Potent transcriptional interference by pausing of RNA polymerases over a downstream promoter

Adam C. Palmer[1,2], Alexandra Ahlgren-Berg[1], J. Barry Egan[1], Ian B. Dodd[1] and Keith E. Shearwin[1]*

[1]School of Molecular and Biomedical Science
University of Adelaide,
Adelaide, South Australia, 5005, Australia.
keith.shearwin@adelaide.edu.au
Phone +61 8 8303 5361
Fax +61 8 8303 4362

[2] Present address
Department of Systems Biology
Harvard Medical School
200 Longwood Avenue
Boston, Massachusetts 02115, USA.

* Author to whom correspondence should be addressed

Running Title: Transcriptional interference by RNAP pausing

SUMMARY

Elongating RNA polymerases (RNAPs) can interfere with transcription from downstream promoters by inhibiting DNA binding by RNAP and activators. However, combining quantitative measurement with mathematical modelling, we show that simple RNAP elongation cannot produce the strong asymmetric interference observed between a natural face-to-face promoter pair in bacteriophage lambda. Pausing of elongating polymerases over the RNAP binding site of the downstream promoter is demonstrated in vivo, and is shown by modelling to account for the increased interference. The model successfully predicts the effects on interference of treatments increasing or reducing pausing. Gene regulation by pausing-enhanced occlusion provides a general and potentially widespread mechanism by which even weak converging or tandem transcription, either coding or non-coding, can bring about strong *in cis* repression.

INTRODUCTION

Widespread transcription of non-coding and/or antisense RNA has been found in the genomes of bacteria, yeast, *Drosophila*, *Arabidopsis*, mouse, and humans (Havilio et al., 2005; Johnson et al., 2005), and the thorough study of 1% of the human genome by the ENCODE pilot project has revealed that complex intercalated transcription is no rare peculiarity but the usual state of the genome (The ENCODE Project Consortium, 2007). There has also been an increasing appreciation that these transcripts possess regulatory functions (Katayama et al., 2005; Prasanth and Spector, 2007) and are often developmentally regulated (Wilhelm et al., 2008). Though a diverse range of functions and mechanisms have been uncovered, the largest class of non-coding RNAs are long transcripts of unknown function (Ponting et al., 2009), which are often antisense to coding transcripts (Katayama et al., 2005; Wilhelm et al., 2008). Transcriptional interference (TI), defined as the suppressive influence of one transcriptional process, directly and *in cis*, on a second transcriptional process (Shearwin et al., 2005), is a proven regulatory role for some of these transcripts, and may well be expected to explain the function of a significant fraction more (Bird et al., 2006; Callen et al., 2004; Hongay et al., 2006; Martens et al., 2004; Petruk et al., 2006) .

TI can occur between convergent (face-to-face), tandem (co-directional), or overlapping arrangements of promoters, where the association and elongation of RNA Polymerases (RNAPs) from one promoter disrupts RNAPs and/or transcription factors at a second promoter. TI provides a clear evolutionary benefit by diversifying the range of functions of transcriptional regulators, permitting repressors and activators of one promoter to indirectly activate or repress a second promoter through the alleviation or enforcement of TI (Callen et al., 2004; Martens et al., 2005). Numerous studies of *in cis* TI in eukaryotes and prokaryotes have arrived at a model where the elongation of RNAP over a target promoter causes TI by displacing and preventing the binding of RNAP or transcription factors to the promoter (Adhya and Gottesman, 1982; Bird et al., 2006;

Callen et al., 2004; Greger et al., 1998; Henderson et al., 1989; Martens et al., 2004). In our previous work we developed a mathematical model of TI by RNA polymerase traffic in *E. coli*, which demonstrated that 'occlusion', prevention of RNAP and activator binding by the passage of elongating RNAPs through a promoter, is incapable of generating a substantial quantity of TI unless either the interfering promoter initiates RNAPs at an extraordinarily rapid rate, or is at least 10× stronger than the target promoter, whose kinetic properties must be attuned to maximise vulnerability to TI (Sneppen et al., 2005). Thus, in our current mechanistic understanding of TI, substantial transcriptional interference can only arise in the action of a strong promoter upon a weak promoter.

To further study the mechanisms of action of TI, our model system is the well characterised P_R and P_{RE} promoter pair from bacteriophage λ, where a face-to-face arrangement (separated by 320bp) and preliminary measurements suggest the presence of TI (Schmeissner et al., 1980; Ward and Murray, 1979). While P_R is constitutive, P_{RE} is dependent upon the tetrameric transcriptional activator λCII. (Shih and Gussin, 1984).

In this combined in vivo and in silico study of TI between P_R and P_{RE}, experimental measurements of TI were made by β-galactosidase assays of LacZ reporter constructs, and mathematical modelling of TI was performed with stochastic simulations. We show in vivo that P_R strongly interferes with P_{RE} but not vice-versa, and that this asymmetric strong TI is inexplicable by our previous mechanistic understanding. We go on to show that this strong interference is attributable to a mechanism that can allow a weak promoter to exert TI, pausing-enhanced occlusion, where pausing of RNAP while positioned over a downstream promoter strongly inhibits transcription.

RESULTS

Induction of P_{RE} by CII and measurement of P_R and P_{RE} transcription

Transcription from P_R and P_{RE} was assayed using a single copy, chromosomal *lacZ* transcriptional fusion system (Figure 1A, Supplemental Figures S1 and S2), inserting a 507 bp fragment of the λ genome in different orientations to measure either P_R or P_{RE}. This reporter system bears an RNase III cleavage site between the promoter fragment and *lacZ*, which has been demonstrated to ensure that *lacZ* message stability and translation is independent of the sequence upstream of cleavage(Linn and St Pierre, 1990). Accordingly, LacZ activity is specifically measuring interference occurring at the level of transcription.

Between P_R and P_{RE} lies *cro*, a transcriptional repressor of P_R. To avoid the complications of this negative feedback, the helix-turn-helix motif of *cro* was mutated (YQ→ER), rendering the protein unable to bind DNA (data not shown). Also located between P_R and P_{RE} is tR1, a Rho-dependent terminator of P_R transcription. Rho termination factor binds to the nascent RNA at the *rut* sequence and leads to transcription termination at a cluster of three sites collectively termed tR1 (Banerjee et al., 2006; Lau et al., 1982). The lambda antiterminator protein, N, blocks termination at tR1 by binding to the *nut* site in the nascent RNA and competing with Rho binding (Vieu and Rahmouni, 2004)(Figure 1A). To investigate the termination efficiency of tR1, a shortened P_R-*lacZ* reporter was constructed, entirely lacking *cro*, tR1, and P_{RE}, such that comparison to full length reporters indicated the efficiency of termination.

The effects of TI on both P_R and P_{RE} were measured over a range of P_{RE} activities, by the use of an IPTG inducible CII expression system. *cII* was encoded on a plasmid under the control of pLac, while a second plasmid expressed *lacI* constitutively from pLacI (Figure 1A). Measurements of non-interfered promoter activity were obtained for P_{RE} by mutagenesis of the -10 hexamer of P_R (400-fold reduction in LacZ units), and for P_R by replacing CII expression vector with empty vector, leaving P_{RE} inactive.

$P_{RE}.(P_R^-)$ demonstrates sigmoidal activation of P_{RE} with respect IPTG (Figure 1B). Curiously, this sigmoid has a Hill coefficient of 5.1 (Figure 1B); since CII binds P_{RE} as tetramers, this number should not exceed 4 for P_{RE} versus CII. Western blotting of cells expressing CII under this IPTG-inducible system revealed an upward nonlinearity in CII versus IPTG (Figure 1C, D). Using this data to transform x-coordinates from IPTG to CII, revealed that the Hill coefficient with respect to CII is 3.1 (Figure 1E), within the expected range for the CII tetramer.

Transcription from P_R severely interferes with P_{RE}

Comparison of $P_{RE}.(P_R^+)$ and $P_{RE}.(P_R^-)$ constructs demonstrated that convergent transcription from P_R reduced P_{RE} transcription 5.5-fold, while the overall shape of the sigmoidal activation curve is altered in neither Hill coefficient nor IPTG concentration required for 50% of maximum transcription (EC50) (Figure 2A).

Transcription from P_{RE} does not substantially interfere with P_R

Measured in the absence of CII, and with *lacZ* positioned downstream of tR1, $P_R.(P_{RE}^+)$ produced 380 LacZ units, which by comparison to the shortened P_R *lacZ* construct (P_R (short): 1100 LacZ units), indicates 66% termination at tR1, comparable to a previous in vivo measurement of 75% (Graham, 2004).

Although transcription from P_R was observed to decrease 1.5-fold with CII activation of P_{RE}, IPTG induction of CII similarly produced a decrease in transcription from P_R (short) and also a reporter of the constitutive pBla promoter (Figure 2B). As both the P_R (short) and pBla constructs lack CII binding sites or a convergent promoter, it appears that CII expression indirectly influences LacZ production; this is unsurprising given the known toxic effect of CII on host DNA replication (Kedzierska et al., 2003). IPTG alone had no effect (data not shown). Whether the reduction of P_R is indirect or due to P_{RE}, the TI of P_R by P_{RE} is less than or equal to 1.5-fold, a minor effect.

Development of a mathematical model of transcriptional interference between P_R and P_{RE}

Previous work in this laboratory developed a mathematical model of TI by RNA polymerase traffic in *E.coli*, with analytical, stochastic, and numerical mean-field implementations (Sneppen et al., 2005). This model describes the interference of a constitutive promoter by a stronger constitutive promoter, which itself is unaffected by the interaction. In the case of P_R and P_{RE}, the similar promoter strengths and mutual interference render the analytical and mean-field models inapplicable (Sneppen et al., 2005), leaving the stochastic model, which was adapted to incorporate an activator dependent promoter (P_{RE}) and a unidirectional terminator (tR1).

The stochastic simulation of TI is a discrete time monte carlo, with time steps dt = (RNAP elongation velocity)$^{-1}$, such that elongating RNAPs advance a single base pair per time step. We model RNAP initiation with three steps: reversible binding of RNAP to DNA to form a closed complex, isomerisation of closed complexes to open complexes, and initiation of elongation by open complexes (Supplemental Figure S3). Binding of RNAP to DNA is treated as an equilibrium process acting on a faster timescale than promoter firing, and is thus governed by the equilibrium binding constant, KB, of RNAP to DNA; binding of CII to DNA is treated similarly. Isomerisation of closed to open complexes, and initiation of elongation by open complexes are governed by rate constants k_{co} and k_{oe}, respectively. These steps received stochastic treatment; occurring with probability k.dt per time step, where k is the relevant rate constant. Parameters values are addressed in the Supplemental Data.

Before considering transcriptional interference, the model must reproduce CII activation of P_{RE}, measured by $P_{RE}.(P_R^-)$.*lacZ* (Figure 1B). *In vitro* studies have shown CII to activate P_{RE} by a 15-fold increase in K_B and a 40-fold increase in k_{co} (Shih and Gussin, 1984). Activation of P_{RE} by CII

was thus modelled by increasing K_B and k_{co} linearly with respect to CII occupation of P_{RE}. Calculating CII occupation of P_{RE} requires measurements of total CII molecules per cell as a function of IPTG (Figure 1C, D), the free energy change of CII tetramerisation, and the affinity of CII for DNA. The affinity of CII for DNA is obtained from the measurement that P_{RE} is 50% activated at 3100 CII molecules per cell (K_D = 3.7μM monomers; from Figure 1E). The free energy change of tetramerisation is the only unknown parameter, which must be fitted to produce the measured Hill coefficient: the best fit to $P_{RE}.(P_R^-)$ data is obtained at $\Delta G° = –22.5$ kcal/mol, which compares remarkably well with an in vitro measurement of $\Delta G° = –23.5$ kcal / mol (Ho et al., 1982).

Turning now to TI, the model contains three mechanisms (Figure 3A): occlusion, where elongating RNAPs block access to a promoter; 'sitting duck' interference, where initiation complexes yet to fire from a promoter are removed by elongating RNAPs; and collisions between RNAPs elongating in opposite directions. A fourth possible mechanism unique to activator dependent promoters is dislodgement of DNA-bound activator by elongating RNAP. This increases the effective dissociation rate of the activator and thereby increases the EC50 for activation, but does not alter promoter activity in the presence of saturating amounts of activator. In the event that an activator's intrinsic dissociation rate is much faster than the rate of dislodgement by RNAP, no effect will be seen (Supplemental Figure S4). As P_{RE} is measured to undergo no change in EC50 with TI (Figure 2A), we conclude that this mechanism is not pertinent to P_{RE}, presumably due to rapid CII dissociation kinetics.

The modelling of TI (further details in Supplemental Data) was essentially the same as in (Sneppen et al., 2005), with the exception that it was assumed here that only one RNAP, rather than both RNAPs, is lost (randomly) after a collision event. This assumption is more consistent with recent AFM imaging of collided *E. coli* elongation complexes suggesting that collisions induce

backtracking of one RNAP (Crampton et al., 2006). The maximal activities of the P_R and P_{RE} promoters were estimated to be 0.17 and 0.12 initiations/sec, respectively, based on LacZ activities compared with a P_{bla}.lacZ reporter and estimates of P_{bla} firing rates (Liang et al., 1999). RNAP velocity was set at 60 bp/sec and termination at tR1, with probability 66%, was assumed to be instant.

A mathematical model of existing mechanisms of transcriptional interference is unable to explain the high interference of P_{RE} by P_R

Applying the above quantitative model to TI between P_R and P_{RE}, it was found that the interference of P_{RE} by P_R is inexplicably high. TI can be enhanced by the adjustment of a promoter's kinetic parameters to maximise interference by dislodgement of open complexes; this 'sitting duck' interference is largest when k_{co} and k_{oe} are similar (Sneppen et al., 2005). Upon selecting kinetic parameters for P_{RE} that maximise TI, only 2.5 fold interference could be simulated; far short of the measured 5.5-fold interference (Figure 3B).

