A Novel Mechanism of Transposon-Mediated Gene Activation

Zhongge Zhang, Milton H. Saier, Jr.*
Division of Biological Sciences, University of California San Diego, La Jolla, California, United States of America

Abstract
Transposable Insertion Sequences (IS elements) have been shown to provide various benefits to their hosts via gene activation or inactivation under stress conditions by appropriately inserting into specific chromosomal sites. Activation is usually due to derepression or introduction of a complete or partial promoter located within the element. Here we define a novel mechanism of gene activation by the transposon IS5 in Escherichia coli. The glycerol utilization operon, glpFK, that is silent in the absence of the cAMP-Crp complex, is activated by IS5 when inserted upstream of its promoter. High-level expression is nearly constitutive, only mildly dependent on glycerol, glucose, GlpR, and Crp, and allows growth at a rate similar to or more rapid than that of wild-type cells. Expression is from the glpFK promoter and dependent on (1) the DNA phase, (2) integration host factor (IHF), and (3) a short region at the 3' end of IS5 harboring a permanent bend and an IHF binding site. The lacZYA operon is also subject to such activation in the absence of Crp. Thus, we have defined a novel mechanism of gene activation involving transposon insertion that may be generally applicable to many organisms.

Introduction

Living organisms possess a variety of mutagenic means to generate genetic diversity, and these depend on environmental conditions and genomic composition [1,2]. One frequently encountered type of mutation results from the insertion of transposable elements, transposons, which when inserted in appropriate locations of the genome, can activate or inactivate critical genes [3,4]. One transposon-mediated mechanism of gene activation involves the formation of a “hybrid promoter” when a small transposon, an Insertion Sequence (IS) element, inserts into the promoter region. In this case, insertion of an IS element results in placing an outwardly directed −35 hexamer in one of the terminal inverted repeats (IRs) of the transposon at the correct distance from a resident −10 hexamer. Such −35 elements have been observed experimentally in several ISs [5]. Activation can also occur by initiating transcription within the transposon, traversing the terminal IR and reading the gene of interest. This second type of mechanism has been observed for IS3 [6] and IS10 [7].

A distinct type of gene regulation by ISs is illustrated by activation of the normally cryptic β-glucosidase (bgl) catalytic operon in E. coli. Activation of this operon can be accomplished in several ways, one of which involves insertion of either IS5 or IS1 upstream or downstream of the promoter [8,9]. For bgl operon activation, IS5 need not be in a specific position and orientation [9,10], and the nucleoid structuring protein H-NS is required [11,12].

The IS5 element has been found to activate the fucAO promoter [13], the flhDC promoter [14], and the ade promoter [15]. In wild type (wt) E. coli cells, expression of the fucAO operon is dependent on FucR, an activator of the fur genes, and the cyclic AMP receptor protein (Crp). Prolonged incubation of wt cells with L-1,2-propanediol (which does not activate FucR) produced mutants that could grow on this carbon source. These mutants express fucAO independently of FucR and harbor IS5 inserted upstream of the fucAO promoter, always in the same orientation with its ins5A promoter distal to fucA [13]. In the case of flhDC, mobility of wt cells on semisolid agar is substantially enhanced when IS5 is inserted at either of two locations (−99.5 and −169.5) upstream of the transcriptional start site. The IS5 orientations proved to be the same, with the ins5A promoter distal to the downstream flhDC promoter [14], as in the case of IS3 insertion in the fucAO promoter. In addition, IS5 has been shown to activate the cryptic ade gene encoding an adenine deaminase that catalyzes deamination of adenine to hypoxanthine in E. coli [15]. In the cases of flhDC and ade, activation is proposed to be due to the relief of HNS-mediated normal repression by IS5 insertion [14,15] but the mechanism of IS5 activation of the fucAO promoter is unknown [13].

