

Specific contacts of the –35 region of the *galP1* promoter by RNA polymerase inhibit GalR-mediated DNA looping repression

Zsolt Csiszovszki¹, Dale E. A. Lewis¹, Phuoc Le¹, Kim Sneppen² and Szabolcs Semsey^{2,*}

¹Laboratory of Molecular Biology, Center for Cancer Research, NCI, NIH, Bethesda, MD, USA and

²CMOL, Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark

Received April 17, 2012; Revised July 28, 2012; Accepted July 30, 2012

ABSTRACT

The *P1* promoter of the galactose operon in *Escherichia coli* is one of the best studied examples of ‘extended –10’ promoters. Recognition of the *P1* promoter does not require specific contacts between RNA polymerase and its poor –35 element. To investigate whether specific recognition of the –35 element would affect the regulation of *P1* by GalR, we mutagenized the –35 element of *P1*, isolated variants of the –35 element and studied the regulation of the mutant promoters by *in vitro* transcription assays and by mathematical modeling. The results show that the GalR-mediated DNA loop is less efficient in repressing *P1* transcription when RNA polymerase binds to the –10 and –35 elements concomitantly. Our results suggest that promoters that lack specific –35 element recognition allow decoupling of local chromosome structure from transcription initiation.

INTRODUCTION

Most *Escherichia coli* promoters transcribed by the major RNA polymerase (RNAP) ($\alpha_2\beta\beta'\sigma^{70}$) contain two conserved sequence elements located at positions 10 and 35 bp upstream of the transcription start point (*tsp*) (1,2). These promoter elements are recognized by the σ^{70} subunit of RNAP. The conserved regions 2.4 and 4.2 of σ^{70} recognize the –10 and –35 promoter elements, respectively (3,4). However, a minor class of promoters lacks a recognizable –35 element. These promoters are contacted with a different pattern, where region 2.4 binds to the –10 hexamer and region 2.5 makes sequence specific contacts with a short sequence located 1 bp upstream of the –10 element (5). This short sequence, the

‘extended –10’ motif, is characterized by the 5'-TRTG-3' consensus sequence (6–9). Recognition of ‘extended –10’ promoters does not require specific contacts between region 4.2 and the –35 element (10,11).

The galactose operon (*gal/ETKM*) of *E. coli* is transcribed from two overlapping promoters, *P1* and *P2* (Figure 1). *P1* is regulated by the galactose repressor (GalR) by two different mechanisms, contact inhibition and DNA looping. GalR dimers can bind to two operator elements, *O_E* and *O_I*, separated by a 113 bp DNA segment comprising *P1* and *P2* (Figure 1) (12–14). Contact inhibition occurs when *O_E*-bound GalR represses *P1* by contacting the C-terminal domain of the α subunit of RNAP (α -CTD), and inhibiting open complex formation (15–17). DNA looping repression of both *P1* and *P2* occurs simultaneously when the *O_E*-bound GalR dimer interacts with the *O_I*-bound GalR dimer to form a tetramer and the intervening DNA loops out, forming a repression complex (known as repressosome) (18–20). Assembly of the Gal repressosome, a higher order nucleoprotein structure containing an antiparallel DNA loop (21,22), requires (i) binding of two dimeric GalR to *O_E* and *O_I*, (ii) negatively supercoiled DNA, (iii) optimal angular orientation of *O_E* and *O_I* and (iv) specific binding of the architectural histone-like protein (HU) to a HU binding site (*hbs*) in the interoperator region (23,24). DNA supercoiling and the binding of HU at a critical position facilitate DNA loop formation over a short distance by decreasing the persistence length of DNA (25).

Previous studies suggested that simultaneous binding of RNAP to the –10 and –35 elements introduces a specific bend in the DNA, as indicated by DNaseI hypersensitivity around position –25 (4,26). In *galP1*, this hypersensitivity was less pronounced in the absence of a specific –35 recognition site, suggesting that the DNA trajectory in RNAP-‘extended –10’ promoter complexes is different from that found in complexes where RNAP specifically recognizes both the –10 and –35 elements (11,26).

*To whom correspondence should be addressed. Tel: +45 249 42613; Fax: +45 353 25425; Email: semsey@nbi.dk

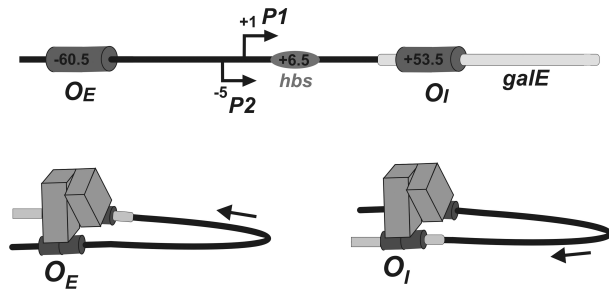


Figure 1. (Top) Schematic map of the *gal* regulatory region containing the *gal* operator sites (O_E and O_I), the *gal* promoters, $P1$ (+1) and $P2$ (-5), and the HU binding site (*hbs*, +6.5). The transcription start point of $P1$ is used as a reference in the numbering system (+1). (Bottom) V-shaped, stacked interaction of operator-bound GalR dimers can lead to two different antiparallel DNA loop trajectories, A1 (left) and A2 (right). Arrows indicate the direction of transcription initiated at the *gal* promoters.

In this work, we asked whether the DNA bend introduced by simultaneous binding of RNAP to the -10 and -35 elements affects the stability of the *gal* DNA loop. We introduced mutations at the -35 hexamer of the $P1$ promoter and monitored transcription regulation of the wild-type (WT) and mutant promoters in the absence and presence of GalR and HU by using *in vitro* transcription assays. Our results are consistent with a model where RNAP binding to the -10 and -35 elements have an inhibitory effect on looping mediated repression of $P1$.

MATERIALS AND METHODS

Strain and plasmid constructions

Plasmid manipulations followed protocols described in Sambrook and Russel (27). Transformations were performed with XL-1 Blue competent cells (Stratagene). Sequencing (ABI Prism) kits were purchased from Applied Biosystems, restriction endonucleases from Fermentas, DNA oligonucleotide primers from Invitrogen and DNA purification kits from Qiagen. DNA sequencing reactions were analysed in a Perkin-Elmer/Applied Biosystems (model 373A) automated sequencer.