With the same kinetic parameters, simulated TI of P_R by P_{RE} is a reasonable match to experimental measurements. The failure of the model to explain the observed TI of P_{RE} suggests that we are missing an asymmetric mechanism of TI, which specifically enhances the ability of P_R to interfere with P_{RE}.

RNAP from P_R pause over P_{RE}

One asymmetry in the interaction of P_R with P_{RE} is the presence of the unidirectional, Rho-dependent terminator tR1. In vitro studies have found termination at tR1 to occur in three clusters located over the P_{RE} promoter (Lau et al., 1982; Roberts et al., 1991)(Figure 4A), and also that Rho-

dependent termination requires pausing at the termination site, allowing time for Rho to travel from its binding site *rut* along the nascent transcript to RNAP to effect termination (Lau et al., 1983; Richardson, 2002; Vieu and Rahmouni, 2004). It appeared possible then that in vivo, RNAPs from P_R pause at tR1, and while so paused occlude the P_{RE} promoter. Indeed, pausing of RNAP from P_R over P_{RE} was confirmed in vivo by potassium permanganate ($KMnO_4$) footprinting, a technique able to detect paused elongation complexes by preferentially modifying single stranded DNA (Figure 4B). Comparison of footprints performed on P_R^- and P_R^+ templates showed a number of bands of increased intensity on the P_R^+ templates, at positions which correspond to the three termination sites previously identified in vitro.

A mutation which enhances pausing of RNAP over P_{RE} increases transcriptional interference

To further investigate if pausing over P_{RE} is a source of TI, P_{RE} was analysed in the presence of a mutation designed to enhance pausing over P_{RE}. Mutations of the Rho-binding site *rut* of the P_R transcript are known to reduce termination at tR1 (Graham, 2004). Weakening the association of Rho and reducing termination at tR1 should result in prolonged occupation of pause sites by RNAP from P_R, since decreased termination implies that: (1) a larger fraction of RNAP will reach the second and third pause sites, and (2) RNAP must be removed from pauses (terminated) less frequently by Rho-factor. We introduced a 3bp mutation to Rho utilisation site A (*rutA*$^-$), which has been demonstrated to reduce termination through modulation of Rho-RNA interactions (Graham, 2004). By LacZ assay of P_R, we confirmed that *rutA*$^-$ reduced termination efficiency, from 66% to 25%.

In vivo pause durations were measured by performing $KMnO_4$ footprinting (Figure 5A, B), following the addition of rifampicin, which prevents reinitiation of RNA synthesis (Hatoum and Roberts, 2008). Pause signals decay as RNAPs leave the pause site but are not replaced, and the rate of loss of the signal gives a measure of pause duration. Scans of pixel intensity down the gel (Figure

5A) clearly show that pausing is only occurring downstream of the *rut*A sequence. Analysis of the most distinctly pR specific pause signals (Figures 4B, 5A) on the wild type template gave pause durations of 33, 33 and 24 seconds at sites I, II and III, which were extended approximately two-fold by *rutA*$^-$ (Figure 5B). Analysis of a number of other bands within tR1 also gave increased pause times on the *rutA*$^-$ template. A second, independent time course experiment (not shown) gave similar pause durations on the wildtype template which were extended ~ 1.3 fold by *rutA*$^-$. These results confirm that mutation of the *rutA* site, and the consequent reduction of termination efficiency, prolongs occupation of the pause sites located over P_{RE}. The RNAP pause times observed are not unusual: a series of *E. coli* promoters have shown pauses of similar magnitude (Hatoum and Roberts, 2008). Comparison of P_{RE}.(*rut*A$^-$ P_R^-) and P_{RE}.(*rut*A$^-$ P_R^+) by LacZ assay revealed that P_R now caused 21-fold TI of P_{RE} (Figure 6A), a major increase in TI relative to P_{RE}.(P_R^+).

'Occlusion by pausing' allows the mathematical model of transcriptional interference to explain the experimental data

Having experimentally confirmed that RNAPs from P_R pause over $P_{RE,}$ and that enhanced pausing increases TI, we next examined whether occlusion by paused RNAPs can quantitatively explain the experimental data using the mathematical model. The model was modified such that RNAPs from P_R pause upon arrival at each of the three tR1 pause/termination sites (Figure 4A). At each site the paused RNAPs prevent RNAP binding to P_{RE}, and face a probability to either terminate or resume elongation at each time-step spent paused. Also included is the ability of a trailing elongating RNAP to force forward translocation of a leading paused RNAP, as has been observed in *E.coli* (Epshtein and Nudler, 2003). Termination probability is defined by the experimentally measured efficiency of termination for wildtype (66%) or *rut*A$^-$ (25%), while the probability to spontaneously resume elongation is the reciprocal of the 'dwell time', or the intrinsic duration of the pause: this

parameter is unknown and was fitted to experimental data (below). The intrinsic duration of the pause is the average time that an RNAP would spend at a pause site in the absence of termination or trailing polymerases. In practice, RNAPs will pause for less time than this 'intrinsic' pause duration, as pauses will be ended prematurely by either termination or forward translocation by a trailing RNAP.

Simulated TI of P_{RE} is seen to increase substantially with increasing intrinsic pause time, while $P_R.(P_{RE}^+).lacZ$ is a reasonable match to experiments with or without RNAP pausing (Figure 3B). The greatest increase in TI of P_{RE} is seen in the presence of *rutA*$^-$, as decreased termination efficiency means that RNAP tend to remain paused for longer. Simulated TI of P_{RE} reaches an asymptote at large intrinsic pause duration, when the probability of spontaneous re-initiation is insignificant relative to the probability of either undergoing termination or of being 'pushed' by a trailing RNAP. At this asymptote, simulated TI of both $P_{RE}.(P_R^+).lacZ$ and $P_{RE}.(rutA^- P_R^+).lacZ$ agree with experimental measurements (Figure 6A). That this asymptote occurs at intrinsic pause durations greater than 20 seconds is in strong agreement with experimentally measured pause durations greater than 30 seconds.

Recalling that pushing by trailing RNAPs causes pauses to be shorter than either intrinsic pause duration or pausing measured in the presence of rifampicin, simulations in the asymptotic case reveal that average RNAP pause times are only 5 seconds, distributed across all three sites. It is striking that 5 seconds of pausing per RNAP is able to quantitatively explain the 5.5-fold interference observed for $P_{RE}.(P_R^+).lacZ$, while in the case of $P_{RE}.(rutA^- P_R^+).lacZ$, average RNAP pause times of 8 seconds can explain the measured 21-fold repression.

The introduction of occlusion by pausing enabled the model to accurately fit the TI of three data sets; $P_R.(P_{RE}^+)$, $P_{RE}.(P_R^+)$ and $P_{RE}.(rutA^- P_R^+)$; by the adjustment of a single parameter: pause

duration (Figure 6A). This fit is not 'fine-tuned', but is satisfied by any pause times greater than 20 seconds, a requirement met by experimentally measured in vivo pause durations. These conclusions are not adversely affected by a different assumption about pushing by trailing polymerases. In the absence of pushing, the model simply selects lower intrinsic pause times because the intrinsic rate of escape through the pause must increase to make up for the lack of push-through. The consequent decrease in pause half-life for a *single* RNAP at a specific site is compensated by the fact that when it moves on, it is replaced by the trailing RNAP.

An obvious additional experimental test of the occlusion by pausing model is provided by considering the effect of a reduction in pausing.

Reduction of pausing by λN reduces interference of P_{RE}

The regulatory λ protein N competitively binds P_R transcripts in place of Rho-factor, and renders the elongating RNAP both termination and pause resistant (Mason et al., 1992). In a model lacking occlusion by pausing, a reduction in termination by N is predicted to increase "normal" occlusion by elongating RNAP and *increase* TI (Figure 6B). Alternatively, if P_{RE} is occluded by paused RNAP, then rendering RNAP from P_R pause resistant should *decrease* the TI of P_{RE}. To test these competing predictions, P_R and P_{RE} were measured in the presence of constitutive chromosomal expression of N. This reduced the termination of P_R transcription at tR1 to 23%, from which it was inferred that approximately half of RNAP are N-modified (Supplemental Data). Measurement of $P_{RE}.(P_R^-)$ showed no influence of N upon P_{RE} in the absence of TI, as expected. Simulations in which 50% of RNAP from P_R were N-modified, and neither paused nor terminated at tR1, produced a 38% increase in P_{RE} activity in the presence of occlusion by pausing, or a 13% decrease in P_{RE} if occlusion by pausing was not present. Experimentally, $P_{RE}.(P_R^+ N^+)$ was 27% stronger than $P_{RE}.(P_R^+)$, confirming the prediction that if paused RNAP occlude P_{RE}, reduced pausing should reduce TI (Figure 6B).

Pausing of RNAP can generate strong occlusion of activator or RNAP binding sites

Transcription has been shown to interfere both with promoter activity and activator binding at sites distal to a promoter (Adhya and Gottesman, 1982; Bird et al., 2006; Callen et al., 2004; Greger et al., 1998; Henderson et al., 1989; Martens et al., 2004). However, strong repression by the passage of elongating RNAPs requires a very high interfering transcription rate, or RNAP flux (Callen et al., 2004; Sneppen et al., 2005). Our study of P_R and P_{RE} shows that TI can be augmented by the pausing of RNAP while passing over a promoter; a result which can in principle also apply to transcription factor binding sites.

An analytical model was constructed to investigate the effect of RNAP elongation and pausing on the availability of a downstream transcription factor binding site (Figure 7A). Considering a 10 bp operator, with a 30 bp long RNAP elongating at a rate of 60 bp/sec, each RNAP occludes the operator for less than 0.7 seconds, and thus even brief pauses can substantially increase occlusion (Figure 7B). In the absence of pausing, very little occlusion ($\leq 40\%$) can be generated even with a flux as high as 1 RNAP per second, while occlusion increases steeply with pause duration even for fluxes as low as 1 RNAP every 20 seconds. If an operator can be occluded by adjacent, non-overlapping pause sites, the fold-interference will be squared: this tactic appears to be utilised at tR1 to obtain strong repression of P_{RE} with brief pauses.

DISCUSSION

When maximally activated by CII, the λP_{RE} promoter reaches 60% of the constitutive strength of λP_R, positioned 320 bp downstream. Transcriptional interference (TI) between these convergent promoters produces no more than a 1.5-fold reduction in P_R activity, but reduces transcription from P_{RE} 5.5-fold. This strong and asymmetric interference is inconsistent with a computational model of TI which predicts moderate mutual TI by the mechanisms of (1) collisions between elongating RNAP, (2) initiation complex dislodgement and (3) promoter occlusion by elongating RNAP. Our acquisition of quantitative data and the use of modelling allowed us to identify a discrepancy between theory and experiment, which demanded a reconsideration of the mechanisms of TI. In vivo footprinting confirmed past in vitro observations: that RNAP from P_R repeatedly pause over P_{RE} in the process of Rho-dependent termination at tR1. Rho-binding site mutations were shown to enhance pausing, and dramatically increased TI of P_{RE} to 21-fold. Expression of the λN antiterminator protein, known to reduce RNAP pausing at tR1 (Mason et al., 1992), partially alleviated TI of P_{RE}. These findings strongly support a new mechanism of TI, occlusion by pausing, in which RNAPs pause over a promoter and sterically hinder the association of RNAP and/or transcription factors. When the experimentally measured RNAP pausing over P_{RE} is introduced to the model, it quantitatively explains all experimental observations.

Though other biological functions for pausing have been described (Landick, 2006; Margaritis and Holstege, 2008), repression of transcription initiation is a new role. Occlusion by paused RNAP provides a mechanism by which the manner of transcription elongation, and by extension elongation regulatory factors, can directly and potently influence the expression of genes encoded on a different transcript. Theoretical analysis of other mechanisms of TI (Sneppen et al., 2005) (Figure 3A), namely collisions, occlusion by elongating RNAP, dislodgement of transcription factors and RNAP (sitting duck), indicates that: (1) adjustment of the quantity of the TI requires

changes to a promoter's activity or kinetic parameters, (2) these mechanisms have a maximal TI that depends upon the strength of the interfering promoter, and (3) extremely strong promoters are necessary to generate more than about 3-fold interference. In contrast, occlusion by pausing can potentially be tuned to any quantity of TI by adjustment of pause duration; demonstrated here by an increase from 5.5 to 21-fold repression due to a doubling of pause times**.** Since the quantity of repression can be widely tuned, independently of promoter activity or kinetics, occlusion by pausing is a powerful and readily evolvable mechanism.

Transcription of antisense or noncoding RNAs has been recently appreciated as a common mechanism of repression. In most eukaryotes these noncoding RNAs can repress their target transcripts with greater than 1:1 efficiency through the catalytic machinery of RNA interference (RNAi). Yet in organisms lacking RNAi such as prokaryotes and *S. cerevisiae*, intergenic and antisense transcription still appears to be a powerful and frequently utilised means of gene regulation, pointing to the existence of other mechanisms (Shearwin et al., 2005). Direct TI, occurring *in cis* by the passage of RNAP across the binding sites of activators and/or RNAP at another promoter, has been observed in bacteria, yeast, flies, HIV and mouse (Adhya and Gottesman, 1982; Bird et al., 2006; Callen et al., 2004; Greger et al., 1998; Henderson et al., 1989; Hongay et al., 2006; Martens et al., 2004; Petruk et al., 2006). We have shown here, in agreement with past work, that strong occlusion of activator and/or RNAP binding sites cannot be obtained even with a very high flux of elongating RNAP (Callen et al., 2004; Sneppen et al., 2005). However, repression of downstream promoters can be drastically improved by the pausing of RNAPs while over activator or RNAP binding sites, producing strong repression with a moderate RNAP flux.