E. coli can use glycerol, glycerol-3-phosphate, or glycerophosphodiester as sole carbon and energy sources. The loss of Crp abolishes its growth on glycerol. This is due to the fact that expression of one operon in the glp regulon, glpFK, encoding proteins essential for glycerol utilization [16,17], is strongly dependent on the Crp-cAMP complex (Figure 1) [18,19]. In addition, glpFK expression is repressed by the binding of GlpR, the glp regulon repressor, to the four operators in the glpFK promoter region [Figure 1] [20]. Repression is relieved in the presence of glycerol-3-phosphate, the inducer of the glp regulon. Our previous research [21] demonstrated that crp deletion mutants can mutate specifically to utilize glycerol (Glp−). The frequencies of such mutations are enhanced by the presence of glycerol and decreased...
by GlpR. Of the four GlpR operators (O1-O4) upstream of glpFK, O1 primarily controls mutation rate while O4 specifically controls glpFK expression. All Glp$^+$ mutants contain an IS5 upstream of the glpFK promoter, always in the same position and orientation [21].

Here, we describe a novel mechanism of gene activation due to IS5 insertion upstream of the promoter of the glpFK operon. Activation is so strong that it overcomes the repression resulting from the absence of Crp. A 177-bp fragment at the 3’ end of IS5, proximal to the downstream promoter, is both necessary and sufficient for full activation. Although for activation, IS5 always inserts into the same position upstream of the glpFK promoter, the active element, encompassing a permanent bend and an IHF binding site, must be present at the correct phase angle relative to the promoter to be effective. Finally, it is shown that in the absence of Crp, the E. coli lactose operon (lacZYA), can be similarly activated by IS5. Although this is the first study to show that unique sequences within a promoter region (IS5), can be similarly activated by IS5. Although this is the first study to show that unique sequences within a promoter, the activating effect is still present at the 10 bp fragment, but not a 5 bp fragment, is inserted between IS5 and the glpFK promoter. Thus, the active element, encompassing a permanent bend and an IHF binding site, must be present at the correct phase angle relative to the promoter to be effective. Finally, it is shown that in the absence of Crp, the E. coli lactose operon can also be activated by upstream IS5 insertion. These results reveal a novel mechanism of gene activation by insertion sequences.

Results

Effects of IS5 Insertion on Expression of glp Genes in Various Genetic Backgrounds

Using real-time PCR, we determined mRNA levels of the five glp operons comparing crp Glp$^+$ cells (with an IS5 insertion upstream of the glpFK promoter) [21] with parental crp cells. No differences were observed in expression levels of four of the five operons, but the glpFK operon showed a dramatic difference. Figure 2A shows expression of glpFK in wt, crp, and Glp$^+$ cells. In the absence of glycerol, glpFK expression was >30 fold higher in crp Glp$^+$ cells than in crp cells (see columns 3 and 5 from the left side), showing that the IS5 insertion led to high level expression of this operon. In the presence of glycerol, glpFK expression increased in all cell types examined, but expression was the highest in crp Glp$^+$ cells, about 15-fold higher than in the crp cells incubated under the same conditions. However, the degree of induction by glycerol in crp Glp$^+$ cells was greatly decreased, suggesting that in these cells, repression by GlpR is weak (~2.5 fold, see columns 5 and 6) compared to that in wt cells (>10 fold, see columns 1 and 2).

When the glpR gene was deleted from each of these three backgrounds, glycerol induction of glpFK expression was abolished (Figure 2B), indicating that increased glpFK expression in wt, crp and crp Glp$^+$ cells in response to glycerol is solely due to the release of GlpR from the control region of the operon.

We measured in vitro glycerol kinase activity in wt, crp and crp Glp$^+$ cells using [14C]glycerol as substrate. The highest levels of glycerol phosphorylation activity were observed in extracts of crp...
Determination of the Start Site of \textit{glpFK} Transcription in \textit{crp} Glp* Mutants

