LeuO is a global regulator of gene expression in Salmonella enterica serovar Typhimurium:

Shane C. Dillon, Elena Espinosa, Karsten Hokamp, David W. Ussery, Josep Casadesús, Charles J. Dorman

LeuO is a global regulator of gene expression in *Salmonella enterica* serovar Typhimurium

Shane C. Dillon[1], Elena Espinosa[2], Karsten Hokamp[3], David W. Ussery[4], Josep Casadesús[2], and Charles J. Dorman[1] *

[1]Department of Microbiology, School of Genetics and Microbiology, Moyne Institute of Preventive Medicine, Trinity College Dublin, Dublin 2, Ireland

[2]Departamento de Genética, Facultad de Biología, Universidad de Sevilla, Apartado 1095, Seville 41080, Spain

[3]Department of Genetics, School of Genetics and Microbiology, Smurfit Institute, Trinity College Dublin, Dublin 2, Ireland.

[4]Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark

*Corresponding author, Email: cjdorman@tcd.ie

Running title: *Salmonella* Typhimurium LeuO regulon

Key words: LeuO, LysR-type transcription regulator, H-NS, gene regulation, RNA polymerase, *Salmonella enterica*

Summary

We report the first investigation of the binding of the *Salmonella enterica* LeuO LysR-type transcription regulator to its genomic targets *in vivo*. Chromatin-immunoprecipitation-on-chip identified 178 LeuO binding sites on the chromosome of *S. enterica* serovar Typhimurium strain SL1344. These sites were distributed across both the core and the horizontally acquired genome, and included housekeeping genes and genes known to contribute to virulence. Sixty-eight LeuO targets were co-bound by the global repressor protein, H-NS. Thus, while LeuO may function as an H-NS antagonist, these functions are unlikely to involve displacement of H-NS. RNA polymerase bound 173 of the 178 LeuO targets, consistent with LeuO being a transcription regulator. Thus, LeuO targets two classes of genes, those that are bound by H-NS and those that are unbound by H-NS. LeuO binding site analysis revealed a logo conforming to the $TN_{11}A$ motif common to LysR-type transcription factors. It differed in some details from a motif that we composed for *Escherichia coli* LeuO binding sites. 1263 and 1094 LeuO binding site locations were predicted in the *S*. Typhimurium SL1344 and *E. coli* MG1655 genomes, respectively. Despite differences in motif composition, many LeuO target genes were common to both species. Thus, LeuO is likely to be a more important global regulator than previously suspected.

Introduction

The LeuO protein belongs to the family of LysR-type transcriptional regulators (LTTRs), the largest family of prokaryotic DNA binding proteins (Momany and Neidle, 2012; Pareja *et al*., 2006) with over 55,000 potential members in the Pfam database (PF00126 HTH_1, LysR_substrate) (Punta *et al*., 2012). The number of different LTTRs in a given species is often very large; for example, *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium) encodes 44 distinct LTTRs (Lahiri *et al*., 2009). LTTRs are typically 300 to 350 amino acids in length and activate the transcription of operons and regulons involved in diverse cellular functions such as nitrogen fixation, the response to oxidative stress and aspects of bacterial virulence (Van Rhijn and Vanderleyden, 1995; Hernandez-Lucas *et al.,* 2008; Lahiri *et al.,* 2008; 2009; Maddocks and Oyston, 2008; O'Byrne and Dorman, 1994; Sheehan and Dorman, 1998). A typical LysR family member consists of an N-terminal DNA binding domain and a C-terminal sensing domain and it activates or represses the transcription of target genes; some, but not all, are known to bind ligands that influence their DNA binding activity (Schell, 1993; Zaim and Kierzek, 2003).

LeuO is a member of the LTTR family and is found in members of the *Enterobacteriaceae*, including *Escherichia coli*, *Salmonella, Shigella* and *Yersinia* spp. (Maddocks and Oyston, 2008). The designation *leuO* was originally used to describe an open reading frame of unknown function located beside the leucine biosynthesis operon (*leuABCD*) (Henikoff *et al*. 1988). The *leuO* gene was later found to be part of a complex *cis*-acting promoter relay system that connects the *leuABCD* and *ilvIH* operons (Chen *et al*., 1992; Wu *et al*., 1995; Wu and Fang, 2003). More recently it has been shown that the *leuO* gene is activated by the transcriptional regulators RcsB and BglJ and is negatively autoregulated (Stratmann *et al*., 2012).

Further work, performed mostly in *E. coli* and *Salmonella* Typhi, has shown that LeuO plays a positive role in the regulation of a number of genes including the *bgl* operon involved in β-glucoside utilization (Ueguchi *et al*., 1998), the *yjjQ-bglJ* operon encoding LuxR-type transcriptional regulators (Stratmann *et al*., 2008), and the *yjcRQP* operon involved in Sulfa drug efflux (Shimada *et al*., 2009). LeuO also activates the *casABCDE* operon (Hernandez-Lucas *et al*. 2008; Westra *et al*., 2010; Medina-Aparicio *et al*., 2011) that is part of the CRISPR/Cas defence system that provides 'immunity' against mobile genetic elements (Karginov and Hannon, 2010). Furthermore LeuO has been implicated in the bacterial response to stress (Majumder *et al*., 2001; Fang *et al*., 2000) and genetic screens have identified LeuO as a *Salmonella* virulence factor (Tenor *et al*., 2004; Lawley *et al*., 2006). LeuO may in part contribute to *Salmonella* virulence by positively regulating the outer membrane porins OmpS1 and OmpS2 (Fernandez-Mora *et al*., 2004; De la Cruz *et al*., 2007; Hernandez-Lucas *et al*., 2008), which are known virulence factors in the mouse model infection system (Rodriguez-Morales *et al*., 2006). LeuO has also been reported to regulate positively *rovA*, a key regulator of virulence in *Yersinia* spp. (Lawrenz and Miller, 2007).

It is clear that LeuO regulates a diverse set of genes and this has been expanded by a recent SELEX analysis in *E. coli* (Shimada *et al*., 2011). However, we do not have a detailed understanding of the genes that are regulated by LeuO in the important pathogen *S*. Typhimurium. On these grounds, we used a global ChIP-chip approach to identify LeuO-regulated genes to obtain comprehensive information about LeuO binding to its genomic targets in living bacterial cells. This approach allowed us to investigate LeuO binding in the context of other DNA binding proteins that are likely to target the same genes, such as RNA polymerase and the H-NS protein. It also allowed us to study LeuO binding to DNA adopting conformations that are natural to the *in vivo* situation, a factor that is known to influence the binding of several DNA binding proteins (Cameron

95 and Dorman, 2012). In our strategy, we exploited knowledge of LeuO protein
96 expression: the *leuO* gene is expressed maximally in bacteria growing in phosphate-
97 limited minimal medium on entry into stationary phase (Fang *et al*., 2000, VanBogelen *et*
98 *al*., 1996). We present the first *in vivo* DNA analysis of LeuO binding to the *S.*
99 Typhimurium chromosome using the ChIP-chip technique, and provide evidence that
100 LeuO is a global regulator in *S.* Typhimurium.
101

Results and Discussion

Identification of LeuO target genes in Salmonella enterica *using a ChIP-chip approach*

In order to understand the function of LeuO in *S.* Typhimurium, we identified the network of LeuO gene targets using a chromatin immunoprecipitation (ChIP) assay. The LeuO protein, tagged with the FLAG epitope, was cross-linked to its DNA targets, and then immunoprecipitated using an anti-FLAG antibody. The DNA targets bound by LeuO were then identified by hybridization to a DNA microarray (Experimental procedures).

Bacterial cultures to be used in the ChIP assay were grown under conditions known to promote maximal LeuO protein expression: growth in a minimal low-phosphate medium (LPM) to stationary phase (equivalent to an OD_{600} value of 1.4) (Fig. 1A) (Fang *et al*., 2000, VanBoegelen *et al*., 1996). The LeuO-bound ChIP DNA fragments were fluorescently labelled with Cy3 dCTP while the genomic DNA control was labelled with Cy5 dCTP. The DNA samples were co-hybridized to a DNA tiling microarray and the intensity of fluorescence of each of the DNA probes was calculated (Fig. 1B). The ChIPOTle peak finding programme (Buck *et al*., 2005) was used to identify LeuO binding sites using a 2-fold cut-off. This procedure identified 261 binding regions common to two biological replicate experiments. However as the ChIP-chip procedure often results in the identification of false positive binding events (Waldminghaus and Skarstad, 2010), a control 'mock' ChIP-chip experiment was also performed, in which normal mouse IgG antibodies were used during a ChIP reaction, to identify any DNA sequences that were non-specifically immunoprecipitated. The ChIPOTle programme identified 83 peaks in the control dataset that were also present in the LeuO dataset; consequently these targets were eliminated from the final analysis. Altogether, 178 LeuO binding sites were identified (Fig.1 and Table S1). Previously characterized LeuO target genes from other bacterial species that were found in our dataset include the CRISPR/Cas operon (Fig. 1),

127 *sdiA, ompN/ompS2*, *dnaE*, *cyoABCDE*, *tesB*, *fimD*, *sdhA*, *add*, *cpsG*, *nuoH*, *tdcD*, *treF*
128 and *phoU* (Table S1**)** (Westra *et al*., 2010; Turnbull *et al*., 2012; Shimada *et al*., 2011).
129 The presence of these targets validated our approach. A large number of new LeuO
130 target genes were also identified in this study, some of which are discussed below (see
131 *Extension of the LeuO regulon*).

132 The majority of the bacterial transcription factors that have been studied by
133 genome-wide location analysis have been found to bind predominantly to non-coding
134 DNA sequences (Grainger *et al*., 2004; 2005; 2006; 2007; Cho *et al*., 2008; Shimada *et*
135 *al*., 2011; Wei *et al*., 2012). This is not surprising because most transcription factors
136 regulate transcription by binding to DNA sites that are located upstream of open reading
137 frames (ORFs). In general, LTTRs bind to several intergenic sites located upstream of
138 their regulated gene(s) (Maddocks and Oyston, 2008). While they often bind close to
139 promoter regions (-55 bp to +20 bp), sites located more than 200 bp upstream of the
140 promoter have been detected, as have binding sites internal to the ORF (+350 bp)
141 (Wilson *et al*., 1995, Viswanathan *et al*., 2007). The evidence that LTTRs can exert their
142 regulatory influence through binding to a wide variety of locations prompted us to
143 examine the location of *S.* Typhimurium LeuO binding sites in detail.

144 LeuO binding sites were classified based on their location, i.e. intergenic or within
145 an ORF. Intergenic and intra-ORF sites were further sub-categorized into Intergenic (I,
146 located upstream of an individual gene), Intergenic Convergent (IC, located between two
147 convergently-transcribed genes), Intergenic Divergent (ID, located between two
148 divergently-transcribed genes), ORF 3' (located within the 3' promoter-distal half of an
149 ORF) and ORF 5' (located within the 5' promoter-proximal half of the ORF) (Fig. 1D).
150 Surprisingly we found that only about 33% of binding sites were located in intergenic
151 regions with the remaining ~66% of binding sites being located within ORFs (Fig. 1D).
152 Thirty six of the 60 intergenic binding events were located upstream of an individual

153 gene, 18 were located between divergently-transcribed genes and 6 binding events were
154 located between convergently-transcribed genes, making target gene predictions based
155 on binding site location difficult. The same number (59) of ORF binding events were
156 distributed equally between the 5' and 3' regions of ORFs. It is possible that a proportion
157 of these intra-ORF binding events have been incorrectly classified in the case of
158 adjacent genes that share short intergenic regions. This is because the resolution
159 capacity of the ChIP-chip method is limited by the average size of the sonicated DNA
160 fragments (~500 bp). However most represent intra-ORF LeuO binding sites of the type
161 that have been documented previously for LeuO and other LysR-like regulators
162 (Shimada *et al*., 2011; Wilson *et al*., 1995; Viswanathan *et al*., 2007). The significance of
163 intra-ORF binding is not clear but because of their location they may have a repressive
164 effect on transcription as in some instances LeuO has been shown to act as a negative
165 regulator (Hernandez-Lucas *et al*., 2008; Shimada *et al*., 2009). Alternatively, as H-NS
166 can bind within coding regions to form transcriptionally repressive nucleoprotein
167 complexes (Nagarajavel *et al*., 2007), these intra-ORF binding events could simply
168 reflect LeuO-mediated antagonism of H-NS binding to internal gene regions. It is also
169 possible that they may reflect an architectural role for the LeuO protein in determining
170 the structure of the nucleoid (Chen and Wu, 2005).
171
172 *Extension of the LeuO regulon*
173 Our ChIP-chip analysis greatly extended the number of known LeuO target genes in *S.*
174 Typhimurium. Prior knowledge of LeuO targets in *Salmonella* extended only to relatively
175 few loci, including genes in *S.* Typhi rather than *S.* Typhimurium (De la Cruz *et al*., 2007;
176 Hernandez-Lucas *et al*., 2008; Medina-Aparicio *et al*., 2011; Turnbull *et al*., 2012). The
177 164 new *S.* Typhimurium regulon members identified here are involved in a variety of
178 cellular processes and include, for example, inner/outer membrane proteins, transport

179 proteins, motility factors, cell division proteins, oxidative stress response proteins and
180 also other LTTRs (Table S1).