How likely is it that occlusion by pausing, which in contrast to direct TI can produce strong repression with a moderate RNAP flux, is a widespread mechanism of gene control in higher organisms? Several lines of evidence suggest that this may the case. There are many common

genetic contexts which are potentially susceptible to occlusion by pausing (Figure 7C). Both basal promoters and their associated proximal or distal transcription factor binding sites can be influenced by polymerase pausing, as can enhancer sequences, where pausing could cause disruption of complexes that activate or repress distant promoters. The transcription which 'delivers' the polymerase may be either convergent or tandem to the interfered promoter, and significantly, this transcription may be either coding or non-coding (Ponting et al., 2009; Shearwin et al., 2005). Secondly, several studies have shown that at least 85% of eukaryotic genomes are transcribed (Birney et al., 2007; David et al., 2006) and that there is a large degree of overlapping transcription (Katayama et al., 2005), suggesting that promoters and their associated elements are likely to be subject to considerable passing polymerase traffic. Both single molecule (Adelman et al., 2002; Neuman et al., 2003) and bulk in vitro studies suggest that there is an abundance of RNAP pause sites in transcribed DNA (Glover-Cutter et al., 2008) and a study of RNA polymerase II dynamics in live human cells in culture indicated that a significant fraction of elongating polymerases pause for cumulatively long periods (>1 minute) (Darzacq et al., 2007). Given the simple evolutionary adjustment required to appropriately locate a pause site, it seems probable that many instances of *in cis* transcriptional interference will feature occlusion by RNAP pausing as a primary mechanism of repression.

EXPERIMENTAL PROCEDURES

Strains and media

NK7049 (Δ*lacIZYA*)X74 *galOP308* StrR Su^- from R. Simons (Simons et al., 1987) was the host for all LacZ assays. DH5α and MC1061 were hosts for recombinant DNA work. Cells were grown at 37°C in Luria Broth with the addition of carbenicillin (100 μg/mL for pZS15) and kanamycin (50 μg/mL for pUHA1).

Constructs

Fusions of the region of λ containing P_R and P_{RE} (λ: 37954–38461)(Supplemental Figure S1) were first formed in the LacZ reporter plasmid pTL61T (Linn and St Pierre, 1990), and then transferred to the LacZ reporter phage λRS45ΔYA for insertion in single copy into the *E.coli* chromosome, as described in (Dodd et al., 2001) and Supplemental Figure S2. The pBla promoter was amplified from plasmid pBR322, in a product extending from -171 to +9, as per (Liang et al., 1999). This reporter system bears an RNase III cleavage site between the promoter fragment and *lacZ*, which reduces contextual differences in LacZ translation, and should prevent any antisense RNA interactions between P_R and P_{RE} transcripts from influencing LacZ activity. When preparing reporters of weakly terminated P_R.*lacZ* (P_R (short), *rut*A^- and N^+), shuttle strains expressing λCI protein were used to repress strong LacZ transcription. P_R^-, cro^-, and *rut*A^- mutations were constructed using Quikchange oligonucleotide mutagenesis (Stratagene).

N expression system

λN was expressed from single copy in the *E.coli* chromosome in the integrating plasmid pAH162 (Haldimann and Wanner, 2001), with the bacteriophage 186pR promoter (186: 2656–2784) transcribing *N* (λ: 35383–35016).

CII expression system

λCII was expressed from the wild-type pLac promoter on the ampicillin resistant plasmid pZS15, which is pZE15 (Dodd et al., 2001) with the colE1 origin replaced by the pSC101 origin (Lutz and Bujard, 1997). The wild-type λ*cII* gene (λ: 38357–38662) was inserted into pZS15 downstream of pLac, with the intervening sequence containing *lacZ* up to stop codons at aa 20 and 21, followed after 15bp by the pET ribosome binding site AGGAGA to drive efficient CII translation. To control CII expression from pLac on pZS15*cII*, Lac repressor was supplied by pUHA1, a p15A plasmid encoding kanamycin resistance and carrying the wild-type *lacI* gene and promoter, obtained from H. Bujard (Heidelberg University, Germany).

LacZ assays

Kinetic LacZ assays were performed in 96-well microtitre plates by a protocol modified from (Dodd et al., 2001). Fresh colonies on selective LB plates were resuspended in LB and used to inoculate 200μL of selective LB in a 96-well microtitre plate, sealed and incubated overnight at 37°C without shaking. Overnight cultures were diluted into LB in proportion to their density, approximately 3-fold, before further diluting 2μL into 98μL of fresh selective LB plus IPTG in a microtitre plate. Cultures were incubated with rotation at 37°C until OD_{600} reached 0.65 – 0.75 (log phase) and were then assayed for LacZ activity (Dodd et al., 2001).

Western blotting of CII

Cultures of NK7049 pUHA1 pZS15*cII* were grown to log phase by the same protocol as for LacZ assays. Cultures were resuspended in 1/25th volume B-PER lysis reagent (Pierce) supplemented with 0.2 mg/mL fresh lysozyme and 2.5 units of Benzonase (Merck), incubated on ice for 1 hour, diluted 1:1 with 2× Novex Tricine SDS sample buffer (Invitrogen), and heated to 85°C for 2 minutes. Pure CII protein was supplied by Pradeep Parrack (Bose Institute, India). For consistency of background bands upon western blotting, pure CII protein was added to cII^- extract prepared as per cII^+ extracts, such that for both pure CII and cII^+ extracts 5μL of cell extract was loaded in a net volume of 10μL for each lane during electrophoresis. For quantitation of IPTG concentrations 100 – 300μM, cII^+ extracts were diluted 1.5μL into 3.5μL of cII^- extract.

Membranes were scanned and images analysed by a Typhoon Trio and ImageQuant (Amersham). Numbers of *E.coli* cells per sample were determined by measurement of culture density prior to harvesting and comparison to colony forming assays.

In vivo permanganate footprinting

In vivo permanganate footprinting was performed on *E. coli* strain NK7049 carrying pTL61T based LacZ reporter plasmids. $P_{RE}.(P_R^-)$ and $P_{RE}.(P_R^+)$ constructs, containing either a wild type or mutated *rut*A site were used. There was no source of CII in these strains, in order that there was minimal transcription from P_{RE}. To determine the location of the presumed pause site(s) of RNAP transcribing from P_R, plasmid DNA prepared from $KMnO_4$ treated cultures were subject to primer extension analysis, as detailed in Supplemental Data.

Stochastic simulations of transcriptional interference

The essential mechanics of the stochastic simulation of TI are described in Results with further details in the Supplemental Data. Simulated promoter activities in Figures 2 and 3 are the result of

simulating 200 hours of transcription at each point. The program, written in FORTRAN, is available on request.

ACKNOWLEDGEMENTS

We thank Kim Sneppen for advice and support, and Pradeep Parrack and Amos Oppenheim for their gifts of pure CII protein and anti-CII antibody. Research in the Shearwin laboratory is supported by the U.S. NIH (GM062976) and the Australian Research Council.

REFERENCES

Adelman, K., La Porta, A., Santangelo, T. J., Lis, J. T., Roberts, J. W., and Wang, M. D. (2002). Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior. Proc Natl Acad Sci U S A *99*, 13538-13543.

Adhya, S., and Gottesman, M. (1982). Promoter occlusion: transcription through a promoter may inhibit its activity. Cell *29*, 939-944.

Banerjee, S., Chalissery, J., Bandey, I., and Sen, R. (2006). Rho-dependent transcription termination: more questions than answers. J Microbiol *44*, 11-22.

Bird, A. J., Gordon, M., Eide, D. J., and Winge, D. R. (2006). Repression of ADH1 and ADH3 during zinc deficiency by Zap1-induced intergenic RNA transcripts. Embo J *25*, 5726-5734.

Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigo, R., Gingeras, T. R., Margulies, E. H., Weng, Z., Snyder, M., Dermitzakis, E. T., Thurman, R. E.*, et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799-816.

Callen, B. P., Shearwin, K. E., and Egan, J. B. (2004). Transcriptional interference between convergent promoters caused by elongation over the promoter. Mol Cell *14*, 647-656.

Crampton, N., Bonass, W. A., Kirkham, J., Rivetti, C., and Thomson, N. H. (2006). Collision events between RNA polymerases in convergent transcription studied by atomic force microscopy. Nucleic Acids Res *34*, 5416-5425.

Darzacq, X., Shav-Tal, Y., de Turris, V., Brody, Y., Shenoy, S. M., Phair, R. D., and Singer, R. H. (2007). In vivo dynamics of RNA polymerase II transcription. Nat Struct Mol Biol *14*, 796-806.

David, L., Huber, W., Granovskaia, M., Toedling, J., Palm, C. J., Bofkin, L., Jones, T., Davis, R. W., and Steinmetz, L. M. (2006). A high-resolution map of transcription in the yeast genome. Proc Natl Acad Sci U S A *103*, 5320-5325.

Dodd, I. B., Perkins, A. J., Tsemitsidis, D., and Egan, J. B. (2001). Octamerization of lambda CI repressor is needed for effective repression of P(RM) and efficient switching from lysogeny. Genes Dev *15*, 3013-3022.

Epshtein, V., and Nudler, E. (2003). Cooperation between RNA polymerase molecules in transcription elongation. Science *300*, 801-805.

Glover-Cutter, K., Kim, S., Espinosa, J., and Bentley, D. L. (2008). RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. Nat Struct Mol Biol *15*, 71-78.

Graham, J. E. (2004). Sequence-specific Rho-RNA interactions in transcription termination. Nucleic Acids Res *32*, 3093-3100.

Greger, I. H., Demarchi, F., Giacca, M., and Proudfoot, N. J. (1998). Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. Nucleic Acids Res *26*, 1294-1301.

Haldimann, A., and Wanner, B. L. (2001). Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. J Bacteriol *183*, 6384-6393.

Hatoum, A., and Roberts, J. (2008). Prevalence of RNA polymerase stalling at Escherichia coli promoters after open complex formation. Mol Microbiol *68*, 17-28.

Havilio, M., Levanon, E. Y., Lerman, G., Kupiec, M., and Eisenberg, E. (2005). Evidence for abundant transcription of non-coding regions in the Saccharomyces cerevisiae genome. BMC Genomics *6*, 93.

Henderson, S. L., Ryan, K., and Sollner-Webb, B. (1989). The promoter-proximal rDNA terminator augments initiation by preventing disruption of the stable transcription complex caused by polymerase read-in. Genes Dev *3*, 212-223.

Ho, Y., Lewis, M., and Rosenberg, M. (1982). Purification and properties of a transcriptional activator. The cII protein of phage lambda. J Biol Chem *257*, 9128-9134.

Hongay, C. F., Grisafi, P. L., Galitski, T., and Fink, G. R. (2006). Antisense transcription controls cell fate in Saccharomyces cerevisiae. Cell *127*, 735-745.

Johnson, J. M., Edwards, S., Shoemaker, D., and Schadt, E. E. (2005). Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments. Trends Genet *21*, 93-102.

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C. C., Suzuki, M., Kawai, J.*, et al.* (2005). Antisense transcription in the mammalian transcriptome. Science *309*, 1564-1566.

Kedzierska, B., Glinkowska, M., Iwanicki, A., Obuchowski, M., Sojka, P., Thomas, M. S., and Wegrzyn, G. (2003). Toxicity of the bacteriophage lambda cII gene product to Escherichia coli arises from inhibition of host cell DNA replication. Virology *313*, 622-628.

Landick, R. (2006). The regulatory roles and mechanism of transcriptional pausing. Biochem Soc Trans *34*, 1062-1066.

Lau, L. F., Roberts, J. W., and Wu, R. (1982). Transcription terminates at lambda tR1 in three clusters. Proc Natl Acad Sci U S A *79*, 6171-6175.

Lau, L. F., Roberts, J. W., and Wu, R. (1983). RNA polymerase pausing and transcript release at the lambda tR1 terminator in vitro. J Biol Chem *258*, 9391-9397.

Liang, S., Bipatnath, M., Xu, Y., Chen, S., Dennis, P., Ehrenberg, M., and Bremer, H. (1999). Activities of constitutive promoters in Escherichia coli. J Mol Biol *292*, 19-37.

Linn, T., and St Pierre, R. (1990). Improved vector system for constructing transcriptional fusions that ensures independent translation of lacZ. J Bacteriol *172*, 1077-1084.

Lutz, R., and Bujard, H. (1997). Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic Acids Res *25*, 1203-1210.

Margaritis, T., and Holstege, F. C. (2008). Poised RNA polymerase II gives pause for thought. Cell *133*, 581-584.

Martens, J. A., Laprade, L., and Winston, F. (2004). Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. Nature *429*, 571-574.

Martens, J. A., Wu, P. Y., and Winston, F. (2005). Regulation of an intergenic transcript controls adjacent gene transcription in Saccharomyces cerevisiae. Genes Dev *19*, 2695-2704.

Mason, S. W., Li, J., and Greenblatt, J. (1992). Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage lambda. J Biol Chem *267*, 19418-19426.

Neuman, K. C., Abbondanzieri, E. A., Landick, R., Gelles, J., and Block, S. M. (2003). Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. Cell *115*, 437-447.

Petruk, S., Sedkov, Y., Riley, K. M., Hodgson, J., Schweisguth, F., Hirose, S., Jaynes, J. B., Brock, H. W., and Mazo, A. (2006). Transcription of bxd noncoding RNAs promoted by trithorax represses Ubx in cis by transcriptional interference. Cell *127*, 1209-1221.

Ponting, C. P., Oliver, P. L., and Reik, W. (2009). Evolution and Functions of Long Noncoding RNAs. Cell *136*, 629-641.

Prasanth, K. V., and Spector, D. L. (2007). Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. Genes Dev *21*, 11-42.

Richardson, J. P. (2002). Rho-dependent termination and ATPases in transcript termination. Biochim Biophys Acta *1577*, 251-260.

Roberts, E. A., Eisenbraun, T. L., Andrews, C. L., and Bear, D. G. (1991). 3'-end formation at the phage lambda tR1 rho-dependent transcription termination site. Biochemistry *30*, 5429-5437.