To determine if the presence of the IS5 element provides a new promoter driving \textit{glpFK} transcription, we measured the transcriptional initiation site of the operon using RNA ligase mediated RT-PCR. The total RNA was treated with tobacco acid pyrophosphatase (TAP) prior to cDNA synthesis. The 5' end region of the \textit{glpFK} cDNA was amplified using a pair of primers (Table S2), one (P\textit{glpFK}-exnt-R) binding to the adaptor sequence and the other (P\textit{glpFK}-exnt-F) binding to a region between 115th and 137th nucleotides downstream of the \textit{glpF} start codon. Two PCR products were obtained for both the wild type and the \textit{crp} Glp* strain (Figure 3A). Using P\textit{glpFK}-exnt-R as primer, DNA sequencing showed that the larger product was non-specific; the smaller one was the 5' end region of the \textit{glpFK} cDNA product (Figure 3B). The junction was found to correspond precisely to the 5' end mapped by primer extension (see Figure 1B) [18]. When the total RNA was not treated with TAP, only the nonspecific PCR product was obtained (see Figure 3A for both the wild type and the \textit{crp} Glp* strains. Activation of transcription by IS5 in \textit{crp} Glp* cells is therefore driven by the native \textit{glpFK} promoter. The start site was the same with or without glycerol (1%) in the growth medium.

Effects of Different Regions within IS5 on \textit{glpFK} Promoter Activity

Using chromosomal \textit{lacZ} fusions, we examined the effect of IS5 and different regions within IS5 on the \textit{glpFK} activity (Figure 1A and Figure 4A). Consistent with the RT-PCR results, the presence of IS5 upstream of \textit{PglpFK} (IS5:\textit{PglpFK}) in \textit{crp} cells dramatically increased the activity of the promoter regardless of the medium used. Both promoter-less IS5 (P-less IS5, in which the 98 bp 5' end region containing the promoter of the transposase gene was deleted), and the 177 bp 3' end region (called "Internal Bend", IB) of IS5 activated P\textit{glpFK} in \textit{crp} cells to an extent comparable to that observed with the intact IS5. However, the 130 bp 3' end region of IS5 (i.e., IB with 47 nucleotides removed from the 5' end) activated the promoter about 40% less efficiently than the 177 bp IB region (data not shown, see below for explanation).

Addition of glycerol to LB medium further increased the promoter activity, in agreement with the conclusion that GlpR weakly decreased IS5: P\textit{glpFK} activity. The increased level (~1.5 fold) of P\textit{glpFK} activity caused by the presence of glycerol in a \textit{crp} genetic background, measured by \textit{β}-galactosidase activity in a \textit{glpF-lacZ} fusion strain, was lower than that observed (~2.5 fold) for \textit{glpFK} mRNA in a \textit{crp} Glp* background (Figure 2A). Such a difference may be due to the fact that \textit{crp} Glp* cells transport and subsequently phosphorylate glycerol more rapidly than \textit{crp} cells. Addition of glucose to the medium slightly reduced the activities of the promoters tested, showing that the strong catabolite repression observed in the wild type strain was largely abolished. Further experiments reported below confirmed and extended this conclusion.

\textbf{IB}:\textit{PglpFK-lacZ} activity was subsequently characterized in other genetic backgrounds. High levels of \textit{β}-galactosidase activity were observed for \textit{wt}, \textit{crp}, \textit{glpR} and \textit{crp glpR} genetic backgrounds with cells grown in LB ± glycerol or ± glucose (Figure 4B). Similar levels of promoter activity were observed in \textit{wt} and \textit{crp} cells in the absence of glycerol, suggesting that the Crp protein does not appreciably influence \textbf{IB}:\textit{PglpFK} activity. Addition of glycerol increased the overall \textbf{IB}:\textit{PglpFK} activity by ~2 fold in \textit{wt} cells (column 1 and column 5 in Figure 4B) and ~1 fold in \textit{crp} cells (columns 2 and 6 in Figure 4B), while addition of glucose slightly
Mechanism of Transposon-Mediated Gene Activation

As demonstrated above, the IB segment has the same ability as the intact IS5 to activate PglpFK in the cbp genetic background. A putative IHF binding site is present in the middle of IB [25]. To determine if the host IHF protein plays a role in activation of PglpFK, we generated a null ihfA mutant in cbp and PglpFK double mutant backgrounds. These mutant strains were compared for growth in M9 minimal medium with glycerol as the sole carbon source (Figure 5A). No obvious difference in growth was found between wt and ihfA mutant cells. However, growth of cbp PglpFK cells was substantially reduced by the ihfA mutation.