Mutations in the *gal* promoter regions were created by polymerase chain reaction (PCR) using the Platinum[®] High Fidelity PCR SuperMix (Invitrogen) and inserted between the EcoRI and PstI sites in plasmid pSA850 (28). The mutated sequences of the *gal* regulatory region in the resulting plasmids were verified.

Protein purification

Expression and purification of the hexahistidine-tagged GalR followed the protocol described by Semsey *et al.* (20). HU protein was purified according to the method described by Aki *et al.* (18). RNAP was purchased from USB. RNAP concentration was specified by the manufacturer. GalR and HU concentrations were measured using the Micro BCA Protein Assay Kit (Pierce). The quality of protein preparations was tested in *in vitro* transcription

reactions using the reference plasmid pSA850. Similar results were obtained to previously published results (22,23).

In vitro transcription assays

In vitro transcription reactions were performed as described in (20). The reaction mixture (50 μ l) contained 20 mM Tris acetate, pH 7.8, 10 mM magnesium acetate, 200 mM potassium glutamate and 2 nM supercoiled DNA template. GalR concentrations vary from 5 to 40 nM as indicated, and HU was used at 80 nM. RNAP (20 nM) was added before incubating the reactions at 37°C for 5 min. Transcription was initiated by the addition of 1.0 mM ATP, 0.1 mM GTP, 0.1 mM CTP, 0.01 mM UTP and 5 μ Ci of [α -³²P]UTP (3000 Ci/mmol). Reactions were terminated after 10 min at 37°C by the addition of an equal volume of transcription loading buffer (0.025% bromophenol blue, 0.025% xylene cyanol, 0.01 M ethylenediaminetetraacetic acid and 90% deionized formamide). After heating at 90°C for 3 min, the samples were loaded onto an 8% polyacrylamide-urea DNA sequencing gel. RNA bands were quantified using the ImageQuant[™] PhosphorImager (Molecular Dynamics, CA). We followed the standard procedure that uses the RNA1 transcript as an internal control between lanes, to decrease the level of potential experimental error (19). The RNA1 transcript is not affected by GalR binding. Band intensities were background corrected as described previously (19). This procedure has <10% error (29). As levels of the studied transcripts relative to the level of the RNA1 transcript may slightly vary depending on the quality of the plasmid DNA preparation, promoter activities in the presence of GalR were expressed relative to the promoter activity in the absence of GalR.

Construction of a mathematical model of the *in vitro* system

In vitro transcription reactions contained a fixed amount of GalR (0–40 nM), RNAP (20 nM) and DNA (2 nM). GalR has two specific binding sites (O_E and O_I) on the DNA molecules used, therefore, the amount of GalR added to the reaction (R_{total}) is the sum of the operator bound molecules (R_B) and free GalR molecules (R_F , not bound to DNA or bound non-specifically). As the DNA template is present at 2 nM concentration, and there are two operators per DNA molecule, $R_B \leq 4$ nM. Based on the possible patterns of operator occupancy, there are five different binding states for GalR: (i) not bound to any operators, (ii) bound to O_E but not to O_I , (iii) bound to O_I but not to O_E , (iv) bound to both operators without DNA loop formation and (v) GalR dimers bound to O_E and O_I form a tetramer, resulting in a DNA loop. At any given GalR concentration, the DNA molecules present in the reaction are distributed between these five states. The ratios of these states can be computed based on the operator binding affinities and on the probability of DNA looping (Table 1).

In each state, RNAP can specifically bind to only the -10 region of the $P1$ promoter, to both the -10 and -35

Table 1. The 15 binding states of the *P1* promoter and regulatory region considered in the model

State	O_E	-35	-10	O_I	Loop	Activity	Relative weight
i	1	-	-	-	-	0	1
	2	-	-	+	-	-	$\frac{P}{K_{10}}$
	3	-	+	+	-	-	$\frac{P}{K_{10} \times F_{35}}$
ii	4	+	-	-	-	0	$\frac{R_F}{K_{RE}}$
	5	+	-	+	-	0.5	$\frac{R_F}{K_{RE}} \times \frac{P}{K_{10}}$
	6	+	+	+	-	0.67	$\frac{R_F}{K_{RE}} \times \frac{P}{K_{10} \times F_{35}}$
iii	7	-	-	-	+	0	$\frac{R_F}{K_{RI}}$
	8	-	-	+	+	-	$\frac{R_F}{K_{RI}} \times \frac{P}{K_{10}}$
	9	-	+	+	+	-	$\frac{R_F}{K_{RI}} \times \frac{P}{K_{10} \times F_{35}}$
iv	10	+	-	-	+	0	$\frac{R_F}{K_{RE}} \times \frac{R_F}{K_{RI}}$
	11	+	-	+	+	0.5	$\frac{R_F}{K_{RE}} \times \frac{R_F}{K_{RI}} \times \frac{P}{K_{10}}$
	12	+	+	+	+	0.67	$\frac{R_F}{K_{RE}} \times \frac{R_F}{K_{RI}} \times \frac{P}{K_{10} \times F_{35}}$
V	13	+	-	-	+	+	$\frac{R_F}{K_{RE}} \times \frac{R_F}{K_{RI}} \times L$
	14	+	-	+	+	+	$\frac{R_F}{K_{RE}} \times \frac{R_F}{K_{RI}} \times L \times \frac{P}{K_{10}}$
	15	+	+	+	+	+	$\frac{R_F}{K_{RE}} \times \frac{R_F}{K_{RI}} \times \frac{P}{K_{10} \times F_{35}} \times L^*$

Parameters used are as follows: R_F , concentration of free GalR; K_{RE} , binding affinity of GalR to O_E ; K_{RI} , binding affinity of GalR to O_I ; L , looping factor; L^* , looping factor when RNAP is bound to both the -10 and -35 elements; P , RNAP concentration; K_{10} , binding affinity of RNAP to the -10 region; F_{35} , factor for increase in RNAP affinity to the promoter resulted from specific contacts at the -35 region. The five GalR-binding categories (i-v) are indicated on the left.