181 LeuO binding has been mapped previously to the CRISPR-associated *casA* and
182 *cas3* promoters in *S.* Typhi (Medina-Aparicio *et al*., 2011); here, we observed high levels
183 of LeuO binding at the promoter for the CRISPR repeats with little or no binding at the
184 *casA* and *cas3* promoters in *S.* Typhimurium (Fig. 1C). However, while P_{casA} and P_{cas3}
185 were not classified as LeuO targets in LPM growth conditions, subsequent LeuO-
186 binding-site motif analysis predicted LeuO binding in the 5' regulatory region of these
187 genes (see *Genome-wide prediction and validation of LeuO binding sites*).

188 Other notable LeuO target genes are *sopA*, encoding an effector protein that is
189 translocated by the *Salmonella* pathogenicity island (SPI) 1 type III secretion system,
190 and *sifA,* the SPI-2 translocated effector gene. These are important virulence
191 determinants of *S.* Typhimurium and their detection is consistent with the previous
192 characterization of LeuO as a *Salmonella* virulence factor required for host-pathogen
193 interactions (Tenor *et al*., 2004).

194 The *rcsA* gene was also identified as a LeuO target. Its product, RcsA, is an
195 auxiliary regulator for the Rcs (regulation of capsular polysaccharide biosynthesis) two-
196 component phosphophorelay system that senses alterations in the outer membrane and
197 the peptidoglycan layer of the cell envelope (Majdalani and Gottesman, 2005).
198 Interestingly, RcsA can form heterodimers with the response regulator RcsB, which
199 activates *leuO* transcription in conjunction with BglJ, counteracting H-NS repression of
200 *leuO* transcription (Stratmann *et al*., 2012).

201 The *rssB* gene, which encodes a response-regulator-like adaptor protein (RssB)
202 for ClpXP proteolytic degradation of the RpoS stress and stationary phase sigma factor
203 (Klauck *et al*., 2001), was found to be a LeuO target. This LeuO-RssB link is significant
204 in the context of an earlier discovery that a Tn*10* insertion that activated *leuO* expression

205 led to a reduction in RpoS levels (Klauck *et al*., 1997). It is possible that in addition to the
206 established effect of LeuO on DsrA regulatory RNA expression, and hence DsrA-
207 sensitive RpoS mRNA translation, LeuO may also influence RpoS protein levels by
208 modulating the expression of the RssB adaptor protein.

209 We also identified the genes coding for three other LysR-like regulators as LeuO
210 targets: *metR*, *yeeY* and *stm2180*. While the functions of YeeY and STM2180 are
211 uncharacterized, MetR is known to be involved in the regulation of methionine
212 biosynthesis (Maxon *et al*., 1989). The possibility that bacteria may link their amino acid
213 biosynthetic pathways through gene regulation is intuitively appealing, since mounting
214 appropriate responses to metabolic challenges is essential for survival. It may be also
215 significant during infection: the *Salmonella* containing vacuole in an infected host cell is a
216 nutrient limiting environment (simulated by growth in LPM), and control of amino acid
217 biosynthesis by LeuO may facilitate survival during the infection process.

218 The observed binding of LeuO to its known target *ompS2*/*ompN* but not to
219 *ompS1*/*ompS* provided an important insight. LeuO is known to induce *ompS2* expression
220 at a lower concentration than required for the induction of *ompS1* (De la Cruz, 2007),
221 consistent with LeuO having a higher affinity for the regulatory region of *ompS2*. The
222 low intracellular concentration of LeuO when cultured in LPM may not allow LeuO to
223 occupy lower affinity sites such as the regulatory region of *ompS1*. To investigate if the
224 genome-wide binding pattern of LeuO was altered upon an increase in the intracellular
225 concentration of LeuO, we used the inducible pBAD system to over-express 3 x FLAG
226 tagged LeuO and monitored its binding pattern using the ChIP-chip technique. This
227 analysis revealed that LeuO bound to 331 chromosomal locations (after removal of any
228 false positives also present in a mouse IgG control ChIP-chip) (Table S1). We observed
229 LeuO binding to other known targets including *ompS1* and *cas3*, which we did not detect
230 previously. This is consistent with LeuO having a lower affinity for these sites so that a

231 higher intracellular concentration of LeuO is required before full binding is achieved.
232 LeuO concentration is very low in exponentially growing cells (~200 molecules/cell) but
233 is believed to reach up to 2000-4000 molecules per cell in stationary phase *E. coli*
234 cultures (Shimada *et al*. 2011). If LeuO binds as a dimer, there may be as many as
235 1000-2000 LeuO binding sites (see *Genome-wide prediction and validation of LeuO*
236 *binding sites*). It is not inconceivable that under appropriate growth conditions
237 *Salmonella* synthesizes large amounts of LeuO to occupy even more binding sites than
238 documented here as we are only beginning to understand the growth conditions under
239 which LeuO plays a regulatory role (Gallego-Hernandez *et al*. 2012).
240
241 *LeuO binding in close proximity to H-NS*
242 Enteric bacteria encode the abundant DNA binding H-NS protein which acts as a global
243 repressor of ~20% of genes in *S.* Typhimurium (Dillon and Dorman, 2010; Dorman,
244 2004). H-NS represses transcription of these genes by binding to curved AT-rich DNA
245 sequences and mediates its repressive function by preventing RNA polymerase binding
246 or by trapping RNA polymerase at promoters (Dame *et al*., 2006; Navarre *et al*., 2006;
247 Lucchini *et al*., 2006; Walthers *et al*., 2011). Counteracting the repressive function of H-
248 NS is important if cells are to express H-NS regulated genes; not surprisingly, a number
249 of H-NS antagonism mechanisms have been described (Stoebel *et al*., 2008). LeuO has
250 recently emerged as an important antagonist of H-NS (Hernandez-Lucas *et al*., 2008;
251 Shimada *et al*., 2009; 2011) and it may exert this function by simply competing with H-
252 NS for binding to DNA (Shimada *et al*., 2011) or acting as a barrier to H-NS
253 polymerization (Chen *et al*., 2003; Chen and Wu, 2005). While LeuO is known to
254 antagonize H-NS, its own gene is repressed by H-NS (Klauck *et al*., 1997; Chen *et al*.,
255 2001; Stratmann *et al*., 2012). Deletion of *hns* in *Salmonella* Typhimurium strain SL1344
256 resulted in a dramatic increase in LeuO protein levels (Fig. 2A), confirming the

257 repressive action of H-NS at the *leuO* gene in SL1344. LeuO binding was observed in
258 close proximity to a previously mapped H-NS binding site at *leuO* (Dillon *et al*., 2010),
259 consistent with LeuO functioning as an H-NS antagonist in *S*. Typhimurium.

260 It was important to determine which of the 178 *S.* Typhimurium LeuO binding
261 events were associated with H-NS and to ascertain whether LeuO mediated its function
262 by displacing H-NS or by another mechanism. To address these questions, H-NS
263 binding to the SL1344 chromosome was examined by ChIP-chip analysis under the
264 same growth conditions that are known to promote *leuO* expression (i.e. grown to
265 stationary phase in LPM) (Fig. 2B). In addition, previously published data (Dillon *et al*.,
266 2010) on H-NS binding under standard laboratory growth conditions in which LeuO is
267 undetectable (i.e. grown to exponential phase in LB) were analysed and the findings
268 were integrated with those from the present investigation.

269 The ChIPOTle algorithm identified 496 H-NS binding regions in SL1344 grown in
270 LPM (Table S2) (456 binding regions were identified in LB, Dillon *et al*., 2010). The
271 locations of these H-NS binding regions were compared with the LeuO binding sites and
272 those LeuO binding sites that overlapped with, or were located within 200 bp of, an H-
273 NS binding region were classified as LeuO+H-NS sites; the remaining sites were
274 classified as LeuO sites (Fig. 2C) (Table S1). We identified 68 LeuO sites that met our
275 criteria for classification as LeuO+H-NS sites; the remaining 110 LeuO sites were not
276 associated with H-NS co-localization and so were designated as LeuO sites.

277 LeuO co-localization with H-NS at 68 locations is consistent with a global H-NS
278 antagonism function. However it is important to consider that LeuO may also repress
279 some of its target genes, perhaps in conjunction with H-NS. This possibility is supported
280 by the observation that LeuO and H-NS are both known to repress the *fimAICDFGH*
281 operon in *E. coli* (Shimada *et al*. 2011). The 68 LeuO+H-NS sites represent only 38% of
282 the total number of LeuO binding events; in contrast, Shimada *et al.* (2011) found that

283 95% of LeuO sites in *E. coli* overlapped with H-NS sites. We then examined average
284 LeuO and H-NS occupancy at the two classes of LeuO binding sites. Average H-NS and
285 LeuO ChIP occupancies were calculated +/- 500 bp with respect to the centre of the
286 LeuO binding sites. These surveys showed that the peak of LeuO binding was offset by
287 100 bp from the ChIPOTle peak centre (Fig. 2D). However an interesting pattern
288 emerged: close analysis of regions of LeuO and H-NS co-occupancy revealed that the
289 LeuO binding peak coincided consistently with a trough in the H-NS binding landscape
290 (Fig. 2D). The significance of this is not clear but may indicate that LeuO functions as an
291 H-NS barrier or antagonist, taking up a position interposed between two consecutive H-
292 NS binding peaks. It is also possible that the intracellular concentrations of LeuO might
293 not be high enough to displace H-NS completely.

294 We wished to know if LeuO could antagonize H-NS binding by competing with
295 and displacing this protein from DNA. We examined H-NS binding at the 110 LeuO-only
296 sites in LB-grown SL1344 and found that none of these 110 locations was occupied by
297 H-NS (Table S1). Therefore the presence of LeuO had not simply resulted in the
298 complete displacement of H-NS in LPM since these locations lacked H-NS binding in the
299 absence of detectable levels of LeuO. Next we examined H-NS occupancy at the 68
300 LeuO+H-NS sites in LB-grown cultures and found that fewer LeuO target genes were
301 bound by H-NS in LB when compared with LPM (41 of the 68 genes were bound by H-
302 NS in LB) (Fig. 3A). Therefore the presence of LeuO appeared to correlate with H-NS
303 binding to more LeuO target genes, which would not be expected if LeuO simply
304 displaced H-NS from its cognate binding sites. However these results did not rule out the
305 possibility that LeuO influenced the pattern of H-NS occupancy without completely
306 displacing H-NS. Therefore we calculated the average H-NS binding levels at LeuO
307 binding sites in both LPM and LB (Fig. 3B). This analysis revealed much higher levels of
308 H-NS binding at LeuO target genes in LeuO-inducing (i.e. LPM) conditions compared to

repressive (LB) conditions. Our data are inconsistent with a simplistic model in which LeuO overcomes H-NS transcriptional repression activity simply by displacing H-NS. They are more suggestive of a subtle remodeling of the nucleoprotein complex that overcomes H-NS-mediated repression without stripping H-NS from the DNA. A mechanism of this kind has been proposed for other H-NS antagonists in *S.* Typhimurium, such as the SlyA protein and OmpR (Perez *et al.,* 2008; Cameron and Dorman, 2012). The ability of LeuO to form DNA-protein-DNA bridges, analogous to those created by LacI, might also allow LeuO to interfere with H-NS activity without removing the protein from the DNA. Significantly, the LacI protein can replace LeuO as an H-NS antagonist at *leuO* (Chen and Wu, 2005), a finding that is consistent with the two proteins operating through a common mechanism. Here, the LeuO/LacI proteins erect a DNA-protein-DNA bridge between two binding sites that contains H-NS behind a LeuO/LacI barrier. This barrier protects a nearby promoter from encroachment by H-NS polymerization without the need to displace H-NS from the DNA.