To determine the effect of the loss of IHF on transcription of IB:PglpFK, an IB:PglpFK-lacZ fusion was transferred into a cbp ihfA double mutant genetic background by P1 transduction. β-Galactosidase assays were performed after growing cells in LB media with or without glycerol or glucose. As shown in Figure 5B, the activity of IB:PglpFK was reduced ~60% by the loss of IHF. Similar results were obtained when M9 + casamino acids + glucose medium was used (data not shown). These results are consistent with the growth data described above and indicate that IHF is required for maximal activation of PglpFK by the IB segment.

To determine if the putative IHF binding site in IB is responsible for IHF-mediated activation, we mutated the site as follows (see Figure 1A): (1) TCAA (~218 to ~221, relative to +1 of
To determine if an IHF binding site alone is capable of stimulating the activity of the glpFK promoter, an IHF binding site was introduced upstream of PglpFK by changing gcttgacagctttoaatcaagcagtta (see Figure 1A). The newly created site was at the same distance as that of the IHF site in IB:PglpFK relative to the promoter (Figure 1B and Figure S3A). Using a chromosomal lacZ fusion, the activity of PglpFK-containing this upstream IHF binding site was examined in wt and crp cells grown in LB medium. The promoter activity was increased by ~40% in wt cells but was not changed in crp cells (Figure S3B). These results show that (1) IHF can bind to the created binding site; (2) the binding of IHF is capable of partially stimulating PglpFK activity but only in the presence of Crp; and (3) IHF stimulation of IB:PglpFK in the absence of Crp is dependent not only on the IHF binding site present in IB but also on the adjacent sequences (e.g., the A-tracts) surrounding the site.

**Dependency of glpFK Operon Activation on the Permanent Bend in IS5**

A characteristic of a permanent DNA bend is the presence of in-phase A-tracts of 3–6 tandem As. The permanent bend in the IB region of IS5 has been shown to be one of the largest bend angles in the E. coli chromosome [25].

To examine the dependency of glpFK expression in a crp Glp+ genetic background on the A-tracts in the IB region of IS5, three different sets of mutated A-tracts were constructed by changing specific As in these tracts to Cs or Gs (see Figure 1A). The mutated strains were then examined for expression using a PglpFK-lacZ reporter gene fusion. All three sets of A-tract mutations lowered the expression level. However, mutating the five upstream A-tracts (A-tracts 4–8, see Figure 1A) had a minimal effect on expression (Figure 5D). Mutation of A-tract 3 only (A-tract 3) also had a minimal effect. However, mutational alteration of three consecutive A-tracts in the downstream region (A-tracts 1–3) had a dramatic effect on gene expression. From these observations and those reported in the previous section, we conclude that (i) the downstream permanent bend is important for gene activation, (ii) the upstream A-tract region is of minimal importance for gene expression, and (iii) IHF binding and the three downstream A-tracts are roughly of equal importance. We can therefore account for IS5 activation of glpFK operon expression by IHF binding and the downstream A-tracts which apparently activate gene expression in an additive fashion, probably by bending the DNA.

To confirm the importance of both the A-tracts and the IHF binding site in IB to IS5 activation of PglpFK, the IB:PglpFK-lacZ fusion containing mutations in A-tracts 1–3 in IB was transferred into the yhA genetic background. As shown in Figure S4, promoter activity was completely abolished. This experiment therefore shows that the activating effects of IHF and the downstream permanent bend are responsible for promoter activation, and that these two effects are additive.