Schmeissner, U., Court, D., Shimatake, H., and Rosenberg, M. (1980). Promoter for the establishment of repressor synthesis in bacteriophage lambda. Proc Natl Acad Sci U S A *77*, 3191-3195.

Shearwin, K. E., Callen, B. P., and Egan, J. B. (2005). Transcriptional interference--a crash course. Trends Genet *21*, 339-345.

Shih, M. C., and Gussin, G. N. (1984). Role of cII protein in stimulating transcription initiation at the lambda PRE promoter. Enhanced formation and stabilization of open complexes. J Mol Biol *172*, 489-506.

Simons, R. W., Houman, F., and Kleckner, N. (1987). Improved single and multicopy lac-based cloning vectors for protein and operon fusions. Gene *53*, 85-96.

Sneppen, K., Dodd, I. B., Shearwin, K. E., Palmer, A. C., Schubert, R. A., Callen, B. P., and Egan, J. B. (2005). A mathematical model for transcriptional interference by RNA polymerase traffic in Escherichia coli. J Mol Biol *346*, 399-409.

The ENCODE Project Consortium (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799-816.

Vieu, E., and Rahmouni, A. R. (2004). Dual role of boxB RNA motif in the mechanisms of termination/antitermination at the lambda tR1 terminator revealed in vivo. J Mol Biol *339*, 1077-1087.

Ward, D. F., and Murray, N. E. (1979). Convergent transcription in bacteriophage lambda: interference with gene expression. J Mol Biol *133*, 249-266.

Wilhelm, B. T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C. J., Rogers, J., and Bahler, J. (2008). Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature *453*, 1239-1243.

FIGURE LEGENDS

Figure 1. Activation of P_{RE} by IPTG induction of CII.

(A) Diagram of single-copy chromosomal *lacZ* reporter and plasmid system for IPTG regulated expression of CII. IPTG regulated expression of CII is achieved with a two plasmid system: pUHA1 contains pLacI-*lacI* and pZS15*cII* contains pLac-*cII*. The region of lambda used in the reporter system is shown as a thicker line. The relative locations of the Rho utilisation site (*rut*), N utilisation site (*nut*) and the tR1 terminator are shown.

(B) Activity of $P_{RE}.(P_R^-).lacZ$ (n=12) as a function of IPTG, which induces expression of CII from pLac in pZS15*cII*. Error bars in this and all subsequent figures are 95% confidence limits. The line connecting points is the Hill function of best fit. The average Hill coefficient is 5.1 ±0.4 (95% confidence limits).

(C) Western blotting of CII from *E.coli* pUHA1 pZS15*cII* grown in a range of IPTG concentrations, quantitated against a calibration curve of pure CII protein added to cII^- *E.coli* extracts. For the IPTG concentrations 100 to 300μM (lower panel), cII^+ extracts were diluted 3.3-fold into cII^- extracts, to give a quantity of CII that lies within the calibration curve. The 100μM IPTG point was measured in both diluted and undiluted form, giving average measurements within 6% of one another.

(D) Quantitation of CII western blotting, showing the results of three independent sets of western blots (filled, open, and grey). Grey circles are the data of (C), and a continuous line shows the average of all data.

(E) Activity of $P_{RE}.(P_R^-).lacZ$ plotted as a function of CII molecules per cell. Using the western blotting of (C) and (D), CII molecules per cell were measured for 7 IPTG concentrations. CII

molecules per cell for those IPTG concentrations not directly measured were determined by interpolation between measured data. The average Hill coefficient is 3.1 ±0.3 (95% confidence limits). Assuming an average cellular volume of 1.4 fL, we can determine from this data that the in vivo affinity of CII for its binding site is K_D = 3.7 μM.

Figure 2. Transcriptional interference between P_R and P_{RE}.

(A) Activity of $P_{RE}.(P_R^-)$.*lacZ* and $P_{RE}.(P_R^+)$.*lacZ* (n=14), demonstrating a 5.5-fold repression of P_{RE} due to TI from P_R. Lines are Hill functions of best fit; the fitted EC50s are 78 μM IPTG for $P_{RE}.(P_R^-)$ and 76 μM IPTG for $P_{RE}.(P_R^+)$.

(B) Activity of P_R (short).*lacZ* (n=6), $P_R.(P_{RE}^+)$.*lacZ* (n=15), and pBla.*lacZ* (n=10), in response to IPTG induction of CII. Only $P_R.(P_{RE}^+)$.*lacZ* contains a CII binding site or CII-activated promoter in between the promoter and *lacZ*.

Figure 3. Simulations of transcriptional interference cannot explain experimental observations using existing mechanisms.

(A) Schematic of different mechanisms of transcriptional interference, in which the left (darker) promoter is interfering with the right (lighter) promoter.

(B) Alignment of experimental measurements of promoter activity with the existing mathematical model of TI. Points are experimental data and lines are the results of stochastic simulations of TI, incorporating the mechanisms of occlusion, 'sitting duck', and collisions.

Figure 4. RNAP from pR pauses at tR1.

(A) Sequence of tR1 and P_{RE}. Above the sequence the three in vitro pause/termination sites at tR1 (Lau et al., 1982) are marked, together with the expected protected region for RNAP paused at each

of these pause sites. Below the sequence is marked the binding region for RNAP at P_{RE}, demonstrating that RNAP paused at tR1 should sterically hinder the association of RNAP to P_{RE}.

(B) In vivo potassium permanganate footprinting was performed to identify RNAP pause sites at tR1, for *rut*A^+ and *rut*A^- templates. To ensure pauses were specific for transcription originating from pR, pR^+ and pR^- templates were compared. Footprints were obtained for both top and bottom strands, and run alongside lanes containing dideoxy sequencing reactions which had been generated using the same primers (only the A lane is shown). The bands which are the most distinctly pR specific are indicated by arrowheads. The indicated bands were observed in several (n=5) independent footprint reactions and are consistent with the known three tR1 termination sites observed in vitro (indicated at the side of figure).

Figure 5. A *rut*A mutation extends the lifetime of the pauses

(A) Measurement of pause durations at the three tR1 sites by in vivo permanganate footprinting following addition of rifampicin. Footprints were obtained for the bottom strand on both *rut*A^+ and *rut*A^- templates and run alongside lanes containing dideoxy sequencing reactions which had been generated using the same primer (only the A and C lanes are shown). The appearance of a strong pR band indicates accumulation of open complexes at pR, showing that rifampicin is blocking further rounds of RNA synthesis. In contrast, the tR1 pause signals decay with time, as the paused RNAP either terminates or resumes elongation. Plots of pixel intensity down the *rut*A^+ lanes of the gel are overlaid for the 15 (blue), 30 (green), 60 (red) and 300 (black) second time points. Black dots indicate the pR-specific bands (Figure 4B) which were used for the estimation of pause durations at sites I, II and III shown in (B).

(B) The average RNAP pause durations for *rut*A^+ and *rut*A^- templates at each of the three tR1 pause sites were estimated by plotting the rate of loss of signal with time following addition of

rifampicin. The average pause durations, calculated as 1/slope of these plots, are indicated within each graph. The intensity of the 15 second time point was used as the initial value, in order to allow time for pR derived polymerases, which were elongating at the time of rifampicin addition, to reach tR1. Average pause durations were consistently increased on the *rut*A$^-$ templates.

Figure 6. Occlusion by paused RNAP can explain the strong transcriptional interference of P_{RE}

(A) Experimental and simulated TI of P_{RE}, in *rut*A$^+$ and *rut*A$^-$ conditions, with simulations incorporating occlusion by paused RNAP. Points are experimental data: P_{RE}.(P_R^-).*lacZ* (blue), P_{RE}.(P_R^+).*lacZ* (red), and P_{RE}.(*rut*A$^-$ P_R^+).*lacZ* (black) (n=7). P_{RE}.(*rut*A$^-$ P_R^+).*lacZ* has been scaled up to normalise P_{RE} activity in the absence of TI, to facilitate comparison of fold-interference against *rut*A$^+$ constructs; this is necessitated by *rut*A$^-$ causing a 27% decrease in P_{RE} LacZ activity in the absence of TI. Lines are stochastic simulations of TI: dotted lines are simulations without RNAP pausing at tR1, and solid lines are simulations with intrinsic RNAP pause durations at tR1 of 20 seconds. Marked along the right of the graph are the predicted maximum activities of P_{RE}.(P_R^-).*lacZ* (red) and P_{RE}.(*rut*A$^-$ P_R^+).*lacZ* (black) with intrinsic pause durations of 0, 2, 5 and 20+ seconds, illustrating how pausing of RNAP progressively increases repression of P_{RE}. The thickness of the 20+ second line spans the range of promoter activity calculated for pause durations from 20 to 1000 seconds.

(B) As for (A), showing the effects of reduced pausing due to λN in green: P_{RE}.(P_R^+ N$^+$).*lacZ* (n=10). Expression of λN had no influence upon P_{RE}.(P_R^-).*lacZ* (data not shown).

Figure 7. Substantial occlusion can only be obtained with RNAP pausing.

(A) Analytical model of occlusion by RNAP pausing. RNAP with an occlusion length of *l* bp, travelling at velocity *v* bp / sec, arrive at an operator of length *m* bp, with flux *f* (RNAP / second).

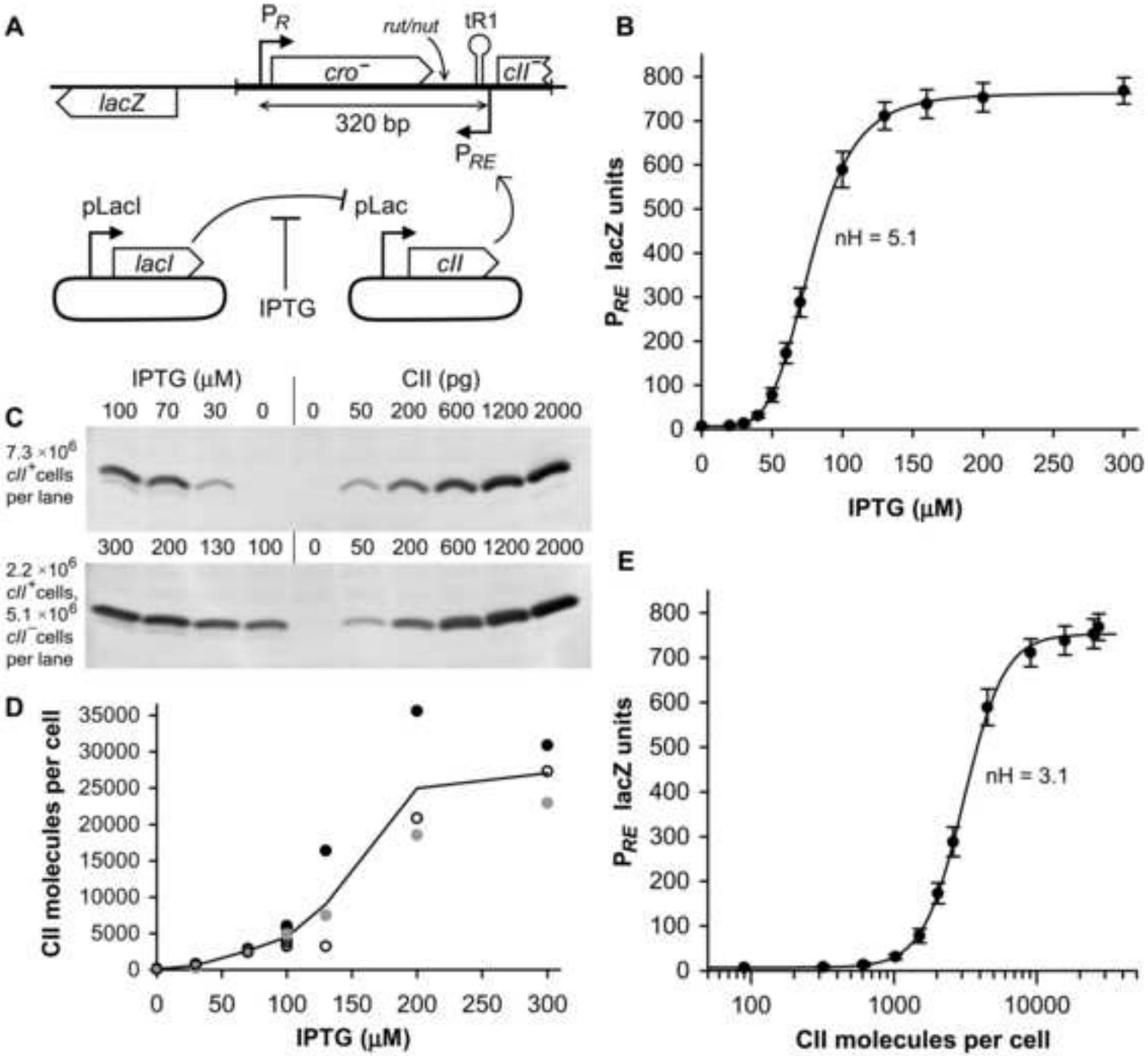
The RNAP move to the end of the operator with rate $v.(l + m)^{-1}$, whereupon they reach a pause site and remain paused for an average duration P sec. Thus with rate P^{-1} the paused state can change to a vacant state, or with rate f another RNAP may arrive at the start of the operator. The fraction of time the operator is occluded is given by the equation shown. This result is independent of whether the pause site is positioned at the start or end of the operator.

(B) Occlusion as a function of pause duration and RNAP flux, calculated for a m = 10bp operator being occluded by l = 30bp long RNAPs which elongate at a rate of v = 60 bp/sec.

(C) Genetic arrangements where occlusion by pausing may regulate gene expression. Transcription may be tandem (upper) or convergent (lower) to the interfered promoter and may be either coding or noncoding. Pausing of RNAP over the promoter itself, or over associated elements such as transcription factor binding sites (TF) or enhancer sequences, may lead to reduction in promoter activity.