RNA polymerase recruitment to LeuO target genes

LTTRs are known to bind sites overlapping or adjacent to the target promoter to repress or activate transcription (Maddocks and Oyston, 2008). It has been suggested that LTTRs activate transcription by interacting with the C-terminal domain of the α subunit of RNA polymerase (αCTD). For example, the LTTR family member OxyR increases RNA polymerase binding to OxyR-dependent promoters, suggesting that OxyR activates transcription partly by recruiting RNA polymerase (Kullik *et al*., 1995). To determine whether LeuO might recruit RNA polymerase, we examined RNA polymerase binding at LeuO target genes under LeuO-inducing growth conditions (LPM) and non-inducing growth conditions (LB). ChIPOTle analysis of both datasets and comparison with the location of LeuO binding sites revealed that 173 of the 178 LeuO binding sites

were also associated with RNA polymerase binding in LPM (co-occupancy was defined as a LeuO binding region located within 200 bp of an RNA polymerase binding region), whereas only 61 of the LeuO binding sites were occupied by RNA polymerase upon growth in LB (Fig. 3A) (Table S1).

This analysis provided three striking observations. First, LeuO appears to be associated with active transcription, as the vast majority of LeuO target genes are associated with RNA polymerase binding in LPM. However it cannot be excluded that some of these co-localization events represent transcriptionally repressive events as LeuO may prevent promoter clearance by trapping RNA polymerase at promoters, a property already described for H-NS (Dame *et al*. 2002). Secondly, the presence of RNAP at 105 LeuO target genes in inducing conditions (and its absence in non-inducing conditions) suggests that LeuO recruits RNA polymerase. Lastly the observation that 63 of the 68 LeuO+H-NS co-occupancy sites are also associated with RNA polymerase binding is intriguing as binding of H-NS and RNA polymerase are believed to be mutually exclusive in *Salmonella* (Lucchini *et al*. 2006). As discussed above this may represent trapping of RNA polymerase by LeuO and H-NS. Alternatively, LeuO may remodel H-NS oligomers and allow RNA polymerase to bind to promoters without the need to displace H-NS. We cannot discount the other possibility that all three factors do not in fact co-localize in individual cells as ChIP quantifies protein occupancy across a population of cells.

Next we plotted the mean RNA polymerase occupancy in both growth conditions with respect to the location of the LeuO binding sites (Fig. 3C). RNA polymerase occupancy was plotted for a distance of +/- 500 bp from the centre of the LeuO binding sites and was found to peak with respect to the centre of LeuO binding sites in both growth media. However, the average RNA polymerase occupancy in LeuO-inducing conditions (LPM) was much higher than in non-inducing conditions (LB) and the peak of

361 binding was also broader. This is consistent with LeuO promoting the recruitment of RNA
362 polymerase to target genes, and the broader peaks of binding are consistent with the
363 detection of elongating RNA polymerase.
364
365 *Identification of an AT-rich LeuO DNA binding motif*
366 The ability of proteins to recognize specific DNA sequences is a key feature of many
367 biological processes. Recognition of a specific DNA sequence by a protein often
368 involves the formation of amino acid and nucleotide-specific hydrogen bonds (Garvie
369 and Wolberger, 2001). For LysR-like proteins, a DNA sequence known as the LTTR box
370 has been identified (Maddocks and Oyston, 2008). The consensus sequence of the
371 LTTR box is T-N_{11}-A and often displays imperfect dyad symmetry (Parsek *et al*., 1994).
372 However this motif is highly degenerate and does not give an accurate understanding of
373 the DNA sequences with which LeuO interacts. Therefore we wanted to determine if
374 LeuO binding was associated with a specific DNA motif in our ChIP-chip binding sites.
375 The recent SELEX study of LeuO binding in *E. coli* (Shimada *et al*., 2011) also gave us
376 information on the location of LeuO binding sites in a related species and we were able
377 to incorporate this information into our analysis.
378 First, the *S*. Typhimurium LeuO+H-NS binding regions described above in which
379 LeuO binding overlapped with or was close to an H-NS binding location were selected
380 for DNA sequence motif analysis. We also created a list of DNA sequences bound by
381 LeuO in *E. coli* by extracting 500-bp DNA sequences that centered on the genomic
382 coordinate provided for each LeuO binding site by Shimada *et al*. (2011). Recall that
383 almost all of the *E. coli* sites were of the LeuO+H-NS type. The details of these datasets
384 are provided in Fig. 4A and in the Experimental procedures. Next we used the unbiased
385 motif-finding algorithm Meme to search the two datasets for significantly over-
386 represented sequence motifs (Bailey *et al*., 2009). Meme identified a 28-bp motif in both

387 datasets (Fig. 4B). Two striking features of the LeuO motifs are their imperfect dyad
388 symmetry and their A+T richness. While some dyad symmetry is discernable in the *E.*
389 *coli* logo, it is much harder to detect in its *S.* Typhimurium counterpart. Furthermore both
390 motifs contain a central region matching the T-N_{11}-A LTTR box motif and alignment of
391 the central T-N_{11}-A motifs of the sequence logos shows significant overlap between the
392 two motifs (Fig. 4B). However, the *E. coli* LTTR box displays a much stronger nucleotide
393 preference at most positions, a sequence divergence that may explain why only 15 of
394 the *E. coli* LeuO target genes were common to *S*. Typhimurium (Table S1). Regulon
395 divergence is not uncommon even in closely related species and this is reflected in
396 differences in the presence and nature of regulatory protein binding sites (Perez and
397 Groisman, 2009). While the *E. coli* and *S*. Typhimurium LeuO proteins are highly related
398 (87% amino acid identity) there are a number of amino acid differences in the N-terminal
399 DNA binding domain which may have altered DNA binding site specificity. Furthermore
400 selective pressure associated with the acquisition and regulatory integration of
401 horizontally-acquired *Salmonella* pathogenicity islands that contain a large number of
402 predicted LeuO binding sites (see following section)~~,~~ may have also altered DNA binding
403 site preference.

404 We know that the LTTR box motif is often associated with dyad symmetry
405 (Schell, 1993; Grob *et al*., 1997; Sheehan and Dorman, 1998) and this property appears
406 to be a general feature of the extended 28 bp motifs identified here, albeit weakly in the
407 case of the *S.* Typhimurium example. The presence of dyad symmetry is consistent with
408 individual LeuO subunits binding to half-sites to form a dimer. However, LTTRs are
409 known to be functionally active as tetramers that protect large regions of DNA (50-60 bp)
410 (Maddocks and Oyston, 2008). Tetramer formation by LeuO would lend itself to DNA-
411 protein-DNA bridging, as it is the case with tetrameric LacI (Chen and Wu, 2005). This
412 would allow LeuO to participate in both short-range and long-range protein-DNA

413 interactions, facilitated by DNA looping. Supporting this hypothesis are data from our
414 genome-wide prediction of LeuO binding sites that identified many examples of clustered
415 binding sites (see the *Genome-wide prediction and validation of LeuO binding sites*
416 section below).

417 The extremely high A+T content of both logos (Fig. 4B) is consistent with the
418 proposed role of LeuO as an H-NS antagonist as H-NS binds to A+T rich sequences
419 (Navarre *et al*., 2006, Lucchini *et al*., 2006, Dillon *et al*., 2010). Furthermore both logos
420 contain a central A-tract at nucleotide positions 15-17. These A-tracts are intriguing
421 because DNA structural studies have determined that A-tracts are associated with
422 narrowing of the DNA minor groove (Beveridge *et al*., 2004; Rohs *et al*., 2009). Variation
423 in DNA shape, in particular DNA minor groove width and DNA twist, is emerging as an
424 important 'indirect' mechanism used by proteins to achieve DNA-binding specificity in the
425 absence of base-specific contacts (Cameron and Dorman, 2012; Rohs *et al*., 2009). In
426 this mechanism the bases are not necessarily involved in contacting the protein but in
427 allowing the DNA to assume a conformation that facilitates protein binding (Rohs *et al*.,
428 2009). These flexible bases are often located in linker sequences that connect two half-
429 sites that are directly bound by protein subunits (Hizver *et al*., 2001; Rohs *et al*., 2009).
430 The quasi-palindromic nature and presence of A-tracts in the LeuO motifs suggests that
431 a combination of direct amino-acid-base-pair interactions and DNA shape may be
432 important features in determining LeuO binding specificity.
433
434 *Genome-wide prediction and validation of LeuO binding sites*
435 The identification of the 28-bp LeuO DNA binding motifs suggested that accurate
436 prediction of LeuO binding sites in *E. coli* and *S.* Typhimurium genome sequences would
437 be possible. We used the Motif Alignment and Search Tool (MAST), which is part of the
438 MEME suite of tools used for motif discovery and searching (Bailey *et al*., 2009), to

439 generate a position-specific scoring matrix (PSSM) from the LeuO sequence logos. This
440 PSSM was used to search for sequence matches in the *S*. Typhimurium SL1344 and *E.*
441 *coli* K-12 MG1655 genome sequences as described in the *Experimental procedures*
442 section. This analysis resulted in the identification of 1263 and 1094 matches in the *S.*
443 Typhimurium and *E. coli* genome sequences, respectively (Fig. 5A, B and Table S3).
444 These predicted sites were often located in intrinsically curved A+T rich regions and are
445 also associated with H-NS binding (Fig. 5A, Table S4). Eight-hundred-and-eighty-nine of
446 the 1263 predicted LeuO binding sites in *S*. Typhimurium were associated with H-NS
447 binding *in vivo*, suggesting that LeuO may function as a more global antagonist of H-NS
448 than previously thought. While the PSSMs used to search for binding sites differed
449 between the species, there were many instances of homologous genes being associated
450 with a predicted binding site(s) in *both* species. For example, predicted LeuO binding
451 sites were identified for the *yjiAXY* operon and in the nearby *yiiQ-bglJ* operon in both
452 species (Fig. 5C, D). LeuO is known to activate *bglJ* in *E. coli* (Stratmann *et al*., 2008)
453 while RcsB-BglJ heterodimers activate *leuO* transcription (Stratmann *et al*., 2012).
454 Furthermore *rcsB* is also associated with a predicted LeuO site in both species,
455 illustrating the complexity of the LeuO regulatory network.
456 In order to validate our genome-wide prediction of LeuO binding sites we
457 searched for sites in other known LeuO-regulated genes that were not identified in our
458 ChIP-chip study. We correctly predicted sites in the 5' regulatory region of *leuO* itself
459 (Chen and Wu, 2005), in the *yjjQ-bglJ* operon (Stratmann *et al*., 2008), in *ompS1*, *assT*
460 (*stm3192*), and in the CRISPR-associated *casA* and *cas3* genes (Westra *et al*., 2010;
461 Medina-Aparicio *et al*., 2011). Furthermore LeuO binding sites have been precisely
462 mapped in the regulatory regions of *Salmonella* Typhi *casA* (Medina-Aparicio *et al*.,
463 2011) and *ompS1* (De la Cruz *et al*., 2007) and our predicted binding sites map to these
464 locations.

To further validate our genome-wide prediction of LeuO binding sites, three *S.* Typhimurium regions were tested for *in vitro* binding of purified LeuO protein by electrophoretic mobility shift assays. The *pipA* and *envR* genes each contain one and two predicted LeuO sites respectively, while *SL3361,* which is located beside *envR* and does not contain a predicted LeuO binding site motif, was used as a negative control (Fig. 6A). Both *pipA* and *envR* DNA probes showed a clear pattern of retarded migration after incubation with increasing concentrations of purified LeuO while LeuO did not bind to the *SL3361* DNA probe (Fig. 6A). The *envR* 5' regulatory region (P_{envR}) contains two predicted LeuO binding sites located in close proximity to each other (81-bp spacing) and displayed a higher affinity for LeuO binding than the *pipA* region. The presence of two LeuO binding sites in close proximity and in helical register along the DNA may lead to LeuO oligomerization and DNA bending (Hryniewicz and Kredich, 1994), which may account for the apparently higher affinity observed for this DNA probe.

Next we used primer extension to resolve DNase I footprints on PCR amplified DNA templates (Cameron and Dorman, 2012). This approach can be used to more accurately map LeuO binding sites and validate the location of predicted LeuO binding sites. A 400 bp DNA probe encompassing two predicted sites in P_{envR} was used as the target in our experiments. We identified five regions that were protected from DNase I digestion by LeuO, two of which overlapped with the location of the predicted binding sites (Fig. 6B). Three other protected sites were identified further upstream of the *leuO* ORF which did not contain a predicted site but were located in a 60 bp region of high A+T content (66%), consistent with LeuO binding to A+T rich sequences.

