microbial biotechnology

Microbial Biotechnology (2010) **3**(3), 259–268 doi:10.1111/j.1751-7915.2009.00114.x

Minireview

Virulence regulation in *Citrobacter rodentium*: the art of timing

Ji Yang,[1,2] Marija Tauschek,[1,2] Emily Hart,[1,2] Elizabeth L. Hartland[1,2] and Roy M. Robins-Browne[1,2]*

[1]*Department of Microbiology and Immunology, The University of Melbourne, Victoria 3010, Australia.*

[2]*Murdoch Childrens Research Institute, Royal Children's Hospital, Parkville, Victoria 3052, Australia.*

Summary

The mouse enteric pathogen *Citrobacter rodentium*, like its human counterpart, enteropathogenic *Escherichia coli*, causes attaching and effacing lesions in the intestinal epithelium of its host. This phenotype requires virulence factors encoded by the locus for enterocyte effacement (LEE) pathogenicity island. For timely expression of these virulence determinants at the site of infection and for efficient delivery of some virulence factors into epithelial cells, *C. rodentium* utilizes a positive regulatory loop involving the LEE-encoded regulatory proteins Ler, GrlA and GrlR to control LEE expression. Several transcription factors not encoded by LEE, some of which respond to specific environmental signals, also participate in this regulatory loop. Recently, we identified a non-LEE encoded, AraC-like regulatory protein, RegA, which plays a key role in the ability of *C. rodentium* to colonize the intestine. RegA functions by activating the transcription of a number of horizontally acquired operons encoding virulence-associated factors, such as autotransporters, fimbriae, a dispersin-like protein and its transporter. In addition, RegA represses transcription of a number of housekeeping genes. Importantly, RegA requires a gut-specific environmental signal, bicarbonate, to exert its effects on gene expression. In our proposed model, when *C. rodentium* senses bicarbonate ions in the gastrointestinal tract, RegA directs the bacterium to reduce the production of proteins involved in normal cellular functions, while enhancing the production of factors required for colonization and virulence.

Received 29 December, 2008; accepted 16 March, 2009. *For correspondence. E-mail r.browne@unimelb.edu.au; Tel. (+61) 3 8344 8275; Fax (+61) 3 8344 8276.

Introduction

Citrobacter rodentium is a Gram-negative enteric bacterium that is a natural pathogen of mice. In its natural host, *C. rodentium* causes colonic epithelial hyperplasia accompanied by mild diarrhoea (for reviews, see Luperchio and Schauer, 2001; Mundy *et al.*, 2005). Its principal importance, however, is that infection of mice with *C. rodentium* provides a convenient small animal model to investigate the molecular and cellular pathogenesis of infections with the human pathogens, enteropathogenic *Escherichia coli* (EPEC) and enterohaemorrhagic *E. coli* (EHEC). This is because all three pathogens produce virtually indistinguishable attaching and effacing lesions in the intestinal epithelium (Frankel *et al.*, 1998) (Fig. 1), due to the fact that they carry the locus for enterocyte effacement (LEE), a highly conserved pathogenicity island (PAI), which is required for the development of these lesions (McDaniel *et al.*, 1995; Elliott *et al.*, 1998; Deng *et al.*, 2001). Thus, studies of the regulation of virulence genes in *C. rodentium* are likely to shed new light on the molecular pathogenesis of EPEC and EHEC, and permit the development of testable hypotheses regarding virulence gene regulation in these important pathogens.

The association of *C. rodentium* with its natural host has led to the evolution of complex regulatory networks, which are the result of recruiting and adapting pre-existing regulatory mechanisms to act in concert with newly acquired regulatory genes carried on mobile elements. These sophisticated regulatory networks bring about precise control of virulence gene expression in response to different environments within and outside the host, and ensure the survival of *C. rodentium* in the environment and the extreme conditions encountered in the stomach and small intestine.

Colonization of the intestinal epithelium by *C. rodentium* is thought to occur in two stages: initial loose attachment

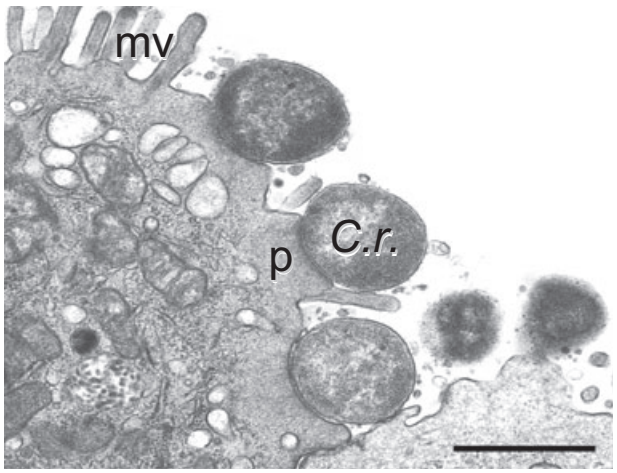

Fig. 1. Transmission electron micrograph of a section of mouse colon 14 days after oral inoculation with *Citrobacter rodentium* (*C.r.*), showing distinctive attaching–effacing lesions characterized by intimately attached bacteria (*C.r.*), disrupted microvilli (mv) and alterations in cellular architecture, such as pedestal formation (p). Scale bar, 1 μm.

of the bacteria to intestinal epithelial cells, followed by tight attachment and epithelial cell effacement. Each stage involves co-ordinated expression of a repertoire of specific virulence factors. The RegA protein is a key transcriptional regulator that is mainly responsible for transcriptional regulation during the early stages of infection, i.e. during transition from the environment to the site of bacterial colonization in the intestine (Hart *et al.*, 2008; Yang *et al.*, 2008). Upon stimulation by a gut-specific effector, bicarbonate, RegA activates transcription of a suite of horizontally acquired genes involved in colonization of the intestine, while inhibiting the expression of several housekeeping genes, thus directing the bacterium's energy and resources away from performing normal cellular functions and towards the production of virulence factors.