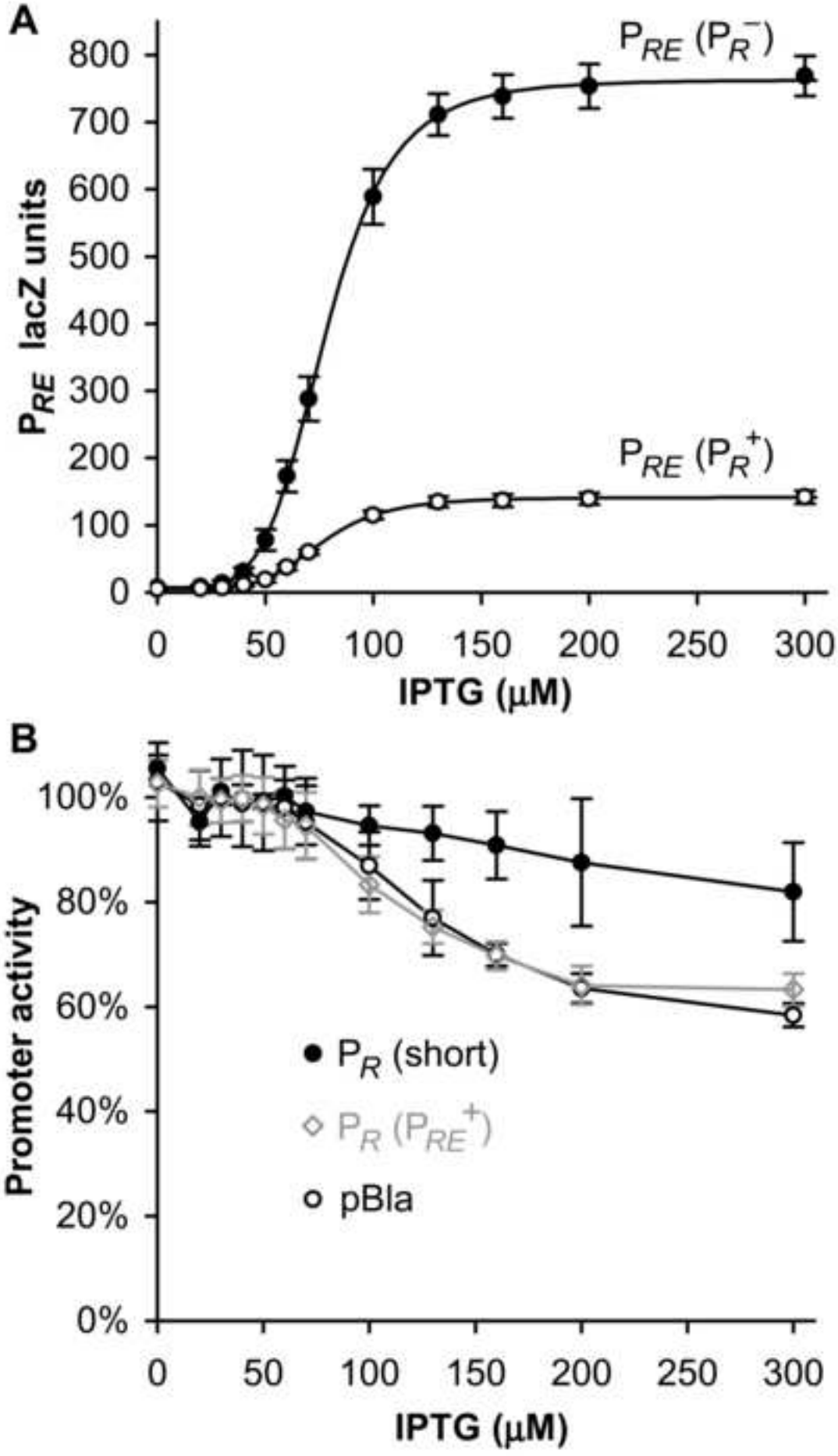
E) Figure

Click here to download high resolution image



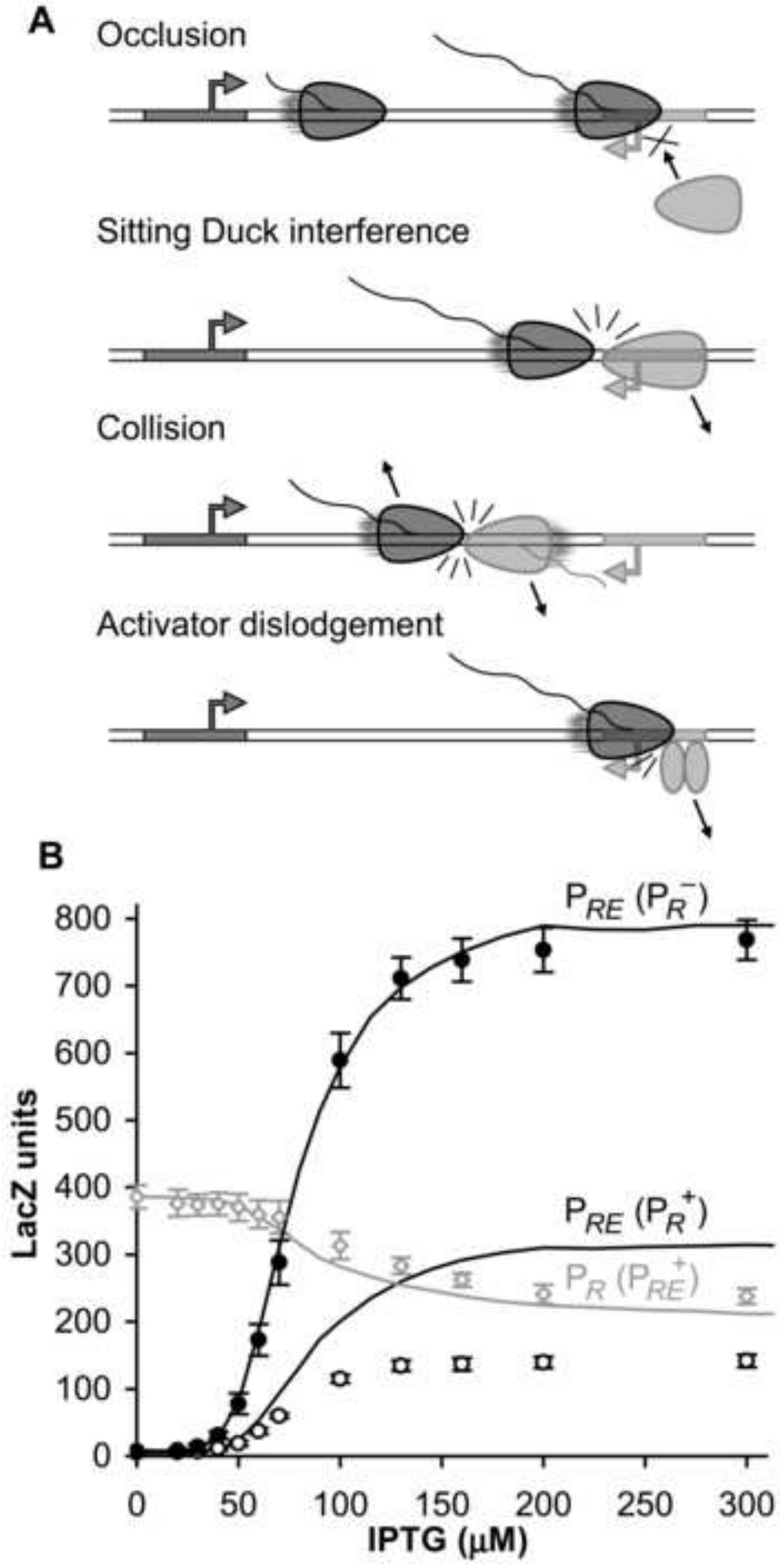
E) Figure

Click here to download high resolution image



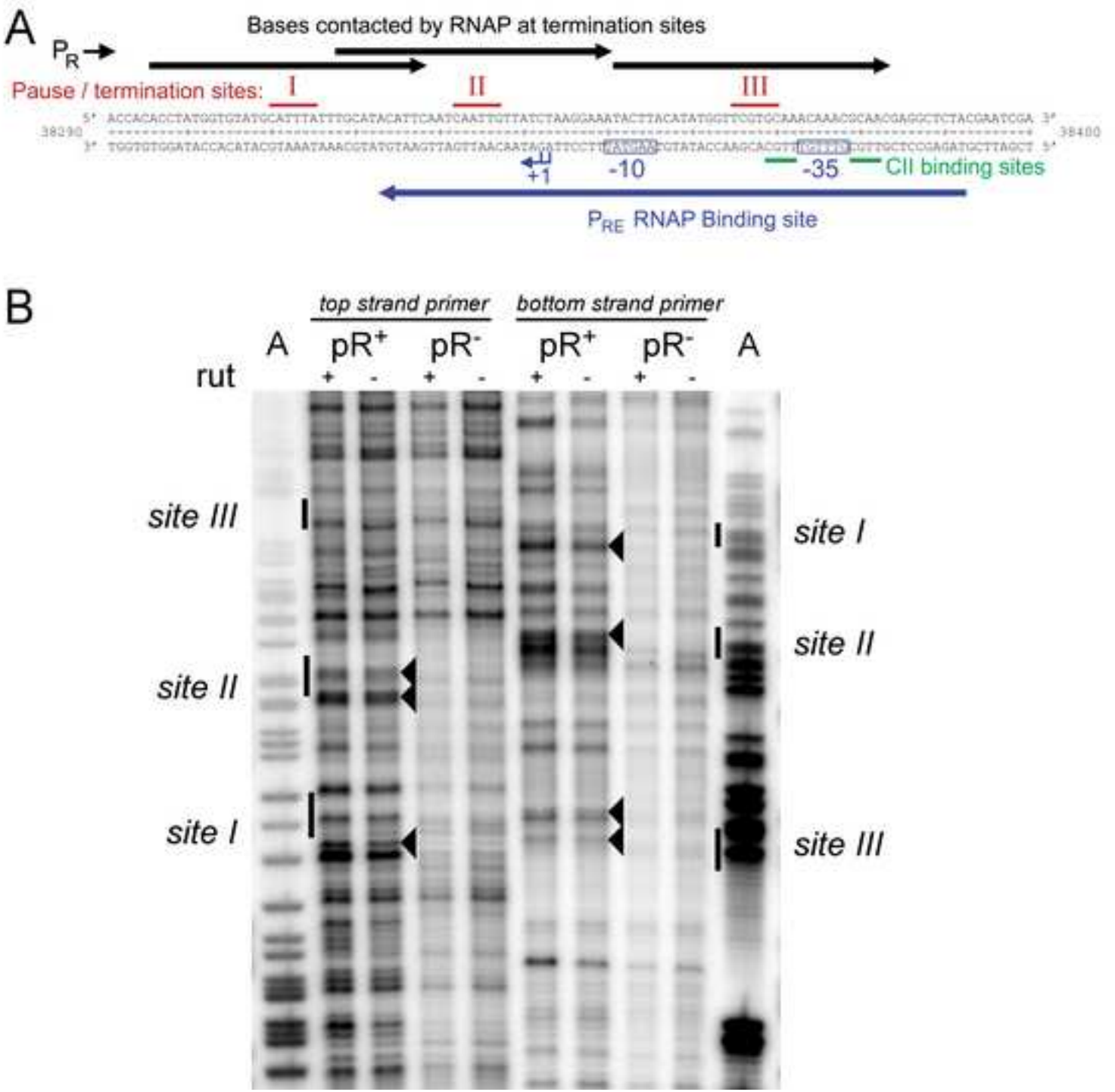
E) Figure

Click here to download high resolution image



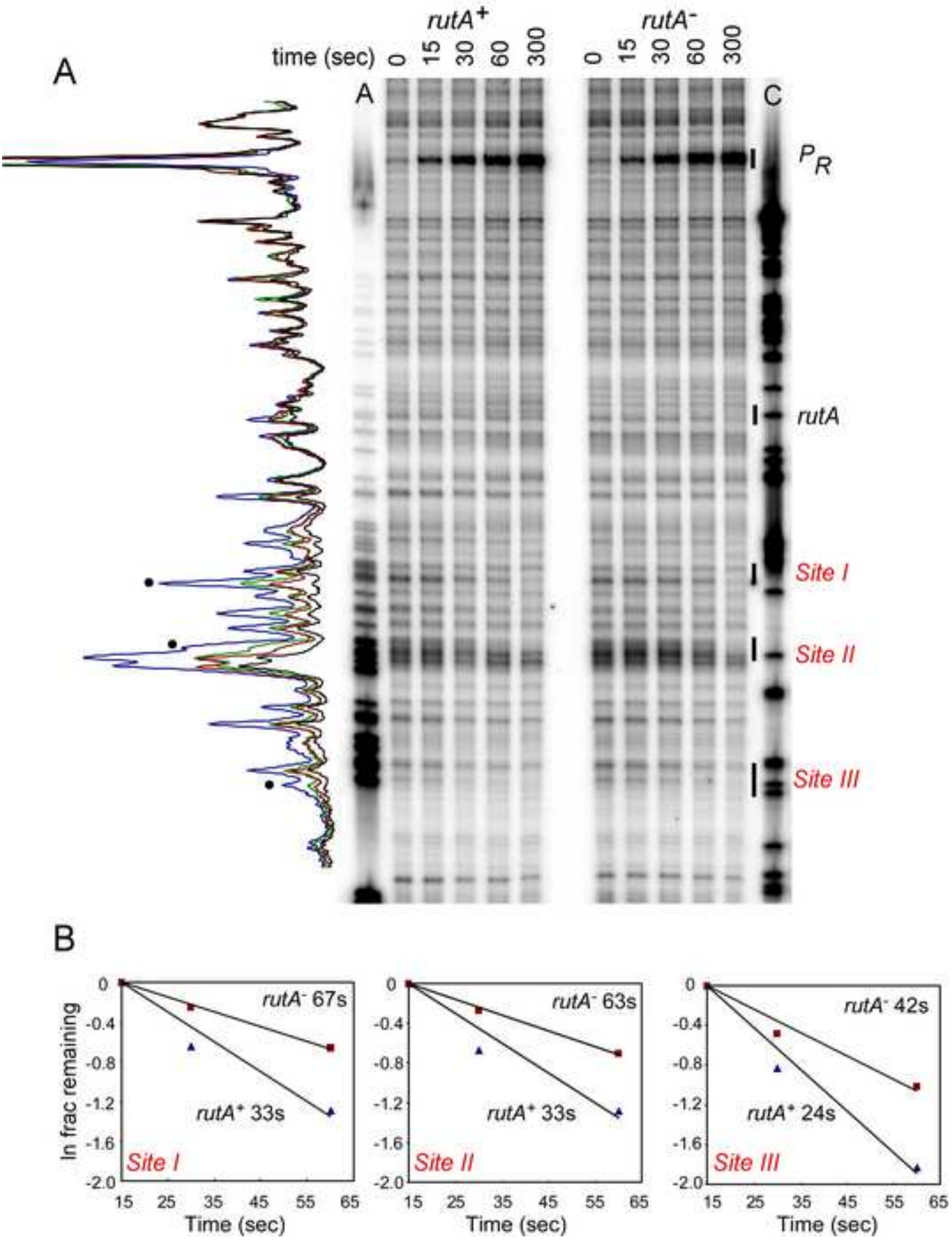
E) Figure

Click here to download high resolution image



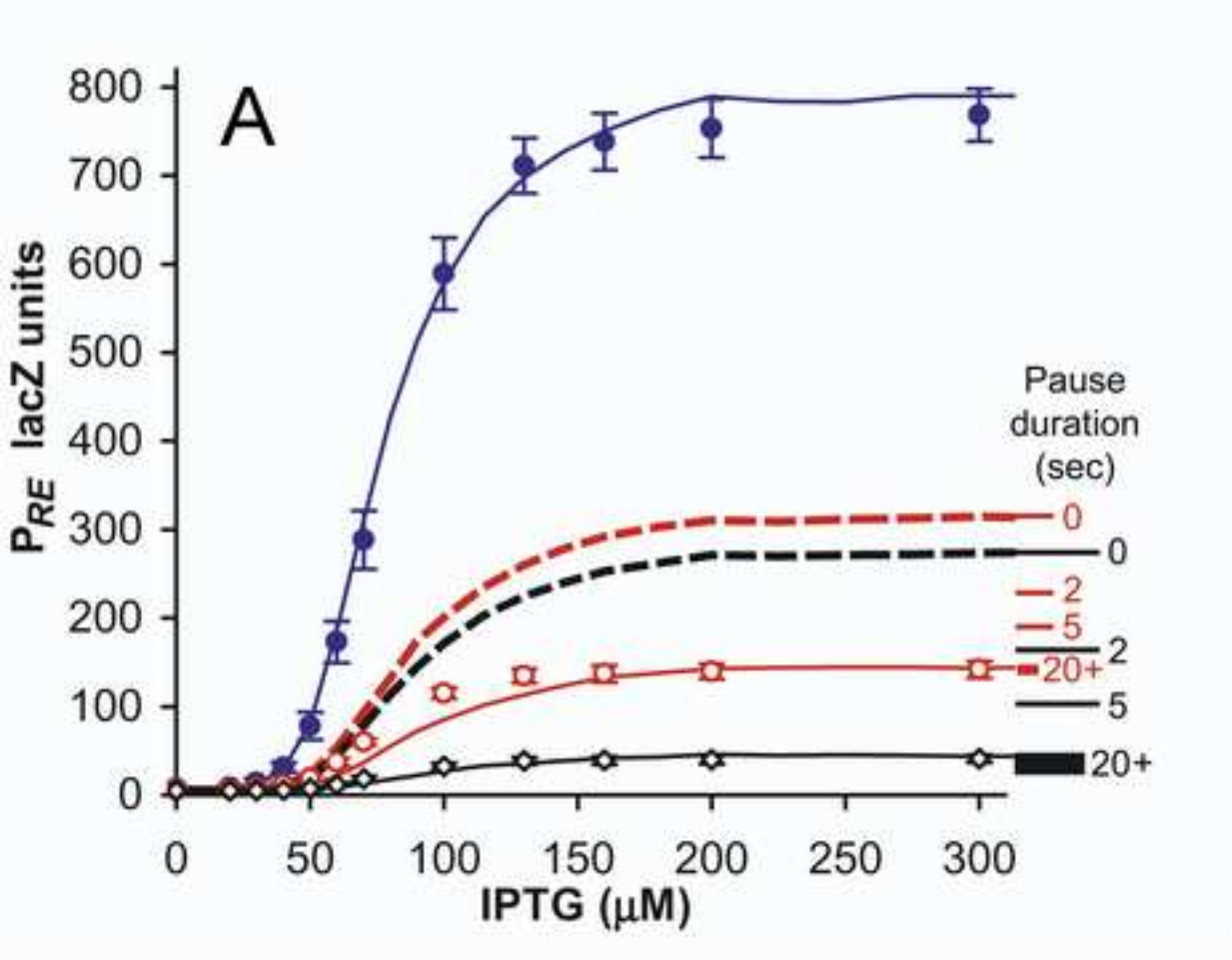
E) Figure

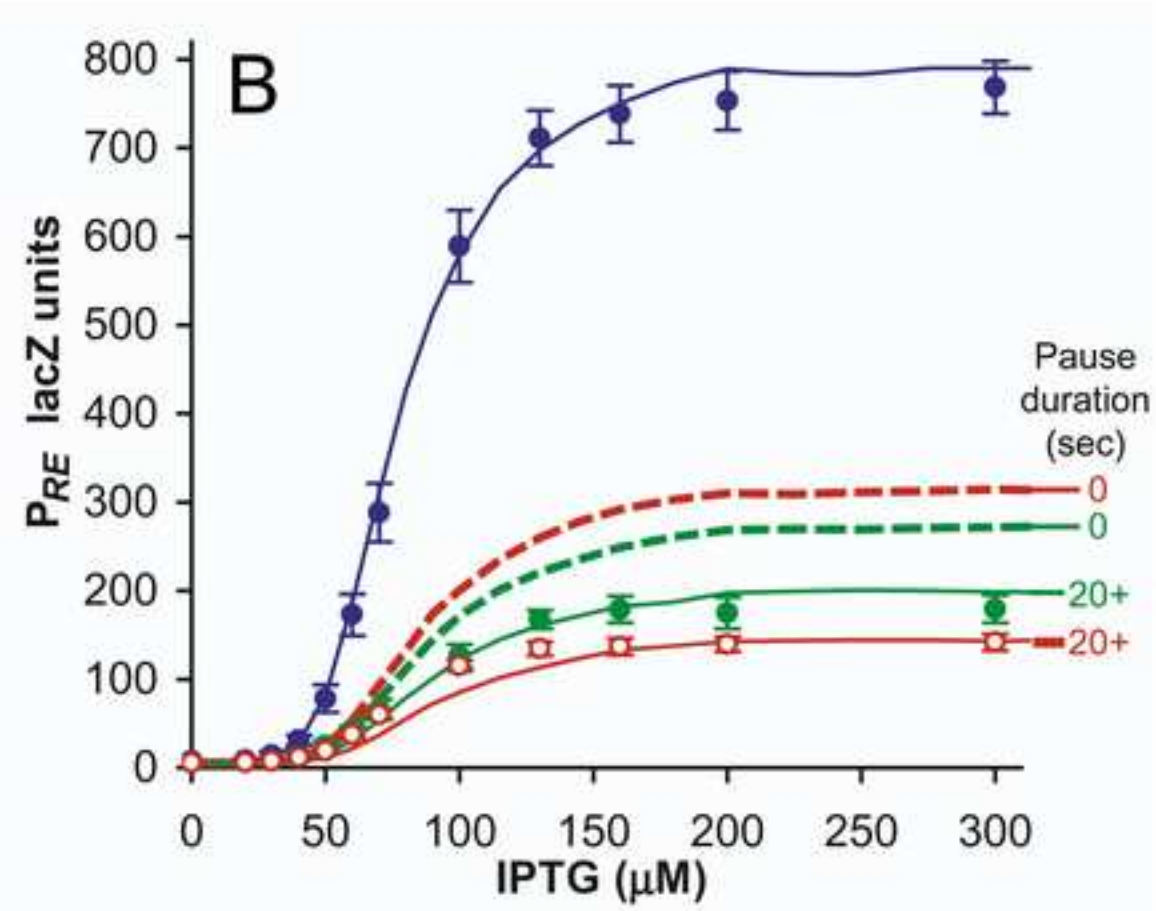
Click here to download high resolution image



E) Figure

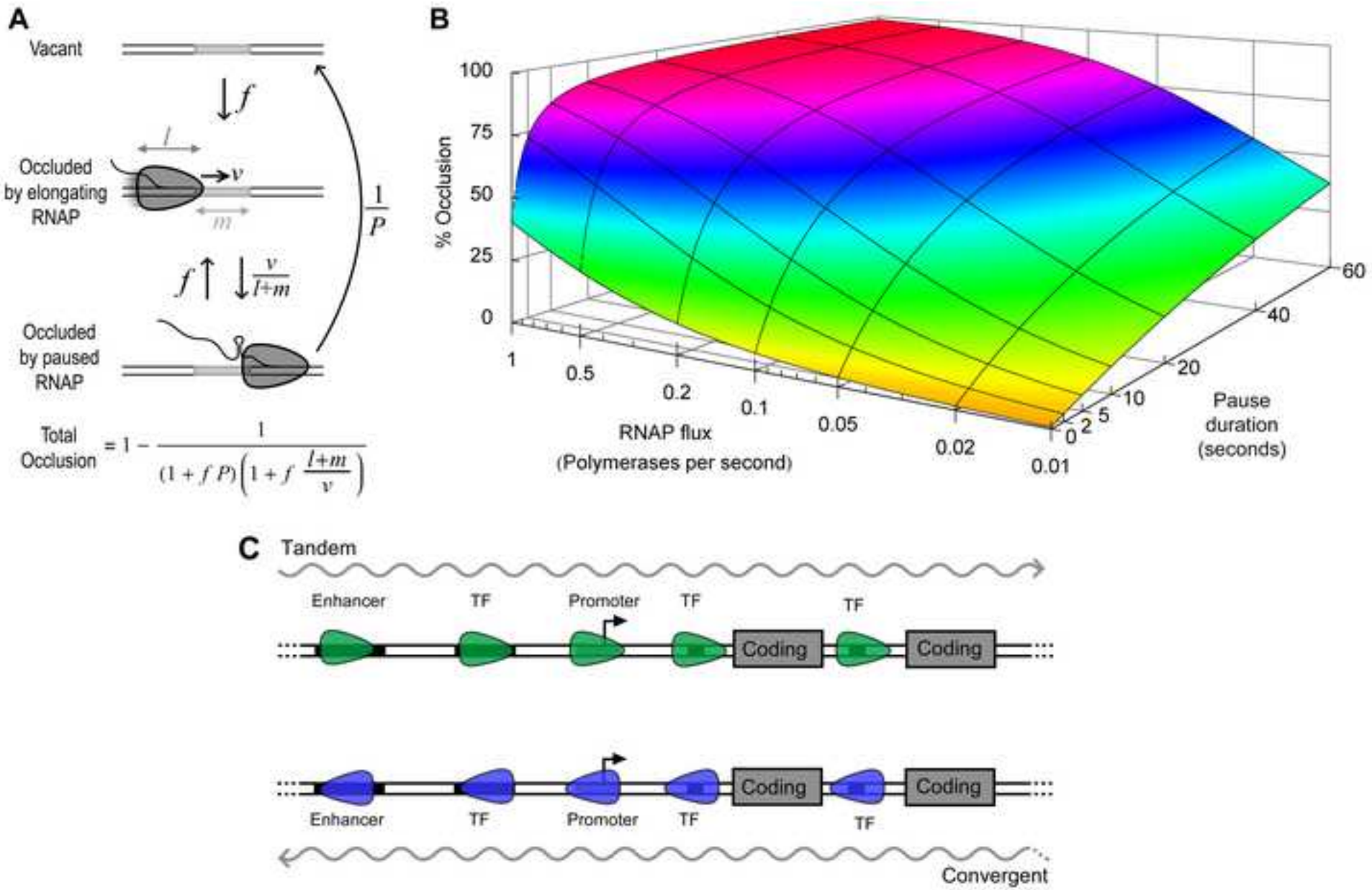
Click here to download high resolution image





E) Figure

Click here to download high resolution image



SUPPLEMENTAL DATA

Potent transcriptional interference by pausing of RNA polymerases over a downstream promoter

Adam C. Palmer, Alexandra Ahlgren-Berg, J. Barry Egan, Ian B. Dodd and Keith E. Shearwin

Modelling the PR-PRE system

Concentration scales

In order to convert between concentrations and numbers of molecules per cell, the volume of an *E. coli* cytoplasm is required. *E. coli* are shaped as cylinders with hemispherical ends, with a ratio of length to width of 4.4, over a range of growth rates (Donachie and Robinson, 1987). The dry weight of an *E. coli* cell is 6.4×10^{-13} g, and *E. coli* are 70% water, with a density of 1.05 g/mL (Bremer and Dennis, 1996). Therefore the total volume of an *E. coli* cell is 2.0×10^{-18} m^3 (2 fL), by which we calculate the cell to be 3.8 μm long and 0.9 μm wide. The volume of the cytoplasm is significantly less than 2 fL, as the periplasm occupies some 20% of the total volume of the cell, and making further subtractions for the volume occupied by the outer and cytoplasmic membranes (each 8 nm thick), the volume of the cytoplasm is estimated to be 1.4 fL (Neidhardt, 1990). In this volume then, one molecule is equivalent to a concentration of 1.2 nM.