Finally we used quantitative RT-PCR to examine the effect of deleting *leuO* on the transcript levels of four predicted LeuO target genes which are also bound by H-NS (Table S4) - *envR*, *pipA*, *sifA* and *sopA*. We examined transcript levels for these genes in (i) a strain that harbours a T-POP transposon (Lee *et al*. 2007) upstream of *leuO* in its

native chromosomal location, so that *leuO* is over-expressed when tetracycline is added to the culture (SL1344 Tpop-*leuO*) and (ii) a strain that harbours a T-POP transposon upstream of *leuO* but the *leuO* gene is deleted (SL1344 Tpop-Δ*leuO*). Deletion of *leuO* dramatically decreased the level of *envR* transcript (Fig. 6C) and increased the levels of *sifA*, *sopA* and *pipA* transcripts (Fig. 6C). The results of this analysis suggest that LeuO activates transcription of *envR*, perhaps by antagonizing H-NS, but appears to function as a repressor at the other target genes. It is possible that LeuO and H-NS function together to repress transcription of these genes. The location of the LeuO binding sites may be important as the predicted binding sites near *envR* are located upstream of the *envR* start codon, consistent with LeuO functioning as a transcriptional activator of this promoter. Conversely the predicted LeuO binding sites for *sifA*, *sopA* and *pipA* are all located downstream of the start codons, where LeuO binding is likely to have a repressive effect on transcription. Thus it appears that LeuO has a dual role as activator and repressor of transcription. Transcriptional activation by LeuO is well documented but LeuO has also been shown to repress the acid stress regulator *cadC*, the small RNA *dsrA* and the *fimAICDFGH* operon in *E. coli* (Shi and Bennett, 1995; Repoila and Gottesman, 2001; Shimada *et al*. 2011). Furthermore, LeuO has a complex relationship with its own gene, antagonizing H-NS-mediated *leuO* repression and antagonizing RcsB-BglJ-mediated *leuO* activation (Stratmann *et al.,* 2012; Chen and Wu, 2005).

Many of the genes on the A+T-rich *Salmonella* pathogenicity islands (SPI) 1 and 2 are repressed by H-NS (Dillon *et al*., 2010) and our analysis identified 25 predicted LeuO sites in SPI1 and 11 in SPI2. Notably, sites were predicted in the promoter regions of key regulators *hilA*, *hilC*, *hilD*, and *ssrAB* (Table S3). This may explain why LeuO was identified as a virulence factor in a *S.* Typhimurium host-pathogen model system (Tenor *et al*., 2004) and in a long-term systemic infection mouse model system (Lawley *et al*., 2006).

517 It is also important to point out that 24 of 44 genes encoding *S.* Typhimurium
518 LTTRs (Lahiri *et al*., 2009) contain one or more predicted LeuO binding site(s) in their
519 regulatory region. These include the genes encoding TdcA, which is involved in the
520 metabolism of L-serine and L-threonine (Kim *et al*., 2009), and NhaR, which regulates a
521 sodium proton antiporter (Rahav-Manor *et al*., 1992). These LTTRs and their
522 neighbouring regulatory targets are also repressed by H-NS (Table S2), suggesting a
523 complex regulatory interplay between LeuO, other LTTRs and H-NS. As LTTRs often
524 auto-regulate their own expression (Maddock and Oyston, 2008), it is possible that LeuO
525 establishes a heterotypic interaction with the corresponding LTTR family member to
526 facilitate this auto-regulation (Knapp and Hu, 2010). The presence of LeuO binding sites
527 at so many LTTR genes shows that LeuO also has the potential to coordinate their
528 expression within a LeuO-dependent regulatory network.
529 Many genes that contain a predicted LeuO binding site were not detected by our
530 ChIP-chip approach. The tetrameric structure of LTTRs permits simultaneous binding to
531 two sites. The nucleoprotein complex upstream of a given target gene can include a
532 variety of distal sites, bridged by the LTTR to a common target-gene-proximal site. Distal
533 site availability may be determined stochastically or may reflect changes in nucleoid
534 structure, dictating which distal sites are likely to be brought into a position adjacent to
535 the target-gene-proximal site such that LeuO-mediated bridging becomes possible.
536 These considerations are not only relevant to a consideration of LeuO as a regulator of
537 transcription, but also apply to its potential role as a nucleoid-structuring element.
538

Experimental procedures

Bacterial strains and culture conditions

The bacterial strains used in this study were *S.* Typhimurium strain SL1344 (*his,* a histidine auxotroph) (Hoiseth and Stocker, 1981); CJD1034scd (SL1344 *leuO*::3xFLAG); CJD1028scd (SL1344 *hns*::3xFLAG); SL1344 *hns*::*kan*; SL1344 *hns*::*kan leuO*::*cat*, SL1344 pBAD*leuO*::3xFLAG; SV7424 (SL1344 Tpop-*leuO*), SV7425 (SL1344 Tpop-Δ*leuO*. *E. coli* K-12 strain BL21 (pIZ1871) was used to overexpress and purify the LeuO protein. *S.* Typhimurium strain SL1344 and its derivatives were grown in 250 ml flasks in low phosphate medium (LPM) (5 mM KCl, 7.5 mM $(NH_4)_2SO_4$, 0.5 mM K_2SO_4, 0.1% casamino acids, 0.34% glycerol, 80 mM MES, after autoclaving 337.5 µM K_2HPO_4/KH_2PO_4 and 10mM $MgCl_2$ were added, pH 7) or Luria Broth at 37°C and at 200 rpm in a C76 water bath shaker (New Brunswick Scientific). Where necessary, antibiotics were used at the following final concentrations: carbenicillin (100 μg ml^{-1}), kanamycin (50 μg ml^{-1}), streptomycin (50 μg ml^{-1}) and chloramphenicol (25 μg ml^{-1}).

Construction of SL1344 derivatives expressing 3xFLAG tagged LeuO and H-NS

The *leuO*::3xFLAG and *hns*::3xFLAG derivatives of *S.* Typhimurium SL1344 were generated using a modified version of the λ Red recombination method (Uzzau *et al.,* 2001). The primers used for these constructions are listed in Table S5. The *leuO*::3xFLAG and *hns*::3xFLAG strains were marked with the kanamycin antibiotic resistance cassette, which was transduced into a clean background using phage P22 HT *int105*. Green agar plates were used to screen for colonies free from phage (Sternberg and Maurer, 1991). The kanamycin resistance cassette was flanked by FRT sites (FLP recombinase recognition targets), which allowed its removal by site-specific recombination using the pCP20 plasmid (Cherepanov and Wackernagel, 1995).

564 *Chromatin immunoprecipitation (ChIP) assay*

565 Cultures of *S.* Typhimurium SL1344, SL1344 *leuO*::3xFLAG, SL1344 *hns*::3xFLAG were
566 inoculated 1:100 into fresh Luria broth or Low phosphate medium (LPM) and grown at
567 37°C until stationary phase. SL1344 pBAD*leuO*::3xFLAG was inoculated 1:100 into fresh
568 Luria broth and grown to exponential phase (OD_{600} ~0.2) and then arabinose was added
569 to a final concentration of 0.2% for 1 hour before harvesting cells. For all experiments 25
570 ml of culture was harvested and re-suspended in 50 ml of PBS. DNA–protein
571 interactions were cross-linked for 30 min by adding formaldehyde (Sigma-Aldrich,
572 catalogue number F8775) to a final concentration of 1%. Glycine was then added to a
573 final concentration of 0.125M to stop the crosslinking. The ChIP assay was then
574 performed as detailed in Dillon *et al*., (2010). The following antibodies were used in this
575 study: mouse monoclonal to the RNA polymerase β' subunit (Neoclone catalogue
576 number W0001), mouse monoclonal to the FLAG epitope (Sigma-Aldrich, catalogue
577 number F3165) and normal mouse IgG (Millipore, catalogue number 12-371).

578

579 *Fluorescent labelling of DNA and microarray hybridization*

580 Fluorescent labelling of ChIP DNA samples for microarray experiments were carried out
581 as described in Dillon *et al*. (2010). The microarrays used in this study were designed
582 and produced by Oxford Gene Technology. The microarrays consisted of 43,453 60-mer
583 oligonucleotides tiled throughout the *S.* Typhimurium SL1344 chromosome and pSLT
584 plasmid. Microarrays were hybridized for 24 h in a hybridization oven (Agilent
585 Technologies) and washed according to instructions provided by Oxford Gene
586 Technology.

587

588 *Microarray data acquisition, analysis and data access*

589 The microarray slides were scanned using an Agilent G2505C scanner. Cy3 and Cy5
590 images were acquired at 3-micron resolution. Scanned images were analysed using
591 Agilent Feature extraction software. This software packages was used to quantify the
592 fluorescent intensities of each spot representing an array element. Background
593 subtracted fluorescence values were reported for each spot in the Cy3 and Cy5
594 channels and used to calculate a background subtracted Cy3/Cy5 ratios. The baseline
595 levels of each dataset were normalized to a value of one, allowing all experiments to be
596 directly compared from this baseline value. The data centering was performed by
597 calculating the median ratio for each experiment and dividing all the Cy3/Cy5 ratios
598 (obtained in that experiment) by this number. The ChIPOTle algorithm (Buck *et al*., 2005)
599 was used to define regions of enrichment in ChIP-on-chip datasets by using a sliding
600 window approach. ChIPOTle calculates the average $\log_2$ ratio within each window and
601 the fold cut-off chosen was $\log_{21}$. A window size of 500 bp and a step size of 125 bp
602 were used for analyzing the datasets, the rationale being that the ChIP procedure
603 produces DNA fragments of approximately 500 bp in size. The raw ChIP-on-chip
604 datasets have been submitted to the GEO database (Accession number GSE35826).
605
606 *Quantitative reverse transcriptase PCR (qRT-PCR)*
607 RNA was extracted from *S*. Typhimurium strain SL1344 stationary phase cultures
608 (O.D.600 ~2) using the SV total RNA isolation system (Promega Co., Madison, WI) as
609 described at http://www.ifr.ac.uk/safety/microarrays/protocols.html. The quantity and
610 quality of the extracted RNA were determined using a ND-1000 spectrophotometer
611 (NanoDrop Technologies, Wilmington, DE). To diminish genomic DNA contamination,
612 the preparation was treated with DNase I (Turbo DNA free, Applied Biosystems/Ambion,
613 Austin, TX). An aliquot of 0.6 µg of DNase-I-treated RNA was used for cDNA synthesis
614 using the High-Capacity cDNA Archive Kit (Applied Biosystems, Foster City, CA).

615 Quantitative RT-PCR reactions were performed in LightCycler 480 II (Roche). Each
616 reaction was carried out in a total volume of 10 µl on a 480 well optical reaction plate
617 (Roche) containing 5 µl SYBR, 0.5 µl DYE II (Takara), 4.6 µl cDNA (1/10 dilution) and
618 two gene-specific primers at a final concentration of 0.2 mM each. Real-time cycling
619 conditions were as follows: (i) 95^{o}C for 10 min, (ii) 40 cycles at 95^{o}C for 15 s, 60^{o}C for 1
620 min. A non-template control was included for each primer set. Melting curve analysis
621 verified that each reaction contained a single PCR product. Gene expression levels were
622 normalized to transcripts of *rfaH* that served as an internal control. Gene-specific primers
623 were designed with PRIMER3 software (http://primer3.sourceforge.net) and are listed in
624 Table S5.
625
626 *Purification of LeuO protein*
627 The *leuO* gene was PCR-amplified using oligonucleotide primers pET21-leuO-BamHI
628 and pET21-leuO-SalI (Table S5). The amplification product was digested with BamHI
629 and SalI, and cloned in plasmid pET21a to yield pIZ1871. For LeuO protein purification,
630 *E. coli* BL21 (pIZ1871) was grown in YT broth, and adding 1 mM IPTG induced LeuO
631 expression. After 4 h of induction, cells were centrifuged and resuspended in lysis buffer
632 (20 mM Tris, 300 mM NaCl, 10 mM imidazole, 1 mM PMSF, 5 µl/ml protein inhibitor
633 cocktail [Sigma Chemical Co., St. Louis, MO], 10% glycerol) and lysed by sonication.
634 The suspension was centrifuged and the supernatant was mixed with Ni-agarose beads
635 (QIA express Type ATG Kit, Qiagen) and incubated overnight with gentle mixing on a
636 rotating wheel at 4^{o}C. Agarose beads were washed with washing buffer (20 mM Tris,
637 300 mM NaCl, 20 mM imidazole, 10% glycerol). Protein elution was performed with 500
638 µl of elution buffer (20 mM Tris, 300 mM NaCl, 200 mM imidazole, 10% glycerol).
639 Imidazole was removed by washing with storage buffer (20 mM Tris, 300 mM NaCl, 10%

glycerol) and centrifuged using Amicon® Ultra centrifugal filters. LeuO-6xHis protein was stored at -80^{o}C.