The LEE PAI, which is required for the attaching and effacing phenotype of *C. rodentium*, carries more than 40 genes that are arranged in five transcriptional units, termed *LEE1* to *LEE5* (Deng *et al.*, 2001). These encode the structural components of a type III secretion system, intimin (an outer membrane protein) and its translocated receptor (a type III-secreted protein known as Tir), as well as several other proteins (Esp) secreted by the type III secretion system. The LEE also codes for the regulatory proteins, Ler and GrlR/A, which play an important role in the transcriptional control of LEE and several virulence determinants not encoded by LEE (Elliott *et al.*, 2000; Deng *et al.*, 2004; Torres *et al.*, 2007; Abe *et al.*, 2008). In addition, RegA appears to modulate expression of LEE by upregulating transcription of *grlR/A*.

The RegA regulon

Involvement of the RegA protein in the virulence of C. rodentium

The *regA* gene was initially identified as a virulence determinant of *C. rodentium* by using signature-tagged transposons to create a library of *C. rodentium* mutants that were screened for their ability to colonize mouse intestine (Kelly *et al.*, 2006; Hart *et al.*, 2008). In a mixed-infection experiment with wild-type *C. rodentium*, a *C. rodentium* mutant, which carried a disrupted *regA* gene, was significantly out-competed by the wild type, indicating an impaired ability to colonize mouse intestine. Additional infection experiments using the *regA* mutant showed that the mean maximum number of the *regA* mutant in faeces was only 0.02% that of the wild type. Fulfilment of molecular Koch's postulates was achieved by trans-complementing the mutant with a plasmid carrying a wild-type *regA* gene that restored its colonizing ability, thus confirming that *C. rodentium* requires *regA* to colonize mouse intestine.

The *regA* gene codes for a protein of 276 amino acids (Hart *et al.*, 2008) which exhibits sequence similarity to AraC-like virulence transcriptional regulators from other intestinal bacterial pathogens. Homologues of RegA include PerA from EPEC (Gomez-Duarte and Kaper, 1995), Rns from enterotoxigenic *E. coli* (Caron *et al.*, 1989), AggR from enteroaggregative *E. coli* (Nataro *et al.*, 1994) and ToxT from *Vibrio cholerae* (Higgins *et al.*, 1992). Adjacent to the *regA* gene on the *C. rodentium* chromosome are two divergently transcribed loci, *adcA* and *kfc*, which encode a putative autotransporter involved in diffuse adherence, and a K99-like fimbrial operon respectively (Hart *et al.*, 2008). Real-time PCR analysis has shown that transcription of the *adcA* gene and the *kfc* cluster is strongly activated by RegA (Yang *et al.*, 2008). Further studies showed that expression of the *adcA* gene in *E. coli* K12 caused the recipient bacteria to self-aggregate and enabled them to adhere to mammalian tissue culture cells *in vitro* (Hart *et al.*, 2008). Although inactivation of *kfc* resulted in a slight, but significant, reduction in the duration of intestinal colonization by *C. rodentium*, deletion of both *adcA* and *kfc* from *C. rodentium* did not lead to the same reduction in colonizing ability as that exhibited by the *regA* mutant alone (Hart *et al.*, 2008). This suggested that there are other RegA-regulated genes, which also contribute to the ability of *C. rodentium* to colonize mice.

Bicarbonate-mediated activation of the adcA *and* kfcC *promoters by RegA*

Interestingly, RegA is the only known AraC-like transcriptional regulator from enteric bacteria that responds

directly to a gut-specific environmental signal, namely bicarbonate ions (Yang *et al.*, 2008). In mammals, bicarbonate is responsible for whole-body homeostasis (Guyton and Hall, 2006). In the gastrointestinal tract, bicarbonate ions are secreted by the pancreas into the duodenum to neutralize stomach acid (Kaunitz and Akiba, 2006). In the duodenum, the concentration of bicarbonate can exceed 50 mM (Meyer *et al.*, 1970). We have shown that *C. rodentium* can sense such bicarbonate-rich environments and respond by activating genes required for intestinal colonization.

In the presence of 40 mM bicarbonate ions, RegA enhances expression of the divergent *adcA* and *kfcC* promoters in an *E. coli* K12 background approximately 100-fold compared with 10-fold when bicarbonate is absent (Yang *et al.*, 2008). This is because in the absence of bicarbonate, RegA binds only weakly to a DNA fragment carrying the regulatory region of *adcA* and *kfcC*, but binding is greatly enhanced when bicarbonate is present. In addition, DNA footprinting analysis has shown that RegA protects an approximately 80 bp portion of the region between the promoters for *adcA* and *kfcC* from digestion by DNase I. The interaction of RegA with this sequence causes a conformational change in DNA, which resembles that observed in eukaryotic nucleosomes (Yang *et al.*, 2008).

The intergenic region between the divergent *adcA* and *kfcC* promoters is extremely AT-rich (~70%) and is predicted to be highly curved. Like many virulence genes of enteric bacteria that were acquired horizontally, transcription of *adcA* and *kfcC* is subject to repression control by the global, histone-like regulatory proteins, H-NS and StpA (Yang *et al.*, 2008). Under inducible conditions, RegA stimulates transcription by counteracting the repressive effect of these proteins.

The RegA protein

By analogy with the AraC protein of *E. coli*, RegA is predicted to constitute two structural domains: an N-terminal domain [amino acid residues (aa) 1–180] and a C-terminal domain (aa 190–276) (Fig. 2). Using secondary structure prediction programmes, we showed that the N-terminal domain includes an α-helical arm (aa 1–32) followed by a region consisting of six β-sheets (aa 34–90). The downstream portion of the N-terminal domain (aa 92–180) is predicted to comprise five α helices. The C-terminal domain contains predominantly α helices, which, like other members of the AraC family, includes two helix–turn–helix, DNA-binding motifs.

We have used mutational analysis to show that the N-terminal arm of RegA is required for its dependence on bicarbonate ions to activate transcription. In particular, deletion of the four amino acids between positions 2 and 5, or changing the asparagine residue at position 16 to an alanine significantly enhance the ability of RegA to activate transcription in the absence of bicarbonate (J. Yang and R. Robins-Browne, in preparation). These observations are consistent with other AraC-like proteins, which also use the N-terminal arm to modulate protein function (Saviola *et al.*, 1998; Basturea *et al.*, 2008; Dominguez-Cuevas *et al.*, 2008). For example, in the absence of the effector 3-methylbenzoate, residues in the N-terminal arm of the XylS protein of *Pseudomonas putida*, interact with its C-terminal domain and inhibit binding of XylS to its DNA targets (Dominguez-Cuevas *et al.*, 2008).