CII tetramerization and DNA binding

The data of Figure 1E, showing P_{RE} activity versus CII molecules per cell, was fitted to the following reaction scheme: $4CII \leftrightarrows CII_4$ with K_{tet} (M^{-3}); $CII_4 + DNA \leftrightarrows CII_4.DNA$ with K_D (M). P_{RE} activity was assumed to be proportional to the fractional occupancy of the single binding site for the tetramer, and as calculated above, cell volume was taken to be 1.4 fL

The optimal values were $K_{tet} = 7.3 \times 10^{15}$ M^{-3} monomers (equivalent to $\Delta G° = -22.5$ kcal/mol of tetramer formed) and K_D = 350 nM tetramers. The measured K_D of CII tetramers for P_{RE} could not

be directly used as CII tetramer affinity in the model, since this would not account for the cooperative binding between CII and RNAP, measured to improve the RNAP binding 15-fold (Shih and Gussin, 1984). Alignment of experimental and theoretical P_{RE} activation curves along the CII axis is achieved with a raw K_D for CII tetramers of 870 nM, strengthened to 58 nM in the presence of cooperative binding to RNAP.

Promoter Kinetics

LacZ units were converted to RNAP initiations per second using the data of (Liang et al., 1999), who measured the RNAP initiation rate from the *E.coli* pBla promoter as a function of growth rate, indicating that under the growth conditions of our LacZ assays (doubling time 42 minutes), each copy of the pBla promoter initiates 0.021 RNAP / second. Our pBla.*lacZ* reporter generated 137 LacZ units, from which we converted LacZ activities for P_R and P_{RE} to RNAP initiations per second. Maximal P_R activity, as measured with the P_R (short) reporter lacking tR1, was 0.17 elongating RNAP per second. P_{RE} ranged from basal (CII^-) activity of 5×10^{-4} to maximal activity (saturated with CII) of 0.12 elongating RNAP per second.

As discussed in the main text, promoter kinetics are characterised by 3 parameters, KB, k_{co}, and k_{oe} (Supplemental Figure S3, Supplemental Table 1). For P_{RE} we have measured both basal and activated rates, and KB and k_{co} have been measured to be enhanced 15× and 40× by CII (Shih and Gussin, 1984), leaving 1 degree of freedom. This degree of freedom is fixed by selecting 'aspect ratio' equal to 1, which is the ratio of the rate of open complex formation and rate of initiation by open complexes. This choice maximises the modelled vulnerability of P_{RE} to TI, since interference by dislodgement of initiation complexes is maximised at aspect ratio 1 (Sneppen et al., 2005). As we are demonstrating that the initial model significantly underestimates the interference of P_{RE} by P_R, our claims are at their most conservative when kinetic parameters maximise interference of P_{RE}. Dislodgement of initiation complexes makes a significant contribution to TI when an interfering

promoter is much stronger than the interfered promoter (Sneppen et al., 2005), and so we expect this to make only a minor contribution to interference between P_R and P_{RE}, which are of similar maximum, non-interfered activities. When kinetic parameters of P_{RE} were changed to reduce aspect ratio by 100-fold (KB (CII^+) = 0.14, k_{co} (CII^+) = 1.1, k_{oe} = 10), simulated interference of P_{RE} by P_R, with 20 second pause times at tR1, changed from 5.6-fold to 4.6-fold. Therefore, provided that the values of KB, k_{co} and k_{oe} are set to produce the experimentally measured P_{RE} activity in the absence of TI (0.12 RNAP initiations/second), the simulated fold-interference is insensitive to alteration throughout KB, k_{co}, k_{oe} parameter space. Similarly, simulated interference of P_{RE} and P_R is insensitive to the specific kinetic parameters of P_R; only net RNAP initiation rate is of importance.

Supplemental Table S1. **Promoter kinetic parameters**

	KB	k_{co} (s^{-1})	k_{oe} (s^{-1})	Net RNAP initiations / second*
P_R	3.3	0.25	100	0.173
P_{RE} (CII^-)	0.018	0.03	0.26	0.0005
P_{RE} (CII^+)	0.27	1.22	0.26	0.123
Activation by CII:	*15 ×*	*40 ×*	*1 ×*	*235 ×*

*Calculation of net RNAP initiations / second from the kinetic scheme will result in slightly faster rates than those listed here, which are slower to account for 'self-occlusion' – an elongating RNAP takes some time to move off of its own promoter and make space for a new RNAP to bind (Sneppen et al., 2005).

Transcriptional interference

RNAP characteristics

The following characteristics of RNAP were using in the modelling:

Extent of DNA occupied by initiating RNAP at the promoter complex = +18 to -55 (Metzger et al., 1989),

Length of elongating RNAP = 30 bp (Epshtein et al., 2003),

Distance from the front of a stalled elongating RNAP to RNA 3' end = 6 bp (Toulme et al., 2000),

Collision of elongating RNAPs was modelled by instantaneous removal of one RNAP.

RNAP elongation speed

Our previous model of transcriptional interference (Sneppen et al., 2005) estimated RNAP velocity to be 40 bp/sec, using data in which changes in the spacing between converging promoters were seen to affect interference (Callen et al., 2004). This estimate was based on the assumption that both RNAPs are lost in a collision, while here we assume only one RNAP is lost, according to AFM imaging of collided *E.coli* elongation complexes demonstrating that collisions induce backtracking of only one RNAP (Crampton et al., 2006). The previous estimate also used a value for the firing rate of the interfering promoter obtained from a different calibration of LacZ units versus initiations per sec. We prefer our current calibration, which gives a 4-fold higher value for initiations per sec, because pBla is more similar in strength to the promoters studied in (Callen et al., 2004) than our past reference promoter. Re-evaluation of the spacing-interference data of (Callen et al., 2004) with these new assumptions gives a revised estimate for RNAP elongation speed of 60 bp/sec, which is used here and is within the range estimates of in vivo RNAP elongation speed obtained by other methods (Sneppen et al., 2005).

Modification of RNAP by N

When an N-modified RNAP passes through tR1, we assumed that it pushes paused, unmodified RNAP out of their pauses, and thus saves them from termination. Thus, in our model N both abolishes termination of modified RNAP, and also reduces the termination rate of unmodified

RNAP. Both mechanisms were taken into account when calculating the percentage of N-modified polymerases based on the observed termination frequency.

Supplemental experimental procedures

$KMnO_4$ footprinting in vivo

In vivo permanganate footprinting was performed on *E. coli* strain NK7049 carrying pTL61T based LacZ reporter plasmids. $P_{RE}.(P_R^-)$ and $P_{RE}.(P_R^+)$ constructs, containing either a wild type or mutated *rut* site were used. There was no source of CII in these strains, in order that there was minimal transcription from P_{RE} .To determine the location of the presumed pause site(s) of RNAP transcribing from P_R, plasmid DNA prepared from permanganate treated cultures were subject to primer extension analysis, as follows.

Overnight cultures were subcultured in 40 ml of fresh M9CAA medium (1x M9 salts, 100 μM $CaCl_2$, 1 mM $MgSO_4$, 0.2% glucose, 2 mg/ml casamino acids, 40 mg/ml thiamine and 100μg/ml ampicillin) and grown at 37°C to an OD_{600} of 0.6. $KMnO_4$ was added to the cultures to a final concentration of 3 mM and left for 2 or 4 minutes before quenching with an equal volume of ice cold STE (50 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA pH 8.0). The bacterial pellets were harvested by centrifugation at 6000 *g* for 25 min at 4°C and then resuspended in buffer P1 (Qiagen Miniprep Kit) with RNaseA added to a final concentration of 200 μg/ml. DNA was prepared using the Qiagen Miniprep Kit according to the standard protocol, with the exception that lysozyme was added to a final concentration of 2 mg/ml for 5 min at room temperature before adding 500 μl of buffer P2 and 700 μl of neutralizing buffer N3. DNA was eluted in 10 mM Tris pH 8.0 preheated to 70°C.

To assess the quality and concentration of the DNA preparations, samples were electrophoresed on a 0.8 % agarose gel, stained with GelRed (Biotium, CA), and imaged in a Typhoon Trio (GE Healthcare). To determine DNA concentrations, gel images were analysed using ImageQuant 5.2 software, using dilutions of a 2-log ladder (New England Biolabs) to generate a standard curve.

Two oligos were used for primer extension analysis; primer cro-Alex (5' GGAAGCGTTTATGCGGAAGAGG 3') located in the *cro* gene facing P_{RE} and primer alt-329 (5' CATGTTTGACAGCTTATCATCGGAGC 3') facing P_R. The oligos were end-labelled using [γ-^{32}P]-ATP (Perkin Elmer) and T4 polynucleotide kinase (New England Biolabs) and purified using MicroSpin G-25 columns (GE Healthcare). 1.2 pmol of either oligo was used for extension of 15 fmol of template DNA allowing for at least a 5-fold excess of primer at the end of the reaction. The extension reactions were performed in a final volume of 50 µl, adding dNTP to a final concentration of 200 mM and 5 Units of Taq DNA Polymerase (New England Biolabs). The extension reactions were run for 15 cycles (95°C 1 min, 57°C 30 sec, 72°C 30 sec) with an initial denaturation step of 95°C for 5 min. Extension products were dried under vacuum and resuspended in 9 µL loading buffer (98 % formamide, 10 mM Tris pH 8.0). Dideoxy sequencing reactions were prepared using the DNA Cycling Sequencing Kit (Jena Bioscience) using the same [γ-^{32}P]-ATP end-labelled primers. Extension products and sequencing reactions were electrophoresed on 6% polyacrylamide gels containing 6M Urea (National Diagnostics) in 1 x TBE (50 mM Tris, 100 mM boric acid and 5 mM EDTA). Dried gels were imaged by a Typhoon Trio (GE Healthcare) and analysed using ImageQuant 5.2 software.

Footprinting reactions performed following rifampicin treatment (0.2 mg/mL) were as described above, except that permanganate treatment was for 1 minute using 6 mM $KMnO_4$, and primer extensions were carried out using only primer alt-329. Gel bands were quantitated using the volume

integration function of Imagequant 5.2. Band intensities were normalised to a group of invariant bands upstream of pR to account for any differences in sample loading between lanes.

Western Blotting

After SDS-PAGE on NuPAGE 16% Tricine (Invitrogen) gels, samples were electrotransferred at 25V for 2 hours onto Hybond-LFP (Amersham). Immunoblotting was performed as per the ECL Plex Western blotting system (Amersham). Primary rabbit antibody to λCII, kindly provided by Amos Oppenheim (Hebrew University-Hadassah Medical School, Israel), was applied at 1/2000 dilution following a 1 hour pre-incubation at 37°C with a 1/60th volume of 10× concentrated, sonicated *cII*$^-$ cell extract. Secondary antibody was Plex Cy-5 anti-rabbit, applied at 1/4000 dilution.

Supplemental References

Bremer, H., and Dennis, P. (1996). Modulation of chemical composition and other parameters of the cell by growth rate. In Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology, F. C. Neidhardt, R. C. III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger, eds. (Washington DC, ASM Press), pp. 1553-1569.

Callen, B. P., Shearwin, K. E., and Egan, J. B. (2004). Transcriptional interference between convergent promoters caused by elongation over the promoter. Mol Cell *14*, 647-656.

Crampton, N., Bonass, W. A., Kirkham, J., Rivetti, C., and Thomson, N. H. (2006). Collision events between RNA polymerases in convergent transcription studied by atomic force microscopy. Nucleic Acids Res *34*, 5416-5425.

Donachie, W. D., and Robinson, A. C. (1987). Cell Division: Parameter Values and the Process. In Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology, F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger, eds. (Washington DC, ASM Press), pp. 1578-1593.

Epshtein, V., Toulme, F., Rahmouni, A. R., Borukhov, S., and Nudler, E. (2003). Transcription through the roadblocks: the role of RNA polymerase cooperation. Embo J *22*, 4719-4727.

Lau, L. F., Roberts, J. W., and Wu, R. (1982). Transcription terminates at lambda tR1 in three clusters. Proc Natl Acad Sci U S A *79*, 6171-6175.

Lau, L. F., Roberts, J. W., and Wu, R. (1983). RNA polymerase pausing and transcript release at the lambda tR1 terminator in vitro. J Biol Chem *258*, 9391-9397.

Liang, S., Bipatnath, M., Xu, Y., Chen, S., Dennis, P., Ehrenberg, M., and Bremer, H. (1999). Activities of constitutive promoters in Escherichia coli. J Mol Biol *292*, 19-37.

Linn, T., and St Pierre, R. (1990). Improved vector system for constructing transcriptional fusions that ensures independent translation of lacZ. J Bacteriol *172*, 1077-1084.

Metzger, W., Schickor, P., and Heumann, H. (1989). A cinematographic view of Escherichia coli RNA polymerase translocation. Embo J *8*, 2745-2754.

Neidhardt, F. C., J.L. Ingraham, and M. Schaechter. (1990). Physiology of the Bacterial Cell: A Molecular Approach (Sunderland MA, Sinauer Associates).