As shown above, when the upstream 47 bps in the 177 bp fragment (IB) were deleted, 60% of the activation was retained. When the upstream A-tracts were disrupted by point mutations, 85% of the activation was retained (Figure 5D). There may therefore be a mild dependency of activation on these upstream sequences, but we do not know exactly why this difference was observed. Possibly, deletion of this upstream region has an indirect contextual effect on the downstream region that plays a dominant role in glpFK promoter activation.
Figure 5. Dependency of glpFK promoter activation on IHF binding to IB (A–C) as well as the A-tract-promoting permanent bend in IB (D). (A) Growth of crp GlpF+ (○), ihfA (□) and crp GlpF− ihfA (□) cells in liquid glycerol (1%) M9 minimal medium. (B) Effect of a host ihfA null mutation on IB:PglpFK activity. ‘None’, ‘single’, and ‘double’ refer to no mutation, mutation of TCAA (−221 to −218 relative to +1 of PglpFK) to GTCT, and mutation of TCAA to GTCT as well as TT (−213 to −212) to GC in the IHF binding site located in IB, respectively (see Figure 1B). (D) lacZ expression measured by β-galactosidase assay for crp cells carrying an IB:PglpFK-lacZ fusion with various A-tract mutations in IB. The strain bearing altered A-tracts 4–8 includes the mutations shown in A-tracts 4–8. ‘none’, ‘A-tracts 4–8’, ‘A-tract 3’ and ‘A-tracts 1–3’ refer to no mutation, mutations in A-tracts 4–8, mutation in A-tract 3 and mutations in A-tracts 1 to 3, respectively (see Figure 1B). doi:10.1371/journal.pgen.1000689.g005

Effects of ISS and IB on the Activity of the lacZYA Promoter (Plac)

To determine if IB can activate a different Crp-dependent promoter in a crp genetic background, IB was placed at positions −126.5 (the same relative position as in IB:PglpFK) and −178.5 (relative to +1 of Plac) upstream of Plac, yielding IB:Plac and IB:Plac′, respectively. These two sites differ by 22 bps, are therefore should be in phase. The difference in activation of Plac would presumably reflect the distance from the activated promoter.

Using chromosomal lacZ fusions, the activities of these constructs were examined in both crp and crp lacI cells grown in M9 medium + 0.66% casamino acids +1% glucose (Figure 6A). In the absence of IPTG, no appreciable activity was observed for Plac with or without the upstream IB sequence in crp cells (columns 1–3 of Figure 6A). However, in the presence of IPTG, IB increased the Plac activity 8 fold when it was located at −126.5, and 2 fold when it was located at −178.5 compared to Plac alone (columns 7–9 of Figure 6A). These two positions are in phase assuming 10.5 bp per turn in DNA. We also measured the activities of IB:Plac and IB:Plac′ in crp lacI double mutant cells. As expected, regardless of the presence of IPTG, they behaved similarly as in crp cells with IPTG (i.e., 8 and 2 fold increased activation for IB:Plac and IB:Plac′, respectively). The entire ISS element showed the same ability as the IB element to activate Plac (data not shown). These results indicate that (i) IS5 or IB can at least partially replace the function of Crp in activating Plac; (ii) a proper location of these fragments is important for activation of gene expression in the absence of Crp; and (iii) in contrast to the glpFK system, activation is fully blocked by operator-bound LacI.

We further measured IB effects on Plac activity in wild type (crp+) cells grown in the same M9 minimal medium as above. In the absence of IPTG, the activities of Plac with or without the upstream IB were extremely low (20 Miller units or less) (Columns 1, 3 and 4 of Figure 6B). In the presence of IPTG, IB increased the lac promoter activity ~2.5 fold when it was located at −126.5 compared to the lac promoter alone. No increased Plac activity was observed when IB was located at −178.5 upstream of the promoter. These results indicate that (i) LacI still blocks the activation of Plac by both Crp and IB, and (ii) when present at an appropriate position, IB is still capable of enhancing the activity of Plac in the presence of Crp. In other words, IB and Crp activate Plac in an additive fashion. This is different from IB activation of PglpFK, in which case, the presence of Crp did not further elevate the activity of IB:PglpFK (see Figure 4B).