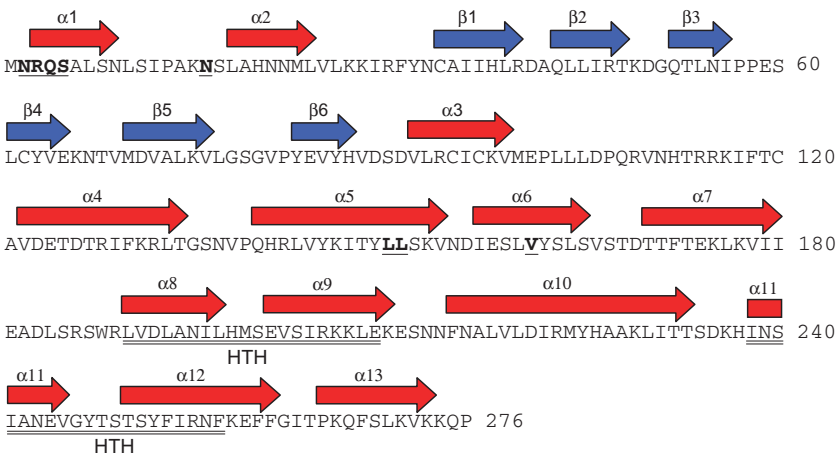

Fig. 2. Amino acid sequence and predicted secondary structure of the RegA protein. The secondary structure was predicted *in silico* by using web-based programs SOPMA, JPRED, PROF and PSIPRED (Geourjon and Deleage, 1995; Cuff *et al.*, 1998; McGuffin *et al.*, 2000; Ouali and King, 2000). Regions predicted to be α-helices and β-sheets are indicated by red and blue horizontal arrows respectively. The residues at positions 2–5 and 16, which are involved in bicarbonate-dependent activation, are underlined and in boldface type. The branched chain amino acid patch (Leu-151, Leu-152 and Val-161), which is implicated in dimerization is also underlined and in boldface. The putative double helix–turn–helix (HTH) DNA binding motif is indicated by double underlining.

Interestingly, RegA also appears to use self-contact inhibition to modulate its DNA-binding activity in response to its effector – bicarbonate. Thus, binding of their respective effectors to XylS and RegA or deletion of their N-terminal arm appears to relieve the inhibition of the DNA-binding activity of these proteins.

Analysis of the crystal structure of the N-terminus of AraC indicates that the arabinose-binding pocket is imbedded within an eight-stranded, antiparallel β-barrel. The corresponding region in RegA, which comprises six successive β-sheets (aa 33–90), is also likely to be involved in effector binding, as disrupting this region by inserting an alanine and a serine into each of the β-sheets completely eliminates bicarbonate-dependent activation via RegA (J. Yang and R. Robins-Browne, in preparation).

Two AraC monomers are able to dimerize via antiparallel coiled-coils. Dimerization requires the packing of three leucine residues at positions 150, 151 and 161 in a knobs-into-holes-like manner (Soisson *et al.*, 1997). As shown in Fig. 2, RegA contains two leucine residues at positions 150 and 151 and a valine at position 161. Replacing leucine-150 with an alanine abolishes the activation function of RegA, indicating the probable involvement of this leucine in dimerization (J. Yang and R. Robins-Browne, in preparation). Thus, although the primary sequence of the N-terminal domain of RegA has diverged considerably during evolution from those of other members of the AraC family, it retains certain structural and functional features, which these proteins employ to control interactions with their target promoters.

Global gene regulation by RegA

Microarray analysis has revealed that RegA is a global transcriptional regulator of *C. rodentium*. RegA strongly activates transcription (between 4- and 126-fold) of 19 open reading frames organized in 11 separate operons at widely different locations on the bacterial chromosome (Table 1). Interestingly, all of the RegA-upregulated operons are flanked by mobile genetic elements such as transposons, insertion sequences and remnants of bacteriophages, indicating that they were horizontally acquired. Several of the RegA-regulated transcriptional units identified in this way code for homologues of known virulence factors of intestinal pathogens, such as the Aap dispersin and its transporter (Aat) from enteroaggregative *E. coli* (Sheikh *et al.*, 2002; Nishi *et al.*, 2003), the CexE virulence factor from enterotoxigenic *E. coli* (Pilonieta *et al.*, 2007) and the SfpA porin (systemic factor protein A) from *Yersinia enterocolitica* (Mildiner-Earley and Miller, 2006). As with *adcA* and *kfc*, activation of these putative virulence determinants of *C. rodentium* by RegA is bicarbonate-dependent and therefore is likely to be directly regulated by this protein.

RegA also effects the transcriptional repression of 24 operons. As with upregulation, repression is also bicarbonate-dependent (Table 1). All of the downregulated genes identified by using microarray analysis are housekeeping genes located on the chromosomal backbone of enteric pathogens and non-pathogenic bacteria alike. These genes encode amino acid transporters, regulatory proteins and enzymes involved in various cellular functions. In contrast to its potent stimulatory effect on the transcription of virulence genes, RegA exerts only a moderate degree of repression on housekeeping genes (less than fivefold). Whether the negative control of these genes is directly or indirectly affected by RegA is yet to be determined. Microarray analysis also revealed that the *flhC-flhD* operon, which codes for a master regulatory complex ($FlhC_2$-D_2), responsible for stimulating the expression of flagella (Liu and Matsumura, 1994), is significantly downregulated by RegA (Table 1). Although *C. rodentium* is non-motile, it nevertheless carries an intact *flhC-flhD* operon, which can affect the expression of a number of gene targets apart from those required for flagella (Pruss *et al.*, 2001).