regions, or be not bound to the *P1* promoter, resulting in 15 binding states (Table 1). The relative statistical weights for the different binding states depend on the probabilities of GalR and RNAP binding to the operators and promoter elements, respectively. The weight of the state where DNA is not bound by any of the proteins is used as a reference (= 1). The weights also depend on the probability of DNA loop formation mediated by tetramer formation of GalR dimers bound to O_E and O_I . Therefore, the relative statistical weights are functions of free protein concentrations and protein binding affinities for the different sites on the DNA (Table 1). Under the quasi-steady-state assumption that GalR and RNAP binding to DNA are sufficiently rapid relative to the transcription rate, the probability of every 1 of the 15 binding states can be calculated (30). The probability of a state i (P_i) is equal to the weight of state i divided by the sum of the weights of all possible states. Promoter activity levels can be given by the following equation:

$$\text{Promoter activity} = \sum_{i=1}^{15} P_i \times \text{activity}_i.$$

Relative promoter activities in each binding state (Table 1) are derived from experimental results. In the absence of RNAP binding, the promoter activity is 0, in the presence of DNA looping a weak promoter activity is observed

(0.05), and contact inhibition by O_E -bound GalR results in 50% repression of WT *P1* where RNAP does not make specific contacts with the -35 region (22) (relative activity = 0.5), and 33% repression in the presence of the consensus -35 box (relative activity = 0.67). The latter value was determined by *in vitro* transcription assays performed on a pSA850 derivative plasmid containing the consensus *P1* -35 element (P1C35), in the presence of 80 nM GalR T322R, a GalR mutant that binds the *gal* operator sites similar to WT GalR but unable to form a DNA loop (19). Promoter activity of 1 was used in the RNAP-bound states where the O_E operator is not bound by GalR. There are several other promoters on the plasmid template (28); hence, it is difficult to estimate the concentration of free RNAP. However, this problem can be overcome by using the term P/K_{10} , where RNAP concentration is expressed relative to its binding affinity to WT *P1*. The upper limit for the value of P/K_{10} is 0.02 because *P1* activity is increased 50-fold with the consensus -35 box mutation (P1C35). We used this maximal value in the model, assuming that the P1C35 promoter is transcribed in all the template molecules in the *in vitro* transcription reactions. Concentration of R_F was computed based on R_{total} and R_B , which were derived from the statistical weights listed in Table 1. The mathematical model was written in FORTRAN.

RESULTS AND DISCUSSION

Effect of mutations in the -35 region on *P1* activity

To study the role of the -35 element on *P1* regulation, a set of *P1* derivatives was created by PCR mutagenesis where the -35 *P1* element was randomized (YWNNNH, TNGNCA and TTGNNN). The mutant promoters were inserted into plasmid pSA850 (28), a reference plasmid used for *in vitro* studies on *P1* and *P2* regulation. From the pool of plasmids containing mutant promoters, we selected 18 plasmids randomly and determined the sequence of the mutant promoter regions. The promoter set did not contain a promoter with a consensus -35 box; therefore, we created this *P1* derivative (TTGACA) separately. Activity and regulation of the WT and 13 mutant promoters were studied using *in vitro* transcription assays (Figure 2). The same batches of RNAP, GalR and HU were used for all the assays to make the results comparable. Mutations in the -35 box substantially increased the intrinsic activity of *P1* and decreased that of *P2*. The promoter containing the consensus -35 sequence showed the highest activity; it was ~50-fold stronger than WT *P1* (Figure 2, lane 22 versus lane 1; also refer Table 2).

Typically the -35 elements of strong promoters are more similar to the consensus sequence than the -35 elements of weaker promoters. However, less deviation from the consensus does not necessarily mean higher activity. For instance, a promoter with only two mismatches compared with the consensus sequence (TcGgCA) does not show substantially stronger activity than the WT promoter with six mismatches (cacttt) (Figure 2, lane 28 versus lane 1).

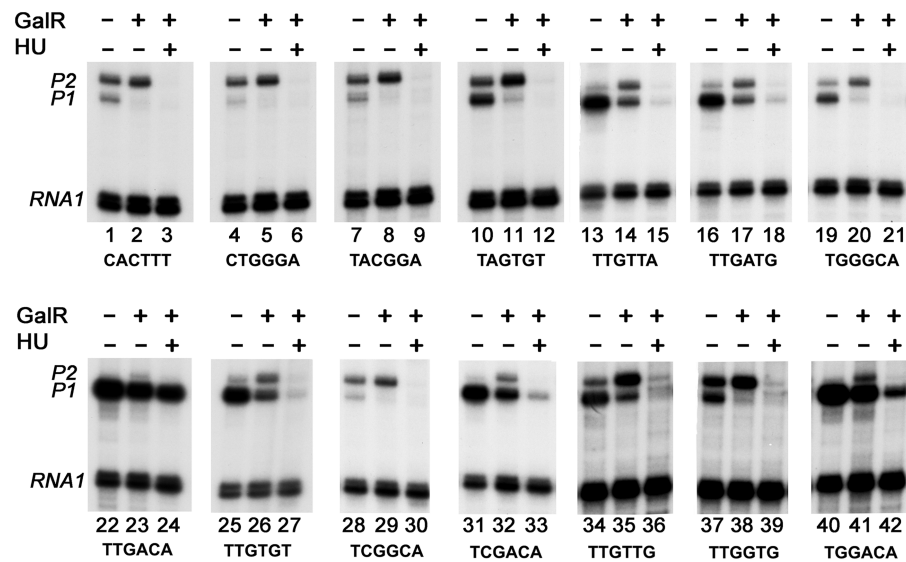


Figure 2. Results of *in vitro* transcription assays performed on WT (lanes 1–3) and mutant –35 derivatives (lanes 4–42) of *P1*. GalR and HU concentrations were 40 and 80 nM, respectively. The RNA1 transcript, which is not affected by GalR and HU binding, serves as an internal control between lanes.