Discussion

We have demonstrated that a transposon, IS5, is capable of activation of the glpFK operon, rendering E. coli cells capable of utilizing glycerol in the absence of Crp. Transposable elements have been found to activate transcription of adjacent genes by introducing complete or partial promoters located within the
Effect on P\textsubscript{lac} grown in M9 minimal medium of transposon-mediated gene activation was of the downstream promoter. In this case, a part of IS5, which is proximal to the adjacent gene and activation by IS5 is distinct from these reported mechanisms. In normally blocks transcription [3,4]. The mechanism of glpFK element or by disrupting or displacing a negative element that plays a role in gene activation although it bends the DNA [33]. In the present study, we showed that the IHF binding site in IS5 plays an important role since an \textit{ihfA} null mutation or alteration of the IHF binding site partially abolished activation to the same degree. No function for the IB region had been recognized prior to our studies. It may have evolved specifically for the purpose of gene activation, for another unrecognized purpose, or for both. We further showed that DNA phasing is important since the insertion of 5 bps (but not 10 bps) between the activating element and IHF abolished activation. Preliminary evidence suggested that the C-terminal domain of the \(\alpha\)-subunit of RNA polymerase is not required for activation (unpublished results), but a DNA looping mechanism is nevertheless proposed.

The activating IB region contains as many as 10 A-tracts, which in general can induce permanent bends in DNA [25,34] and can increase promoter activity [35,36]. In one case, transcriptional activation involves binding of the DNA to the C-terminal domain of the RNA polymerase \(\alpha\)-subunit [35]. In view of our results, it seems likely that the permanent bend in the downstream region of IS5 together with bound IHF activates glpFK promoter activity at least in part by bending the upstream DNA. A direct interaction of IHF with the transcriptional initiation complex is possible [37]. Although this is the first study to show that unique sequences inside a transposon are necessary and sufficient to activate a downstream silent promoter, similar mechanisms of gene activation may occur for other operons (see Figure 6 and unpublished results). These include the \textit{lacZYA}, \textit{fucAO} and \textit{fhlDC} operons [13,14 and observations reported here]. On the other hand, the isolation and analysis of mutations that allow \textit{E. coli} \textit{crp} mutants to grow on several other \textit{crp}-dependent carbon sources [38] indicate that IS-mediated gene activation is not the only mechanism available to \textit{E. coli}. We propose that IS5 insertion under the control of a host regulatory protein represents just one of many mechanisms of operon adaptive activation that will prove to occur under stressful conditions.

**Materials and Methods**

**Bacterial Strains and Growth Conditions**

Strains and oligonucleotides used in this study are described in Table S1 and Table S2, respectively. The \textit{crp}, \textit{glpR}, \textit{crp} \textit{Glp}\textsuperscript{+} and \textit{crp} \textit{glpR} \textit{Glp}\textsuperscript{−} mutants, derived from \textit{E. coli} K-12 strain BW25113, were constructed previously [21]. The \textit{crp lacI} double mutant was made.
by transferring the lacI insertion mutation from MG1655 [39] to a BW25115 crp deletion background using P1 transduction. The hns and cI52 isogenic deletion mutant was derived from the same wt strain BW25113 using the method of Datsenko and Wanner [40]. Double and triple mutants, crp hns, crp GlpK hns, crp cI52 and crp GlpK cI52, were made by P1 transduction. The detailed procedures for generating single or double mutants are described in the reference [17]. Strain BW25113 is deleted for the lacZ gene and the araBAD operon encoding proteins required for L-arabinose metabolism [40]. All mutants were verified by PCR. The strains were cultured in LB or minimal M9 media with various carbon sources at 37 °C. When appropriate, kanamycin [25 μg/ml], or ampicillin (Ap; 100 μg/ml) was added to the media.

Glycerol Kinase Activity Assay

The activity of glycerol kinase encoded by glpK was determined using [1,3,5-3H]glycerol as substrate. Cells were cultured in LB with or without 1% glycerol, and cellular extracts were prepared using a French press and the subsequent ultracentrifugation. Glycerol phosphorylation by GlpK in the extracts was quantitated as described previously [41].