Autogenous regulation of regA

The *regA* gene is expressed from multiple promoters (Fig. 3). Promoter P1 is located immediately upstream of the *regA* structural gene and is constitutively expressed (A. Tan and R. Robins-Browne, in preparation). Promoters P2 and P3 are located in the regulatory region of an open reading frame of unknown function (*orfU*), situated immediately upstream of *regA* (Fig. 3). Microarray and transcriptional analysis using promoter–*lacZ* fusions have shown that RegA activates transcription of P2 and P3 in the absence of bicarbonate (A. Tan and R. Robins-Browne, in preparation). In addition, bicarbonate enhances RegA-mediated activation of *orfU* from 15- to 50-fold (Table 1). A fourth promoter, P4, which also contributes to *regA* expression, is the promoter of the *adcA*

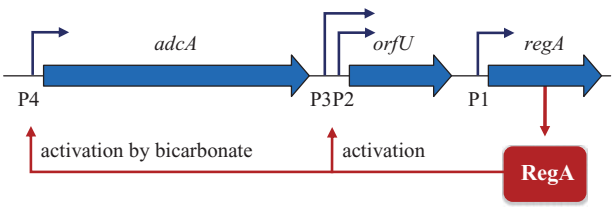

Fig. 3. Schematic representation of the double feed-back auto-activation loop of *regA* transcription. The four promoters (P1, P2, P3 and P4) which are responsible for *regA* expression are indicated by hooked horizontal arrows. P1 is a constitutive promoter, whereas transcription of P2 and P3 can be activated by RegA alone and further stimulated by bicarbonate. Transcription of the P4 promoter is also stimulated by RegA and activation is largely bicarbonate-dependent.

Table 1. Genes activated/repressed by RegA identified by microarray analysis.

ORF[a]/gene	Product	Fold increase/ decrease[b]
ROD3421	Homologue of Aap (dispersin) of enteroaggregative *E. coli*, and virulence factor CexE of enterotoxigenic *E. coli*	126.4
ROD3431-81	Homologue of the Aat transporter of enteroaggregative *E. coli*	12.4
ROD15971	65% identity : 83% similarity to unknown protein, Z2283, of *E. coli* O157:H7	5.0
ROD16181	69% identity : 80% similarity to unknown protein, Z0957, encoded by prophage CP-933K *E. coli* O157:H7	10.7
ROD16201	Homologue of the porin protein SfpA of *Yersinia enterocolitica*	23.4
grlA	GrlA	4.6
grlR	GrlR	4.6
ROD41031-51	Gene cluster encoding an unknown protein, and homologues of HlyD and HlyB secretion proteins of *Shewanella woodyi*	7.9
kfcC-H	Kfc, K99-like fimbriae	16.4
adcA	AdcA, autotransporter	19.6
orfU	Unknown ORF encoded immediately upstream of *regA*	56.0
ROD50001	Homologue of a putative virulence-related PagC-like membrane protein of *E. coli* O157:H7	4.6
puuB	gamma-Glu-putrescine oxidase, FAD/NAD(P)-binding	–2.1
ybdR	predicted oxidoreductase, Zn-dependent and NAD(P)-binding	–2.3
gltL	ATP-binding component of an ABC superfamily glutamate/aspartate transporter	–2.6
tK	Membrane component of an ABC superfamily glutamate/aspartate transporter	–2.4
gltI	Periplasmic-binding component of an ABC superfamily glutamate/aspartate transporter	–2.8
sucA	2-oxoglutarate decarboxylase, thiamin-requiring	–2.0
KPN_00795	Putative urocanase	–2.5
putA	Multifunctional proline dehydrogenase and DNA-binding transcriptional regulator	–2.1
putP	Proline : sodium symporter	–2.8
phoH	Conserved protein with nucleoside triphosphate hydrolase domain	–3.4
flgB	Flagellar component of cell-proximal portion of basal-body rod	–4.9
aldA	Aldehyde dehydrogenase A, NAD-linked	–2.3
treA	Periplasmic trehalase	–2.3
flhC	DNA-binding transcriptional dual regulator with FlhD	–2.0
flhD	DNA-binding transcriptional dual regulator with FlhC	–2.2
argT	Periplasmic-binding component of an ABC superfamily lysine/arginine/ornithine transporter	–2.5
fimA	Major type 1 subunit fimbrin (pilin)	–2.3
ytfQ	Predicted sugar transporter subunit	–2.2
actP	Acetate permease	–2.5
Acs	Acetyl-CoA synthetase	–2.3
fadB	Fatty acid oxidation complex subunit alpha	–2.5
fadA	3-ketoacyl-CoA thiolase (thiolase I)	–2.6
lldD	L-lactate dehydrogenase, FMN-linked	–3.1
dppA	Periplasmic-binding component of an ABC superfamily dipeptide transporter	–2.2

a. ORF names are taken from http://www.sanger.ac.uk/Projects/C_rodentium/C_rod_genome_CDS.tab and gene names and product descriptions are taken from EcoCyc (http://www.ecocyc.org). ORF/Gene names are listed in accordance to their predicted positions on the *C. rodentium* chromosome.

b. Fold increase/decrease derived from the average $\log_2$ ratio of transcript levels for a RegA+ strain to those for an isogenic RegA– strain, in the presence of 45 mM bicarbonate (a value of 1 indicates no change). The cut-off value for upregulation is $\geq$ 4-fold and for downregulation is $\geq$ 2-fold. All changes are statistically significant ($P < 0.05$).

gene (Fig. 3), which is strongly upregulated by RegA in the presence of bicarbonate.

Constitutive expression from the P1 promoter ensures that trace amounts of RegA are produced when bicarbonate is absent. Under these conditions, RegA would exert little effect on the expression of genes of the RegA regulon. Once *C. rodentium* reaches the proximal small intestine of its host, however, the presence of bicarbonate leads to induction of RegA synthesis through stimulation of promoter P4. The resultant elevated levels of RegA lead to further RegA synthesis from promoters P2 and P3, which results in the synchronous activation of multiple virulence genes and the downregulation of some housekeeping genes.

Transcriptional regulation of the LEE PAI

The Ler protein

The Ler regulatory protein plays a central role in controlling transcription of the entire LEE PAI (Mellies *et al.*, 1999; 2007). Ler proteins from EPEC, EHEC and *C. rodentium* are highly homologous in terms of their structure, function and mechanism of action. Although there was some uncertainty about which of the three methionine codons at the N-terminus is the start site of the *ler* structural gene, this is now known to be the methionine at position 7 of the previously predicted Ler open reading frame, reducing the predicted length of Ler from 127 to 121 amino acids (Yerushalmi *et al.*, 2008).