Table 2. Relative activities of WT and mutant promoters in the presence of GalR and 80 nM HU

Promoter	Strength (P/RNA1)	Relative activity GalR				
		0 nM	5 nM	10 nM	20 nM	40 nM
<i>P1</i>	0.2	1	0.27	0.09	0.08	0.06
<i>P1C35</i>	10.0	1	0.71	0.36	0.22	0.15
<i>P1S35G</i>	4.1	1	0.72	0.38	0.16	0.06
<i>P1S35C</i>	2.6	1	0.74	0.42	0.20	0.07
<i>P1C10</i>	2.5	1	0.52	0.18	0.09	0.06
<i>P2</i>	0.3	1	0.54	0.17	0.09	0.07
<i>P2C35</i>	1.9	1	0.89	0.50	0.18	0.05

Promoter strength is expressed as the ratio of intensities of bands corresponding to the unregulated promoter and the RNA1 transcript. Promoter activities were determined based on the results of *in vitro* transcription assays presented in Figure 3, as described in the *Materials and Methods* section. Concentrations are calculated for GalR dimers.

Effect of mutations in the –35 region on the regulation of *P1*

Mutations in the –35 box of *P1* did not result in a qualitative change in the regulation of promoters. Similar to the WT promoters, mutant *P1* promoters were repressed in the presence of GalR, while *P2* promoters were activated. In the presence of both GalR and HU, both the *P1* and the *P2* promoters were repressed as a result of DNA loop formation (Figure 2). The results suggested that the efficiency of *P1* repression depends on the –35 sequence. HU-independent GalR-mediated transcription inhibition, which is a combined result of contact inhibition and non-assisted DNA looping (22,31), was least efficient in the case of the promoter containing the consensus (TTGACA) –35 sequence. This promoter retained 50% activity

in the presence of GalR, compared with 20% activity observed in the case of the WT promoter (Figure 2, lane 23 versus lane 22 and lane 2 versus lane 1). Promoters with TgGgCA and TTGttg –35 hexamers showed similar level of GalR-mediated repression as the WT promoter (Figure 2, lanes 19–20, 34–35). Results obtained in the concomitant presence of GalR and HU suggested that Gal repressosome-mediated repression can also be affected by certain mutations in the –35 boxes. Repression of most of the mutant promoters by repressosome-mediated DNA looping was comparable with the WT promoter. However, in some cases we observed more incomplete repression, which was most prominent in the case of the promoter containing the consensus –35 sequence (TTGACA) (Figure 2, lane 24 versus lane 22). Two other promoters, which had only one mismatch in the –35 hexamer compared with the consensus sequence (TgGACA and TcGACA), also showed incomplete repression (lane 42 versus lane 40, and lane 33 versus lane 31).

Looping repression inhibits a step prior the formation of the first phosphodiester bond (32). Torsional and lateral inflexibility of the short looped DNA may impede RNAP binding or isomerization from close to open complex (33). The lifetime of the *gal* DNA loop is relatively short and decreases with increased tension in the DNA. The time spent in the unlooped state is also tension dependent and increases with tension (34). To obtain efficient repression, the time interval when the DNA is in the unlooped state must be shorter than the time RNAP needs to find the promoter (35) and initiate transcription. Therefore, the DNA loop should form faster if DNA looping inhibits only open complex formation and not RNAP binding to the DNA.

The length of the time interval spent in the unlooped state is not known in our *in vitro* transcription assays but

perhaps RNAP is allowed to bind to the *gal* promoters in the unlooped intervals. This assumption is based on the observation that the A2 loop trajectory (Figure 1, bottom right), which has similar thermodynamic properties as the A1 loop (formed on WT DNA), fails to repress transcription, likely because the RNAP binding site overlaps with the HU binding site in the A2 loop (22,34). In this case, RNAP binding inhibits reclosing of the loop with the A2 trajectory. As the A1 and A2 trajectories have similar lifetimes of looped and unlooped states (the A2 loop forms and breaks down slightly more frequently) (34), we can assume that RNAP can bind to the *gal* promoters in the unlooped intervals of the WT setup used in the *in vitro* transcription assays.

There are at least three different possible explanations for the incomplete looping repression: (i) mutations alter the deformability of DNA; (ii) stronger promoters need shorter times for initiation of transcription and (iii) the bending introduced by concomitant binding of RNAP to the -35 and -10 elements interferes with DNA loop formation. In regard to (i), it is unlikely that the altered -35 sequence would substantially change DNA deformability. The predicted twisting difference (36) caused by the TTG ACA substitution is very minor (~0.2°), and the substitution is located in a region which is not deformed in the DNA loop (22). Therefore, in the next part of the manuscript, we focus on the other two scenarios. To study the second possibility (ii), we tested whether increasing the promoter strength by enhancing the -10 element has similar effect as observed in the case of -35 element mutations. The third possibility (iii) was tested by a computational approach, asking whether the data obtained on the WT and on the -35 mutant (TTGACA) *P1* promoters can be fitted to the model using the same GalR binding and looping parameters.

The effect of promoter strength on looping repression

As described earlier, looping repression becomes inefficient if the unlooped time is shorter than the time required for transcription initiation at the promoter located inside the DNA loop. To test whether incomplete repression obtained for some of the mutant promoters (Figure 2) can be explained by the increased promoter strength (faster initiation), we created a mutant, which has a consensus -10 hexamer (Figure 3, PIC10). Repression of the PIC10 mutant was compared with repression of the -35 hexamer mutants, which showed incomplete looping repression, TTGACA (PIC35), TgGACA (P1S35G) and TcGACA (P1S35C) at different GalR concentrations (Figure 3 and Table 2). We did not find a strong correlation between promoter strength and efficiency of repression. The PIC10 promoter has similar intrinsic activity to the P1S35C promoter; however, repression of P1S35C is less efficient at intermediate GalR concentrations. Also, while the intrinsic activity of P1S35G is somewhat stronger than that of P1S35C, repression of these promoters showed similar dependence on GalR concentration. A mutant *P2* promoter having a consensus -35 box was also weaker than PIC10 but was less efficiently repressed by GalR than PIC10 or the WT *P2* promoter (Table 2). Therefore, we concluded that faster transcription initiation by itself is not sufficient to explain the behavior of the -35 mutant promoters.