Site-Directed Mutagenesis

Using the quick-change site-directed mutagensis kit (Stratagene), the following modifications were made in IS5, its IB region or the glpFK promoter region: (i) removal of the EcoRI site in IS5 by changing A (135 bp from the 5’ end) to G; (ii) insertion of a 3-bp (tacct) or a 10-bp (tacctacct) oligonucleotide at -117.5 (relative to +1 of PglpFK) in the IB:PglpFK junctional region (see Figure 1B); (iii) mutation of the IHF binding site in IB by changing TCAA to TT at +100 of PglpFK [44]. The entire IS5 or the deleted IS5 site of the plasmid so that IS5 or the deleted IS5 site was located at -126.5 or -178.5, relative to +1 of PlacZ[2]. The fusions of Plac with or without IS5 to lacZ were integrated to the chromosome of crp cells. Primers used for RT-PCR and lacZ fusion construction are listed in Table S2.

β-Galactosidase assays, conducted after growth in either LB or M9 minimal media ± 1% glycerol, 1% glucose or 0.66% casamino acids (CAA), were as described by Miller [45]. CAA was added to improve growth of crp cells in minimal media.

DNA Ligase Mediated PCR

RNA ligase mediated PCR was used to determine the transcriptional start site of the glpFK operon as described by Bensig et al. [46]. Cells were grown in LB with or without 1% glycerol. The total RNA was prepared using a BioMiga RNA purification kit and treated with or without tobacco acid pyrophosphatase (TAP). The RNA oligonucleotide adaptor (Table S2) was ligated to all RNAs, and the 5’-end of the glpFK cDNA was synthesized using an oligonucleotide [PglpFK-extn-R] complementary to nucleotides 118-139 (Table S2). Downstream of the ATG start codon, cDNA was amplified using primers PglpFK-extn-F and PglpFK-extn-R, and the PCR product was sequenced using the same primer (PglpFK-extn-R) as for cDNA synthesis. The transcriptional start site, i.e., the junction between the cDNA and the RNA oligonucleotide, was determined by sequencing.

Supporting Information

Figure S1 Real-time PCR analysis of effects of the hns mutation on glpFK expression in wt, crp and crp GlpK” backgrounds. Cells were grown in LB liquid medium. Found at: doi:10.1371/journal.pgen.1000689.s001 (0.10 MB TIF)

Figure S2 Effects of IS5 and various regions within IS5 on expression of the downstream glpFK promoter in crp cells lacking the IS5, ‘Ples’, ‘I’, ‘B’, ‘none’, and ‘178’, refer to transcriptional lacZ fused for IS5: PglpFK, promoter-less IS5: PglpFK, IB: PglpFK, native PglpFK, and 178bp PglpFK, respectively (see Figure 1A). E. coli strain B has been reported to lack IS5 in its genome [24]. The crp mutation was transferred to strain B by P1 transduction. The promoter lacZ fused described above were individually transferred to strain B crp cells from BW25113 by P1 transduction. For β-galactosidase assays, E. coli strain B crp cells containing these lacZ fusion constructs were grown in LB with shaking. Found at: doi:10.1371/journal.pgen.1000689.s002 (0.12 MB TIF)

Figure S3 Effect of addition of an IHH binding site upstream of PglpFK on promoter activity. (A) The glpFK promoter region showing that an IHH binding site is added upstream of PglpFK by changing gtctgtagcagatt (−222 to −210) to ataacaagcagattga (−222 to −210) to ataacaagcagattga. The newly added IHH binding site is located at the same relative distance as IS5. (B) Effect of the added IHH site on PglpFK activity in wt and crp cells grown in LB medium. PglpFK and PglpFK-IHH refer to the transcriptional lacZ fusions for native PglpFK and the same promoter with an upstream IHH binding site, respectively.
References