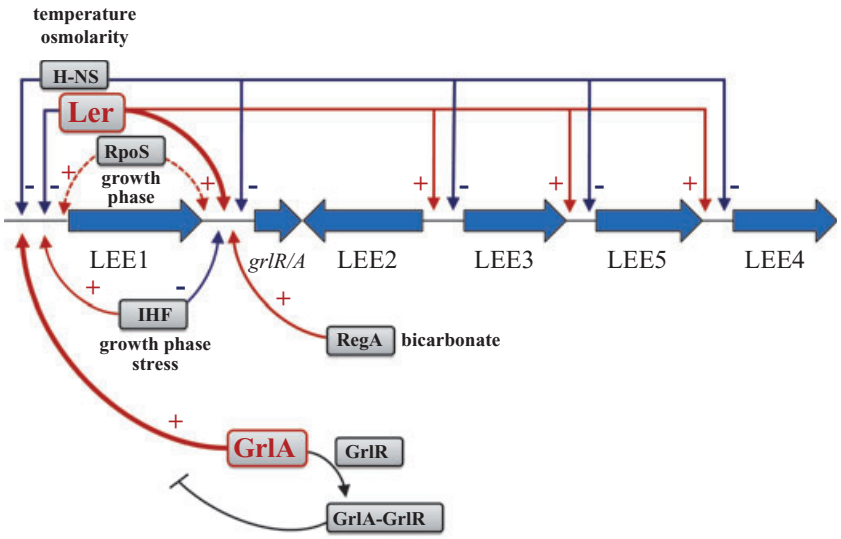

Fig. 4. Transcriptional regulation of the LEE PAI of *C. rodentium*. The key transcriptional regulators of LEE expression, Ler and GrlA, are shown in red. All of the other regulatory proteins are shown in black. The environmental factors and conditions associated with the functions of these proteins are indicated. Positive regulation is shown by red arrows and '+' signs and downregulation is indicated by dark blue arrows and '–' signs. The thick red arrows show the positive regulatory loop. Whether the effect of RpoS on *grlR/A* and *LEE1* expression (shown in dashed arrows) is direct or indirect is yet to be determined.

In silico analysis and mutational studies of Ler from EPEC have shown that Ler contains three putative functional domains (Sperandio *et al.*, 2000; Yerushalmi *et al.*, 2008). The N-terminal domain includes a coiled-coil region that may be involved in high-order oligomerization, whereas the central and C-terminal domains are responsible for homo-oligomerization and DNA binding respectively.

Ler activates transcription of the *LEE2*, *LEE3* and *LEE5* operons, and to a lesser extent the *LEE4* operon, by disrupting H-NS-mediated transcriptional silencing (Elliott *et al.*, 2000; Bustamante *et al.*, 2001; Haack *et al.*, 2003). In *C. rodentium*, Ler also upregulates expression of a small transcriptional unit, located between *LEE1* and *LEE2*, encoding the GrlR and GrlA proteins (Fig. 4) (Deng *et al.*, 2004).

Ler belongs to the H-NS family of DNA-binding proteins. Like H-NS, Ler binds to AT-rich DNA sequences, but so far, no consensus DNA binding sequence has been identified. Because of this, it has been suggested that Ler may recognize a DNA structural motif rather than a distinct sequence (Haack *et al.*, 2003).

Some investigators have reported that transcription of the EHEC *LEE1* promoter is autorepressed by Ler when *ler* expression is induced from a *tac* promoter by isopropyl-β-D-thiogalactopyranoside (Berdichevsky *et al.*, 2005; Yerushalmi *et al.*, 2008). However, comparison of promoter activity using a *LEE1-cat* fusion in wild-type *C. rodentium* and an isogenic *ler* mutant failed to demonstrate any negative effect of Ler on expression of the *LEE1* promoter (Deng *et al.*, 2004). This suggests that the repressive effect of Ler on *LEE1* occurs only when Ler is present at particularly high concentrations.

The LEE1 *promoter*

The *LEE1* promoter, which is responsible for the transcription of *ler*, is a key regulatory check-point of the LEE PAI. Within the *LEE 1* regulatory region of EHEC are two possible locations that contain promoter sequences, whereas in *C. rodentium* there is only one. The transcriptional start site of the *LEE1* operon of *C. rodentium* has been mapped to a position 174 bp upstream of the start codon of Ler (Deng *et al.*, 2004) that corresponds to the distal *LEE1* promoter (P1) of EPEC and EHEC (Porter *et al.*, 2005). In *C. rodentium*, this is the only *LEE1* promoter as nucleotide differences in the region corresponding to the EHEC-specific proximal promoter (P2) abolish both the –35 and extended –10 sequences of this promoter (Deng *et al.*, 2004).

Transcription from the *LEE1* promoters of EHEC, EPEC and *C. rodentium* is regulated by multiple transcription factors. The nucleotide sequences of the *LEE1* promoter region from these strains are highly homologous and the promoters share a number of common regulators. For example, each of the *LEE1* promoters is positively regulated by the integration host factor (IHF) and GrlA, and negatively regulated by H-NS. However, unlike the situation with EPEC and EHEC where the *LEE1* promoters are upregulated by a plasmid-encoded PerC protein (Gomez-Duarte and Kaper, 1995; Mellies *et al.*, 1999) and chro-

mosomally encoded PerC-like proteins (PchA, B and C) (Iyoda and Watanabe, 2004; Porter *et al*., 2005) respectively, no homologue of PerC is evident in the *C. rodentium* genome. Expression of the *LEE1* promoter of EHEC is also subject to regulation via quorum sensing (QS) (Russell *et al*., 2007). This is mediated by the QseA protein, which acts directly on the distal P1 promoter of *LEE1*. Although there is a QseA homologue in *C. rodentium*, its involvement in the regulation of the *LEE1* promoter remains to be tested.

Coulthurst and colleagues (2007) have reported the presence of an *N*-acylhomoserine lactone (AHL) QS system in *C. rodentium*. This system is directed by CroI, a LuxI homologue. Compared with the wild type, a *croI* mutant strain adhered less efficiently to an abiotic surface *in vitro* and showed enhanced virulence for mice. The gene targets of this QS system in *C. rodentium* have not yet been identified.

The grlR/A *operon*

Systematic mutational analysis of the LEE PAI of *C. rodentium* by Deng and colleagues (2004) has revealed that the GrlR and GrlA proteins encoded by *orf10* and *orf11*, respectively, are involved in transcriptional regulation of the LEE PAI. Deletion of *grlA* (*orf11*) from *C. rodentium* results in a major reduction in transcription of *LEE1*, *LEE2* and *LEE5*, whereas deleting *grlR* (*orf10*) leads to a small increase in transcription of *LEE1*, indicating that GrlA and GrlR are positive and negative regulators of LEE PAI transcription respectively. Barba and colleagues (2005) have identified a *cis*-acting element upstream of the *C. rodentium LEE1* promoter that is responsible for GrlA-mediated activation, indicating that GrlA interacts with the *LEE1* promoter directly. In addition, Huang and Syu (2008) have used a GrlA–GST fusion protein in an electrophoretic mobility shift assay to show that GrlA from EHEC forms complexes with DNA fragments carrying the *LEE1* promoter region.