An important difference between the PIC10 and the -35 mutant *P1* promoters is that RNAP binding to the -35 mutant promoters results in a specific bend in the DNA. Stability of short DNA loops is often increased by architectural proteins which bind to the looped region and introduce a bend with a characteristic stereo specificity (25). However, introduction of a bend at an improper position can inhibit DNA loop formation (37). The bend

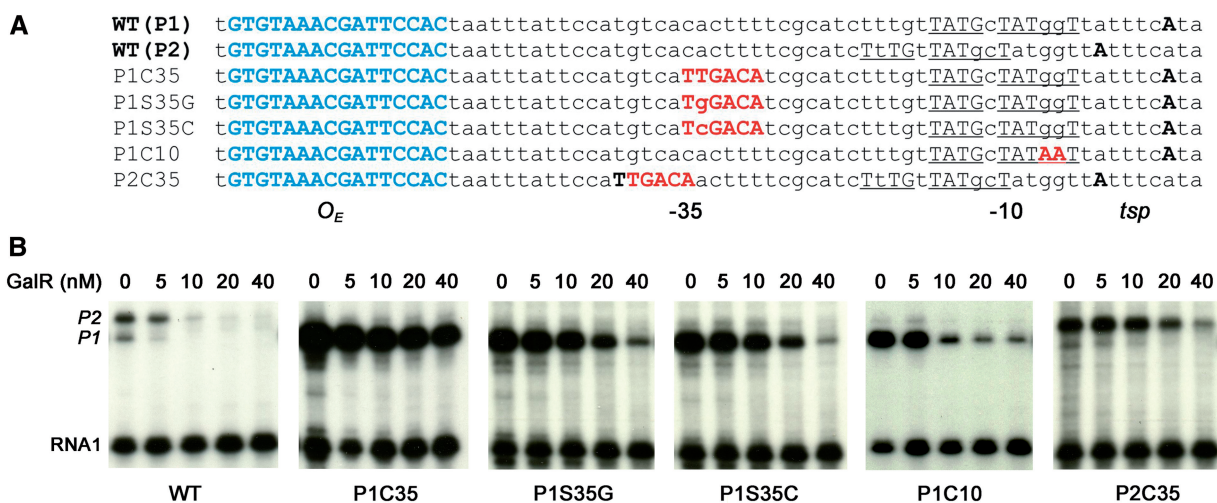


Figure 3. (A) DNA sequences of *P1*, *P2* and their mutant derivatives. The O_E operator sequence (blue) was used to align the sequences. Transcription start points (*tsp*) are typed boldface. Extended -10 and -10 promoter elements are underlined. Red sequences indicate mutations. Positions matching the consensus -10 and -35 sequences are capitalized. (B) Results of *in vitro* transcription assays performed on plasmids carrying the indicated promoter sequences in the presence of 0–40 nM GalR. HU (80 nM) was present in all of the reactions. The same batches of RNAP, GalR and HU were used for all the assays to make the results comparable. These batches were different from the ones used in the experiments shown in Figure 2. Band intensities were quantified and are shown in Table 2.

introduced by RNAP in the case of PIC35 (11) most likely interferes with DNA loop formation because it affects a region which is not deformed in the GalR-mediated DNA loop formed on WT DNA. Inhibition of DNA loop formation allows longer time for RNAP binding and transcription initiation, resulting in failure of DNA loop mediated repression as it was observed in the A2 loop geometry, where RNAP binding to the *P2* promoter inhibits HU binding and stabilization of the DNA loop (22,34).

Mathematical model of the *in vitro* system

We have constructed a mathematical model to simulate the effects of specific RNAP contacts in the -35 region. The model is described in the 'Materials and Methods' section. By fitting this model to the experimental data on WT *P1* and *P2* (Table 2), we obtained estimates on the parameters for the binding strengths of GalR to O_E (K_{RE}) and O_I (K_{RI}), and also for the looping factor (L). The fitting criterion used was that the predicted values must not be $>20\%$ different from the corresponding measured values (Figure 4). This interval is at least twice as big as the expected experimental error (29). F_{35} was set to 10000, assuming negligible contribution of specific RNAP binding to the -35 region. Parameter sets (K_{RE} , K_{RI} and L) were chosen randomly and the sets fulfilling the fitting criterion were recorded. In the parameter sets obtained, K_{RE} was between 1.1 and 7.6 nM. Previous estimates of K_{RE} by DNaseI footprinting (1.3 nM) (38) and by fluorescence anisotropy (4.2 nM) (39) fall in this interval. The smallest values obtained for K_{RI} and L were 36 nM and 137, respectively. Based on these values, we can conclude that the states where GalR is bound to O_I but not to O_E or bound to both operators without DNA loop formation have little importance because $R_F/K_{RE} \gg R_F/K_{RI}$, and $L/K_{RI} \gg 1/K_{RI}$. In the rest of the GalR binding states, the parameters L and K_{RE} appear only in the looped state as $\frac{R_E^2 \times L}{K_{RE} \times K_{RI}}$. Therefore, in the parameter sets that satisfy the fitting criteria the L/K_{RI} ratio is constrained, falling between 2.3 and 4.1 nM^{-1} .