Gel mobility shift assays

DNA probes labelled with 6-caroxyfluorescein (6-FAM) were prepared by PCR amplification and primer pairs used are listed in Table S5. PCR products were purified with the Wizard® SV Clean-Up-System (Promega, Madison, WI). For gel shift assays, 50 ng of each FAM-labelled probe was incubated at room temperature for 30 minutes with increasing concentrations of LeuO-6xHis in a final volume of 20 µl. The binding buffer L10x contained 20 mM HEPES, 100 mM KCl, 2 mM $MgCl_2$, 0,1 mM EDTA and 20% glycerol (De la Cruz *et al.,* 2007). Protein-DNA complexes were subjected to electrophoresis at 4^{o}C in a 6% nondenaturing acrylamide:bisacrylamide (29:1) gel in 0.5 Tris-borate-EDTA buffer. DNA fragments were visualized with a FLA-5100 Imaging system (Fujifilm, Tokyo, Japan).

DNase I footprinting of LeuO binding to PCR-amplified DNA

The P_{envR} DNA probe labelled with 6-caroxyfluorescein (6-FAM) was prepared by PCR amplification and the primers used are listed in Table S5. DNase I footprinting was performed as described in Cameron and Dorman (2012) with some small changes. The DNase I footprinting reactions were conducted in 15 µl reaction volumes containing 1x DNase I buffer (Roche)(40 mM Tris-HCl, 10 mM NaCl, 6 mM $MgCl_2$, 1 mM $CaCl_2$, pH 7.9) 0.01 mM DTT, 100 ng/ µl BSA, 50 nM bait DNA, and 50 nM LeuO-6xHis. LeuO-DNA binding was allowed to equilibrate at 37ºC for 20 minutes, then 1 µl (0.04 units) of pre-warmed DNase I was added and mixed gently, then incubated at 37ºC for 5 minutes. Reactions were stopped by addition of 2 µl EDTA (100 mM) followed by vigorous vortex

665 mixing and heat denaturation at 95ºC for 10 minutes. Digestion products were desalted
666 using MicroSpin G-25 columns (GE Healthcare) and were analyzed on an ABI 3730
667 DNA Analyzer along with GeneScan 500-LIZ size standards (Applied Biosystems).

668 *Western immunoblotting*

669 Total proteins extracts were prepared from bacterial cultures grown at 37^{o}C in LPM or
670 LB medium until stationary phase was reached ($O.D._{600}$ ~1.4 and $O.D._{600}$ ~2
671 respectively). Bacterial cells contained in 1 ml of culture were collected by centrifugation
672 (16,000 g, 2 min, 4^{o}C) and suspended in 50 µl of Laemmli sample buffer [1.3% SDS,
673 10% (v/v) glycerol, 50 mM Tris-HCl, 1.8% β-mercaptoethanol, 0.02% bromophenol blue,
674 pH 6.8]. Proteins were resolved by Tris-Tricine-PAGE, using 12% gels. Conditions for
675 protein transfer have been described elsewhere (Balbontin *et al*., 2006). Primary
676 antibodies were anti-FLAG M2 monoclonal antibody (1:5,000, Sigma Chemical Co, St.
677 Louis, MO), anti-DnaK monoclonal antibody (1:5,000, MBL International, MA), and anti-
678 GroEL polyclonal antibody (1:10,000, Sigma Chemical Co, St. Louis, MO). Goat anti-
679 mouse horseradish peroxidase-conjugated antibody (1:5,000, BioRad, Hercules, CA)
680 was used as secondary antibody. Proteins recognized by the antibodies were visualized
681 by chemiluminiscence using luciferin-luminol.
682
683 *DNA sequence analysis*
684 To test for the presence of over-represented motifs in DNA sequences bound by LeuO in
685 *E. coli* and *S*. Typhimurium, the SELEX screening (Shimada *et al*., 2011) and ChIPOTle
686 datasets were manually curated to define short binding regions that could be analyzed
687 by the Meme motif-finding program. For the SELEX data this involved extracting 500 bp
688 of DNA sequence centred on the genomic coordinate presented for each LeuO binding
689 site in Shimada *et al*. (2011). For *S*. Typhimurium this involved selecting the highest-

690 scoring probes from the broader binding regions identified by ChIPOTle and the
691 corresponding DNA sequences were extracted using the Artemis genome viewer
692 (Rutherford *et al.,* 2000). The unbiased motif-finding program Meme (Bailey *et al.,* 2009)
693 was used to search the curated datasets. Meme parameters were set as follows: motifs
694 could range in size from 10 to 50 bp, each DNA sequence could contain multiple or no
695 motif sites, and both palindromic and non-palindromic motifs could be found. The MAST
696 program (Bailey *et al*., 2009) was used to generate position specific scoring matrix
697 (PSSM) from the *E. coli* LeuO and *S*. Typhimurium motifs. The PSSMs were used to
698 scan the *E. coli* K-12 MG1655 and *S*. Typhimurium SL1344 genome sequences for
699 matches with an E-value < 0.1 and a position p-value < 0.0001. The matching
700 sequences are listed in Table S3.
701
702 **Acknowledgments**
703 This work was supported by Science Foundation Ireland grant 07/IN.1/B918 (Dorman
704 Laboratory) and by grants BIO2010-15023 and CSD2008-00013 from the Ministerio de
705 Ciencia e Innovación of Spain (MICINN) and the European Regional Fund (Casadesús
706 Laboratory). Elena Espinosa holds a FPI Ph.D. fellowship from the MICINN. We thank
707 Andrew Cameron for advice on DNase I footprinting and the Trinity Centre for High-
708 Performance computing for providing computational resources and Saccha Lucchini and
709 Jay Hinton for strain SL1344 *hns*::*kan,* and for scientific discussions. We also thank
710 Modesto Carballo, Laura Navarro, and Cristina Reyes of the Servicio de Biología
711 (CITIUS, Universidad de Sevilla) for help in experiments performed at the facility. The
712 authors declare that they have no conflicts of interest.
713
714

References

Bailey, T. L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., *et al.* (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* **37**: W202-208.

Balbontin, R., Rowley, G., Pucciarelli, M.G., Lopez-Garrido, J., Wormstone, Y., Lucchini, S., *et al.* (2006) DNA adenine methylation regulates virulence gene expression in *Salmonella enterica* serovar Typhimurium. *J Bacteriol* **188:** 8160-8168.

Beveridge, D.L., Dixit, S.B., Barreiro G., and Thayer, K.M. (2004) Molecular dynamics simulations of DNA curvature and flexibility: helix phasing and premelting. *Biopolymers* **73**: 380-403.

Buck, M.J., Nobel A.B., and Lieb, J.D. (2005) ChIPOTle: a user-friendly tool for the analysis of ChIP-chip data. *Genome Biol* **6**: R97.

Cameron, A.D.S., and Dorman, C.J. (2012) A fundamental regulatory mechanism operating through OmpR and DNA topology controls gene expression in *Salmonella* pathogenicity islands SPI-1 and SPI-2. *PLoS Genetics* **8:** e1002615.

Chen, C.C., Chou, M.Y., Huang, C.H., Majumder, A., and Wu, H.Y. (2005) A *cis*-spreading nucleoprotein filament is responsible for the gene silencing activity found in the promoter relay mechanism. *J Biol Chem* **280**: 5101-5112.

Chen, C.C., Fang, M., Majumder A., and Wu, H.Y. (2001) A 72-base pair AT-rich DNA sequence element functions as a bacterial gene silencer. *J Biol Chem* **276**: 9478-9485.

Chen, C.C., Ghole, M., Majumder, A., Wang, Z., Chandana S., and Wu, H.Y. (2003) LeuO-mediated transcriptional derepression. *J Biol Chem* **278**: 38094-38103.

738 Chen, C.C., and Wu, H.Y. (2005) LeuO protein delimits the transcriptionally active and
739 repressive domains on the bacterial chromosome. *J Biol Chem* **280**: 15111-
740 15121.
741 Chen, D., Bowater, R., Dorman, C.J., and Lilley, D.M.J. (1992) Activity of a plasmid-
742 borne *leu500* promoter depends on the transcription and translation of an
743 adjacent gene. *Proc Natl Acad Sci USA* **89:** 8784-8788.
744 Cherepanov P.P., and Wackernagel, W. (1995) Gene disruption in *Escherichia coli*: Tc^r
745 and Km^r cassettes with the option of Flp-catalyzed excision of the antibiotic-
746 resistance determinant. *Gene* **158:** 9-14.
747 Cho, B.K., Knight, E.M., Barrett, C.L., and Palsson, B.O. (2008) Genome-wide analysis
748 of Fis binding in *Escherichia coli* indicates a causative role for A-/AT-tracts.
749 *Genome Res* **18**: 900-910.
750 Dame, R.T., Wyman, C., Wurm, R., Wagner, R., and Goosen, N. (2002) Structural basis
751 for H-NS-mediated trapping of RNA polymerase in the open initiation complex at
752 the *rrnB* P1. *J Biol Chem* **277**: 2146-2150.
753 Dame, R.T., Noom, M.C., and Wuite, G.J. (2006) Bacterial chromatin organization by H-
754 NS protein unraveled using dual DNA manipulation. *Nature* **444:** 387-390.
755 De la Cruz, M.A., Fernandez-Mora, M., Guadarrama, C., Flores-Valdez, M.A.,
756 Bustamante, V.H., Vazquez A., *et al*. (2007) LeuO antagonizes H-NS and StpA-
757 dependent repression in *Salmonella enterica ompS1*. *Mol Microbiol* **66**: 727-743.
758 Dillon, S.C., Cameron, A.D., Hokamp, K., Lucchini, S., Hinton, J.C., and Dorman, C.J.
759 (2010) Genome-wide analysis of the H-NS and Sfh regulatory networks in
760 *Salmonella* Typhimurium identifies a plasmid-encoded transcription silencing
761 mechanism. *Mol Microbiol* **76**: 1250-1265.
762 Dillon, S.C., and Dorman, C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid
763 structure and gene expression. *Nat Rev Microbiol* **8**: 185-195.

764 Dorman, C.J. (2004) H-NS, a universal regulator for a dynamic genome. *Nature Rev*
765 *Microbiol* **2:** 391-400.
766 Fang, M., Majumder, A., Tsai K.J., and Wu, H.Y. (2000) ppGpp-dependent *leuO*
767 expression in bacteria under stress. *Biochem Biophys Res Commun* **276**: 64-70.
768 Fernandez-Mora, M., Puente, J.L., and Calva, E. (2004) OmpR and LeuO positively
769 regulate the *Salmonella enterica* serovar Typhi *ompS2* porin gene. *J Bacteriol*
770 **186**: 2909-2920.
771 Gallego-Hernandez, A. L., Hernandez-Lucas, I., De la Cruz, M. A., Olvera, L., Morett, E.,
772 Medina-Aparicio, L., *et al*. (2012) Transcriptional regulation of the *assT-dsbL-*
773 *dsbI* gene cluster in *Salmonella enterica* serovar Typhi IMSS-1 depends on
774 LeuO, H-NS, and specific growth conditions. *J Bacteriol* **194**: 2254-2264.
775 Garvie, C.W., and Wolberger, C. (2001) Recognition of specific DNA sequences. *Mol*
776 *Cell* **8**: 937-946.
777 Grainger, D.C., Aiba, H., Hurd, D., Browning, D.F., and Busby, S.J.W. (2007)
778 Transcription factor distribution in *Escherichia coli*: studies with FNR protein.
779 *Nucleic Acids Res* **35**: 269-278.
780 Grainger, D.C., Hurd, D., Goldberg, M.D., and Busby, S.J.W. (2006) Association of
781 nucleoid proteins with coding and non-coding segments of the *Escherichia coli*
782 genome. *Nucleic Acids Res* **34**: 4642-4652.
783 Grainger, D.C., Hurd, D., Harrison, M., Holdstock, J., and Busby, S.J.W. (2005) Studies
784 of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase
785 along the *E. coli* chromosome. *Proc Natl Acad Sci USA* **102**: 17693-17698.
786 Grainger, D.C., Overton, T.W., Reppas, N., Wade, J.T., Tamai, E., Hobman, J.L., *et al.*
787 (2004) Genomic studies with *Escherichia coli* MelR protein: applications of
788 chromatin immunoprecipitation and microarrays. *J Bacteriol* **186**: 6938-6943.