Although the GrlR/A proteins from EHEC, EPEC and *C. rodentium* are highly conserved, they display a different regulatory pattern, in that GrlA of *C. rodentium* activates *LEE1* transcription in both *C. rodentium* and *E. coli* K12 (Barba *et al*., 2005), whereas GrlA from EHEC can only activate *LEE1* expression in its native background (Russell *et al*., 2007). These findings suggest that GrlA-mediated activation of the EHEC *LEE1* promoter requires one or more additional factors.

A positive regulatory loop

Transcription of the *C. rodentium grlR/A* operon is driven by a σ^{70} promoter that is repressed by IHF (M. Tauschek and R. Robins-Browne, in preparation) and H-NS but activated by Ler (Barba *et al*., 2005). Furthermore, Dong and colleagues (2009) have provided evidence that RpoS positively influences expression of the LEE, possibly through activation of the *grlR/A* and *ler* genes. The reciprocal activation of the *ler* (*LEE1*) promoter by GrlA and the *grlR/A* promoter by Ler allows a rapid build-up of the two activators that are required to initiate virulence gene expression (Fig. 4). Expression of this positive regulatory loop is further modulated by a number of proteins acting on the *LEE1* and *grlR/A* promoters. Some of these proteins such as H-NS and IHF are able to respond directly to different environmental signals. In *C. rodentium*, transcription of the *grlR/A* operon is also upregulated by RegA in the presence of bicarbonate (Table 1). Transcriptional analysis using a *grlR/A–lacZ* fusion in *E. coli* K12 and electrophoretic mobility shift assays have confirmed that the RegA-mediated upregulation results from direct binding of the RegA protein to the *grlR/A* promoter region (M. Tauschek and R. Robins-Browne, in preparation). Thus, through its connection with RegA, the Ler-GrlA positive regulatory loop can be further stimulated by bicarbonate ions, which serve as a gut-specific signal for virulence gene activation.

Although attaching–effacing pathogens must be ready to activate LEE-encoded genes when they reach the site of infection, they also need to maintain appropriate levels of Ler and GrlA within the cell to reduce the metabolic burden that would be imposed by their overexpression. Two mechanisms that may be involved in preventing runaway expression of Ler and GrlA, include (i) concentration-dependent autogenous repression by Ler at the *LEE1* promoter (see above) (Berdichevsky *et al*., 2005), which ensures that transcription of *ler* is blocked when the Ler concentration reaches a certain threshold, and (ii) inhibition of GrlA via the formation of GrlR-GrlA heterodimers (Jobichen *et al*., 2007) (Fig. 4).

Conclusion

The mouse model of infection with *C. rodentium* has proved invaluable in elucidating key features of the pathogenesis of infections with attaching and effacing enterobacteria in general. In addition, studies of gene regulation in *C. rodentium* have provided important new insights into how some intestinal pathogens can respond to specific environmental stimuli by producing essential virulence determinants at the appropriate stage of infection. The key role played by bicarbonate sensing in the expression of virulence by *C. rodentium* has parallels in EPEC and EHEC (Kenny *et al*., 1997; Yoh *et al*., 2003), but in *E. coli* these pathways have not been fully explored.

Although there are some differences in the transcriptional control of the LEE PAIs of EPEC, EHEC and *C. rodentium*, the principal regulatory mechanisms are the same. Specifically, in all of these bacteria, the evolution of

sophisticated regulatory networks has ensured the rapid production of a key regulatory protein, Ler, which in turn activates other operons within the LEE PAI.

Notwithstanding its essential role in virulence, the LEE is not sufficient on its own for *E. coli* or *C. rodentium* to cause disease. In *C. rodentium*, the RegA regulon, which comprises at least 11 putative virulence operons and 24 genes not directly associated with virulence, also plays a central role in infection. Genes of the RegA regulon appear to be particularly important during the early stages of infection. In addition, through its action on *grlA*, RegA provides a key link between the sensing of environmental bicarbonate and the expression of virulence determinants encoded by the LEE.

Despite recent advances in our understanding of several important aspects of virulence gene regulation in *C. rodentium*, a number of key questions remain to be resolved, including how bicarbonate ions bind to the RegA protein and what structural changes these ions cause? Is there a consensus nucleotide sequence for RegA binding? What are the gene targets of the AHL QS system in *C. rodentium*? What are the molecular mechanisms of transcriptional and post-transcriptional regulation of the *nle* genes (non LEE-encoded effectors) (Deng *et al.*, 2004; Kelly *et al.*, 2006; Roe *et al.*, 2007; Wickham *et al.*, 2007; Garcia-Angulo *et al.*, 2008)? What role does each of the newly identified putative virulence genes play in infection? Ongoing investigations into virulence gene regulation in *C. rodentium* will address these specific questions and provide further insights into how pathogenic bacteria are able to persist under widely differing environmental conditions inside and outside their hosts and cause disease.

Acknowledgements

We are grateful to A. Tan for allowing us to use her unpublished results. Work in the authors' laboratories is supported by research grants from the Australian National Health and Medical Research Council and the Australian Research Council.

References

Abe, H., Miyahara, A., Oshima, T., Tashiro, K., Ogura, Y., Kuhara, S., *et al.* (2008) Global regulation by horizontally transferred regulators establishes the pathogenicity of *Escherichia coli*. *DNA Res* **15:** 25–38.

Barba, J., Bustamante, V.H., Flores-Valdez, M.A., Deng, W., Finlay, B.B., and Puente, J.L. (2005) A positive regulatory loop controls expression of the locus of enterocyte effacement-encoded regulators Ler and GrlA. *J Bacteriol* **187:** 7918–7930.

Basturea, G.N., Bodero, M.D., Moreno, M.E., and Munson, G.P. (2008) Residues near the amino terminus of Rns are essential for positive autoregulation and DNA binding. *J Bacteriol* **190:** 2279–2285.