To fit the model to the data on the PIC35 promoter, we used $F_{35} = 0.0001$ (assuming concomitant -10 and -35 contacts), K_{RE} values from the interval obtained on the WT promoters (1.1–7.6 nM), and $K_{RI} > 36$ nM. We again accepted parameter sets that give activities within 20% from the measured values (Figure 4). In the parameter sets obtained, K_{RI} values were between 36 and 500 nM, L values were between 1 and 136, whereas L/K_{RI} ratios were between 0.021 and 0.271 nM^{-1} . Data could be fitted using F_{35} values that were smaller than 0.001. However, L values were always <136 , and L/K_{RI} ratios were $<0.52 \text{ nM}^{-1}$. Therefore, it was not possible to fit the data obtained on the WT and on the -35 mutant (TTG ACA) *P1* promoters using the same GalR binding and looping parameters. Results of the simulations indicate that DNA loop formation is less probable in the case of the PIC35 promoter, i.e. DNA loop formation is inhibited. Unlike the PIC35 data, results obtained on the PIC10 promoter could be explained by weaker repression

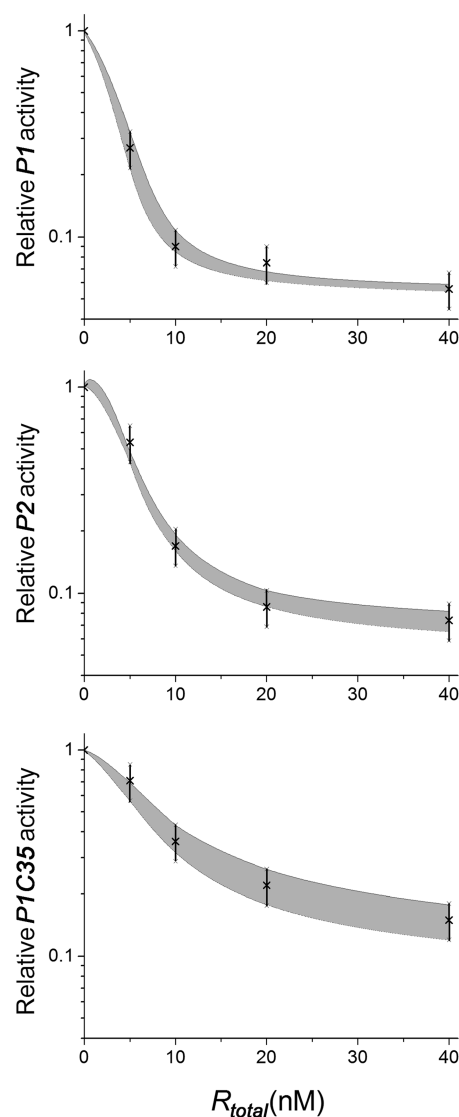


Figure 4. Fitting of the mathematical model to the experimental data obtained on WT *P1* (top), *P2* (middle) and PIC35 (bottom). Crosses indicate experimental measurements (shown also in Table 2) and vertical bars show the $\pm 20\%$ acceptance intervals. Gray areas indicate results of simulations. R_{total} indicates the concentration of GalR dimers (sum of the free and operator bound forms).

by the O_E -bound GalR dimer. These data could be fitted using $L/K_{RI} > 2$ and 33% repression by contact inhibition. Therefore, it is likely that specific RNAP binding to the -35 sequence is required for inhibition of DNA loop formation.

Using the parameters obtained, we computed the distribution of the different binding states as a function of total GalR concentration in the reactions. The left panel of Figure 5 shows how the DNA molecules carrying the WT *gal* regulatory region are distributed between the different GalR-bound states. At low GalR concentrations the unbound state dominates, whereas at higher GalR concentrations, most of the DNA molecules are in the looped state. The concentration of single operator bound states shows non-monotonic behavior because of

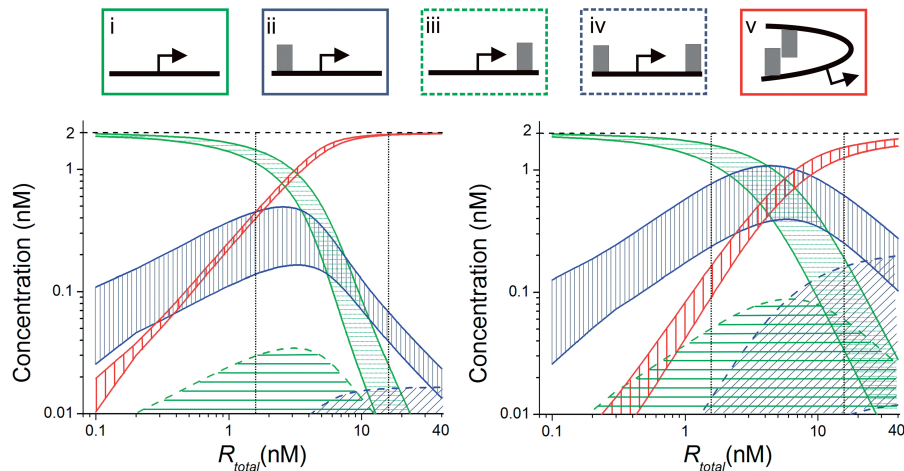


Figure 5. Modeling the distribution of DNA molecules containing the WT *PI* promoter (left) and the PIC35 promoter (right) between the five different binding states of GalR as a function of GalR concentration (R_{total}). Hatched areas between the boundary lines indicate the concentration ranges of DNA molecules present in the different binding states. The five binding states of GalR are the following: (i) not bound to any operators (solid green boundary lines), (ii) bound to O_E but not to O_I (solid blue boundary lines), (iii) bound to O_I but not to O_E (dashed green boundary lines), (iv) bound to both operators without DNA loop formation (dashed blue boundary lines) and (v) bound to both operators and forming a DNA loop (solid red boundary lines). Same colors mean same regulatory outcome: no repression (green), contact inhibition by the O_E -bound GalR dimer (blue), and repression by DNA looping (red). Vertical dotted lines at 1.5 and 15 nM GalR concentrations are shown for guidance. In the schematic drawings of the binding states, (top) gray boxes indicate GalR, and the arrow indicates the direction of transcription initiated at the promoter studied.

the cooperative interactions in DNA loop formation. The state where GalR binds to only O_I is negligible, whereas ~ 5 – 20% of the DNA molecules are in the only O_E -bound state at 3 nM GalR concentration (Figure 5). In the absence of D-galactose, the steady-state level of intracellular free GalR (R_F) is ~ 15 nM (40). At the corresponding GalR concentration ($R_{\text{total}} \sim 19$ nM) $>95\%$ of the molecules are in the looped state.

