured and compared to those of other proteins. The impact on the cell and on ICE*Bs1* of ImmA degradation could be investigated.

The degradation of ImmA is also interesting independently of its role in regulating ICE*Bs1*. Signals for degradation by ClpXP have been explored in detail in *E. coli*, but not in *B. subtilis*. Some C-terminal sequences work similarly in both bacteria, but others are quite different (KG and ADG, unpublished observations). The C-terminal sequence of ImmA doesn't resemble C-terminal sequences known to be recognized by ClpXP in *E. coli* (Flynn *et al.*, 2003). Additionally, it seems that *lonA* and/or *lonB* might play a role in ImmA degradation. Further study of proteolysis of ImmA might provide new information

about how one or more of the ATP-dependent proteases in *B. subtilis* recognizes substrates.

ImmR

ImmR's N terminus has a predicted helix-turn-helix (HTH) DNA-binding domain, but the rest of ImmR could be characterized further. This portion of ImmR contains the cleavage site and the residues flanking it that can be mutated to produce *immR(ind-)* alleles. Also, the C-terminal portion of ImmR (without the HTH) can interact with itself and with ImmA in yeast two-hybrid assays. Perhaps this portion folds independently, forming a domain for protein-protein interactions. If so, it is likely to be conserved, since there are many homologs of ImmR. The structure of the C-terminal part of ImmR could be determined. Perhaps hybrid repressors could be made, e.g. that repress P*xis* but autoproteolyze in response to different signals, or that repress different promoters but respond to ImmA.

The hunt for *immR(ind-)* alleles and characterization of those that have been found could be continued. Two of the isolated mutants are not cleaved in vivo, and the other two are cleaved less well. The efficiency with which these mutants are cleaved in vitro should be determined. Targeted mutations flanking the cleavage site could be made to figure out which residues are important for recognition and proteolysis. Since the four mutants closely flank the cleavage site, perhaps a short peptide bearing the ImmR sequence around the cut site will be susceptible to cleavage by ImmA.

ImmA and ImmR homologs in other systems

ImmA and ImmR homologs have been identified in other systems, in association with other known or putative mobile elements. For the phage Φ105, we showed that its

regulatory proteins behave similarly to those of ICE*Bs1*. Its ImmR homolog is necessary and sufficient for repression of transcription from the promoter in the position of P*xis*, and its ImmA homolog is required for induction in response to MMC. Phage Φ105, like ICE*Bs1*, resides in *B. subtilis*. The interplay between these two elements could be studied. It's possible that one affects the other's ability to move into a new host, or to be transferred out of one. Perhaps there is some interplay between their regulatory proteins, or perhaps there is no cross-talk. Either way, the specificity of the ImmA-ImmR system of each could be studied.

We think that the ImmA-ImmR mechanism of regulation is conserved in other systems carrying homologs of these proteins. Some of the other mobile elements with ImmA and ImmR homologs are known to be induced by the SOS response and/or have a Rap and Phr system. A tempting idea in thinking about the conservation of this form of regulation is that ImmA homologs help elements respond to different types of signals (e.g. cell-cell signaling and DNA damage). Perhaps having an anti-repressor in addition to a repressor allows single elements to respond to a wider variety of signals.

References

Auchtung, J.M., Lee, C.A., Monson, R.E., Lehman, A.P., and Grossman, A.D. (2005) Regulation of a Bacillus subtilis mobile genetic element by intercellular signaling and the global DNA damage response. *Proc Natl Acad Sci U S A* **102**: 12554-12559.

Auchtung, J.M., Lee, C.A., Garrison, K.L., and Grossman, A.D. (2007) Identification and characterization of the immunity repressor (ImmR) that controls the mobile genetic element ICEBs1 of Bacillus subtilis. *Mol Microbiol* **64**: 1515-1528.

Bose*, B., Auchtung*, J.M., Lee, C.A., and Grossman, A.D. (2008) A conserved anti-repressor controls horizontal gene transfer by proteolysis. *Mol Microbiol* **70**: 570-582. * co-first authors.

Flynn, J.M., Neher, S.B., Kim, Y.I., Sauer, R.T., and Baker, T.A. (2003) Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol Cell* **11**: 671-683.

Miyoshi, S., and Shinoda, S. (2000) Microbial metalloproteases and pathogenesis. *Microbes Infect* **2**: 91-98.

Rawlings, N.D., and Barrett, A.J. (1995) Evolutionary families of metallopeptidases. *Methods Enzymol* **248**: 183-228.