Berdichevsky, T., Friedberg, D., Nadler, C., Rokney, A., Oppenheim, A., and Rosenshine, I. (2005) Ler is a negative autoregulator of the *LEE1* operon in enteropathogenic *Escherichia coli*. *J Bacteriol* **187:** 349–357.

Bustamante, V.H., Santana, F.J., Calva, E., and Puente, J.L. (2001) Transcriptional regulation of type III secretion genes in enteropathogenic *Escherichia coli*: Ler antagonizes H-NS-dependent repression. *Mol Microbiol* **39:** 664–678.

Caron, J., Coffield, L.M., and Scott, J.R. (1989) A plasmid-encoded regulatory gene, *rns*, required for expression of the CS1 and CS2 adhesins of enterotoxigenic *Escherichia coli*. *Proc Natl Acad Sci USA* **86:** 963–967.

Coulthurst, S.J., Clare, S., Evans, T.J., Foulds, I.J., Roberts, K.J., Welch, M., *et al.* (2007) Quorum sensing has an unexpected role in virulence in the model pathogen *Citrobacter rodentium*. *EMBO Rep* **8:** 698–703.

Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M., and Barton, G.J. (1998) JPred: a consensus secondary structure prediction server. *Bioinformatics* **14:** 892–893.

Deng, W., Li, Y., Vallance, B.A., and Finlay, B.B. (2001) Locus of enterocyte effacement from *Citrobacter rodentium*: sequence analysis and evidence for horizontal transfer among attaching and effacing pathogens. *Infect Immun* **69:** 6323–6335.

Deng, W., Puente, J.L., Gruenheid, S., Li, Y., Vallance, B.A., Vazquez, A., *et al.* (2004) Dissecting virulence: systematic and functional analyses of a pathogenicity island. *Proc Natl Acad Sci USA* **101:** 3597–3602.

Dominguez-Cuevas, P., Marin, P., Busby, S., Ramos, J.L., and Marques, S. (2008) Roles of effectors in XylS-dependent transcription activation: intramolecular domain derepression and DNA binding. *J Bacteriol* **190:** 3118–3128.

Dong, T., Coombes, B.K., and Schellhorn, H.E. (2009) Role of RpoS in the virulence of *Citrobacter rodentium*. *Infect Immun* **77:** 501–507.

Elliott, S.J., Wainwright, L.A., McDaniel, T.K., Jarvis, K.G., Deng, Y.K., Lai, L.C., *et al.* (1998) The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic *Escherichia coli* E2348/69. *Mol Microbiol* **28:** 1–4.

Elliott, S.J., Sperandio, V., Giron, J.A., Shin, S., Mellies, J.L., Wainwright, L., *et al.* (2000) The locus of enterocyte effacement (LEE)-encoded regulator controls expression of both LEE- and non-LEE-encoded virulence factors in enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infect Immun* **68:** 6115–6126.

Frankel, G., Phillips, A.D., Rosenshine, I., Dougan, G., Kaper, J.B., and Knutton, S. (1998) Enteropathogenic and enterohaemorrhagic *Escherichia coli*: more subversive elements. *Mol Microbiol* **30:** 911–921.

Garcia-Angulo, V.A., Deng, W., Nikhil, A.T., Finlay, B.B., and Puente, J.L. (2008) Regulation of expression and secretion of NleH, a new non-locus of enterocyte effacement-encoded effector in *Citrobacter rodentium*. *J Bacteriol* **190:** 2388–2399.

Geourjon, C., and Deleage, G. (1995) SOPMA: significant improvements in protein secondary structure prediction by

consensus prediction from multiple alignments. *Comput Appl Biosci* **11:** 681–684.

Gomez-Duarte, O.G., and Kaper, J.B. (1995) A plasmid-encoded regulatory region activates chromosomal *eaeA* expression in enteropathogenic *Escherichia coli*. *Infect Immun* **63:** 1767–1776.

Guyton, A.C., and Hall, J.E. (2006) Regulation of acid-base balance. In *Textbook of Medical Physiology*. Schmitt, W., Gruiliow, R., and Norwitz, A. (eds). Philadelphia, PA, USA: Elsevier Saunders, pp. 383–387.

Haack, K.R., Robinson, C.L., Miller, K.J., Fowlkes, J.W., and Mellies, J.L. (2003) Interaction of Ler at the LEE5 (tir) operon of enteropathogenic *Escherichia coli*. *Infect Immun* **71:** 384–392.

Hart, E., Yang, J., Tauschek, M., Kelly, M., Wakefield, M.J., Frankel, G., *et al.* (2008) RegA, an AraC-like protein, is a global transcriptional regulator that controls virulence gene expression in *Citrobacter rodentium*. *Infect Immun* **76:** 5247–5256.

Higgins, D.E., Nazareno, E., and DiRita, V.J. (1992) The virulence gene activator ToxT from *Vibrio cholerae* is a member of the AraC family of transcriptional activators. *J Bacteriol* **174:** 6974–6980.

Huang, L.H., and Syu, W.J. (2008) GrlA of enterohemorrhagic *Escherichia coli* O157:H7 activates LEE1 by binding to the promoter region. *J Microbiol Immunol Infect* **41:** 9–16.

Iyoda, S., and Watanabe, H. (2004) Positive effects of multiple pch genes on expression of the locus of enterocyte effacement genes and adherence of enterohaemorrhagic *Escherichia coli* O157:H7 to HEp-2 cells. *Microbiology* **150:** 2357–2571.

Jobichen, C., Li, M., Yerushalmi, G., Tan, Y.W., Mok, Y.K., Rosenshine, I., *et al.* (2007) Structure of GrlR and the implication of its EDED motif in mediating the regulation of type III secretion system in EHEC. *PLoS Pathog* **3:** e69.

Kaunitz, J.D., and Akiba, Y. (2006) Review article: duodenal bicarbonate-mucosal protection, luminal chemosensing and acid-base balance. *Aliment Pharmacol Ther* **24** (Suppl. 4)**:** 169–176.

Kelly, M., Hart, E., Mundy, R., Marches, O., Wiles, S., Badea, L., *et al.* (2006) Essential role of the type III secretion system effector NleB in colonization of mice by *Citrobacter rodentium*. *Infect Immun* **74:** 2328–2337.