789 Grob, P., Kahn, D., and Guiney, D.G. (1997) Mutational characterization of promoter
790 regions recognized by the *Salmonella dublin* virulence plasmid regulatory protein
791 SpvR. *J Bacteriol* **179**: 5398-5406.
792 Henikoff, S., Haughn, G.W., Calvo, J.M., and Wallace, J.C. (1988) A large family of
793 bacterial activator proteins. *Proc Natl Acad Sci USA* **85**: 6602-6606.
794 Hernandez-Lucas, I., Gallego-Hernandez, A.L., Encarnacion, S., Fernandez-Mora, M.,
795 Martinez-Batallar, A.G., Salgado, H., *et al*. (2008) The LysR-type transcriptional
796 regulator LeuO controls expression of several genes in *Salmonella enterica*
797 serovar Typhi. *J Bacteriol* **190**: 1658-1670.
798 Hizver, J., Rozenberg, H., Frolow, F., Rabinovich, D., and Shakked, Z. (2001) DNA
799 bending by an adenine-thymine tract and its role in gene regulation. *Proc Natl*
800 *Acad Sci USA* **98**: 8490-8495.
801 Hoiseth, S.K., and Stocker, B.A. (1981) Aromatic-dependent *Salmonella typhimurium*
802 are non-virulent and effective as live vaccines. *Nature* **291:** 238-239.
803 Hryniewicz M.M., and Kredich, N.M. (1994) Stoichiometry of binding of CysB to the
804 *cysJIH*, *cysK*, and *cysP* promoter regions of *Salmonella typhimurium*. *J Bacteriol*
805 **176**: 3673-82.
806 Karginov, F.V., and Hannon, G.J. (2010) The CRISPR system: small RNA-guided
807 defense in bacteria and archaea. *Mol Cell* **37**: 7-19.
808 Kim, M., Lim, S., Kim, D., Choy, H.E., and Ryu, S. (2009) A *tdcA* mutation reduces the
809 invasive ability of *Salmonella enterica* serovar Typhimurium. *Mol Cell* **28**: 389-
810 395.
811 Klauck, E., Bohringer, J., and Hengge-Aronis, R. (1997) The LysR-like regulator LeuO in
812 *Escherichia coli* is involved in the translational regulation of *rpoS* by affecting the
813 expression of the small regulatory DsrA-RNA. *Mol Microbiol* **25**: 559-569.

814 Klauck, E., Lingnau, M., and Hengge-Aronis, R. (2001) Role of the response regulator
815 RssB in sigma recognition and initiation of sigma proteolysis in *Escherichia coli*.
816 *Mol Microbiol* **40**: 1381-1390.
817 Knapp, G.S., and Hu, J.C. (2010) Specificity of the *E. coli* LysR-type transcriptional
818 regulators. *PLoS One* **5**: e15189.
819 Kröger, C., Dillon, S.C., Cameron, A.D.S., Papenfort, K., Sivasankaran, S.K., Hokamp,
820 K., *et al.* (2012) The transcriptional landscape and small RNAs of *Salmonella*
821 *enterica* serovar Typhimurium. *Proc Natl Acad Sci USA* **109:** e1277-1286.
822 Kullik, I., Stevens, J., Toledano, M.B., and Storz, G. (1995) Mutational analysis of the
823 redox-sensitive transcriptional regulator OxyR: regions important for DNA binding
824 and multimerization. *J Bacteriol* **177**: 1285-1291.
825 Lahiri, A., Das, P., and Chakravortty, D., (2008) The LysR-type transcriptional regulator
826 Hrg counteracts phagocyte oxidative oxidative burst and imparts survival
827 advantage to *Salmonella enterica* serovar Typhimurium. *Microbiol* **154:** 2837-
828 2846.
829 Lahiri, A., Das, P., and Chakravortty, D. (2009) *Salmonella* Typhimurium: insight into the
830 multi-faceted role of the LysR-type transcriptional regulators in *Salmonella*. *Int J*
831 *Biochem Cell Biol* **41**: 2129-2133.
832 Lawley, T.D., Chan, K., Thompson, L.J., Kim, C.C., Govoni, G.R., and Monack, D.M.
833 (2006) Genome-wide screen for *Salmonella* genes required for long-term
834 systemic infection of the mouse. *PLoS Pathog* **2**: e11.
835 Lawrenz, M.B., and Miller, V.L. (2007) Comparative analysis of the regulation of *rovA*
836 from the pathogenic yersiniae. *J Bacteriol* **189**: 5963-5975.
837 Lee, C.C., Wozniak, C., Karlinsey, J.E., and Hughes, K.T. (2007) Genomic screening for
838 regulatory genes using the T-POP transposon. *Methods Enzymol* **421**: 159-167.

839 Lucchini, S., Rowley, G., Goldberg, M.D., Hurd, D., Harrison, M., and Hinton, J.C. (2006)
840 H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog*
841 **2**: e81.
842 Maddocks, S.E., and Oyston, P.C. (2008) Structure and function of the LysR-type
843 transcriptional regulator (LTTR) family proteins. *Microbiology* **154**: 3609-3623.
844 Majdalani, N., and Gottesman, S. (2005) The Rcs phosphorelay: a complex signal
845 transduction system. *Annu Rev Microbiol* **59**: 379-405.
846 Majumder, A., Fang, M., Tsai, K.J., Ueguchi, C., Mizuno, T., and Wu, H.Y. (2001) LeuO
847 expression in response to starvation for branched-chain amino acids. *J Biol*
848 *Chem* **276**: 19046-19051.
849 Maxon, M.E., Redfield, B., Cai, X.Y., Shoeman, R., Fujita, K., Fisher, W., *et al*. (1989)
850 Regulation of methionine synthesis in *Escherichia coli*: effect of the MetR protein
851 on the expression of the *metE* and *metR* genes. *Proc Natl Acad Sci USA* **86**: 85-
852 89.
853 Medina-Aparicio, L., Rebollar-Flores, J.E., Gallego-Hernandez, A.L., Vazquez, A.,
854 Olvera, L., Gutierrez-Rios, R.M., *et al*. (2011) The CRISPR/Cas immune system
855 is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein
856 in *Salmonella enterica* serovar Typhi. *J Bacteriol* **193**: 2396-2407.
857 Momany, C., and Neidle, E.L. (2012) Defying stereotypes: the elusive search for a
858 universal model of LysR-type regulation. *Mol Microbiol* **83:** 453-456.
859 Nagarajavel, V., Madhusudan, S., Dole, S., Rahmouni, A.R., and Schnetz, K. (2007)
860 Repression by binding of H-NS within the transcription unit. *J Biol Chem* **282**:
861 23622-30.

862 Navarre, W.W., Porwollik, S., Wang, Y., McClelland, M., Rosen, H., Libby, S.J., *et al*.
863 (2006) Selective silencing of foreign DNA with low GC content by the H-NS
864 protein in *Salmonella*. *Science* **313**: 236-238.
865 O'Byrne, C.P., and Dorman, C.J. (1994) The *spv* virulence operon of *Salmonella*
866 *typhimurium* LT-2 is regulated negatively by the cAMP-CRP system. *J Bacteriol*
867 **176:** 905-912.
868 Pareja, E., Pareja-Tobes, P., Manrique, M., Pareja-Tobes, E., Bonal, J., and Tobes, R.
869 (2006) ExtraTrain: a database of Extragenic regions and Transcriptional
870 information in prokaryotic organisms. *BMC Microbiol* **6**: 29.
871 Parsek, M.R., Ye, R.W., Pun, P., and Chakrabarty, A.M. (1994) Critical nucleotides in the
872 interaction of a LysR-type regulator with its target promoter region: *catBC*
873 promoter activation by CatR. *J Biol Chem* **269**: 11279-11284.
874 Perez, J.C., Latifi, T., and Groisman, E.A. (2008) Overcoming H-NS-mediated
875 transcriptional silencing of horizontally acquired genes by the PhoP and SlyA
876 proteins in *Salmonella enterica. J Biol Chem* **283:** 10773-10783.
877 Perez, J.C., and Groisman, E.A. (2009) Evolution of transcriptional regulatory circuits in
878 bacteria. *Cell* **138**: 233-244.
879 Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., *et al*. (2012)
880 The Pfam protein families database. *Nucleic Acids Res* **40**: D290-301.
881 Rahav-Manor, O., Carmel, O., Karpel, R., Taglicht, D., Glaser, G., Schuldiner, S., *et al*.
882 (1992) NhaR, a protein homologous to a family of bacterial regulatory proteins
883 (LysR), regulates *nhaA*, the sodium proton antiporter gene in *Escherichia coli*. *J*
884 *Biol Chem* **267**: 10433-10438.
885 Repoila, F., and Gottesman, S. (2001) Signal transduction cascade for regulation of
886 RpoS: temperature regulation of DsrA. *J Bacteriol* **183**: 4012-23.

887 Rodriguez-Morales, O., Fernandez-Mora, M., Hernandez-Lucas, I., Vazquez, A., Puente,
888 J.L., and Calva, E. (2006) *Salmonella enterica* serovar Typhimurium *ompS1* and
889 *ompS2* mutants are attenuated for virulence in mice. *Infect Immun* **74**: 1398-
890 1402.
891 Rohs, R., West, S.M., Sosinsky, A., Liu, P., Mann, R.S., and Honig, B. (2009) The role of
892 DNA shape in protein-DNA recognition. *Nature* **461**: 1248-1253.
893 Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., and
894 Barrell, B. (2000) Artemis: sequence visualization and annotation. *Bioinformatics*
895 **16:** 944-945.
896 Schell, M.A. (1993) Molecular biology of the LysR family of transcriptional regulators.
897 *Annu Rev Microbiol* **47**: 597-626.
898 Sheehan, B.J., and Dorman, C.J. (1998) *In vivo* analysis of the interactions of the LysR-
899 like regulator SpvR with the operator sequences of the *spvA* and *spvR* virulence
900 genes of *Salmonella typhimurium. Mol Microbiol* **30:** 91-105.
901 Shi, X., and Bennett, G.N. (1995) Effects of multicopy LeuO on the expression of the
902 acid-inducible lysine decarboxylase gene in *Escherichia coli*. *J Bacteriol* **177**:
903 810-814.
904 Shimada, T., Bridier, A., Briandet R., and Ishihama, A. (2011) Novel roles of LeuO in
905 transcription regulation of *E. coli* genome: antagonistic interplay with the
906 universal silencer H-NS. *Mol Microbiol* **82**: 378-397.
907 Shimada, T., Yamamoto, K., and Ishihama, A. (2009) Involvement of the leucine
908 response transcription factor LeuO in regulation of the genes for sulfa drug efflux.
909 *J Bacteriol* **191**: 4562-4571.
910 Sternberg, N.L., and Maurer, R. (1991) Bacteriophage-mediated transduction. *Methods*
911 *Enzymol* **204:** 18-43.

912 Stoebel, D.M., Free, A., and Dorman, C.J. (2008) Anti-silencing: overcoming H-NS-
913 mediated repression of transcription in Gram-negative enteric bacteria.
914 *Microbiology* **154**: 2533-2545.
915 Stratmann, T., Madhusudan, S., and Schnetz, K. (2008) Regulation of the *yjjQ-bglJ*
916 operon, encoding LuxR-type transcription factors, and the divergent *yjjP* gene by
917 H-NS and LeuO. *J Bacteriol* **190**: 926-935.
918 Stratmann, T., Pul, U., Wurm, R., Wagner R., and Schnetz, K. (2012) RcsB-BglJ
919 activates the *Escherichia coli leuO* gene, encoding an H-NS antagonist and
920 pleiotropic regulator of virulence determinants. *Mol Microbiol* **83:** 1109-1123.
921 Tenor, J.L., McCormick, B.A., Ausubel, F.M., and Aballay, A. (2004) *Caenorhabditis*
922 *elegans*-based screen identifies *Salmonella* virulence factors required for
923 conserved host-pathogen interactions. *Curr Biol* **14**: 1018-1024.
924 Turnbull, A. L., Kim W., and Surette, M.G. (2012) Transcriptional regulation of *sdiA* by
925 cAMP-receptor protein, LeuO, and environmental signals in *Salmonella enterica*
926 serovar Typhimurium. *Can J Microbiol* **58**: 10-22.
927 Ueguchi, C., Ohta, T., Seto, C., Suzuki, T., and Mizuno, T. (1998) The *leuO* gene
928 product has a latent ability to relieve *bgl* silencing in *Escherichia coli*. *J Bacteriol*
929 **180**: 190-193.
930 VanBogelen, R.A., Olson, E.R., Wanner, B.L., and Neidhardt, F.C. (1996) Global
931 analysis of proteins synthesized during phosphorus restriction in *Escherichia coli*.
932 *J Bacteriol* **178**: 4344-4366.
933 Van Rhijn, P., and Vanderleyden, J. (1995) The Rhizobium-plant symbiosis. *Microbiol*
934 *Rev* **59**: 124-142.
935 Viswanathan, P., Ueki, T., Inouye, S., and Kroos, L. (2007) Combinatorial regulation of
936 genes essential for *Myxococcus xanthus* development involves a response
937 regulator and a LysR-type regulator. *Proc Natl Acad Sci USA* **104**: 7969-7974.