1. Friedberg EC, Walker GC, Siede W (1995) DNA Repair and Mutagenesis. Washington, DC: ASM Press.
2. Hengge-Aronis R (2000) The general stress response in Escherichia coli. In Storz G, Hengge-Aronis R, eds. Bacterial Stress Response. Washington, DC: ASM Press. pp: 161–178.
3. Chandler M, Mahillon J (2002) Insertion sequence revisited. In Craig NL, Craige R, Gellett M, Lambowitz AM, eds. Mobile DNA II. Washington, DC: ASM Press. pp: 305–366.
4. Mahillon J, Chandler M (1998) Insertion sequences. Microbiol Mol Biol Rev 62: 725–774.
5. Galas DJ, Chandler M (1989) Bacterial insertion sequences. In Berg D, Howe M, eds. Mobile DNA. Washington, DC: American Society for Microbiology. pp: 199–162.
6. Charlier D, Piette J, Glaudonf N (1982) IS can function as a mobile promoter in E. coli. Nucleic Acids Res 10: 5935–5948.
7. Giampi MS, Schmid MB, Roth JR (1982) Transposon Tn10 provides a promoter for transcription of adjacent sequences. Proc Natl Acad Sci USA 79: 5016–5020.
8. Reynolds AE, Felton J, Wright A (1981) Insertion of DNA activates the cryptic bgl operon in E. coli K12. Nature 293: 625–629.
9. Schnetz K, Rak B (1992) IS5: a mobile enhancer of transcription in E. coli. EMBO J 11: 3157–3160.
10. Dole S, Nagarajavel V, Schnetz K (2004) The histone-like nucleoid structuring protein H-NS represses the bgl operon downstream of the promoter. Mol Microbiol 52: 589–600.
11. Makr R, Mahadevan S (1997) Characterization of the negative elements involved in silencing the bgl operon of Escherichia coli possible roles for DNA gyrase, H-NS, and CRP-cAMP in regulation. Mol Microbiol 24: 617–627.
12. Chen YM, Lu Z, Lin ECC (1988) Constitutive activation of the lacO operon and silencing of the divergently transcribed lux operon by an IS5 element in Escherichia coli mutants selected for growth on L-1,2-propanediol. J Bacteriol 171: 6097–6103.
13. Barker CK, Proh BM, Matsunara P (2004) Increased motility of Escherichia coli by insertion element sequence integration into regulatory region of the bgl operon. J Bacteriol 186: 7529–7537.
14. Petersen C, Moller LB, Valentin-Hansen P (2002) The cryptic adenine diminase gene of Escherichia coli. J Biol Chem 277: 31573–31580.
15. Lin ECC (1976) Glycerol dissimulation and its regulation in bacteria. Annu Rev Microbiol 30: 535–578.
16. Zhao N, Oh W, Trybul D, Thraisher KS, Kingsbury TJ, et al. (1994) Characterization of the interaction of the glpF repressor of Escherichia coli with single and tandem glp operator variants. J Bacteriol 176: 2393–2397.
17. Freedberg WB, Lin ECC (1973) Three kinds of controls affecting the expression of the glp regulon in Escherichia coli. J Bacteriol 115: 816–823.
18. Weissenborn DL, Witterkind N, Larson TJ (1992) Structure and regulation of the glpF operon encoding glycerol diffusion facilitator and glycerol kinase of Escherichia coli K12. J Biol Chem 267: 6122–6131.
19. Zeng G, Ye S, Larson TJ (1996) Repressor for the sgl-glucose 3-phosphate regulon of Escherichia coli K-12: primary structure and identification of the DNA-binding domain. J Bacteriol 178: 7089–7089.
20. Zhang Z, Saier MH, Jr (2009) A mechanism of transposon-mediated directed mutation. Mol Microbiol 74: 29–43.
21. Kroger M, Holm G (1982) Structural analysis of insertion sequence IS5. Nature 297: 159–162.
22. Rak L, Lusky M, Hable M (1982) Expression of two proteins from overlapping and oppositely oriented genes on transposable DNA insertion element IS5. Nature 297: 124–128.
23. Schneider D, Duperchy E, Depe vrot J, Courange E, Lenski RG, et al. (2002) Genomic comparisons among Escherichia coli strains B, K-12 and O157:H7 using elements as molecular markers. BMC Microbiol 2: 6 pgs.