Kenny, B., Abe, A., Stein, M., and Finlay, B.B. (1997) Enteropathogenic *Escherichia coli* protein secretion is induced in response to conditions similar to those in the gastrointestinal tract. *Infect Immun* **65:** 2606–2612.

Liu, X., and Matsumura, P. (1994) The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class II operons. *J Bacteriol* **176:** 7345–7351.

Luperchio, S.A., and Schauer, D.B. (2001) Molecular pathogenesis of *Citrobacter rodentium* and transmissible murine colonic hyperplasia. *Microbes Infect* **3:** 333–340.

McDaniel, T.K., Jarvis, K.G., Donnenberg, M.S., and Kaper, J.B. (1995) A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc Natl Acad Sci USA* **92:** 1664–1668.

McGuffin, L.J., Bryson, K., and Jones, D.T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics* **16:** 404–405.

Mellies, J.L., Elliott, S.J., Sperandio, V., Donnenberg, M.S., and Kaper, J.B. (1999) The Per regulon of enteropathogenic *Escherichia coli*: identification of a regulatory cascade and a novel transcriptional activator, the locus of enterocyte effacement (LEE)-encoded regulator (Ler). *Mol Microbiol* **33:** 296–306.

Mellies, J.L., Barron, A.M., and Carmona, A.M. (2007) Enteropathogenic and enterohemorrhagic *Escherichia coli* virulence gene regulation. *Infect Immun* **75:** 4199–4210.

Meyer, J.H., Way, L.W., and Grossman, M.I. (1970) Pancreatic bicarbonate response to various acids in duodenum of the dog. *Am J Physiol* **219:** 964–970.

Mildiner-Earley, S., and Miller, V.L. (2006) Characterization of a novel porin involved in systemic *Yersinia enterocolitica* infection. *Infect Immun* **74:** 4361–4365.

Mundy, R., MacDonald, T.T., Dougan, G., Frankel, G., and Wiles, S. (2005) *Citrobacter rodentium* of mice and man. *Cell Microbiol* **7:** 1697–1706.

Nataro, J.P., Yikang, D., Yingkang, D., and Walker, K. (1994) AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli*. *J Bacteriol* **176:** 4691–4699.

Nishi, J., Sheikh, J., Mizuguchi, K., Luisi, B., Burland, V., Boutin, A., *et al.* (2003) The export of coat protein from enteroaggregative *Escherichia coli* by a specific ATP-binding cassette transporter system. *J Biol Chem* **278:** 45680–45689.

Ouali, M., and King, R.D. (2000) Cascaded multiple classifiers for secondary structure prediction. *Protein Sci* **9:** 1162–1176.

Pilonieta, M.C., Bodero, M.D., and Munson, G.P. (2007) CfaD-dependent expression of a novel extracytoplasmic protein from enterotoxigenic *Escherichia coli*. *J Bacteriol* **189:** 5060–5067.

Porter, M.E., Mitchell, P., Free, A., Smith, D.G., and Gally, D.L. (2005) The *LEE1* promoters from both enteropathogenic and enterohemorrhagic *Escherichia coli* can be activated by PerC-like proteins from either organism. *J Bacteriol* **187:** 458–472.

Pruss, B.M., Liu, X., Hendrickson, W., and Matsumura, P. (2001) FlhD/FlhC-regulated promoters analyzed by gene array and *lacZ* gene fusions. *FEMS Microbiol Lett* **197:** 91–97.

Roe, A.J., Tysall, L., Dransfield, T., Wang, D., Fraser-Pitt, D., Mahajan, A., *et al.* (2007) Analysis of the expression, regulation and export of NleA-E in *Escherichia coli* O157:H7. *Microbiology* **153:** 1350–1360.

Russell, R.M., Sharp, F.C., Rasko, D.A., and Sperandio, V. (2007) QseA and GrlR/GrlA regulation of the locus of enterocyte effacement genes in enterohemorrhagic *Escherichia coli*. *J Bacteriol* **189:** 5387–5392.

Saviola, B., Seabold, R., and Schleif, R.F. (1998) Arm-domain interactions in AraC. *J Mol Biol* **278:** 539–548.

Sheikh, J., Czeczulin, J.R., Harrington, S., Hicks, S., Henderson, I.R., Le Bouguenec, C., *et al.* (2002) A novel dispersin protein in enteroaggregative *Escherichia coli*. *J Clin Invest* **110:** 1329–1337.

Soisson, S.M., MacDougall-Shackleton, B., Schleif, R., and Wolberger, C. (1997) Structural basis for ligand-regulated oligomerization of AraC. *Science* **276:** 421–425.

Sperandio, V., Mellies, J.L., Delahay, R.M., Frankel, G., Crawford, J.A., Nguyen, W., and Kaper, J.B. (2000) Activation of enteropathogenic *Escherichia coli* (EPEC) LEE2 and LEE3 operons by Ler. *Mol Microbiol* **38:** 781–793.

Torres, A.G., Lopez-Sanchez, G.N., Milflores-Flores, L., Patel, S.D., and Rojas-Lopez, M., Martinez de la Pena, C.F., *et al.* (2007) Ler and H-NS, regulators controlling expression of the long polar fimbriae of *Escherichia coli* O157:H7. *J Bacteriol* **189:** 5916–5928.

Wickham, M.E., Lupp, C., Vazquez, A., Mascarenhas, M., Coburn, B., Coombes, B.K., *et al.* (2007) *Citrobacter rodentium* virulence in mice associates with bacterial load and the type III effector NleE. *Microbes Infect* **9:** 400–407.

Yang, J., Hart, E., Tauschek, M., Price, G.D., Hartland, E.L., Strugnell, R.A., and Robins-Browne, R.M. (2008) Bicarbonate-mediated transcriptional activation of divergent operons by the virulence regulatory protein, RegA, from *Citrobacter rodentium*. *Mol Microbiol* **68:** 314–327.

Yerushalmi, G., Nadler, C., Berdichevski, T., and Rosenshine, I. (2008) Mutational analysis of the locus of enterocyte effacement-encoded regulator (Ler) of enteropathogenic *Escherichia coli*. *J Bacteriol* **190:** 7808–7818.

Yoh, M., Bi, Z., Matsuyama, J., Nagayama, K., and Honda, T. (2003) Effect of environmental conditions on proteins secreted by enterohemorrhagic *Escherichia coli* O26:H11. *Microbiol Immunol* **47:** 1–6.