938 Waldminghaus, T., and Skarstad, K. (2010) ChIP on Chip: surprising results are often
939 artifacts. *BMC Genomics* **11**: 414.
940 Walthers, D., Li, Y., Liu, Y., Anand, G., Yan, J., and Kenney, L.J. (2011) *Salmonella*
941 *enterica* response regulator SsrB relieves H-NS silencing by displacing H-NS
942 bound in polymerization mode and directly activates transcription. *J Biol Chem*
943 **286:** 1895-1902.
944 Wei, Q., Le Minh, P.N., Dotsch, A., Hildebrand, F., Panmanee, W., Elfarash, A., *et al*.
945 (2012) Global regulation of gene expression by OxyR in an important human
946 opportunistic pathogen. *Nucleic Acids Res* **40:** 4320-4333.
947 Westra, E.R., Pul, U., Heidrich, N., Jore, M.M., Lundgren, M., Stratmann, T., *et al*. (2010)
948 H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12
949 can be relieved by the transcription activator LeuO. *Mol Microbiol* **77**: 1380-1393.
950 Wilson, R.L., Urbanowski, M.L., and Stauffer, G.V. (1995) DNA binding sites of the LysR-
951 type regulator GcvA in the *gcv* and *gcvA* control regions of *Escherichia coli*. *J*
952 *Bacteriol* **177**: 4940-4946.
953 Wu, H.Y., Tan, J., and Fang, M. (1995) Long-range interaction between two promoters:
954 activation of the *leu500* promoter by a distant upstream promoter. *Cell* **82:** 445-
955 451.
956 Wu, H.Y., and Fang, M. (2003) DNA supercoiling and transcription control: a model from
957 the study of suppression of the *leu-500* mutation in *Salmonella typhimurium topA*$^{-}$
958 strains. *Prog Nucleic Acid Res Mol Biol* **73**: 43-68.
959 Zaim, J., and Kierzek, A.M. (2003) The structure of full-length LysR-type transcriptional
960 regulators. Modeling of the full-length OxyR transcription factor dimer. *Nucleic*
961 *Acids Res* **31:** 1444-1454.
962
963

Figures

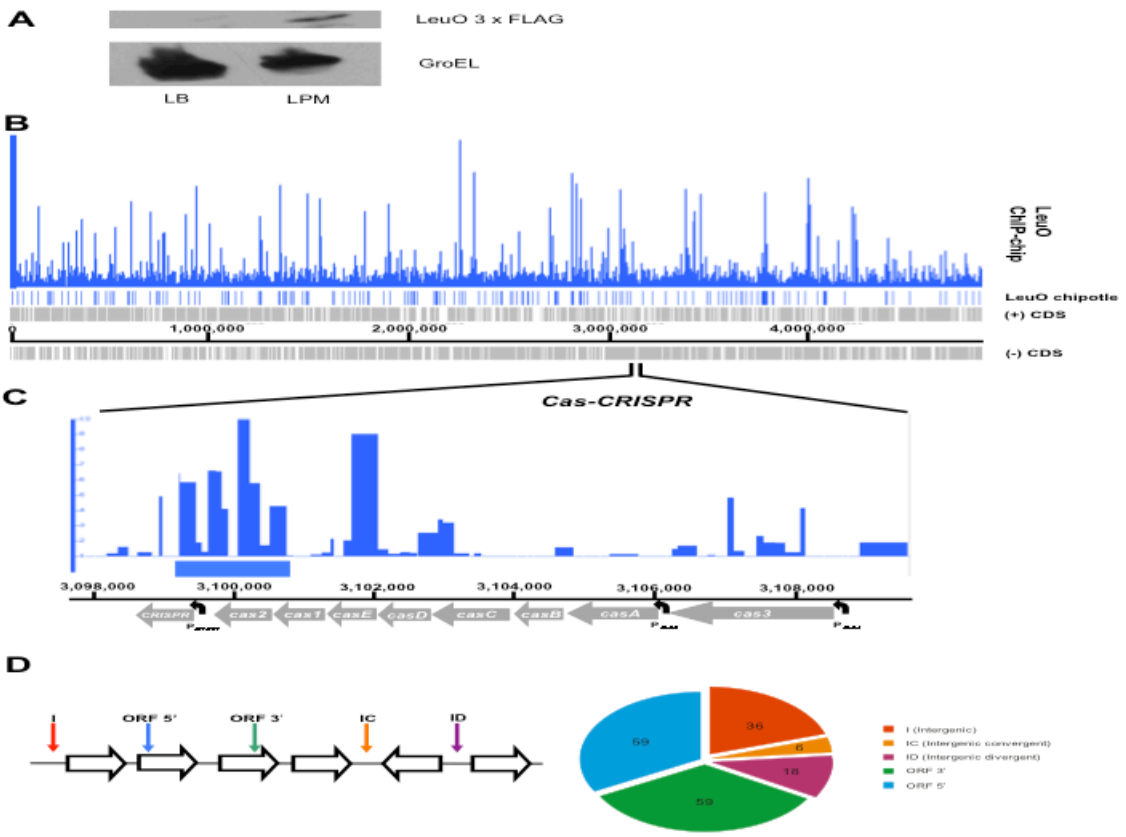


Fig. 1: Global identification and categorization of LeuO binding sites by ChIP-chip

(A) Western Immunoblot analysis of LeuO protein levels in SL1344 cells grown to stationary phase in LB or LPM are shown in the top panel. GroEL loading controls are indicated in the bottom panel.

(B) Visualization of LeuO ChIP-chip data using the Integrated Genome Browser (IGB) for *S*. Typhimurium SL1344. The location of LeuO binding sites, as defined by the ChIPOTle algorithm, are indicated by horizontal bars in the LeuO ChIPOTle track. The locations of known coding sequences (CDS) on the plus (+) and minus (-) DNA strands and SL1344 chromosome coordinates are also shown.

(C) Detailed view of the *S*. Typhimurium CRISPR/Cas locus. LeuO ChIP-chip data are presented quantitatively, with enrichment ratios shown on the y-axis. The binding site

identified by ChIPOTle is depicted by the blue rectangle and known promoter locations are indicated by bent arrows. The *cas3* transcription start site was determined by Kröger *et al*. (2012).

(D) Schematic representation of LeuO binding site classification as outlined in the text. The pie chart shows the relative distribution of LeuO binding sites among the location classes summarized in the genetic diagram on the left, with horizontal open arrows used to represent ORFs and their relative orientations. The colours of the vertical arrows correspond to the colours used in the pie chart segments.

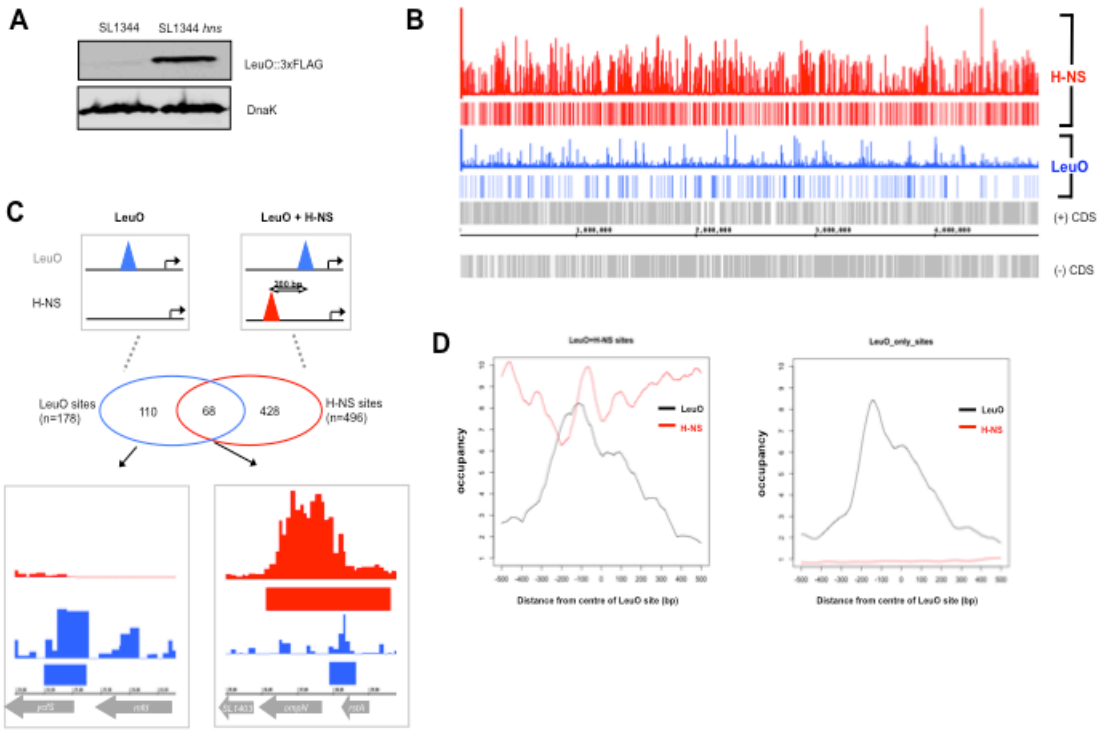


Fig. 2: Integration of H-NS and LeuO genome-wide binding data

(A) Western Immunoblot analysis of LeuO protein levels in wild type SL1344 and SL1344 *hns* are shown in the top panel. DnaK loading controls are indicated in the bottom panel.

(B) Visualization of H-NS (red) and LeuO (blue) ChIP-chip data in the Integrated Genome Browser with ChIPOTle identified binding sites depicted below each lane as horizontal bars. The locations of known coding sequences (CDS) on the plus (+) and minus (-) DNA strands and SL1344 chromosome coordinates are also shown.

(C) Schematic representation of LeuO and H-NS overlap analysis. The Venn diagram illustrates the number of LeuO sites that did (LeuO+H-NS) and did not (LeuO) overlap with an H-NS binding site. Quantitative LeuO (blue) and H-NS (red) ChIP-chip data are shown for representative examples of LeuO and LeuO+H-NS binding sites.

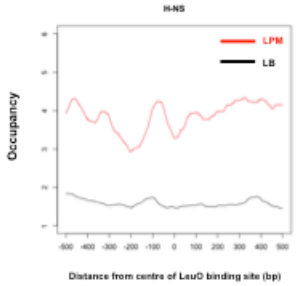
(D) Average plots of LeuO and H-NS occupancy (ChIP-chip enrichment ratios) at LeuO+H-NS and LeuO-only sites. Averaged LeuO and H-NS data were plotted in 100-bp windows with respect to the center of ChiPOTle-defined LeuO binding sites.

A

LeuO LPM target genes (n=178)

	H-NS only	H-NS + RNAP	RNAP only	None
LPM	5	63	105	5
LB	30	11	50	46

B



C

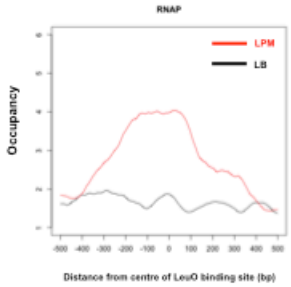


Fig. 3: LeuO does not simply displace H-NS from DNA and may be involved in the recruitment of RNA Polymerase to target genes

(A) The number of H-NS, RNA Polymerase (RNAP) and LeuO co-localization events in LeuO-inducing conditions (LPM) is indicated in the first row of the table. The second row indicates the number of H-NS and RNAP binding events at the LPM defined LeuO sites in non-inducing conditions (LB). (B) An average plot of H-NS occupancy with respect to LeuO binding site location in LeuO inducing conditions (LPM) and non-inducing conditions (LB). (C) An average plot of RNAP occupancy with respect to LeuO binding site location in LeuO inducing conditions (LPM) and non-inducing conditions (LB).