We performed a similar analysis of states in the case of the PIC35 promoter. The curves had similar trajectories to the corresponding ones obtained with the WT promoter, however, they were quantitatively different (Figure 5, right panel). At $R_F = 15$ nM ($R_{\text{total}} \sim 18.5$ nM) only $\sim 75\%$ of the molecules were in the looped state, and ~ 10 – 20% of molecules were in the state where only O_E is bound. The highest abundance of the latter state (15–50%) was observed at R_{total} of 4–5 nM.

The weaker loop formation in the presence of specific RNAP contacts with the -35 region may have important consequences *in vivo*. The *gal* DNA loop is not able to repress transcription after elongation is initiated. Therefore, a single elongation initiation results in dissociation of the O_I -bound GalR dimer and therefore increases the time spent in the unlooped state. This event can facilitate multiple initiations because the on-rate of intracellular free RNAP (present at few hundred nanomolarity concentration) (41) is much higher than the on-rate for GalR binding to O_I and then forming a loop. This effect can become more pronounced when D-galactose is present at low levels, decreasing the active fraction of free GalR. Besides inhibiting DNA binding, the presence of D-galactose also interferes with GalR tetramerization resulting in an increase of the fraction of time spent in the unlooped state (42). Interaction of GalR dimers is more sensitive to

D-galactose concentration than operator binding by GalR. For instance, at 0.1 mM D-galactose GalR tetramerization is completely inhibited, whereas $\sim 85\%$ of the O_E and $>50\%$ of the O_I operator DNA can be bound by GalR (42). Therefore, burstiness of transcription initiation resulted from the slow on-rate of a transcription regulatory protein (35) would likely modulate the behavior of the feedbacks circuits in the galactose utilization system (43,44) in a D-galactose-dependent manner. There are two mechanisms that act against such burstiness of transcription initiation at the *galPI* promoter: lack of specific RNAP contacts at the -35 region, and inhibition of open complex formation by the O_E -bound GalR dimer (17). The first mechanism reduces the interference of RNAP binding and DNA loop formation, whereas the second mechanism decreases the probability of transcription initiation, and thus the chance of removal of GalR from O_I by the elongating RNAP.

CONCLUSIONS

The extended -10 motif is present in $\sim 20\%$ of all *E. coli* promoters (6), and 26% of these promoters have weak -35 boxes, with not more than two matches to the consensus sequence (8). Promoters belonging to this minor class of promoters lacking the specific -35 element recognition are utilized to transcribe the *gal* operon. RNAP binding to these promoters does not interfere with formation of the short *gal* DNA loop, suggesting that it has little if any impact on DNA flexibility. However, further studies are needed to investigate whether extended -10 promoters were generally evolved to decouple local chromosome structure from transcription initiation.

ACKNOWLEDGEMENTS

The authors thank their colleagues for various inputs in the project, in particular Sankar Adhya, Victor Zhurkin and Namiko Mitarai for valuable suggestions.

FUNDING

Danish Council for Independent Research, Natural Sciences, the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research and the Danish National Research Foundation. Funding for open access charge: Danish National Research Foundation.

Conflict of interest statement. None declared.