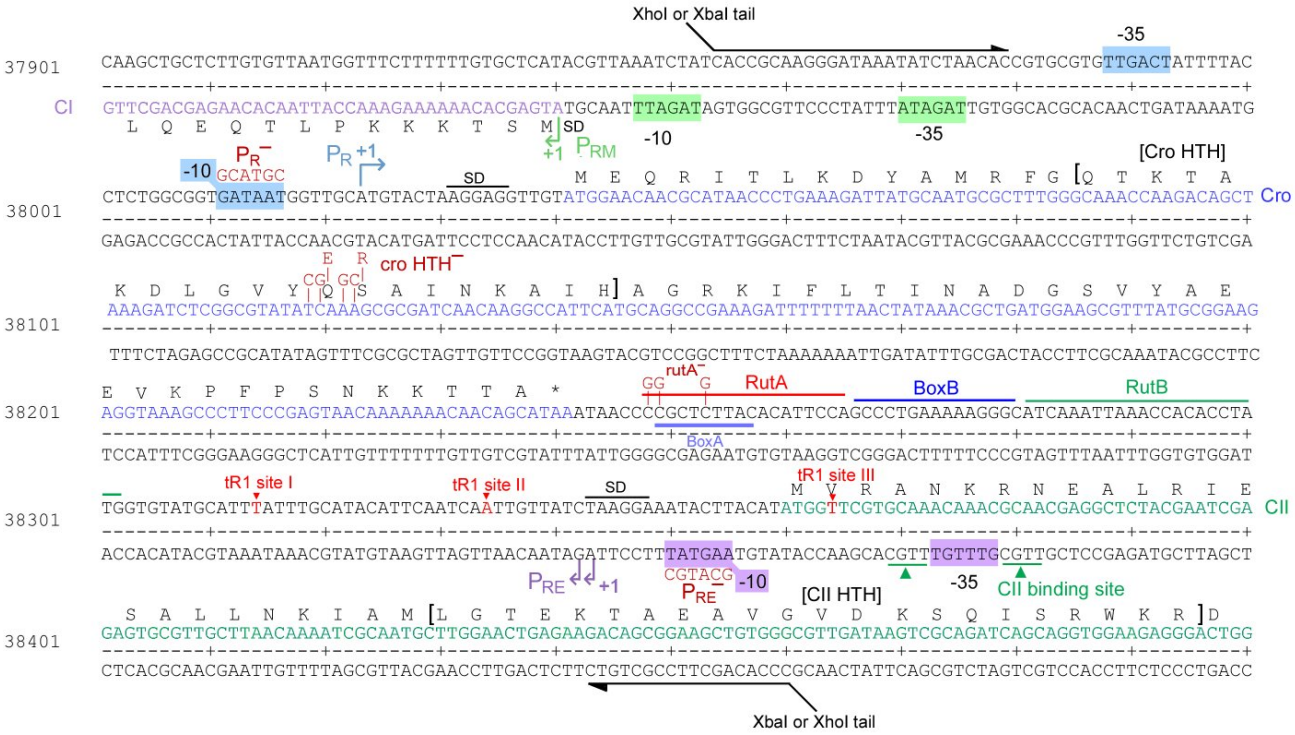
Powell, B. S., Rivas, M. P., Court, D. L., Nakamura, Y., and Turnbough, C. L., Jr. (1994). Rapid confirmation of single copy lambda prophage integration by PCR. Nucleic Acids Res *22*, 5765-5766.

Shih, M. C., and Gussin, G. N. (1984). Role of cII protein in stimulating transcription initiation at the lambda PRE promoter. Enhanced formation and stabilization of open complexes. J Mol Biol *172*, 489-506.

Sneppen, K., Dodd, I. B., Shearwin, K. E., Palmer, A. C., Schubert, R. A., Callen, B. P., and Egan, J. B. (2005). A mathematical model for transcriptional interference by RNA polymerase traffic in Escherichia coli. J Mol Biol *346*, 399-409.

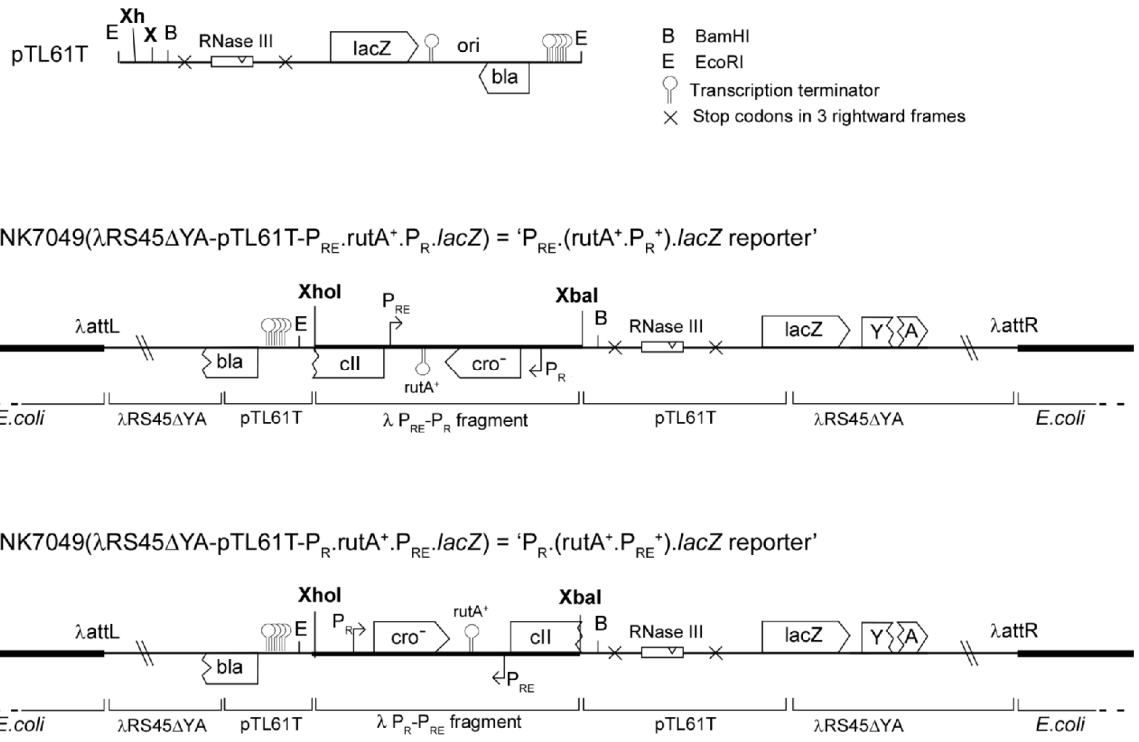
Toulme, F., Mosrin-Huaman, C., Sparkowski, J., Das, A., Leng, M., and Rahmouni, A. R. (2000). GreA and GreB proteins revive backtracked RNA polymerase in vivo by promoting transcript trimming. Embo J *19*, 6853-6859.

Supplemental Figures



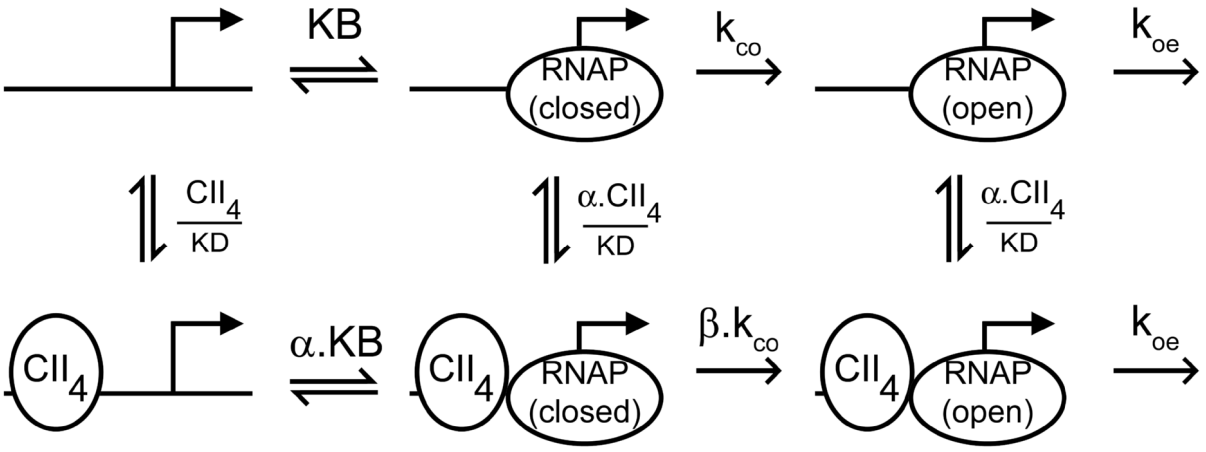
Supplemental Figure S1. **Sequence and features of bacteriophage lambda P_R to P_{RE} region.** The -10 and -35 regions of P_{RM}, P_R and P_{RE} are shown shaded in green, blue and purple, respectively. Protein coding sequences are shown in light blue, and their ribosome binding sites (SD) are overlined. The *rut*A and *rut*B sequences (between the *cro* and *c*II genes), which together make the Rho utilisation site (*rut*) are shown, as are the BoxA and BoxB sequences, which form the *nut*R site.

The PCR primers used to amplify the P_R to P_{RE} region for insertion into the *Xho*I and *Xba*I sites of pTL61T (Linn and St Pierre, 1990) are shown, as are the P_R^- and P_{RE}^- mutations, the Cro$^-$ amino acid changes within the Cro helix-turn-helix motif (HTH), and the three base-pair changes which together make the *rut*$^-$ mutation. The three termination sites at tR1 (bases shown in red) are indicated as tI, tII and tIII (Lau et al., 1982; Lau et al., 1983). The CII binding site, consisting of two direct repeats either side of the P_{RE} -35 sequence are indicated as green lines.



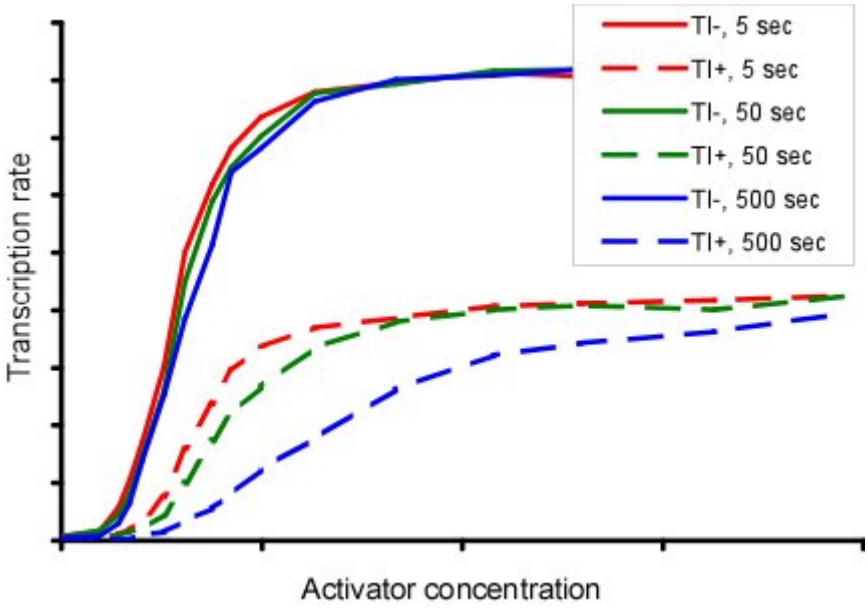
Supplemental Figure S2. **Construction of lacZ reporters**

Promoter fragments to be assayed (sequence positions 37954 to 38461) were amplified by PCR with the proofreading Pfu DNA polymerase with primers adding *Xho*I and *Xba*I ends; 5' *Xho*I and 3' *Xba*I for pR to control lacZ, 5' *Xba*I and 3' *Xho*I for P_{RE} to control lacZ (where 5'→ 3' is from P_R towards P_{RE}). These PCR products were ligated into the polylinker of pTL61T; both junctions and the insert were checked by sequencing. This was recombined in vivo with the modified lambda phage, λRS45ΔYA. Reporter phage were purified and lysogenised in NK7049, and single lysogens were isolated (Powell et al., 1994). These lysogens contained a single copy of the reporter fragment at the λ attachment site in the *E. coli* chromosome, with either P_R or P_{RE} directing transcription of the lacZ gene. Between promoter fragments and the lacZ gene is an RNaseIII cleavage site, which causes lacZ RNA to be separated from RNA transcribed within the reporter fragment; this is observed to reduce potential context effects from different constructs (Linn and St Pierre, 1990).



Supplemental Figure S3. **Kinetic scheme of promoters in simulations of TI**

P_{RE} is modelled with this full kinetic scheme, while for P_R only the top row of states (lacking CII) are relevant. KB is the dimensionless equilibrium binding constant of RNAP to a promoter, treating RNAP concentration as a constant which is absorbed into the value of KB. k_{co} (sec^{-1}) and k_{oe} (sec^{-1}) are the rate constants for closed to open complex isomerisation, and initiation of elongation from an open complex. KD (CII tetramers per cell)$^{-1}$ is the equilibrium dissociation constant of CII tetramers for their binding site at P_{RE}. $\alpha = 15$ is the fold enhancement of KB in the presence of CII, and $\beta = 40$ is the fold enhancement of the rate constant for closed to open complex isomerisation (Shih and Gussin, 1984). As only the slow, non-equilibrium reactions k_{co} and k_{oe} are treated stochastically, this reaction scheme reduces to two irreversible steps: an effective isomerisation rate, and the rate of initiation of elongation (k_{oe}). For P_R the effective isomerisation rate is $k_{co}.KB.(1 + KB)^{-1}$, while for P_{RE} the effective isomerisation rate is $k_{co}.(KB + \alpha.\beta.KB.CII_4.KD^{-1}).(1 + KB + CII_4.KD^{-1} + \alpha.KB.CII_4.KD^{-1})^{-1}$.



Supplemental Figure S4. **Dislodgement of a DNA-binding activator increases the activator's EC50**

Simulations are of an activator-dependent promoter, where the activator spans a range of association/dissociation kinetics; times listed in legend are (dissociation rate constant)$^{-1}$. Variable speeds of association/dissociation reactions have no effect upon a non-interfered (TI–) promoter, as only the equilibrium binding constant, here held constant, influences the EC50. In the presence of interference (TI+), slow association/dissociation kinetics leave the promoter vulnerable to interference by activator dislodgement, with transcription initiation rate decreasing more with slower kinetics. As effective association rate constant increases with activator concentration, the effects of activator dislodgement are eventually overcome with a sufficiently high concentration of activator, and so TI by activator dislodgement manifests as an increase in EC50, rather than a decrease in maximum possible transcription rate.