A

Species	Dataset	Number of Sequences	Average Size (bp)	Total Size (bp)	% A+T	Source
Salmonella Typhimurium	LeuO+ HNS	64	653	41825	50.5	This study
Escherichia coli	E. coli LeuO	119	501	59619	59.9	Shimada *et al.* 2011

B

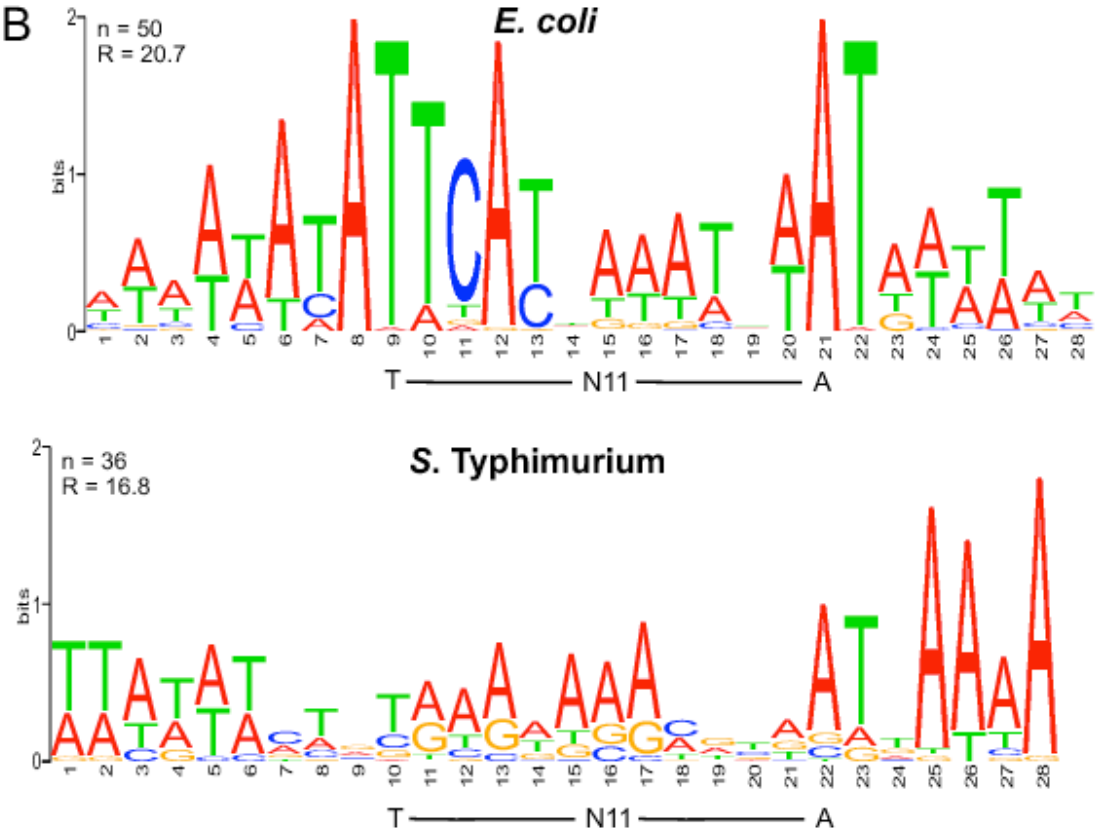
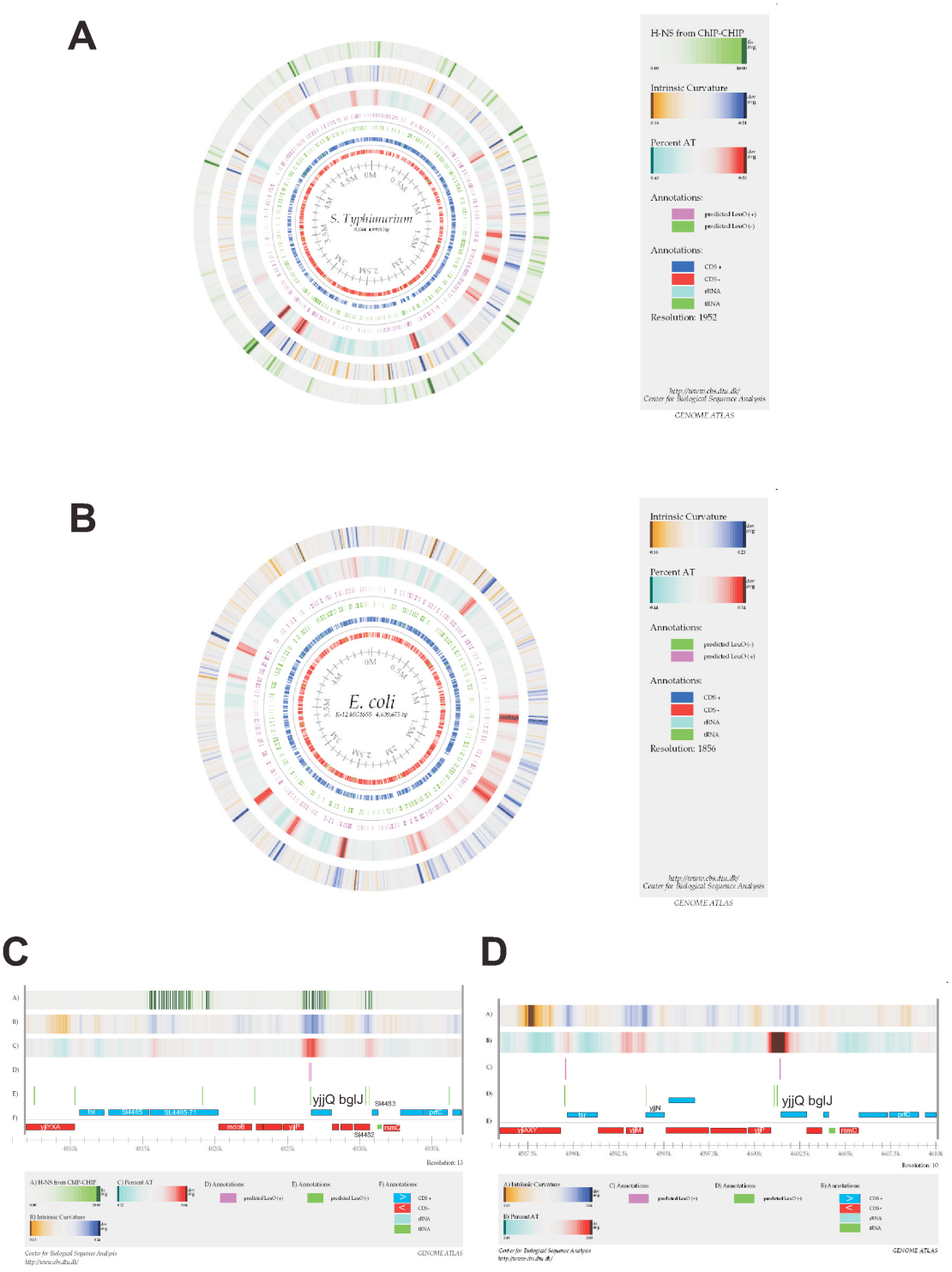


Fig. 4: Similarities and differences in the LeuO binding site motifs of S.

Typhimurium and E. coli

(A) Details of the manually curated *S*. Typhimurium and *E. coli* LeuO datasets used to derive the LeuO binding motifs. (B) Alignment of sequence logos illustrating the *E. coli* and *S*. Typhimurium motifs identified by Meme. The location of the LTTR T-$_{N11}$-A box is indicated below each Logo.



1024
1025 **Fig. 5: Predicted LeuO binding sites in the context of genome atlases for the S.**

Typhimurium and *E. coli* chromosomes

(A+B) The locations of predicted LeuO binding sites on both strands of the *S.* Typhimurium SL1344 (A) *and E. coli* MG1655 (B) chromosomes are shown using coloured concentric circles. In addition, we show published H-NS ChIP-chip data from SL1344 grown in LPM together with DNA intrinsic curvature measurements and percentage A+T content (http://www.cbs.dtu.dk/services/GenomeAtlas). The colour code used in each concentric circle is explained in the key at the right of each diagram. Numbers on the inside of the innermost circle are locations relative to position zero measured in millions of base-pairs (Mbp). (C+D) A detailed view of a predominantly syntenic region of the *S*. Typhimurium (C) and *E. coli* (D) chromosomes illustrating the similarities and differences of predicted LeuO binding sites in the two species. The locations of predicted LeuO binding sites on the plus and negative strands are indicated by purple and green bars in tracks D and E. The locations of CDS on the plus and negative strands are indicated by blue and red rectangles, respectively.

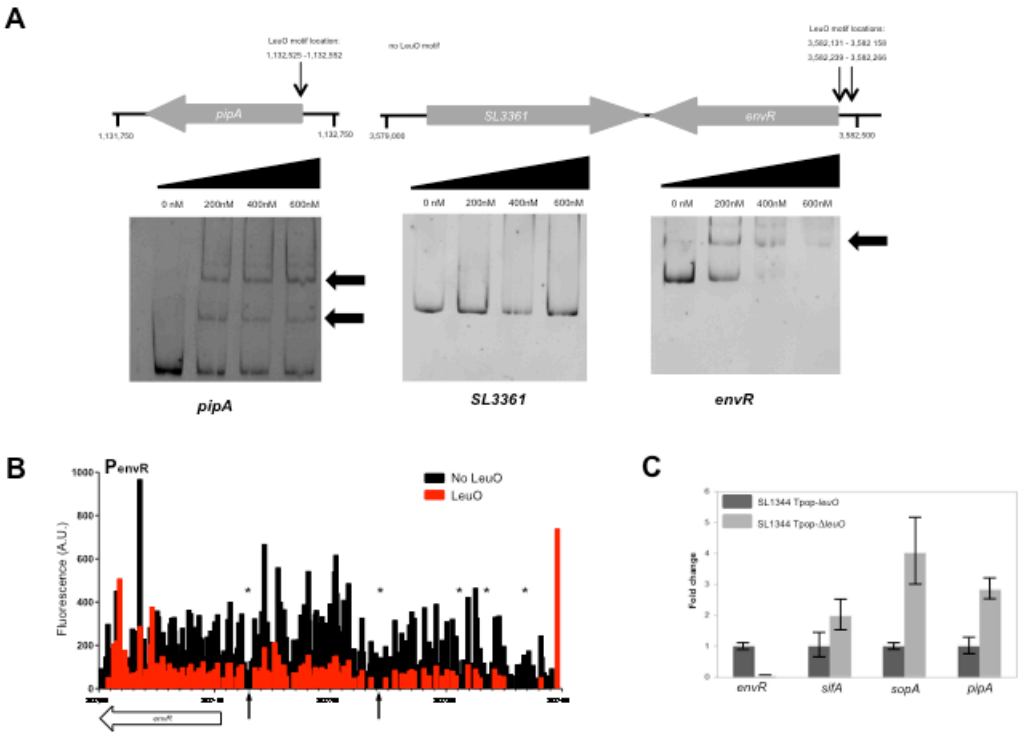


Fig. 6: Validation of predicted LeuO binding sites

(A) EMSA analysis was used to validate predicted LeuO binding sites. *pipA* and *envR* were associated with one and two predicted binding sites respectively (indicated by thin black arrows) while *SL3361* was not associated with a predicted site. DNA probes were incubated with increasing amounts of LeuO and complex formation is indicated by thick black arrows. (B) DNase I footprinting of LeuO binding to P_{envR} using end labelled DNA fragments. The size and quantity of 6-FAM-labelled DNase I digestion products were measured by fluorescent DNA sequencing with capillary electrophoresis. The locations of predicted LeuO binding sites are indicated by arrows and protected regions are indicated by asterisks. SL1344 chromosome coordinates are indicated on the x-axis. (C) qRT-PCR was used to monitor transcript levels for *envR, sifA, sopA and pipA* following *leuO* over-expression (SL1344 *Tpop-leuO*) and deletion of *leuO* (SL1344 *Tpop-ΔleuO*).

1054 Fold changes in transcript levels are expressed relative to strain SL1344 *Tpop-leuO*,
1055 which is normalized to 1.