REFERENCES

- Hawley, D.K. and McClure, W.R. (1983) Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.*, **11**, 2237–2255.
- McClure, W.R. (1985) Mechanism and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.*, **54**, 171–204.
- Gross, C.A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J. and Young, B. (1998) The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harb. Symp. Quant. Biol.*, **63**, 141–155.
- Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O. and Darst, S.A. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science*, **296**, 1285–1290.
- Barne, K.A., Bown, J.A., Busby, S.J. and Minchin, S.D. (1997) Region 2.5 of the *Escherichia coli* RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J.*, **16**, 4034–4040.
- Burr, T., Mitchell, J., Kolb, A., Minchin, S. and Busby, S. (2000) DNA sequence elements located immediately upstream of the -10 hexamer in *Escherichia coli* promoters: a systematic study. *Nucleic Acids Res.*, **28**, 1864–1870.
- Keilty, S. and Rosenberg, M. (1987) Constitutive function of a positively regulated promoter reveals new sequences essential for activity. *J. Biol. Chem.*, **262**, 6389–6395.
- Mitchell, J.E., Zheng, D., Busby, S.J. and Minchin, S.D. (2003) Identification and analysis of 'extended -10' promoters in *Escherichia coli*. *Nucleic Acids Res.*, **31**, 4689–4695.
- Ponnambalam, S., Webster, C., Bingham, A. and Busby, S. (1986) Transcription initiation at the *Escherichia coli* galactose operon promoters in the absence of the normal -35 region sequences. *J. Biol. Chem.*, **261**, 16043–16048.
- Kumar, A., Malloch, R.A., Fujita, N., Smillie, D.A., Ishihama, A. and Hayward, R.S. (1993) The minus 35-recognition region of *Escherichia coli* sigma 70 is inessential for initiation of transcription at an "extended minus 10" promoter. *J. Mol. Biol.*, **232**, 406–418.
- Minakhin, L. and Severinov, K. (2003) On the role of the *Escherichia coli* RNA polymerase sigma 70 region 4.2 and alpha-subunit C-terminal domains in promoter complex formation on the extended -10 galP1 promoter. *J. Biol. Chem.*, **278**, 29710–29718.
- Adhya, S., Geanakopoulos, M., Lewis, D.E., Roy, S. and Aki, T. (1998) Transcription regulation by repressosome and by RNA polymerase contact. *Cold Spring Harb. Symp. Quant. Biol.*, **63**, 1–9.
- Irani, M.H., Orosz, L. and Adhya, S. (1983) A control element within a structural gene: the gal operon of *Escherichia coli*. *Cell*, **32**, 783–788.
- Majumdar, A. and Adhya, S. (1984) Demonstration of two operator elements in gal: in vitro repressor binding studies. *Proc. Natl Acad. Sci. USA*, **81**, 6100–6104.
- Choy, H.E. and Adhya, S. (1992) Control of gal transcription through DNA looping: inhibition of the initial transcribing complex. *Proc. Natl Acad. Sci. USA*, **89**, 11264–11268.
- Choy, H.E., Park, S.W., Aki, T., Parrack, P., Fujita, N., Ishihama, A. and Adhya, S. (1995) Repression and activation of transcription by Gal and Lac repressors: involvement of alpha subunit of RNA polymerase. *EMBO J.*, **14**, 4523–4529.
- Roy, S., Semsey, S., Liu, M., Gussin, G.N. and Adhya, S. (2004) GalR represses galP1 by inhibiting the rate-determining open complex formation through RNA polymerase contact: a GalR negative control mutant. *J. Mol. Biol.*, **344**, 609–618.
- Aki, T., Choy, H.E. and Adhya, S. (1996) Histone-like protein HU as a specific transcriptional regulator: co-factor role in repression of gal transcription by GAL repressor. *Genes Cells*, **1**, 179–188.
- Geanakopoulos, M., Vasmatzis, G., Lewis, D.E., Roy, S., Lee, B. and Adhya, S. (1999) GalR mutants defective in repressosome formation. *Genes Dev.*, **13**, 1251–1262.
- Semsey, S., Geanakopoulos, M., Lewis, D.E. and Adhya, S. (2002) Operator-bound GalR dimers close DNA loops by direct interaction: tetramerization and inducer binding. *EMBO J.*, **21**, 4349–4356.
- Geanakopoulos, M., Vasmatzis, G., Zhurkin, V.B. and Adhya, S. (2001) Gal repressosome contains an antiparallel DNA loop. *Nat. Struct. Biol.*, **8**, 432–436.
- Semsey, S., Tolstorukov, M.Y., Virnik, K., Zhurkin, V.B. and Adhya, S. (2004) DNA trajectory in the Gal repressosome. *Genes Dev.*, **18**, 1898–1907.
- Aki, T. and Adhya, S. (1997) Repressor induced site-specific binding of HU for transcriptional regulation. *EMBO J.*, **16**, 3666–3674.
- Lewis, D.E. and Adhya, S. (2002) In vitro repression of the gal promoters by GalR and HU depends on the proper helical phasing of the two operators. *J. Biol. Chem.*, **277**, 2498–2504.
- Semsey, S., Virnik, K. and Adhya, S. (2005) A gamut of loops: meandering DNA. *Trends Biochem. Sci.*, **30**, 334–341.
- Ozoline, O.N. and Tsyganov, M.A. (1995) Structure of open promoter complexes with *Escherichia coli* RNA polymerase as revealed by the DNase I footprinting technique: compilation analysis. *Nucleic Acids Res.*, **23**, 4533–4541.
- Sambrook, J. and Russel, D.D. (2001) *Molecular Cloning*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Lewis, D.E. (2003) Identification of promoters of *Escherichia coli* and phage in transcription section plasmid pSA850. *Methods Enzymol.*, **370**, 618–645.
- Liu, M., Garges, S. and Adhya, S. (2004) lacP1 promoter with an extended -10 motif. Pleiotropic effects of cyclic AMP protein at different steps of transcription initiation. *J. Biol. Chem.*, **279**, 54552–54557.
- Ackers, G.K., Johnson, A.D. and Shea, M.A. (1982) Quantitative model for gene regulation by lambda phage repressor. *Proc. Natl Acad. Sci. USA*, **79**, 1129–1133.
- Perez, N., Rehault, M. and Amouyal, M. (2000) A functional assay in *Escherichia coli* to detect non-assisted interaction between galactose repressor dimers. *Nucleic Acids Res.*, **28**, 3600–3604.
- Choy, H.E. and Adhya, S. (1992) Control of gal transcription through DNA looping: inhibition of the initial transcribing complex. *Proc. Natl Acad. Sci. USA*, **89**, 11264–11268.
- Choy, H.E., Park, S.W., Parrack, P. and Adhya, S. (1995) Transcription regulation by inflexibility of promoter DNA in a looped complex. *Proc. Natl Acad. Sci. USA*, **92**, 7327–7331.
- Lia, G., Semsey, S., Lewis, D.E., Adhya, S., Bensimon, D., Dunlap, D. and Finzi, L. (2008) The antiparallel loops in gal DNA. *Nucleic Acids Res.*, **36**, 4204–4210.
- Nakanishi, H., Mitarai, N. and Sneppen, K. (2008) Dynamical analysis on gene activity in the presence of repressors and an interfering promoter. *Biophys. J.*, **95**, 4228–4240.
- Olson, W.K., Gorin, A.A., Lu, X.J., Hock, L.M. and Zhurkin, V.B. (1998) DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl Acad. Sci. USA*, **95**, 11163–11168.
- Dworkin, J., Ninfa, A.J. and Model, P. (1998) A protein-induced DNA bend increases the specificity of a prokaryotic enhancer-binding protein. *Genes Dev.*, **12**, 894–900.

38. Geanakopoulos, M. and Adhya, S. (1997) Functional characterization of roles of GalR and GalS as regulators of the gal regulon. *J. Bacteriol.*, **179**, 228–234.
39. Chatterjee, S., Zhou, Y.N., Roy, S. and Adhya, S. (1997) Interaction of Gal repressor with inducer and operator: induction of gal transcription from repressor-bound DNA. *Proc. Natl Acad. Sci. USA*, **94**, 2957–2962.
40. Krishna, S., Orosz, L., Sneppen, K., Adhya, S. and Semsey, S. (2009) Relation of intracellular signal levels and promoter activities in the gal regulon of *Escherichia coli*. *J. Mol. Biol.*, **391**, 671–678.
41. Shepherd, N., Dennis, P. and Bremer, H. (2001) Cytoplasmic RNA polymerase in *Escherichia coli*. *J. Bacteriol.*, **183**, 2527–2534.
42. Semsey, S., Virnik, K. and Adhya, S. (2006) Three-stage regulation of the amphibolic gal operon: from repressosome to GalR-free DNA. *J. Mol. Biol.*, **358**, 355–363.
43. Krishna, S., Semsey, S. and Sneppen, K. (2007) Combinatorics of feedback in cellular uptake and metabolism of small molecules. *Proc. Natl Acad. Sci. USA*, **104**, 20815–20819.
44. Semsey, S., Krishna, S., Sneppen, K. and Adhya, S. (2007) Signal integration in the galactose network of *Escherichia coli*. *Mol. Microbiol.*, **65**, 465–476.